Copyright Infringement Detection of Multimedia Objects using K–L Transform

Stuti Pandey[1], Nihar Ranjan Pradhan[1], Akhilendra Pratap Singh[1], Dharmender Singh Kushwaha[2]
Department of Computer Science and Engineering, National Institute of Technology Meghalaya[1]
Department of Computer Science and Engineering, Motilal Nehru National Institute of Technology Allahabad[2]
p21cs011@nitm.ac.in[1], niharpradhan@nitm.ac.in[1], akhilendra.singh@nitm.ac.in[1], dsk@mnnit.ac.in[2]

2022 4th International Conference on Energy, Power and Environment (ICEPE) | 978-1-6654-7124-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICEPE55035.2022.9798179

Abstract— **The copyright infringement is a critical issue for original copyright owners of multimedia objects like videos, photographs and graphic designs which are uploaded on various peer-to-peer image sharing systems and websites. Copyright infringement negatively impacts copyright holders' interest by performing some extent of modifications in the original uploaded content. Various frame modification and image processing techniques like frame compression, frame speed manipulation, grayscale conversion, cropping and rotation are used by the copyright violators for modification of the original multimedia content. This paper proposes a perceptual hashing technique using Karhunen-Loeve Transform (K-L Transform) for detecting the copyright violations of the multimedia content, so the work acknowledgement should be provided to the original copyright holders.**

Keywords— copyright, infringement detection, multimedia objects, K–L transform

I. Introduction

In the digital era of Internet, copyright violation is one of the major problems faced by the photographers, video providers, graphic designers who sell their content through stock photo and video agencies. The definition for the term copyright violation is: "without being the original author getting acknowledged for the content by just making little changes in the video or images" [12].

There is a continuous expansion in the distribution of data which comes in the formats of video, audio, image, texts on information-sharing portals and websites [9, 14]. In the digital platform of Internet, they are frequently vulnerable to stealing and tempering [8] which is a crucial concern in digital forensics as copyright violation.

Somebody may use the data for purposes like getting the economic benefits [22] which can actually be get secured in some of the form. Under the digital forensic copyright protection act which remains in different form globally, using other's data without their agreement is a criminal offense. Thus, different practices are getting formed as a solution by the researchers worldwide to this existing problem, which in real scenario has accomplished few success [20].

Some conventional cryptographic hash matching methods like SHA-1 and MD5 are used for multimedia content matching by certain websites to ignore the copyright violation with the existing database record. But they suffer from avalanche effect. Thus, certain changes in the original content would come out as an extreme value change in the resultant hash [16]. So, if there is a slight change made in the multimedia content like images and video keyframes, the conventional cryptographic hashing methods are less likely to identify them. This will totally avoid the exposure of copyright violation [3].

For this reason, perceptual hash function is a major solution to the appeared problems in copyright infringement detection. The feature it holds is the correspondence between the hash of the marginally tempered input and the original input [13]. So, the perceptual hash is effective in recognizing the modified images and videos frames which in real does not change their appearance.

In this paper, we propose an approach to figure out the copyright infringement issues for multimedia objects using K–L Transform. K–L Transform has some excellent properties which makes it different from other transforms. These properties are presented in figure 1.

The next section describes the related works. The proposed methodology for copyright infringement is described in section 3. At the last, section 4 describes the conclusion and future scope of the proposed methodology.

II. Related Work

Copyright infringement is the main focus of this part along with the previous studies (state-of-the-art) and approaches. To extract the visible features, Watson's model has been used in [18] which combined the approaches of image-block-based and key-point-based features and results the computation of perceptual hash. The method is useful in precisely detecting the locally tempered regions and plays a crucial role in perceiving the image content. The hash length soars to tens of thousands is the major shortcoming of the model.

In [22], based on the shape of the image and Hu invariants moments, perceptual image hashing is given. But the Hu invariant moments are not convenient for translation, rotation and scale kind of operations in transformation matching.

Discrete Cosine Transform (DCT) technique is used in [20] which illustrates an approach based on perceptual hash in image for feature extraction. Matching the perceptual hash of target object frame and Hamming distance is minimized at the same time in the approach to get the similarity in the nearest frame.

The methods given in [3] converts the RGB image into grayscale image which gets divided then into 4×4 and 8×8 blocks. Then the 2D coefficients is calculated for all the blocks. Also, the Euclidean distance is used for identification of similar kind of blocks. Due to the discard of the high frequency of pixel values, this approach is not adequate.

The work in [15] is to extract texture features for the encoding of a dual-cross pattern (DCP). In this salient structural feature is applied for calculation of image hash and

reduction of texture features. For classification purpose of the tempered image, the derived hash is secure but not desirable.

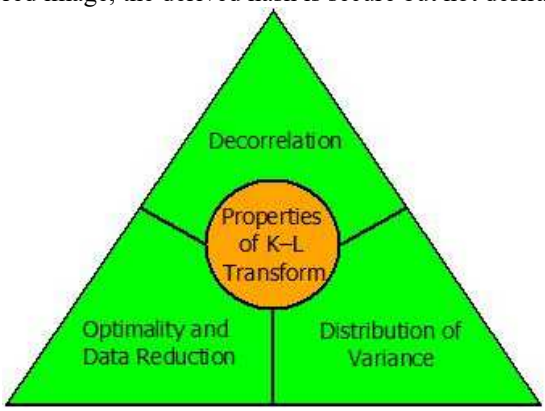

Figure 1. Properties of K–L Transform

For detection of forgeries, an image's hue modification has been used in [7] which is based on Photo Response Non-Uniformity (PRNU) approach. The tempering localization detection i.e., detection of tempered region is the more focus point in a real scenario.

In [1, 4] IPFS is used for watermarking details storage of an image. The blockchain-based decentralization concepts are used here which applies a cryptographic hash technique for generation of the digital signature and to establish the content authentication of an image. The avalanche effect can affect the conventional cryptographic approaches and thus the perceptual feature extraction of image cannot be done properly. The image data with the corresponding hash values are stored in blockchain structure but still the concept is lagging behind in authentication of the duplicate and modified images on the IPFS and blockchain structure.

The authors in [5] proposed the Least Significant Bit (LSB) techniques benefit for authentication of content of an image. But when the noise has been added in the image, it fails to perceive the image exactly and it is appropriate only for grayscale images.

To secure the attributes of an image correctly, a Discrete Fourier transform (DFT) is used in [19] for perceptual hash generation. The resultant hash is of 144 bits and it is actually larger than the original perceptual hash of 64 bits.

For identifying the similar images, a histogram-based perceptual hash (fingerprint) is used in [6]. In this work, for video frame detection the same algorithm is applied i.e., the frame which is most similar with the destined frame. Also, the way creates a hash of variable length which is based on length of the videos.

For the keyframe extraction purpose, a high dimensional space is constructed in [11]. Representation of the individual frame can be calculated in an efficient manner by the keyframes. For a video sequence, the keyframe can lose the dynamic effect but the approach of keyframe selection can either represent or make a matching of the originally generated video content.

The authors in [21] has used the perceptual hash technique for computation of the frame's hashes. In the work, an edge matching also known as brightness technique is used to discard the extra frames of a video, though the technique is sensitive for the frame's brightness. It is a golden section method to get the candidate frames and for frames hash values matching, the Hamming distance is utilized. Here the frame's entropy is set to local maximum.

Scale Invariant Feature Transform (SIFT) approach is used in [2, 10, 23, 24] in an image to identify the analogous features, although it only identifies the region of the image i.e., partial-duplicate image detection. It also fails in image matching when a lot feature changes occur and lags in identification of discriminating pixels when occurred a blurred operation.

III. Proposed Methodology

The main objective of the proposed methodology is to propose a secure distributed network framework where copyright infringement for multimedia objects can be mitigated effectively. Figure 2 presents the proposed framework. The proposed framework is divided into two parts to redress the problem of copyright infringement.

Uploading of multimedia objects: In this part, any person can upload their images and videos with its details. They do so to get the credits of their work.

Generation and matching of Perceptual hash: Proposed framework generate perceptual hash of the uploaded multimedia objects in this part. Matching of generated hash function with the existing hash function is also performed here.

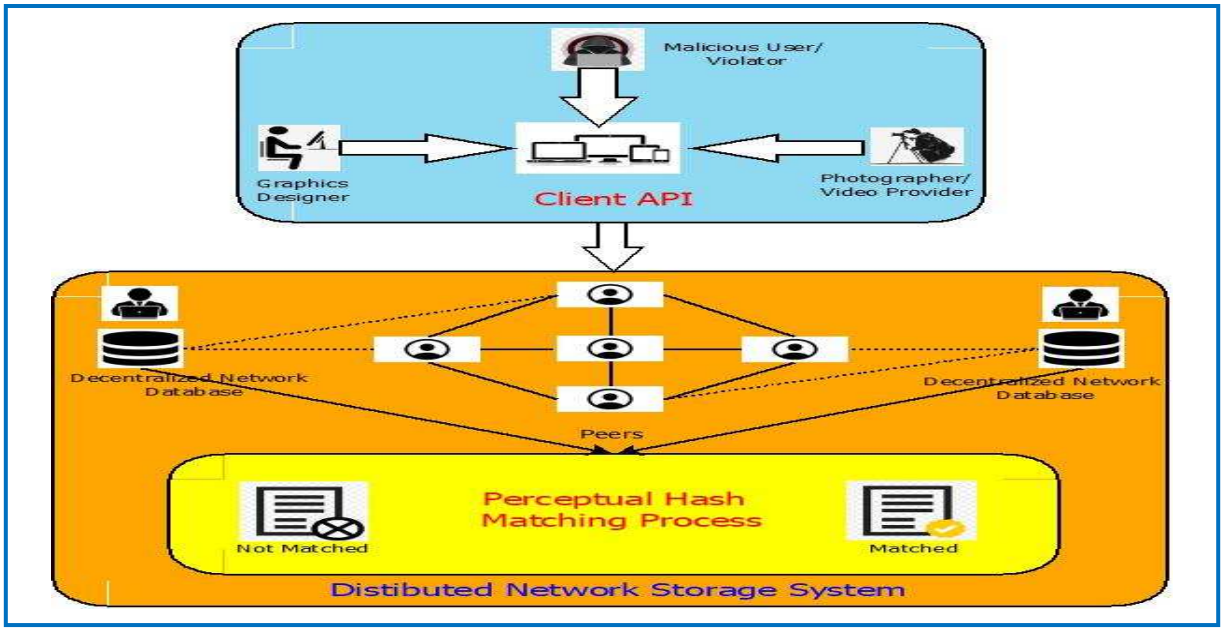

Figure 2. Proposed Framework for Copyright Infringement

The working of proposed methodology for copyright infringement detection is presented with help of flowchart in figure 3.

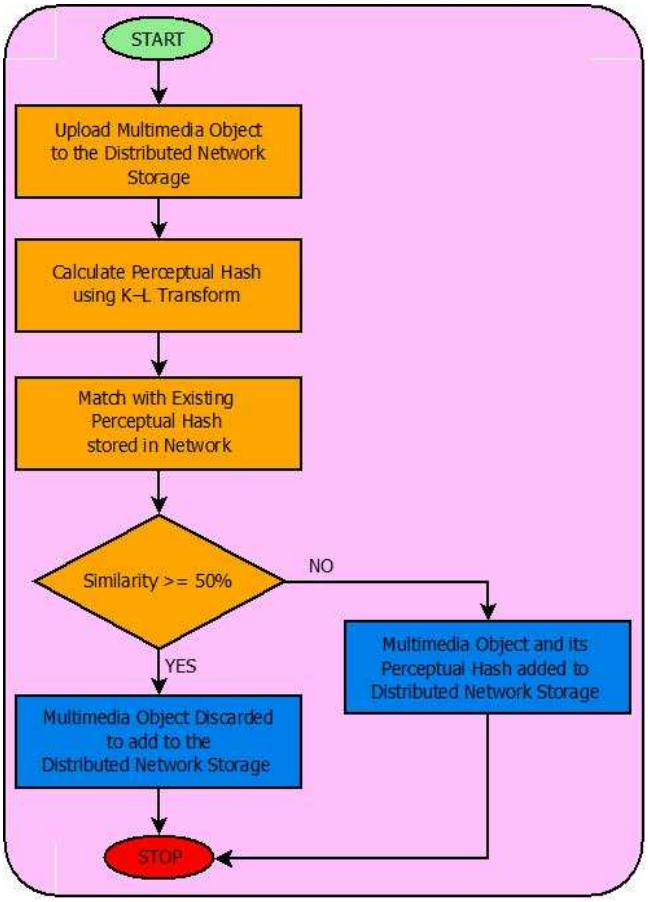

Figure 3. Flowchart of Proposed Methodology

The steps involved in proposed methodology are:

1. In the first step, uploading of multimedia objects to the distributed network is performed by users. This user can be anyone who wants to upload their genuine work. If not, they can be a copyright violator who modifies the images by using various frame modification and image processing techniques like frame compression, frame speed manipulation, grayscale conversion, cropping and rotation. The aim of this malicious user is to claim his ownership on marginally modified images.
2. The second step is the calculation of Perceptual Hash of the multimedia objects by the system using K–L Transform.
3. In this step, matching of Perceptual Hash is performed with existing hash function of image/video which is already in the storage.
4. After matching of Perceptual Hash, if similarity of Perceptual Hash is more than or equal to fifty percent, then copyright violation is found and the object is rejected for uploading to the distributed network.
5. Contrarily, if the similarity of Perceptual Hash is less than fifty percent, then uploaded multimedia objects will be added to the distributed network storage. The Perceptual Hash of the uploaded object is also added to the network as metadata. Proposed methodology has considered 50% similarity as threshold value due to existing work [17] for evaluation of text similarity.

A. The Desirables for Image Transforms

The reason to choose K–L Transform for calculating Perceptual Hash is its excellent properties that is shown in figure 1. Calculating perceptual hash needs image transformation which can be done by many different ways. Therefore, a good image transformation technique has some desirables which are shown in figure 4. The K–L Transform also fulfills these desirables for image transforms.

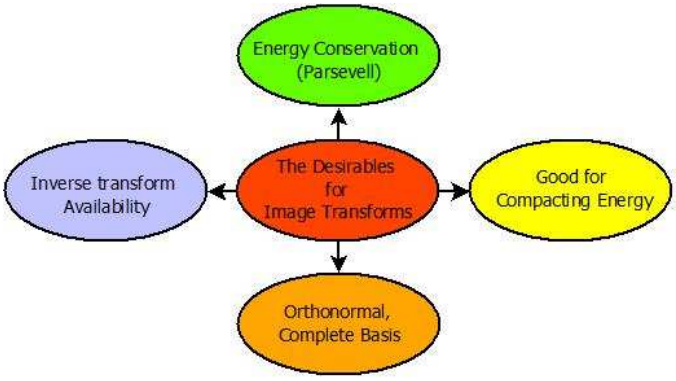

Figure 4. Desirables for Image Transforms

B. 2–D K–L Transform of Images

For a 2–D random process {x(m, n)}, m, n ∈ [0, N − 1] which has zero mean, the 2–D KL transform of image matrix x can be calculated as:

$$X = \Psi_1^{*t} \, x \, \Psi_2^{*} \qquad (1)$$

where Ψ_1^{*t} and Ψ_2^{*} are 1–D KL matrices applied to columns and rows of the image, respectively. The inverse KL transform can be calculated as:

$$x = \Psi_1 \, X \, \Psi_2^{t} \qquad (2)$$

The eigen-images (or basis images) of 2–D KL transform, are $K(k,l) = \xi_{1k}\xi_{2l}^{t}$, (k, l) ∈ [0, N − 1] where ξ_{1k} is kth column of Ψ_1 and ξ_{2l}^{t} is lth row of Ψ_2^{t}.

Image x is decomposed as a linear combination of these eigen-images with the KL coefficients, X(k, l)s, i.e.

$$x_{k,l} = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} X(k,l)K(k,l) \qquad (3)$$

However, this requires finding two KL matrices Ψ_1 and Ψ_2.

C. Perceptual Hash using K–L Transform

The generation of Perceptual Hash involves following steps:

(a) In the first step, original image size is reduced into 8 × 8 pixels.

(b) The second step is grayscale conversion of image.

(c) In this step, K–L Transform (KLT) and calculation of KLT coefficient using equation (3) is performed.

(d) In the fourth step, computation of average value a is performed using equation (4)

$$a = \frac{1}{64} \sum_{r=1}^{8} \sum_{c=1}^{8} X_{r,c} \qquad (4)$$

(e) Using mean value a, comparison of 64 KLT coefficients is performed. 1 and 0 is placed for higher and lower intensity respectively. It is mentioned in equation (5).

$$h_i = \begin{Bmatrix} \{0, x_{k,l} < a\}, \{1, x_{k,l} \geq a\}, \\ \forall \, (k,l) \in [0,63] \end{Bmatrix} \qquad (5)$$

(f) In the last step, 64 bits perceptual hash is generated by applying equation (5).

The obtained perceptual hash is used to compare the similarity of two different images. For this purpose hamming distance (H_d) is used which is mentioned in equation (6).

$$H_d = \sum_{i=1}^{8} \sum_{j=1}^{8} \left| h_{i,j}^1 - h_{i,j}^2 \right| \quad (6)$$

IV. Conclusions And Futute Scope

Copyright infringement is a leading issue in reference to multimedia objects and its detection is very challenging. In this paper, proposed work presents an approach for copyright infringement by perceptual hashing using K–L transform. In future, blockchain and IPFS will be employed for sharing of multimedia objects among peers. To avoid copyright infringement, perceptual hashing using K–L transform will be used.

References

[1] F. Ahmad, L.-M. Cheng, Authenticity and copyright verification of printed images, Signal Process. 148 (2018) 322–335.

[2] S.L. Al-khafaji, J. Zhou, A. Zia, A.W.-C. Liew, Spectral-spatial scale invariant feature transform for hyperspectral images, IEEE Trans. Image Process. 27 (2) (2017) 837–850.

[3] M.H. Alkawaz, G. Sulong, T. Saba, A. Rehman, Detection of copy-move image forgery based on discrete cosine transform, Neural Comput. Appl. 30 (1) (2018) 183–192.

[4] J. Benet, IPFS-content addressed, versioned, P2P file system, 2014, arXiv preprint arXiv:1407.3561.

[5] U.K. Das, S.G. Samaddar, P.K. Keserwani, Digital forensic enabled image authentication using least significant bit (LSB) with tamper localization based hash function, in: Intelligent Communication and Computational Technologies, Springer, 2018, pp. 141–155.

[6] K. Hamon, M. Schmucker, X. Zhou, Histogram-based perceptual hashing for minimally changing video sequences, in: 2006 Second International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution, AXMEDIS'06, IEEE, 2006, pp. 236–241.

[7] J.-U. Hou, H.-K. Lee, Detection of hue modification using photo response nonuniformity, IEEE Trans. Circuits Syst. Video Technol. 27 (8) (2016) 1826–1832.

[8] C. Iwendi, Z. Jalil, A.R. Javed, T. Reddy, R. Kaluri, G. Srivastava, O. Jo, Keysplitwatermark: Zero watermarking algorithm for software protection against cyber-attacks, IEEE Access 8 (2020) 72650–72660.

[9] M. Kamal, G. Srivastava, M. Tariq, Blockchain-based lightweight and secured V2V communication in the internet of vehicles, IEEE Trans. Intell. Transp. Syst. (2020).

[10] S.-A. Li, W.-Y. Wang, W.-Z. Pan, C.-C.J. Hsu, C.-K. Lu, FPGA-based hardware design for scale-invariant feature transform, IEEE Access 6 (2018) 43850–43864.

[11] X. Li, B. Zhao, X. Lu, Key frame extraction in the summary space, IEEE Trans. Cybern. 48 (6) (2017) 1923–1934.

[12] Z. Meng, T. Morizumi, S. Miyata, H. Kinoshita, Design scheme of copyright management system based on digital watermarking and blockchain, in: 2018 IEEE 42nd Annual Computer Software and Applications Conference, COMPSAC, Vol. 2, IEEE, 2018, pp. 359–364.

[13] X.-m. Niu, Y.-h. Jiao, An overview of perceptual hashing, Acta Electron. Sin. 36 (7) (2008) 1405–1411.

[14] D. Połap, G. Srivastava, K. Yu, Agent architecture of an intelligent medical system based on federated learning and blockchain technology, J. Inform. Secur. Appl. 58 (2021) 102748.

[15] C. Qin, X. Chen, X. Luo, X. Zhang, X. Sun, Perceptual image hashing via dual-cross pattern encoding and salient structure detection, Inform. Sci. 423 (2018) 284–302.

[16] M. Steinebach, Robust hashing for efficient forensic analysis of image sets, in: International Conference on Digital Forensics and Cyber Crime, Springer, 2011, pp. 180–187.

[17] O. Uzuner, R. Davis, B. Katz, Using empirical methods for evaluating expression and content similarity, in: 37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the, IEEE, 2004, pp. 8–pp.

[18] X. Wang, K. Pang, X. Zhou, Y. Zhou, L. Li, J. Xue, A visual model-based perceptual image hash for content authentication, IEEE Trans. Inf. Forensics Secur. 10 (7) (2015) 1336–1349.

[19] L. Weng, B. Preneel, A secure perceptual hash algorithm for image content authentication, in: IFIP International Conference on Communications and Multimedia Security, Springer, 2011, pp. 108–121.

[20] S. Yu, Z. Jiang, Visual tracking via perceptual image hash from a mobile robot, in: 2015 IEEE International Conference on Information and Automation, IEEE, 2015, pp. 1612–1616.

[21] M. Zhang, L. Tian, C. Li, Key frame extraction based on entropy difference and perceptual hash, in: 2017 IEEE International Symposium on Multimedia, ISM, IEEE, 2017, pp. 557–560.

[22] B. Zhang, Y. Xin, X.-X. Niu, Image perceptual hash algorithm based on target character, in: 2011 IEEE 13th International Conference on Communication Technology, IEEE, 2011, pp. 397–401.

[23] J. Zhao, N. Zhang, J. Jia, H. Wang, Digital watermarking algorithm based on scale-invariant feature regions in non-subsampled contourlet transform domain, J. Syst. Eng. Electron. 26 (6) (2015) 1309–1314.

[24] Z. Zhou, Y. Wang, Q.J. Wu, C.-N. Yang, X. Sun, Effective and efficient global context verification for image copy detection, IEEE Trans. Inf. Forensics Secur. 12 (1) (2016) 48–63.