Security Protection of Video-GIS Data Based on Data Encryption and Digital Watermarking

Yinguo Qiu[1], Jiaying Long[2], Cong Ma[2], Juhua Luo[1], Qitao Xiao[1]

[1] Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China – ygqiu@niglas.ac.cn; jhluo@niglas.ac.cn; qtxiao@niglas.ac.cn.
[2] School of Marine Technology and Geomatics, Jiangsu Ocean University, Lianyungang 222005, China - 13277889589@163.com; 13863720702@139.com.

Keywords: Video-GIS Data, Security Protection, Data Encryption, Sensitive Information Recognition.

Abstract

Video-GIS data, a novel generation of multimedia produced through the integration of geographic videos and GIS, has evolved into an innovative form of geospatial data representation. As a kind of digital products, Video-GIS data is vulnerable to unauthorized replication, dissemination, and usage. The potential for large-scale leakage of sensitive or confidential information, including commercial, political, military, and personal privacy data, poses a significant threat to the rights of data producers and may even compromise national security. Given this context, the primary concern is ensuring the security protection of video data in an open and shared data environment. To address this issue, we propose a novel scheme for the security protection of Video-GIS data in this paper, which collaboratively utilizes digital watermarking and data encryption. In the proposed scheme, Video-GIS data is divided into two categories, i.e., general and confidential, based on the presence or absence of sensitive information within the image frames. Appropriate security measures were meticulously designed for each data type to ensure effective prevention and control. Experimental results indicate that this watermarking algorithm exerts minimal influence on the quality of the Video-GIS data, while simultaneously maintaining optimal invisibility and robustness. The employed audio denoising method is characterized by its low complexity, simplicity, and feasibility, thereby ensuring the security of audio signal storage and transmission. This approach effectively realizes the dual protection objective for Video-GIS data, audio, and video.

1. Introduction

Within the burgeoning realm of Geographic Information Systems (GIS), traditional geospatial data formats are becoming increasingly insufficient to meet the escalating, multifaceted demands of users. These conventional data types, such as vector data, raster data, digital elevation models (DEM), and remote sensing images, are progressively revealing limitations in large-scale data analysis, intricate spatial relationship modelling, and detailed geographic information representation. As a result, there is an urgent need for innovative approaches and technological exploration to redefine the expression and storage methods of geospatial data to better align with user requirements (Qiu et al., 2022a). Video-GIS data, a novel form of geographic spatial data representation, is gaining increasing prominence and widespread application due to its ability to present geographic spatial scenes (Feng et al., 2014), such as terrain and landforms, in a more realistic and vivid manner.

Compared to conventional video data, Video-GIS data is characterized by its rich geographic spatial attributes (Xie et al., 2017), including land and resource monitoring, as well as water environment surveillance (Liu et al., 2022; Qiu et al., 2022b; Wang et al., 2023; Liu et al., 2024). By integrating video elements, Video-GIS can render real-world geographic information in a more realistic and vivid manner. This kind of method facilitates a three-dimensional (3D), intuitive presentation of geographic data, providing users with a unique "side-view" perspective. Furthermore, it exhibits a robust capability for measuring and analysing 3D spatial elements. However, given that video data is a digital product, it can be easily replicated, disseminated, and utilized without authorization (Artru et al., 2019). The unauthorized disclosure of large volumes of sensitive or confidential information, including commercial, political, military, and personal privacy data, could pose significant threats to the rights of data producers and even national security. Consequently, ensuring the security protection of video data within an open and shared data environment has become an urgent issue that requires immediate attention.

Traditional data protection methods utilize encryption to secure data during transmission, ensuring it remains encrypted within the channel (Potdar et al., 2009). Upon receipt, the data is decrypted into its original readable format. However, advancements in decryption technology and the detectability of encrypted information have led to concerns about data security and copyright infringement (Jiao et al., 2023). Digital watermarking technology addresses these issues by effectively protecting copyright and authenticating data content without compromising the quality of the original carrier data (Qu et al., 2024). A significant aspect of digital watermarking is video watermarking, which embeds crucial information (such as copyright or user details) into the carrier video, integrating seamlessly with the carrier data. This allows for the effective extraction of this information as evidence of copyright ownership when necessary (Peng et al., 2020). Video watermarking must be invisible and robust to counteract various attacks on video data, including compression, noise, filtering, and rotation attacks, to ensure secure sharing and copyright protection. Despite digital watermarking's ability to support copyright recognition and accountability for infringement, it cannot inherently prevent unauthorized use of original data.

Current methodologies for Video-GIS data security prevention and control predominantly utilize traditional video data security measures, with a primary focus on safeguarding the video image frames. These approaches often overlook the security concerns related to the audio component of the video, thereby presenting significant limitations. In practical applications, audio frequently contains sensitive information such as voice fingerprints, ID card numbers, and geographical coordinate data (Wang et al., 2017; Qiu et al., 2023). Moreover, Video-GIS

data, unlike traditional video data, encompasses a substantial amount of confidential or sensitive information, including military buildings, road signs, and license plates, which are not appropriate for public sharing or arbitrary distribution. However, traditional research has not adequately addressed this issue, leading to considerable security risks for Video-GIS data.

This paper presents a novel Video-GIS data security protection scheme that combines digital watermarking and data encryption technologies to address the challenges discussed. The uniqueness of this approach lies in three key aspects: (1) The system is capable of intelligently identifying classified and sensitive information within Video-GIS data, thereby establishing a basis for data classification and security protection. (2) For standard Video-GIS data, excluding confidential or sensitive information, it is possible to simultaneously achieve copyright protection of image frames and security control of audio signals. (3) In the context of sensitive Video-GIS data, including classified or sensitive information, it is feasible to concurrently achieve encryption protection for sensitive elements within image frames and security prevention and control of audio signals.

2. The Proposed Scheme

2.1 Main Idea

The main idea of the proposed scheme is shown in Figure 1.

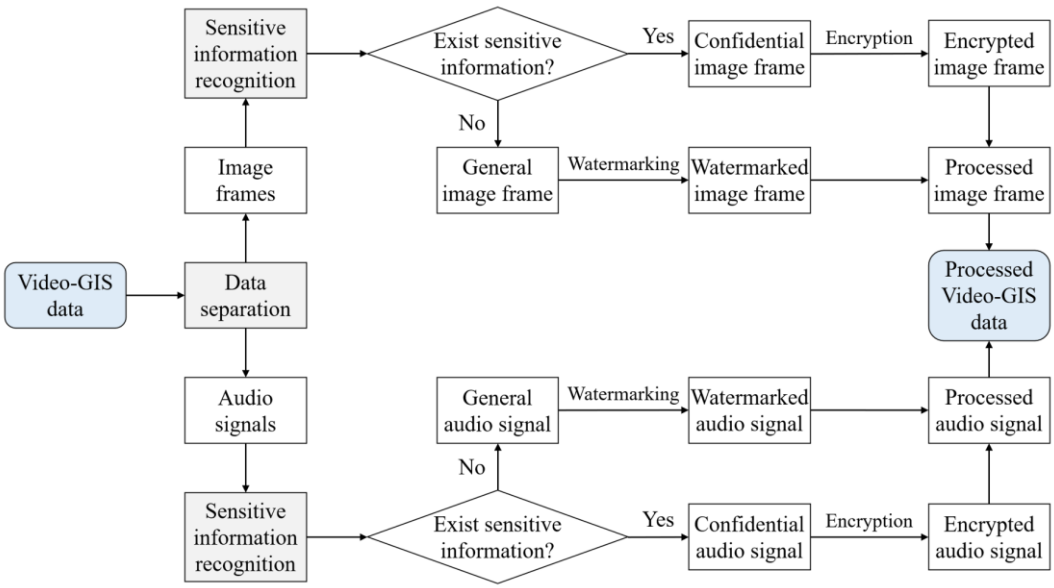


Figure 1. The main idea of the proposed scheme.

In the proposed methodology, data separation is initially performed on the Video-GIS data set to be processed. This process yields image frames and audio signals that constitute the original Video-GIS data. Subsequently, sensitive information recognition is conducted within both the image frames and audio signals. The Video-GIS data is then processed based on the outcomes of this sensitive information recognition. Ultimately, the processed image frames and audio signals are amalgamated to generate the final processed Video-GIS data.

2.2 Sensitive Information Recognition within Video-GIS Data

In this study, we utilize license plate information and geographical indication plate data as exemplars for the recognition of sensitive information within image frames. Furthermore, we designate several words and phrases as sensitive content in audio signals. In real-world applications, users can tailor the identification of sensitive information to their specific requirements, employing the same recognition methodology as presented herein.

The EAST algorithm (Feng et al., 2020), an efficient and precise scene text detection method based on deep neural networks, has garnered significant attention and practical application. When compared to conventional network models, the EAST algorithm demonstrates notable advancements in both accuracy and speed. This algorithm is capable of generating word-level and line-level prediction results with flexibility, thereby achieving end-to-end text detection. This implies that the entire process, from the input image to the output text detection result, is directly predicted by a single neural network. This network predicts quadrilateral text lines in any direction within the image. The network consists solely of the FCN stage and NMS stage, rendering it concise and elegant, and truly realizing end-to-end text detection.

Concurrently, a comprehensive database of sensitive vocabulary was developed, and an algorithm for detecting such terms in audio signals was formulated using machine learning techniques. During the detection process, the presence of sensitive content within the audio signal is ascertained by examining its entirety.

Upon the identification of sensitive information, Video-GIS data can be categorized as either general or confidential. General data refers to instances where sensitive information is present within image frames or audio signals, while confidential data denotes instances where such information is absent. The security prevention and control methods for both types of Video-GIS data will be elaborated upon in Sections 2.3 and 2.4, respectively.

2.3 Security Prevention and Control of General Video-GIS Data

The security prevention and control process for general Video-GIS data is depicted in Figure 2. This process encompasses three primary steps: the embedding of watermark data into image frames, the addition of noise data to audio signals, and the extraction of watermark data from the watermarked image frames.

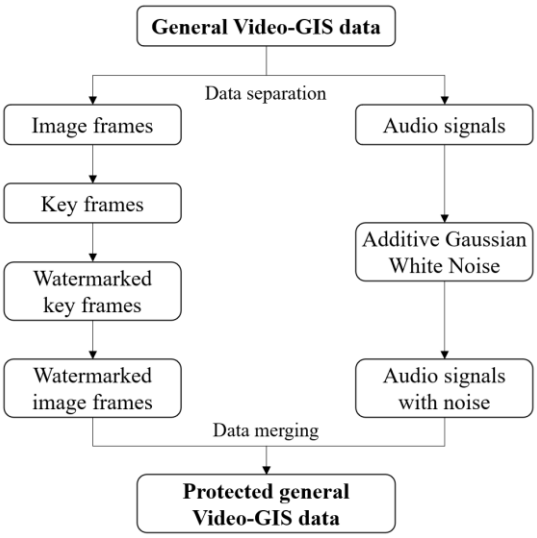


Figure 2. The process of securing general Video-GIS data.

2.3.1 Embed Watermark Data into the Image Frames: The watermark embedding process, given the original Video-GIS data (denoted as V) and the watermarking image (denoted as W), can be distilled into eight distinct sub-steps (Assume that the size of each image frame is $M_1 \times N_1$):

① Process watermarking image. Select a binary image of size $M_2 \times N_2$ to serve as the watermarking image, denoted as W. Utilize the chaotic sequence produced by Logistic mapping to process this watermark, resulting in the encrypted watermarking image, represented as W_1.

② Extract image frames from V. Securely retrieve each frame of the original carrier video utilizing a data separation method (Song et al., 2014).

③ Extract key image frame from the obtained image frames. The inter-frame difference for each frame of the carrier video should be calculated, and the frames exhibiting the local maximum average inter-frame difference intensity should be selected as the key frames of V.

④ The task involves transitioning the colour of keyframe images from the RGB space to the YUV matrix format.

⑤ Decompose the Y component of key frames. Perform DWT decomposition (Ren et al., 2021) on the Y component of keyframes to obtain Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH) sub bands, and then perform secondary DWT decomposition on the obtained LL sub bands to obtain LL_2, LH_2, HL_2, and HH_2.

⑥ Watermark Embedding. The process involves block-wise processing of the acquired LL_2 sub-bands. Each block undergoes Singular Value Decomposition (SVD) to yield a singular value matrix, within which the watermark is embedded into the first singular value. This embedding method is represented in Eq. (1):

$$\begin{cases} if\ W'_{(i,j)}=1,\ \lambda'_{i1}=\begin{cases}\lambda_{i1}-\mathrm{mod}\ (\lambda_{i1},\ \alpha)+3\cdot\alpha/4,\ \mathrm{mod}\ (\lambda_{i1},\ \alpha)\geq\alpha/4\\ \lambda_{i1}-\mathrm{mod}\ (\lambda_{i1},\ \alpha)-\alpha/4,\ \mathrm{mod}\ (\lambda_{i1},\ \alpha)<\alpha/4\end{cases}\\ if\ W'_{(i,j)}=0,\ \lambda'_{i1}=\begin{cases}\lambda_{i1}-\mathrm{mod}\ (\lambda_{i1},\ \alpha)+5\cdot\alpha/4,\ \mathrm{mod}\ (\lambda_{i1},\ \alpha)\geq 3\cdot\alpha/4\\ \lambda_{i1}-\mathrm{mod}\ (\lambda_{i1},\ \alpha)+\alpha/4,\ \mathrm{mod}\ (\lambda_{i1},\ \alpha)<3\cdot\alpha/4\end{cases}\end{cases} \quad (1)$$

where α = the strength for watermark embedding.

⑦ Generate the watermarked key frames. Firstly, the original singular value is replaced with the singular value derived from the embedding formula. Subsequently, Singular Value Decomposition (SVD) inverse decomposition is performed. This is followed by a secondary Discrete Wavelet Transform (DWT) inverse transformation. Finally, the matrix is reconstructed to synthesize keyframes that contain the watermark.

⑧ Secure the Video-GIS data. Replace all keyframes embedded with watermarks with their counterparts devoid of watermarks, and subsequently amalgamate them with other non-key frames to construct a comprehensive video.

2.3.2 Add Noise Data into the Audio Signals: Given the original Video-GIS data (denoted as V) and the noise data (denoted as N), the process of adding noise can be broken down into the following three distinct sub-steps.

① Adjust the length of N to match that of the original audio signal, denoted as S. Process Gaussian white noise to produce a white noise sequence of the same duration as the original audio. Let N_1 represent the resulting noise data.

② Reset the parameters of N_1. Collect the number of channels, sampling rate, sampling width, and sampling points from the original audio. Process these parameters to generate noise that mirrors those of the original audio. Let N' represent this newly generated data.

③ The process involves generating audio-protected Video-GIS data. Subsequently, the final noisy data is overlaid with the original audio to produce noisy audio data. The method is shown in Eq. (2), where β is the noise intensity, and S' is the protected Video-GIS data.

2.3.3 Extract watermark data from the watermarked image frames: This process is the inverse one of Section 2.3.1, which can be broken down into four distinct sub-steps:

① Extract key image frame from the obtained image frames, using the method introduced in sub-step ③ of Section 2.3.1.

② The task involves transitioning the colour of keyframe images from the RGB space to the YUV matrix format.

③ The Y component of keyframes should be decomposed. Subsequently, a second level Discrete Wavelet Transform

(DWT) decomposition is performed on the Y component of the image to acquire the LL_2 sub-bands for block partitioning. Each block is then subjected to SVD, resulting in a singular value matrix. The first singular value of this matrix is extracted. The watermark information is then derived using Eq. (2):

$$W'_{(i,j)} = \begin{cases} 1, \ \mathrm{mod}\ (\lambda'_{i1}, \ \alpha) > \alpha / 2 \\ 0, \ \mathrm{mod}\ (\lambda'_{i1}, \ \alpha) \le \alpha / 2 \end{cases} \quad (2)$$

where α = the strength for watermark embedding

④ The process involves generating a copyrighted watermarking image, followed by the use of a key to recover the watermark series. This step is crucial in obtaining the extracted watermark image.

2.3.4 Eliminate Noise Data from the Encrypted Audio Signals: This process is the inverse operation of Section 2.3.2, which can be broken down into four distinct sub-steps:

① Extract the noisy audio from S' and input a specified Gaussian white noise N.

② Determine the length of the noisy audio and process Gaussian white noise to match this length, resulting in a uniform white noise. Let N' represent the resulting noise data.

③ The process begins by determining the number of channels, sampling rate, sampling width, and sampling points of the noisy audio. These parameters are then processed as noise, maintaining consistency with the original audio signal's parameters. The resultant audio signal is denoted as N''.

④ The original audio data can be obtained by subtracting the noisy audio (S') from the final noisy data (N').

$$S = S' - \beta \cdot N'' \quad (3)$$

2.4 Security Prevention and Control of Confidential Video-GIS Data

The methodology for the security prevention and control of confidential Video-GIS data is depicted in Figure 3. Given the original Video-GIS data (denoted as V) and a watermarking image with dimensions $M_1 \times N_1$, the security prevention and control approach can be distilled into two distinct steps which will be introduced in detail in Section 2.4.1 and Section 2.4.2, respectively.

2.4.1 Watermark Embedding and Data Encryption of Both Image Frames and Audio Signals: This process can be broken down into the following ten distinct sub-steps:

① Upon examining each frame of the original carrier video, one must ascertain whether it contains confidential information. Subsequently, all frames that do contain such information should be selected. The watermark information is then repeatedly embedded into these confidential image frames.

② The process involves converting confidential video frame images from the RGB format to the YUV matrix format.

③ In the process of image classification, block the Y component of each frame. Subsequently, perform a Discrete Wavelet Transform (DWT) decomposition on each block to derive the Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH) sub-bands. Following this, conduct a SVD on the LL sub-bands to obtain singular value matrices.

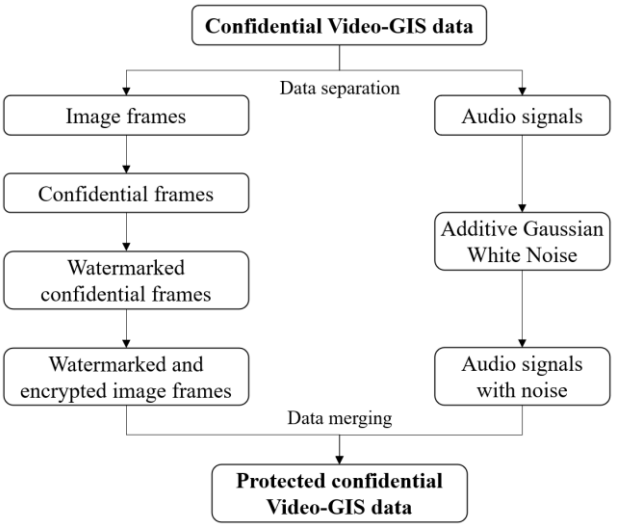


Figure 3. The process of securing confidential Video-GIS data.

④ Embed the watermark onto the maximum singular value within the singular value matrix. The function for watermark embedding is shown in Eq. (4):

$$\lambda' = \begin{cases} (\alpha - \frac{1}{2}) \cdot \mathrm{q} & \mathrm{mod}(W_{(i,j)} + \alpha, 2) = 1 \\ (\alpha + \frac{1}{2}) \cdot \mathrm{q} & \mathrm{mod}(W_{(i,j)} + \alpha, 2) = 0 \end{cases} \quad (4)$$

where α = the strength for watermark embedding

⑤ Iterate through all small blocks and embed the watermark information within each.

⑥ The process begins by substituting the original maximum singular value with the one derived from the embedding formula. This is followed by performing SVD inverse decomposition, then applying Discrete Wavelet Transform (DWT) inverse transformation. The final step involves reconstructing the matrix to generate a watermark that incorporates a confidential video frame.

⑦ The R component of confidential video frames, which contain watermarks G, undergoes logistic chaotic encryption applied to each colour channel. This process generates a watermarked ciphertext image frame.

⑧ The process involves substituting all watermarked ciphertext frames with their original, classified video counterparts. These are then amalgamated with non-classified image frames to generate comprehensive image frames.

⑨ The audio of the video should be extracted and a Gaussian white noise is read. Then, the same method as in Section 2.3 for audio signal processing is used to add the noise into the audio.

If there are sensitive information in the audio, the intensity of the noise will be much higher.

⑩ Integrate noisy audio data S' with the encrypted video data to construct a comprehensive Video-GIS dataset.

2.4.2 Watermark Extraction and Data Decryption of Image Frames and Audio Signals: This process is the inverse one of Section 2.4.1, which can be described as follows:

① Upon examining each frame of the original carrier video, compute the information entropy for each individual frame image. Subsequently, extract the watermark that contains the ciphertext frame image where the information entropy exceeds a threshold value of 'n'.

② The R component of a ciphertext frame image, which contains watermarks G, undergoes B-component logistic decryption. This process is applied to each colour channel to yield a video frame that encompasses both the watermarks and classified information.

③ The task at hand involves the conversion of watermarked confidential video frame images from the RGB format to the YUV matrix format.

④ The Y component of classified video frames should undergo $n \times n$ block processing. Subsequently, each block should be subjected to Discrete Wavelet Transform (DWT) decomposition to yield the Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH) sub-bands. The LL sub-bands obtained from this process should then be further subjected to SVD to procure singular value matrices.

⑤ The watermark information is extracted from the maximum singular value within the singular value matrix. The watermark extraction function is shown in Eq. (5):

$$\begin{cases} \beta = floor\left(\dfrac{\lambda'}{q}\right) \\ W'_{(i,\,j)} = \begin{cases} 1 & \mathrm{mod}(\beta, 2) = 1 \\ 0 & \mathrm{mod}(\beta, 2) = 0 \end{cases} \end{cases} \tag{5}$$

where β = the quantitative value of λ'

q = the quantization step

$W'_{(\mathrm{i},\,j)}$ = the extracted watermark value

⑥ It is necessary to traverse all small blocks in order to extract the complete watermark information.

⑦ Extract the noisy audio S' from the video, input the designated Gaussian white noise N , and utilize the methodology identical to that employed in Section 2.3 for the purpose of audio denoising.

3. Experiments and Discussion

Experiments were executed on a PC equipped with an Inter Core i7-8700 CPU operating at 3.20 GHz, 64 GB of RAM, and running Windows 10 Education 64-bit OS. The software used included ArcGIS Engine 10.2 and Visual Studio 2010.

3.1 Experiments of General Video-GIS Data

In this simulation experiment, we selected mp4 format video GIS data as the experimental Video-GIS data, which has a frame rate of 572 and a frame size of 1920 × 1056 pixels. We extracted ten keyframes from the carrier video using the inter-frame difference method, with an embedded strength set at 0.3. The watermark image measured 120 × 66 pixels in size.

3.1.1 Invisibility Analysis of Watermarked Image Frames: The invisibility of the experiment is quantified using the Peak Signal-to-Noise Ratio (PSNR). A higher PSNR value indicates a smaller discrepancy between the watermarked video frame and the original video frame. The experimental outcomes are presented in Table 1 (taking ten key frames as example), which demonstrates that the algorithms exhibit commendable invisibility.

Key frame	PSNR/(dB)
frame_01	43.9161
frame_02	43.8413
frame_03	43.8445
frame_04	43.9170
frame_05	44.0051
frame_06	43.7269
frame_06	43.8699
frame_07	43.9188
frame_08	43.8144
frame_09	43.9116
frame_10	43.9161
Average	43.8765

Table 1. The experiment results of invisibility analysis.

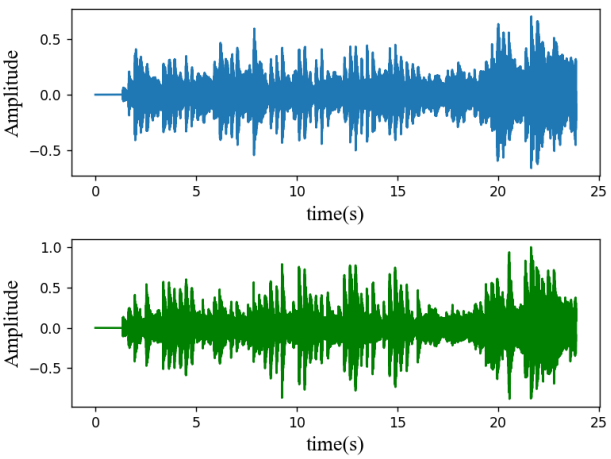
3.1.2 Robustness Analysis of Watermarked Image Frames: To assess the robustness of the watermark algorithm presented in this chapter, simulation experiments were executed on video frame images embedded with watermarks. These tests employed various attack methods, including compression, filtering, and noise. The normalized correlation (NC) value between the extracted watermark and the original was utilized as a metric to evaluate the algorithm's robustness. A closer NC value to 1 indicates a higher similarity between the extracted and original watermarks, thereby signifying superior algorithmic robustness. In these experiments, diverse attacks and intensities were applied to the video frames containing the watermarks. The results are detailed in Table 2.

3.1.3 Audio Signal Analysis: The audio noise experiment selected WAV audio extracted from the video, with a channel count of 2, a sampling byte length of 2, a sampling frequency of 44100, a total audio frame rate of 1051785, and a stacking strength of 2.3.

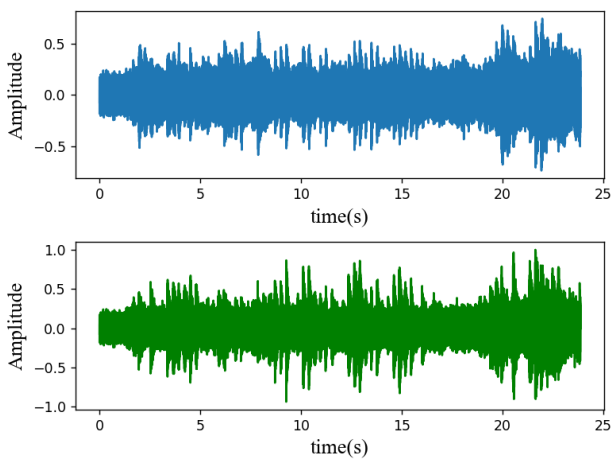
The result of audio signal analysis is shown in Figure 4. Here, (a) is the waveform of the original audio signal, (b) is the waveform of the audio signal containing noise data, (c) is the waveform of the audio signal obtained after correct denoising, and (d) is the waveform of the audio signal obtained after erroneous denoising.

Attack type	NC
JPEG compression (70)	0.9883
JPEG compression (50)	0.9342
JPEG compression (30)	0.9012
Clipping (1/16)	0.9851
Clipping (1/8)	0.9588
Clipping (1/4)	0.8483
Median filtering (1*1)	0.9999
Median filtering (3*3)	0.8215
Mean filtering (1*1)	0.9999
Mean filtering (3*3)	0.8803
Gaussian low-pass filtering (3*3)	0.9907
Gaussian low-pass filtering (5*5)	0.9902
Salt-and-pepper noise (0.001)	0.9983
Salt-and-pepper noise (0.002)	0.9895
Salt-and-pepper noise (0.003)	0.9768
Salt-and-pepper noise (0.005)	0.9151
Gaussian noise (0.001)	0.9948
Gaussian noise (0.003)	0.8877
Gaussian noise (0.005)	0.7047
Multiplicative noise (0.003)	0.9833
Multiplicative noise (0.005)	0.9418
Multiplicative noise (0.01)	0.8178
Poisson noise	0.9409

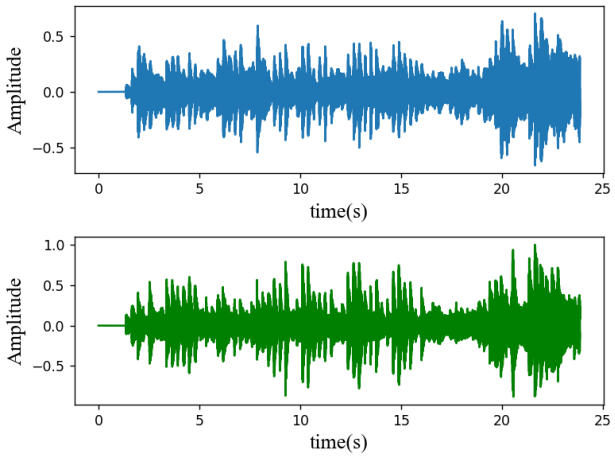
Table 2. The experiment results of robustness analysis.



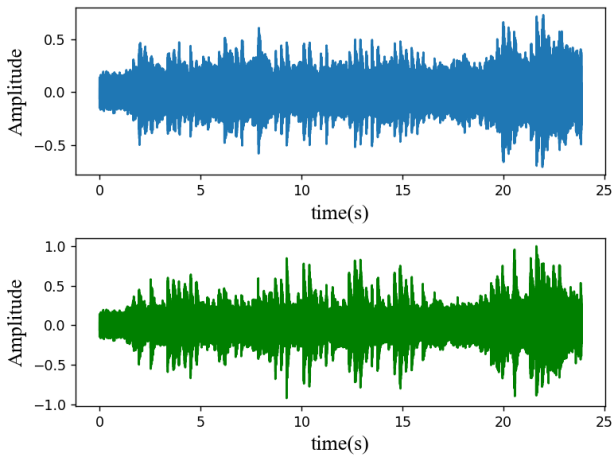
(a) The waveform of the original audio signal



(b) The waveform of the audio signal containing noise data



(c) The waveform of the audio signal obtained after correct denoising



(d) The waveform of the audio signal obtained after erroneous denoising

Figure 4. The result of audio signal analysis.

3.2 Experiments of Confidential Video-GIS Data

3.2.1 Invisibility Analysis: The invisibility of the experiment is quantified using the peak signal-to-noise ratio (PSNR). A higher PSNR value indicates a smaller discrepancy between pre- and post-watermark embedding, suggesting superior performance. Two frames from the classified video images are selected, one containing license plate information and the other geographic identification plate data. Figures 5 present the original classified frames with license plate details and the frames identified through license plate recognition from these video frames. Figures 6 depict the original classified frames alongside the geographic information frame images discerned from the video frames based on geographic identification.

(a) Confidential frame_01 (b) Sensitive information of (a)

(c) Confidential frame_02 (d) Sensitive information of (c)

Figure 5. The original video frames, which were classified and contained license plate information, and the frame that displayed the recognized license plate information.

The results of invisibility analysis are shown in Table 3. It can be seen from Table 3 that the algorithms exhibit commendable invisibility.

Sensitive image frame	PSNR/(dB)
Confidential frame_01	42.0288
Confidential frame_02	41.9943
Confidential frame_03	41.8643
Confidential frame_04	42.0298
Confidential frame_05	41.9762
Confidential frame_06	41.9107
Confidential frame_07	41.8767
Confidential frame_08	41.9313
Confidential frame_09	41.7970
Confidential frame_10	41.9210

Table 3. The experiment results of invisibility analysis.

3.2.2 Security Analysis: Conventional cryptographic systems are vulnerable to differential attacks when utilizing similar keys, thereby highlighting the importance of key sensitivity in ensuring the security of cryptographic algorithms. In this experiment, the sensitivity of the key was assessed by decrypting the ciphertext image with a variation of one thousandth of the key.

In this section, the initial key, denoted as K_1, is selected to encrypt two of the frames. Subsequently, K_2 and K_3 are introduced as slightly varied error keys, derived from altering one of the original keys.

$$\begin{cases} K_1 = \{x = 0.13,\ u = 3.98\} \\ K_2 = \{x = 0.1301,\ u = 3.98\} \\ K_3 = \{x = 0.13,\ u = 3.9801\} \end{cases}$$

The decryption of the encrypted image K_1 is achieved using K_2 and K_3, with the results depicted in Figure 6(c) and Figure 6(d) respectively. As observed from the charts, even a minute difference of one thousandth between the incorrect and correct keys prevents the acquisition of the accurate frame image. This underscores the high key sensitivity and robust security of the algorithm.

(a) Confidential frame (b) Decryption result using K_1

(c) Decryption result using K_2 (d) Decryption result using K_3

Figure 6. The decryption result of confidential frame_01 using slightly different keys.

3.2.3 Robustness Analysis: In order to ascertain the robustness of the watermark algorithm discussed in this paper, a series of simulation experiments were executed on encrypted video frames embedded with watermarks. These experiments employed various attack methods, including compression, filtering, and noise. The experimental results are shown in Table 4. The algorithm demonstrates robustness against prevalent attacks, including compression, noise, and filtering.

Attack type	NC
JPEG compression (70)	1.0000
JPEG compression (50)	1.0000
JPEG compression (30)	0.9996
Median filtering (3*3)	0.9995
Median filtering (5*5)	0.9772
Mean filtering (3*3)	0.9975
Mean filtering (5*5)	0.9225
Gaussian low-pass filtering (3*3)	0.9999
Gaussian low-pass filtering (5*5)	0.9997
Salt-and-pepper noise (0.001)	0.9972
Salt-and-pepper noise (0.003)	0.9787
Salt-and-pepper noise (0.005)	0.9477
Gaussian noise (0.001)	0.9969
Gaussian noise (0.003)	0.9324
Gaussian noise (0.005)	0.8128
Multiplicative noise (0.001)	0.9964
Multiplicative noise (0.003)	0.9514
Multiplicative noise (0.005)	0.8970
Poisson noise	0.9840

Table 4. The experiment results of robustness analysis.

3.3 Discussion

Video-GIS data, a novel form of information, holds significant value across various sectors of the national economy. Concurrently, the security concerns associated with this data are increasingly gaining attention. Traditional solutions predominantly treat Video-GIS data as general video or image data (frames), often neglecting sensitive information. To address this gap, we have developed a method for recognizing sensitive information in Video-GIS data and subsequently implemented distinct security protection measures for both general and confidential data. In our study, sensitive information is defined to include license plate details and geographical indication plate data within image frames, as well as specific words and phrases in audio signals. It is important to note that these elements serve merely as examples; users should define their own specific elements based on their applications.

Simulation experiments have substantiated the efficacy of this scheme, suggesting its potential to offer robust security protection for various types of data in a shared environment. These include traffic monitoring data, scenic area surveillance data, and wetland monitoring data.

4. Conclusions

This paper presents a thorough investigation into the security protection of Video-GIS data. Initially, a strategy for classifying and protecting Video-GIS data was formulated. Subsequently, sensitive information within the Video-GIS data was identified, leading to a division of the data into two categories: general and confidential. This classification is based on the presence or absence of sensitive information in the image frames. Finally, specific security measures were meticulously designed for each category to ensure effective prevention and control. The experimental outcomes have substantiated the efficacy of the proposed scheme.

In the future work, there are two aspects worth further in-depth research. Firstly, in this work, the robustness of watermark algorithms is not sufficiently complementary, and the attack set used to simulate attacks on video frames embedded with watermarks only includes select representative methods. This does not encompass all potential attacks that may occur during image transmission. The subsequent step should involve expanding the attack set and exploring more superior and complementary digital watermark algorithms to better resist these attacks. Moreover, when embedding watermarks into confidential image frames within Video-GIS data, it is imperative to initially undertake block processing and subsequently apply DWT to each block. However, when the frame size of the classified video is excessively large during SVD transformation, issues such as slow computation speed and extended computation time may arise. Consequently, further optimization and enhancement can be implemented to render the watermark algorithm more user-friendly, increase operational speed, and reduce execution time.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (Grant No. 42101433) and the National Key Research and Development Program (Grant No. 2023YFD1702104-04).

References

Artru, R., Gouaillard, A., Ebrahimi, T., 2019: Digital watermarking of video streams: Review of the state-of-the-art. *arXiv:1908.02039*. doi.org/10.48550/arXiv.1908.02039

Feng, J., Song, H., 2014: Analytical method for mobile elements in geo-video using random graph grammar. *Geomatics and Information Science of Wuhan University*, 39: 206-209. doi.org/10.13203/j.whugis20120711

Feng, X., Liu, Q., Wang, Z., 2020: Port container number detection based on improved EAST algorithm. *Journal of Physics Conference Series*, 1651, 012088. doi.org/ 10.1088/1742-6596/1651/1/012088

Jiao, Y., Ma, C., Luo, J., Qiu, Y., 2023: Security protection of 3D models of oblique photography by digital watermarking and data encryption. *Applied Sciences-Basel*, 13, 13088. doi.org/10.3390/app132413088

Liu, J., Xia, C., Xie, H., Wang, X., Qiu, Y., 2022: Accurate Monitoring of Algal Blooms in Key Nearshore Zones of Lakes and Reservoirs Using Binocular Video Surveillance System. Water, 14, 3728. doi.org/10.3390/w14223728

Liu, F., Qiu, Y., Zhang, C., Hou, J., Wang, J., Liu, J., Long, J., Ma, C., Zhang, M., 2024: Real-time quantitative monitoring method of cyanobacterial blooms in lake riparian zones based on video surveillance and application research. *Chines Journal of Environmental Engineering*, 18, 614-622. doi.org/10.12030/j.cjee.202309058

Peng, F., Zhang, X., Lin, Z.X., Min, L., 2020: A tunable selective encryption scheme for H. 265/HEVC based on chroma IPM and coefficient scrambling. *IEEE Transactions on Circuits and Systems for Video Technology*, 30, 2765-2780. doi.org/ 10.1109/TCSVT.2019.2924910

Potdar, U., Talele, K., Gandhe, S.T., 2009: Comparison of MPEG video encryption algorithms. *International conference on advances in computing, communication and control 2009*, pp. 266-271. doi.org/10.1145/1523103.1523163

Qiu, Y., Duan, H., Xie, H., Ding, X., Jiao, Y., 2022a: Design and development of a web-based interactive twin platform for watershed management. *Transactions in GIS*, 26, 1299-1317. doi.org/10.1111/tgis.12904

Qiu, Y., Duan, H., Wan, N., Gao, R., Huang, J., Xue, K., Peng, Z., Xiao, P., 2022b: Design and practice of a platform for monitoring, early-warning and simulation of algal blooms in Lake Chaohu. *Journal of Lake Sciences*, 34, 38-48. doi.org/10.18307/2022.0102

Qiu, Y., Liu, H., Liu, F., Li, D., Liu, C., Liu, W., Huang, J., Xiao, Q., Luo, J., Duan, H., 2023: Development of a collaborative framework for quantitative monitoring and accumulation prediction of harmful algal blooms in nearshore areas of lakes. *Ecological Indicators*, 156, 111154. doi.org/10.1016/j.ecolind.2023.111154

Qu, C., Du, J., Xi, X., Tian, H., Zhang J., 2024: A hybrid domain-based watermarking for vector maps utilizing a complementary advantage of discrete fourier transform and singular value decomposition. *Computers & Geosciences*, 183, 105515. doi.org/10.1016/j.cageo.2023.105515

Ren, N., Zhao, Y., Zhu, C., Zhou, Q., Xu, D., 2021: Copyright Protection Based on Zero Watermarking and Blockchain for Vector Maps. ISPRS Int. J. Geo-Inf., 10, 294.

Song, H., Liu, X., Lv, G., Zhang, X., Wang, F., 2014: Video Scene Invariant Crowd Density Estimation using Geographic Information Systems. *China Communications*, 11, 80-89. doi.org/ 10.1109/CC.2014.7004526

Wang, Z., Liu, X., 2017: Analysis of burglary hot spots and near-repeat victimization in a large Chinese city. *ISPRS International Journal of Geo-Information*, 6, 148. doi.org/10.3390/ijgi6050148

Wang, Z., Wang, C., Liu, Y., Wang, J., Qiu, Y., 2023: Real-time identification of cyanobacteria blooms in lakeshore zone using camera and semantic segmentation: A case study of Lake Chaohu (Eastern China). *Sustainability*, 15, 1215. doi.org/10.3390/su15021215

Xie, Y., Wang, M., Liu, X., Wu, Y., 2017: Surveillance video synopsis in GIS. *ISPRS International Journal of Geo-Information*, 6, 333. doi.org/10.3390/ijgi6110333