Expert Systems with Applications 40 (2013) 3055–3064

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

High payload data-hiding in audio signals based on a modified OFDM approach

Jose Juan Garcia-Hernandez [a,*], Ramon Parra-Michel [b], Claudia Feregrino-Uribe [c], Rene Cumplido [c]

[a] The Information Technologies Laboratory, CINVESTAV-IPN, Tamaulipas, Mexico
[b] The Department of Electrical Engineering, Communications Section, CINVESTAV-IPN, Guadalajara, Mexico
[c] The National Institute for Astrophysics, Optics and Electronics, Puebla, Mexico

ARTICLE INFO

Keywords:
Data-hiding
High payload
Audio signal
OFDM
MPSK

ABSTRACT

This paper presents a high payload data-hiding scheme for audio signals following the OFDM concept. It is based on changing the phase component of the audio signal via a reduced-arc *M*-order Phase Shift Keying (MPSK) modulator on selected frequency samples of the audio signal. This approach allows the data-hiding payload and perceptual distortion to be controlled by the number of altered frequency components of the audio signal and the severity of this alteration. The data-hiding payload can be estimated *a priori* independently of the audio signal features. In addition, due to the relation of the proposed scheme with the traditional use of MPSK modulation, the BER of the encoded data is analytically derived, therefore, it is known in advance. This allows the transmission of reliable hidden data, in contrast to previous works. Results obtained with representative test cases show that the proposed algorithm provides an increase of one order of magnitude in data payload compared with state-of-the-art works.

1. Introduction

Data-hiding is a technique that embeds an imperceptible signal to the digital content. Within this technique, digital watermarking and steganography are the main concepts devoted to exploit the hidden data for different purposes. In one hand, watermarking is devoted to solving problems such as unauthorized copying and distribution of digital materials (Furon, 2005). In order to be considered suitable for practical applications, watermarking algorithms must satisfy some requirements, such as imperceptibility of the embedded signal (watermark) and robustness to some common intentional and/or non intentional attacks. These requirements have limited the payload to a few bits of hidden data. On the other hand, steganography is a kind of secret communication using digital multimedia as the communication channel where the main demand is for statistical undetectability, high payload and high perceptual transparency. Contrary to watermarking algorithms, in steganographic systems the robustness is not an issue (Petitcolas, Anderson, and Kuhn, 1999). In fact, an ideal steganographic scheme should have excellent statistical undetectability, a large embedding capacity and excellent perceptual transparency (Fridrich, 2009; Luo, Wu, Wang, Lin, & Tsai, 2010).

Among the variety of data where information can be hidden, digital images and audio signals are preferred, where a lot of work has been carried out using digital images as cover signals (Cox, Kilian, Leighton, & Shamoon, 1997; Ouhsain & Hamza, 2009; Yang, Weng, Tso, & Wang, 2011) and to a lower extent using audio signals. This work is focused on high payload data-hiding applications using audio signals, such as additional content delivery in the audio signal within a digital movie, where high payload and low perceptual distortion are pursued. During the last decade, several audio data-hiding algorithms have been proposed in the literature (Arnold, 2002; Kim & Choi, 2003; Ko, Nishimura, & Suzuki, 2005; Oh, Seok, Hong, & Youn, 2001; Seok, Hong, & Kim, 2002; Yeo & Kim, 2003). Within the main approaches, it can be cited the method of echo hiding as one of the most successful temporal domain methods (Kim & Choi, 2003; Ko et al., 2005; Oh et al., 2001). It embeds data bits using echo signals with different delays with the property of being highly imperceptible, but the achieved payload is low. Other methods, based on spread-spectrum, embed a pseudorandom sequence using a secret owner's key into some frequency bands of the audio signal (Cox, Miller, & Bloom, 2003; Kirovski & Malvar, 2003; Seok et al., 2002). In spread-spectrum method, an audio signal is transformed by Discrete Cosine Transform (DCT), Discrete Fourier Transform (DFT), or Discrete Wavelet Transform (DWT), and this pseudorandom sequence is embedded in the frequency domain considering the imperceptibility requirements. Implementation of this method is generally simple; however, the achieved payload is very low. Other algorithm, recently proposed, is based on spline interpolation and has shown one of the highest bitrates and human transparencies reported to date (Fallahpour & Megias, 2009).

* Corresponding author. Tel.: +52 834 1070220.
E-mail addresses: jjuan@tamps.cinvestav.mx (J.J. Garcia-Hernandez), rparra@gdl.cinvestav.mx (R. Parra-Michel), cferegrino@ccc.inaoep.mx (C. Feregrino-Uribe), rcumplido@ccc.inaoep.mx (R. Cumplido).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.12.021

High performance audio data-hiding algorithms are not easy to develop, because the human auditory system is very sensitive to small changes to the audio signal due to the secret data embedding (Arnold, 2002).

Modifying the phase of several frequency components is another approach used to hide data in media (Ansari, Malik, & Khohkar, 2004; Bender, Gruhl, & Morimoto, 1996). When the frequency components are obtained using an orthogonal base, these data-hiding schemes can be analyzed as Orthogonal Frequency-Division Multiplexing (OFDM) systems. OFDM is an scheme utilized as a digital multi-carrier modulation method where a large number of orthogonal sub-carriers are used to carry data. The data are divided into several parallel data streams or channels, one for each sub-carrier. Each sub-carrier is modulated with a conventional modulation scheme such as quadrature amplitude modulation (QAM) or phase-shift keying (PSK). In Tie-sheng, Jian-sheng, Wei-hong, and Da-peng (2008), the OFDM approach was used in combination with Quantization Index Modulation (QIM) in order to hide a secret image into a host image reporting a payload of 1 bit/pixel. In Lin, Pan, Shieh, and Shi (2006) the data are hidden into the physical layer of OFDM-based wireless systems. This scheme is not appropriate for high payload data-hiding purposes due to data are not hidden in the media signal but in the OFDM carriers of the wireless systems. The OFDM approach using frequency components of audio signals as carriers was presented as Acoustic OFDM in Matsuoka, Nakashima, and Yoshimura (2009). That scheme is focused on applications to transmit short messages to mobile handheld devices over aerial audio links. Acoustic OFDM uses Differential Binary Phase Shift Keying (D-BPSK) modulation. Although the robustness in analog scenarios is satisfactory, the payload reported by Acoustic OFDM is low and inappropriate for high payload data-hiding applications. The bitrate of the schemes based in OFDM concept are dependent on the audio clip used as host signal, in addition, the Objective Difference Grade (ODG) estimation is carried out after the data-hiding process. These characteristics limit the flexibility of the data-hiding process in the sense that it is not possible to know *a priori* the embedding capacity of the medium. In addition, none of the reported previous works have discussed the BER obtained trough their approaches, letting the user/system unaware of the reliability of the transmitted data.

In this paper, we propose a generalization of a data-hiding scheme for audio signals that uses the concept of OFDM and results in the highest payload capacity when compared with previous commented studies. In addition, the proposed scheme allows the hiding payload and perceptual distortion to be controlled by the number of altered frequency components of the audio signal and the severity of this alteration, in contrast to the previous OFDM-based data-hiding schemes where the payload is a function of the audio signal being used. Few parameters are defined which control, using linear equations, the number of channels per second available for transmission and the space available for the used constellation. From experiments, the border values for these parameters are obtained in order to keep the perceptual transparency at an acceptable range. In the proposed scheme, the constellation order does not affect the perceptual distortion, therefore, the hiding payload as well as the achieved BER can be defined *a priori*, independently of the audio signal features.

The proposed data-hiding scheme can be classified as a non-blind scheme because the original phases are necessary for extraction of the hidden data. Non-blind schemes are useful in network intrusion detection (Houmansadr, Kiyavash, & Borisov, 2009), anti-collusion digital fingerprinting (Kuribayashi, 2011, 2012) and to protect digital data through the Internet (El-Taweel, Onsi, Samy, & Darwish, 2005). Although non-blind condition could seem restrictive, this approach continues to be of interest to the scientific community (Bahi, Couchot, & Guyeux, 2011; Dharwadkar & Amberker, 2010; Dharwadkar, Amberker, & Gorai, 2011) because the potential applications listed above. However, in order to allows the proposed hiding algorithm to be used in blind applications, (where the original signal is not necessary for data extraction) an extension to this scheme is also provided in this paper. This blind-scheme extension maintains all the results of the original algorithm, at the expense of payload reduction.

This paper is organized as follows: Section 2 presents a background about OFDM and MPSK to ease explanation of the proposed scheme. Section 3 describes the proposed non-blind data-hiding scheme. The experiments carried out to establish the parameters of the algorithm are described in Section 4. Results and discussions from experiments are presented in Section 5. The extension for the blind approach is discussed in Section 6 while the conclusions are presented in Section 7.

2. Background

2.1. Orthogonal Frequency Division Multiplexing

Orthogonal Frequency Division Multiplexing (OFDM) is one of the most recently used modulation techniques used to combat the frequency-selective fading of the multi-path channels, permitting high data rates by reducing intersymbol interference (ISI). In a conventional data transmission system, the symbols are transmitted sequentially, one by one, with the frequency spectrum of each data symbol allowed to occupy the entire available bandwidth. A high rate data transmission supposes a very short symbol duration, implying large bandwidth requirement of the modulation symbol. There is a high likelihood that the frequency selective channel response affects in a very distinctive manner the different spectral components of the data symbol, hence introducing the ISI. It is possible to assume that the frequency selectivity of the channel can be mitigated if, instead of transmitting a single high rate data stream, the data are transmitted simultaneously, on several narrow-band subchannels (with a different carrier corresponding to each subchannel). Hence, for a given overall data rate, increasing the number of carriers reduces the data rate that each individual carrier must convey, therefore lengthening the symbol duration on each subcarrier. Slow data rate (and long symbol duration) on each subchannel merely means that the effects of ISI are greatly reduced. This is in fact the basic idea that lies behind OFDM: transmitting the data among a large number of closely spaced subcarriers. The orthogonality allows for efficient modulator and demodulator implementation using the DFT algorithm on the receiver side, and inverse DFT on the transmitter side. Although the principles and some of the benefits have been known since the 60 s, OFDM is popular for wideband communications today by way of low-cost digital signal processing components that can calculate the Fast Fourier Transform (FFT), which is an efficient DFT implementation. Figs. 1 and 2 show a simplified scheme of an OFDM transmitter and receiver, respectively, using the FFT algorithm.

In the transmitter, $s[n]$ is a serial stream of binary data, these are demultiplexed into N parallel streams $X_0, X_1, \ldots, X_{N-1}$, and mapped to a symbol stream using a symbol constellation. An inverse FFT is computed on each set of symbols and then quadrature-mixed to passband in the standard way. In the Figs. 1 and 2, $\Re e$ and $\Im m$ are the phase and quadrature components of the transmitted/received signals. The resulting $s(t)$ signal is transmitted over the air as electromagnetic waves.

The receiver picks up the signal $r(t)$ and subsequently downconverts it via a quadrature receiver by using cosine and sine waves at the carrier frequency followed by low-pass filters. After Analog-to-Digital Conversion (ADC), an FFT is performed to obtain the orthogonal components. This returns N parallel streams, each of them is

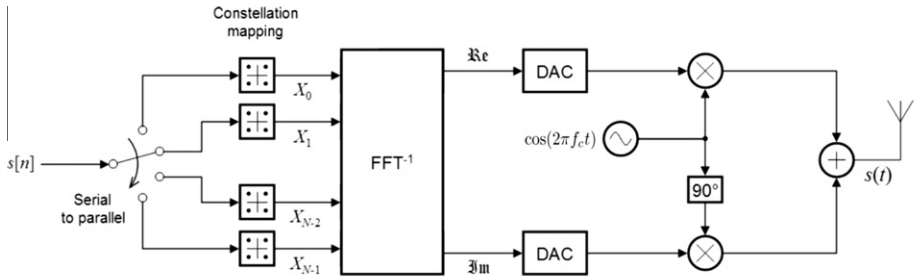

Fig. 1. OFDM transmitter.

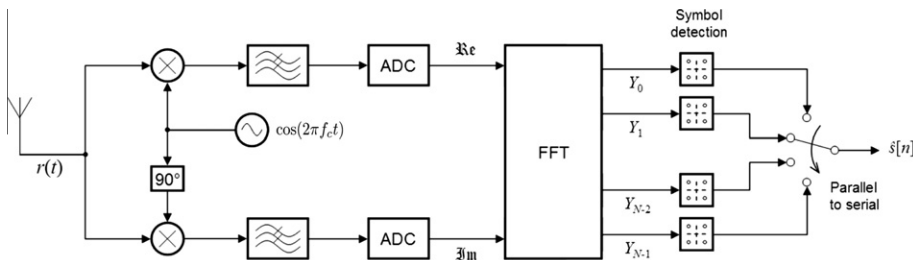

Fig. 2. OFDM receiver.

converted to a binary stream using the appropriate symbol detector. These streams are then re-ordered into a serial stream, $\hat{s}[n]$, which is an estimate of the original binary stream sent by the transmitter.

The data to be transmitted in an OFDM scheme must be modulated using a given constellation. For audio signals, the human auditory is less sensitive to changes in the phase components that changes in the magnitude components. Due that, in the present work we use MPSK modulation, that produces high perceptual transparency in the stego audio, instead of other modulation schemes such as QAM, that modifies the magnitude components. MPSK modulation is described in the next section.

2.2. MPSK modulation

M-order Phase Shift Keying is a digital modulation scheme that conveys data by changing, or modulating, the phase of a reference signal. A convenient way to represent MPSK modulations is by using a constellation diagram. In MPSK, the defined constellation points are usually positioned with uniform angular spacing around a circle. This gives maximum phase-separation between adjacent points and thus the best noise immunity. MPSK symbols are positioned on a circle so that they can all be transmitted with the same energy. Fig. 3 shows a 8-PSK constellation that is typically used in OFDM systems.

3. Proposed scheme

In this paper we utilize the OFDM concept in the sense that each frequency component of the audio signal can be interpreted as each of the carriers in the multi-carrier system. Besides, the information of each carrier is modulated using a reduced-arc MPSK modulation. That is, the original audio signal has not been eliminated, but just changed with the modulated scheme. In addition, by limiting the modifiable phase-space, the transparency of digital media can be controlled. This choices give raise to a modified version of an OFDM system, and the strength of the proposed scheme over previous approaches. The scheme is explained in two parts: insertion and extraction process.

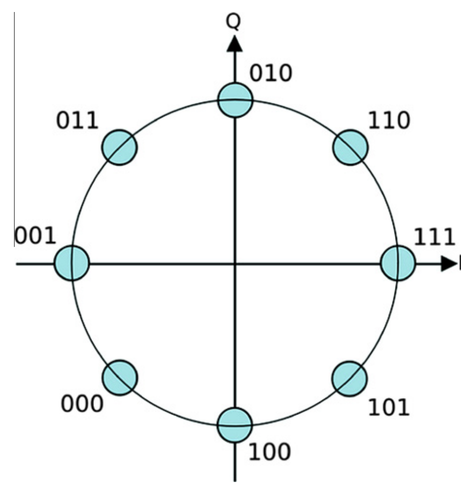

Fig. 3. 8-PSK constellation.

3.1. The insertion process

The insertion process of the proposed data-hiding scheme consists of an orthogonal transformation, and an inverse orthogonal transformation. The audio signals are transformed in orthogonal components using a DFT and those components are modified according to the reduced-arc MPSK-modulated data. In order to obtain the stego audio signal, the modified components are transformed back to the time domain using an inverse DFT.

Fig. 4 shows the general insertion process; by comparing this figure with Fig. 1 it is clear that the insertion algorithm can be interpreted as an OFDM-kind modulator, due to the parallelized data is hidden into a set of orthogonal components (the DFT components of the audio signal). This has been one of the main motivations of this research.

Each reduced-arc MPSK constellation point is added to the original phase values, P_n, in an offset fashion. Previous to the insertion process, it is necessary to define which and how many frequency values of the digital audio signal are modified in order to maintain

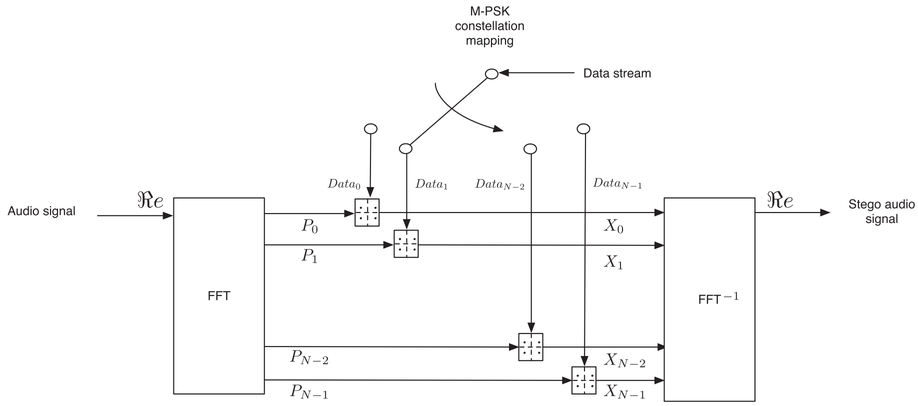

Fig. 4. The insertion process.

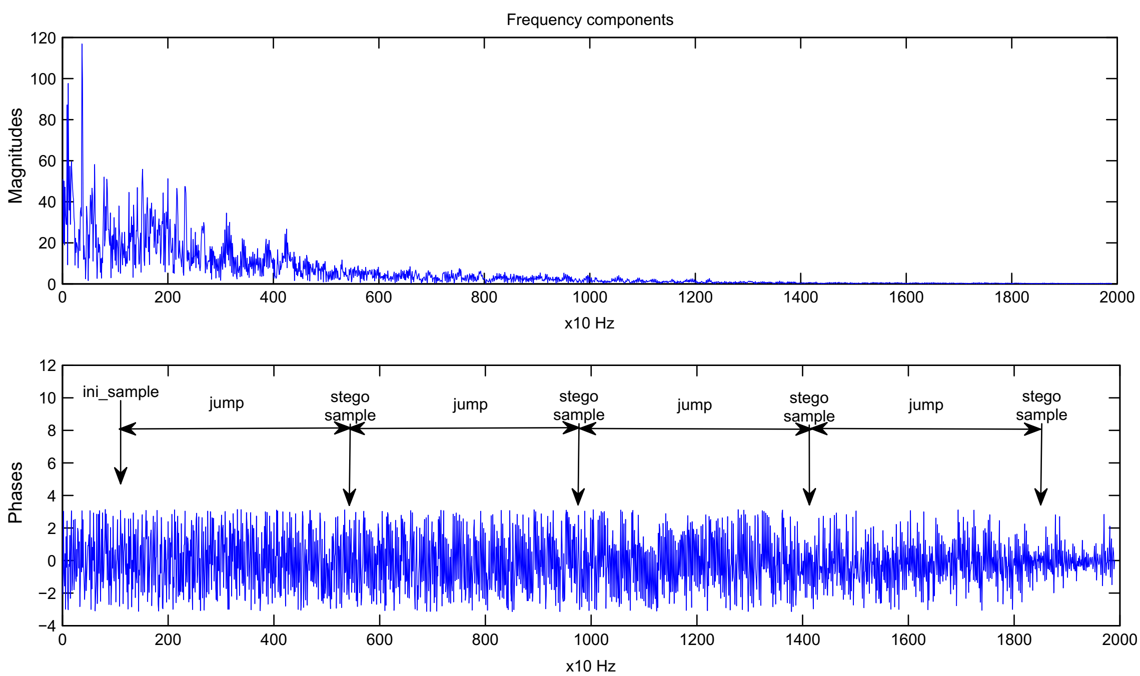

Fig. 5. Meaning of the *jump* parameter.

acceptable values of the perceptual transparency in terms of the ODG. An ODG value between 0 and −1 is considered a good perceptual transparency (Thiede et al., 2000). To define these *Phase* values, two parameters are required: *jump* and *arch*. The *jump* parameter represents the separation, in number of phase samples, between two consecutive phase samples to be modified. For example, if *jump* = 3 and the first modified phase sample starts at the one hundred sample, denoted as *Phase*[100], the next modified phase samples will be *Phase*[103], *Phase*[106], *Phase*[109], etc., until the whole frequency spectrum is covered. Fig. 5 shows a graphical meaning of the *jump* parameter.

The *arch* parameter represents the available space, in terms of arc degrees, to be used by the reduced-arc MPSK constellation. For example, if *arch* = 5° it means that the available space for the constellation is between −5° and 5°, i.e. a 10° arc. In the proposed scheme the MPSK constellation is mapped only to the space defined by *arch* instead of using the whole phase space (360°) as in communications systems, because the perceptual transparency is kept at acceptable ranges when no drastic changes are carried out on the phase components. Fig. 6 shows a graphical meaning of the *arch* parameter. As will be shown in the results, the order of the reduced-arc MPSK constellation does not affect the ODG

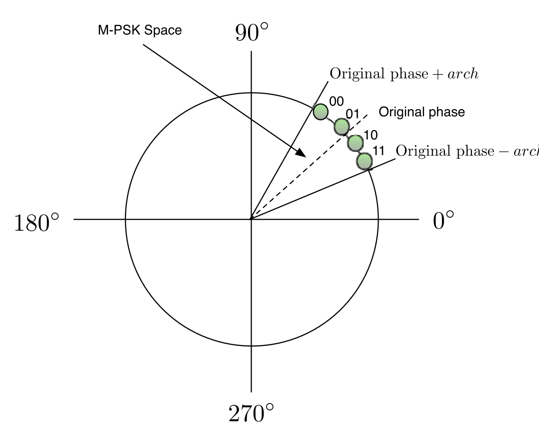

Fig. 6. Meaning of the *arch* parameter.

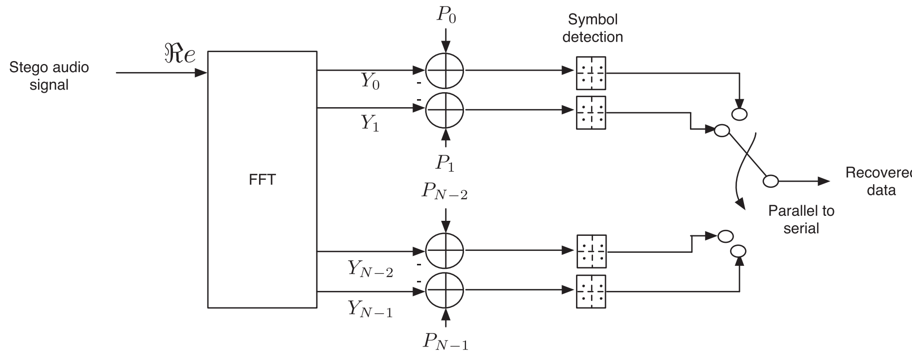

Fig. 7. The extraction process.

value for previously defined *jump* and *arch* values. Therefore, it is possible to transmit high bitrates, only limited by the allowable error probability associated with the used reduced-arc MPSK constellation and the target application.

The number of channels per second available for data-hiding is a function of the *jump* parameter as follows:

$$n\_chan = \frac{fs}{n\_samp}\left(\frac{.5n\_samp - ini\_samp}{jump}\right), \quad (1)$$

where *n_chan* is the number of channels per second available for data-hiding, *fs* is the sampling frequency of the digitalized audio signal, *n_samp* is the number of audio samples to be transformed by the DFT, and *ini_samp* is the first phase sample to be modified. In this paper a CD-quality audio signal is considered, therefore, $fs = 44100$ Hz, and $n\_samp = 4096$. The rest of the parameters are obtained using experimental results. With these values Eq. (1) becomes:

$$n\_chan = 10.7\left(\frac{2048 - ini\_samp}{jump}\right). \quad (2)$$

The bitrate is calculated as follows:

$$bit\_rate = n\_chan \log_2 M, \quad (3)$$

where *bit_rate* is the bitrate in bits per second (bps) and *M* is the order of the used reduced-arc MPSK constellation. From Eqs. (2) and (3) it can be seen that for a fixed *jump* value (*ini_samp* is a constant as it will be shown in the next section) the amount of data to be hidden is only a function of the reduced-arc MPSK order. That conclusion suggests that the proposed scheme is very flexible in terms of the payload desired to be inserted in the data.

3.2. The extraction process

The extraction process is conformed by an orthogonal transformation, the same used in the insertion process, and a reduced-arc MPSK demodulation. Fig. 7 shows the general extraction process.

The audio signals are transformed in orthogonal components using a DFT and those components are used to recover the reduced-arc MPSK-modulated data by subtracting them from the original phase components. After that, the symbol extraction is carried out by using the appropriate reduced-arc MPSK demodulator.

In what follows, the BER of the proposed scheme will be analyzed, while next section describes the experiments carried out to obtain the border values for *jump* (separation between two close phase components to be modified, measured in number of phase samples) and *arch* (available phase space for the reduced-arc MPSK constellation used for the data-hiding process) parameters in order to keep the human transparency in an acceptable range.

3.3. Bit-error probability estimation

One of the main strengths of the proposed scheme lies in its ability to predefine the bit error rate (BER) achieved in the hidden data after it is decoded. This permits the design of the associated channel coder to guarantee that what is extracted from the hidden signal is reliable data. This is opposed to previous approaches, where medium payloads are pursued without considering if hidden data are worthy to be detected, therefore, there is an uncertainty about the true performance and implementability of those schemes.

The following analysis considers the estimation of BER for the proposed OFDM data-hiding scheme and models the quantization error as additive white Gaussian noise (AWGN), a consideration commonly employed.

Firstly, one parameter, α,[1] is defined as:

$$\alpha = \frac{arch}{\pi}. \quad (4)$$

The α parameter represents the percentage of circumference being used in MPSK modulation.

Let us first recall the well-known formula for symbol-error probability in MPSK modulation systems, $P_{e,symbol}$, defined as Proakis (1983):

$$P_{e,symbol} = \frac{1}{M}\sum_{m=1}^{M} 2Q(\cdot)_{\mathrm{MPSK}}, \quad (5)$$

where $Q(\cdot)_{\mathrm{MPSK}}$ is the BER produced by each of the *M* MPSK symbols and $Q(\cdot)$ is the *Q*-function defined as:

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} \exp(-t^2/2)dt. \quad (6)$$

For Gray coding the bit-error probability, $P_{e,bit}$, is related to $P_{e,symbol}$ with:

$$P_{e,bit} = \frac{1}{k}P_{e,symbol}, \quad (7)$$

where $k = \log_2 M$. For a symetric MPSK modulation, the *M* terms in Eq. (5) are equal, and the $P_{e,bit}$ is readily reduced to Proakis (1983):

$$P_{e,bit} \approx \frac{2}{k}Q\left(\sqrt{2k\mathrm{SNR}_{Eb}}\sin\left(\frac{\pi}{M}\right)\right), \quad (8)$$

where SNR_{Eb} is the Signal-to-Noise Ratio per bit.

The BER of the proposed scheme can be calculated from the previous equations by noticing from Fig. 6 that the α parameter has the effect of shrinking the original MPSK constellation, and therefore, the constellation symbols get closer, with the exception of

[1] Radian units are used instead arc degrees.

the symbols near π. Hence, Eq. (5) should be divided into two components, one with two Gaussian tails and the other with $M-1$ Gaussian tails. Under the commented considerations and after simple manipulations, the $P_{e,bit,ra}$ of the proposed scheme results in:

$$\begin{aligned} P_{e,bit,ra} &\leqslant \frac{M-1}{M}\left(\frac{2}{k}\right)Q(\cdot)_{\alpha} + \frac{2}{Mk}Q(\cdot)_{1-\alpha}, \\ &\leqslant \frac{M-1}{M}\left(\frac{2}{k}\right)Q\left(\sqrt{2k\mathrm{SNR}_{Eb}}\sin\left(\frac{\pi\alpha}{M}\right)\right) \\ &\quad + \frac{2}{Mk}Q\left(\sqrt{2k\mathrm{SNR}_{Eb}}\sin\left((1-\alpha)\pi + \frac{\pi\alpha}{M}\right)\right), \end{aligned} \tag{9}$$

where the inequality is required due to the fact that the considerations for deriving Eq. (8) have been stressed. Nevertheless, this approximation is valid and close enough when $M \geqslant 4$. When M gets bigger, it is possible to dismiss the second component and to simplify Eq. (9) as follow:

$$P_{e,bit,ra} \approx \frac{M-1}{M}\left(\frac{2}{k}\right)Q\left(\sqrt{2k\ \mathrm{SNR}_{Eb}}\sin\left(\frac{\pi\alpha}{M}\right)\right). \tag{10}$$

From Eq. (10) it is possible to see that if α goes smaller and/or M goes bigger then, it is necessary to rise the SNR_{Eb} value in order to keep BERs small enough, in other words, if the percentage of circumference used by MPSK modulation is smaller. Nevertheless, due to the high Signal-to-Quantization-Noise-Ratio (SQNR) in audio digital signals, the resulting BER is significantly low, as will be discussed below.

4. Experiments

In order to determinate the best values for *ini_samp*, *jump* and *arch* parameters, several experiments were carried out. Due to one of the aims of the proposed scheme is to keep the perceptual transparency at an acceptable range, only medium and high frequencies are modified and the low frequencies are kept unaltered, as the human hearing sensitivity is higher in low frequency bands than in medium and high frequency bands, according to the human auditory system (HAS) (Katzenbeisser & Petitcolas, 2000). The cut-off frequency for the data-hiding process is represented by the *ini_samp* parameter and was estimated via the following simulation settings:

- For 10,000 repetitions, an audio signal was divided in blocks of 4096 samples (that block size provides enough resolution for audio processing purposes). Each block was transformed to the frequency domain by the FFT algorithm. From the 2048 frequency components 64 phase samples[2] were modified, with *ini_samp*, *jump* and *arch* parameters being random numbers in order to simulate different configurations of available constellation space and amount of phase samples being modified.
- Due to the ODG value is the unique perceptual transparency indicator, for each repetition the ODG was calculated and registered.
- For ODG values within the range $[0,-1]$, the highest *ini_samp* value was selected because that range guarantees perceptual transparency.

From this procedure, the *ini_samp* value obtained was 120, which corresponds to the 1.3 kHz component. This means that the frequency components below of 1.3 kHz will not be modified during the data-hiding process.

[2] Preliminary experiments shown that a lower or higher number of components being modified does not significantly vary the *ini_samp* estimation.

In order to know the combinations of *jump* and *arch* parameters that guarantee the perceptual transparency determined for the ODG value, the next experiment was carried out:

- The *ini_samp* parameter was set to 120 as it was obtained from the last experiment.
- The *arch* parameter was varied in a random fashion (simulating the points of an MPSK constellation) from ±1° to ±178° for each *jump* value from 1 to 30. Note that if the number of unmodified phase samples between two samples being modified is higher that 30, the number of available channels for data-hiding is not attractive for high payload data-hiding applications because it would be similar to the one obtained in previous works, therefore the bitrate will be too low.
- The ODG value was calculated for each pair, and recorded in a matrix **ODG** [*arc*,*jump*].
- The procedure was repeated for five different songs and the ODG matrices were averaged in order to merge the characteristics of each song in the general estimation. Each song belongs to different sorts of music such as pop, classic, latin, big band and rock music in order to stress the performance of the proposed scheme and validate its suitability for real applications.

The border values for *jump* and *arch* parameters are defined in function of ODG values in the range $[0,-1]$ because that range of the ODG values is the acceptable in terms of perceptual transparency. This means that only the combinations of available constellation space and amount of phase samples that guarantee perceptual transparency will be used.

5. Experimental results and discussion

Fig. 8 shows the ODG values obtained after performing the experiments, described on the previous section, for several amounts of separation between two samples to be modified. As it is expected, for high separation between modified samples the ODG improves because a smaller number of phase samples are modified in comparison with a low separation according to Eq. (2). This happens because when a low number of phase samples is modified the introduced distortion is minor than when a higher number of phase samples is altered.

Table 1 shows the *jump* and *arch* useful values (those that produce ODG values within the acceptable range $[0,-1]$) and their *n_chan* and ODG associated values. From this table, it can be seen that for *jump* = 2, the highest *arch* value for ODG in the range $[0,-1]$ is 25°. On the other hand, for *jump* = 30, the highest *arch* value for ODG in the range $[0,-1]$ is 89°.

From Fig. 8 it is possible to appreciate that each ODG curve (for each *jump* value) could be approximated by a linear equation. A frequently used technique for approximating a phenomena by a linear equation is using the multiple regression method (Draper & Smith, 1998). In order to obtain an ODG equation, an online multiple regression computer (Wessa, 2010) was used. The resulting ODG equation is:

$$\begin{aligned} ODG[arch, jump] = 0.30592979 &- 0.00601405\,jump \\ &- 0.02242764\,arch + 0.00030866\,(jump) \\ &\times (arch). \end{aligned} \tag{11}$$

Eq. (11) allows us to determine the ODG value that will be obtained for a pair of defined *jump* and *arch* values. Moreover, it is possible to define ODG = -1 and to select the *arch* and *jump* values according to a specific application. For example, in an application where the goal is to transmit at high bitrates, the separation between two samples to be modified (*jump* value) must be low in order to use a high

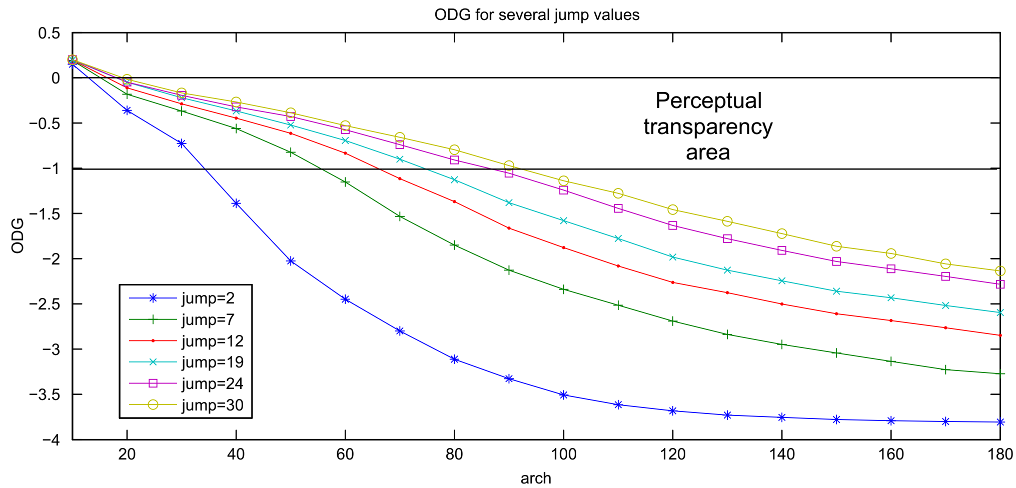

Fig. 8. ODG plot for several *jump* values.

Table 1
jump and *arch* useful values (ODG in the range [0, −1]).

jump	*arch* (°)	Available space (°)	*n_chan*	ODG
2	25	50	10314.8	−0.97335
3	31	62	6869.4	−0.99343
4	36	72	5157.4	−0.98230
5	40	80	4119.5	−0.99798
6	44	88	3434.7	−0.99825
7	46	92	2942.5	−0.99072
8	49	98	2578.7	−0.99869
9	51	102	2289.8	−0.97790
10	53	106	2054.4	−0.97532
11	55	110	1872.5	−0.97306
12	57	114	1712.0	−0.99522
13	59	118	1583.6	−0.98800
14	61	122	1465.9	−0.96209
15	60	120	1369.6	−0.96558
16	64	128	1284.0	−0.98832
17	64	128	1209.1	−0.98328
18	69	138	1144.9	−0.97832
19	69	138	1080.7	−0.99054
20	72	144	1027.2	−0.98291
21	75	150	973.7	−0.99706
22	76	152	930.9	−0.99426
23	77	154	888.1	−0.99413
24	80	160	856.0	−0.99337
25	80	160	823.9	−0.99960
26	83	166	791.8	−0.97762
27	83	166	759.7	−0.99201
28	87	174	727.6	−0.99507
29	86	172	706.2	−0.95715
30	89	178	684.8	−0.97239

number of channels for transmission. State-of-the-art data-hiding schemes such as Fallahpour and Megias (2009, 2010b) transmit about 3 kbps and 5 kbps, respectively without being able to declare the BER their studies achieve. In the proposed scheme, if *jump* = 2 and the modulator is reduced-arc 256-PSK it is possible to transmit about 82.5 kbps with a good perceptual transparency. Moreover, in a CD-quality audio signal, the BER is ≈0 due to the audio signal' SQNR is about 93 dB. Theoretically, if the precision used in the implementation of the proposed scheme were equal to infinite, it would be possible to hide bit streams using a reduced-arc MPSK constellation of order equal to infinite. However, during this research, utilizing double precision arithmetic in the implementation, the maximum order of the reduced-arc MPSK constellation used is 256, which results in up to 8 bits per channel. Under those conditions the minimum payload achieved is 5.47 kbps (*jump* = 30, *arch* = 89° and reduced-arc 256-PSK constellation) and the maximum one is 82.5 kbps (*jump* = 2, *arch* = 25° and reduced-arc 256-PSK constellation).

Table 2 shows a comparison with several state-of-the-art data-hiding schemes in audio signals. From this results, it is possible to see that the proposed scheme, with the values used in this paper (*jump* = 2, *arch* = 25° and reduced-arc 256-PSK constellation), can hide about 9.4 times more bits than the highest payload of a data-hiding scheme reported to date, while at the same time, the noise immunity of the payload is known.

At this point, for the purposes of testing the validity of the theoretical analysis for the presence of AWGN through experimentations. Fig. 9 shows, in BER terms, the behavior of the proposed scheme with *jump* = 30, *arch* = 89° and reduced-arc B-PSK constellation. The label Analytical corresponds to the plot obtained using Eq. (10). It can be noted that the match between the analytical results and the experimental ones is very close.

Finally, it is worth noting that the proposed scheme was devoted to the high payload data-hiding application. However, due to its high payload, it could be possible to apply the proposed OFDM-based data-hiding approach in watermarking applications where robustness is a major concern. In watermarking systems, robustness means to keep BER low in a noisy channel. By using Error Correcting Codes (ECC) it is possible to mitigate BER at bitrate cost. In order to find out which class of ECC is the adequate for achieving a satisfactory robustness for a given scenario, it is necessary to carry out further research and it constitutes an open research issue. However, because the high payloads achieved by the proposed scheme, it is possible to expect that, despite bitrate reduction by ECC, the payload will be high and still attractive for both watermarking and high payload data-hiding applications.

6. Extension to blind data-hiding scheme

Although it is clear that non-blind approaches continue to be of interest to the scientific community (Dharwadkar & Amberker, 2010; Bahi et al., 2011; Dharwadkar et al., 2011), over the last few years blind watermarking has been preferred for practical applications. The proposed scheme can be classified as a non-blind watermarking scheme as it requires the original signal to extract the hidden data, so, to cover applications where the original signal is not available during the extraction process, in this section an extension to the scheme is proposed in order to evolve it to a blind scheme at the expense of payload. The modification consists in

Table 2
Comparison of different data-hiding schemes.

Algorithm	Payload (bps)	ODG	BER known *a priori*
Ansari et al. (2004)	1000	Not reported	Not reported
Lin and Abdulla (2007)	315	Not reported	Not reported
Garcia-Hernandez et al. (2008)	689	Not reported	Not reported
Fallahpour and Megias (2009)	2996	−0.60	Not reported
Matsouka et al. (2009)	240	Not reported	Not reported
Bhat et al. (2010)	46	Not reported	Not reported
Fallahpour and Megias (2010a)	5501	−0.50	Not reported
Fallahpour and Megias (2010b)	1478–8719	−0.18 to −0.78	Not reported
Fallahpour and Megias (2011)	11,000	−0.4 to −0.8	Not reported
Wu et al. (2011)	6500	Not reported	Not reported
Lei et al. (2012)	170	Not reported	Not reported
Proposed	5472–82,518	0.0 to −0.97	Provided by Eq. (10)

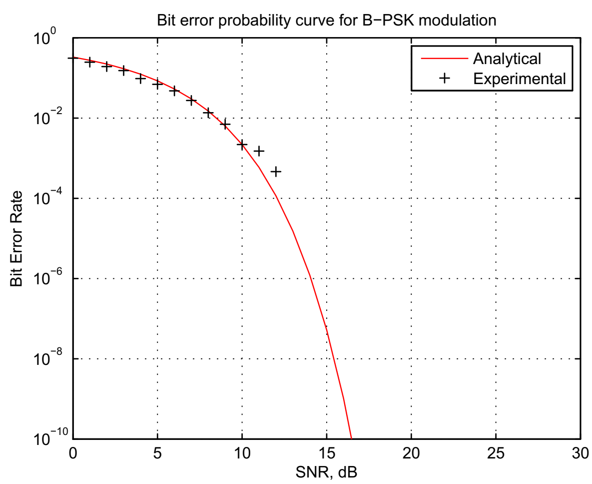

Fig. 9. BER versus SNR for *jump* = 30, *arch* = 89° and reduced-arc B-PSK constellation.

knowing the real value (the actual audio phase value) of the signal where the hidden data is to be inserted. This can be performed by taking this value and inserting it into a set of subcarriers. Therefore, for each carrier to be modified, a second phase carrier should be modified for holding the actual value of the first carriers. This modification permits blinded transmission scheme at the expenses of reducing the payload at the half. The insertion procedure is subsequently discussed in detail.

6.1. Insertion process

For insertion the strategy is as follows:

- Using a linear interpolator compute the interpolated values P'_j for a half of the available channels, $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$.
- Compute the interpolation-error e_j using the original P_j and interpolated phase P'_j values for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$ as follows:

$$e_j = P_j - P'_j. \tag{12}$$

- Add the interpolation-error information to the adjacent channels as follows:

$$P_{j+\frac{n\_chan}{2}} = \text{round}\left(P_{j+\frac{n\_chan}{2}}\right) + \frac{e_j}{1000} \tag{13}$$

for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$, where round $(\cdot)$ returns the rounded value of the argument in degree units.

- Embed data into P_j for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$ using the proposed scheme as described in Section 3.1.

6.2. Extraction process

For data extraction the procedure is as follows:

- With the linear interpolator used in the insertion process compute the interpolated values P'_j for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$.
- Extract the interpolation-error information from the adjacent channels as follows:

$$e_j = 1000 * \left(P_{j+\frac{n\_chan}{2}} - \text{round}\left(P_{j+\frac{n\_chan}{2}}\right)\right) \tag{14}$$

for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$.

- Using P'_j and e_j calculate the original P_j values as follows:

$$P_j = P'_j + e_j \tag{15}$$

for $j = 1, 2, 3, \ldots, \frac{n\_chan}{2}$.

- Finally, extract hidden data as described in Section 3.2.

Converting the proposed non-blind scheme to a blind scheme is done at the expense of half the payload capacity, because a pair of channels is necessary to embed a constellation symbol; one channel is used for embedding the constellation symbol and the other carries necessary information for detection as it was described above. Perceptual transparency is unaltered since distortion due to the insertion of interpolation-error information is very low. Distortion due to round $(\cdot)$ is ±0.5° at most and, in the worst scenario, $\frac{e_j}{1000} = \pm 0.18°$. Therefore, the maximum distortion due to the insertion of interpolation-error information is ±0.68°. From Table 1 it is possible to observe that the minimum value of *arch* that guarantees the original perceptual transparency after the extension to the blind scheme is 25° which is almost two orders of magnitude than the required modification for the blind scheme. The second main modification is the achieved BER. The multiplication times 1000 in (14) for the mean value of the data subcarriers, directly enhances the noise by 30 dB. Adding this noise to the data subcarriers has the effect of reducing the actual SNR_{Eb} by 30 dB. Therefore, by changing SNR_{Eb} by $SNR_{Eb} - 30$ dB in Eq. (10), the BER for the blind scheme can be evaluated *a priori* to determine the modifications in the parameters of the algorithm for any given application. The $P_{e,bit,ra,blind}$ equation for the blind scheme becomes:

$$P_{e,bit,ra,blind} \approx \frac{M-1}{M}\left(\frac{2}{k}\right)Q\left(\sqrt{2k(\text{SNR}_{Eb} - 30\text{ dB})}\,\sin\left(\frac{\pi\alpha}{M}\right)\right). \tag{16}$$

6.3. Experimental results

In order to verify that the proposed extension maintains the ODG values according to the estimates for non-blind scheme, several audio clips were watermarked and evaluated in terms of perceptual transparency. Fig. 10 shows the ODG values for 25

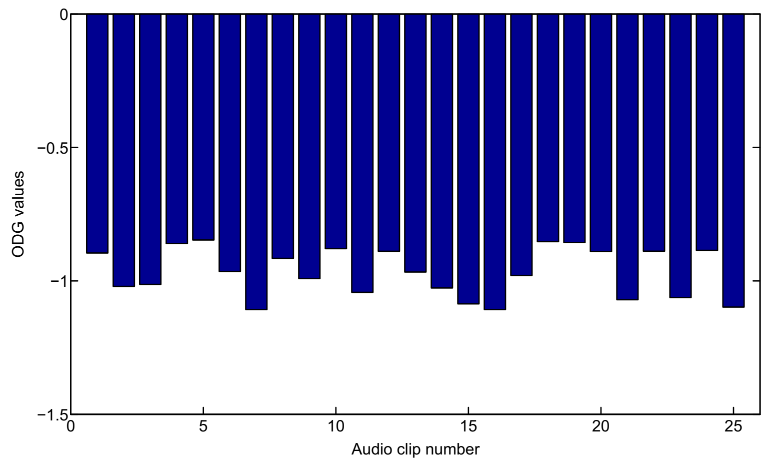

Fig. 10. ODG values for 25 audio clips using the blind extension.

CD-quality audio clips using the blind extension of the proposed algorithm. The system values for these experimentations are: *jump* = 2, *arch* = 25° and *M* = 256.

As it can be seen in Fig. 10, ODG values are contained in the vicinity of -1 according to expectations in Table 1. These results confirm that only the payload is affected.

The extension presented in this section increases the versatility of the proposed data-hiding scheme since it is suitable for applications where the original signal is not available. Despite the reduction in payload and the added 30 dB margin, due to the extension of the proposed scheme, it is still higher than those reported in the state-of-the-art if the appropriate parameters, *jump*, *arch* and *M* are chosen while at the same time, the achieved BER is also included at the first time in this area.

7. Conclusions

In this paper, a data-hiding scheme for digital audio signals based on the OFDM concept was presented. Two parameters were defined in order to regulate the payload and the perceptual transparency. Those parameters represent the number of samples to be modified and the severity of the modification. In order to guarantee the perceptual transparency, after some experiments it was possible to determine the border values for modification such that ODG is kept within the acceptable range $[0, -1]$. Moreover, an ODG equation depending of the relevant parameters is provided by using the multiple regression strategy. That equation lets the user know the perceptual distortion independently of the audio clip. In the same form, the order of the utilized reduced-arc MPSK modulator in the scheme does not affect the perceptual transparency. The bitrate can be estimated *a priori* and does not depend of the audio signal. The proposed data-hiding scheme has shown a payload about 10x higher that the highest payload in an audio data-hiding scheme reported to date. Finally, an extension to the proposed scheme is presented in order to evolve it for blind scheme applications keeping a high payload and *a priori* BER knowledge.

Acknowledgment

The authors thank CONACyT for financial support under Grants CB-2007-1-84668 and CB-2010-1-50910.

References

Ansari, R., Malik, H., & Khohkar, A. (2004). Data hiding in audio using frequency-selective phase alteration. *International conference on acoustics, speech, and signal processing* (Vol. 5, pp. 389–392). IEEE.

Arnold, M. (2002). Audio watermarking: Features, applications and algorithms. *International conference on multimedia and expo* (Vol. 2, pp. 49–54). IEEE.

Bahi, J., Couchot, J., & Guyeux, C. (2011). Steganography: A class of algorithms having secure properties. *Seventh international conference on intelligent information hiding and multimedia signal processing* (Vol. 7, pp. 209–212). I.C. Society.

Bender, W., Gruhl, D., & Morimoto, N. (1996). Techniques for data hidding. *IBM Systems Journal, 35*, 313–336.

Bhat, V., Sengupta, I., & Das, A. (2010). An adaptive audio watermarking based on the singular value decomposition in the wavelet domain. *Digital Signal Processing, 20*, 1547–1558.

Cox, I., Kilian, J., Leighton, T., & Shamoon, T. (1997). Secure spread spectrum watermarking for multimedia. *IEEE Transactions on Image Processing, 6*, 1673–1687.

Cox, I., Miller, M., & Bloom, J. (2003). *Digital watermarking*. Morgan Kaufman.

Dharwadkar, N., & Amberker, B. B. (2010). An efficient non-blind watermarking scheme for color images using discrete wavelet transformation. *International Journal of Computer Applications, 2*, 60–66.

Dharwadkar, N., Amberker, B. B., & Gorai, A. (2011). Non-blind watermarking scheme for color images in rgb space using dwt-svd. In *International conference on communications and signal processing* (pp. 489–493). Kerala, India: IEEE.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis*. Wiley-Interscience.

El-Taweel, G. S., Onsi, H. M., Samy, M., & Darwish, M. G. (2005). Secure and non-blind watermarking scheme for color images. *International Journal on Graphics, Vision and Image Processing, 5*, 1–5.

Fallahpour, M., & Megias, D. (2009). High capacity audio watermarking using fft amplitude interpolation. *IEICE Electronics Express, 6*, 1057–1063.

Fallahpour, M., & Megias, D. (2010a). Dwt-based high capacity audio watermarking. *IEICE Transactions on Fundamentals, E93*, 331–335.

Fallahpour, M., & Megias, D. (2010b). Robust high-capacity audio watermarking based on fft amplitude modification. *IEICE Transactions on Information and Systems, E93*, 87–93.

Fallahpour, M., & Megias, D. (2011). High capacity audio watermarking using the high frequency band of the wavelet domain. *Multimedia Tools and Applications, 52*, 485–498.

Fridrich, J. (2009). *Steganography in digital media: Principles, algorithms, and applications*. Cambridge University Press.

Furon, T. (2005). A survey of watermarking security. In *International workshop on digital watermarking. Lecture notes on computer science* (Vol. 3710, pp. 201–215). Springer.

Garcia-Hernandez, J. J., Nakano, M., & Perez, H. (2008). Data hiding in audio signals using rational dither modulation. *IEICE Electronics Express, 5*, 217–222.

Houmansadr, A., Kiyavash, N., & Borisov, N. (2009). Rainbow: A robust and invisible non-blind watermark for network flows. In *16th Annual network and distributed system security symposium*. The internet society.

Katzenbeisser, S., & Petitcolas, F. (2000). *Hiding techniques for steganography and digital watermarking*. Boston, MA: Artech House, Inc..

Kim, H. J., & Choi, Y. H. (2003). A novel echo hiding algorithm. *IEEE Transactions on Circuits and Systems for Video Technology, 13*, 885–889.

Kirovski, D., & Malvar, H. (2003). Spread spectrum watermarking of audio signals. *IEEE Transactions on Signal Processing, 51*, 1020–1033.

Ko, B. S., Nishimura, R., & Suzuki, Y. (2005). Time spread echo method for digital audio watermarking. *IEEE Transactions on Multimedia, 7*, 212–221.

Kuribayashi, M. (2011). Hierarchical spread spectrum fingerprinting scheme based on the cdma technique. *EURASIP Journal on Information Security*, 2011.

Kuribayashi, M. (2012). Interference removal operation for spread spectrum fingerprinting scheme. *IEEE Transactions on Information Forensics and Security, 7*, 403–417.

Lei, B., Soon, I. Y., Zhou, F., Li, Z., & Lei, H. (2012). A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition. *Signal Processing*. http://dx.doi.org/10.1016/j.sigpro.2011.12.021.

Lin, Y. Q., & Abdulla, W. H. (2007). Multiple scrambling and adaptive synchronization for audio watermarking. In *IWDW '07 proceedings of the sixth international workshop on digital watermarking. Lecture notes on computer science* (Vol. 3304, pp. 456–469). Springer.

Lin, C., Pan, J., Shieh, C., & Shi, P. (2006). An information hiding scheme for ofdm wireless networks. In *International conference on intelligent information hiding and multimedia signal processing (IIH-MSP'06)* (pp. 51–54).

Luo, D., Wu, N., Wang, C., Lin, Z., & Tsai, C. (2010). A novel adaptive steganography based on local complexity and human vision sensitivity. *The Journal of Systems and Software, 83*, 1236–1248.

Matsouka, H., Nakashima, Y., & Yoshimura, T. (2009). Acoustic ofdm system and its extension. *The Visual Computer: International Journal of Computer Graphics, 25*, 3–12.

Oh, H. O., Seok, J. W., Hong, J. W., & Youn, D. H. (2001). New echo embedding technique for robust and imperceptible audio watermarking. In *IEEE international conference on acoustics, speech and signal processing, Vol. 3* (pp. 1341–1344).

Ouhsain, M., & Hamza, A. B. (2009). Image watermarking scheme using nonnegative matrix factorization and wavelet transform. *Expert Systems with Applications, 36*. 2131-2129.

Petitcolas, F., Anderson, R., & Kuhn, M. (1999). Information hiding – a survey. In *Proceedings of the IEEE: Special issue on identification and protection of multimedia content, Vol. 87* (pp. 1062–1078).

Proakis, J. G. (1983). *Digital Communications*. McGraw Hill.

Seok, J., Hong, J., & Kim, J. (2002). A novel audio watermarking algorithm for copyright protection of digital audio. *ETRI Journal, 24*, 181–189.

Thiede, T., Treurniet, W., Bitto, R., Schmidmer, C., Sporer, T., Beerens, J., et al. (2000). Peaq – the itu standard for objective measurement of perceived audio quality. *AES, 48*, 3–29.

Tie-sheng, F., Jian-sheng, X., Wei-hong, X., & Da-peng, Q. (2008). Ofdm information hiding method by preposition embedded qim. In *International conference on computational intelligence for modelling control & automation* (pp. 19–24). I.C. Society.

Wessa, P. (2010). Free Statistics Software, version 1.1.23-r5. Technical report office for research development and education. <http://www.wessa.net/>.

Wu, T. S., Lin, H. Y., Hu, W. C., & Chen, Y. S. (2011). Audio watermarking scheme with dynamic adjustment in mute period. *Expert Systems with Applications, 38*, 6787–6792.

Yang, C., Weng, C., Tso, H., & Wang, S. (2011). A data hiding scheme using the varieties of pixel-value differencing in multimedia images. *The Journal of Systems and Software, 84*, 669–678.

Yeo, I.-K., & Kim, H. J. (2003). Modified patchwork algorithm: A novel audio watermarking scheme. *IEEE Transactions on Speech and Audio Processing, 11*, 381–386.