IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 12, December 2018

An Evaluation Approach for Image Copy Detection Based on Verification of Scale invariant Feature Transform (SIFT) Matches

Sushma.B[1], Nagaveni.B.Biradar[2]

M. Tech, 3rd Sem. Dept. of Computer Science & Engineering, RYMEC, Ballari, India[1]

Assistant Professor, Dept. of Computer Science & Engineering, RYMEC, Ballari, India[2]

Abstract: The So-called image copy detection based on an image and pasting in another location of the same image is common way to manipulate the image content and difficult to detect illegal copies of copyrighted images,. In this paper the existing system is based on matching pairs by descriptors mainly by using visual words for given image matching which is hard to distinguish between images. This technique is called Bag-Of-Words (BOW) Quantization which is used cannot solve the problem well. To address this problem Scale-Invarient Feature Transform (SIFT) matches between the images where the technique used is mainly based on BOW Quantization where global context regions is done to filter false images. And thus comparatively gives rich performance of encoding between the two images.

Keywords: Bag - Of - Words (BOW), Scale - Invarient Feature Transform (SIFT), Geometrical Consistency, Difference-of-Gaussians (DoG)

I INTRODUCTION

With rapid development of network technology digital multimedia like (images, audios, videos) has become easy for replication, modification of data. To protect it from unauthorized duplication it must become a basic requirement[5]. The duplicate image detection is done comparing with original image which helps to get the similarity between the query images. The similarity factor may mainly depend on various geometrical consistencies factors such as viewpoints ,illumination, scaling, cropping, rotation, affine transformation which depends on image pair samples.
To tackle huge number of duplicate of false image show ever first BOW model was widely used which unsupervised feature extraction method for variety of computer vision applications.

The bag of words is an important task for unsupervised representation of visual based data on local feature descriptors based on which lead to many false images, thus had two drawbacks, the first is that local features do not encode enough spatial information and thus have limited discriminability. The second is that the BOW quantization errors will further degrade the discriminability of local features. Both the two drawbacks will lead to many false matches when there are many visually similar local patches, typically found between the similar images. To overcome this SIFT has been developed where matching of SIFT features are done based on Difference-of-Gaussians (DoG). Note that the two corresponding local patches of each match are visually quite similar but not copies of each other, and thus all of these SIFT matches can be regarded as false matches. Thus this paper focuses on how effectively and efficiently the original and the query image is matched for a large database which have large similar images in it which explores large context matching based on SIFT features to identify and remove false images most specifically based on Difference-of-Gaussians (DoG).

II RELATED WORK

The literature, many copy detection methods based on local features have been proposed. To the best of our knowledge these methods are generally investigate the local features which is given by two methodologies respectively. In [1], local features are extracted from small local patches and thus encode less spatial context information, they have limited discriminability. As a result, many false matches will occur when there are many visually similar local patches, typically found between the similar images As a result, it quantizes the extracted local features into visual words and then indexes images using an inverted file structure for image search some similar images will be falsely detected as image copies, which will significantly affect detection accuracy. Although BOW model was proposed they can achieve efficiency, more false matches will occur between similar images, since the BOW quantization errors will further degrade the discriminability of local features. Consequently, the accuracy of copy detection will be further decreased.

[2] Scale Invariant Feature Transform (SIFT) which is the most renowned feature-detection-description algorithm SIFT detector is based on Difference-of-Gaussians (DoG) operator where it is an Feature-points are detected by searching scale space extrema using DoG at various scales of the subject images SIFT is robustly invariant to geometric consistency such as image rotations, scale, and limited affine variations.

III METHODOLOGY

In the past few years Bag-Of-words(BOW) model [3] has become very popular for the large scale image retrievals. This method generally works as the representation in the fig., where this can be obtained by using three steps such as feature detection based on feature description, Codebook generator ,Content Based image Indexing and Retrieval(CBIR) such that local features are extracted from images and are quantized into visual words and then indexed with inverted file structure for image. Feature representation: After feature detection ,each image is abstracted by several local patches. Feature representation methods figure 1(a) and figure 1(b) deals with how to represent the patches as numerical vectors. These values are called feature descriptors. Codebook generation and Content Based image Indexing and Retrieval(CBIR)are used in the next step of BOW model is to convert certain code words through grouping process and thus the image can be represented by frequencies of code words. The BOW model has major two drawbacks when applied to copy detection. The first is local features do not encode enough spatial information. The second is BOW quantization errors will further degrade the discriminability of local features which results in many false images. To overcome this good descriptor should have ability to handle the geometric consistency such as rotation, scaling and various affine transformation to some extent. One of the most famous descriptor is Scale-invariant feature transform(SIFT).

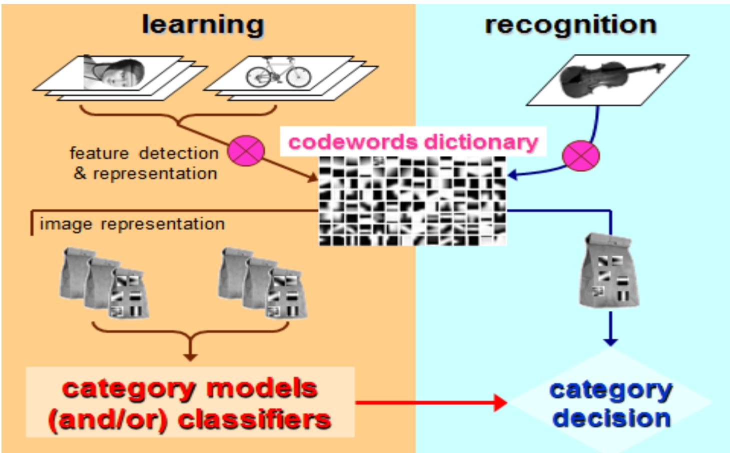

Figure 1 (a) It represents the Bag-Of Words (BOW)

Model where in the image recognition of the feature extraction and category classification is done based on visual vocabulary and is Code words Dictionary.

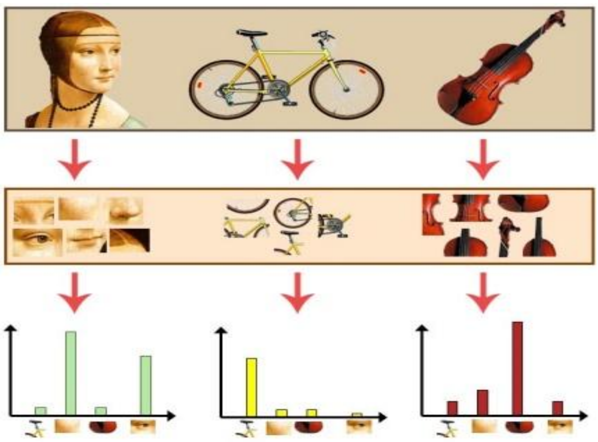

Figure 1(b) represents the Bag-Of-Words (BAG) model based on the values given by Codebook Generator and represents based on images of frequencies of visual words.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 12, December 2018

3.1 Proposed System

Therefore we propose effective and efficient Global Context Verification Scheme which explores the global context information by matching SIFT feature[4] and these which results in global context regions which almost covers the whole image by expanding the local features proportionally. We can clearly say that although the corresponding local patches of two falsely matched where SIFT features are visually similar there global context is somewhat different. SIFT feature obtains SIFT matches between the images based on BOW quantization technique and we extract overlapping region based on global Context descriptors for matched SIFT features to describe its global context information. The global context descriptor for verification encodes relatively rich global context information of each match SIFT feature but also has good robustness and efficiency.

A. Technique for SIFT feature matching

We extract hundreds of SIFT features based on SIFT algorithm, as shown in Figure 2 and these keypoints extract SIFT features are then efficiently matched using Difference-of-Gaussians (DoG). These features extracted mainly based on relative spatial positions based on original scene which can robustly identify even among the groups or even under the occlusion mainly because of SIFT feature descriptor which is invariant to scaling, rotation, orientation, illumination changes and is partial invariant to affine distortion. After efficient match codebook includes lot of visual words generated through group of sample set of SIFT features keypoints by hierarchy of visual vocabulary approach. According to their visual words to obtain inverted index file the indexed features records the ID of the image which belongs to domain orientation Φ, characteristic scale belongs s and coordinated (x,y) which helps in extortion of overlapping global context descriptors for SIFT matches verification. By using the inverted index file, any two SIFT features from two images quantized to the same visual word are regarded as a local match between the two images.
To achieve robustness for geometric transformation which includes rotation, scaling, cropping we construct overlapping regions of images for global context descriptor extraction. To get each pair of matched SIFT features between the two images we depend on domain orientation and characteristic scale to adjust the orientation and characteristic scales to make them consistent.

Figure - 2: Comparison between the two Images using Keypoints for Image match using SIFT Feature

We can represent their keypoint coordinates to compute the overlapping regions between the images i.e., the two matched SIFT features f_A and f_B are a correct match between an original image A and its copy B, which is generated by transforming image A with a combination of the three geometric attacks, which are rotation, scaling, and cropping. We denote the corresponding keypoints of the two features as P_a and P_b and, their orientation angles as Φ_a Φ_b and scales s_A and s_B.
SIFT features are extracted based on image properties, their property values the dominant orientations and characteristic scales change covariantly with the rotation and scaling transformations. we can adjust the orientations and scales of images A and B to be consistent according to the dominant orientations and characteristic scales of their features f_A and f_B. Taking f_B as reference point and setting the keypoint coordinates of feature f_A as the origin image in A, we first rotate the image A by aligning the orientation of f_A to the same of the f_B. Then, the scale of image A is adjusted to be the same as that of image B by multiplying with the ratio between scales s_B and s_A. The transformation can be represented by

$$X_A' = s_B / s_A \begin{pmatrix} \cos(\phi_B - \phi_A) & -\sin(\phi_B - \phi_A) \\ \sin(\phi_B - \phi_A) & \cos(\phi_B - \phi_A) \end{pmatrix} X_A$$

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
IJARCCE
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 7, Issue 12, December 2018

X_A denotes the coordinates of any pixel point in image A, and $X_{A'}$ is the adjusted coordinates of the point in the transformed image.

Scale-space extrema detection

By detecting points of interest, which are termed keypoints in the SIFT framework. The image is convolved with Gaussian filters at different scales, and then the difference of successive Gaussian-blurred images are taken. Keypoints are then taken as maxima/minima of the Difference of Gaussians (DoG) that occur at multiple scales. Specifically, a DoG image D(x,y,σ) is given by

$$\mathbf{D(x,y,\sigma)=L(x,y,k_i\,\sigma)-L(x,y,k_j\,\sigma)}$$

Where L(x,y,k σ) is the convolution of original Image I(x,y) at the scale k σ i.e.,

$$\mathbf{L(x,y,k\,\sigma) = G(x,y,k\,\sigma) * I(x,y)}$$

Hence a DoG image between scales k_i σ and k_j σ is just the difference of the Gaussian-blurred images at scales k_i σ and k_i σ . For scale space extrema detection in the SIFT algorithm, the image is first convolved with Gaussian-blurs at different scales.

Scale-space extrema detection produces too many keypoints , some of which are unstable. The next step is to perform a detailed fit to the nearby data for accurate location.

First, for each keypoint, interpolation of nearby data is used to accurately determine its position. The approach is to just locate each keypoint at the location and scale of the given keypoint. The approach calculates the interpolated location of the extreme. The D(x,y,σ) keypoint as the origin is given by keypoint as the origin

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}}\mathbf{x} + \frac{1}{2}\mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2}\mathbf{x}$$

where D and its derivatives are evaluated at the candidate keypoint and x = (x,y, σ) is the offset from this point.

Orientation assignment

Here each keypoint is assigned one or more orientations based on local image gradient directions. This is the key step in achieving invariance to rotation as the keypoint descriptor can be represented relative to this orientation and therefore achieve invariance to image rotation.

$$\theta(x,y) = \mathrm{atan2}\left(L(x,y+1) - L(x,y-1), L(x+1,y) - L(x-1,y)\right)$$

$$m(x,y) = \sqrt{\left(L(x+1,y) - L(x-1,y)\right)^2 + \left(L(x,y+1) - L(x,y-1)\right)^2}$$

For the image L(x,y, σ) at the keypoint's scale σ, is taken so that all computations are performed in a scale-invariant manner. For an image sample L(x,y) at the scale σ the gradient magnitude m(x,y) and orientation Θ(x,y) is given by the magnitude and direction calculations for the gradient are done for every pixel in a neighboring region around the keypoint. Therefore, SIFT descriptors are invariant to minor affine changes. To test the distinctiveness of the SIFT descriptors, matching accuracy is also measured against varying number of keypoints in the testing database, and it is shown that matching accuracy decreases only very slightly for very large database sizes, thus indicating that SIFT features are highly distinctive.

CONCLUSION

In this paper we present efficient and efficiency global context verification scheme for image copy detection using SIFT feature Descriptor . Hence SIFT features is explored for verification of SIFT matches to remove false matches and helps in improve the detection performance. Hence we can conclude that our method has significant in copyright protection and is also quite useful in many other tasks such as automatic annotating, content based web links, creation and redundancy elimination.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 7, Issue 12, December 2018

REFERENCES

[1]. K. Yan and L. Huston, "Efficient near-duplicate detection and subimage retrieval," in Proc. 12th ACM Int. Conf. Multimedia, pp. 869-876, Jan. 2004.
[2]. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.
[3]. J Sivic, and A. Zisserman, "Video Google: a text retrieval approach toobject matching in videos," in Proc. 9th IEEE Int. Conf. Comput. Vis.,2003, pp. 1470-1477.
[4]. Dr. K. Velmurugan "A Survey of Content-Based Image Retrieval Systems using Scale-Invariant Feature Transform (SIFT)" International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 1, January 2014.
[5]. Zhili Zhou, Yunlong Wang, Q.M. Jonathan Wu, Member, IEEE, Ching-Nung Yang and Xingming Sun "Effective and Efficient Global Context Verification for Image Copy Detection" IEEE Transaction on Imag/e Process, Volumn 15, 2017.
[6]. E. Mortensen, H. Deng, and L. Shapiro, "A SIFT descriptor with global context," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., 2005. pp. 184-190.
[7]. J. Liu and G. Zeng, "Improved global context descriptor for describing interest regions," J. Shanghai Jiaotong Univ., vol. 17, no. 2, pp.147 152, Apr. 2012.