Robust Perceptual Video Hashing Based on Histogram Equalization and Sift

D.R. Denslin Brabin*, M. Maharajan
Department of Computer Science and Engineering, Kings Engineering College, India
***Corresponding author: E-Mail: denscse@gmail.com**

ABSTRACT

Securing data has become more important especially in the applications related to digital forensics, military, medical etc. Various rising technologies are used for providing integrity of data. In the past decades attack against integrity of digital video content during transmission is increased. Scalable video code stream contains the base layer and enhancement layer. In this paper a method of Key Frame Extraction and Histogram equalization is used to produce improved video frames and then use a method called Scale Invariant Feature Transformation (SIFT) to extract the visual features from the enhanced frames. The visual features are used to create the hash code of all the frames. Finally combine all these hash codes into single hash code which can be stored or transferred to check the integrity of the video.

KEY WORDS: Key Frame Extraction, Color Histogram Equalization, SIFT, SHA.

1. INTRODUCTION

Network security is a mean to protect the network from unlawful access and security violations. The role of network administrators is to deploy security measures to prevent security attacks and potential security threats against their networks. Computer networks that are used in normal transactions and communication within the individuals, organization, institutions or government require security. The most simple and common way of preventing unauthorized access is by assigning it a unique name or identifier and a password. This simple technique is not enough to preserve the integrity of any important data. Security services enhance the security of the information transfers of an organization and the data processing systems. There are five categories of services according to X.800 (Stalling, 2011): Authentication, Access Control, Data Confidentiality, Data Integrity, and Nonrepudiation.

Data integrity is an essential characteristic of network security and storage security. With the arrival of new technologies, challenges arise in ensuring data integrity during storage and transmission of data. Data Integrity means providing assurance to the receiver that the message received is exactly same as sent by the sender. That is the message is not modified by adversary or otherwise. Integrity, in terms of data security, is nothing but the guarantee that data is not accessed or modified by unauthorized parties. In simple word, it is process of verifying originality of data. As data integrity gives the guarantee that data is of high quality, correct and unmodified.

This paper deals with creation of hash code to check the integrity of video that is transferred from the sender to the receiver. The receiver checks the hash code by repeating the process that is performed by the sender in which receiver compare his hash code that is generated currently with the hash code that is obtain from the sender. Thus the integrity of the message is verified. A design technique to generate hash code for video data is proposed in this paper.

Literature Survey: In Scalable Video Coding (SVC), code stream contains of one fundamental layer called as base layer and several enhancement layers. The base layer is the foundation of SVC which comprises the low quality and resolution images, and must be transferred to recipients. But the enhancement layers comprise richer contour/texture of images in order to boost the base layer in terms of quality, resolution and temporal scalabilities. Wei (2014), proposed a novel Hybrid Authentication (HAU) scheme. HAU scheme contains both content-based authentication and cryptographic authentication techniques to ensure authenticity and integrity of the SVC code streams.

Sreedhar and Panlal (2012), proposed a morphological transformation method. It is used to enhance the poor contrast images and to identify the background clearly. Image enhancement has been carried out by the following two methods. The first method makes use of the information from image background analysis and the second transformation method uses the opening operation and closing operation. Image background approximation by means of block analysis in combination with transformations enhances the images with poor lighting.

A method for key-frame extraction was proposed by Sheena and Narayanan (2015), which uses threshold of absolute difference between histogram of consecutive frames of video data. Summarization of videos for different applications like, video retrieval, video archival, video object recognition and classification is an interesting research part of computer vision. Extraction of key-frames from a video is an important method to summarize video data.

Kapoor and Aurora (2015), proposed histogram equalization method for gray-level images and extended it for color images. Histogram equalization method is used to enhance the brightness of a gray scale image by adjusting the contrast of an image using its histogram. The increased brightness is different from the average brightness of the original image. In this technique, first the obtained image is converted into Hue, Saturation and Value (HSV). Then the image is divided into two parts based on the disclosure threshold and then equalized them independently. Clipping threshold is used to control the over enhancement. The performance of the enhanced image is measured by calculating the contrast and entropy.

Khelifi and Jiang (2010), proposed a hashing technique based on virtual watermark detection which is more secure and robust. It uses the property that for perceptually close images the watermark detector responds in the same way using a watermark. Color vector angle is widely used in image retrieval and edge detection. Tang (2014), used the color vector angle in image hashing and proposed a robust hashing algorithm combining color vector angles with Discrete Wavelet Transform (DWT).

A hash code generation technique for digital images was proposed by Brabin and George (2016), based on ring partition and Nonnegative Matrix Factorization (NMF). This technique is more robust because the same hash code is generated even if the image is rotated. Zhu (2017), proposed a hashing technique named as Semantic-Assisted Visual Hashing (SAVH). It is an unsupervised visual hashing method and different from supervised hashing approach. The primary idea of SAVH is effectively extract the semantics implanted in auxiliary parts of images to improve the efficiency of visual hashing without any clear semantic tags. To implement this idea, a joined unsupervised framework was proposed to study the hash codes by concurrently safeguarding visual similarities of images.

Nowadays many cryptography techniques, authentication mechanism and security protocols use hash functions as a primitive element. Thulasimani and Muthusamy (2012), uses Secure Hash Algorithm (SHA-192) to check integrity of data. The basic design of SHA- 192 is to have the output length of 192. The SHA-192 was intended to suit the different stages of security enhancements and to resist the complex SHA attacks. In the proposed technique, the Digital Signature Algorithm (DSA) (NIST standard) has been modified using SHA-192. Using the proposed SHA-192 can be used in many applications such as digital signature scheme, authentication mechanism public key cryptography, random key generation and in security architecture of wireless devices. In video or image bit based hashing techniques like SHA and MD5 cannot be applied for many applications. In this paper a content based hashing technique is proposed to generate hash code for video based on key frame extraction, frame enhancement and feature extraction.

2. PROPOSED TECHNIQUE

Proposed system generates strong hash code from video input based on key frame extraction, frame enhancement and feature extraction. Key Frame Extraction is the first operation in the proposed system. The key frames represent the information of the entire video with compact frame set. In videos, key frames indicate a limited frame subset to represent the major contents of video sequences. The use of key frames greatly reduce the amount of video information, achieve video data transmission at a low bit rate channel, reduce the physical memory space, and provide convenience for the users to view the main content of the video.

Second operation is Enhanced Frame Creation which improves the interpretability or perception of information in images. For automated image processing system it provides better input. It helps to examine background information that is critical to understand object behavior without requiring manual inspection.

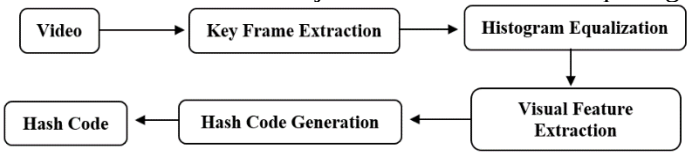

Figure.1. System Architecture

Third operation is Visual Feature Extraction. The small set of extracted visual features describes a large set of video data. Number of variables involved is a key characteristic in performance analysis of complex data. When large number of variables is used in analysis, large amount of computational power and memory are required. And also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. Fourth hash code generation which gives the strong hash code. The important operations in the projected system represented in Figure.1.

Key Frame Extraction: The image frame in the video sequence which is representative and able to reflect the summary of a video content is called as key frame. It is able to convey the core content of video data in clear manner by using the key-frame and it reduces the amount of memory needed for video data processing and complexity greatly. So it is necessary to make the storage organization, retrieval and recognition of video information more appropriate and efficient.

The input video in any format such as .avi, .mp4, .wmv etc. is converted into frames. Then key frame extraction can be achieved by applying frame by frame comparison. Histogram, mean and standard deviation of each frame is used to produce highly different frames which are called as key frames or distinct frames. The algorithm for key frame extraction is given below:

Algorithm:

- Extract frames one by one from the input video.
- Compute histogram difference between each two consecutive frames
- Calculate mean and standard deviation of absolute difference.

- Compute threshold Th.
- Compare the difference with Th and if it is greater than Th, select it as a key-frame else go to step 2.
- Continue until end of the frames in video.

Histogram Equalization for Color Image: Histogram equalization is used for improving the contrast of the key frames. The intensity values of the original frame are mapped to new intensity to give a uniform histogram of intensity values. Intensity refers to the amount of light or the numerical value of a pixel. This method cannot be directly applied to the color image. First the RGB component of color image is converted into HSV component and then the V component is extracted from the HSV component of the image. Then the histogram equalization is applied to the enhanced color image. Finally combine the enhanced V component with the H and S components and then convert the HSV component into RGB component.

This method is a finest technique for maintaining the brightness of frames. It also gives high entropy. Underexposed image parts are exposed by sub image exposure scheme. It controls the over enhancement by histogram clipping. Histogram equalization method increases the color contrast and quality of the image.

Visual Feature Extraction: Visual feature extraction is a crucial method of computer vision for intelligent video processing. Among the important extraction techniques, SIFT is the most widely adopted approach that provides stable visual feature Extraction from images. We propose a modified SIFT Algorithm for extracting the RGB color features and texture features. First extract the RGB color features from the enhanced frames and store it in one matrix such as A-Matrix and then extract the Texture features from the enhanced frames and store it in another matrix such as B-Matrix. Then combine these two matrices A-Matrix and B-Matrix and convert into binary digits. The algorithm steps are given below:

Algorithm:

- Input the key frames
- Extract visual features
- Calculate the texture details of the image A = [x, y, z]
- Calculate the RGB color details of the image B = [r, g, b]
- Get the mathematical vector which contains the measured features. V = [A, B]

The combined hash code can be obtained by combining feature vector V of all the key frames or by applying SHA algorithm on the feature vector of all the key frames. SHA-256 algorithm generates an almost-unique, fixed size 256-bit (32-byte) hash code. This hash code can be used to check the integrity of the input video.

3. RESULTS AND DISCUSSION

The proposed technique is implemented in MATLAB and tested with several video clips. Output of the sample video “Batman2.wmv” is presented here. The sample video consists of 207 frames in it. Some of the frames are shown in Figure.2.

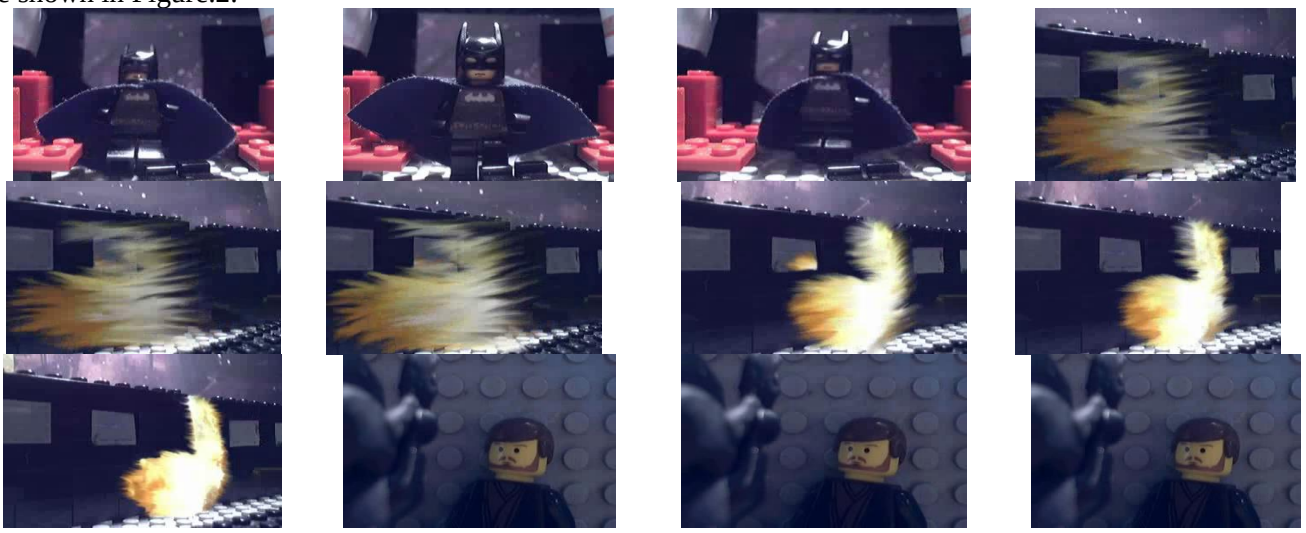

Figure.2. Sample Frames

Figure.3. Key Frames


Figure.4. Enhanced frame creation

By applying frame by frame comparison the key frames shown in Figure.3, are extracted from the 207 frames in the video. Then the key frames are enhanced by applying histogram equalization. A sample enhanced frame is shown in Figure.4. The extracted features from the enhanced sample 4 frames are given below:

1.675916663270638e+05 1.615732465517242e+05
1.835841983150470e+05 1.775732356452456e+05
1.765932258751306e+05 1.655733391588297e+05
1.455733870873102e+05 1.605732886907028e+05
1.805734443840733e+05 1.845947943181818e+05
1.115731565830721e+05 1.455935502089864e+05
1.665731564263323e+05 1.905936233803553e+05
1.325731995036573e+05 1.655734322171983e+05
1.765734195355189e+05 1.675734441116193e+05
1.325932552507837e+05 1.905731362460815e+05
1.805917907654127e+05 1.765573133803552 8e+05
1.655952992032393e+05 1.645731952716823e+05

These values can be used as hash code or SHA algorithm can be applied on these values to generate fixed length hash code. Before applying SHA algorithm, the extracted feature values are converted into binary digits. A sample SHA hash code is given below:

c710b41076768b6105d24d9a9c4f30057b834dfc13881a2164091682f10d9a1861e9109544fadf56ce7a38d286b5bab4a1e244dc791155d16423dc4c534e6e1f.

4. CONCLUSION

We have proposed a video hashing based on Histogram Equalization and SIFT, which is a robust video hashing against modification of video. Content based hashing is achieved for video by the extraction of key frames and the extraction of visual features of the key frame. The generated hash code provides strong data integrity while transmission of data and storage of data.

REFERENCES

Brabin DRD, George, JAM, Rotation Invariant Robust Perceptual Image Hashing Based on Ring Partition NMF and Steganography, Middle-East Journal of Scientific Research, 24, 2016, 1524-1529.

Kapoor K, Aurora S, Color Image Enhancement Based On Histogram Equalization, Electrical & Computer Engineering: An International, 4, 2015, 73-82.

Khelifi F, Jiang J, Perceptual Image Hashing based on virtual watermark detection, IEEE Transaction on Image Processing, 19, 2010, 981-994.

Sheena CV, Narayanan NK, Key-Frame Extraction by Analysis of Histograms of Video Frames using Statistical Methods, Procedia Computer Science, 70, 2015, 36-40.

Sreedhar K, Panlal B, Enhancement of Images using Morphological Transformations, International Journal of Computer Science and Information Technology, 4, 2012, 33-49.

Stalling W, Cryptography and Network Security, Fifth Edition, Pearson Education, India, 2011.

Tang Z, Dai Y, Zhang X, Huang L, Yang F, Robust Image Hashing via Colour Vector Angles and Discrete Wavelet Transform, IET Image Processing, 8, 2014, 142-149.

Thulasimani L, Muthusamy M, A Novel Secure Hash Algorithm for Public Key Digital Signature Schemes, The International Arab Journal of Information technology, 9, 2012, 262-267.

Wei Z, Wu Y, Deng RH, Ding X, A Hybrid Scheme for Authenticating Scalable Video Code streams, IEEE Transactions on Information Forensics and Security, 9, 2014, 543-553.

Zhu L, Shen J, Xie L, Cheng Z, Unsupervised Visual Hashing With Semantic Assistant for Content-Based Image Retrieval, IEEE Transactions on Knowledge and Data Engineering, 29, 2017, 472-486.