AUDIO STEGANOGRAPHY BASED ON SIGNAL MODULATION IN WAVELET DOMAIN

Loay E. George, * Ghassan A. Mahmood
Unit of Information Technology, College of Science, University of Baghdad. Bagdad-Iraq.
* Department of Computer, College of Science, University of Al-Nahrain. Bagdad-Iraq.

Abstract

In this paper, an audio steganography system is proposed to hide a sequence of binary digits (bits) in digital audio data. The hiding method is based on applying amplitude modulation technique on wavelet transform coefficients. The embedding schema implies partitioning the audio signal (cover) into a number of non overlap slices, then the average of AC energy of all slices are examined; and the slices have the highest energy are considered as good hosts for secret bit. The shift in coefficients values, caused by embedding process, may cause modification in the energy status of the host slice; this problem is avoided by applying an algebraic manipulation so as to ensure that the extraction detector can to correctly identify the audio host slices. Bits insertion is done by transforming the selected host slices into wavelet domain using the biorthogonal tap (9\7) wavelet filters. Then, some of the highly valued coefficients of the high frequency subband are quantized to use them to host the secret bits {0, 1} by adding or subtracting a modulation step. The test results indicated that a perfect retrieval of secret bits could attained, while the hiding rate is small and the quality of stego objects is high.

الاخفاء في الصوت باستخدام تضمين الاشارة في المجال المويجي

لؤي ادور جورج، *غسان عبد الحكيم محمود
وحدة تكنولوجيا المعلومات، كلية العلوم، جامعة بغداد. بغداد-العراق.
*قسم الحاسبات، كلية العلوم، جامعة النهرين. بغداد-العراق.

الخلاصة

في هذا البحث، تم اقتراح منظومة لاخفاء سلسلة من البتات في بيانات الصوت الرقمية. تستند طريقة الاخفاء على تطبيق تقنية التضمين السعوي على معاملات التحويل المويجي. يتضمن هيكل طريقة الاخفاء عملية تجزأة اشارة الغطاء الصوتية الى عدد من الشرائح الغير متداخلة، بعد ذلك يتم حساب معدل الطاقةللمركبة المتناوبة لجميع الشرائح يجري حسابه: ليتم بعدها اختيار الشرائح التي تمتلك الطاقة المتناوبة العليا كوسط لاخفاء البتات السرية. ان الزحف الذي سيحصل بسبب عملية الاخفاء ممكن ان يتسبب في تغيير حالة الطاقة للشريحة المضيفة، ان هذه المشكلة تم تجنبها من خلال تطبيق معالجة جبرية لضمان نجاح كاشف الاستعادة من تحديد الشرائح المضيفة. تجري عملية اخفاء البتات بعد تحويل بيانات الشريحة المضيفة الى المدى المويجي (باستخدام مرشحات التحويل المويجي الثنائي التعامد Tap 9/7). بعد ذلك، يتم انتقاء بعض معاملات التحويل المويجي ذات القيمة العالية والتي تقع ضمن حزمة التردد العالي لتجري عليها عملية التكميم ومن ثم استخدامها لاستضافة قيم البتات{0, 1} وذلك باجراء عملية جمع او طرح لقيمة خطوة التضمين. اشرت نتائج الاختبار بامكانية تحقيق استعادة تامة للبتات المخفية وان معدل الاخفاء صغير في الوقت الذي فيه ان جودة بيانات صوت الغطاء تبقى عالية.

1. Introduction

With the development of information technology, people have paid more and more attention on the information security. Information hiding technology can embed secret information into a digital media without impairing the perceptual quality of that source, such that other people can't feel this secret information.

Audio file can be used to hide information. Steganography is often used to copyright audio file to protect the rights of music artists. Techniques like least significant bit insertion, phase coding, spread spectrum coding, and echo hiding can be used to protect the content of audio file. The biggest challenge face all these methods is the sensitivity of human auditory system (HAS), it is so sensitive, such that people can often pick up randomly added small noise, and this making hard to success-fully hide data within audio data [1-3]. During the last decade, the use of wavelet transform for hiding covert data in digital audio and video was rapidly grown due to robust characteristics of wavelet. Various types of wavelet families (like, Haar, Daubetchies, …) and insertion mechanisms in wavelet domain (like, LSB, addition, selective,…) were introduced and investigated [4-6]. The hiding performance of these introduced hiding methods is different; mainly it depends on which aspect (i.e., hiding rate, impercibility and robustness) is going to be preserved.

2. The Proposed System

It consists of two modules, the embedding and extraction. These two modules are explained below.

2.1. Embedding Phase

The embedding phase includes many stages. Figure 1. shows the involved processes in this module, they are:

A. Load Wave File

First, the wave file content is loaded, it consists of header and data section. Header contains information about the audio file attributes (like, sample rate, no. of channels, bits per channel, … etc), while the data section holds the values of audio samples within the wave [7].

B. Determination of Mean

The mean value of all loaded samples is calculated using the following equation [8]:

$$m = \frac{1}{k}\sum_{i=1}^{n} s(i), \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where, n is total number of samples s(i) is the i^{th} sample value.

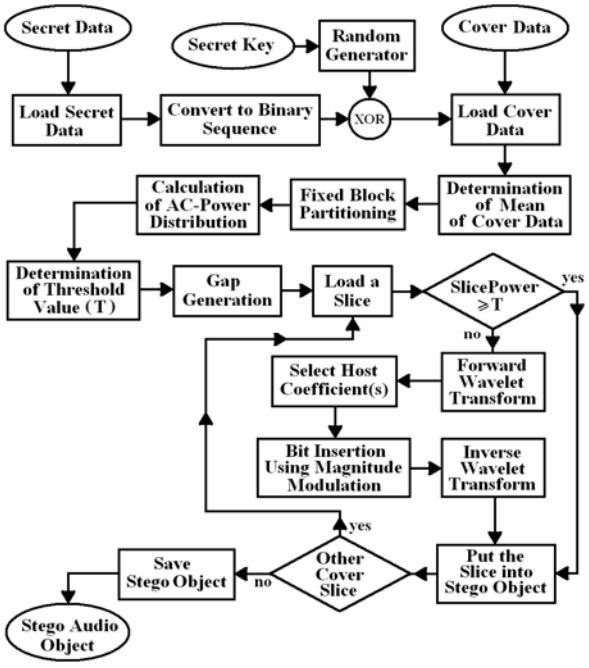


Figure 1: Block diagram of the embedding module

C. Audio Signal Partitioning

The audio signal data is partitioned into non overlapped slices of equal length. The slice length is predefined by the authenticator. Then the number of slices of the audio array is determined. Small slices are preferred so as to increase the complexity of secret bits extraction for the attacker. It is found that slices with length less than 20 samples are less, subjectively, distorted by quantization and bit insertion operation.

D. AC-Energy Calculation

The adopted embedding operation is based on the fact that the HAS can hardly recognize the variations (up to some extent) in highly valued coefficients of the high frequency subband; while it easily recognize the small variations might occurred in small valued coefficient. This fact is utilized to decide whether the candidate audio block could be used as a host for embedding secret bits or not. The following formula is used to determine the AC-energy [9, 10]:

$$P(j)=\sum_{i=0}^{k-1}(s(i-kj)-m(j))^2,\ldots\ldots\ldots\ldots(2)$$

Where, p(j) is energy magnitude for the j^{th} slice, s(i) is the magnitude of the $(i\text{-}kj)^{th}$ sample within jth slice, M(j) is the mean of j^{th} slice.

E. Histogram Calculation

The histogram of the energy of audio slices is determined (i.e., the number of slices have distinct energy). The hiding process is run out only on the slices belong to right most part of the histogram, which means that the most energetic slices are chosen as a hosts for secret bits (see Figure 2).

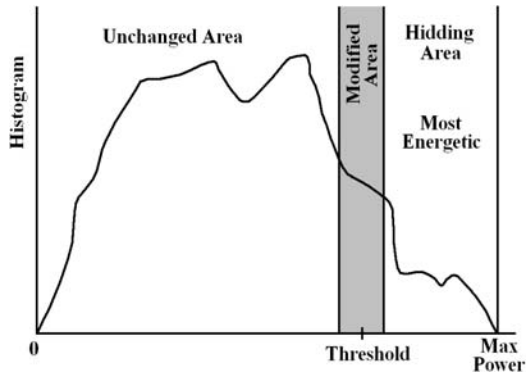


Figure 2: Using the right most part of histogram

The number of slices used to host sec-ret bits depends on the total number of secret bits and on the number of bits would be hosted in each slice. In this work the applied system uses k slices to hide k bits (i.e., each slice hosts one bit). In general, the number of required host slices is determined using:

$$n_b=\frac{n_s}{r_h},\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Where, n_s is the number of secret bits,

r_h is the number of bits hidden in each slice,

n_b is the number of audio cover blocks used as host.

The value of nb should be less than or equal to the actual total number of slices. In case of nb less than the actual number of slices, then the most energetic slices should used as hosts; in such case, the minimum energy threshold value is needed to be determined. This threshold value (T) could easily determine by applying the following condition on the energy histogram array:

$$\sum_{i=T}^{\max}H(i)\geq n_s,\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Where, *H(i)* is the number of slices have energy (i), *Max* is the maximum registered energy of blocks.

After the determination of threshold (*T*), then each slice have energy greater than or equal to (*T*) is selected as host slice for hiding secret bit*s*. This value is very important for the detector to recognize the host slices [11, 12].

F. Gap Generation

Till this step, no modification happened on the samples values of host slices, but in this and next steps, a real change is occur on host slices, which is due to the bit insertion (it implies the stages: wavelet, quantization, and amplitude modulation). Due to the changes occurred on samples of host slices, the host slices whose energy close to threshold values may changed such that its new energy lowered and become less than (*T*), and in such case they could not be recognized as host slice during the extraction phase. To avoid the occurrence of this case, the samples values of each of these host slices should slightly increased to set its energy far, to some extent, from threshold value, as shown in Figure 3. Without applying this step, a serious error will occur during the extraction step, leading to process failure.

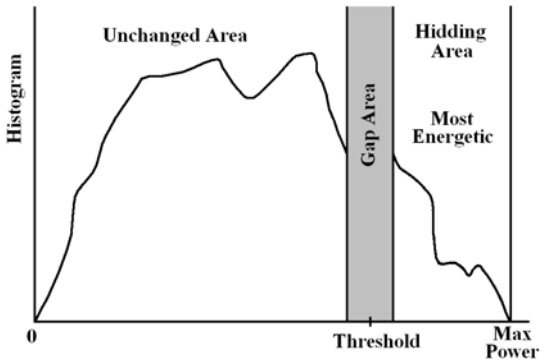


Figure 3: Shape of histogram after applying gap generation step.

In this work the following expanding equation was applied to the slices whose energy are close to (*T*):

$$s'(i)=\alpha\big(s(i)-m\big)+m,\ldots\ldots\ldots\ldots\ldots(5)$$

Where, *s'(i)* is the new value of i^{th} sample in the slice, *s(i)* is the original value of i^{th} sample in the slice, *m* is the mean value for the slice, α is the shift ratio factor.

This operation guarantees that no slice will change its identity from host to non-host slice.

The effectiveness of this operation depends on α value. The value of α is set greater than (1) for shifting the host slices, and it is set less than (1) when shifting the non-host slices (if there is a need to do that). Although the gap generation operation causes some changes in values of slices samples; but it leads to a great reduction in error ratio of the retrieved secret bits.

G. Embedding Stage

The embedding stage implies many steps, as stated below:

G1. Wavelet Transform

The proposed system uses the wavelet dom-ain for bit hiding. The selected host slices are transformed by applying the wavelet transform. The use of wavelet domain, instead of spatial domain, adds more robustness to the hiding process. Tap (9\7) biorthogonal filter is selected for wavelet transform. Only one wavelet pass is applied; which leads to two subbands (i.e., low and high) [13, 14]. High frequency coefficients are treated as the host for the secret bit, while the low coefficients are kept unchanged.

G2. Bits Embedding

In this stage, two issues are taken into consideration, they are: the way of host coefficients selection, and the way of bit insertion. In this work, from each slice one of the coefficients of the high frequency subband is selected for hiding a secret bit, the selection of this coefficient is done using a random generator (whose output is used to assign the number of the host high subband number). The use of random generator as a coefficient selector adds extra security level to the hiding system.

The bit insertion is done by shifting the quantized value of the selected coefficients according to the following formula:

$$C_h(i) = \left\lfloor \frac{C(i)}{\beta} \right\rfloor \beta + \frac{\beta}{\rho}(2s-1), \ldots\ldots(6)$$

Where, $C_h(m)$ is the coefficient value after embedding, $C(m)$ is the original value of the selected host coefficient, β is the gap step (its value > 1), ρ is the modulation factor (its value > 2), and s is the secret bit value{0,1}.

The first term implies the quantization step. β is the quantization step, it specifies the bins size, which represent the differences between adjacent quantized values.

G3. Inverse Transform

After bits embedding, the host slices should transformed back to the spatial domain using inverse wavelet transform (Tap 9/7). Then these slices should be returned to their locations in audio signal to get a stego cover audio object.

2.2. Extraction Phase

Most of the steps of the embedding module are repeated in the extraction phase (see Figure 4) except the bit embedding stage, which is replaced by the bit extraction step. The decision of whether the extracted secret bit is 0 or 1 is taken according to following criterion:

$$S(i) = \begin{cases} 1 & \text{if } C_h > round(Ch/\beta)\beta \\ 0 & \text{if } C_h < round(Ch/\beta)\beta \end{cases}, \ldots(7)$$

Where, $S(i)$ is the extracted bit array of sequence i, C_h is the wavelet host coefficient.

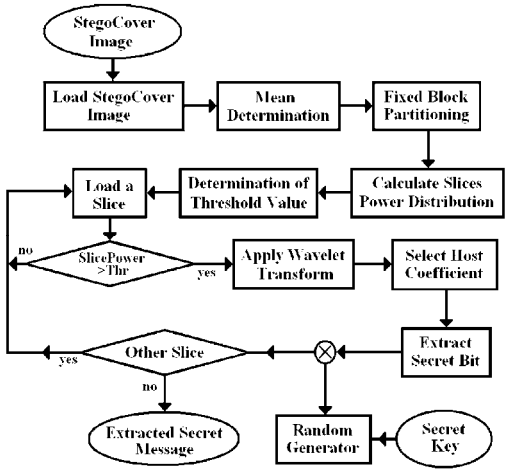


Figure 4: Extraction phase

3. Experimental Results

The performance of the proposed system was using five audio files (each has specifications: PCM, mono, 8 bits/sample, 11024 sample/ sec, the size is 500KB). During tests the follow-ing parameters have been use as performance indicators:

1. Peak Signal to noise ratio (PSNR) of the stego cover audio object, where:

$$PSNR = 10\log\left(\frac{255^2}{\frac{1}{n}\sum_{i=0}^{n-1}(s'_i - s_i)}\right), \ldots\ldots(8)$$

n is the number of cover audio samples.
s_i is the original i^{th} audio sample.
s'_i is the value of i^{th} stego sample.

2. The Rate of Wrong Retrieved Bits (RWRB), which is determined using the following equation:

$$RWRB = \frac{n_w}{n_s} \times \%100, \ldots\ldots\ldots\ldots(9)$$

Where, n_s is then number of secret bits and n_w is the number of wrong retrieved bits.

The effectiveness of the system parameters (ρ, β, α) was tested by assigning various values to each one and, then, the parameters PSNR and RWRB are determined. During the test the amount of secret bits was taken 12000 bits. Due to nature of signal variability the proposed system determines the required threshold (T) needed to keep hiding only in the most energetic slices of the audio cover signal.

Before testing the system performance, some of the first order statistical parameters (i.e. mean, standard deviation and entropy) of the cover audio files were determined. The statistical parameters could be used as indicators to characterize the signal variability in each cover object. Table (1) shows the values of the statistical parameters. The low values of standard deviation ($\sigma = \overline{s^2} - \bar{s}^2$) and entropy ($E = -p_i \times (log\ p_i$, where p_i is the probability of occurrence of i^{th} value of the cover audio samples) for b_wave file indicates that the slices of this object are mostly low energetic. The highest determined standard deviation value is for d_ wave object, which indicates that the slices of this file are more energetic than those belong to other audio objects.

Table (2) illustrates the effectiveness of modulation factor (ρ) on SNR. The values of other system parameters were taken: slice length SL=8, β= 4, α =1.03. During the tests it

Table 1: The values of mean, standard deviation (σ) and entropy (E)

Cover	Mean	Standard Deviation	Entropy
a_wave	127.986	17.156	6.0975
b_wave	128.020	13.221	5.6551
c_wave	127.989	27.986	6.8244
d_wave	127.989	28.800	6.8691
e_wave	127.985	23.250	6.5616

Table 2: The effect of ρ on SNR

Cover	Thresh-Old	SNR			
		ρ=2	ρ =3	ρ =5	ρ =7
a.wav	3211	40.87	41.26	41.47	41.52
b.wav	3776	38.51	38.94	39.17	39.24
c.wav	12151	41.47	41.66	41.75	41.79
d.wav	8316	42.94	43.15	43.27	43.30
e.wav	5777	41.76	42.03	42.15	42.19

was noticed that the applied threshold value depends mainly on the number of secret bits and slice length.

(Figure 5) illustrates the effect of the modulation parameter (ρ) on RWRB, the values of other system parameters were taken (SL=10, β=6, α =1.03). It is obvious that the increase in value of ρ increases the probability of correct retrieval of secret bits. As shown in the figure the proper value of ρ is (>2.9). (Figure 6) illustrates the effect of slice length (SL) on both SNR and RWRB.

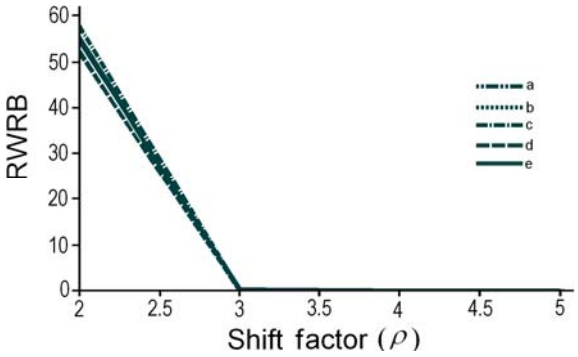
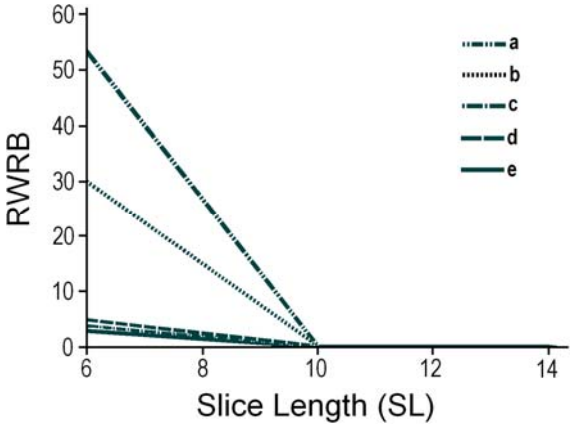
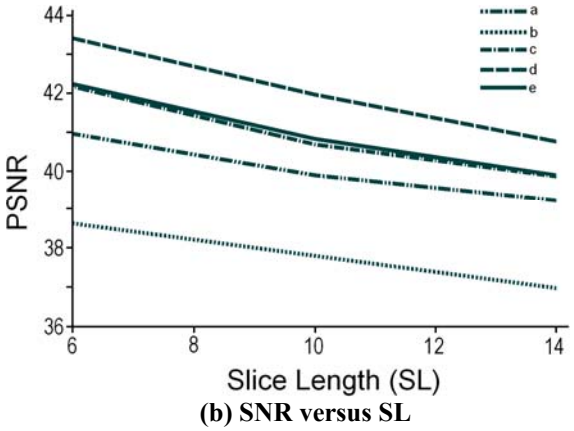


Figure 5: The effect of ρ on RWRB



(a) RWRB versus SL



(b) SNR versus SL

Figure 6: The effect of slice length (SL) on both RWRB and SNR

The values of other system parameters were taken β =6, ρ =5, α =1.03. The results presented in the figure indicate that the increase of SL significantly reduces the probability of

retrieving wrong bits, but at same time the quality of cover signal is slowly reduced.
Tables (3) and (4) present the effects of quantization step factor (β) on SNR and RWRP, respectively. The values of other system parameters were taken: SL=10, ρ =5, α =1.03.

Table 3: The effect of β on RWRB

Cover	RWRB		
	β=3	β =4	β =5
a.wav	11	2	0.00
b.wav	8.58	2.08	0.00
c.wav	9.416	2.75	0.00
d.wav	9.75	1.91	0.00
e.wav	11.75	2.41	0.00

Table 4: The effect of β on SNR

Cover	SNR		
	β=3	β =4	β =5
a.wav	41.21	40.93	40.51
b.wav	39.12	38.74	38.33
c.wav	41.28	41.10	40.96
d.wav	42.66	42.45	42.23
e.wav	41.69	41.47	41.20

Figure 7. illustrates the effectiveness of shift ratio factor (α) on RWRB. The values of the parameters were set: β=5, ρ=5, SL=10. The shown test results indicate that the proper values of α is (>1.02), at this range of values the probability of wrong retrieved bits is significantly reduced.

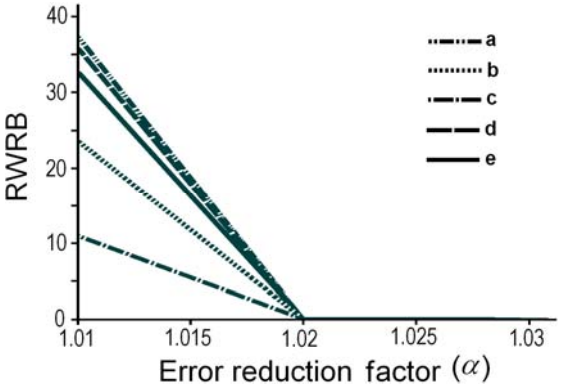


Figure 7: The effect of α on RWRB

4. Conclusions

1. The test results indicate the effectiveness of the system parameters on RWRB and SNR, so the values of these control parameters should be chosen carefully.
2. The threshold value is exclusively affected by slice size and number of secret bit, no effect is noticed when the values of other system parameters (i.e., α, β, ρ) is changed.
3. The best value of modulation factor (ρ) is 4, because it leads to perfect retrieval of secret bits (i.e., RWRB=0).
4. The best value of quantization step (β) is 5, because it leads to low RWRB with acceptable SNR value.
5. The best slice length (SL) is 10 samples; it leads to low MBEP and high SNR.
6. The best value of shift ratio factor (α) is 1.02 because it leads to perfect secret bits retrieval with acceptable value of SNR.
7. The attained hide rate is smaller than that attained by LSB technique; this is due to the exclusion of unvoiced segments and low frequency coefficients from being part of host media. But, the attained hiding rate is nearly similar to those reached by other wavelet-based hiding methods. Here, we have to take into consideration the trade-off between quality and hiding rate, the increase in quality will associated with a decrease in hiding rate.
8. Also, the noise caused by the suggested method is subjectively less than that caused by other hiding methods which use all cover data as host media.

References

1. AL-Kawaz, H. M**. 2006**. Low rate hiding in audio data using phase domain, M.Sc. thesis, Department of Computer Science, College of Science, Universit of AL-Nahrain y, Iraq.
2. Xiuhui, G.; Renpu, J.; Hao, T.; Jiazhen, W. **2006**. Research on information hiding, Research, Hebei University of Economics and Business/ Ordnance Engineering College, *US-Chain Education Review*, ISSN1548-6613, **5**(3):77-81
3. Katezbeisser, P. and Fabian, A. **2000**. Information hiding techniques for steganography and digital watermarking, Artech House Inc., Norwood University.
4. Cvejic, N. and Seppanen, T. **2002**. A Wavelet domain LSB insertion algorithm for high capacity audio steganography, Digital Signal Processing, A Workshop-2002, and the 2nd signal processing education workshop. *Proceedings of 2002 IEEE*, **10th**:53-55.
5. Delforouzi, A. and Poyaan, M. **2008**. Adaptive digital audio steganography based on integer wavelet transform; Circuits, *Systems, Signal Process*, **27(**2):247-259.
6. Santosa, R. A.; Bao, P. **2005**. *Audio to image wavelet transform based on audio*

steganography; ELMAR, 47th International Symposium, P.124.

7. Microsoft Developer Network library, MSDN, Web site: http://msdn.microsoft .com /en-us/library /default.aspx
8. Xue-min, R.; Ting, Z.; and Fei, W. **2006**. Audio steganalysis based on negative resonance phenomenon caused by steganographic tools, *Journal of ZhejJang University Science-A*, ISSN 1009-3095 ,**7**(4):577-583.
9. Nedeljko, C. **2004.** Algorithms for Audio Watermarking and Steganography, Academic Dissertation, University of Oulu, Finland.
10. Su, J.; and Girod, B. **2002**. Power-spectrum condition for energy-efficient watermarking, *IEEE Transactions on Multimedia*, **4(**4):551–560.
11. Kunio, K.; Kurozumi, T.; Murase, H. **2003**. A quick search method for audio and video signals based on histogram pruning, *IEEE Transactions on Multimedia*, **5(**3) September.
12. Smith, S. **2007**. *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, USA.
13. Carrion, P.; De Oliveira, H.; De Souza, R. **2008**. *A low-throughput wavelet-based Steganography audio scheme*, VIII Simposio Brasileiro em seguran ca da informancao e de sistemas computac-ionais, pp.259-262.
14. Parker, K. **2007**. Watermarking with wavelet transforms, M.Sc. thesis, University of Mississippi State, USA.