Computers and Electrical Engineering 68 (2018) 14–25

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Hiding information in videos using motion clues of feature points☆

Mahdi Hashemzadeh

Faculty of Information Technology and Computer Engineering, Azarbaijan Shahid Madani University, Tabriz, Iran

ARTICLE INFO

Keywords:
Information hiding
Data hiding
Video steganography
Adaptive steganography
Feature points
Motion analysis

ABSTRACT

Motivated by the weaknesses of human visual system in understanding the changes in dynamic scenes, a new video steganography method is proposed. There are two main ideas in this method: (1) detecting highly dynamic regions in video scenes and use them to hide the data, and (2) determining the right amount of data to be embedded in the selected regions. Motion clues of the feature points are used to study the dynamics of the scenes, and then, select the regions of interest accordingly. To determine the embedding capacity for each pixel, some statistical indicators extracted from behaviors of the feature points are used. The experimental results on a large database show that the proposed method achieves a substantial performance improvement over the existing methods; the average embedding capacity and the perceptual invisibility rate values are 0.52 bpp and 50.66 dB respectively.

1. Introduction

With the growth of digital information and the increasing need for computer networks for their exchange, their security and privacy have also become highly important. The Internet is being developed dramatically and has become a popular medium for transferring information. However, as a completely open system, it may pose several risks and losses. Eavesdropping and data theft by hackers or hostile individuals can have disastrous consequences for most of today's information systems. In certain organizations, such as the military or commercial institutions, in which a lot of confidential information is exchanged, the scope of these attacks is broader and their consequences are worse.

One of the approaches researchers use now to provide solutions to secure the exchange of secret information is the steganography technique [1]. In this approach, the secret information to be exchanged between two parties is embedded inside other data, which is called the "cover media," so that such hidden messages are invisible to a third party. The cover media containing the hidden messages is called "stego media" that is sent to the receiver. Fig. 1 shows the general block diagram of the steganography concept.

The perceptual invisibility of a hidden message and the embedding capacity are two important factors to evaluate different steganography methods [1]. The primary goal of any steganography algorithm is to maximize the embedding capacity and minimize the distortions in the cover media. The cover media can be a text, an image, an audio or a video clip. Video streams usually have a large number of consecutive frames including many redundant bits that by replacing them with the data of the secret message, usually minor changes are made in the scenes. Thus, this media is considered as convenient cover media for hiding data.

Hiding information in a video clip can be viewed as an extension of hiding information in an image. A video file consists of a series of images (frames). Each of them can be used to hide a part of the secret message. Therefore, various techniques, such as [2,3], which have been proposed for image steganography, can be used in video steganography as well. The most common approach used for image steganography is the least significant bit (LSB) substitution technique [1]. It replaces some LSBs of pixel values of the cover

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Area Editor Dr. E. Cabal-Yepez.

E-mail address: hashemzadeh@azaruniv.ac.ir.

https://doi.org/10.1016/j.compeleceng.2018.03.046
Received 5 July 2017; Received in revised form 23 March 2018; Accepted 26 March 2018

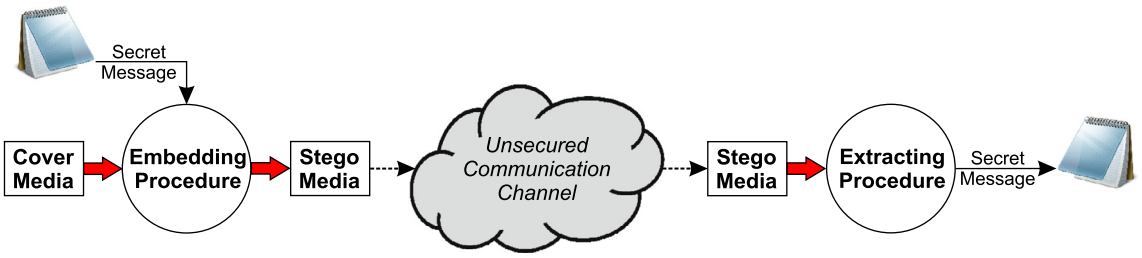

Fig. 1. General block diagram of the steganography concept.

image with the secret message bits [4]. Despite its simplicity, it can hide a large amount of data in the images. Hence, most of the existing video steganography techniques, such as [5–8], have been inspired by this technique [1].

Compared with images, there is much more potential in videos. Videos provide us with a dynamic environment, including new dimensions, such as the time and motion components to be used for hiding secret information. Because of the dynamic content of this media, the possibility of detection and extraction of the hidden message in it is much lower compared with images. Thus, in recent years, various video steganography methods have been proposed in the literature [5–14]. Although promising results have been achieved, operations such as finding the appropriate image regions for hiding the information, or investigating the motion components to enhance the system performance, are rarely performed in the majority of existing approaches [1].

Recently, a new class of substitution-based approaches, namely the adaptive video steganography methods, has been proposed. They are sometimes referred to as “Masking” or “Statistics-aware embedding” [1]. An adaptive approach usually works by studying the statistical features of the cover video before embedding the secret data [15]. This process helps in identifying the appropriate regions of the images to hide data in. These are called the Regions-Of-Interest (ROIs). For example, in the method proposed in [16], only some video frames, whose image histogram had certain changes, were used to embed the information. Or, in the approach presented in [5] only foreground regions of images were used to hide information. The method in [15] selected the pixels from the regions with a non-homogeneous texture to embed information. This is because these regions are originally noisy and it is very difficult to detect the distortions caused by the embedding of extra information. In [17], a video steganography technique was proposed which uses the frames containing scene-changes to hide the data. Scene changes were detected using the discrete cosine transform coefficients of the video sequences. Method in [18], used only the human face in the video to embed the secret information. The idea behind this approach was that the dynamic features that exist in the human face make it difficult to detect the distortions in these regions. However, we believe that in the case of video steganography systems, we can enjoy a lot more benefits from the weaknesses of the human visual system (HVS) in understanding the changes within the scenes.

Research has proved that the HVS is more sensitive to changes in the flat regions of images than other regions, such as the highly textured parts, the dynamic and moving regions, and the edges or corners [1,19]. Considering these limitations of the HVS, we propose a new video steganography method. The main idea in our method is to detect dynamic foreground regions in video scenes and use them to hide the information. To this end, we propose to use the motion clues of feature points (FPs) in the images to study the dynamics of the scenes and select the regions of interest. These points are easy to detect and track along consecutive frames, while they can give us very useful information from the nature of the scene. Example applications are presented in [20–22].

FPs are usually selected in the regions of the image which have a complex texture, such as the body, clothes, faces, a crowd, the leaves of trees, the waves of the sea; or, they are selected in parts of the image with very high dynamics, like a waterfall, or fast-moving objects, or even static elements, like the background of a natural scene, such as a field of flowers, grass, mountains, and even unnatural objects, such as room decoration, flooring, and many such things that can be found in almost any video. In our opinion, we can take advantage of this nature of the FPs to identify the salient, dynamic, and moving regions of video images, and hide information in these regions. Since the texture and the dynamics of these points and the surrounding regions are usually high, the changes made to the gray (or color) level of their pixels, due to embedded information, can be rarely perceptible by HVS. This is exactly the goal we pursue in the video steganography. Moreover, the second idea behind our method is to utilize motion information of FPs to determine the right amount of embedding for each detected region. This helps us to efficiently maximize the embedding capacity of our steganography method.

In the proposed method, in particular frames of the video, FPs existing in the scene are detected and tracked in the next frame. Motion vectors of FPs together with a simple morphological operation are used to extract the ROIs. Also, some statistical indicators, which can determine the number of bits replaceable in every pixel of the detected regions, are calculated from the spatial and temporal behaviors of FPs. Finally, the LSB algorithm is used to embed the secret data into the extracted regions based on the determined capacities. The proposed method is tested and evaluated on a large database comprising 12 videos. The results show that the proposed method performs well and has a larger embedding capacity than the existing methods.

The rest of this paper is organized as follows: Section 2 describes the details of the proposed method. The results of our experiments are presented and compared with other approaches in Section 3. Finally, in Section 4, we draw the conclusions and present the ideas for future work.