Control and Cybernetics

vol. **40** (2011) No. 4

Steganographic algorithm of hiding information in sound based on Fourier transform and masking*†

by

Grzegorz Kozieł

Institute of Computer Science, Lublin University of Technology
Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract: A new steganographic method based on Fourier transform is proposed. Data is attached in the frequency domain within the limits of the audible band. Masked frequencies are used to conceal information in order to make the changes inaudible for human ear. This approach allows for modification in the sound frequency band only, and does not require any additional modifications. It also allows for concealing information efficiently, minimizing the number of changes and gaining robustness with respect to popular sound transformations.

Keywords: keywords steganography, hidden communication, Fourier transform.

1. Introduction

Steganography is a discipline interested in information protection. It becomes more and more efficient and it competes with omnipresent cryptography, due to a fresh approach to the problem of information protection. In distinction to traditional protection, the anonymity of both communicating parties is offered. The sender places data in a carrier called container by using a fixed algorithm and a previously negotiated key. This results in obtaining of a *stegocontainer*, which is a file containing the added data. It is important that the original file and the stegocontainer cannot be told apart without expert analysis. The so prepared carrier is subsequently put in a public place and this allows for both communicating parties to stay anonymous (Wayner, 2002).

There are a lot of different steganographic methods using different types of containers. The Least Significant Bits methods are characterised by the highest capacity. These methods involve replacing the least significant bits of each sample of the container with hidden data bits. They can work in the time

*Submitted: September 2010; Accepted: September 2011

†This is an extended and amended version of the paper, presented at the 5th Congress of Young IT Scientists (Międzyzdroje, 23-25.IX.2010).

or frequency domains, or any other which allows for the full reversal of the operation (Agaian et al., 2005; Cvejic and Seppanen, 2002a,b; Delforouzi and Pooyan, 2008). These techniques are often combined with other ones, such as minimum error replacement (Cvejic and Seppanen, 2002b) or error dispersal (Cvejic and Seppanen, 2004). They can also use masking effect (Cvejic and Seppanen, 2002a). These methods are not resistant to data corruption. Some authors implement in their algorithms solutions, which result in increased resistance to data corruption. However, due to the high steganographic capacity and utilization of the least change-resistant part of the container, only a minor increase of resistance can be achieved (Cvejic and Seppanen, 2004; Gopalan, 2003, 2005).

In order to avoid the introduction of additional signal deformation, modifications may be performed on the existing noise. Some authors propose the use of adequately modified noise removal algorithms, which enables improvement of the signal quality during the steganographic data hiding (Katzenbeisser and Petricolas, 2000; RLE, 1999).

To obtain a higher resistance of the hidden data to signal modification we need to use specific transformations, which frequently introduce significant interference. It is necessary then to take advantage of the human auditory system (HAS) imperfections to mask the introduced changes.

A watermarking technique using Fourier transform and masking effect is presented in Tachibana et al. (2001).

Echo introducing steganographic methods gained a widespread popularity. Various forms of this approach (Bender et al., 1996; Dymarski, 2006; Dymarski et al., 2003; Garay, 2002; Gruhl and Lu, 1996; Johnson and Katzenbeisser, 2000; Kim, Kwon and Bae, 2004) allow for obtaining good resistance to the hidden data damage. The subband filtration method presented in Dymarski (2006) also provides a high resistance level, but in some cases it may generate audible interference. High resistance can be achieved with techniques using phase coding or phase modulation of the sound (Bender et al., 1996, Johnson and Katzenbeisser 2000; Matsuka 2006). These methods are characterised by low steganographic capacity. It is much more difficult to obtain good resistance in the time domain. It was possible thanks to hiding data in histogram modifications (Xiang, Huang and Yang, 2007; Xiang, Kim and Huang, 2008) and through distance modification between significant signal points.

Techniques for hiding data in an image are proposed in Bao and Ma (2004), Santosa and Bao (2005). In this case an audio signal is transformed into an image, which serves the purpose of hiding information. The obtained stegocontainer is then transformed back into sound. This approach allows for achieving resistance to mp3 compression.

Solutions found in literature show that better results in the field of resistance to damage are possible when using transformation into other signal representation domains. No descriptions of any effective communication method, which uses Fourier transform exist in the literature. Research into these types of

methods has been stopped due to serious problems caused by generating audible interference. Solution to this problem was found during research conducted by the present author. Methods used and results obtained are presented in this article.

2. Fourier transform and Fourier transform-based steganographic methods

Transform methods are based on transforming the traditional sound recording in the frequency domain. This is done by using a chosen transform. Adding extra data to the signal is possible by changing the transform coefficients. Inverse transform allows for returning to the time domain. For a transformation to be used in steganography it must be fully reversible.

Attaching the data in the frequency domain is more reliable than substitution, as the introduced changes are dispersed within the whole signal recording, also by using most significant bits of the recording. Fourier transform can be used for information concealment purposes. This transformation approximates the signal by assembling numerous sine and cosine functions (Czyżewski, 2001). Discrete Fourier transform (DFT, equation (1)) of a linear signal is used for digital signal processing (Izydorczyk, Płonka and Tyma, 2006; Zieliński, 2005):

$$X(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{\frac{j2\pi kn}{N}}, \tag{1}$$

where: k – base function index, N – number of signal samples.

As a result of DFT we obtain the signal spectrum which is a set of transform coefficients (strips). Each strip represents a defined frequency and its value defies the extent of that frequency in the sound.

In order to change the frequency signal representation back to the time form, inverse discrete Fourier transform (IDFT, equation (2)) should be used (Izydorczyk, Płonka and Tyma, 2006; Zieliński, 2005):

$$x(n) = \sum_{k=0}^{N-1} X(k) e^{\frac{j2\pi kn}{N}}, \tag{2}$$

where: k – base function index, n – number of signal sample, $X(k)$ – DFT sample value, N – number of signal samples (Czyżewski, 2001).

Fourier transform-based steganographic methods use modification of chosen spectrum strips values or the sound phase in order to hide the additional information. When one DFT coefficient amplitude is modified, the presence of this component corresponds to the hidden binary value of *1* and lack of it – corresponds to the binary value of *0*. When two DFT coefficients are used, the rate of their values is analyzed. The higher ratio of f_1 frequency is equal to encoding of *1*, and the higher ratio of f_2 frequency is equal to encoding of *0*.

One bit can be concealed in the processed signal using the presented method. It is necessary to divide the signal into blocks (signal fragments having definite number of samples) in order to gain a higher steganographic capacity. Each block carries one bit of information. To start with, the block is transformed, using DFT, to frequency representation and the values of the chosen spectrum strips are determined. Those values are modified according to the established algorithm in the next step. The prepared spectrum is retransformed into the time form using IDFT. The processed blocks are combined into one signal.

However, human ear is very sensitive to frequency changes. To achieve distortion inaudibility in the presented method, the attached signal must have very low power. This, in turn, results in very low robustness and only allows for the use of frequencies higher than 16kHz, as they are inaudible for the human ear. But this frequency band is often omitted during lossy compression or lowering of sampling frequency. Moreover, the usual sound frequency range is between 20Hz to 10kHz. During signal spectrum analysis it is very easy to notice significant increase in the values of the spectrum strips representing frequencies above 10kHz. This results in the presence of concealed information being easily detectable. The method of frequency modification in sound was rejected due to this fact (Johnson and Katzenbeisser, 2000).

It is possible to modify the presented method of spectrum change introduction. Instead of attempting to achieve the form in which frequency f_1 has a bigger share in the signal than frequency f_2, it is also possible to hide data using only the difference of values. The method proposed by the author relies on achieving a difference between two spectrum strip values, which is equal R in the binary *1* coding, or *0* in the binary *0* coding. The changes are introduced so as to minimize their strength and impact on the sound timbre. The same value changes are introduced to spectrum strips while coding the binary *0*. Consequently, we have the processed strip values at the level of their previous mean value. If there is a need to make a change during the binary *1* coding, we first reduce the strip having smaller value. If this is not enough, then we increase the one having bigger value. The introduced interference audibility tapering is possible by minimizing changes introduced to the strip having bigger value while introducing bigger changes to the strip having smaller value. Consequently, the strip having bigger value additionally masks modifications made in the one having smaller value.

Investigations conducted by the author proved that achieving the same level of the hidden data robustness is possible by using methods described by other authors and the proposed one as well. The method presented allows for reducing the strength of the introduced changes and significantly decreases the emerging noise audibility.

3. Masking

Masking is a phenomenon of human hearing not being able to register some of the sound (masked), because they are 'jammed' by other sounds (maskers). There are two types of masking:

1. simultaneous,
2. non-simultaneous.

Non-simultaneous masking is based on blocking of signal perception by other, louder signal appearing not later than 40ms after or up to 200 ms before the masked signal (Jorasz, 1999).

Frequency masking (simultaneous) is based on masking the quieter sound by a simultaneously appearing louder sound with similar frequency. The masked sound has to be below the masking threshold. The shape of the masking threshold depends on the frequency and character of the masked tone and the masker (whether it is a pure sound or low-range noise). This dependence is shown in Fig. 1. Sounds fulfilling the above conditions may be cut and added without losing the quality of audio signal through non-simultaneous masking. This is used in compression algorithms (as it allows to reduce the amount of data in the sound track), as well as in steganography (where additional data is hidden by encoding with the use of added sounds). It should be noted that data may be removed by compression algorithms using the same dependencies (Matsuka, 2006; RLE, 1999).

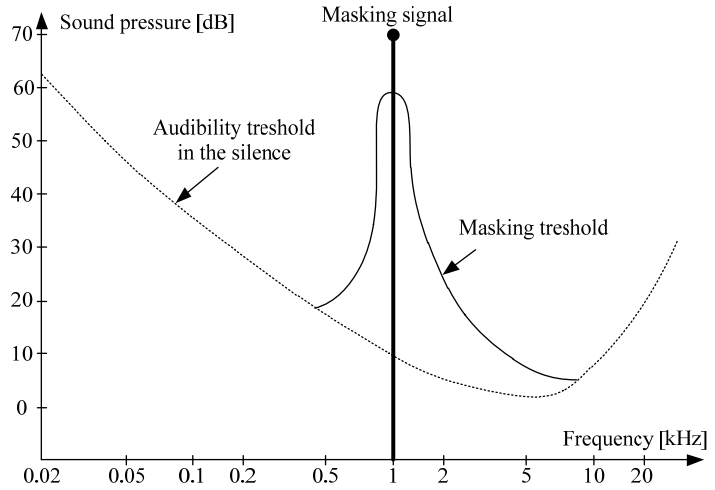

Figure 1. Simultaneous masking threshold for 1kHz sine masker signal while masking 'pure' sound (Nedeljko, 2004)

4. The new Fourier transform-based method

Fourier transform-based methods which hide information by changing frequency are successfully used in cases where an image is the data carrier. Due to high sensitivity of the human ear to frequency changes, the development of these methods, where sound is the data carrier was stagnant, as it was very hard to hide the presence of changes. Masked frequencies can be used for steganographic purposes to avoid these difficulties. This solution was applied in the previous version (referred to as MF_{v1} of the presented method, described in Kozieł (2009). This article presents the results of working over the refinements of the MF_{v1} method.

To start the hiding process, the chosen sound fragment is transformed using DFT. The sound amplitude spectrum is computed from the result. This spectrum is analysed in order to find the strip with the highest value (a salient point, Kim, 2005). This strip is marked $p_{\max}$, while its value is marked $W_{\max}$. This strip corresponds to frequency $f_{\max}$ having the highest ratio in the signal, so it can be treated as a masker signal. Two spectrum strips are chosen from the masked range. Those strips are marked p_1 and p_2, their values w_1 and w_2 their frequencies f_1 and f_2 respectively, and will be used for hiding a bit of information. The choice of strips depends on the steganographic key. For the strip to carry information it must fulfil specific conditions:

1. to be placed in the distance range (F_{dif1}, F_{dif2}) from $f_{\max}$,
2. to have a value not greater than determined by the dependence (3):

$$W_n \leq \left(a - \frac{f_{\max} - (f_n)^2}{b} \right) * W_{\max}. \tag{3}$$

Values F_{dif1}, F_{dif2}, a, b, come from the steganographic key, f_n is the frequency corresponding to the n-th strip.

Strips that fulfil the above conditions are placed in the table according to the sequence, determined by the key. Then, in sequence, they are checked to choose a pair which allows for hiding the determined binary value. The chosen pair is used for attaching the bit of information. In case that there are some neglected strips in front of the chosen one in the table, their values are modified in such a way that does not fulfil the dependence (3).

Next, the difference of values for strips p_1 and p_2 is calculated in order to determine whether it is as expected or any modification is needed. The expected value difference (R) is calculated on the basis of the steganographic key, which contains R_p value determining the ratio of R to the maximum value $W_{\max}$. Value of R is calculated according to equation (4):

$$R = W_{\max} * R_p, \tag{4}$$

This solution allows for adapting the power of modification to the signal strength and enables to use all signal blocks. Hiding of the bit b in the signal

means signal transformation in a way which fulfil dependence (5).

$$\begin{cases} |w_1 - w_2| \geq R, & \text{for } b = 1, \\ |w_1 - w_2| \leq \beta, & \text{for } b = 0, \end{cases} \tag{5}$$

where: β is the value in the key determining the maximum range of the random value added to the calculated strip value.

Once the above dependence is met, the fragment is transformed back into the time domain with the use of the inverse Fourier transform (IDFT) and placed in the signal instead of the original fragment.

The following algorithm describes how the bit of information $b = 1$ is being hidden in the signal:

1. DFT is used to transform the signal, resulting in obtaining the vector of complex values Y_c,
2. absolute value of vector $Y_r = |Y_c|$ is computed,
3. maximum value $W_{\max} = \max(Y_r)$ is determined in Y_r,
4. difference $R = W_{\max} * R_p$ is computed,
5. positions of $f_{\max}$ of the max value strip $p_{\max}$ and the strips meeting the conditions for carrying of hidden information are calculated,
6. strips p_1 and p_2, meant to carry the hidden data are chosen and their values w_1 and w_2 are determined,
7. on the basis of the key the maximum allowed value for the each of the two chosen strips is calculated: $w_{1allowed}$ and $w_{2allowed}$,
8. if $|w_1 - w_2| \geq R$ then we finish the algorithm (values of both strips are correct),
9. if $|w_1 - w_2| \leq R$ then we determine, which of the strips has the greater value and which one has the smaller value and we mark them, respectively w_w and w_m. After this, we calculate target values of the strips:
   a. if $w_m/\theta_{\max} + R \leq w_w$ then $w_m = w_w - R - rnd(\beta)$, ($rnd(\beta)$ is the function that returns the random value from the range $< -\beta, \beta >$, $\theta_{\max}$ is the maximum value allowed to be used to divide the strip value during its reduction),
   b. if $w_m/\theta_{\max} + R > w_w$ then $w_m = w_m/\theta_{\max}$, $w_w = w_m + R + rnd(\beta)$.
10. next, the Y_c vector is updated on the basis of calculated values (the phase is preserved as in the original signal),
11. the updated vector Y_c is transformed into a signal in the time form by IDFT.

When it is necessary to change the values of the strips, the strip of the lesser value is modified first. Consequently, the amplification of the second strip is reduced. Thus, the power of the second amplified frequency is reduced, which results in distortion that can be masked in an easier way. Due to the fact that the zero values almost do not appear in the signal spectrum, the author decided that the value of strip having smaller value reduction will be done by dividing it

by a value from the range $(1, \theta_{\max})$. $\theta_{\max}$ is a value defined in the steganographic key. It allows for avoiding introduction of zero value strip.

In order to obtain higher steganographic capacity of the signal, it should be divided into blocks. Successive bits of concealed information should be attached to those blocks. In the next step, the blocks should be combined.

There are often discontinuities at the block connections, which introduce interference in the form of cracks. Smoothing out of the discontinuities is necessary for eliminating the cracks. The author proposes to use connection blocks put between information-carrying blocks.

In the following algorithm, information-carrying blocks are marked as fr_1, fr_{r2}, connecting block as s, the number of signal samples in s block as k.

Blocks are placed as following: fr_1, s, fr_2.

The algorithm of the connection:

1. signal is divided into blocks in the following order: fr_1, s, fr_2, s, ...
2. information is concealed in blocks fr_1 and fr_2,
3. the difference r_1 between the last sample of block fr_1 and the corresponding sample of the original signal is computed,
4. the difference r_2 between the first sample of block fr_2 and the corresponding sample of the original signal is computed,
5. s block is copied to block s_1, r_1 is subtracted from all samples of s_1,
6. value of each sample in block s_1 is multiplied by $w = k/sampleindex$, indexes are the subsequent numbers starting from 0,
7. block s is copied to block s_2, r_2 is subtracted from all samples of s_2,
8. value of each sample is multiplied by its index,
9. values of samples of connection block s are computed according to formula:

$$s(i) = \frac{s_1(i) + s_2(i)}{k}, \quad i = 0, 1, 2, \ldots, k, \tag{6}$$

   where: i – sample number,
10. blocks fr_1, s, fr_2 are put into the output signal.

The use of connecting blocks allows for acquiring solid result signal. The cracks are completely removed. It is also possible to avoid cracks by shaping a signal inside the processed block in the way that allows for obtaining the continuous signal at the block connections as presented in Tachibana (2001).

In the *MF* method the original carrier is not needed for reading the concealed data. Only the knowledge of the steganographic key is essential. This key contains such data as:

1. placement and size of information-carrying blocks,
2. connecting blocks size,
3. the hiding strength (R_p), described as the biggest strip percentage,
4. a, b, F_{dif1}, F_{dif2} coefficients values,
5. the sequence for strip placement in the table, which schedules the sequence for choosing them as data carrying strips,

6. the β parameter describing the range $< -\beta, \beta >$, which is used to draw the value to add to the chosen strip value to avoid the constant difference between two processed strips,
7. the $\theta_{\max}$ coefficient value, corresponding to the maximum denominator value that can be used in reducing the strip value,
8. hidden data length.

In order to read the hidden information, the positions of the stegocontainers have to be determined. Each stegocontainer should be transformed using DFT. The positions of modified strips in the result spectrum should be determined. The difference R' of their values should be tested. The value of the maximum strip $W_{\max}$ should be determined as well. In the next step, the value of the hidden bit b can be read according to dependence (7):

$$\begin{cases} b = 1, & \text{when } R \geq (R_p - M) * W_{\max} \\ b = 0 & \text{when } R \leq M * W_{\max} \\ b = -1 & \text{when } M * W_{\max} < R < (R_p - M) * W_{\max} \end{cases}, \tag{7}$$

where M is a value defined in the key. $b = -1$ means that the obtained value is treated as undefined.

As multichannel recording is currently used most frequently, additional channels can be used to introduce error correction codes. Considering two-channel signal, the copy of the information can be put in the second channel in order to make the correct reading of the hidden data possible in case of damaging one of the copies. Multichannel recording gives more opportunities, as one error correction code can be put in one channel, while the remaining channels can be used to carry information.

5. Research results

In this section an evaluation is provided of the presented method. To show the advantages of the method, it was compared with the previous method based on Fourier transform (denoted TF) and MF_{v1}. The previous method based on Fourier transform hides data by changing values of the chosen spectrum strips. This method divides the signal into blocks. In each block, it modifies the same spectrum strips to hide one additional bit of data. In the study, performed for this article, two consecutive strips corresponding to frequencies over 300Hz were used. If the first strip value was bigger than the second strip value, the hidden bit value was *1*, in the inverse situation the hidden bit value was *0*. The value of the introduced change was the same in all fragments and equal to average change applied in the *MF* method. The used block size was 2000 signal samples.

The MF_{v1} method is the earlier version of the *MF*. It hides data in a very similar way. The differences are: hiding by changing strips values in a way to obtain the first strip having bigger value than the second one when hiding the binary *1*, or to obtain the first strip having smaller value than the second one

when hiding the binary *0*. This method does not use the masking effect, but only limits the introduced change values.

To conduct the tests, the sound signal with 44,1 kHz sampling frequency was used. In all methods blocks composed of 2000 signal samples were used, allowing for obtaining the steganographic capacity of 42 bps in each channel (44 bps in TF method). In MF_{v1} and MF methods the R_p value was set to 15%.

The efficient concealment of the fact that additional data has been attached is the most important feature in case of steganographic communication systems. Moreover, this data should be attached directly to the signal, so that it would be impossible to separate them. The proposed method fulfils these requirements as all of the modifications are done in the audible bandwidth. Fig. 2 shows the sample signal spectrum and changes made by the presented algorithm. Information dispersion in the audible band can be noticed. The algorithm does not introduce additional frequencies to the signal that might be easily detected by a stegoanalyst.

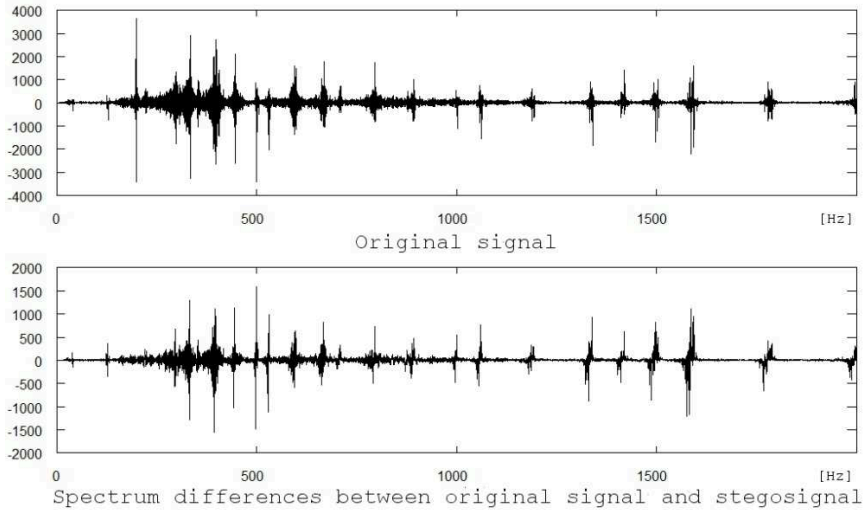

Figure 2. Spectrum differences introduced during data hiding and original signal spectrum

The following measures were used to determine the introduced interference level according to Rahman and Syed (2008):

1. mean square error (MSE),
2. normalized mean square error (NMSE),
3. signal to noise ratio (SNR),
4. peak signal to noise ratio (PSNR),
5. maximum difference between the original and modified signal (MD),

6. average absolute difference between signals (AD),
7. normalized average absolute difference between signals (NAD),
8. watermark transparency (AF).

The results obtained are presented in Table 1.

Table 1. The distortion level introduced during data hiding in the sound.

R_p	MSE	NMSE	SNR[dB]	MD	AD	NAD	AF
1%	7E-6	13E-4	28.95	0.06	10E-4	0.02	1
10%	9E-6	17E-4	27.6	0.06	16E-4	0.03	1
15%	14E-6	27E-4	25.66	0.06	20E-4	0.04	1
20%	23E-6	43E-4	23.62	0.06	28E-4	0.06	1
30%	56E-6	1E-3	19.77	0.06	43E-4	0.09	0.99
40%	1E-4	3E-3	16.98	0.08	58E-4	0.12	0.98

Results in Table 1 show that the proposed method introduces a very low level of signal distortion. The high SNR value confirms this. This value meets very strict demands put on watermarks, where the SNR value has to be bigger than 22dB. Additionally, the watermark transparency measure assumes the biggest possible value within a wide range.

To verify the real audibility of the introduced interference the double blind listening test was conducted. Clips were presented to listeners in pairs containing a modified one and a non-modified one. Listeners were asked to evaluate the difference between the two presented recordings. The difference grade could be given in the 5 step scale:

1. no differences between recordings,
2. not sure if differences exist,
3. very weak,
4. weak,
5. clearly noticeable.

Tests were conducted in a group of 30 people. Each tester rated records on his own using Behringer HPS3000 headphones having the bandwidth of 20Hz - 20000Hz and the efficiency equal to 110dB for the 1kHz frequency. The headphones were connected to a 16 bit sound card.

The listening test was run in order to determine the minimum length of connecting block, which was essential for removal of the interferences emerging on the connections of the information-carrying blocks. The test results are shown in Table 2.

Additionally, robustness tests were run in order to determine whether attached data is immune to damage during popular audio transformations. Stereo recording in "wav" format with the sampling frequency of 44100 Hz and the resolution of 16 bits per sample was tested. During processing, the $R_p = 20\%$ was

Table 2. Influence of length of connecting blocks on interference emerging on connections of blocks

Kind of sound and hiding power	Connection block length (number of samples)	Noise strength
Pop, R_p=40%	4	Weak
Pop, R_p=40%	10	Very weak
Pop, R_p=40%	over 40	Imperceptible
Ballad, R_p=10%	4	Weak
Ballad, R_p=10%	10	Very weak
Ballad, R_p=10%	20	Very weak
Ballad, R_p=10%	over 40	Imperceptible
Speech, R_p=50%	4	Very weak
Speech, R_p=50%	over 10	Imperceptible
Piano, R_p=50%	2	Very weak
Piano, R_p=50%	over 4	Imperceptible
Rock, R_p=50%	4	Weak
Rock, R_p=50%	20	Very weak
Rock, R_p=50%	over 40	Imperceptible

Table 3. Robustness of attached data to damage during popular audio carrier transformations.

	The metod proposed			MF_{v1}
Operation or used sound format	Correctly read bits [%]	Incorrectly read bits [%]	Unsteadily read bits [%]	Correctly read bits [%]
ogg	98.6	0.7	0.7	96.4
mp3 64kbit/s	97.9	1.4	0.7	80.1
mp3 128kbit/s	98.6	0.7	0.7	86.0
aac 64kbit/s	96.4	2.9	0.7	84.3
aac 128kbit/s	96.4	1.4	2.1	90.0
wma quality=98	100.0	0.0	0.0	99.1
wma quality=50	97.1	2.9	0.0	91.4
SFR to 22kHz	98.6	0.7	0.7	100.0
SFR o 11kHz	98.1	1.1	0.8	100.0
SRR to 8 bits	98.6	0.7	0.7	97.6
BF 1Hz-10kHz	100.0	0.0	0.0	98.2

used. Signal with attached data was subject to those transformations and it was reconverted to the initial form. The results are presented in Table 3. The last column of this table contains results obtained by the MF_{v1} method. Abbreviations used in the table are: "SFR" – sampling frequency reduction, "SRR" – signal sample resolution reduction, "BF" – bandpass filtering.

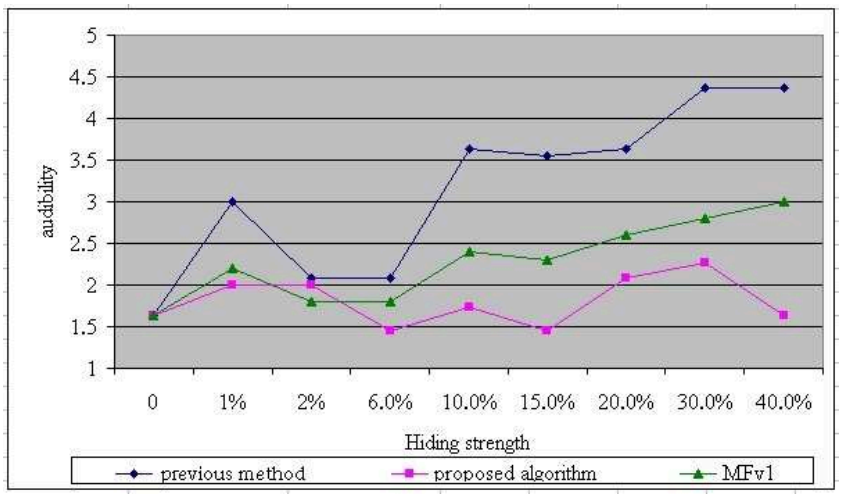

Figure 3. Comparison of interference levels introduced by different methods

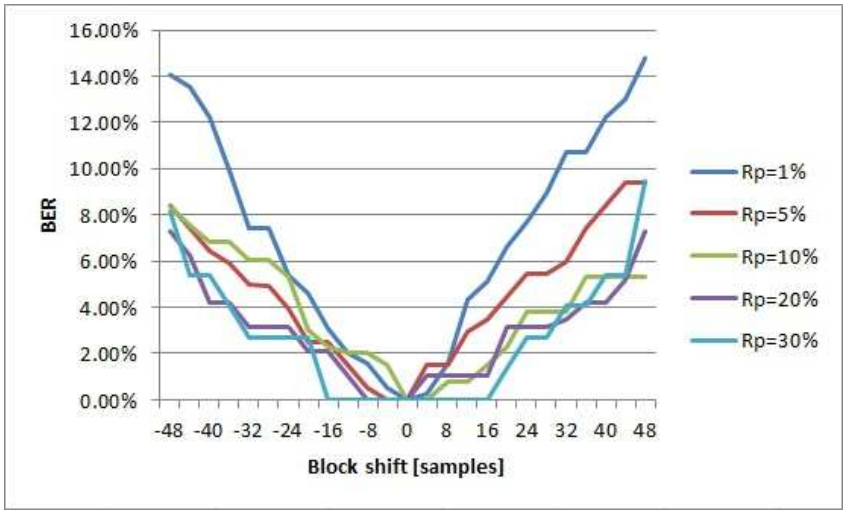

Figure 4. Influence of the length of block shift

In comparison to the MF_{v1} version a higher transparency level was achieved in the here proposed method. The results of the listening tests, conducted to compare interference audibility, inserted by the new method, the MF_{v1} and previous Fourier method, are shown in Fig. 3. In this figure the used value of parameter R_p was shown on the x-axis. The $R_p = 0\%$ value means that the two identical copies of original recording were compared. Original signal evaluation was placed in the test to obtain the reference level to compare with grades achieved by the modified fragments.

Of course, some compression methods introduce a shift in the signal because of removing silence or adding some samples at the beginning. The method proposed is successful even with a small shift, if a big enough R_p value is used. The *MF* sensitivity for the incorrect block placing is presented in Fig. 4. If the method has to be robust for the shift bigger than a few samples it is necessary to add the marker at the beginning of the hidden data. The marker is a previously established sequence of bits. At the beginning of the read process the initial part of the signal is read by using different block shifts. In the end, the results obtained from different reads are examined to find the initial sequence and determine the shift introduced by other conversions.

6. Summary

The here presented new method can be successfully used for steganographic information concealment. It allows for attaching of additional data directly to the signal. It does not introduce noticeable changes to the container signal in the examined range of parameter R_p. Using the audible bandwidth guarantees that additional information will not be removed accidentally and it allows for introducing changes in a way that is unnoticeable and hard to detect even by a stegoanalyst. The use of the audible frequency range prevents a simple removal of attached data by filtering one component frequency as in the previous Fourier method. The algorithm also allows for gaining steganographic capacity of 84 bits per second of stereo recording, which meets the requirements of steganographic methods.

Because the key lacks the definite values, as it contains only a method of calculating them on basis of the processed fragment signal parameters, it is possible to obtain high steganographic security level. Using Fourier transform and a wide range of frequencies, among which information is dispersed, allowed for obtaining high robustness to hidden data damage.

The use of masking, minimization of the introduced changes possible due to the use of only the value difference between strips, adapting the hidden data attaching strength to the signal strength and reduction of modifications introduced in the strip having bigger value, made it possible to achieve very good results in the transparency of the introduced changes. This was verified by numerical measures and listening tests, which showed that the modified signals could be assessed as similar to the original one. The comparison of the proposed

method with other methods based on the Fourier transform shows significant differences. This is especially evident in the audibility of the introduced changes.

The here presented method, due to its low computational complexity and the fact that only a fraction of a second of the signal is used to hide information, can be used in real-time systems.

References

Agaian, S., Akopian, D., Caglayan, O. and D'Souza, S. (2005) Lossless adaptive digital audio steganography. *Proc. IEEE Int. Conf. Signals, Systems and Computers*. IEEE, 903-906.

Bao, P. and Ma, X. (2004) MP3-resistant music steganography based on dynamic range transform. *IEEE Int. Sym. Intelligent Signal Processing and Communication Systems*, 266-271.

Bender, W., Gruhl, D., Morimoto, N. and Lu, A. (1996) Techniques for data hiding. *IBM Systems Journal*, **35**, 3&4, 313-336.

Cvejic, N. and Seppanen, T. (2002a) A wavelet domain LSB insertion algorithm for high capacity audio steganography. *Proc. IEEE Digital Signal Processing Workshop*. IEEE, 53-55.

Cvejic, N. and Seppanen, T. (2002b) Increasing the capacity of LSB-based audio steganography. *IEEE Workshop on Multimedia Signal Processing*. IEEE, 336-338.

Cvejic, N. and Seppanen, T. (2004) Increasing robustness of LSB audio steganography using a novel embedding method. *Proc. IEEE Int. Conf. Info. Tech. Coding and Computing*, 2. IEEE, 533-537.

Czyżewski, A. (2001) *Dźwięk cyfrowy* (*Digital sound*; in Polish). Exit, Warszawa.

Delforouzi, A. and Pooyan, M. (2008) Adaptive Digital Audio Steganography Based on Integer Wavelet Transform. *Circuits Syst Signal Process*, **27**, 247-259.

Dymarski, P. (2006) Filtracja sygnałów dźwiękowych jako metoda znakowania wodnego i steganografii (Filtration of sound signals as a metod of watermarking and steganography; in Polish). Krajowe Sympozjum Telekomunikacji '06, Bydgoszcz. Akademia Techniczno-Rolnicza w Bydgoszczy 12-23.

Dymarski, P., Pobłocki, A., Baras, C. and Moreau, N. (2003) Algorytmy znakowania wodnego sygnałów dźwiękowych (Algorithms of watermarking of sound signals; in Polish). *Krajowe Sympozjum Telekomunikacji '03*, Bydgoszcz. Akademia Techniczno-Rolnicza w Bydgoszczy, 26-34.

Garay, A. (2002) Measuring and evaluating digital watermarks in audio files. A Thesis submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University, Washington.

Gopalan, K. (2003) Audio steganography using bit modification. *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 2. IEEE, 421-424.

GOPALAN, K. (2005) Audio steganography by cepstrum modification. *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, **5**, 481-484.

GRUHL, D. and LU, A. (1996) Echo Hiding. *Information Hiding Workshop*, Cambridge University, U.K. Cambridge University Press, 295-315.

IZYDORCZYK, J., PŁONKA, G. and TYMA, G. (2006) *Teoria sygnałów* (*Theory of signals*; in Polish). Helion, Gliwice.

JOHNSON, N. and KATZENBEISSER, S. (2000) *A survey of steganographic techniques, Information hiding: Techniques for steganography and digital watermarking.* Artech House, London

JORASZ, U. (1999) *Selektywność ludzkiego słuchu* (*Selectivity of human hearing*; in Polish). Wydawnictwo Naukowe UAM, Poznań.

KATZENBEISSER, S. and PETRICOLAS, A. (2000) *Information Hiding.* Artech House, London.

KIM, H. (2005) Audio watermarking techniques, *Proc. Pacific Rim. Workshop Digital Steganography.* Kyushu Institude of. Technology, 1-17.

KIM, S., KWON, H. and BAE, K. (2004) Modification of polar echo kernel for performance improvement of audio watermarking. *International Workshop on Digital Watermarking No.* 2, **LNCS 2939**. Springer, 456-466.

KOZIEŁ, G. (2009) Method of concealing information in sound based on Fourier transform and Masking. *Polish Journal of Environmental Studies*, 18, 3B.

MANSOUR, M. and TEWFIK, A. (2001) Audio Watermarking by Time-Scale Modification. *IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings* 3. IEEE, 1353-1356.

MATSUKA, H. (2006) Spread spectrum audio steganography using subband phase shifting. *IEEE Int. Conf. Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP'06).* IEEE, 3-6.

NEDELJKO, C. (2004) *Algorithms for audio watermarking and steganography.* Oulu University Press, Oulu.

RAHMAN, S. and SYED, M.(2008) *Multimedia Technologies: Concepts, Methodologies, Tools, and Applications.* Information Science Reference, London.

RLE (1999) http://rleweb.mit.edu/Publications/currents/cur111/111watermark.htm, Massachusetts.

SANTOSA, R. and BAO, P. (2005) Audio-to-image wavelet transform based audio steganography. *IEEE Inernational Symosium.* IEEE, 209-212.

TACHIBANA, R., SHIMIZU, S., NAKAMURA, T. and KOBAYASHI, S. (2001) An audio watermarking method robust against time and frequency fluctuation. *Proc. of SPIE Int. Conf. on Security and Watermarking of Multimedia Contents III,* Security Professionals Information Exchange, **4314**, 104-115.

WAYNER, P. (2002) *Disappearing Cryptography.* Morgan Kaufmann, Massachusetts.

XIANG, S., HUANG, J. and YANG, R. (2007) Time-Scale Invariant Audio Watermarking Based on the Statistical Features in Time Domain. *Artificial*

Intelligence and Lecture Notes in Bioinformatics, Springer, 93-108.

Xiang, S., Kim, H., Huang, J. (2008) Audio watermarking robust against time scale modification and MP3 compression. *Signal Processing*, **88**,10, 2372-2387.

Zieliński, T. (2005) *Przetwarzanie sygnałów cyfrowych* (*Processing of digital signals*; in Polish). Wydawnictwa Komunikacji i Łączności, Łódź.