ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Content Based Image Authentication and Retrieval

Deepashree Deshmukh[1], Divya Holkar[1], Sayli Mangle[1], Nikita Navale[1], Pranjali Kuche[2]

B. E Students, Department of Information Technology, MMCOE, Karvenagar Pune, Savitribai Phule Pune University, Pune, India.

Professor, Department of Information Technology, MMCOE, Karvenagar Pune, Savitribai Phule Pune University, Pune, India

ABSTRACT: Perceptual image hash has been widely investigated in an attempt to solve the problems of image content authentication and content-based image retrieval. This paper presents a perceptual image hashing algorithm which can distinguish maliciously attacked images from authentic ones. We combine statistical analysis methods and visual perception theory to develop a real perceptual image hash method for content authentication. We generate robust perceptual hash code by combining image-block-based features and key-point-based features. A secure hashing scheme is developed for image authentication and finding the forgery regions. The hash can be used to find similar, forged, and different images. The hash is very sensitive to malicious tampering. The hash of a test image is compared with the reference image. The proposed method achieves a tradeoff between perceptual robustness to tolerate content-preserving manipulations and a wide range of geometric distortions and perceptual sensitivity to detect malicious tampering. Furthermore, it has the functionality to detect compromised image regions. Moreover, compared with some other image hashing algorithms, the proposed approach also achieves better performance even in the aspect of robustness, which is more important in some image hashing applications.

KEYWORDS: Perceptual image hash, content authentication, tampering detection, tampering localization.

I. INTRODUCTION

With the widespread use of image editing software, more and more digital media products are easily available to distribute illegal copy. With the advances in multimedia and networking technologies, it has become easy to copy original material completely and distribute illegal copies rapidly over the Internet. In order to trace the unauthorized use of digital contents, media hashing technologies have been applied to digital content management.

a. Background

Unfortunately, it is the sensitivity that makes these functions not applicable to digital images. Since images will also be considered as the identical one even if they have undergone some content preserving manipulations, such as image compression, noising, and filtering. Image hashing is a technique that extracts a short sequence from the image to represent its contents, and therefore can be used for image authentication. . Consequence, it is expected that images which look like the same or very similar should have the same or very similar hash codes, while images which differ from each other should have distinct hash codes. Perceptual image hashing has been therefore presented to provide the content-based authentication, copyright verification and some other protections for digital images. The core idea of perceptual image hashing is to construct the hash by extracting characteristics of human perception in images, and use this constructed hash to authenticate or retrieve an image without considering the various variables or formats of this image. Recently, perceptual image hash has been developed as a frontier research topic in the field of digital media

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

content security and multimedia applications. The generation of a perceptual image hash is based on well-designed image features that are in accordance with the perceptual characteristic of the human visual system.

b. Motivation

Encryption and randomization are used to create the last hash code. The proposed structure gives a better content authentication protocol than the existing ones. This is possible by creating a perceptual picture hash. Perceptual picture hash has been proposed as a primitive technique to take care of issues of picture content authentication. The perceptual picture hash is created by utilizing the perceptual components that are as per a human's visual qualities. It is required to have the capacity to survive unexpected contortion and reject malignant altering. Therefore it gives a more productive way to deal with examining the changes made in a picture.Hash functions are essential mathematical tools that are used to translate data of arbitrary size into a fixed sized output. There are many different kinds of these functions, each with their own characteristics and purpose. For example, cryptographic hash functions can be used to map sensitive information into hash values with high dispersion, causing even the slightest changes in the source information to produce wildly different hash results. Because of this, two cryptographic hashes can (usually) only be compared to determine if they came from the exact same source. We cannot however measure the similarity of two cryptographic hashes to ascertain the similarity of the sources. Perceptual hashes are another category of hashing functions that map source data into hashes while maintaining correlation. These types of functions allow us to make meaningful comparisons between hashes in order to indirectly measure the similarity between the source data.

II. LITERATURE SURVEY

Sr No	Paper Name	Technique	Advantage/Disadvantage	Point Referred
1.	Cl´eoBaras and Franc¸oisCayre,"2D BAR-CODES FOR AUTHENTICATION: A SECURITY APPROACH", 20th European Signal Processing Conference (EUSIPCO 2012)	A simple estimator of the 2D-BC based on copies averages is proposed, letting the opponent print a fake 2DBCwhich aims at being declared as genuine by the system detector.	Advantage: 1.Create a fake 2D-BCs declared as genuine by the detector. Disadvantage:- 1.Require additional noise to generate fake barcode.	1.Generating fake 2D QR code declared as original by QR code reader.
2.	Thach V. Bui, Nguyen K. Vu, Thong T.P. Nguyen∗,,"Robust Message Hiding for QR Code",2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing	Proposes method for Hiding secret information based onbit technique is so fragile to modification attack.	Advantage:- 1.For attacker difficult to find original data 2. Reed-Solomon codes and List Decoding to overcome this problem.	1.We refer Reed-Solomon codes for encoding content in qr code.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

3.	Tobias Langlotz and Oliver Bimber," Unsynchronized 4D Barcodes Coding and Decoding Time-Multiplexed 2D Colorcodes", Bauhaus-University Weimar	Proposes no direct connection between devices can exist. Time-multiplexed, 2D color barcodes are shown on screen &recorded with camera embed mobile phones.	Advantage:- 1. Maximizes the data throughput and the robustnessof the barcode recognition,	1.We refer Time-multiplexed, 2D color barcodes
4.	Ching-Yung Lin and Shih-Fu Chang," Distortion Modeling and Invariant Extraction for Digital Image Print-and-Scan Process",	Properties of the discretized, rescanned image in both the spatial and frequency domains,then further analyze the changes in the Discrete Fourier Transform (DFT) coefficients.	Advantage:- 1. Authentication by image watermarking. Disadvantage:- 1. Uses watermarking based authentication	1.Discrete Fourier Transform coefficients and image based authentication
5.	Pei-Yu Lin, Yi-Hui Chen," Secret Hiding Mechanism Using QR Barcode",	Concealthe secret data into the cover QR code without distorting thereadability of QR content. That is, general browsers canread the QR content from the marked QR code for the sakeof reducing attention.	Advantage:- 1. Only the authorized receiver canencrypt and retrieve the secret from the marked QR code.	1.Secret hiding mechanism for QR code by encrypted payload.
6.	M. QUERINI, A. GRILLO, A. LENTINI and G.F. ITALIANO,"2D COLOR BARCODES FOR MOBILE PHONES"	use colors to increase the barcode data density. The introductionand recognition of colored modules poses some new and non-trivial computer vision challenges, such as	Advantage:- 1. Prototype for generatingand reading the HCC2D code format, both on desktops	1.we refer 2D COLOR BARCODES

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

		handling the color distortions		
7.	KaushalSolanki, Member, IEEE, UpamanyuMadhow, Fellow, IEEE, B. S. Manjunath,"Print and scan' resilient data hiding in images", IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 1, NO. 4, DECEMBER 2006	quantization indexmodulation scheme embeds information in phase spectrum of images by quantizing the difference in phase of adjacent frequencylocations.	Advantage:- 1. Hiding information intoimages in a manner that is robust to printing and scanning.	1.Referred point is secret hiding in images.
8.	R. Villan, S. Voloshynovskiy, O. Koval, F. Deguillaume, and T. Pun," Tamper-proo¯ng of Electronic and Printed Text Documents via Robust Hashing and Data-Hiding"	two modern text data-hiding methods,namely color index modulation (CIM) and location index modulation (LIM),	Advantage:- 1. robustness against typical intentional/unintentional document Disadvantage: 1. Does not produce consistent results.	1.We refer color index modulation (CIM) and location index modulation
9.	Kuo-Chen Wu and Chung-Ming Wang, Member, IEEE, "Steganography Using Reversible Texture Synthesis", Ieee Transactions On Image Processing Vol: 24 No: 1 Year 2015	By using texture synthesis process into steganography to hide secret messages	Advantage:- 1. Hide secret message to kernel block 2. Block encoding scheme. 3. Index table generation for patch entry	.1.We refer patch based data encoding 2.Reverse texture synthesis for QR coding hiding as secret message.

III. EXISTING APPROACH

To develop a new and more secured way so that we can save and authenticate the user's image data on the basis of perceptual hash. We are proposing this client based application which can be installed on the system and performs operations like Uploading Image to Server, Generating perceptual hash and identify changes made to an image. The application will generate advance perceptual hash and save them at local storage for every image uploaded and Authenticate content and detect forgery. The application on downloading images back will perform content authentication on every images using locally stored hash information. So that we provide the way to keep image data safe as well as find out whether the server is safe or not.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

IV. ALGORITHM USED

1. MD5 ALGORITHM

MD5 belongs to the MD family of hash functions. It is a simple, fast and most widely used hashing method which generates 128 bit hash or message digest. It is the successor of MD4 method of hashing of the same MD family. Some research proved that MD4 is not that much secure and can be easily attacked, so MD5 was invented in 1991. However, in 2004, MD5 was also found to be insecure, as it can produce same hash value for 2 different messages resulting in collision, because of the small length of the generated message digest (128 bits).

2. Scale Invariant Feature Transform(SIFT) Scale-Invariant Feature Transform (SIFT) is an algorithm used to discover and describe local features in images. SIFT can easily identify objects even with noise and under partial obstructions, as the SIFT feature descriptor is invariant to uniform scaling, orientation, brightness changes. SIFT keypoints of objects are discovered from a set of images and stored in a database. An object is recognized in a new image by comparing every feature from the new image to images in the database and finding matching features based on the Euclidean distance of their feature vectors.

V. FLOW OF THE SYSTEM

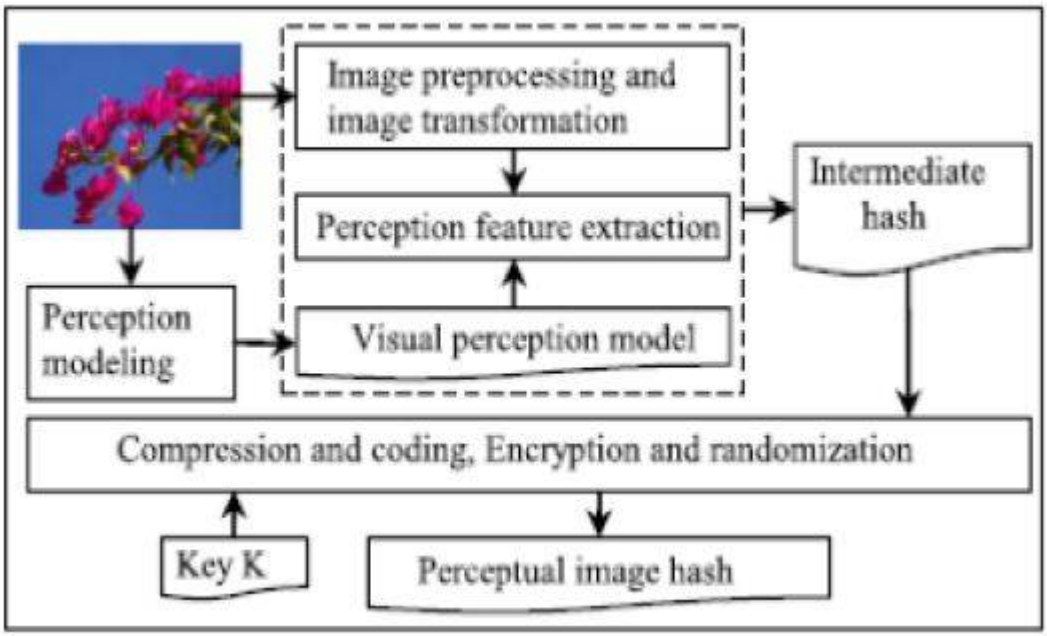

Fig 01 System Flow

The proposed perceptual image hash scheme includes a hash generation algorithm, a tampering detection algorithm, and a tampering localization algorithm. The hash generation algorithm consists of three stages: feature extraction, encryption and randomization, and compression and coding. Considering that the SIFT (Scale Invariant Feature Transform) features are invariant to translation, rotation and scaling transformations in image domain and robust to moderate perspective transformations and illumination variations, we begin our work with the SIFT feature extraction. Regarding image content changing, it is difficult to define a clear boundary between perceptually insignificant distortion and malicious tampering because some content-preserving manipulations such as JPEG compression are

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

lossy. This results in an intriguing question, that is, the trade-off between robustness to tolerate content-preserving manipulations and sensitivity to malicious tampering. In our work, tampering detection and tampering localization are realized by comparing a distance metric to measure the similarity between hash values. Image tampering will cause the change of feature points, that is, if an image hasbeen tampered with, then the feature points detected from the compromised regions will be different from the feature points defined in its original version. As a result, the distance between feature points defined in the original version and those detected will be greater. We can use the distance change to detect changed image regions. That is, an image block can be considered as a tampered region if it contains changed feature points, and the change of feature points can be measured via distances between the original feature points and detected feature points in the corresponding image region. Because distance values are clearly distinguishable, we can estimate the cut-off values of the distance change by using the statistical values that are obtained from experimental results. In general, a perceptual image hash system consists of four stages: image preprocessing and transformation, perceptual Feature extraction and description, compression and coding, and encryption and randomization. The general framework is shown in Fig. 1. The purpose of image preprocessing and transformation is to eliminate irrelevant information, recover useful information and enhance image features that are important in subsequent processing. To ensure perceptual robustness and perceptual sensibility, the selection and extraction of perceptual features are very important. Perceptual feature extraction is based on the human visual perception model that is established by the cognitive science theory. It is accomplished via signal processing methods that remove redundant data but retain perceptually significant features. Moreover, to reduce hash Length and improve convenience for storage and hardware implementation, post-processing such as compression and Coding is necessary. Encryption and randomization are used to reduce hash collisions to improve the security of the algorithm

VI. PROPOSED SYSTEM ARCHITECTURE

In this paper we are proposing content authentication, tampering detection and security of server for the users who wish to save the image data on server like cloud, LAN. So that will make the storage more secure. Currently there is tremendous amount of image data being generated over internet, this data is very vulnerable to tampering, and hence it is necessary to provide an authentication process for image dataset as images are many times used as proofs. Here a real perceptual image hash method is proposed. Based on this hash, an image tampering detection is presented. The proposed method is sensitive to changes caused by malicious attacks, and it achieves a trade-off between robustness against geometric distortion and tampering localization. The proposed method can be used for content-based image authentication and for image retrieval and matching in large-scale image databases. A perceptual image hash is expected to be able to survive unintentional distortion and reject malicious tampering within an acceptable extends .It involves proposing content based image authentication, forge detection. Additionally, it achieves a trade-off between robustness to tolerate tampering. Develop a client based application which can be installed on the system. The application will upload images to cloud server. The application will generate advance perceptual hash and save them at local storage for every image uploaded. The application, on downloading images back will perform content authentication on every images using locally stored hash information.

A. SIFT ALGORITHM WORKING.

1. Create internal representations of the original image to ensure scale invariance. This is done by generating a "scale space".
2. Find key points in an image and approximate it using the representation created earlier.
3. We now try to find key points. These are maxima and minima in the Difference of Gaussian image we calculated in step 2
4. Get rid of bad key points: Edges and low contrast regions are bad key points. Eliminating these makes the algorithm efficient and robust.
5. Assigning an orientation to the key points: An orientation is calculated for each key point. Any further calculations are done relative to this orientation. This effectively cancels out the effect of orientation, making it rotation invariant.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

6. Generate SIFT features: finally, with scale and rotation invariance in place, one more representation is generated. This helps uniquely identify features. Let's say you have 50,000 features. With this representation, you can easily identify the feature you're looking for (say, a particular eye, or a sign board).

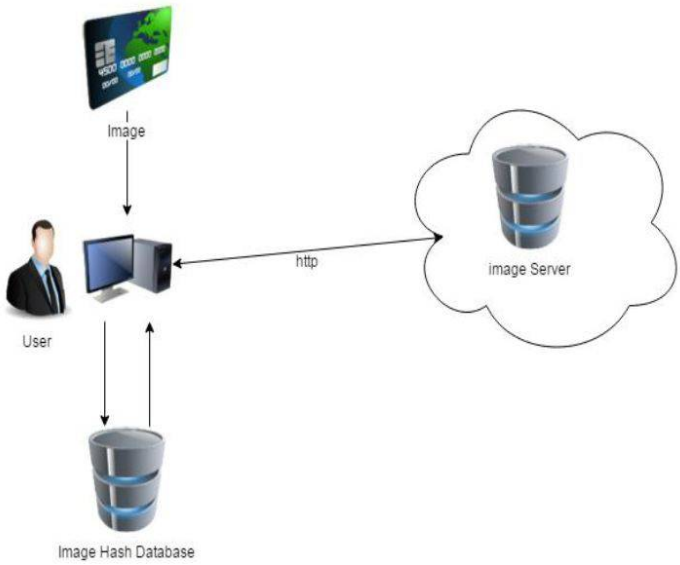

Fig.2. System Architecture

It shows the architecture of the proposed system. Architecture contains User in the mid, who will upload or download image. Whenever user uploads or downloads the image from server,hash will be generated and stored in hash database. So that the comparison of hash values can be done.

VII. IMPLEMENTATION AND DISCUSSION

1. **User Image Upload Module:** This module allows logged user to upload image file to cloud space. It will only allow image based file extension e.g. jpg, jpeg, png, bmp etc. i. Download Module**:** This module includes all functionalities which allows user to select one of the uploaded files and download it. ii. Block Feature Extraction: This module has all functionalities and algorithm required to calculate block wise hash values for input image. First image is split into multiple non overlapping blocks. Features are calculated using SIFT algorithm for each block. This hash is stored at local machine in an index file and not at cloud. iii. Block Feature Matching: Here all these hashes are matched which will be a unique signature for all the images.
2. **Admin module:** It contains functionality to compare two images. It reads the downloaded image first and performs block feature extraction. Using this field it will get previously calculated hashes from index file. Now it will compare the hashes block by block. If all blocks match, then image is authenticated, no modification has taken place on image.
3. **Hacker Module-** The main tampering part is done by the Hacker. The hacker may change the properties of the image by altering a small or large part of the image. It may not even be visible to the human eye. But even slightest change affects the authenticity. The hacker changes some details and uploads the modified image and makes it seem like that was the original picture.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Table No: 1. Sensitivity Analysis and Performance Comparison

Parameter	Proposed	Existing
Tampering rate	60	70
Hash generation	72	60

The tampering manipulations include adding image objects, copy-move attack, object replacement attack, hiding image object attack. Various tampering methods were put to use for the purpose of checking the reliability and accuracy of this system. A large database of image files was used for testing. The system can easily detect most alterations made in the images. Changes that may not be visible to the naked eye were also detected by this system. It could accurately display the sections of the image where tampering was done.

Fig. 3. Detection results for object replacement attack.

The detected results are indicated by displaying only the part of the image on which tampering was done . As shown by the experimental results, the proposed method can detect the locations of compromised image regions. It is valid for detecting compromised images that have undergone geometric distortions.

VIII. CONCLUSION

From the consideration of all the above points we conclude that, a genuine perceptual picture hash strategy is proposed. Taking into account this hash, a picture altering discovery what's more, altering restriction system is displayed. As a device for picture content verification, the proposed technique is sensitive to geometric distortions and content preserving methods, for example, JPEG pressure, including sifting, and others. It can easily detect changes created by destructive assaults, and it accomplishes an tradeoff between strength against geometric contortion and altering limitation. The experiments performed by the authors demonstrate the adequacy and the accessibility of the proposed experiment for various altering assaults at three execution levels: image tampering detection (detection precision), promised region localization (visual impact), and localization accuracy (recognition rate at the pixel level). The proposed system can be utilized for content based picture authentication and for picture recovery and coordinating in large scale image databases.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

REFERENCES

1. C. Baras and F. Cayre, "2D bar-codes for authentication: A security approach," in Proc. 20th Eur. Signal Process. Conf. (EUSIPCO), Aug. 2012, pp. 1760–1766.
2. T. V. Bui, N. K. Vu, T. T. P. Nguyen, I. Echizen, and T. D. Nguyen, "Robust message hiding for QR code," in Proc. IEEE 10th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process. (IIH-MSP), Aug. 2014, pp. 520–523.
3. T. Langlotz and O. Bimber, "Unsynchronized 4D barcodes," in Proc. 3rd Int. Symp., ISVC 2007, Lake Tahoe, NV, USA, Nov. 26–28, 2007, pp. 363–374.
4. C.-Y. Lin and S.-F. Chang, "Distortion modeling and invariant extraction for digital image print-and-scan process," in Proc. Int. Symp. Multimedia Inf. Process., 1999, pp. 1–10.
5. P.-Y. Lin, Y.-H. Chen, E. J.-L. Lu, and P.-J. Chen, "Secret hiding mechanism using QR barcode," in Proc. IEEE Int. Conf. Signal-Image Technol. Internet-Based Syst. (SITIS), Dec. 2013, pp. 22–25.
6. M. Querini, A. Grillo, A. Lentini, and G. F. Italiano, "2D color barcodes for mobile phones," Int. J. Comput. Sci. Appl., vol. 8, no. 1, pp. 136–155, 2011.
7. K. Solanki, U. Madhow, B. S. Manjunath, S. Chandrasekaran, and I. El-Khalil, "'Print and scan' resilient data hiding in images," IEEE Trans. Inf. Forensics Security, vol. 1, no. 4, pp. 464–478,Dec. 2006.
8. R. Villán, S. Voloshynovskiy, O. Koval, F. Deguillaume, and T. Pun, "Tamper-proofing of electronic and printed text documents via robust hashing and data-hiding," in Proc. SPIE, vol. 6505, p. 65051T, Feb. 2007.
9. Kuo-Chen Wu and Chung-Ming Wang, Member, IEEE, "Steganography Using Reversible Texture Synthesis", Ieee Transactions On Image Processing Vol: 24 No: 1 Year 2015
10. S. V. Voloshynovskiy, O. Koval, F. Deguillaume, and T. Pun, "Visual communications with side information via distributed printing channels: Extended multimedia and security perspectives," Proc. SPIE, vol. 5306, pp. 428–445, Jun. 2004.