A Survey on Collaborative Joint Search On Mobile Device with Location and Maps

Smita M. Dahegaonkar, Prof. Alka Khade, Dr. Seema Biday
Student In Master of Engineering, Department of Computer Engineering, Terna College of Engineering, Nerul, Navi Mumbai, Maharashtra, India
Professor, Dept. of C.E, Terna College of Engineering, Nerul, Navi Mumbai, Maharashtra, India
HOD, Dept. of Electronics, Terna College of Engineering, Nerul, Navi Mumbai, Maharashtra, India

ABSTRACT: The system presented in this paper depicts a novel multimodal interactive image search system on mobile device. It is intended for clients who as of now have pictures in their minds yet have no exact depictions or names to address them. By depicting it using speech and after that refining the perceived question by intuitively creating a visual inquiry using exemplary images, the client can without much of a stretch discover the desired pictures through a couple of regular multimodal connections with his/her cell phone: SIFT features and image annotations (tags) as well as the combination of SIFT.

KEYWORDS: SIFT, CCV, Texture, Gray Scale, SURF.

I. INTRODUCTION

Present day time cellular telephones and tablets have developed into effective picture and feature transforming gadgets, prepared with high-determination cams, color displays, and hardware-accelerated graphics. They are too outfitted with area sensors (GPS recipient, compass, and gyrorator), and associated with broadband remote systems, permitting quick data transmission and empowering a class of applications that use the telephone’s inherent cam to launch search queries about objects in the client’s visual nearness.

Now-a-days image search is an intriguing issue in both machine vision and information retrieval with numerous applications. The conventional desktop image search frameworks with text queries have ruled the client conduct for a long stretch. Contrasted and content pursuit, guide hunt, and pho-to graph to inquiry, visual (picture and feature) inquiry is still not extremely famous on the telephone, however image search has turned into a typical instrument on the PC since 10 years prior, with which the client can include content question to recover pertinent pictures. A primary motivation behind why such picture seek applications are not prominent on cell phone is that the current picture search applications don’t splendidly suit to the portable and nearby situated client aim. Because of this, the indexed lists are seldom valuable and the client encounter on the telephone is not generally pleasant. Most importantly, writing is a dull employment on the telephone regardless of whether a little console or a touch screen is utilized. Despite the fact that voice queries are accessible on a few gadgets, there are still numerous cases that semantic and visual aim can barely be communicated by these descriptions for search. For instance, in a typical picture pursuit under-taking, the client may have officially imagined the general thought of expected pictures, for example, color designs and arrangements. Nonetheless, the client normally needs to get perfect pictures in the midst of substantially more unimportant results.

II. LITERATURE REVIEW

Bernd Girod et al. gave an extensive outline of photograph to-search in [1], from the structural planning of an effective versatile framework to the system of a image recognition algorithm. Robust local image features achieve a high degree of invariance against scale changes, rotation, as well as changes in illumination and other photometric conditions. The BoW approach offers resiliency to partial occlusions and background clutter, and allows design of efficient indexing schemes. In case of visual representation of images, the images are over segmented before color
III. RELATED WORK

Bernd Girod et al. gave an extensive outline of photograph to-search in [1], from the structural planning of an effective versatile framework to the system of a image recognition algorithm. In case of visual representation of images, the images are over segmented before color feature extraction using algorithm by Felzenszwalb and Huttenlocher [2]. Affinity Propagation (AP) algorithm [4] is received to group the applicant pictures into a few gatherings. Speech recognition is now a much more mature technique than image recognition. Especially for speech to text technique, the state-of-art accuracy achieves an accuracy of 98% in quiet environment using Hidden Markov Models [5]. Tuytelaars, Bay and Van Gool presents a novel scale and rotation invariant detector and descriptor, coined SURF (Speeded-Up Robust Features) [6]. In paper [7], a multimodal mobile search framework is designed intended to do visual search. In [10] the creators fabricate a Sketch2photo framework that uses basic content explained line representation to naturally combine practical pictures. They additionally utilize content and representation to hunt down formats which are then sewed on a foundation to create a montage. Then again, their work concentrates on picture making rather out of picture retrieval. As the speech recognition got to be develop, telephone applications utilizing speech recognition quickly becomes as of late. The most illustrative application is Apple Siri [9], which joins speech recognition; natural language understanding and learning based searching procedures. In scholastic circles, there is not an enormous contrast from industry. Researchers for portable pursuit likewise concentrate fundamentally on photograph to-inquiry methods. Conventional features, for example, SIFT [8] and Speeded Up Robust Feature (SURF) [6] are broadly utilized as a part of such visual search frameworks due to their invariance to illumination, scale and rotation. Chandrasekhar et al. talked about their compres-sion and in addition proposed another feature of Compressed Histogram of Gradients (CHoG) [12]. It can quantize and encode inclination histogram with Huffman and Gagie trees to deliver low bit-rate descriptors. They then develop an efficient segmentation algorithm based on this predicate, and show that although this algorithm makes greedy decisions it produces segmentations that satisfy global properties.

On Flickr [3], everyone gets 1000GB of free storage, enough space for more than 500,000 photos. Their powerful search technology means you can find them anytime you want. Its the world’s largest photography community. Affinity Propagation (AP) algorithm [4] is received to group the applicant pictures into a few gatherings. Frey and Dueck devised a method called affinity propagation, which takes as input measures of similarity between pairs of data points. Real-valued messages are exchanged between data points until a high-quality set of exemplars and corresponding clusters gradually emerges. Speech recognition is now a much more mature technique than image recognition. Especially for speech to text technique, the state-of-art accuracy achieves an accuracy of 98% in quiet environment using Hidden Markov Models Tuytelaars, Bay and Van Gool presents a novel scale and rotation invariant detector and descriptor, coined SURF (Speeded-Up Robust Features) [6]. SURF approximates or even outperforms previously proposed schemes with respect to repeatability, distinctive- ness, and robustness, yet can be computed and compared much faster. This is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors; and by simplifying these methods to the essential. This leads to a combination of novel detection, description, and matching steps. In paper [7], a multimodal mobile search framework is designed intended to do visual search. In [10] the creators fabricate a Sketch2photo framework that uses basic content explained line representation to naturally combine practical pictures. They additionally utilize content and representation to hunt down formats which are then sewed on a foundation to create a montage. Then again, their work concentrates on picture making rather out of picture retrieval. As the speech recognition got to be develop, telephone applications utilizing speech recognition quickly becomes as of late. The most illustrative application is Apple Siri [9], which joins speech recognition; natural language understanding and learning based searching procedures.
The entire search interaction starts with the natural language understanding. In the framework, a basic entity extraction technique is embraced to handle this issue. The entity extraction from speech can be separated into two steps: 1) speech recognition, and 2) key entity extraction. After the voice is caught by the phone, it is first sent to a commercial online speech recognition engine to interpret the recorded sound stream into a text string. A Hidden Markov Model (HMM) and N-gram-based engine has the capacity to handle both natural sentences and phrase fragments. In the framework, once the content is accessible, the core entities inside the content will be extracted. Typically, such word extraction is carried out by either key word extraction or unimportant word filter. These techniques are troublesome and questionable. A basic approach to extract just meaningful words in the sentence is adopted. A word reference can be first naturally settled by WordNet, which is famous distributed various leveled word database unveil the hierarchical relationships among English words.

Visual Query Formulation

Not the same as traditional text based image search framework, in the image search framework, the client can utilize a visual query to further expressive the client’s visual expectation. In the framework, to form a visual inquiry, the client require just to pick one or more exemplary images in our display to make a composition as per the client’s visual goal.

1) Visual Exemplar Generation: A visual exemplar database ought to be prepared to give exemplary images to the client as indicated by their key words. Accordingly, a lot of exemplary pictures are produced offline. Given a keyword, a set of pictures are initially downloaded from the Internet by asking a image search engine. Normally, top results from an Internet web index for straightforward ideas are dependable than numerous different assets (e.g., Flickr). It is observed that the top results constantly well cover the ideas, so the main 200 pictures are utilized to abstract representative pictures. These pictures are then clustered into a few groups as indicated by their visual content. The impact of the background in pictures can truly harm the execution of the content based clustering, so the background ought to be uprooted before clustering. Since the candidate visual exemplars are queried with single idea, it is expected that these pictures are with single objects and straightforward backgrounds. Thus, a basic salient region detector is applied to these candidate exemplars to spot the object area. To further uproot the impact of background and to exactly select the forefront part, Grabcut is received to section the frontal area object from the picture. To describe the foreground object, Bag-of-word (BOG) model is utilized. At long last, a visual words histogram is created for a picture utilizing the Bag-of-word model.
IV. ACKNOWLEDGEMENT

We would like to take this opportunity to express my sincere gratitude to my Project Guides Prof. Alka Khade (Assistant Professor, Computer Engineering Department) and Dr. Seema Biday (HOD, Electronics Department) for their encouragement, guidance, and insight throughout the research and in the preparation of this dissertation. They are truly exemplifies the merit of technical excellence and academic wisdom.

V. CONCLUSION

In this paper, an interactive mobile visual search framework is introduced which permits the clients with form their search purpose through natural multimodal interactions with cell phones. The system represents the first study on mobile visual search by taking the points of interest of multimodal and multi-touch functionalities on the telephone. The gives a game-like interactive image search plan with composition of various exemplars. The visual query created by the client can be effectively used to retrieve similar images by this system.

REFERENCES

4. B. Frey and D. Dueck, "Clustering by passing messages between data points", Department of Electrical and Computer Engineering, University of Toronto, 10 Kings College Road, Toronto, Ontario M5S 3G4, Canada., 2007.