Improving Content Based Image Retrieval using Scale Invariant Feature Transform

Mamta Kamath, Disha Punjabi, Tejal Sabnis, Divya Upadhyay, Seema Shrawne

Abstract: - Content-Based Image Retrieval (CBIR) is a challenging task. Common approaches use only low-level features. Notwithstanding, such CBIR solutions fail on capturing some local features representing the details and nuances of scenes. Many techniques in image processing and computer vision can capture these scene semantics. Among them, the Scale Invariant Features Transform (SIFT) has been widely used in a lot of applications. This approach relies on the choice of several parameters which directly impact its effectiveness when applied to retrieve images. In this paper, we attempt to evaluate the application of the SIFT to refine CBIR.

Keywords: - Content Based Image Retrieval (CBIR), Difference of Gaussian (DOG), Nearest Neighbour Search (NNS), Scale Invariant Feature Transform (SIFT).

I. INTRODUCTION

Despite the many research efforts, the retrieval accuracy of today's CBIR algorithms is still limited and often worse than keyword based approaches. The problem stems from the fact that visual similarity measures, such as color histograms, in general do not necessarily match perceptional semantics and subjectivity of images. In addition, each type of image features tends to capture only one of many aspects of image similarity and it is difficult to require a user to specify clearly which aspect exactly or what combination of these aspects he/she wants to apply in defining a query. To address these problems, we propose the utilization of the Scale Invariant Feature Transform algorithm. 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. Matching features across different images in a common problem in CBIR. SIFT extracts distinctive invariant features from an image which can be used for performing reliable matching between different views of an object or scene. These features are scale and rotation invariant and have shown to provide robust matching across many different affine distortions, changes in 3D viewpoints, addition of noise and change in illumination. A single feature can be matched correctly with a very high probability against a large database of features of many images.

Manuscript Received on June 22, 2012.
Mamta Kamath, Computer Technology Department, Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India
Disha Punjabi, Computer Technology Department, Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India
Tejal Sabnis, Computer Technology Department, Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India
Divya Upadhyay, Computer Technology Department, Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India
Seema Shrawne, Computer Technology Department, Veermata Jijabai Technological Institute, Mumbai, Maharashtra, India

The most obvious drawback with SIFT is the time it takes to compare two images. The running time of an NNS search is so large that it effectively renders SIFT useless for a system like image search engine. We propose a system that uses SIFT algorithm to refine the search results returned by a standard content based image retrieval algorithm. This combines the best features of both algorithms providing speed as well as better precision.

II. BACKGROUND

A. Scale Invariant Feature Transform (SIFT)

Image matching is a fundamental aspect of many problems in computer vision, including Content based image retrieval. For any object in an image, interesting points on the object can be extracted to provide a "feature description" of the object. This description, extracted from a training image, can then be used to identify the object when attempting to locate the object in a test image containing many other objects. To perform reliable recognition, it is important that the features extracted from the training image be detectable even under changes in image scale, noise and illumination. Such points usually lie on high-contrast regions of the image, such as object edges. SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of keypoints that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches.

B. Content-based image retrieval (CBIR)

Content-Based Image Retrieval, also known as query by image content and content-based visual information retrieval is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. Content based means that the search makes use of the contents of the images themselves, rather than relying on human-input metadata such as captions or keywords. A content-based image retrieval system (CBIR) is a piece of software that implements CBIR. Content based image retrieval plays a central role in the application areas such as multimedia database systems in recent years. The work focused on using low-level features like color, texture, shape and spatial layout for image representation. Among all the visual features, color is perhaps the most distinguishing one in many applications. It may be represented by a color histogram in 2D or 3D.CBIR is desirable because most web based image search engines rely purely on metadata and this produces a lot of garbage in the results.

Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results.

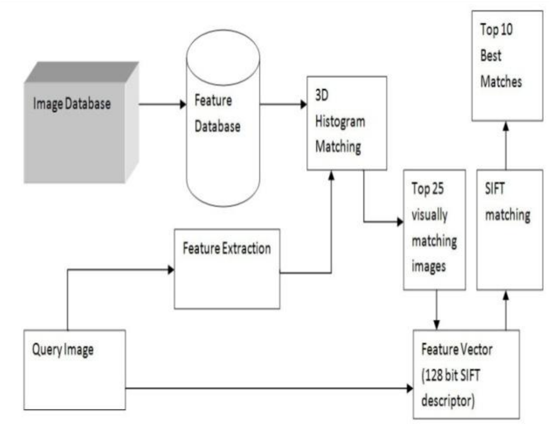
The goal of CBIR systems is to support image retrieval based on content e.g., shape, color, texture. CBIR uses features of images for image comparison, searching and matching. The commonly used features are Color, Shape, Texture and spatial positioning.

III. EXPERIMENTATION

A. The SIFT Approach

The following figure indicates how SIFT can be used for the purpose of CBIR -

A database of histograms is maintained along with a database of images. This database of features is computed offline, whereas the features of query image are computed online.



In order to retrieve M (user-defined) query results using CBIR algorithm, the following steps are executed:

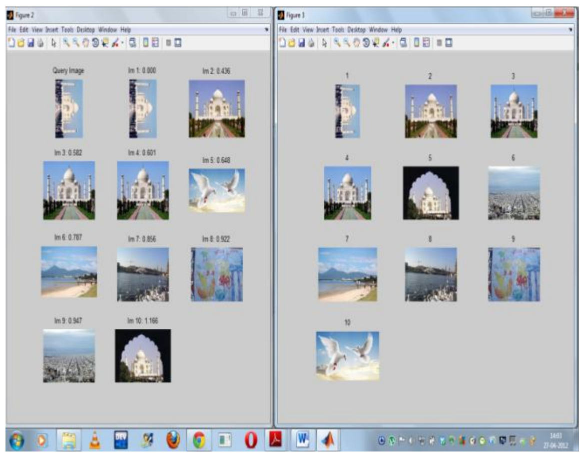
1. The 3D (HSV) histogram of the query image is computed. Then, the number of bins in each direction (i.e., HSV space)is duplicated by means of interpolation.
2. For each image i in the database:
(i) Load its histogram Hist(i).
(ii) Use interpolation for duplicating the number of bins in each direction.
(iii) For each 3-D hist bin, compute the distance (D) between the hist of the query image and the i-th database image.
(iv) Keep only distances (D2) for which, the respective hist bins of the query image are larger than a predefined threshold T (let L2 the number of these distances).
(v) Use a 2nd threshold: find the distance (D3) values which are smaller than T2, and let L3 be the number of such values.
(vi) The similarity measure is defined as: S(i) = L2 * average(D3) / (L3^2).
3. Sort the similarity vector and prompt the user with the images that have the M smaller S values.

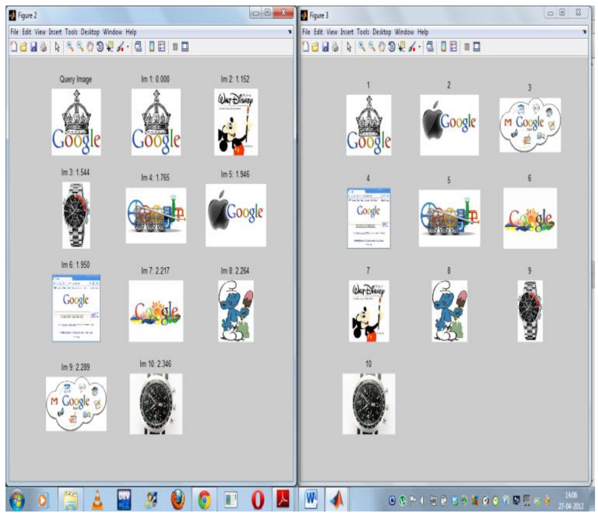
The resulting images are written back to a temporary database and SIFT function is called for refining and ranking the result images.

SIFT follows the following steps:

1. The algorithm SIFT first computes the a descriptor for each query image in the database.
   Constructing a scale space is the initial preparation. Internal representations of the original image are created to ensure scale invariance.
2. SIFT computes the descriptor for each query image in the database. Unwanted keypoints like edges and low contrast regions are eliminated by DOG.
3. An orientation is calculated for each key point. Any further calculations are done relative to this orientation, making it rotation invariant.
4. SIFT matches the keypoints of query image with the keypoints of each of the query results of the CBIR using NNS. They are then ranked as per number of matching keypoints.

IV. RESULTS





The LHS in the above figures show the results returned by CBIR (along with the query image). On the RHS we have the rectified ranking of the images done by SIFT (Applied only on the images returned by CBIR).

V. ANALYSIS

We have used images captured using different viewpoints, illumination and scale. The CBIR algorithm will make use of the pre-computed histograms of the database images. CBIR search is based on image intensity, texture and histogram, because of which some irrelevant images are returned, thus resulting in faulty ranking. In SIFT, nearest neighbour search (NNS) using Euclidean distance being expensive, an approximation of that is used to improve efiiciency. SIFT calculates the keypoints of query just once and uses number of matching keypoints with the query image to rank them. We use the location, scale and orientation of the images to match key-points of the query image and the images returned by CBIR. The algorithm SIFT does the job of eliminating the unwanted/non-matching images computed by CBIR. To speed up the process of image search and retrieval, we use CBIR over the database initially. Next, we apply SIFT over the images returned by CBIR, which saves a lot of time.

We have thus utilized the best features of both the algorithms to speed up the search and improve the accuracy.

VI. CONCLUSIONS

Many sophisticated algorithms have been designed to describe colour, shape, and texture features. But these algorithms cannot adequately model image semantics. They have many limitations when dealing with broad content image databases. Extensive experiments on CBIR systems show that low-level contents often fail to describe the high level semantic concepts in user's mind. Therefore, the performance of CBIR is still far from user's expectations. But, owing to SIFT's excellent performance in image matching, we can use our technique for exact image matching, say, looking for the almost same objects or scenes in a huge image database. Since the SIFT feature extracting and the high-dimensional vector indexing are both costly computational procedures, parallel processing using multiprocessors can be used in some applications where the runtime is critical.

REFERENCES

1. David G. Lowe , Distinctive Image Features from Scale-Invariant KeypointsComputer Science Department,University of British Columbia, Vancouver, B.C., Canada, lowe@cs.ubc.ca
2. Xu Wangming, Wu Jin, Liu Xinhai, Zhu Lei, Shi Gang, Application of Image SIFT Features to The Context of CBIR, Engineering Research Center of Metallurgical Automation and Measurement Technology，Ministry of Education, College of Information Science and Engineering, Wuhan University of Science and Technology Wuhan, China
3. P.S.Suhasini, Dr. K. Sri Rama Krishna, Dr. I.V. Murali Krishna , Cbir Using Color Histogram Processing , Asst. Professor in ECE , DMSSVH College of Engineering , Machilipatnam, A.P., India - 521001 Professor & Head, ECE Department, VR Siddhartha Engineering College, Vijayawada, A.P., India–520007
4. Ajay B. Kurhe, Suhas S. Satonka, Prakash B. Khanale, Color Matching Of Images By Using Minkowski- Form Distance , University Grant Commission

AUTHORS PROFILE

Mamta Kamath is a final year Btech student at Veermata Jijabai Technological Institute, Mumbai, majoring in Computer Science. She is keenly interested in DBMS and Image processing.

Disha Punjabi is a final year Computer Science, Btech Student at Veermata Jijabai Technology Institute, Mumbai. She is interested in software design, Multimedia, Security and would like to work with different technologies.

Tejal Sabnis is a final year Btech student at Veermata Jijabai Technological Institute (Mumbai) majoring in Computer Science. Her prime areas of interest are data structures and analysis of algorithms.

Divya Upadhyay is a final year Btech student at Veermata Jijabai Technological Institute, Mumbai. She likes programming and has a keen interest in technology in general, and DBMS,system security,front end designing in particular.

Mrs. Seema Shrawne graduated with a Bachelor's Degree in Engineering in computer science from Government College of Engineering, Amravati, in 1992.She has been a Senior lecturer in the computer technology department at Veermata Jijabai Technological Institute, Mumbai (India) for the last 11 years