ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Web Image Re-Ranking Using Query Specific Semantic Signature

Vishakha Khair[1], Nerika Kodare[2], Rutuja Mali[3], Mamta Parab[4], Snehal Mangale[5]

Student, Department of IT, RMCET, Devrukh, India[1234]

Lecturer, Department of IT, RMCET, Devrukh, India[5]

ABSTRACT: The Image search has become an important feature of multimedia now a days. Some Image search results are satisfactory and some are unsatisfactory. Image search engine mostly use keyword and they rely on surrounding for searching images. By specific keyword as a query it is difficult for search engine to understand the intention of the user. Another major challenge is to get relevant result without online training on low level features.In proposed system, we review the framework in which, the pool of images are retrieved based on given query keyword. At offline stage framework of image re-ranking automatically learns different semantic space for different query keywords using keyword expansion. Visual and textual features of the images are then projected into their related semantic spaces, for getting the semantic signature of images. At online stage, semantic signatures of images which are obtained from the query keyword and visual similarities are compared with the query images are re-ranked. This concept improves the efficiency and accuracy of the re-ranking process.

KEYWORDS: Image search, Semantic space, Semantic signature, Image re-ranking, Keyword expansion.

I. INTRODUCTION

Web image re-ranking is an effective way to improve the accuracy and efficiency. To describe the accurately query image by using query key word is difficult for user. Hence image search should be based on the text format from the ambiguity of query key word.Example like “APPLE” is a query keyword, then different categories of images are retrieved, which may be apple logo, apple fruits ,apple laptop etc. are displayed. Hence result should be search as per user interaction.

The disadvantages of text-based image retrieval are: by using key word the important features of image cannot be described completely. Thisdisadvantage is overcome by using CBIR technique. It is more efficient technique for retrieved the images. CBIR technique is helps to classify the digital image retrieved according to their visual content from the large database by user interaction. The use of CBIR technique is find similarity in images like shape, color andtexture etc. to find the closeness of the different region of image. Also using this technique the low-level image features can be extracted from the image. In this by using query key word firstly the pool of the images are retrieved, then according to user intention he/she selects the image from pool of images. Also remaining images are ranked based on their visual similarities.

II. LITERATURE SURVEY

A Ramchandran,Msaikumar,Dr.C.Nalini (3,March 2015) discovered Image re-Ranking, is an effective way to improve the result of web-based image search. Also at offline stage framework of image re-ranking automatically learns different semantic space for different query keywords using keyword expansion. Visual and textual features of the images are then projected into there related semantic spaces, for getting the semantic signature of images. And at online stage, semantic signatures of images which are obtained from the query keyword and visual similarities are compared with the query images are re-ranked. This concept improves the efficiency and accuracy of the re-ranking process.
Xiaogangwang,ShiQiu,keLiu,andxiaoou tang(VOL.36,APRIL 2014) discovered Query Specific Semantic Signature using CBIR Technique extract visual features or low-level features/match the visual similarities. CBIR technique is

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

helps to classify the digital image retrieved according to their visual content from the large database by user interactionand improved the accuracy and efficiency of image which is stored in database.

III. PROBLEM DEFINITION

To re-rank the images by comparing their semantic signature obtained from the semantic spaces ,to improve the accuracy and efficiency of image retrieval. Content based image retrieval uses visual features such as color, shape, image height, image width, intensity of color to represent and rank the images. It also retrieves the images by automatic segmentation technique.

IV. EXISTING SYSTEM

WEB image search engines frequently use keywords as queries and depend on surrounding text to search the particular images. They suffer from the irrelevant results of query keywords, because it is difficult for user to accurately describe the visual features of images in database only using query keywords. For example, using "Sonata" as a query keyword, that retrieves the images belonging to different categories, such as "sonata car", "sonata watch" etc.The limitations of existing system are: 1) Some popular visual features are in high dimensions and efficiency is not satisfactory if they are directly matched. 2) Another major challenge is that, without online training the similarities of low-level visual features could not wellmatchwithimageswhichhavehigh-level semantic space or meanings which interpret users search intention.

V. PROPOSED SYSTEM

5.1 Flow Chart:

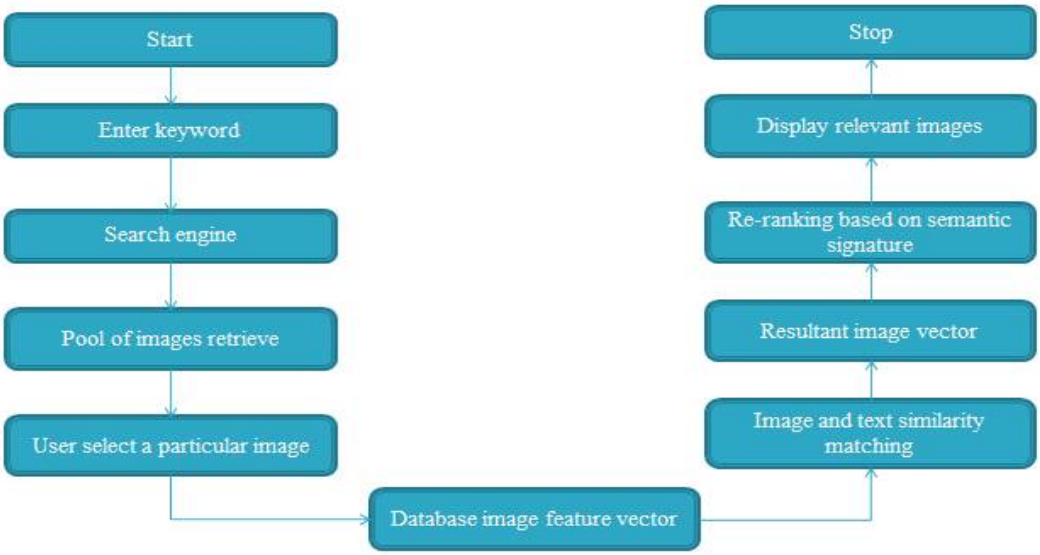

Fig1. Flow Chart

At the starting stage, user enters the query keyword in search boxas per his/her intention. Depending on that query keyword the search engine retrieve the result in the form of pool of images. From that pool of images user will select particular image which is query image. This query image and the images in the database are matched on the basis of their features (i.e. semantic signature).The Re-ranking is done by number of clicks on particular image. The images

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

which get matched with query image selected by the user only that images are retrieved and user will get the final result according to the working of CBIR.

Methodology:

1. Keyword Expansion-

When a user enters the query keyword for searching a particular image then pool of images is retrieve which are relevant to user query keyword. But the query enter by the user is not sufficient or that query not fulfills the users intention because of us lack of knowledge about giving the specific query for search image. Because of this search performance is also decreased. To overcome this keyword Expansion methodology is used.The"keyword expansion" means attaching new keyword to query keyword i.e. keyword automatically discovered. In this the extra information of the query keyword is need to be capture using Keyword expansion. Some time user not clear about his/her query. By using this methodology search engine can predict the user intention. For example when user wants some particular product of Samsung but user is confused about it. Then user enter Samsung query keyword and by methodology Samsung keyword get expand and give result of Samsung mobile, Samsung laptop, Samsung camera etc.

2**.** Semantic Signature:

There is r reference class, user keyword is K and their training images. Trained images are obtain by applying multi-class classifier on the basis of visual features of images. P indicating probabilities of new image N, which is belonging to different reference classes i.e. semantic signature of N. The distance between two image x & y i.e. measured as the M1. The distance between semantic signature of their image Px and Py.

$$\text{Distance } (P^x, P^y) = \| P^x - P^y \|$$

3. One Click Feedback:

Online re-ranking of image which limits user's effort to just one-click feedback, which is an effective way to improve search results of users query and the interaction between the user and web is very easy. Major web image search engines have used this strategy. The query keyword is given by the user. After this a pool of images which are relevant are fetched. When the user clicks on particular image from the pool of images, the count of that image will be increased by one and remaining re-ranked of image based on the count of each image. The highest count of image will be displayed first and that it may match with the user's search intention.

5.2 ***Algorithm:***

Content Based Image Retrieval:-

CBIR is the process or the technique which is used for retrieving the images from a large database or a digital library according to the visual content of the images. In other words , it extracts or retrieve the visual attributes or low level features like color texture or shapes. It is identifying the images by extracting the syntactical features and indexes accordingly.

Steps involved in CBIR:-

-increasing the database by extracting and storing the features for images I_1 to I_N and,

-the features are extracted by processing the query image Q which is specified by the user and performing the similarity comparison for retrieval of images from the database.

Color Features-

Color is one of the most important feature used in CBIR products. The popularity of this attribute is to easily implement and differentiate the colors. Color can be represented by different models such as HIS,YIQ,CMY,RGB.

Shape Features-

The common low-level feature used to describe the geometric characteristics of image is shape. It is an important and basic visual feature used to describe image content. However, the shape representation and description is a difficult task.The shape parameters used are Mass, Center of gravity(Centroid), Mean ,Variance, Dispersion, Axis of least inertia, etc. Some of these are described as follows,

Mass is the no. of pixels contained in one class ,.

mass= $\sum_{xy} \boldsymbol{h}(\boldsymbol{x}, \boldsymbol{y})$

where

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

$$h=\begin{Bmatrix} 1 & if \;\; s(x,y) \in c \\ 0 & if \;\; s(x,y) \notin c \end{Bmatrix}$$

Centroid is also called as the center of mass; h is a mask of cluster c over image s(x,y). The co-ordinates (x_c ,y_c) of the centroid are defined as:

$$x_c = \frac{\sum_{xy} x*h(x,y)}{mass}$$

and

$$y_c = \frac{\sum_{xy} x*h(x,y)}{mass}$$

The mean and variance features of the class c are computed over the original image I considers the resulting segmentation s of the images, & they are respectively denoted by μ_c

$$\mu_c = \frac{\sum_{xy} x*h(x,y)}{mass}$$

and

$$\sigma_c^2 = \frac{\sum_{xy} x*h(x,y)}{mass}$$

Dispersion is the sum of the distances of each region of a class from the class centroid. The distance is calculated by Euclidean distance formula. The dispersion can be given as

$$Disp=\sum_i dist(O_c , O_{i,c})$$

Where, $\mathbf{dist(O_c,O_{i,c})}$ is the Euclidean distance

$\mathbf{O_c}$=centroid of the class c

$\mathbf{O_{i,c}}$=centroid of region I of class c

Color Retrieval-

Color retrieval system works on two stages.

i. In the first stage, Histogram based comparison is done and matching images are shortlisted.
ii. In the second stage, the Color Coherence Vector of the short listed images (stage 1) are used to refine the results.

The Euclidean distance is calculated between two matching histogram h and h'-

$$d=\sqrt{\sum_{i=1}^{n}(hi - h'i)^2}$$

Shape Retrieval-

The Proposed shape retrieval system based on the automatic segmentation process to get approximate information about the shape of an object.It begins by segmenting the image into 5 classes depending on their brightness.

Similarity Measure-

In this algorithm we have proposed that the matching of the images is done on color by color basis. By analyzing the particular histograms, we first calculate the number of color in both the query image and database image . Then both images are matched by observing if the proportions of a particular color in both the images are comparable and matching. Image retrieval result is not just a single image but a list of images ranked by their similarities since, CBIR is not based on only exact matching.

$$\text{Recall}=\frac{relevant\ retrieved\ images}{all\ relevant\ images\ retrieved}$$

$$\text{Precision}=\frac{relevant\ retrieved\ images}{all\ retrieved\ images}$$

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

VI. IMPLEMENTATION AND RESULT

The proposed system has following stages:

- User enter a query keyword for searching purpose.
- Images displayed according to query keyword enter by user.
- Ranking of images according to number of click by user.
- Feature Extraction and Similarity matching of images using CBIR.
- Images are ranked and displayed as per the matching of features.

GUI of proposed system:

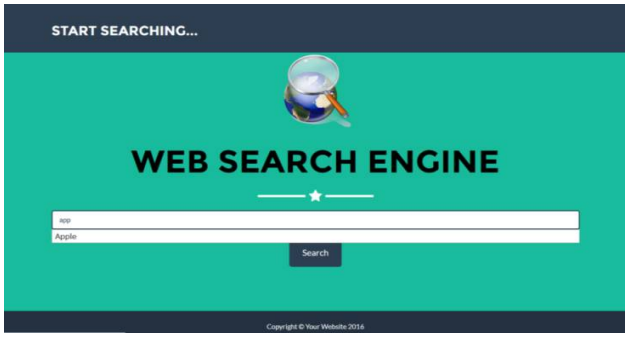

Fig1. GUI of proposed system

Images Displayed according to the query Keyword:

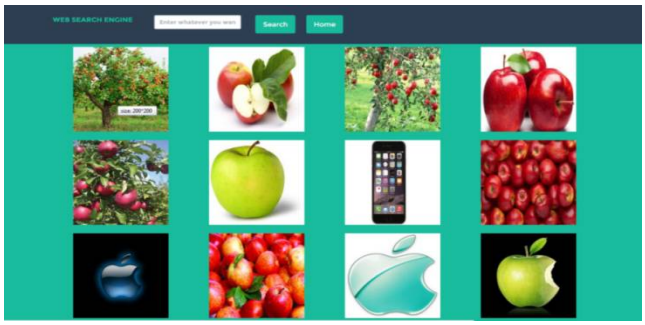

Fig2. Images Displayed according to the query Keyword

Feature Extraction and Similarity matching using CBIR

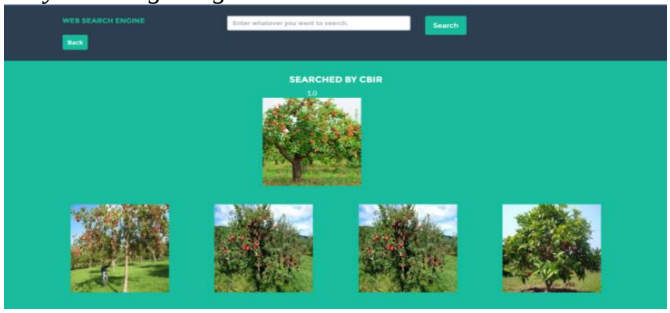

Fig4. Result of Feature Extraction and Similarity matching using CBIR for keyword “Apple”

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Comparison between Proposed and Existing System:

Sr.no	Parameters	Existing System	Proposed system
1.	Approach	Online	Offline
2.	Feature extraction through	Textual information	Click feature of image
3.	Input for search	Only Keyword	Keyword with query image
4.	Accuracy	Moderate	More
5.	Relevance	Less	More
6.	Search time	More	Less

Fig5. Proposed v/s Existing System

VII. ADVANTAGES

1. Using semantic signature we can improve the accuracy and efficiency of web search image.
2. The visual features of images are deliberate into their related semantic spaces automatically and learned through keyword expansions offline.
3. Our experiments show that the semantic space of a query keyword can be described by reference classes. Therefore the image re-ranking using semantic signature becomes extremely efficient.

VIII. CONCLUSION

This image re-ranking system is proposed for generating relevant images over an internet through single click only. This will reduce manual work by integrating visual and textual similarities of images and provide more efficient image retrieval. The purpose of this paper was to improve the accuracy(precision and recall)of CBIR application by retrieving more images related to query image.

ACKNOWLEDGEMENT

It is an opportunity of immense pleasure for us to present the paper “Web Image Re-Ranking Using Query Specific Semantic Signature” expressing our heart left gratitude to all those who have generously offered their valuable suggestion towards the completion of paper. The credit goes to our Prof. Mangale S.R. (RMCET, Ambav, Ratnagiri) whose positive attitude; moral support and encouragement lead to the success of the paper

REFERENCES

[1] A Ramchandran, M sai Kumar, Dr. C. Nalini, ”Learning Image Re-Rank: Query-Dependent Image Re-Ranking Using Semantic Signature”, International Journal Of Innovative Research in Science, Engineering and Technology, Vol.4, Issue 3, March 2015.
[2] Xiaogang Wang, Shi Qiu, Ke Liu, and Xiaoou Tang, ”Web Image Re-Ranking Using Query-Specific Semantic Signature”,IEEE transaction on pattern analysis and machine intelligence, Vol.36, no.4,April 2014.

BIOGRAPHY

Prof. Mangale S.R. I have completedBachelors in Computer Engineering (BE) fromRMCET college, Mumbai university andcurrently pursuing ME in computers from SKNSITS, Lonavala.My research interestsare Data mining, Database technologies, Software Project Management, Software Testing and Software Engineering.

ISSN(Online): 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 3, March 2016

Ms.VishakhaKhair. I am pursuing my B.E. in Information Technology, RMCET, Mumbai University. I am member of ISTE. My area of interest is Programming and Database.

Ms.Rutuja Mali. I am pursuing my B.E. in Information Technology, RMCET, Mumbai University. I am member of ISTE. My area of interest is Database and Testing.

Ms.NerikaKodare. I am pursuing my B.E. in Information Technology, RMCET, Mumbai University. I am a member of ISTE. My area of interest is database and Testing.

Ms.MamtaParab. I am pursuing my B.E. in Information Technology, RMCET, Mumbai University. I am a member of ISTE. My area of interest is Database and Testing.