International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

APPLICATIONS OF IMAGE AND VIDEO DEDUPLICATION: A SURVEY

Nivetha.L[1], Gowri.V[2], Dr.J.Jane Rubel Angelina[3]

[1]UG Student, Dept. of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India
[2]UG Student, Dept. of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India
[3]Assistant Professor, Dept. of Computer Science and Engineering, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

---***---

Abstract - *In this digitizing technology world, people are curious about uploading and sharing information mainly in picture and video format. This growth has resulted in an increase in storage capacity that includes a lot of redundant multimedia data. Deduplication is one of the new techniques dealing with redundant data which is stored at different storage locations. Deduplication is one of the new techniques dealing with redundant data which is stored at different storage locations. In deduplication, a single copy is preserved when more than one copy of the same data is identified, and the other data is replaced with pointers pointing to the preserved copy. Storage may effectively be used to store a lot of other data. While there are various types of techniques for deduplication, researchers concentrate a lot on image and video deduplication techniques and their implementations as it is a difficult one. This article discusses the implementations of the real-time system image and video deduplication techniques.*

Key Words: **Image deduplication, Video deduplication, Storage optimization.**

1. INTRODUCTION

The planet is being packed with various forms of multimedia data in the last few years. We've experienced the explosion of video content over the Internet in recent years as billions of videos are stored and exchanged in the cloud, a large portion of which consists of near-duplicate video copies. Rapid advances in multimedia technologies fueled this growth. So, deduplication is needed to eliminate the occupied storage space that could be used efficiently for some other purpose. Deduplication has been commonly used in backup and archival systems to significantly increase storage use.

There are different types of deduplication techniques like data deduplication, image deduplication, and video deduplication. There are many methods to manage file deduplication out of the above forms. Yet removing image and video duplicates is a difficult task. That is because the data will be exactly replicated in case of file deduplication. But in case of an image or video, one should also think about different formats in which the same image or video could be stored. So, handling images and videos is a quite challenging task.

1.1 Image deduplication

Image deduplication recognizes duplicate images, removes all images but maintains one copy to create local references to the information that users can access without any problem. The main concept in the deduplication of images is to delete duplicate images by five stages [1] comprising feature extraction, high-dimensional indexing (scaling, changing storage format), accuracy optimization by extracting image edge information, centroid selection using cluster formation, and fixing the centroid image and deduplication evaluation followed by comparison. When a new picture comes in, the client first extracts the fingerprinter and sends the fingerprint to the sides of the storage. When the storage side cannot locate the corresponding fingerprinter in index tables, the next image will be processed and the image fingerprinter moved to index tables. Alternatively, the side of the storage will maintain the higher image perception-quality as the centroid image, and create a point pointing to the centroid image.

1.2 Video deduplication:

Video data is an important part of the today's Internet data and it creates a lot of flexibility in storage and distribution. Video deduplication is the repetition of scenes or frames within the video and also among other videos. Consider two similar videos, for example, where the action performed in the first frame of the first video is replicated in the last frame of the second video. Then the last frame of the second video is redundant and should be replaced with a pointer. This not only saves storage space effectively, but also greatly increases the accuracy of duplicate images to be retrieved. The original videos have to be chunked into frames for further processing to perform video deduplication. In the case of variable-sized frames the frames are of the same fixed size (in fixed-sized frames) or different sizes. Similar to the technique of image deduplication, fingerprints are created for each frame and compared with the other ones of same or different videos. It can be indexed using data structures such as inverted index, B+ tree etc to maintain the order of the frames. So that the videos can be recovered in order and viewed by the users without any inconvenience.

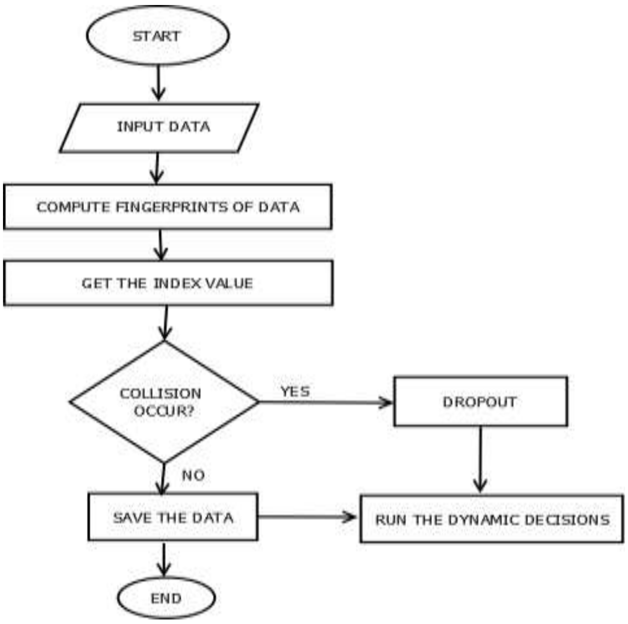

Fig -1: Process of Deduplication

2. Literature Review:

Researchers have been proposed certain applications for image and video deduplication in real time. Let us discuss a few of those applications using different types of methods that would deliver a better result in day-to-day activities.

Ashish Kumar Layeket al [2] proposed a novel methodology for identifying and grouping near-duplicates from image streams posted on online social media. This is viewed as the problem of clustering the images in the stream, such that each cluster consists of a series of near-duplicate images. The proposed technique is a hierarchical one consisting of two levels – a global feature descriptor-based similarity check at the top level to minimize search space, followed by Locality Sensitive Hashing (LSH)-based similarity checking at the next level to detect similarity with only a few images, thus maximizing both precision and processing time. The technique is composed of two distinct stages. The first stage is the training process in which a wide set of random images shared on social media sites are used to create a training dataset. This process also involves clustering the collection of random images using moment-based features as they are invariant to translation, rotation and scaling and generation of visual vocabulary generation. The second stage is an online near-duplicate process of detection. Here, some candidate images are to be found such that each candidate image is a representation of a distinct set (cluster) of near-duplicate images and then LSH-based similarity check between the test image and the identified candidate images is performed.

D Wen et al [4] designed a face spoof detection system that finds the identical images on various types of images for the same person. The person's face is identified in this system, and is normalized. This is focused on an Study of Image Distortion Analysis (IDA).Four forms of IDA features (specular reflection, blurriness, color moments and variation of color) were designed to catch image distortion in the spoof face images . The normalized image features are extracted and compared with the classifiers learned from different training data groups. Ultimately, if the image matches all of the images in the classifier then the program determines that the picture is a spoof or fake image.

Zhili Zhou et al [5] presented an effective and efficient global context verification scheme for image copy detection. The framework consists of three main components, namely Scale Invariant Feature Transform (SIFT) feature matching, OR-GCD extraction, and verification of SIFT matches. To obtain initial SIFT matches, the first one matches the SIFT features between images based on the Bag of Words (BOW) quantization. The global context of every matched SIFT function is defined in the OR-GCD extraction. Then, by comparing their corresponding OR-GCDs, the matched SIFT features are checked by extracting and concatenating two kinds of binary vectors — one based on intensities and the other based on gradients, and then the verification result can be further used to calculate image similarities for copy detection. The approach suggested is expanded to cover the function of partial-duplicate image detection. Potential duplicated region location is a crucial stage for partial-duplicate image detection. SGV strategy is suggested that takes into account not only the relative positions of local features but also their characteristic relationships to filter geometrically inconsistent matches for potential duplicated region location.

Z Zhou et al [6] designed to eliminate the near-duplicate image quickly and accurately for visual sensor networks. This is because visual sensor nodes, that is the camera nodes, produces a lot of visual data, such as digital images and videos, that gets transmitted on visual sensor networks. Among those, there are many near-duplicate images, which cause a serious wastage in storage. Here the images are split into a 3*3 matrix and converted to hash code. Near-duplicate clustering is the first and crucial step towards the removal of near duplicates. Next, they suggest a novel seed image selection method based on the PageRank algorithm, which can pick the most appropriate images accurately as seed images and delete the other redundant images to complete the near-duplicate elimination. Locality Sensitive Hashing (LSH) is a famous inverted index file building technique for matching high-dimensional features. The hash values of a given query function are looked up from inverted index file entries, and then an additional comparison is applied between the original features to further validate whether or not two features are a match.

Etienne Gadeski et al [7] suggested a solution to the issue of indexing and finding duplicates of images while streaming visual data in a social media intelligence scenario. Rescaling is done in the proposed method to transform the image into a grayscale and to reduce its size to pixels. Based on comparisons of the pixel value, hash is determined, and each hash value is encoded to construct the descriptor. The hamming distance is then used to calculate the distance. Therefore the Hamming distance counts the number of bits that vary between two values. That can recognize duplicates that have been flipped. Due to the efficient use of the Hamming distance we can calculate a descriptor for one image and scan a large set of images for duplicates on a single core processor in less than a second. This has resulted in an effective and robust description of the image that is well suited for indexing and searching large visual data streams.

NP Ramaiah et al[8] developed a similarity-based image deduplication with photographs of family cards from various districts. The authors invented a Content-Based Image Retrieval (CBIR) model in this process. This model consists of three steps: feature extraction, clustering and deduplication. In feature extraction, the color content and texture content are extracted with the help of histogram refinement technique. The extracted features are divided into 23 clusters for each district. The clusters are further clustered based on similarities using the k-means clustering algorithm into separate groups. In the deduplication process, the histogram of the image is determined and compared with the other images. The images which are less identical than the empirical threshold are viewed and manually checked to avoid false positive images. The new approach removes up to 0.35 million redundant images.

Javed Akhtar Khan [9] suggested a method for the identification of duplicate frames in Closed Circuit Television(CCTV) DVR by correlation coefficient factor. Unused frame identification is filtered by correlation coefficient and finds similar frame images which help to save storage space.

Q Chen et al [3] proposed a novel face deduplication system in Video Surveillance. This system consists mainly of two parts: face detection and face quality evaluation. The key method is to locate all facial images through facial detection, obtain multiple faces belonging to the same person as a face subset, and then pick a limited number of high-quality faces using image deduplication. For face recognition, Normalized Pixel Difference (NPD) features are extracted from each of the frame image in the video sequences from the surveillance videos. Some of the most discriminative features are selected from all the features using gentle AdaBoost algorithm to create classifiers. Now, using soft-cascade structure, strong classifier is created. It will give the face confidence as the output. (The greater confidence means greater possibility of being a human face will be greater). Assessment parameters such as facial confidence, image sharpness and resolution are used for calculating the facial quality scores in the face quality assessment. The higher the score, the better is the quality of a face. The aim of deduplication is to ignore the low-quality faces that have been identified and to retain the high-quality faces for face recognition to be more accurate.

Jayashree Kharat et al [10] developed a novel passive blind forgery detection to identify frame duplication attack in Moving Picture Experts Group-4 (MPEG-4) coded videos. The frame duplication attack is nothing but one or more frames are copied and the same gets pasted in the same video at different location to hide particular operation. The method proposed is composed of two phases. Firstly, the suspicious frames (most likely faked frames) are marked by measuring the motion vector from the test video for all frames. Then, Scale Invariant Feature Transform (SIFT) is a computer vision algorithm used to detect and extract descriptors of local features as the forger may adjust the illumination or add noise to the frame before pasting it at another location. When the frames are a duplicate of each other, then both frames will have the same sift key-points. The SIFT features are used to compare suspicious frames, and Random sample consensus (RANSAC) is eventually used to identify and find duplicate frames for identification of forgeries.

Donghyeok Lee et al[11] developed a secure and efficient model based on blockchain and deduplication technology for Intelligent CCTV surveillance. The proposed approach ensures data integrity of CCTV video data based on edge blockchain region and by deduplication, it increases the efficiency of video data transmission. Here a method of building the blockchain separately is proposed in the trusted edge blockchain area. With this, the object's privacy can be secured against CCTV image data, and the image information can be safely submitted to the cloud server. Also, through the event-based deduplication technology,the event analysis is actually carried out to compare the image blocks in the file byte unit, even if they have several different contents in the file data, the comparison value for the outcome of the event analysis is regarded as the same image and can be defined as duplicates. This helps the bandwidth to be reduced at the time of transmission and a safe and effective video surveillance environment can be constructed.

Hongyang yan et al [12] build a centralized privacy-preserving duplicate removal video storage system in IoT. Because of the data redundancy, the large volume of data generated by IoT devices results in the use of communication bandwidth and storage space. All the operations and computations must be carried out in encrypted form to protect data privacy. In video files there is a high level of redundancy due to the adjacent video frames which are strongly correlated with each other. In the proposed block-level deduplication method, deduplication is not only performed on the blocks of a single image to perform deduplication on video data; but also to increase the performance on the blocks of different images. It is found that the duplicated privacy-preserving removal system will be helpful with enhanced deduplication capability and increased protection for IoT's centralized storage environment.

Heng Tao Shen et al [13] developed a prototype framework for online near-duplicate video clip (NDVC) detection, called UQLIPS. It's a particular content-based video retrieval (CBVR) issue that looks for the near-duplicates of a query clip. UQLIPS core consists of two novel complementary schemes for the identification of NDVCs. Bounded Coordinate System (BCS), a compact representation model that ignores temporal information aimed at capturing the video's dominant content and evolving content trends, summarizes each video globally to a single vector that captures the dominant content and evolving content trends in each clip.

The other idea, called FRAme Symbolisation (FRAS), maps each clip to a sequence of symbols, and takes video sequences in temporary order. Next, FRAS generates a symbol dictionary by conducting hierarchical clustering (for ex, k-means) through the entire dataset of frames. This description is more compact, since every frame is represented by a cluster id. The similarity between video clips can be calculated by Probability-based Edit Distance (PED), which extends string edit distance and thus the whole system leads to more actuate NDVC search action.

Table -1: Comparison of different applications in image and video Deduplication

Type of deduplication	Application	Technique used	Analysis based on	Performance analysed
Image Deduplication	Fast and accurate near-duplicate image elimination for visual sensor network [6]	Improvised k-means clustering, simple random sampling and systematic sampling	Deduplication Rate	Deduplication rate of 89.94% has been achieved
Image Deduplication	De-duplication of photograph images using histogram refinement [8]	k-means clustering	Accuracy	90% accuracy has been achieved
Image Deduplication	A passive blind forgery detection technique to identify frame duplication attack [10]	Scale Invariant Feature Transform (SIFT) algorithm	Simulation time	33 seconds
Video Deduplication	Face deduplication in video surveillance [3]	Ada Boost Algorithm	Accuracy	Accuracy increases when $N \leq 16$ (N = No. of faces)
Video Deduplication	Centralized Duplicate Removal Video Storage System with Privacy Preservation in IoT [12]	Encrypted deduplicated data (Block-level deduplication)	Deduplication rate	84.04% of deduplication rate has been achieved

3. CONCLUSION

This article deals with the various applications that can be used for the multimedia data deduplication in real time. We have explained, analyzed and compared the several applications suggested by different authors in order to improve the efficiency, rate of deduplication and accuracy. Measures could be taken to improve the performance of these applications without disturbing the other features.

REFERENCES

[1] Chen, M., Wang, S., & Tian, L. (2013). A High-precision Duplicate Image Deduplication Approach. JCP, 8(11), 2768-2775.

[2] Layek, A. K., Gupta, A., Ghosh, S., & Mandal, S. (2016, December). Fast near-duplicate detection from image streams on online social media during disaster events. In *2016 IEEE Annual India Conference (INDICON)* (pp. 1-6). IEEE.

[3] Chen, Q., Yang, L., Zhang, D., Shen, Y., & Huang, S. (2018). Face deduplication in video surveillance. *International Journal of Pattern Recognition and Artificial Intelligence*, *32*(03), 1856001.

[4] Wen, D., Han, H., & Jain, A. K. (2015). Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, *10*(4), 746-761.

[5] Zhou, Z., Wang, Y., Wu, Q. J., Yang, C. N., & Sun, X. (2016). Effective and efficient global context verification for image copy detection. *IEEE Transactions on Information Forensics and Security*, *12*(1), 48-63.

[6] Zhou, Z., Wu, Q. J., Huang, F., & Sun, X. (2017). Fast and accurate near-duplicate image elimination for visual sensor networks. International Journal of Distributed Sensor Networks, 13(2), 1550147717694172.

[7] Gadeski, E., Le Borgne, H., & Popescu, A. (2015, June). Duplicate image detection in a stream of web visual data. In 2015 13th International Workshop on Content-Based Multimedia Indexing (CBMI) (pp. 1-6). IEEE.

[8] Ramaiah, N. P., & Mohan, C. K. (2011, September). De-duplication of photograph images using histogram refinement. In 2011 IEEE Recent Advances in Intelligent Computational Systems (pp. 391-395). IEEE.

[9] Khan, J. (2018, February). Identification of Duplicate Frame In CCTV DVR by Correlation Coefficient Factor. In 2018 Second International Conference on Computing Methodologies and Communication (ICCMC) (pp. 235-239). IEEE.

[10] Kharat, J., & Chougule, S. (2020). A passive blind forgery detection technique to identify frame duplication attack. Multimedia Tools and Applications, 1-17.

[11] Lee, D., & Park, N. (2019). CCTV Video Data Deduplication Method using Blockchain and Big Data Analysis. *International Journal of Advanced Science and Technology*, *28*(3), 267 - 275. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/328

[12] Yan, H., Li, X., Wang, Y., & Jia, C. (2018). Centralized duplicate removal video storage system with privacy preservation in IoT. *Sensors*, *18*(6), 1814.

[13] Shen, H. T., Zhou, X., Huang, Z., Shao, J., & Zhou, X. (2007, September). UQLIPS: a real-time near-duplicate video clip detection system. In Proceedings of the 33rd international conference on Very large data bases (pp. 1374-1377).