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. 3, Issue 10, October 2015

Faster Mobile Image Retrieval Using Android Smartphones

Bhagyashri Parate[1], Bhagyashree Tingare[2], Parvathi Saxena[3], Pooja Patil[4], AshwiniBavage[5], Shanti Guru[6]

B.E. Students, Dept. of CSE, D.Y Patil College of Engineering, Akurdi, Pune, India[1,2,3,4]

Asst. Professor, Dept. of CSE, D.Y Patil College of Engineering, Akurdi, Pune, India [5,6]

ABSTRACT*:* Image retrieval is a growing field in the image processing domain and it has covered so much more areas including the smartphones. As the days pass by great technologies are introduced every day, smartphones have become more and more important in our lives. So, many tasks can be done using a smartphone. It is almost like carrying your personal desktop with you. Images can be taken and shared instantly, messages can be recorded and sent overseas, office work can be delayed and saved and so much more can be done with smartphones. All these can be done quickly as it's just a click away. In this paper, an approach for image retrieval is discussedusing an Android smartphone by creating an application to enable faster retrieval of images on the phone. The approach proposes a method to retrieve images based on the query image and how relevant its features are to improve the retrieval process. Moreover this new approach has been discovered to reduce the amount of time for retrieving images in less time i.e. by ≤ 5% which is better than the existing systems. We show that our approach can also be integrated to improve its retrieval accuracy.

KEYWORDS: Mobile image retrieval, salient visual words, scalable image retrieval, multiple queries

I.INTRODUCTION

Statistics show thatin the worldabout 91% people use PC/Laptop, 80% smartphone, 47% tablet and 37% games console users for the year 2015. Smartphones are indispensable they are a necessity today. Everydayactivities depend on it. They use them for clicking pictures, GPS tracking, surfing internet, playing games and so much more. It is like everything and anything can be done using it. People are sharing their day to day details online for their friends to view and comment. This is making image retrieval a major concern on the user end's side i.e. the smartphone end. Image retrieval onsmartphones isn't a tedious task but it is a time consuming task. Many factors such has bandwidth, sound disturbances and so on are to be looked out for. Nobody realizes it until it takes too long to show results probablybecause no one likes to wait. The results are always relevant to the query image passed. All most all thetime the results are somewhat close to the query image.

For image retrieval on smartphones is purely wireless based retrieval process. But poor conditions of the wireless channel also affect the retrieval process. Other concerns also include limited bandwidth andinstability of channel need to be looked out for. Normally, mobile image retrieval isbased on text. Some search engines, like Google, which handles 3 million searches in a day can providethousands of relevant images successfully when a userinputs a textual query. But text based image retrieval dependson the image tags.However, several images may not have tags, and thetext description is not always truthful. This can cause the retrieved images to be less relevant. Smartphone cameras adopt content based image retrieval. Most of its approaches focus of compressing the BoW histogram[1]–[2]. But the BoW has many deficiencies like it can carry quantization loss which may result polysemy phenomenon. It also neglects the spatial relationship among visual words. It doesn't have descriptive power and it cannot describe the local region in images exactly.

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. 3, Issue 10, October 2015

The existing retrieval systems usually require a single queryimage [3]–[2]. But these systems rarelyachieve good performance i.e. if an object in the query cannotbe seen clearly or there are too many localinterest points (such as SIFT feature points) foundon that image. If a large part of those interest points are noise or irrelevant to the object in the image can increase thecomputational complexity. And unstable noisy local featureshave indeed negative effect on the retrieval performance. Queryexpansion (QE) [4] is another approach that will update the original query by combiningit with retrieval results to expand the query and weaken thedisturbance of noise. The goal of QE is to explore mostrelevant images iteratively to eliminate the deficiency of singleimage based retrieval.

Normally we click pictures at a new place we tend to click more than one picture. We upload an image that is relevant to our naked eyes in the sense that, that particular image relevant to our eyes is based on the clarity and its resolution. But this isn't always the based case for retrieval process. The proposed work is based on an approach to retrieve images based on the query image. It possesses some salient features [15] which will be considered for the retrieval process. Our approach consists of following 2 steps, 1) Mining relevant photos and exploring contextual saliency. On the user's smartphone i.e. the user's end may contain more than one image for the query image being passed for retrieval process. Images will be mined based on their salient features. 2) Ranking salient features. Those relevant images are ranked based on the extracted salient features.

This paper is organized as follows, Section 2 reviews the related work. Section 3 depicts the overview of the proposed system, whereas Section 4 mentions how we can mine relevant images based on their contextual information and Section 5 tells how we will sort those images based on those features. Section 6 illustrates the architecture of the proposed work and conclusions are drawn in Section 7.

II. RELATED WORK

In the recent years there has been a rapid development seen in the field of image retrieval due to BoW representation [5] and local features like SIFT [6] and so on.

The idea of hierarchical vocabularytree [4] accelerated the speed of clustering and quantizingof large scale retrieval of images. But it suffers from many factors affecting its efficiency. Many works madeup remedies forthese defects, such as introducingspatial verification [5], [13], using multiple queries [7],[8]–[9], [10], [11], [12] and compact descriptors [16]-[17].

A. *Spatial Verification*

Query expansion (QE) is one of the standard methods for improving performance in text retrieval applications [9]. A number of the highly ranked documents from the original query are re-issued as a new query. In this way, additional relevant terms can be added to the query.QE methods focus on enriching the query model by adding spatially verified features. Retrieval with the "expanded" query follows. It has been observed that if the shortlist has enough true positives, the spatial verification re-ranking almost always correctly identifies relevant images, and, consequently, results for the expanded query are significantly better than the original single image query.

Zhang [14] *et al,* extracted local feature group from images,and measure the spatial contextual similarity between groupsto find the best matched order by which it is used to calculate thegroup distance for NDIR and topic based image re-ranking.The approach of visual synonymsstill does not use geometry relation inretrieval stage. But this geometry relation is a must to restrict the spatial consistency in images but this could mislead the matched visual phrase separately [5].

B. *Using Multiple queries*

Distractors are the regions of an image that draw attention away from the main subjects and reduce the overall image quality. Learning synonyms and hierarchical visual word path provide multiple choices for feature matching. Among the massive images available how to find images satisfying user's interest becomes more and more necessary. By providing multiple images as query would probably reduce the distractors, reduce quantization

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. 3, Issue 10, October 2015

loss and will help in learning visual synonyms. In order to reduce quantization loss different image patches are mapped to the same visual word and for visual synonyms two visual words different in descriptor space may correspond to the same object in real word. Visual synonyms as a set of hierarchical word paths correspond to the same real world object are extracted as pairs of independent visual words. Hierarchical k-means clustering is used to build vocabulary tree. Hierarchical quantization maps each descriptor to a set of words to reduce quantization loss. Spatial layout of visual words using geometry constraints in visual synonyms learning is used to verify visual word paths. For finding visual word paths which represent the same real word object visual synonyms are detected and visual synonyms expansion is used to estimate the probability weight.

C. *Compact Descriptors*

Scale-invariant feature transform (or SIFT) is an algorithm describes and detects local features in images that can help in object recognition. They use particular points to detect their features which are invariant to scaling and rotation. They can be easily matched to a larger database of local features. The problems like key localization, scale and rotation can be solved by DoG which helps in increasing accuracy and stability. The problem of geometric distortion can be eliminated by using blurring which in return increases the stability.

III. SYSTEM OVERVIEW

As shown in Fig. 1, the proposed mobile imageretrieval approach is based on mobile-server architecture. The mobile side of the architecture will represent the user’s end and the server side represents the database admin because it will be responsible for showing results for the image query being passed. In simple words the server is responsible for the image repository. Generally a user will upload an image and retrieve based on their query image itself. But in the proposed work the image will be chosen from the mobile album i.e. user’s gallery irrespective of what image query is passed. Firstly similar images to the query image will be mined and their salient features will be extracted. And secondly those extracted salient features will be ranked based on how relevant its content is to the bandwidth used for communicating to the server end.

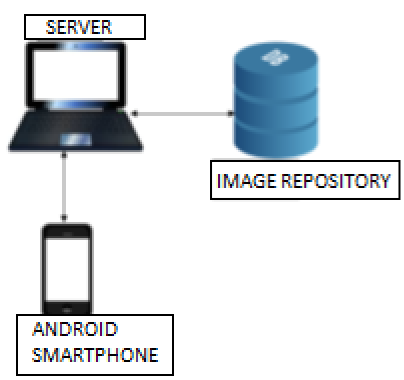

Fig. 1: System Architecture

IV. MINING RELEVANT PHOTOS AND EXPLORING CONTEXTUAL SALIENCY

There is a possibility that there exists several relevant photos to the query image in the user’s gallery. Our aim is find contextual saliency of those images to carry out the retrieval process. Mining relevant photos will help us to reduce the number of the transmitted features which will help to improve the image retrieval process faster and efficiently.It consists of the following two steps:1) feature extraction and quantization; and 2) multiple relevantphotos mining.

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. 3, Issue 10, October 2015

A. *Feature extraction and Quantization*

By difference of Gaussian (DoG) detector one can find points $< x, y, scale, orientation >$ and then describe these points using 128-dimension SIFT descriptor. By performing hierarchical clustering on the SIFT descriptors hierarchical vocabulary tree can be generated. This tree is used to quantize local descriptors and save its features in the vocabulary tree from root to leaf so that images can be represented with a set of visual words paths

B. *Multiple relevant photo mining*

Mining relevant images is find similar contextual information in those images for faster retrieval process. But contextual information could include information about the location, time and visual information. But geographical information will not be considered because on the mobile end the GPS facility may not be enabled while those images where clicked. So much factors are to be looked out for.

In this paper we propose to achieve this mining process using the following steps i.e. 1) searching relevant/similar images from the user gallery and 2) eliminating noisy images.

1) Searching relevant/similar images from the user gallery: Using an histogram generation approach using BOW histogram [15] we will measure their similarity of the query image to the images present on the mobile end. This will be done creating a block distance $D(k)$ where,

$$D(k) = \exp(-\left|\left|h_q - h_m(k)\right|\right|_1) \qquad (1)$$

2) Eliminating noisy images: Even after finding relevant images they can have some amount of noise in them. This noise could disintegrate the performance. This is made possible by establishing a threshold value, say I_{th} which will help to eliminate them based on how they score. If an image scores too high or too small compared to I_{th} we will eliminate them.

V. RANKING SALIENT FEATURES

Yang [15] *et al.* mined relevant images based on an approach to find Invariant Salient Points (ISP) from images and prominent ISP were called Salient Visual Word (SVW). Then neighboring SVWs were paired together to form Salient Visual Pairs (SVP). Because multiple images are used for checking, there is a possibility that two or more images can have same SVWs. Those SVWs are merged to reduce the transmitted data and to measure the stability of the merged SVPs.

Ranking of these relevant images will be based on their contextual informed that was used in the mining process. This can also be done following Yang [15] *et al.* which can either follow any two ways, either based on frequency of occurrence or based on stability of those images.

VI. SYSTEM ARCHITECTURE

The proposed work is designed to avoid noise disturbances and retrieve most relevant images first. The system is designed to take the input image and perform Gaussian blur. Gaussian blur is the fastest among all the other blurs, except the box blur. But Gaussian is faster than doing the 2D convolution.

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. 3, Issue 10, October 2015

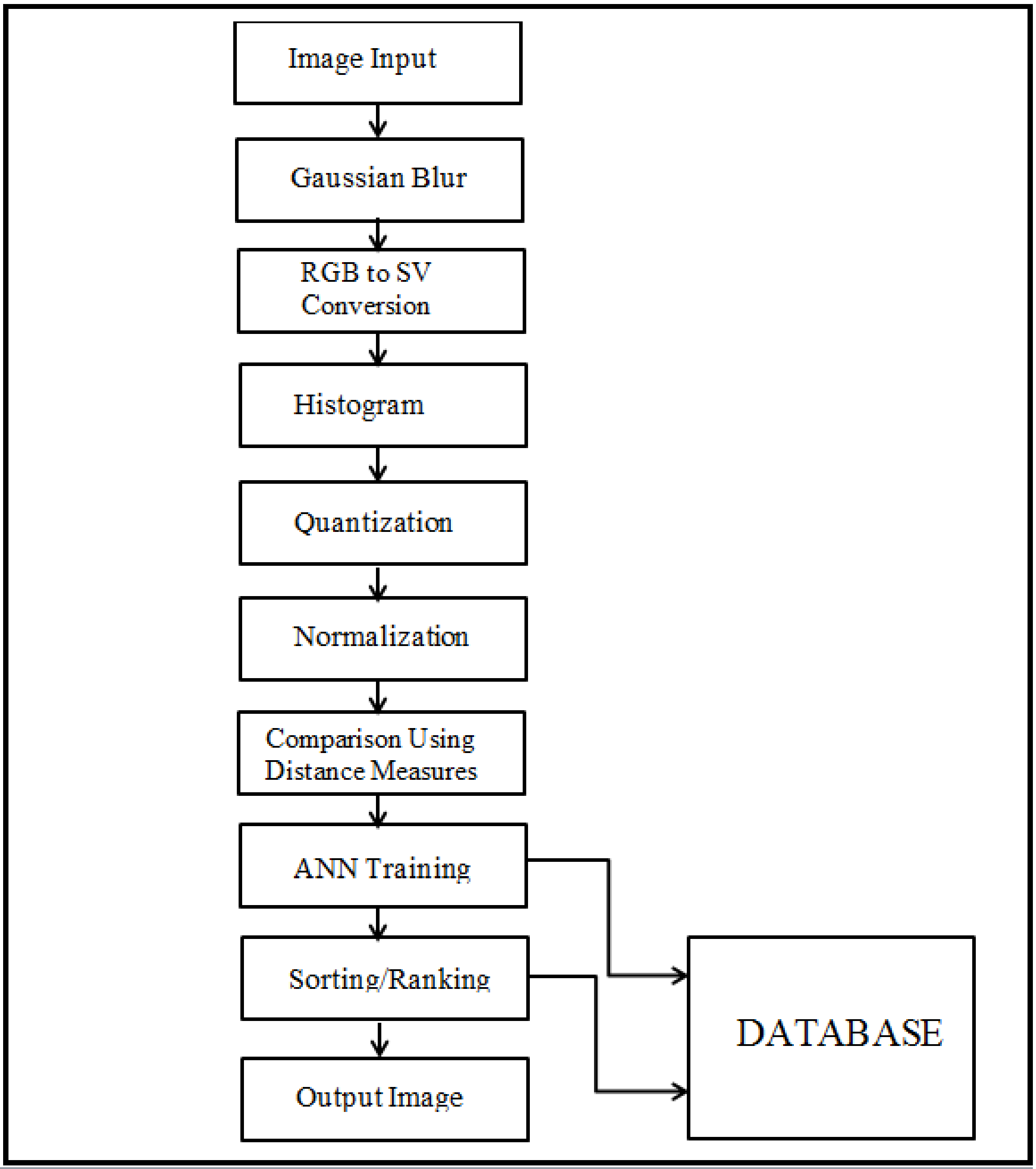

Figure: Flow of System

Many devices still use HSV image specifications. HSV colour model is preferred over the RGB colour model. This model describes colors similarly to how the human eye tends to perceive color. RGB defines color in terms of a combination of primary colors, whereas, HSV describes color using more familiar comparisons such as color, vibrancy and brightness. The

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. 3, Issue 10, October 2015

HSV generation will lead to cylindrical coordinates. With this information histogram will be generated. The histogram will give the tonal information i.e. lightness (also called tone). In the quantization stage, the size of the image will be compressed for easy transmission. Quantization technique is achieved by compressing a range of values into a single quantum value. For representing a digital image the colours of the image are reduced so that the images gets compressed. Quantization is a lossycompression technique, so it provides options for compression that could include low compression (high quality) or high compression (low quality). Some images may have poor contrast maybe because of glare. Quantization loss was a major concern in the retrieval process. So the proposed work would help in selecting the right image using their contextual saliency and improve the efficiency in retrieving the right image for a specified bandwidth. Normalization will deal with such images. In this process the pixel values are changed but it is done within some fixed range. Even after the pixels are maintained under a range their value may or may not increase. In any case we are looking for transferring

VII. CONCLUSION

In this paper, a novel SVP ranking algorithm is used to retrieve images in mobile at faster rate. The performance of this approach can be further improved by increasing the bandwidth at mobile as well as server side. This new approach has been discovered to reduce the amount of time for retrieving images in less time i.e. by ≤ 5% which is better than the existing systems. While reducing quantization loss, this method is more efficient to achieve our goal.

REFERENCES

[1] D. Chen, S. S. Tsai, V. Chandrasekhar, G. Takacs, J. Singh, andB. Girod, “Tree histogram coding for mobile image matching,” in *Proc.DCC*, 2009, pp. 143–152.
[2] R. Ji, L.-Y. Duan, J. Chen, H. Yao, Y. Rui, and W. Gao, “Locationdiscriminative vocabulary coding for mobile landmark search”,*Int.J. Comput. Vis*., vol. 96, no. 3, pp. 290–314, Feb. 2012.
[3] V. Chandrasekhar, G. Takacs, D. Chen, and S. Tsai, “CHoG: Compressedhistogram of gradients—A low bit-rate feature descriptor,” in *Proc. IEEEConf.* CVPR, Jun. 2009, pp. 2504–2511.
[4] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, “Total recall:Automatic query expansion with a generative feature model for objectretrieval”, in *Proc. IEEE 11th ICCV, Oct.* 2007, pp. 1–8.
[5] W. Zhou, Y. Lu, H. Li, Y. Song, and Q. Tian, “Spatial coding for largescale partial-duplicate web image search”, in *Proc. Int. Conf.* MM, 2010,pp. 511–520.
[6] D. G. Lowe, “Distinctive image features from scale-invariant keypoints”,Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.
[7] Y. Xue, X. Qian, and B. Zhang, “Mobile image retrieval using multiphotosas query”, in *Proc. IEEE ICMEW*, Jul. 2013, pp. 1–4.
[8] Y. Xue and X. Qian, “Visual summarization of landmarks via viewpointmodeling”, in *Proc. 19th IEEE ICIP*, Sep./Oct. 2012, pp. 2873–2876.
[9] O. Chum, A. Mikulík, M. Perdoch, and J. Matas, “Total recall II:Query expansion revisited”, in *Proc. IEEE Conf.* CVPR, Jun. 2011,pp. 889–896.
[10] B. Fernando and T. Tuytelaars, “Mining multiple queries for imageretrieval: On-the-fly learning of an object-specific mid-level representation”, in *Proc. IEEE ICCV*, Dec. 2013, pp. 2544–2551.
[11] X. Qian, Y. Xue, X. Yang, Y. Y. Tang, X. Hou, and T. Mei,“Landmark summarization with diverse viewpoints”,*IEEE Trans.*Circuits *Syst.* Video Technol., to be published. [Online]. Available:http://dx.doi.org/10.1109/TCSVT.2014.2369731
[12] X. Yang, X. Qian, and T. Mei, “Learning salient visual word for scalablemobile image retrieval”,*Pattern Recognit*., to be published. [Online].Available: http://dx.doi.org/10.1016/j.patcog.2014.12.017
[13] R. Ji, H. Yao, W. Liu, X. Sun, and Q. Tian, “Task-dependent visualcodebookcompression”, IEEE Trans. Image Process., vol. 21, no. 4,pp. 2282–2293, Apr. 2012.
[14] S. Zhang, Q. Huang, G. Hua, S. Jiang, W. Gao, and Q. Tian, “Buildingcontextual visual vocabulary for large-scale image applications”, in Proc.Int. Conf. MM, 2010, pp. 501–510.
[15] X. Yang, X. Qian and Y. Xue, “Scalable Mobile Image Retrieval by Exploring Contextual Saliency”, IEEE Trans. Image Process., vol. 24, no. 6, June 2015.
[16] R. Ji, L.-Y. Duan, J. Chen, and W. Gao, “Towards compact topicaldescriptors”, in *Proc. IEEE Conf. CVPR*, Jun. 2012, pp. 2925–2932.
[17] R. Ji, L. Duan, J. Chen, H. Yao, and W. Gao, “A lowbit rate vocabularycoding scheme for mobile landmark search”, in *Proc. IEEE ICASSP*,May 2011, pp. 2316–2319.