Journal of Applied Sciences 10 (15): 1644-1649, 2010
ISSN 1812-5654

Hiding Data in Video File: An Overview

[1]A.K. Al-Frajat, [1]H.A. Jalab, [1]Z.M. Kasirun, [2]A.A. Zaidan and [2]B.B. Zaidan
[1]Department of Software Engineering, Faculty of Computer Science and Information Technology, University Malaya, 50603 Kuala Lumpur, Malaysia
[2]Department of Computer System and Technology, Faculty of Engineering, Multimedia University, Jalan Multimedia, 63100 Cyberjaya, Malaysia

Abstract: The aim of this review is to study the methods of steganography using the video file as a cover carrier. The steganography is the art of protecting the information through embedding data in medium carrier, for instants this study illustrate historically this art, as well as the study describes methods as a review for this art in the video file. The video based steganography can be used as one video file, separated images in frames or images and audio files. Since that, the use of the video based steganography can be more eligible than other multimedia files. As a result of this study, the video based steganography has been discussed and the advantages of using the video file as a cover carrier for steganography have been proposed.

Key words: Data hidden, steganography, video file, hiding in MPEG, hiding in video

INTRODUCTION

The term of hide information is the process of covering some private or secrete data to make sure that there is no another party can disclose or altering it (Al-Azawi and Fadhil, 2010). Under this topic we can drive two techniques which are used to hide information one is the digital watermarking this technique used for two main things the first is for authentication purpose and the second is to demonstrate the intellectual property rights, the another technique is the steganography which is used mainly for hiding information from any unauthorized party, in this case the aim is to prevent the message being detected by any other party (Kawaguchi and Eason, 1998; Majeed *et al*., 2009).

Steganography and encryption are both used to ensure data confidentiality. However, the main difference between them is that with encryption anybody can see that both parties are communicating in secret. Steganography hides the existence of a secret message and in the best case nobody can see that both parties are communicating in secret. This makes steganography suitable for some tasks for which encryption are not, such as copyright marking. Adding encrypted copyright information to a file could be easy to remove but embedding it within the contents of the file itself can prevent it being easily identified and removed. The formal definition for the steganography is the art and science of communicating in such a way that the presence of a message cannot be detected. Simple steganographic techniques have been in use for hundreds of years, but with the increasing use of files in an electronic format new techniques for information hiding have become possible (Shirali-Shahreza and Shirali-Shahreza, 2008).

Information hiding is a technique of hiding secret using redundant cover data such as images, audios, movies, documents, etc. This technique has recently become important in a number of application areas. For example, digital video, audio and images are increasingly embedded with imperceptible marks, which may contain hidden signatures or watermarks that help to prevent unauthorized copy (Jalab *et al*., 2009). It is a performance that inserts secret messages into a cover file, so that the existence of the messages is not apparent. Research in information hiding has tremendous increased during the past decade with commercial interests driving the field (Jalab *et al*., 2009).

HIDING DATA

The steganography has been used for long time before. The main use for it was for military and government messages, nowadays; the approaches of steganography become widely used for many purposes (Zaidan *et al*., 2009). Anyway, the researchers provide and found out many approaches while others enhanced

Corresponding Author: Ali K. Al-Frajat, Department of Software Engineering, Faculty of Computer Science and Information Technology, University Malaya, 50603 Kuala Lumpur, Malaysia

the methods and the approaches of the steganography in order to improve the steganographic applications, in this part of the thesis a discussion about the existent methods or approaches for the steganography and how the researchers improve it within more than 10 years will be presented.

One of the latest techniques that have been used in this area by researchers at the Mount Sinai School MOUNT SINAI Medical in New York in 1999, as they managed to hide the secret texts in chromosome strand human DNA by using a technique called genetic system coverage (Genomic Steganography) and by placing signs resolution to be agreed upon in the nuclei chromosomes and then integrate these with millions sentences and sent to the other end. To extract the secret message is soaking get special distinction sentences used on the other and then placed under the microscope to extract the required text.

The oldest authentications on steganography taken from the legendary stories Greeks Herodotus and then back to the fifth century BC, these sources indicate that they felt they fly head of the Messenger and then write the secret letter in the head, leaving hair to grow then be sent to the required which is a re-extraction letter. (Johnson and Jajodia, 1998; Zaidan *et al*., 2008).

Authentications and other writing secret messages on the wood panels and then covered wax and will be hid those writing panels appear free of anything. And they were killing their animals as rabbit example corner confidential letter inside it.

Other means that the common use since the first century AD, invisible inks Invisible Inks, which was able to write a confidential letter with any other non-value-confidential and usually write between lines, for example those rabbis some fruit juices Fruit Juices, milk, urine, vinegar and all these species become dark and visible when exposed to heat the written document.

Then these kinds of inks evolved with the evolution of science chemical was used vehicles carrying chemical characteristics of the same old species with a more accurate and efficient have been used during the first and second world wars in the military secrecy of correspondence. Other technical methods has been used during world war II is sending a message hidden within another message is not relevant and based on the idea of a nomination letters every word of the letter counterfeit representation of characters from the characters letter requested confidentiality (Johnson and Jajodia, 1998).

The earlier application of text based steganography founded during world war II; the Germans would hide data as microdots. This involved photographing the message to be hidden and reducing the size so that it could be used as a period within another document. FBI director J. Edgar Hoover described the use of microdots as the enemy's masterpiece of espionage.

A message sent by a German spy during World War II read:

Apparently neutral's protest is thoroughly discounted and ignored. Isman hard hit. Blockade issue affects for pretext embargo on by-products, ejecting suets and vegetable oils.

By taking the second letter of every word the hidden message Pershing sails for NY June 1 can be retrieved.

More recent cases of steganography include using special inks to write hidden messages on bank notes and also the entertainment industry using digital watermarking and fingerprinting of audio and video for copyright protection (Naji *et al*., 2009).

Moreover, there are numerous ideas for the same method is used to be more than characters, or take certain words or phrases within the text fake and leaving the rest. Finally, it should be noted that the senior researcher in the area of concealment and science-based organization itself is German Johannes Trithemius)) between 1462-1526 and the oldest books in the area of coverage Posted by Gaspari Schotti)) in 1665 in the name of (Steganographyica) and (400) contains a page where all the ideas included (Trithemius).

Nowadays, the steganography application for sharing secure message used the multimedia files as a cover carrier for the secure message, since that many approaches has been proposed to use different type of the covers to send the secure message.

The text-based steganography is so limited in capacity and it is the easiest approach to be altered even accidently since it use the visible text to hide the secure message. In addition, there is no measurement can be used for the text-based steganography to assure the confidentiality of the secure message.

The image based-steganography tried to improve the capacity where in the literature more than 50% of the original image size has been used to hide the secure messages. Since, there is a limitation on how much information can be hidden into an image (Chang *et al*., 2002), making difficult to use the image methods, then in order to help to increase collaborative documents security. The video based steganography has been found to overcome the capacity problem, the video consist of a number of images placed in a frames to be presented in sequence one after the others. We can use any image within the video to hide the secure message with it, the use of video based steganography has another advantage as the discloser will facing a problem to attack the image since the sequence of the image within the video is

unknown for the attacker, so the attacker need to check all the images within the video which make it more difficult to attack the secure message. As well as the video-based steganography has the lowest chances of being suspicious because of the quickly displaying of the frames so it's become harder to be suspected by the human vision system.

The capacity problems overcame by Noda *et al.* (2004) the proposed method was based on wavelet compression for video data and Bit Plane Complexity Segmentation (BPCS) steganography. The author enhanced the capacity problems by using the video-based steganography.

A video error correction using steganography has been proposed in (Robie and Mersereau, 2002) since the transmission of any data is always subject to corruption due to errors, then the video transmission (because of its real time nature) must deal with these errors without retransmission of the corrupted data. However, the study proposed another application for the steganography rather than for security purposes.

On another hand, Jalab *et al.* (2009) proposed collaborate approach for select frame using Bit Plane Complexity Segmentation (BPCS) for hiding data within MPEG Video. The proposed approach invented a high secure data hidden using select frame from MPEG Video. However, the proposed approach achieved a high capacity using video as a cover carrier but the steganography alone unable to achieve high secure system for message sharing purposes.

Eltahir *et al.* (2009) proposed approach for video steganography using the Least Significant Bits (LSBs). The method considered the digital video file as separated frames and changed the output image displayed on each video frame by hidden data that does not visually change the image. With this technique, one can apply hidden information with more space better than other steganography media. However, the authors used the video-based steganography to enhance the capacity of the hidden message but the security requirements such as data integrity has not appeared in the study (Eltahir *et al.*, 2009).

While Bhaumik *et al.* (2009) suggested another video-based steganography. By using AVI videos which are large in size but still can be transmitted from source to target over network after processing the source video by using the data hiding and extraction procedure securely. There are two different procedures, which are used in this study at the sender and receiver sides respectively. The procedures are used as the key of data hiding and extraction processes. However, the suggested key in this study does not meet the security standards and it is unable to ensure the authentication and integrity of the data or the message.

Wu *et al.* (2003) applied multilevel embedding to allow the amount of embedded information that can be reliably extracted to be adaptive with respect to the actual noise conditions. The authors proposed strategies for handling different embedding capacity from region to region within a frame as well as from frame to frame. The authors also embed control information within the video to facilitate the accurate extraction of the user data payload and to combat such distortions.

However, the enhancements of the security in steganography approaches can be achieved by integrate the steganography with other techniques. In (Chae and Manjunath, 1999) a video-based steganography proposed in which the embedded signature data is extracted without knowing the original host video. The proposed method enables high rate of data embedding. Embedding is based on texture masking and utilizes a multi-dimensional lattice structure for encoding signature information. Signature data is embedded in individual video frames using the block DCT. However, the author claim that this approach is robust for motion compensated codes, such as MPEG-2, without showing the proof for the robustness.

While in (Socek *et al.*, 2007) the author proposed video encryption algorithm designed for both lossless and lossy low-motion spatial-only video codec. The proposed encryption method can thus be performed before compression at the encoder side and after decompression at the decoder side. As well the author introduced a new type of steganography as an extension to the encryption approach. The proposed steganographic scheme enables disguising a video with another video, which is a new concept in video-based steganography. However, the encryptions usually increase the size of the cipher-text which is affecting the capacity of the secret message.

On the other hand, Westfeld and Wolf (1998) presented the steganography in a video conferencing, the video conference used for the implementation of the steganographic system presented in this study works on the H.261 standard. However, the H.261 is not adequate to be used on the Internet, where just low bit-rate is available (De Oliveira, 1997).

Zaidan and Zaidan (2009) proposed a collaborate approach between steganography and cryptography. The approach provided a high secure data hidden using Public Key Infrastructure (PKI) method. However, the security aspect has been considered in this study, although the size of the cipher-text is a genuine problem for steganography. Furthermore, PKI encryption has been proposed for the purpose of integrity. A

solution of using hash function instead of PKI can give faster processing, less size for authentication of the message.

The use of video as a carrier cover for the secure message is overcame the capacity problem and added small enhancement to the security aspects. The integration of steganography and cryptography techniques provided powerful systems for sharing secure messages.

HIDING DATA IN VIDEO FILE

Steganography in video files based on exploiting the YCbCr colour space: YCbCr or Y'CbCr is a family of colour spaces used as a part of the Color image pipeline in video and digital photography systems. YCbCr represents colours as a combination of three values:

- Y = The luminosity (roughly the brightness)
- Cb = The chrominance (roughly colour) of the blue primary
- Cr = The chrominance (roughly colour) of the red primary (Green is achieved by using a combination of these three values)

This technique is based on YCbCr. YCbCr space that can remove the correlation of R, G and B in a given image, as less correlation between colours means less noticeable distortion. In this study they concentrate on human video images, more specifically on human skin tones or colours for data hiding, note that human skin colour can range from almost black to nearly colourless-appearing reddish white due to the blood vessels under the skin-in different people. The objective of this study is to combat the use of forged passport documents or national identity cards, a security measure would be to embed individuals' information in their photos (Su *et al*., 2008).

Transform domain embedding: This is a more complex way of hiding information in an image. Various algorithms and transformations are applied on the image to hide information in it. DCT (Direct Cosine Transformation) is one such method, which is used in JPEG compression algorithm to transform successive 8×8 pixel blocks of the image, into 64 DCT coefficients each (Umbaugh, 1997).

DCT helps separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality), embedding in DCT domain is simply done by changing DCT coefficients, for example by changing the least significant bit of each coefficient. One of the constraints of embedding in DCT domain is that many of the 64 coefficients are equal to zero and changing too many zeros to non-zero values will have an effect on the compression rate. That is why the number of bits one could embed in DCT domain, is less that the number of bits one could embed by the LSB method (Ahmed *et al*., 2010).

Least significant bit insertion (Doerr and Dugelay, 2004): This is a very popular method because of its simplicity, in this method; the LSB bit of 1 byte in the image is used to store the secret data. The resulting changes are too small to be recognized by the human eye. Since, this technique uses each pixel in an image, a lossless compression format like bmp or gif has to be used for the image. If lossy compression technique is used, some of the hidden information might be lost.

Considering video as separate images (Shirali-Shahriza, 2006; Kharrazi *et al*., 2004): In this method, each video frame is considered as a separate image, in which information is hidden. The main advantage of this method is the possibility of using the algorithms used in image steganography and watermarking for video, but it requires a large amount of computation. The algorithm we have suggested is almost similar to this method.

Real-time video steganography (Shirali-Shahriza, 2006): This technique involves hiding the information on the output image of the instrument (such as image displayed by an electronic advertising billboard). This method considers each frame that is shown by the machine at any moment, irrespectively of whether it is photo, text, or else, as an image. Then the system divides the image into small blocks. If the pixel colours of the blocks are similar, it changes the colour characteristics of a number of these pixels to a certain extent so data information is hidden in the image.

In the following section the embedding of data into video based steganography. The embedding of data within video file start by selecting the desired video, after selecting the video, the system should read all the video frames and assign frame number to each frame and then the desired frames can be selected for further processing. Note, by selecting different frames sequence at every time provide more security and being harder to the attacker from attacks the frame due to unknown frame sequence.

The extraction of the data from the video file illustrated by read the video frames, then the desired frame can be identified by it sequence number, after identifying the frame number the extraction function can be started.

DISCUSSION AND ADVANTAGES

In this study, the review has shown above presents the steganography approaches and how some researchers tried to enhance the limitation of steganography. Before 10 years the capacity of the secure image was limited (Moskowitz *et al*., 2000) while now some researchers provide new approaches which can embed secure message or image within more than 50% of the original image size. As we shown in the previous work on steganography there are many use for the steganography even not for pure security where sometimes it is usable for error correction like in (Robie and Mersereau, 2002). For the capacity purpose there is a limitation on how much information can be hidden into an image, making difficult to use the image methods (Chang *et al*., 2002). In the video steganography we have a flexibility of make a selective frame steganography to higher the security of the system or using the whole video for hiding a huge amount of data (Zaidan and Zaidan, 2009). That reason makes the use of video-based steganography more eligible. As well as, the confidentiality need to be tested through testing the quality of steganographic object, as we will present in the second part of this chapter the quality of steganographic object metrics are not sufficient for these kind of test.

There are some techniques facing the capacity limitation problem, in general, the image has limited capacity when use it as a cover medium and it's easier to being tested by the attacker it than the video where the video consist of a set of images and can use more than one image within the video as a carrier cover and more difficult to test it because the sequence of the stego image within the video is unknown by the attacker.

The video based steganography can be used as one video file, separated images in frames or images and audio files. Since that, the use of the video based steganography can be more eligible than other multimedia files.

CONCLUSION

In this study, we have clarified the knowledge of data hidden field. Furthermore, we have presented the history of the Steganography since ancient times until the present day. One of the challenges in this article reviewing the most important methods in the video file that used in this field. Steganography embedding types in video has been illustrated in this study. In addition, we have proposed the video architecture and how we can make use of the internal structure of the video to hide secure data. Finally, we demonstrate the advantage points of the hiding in video. Further direction can be done by using the audio part from the video file to hide a key for encryption and decryption purpose and the video frames as still image to hide the encrypted key.

ACKNOWLEDGMENTS

This research has been funded by the University of Malaya, under the grant number (P0033/2010A). The author would like to take this opportunity to thank and acknowledge his supervisors: Dr. Hamid Jalab and Dr. Zarinah Mohd Kasirun, for having rendered their ceaseless and unconditional support throughout the entire duration of the study. The author would also like to extend his heartfelt gratitude to all his friends and associates who had offered him the much needed assistance and encouragement from the start to the end of the research period.

REFERENCES

Ahmed, M.A., M.L.M. Kiah, B.B. Zaidan and A.A. Zaidan, 2010. A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm. J. Applied Sci., 10: 59-64.

Al-Azawi, A.F. and M.A. Fadhil, 2010. Arabic text steganography using kashida extensions with huffman code. J. Applied Sci., 10: 436-439.

Bhaumik, A.K., M. Choi, R.J. Robles and M.O. Balitanas, 2009. Data hiding in video. Int. J. Database Theory Appl., 2: 9-16.

Chae, J.J. and B.S. Manjunath, 1999. Data hiding in video. Proceedings of the 6th IEEE International Conference on Image Processing, Oct. 1999, IEEE, pp: 243-246.

Chang, C.C., T.S. Chen and L.Z. Chung, 2002. A steganographic method based upon JPEG and quantization table modification. Inform. Sci., 141: 123-138.

De Oliveira, J., 1997. A Java H. 263 decoder implementation. Electrical and Computer Engineering Department, University of Ottawa. http://www.lncc.br/~jauvane/papers/H263JavaDecoder.pdf.

Doerr, G. and J. Dugelay, 2004. Security pitfalls of frame by frame approaches to video watermarking. IEEE Trans. Signal Proc., 52: 2955-2964.

Eltahir, M.E., M.L.M. Kiah, B.B. Zaidan and A.A. Zaidan, 2009. High rate video streaming steganography. Proceedings of the 2009 International Conference on Future Computer and Communication, April 03-05, IEEE Computer Society, Kuala Lumpur, Malaysia, pp: 550-553.

Jalab, H., A. Zaidan and B.B. Zaidan, 2009. Frame selected approach for hiding data within MPEG video using bit plane complexity segmentation. J. Comput., 1: 108-113.

Johnson, N.F. and S. Jajodia, 1998. Steganalysis: The investigation of hidden information. Proceedings of IEEE Information Technology Conference, Sept. 1-3, New York, USA., pp: 113-116.

Kawaguchi, E. and R.O. Eason, 1998. Principle and applications of BPCS-steganography. Proc. SPIE, 3528: 464-473.

Kharrazi, M., H.T. Sencar and N. Menon, 2004. Image steganography: Concepts and practice. Lecture Notes Comput. Sci., 39: 204-211.

Majeed, A., M.L.M. Kiah, H.T. Madhloom, B.B. Zaidan and A.A. Zaidan, 2009. Novel approach for high secure and high rate data hidden in the image using image texture analysis. Int. J. Eng. Technol., 1: 63-69.

Moskowitz, I.S., G.E. Longdon and L. Chang, 2000. A new paradigm hidden in steganography. Proceedings of the 2000 Workshop on New Security Paradigms, Sept. 12-22, ACM, Ballycotton, County Cork, Ireland, New York, pp: 41-50.

Naji, A.W., A.A. Zaidan and B.B. Zaidan, 2009. Challenges of hidden data in the unused area two within executable files. J. Comput. Sci., 5: 890-897.

Noda, H., T. Furuta, M. Niimi and E. Kawaguchi, 2004. Application of BPCS steganography to wavelet compressed video. Proc. Int. Conf. Image Process., 4: 2147-2150.

Robie, D.L. and R.M. Mersereau, 2002. Video error correction using steganography. EURASIP J. Applied Signal Process., 2002: 164-173.

Shirali-Shahriza, M., 2006. A new method for real-time steganography. Int. Conf. Signal Proc., 4: 16-20.

Shirali-Shahreza, M. and S. Shirali-Shahreza, 2008. High capacity persian/arabic text steganography. J. Applied Sci., 8: 4173-4179.

Socek, D., H. Kalva, S.S. Magliveras, O. Marques, D. Culibrk and B. Furht, 2007. New approaches to encryption and steganography for digital videos. Multimed. Syst., 13: 191-204.

Su, Y., C. Zhang, L. Wang and C. Zhang, 2008. A new video steganalysis based on mode detection. Proceedings of the International Conference on Audio, Language and Image Processing, July 7-9, Shanghai, pp: 1507-1510.

Umbaugh, S.E., 1997. Computer Vision and Image Processing: A Practical Approach Using Cviptools. Prentice Hall PTR Upper Saddle River, NJ, USA.

Westfeld, A. and G. Wolf, 1998. Steganography in a Video Conferencing System. Springer, New York.

Wu, M., H. Yu and B. Liu, 2003. Data hiding in image and video: Part II-designs and applications. IEEE Trans. Image Process., 12: 696-705.

Zaidan, B.B., A.A. Zaidan and F. Othman, 2008. Enhancement of the amount of hidden data and the quality of image. Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia.

Zaidan, A. and B. Zaidan, 2009. Novel approach for high secure data hidden in MPEG video using public key infrastructure. Int. J. Comput. Network Security, 1: 1985-1993.

Zaidan, A., B. Zaidan and F. Othman, 2009. New technique of hidden data in pe-file with in unused area one. Int. J. Comput. Electrical Eng., 1: 1793-8198.