http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 10 (10): 1983-1988, 2011
ISSN 1812-5638 / DOI: 10.3923/itj.2011.1983.1988

A Huffman Coding Section-based Steganography for AAC Audio

Jie Zhu, Rang-Ding Wang, Juan Li and Di-Qun Yan
CKC Software Laboratory, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract: Steganography techniques can be used to embed secret information into audio signals. A Huffman coding section-based steganographic scheme for MPEG-2/4 Advanced Audio Coding (AAC) audio is proposed in this study. Based on the characteristics of Huffman coding section, the scheme hides secret information by modifying the sections of Huffman coding without affecting its normal process of encoding. Experimental results are given to show that the proposed scheme not only has a larger capacity, 15 bits per frame averagely but also does not change the statistical properties of the audio carrier which means having certain undetectability. In addition, the scheme does not affect the quality of audio carrier, having good imperceptibility.

Key words: Advanced audio coding, steganography, capacity, undetectablity, audio quality

INTRODUCTION

Steganography (Rabah, 2004) is the art and science of hiding secret information into some innocuous cover-objects (digital images, audios and videos for instance). In such a way, no one apart from the sender and receiver, suspects the existence of the information. Steganography for audio is divided into two categories, compressed domain and non- compressed domain. Steganography for non-compressed audio is mature, such as the Least Significant Bit (LSB) algorithm (Lemma *et al*., 2002), echo hiding algorithm (Li *et al*., 2003), spread spectrum algorithm (Kirovski and Malvar, 2001), Quantization Index Modulation (QIM) algorithm (Li *et al*., 2011) and so on. However, with the development of multimedia technology and internet, MPEG-1 Layer III (MP3), MPEG-2/4 Advanced Audio Coding (AAC) and other compression technologies widely used. Therefore, people also pay more and more attention on steganography for compressed audio.

At present, most of the steganographic algorithms for compressed audio are based on the MP3. Wang *et al*. (2004) uses the Modified Discrete Cosine Transform (MDCT) coefficients at low frequency to hide secret information, while, Yan *et al*. (2009) adjusts the parity of quantization step to hide secret data. Otherwise, Quan and Zhang (2006) achieve steganograhpy with the modulation of quantization step based on wet paper coding strategy. The most famous steganographic scheme is modifying terminating condition of the inner loop in MP3 encoding which is called MP3Stego (Petitcolas, 2002). Nevertheless, as a new generation of audio compression technology, AAC has higher compressed ratio than MP3 and may become the most popular compression coding technology instead of MP3. So the study of steganographic algorithm for AAC is kind of forward-looking.

Compared with MP3, the steganographic algorithms for AAC are much fewer. Tachibana (2004) and Neubauer and Herre (2000) both use MDCT coefficients to hide secret information without affecting the subjective effects of the AAC audio. As MDCT coefficients own characteristic, the hiding capacity is quite limited. Based on the characteristic of quantization, Xu and Zhang (2009) hides secret data into the redundant bits obtained by modifying the quantization factor. But the scheme hides secret information in the process of quantification, so the AAC audio must be decoded deeply and its complexity is unacceptable. Wang *et al*. (2009) gets secret messages hidden on the base of the mechanism called escape coding which is used by the encoder to achieve lossless coding when MDCT quantified coefficient in AAC is greater than 15. However, the hiding capacity does not keep stable while the style of audio carrier is changed.

Based on the characteristic of Huffman coding, a new steganographic scheme for AAC audio is proposed in this study. We first extract the Huffman coding sections and then modify them according to the secret information, so as to achieve the purpose of steganography. The proposed scheme has high capacity, good imperceptibility and certain undetectability.

Corresponding Author: Jie Zhu, CKC Software Laboratory, Ningbo University, Ningbo, Zhejiang 315211, China

THE PROPERTY OF HUFFMAN CODING SECTION

The basic unit of Huffman coding in AAC is called section. The encoder divides the set of 1024 quantized spectral coefficients into several sections, each of which uses a single Huffman codebook to code. Section boundaries can only be at scale factor band boundaries for reasons of coding efficiency, so that each section contains at least one scale factor band. The scale factor band is fixed while section is dynamic and typically varies from block to block. In this way, the number of bits needed to represent the full set of quantized spectral coefficients is minimized.

Finding the proper sections is a continuous process of trying. Actually, a greedy merge algorithm is used to get it done. It starts with the maximum possible number of sections each of which used the Huffman codebook with the smallest possible index. Then, if the resulting merged section results in a lower total bit count, sections are merged, with merges that yield the greatest bit count reduction done first. When the sections to be merged do not use the same Huffman codebook, the codebook with the higher index must be used. Figure 1 shows a simple example of the merging process. S0 and S1 in the figure present two sections, assuming the used Huffman codebook of them are Codebook (S0) and Codebook (S1), respectively. First, the encoder determines whether the merging of S0 and S1 can lead to use fewer bits for Huffman coding. If so, the two sections are merged and coded using the Huffman codebook of MAX (Codebook(S0), Codebook(S0)). Otherwise, each section still uses the original Huffman codebook to code.

After the merging of sections, a number of scale factor bands may use a single codebook for the Huffman coding. If this probability is quite close to 50%, it is possible to hide secret information. We select six different styles of AAC audios, whose number is 1,200 totally and calculate the probability of the scale factor bands being coded by a single Huffman codebook. For all of experimental audios, the probability can be calculated by the total number of groups which may be combined by two or three scale factor bands and the number of groups whose scale factor bands use a single Huffman codebook, dividing the latter by the former. The results are shown in Table 1, where first row and second row are the probability when two and three scale factor bands are combined as a group, respectively. For example, the value in first row and first column, 52.54, means the probability of two scale factor bands using a single codebook for audios of blues style is nearly 52.54%. It can be seen from the table that whether combining two or three scale factor bands as a group, the statistical results cannot meet the requirement of 50%. Therefore, the statistical properties of AAC audio will be affected and the undetectability of steganography cannot be guaranteed, if the scale factor bands are forced to change the Huffman codebook for coding. However, if the steganography is achieved by two and three scale factor bands as a group alternately, the probability of scale factor bands using a single Huffman codebook may not have a great change after steganography. As a result, the undetectability of steganography could be ensured.

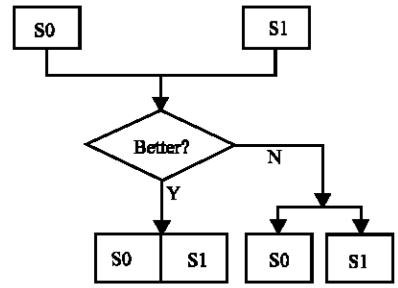

Fig. 1: Section merging

Table 1: The probability of scale factor bands using a single codebook (%)

No.	Blues	Classical	Country	Folk	Jazz	Pop
1	52.54	53.57	54.67	52.63	56.25	53.87
2	37.97	35.90	36.17	34.11	38.58	33.75

PROPOSED SCHEME

Based on the characteristics of Huffman coding section, this study proposed a novel steganography which hides secret messages without affecting the normal encoding process. The scheme is described as follows.

Steganography: Assuming that the secret information is a two-dimensional image, the steganography steps are as follows:

Step 1: Preprocess the two-dimensional image and convert it into one-dimensional binary sequence $\omega = \{\omega_1, \omega_2, \omega_3,..., \omega_i,...\}$, $\omega_i \in \{0, 1\}$

Step 2: Generate one-dimensional binary sequence $k = \{k_1, k_2, k_3,..., k_j,...\}$, $k_j \in \{0, 1\}$ randomly and send it to the extraction client as a key

Step 3: Decode the j frame of original AAC audio to the step of Huffman coding, if $k_j = 0$, then num = 2, otherwise, num = 3. Take num scale factor bands as a group and extract the Huffman codebook used by each scale factor band in the group. If all of them use a single Huffman codebook, set the

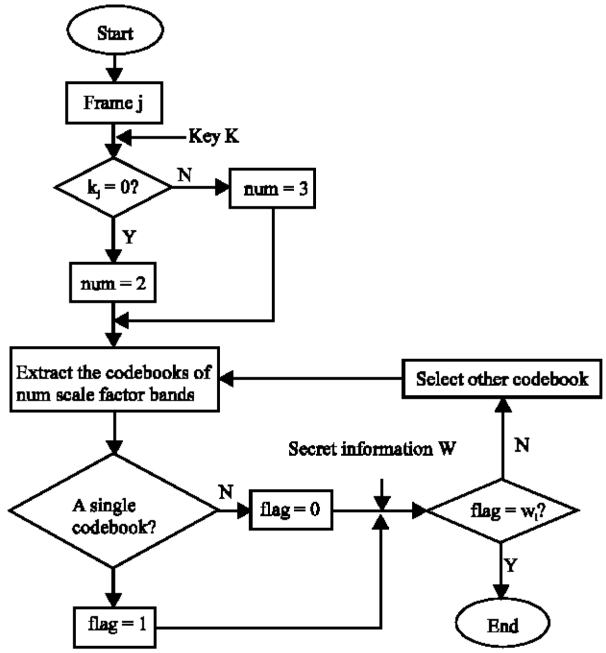

Fig. 2: The process of hiding one bit secret information

flag variable flag = 1, or else, flag = 0. Assuming that the hidden secret information is w_i, if w_i = flag, then the steganography of w_i is finished. Otherwise, select other Huffman codebooks for scale factor bands to meet w_i = flag. As the scale factor bands in high frequency band usually use No. 0 Huffman codebook to code, the proposed scheme hides secret messages in all scale factor bands except the exception ones in high frequency band, where:

$$\text{exception} = \begin{cases} 2 & \text{short\_block} \\ 7 & \text{else} \end{cases}$$

Step 4: Code this frame and format one frame of AAC bitstream

Step 5: Continue decoding the next frame, repeat Step 3 and Step 4 until all the secret messages are hidden completely

The process of hiding one bit secret information is shown in Fig. 2.

Extraction: The extraction process is:

Step 1: Get the key $k = \{k_1, k_2, k_3,..., k_j,...\}$, $k_j \in \{0, 1\}$ sent from the steganography client

Step 2: Extract the side information of AAC audio's j frame, if $k_j = 0$, then num = 2, or else, num = 3. Take num scale factor bands as a group and extract the Huffman codebook used by each scale factor band in the group. If all of them use a single Huffman codebook, the hidden information $w_j = 1$, otherwise, $w_j = 0$. The proposed scheme hides secret messages in all scale factor bands except the exception ones in high frequency band, so the exception ones in high frequency band are not extracted

Step 3: Go back to Step 2 to continue decoding the side information of next frame until all of the secret information is attained

Step 4: Anti-preprocess the extracted secret information $\omega = \{\omega_1, \omega_2, \omega_3,..., \omega_i,...\}$, $\omega_i \in \{0, 1\}$ and attain the final secret information

The process of extracting one bit secret information is shown in Fig. 3.

EXPERIMENTAL RESULTS AND ANALYSIS

The proposed scheme hides secret data based on the Huffman coding section. The whole steganographic process occurs in the part of Huffman coding which belongs to the lossless coding. Therefore, the steganography doesn't have any impact on the quality of AAC audio. Because of this, here we only analyse the capacity as well as the security of proposed scheme. In this study, we select six different styles of AAC mono

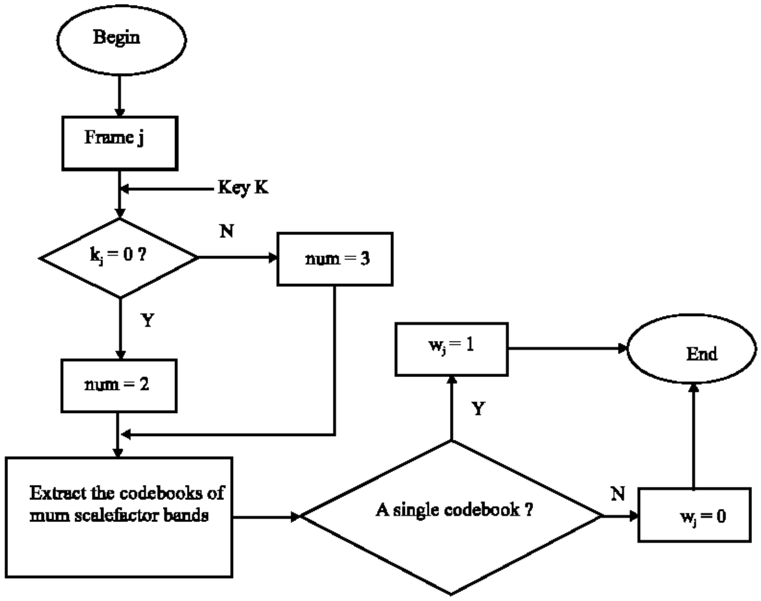

Fig. 3: The process of extracting one bit secret information

Table 2: Capacities of different styles of audios (bits/frame)

AAC audio	Blues	Classical	Country	Folk	Jazz	Pop
Capacity	15.42	16.08	14.97	15.54	15.95	13.97

audios as the covers, all of which are about 10 sec long and sampled at 44100 Hz with 16 bits resolution.

Capacity: The proposed scheme hides secret messages by modifying the Huffman coding sections, so the capacity of the steganography depends on the number of scale factor bands, sfb which is not excluded in one AAC frame and the number of scale factor bands in each group, num. The hiding capacity is sfb/num bits per frame.

Table 2 shows the capacities of different styles of AAC audios. Known from the data in the Table, different styles of AAC audios have a similar capacity which is from 14 to 16 bits per frame. No matter what style of AAC audio, it almost coded by the long block which affects the capacity of steganography directly. Therefore, the capacity of proposed scheme keeps stable when the style of AAC audio is changing. In the information hiding method for AAC audio based on the spread spectrum modulation proposed by Cheng *et al.* (2002), the embedded bits rate is commonly 30 bits per sec and 0.7 bits per frame while converted into the average embedded bits per frame which is much less than the capacity of proposed scheme in this study. In a word, the capacity of proposed scheme is not only high but also stable for different styles of AAC audios.

Experimental results evaluate that the value of SNR of the algorithm is larger, it means the imperceptibility of the audio aggregation which is got by the method in this study is better.

Security: The assessment of steganographic scheme's security mainly uses steganalysis technique, one important method of which is the detection based on statistical characteristics. Attackers can determine whether the detected object is the one containing secret information according to the differences of statistical characteristics, since hiding the secret information would change some statistical characteristics of the cover object, such as distribution of MP3 block length (Westfeld, 2002), etc.

The proposed scheme modifies the Huffman coding section to hide secret messages, so we must pay attention to them, whose probability may be changed after the secret information hidden. The probabilities of scale factor bands using a single Huffman codebook before and after steganography are shown in Fig. 4, where Fig. 4a and b are the results of 2 and 3 scale factor bands as a group, respectively. It can be seen from the figure that, the probability of scale factor bands using a single Huffman codebook is hardly changed after steganography which means the proposed scheme has good security to resist some attack based on the statistical characteristics.

In addition, after modifying the Huffman coding sections, the proposed scheme may code the MDCT quantized coefficients using another Huffman codebook which makes the size of original AAC cover changed directly. Table 3 shows the changes of AAC audios' size before and after steganography. From the Table, we can see that the size of AAC audios increases slightly compared with the ones before steganography, generally between 1 and 3%. However, due to the attacker cannot attain the original AAC audio files, a slightly change of

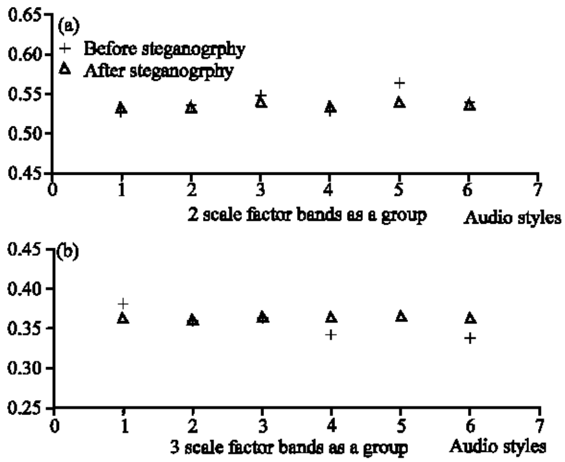

Fig. 4(a-b): The probability of scale factor bands using a single Huffman codebook before and after steganography

Table 3: The changes of AAC audios' size before and after steganography (%)

AAC audio	Blues	Classical	Country	Folk	Jazz	Pop
Changes	1.23	2.23	2.00	2.41	2.49	1.64

file size do not have much effect on the security of steganography.

CONCLUSION

In this study, we present a novel AAC audio steganography based on the Huffman coding section. The scheme is simple, real time and practical because the whole hiding process is completed within the process of Huffman coding. In addition, the steganography only modifies the codebook used by Huffman coding which is lossless, so the quality of AAC audio is not affected at all, having high undetectability. The experimental results evaluate that the proposed scheme has certain stability, high capacity and good performance for statistical properties of AAC audio, resisting to the corresponding steganalysis attack. However, the shortage of proposed scheme is the AAC file will be a little larger after hiding the secret messages. Although, the security of steganography may not be affected under the condition of blind steganalysis, we will carry out related work in order to reduce the increase of file size and improve the performance of steganography in the future work.

ACKNOWLEDGMENTS

This study is supported by the National Natural Science Foundation of China (60873220), Doctoral Fund of Ministry of Education of China (20103305110002), Zhejiang Natural Science Foundation of China (Y108022, Z1090622, Y1090285), Zhejiang Science and Technology Preferred Projects of China (2010C11025), Zhejiang Province Education Department Key Project of China (ZD2009012), Ningbo Science and Technology Preferred Projects of China (2009B10003), Ningbo Key Service Professional Education Project of China (2010A610115), Ningbo Natural Science Foundation (2009A610085) and Ningbo University Foundation (XYL10002, XK1087). In addition, our programs are supported by High School Special Fields Construction of Computer Science and Technology (TS10860), Zhejiang Province Excellent Course Project (2007), Zhejiang Province Key Teaching Material Construction Project (ZJB2009074).

REFERENCES

Cheng, S., H. Yu, Heather and Z. Xiong, 2002. Enhanced spread spectrum watermarking of MPEG-2 AAC audio. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May 13-17, Orlando, FL. pp: IV-3728-IV-3731.

Kirovski, D. and H. Malvar, 2001. Robust spread-spectrum audio watermarking. Proc. IEEE Int. Conf. Acoustics, Speech Signal Process., 3: 1345-1348.

Lemma, A.N., J. Aprea and W. Oomen, 2002. A robustness and audibility analysis of a temporal envelope modulating audio watermark. Proceedings of the 10th IEEE Digital Signal Processing Workshop and the 2nd Signal Processing Education Workshop, Oct. 13-16, Bordeaux, France, pp: 62-67- 10.1109/DSPWS.2002.1231077.

Li, W., X. Xue and P. Lu, 2003. A novel feature-based robust audio watermarking for copyright protection. Proceedings of the International Conference on Information Technology: Computers and Communications, April 28-30, IEEE Computer Society, Washington, DC., USA., pp: 554-558.

Li, J., R.D. Wang and J. Zhu, 2011. A watermark for authenticating the integrity of audio aggregation based on vector sharing scheme. Inform. Technol. J., 10: 1001-1008.

Neubauer, C. and J. Herre, 2000. Audio watermarking of MPEG-2 AAC bit streams. Proceedings of the AES 108th Convention, Feb. 19-22, Paris, pp: 1-19.

Petitcolas, F., 2002. mp3stego. http:// www.petitcolas.net/ fabien/steganography/mp3stego/index.html

Quan, X. and H. Zhang, 2006. Data hiding in MPEG compressed audio using wet paper codes. Proceedings of the 18th International Conference on Pattern Recognition, Aug. 20-24, HongKong, pp: 727-730.

Rabah, K., 2004. Steganography: The art of hiding data. Inform. Technol. J., 3: 245-269.

Tachibana, R., 2004. Two-dimensional audio watermark for MPEG AAC audio. Proc. SPIE., 5306: 139-150.

Wang, C.T., T.S. Chen and W.H. Chao, 2004. A new audio watermarking based on modified discrete cosine transform of MPEG/audio layer III. Proc. IEEE Int. Conf. Network. Sens. Control, 2: 984-989.

Wang, J., Y. Yang and R. Xiao, 2009. Audio watermarking based on MPEG-2 AAC. J. Univ. Sci. Technol., 31: 525-529.

Westfeld, A., 2002. Detecting low embedding rates. Proceedings of the Information Hiding Workshop, Oct. 07-09, Springer-Verlag, London, UK., pp. 324-339.

Xu, S.Z. and P. Zhang, 2009. Performance analysis of data hiding in MPEG-4 AAC audio. Tsinghua Sci. Technol., 14: 55-61.

Yan, D., R. Wang and L. Zhang, 2009. Quantization step parity-based steganography for MP3 audio. Fundam. Inform., 97: 1-14.