Tech Science Press

Coverless Video Steganography Based on Frame Sequence Perceptual Distance Mapping

Runze Li[1], Jiaohua Qin[1,*], Yun Tan[1] and Neal N. Xiong[2]

[1]College of Computer Science and Information Technology, Central South University of Forestry & Technology, Changsha, 410004, China
[2]Department of Mathematics and Computer Science, Northeastern State University, Tahlequah, 74464, OK, USA
*Corresponding Author: Jiaohua Qin. Email: qinjiaohua@163.com

Abstract: Most existing coverless video steganography algorithms use a particular video frame for information hiding. These methods do not reflect the unique sequential features of video carriers that are different from image and have poor robustness. We propose a coverless video steganography method based on frame sequence perceptual distance mapping. In this method, we introduce Learned Perceptual Image Patch Similarity (LPIPS) to quantify the similarity between consecutive video frames to obtain the sequential features of the video. Then we establish the relationship map between features and the hash sequence for information hiding. In addition, the MongoDB database is used to store the mapping relationship and speed up the index matching speed in the information hiding process. Experimental results show that the proposed method exhibits outstanding robustness under various noise attacks. Compared with the existing methods, the robustness to Gaussian noise and speckle noise is improved by more than 40%, and the algorithm has better practicability and feasibility.

Keywords: Coverless steganography; frame sequence; perception distance

1 Introduction

As the most popular communication medium today, digital video is gradually replacing image as the most influential communication medium in today's society with rich visual performance and huge information-carrying capacity. As an ideal steganographic communication carrier, digital video has attracted widespread attention from researchers in information hiding and has become one of the research hotspots in this field.

In the traditional video steganography algorithm [1–3], researchers mainly use the redundant characteristics of the carrier itself and the insensitive characteristics of human vision to embed meaningful secret information in it to achieve the effect of information hiding. However, this will inevitably leave traces of modification, which cannot cope with the detection of steganalysis tools [4,5].

To fundamentally resist the detection of steganalysis tools, Zhou et al. [6] proposed a new concept of "coverless" in May 2014. "Coverless" does not mean that no carrier is used, but not making any modifications to the carrier. The sender and the receiver hide and extract the secret information of the carrier data through the shared mapping rules [7,8], avoid the process of carrier modification, and therefore can completely resist steganalysis.

The breakthrough progress of deep learning in computer vision [9,10] also brings new ideas to coverless steganography and is widely used in text [11,12], image and video. Xiang et al. [13] proposed a robust text coverless information hiding method based on multi-index, which improved the algorithm's robustness. In the method based on the image, Luo et al. [14] proposed image coverless steganography based on multi-object recognition, which improved the hiding capacity and hiding rate. Liu et al. [15] proposed an image coverless hiding algorithm based on DenseNet feature image retrieval and DWT sequence mapping, which has better robustness and security against image attacks. Cao et al. [16] proposed a coverless information hiding method based on animation character generation, significantly increasing the hiding capacity. Pan et al. [17] proposed a video coverless steganography algorithm based on semantic segmentation in the single-frame video coverless steganography scheme. This is the first time that the coverless steganography solution has been applied to video carriers, bringing new ideas to the field of coverless steganography. Zou et al. [18] proposed a steganography scheme based on the combination of frames, which further improved the video coverless information hiding capacity. Meng et al. [19] proposed a coverless steganography algorithm based on the maximum DC coefficient, which has good capacity, robustness, and security performance. However, these steganography schemes do not effectively utilize the unique sequential features of the video, and their robustness is poor. Tan et al. [20] proposed a video coverless steganography algorithm based on video optical flow analysis in the scheme of multi-frame video coverless steganography. This algorithm has a good compromise between the hidden information capacity and robustness and has a higher hiding success rate and a lower transmission load. However, this scheme's optical flow calculation efficiency is low, and the time cost is high.

We have noticed that most of the existing work on video coverless information hiding only uses the characteristics of a specific frame of the video for normal mapping and does not use the rich temporal and spatial continuity of the video. In addition, there is still massive room for improvement in terms of the robustness of the steganography algorithm and the efficiency of hash generation.

In order to improve the robustness, we propose a coverless video steganography based on frame sequence perceptual distance mapping. The main contributions of this paper are summarized as follows:

1) We apply LPIPS algorithm to obtain the sequential features of the video, and to establish the relationship map between the feature and the hash sequence for information hiding. The robustness of the algorithm has been significantly improved.
2) The MongoDB database is used to store the mapping relationship and speed up the index access speed in the information hiding process, which improves the execution efficiency of the algorithm significantly.

The rest of this article is arranged as follows: Section 2 introduces related research, Section 3 detailly introduces the proposed method and shows the secret information transfer process, Section 4 gives experimental results and comparisons, and finally, we summarize this article in Section 5.

2 Related Work

There are many distance comparisons in the field of computer science, such as comparing the Hamming distance of binary strings [21], the edit distance of text files [22], and the Euclidean distance of vectors [23]. But it is difficult to measure the pixel level of two pictures because of human subjectivity.

Zhang et al. [24] proposed using the features mentioned in the neural network convolutional layer to calculate the cosine distance in the channel and cross-space dimensions. The computing framework is shown in Fig. 1. This distance quantifies the perceptual distance between images and reflects the similarity of the images.
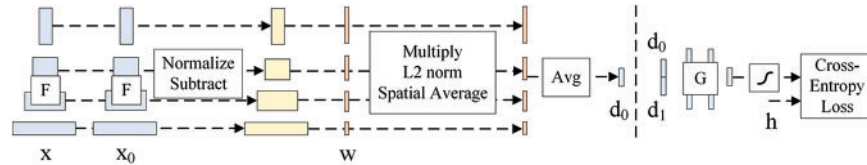

Figure 1: Computing distance

For a given network f, the perception distance d_0 of two similar blocks x, x_0 must be calculated. First, extract the features from the L layer, normalize the channel dimensions, and then scale to each channel by the vector ω and take the distance l_2. Then the spatial and network layer dimensions are averaged. Finally, a small network g is used to predict the distance h. The calculation formula is as follows:

$$d(x, x_0) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} ||w_l \odot (\hat{y}_{hw}^l - \hat{y}_{0hw}^l)||_2^2 \qquad (\hat{y}^l, \hat{y}_0^l \in R^{H_l \times W_l \times C_l}, \omega^l \in R^{C_1}) \tag{1}$$

3 The Proposed Coverless Steganography Scheme

This section will introduce the proposed video coverless information hiding algorithm, mainly composed of five parts: 1) Coverless video steganography framework based on frame sequence perception distance mapping. 2) Hash sequence generation algorithm. 3) The establishment of a video index database. 4) Secret information hiding algorithm. 5) Secret information extraction algorithm.

3.1 Coverless Video Steganography Framework Based on Frame Sequence Perception Distance Mapping

Fig. 2 shows the framework of the proposed coverless video steganography system. In this scheme, we use LPIPS algorithm to extract the sequential features from video database on cloud. The perceptual distance between multiple frames is mapped to a hash code through a hash sequence generation algorithm. We save the hash map results in an index database, which will be used to hide secret information. Then the sender splits the secret messages and matches them in the index database. The matching result is recorded as auxiliary information, which is sent by the sender. Finally, the receiver locates the video based on the auxiliary information to restore the secret information.
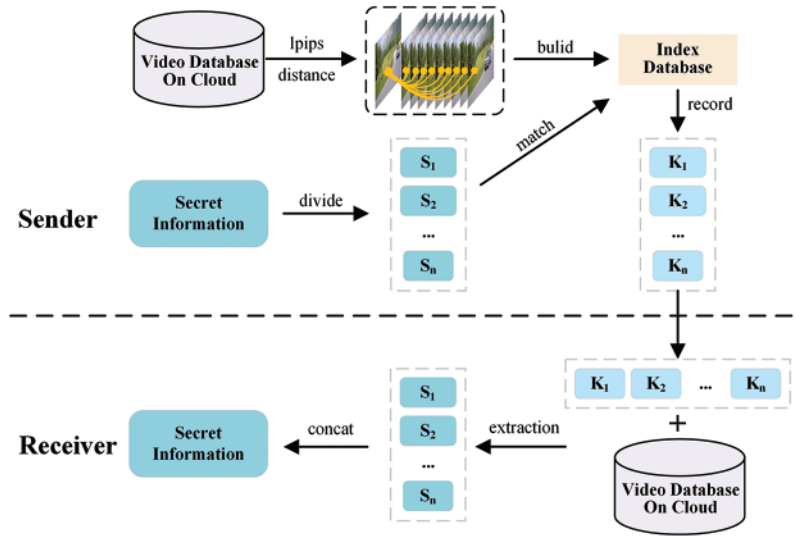
Figure 2: The framework of the proposed coverless video steganography system

3.2 Hash Sequence Generation Algorithm

The depth perception distance (LPIPS) can quantify the similarity of several adjacent video frames. In Fig. 3, the algorithm proposed in this paper fixes the first frame as the keyframe. After applying the depth perception distance (LPIPS) to generate M pairs of perception distance, it compares the consecutive M frames.

$$D = \{d_1, d_2, \ldots, d_M\} \tag{2}$$

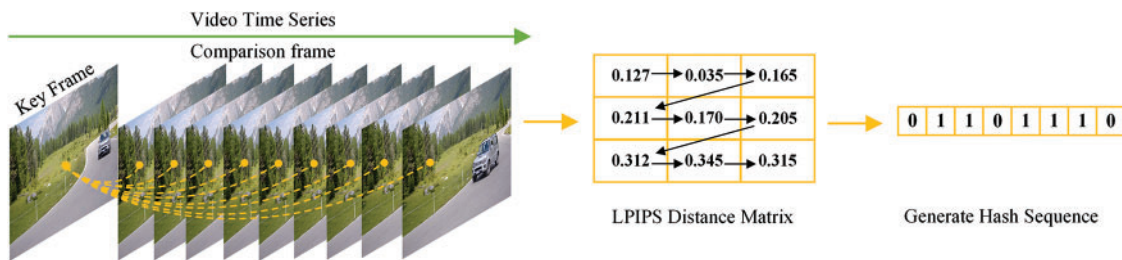**Figure 3:** Hash sequence generation

After feature mapping rules and hash formulas, a hash sequence of $M - 1$ bits is generated.

$$H = \{h_1, h_2, \ldots, h_{M-1}\} \tag{3}$$

Among them, h_i $(1 \leq i \leq M - 1)$ is obtained by the following hash generation formula.

$$h_i = \begin{cases} 0, & d_i > d_{i+1} \\ 1, & d_i \leq d_{i+1} \end{cases} \quad (1 \leq i \leq M - 1) \tag{4}$$

Among them, d represents two frames' depth perception distance (LPIPS) value.

Algorithm 1: Hash sequence generation

Input: single video data v, the index of the keyframe K, the number of frames for comparison M.
Output: hash sequence H.
1: function HashGeneration(v, K, M)
2:    Decompose video to pictures: P $\{p_1, p_2, \ldots, p_n\}$ = VideoToFrames(v)
3:    for $i = (K+1)$ to $(K+1+M)$ do
4:       Calculate the LPIPS distance between p_K with p_i: d_i = LPIPS(p_K, p_i)
5:       D.append(d_i)
6:    end for
7:    H = CalcHash(D)
8:    return H
9: end function
10:
11: function CalcHash(D)
12:    for $i = 1$ to length(D) -1 do
13:       if $D_i > D_{i+1}$ then
14:          H append $h_i =$ "0"
15:       else
16:          H append $h_i =$ "1"
17:       end if
18:    end for
19:    return H
20: end function

3.3 The Establishment of Video Index Database

To match the carrier more accurately and quickly, it is necessary to establish a video index database. This paper uses the method in 3.2 to obtain the depth perception distance of consecutive frames for hash mapping. In particular, we use a cyclic comparison algorithm, that is, to fix a specific frame of the video and compare the depth perception distance of the keyframe and the subsequent frame in the sequential sequence. For example, the video Bear.avi obtains 90 effective video frames through framing. After fixing the first frame and the following 9 consecutive frames for comparison, it is then fixing the second frame and the subsequent 9 consecutive frames for comparison. By analogy, a total of 80 consecutive comparisons are performed in the bear video set to obtain 80 sets of hash sequences.

In Fig. 4, we are accustomed to dividing the secret information into information fragments of equal length and then carrying out the matching selection of the carrier. The choice of a carrier is a very time-consuming operation. We use the MongoDB distributed storage database as the video index database to solve this problem. Its data format is similar to JSON objects and consists of key-value pairs.

Among them, *Hash Sequence (_id)* represents the primary key of the generated hash sequence, *IndexID* represents the index, *VideoID* represents the video information in the video collection, *FrameID* represents the keyframe, *SequenceNum* represents the number of consecutive frames to be compared.
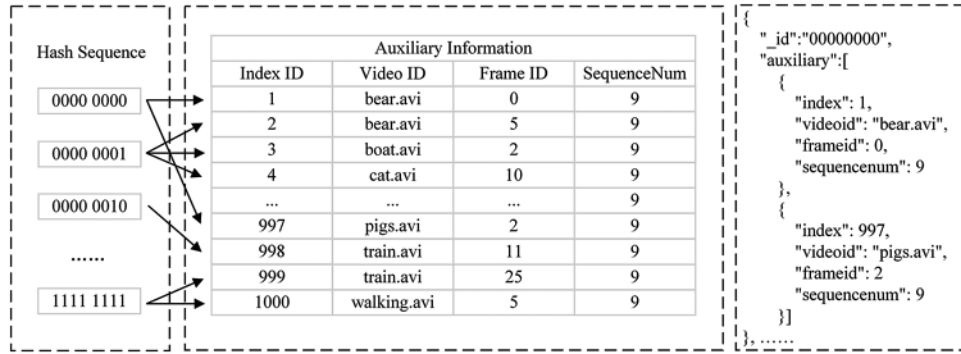
Figure 4: The establishment of video index database

Algorithm 2: The establish of video index database

Input: video datasets $V = \{v_1, v_2, \ldots, v_n\}$, the number of frames for comparison M.
Output: mapping the index $I = \{Ind_1, \ldots, Ind_j\}$.
1: function IndexDatabases(V, K, M)
2:    for $i = 1$ to length(V) do
3:       Decompose video to pictures: $P \{p_1, p_2, \ldots, p_m\}$ = VideoToFrame(v_i)
4:       for $j = 1$ to length(P)-M do
5:          H = HashGeneration(v_i, j, M)
6:          Update video index database in MongoDB: $H \rightarrow ind_j = \{indexID, v_i, j, M\}$
7:       end for
8:    end for
9: end function

3.4 Secret Information Hiding Algorithm

In coverless steganography, secret information hiding mainly uses the mapping carrier to find the appropriate secret information (hash sequence) from the carrier database. This section will introduce the secret information hiding algorithm in detail.

Step 1: Assuming that the secret information is divided into binary information fragments of equal length, the formula is as follows:

$$N = \begin{cases} \dfrac{L}{M}, & if \ L\%M = 0 \\ \dfrac{L}{M} + 1, & otherwise \end{cases} \tag{5}$$

where L represents the total length of the secret information S, M represents the specified fragment length, and N represents the number of fragments. When the secret information L cannot divide M, it will be completed in the form of 0, and the number of 0 will be recorded.

Step 2: After the fragmented secret information is mapped and matched in the video index database, the corresponding carrier is selected and recorded as the key information.

Step 3: Repeat Step 2 until all secret information is matched and combined to construct auxiliary information.

Step 4: Append the number information of the filled 0 to the end of the auxiliary information and send it to the receiver.

Algorithm 3: Information hiding

Input: mapping the index $I = \{ind_1, ind_2, \ldots, ind_j\}$, secret information S.
Output: auxiliary information $I' = \{ind_1, ind_2, \ldots, ind_N\}$.
1: padding secret information bits: $S'\{s_1, s_2, \ldots, s_m\} = \text{Padding}(S)$
2: divide S' into N segment:
3: for $i = 1$ to N do
4:    match s_i with I
5:    record the index and set auxiliary information $I' = \{ind_N\}$
6: End for
7: Send auxiliary information I' to the receiver

3.5 Secret Information Extraction Algorithm

On the receiver side, the receiver performs the restoration of the secret information in the following order.

Step 1: The receiver locates the cover carrier according to the auxiliary information.

Step 2: For each pair of keyframe combinations, extract the depth perception distance (LPIPS) from the cover carrier, and generate a hash sequence according to the algorithm rules in 3.2.

Step 3: Combine and connect all the hash sequences in order

Step 4: Cut according to the records filled with 0, obtain the secret information S.

Algorithm 4: Information extraction

Input: Video database $V = \{v_1, v_2, \ldots, v_n\}$, auxiliary information $I' = \{ind_1, \ldots, ind_N\}$.
Output: Secret information $S = \{s_1, s_2, \ldots, s_N\}$.
1: For $i = 1$ to N do
2:    Get VideoID v_i, FrameID f, SequenceNum M from ind_i
3:    $H_i = \text{HashGeneration}(v_i, f, M)$
4: end for
5: connect all the segments as: $H\{H_1, H_2, \ldots, H_M\}$
6: remove padding bits
7: secret information bits stream is recovered as $S = \{s_1, s_2, \ldots, s_M\}$

---

4 Experimental Results and Analysis

4.1 Experimental Environment

Experimental environment: Intel(R) Core (TM) i7-7800xCPU @3.50 ghz, 64.00GB RAM, two NVIDIA GeForce GTX 2080Ti GPU graphics cards. The experiments in this paper are all done on MatLab2016a and PyCharm engineering platforms.

4.2 Capacity

In the method proposed in this paper, the number of bits of the generated hash sequence is determined by the number of consecutive frames selected in the cyclic comparison algorithm. Therefore, for each video, its hidden capacity formula can be expressed as:

$$C = M - 1 \tag{6}$$

C represents the hidden capacity, that is, bits; M represents the number of consecutive frames selected for comparison.

Assuming that 10 consecutive frames are selected (1 keyframe, 9 consecutive frames), according to the algorithm proposed in this paper, the number of bits that can be hidden is 8 bits.

The comparison of hidden bits of information is shown in Tab. 1. It can be seen that in the proposed scheme, the capacity of information hiding is determined by the number of consecutive frames selected, and the larger the number of successive frames M selected, the greater the hidden capacity of the video hiding. This paper compares the proposed algorithm with existing schemes.

Table 1: Hidden bits number comparison

Methods	Pan's [17]	Zou's [18]	Tan's [19]	Proposed (M = 9)
Capacity	8 bit	16 bit	32 bit	8 bit

In the comparison method, $M = 9$ is used in this article; 9 consecutive frames are used to generate 8-bit hidden information. Compared with the existing methods, the algorithm is not high in capacity. This is also to consider finding a balance between the success rate of information hiding and the algorithm's robustness to improve the practicability and feasibility of the scheme.

4.3 Robustness Analysis

In the field of coverless steganography, robustness refers to whether the receiver can correctly extract secret information from the carrier after external attacks. Based on the *DAVIS-2017* and *UCF101* data sets, this paper adds different factors such as Gaussian noise, salt and pepper noise, speckle noise, and compressed images to test the robustness. The accuracy rate is calculated as:

$$ACC = \frac{\sum_{i=1}^{m} f(i)}{m}, f(i) = \begin{cases} 1, & if \ B_i = B_i' \\ 0, & if \ B_i \neq B_i' \end{cases} \tag{7}$$

Among them, B_i is the hash sequence generated by the original data set, B_i' is the hash sequence generated by the attacked data set, and m represents the number of frames of the video.

Most existing video coverless steganography algorithms use a single video frame for information mapping. Once the carrier is attacked, it is difficult to recover the secret information. Tab. 2 shows the single-byte extraction accuracy results for different types of attacks. It can be seen that compared with the current video coverless information hiding algorithm, the steganography scheme proposed in this paper has dramatically improved the robustness, and the anti-attack effect is balanced.

Table 2: Single byte accuracy with different attacks

Attack	Pan's [17]	Zou's [18]	Proposed
Gauss noise ($\sigma = 0.001$)	30.6%	50.9%	71.2%
Gauss noise ($\sigma = 0.005$)	29.8%	51.7%	71.4%
Speckle noise ($\sigma = 0.01$)	49.6%	52.5%	81.5%
Speckle noise ($\sigma = 0.05$)	30.0%	51.5%	71.9%
Salt & pepper noise ($\sigma = 0.001$)	83.3%	–	81.5%
Salt & pepper noise ($\sigma = 0.005$)	61.7%	–	74.8%
JPEG compression (Q(70))	77.2%	–	92.4%
JPEG compression (Q(90))	88.3%	85.9%	96.1%

The depth perception distance of the global feature used in the algorithm proposed in this paper is inherently robust. More importantly, the algorithm uses the similarity between several consecutive frames for hash mapping. Even if the carrier is attacked, the attack is also a constant attack, which is an attack on the whole video. The similarity relationship between them will not be significantly affected, so the proposed algorithm has strong robustness.

In addition, this paper also makes an experimental comparison for the generation of a single bit. The calculation formula of single-bit robustness is as follows.

$$ACC\_bit = \frac{\sum_{i=1}^{m} f(i)}{m}, f(i) = \begin{cases} 1, & if \ b_i = b'_i \\ 0, & if \ b_i \neq b'_i \end{cases} \tag{8}$$

Among them, b_i is the bit in the hash sequence generated by the original data set, b'_i is the bit in the hash sequence generated by the attacked data set, and m represents the number of frames of the video.

In Tab. 3, this paper compares the robustness of extracting single-bit secret information. Compared with the existing multi-frame-based video coverless information hiding methods, the proposed algorithm also has a good performance in single-bit robustness. Regarding image type attacks and video compression type attacks, the extraction accuracy of a single bit of the algorithm is maintained above 85%, and the robustness is high and balanced.

Table 3: Single bit accuracy with different attacks

Attack	Tan's [20]	Proposed
Salt & pepper noise ($\sigma = 0.001$)	99.9%	89.5%
Salt & pepper noise ($\sigma = 0.005$)	99.2%	87.0%
Salt & pepper noise ($\sigma = 0.01$)	98.8%	86.4%
Gauss noise ($\sigma = 0.001$)	70.1%	86.4%
Gauss noise ($\sigma = 0.005$)	64.9%	86.8%
Gauss noise ($\sigma = 0.01$)	62.0%	86.8%
Speckle noise ($\sigma = 0.001$)	82.4%	89.6%

(Continued)

Table 3: Continued

Attack	Tan's [20]	Proposed
Speckle noise ($\sigma = 0.005$)	81.0%	89.6%
Speckle noise ($\sigma = 0.01$)	80.0%	92.9%
Compressed MPEG-4 file with H.264	95.9%	90.5%
Compressed motion JPEG 2000 files (.mj2 file)	84.8%	90.1%

4.4 Hiding Success Rate Analysis

In coverless steganography, the capacity is judged by the number of bits of secret information that the same carrier can hide. Still, the success rate of hiding is considered by the adequate number of capacities that the same data set can represent. For the video coverless information hiding algorithm based on the feature mapping method, the extracted feature sequence ensures robustness and reflects the difference of features. This is a critical and challenging indicator for short video carriers and an essential indicator for measuring the feasibility and practicability of secret information transmission schemes. The calculation formula for hiding the success rate is as follows.

$$R = \frac{C_{eff}}{2^C} \tag{9}$$

where C represents the hiding capacity of the algorithm, C_{eff} represents the effective bit sequence count, and R represents the hiding success rate of the hiding algorithm.

In this paper, 200 videos are randomly selected from the *UCF101* and *HMDB51* video sets; the hiding success rate experiment is carried out in the same experimental environment and compared with the method based on semantic segmentation [20]. The experiment selected 9 consecutive frames to generate the hash sequence in the test.

The experimental results are shown in Fig. 5. It can be seen that the hiding success rate of the proposed method increases with the increase of the number of videos, and the hiding success rate can reach more than 90% when the number of videos is about 200. But under the limited number of videos, the hiding success rate is not as good as the method proposed by Tan et al. [20]. This is mainly because this scheme chooses a balance between robustness and embedding capacity to ensure the practicability and feasibility of the algorithm. Secondly, this scheme is based on the generation of hashes based on multiple frames of continuous video. Compared with generating hashes based on single-frame images, there are fewer materials available. Therefore, under the premise of a fixed number of videos, the success rate of hiding is relatively low.

4.5 Efficiency Analysis

This indicator mainly measures the algorithm's efficiency when generating a hash sequence based on the video carrier and constructing an index database. The video index database only needs to be once built at the sender. If a new carrier video is added, the sender only needs to update the database. In the process of hiding secret information and proposing secret information, the main time consumption is generating a hash sequence from the feature extraction of the video carrier. Therefore, the experiment compares the time cost of different methods in hiding the same length of information. The experimental results are shown in Tab. 4.
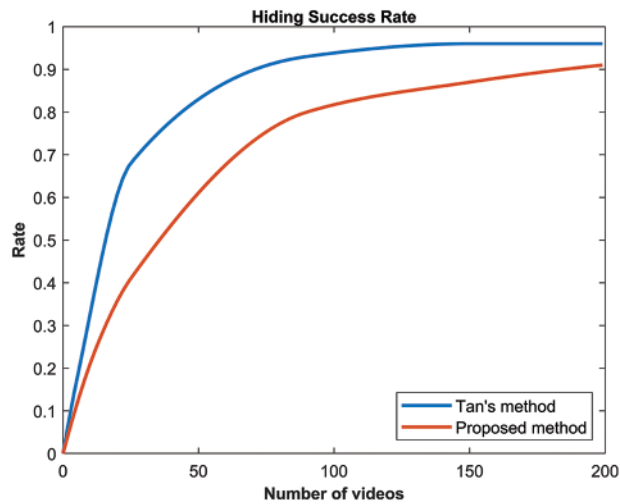
Figure 5: Hiding success rate comparison

Table 4: Time cost comparison

Method	Pan's [17]	Tan's [19]	Proposed
Time cost	1.3769 s/B	0.7416 s/B	0.0766 s/B

Compared with the video coverless information hiding algorithm based on semantic segmentation and optical flow calculation, the algorithm proposed in this paper has achieved good results in the efficiency of hash generation. This is mainly due to the use of cyclic comparison, which reduces the process of grouping and comparing the video after framing. The algorithm uses the overall features of the video frame for similarity quantification. The algorithm does not use the standard image block comparison scheme, which further improves the efficiency of hash sequence generation and reduces the time cost.

5 Conclusion

This paper proposes a video coverless information hiding algorithm based on frame sequence perceptual distance mapping. The algorithm introduces depth perception distance to quantify the similarity between consecutive video frames, selects keyframes to compare with time series frames, and builds a hash map index database. The receiver quickly locates the video sequence and calculates the perception distance based on the auxiliary information to extract the secret information. Compared with the existing video coverless steganography algorithm, the steganography algorithm proposed in this paper has significantly improved robustness. The algorithm can resist most attacks, especially Gaussian noise and speckle noise. At the same time, the complexity of the hash generation algorithm has also been significantly reduced, and the generation efficiency has been increased by more than 10 times, which has reasonable practicability and feasibility. However, the algorithm proposed in this paper does not have a high hiding success rate under the limited number of videos, which will be further optimized in the follow-up work.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Y. X. Liu, S. Y. Liu, Y. H. Wang, H. G. Zhao and S. Liu, "Video steganography: A review," *Neurocomputing*, vol. 335, pp. 238–250, 2019.

[2] Y. Z. Yao and N. H. Yu, "Motion vector modification distortion analysis-based payload allocation for video steganography," *Journal of Visual Communication and Image Representation*, vol. 74, pp. 102986, 2021.

[3] D. C. Nguyen, T. S. Nguyen, F. R. Hsu and H. Y. Hsien, "A novel steganography scheme for video H. 264/AVC without distortion drift," *Multimedia Tools and Applications*, vol. 78, no. 12, pp. 16033–16052, 2019.

[4] G. Feng, X. Zhang, Y. Ren, Z. Qian and S. Li, "Diversity-based cascade filters for JPEG steganalysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 2, pp. 376–386, 2020.

[5] M. Boroumand, M. Chen and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 5, pp. 1181–1193, 2019.

[6] Y. Cao, Z. L. Zhou, X. M. Sun and C. Z. Gao, "Coverless information hiding based on the molecular structure images of material," *Computers, Materials & Continua*, vol. 54, no. 2, pp. 197–207, 2018.

[7] Q. Liu, X. Y. Xiang, J. H. Qin, Y. Tan and Q. Zhang, "A robust coverless steganography scheme using camouflage image," *IEEE Transactions on Circuits and Systems for Video Technology*, pp. 1–1, 2021. https://doi.org/10.1109/TCSVT.2021.3108772.

[8] Z. L. Zhou, H. Y. Sun, R. H. Harit, X. Y. Chen and X. M. Sun, "Coverless image steganography without embedding," *International Conference on Cloud Computing and Security*, vol. 9483, no. 1, pp. 123–132, 2015.

[9] C. L. Wang, Y. L. Liu, Y. J. Tong and J. W. Wang, "GAN-GLS: Generative lyric steganography based on generative adversarial networks," *Computers, Materials & Continua*, vol. 69, no. 1, pp. 1375–1390, 2021.

[10] T. F. Zheng, Y. C. Luo, T. Q. Zhou and Z. P. Cai, "Towards differential access control and privacy-preserving for secure media data sharing in the cloud," *Computers & Security*, vol. 113, pp. 102553, 2022. https://doi.org/10.1016/j.cose.2021.102553.

[11] S. Zhao, M. H. Hu, Z. P. Cai, Z. J. Zhang, T. Q. Zhou *et al.,* "Enhancing Chinese character representation with lattice-aligned attention," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021. https://doi.org/10.1109/TNNLS.2021.3114378.

[12] S. Zhao, M. H. Hu, Z. P. Cai and F. Liu, "Dynamic modeling cross-modal interactions in two-phase prediction for entity-relation extraction," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2021. https://doi.org/10.1109/TNNLS.2021.3104971.

[13] L. Xiang, J. H. Qin, X. Y. Xiang, Y. Tan and N. N. Xiong, "A robust text coverless information hiding based on multi-index method," *Intelligent Automation & Soft Computing*, vol. 29, no. 3, pp. 899–914, 2021.

[14] Y. J. Luo, J. H. Qin, X. Y. Xiang and Y. Tan, "Coverless image steganography based on multi-object recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 7, pp. 2779–2791, 2021.

[15] Q. Liu, X. Y. Xiang, J. H. Qin, Y. Tan, J. S. Tan *et al.,* "Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping," *Knowledge-Based Systems*, vol. 192, no. 2, pp. 105375–105389, 2020.

[16] Y. Cao, Z. L. Zhou, Q. M. Wu, C. S. Yuan and X. M. Sun, "Coverless information hiding based on the generation of anime characters," *EURASIP Journal on Image and Video Processing*, vol. 2020, no. 36, pp. 1–15, 2020.

[17] N. Pan, J. H. Qin, Y. Tan, X. Y. Xiang and G. Hou, "A video coverless information hiding algorithm based on semantic segmentation," *EURASIP Journal on Image and Video Processing*, vol. 2020, no. 1, pp. 1–18, 2020.

[18] L. Zou, W. Wan, B. Wei and J. Sun, "Coverless video steganography based on inter frame combination," *Communications in Computer and Information Science*, vol. 1386, pp. 134–141, 2021.

[19] L. Meng, X. Jiang, Z. Zhang, Z. Li and T. Sun, "Coverless video steganography based on maximum DC coefficients," arXiv preprint arXiv:2012.06809, 2020.

[20] Y. Tan, J. H. Qin, X. Y. Xiang, C. H. Zhang and Z. D. Wang, "Coverless steganography based on motion analysis of video," *Security and Communication Networks*, vol. 2021, pp. 16, 2021.

[21] S. Grabowski and T. M. Kowalski, "Algorithms for all-pairs hamming distance based similarity," *Software: Practice and Experience*, vol. 51, no. 7, pp. 1580–1590, 2021.

[22] M. Boroujeni, M. Ghodsi and S. Seddighin, "Improved MPC algorithms for edit distance and Ulam distance," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 11, pp. 2764–2776, 2021.

[23] R. J. Zhou, X. Wang, J. Wan and N. X. Xiong, "EDM-fuzzy: An Euclidean distance based multiscale fuzzy entropy technology for diagnosing faults of industrial systems," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4046–4054, 2021.

[24] R. Zhang, P. Isola, A. A. Efros, E. Shechtman and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, New York, NY, USA, pp. 586–595, 2018.