Applied Soft Computing 34 (2015) 744–757

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Fast retrieval of hidden data using enhanced hidden Markov model in video steganography

Mritha Ramalingam, Nor Ashidi Mat Isa*

Imaging and Intelligent System Research Team (ISRT), School of Electrical and Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, Malaysia

ARTICLE INFO

Article history:
Received 3 October 2014
Received in revised form 10 May 2015
Accepted 22 May 2015
Available online 17 June 2015

Keywords:
Video steganography
Data hiding
Enhanced hidden Markov model
Conditional states
Fast data retrieval

ABSTRACT

In the digital world, secure data communication has an important role in mass media and Internet technology. With the increase in modern malicious technologies, confidential data are exposed at a greater risk during data communication. For secured communication, recent technologies and the Internet have introduced steganography, a new way to hide data. Steganography is the growing practice of concealing data in multimedia files for secure data transfer. Nowadays, videos are more commonly chosen as cover media than other multimedia files because of the moving sequence of images and audio files. Despite its popularity, video steganography faces a significant challenge, which is a lack of a fast retrieval system of the hidden data. This study proposes a novel video steganography technique in which an enhanced hidden Markov model (EHMM) is employed to improve the speed of retrieving hidden data. EHMM mathematical formulations are used to enhance the speed of embedding and extracting secret data. The data embedding and retrieving operations were performed using the conditional states and the state transition dynamics between the video frames. The proposed EHMM is extensively evaluated using three benchmark functions, and experimental evaluations are conducted to test the speed of data retrieval using differently sized cover-videos. Results indicate that the proposed EHMM yields better results by reducing the data hiding time by 3–50%, improving the data retrieval rate by 22–77% with a minimum computational cost of 20–91%, and improving the security by 4–77% compared with state-of-the-art methods.

1. Introduction

Steganography is the process of hiding the existence of data in a carrier (or a cover-medium) such as text, image, audio, or video. Modifying a cover-medium to embed data into it facilitates secret communication. In digital steganography, the concealment of secret data is performed with a minimal effect on the interpretation of the original medium. Video steganography is an emerging sub-field of digital steganography. Many video steganography techniques have been proposed in recent years.

Fig. 1 shows the generic form of the video steganography process. The cover-video is the carrier of the secret message. Using a data hiding algorithm, the secret message is embedded in the cover-video, producing a stego-video. The stego-video is transmitted over the Internet. The receiver uses the data retrieval algorithm to extract the hidden message from the stego-video.

This study proposes a video steganography method that focuses on the fast retrieval of secretly embedded data. The rest of this paper is organized as follows. Section 2 reviews the related literature on the existing methods. Section 3 presents the proposed video steganography system using enhanced hidden Markov model (EHMM). Section 4 discusses the EHMM algorithm used for data hiding and retrieval. Also described in this section are the experimental studies that illustrate the efficiency of the proposed approach. Section 5 provides the concluding remarks.

2. Literature review

Several authors have proposed different steganography methods to maintain secrecy by hiding data in multimedia files such as text, audio, image, and video [1]. A considerable amount of data can be hidden in multimedia files through controlled alterations. The perceptual quality of the cover-medium, in particular video, is preserved after these alterations. Therefore, video is a better

* Corresponding author. Tel.: +60 45996051; fax: +60 45941023.
E-mail addresses: mritha2011@gmail.com (M. Ramalingam), ashidi@usm.my (N.A.M. Isa).

http://dx.doi.org/10.1016/j.asoc.2015.05.040
1568-4946/

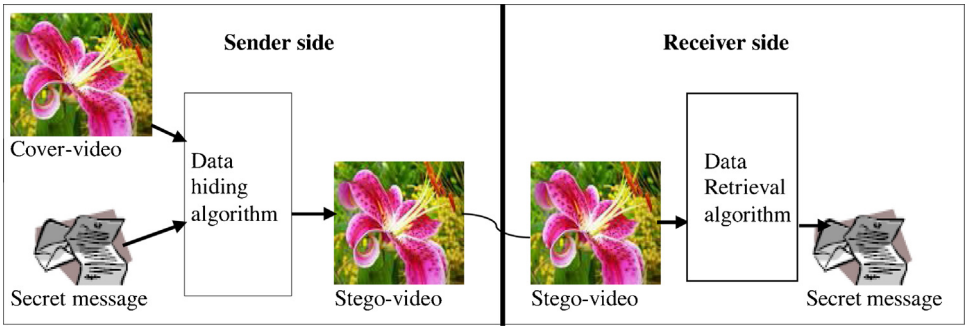

Fig. 1. General video steganography process.

cover-medium of choice for transporting large amounts of secret data [2].

A recent video steganography method conceals secret data in moving picture experts group (MPEG) videos using multivariate regression and flexible macro-block ordering (MRFB) to examine the accuracy of secret message retrieval and the imperceptibility of the data in the resulting videos [3]. In [3], two data hiding approaches were proposed. The first approach described hiding the secret data bits by adopting a quantization level of a steady bit rate video. The second approach used flexible macro-block properties to conceal the data bits.

Another digital steganography approach hides secret data in selected frames of MPEG videos using bit-plane complexity segmentation (BPCS) [4]. The BPCS method uses a highly secured data embedding process for data hiding. This method has a high embedding rate. However, the method encounters distortions in the resulting videos. The data hiding method in [6] hides the secret data in motion vectors of compressed video based on their associated prediction error. Other steganography techniques [5–7] extract data from frames with minimum distortion.

A data embedding method [8] embeds secret data in the vector quantization indices of a single set of compressed image blocks using the Chinese remainder theorem. This method performs better in protecting data from attacks. Hidden messages can be extracted without using the original image host [9,10], achieving enhanced data recovery. The effectiveness of video steganalysis in determining hidden data has been examined over the spatial and temporal domains by Tasdemir and Kurugollu [11]. An adaptive integer wavelet transforms based data hiding method provided rendering payload to increase the absolute visual quality of the stego-image [12]. Alternatively, another steganalysis method uses the Markov random process to detect the presence of hidden messages that uses the altered statistical properties of the image [13].

Hidden Markov models (HMM) are based on Markov chains. A Markov chain is a representation of a stochastic process. A Markov chain-based approach evaluates both inter- and intra-block correlation properties for feature extraction in the steganalysis process [14]. HMM have two important properties: (1) a Markov chain does not use memory for any state and (2) the conditional probabilities of all states do not depend on the position (time) in the sequence [15]. HMM are useful in formulating the conditional states that are used in steganographic systems [16].

The use of the Markov process has increased in the last several years because of two reasons: first, the mathematical structure of HMM is rich and therefore forms the theoretical basis for many applications; second, HMM are suitable for several important applications that do not require memory space [17]. A text-based steganography method hides short information in online communications using Markov chains [18]. A Markov chain model that provides an analytically detectable structure for data hiding with spatial dependencies is proposed in [19].

Wang and Yu [20] used a Markov model in a reversible data hiding scheme based on the histogram modification technique. This scheme had greater hiding capacity and less image distortion, but the computation time and speed of retrieval of the hidden data need to be improved.

To the best of authors' knowledge, only Wang et al. [21] have reported the implementation of a fast video steganographic system. All other steganography techniques [22–25] focus on increasing the embedding capacity. In addition, a low-distortion data embedding method using pixel-value differencing is proposed in [26], offering large payloads.

A Markov chain-based steganography method analyzes the steganographic capacity of Markov covers using square root law (SRL) [27]. The method has significant inferences in steganography and steganalysis, but the measure of payload is not resolved. However, according to these studies, the experimental results had considered the quantity of the hidden data, accuracy of the extracted data, and imperceptibility of the resulting stego-images.

In contrast to the classical approaches, the proposed method emphasizes fast data embedding and extracting using EHMM. The objective of the proposed system is to improve the speed of data hiding and extracting processes in video steganography by integrating the computing techniques of conditional states and state transition dynamics.

3. Enhanced hidden Markov model

This section first discusses the motivation for the use of Markov models in the proposed algorithm. Second, the section describes the design of EHMM, demonstrates HMM, and discusses conditional states and state transition dynamics.

3.1. Motivation

The design of the Markov model is based on the Markov chains that are formed during the extraction of hidden data. To minimize the overhead complexity, the Markov model simplifies the process by considering all possibilities from a given state to another. Therefore, the quality of the data generated by the Markov chain is improved considerably [28].

The HMM performs data embedding by tracking colored objects and applying mathematical tools to model these objects in the spatial domain [29]. The data retrieval speed of most existing transform domain-based video steganography systems is significantly lower because more time is consumed in data embedding and extraction [8]. Thus, the HMM is used in the proposed work to improve the data retrieval rate. The main objective of the proposed EHMM video steganography system is to enhance the speed of the data retrieval process.