

Content-Based Image Retrieval Based On Key Point Matching Using Scale Invariant Features Transform(Sift) Algorithm

MsPranjaliTayade¹, Prof Mr. D. R. Dhotre².

PG Student ¹, Guide ² SSGMCE College of Engineering Shegaon

Abstract—

Advances in data storage, data transmission, and image acquisition led to the creation of large image datasets. It became the motivation for systems that are able to efficiently retrieve images from these collections. This task has been addressed by the so-called Content-Based Image Retrieval (CBIR) systems. There are some quite powerful image descriptors designed to represent global features of images. These approaches have been widely used in image retrieval due to their usually low computational costs and acceptable effectiveness. However, such CBIR solutions fail on capturing some local features representing the details and nuances of the 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.

Keyword CBIR SIFTSURFetc

I INTRODUCTION

Due to rapid changes in digital technologies, in the recent years many people wish to publish their digital information on the Web such as text, image, video, audio etc. Hence, it requires effective indexing and searching tools for Web. Many researchers have been involved in developing the system to retrieve the set of similar Web images, for the given query image. The contents of an image have been used to represent the image semantically. The derived image features are used to retrieve relevant images semantically from the Web. Image processing now a day finds its application in all fields around us. Database related to images is on increasing. SIFT is an image processing algorithm which can be used to detect distinct features in an image. Once features have been detected for two different images, one can use these features to answer questions like “are the two images taken of the same object?” and “given an object in the first

image, is it present in the second image?” Thus, the feasibility of SIFT algorithm for CBIR is tested and used in image retrieval process. address a user. # Tags are used for trending a topic. RT shows that the tweet is retweeted[4][6]. Content-based image retrieval uses the visual contents of an image such as colour, shape, texture, and spatial layout to represent and index the image. In typical content-based image retrieval systems (Fig 1), the visual contents of the images in the database are extracted and described by multidimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarity distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image.

II LITERATURE SURVEY

Zhang *et al.* [2] proposed a bag of images for CBIR schemes. They supposed that the image collection composed of image bags rather than independent individual images. They contain some relevant images that have same perceptual meaning. The image bags were built before image retrieval. In addition, a user’s query is an image bag, named query image bag. In this condition, all image bags in the image collection are sorted according to their similarities to the query image bag. It hypothetically represented that the new idea can enhance the image retrieval process. However, this work needs to develop more efficient ways to measure the dissimilarity between two image bags.

Ponitzet *al.* [3] attempted to solve the problem of detecting images limitations in huge scale image databases. They decide to enhance the methodology

of BoVW by improving the distance measure between image signatures to avoid the occurrence of vague features. They utilized SIFT algorithm for local visual features acquisition. Only 60% of all images were randomly chosen, and their features utilized for clustering. These features were then quantized. 100 random images are selected as input images. The images were changed with mounting distortion to test the robustness of the application. It needs more discrimination force of the actual image description.

Liu [4] reviewed BoVW model in image retrieval system. He provided details about BoVW model and explained different building strategies based on this model. First, he presented several procedures that can be taken in BoVW model. Then, he explained some popular keypoint detectors and descriptors. Finally, he looked at strategies and libraries to generating vocabulary and do the search.

FEATURE EXTRACTION

The Scale-Invariant Feature Transform (SIFT) extracting keypoints are invariant to rotation, scaling, and translation. These keypoints are used to detect distinctive edges and textures in an image. Each keypoint has 132 dimensions: 128 spatial orientations bins, plus coordinates, scale, and rotation of the keypoints. We will also discuss to extract global image descriptors and represent images as a collection of local properties and calculate from a set of small sub-images called blocks. Image features are extracted from a set of collected images and stored in a database. According to David Lowe [4], SIFT gives a good result to recognize/retrieve images. This approach achieves scale invariance and also less sensitive to local image distortions such as 3D viewpoint change. The feature vector consists of SIFT features computed on a regular grid across the image Dense SIFT (DSIFT). These feature vectors are quantized into visual words to build visual dictionary using proposed quantization algorithm.

SCALE INVARIANT FEATURE TRANSFORM (SIFT)

Lowe [3] developed SIFT as a continuation of his previous work on invariant feature detection. It has four computational phases:

- (a) extrema detection
- (b) keypoint localization

(c) orientation assignment

(d) keypoint description.

The first phase examines the image under different octaves and scales to isolate points of the image that are different from their surroundings. These points, which are called extrema, are potential candidates for image features. In keypoint localization phase, it selects some of extrema points to be keypoints. Candidate keypoints are refined by reject extrema points that are caused by edges and by low contrast points. In the orientation assignment phase, it represents every keypoint and neighbors as a set of vectors using the magnitude and the direction. In the last phase, it takes a collection of vectors in the neighborhood of every keypoint and combines this information with a set of eight vectors called the descriptor. The neighborhood is divided into 4×4 regions, in each region the vectors are histogrammed in eight bins. SIFT provides a 128 element of the keypoint descriptor.

MOTIVATION

In the past few years, many researchers have been involved in the area of Content-Based Image Retrieval (CBIR) system to develop techniques to retrieve unannotated images [1]. Today many people use a digital images and video libraries as the main source of visual information. Hence it is an open challenge for the research community to develop cost effective technologies for retrieving, managing and browsing the images in the Web.

CONTENT BASED IMAGE RETRIEVAL

Content-based image retrieval uses the visual contents of an image such as colour, shape, texture, and spatial layout to represent and index the image. In typical content-based image retrieval systems (Figure 1), the visual contents of the images in the database are extracted and described by multidimensional feature vectors. The feature vectors of the images in the database form a feature database. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarity distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing

scheme. The indexing scheme provides an efficient way to search for the image database.

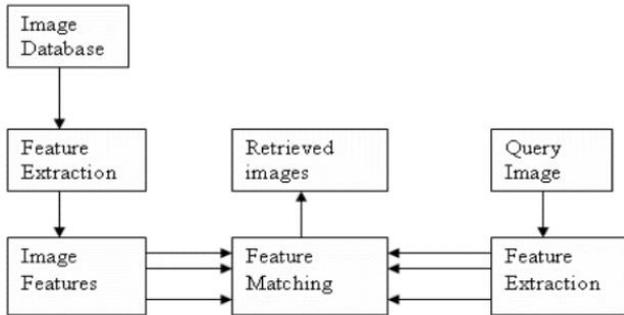
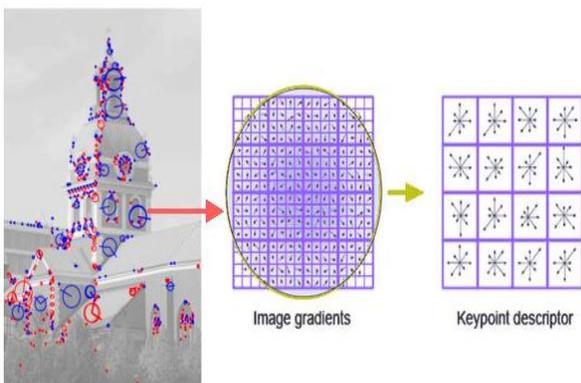


Fig 1 CBIR SYTEM

SIFT - Scale Invariant Feature Transform

Point like features are very popular in many fields including 3D reconstruction and image registration. A good point feature should be invariant to geometrical transformation and illumination. A point feature can be a blob or a corner. SIFT is one of most popular feature extraction and description algorithms. It extracts blob like feature points and describe them with a scale, illumination, and rotational invariant descriptor.



PROPOSED SIFT ALGORITHM

The SIFT algorithm identifies features of an image that are distinct, and these features can in turn be used to identify similar or identical objects in other images. SIFT has four computational phases. The reason for this being that some computations performed by SIFT are very expensive. The cost of extracting the keypoints is minimized by the cascading approach of SIFT. The more expensive operations are only applied on locations that pass an

initial, cheaper test. The output of the SIFT algorithm is a set of keypoint descriptors. Once such descriptors have been generated for more than one image, one can begin image matching (*Figure 1.1*). The image matching, or object matching, is not part of the SIFT algorithm. For matching we use a Nearest Neighbour Search (NNS), an algorithm that is able to detect similarities between keypoints. Thus, SIFT only makes matching possible by generating the keypoint descriptors.

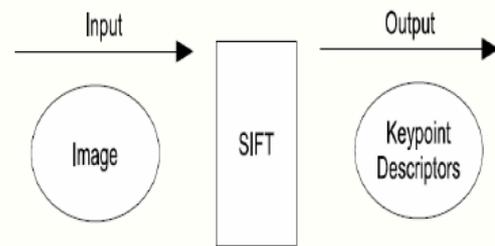


Figure 1.1 SIFT takes as input an image, and generate a set of keypoint descriptors. The keypoint descriptors may then be stored in a separate file.

Although the matching process is not part of SIFT, the results from image matches are used as an indicator of how well the SIFT algorithm is suited for image matching. (*Figure 1.2*) shows that in an image match checking, the two sets of keypoint descriptors are given as input to a nearest neighbour search algorithm. The output of the algorithm is a set of keypoint descriptors found to be very similar

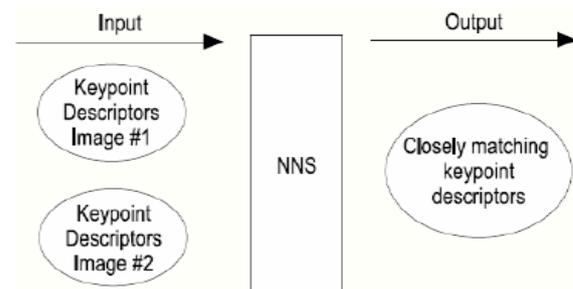


Figure 1.2 In an image match checking, the two sets ofkeypoint descriptors

Image Matching

The three possible image matches are,

1.A match where the whole of one image matchesthe whole of another image.

2. Part of one image matches the whole of another image.

3. Part of one image matches the part of another image.

The different matches will have different characteristics. If there is a match in case

1, a fairly large percentage of all keypoints are matched. In case

2, there is a large percentage match in the image with a whole match, and a small percentage match in the image partly matching.

In case 3, there is a fairly low percentage of keypoints matching in both images.

1) BAG-OF-FEATURE (BOF) DESCRIPTOR

BoF is one of the popular visual descriptors used for visual data classification. BoF is inspired by a concept called Bag of Words that is used in document classification. A bag of words is a sparse vector of occurrence counts of words; that is, a sparse histogram over the vocabulary. In computer vision, a bag of visual words of features is a sparse vector of occurrence counts of a vocabulary of local image features. BoF typically involves in two main steps. First step is obtaining the set of bags of features. This step is actually offline process. We can obtain set of bags for particular features and then use them for creating BoF descriptor. The second step is we cluster the set of given features into the set of bags that we created in first step and then create the histogram taking the bags as the bins. This histogram can be used to classify the image or video frame.

2) BAG-OF_FEATURES WITH SIFT

we build BoF with SIFT features.

1. Obtain the set of bags of features.

1. Select a large set of images.

2. Extract the SIFT feature points of all the images in the set and obtain the SIFT descriptor for each feature point that is extracted from each image.

3. Cluster the set of feature descriptors for the amount of bags we defined and train the bags with clustered feature descriptors (we can use the K-Means algorithm). Obtain the visual vocabulary.

• 2. Obtain the BoF descriptor for given image/video frame.

5. Extract SIFT feature points of the given image.

6. Obtain SIFT descriptor for each feature point.

7. Match the feature descriptors with the vocabulary we created in the first step Build the histogram.

III PROPOSED ARCHITECTURE

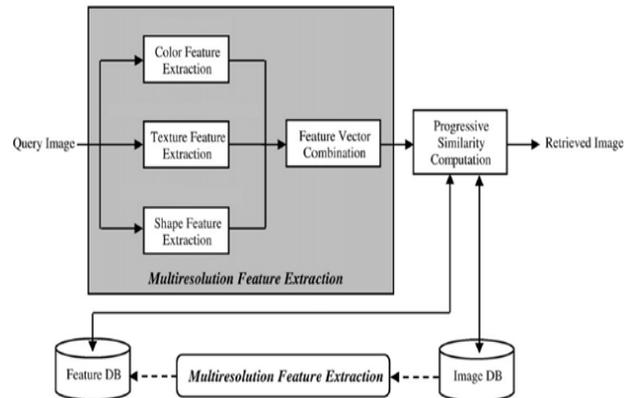
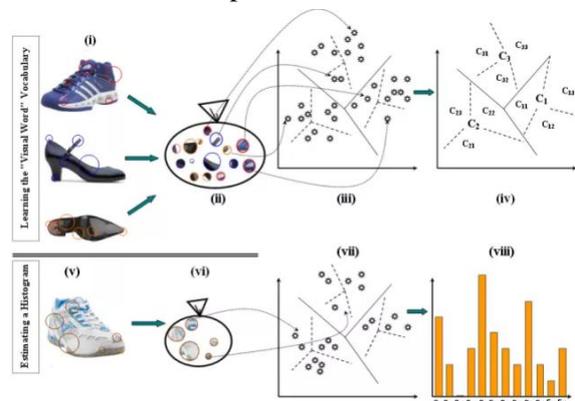


Fig 3 Proposed Architecture

EXPERIMENT

We study the extent to which a standard 'bag of visual words' image classifier can be used to tag products with useful information, such as whether a sneaker has laces or velcro straps. Using Scale Invariant Feature Transform (SIFT) image descriptors at random keypoints, a hierarchical visual vocabulary, and a variant of nearest-neighbor classification, we achieve accuracies between 66% and 98% on 2- and 3-class classification tasks using several dozen training examples. We also increase accuracy by combining information from multiple views of the same product.



RESULTS

The result of the project is classified into three different demos.

- i) Keypoint generation for a particular image
- ii) Matching of similar keypoints between two images
- iii) Retrieval of images based on the content of the query image. Retrieval of images based on the content of the query image

IV CONCLUSION

In this paper proposed semantic based image retrieval system to retrieve set of relevant images for the given query image from the Web. We have used global color space model and SIFT feature extraction technique to generate visual dictionary using proposed quantization algorithm. The images are transformed into set of features. These features are used as inputs in our proposed Quantization algorithm for generating the code word to form visual dictionary. These code words are used to represent images semantically to form visual labels using Bag-of-Features (BoF). The Histogram intersection method is used to measure the distance between input image and the set of images in the image database to retrieve similar images. The experimental results are evaluated over a collection of 1000 generic Web images to demonstrate the effectiveness of the proposed system.

V REFERENCES

[1] Takeshi Sakaki, Makoto Okazaki, Yutaka Matsuo, „Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors“, April 26-30, 2010, Raleigh, North Carolina. 2010

[2] Zhang, J., & Ye, L. “Image retrieval based on a bag of images“. In Image Processing (ICIP), 2009 16th IEEE International Conference on (pp. 1865-1868). IEEE. 2009.

[3] Pönitiz, T., Stöttinger, J., Donner, R., & Hanbury, “A. Efficient and Distinct Large Scale Bags of Words“. Austrian Research Promotion Agency (FFG). 2010.

[4] Alfanindya, A., Hashim, N., & Eswaran, C. “Content-Based Image Retrieval and Classification using speeded-up robust features (SURF) and grouped bag-of-visual-words (GBoVW) “. In Technology, Informatics, Management, Engineering, and Environment (TIME-E), 2013 International Conference on (pp. 77-82). IEEE. 2013.

[5] Rasim M A, Ramiz M A, and Saadat A N. Classification of Textual E-mail spam using Data Mining Techniques [J]. In the Journal of Applied Computational Intelligence and Soft Computing, 2011. [6] Wang, A.H.: Don't follow me: Spam detection in twitter. In: Security and Cryptography (SECRYPT), Proceedings of the 2010 International Conference on, IEEE (2010) 1–10

[7] Gao, H., Chen, Y., Lee, K., Palsetia, D., Choudhary, A.: Towards online spam filtering in social networks. In: Symposium on Network and Distributed System Security (NDSS). (2012)

[8] Biro I, Szabo J, Benczur A, and Siklosi D. Linked Latent Dirichlet Allocation in Web Spam Filtering [A]. In Proceedings of the 4th International Workshop on Adversarial Information Retrieval on the Web (AIR Web), Madrid, Spain, 2009.

[9] Perkins A. The classification of search engine spam. <http://www.ebrandmanagement.com/whitepapers/spamclassification>, 2001.

[10] Paulo C, Clotilde L, Pedro S. Symniotic data mining for personalized spam filtering