International Journal on Future Revolution in Computer Science & Communication Engineering
Volume: 5 Issue: 1
ISSN: 2454-4248
53 – 56

Watermarking Based Image Authentication for Secure Color Image Retrieval in Large Scale Image Databases

Piyusha Narayan Chaudhari
Department of E&TC
J T Mahajan College of Engineeing
Faizpur (MS)
piyuchaudhari26@gmail.com

Kanchan Subhash Bhagat
P.G.Coordinator,
Department of E&TC
J T Mahajan College of Engineeing
Faizpur (MS)
ksbhagat80032@gmail.com

Dr. J. P. Chaudhari
Associate Professor,
Charotar University of
science and technology ,at
and post changa,Gujarat
jitendrachaudhari.ec@charusat.ac.in

Abstract— An important facet of traditional retrieval models is that they retrieve images and videos and consider their content and context reliable. Nevertheless, this consideration is no longer valid since they can be faked for many reasons and at different degrees thanks to powerful multimedia manipulation software. Our goal is to investigate new ways detecting possible fake in social network platforms. In this paper, we propose an approach that assets identification faked images by combining standard content-based image retrieval (CBIR) techniques and watermarking. We have prepared the wartermarked image database of all images using LSB based watermarking. Using gabor features and trained KNN, user is able to retrieve the matching query image. The retrieved image is authenticated by extracting the watermark and matching it again with the test image.

Keywords—*watermarking, LSB watermarking, KNN, Gabor features. CBIR.*

*****

I. INTRODUCTION

Content-based image retrieval (CBIR) are easy to manipulate and edit due to availability of powerful image processing and cloud computing environment. With the development of the imaging devices, such as digital cameras, smartphones, and medical imaging equipment's, our world has been witnessing a tremendous growth in quantity, availability, and importance of images. The needs of efficient image storage and retrieval services are reinforced by the increase of large-scale image databases among all kinds of areas. Meanwhile, after more than twenty years of development, CBIR techniques show the potential of usefulness in many real-word applications. For example, clinicians can use CBIR to find similar cases of patients and facilitate clinical decision-making processes.

Large image database usually consists of millions of images. Therefore, CBIR services typically incur high storage and computation complexities. Cloud computing offers a great opportunity for the on-demand access to ample computation and storage resources, which makes it an attractive choice for the image storage and CBIR outsourcing. By outsourcing CBIR services to the cloud server, the data owner is relieved from maintaining local image database and interacting with database users online. Despite the tremendous benefits, image privacy becomes the main concern with CBIR outsourcing. For example, patients may not want to disclose their medical images to any others except to a specific doctor in medical CBIR applications. To formulate the problem, this paper considers watermarking based content image retrieval. The watermarked image database is prepared using LSB based watermark. The KNN classifier is trained using the gabor features extracted from the watermarked image dataset. The watermarked image extracted from the recognized image is matched with the test image watermark for authentication.

II. EASE OF USE

We intent to develop a system that enables to make sense about the images fake. Such system needs to be able to manage the image content as well as users security. We believe that content-based image retrieval (CBIR) techniques and watermarking based image security is the appropriate solution. In our work, we consider the image as faked if the extracted watermark is not matching with the original inserted watermark. Since a faked image is generated based on an original image, the requirements of the image fakery is that it retains similar visual content to the original image. Therefore, the faked image shares lots of information with the original enough to be distinguished from any other images. Based on this fact, detecting faked images requires robust and efficient visual features to be managed. To fulfill the above requirements and to achieve the interoperability, well established gabor features will be used to create the image signature. The low level image features do not provide a comprehensive information about the image fake, but they can be support the community decision. On the other hand, we also intent to deploy security aspects in our fake detection approach. The key idea is behind the fact that most of discovered faked images were identified by some one who knows the original and then those images have been analyzed against the fake.

III. RELATED WORK

Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun and Kui Ren all in their paper titled "A Privacy-preserving and Copy-deterrence Content-based Image

 ISSN: 2454-4248
53 – 56

Retrieval Scheme in Cloud Computing" focused on preserving privacy and doing copy deterrence while retrieving content based images in cloud computing environment. In this system, feature vectors are extracted to represent the corresponding images. For preserving privacy, access to unauthorized users get prevented. Watermark certification authority is provided in cloud computing environment for making images more secure [1].

Kui Ren, Zhihua Xia, Zhan Qin, and Yi Zhu, all in their paper titled Towards Privacy-preserving Content-based Image Retrieval in Cloud Computing focused on Content-based image retrieval (CBIR) applications which are developed very fastly along with the improvement in the availability, quantity and importance of images which are present in daily life. In this system, privacy is get preserved of retrieval process to control the access of images by authorized users only. There is data owner is present who send the CBIR service and image database to the cloud, without giving any idea about the original contents of the image database to the server [2].

Xinhui Wang, Zhihua Xia, Zhan Qin, Liangao Zhang, Kui Ren and Xingming Sun, all in their paper titled A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing focused on retrieval process of content based image with copy deterrence and preserving privacy in cloud computing. For preserving privacy, sensitive images, like personal and medical images, required to convert in encrypted form before outsourcing, because of this CBIR technologies present in plaintext domain becomes unusable. Moreover, secure kNN algorithm is used to protect the feature vectors, and standard stream cipher encrypt the image pixels [3].

Nasir Memon, K. Gopalakrishnan, Poorvi L. Vora, all in their paper titled Protocols for Watermark Verification focused on adding a watermark signal into the digital image which is later be detected or extracted for making an assertion about the particular image. There are two catagories of watermarks present: invisible and visible. Conspicuously company logos or visible messages are present in visible watermarks which indicates the image ownership. On the other hand, Invisible watermarks contains unobtrusive modifications to the image and the invisibly watermarked image which visually appears very similar to the original image [4].

Jens-Rainer Ohm, B. S. Manjunath, Akio Yamada and Vinod V. Vasudevan all in their paper titled Color and Texture Descriptors focused on presenting color overview and texture descriptors that get approved for the MPEG-7 standard Final Committee Draft. Histogram descriptor present in this color descriptors standard are get coded with the help of Haar transform, a dominant color descriptor, a color structure histogram, and a color layout descriptor. These all texture descriptors contains the one which make characterization of homogeneous texture regions. It also contain another which provide representation of local edge distribution [5].

X. Wang, Z. Xia, Q. Wang and X. Sun, all in their paper titled A secure and dynamic multi-keyword ranked search scheme over encrypted cloud data focused on a special index structure created for tree based images and uses a Greedy Depth-first Search algorithm for providing efficient ranked search of multi-keyword. For encrypting the index and query vectors, secure kNN algorithm comes in account, and this algorithm ensure accuracy of relevance score calculation in between query vectors and encrypted index [6].

P. T. Boufounos and S. Rane all in their paper titled Privacy preserving nearest neighbor methods: comparing signals without revealing them, focused on the privacy-preserving NN (PPNN) method, in which the reader will come to know that it convenient to make dividation of this in two different problems: privacy-preserving minimum finding method follow another method that is privacy-preserving distance computation. Under certain considerations, privacy model dictate that which mathematical tools should be useful for PPNN. It also define the complexity and structure of resulting protocols . These models makes assumptions upon requirement, behavior,sharing. These assumptions have main focus on participating entities behavior, the amount of possible information that get shared among participants and privacy requirements. [7].

IV. IMAGE WATERMARKING USING LSB

There are many algorithms available for invisible digital watermarking. The simplest algorithm is Least Significant Bit (LSB) Insertion, in which each 8-bit pixel's least significant bit is over written with a bit from the watermark. In a digital image, information can be inserted directly into every bit of image information or the more busy areas of an image can be calculated so as to hide such messages in less perceptible parts of an image. This method is based on the pixel value's Least Significant Bit (LSB) modifications. The principle of embedding is fairly simple and effective. If we use a grayscale bitmap image, which is 8- bit, we would need to read in the file and then add data to the least significant bits of each pixel, in every 8-bit pixel. In a grayscale image each pixel is represented by 1byte consist of 8 bits. It can represent 256 gray colors between the black which is 0 to the white which is 255. The principle of encoding uses the Least Significant Bit of each of these bytes, the bit on the far right side. If data is encoded to only the last two significant bits (which are the first and second LSB) of each color component it is most likely not going to be detectable; the human retina becomes the limiting factor in viewing pictures [8].For the sake of this example only the least significant bit of each pixel will be used for embedding information. If the pixel value is 138 which is the value 10000110 in binary and the watermark bit is 1, the value of the pixel will be 10000111 in binary which is 139 in decimal. In this example we change the underline pixel. [9].

Features of LSB (Least-Significant-Bit)

 ISSN: 2454-4248
53 – 56

a. It is simple to understand
b. Easy to implement
c. It results in stego-images that contain hidden data yet appear to be of high visual fidelity

V. SYATEM ARCHITECTURE

Watermarked image database information is shared among image owner and users. Using these information, watermark insertion and extraction is done at image owner and users level. Robustness is considered while developing this system model. We are performing strong literature survey which helps in developing robustness and privacy of retrieval using various algorithms, at the same time deal with the problems. When image users get search results, watermark extraction get performed for verifying that search results are satisfactory or not. If it is not then alert message is sent to appropriate image owner. These watermarks are get embedded in image by image owner. Watermarking information is shared among image users and image owner which is responsible for maintaining privacy.

1. Image owner embeds watermark images, index and saves them on server along with user authentication information.
2.Image users generate trapdoors and then fire required query. According to query, server reply with proper results.
3. Watermark extraction algorithm is used for verification purpose.
4. If the results are satisfactory to image users, process terminates otherwise alert message is sent back to image owner.

This system is useful in similar face finding or similar pattern finding mechanism, as kNN algorithm is used for similarity matching. K nearest neighbor algorithm is used to map which images form database are closely similar to input query image. It is also useful in finding geographical information as per image users requirements, and if no requirements match then image user have rights to send acknowledgement to cloud. In medical image database, for grouping patients according to their health issues, this criteria used. That means the patients having similar health problems are grouped together. After some time if any doctor want to find that patients report, the he will fire the query along with its specification for getting results.

A. Steps to perform Watermark Embeding:
1. Generate the watermark image for each image class.
2. For each image in the database, embed the watermark using LSB watermarking.
3. Prepare the watermarked image database.

B. Steps to perform offline database of features and training:
1. Select the watermarked image database.
2. For each image in the watermarked database, perform the noise removal using median filter.
3. Extract the gabor features from the images
4. Train KNN classifier using the extracted features.

C. Steps to perform image recognition:
1. Select the test image from watermarked image dataset.
2. Extract the watermark from the test image using LSB method.
3. Perform noise removal from test image using median filter.
4. Extract gabor features from test image.
5.Use the KNN classifier for recognizing similar image from the watermarked image dataset

D.Steps to perform Image authentication:
1. Extract watermark from recognized image using LSB method.
2. Match the extracted image watermark with the watermark from test image using SSIM image similarity matching.
3. If the watermarks are matched, then image is authenticated.

VI. EXPERIMENTAL EVALUATION

The experiments are conducted on the COIL100 image dataset[10]. COIL100 image is successfully processed at each step in image retrieval system and the desired result is obtained. Watermarked image database is prepared from COIL 100 image dataset. LSB watermarking is used for watermarking. This watermarked image dataset is used for feature extraction, training and retrieval. The watermarked images are undergone through median filtering for noise removal. Gabor features are extracted from these preprocessed images. Using these features, KNN classifier is trained. At the other side, for recognition, test image is selected. After applying median filter on test image, gabor features are extracted from the test image. Using KNN classification, the matching image is extracted from the dataset. From this recognized image, watermark is extracted. This extracted watermark is matched against the watermark of the test image. If the watermarks are matched, then the retrieved image is treated as authenticated. Obtained results are shown in Fig. 1

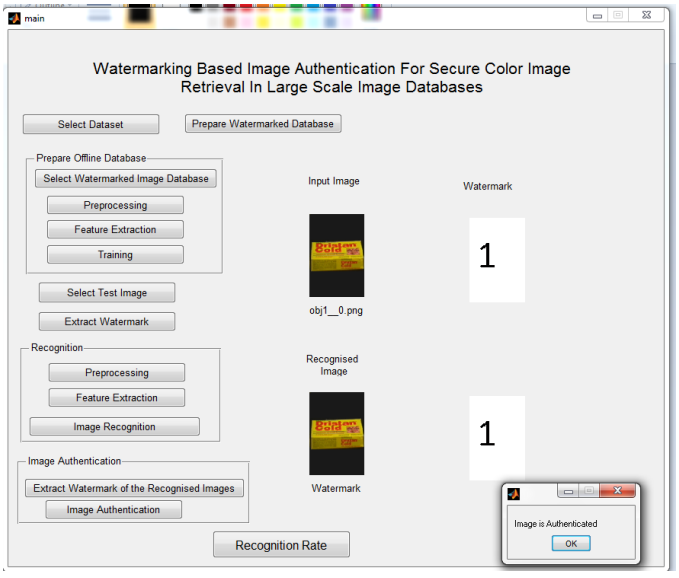

Figure 1: Watermarking based image authentication for CBIR

The test image is classified as one of the appropriate category

 ISSN: 2454-4248
53 – 56

from 100 COIL dataset. Using KNN we get the accuracy of 100 % for image recognition and authentication.

VII. CONCLUSION

In this paper, we presented an overview of our proposed approach to detect faked images. We believe that with combining CBIR with emerging watermarking concepts the fake detection problem in the large scale community image sharing platforms can be solved. Prototype to prove the concepts described above is implemente. In this paper, watermark embeding on images is done before outsourcing and with copy deterrence , aceess from unauthorized users are get prevented. The Performance of this system based on factors like how efficiently the watermark is get embedded and extraxted from the image and the rate of verification. Comparitively, previous systems does not have verification and audit generation. Because of this unavailability, the accuracy of result is less. This system focus on improving the result accuracy with minimum CPU utilization.In short, Efficiency of project can be calculated based on

1. Time consumption of the watermarked image embedding
2. Time consumption of the feature extraction and training.
3. Time consumption of the search operation.
4. Time consumption of the watermark extraction..
5. Storage consumption of the watermarked image database
6. Watermark extraction accuracy

In the future work we plan also to investigate on more types of fake issues. We also intend to test some combinations of different feature sets, not only to detect the manipulated images but also to detect the similar fake.

ACKNOWLEDGMENT

I would like to express greatfulness to P.G.Dept. of Electronics and Telecommunication, J T Mahajan College of Engineering,Faizpur.

REFERENCES

[1] Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun and Kui Ren, A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing, IEEE TRANSCATION ON INFORMATION FORENSIC AND SECURITY, vol.11, 2016, pp. 2594 - 2608. [2] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, Towards privacy-preserving content-based image retrieval in cloud computing, IEEE Transactions on Cloud Computing, vol. PP, no. 99, 2015, pp. 1-1.

[2] Z. Qin, J. Yan, K. Ren, C. W. Chen, and C. Wang, Towards efficient privacy-preserving image feature extraction in cloud computing, in ACM International Conference on Multimedia. ACM, 2014, pp. 497-506.

[3] Zhihua Xia, Xinhui Wang, Liangao Zhang, Zhan Qin, Xingming Sun, and Kui Ren, A Privacy-preserving and Copy-deterrence Content-based Image Retrieval Scheme in Cloud Computing, IEEE TRANSCATION ON INFORMATION FORENSIC AND SECURITY, VOL. , NO. , SEPTEMBER 2016.

[4] K. Gopalakrishnan, N. Memon, and P. L. Vora, Protocols for watermark verification, IEEE MultiMedia, no. 4, pp. 66-70, 2001.

[5] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, Color and texture descriptors, IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, 2001, pp. 703-715.

[6] N. Cao, C. Wang, M. Li, K. Ren, and W. Lou, Privacy-preserving multikeyword ranked search over encrypted cloud data, IEEE Transactions on Parallel and Distributed Systems, vol. 25, no. 1, 201, pp. 222- 233.

[7] S. Rane and P. T. Boufounos, Privacy-preserving nearest neighbor methods: comparing signals without revealing them, IEEE Signal Processing Magazine, vol. 30, no. 2, 2013, pp. 18-28.

[8] Robert, L., and T. Shanmugapriya, “A Study on Digital Watermarking Techniques ”, International Journal of Recent Trends in Engineering, vol. 1, no. 2, pp. 223-225, 2009.

[9] H.Arafat Ali, “Qualitative Spatial Image Data Hiding for Secure Data Transmission”, GVIP Journal,Volume 7,Issue 2 , pages 35

[10] http://www.cs.columbia.edu/CAVE/software/softlib/coil-100.php