International Journal of
Scientific Engineering and Technology Research

www.ijsetr.com

ISSN 2319-8885
Vol.05,Issue.07,
March-2016,
Pages:1384-1389

Interactive Multitudinous Visual Search on Mobile Device

ROSHANI E BAGUL[1], P. P. ROKADE[2]

[1]PG Scholar, Dept of Computer Engineering, S.N.D.C.O.E.R.C. Yeola, Nashik, SPPU, Maharashtra, AP, India,
E-mail: roshanibgl1@gmail.com.

[2]Professor, Dept of Computer Engineering, S.N.D.C.O.E.R.C. Yeola, Nashik, SPPU, Maharashtra, AP, India,
E-mail: prakashrokade2005@gmail.com.

Abstract: This paper describes the robust interactive multitudinous visual search system on Android mobile device. This system is able to take any type of input e.g. text, image, voice and perform visual search on mobile device. It gives the relevant results what an user want by taking the full benefit of interactive nature of mobile device. The main focus is on color features, the searching method searches similar images from the image database by matching the feature vector to locate object that matches query object perfectly. Proposed system makes the use of Naïve-Bayes algorithm which classifies the object according to their features and predicts the new label to an image. Main contribution of this paper is Multimedia search, in which image is given as input and user get similar youtube links, web links. This is designed for the users who have number of pictures in their mind but unable to describe and address them, and cannot achieve expected results for given query. The proposed system will differ from existing system by having following features- 1) Provide better features to user 2) Performs automated image prediction 3) Increased search performance.

Keywords: Multimodal Visual Search, Relevant Images, Feature Vector, Automated Image Annotation.

I. INTRODUCTION

Visual Search is becoming difficult now a days but it is one of the most popular applications on mobile devices. People are more and more addicted for mobile search. It is examined that one-third of search queries comes from mobile phones by 2014. It is most difficult task to type a text as query on desktop devices for that the existing system will not offer best user interaction for search process. This paper describe about the search system which is designed for the users who already have number of imaginations in their minds but not able to describe them. So by describing them using text, photo, speech, Image, their database. Matched images are then displayed to user; if these feature vectors are not available in database then such images are searched in Google's database using Google Rest API. The proposed system gives better user interaction; this application can be a better featured application of android mobile as user can search for something just while on the way.

A. Motivation Of The Proposed Approach Definition

Consider an example when a user goes to unfamiliar places and takes snacks from one unknown restaurant. When next time to visit the same restaurant 1) He simply takes a photo of that one. 2) Another situation is that he forgets to take photo of that restaurant. After returning to home he remembered only the appearance of the restaurant. Also he has no idea of the name of restaurant, but he can describe its particular appearance such as "A 3 star restaurant with black door, three blue lamps lions, and many purple pillars in front". Proposed system can handle these two situations. In first case, system uses captured restaurant's photo as an image query and start searching process and retrieve similar images from the web. In second case, user doesn't have an existing image but the user can generate an image query by giving speech as an input query to the system, which represents the picture described in the user's mind.

II. LITERATURE SURVEY

Number of system deployed on a real-world phone system, evaluated the performance on one million images, As a mobile visual search system, the most related works include many multimedia search apps available on mobile devices.

A. Traditional Text-Based Search

Google and Bing which are still available on mobile devices for text based image search. Lengthy text queries are neither user-friendly, nor machine-friendly for search engine. It can hardly express their search intent. For text based search tag should be given to each image for search process, and due to manual annotation it affects the search performance, and requires lots of man power. Table 1 summarizes the recently developments in mobile phone applications for multimedia search.

B. Traditional Speech Based Search

The most representative application like Apple Siri [11], combines speech recognition, natural language understanding and knowledge- based searching techniques. The user is enabled to make a speech conversation with the phone and get information and knowledge from it. It provides unprecedented ask-and answer user experiments. The user can ask the phone for anything by only speech and get multimedia answers.

C. Photo-To-Search

Photo-to-search applications also became popular on mobile phones. Such applications enable users to search for what they see by taking a photo on the go. As we have summarized in TABLE I, Google Goggles[13], Point and Find[14], and Snaptell[15] are good examples in this field. 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. In academic circles, there is not a big difference from industry. Researchers for mobile search also focus mainly on photo-to-search techniques. The research topics include visual feature design, database indexing, and transmission.

D. Content Based Search

Regarding content-based image search, one kind of famous products, including Google Image, TinEye[12] on PC, and GoogleGoggles[13] on mobile phone, can accept single images as search queries, and return to the user similar images or even with information mined from their databases. With very large databases, these engines are able to achieve impressive results. However, to initiate such a visual search, the user must have an existed image on hand as a query. Moreover, it needs partially duplicate images or exact the same thing existing in the database.

TABLE I: Image Search Applications

Apps	Features
Goggles	Product, Cover, Landmark, Name Card
Digimarc Discover	Print, Article, Ads
Point And Find	Place, 2D Barcode
Snapnow	MMS, Email, Print, Broadcast
Kooaba	Media Cover, Print

E. Sketch Based Search

Another kind of image search engines designed for desktop, including GazoPa[16] and some other sketch-based image search researches, use hand-drawn sketches to search for satisfied images. Though 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. MindFinder[5] and Concept Map[6] also provide visual aids to search for images. In these works, visually and semantically similar images are retrieved by multiple exemplary image patches. Apps Features Goggles Product, cover, landmark, name card Digimarc Discover Print, article, ads Point and Find Place, 2D barcode Snapnow MMS, email, print, broadcast Kooaba Media cover.

III. PROPOSED SYSTEM

With the rich interactions and visual techniques, The proposed image search system enables the user to conduct a image search with visual aids. The objective is to design a efficient visual aided image search application on mobile phone combined with local spot and scene search. The proposed system contain three modules, text to image search, photo to image search, speech to image search, The proposed system handles following situations. In first case text is used as query to retrieve an image, input text is parsed and stop words are removed using porter stemmer algorithm, resulted keywords are matched with Image database by matching their feature vector, In second case image is given as input query such as a captured photo of the railway and firstly in training phase feature vector of entered image is created in testing phase it is compared with the feature vector of all database images , images are matched by cosine similarity, if calculated value is greater than threshold, then matched images are displayed on screen and feature vector of input image with new label is stored in database. If the input image does not match with any database image then it is searched in Google Image database, input query is passed as URL to Google through web service Using Naïve Bayes algorithm, during classifications and clustering it predict the label for input image and assign the label to feature vector of new image, which increases the search performance. In the third case, user's doesn't have an existing image, but the user can generate an image query by giving speech input to the system, that represent picture described in the user's mind. In forth case, user gives image as an input query and user get various video links related to input query as shown in Fig.1.

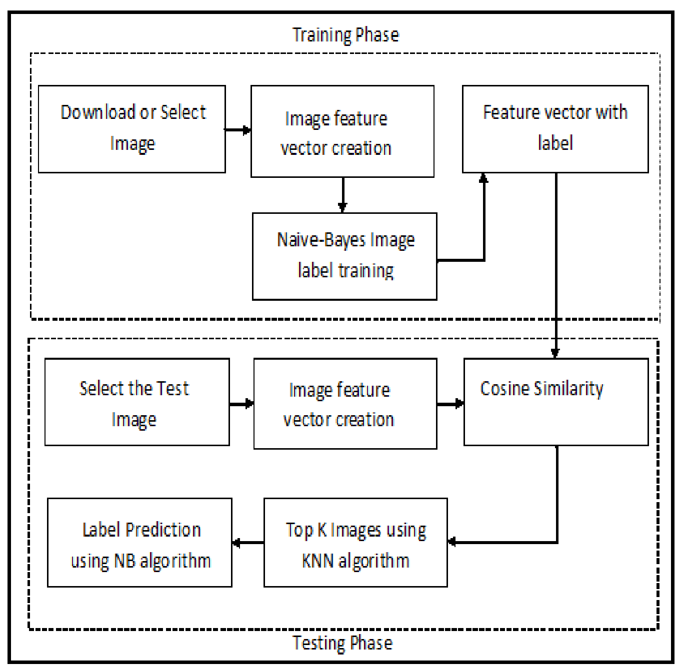

Fig.1. Block diagram of system architecture.

IV. IMPLEMENTATION DETAILS

A. Pre-Processing

- 1000 Image's feature vector creation and store these in Database,
- Create RGB feature vector of each image.
- Assign the label to each feature vector.
- Store that image with assigned label in Database.

B. Text to Image Search Mode

- Input text query e.g. "find the sun and grass" are processed.
- Stop words e.g. find, the, and are removed.
- Resulted keywords are searched in database.
- Database contains labels of 100 image's feature vector.
- If labels are matched with the label of database image's feature vector.
- Matched Images are displayed.

C. Image to Image Search Mode

- Image is given as input query.
- Its RGB feature vector is created.
- Its feature vector is compared with database image's feature vector, by using cosine similarity formula.
- If similarity value is greater than Threshold, that image in the database is selected.
- Continue the process for all database images till the Threshold value is less than distance value.
- Filter out the output images of cosine similarity by classification using Naïve Bayes algorithm , which gives most relevant images.
- Feature vector of input image is stored in database with new label.
- Display the retrieved images.

D. Speech to Image Search Mode

- Speech is converted into text by Android Intent feature of mobile.
- Resulted keywords are then matched with the labels of database's images.
- Matched images are then displayed.

E. Multimedia Search Mode

- Image is uploaded on image web server.
- Features of input image is get extracted.
- Cosine similarity distance measure between input image and database image is performed.
- Prediction the image label is performed using ID3 algorithm.
- Rest API is used to link to youtube for video links and ID3 algorithm match the image labels with label of video links.
- Similar video links and web links are displayed.

V. ALGORITHMS AND EQUATIONS

A. Porter Stemmer Algorithm

In information retrieval, porter stemmer plays important role by stemming entered sentence. It groups the words into semantically similar sets.

Algorithm:

- Gets rid of plurals and –ed or –ing suffixex.
- Turns terminal y to I when there is another vowel in the stem
- Maps double suffixex to single ones: -izationational, etc. Deals with suffixes, -full, -ness etc.
- Takes off –ant, -ence, etc.
- Removes a final –e.

B. Naive Bayes Algorithm

The Bayesian Classification represents a supervised learning method as well as a statistical method for classification. It is based on the Bayesian theorem. It is particularly suited when the dimensionality of the inputs is high. Parameter estimation for naive Bayes models uses the method of maximum likelihood. It requires a small amount of training data to estimate the parameters. Nave Bayes is a simple density estimation method for construction of a classification method. Based on prior knowledge, a Bayesian classifier classifies patterns to class C, to which it in all probability belongs. Knowledge of K, prior probability of each class and class conditional probability density functions, where X is a feature vector should be available initially to compute the likelihood. Then, Bayes theorem calculates the probability that X belongs to class C given by the following Equation.

$$P(C/X) = \frac{P(C)\ P(X/C)}{P(X)} \quad \textbf{(1)}$$

Where,
P(C) : Prior probability of Class C having given image.
P(X) : Prior probability of training data X
P(C/X) : Probability of C given X
P(X/C) : Probability of X given C

Algorithm:

Input: Image, Class C, Trained dataset

Output: Images which belongs to Class C

Steps:

- Extract the features of input image.
- Find out the probability P(C) of Class C that may contain the features of input image, called as prior probability.
- Find out the probability P(X/Ci), probability of occurrence of input image in given Class C, (Likelihood).
- Find out the probability P(X) , probability of occurrence of input image among all classes. (Evidance).
- Find out the probability P(Ci/X), Probability of Class Ci that contain given input image X, is the possibility of that X can be labeled Ci.
- Repeat the step 2 to 5 for all Classes.
- Assign the label of class to Input image, who has the maximum posterior probability among all classes.

Here, the posterior probability is calculated for all classes, and the class with the highest probability will be the instances label.

C. ID3 Algorithm

This algorithm is used to predict the label of input image and classify the results such as youtube links and web links according to the label of input image. ID3 builds a decision tree from a fixed set of examples. The resulting tree is used to classify future samples. The example has several attributes and belongs to a class (like yes or no). The leaf nodes of the decision tree contain the class name whereas a non-leaf node is a decision node. The decision node is an attribute test with each branch (to another decision tree) being a possible value of the attribute. ID3 uses information gain to help it decide which attribute goes into a decision node as shown in Fig.2.

Fig.2. The user interface of proposed system deployed on android device. The left is the operation interface and the right is the search result page.

Algorithm:

Input: Image, Class C, Trained dataset

Output: Image label, web links, youtube links that are matching with image label.

Steps:

- Establish Classification Attribute Training Set using Image Feature vector and Label.
- Compute Classification Entropy of Image Label
- For each attribute in R, calculate Information Gain using classification attribute.
- Select Attribute with the highest gain to be the next Node in the tree (starting from the Root node).
- Remove Node Attribute, creating reduced table tree.
- Repeat steps 3-5 until all attributes have been used, or the same classification value remains for all rows in the reduced table.

VI. RESULT ANALYSIS

Using metrics we can evaluate the effectiveness of the given system forming methods, and can compare the different applied methods. Thus, we can determine how the system can work better. The different category of images e.g. sky, animal, grass, moon, apple, sun are tested using the proposed system, calculated the precision and recall value for image of each category every time, and then obtained their average value. Which is showing accuracy and performance of proposed system as shown in Figs.3 to 6. L denotes length of retrieval list, from which R pieces matter as relevant results and E denotes the number of expected relevant hits. If we know this information, the following metrics can be calculated.

$$\text{recall} = \frac{\text{relevent hits (R)}}{\text{expected hits (L)}} \qquad (2)$$

Where the precision gives information about the relative effectiveness of the system.

$$\text{precision} = \frac{\text{relevent hits (R)}}{\text{All hits (L)}} \qquad (3)$$

Where the recall gives information about the absolute accuracy of the system. The number of all and expected hits is determined in each case of testing methods.

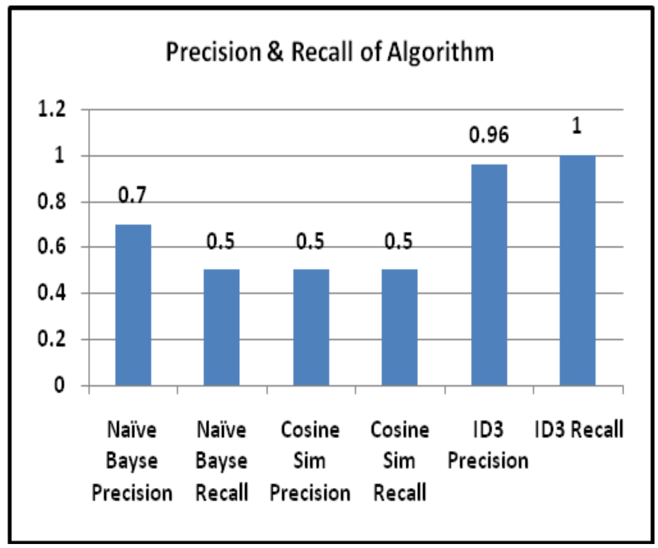

Fig.3. Precision and recall value for different algorithm.

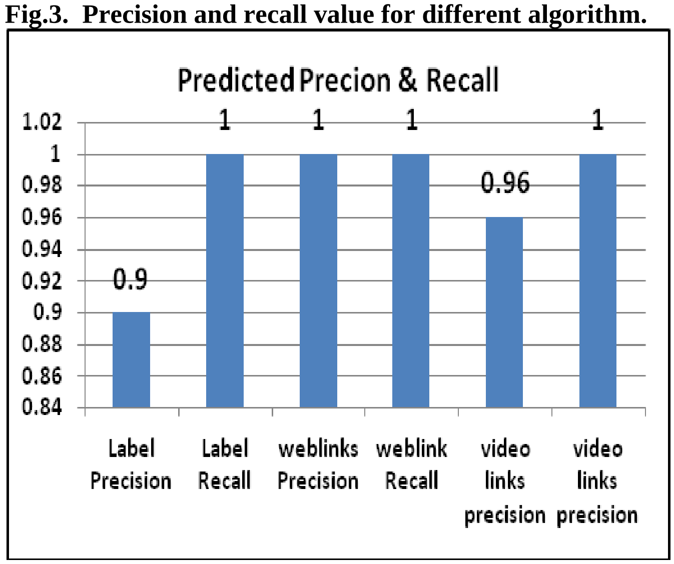

Fig.4. Predicted precision value and predicted recall value.

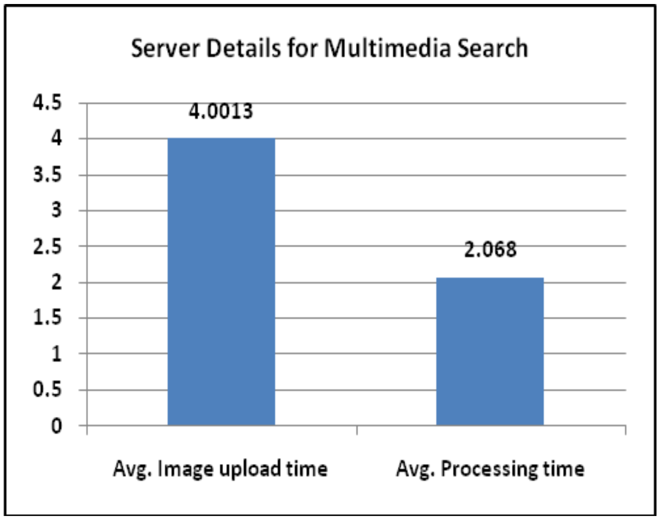

Fig.5. Process response time in Ms for Multimedia Image Search.

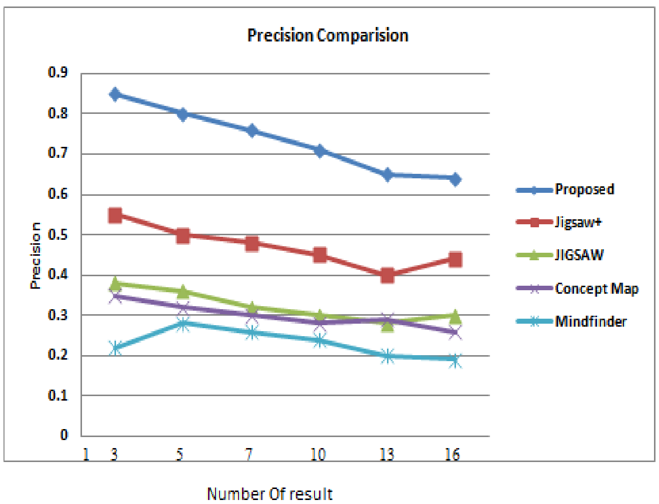

Fig.6. Comparison of Proposed & Existing systems.

VII. CONCLUSION

The proposed interactive multitudinous visual image search system on mobile device will helps the users to express their needed information implicitly and explicitly. The visual query generated by the user can be effectively used to retrieve similar images by the proposed method. Proposed system gives a better way to express their visual intent than other existing systems; it provides more relevant search results, especially in case where users can have a partial picture description in mind. Thus, the user's search experience on mobile device is significantly improved by this image search system. Future works, can focus on content based image search application.

VIII. ACKNOWLEDGMENT

We are very much thankful to all authors; those are mentioned in the references. Also to Prof. S. R. Durugkar, Head of Department, Computer Engineering, S.N.D. College of Engineering & Research Center, Yeola and Prof. V. R. Sonawane , PG coordinator of computer department, S.N.D. College of Engineering & Research Center, Yeola and all the respected peoples who helped us for designing and development of our work.

IX. REFERENCES

[1]Houqiang Li, Yang Wang, Tao Mei, Jingdong Wan, Shipeng Li, ”Interactive Multimodal Visual Search on Mobile Device”, IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 15, NO. 3, APRIL 2013.

[2]M. Porter, ”An algorithm for suffix stripping,” Program: Electron. Library Inf. Syst., vol. 14, no. 3, pp. 130-137, 1993.

[3]J. Shi and J. Malik, ”Normalized cuts and image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888-905, Aug. 2000.

[4]C. Rother, V. Kolmogorov, and A. Blake, ”GrabCut: Interactive foreground extraction using iterated graph cuts,” ACM Trans. Graph., vol. 23 no. 3, pp. 309-314, Aug. 2004.

[5]F. Jing, M. Li, H.-J. Zhang, and B. Zhang, ”An efficient and effective region-based image retrieval framework,” IEEE Trans. Image Process.,vol. 13, no. 5, pp. 699-709, 2004.

[6]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.

[7]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.

[8]M. Gales and S. Young, ”The application of hidden Markov models in speech recognition,” Foundations and Trends in Signal Process., vol. 1, no. 3, pp. 195-304, 2008.

[9]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.

[10]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.

[11]H. Xu, J. Wang, X. Hua, and S. Li, ”Image search by concept map,” in Proc. ACM SIGIR, 2010, pp. 275-282.

[12]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.

[13]Y.Wang, T. Mei, J. Wang,H. Li, and S. Li, ”JIGSAW: Interactive mobile visual search with multimodal queries,” in Proc. ACMMultimedia,2011, pp. 73-82.

[14]Anushma C R,” Multimodal Image Search System on Mobile Device”, International Journal of Computer Organization Trends - Volume 7 Number 1 - Apr 2014 ISSN: 2249-2593.

[15]Siri. [Online]. Available: http: //www.apple.com /iphone/ features/ siri.html.

[16]TinEye. [Online]. Available: http://www.tineye.com/.

[17]GoogleGoggles[Online]Available:http://www.google.com/ mobile/goggles/.
[18]NOKIA Point and Find [Online]Available: http://pointandfind. nokia.com/.
[19]SnapTell. [Online]. Available: http://www.snaptell.com/.
[20]GazoPa. [Online]. Available: http://www.gazopa.com/.

Author's Profile:

Roshani E. Bagul, Student, Department of Computer Engg, S.N.D.C.O.E.R.C. Yeola, Nasik, Savitribai Phule Pune University, Maharashtra, India.
Email: roshanibgl1@gmail.com.

P. P. Rokade, Professor, Department of Information Technology, S.N.D.C.O.E.R.C. Yeola, Nasik, Savitribai Phule Pune University, Maharashtra, India.
Email: prakashrokade2005@gmail.com.