



I-MIS: Intelligent Mobile Image Searching Technique using Image Composition and ANN

Yojana S. Pawale, Prof. Prakash Devale

Research Scholar, Dept. of IT, Bharati Vidyapeeth Deemed University, College of Engineering, Pune City, India

Professor, Dept. of IT, Bharati Vidyapeeth Deemed University, College of Engineering, Pune City, India

ABSTRACT: The system provides a powerful technique for image searching called 'I-MIS: Intelligent Mobile Image Searching Technique using Image Composition and ANN'. The system uses multimodal mechanism to find images through android mobile phones where users can provide input in text, image or voice form. The system is divided into three phases- Image Composition, Image Processing and ANN. So that it provides an effective way to find most relevant images to the user. The best thing of it is that user can again provide one of the resulted images as an input to the system to find more relevant images. It can be also possible for users to provide speech input to the system where system first converts voice query into text form and then applies text searching mechanism on it. So that this system provides user friendly mechanism to the users by providing multiple choices for providing input to the system to find required images.

KEYWORDS: Image Search, ANN, LAB image, Image Processing, Multimodal Input.

I. INTRODUCTION

Image searching is a process of finding, retrieving and then providing desired images to the users. One database is maintained at server side which contains number of images. When user provides any query to find images to the system then it's a responsibility of that system to provide desired images to the users. It can also be possible to perform image searching on mobile phones where users can be able to find desired images at anytime from anywhere. Hence, this technique is more populated on mobile phones. But it is not an easy process. There are many problems faced by the system during image searching process because of following reasons- It is very difficult to display full image on mobiles because of their screens. It is also not be easy to provide input to the system from mobile phones.

II. RELATED WORK

Various techniques are used for image searching. Text based technique is mostly used for image searching in which system accepts text query from user and provides relevant images to them. But system is not works well while dealing with longer queries. Basically, mobile users use only 2.6 terms on average for search [2] which is hardly to express their intent. Some more techniques are also used for image searching. Photo- to- Search technique is one of them where user finds desired images from the system by providing any other relevant image as an input to the system. This technique is used by Google Googles[4], Point and Find[5], Snaptell[6] where system provides some extra information to users with resulted image. But this technique is only used for landmarks, CD covers etc. One more limitation of it is that user should have relevant image to find targeted images from the system. Apple Siri[3] is an one application used on mobile phones which uses speech recognition and language understanding approaches to provide knowledge based searching mechanism. JIGSAW+ [1] is a one more technique used for image searching which also provides multimodal mechanism to users but it does not considers user's intent behind image searching. For example, if user provides a query as 'Head', then whether users want images related to 'human head' or 'head of department' is not considered by the system.

Hence we have implemented one more approach for image searching called 'I-MIS: Intelligent Mobile Image Searching Technique using Image Composition and ANN' where we have combined three techniques- text, image and speech based technique to provide a powerful mechanism for image searching. Section 2 of this paper introduces

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

proposed system. System architecture is represented in section 3. It also describes each module in detail. Section 4 contains experimental work and finally last section concluded this concept.

III. PROPOSED SYSTEM

System works in three different phases as Image Composition, Image Processing and ANN. System first searches and provides composed images to users according to the user's query. Then user selects some relevant images among them and again provides them as an input to the system. Then system applies image processing technique on that composed images to process them. Finally, with the help of ANN technique, system provides most relevant images to the users as a final output. Here, user can again provide resulted image to the system to find some more relevant images. Fig. 1 shows basic idea of proposed system-

Fig. 1(a) represents voice/text based technique where system accepts voice input and provides relevant images. Fig. 1(b) represents image based technique where system provides desired images by accepting image input. Fig. 1(c) represents proposed system where system combines all these three techniques and provides more effective way for image searching. The proposed system provides three options for users to give input to the system and finds relevant images from it.

IV. DETAILED SYSTEM DESCRIPTION

Fig. 2 represents system architecture. Proposed system works in three phases as Image Composition, Image Processing and ANN. System first provides composed images to the user according to their query where user can send the query in the form of text, image or voice. Then, user selects some composed images and sends to the system as an input. Then system processes that images to extract correct features from it and applies ANN mechanism to find relevant images present in the database. Lastly, system provides these relevant images to the user.

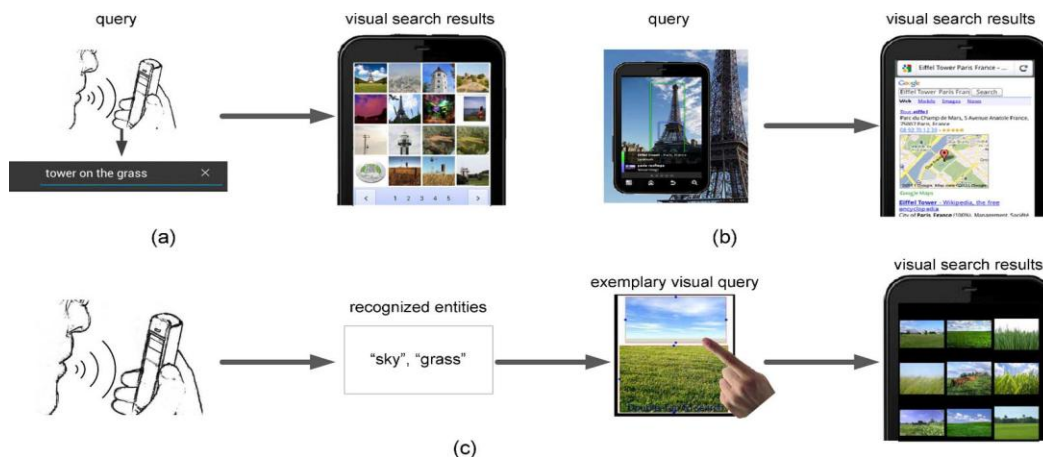


Fig. 1 Basic Idea of Proposed System

A. IMAGE COMPOSITION

Image Composition is the first phase of the system where system accepts input from the user in the form of text, image or speaker sends voice input, system uses android's speech to text API to convert it into text form. Then system divides text query into different keywords and provides number of composed images to the user. After that user send some composed images among them to find more images.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

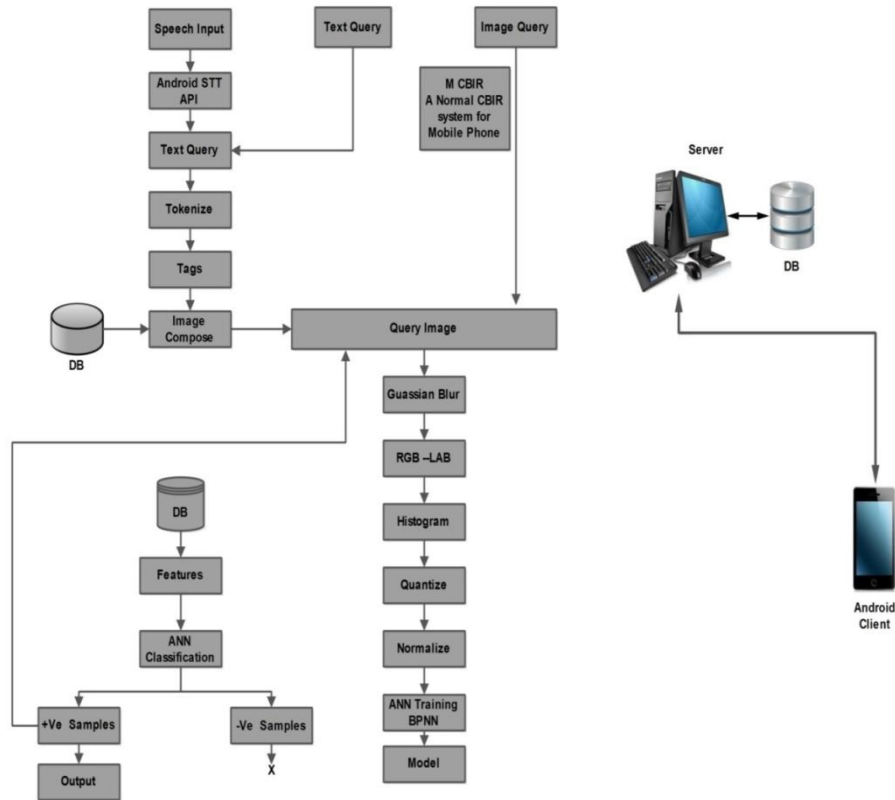


Fig. 2 Architecture of Proposed System

B. IMAGE PROCESSING

In the second phase, system converts image provided by user into LAB form for further processing. Then system performs image processing to extract features from it where it creates histogram of an image to represent its features.

C. ANN

Lastly, ANN technique is applied on image where BPNN mechanism is used that compares ANN features of an image present in database with result of BPNN. Finally system gives most relevant images to the user as a final output.

V. EXPERIMENTAL WORK

Proposed system has provided three options to users as shown in fig. 3



Fig. 3 Three Options

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

If user selects an option- text search then system will provide composed images to the user according to that query as shown in fig. 4

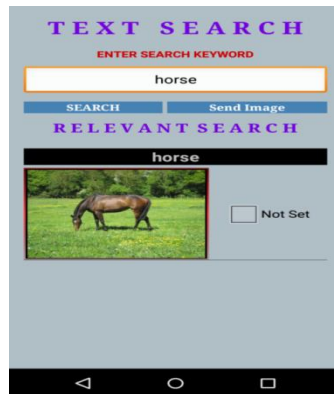


Fig. 4 Composed Image from Text search

Then if user selects some images and clicks on send image option. Then system accepts these images as an input and provides final relevant images to the user as shown in fig. 5.



Fig. 5 Resulted Images from Text Search

Here, user can again selects number of images within resulted image set and can sends as an input to the system to find more relevant images. As shown in fig. 6, if user selects and sends 'grass' image then system will provide images related to 'grass'.



Fig. 6 Images of Grass

In the same way user can send voice or image query to the system to find targeted images.

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

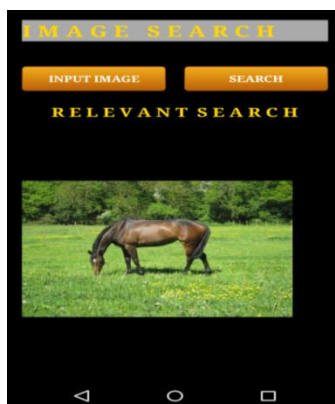


Fig. 7 Input Image for Image Search

As shown in fig. 6, if user sends 'Horse' image as an input to the system then system will provide images related to 'Horse' as shown in fig. 7.

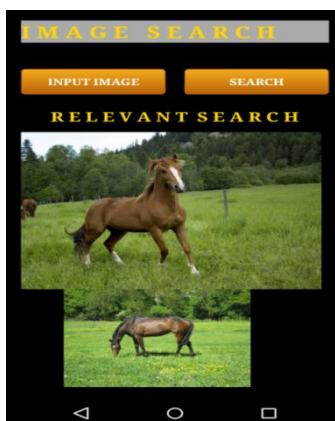


Fig. 8 Result of Image Search

VI. CONCLUSION

The system provides a new mechanism for image searching through mobile phones where user can give an input query in the form of text, image or voice to the system to find relevant images. Hence, this system is also called as multimodal system. The best part of this system is that it makes a use of ANN mechanism to provide most relevant images to the user. System also uses LAB image form for processing an input image by converting original RGB image into LAB form which helps to extract accurate features from an image. If user does not wants to type any text query then by simply providing voice input to the system user can find desired images from the system. It is also possible to provide any other relevant image as an input to the system for retrieving desired images. Hence this system provides user friendly mechanism to the user by providing multiple options to give input to the system to find required images. System also provides most relevant images to the user by using ANN technique.

REFERENCES

1. Houqiang Li, Senior Member, IEEE, Yang Wang, Tao Mei, Senior Member, IEEE, Jingdong Wang, Senior Member, IEEE, and Shipeng Li, Fellow, IEEE "Interactive Multimodal Visual Search on Mobile Device," IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 3, APRIL 2013.
2. 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.
3. Siri. [Online]. Available: <http://www.apple.com/iphone/features/siri.html/>
4. Google Goggles. [Online]. Available: <http://www.google.com/mobile/goggles/>
5. NOKIA Point and Find. [Online]. Available: <http://pointandfind.nokia.com/>



International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 8, August 2015

6. SnapTell. [Online]. Available: <http://www.snaptell.com/>
7. H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded-up robust features," in Proc. ECCV, 2008, vol. 110, no. 3, pp. 346–359.
8. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in Proc. IEEE Int. Conf. Comput. Vis., Nov. 2011, pp. 2564–2571.
9. 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.
10. M. Jia, X. Fan, X. Xie, M. Li, and W. Ma, "Photo-to-search: Using camera phones to inquire of the surrounding world," in Proc. Mobile Data Manag., 2006.
11. S. Tsai, H. Chen, D. Chen, G. Schroth, R. Grzeszczuk, and B. Girod, "Mobile visual search on printed documents using text and low bitrate features," in Proc. IEEE Int. Conf. Image Process., Sep. 2011, pp. 2601–2604.
12. 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.
13. X. Liu, Y. Lou, A.W. Yu, and B. Lang, "Search by mobile image based on visual and spatial consistency," in Proc. IEEE Int. Conf. Multimedia Expo, 2011, pp. 1–6.
14. B. Girod, V. Chandrasekhar, D. Chen, N. Cheung, R. Grzeszczuk, Y. Reznik, G. Takacs, S. Tsai, and R. Vedantham, "Mobile visual search," IEEE Signal Process. Mag., vol. 28, no. 4, pp. 61–76, 2011.
15. S. Tsai, H. Chen, D. Chen, G. Schroth, R. Grzeszczuk, and B. Girod, "Mobile visual search on printed documents using text and low bitrate features," in Proc. IEEE Int. Conf. Image Process., Sep. 2011, pp. 2601–2604.
16. S. S. Tsai, D. Chen, H. Chen, C.-H. Hsu, K.-H. Kim, J. P. Singh, and B. Girod, "Combining image and text features: a hybrid approach to mobile book spine recognition," in Proc. ACM Multimedia, 2011, pp. 1029–1032.
17. Z. Zha, L. Yang, T. Mei, M. Wang, and Z. Wang, "Visual query suggestion," in Proc. ACM Multimedia, 2009, pp. 15–24.