An Improvised Approach to Content based Image Retrieval

K.Lakshmi Sudha
Research Scholar, Sathyabama University
Assistant Professor, S.I.E.S. GST
Mumbai, India

Megha Redkar
Department of Information Technology S.I.E.S. GST
Mumbai, India

Anitha Ranganathan
Department of Information Technology S.I.E.S. GST Mumbai, India

Karishma Upadhyay
Department of Information Technology S.I.E.S. GST Mumbai, India

ABSTRACT
Image matching is a fundamental aspect of many problems in computer vision, solving for 3D structure from multiple images, stereo correspondence, and motion tracking. An image may have features that have properties making them suitable for matching images. There have been various algorithms and optimizations for Content Based Image Retrieval. A few algorithms include Simple Harris, SIFT. Feature detectors and high matching consuming creates a low automation problem.

To overcome these issues there have also been papers proposing optimized algorithms on Harris and SIFT [1]. This algorithm also has several flaws. The optimized algorithm uses Harris for feature extraction and description but Harris has a constraint that Harris detectors detect points only on black and white events. SIFT is flawed in itself since it is inefficient for poor resolution images and is also a time consuming algorithm [5]. The nearest neighbor search used as a matching algorithm is also time consuming and results in random overhead of outcomes. To overcome these shortcomings this paper proposes an algorithm that combines the advantages of Harris, SIFT and the matching algorithms. Color saliency is used along with Harris improvising its efficiency [6]. SIFT matching technique along with the nearest neighbor algorithm is supplemented with an epipolar concept to tender accurate results with lesser discrepant values.

General Terms
Feature vectors, feature descriptor, visual content descriptor.

Keywords
Scale invariant features, epipolar constraint, SIFT.

1. INTRODUCTION
Since the 1990s, images have been extracted from large image databases in accordance with the technique that uses visual contents following the user's interests. This area of research has been flourishing ever since. During the past decade, remarkable progress has been made in both theoretical research and system development. In spite of the tremendous research, the accuracy and efficiency of system has attracted a lot of researchers over time. Early work on image retrieval can be traced back to the late 1970s. A conference, in

1979 on Database Techniques for Pictorial Applications was held in Florence. Since then, the application potential of image database management techniques has attracted the attention of researchers.

Images were tagged with keywords and textual annotations earlier. In other words, images were first annotated with text and then searched using a text-based approach from traditional database management systems [8, 9].

Easy navigation and browsing was facilitated by organizing the text descriptions and images by topical or semantic hierarchies based on standard Boolean queries. However, since automatically generating descriptive texts for a wide spectrum of images is not feasible, most text-based image retrieval systems require manual annotation of images. Also, annotating images manually is a cumbersome and expensive task for large image databases, and is often subjective, context-sensitive and incomplete. As a result, the traditional text based methods were incompetent to support the variety of queries. In the early 1990s, new digital image sensor technologies, the volume of digital images produced by scientific, educational, medical, industrial, and other applications available to users increased dramatically. The difficulties faced by text-based retrieval became more and more severe. A more efficient and intuitive way to represent and index visual information was required would be based on properties that are inherent in the images themselves. The traditional system was eventually replaced by a modern method involving or feature descriptors that works with the visual content of the images. The working of an ideal modern methodology is given in Figure 1 [10].

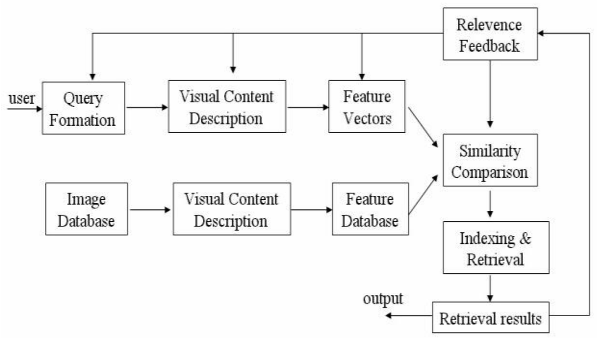

Figure 1: Flow diagram for an ideal CBIR system

2. RELATED WORK
Harris algorithm has an operator that uses image break and neighboring point eliminating method. The Harris detector is

used to extract feature points for the purpose of matching and the interesting values of detected corner points are recorded at the same time [1].Further, using the method of fragmental image processing, the images are processed. A histogram of Gaussian for 16*16 pixels for the region of interest is computed using Prewitt or Sobel masks followed by binning. The merging algorithm is used to reject neighboring points. Figure 2 shows the detected interest points of a cat image [6].

Figure 2: Interest points on an image using Harris detector

SIFT extracts distinctive invariant features from an image which can be used for performing reliable matching between different views of an object or scene [2, 13, and 14]. These features are scale and rotation invariant and have shown to provide robust matching. The SIFT has been demonstrated to be very suitable for object detection in images with high resolution. However, SIFT performs poorly when it is faced with images of poor resolution [1].

The conventional epipolar constraint is a powerful tool for matching keypoints (salient local features) between pairs of images, but in its standard form it treats the detected keypoints point-like entities without intrinsic scales [3,4]. In contrast, recent keypoint detectors typically associate a scale (for multiscale detectors) or a full affine frame (for affine-invariant detectors) to each detection.

Salient points are locations in an image where there is a significant variation with respect to a chosen image feature. Since the set of salient points in an image capture important local characteristics of that image, they can form the basis of a good image representation for content-based image retrieval (CBIR). Salient features are generally determined from the local differential structure of images. They focus on the shape saliency of the local neighborhood. Most of these detectors are luminance based which have the disadvantage that the distinctiveness of the local color information is completely ignored in determining salient image features. To fully exploit the possibilities of salient point detection in color images, color distinctiveness should be taken into account in addition to shape distinctiveness [7].

3. PROPOSED WORK

The paper proposes an optimization matching algorithm based on Harris and SIFT algorithm. The disadvantages of Harris are overcome by enhancing the overall working using color saliency. The color and texture information around these points of interest serve as the local descriptors of the image. In addition, the shape information is captured in terms of edge images computed using Gradient Vector Flow fields. Invariant moments are then used to record the shape features. The combination of the local color, texture and the global shape features provides a robust feature set for image retrieval.

The efficiency of salient point detection depends on the distinctiveness of the extracted salient points. Color is also considered to play an important role in attributing image saliency. In our approach, the color saliency is based on the work reported in [12]. To achieve the color saliency, the color axes are rotated followed by a rescaling of the axis and, the oriented ellipsoids are transformed into spheres. Thus, the vectors of equal saliency are transformed into vectors of equal length. The difference between a simple Harris and color saliency is demonstrated in Figure 5. The salient points are circled in yellow and blue for Harris detector and the proposed method, respectively. The Harris detector detects points based on black and white events, while the proposed method uses color saliency to detect the events. It can be seen from the figure that the Harris detector detects salient points that typically cluster around textured areas, while the proposed method spreads them according to color saliency. Figure 3 shows us how a particular interest point is categorized into 3x3 windows for color and 9x9 windows for texture [12].

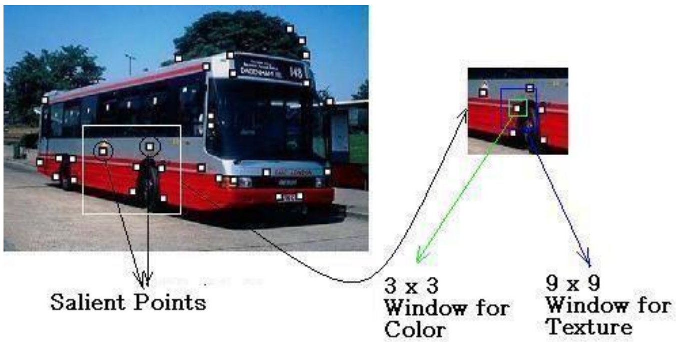

Figure 3: Feature computaion process

An enhanced Harris corner detection is followed by construction of the SIFT descriptor for image feature escription. The matching result is finally obtained by the nearest neighbor matching algorithm on the condition that feature points are well-proportioned distributing. In addition, it applies the knowledge of analytic geometry to calculate the distance between matching point and epipolar line to reduce the error matching. The experimental results prove that the combination of those algorithms is effective. This algorithm wins high matching accuracy and matching time- consuming cuts down.

Scale Invariant Feature Transform (SIFT), is an important aspect as it transforms the image data into scale-invariant coordinates relative to local features. Scale invariant features are those features that do not affect the algorithm upon modifying the scale. An important feature of this approach is that it generates large numbers of features that densely cover the image over the full range of scales and locations. The quantity of features is particularly important for object recognition, where the ability to detect small objects in cluttered backgrounds requires that at least 3 features be correctly matched from each object for reliable identification.

For image matching and recognition, SIFT features are first extracted from a set of reference images and stored in a database. A new image is matched by individually comparing each feature from the new image to this previous database and finding candidate matching features based on Euclidean distance of their feature vectors.

The keypoint descriptors are highly distinctive, which allows a single feature to find its correct match with good probability in a large database of features. The correct matches can be filtered from the full set of matches by identifying subsets of keypoints that agree on the object and its location, scale, and orientation in the new image using the epipolar constraints. Geometric relations between 3-dimensional points and their corresponding 2-dimentional projections lead to constraints between the image points. The determination of these consistent clusters can be performed rapidly by using an efficient hash table implementation of the generalized Hough transform. Figure 4 represents epipolar constraints in 2D form [4].

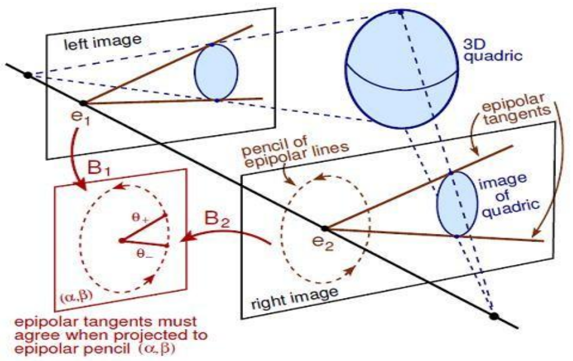

Figure 5: Projection of a 3D quadric to two image conics, and then projection via B; B0 of the pair of epipolar lines tangent to each conic to the epipolar pencil.

4. PROPOSED RESULTS

The expected test results are mentioned in the following discussion. The Harris corner detector was confirmed as a stable feature point extractor. At the same time the detector has some shortage. Figure 5 shows the salient points detected using the Harris corner detector and the proposed method [7]. The salient points are circled in yellow and blue for Harris detector and the proposed method, respectively [1].

Fig 1: If necessary, the images can be extended both columns

Figure 5b: Uploaded input image

Figure 5c: Results

In the figure 6 (a), 6 (b), the corner detection method avoids us to set up the threshold, makes the corners to be well-proportioned distributing and reject neighboring points. Based on the improved detector, the number of matches found is suitable, but not much greater than the amount needed for the considered application. This final step follows-up feature matching on the condition that the feature points are well-proportioned distributing.

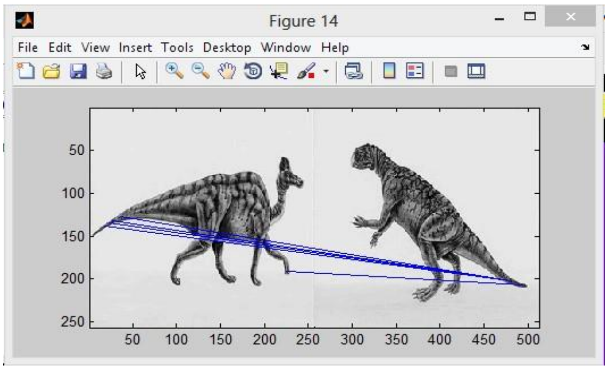

Figure 6 (a): Match result of Harris-SIFT algorithm

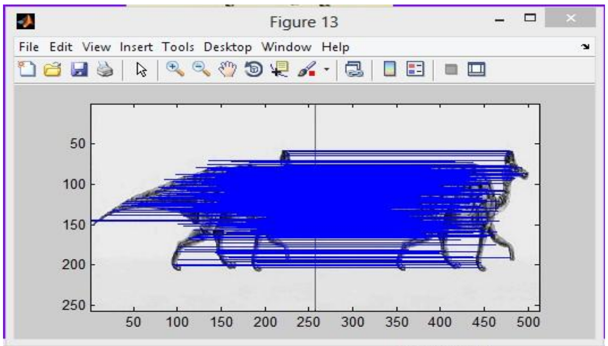

Figure 6 (b): Match result of Harris-SIFT algorithm

In the Table 1, The paper utilizes the NN (nearest neighbor) algorithm to obtain 53 matching points firstly, and then epipolar constraint presented in the work removes 7 false matching points [1].

Table 1 Comparison of Data in Image Matching

EXPERIMENT ALGORITHM	POINT AMOUNT MATCH ED	MATCHIN G RATE	MATCHIN G COMPUTE R TIME
Harris-SIFT	46	0.938	71.29s
SIFT algorithm	102	0.895	137.35s

5. CONCLUSION

Content based image retrieval is a technology that has been used in various search engines for several applications. The speed and accuracy of an algorithm is what determines the efficacy of the system that deploys the technique of CBIR. The idea presented in this paper is not only an optimization of two distinctively different and efficient algorithms but is also a simple optimization of both the algorithms where both their flaws are replaced with better techniques. These algorithms have been implemented to obtain better results separately in different areas. Our motivation for this innovation has been to produce a better and least faulty system for content based image retrieval.

6. ACKNOWLEDGMENTS

We have immense pleasure in presenting the synopsis report of our project entitled “Content Based Image Retrieval”. We would like to take this opportunity to express our gratitude to a number of people who have been sources of help and encouragement during the course of this project. We would like to place on record our sincere gratitude to our esteemed Principal Dr. Alka Mahajan for having given us this opportunity to prepare this report on “Content Based Image Retrieval” as a part fulfillment of requirements for the Degree of Bachelor of Engineering (B.E) in Information Technology. We also extend our gratitude to Prof. Leena Ladge, our H.O.D., for her valuable support and impartial opinion on our project. We are indebted to our guide Prof. K Lakshmi Sudha, for her valuable guidance and support and for all the time she could spare us throughout our Endeavour. We recognize the support and assistance given by the college staff and lab in-charges for providing us with the required facilities which enabled us to complete this project on time. Last but not the least; our families do require a special mention for their priceless and experienced advice and everlasting support at home. A ‘Thank You’ to just about everyone involved, to help make this project possible. That includes group members.

7. REFERENCES

[1] Jie Zhao1, Li-Juan Xue1, Guo-Zun Men. Optimization Matching Algorithm Based on improvement Harris and SIFT. College of Electronics and Information Engineering of Hebei University, Baoding 071002 at China College of Economics, Hebei University, Baoding, 071002, China 11-14th July 2010.

[2] David G. Lowe , Distinctive Image Features from Scale- Invariant Keypoints Computer Science Department, University of British Columbia, Vancouver, B.C., Canada, lowe@cs.ubc.ca 2004.

[3] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. Int. J. Computer Vision, 60(1):63–86, 2004.

[4] P. Bendale, B. Triggs, and N. Kingsbury. Multiscale keypoint analysis based on complexwavelets. In British Machine Vision Conference, August 2010.

[5] Mamta Kamath, Disha Punjabi, Tejal Sabnis, Divya Upadhyay, Seema Shrawne. Improving Content Based Image Retrieval usingScale Invariant Feature Transform. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 –8958, Volume-1, Issue-5, June 2012.

[6] K.Velmurugan, Lt. Dr. S.Santhosh Baboo. Image Retrieval using Harris Corners and Histogram of Oriented Gradients. International Journal of Computer Applications (0975 – 8887) Volume 24– No.7, June 2011.

[7] Hiremath P.S. and Jagadeesh PujariContent Based Image Retrieval using Colour Boosted SalientPoints and Shape features of an image.International Journal of Image Processing, Volume (2): Issue (1) 10.

[8] K. Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant Fourier descriptors to recognition of 3D objects," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, pp. 640-647, 1990.

[9] E. M. Arkin, L.P. Chew, D.P. Huttenlocher, K. Kedem, and J.S.B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," IEEE Trans. Pattern Analysis and MachineIntelligence, vol. 13, no. 3, pp. 209- 226, 1991.

[10] Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng. Fundamentals of content-based image retrievals. 2004.

[11] Faraj Alhwarin, Chao Wang, Danijela Risti -Durrant, Axel Gräser. Improved SIFT-Features Matching for Object Recognition. BCS International Academic Conference 2008 – Visions of Computer Science.

[12] J. van de Weijer, Th. Gevers, J-M Geusebroek, "Boosting Colour Saliency in Image Feature Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27 (4), April 2005. ," in CVPR, 2000, pp. 612–618.

[13] David G. Lowe. : Object recognition from local scale- invariant features, International Conference on Computer Vision, Corfu, Greece (September 1999), pp. 1150-1157.

[14] David G. Lowe. : Local feature view clustering for 3D object recognition, IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii (December 2001), pp. 682-688.