RESEARCH PAPER

A Review on Perceptual Hash based Video Shot Boundary Detection

Debjani Bhowmik*[1] and Saptarshi Chakraborty[1]
*ICFAI University of Tripura, Agartala, Computer Science and Engineering Department,
debjani.ece@gmail.com, saptacse@gmail.com

Abstract: The consumption and production of multimedia data have grown significantly in the modern world due to advancement in the technology and internet speed. These factors necessitate the use of effective and efficient multimedia data especially video that is a collection of frames. In consecutive frames, specific frame rate with corresponding audio track represents a video signal. Shot boundary detection (SBD) is a temporal video partitioning, video retrieval and video browsing process. The method can detect shot transition and their boundaries between consecutive video shots. This study introduces a robust method leveraging KAZE feature descriptors and perceptual image hashing for accurate detection. For addressing challenges in multimedia data indexing and retrieval, KAZE based SBD generates unique hashes for video frames, enabling precise transition identification. Algorithms developed for shot boundary detection can be improved by enabling it to determine correct transition in a video frame sequence and also by reducing computational cost of the algorithm. A brief literature survey on the fundamental concepts and challenges of KAZE based video shot boundary detection is presented. The advantages and difficulties associated with each of the approaches are surveyed. Challenges associated with few developed KAZE based SBD algorithms are discussed.

Keywords: Shot Boundary Detection, Video retrieval, Video browsing, KAZE Feature descriptor, Shot transition, Perceptual image hashing.

How to cite this article: Bhowmik D, Chakraborty S. A Review on Perceptual Hash based Video Shot Boundary Detection. Int J Drug Deliv Technol. 2026;16(62s): 335-342. DOI: 10.25258/ijddt.16.62s.40

Source of support: Nil.

Conflict of interest: None.

1. Introduction

Multimedia content including pictures, videos, audio, and other digital content that is shared online, has grown significantly in the modern period. Multimedia applications can alter or modify how data is represented. Numerous risks to legitimate content result from this Multimedia data indexing and retrieval can lead to issues. For multimedia data, it is crucial to provide an effective indexing or classification system with a suitable recovery mechanism. The hashing approach creates a hash value with a set size for the generated hash value and a variable length for the input[1]. Techniques for compression and quantization are typically employed to reduce storage, improve algorithm performance, and shorten the time needed to generate a hash value. Hash values and similarity measures are comparable. Hash values can be used to identify how similar or dissimilar two photos' visual components are. If a hashing method possesses both resilience and discrimination, it is deemed adequate. If the images are clearly identical, the resulting hash sequences are comparable and independent of one another. The algorithm's hash sequence for the same image can vary greatly depending on the supplied key conditions [2]. Robustness increases when the threshold is raised. The ability to discriminate may be diminished by increased robustness and vice versa. To preserve the qualities of robustness and discrimination capability, the threshold value should be balanced [3]. This hashing method is not the same as cryptographic hash functions, which react to small changes in the digital data. Perceptual image hashing has numerous applications, including digital watermarking, tamper detection, picture validation, image indexing and recovery, and more [1]. The KAZE feature descriptor, which has a discriminating property towards single and multiple sets of manipulations, is used in this methodical and forceful scheme of picture hashing [3]. The method in the present work proposes using fixed key points using the KAZE feature descriptor. Out of all the fixed key points, key points are extracted from the image and used to construct features. To create the hash value for the current function, the first key point is selected. The Hamming distance is used to assess the perceptibility of pictures. Using the KAZE feature detector, it creates hashes for each video frame, we suggest an image hashing method that can assess whether or not images extracted from video are perceptible. Based on the Hamming distance between the hash values, the suggested method may categorize abrupt transition frames.

2. Perceptual Image Hashing

Detecting changes in multimedia data information is essential for many applications, such as image databases in the fields of journalism, medicine, or art. Multimedia data security, authenticity, and integrity can be verified using the perceptual image hashing technique and its reliable algorithms. The use of perceptual image hashing

*Author for Correspondence: debjani.ece@gmail.com

in digital communication, telemedicine, multimedia systems, and medical imaging is well acknowledged. For the following reasons, a conventional cryptosystem is not a suitable solution when encrypting multimedia data. Because multimedia data is so large, it takes longer for typical cryptosystems to transform it into an encrypted format. Multimedia data must match the original multimedia data after decryption, which is another problem. It is not required to preserve the same data for images or videos before and after encryption. Perceptual image hashing is based on human perception; multimedia data with slight distortion is likewise acceptable. Perceptual image hashing is also applied in the watermarking industry. A watermark, which contains copyright information in multimedia data, is used to safeguard intellectual property rights. In order to retrieve the identity of the ownership of the host multimedia data when needed, the watermark needs to be robust and impervious to security threat [4].

Perceptual image hashing, also known as a perceptual hash function, is a technology introduced to authenticate data in image applications for multimedia data security purposes. Perpetual hash functions have the following characteristics:

I. When two images are consistently similar, their hash values are nearly identical. This special feature is considered as robustness to perception.

II. Various photos will produce various hash values when their visual contents differ and they are always unique. This is called capability to discriminate. Perceptual hashing has a few minor accuracy problems. When working through blocks, several developed image hashing algorithms are unable to handle major angle reaction. Although histogram-based image hashing methods are resilient and efficient, they are less able to differentiate images due to luminance [5]. Certain positives and negatives are mistakenly recognized in perceptual image hashing. Accurately maintaining the integrity of databases containing photographs and videos is challenging. Cyber attacks may have an impact on perceptual hashing. To avoid observing dangerous content (false negatives) or incorrectly classifying beneficial content (false positives) in perceptual hashing, images and videos must be distributed precisely. A number of current image hashing techniques are designed to prevent content manipulation, but they are not resilient to multiple sets of changes [6]. A crucial component of image hashing techniques used in authentication is feature extraction.

Perceptual image hashing classification

The four types of perceptual image hashing classification are as follows.

I. **Statistical strategy**: Extracted hash features are used to calculate image statistics such as mean, image blocks, variance, histogram, etc.

II. **Relation-based strategy:** These methods use the coefficient's invariant relationships in the Discrete Wavelet Transform (DWT) or Discrete Cosine Transform (DCT) to extract hash features.

III. **Coarse representation strategy**: An image is used to generate coarse information, which is then used to create perceptual hashes. The geographical distribution of important wavelet coefficients and low-frequency Fourier transform coefficients is comparable to coarse information.

IV. **Low level feature-based strategy**: Upon identifying the salient picture feature points, hashes are extracted. In order to use the coefficients to determine the final value of the hash, these algorithms first apply the DWT or DCT transform to an image. The perceptual properties of the photographs are never affected by these types of hash values, which are sensitive to both local and global aberrations.

3. Framework of Perceptual Image Hashing

The stages of perceptual image hashing are following

Transformation of Image

Feature Extraction

Quantization Phase

Compression and Encryption Stage

In order to relate to image frequency coefficients and base extracted characteristics on the pixel values of the image, spatial or frequency transformation is first applied. The system extracts characteristics from the input image and generates hash vectors sequentially. Continuous hash vectors are quantized into discrete hash vectors during the quantization stage. Next, perceptual hash string values of binary values are created from discrete vectors of hashes. A compression stage is applied to the binary hash string values of the perceptual hash. The string is transformed into a final hash value using the encryption algorithm [6].

Transformation of Image: The images' dimensions, expressed in bytes as x by y are taken into account as input for the transformation stage. Smoothing, color transformation, affine transformation, and other spatial and frequency transformations, such as Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT), are applied to the image. In this kind of modification, the features of the retrieved frames are determined by the values of the picture pixels or in the case of frequency space, the features of the extracted frames are altered depending on the frequency coefficients of the image [7].

Feature Extraction: In the stage of extracting features, image features are extracted by the system from the transformed image and produce produces the feature vector of L features where $L \ll X \times Y$.

Quantization Phase: In the quantization phase, a hash vector is quantized and used with $L \times P$ byte elements. A consistent quantization is applied to each of the modules of the uninterrupted perceptual hash vector for quantization. Quantization with adaptive estimation is very much useful for image hashing.

Compression and Encryption Stage: The perceptual image hashing system concludes with the compression and encryption stage. Through compression and encryption, the intermediate binary string of the perceptual hash is transformed into a brief, fixed-size perceptual hash of one byte. It is the last hash generated that enables image authentication and verification on the other end, or the receiver side. A cryptographic hash function can ensure this penultimate and final step of the perceptual image hashing system [8].

4. Metrics used for a perceptual image hashing

Two types of perceptual hash functions are in use. With key value, perceptual hash functions produce a value of hash h, where x is considered as an input value of arbitrary length and k is a secret key h= (H(x;k)). A perceptual hash function without key produces a value of hash h from an input value of x and thus hash function H(x)=h. Design of robust and efficient techniques for perceptual image hashing is very challenging because of several conflicting requirements [9]. If we consider P as probability and H( ) as perceptual hash function which generates a string of binary values of length l. If an image is considered as I and the modified version of the image is considered as I_{alike} which is distinguishably similar to I. I_{unlike} represents an image that is distinctly not similar as I. Hash values h_1 represents the hash value of the original image I and h2 represents the hash value of the perceptually non similar image I_{unlike} from I. Binary string of length consists of 0 and 1. Necessary properties of hashing functions for perceptually similar images can be summarized as follows [7].

1. Hash values follow equal distribution property.

$$P_r(H(I) = h_1) \approx \frac{1}{2^l}, \forall\ h_1 € \{0,1\}^l \quad (1)$$

2. Image I and perceptually different image I_{unlike} are independent of each other.

$$P_r\left(H(I) = \frac{h_1}{H(I_{new})} = h2\right) M \approx\ P_r(H(I_{alike}) = h_1\ ), \forall\ h_1, h_2 € \{0,1\}^l \quad (2)$$

3. In-variance property for Image I and perceptually similar image $I_{alike.}$

$$P_r\,(H(I)) = \left(H(I_{alike})\right) \geq 1 - \Theta_1, for\ a\ given\ \Theta_1 \approx 0 \quad (3)$$

4. Distinction property of Image I and perceptually different Image $I_{unlike.}$

$$P_r(H(I) \neq) H(I_{unlike})) \geq 1 -\ \Theta_2\,, for\ a\ given\ \Theta_2 \approx 0. \quad (4)$$

For equation no (3) in perceptual hash functions features are extracted from images so that features are invariant for modifications globally like enhancement and compression [10]. In equation no (4), for any image I, it is not possible for an opponent to establish a perceptually non-identical image I_{unlike} such that $H(I) = H(I_{unlike.})$

Perceptual Hash functions which are published are known to all and this property is very difficult to achieve.

5. Related Work

Perceptual image hashing has been the subject of numerous experiments in recent years. Below is a discussion of some of the significant and noteworthy ones. For data security in surveillance systems, a hashing is suggested. A robust process that reduces SVD dimensions and generates hash values with a smaller set of dimensions was proposed by Neelima et al. Rotation, translation, and scaling are obtained by combining the benefits of SIFT's invariant features with the feature points produced by SVD [9]. A Scale Invariant Feature Transform (SIFT) technique that extracts feature points from the image was proposed by Arambam et al. Values are taken from each of the many blocks that make up the image. A more reliable result is obtained using the SIFT approach.

However, it becomes less effective and needs more processing resources when paired with Singular Value Decomposition (SVD) [10]. Zernike moments, which can handle rotation attack for small rotation angles but not for large rotation angles, are used in to state a novel methodology [11]. Weber Local Binary Pattern (WLBP) is utilized in to extract characteristics that integrate color, texture, and angles. Pixels in the center and surrounding areas are contrasted [12]. When principal component analysis is used, the difference decreases and a histogram is produced. The approach is used on the pictures, and it is not as reliable. Gharde et al. used the fuzzy histogram to create a reliable perceptual hash algorithm.

A normalized fuzzy histogram is the source of hash values. This approach is reliable and has the same level of discrimination as human perception. The fuzzy color histogram can be normalized to make the system scale invariant. The normalized FUZZY color histogram is used to create hash values [13]. The Canny operator can create a single image, and then DCT coefficients are extracted from each block. Differentiation capability is minimal, making it an ineffective method. An image hashing method makes advantage of the Zernike moment. Zernike moments are a reliable and efficient way to depict the features in a picture. This approach is also strong and effective against content manipulation. Attacks up to 20º rotation can be handled by this method [14].

The picture hashing method uses local entropies and DWT. Block entropies are altered linearly when content modifications are maintained. A hash is produced by applying 2D discrete wavelet transforms to each block entropy. The current DWT-based picture hashing method is methodical and robust against different threats to protect contents. Minimal rotation can be handled by this picture hashing [10]. The grey-level co-occurrence matrix (GLCM) is used to create a hashing technique for texture differentiation. Cosine transform coefficients of local discrete type are utilized as parameters in order to remove discriminative limitation. The final hash value is created by combining local and global features [15]. Edge information is captured using a color vector.

The color vector is compressed using DWT in order to produce a hash. The stated method may withstand several attacks, including a 5º angle rotation attack [16]. A Laplacian pyramid-based method is suggested. It generates an image's multi-scale properties based on the difference. Features are extracted and fixed points are found. Although the method is effective and resilient to some attacks, it is unable to withstand geometric attacks [17]. A novel picture hashing method utilizing several form features has been proposed by Tang et al. (2013). Edge detection is used in shape-based features image hashing to capture the angular transformation and produce a hash. With the exception of the angular rotation characteristic [18] the suggested method is resilient to a number of attacks. A hash value is created using a random transform. When it comes to picture retrieval and image attack with rotation, the system as a whole is robust and effective.

6. Video and types of shot transition

There are several scenes, shots, and frames in a video. A scene is composed of multiple shots. A sequence of frames captured by a single camera can be referred to as a shot. A single shot's frames provide similar or identical features and information with temporal variations. A single shot's frame is its unit. [19]. Video frames are interdependent and momentarily arranged. SBD techniques are used to extract fundamental shots from videos. There are two kinds of shots: abrupt or cut transitions and gradual or gentle transitions.

Figure 1: Frames of figure 1(A) depicts Abrupt shot transition from Video D2 of TRECVid2001 dataset. Frames of figure 1(B) depict Abrupt shot transition from Video D6 of TRECVid2001 dataset.

When there is an abrupt difference between two consecutive shots, it is referred to as a hard transition or cut transition. When the camera is stopped or restarted, it could also happen in a single frame. Figure1 shows an example of a hard transition that occurs between the final frame of the previous shot and the beginning frame of the subsequent shot.

Figure 2: Frames of figure 2(A) depicts Gradual shot transition from Video D2 of TRECVid2001.Frames of figure 2(B) depicts Gradual shot transition from Video D2 of TRECVid2001.

When there are slight variations between two successive shots, this is known as a soft transition. Edit effects may sometimes be the cause of a soft transition. several kinds of soft transitions, such as wipe, dissolve, and fade transitions. A blank frame precedes the fade-in transition, which concludes with a brilliant frame. Every frame has a progressive increase in intensity. However, in a fade-out transition, the intensity progressively drops from a brilliant frame to a blank frame. One shot gradually replaces the subsequent shot during the dissolve transition. When the next shot comes, the first one vanishes, and the shots may briefly overlap. One shot can take the place of another during a wipe transition. The second shot is pushed off the screen by the first. The shot can travel vertically, horizontally, obliquely (from one corner to the next), starting from the center of the screen or advancing in that direction.

7. Feature Extraction

The technique of feature extraction aids in decision-making for object classification or recognition, pattern recognition, etc. Both low-level and mid-level information may be included in the retrieved video frame. The minor aspects that can be retrieved from films include the intensity value, RGB value, and histogram [20]. A feature point detector is used to identify objects in a picture.

Accurate hard or soft transitions can be obtained with the aid of shot boundary detection (SBD) techniques. Below is an introduction to a survey on several SBD techniques.

A crucial component of the SBD technique is feature extraction and selection [21], which is crucial for enhancing the SBD system's overall performance. Better performance during transitions can be achieved by carefully choosing features. Local features are in charge of the diversity of features inside a shot. Global characteristics are unable to differentiate between intra-shot and inter-shot frames. Therefore, choosing the right frame characteristic is crucial to enhancing SBD performance.

8. Approaches of Shot Boundary Detection

More precise and effective shot boundaries are produced by more dependable and accurate algorithms. A number of popular techniques for handling cut transition and soft transition are covered.

Detection based on Pixel: This method compares the pixel intensity for two consecutive video frames. The percentage of pixels that have changed over the course of two successive frames is used to assess changes in pixel intensity. The chosen threshold value is compared with the total pixel difference. If the pixel value's intensity exceeds the chosen threshold, shot modifications are taken into account [22]. When the camera is moving and an item is moving quickly, this method is sensitive. In this situation, the threshold is set by hand.

Detection based on Histogram: Color histograms can be used to determine an image's color distribution. Regardless of the location within the image, the histogram represents the distribution of texture and shape. Histogram difference determines how two consecutive frames' histograms differ from one another. Histogram similarity can be used to assess if two consecutive frames are nearly same or not. Histograms are taken from two successive video frames, and the difference between the histograms is then computed. A shot change occurs when the measured deviation exceeds the chosen threshold [23]. For RGB color spaces, a block-based histogram difference approach with an automatically established threshold in Abrupt Shot Transition is shown [24]. Histogram-based SBD techniques are widely used and effective for measuring performance metrics.

Detection based on Edge: In order to identify hard or soft transition, edge based information offers an additional set of features of an object's image boundary that distinguishes from the background and depicts its edge. Transition can be announced in edge based techniques when the position of the current frame's edge exhibits a large gap with the position of the former frame's edge that has disappeared [25]. Finding the ration of edge change and motion compensation are essential procedures for accessing edge change in the edge-based method.

Detection based on Motion: One crucial aspect of video frame is motion. Shots may be wrongly categorized as gradual transitions when the camera is moving. Shot boundary detection is more precise when using motion based methods [25]. Motion vectors are calculated by using the Block Matching Algorithm (BMA) to match the blocks of the current frame with the blocks of the subsequent frames. As a result, it is possible to distinguish between a video shot's transition and camera movements. Zoom operation is used improve the accuracy of the SBD algorithm. Hard transition is detected if the block match value derived from the motion vector is greater than the threshold [26].

Detection based on Deep Learning: Convolution Neural Networks, or CNNs, are a significant Deep Learning technique that can extract significant high level video frame characteristics [27]. A CNN based approach that can extract interpretable video frame information was proposed by Tong et al. [28]. Both hard and soft transition limits can be found using this technique. The pre-processing step for choosing a collection of video frames is thought to be adaptive threshold method. The appropriate output is chosen using a probability distribution in 100 classes, with one frame serving as the input.

9. Challenges in Shot Boundary Detection

While determining video shot boundaries it needs to solve all challenging issues. To solve these challenging issues few more research work needs to be contributed from different aspects. Video frame features, similarity measures and threshold are the parameters that have a remarkable role in detecting video shot boundaries. Camera and object motion, illumination changes and dim lighting frames are few of the major challenging issues of shot boundary detection. Due to the light effects and flash light effects, there can be sudden changes of illumination. Several research works conducted to handle lighting effect issue [28]. Chen et al. describes the method where it deals with changing of flash lights for short duration video where contents are similar for before and after the happening of lighting flash. Dim lighting frames can significantly affect video shot boundary detection techniques. Due to dim lighting problems in different videos like short duration video, movies and news if consecutive frames are almost similar then gradual transition is considered. Sometimes it is difficult to differentiate between hard transition and soft transition due to dim lighting effects [29]. Generally, SBD algorithms are unable to distinguish smooth object-camera motion and soft transition. When there is small region changes exist in frames of video like movies, news and advertisements then frames are not considered as transition frames as consecutive two frames are almost similar due to small changes [30]. Performance can be significantly improved when motion effect and abrupt light variation effect is present with the automatic SBD method. This motion sensitive method is also sensitive to the intensity of flash light which can affect the whole frame [31].

10. Evaluation Metrics on SBD and Comparative study on several techniques

The SBD algorithm's performance can be estimated using two future metrics. Accuracy and computational complexity are the two assessment metrics. Processing time is correlated with computational complexity.

Reduced time complexity results in shorter processing times. Accuracy and Time complexity can be achieved by computing Recall, Precision and F1 score.

Precision

Precision is a measure of accuracy used in video frame border detection.

It is the percentage of correctly reported transitions as well as the sum of all transitions (both true and false).

$$Precision = \frac{N_C}{N_C+N_F}100 \quad (1)$$

Recall

The percentage of correctly reported transitions and missing transitions is known as recall.

$$Recall = \frac{N_C}{N_C+N_M}100 \quad (2)$$

F1 Score

The F1 Score is the metric used to assess a system's effectiveness. Combining recall and precision results in this metric. The range of its value is [0,1].

When a system's F1 Score is 1, it is at its most efficient.

$$F_1Score = \frac{2 * REC * PRE}{REC + PRE}100 \quad (3)$$

Table 1: Comparison of several techniques for determining Video Shot Boundary

S. No.	Author, Year	Database	Methods of Algorithms	Evaluation Metrics	Remarks
1	[31],2014	TRECVid 2001 and 2005 test data	Similarity of edge-based features	P-96% R-98% F-97%	Performance is affected when multiple transitions exist in a single window.
2	[32],2016	TRECVid 2001 test data	Convolution Neural Net-work based on deep learning	F1 for C.T-98.8%, F1 for G.T-96.8%	For highly similar cases feature vector does not change and unable to locate soft transitions accurately.
3	[33],2017	TRECVid 2001 and 2007 video test dataset	SBD using Gist and MSER	R-89% P-91%	A common behavior is shared for soft transition but wipe transition is not considered.
4	[34],2017	TRECVid 2001 test data	GOFD index based adaptive threshold	P-91.8% R-89.8% F1-90.3%	Based on the total number of used frames threshold is selected automatically.
5	[35],2017	CSE2009 test functions	Grasshopper Optimization Algorithm (GOA), Multi Objective GOA.	Multi objective GOA	For solving multi objective algorithms effectiveness of different constraint handling techniques is less
6	[36],2019	TRECVid 2001 and 2007 video test dataset	Wipe transition detection using mean luminance pattern is nor included	R -94.3% P-90.1% F1-92.1%	Missing information about wipes transition detection to train the FNN for video analysis.
7	[37],2019	TRECVid 2001 and 2007 video test dataset	Gravitational Search Algorithm (GSA), PSO-GSA	R -94.1% P-79.2% F1-86%	Missing information about wipes transition detection to train the FNN for video analysis
8	[38],2022	TRECVid 2001 and 2007 video test dataset	Ant Lion Optimizer (ALO)	R -96.3% P-96.3% F1-96%	Missing information about wipes transition detection to train the FNN for video analysis
9	[49],2022	TRECVid 2001	Color histogram and SURF local feature	R=94.6 % P=95.1 % F1=94.8%	Local adaptive threshold is calculated for each primary segment boundary

11. Research Gap

Following a thorough assessment of the literature, signi ficant research gaps were found and are listed below.It is necessary to develop a technique that can extract the features from a potential video segment in order to save computational time and expense, as opposed to extract the features from an entire video.It must determine man y thresholds because a single threshold might not be suf ficient to detect both sudden and slow shot transitions[3 3].It is necessary to find ways to identify false transitio ns caused by changes in illumination, flash light, or mo tion.

12. Conclusion

Multimedia is the most crucial means of communicating with a group of people in the modern world. Accessing the needed data demands more processing time due to the growing volume of multimedia material. Video segmentation converts videos into smaller sizes and that is insufficient to address the problem. To handle that enormous amount of multimedia data, video indexing and retrieval are required. For such a large amount of multimedia data, manual indexing is ineffective. The application of KAZE based SBD has a significant impact in this situation. Hash based algorithms and methods can generate different Hash values for each of the detected cuts. To show the robustness of the proposed shot boundary detection system, RGB color histograms are created for each of the detected cuts. After analyzing the performance of the KAZE Hash-based abrupt transition detection system it becomes clear that the approach is efficient to use short videos with less variation. The approach is suitable for a dataset containing a set of short videos.

Bibliography

3. Roy M, Thounaojam D M, Pal S, Perceptual hashing scheme using KAZE feature descriptors for combinatorial manipulations, Multimedia Tools and Applications (2022) 81:29045–29073.
4. Kr. Biswas S, Gharde1 N D, Thounaojam D M, Soni1 B, Robust perceptual image hashing using fuzzy color Histogram. Multimedia Tools Appl (2018).
5. Sajjad M, Haq IU, Lloret J, Ding W, Muhammad K (2019) Robust image hashing based efficient authentication for smart industrial environment. IEEE Trans Ind Inf 15(12):6541–6550.
6. Neelima A, Singh KM (2016) Perceptual hash function based on scale-invariant feature transform and singular value decomposition. Comput J 59(9):1275–128.
7. Arambam N, Singh KM (2016) Perceptual hash function based on scale-invariant feature transform and singular value decomposition. Comput J 59:1275–1281.
8. Singh KM, Neelima A, Tuithung T, Singh KM (2019) Robust perceptual image hashing using sift and svd. Curr Sci 117(8):1340.
9. Tang Z, Dai Y, Zhang X, Huang L, Yang F (2014) Robust image hashing via colour vector angles and discrete wavelet transform. IET Image Process 8:3:142–149.
10. Tang Z, Dai Y, Zhang X, Huang L, Yang F (2014) Robust image hashing via colour vector angles and discrete wavelet transform. IET Image Process 8:3:142–149.
11. Zhao Y, Wang S, Zhang X, Yao H (2013) Robust hashing for image authentication using Zernike moments and local features. IEEE Trans Inf ForensSecur 8:1:55–63.
12. Tang ZJ, Zhang XQ, Dai YM, Lan WW (2013) Perceptual image hashing using local entropies and DWT. Imag Sci J 61:2:241–251.
13. Huang Z, Liu S (2020) Perceptual image hashing with texture and invariant vector distance for copy detection. IEEE Transactions on Multimedia.
14. Qin C, Hu Y, Yao H, Duan X, Gao L (2019) Perceptual image hashing based on weber local binary pattern and color angle representation. IEEE Access 7:45460–45471.
15. Gharde ND, Thounaojam DM, Soni B, Biswas SK (2018) Robust perceptual image hashing using fuzzy color histogram. Multimed Tools Appl 77(23):30815–30840.
16. Hamid H, Ahmed F, Ahmad J (2020) Robust image hashing scheme using laplacian pyramids. ComputElectr Eng 84:106648.
17. Hernandez, R.A.P., Miyataki, M.N. and Kurkoski, B.M. (2011) Robust Image Hashing using Image Normalization and SVD Decomposition. Proc. IEEE Int. Symp. Circuits and Systems (MWSCAS), Seoul, August 7–10, pp. 1–4. IEEE.
18. Jie Z (2013) A novel Block-DCT and PCA based image perceptual hashing algorithm. IJCSI Int J Comput Sci 10:399–403.
19. Tang Z, Wang S, Zhang X, Wei W, Su S (2008) Robust image hashing for tamper detection using non-negative matrix factorization. J Ubiquitous Converg Technol 2:18–26.
20. Zizhuo, W. A. N. G., et al. "Robust blind image watermarking based on interest points." Virtual Reality \& Intelligent Hardware 6.4 (2024): 308-322.
21. Lin, Chuchao, Changjun Zou, and Hangbin Xu. "SCNet: A Dual-Branch Network for Strong Noisy Image Denoising Based on Swin Transformer and ConvNeXt." Computer Animation and Virtual Worlds 36.3 (2025): e70030.
22. FFFN: Frame-by-frame feedback fusion network for video super-resolution. IEEE Transactions on Multimedia 25 (2022): 6821-6835.
23. Tang Z, Zhang X, Dai X, Yang J, Wu T (2013) Robust image hash function using local color features. Int J Electron Common (AEU) 67:717–722.
24. Tang Z, Yang F, Huang L, Zhang X (2014) Robust image hashing with dominant DCT Co-efficients. Optik-Int J Light Electron Opt 125:5102–5107.
25. Xiang S, Kim HJ, Huang J (2007) Histogram-based image hashing scheme robust against geometric deformations. In: Proc Int conf multimedia \& security. Dallas, pp 121–128.
26. Jie, Z. (2013) A novel block DCT and PCA based image perceptual hashing algorithm. Int J. Comput. Sci., 10, 399–403.
27. Mohtavipour, S.M., Saeidi, M. &Arabsorkhi, A. A multi-stream CNN for deep violence detection in video sequences using handcrafted features. The Vis Comput 38, 2057–2072 (2022). https://doi.org/10.1007/s00371-021-02266-4.

28. Nozaripour, A., Soltanizadeh, H. Image classification via convolutional sparse coding. Vis Comput 39, 1731–1744 (2023). https://doi.org/10.1007/s00371-022-02441-1.
29. Monga, V., and Evans, B. Perceptual image hashing via feature points: Performance evaluation and tradeoffs. Image Processing, IEEE Transactions on 15, 11 (Nov. 2006), 3452 -3465.
30. Tang Z, Wang S, Zhang X, Wei W, Su S. Robust image hashing for tamper detection using nonnegative matrix factorization. J Ubiquitous Convergence Technol 2008; 2:18–26.
31. Weng L, Preneel B (2009) Shape-based features for image hashing. Int Conf Multimed Expo, 1074–1077.
32. Ahmed F, Mohammed YS, Abbas VU (2010) A secure and robust hash-based scheme for image authentication. Signal Process 90:1456–1470.
33. Singh,A.,Thounaojam,D. M. &Chakraborty,S.: A novel automatic shot boundary detection algorithm: robust to illumination and motion effect,Signal, Image and Video Processing, 1–9, 2019.
34. Dutta, D., Saha, S. K., Chanda, B.: A shot detection technique using linear regression of shot transition pattern. Multimedia Tools and Applications, 75, 93-113, 2016.
35. Xu, J., Song, L. & Xie, R.: Shot boundary detection using convolutional neural networks. In: Visual Communications and Image Processing (VCIP), 1–4, 2016.
36. Thounaojam, D.M., Bhadouria, V.S., Roy, S. & Singh, K.M.: Shot boundary detection using perceptual and semantic information. Int J Multimed Inf Retr 6:167–174, 2017.
37. Kar, T. & Kanungo, P.: A motion and illumination resilient framework for automatic shot boundary detection. Signal Image Video Process, 11:1237–1244, 2017.
38. Chakraborty, S., Thounaojam, D.M.&Sinha, N.: A Shot boundary Detection Technique based on Visual Colour Information. Multimed Tools Appl, 80, 4007–4022, 2021.
39. Chakraborty, S.&Thounaojam, D.M.: A novel shot boundary detection system using hybrid optimization technique. Appl Intell, 49, 3207–3220, 2019.
40. Chakraborty, S., &Thounaojam, D. M.: A Novel Shot Boundary Detection System Using Hybrid Optimization Technique. Appl. Intell., 49, 3207–3220, 2019.
41. Suguna, R. M., Kalaivani, A., & Anusuya, S.: The Detection of Video Shot Transitions Based on Primary Segments Using the Adaptive Threshold of Colour-Based Histogram Differences and Candidate Segments Using the SURF Feature Descriptor. Symmetry, MDPI, 14, 1705–1720, 2022.
42. Chakraborty, S., Thounaojam, D. M., Singh, A., & Pal, G.: ALO-SBD: A Hybrid Shot Boundary Detection Technique for Video Surveillance System, 2022.