



HAL
open science

An Informed Synchronization Scheme for Audio Data Hiding

A. Loboguerrero, Patrick Bas, J. Lienard

► **To cite this version:**

A. Loboguerrero, Patrick Bas, J. Lienard. An Informed Synchronization Scheme for Audio Data Hiding. Proc of SPIE Electronic Imaging, Security and Watermarking of Multimedia Content VI, Jan 2004, San-Jose, United States. pp.116-126. hal-00166580

HAL Id: hal-00166580

<https://hal.science/hal-00166580>

Submitted on 7 Aug 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Informed Synchronization Scheme for Audio Data Hiding

LoboGuerrero A., Bas, P. and Liénard J.*

Laboratoire d'Images et Signaux, 961 rue de la Houille Blanche
Domaine Universitaire B.P. 46 38402 Saint Martin d'Herès cedex France

ABSTRACT

This paper deals with the problem of synchronization in the particular case of audio data hiding. In this kind of application the goal is to increase the information of an audio data set by inserting an imperceptible message. An innovating synchronization scheme that uses informed coding theory is proposed. The goal is to realize a complementary approach from two different techniques in order to obtain an enhanced synchronization system. To that end, the analysis of the classical spread spectrum synchronization is done and this classical scheme is improved by the use of side information. Informed coding theory is presented and revisited taking into account the problem of synchronization to enable the selection of signal realizations called Feature Time Points (FTP) which are correlated with a code. Such considerations yield to the definition of informed synchronization. The proposed scheme and the definition of FTP are after presented taking into account the robustness criterion. Finally, results and comparison with classical spread spectrum synchronization schemes are presented.

Keywords: Audio Data Hiding, Informed Synchronization, Spread Spectrum

1. INTRODUCTION

Audio data hiding^{1,2} is a method that allows the insertion of an imperceptible mark in a signal data set. Although this technique is mainly used to guarantee intellectual property, applications in different areas are constantly been developed. This is the case of data hiding for data transmission. Several applications are possible in this particular context. For example, if the priority information which is transmitted is the speech voice, then it may be useful to transmit video analysis information such as lips moving parameters, the speaker's face or any other information identifying the speaking face. Such an application is addressed in the project ARTUS^N which aims to embed animation parameters into audio and video contents. When implementing audio data hiding schemes the designer has to take into account not only coding but also synchronizations constraints. These constraints are:

- Minimum acquisition delay and complexity: The procession time devoted for the synchronization task is wasted and can not be used for decoding the hidden message;
- Temporal derive management: due to D/A A/D conversions, the global system has to cope with re-sampling and temporal derive;
- Robustness to noise and compression: those processes are part of the transmission channel;
- Minimal audible distortion: the hidden information has to remain unnoticeable;
- Maximal transmission rate of the inserted message: the capacity of the scheme has to be as important as possible.

Watermarking synchronization for audio is classically achieved using spread spectrum (S.S.) synchronization theory³ or audio characteristics as a beat⁴. In this article the S.S. approach is completed by informed embedding analysis yielding

* Further author information : (Send correspondence to Alejandro Lobo Guerrero)

A.L. : E-mail: alobogue@lis.inpg.fr, Telephone : 00.33. 4.76.82.62.58

P.B.: E-mail: Patrick.Bas@lis.inpg.fr, Telephone : 00.33. 4.76.82.64.21

J.L. : E-mail: Joel. Lienard@lis.inpg.fr Telephone : 00.33. 4.76.82.68

^N This project was funded by French RNRT Artus Project

an informed synchronization scheme that introduces an improved detection criteria and aims to maximize robustness while minimizing acquisition delays.

This paper is organized as follows:

At first, a brief introduction of the existing audio data hiding synchronization techniques is done. Special interest is focused on spread spectrum synchronization techniques. The related problems of the latter are outlined and the possible improvements are derived. In the next section, a description of the informed data hiding concepts is presented. Then, the proposed scheme based on the precedents presentations is introduced. The last section outlines the improvement by comparing the synchronization behaviors of the presented scheme.

2. PREVIOUS WORKS

Few authors have proposed solutions to the audio data hiding problem. In this section, the main synchronization schemes that have been developed for audio data-hiding are exposed. They are based on correlation techniques or signal analysis.

- Kirovsky and Malvar have proposed an audio watermarking scheme that uses audio analysis for extracting instants that enable synchronization and avoid security issues inherent with classical spread spectrum synchronization techniques⁴. The synchronization scheme is based on beat detection of musical contents that lead to a natural cutting of temporal axis. The proposed algorithm first estimates the mean length of the beat and secondly estimates the exact occurrence of each beat. The audio signal is then after time-warped in such a way that the period between each beat remains constant and equal to the length of message code. The signal is after watermarked using a classical spread spectrum scheme and an inverse warping is also performed. Such a scheme is robust to time-warping attack and time translation but is not suitable for signals that did not contains beats such as speech signals.

- Haistma et al. have presented an audio watermarking scheme that uses a modified correlation function to decode the mark and enable synchronization⁵. The matching function, called SPOMF for Symmetrical Phase only Matched Filtering, is performed in the Fourier domain and the translation recovery is done calculating:

$$Corr_p(S_1, S_2) = TF^{-1} \left[\frac{TF(S_1) \overline{TF(S_2)}}{|TF(S_1) TF(S_2)|} \right]$$

The use of phase information improve the reliability of the detection of the synchronization peak (if $S_1(t) = S_2(t-\tau)$) the $Corr_p(S_1, S_2) = \delta(t)$. The behavior if this system will be studied and compared with the proposed scheme in the section devoted to results.

The following sections present the main synchronization scheme that is used in audio data-hiding and the concepts of the informed theory. This will establish the context to introduce the informed synchronization scheme. The related advantages and disadvantages are outlined.

3. FROM SPREAD SPECTRUM SYNCHRONIZATION TO INFORMED SYNCHRONIZATION

The present work is related to two different approaches which are presented and the advantages and disadvantages are outlined. The goal is to develop a scheme that aims to profit from the best characteristics reducing the inconvenients of each one. Theses approaches are the Spread Spectrum and Informed approaches.

3.1. Classical Spread Spectrum Approach

One of the most popular synchronization system in the field of telecommunication is based on the classical Spread Spectrum theory³. This system uses a specific code (a pseudo random signal) for guaranteeing synchronization. This code is added to the host signal in the transmission stage. The synchronization at the receiver may be implemented by using an sliding correlator matched with the received code. Then, the resulting signal is

$$y(n) = \{s(n) + Ac(n)\} \otimes c^*(n-m) \quad (1)$$

Where $s(n)$ represents the host audio signal, $c(n)$ the inserted code which power is determined by the coefficient A . $C(n)$ is a white code and by consequence its autocorrelation is ideally an unit impulse. The idea is then to detect the presence of this peak by computing the correlation between the received code and a local copy. The correlator output is then compared with a threshold to determine the phase of the received synchronization code. Note that when using this technique in audio data-hiding applications the host signal $s(n)$ is treated as a “noise” and consequently the presence of $s(n)$ perturbs the detection of the peak (figure No. 1).

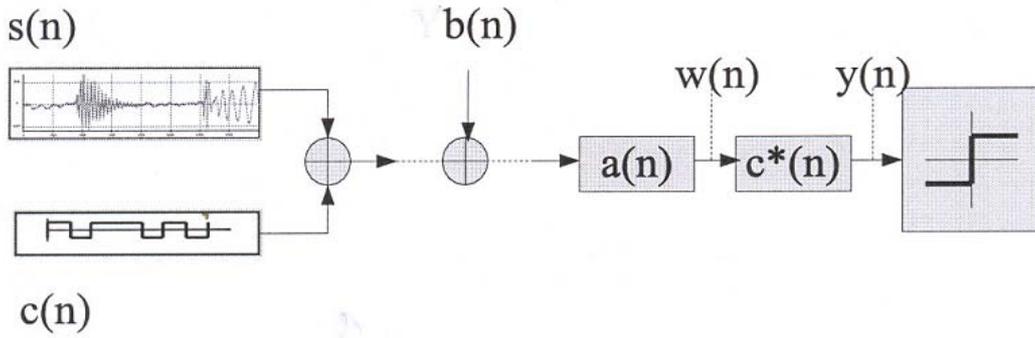


Figure 1. Classical Spread Spectrum Synchronization

When the SNR is not large enough (e.g. the power of the code is too weak), the localization of the peak may fail because the intercorrelation between the code and the host signal is too important. In this case, before evaluating the threshold, a supplementary filter can be applied to reduce the noise perturbation. One classical solution to solve this drawback is to combine the use of a periodic code and an average filter. This average filter $a(n)$ is expressed by the next equation:

$$w(n) = \frac{1}{K} \sum_{k=0}^{K-1} r_k(n) \quad (2)$$

In this filter, the received signal $r(n)$ (host signal, codes and channel noise) has been divide in blocks noted with the sub index k . The number of averaged blocks is determined by the constant K . This filter represents an average system that aims to reduce the noise effect (this is composed by $s(n)$ plus channel noise). The signal $y(n)$ is then expressed as:

$$y(n) = w(n) \otimes c^*(n-M) \quad (3)$$

Replacing $w(n)$ as in equation 2, $r_k(n)$ by $s(n)+n(n)$ and regrouping, it follows:

$$y(n) = \frac{A}{K} \left\{ \sum_{k=0}^{K-1} c_k(n) \otimes c^*(n-M) \right\} + \frac{A}{K} \left\{ \sum [s_k(n) + n_k(n)] \otimes c^*(n-M) \right\} \quad (4)$$

Ideally, when K tends towards infinity the effect of the noise perturbing signals $s(n)$ and $n(n)$ will be reduced and the codes autocorrelation will be preserved.

Such a synchronization process leads to several disadvantages when applied to audio data hiding. A spread spectrum based synchronization is defined by 3 different parameters that will enable to improve the detection step:

1. To increase the power of the inserted codes (A in equation 4);
2. To increase the length codes (M in equation 1);
3. To increase K (in equation 4).

Nevertheless, each one of these strategies presents important drawbacks. Strategy 1 is often discarded because of the high probability of adding audible distortion. In fact, the constraint of inaudibility imposes a very weak SNR. The second and third options imply main disadvantages:

- Important acquisition delays;
- Unsuitable in presence of temporal clock derive.

We propose in this paper to use informed coding theory to improve the above issues. In the following, the principles of this technique are introduced and the new informed synchronization scheme is proposed.

3.2. Informed Synchronization Approach

In this section, the main concepts of the informed theory are introduced including definition, robust synchronization and a more general presentation of informed synchronization.

3.2.1. Definition

General informed coding theory⁶ proposes to use the knowledge of the host signal in the transmission stage (so it is no longer treated as a random signal but as a deterministic one) to select optimal codes. Basically, using side-information for coding enables a reduction of the embedding power while it is possible to achieve the same performance in term of robustness. In this kind of analysis, the host signal and the code are often represented as vectors. Informed coding is then equivalent to select a codeword from a codebook representing the code-word that is the nearest⁺ to the host vector. Such techniques first presented by Costa⁷, and used in the field of image watermarking by several authors^{8,9}, lead to an increasing of the capacity while preserving the robustness

Informed synchronization can be seen as an analogous strategy where the goal is to select *signal vectors that are the nearest to one codeword*. The selected host vectors denoted in this paper by Feature Time Points (FTP's) will be used for synchronization purposes (see section 4). Thus, Informed synchronization can be seen as an extension of synchronization techniques which relies on the content of the host signal to provide invariant feature points or areas. Moreover, as it is addressed further in the paper, the use of informed theory enable to improve the robustness of the scheme to synchronization processes compared with classical S.S. techniques.

Figure 2 presents the geometrical interpretation of informed synchronization and is compared with a classical S.S. synchronization scheme. The vertical axis represents the linear correlation between the host vector $s(t)$ and the codeword c . The horizontal axis represents the temporal axis, it is equivalent in the informed coding paradigm with the index of the codebook. In this correlation/temporal plan, different realizations of the host signal are represented by crosses. In (a), the arrows represent the effect in a classical S.S. scheme of a constant embedding power. In this case, the code word is added on each host vector to create a synchronized signal and the coding of the synchronization information does not depend on the host signal. Ideally, in absence of the host signal, the effect of the added codes yields a correct detection where all synchronization points are over a predefined threshold (denoted with a gray region). When the host signal is present, the resulting marked signals may all lead to present a significant correlation[±]: this is why classical S.S. synchronization averages vectors before computing linear correlation.

⁺ The distance is defined by a correlation measure

[±] Correlation that is above a threshold or maximum of the correlation function

In (b), the figure illustrates the effect of informed synchronization which can also be seen as feature detection: only the most highly correlated vectors are kept for synchronization purposes (noted with dark crosses). They represent singular points such as local maximum of the correlation considering a temporal neighboring window $V_i(t)$ (cf. $S(t_1)$ and $S(t_2)$). It is important to note in this case that the synchronization can be achieved without altering the original signal (no distortion introduced) following the condition that the robustness is not guaranteed.

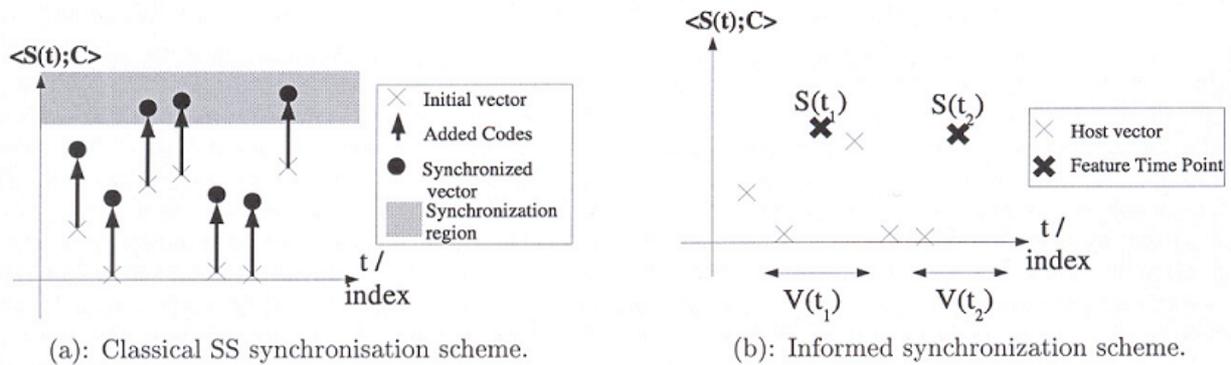


Figure 2. Geometrical comparison of classical and informed synchronization schemes.

The extraction of FTP's will be described further in this paper: the main idea is to implement an invariant time filtering system. Message coding is achieved by adding the message codes in the instants corresponding to the Feature Time Points. In reception stage, the message will be decoded accordingly with the instants determined by a local copy of the feature time detector.

3.2.2. Controlled Robustness Informed Synchronization

Because considering noise in the model enables to deal with robustness, this perturbation has been taken into account to lead to a Controlled Robustness Synchronization Scheme. The principle of controlled robustness is depicted on figure 3.: The chosen maximums of the correlation (FTP's) can be modified to bring these points farther from the other points in the window neighborhood $V_i(t)$ (cf. $S(t_1)$ and $S(t_2)$). This operation will increase the probability that a local maxim is preserved in noise presence. Controlled Robust Synchronization can be also used when there is not FTP's detected over the host signal (e.g. the host signal power is very weak). In this case, one point (vector) of the host signal can be modified in such a way that the correlation is reinforced and it becomes an FTP's. It means that a FTP is forced (cd $S(t_3)$). From a vector point of view, this means that a chosen vector is modified by an addition of a from a code vector to transform it into a FTP. The robustness is represented by the distance from the FTP to the nearest vector and is function of the probability of missing or creating a new FTP.

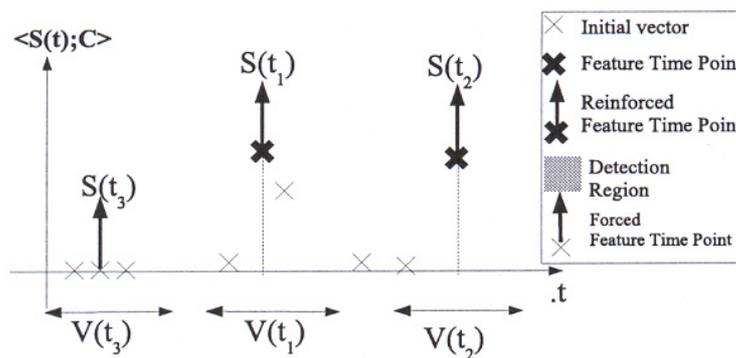


Figure 3. Controlled Robustness Informed Synchronization scheme.

3.2.3. General Informed Synchronization

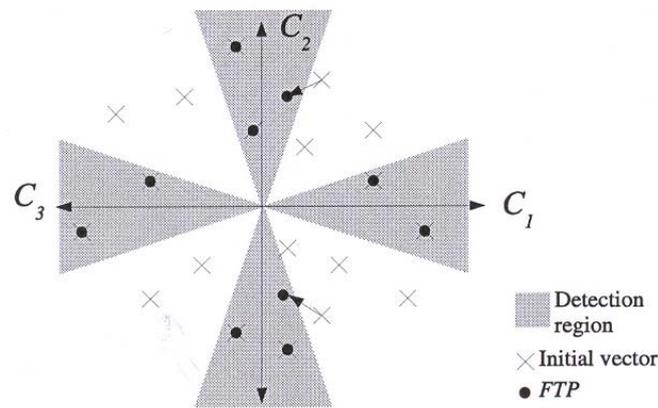


Figure 4. General informed synchronization scheme.

In the previous section the informed synchronization has been presented from a vector point of view where the selection of the closest pair (host vector, codeword) is done only considering a set of different host vectors. Nevertheless, the theory of informed coding can be extended by using a codebook, of several code words that will permit to reduce the distance between a host vector and a code word. In this context, the synchronization has to deal with two different codebooks. The first is provided by the host signal, the second is a predefined codebook. Figure 4 illustrates the representation of a general informed synchronization scheme. In this case, four code words have been used and a normalized correlation is computed to lead to a detection region representing a cone¹¹. Note that if on one hand, an important codebook will lead to a statistical good matching between the code word and the host signal, on the other hand, the complexity of such a scheme will be multiplied by the number of code words present in the codebook[§].

4. PROPOSED SYNCHRONIZATION SYSTEM

Once the classic synchronization systems and the basis of informed embedding have been introduced, an implementation of such a system is proposed. This synchronization system has been developed taking into account the main advantages of the S.S. and informed systems and looking for minimizing the existing disadvantages. Such a scheme must fulfill few requirements:

- Informed System: $s(n)$ is not analyzed as a noise;
- Detection of FTP's should be a simple process;
- FTP's must avoid time overlapping of consecutive message codes: A minimum time separation is required;
- The system must not include important average delays;
- The system has to integrate a noise robustness control system.
- The acoustic distortion must be taken into account.

The figure 5 shows the proposed informed synchronization scheme.

This system is composed by three stages. In the transmission stage, a feature instants detector analyses the host signal $s(n)$ in order to find and reinforce FTP's. How the FTP's are defined and reinforced is presented in section 4.1.

[§] This is present just as a theoretical approach. In the present work, only one code has been used. Future developments will be focused to the usage of a codebook.

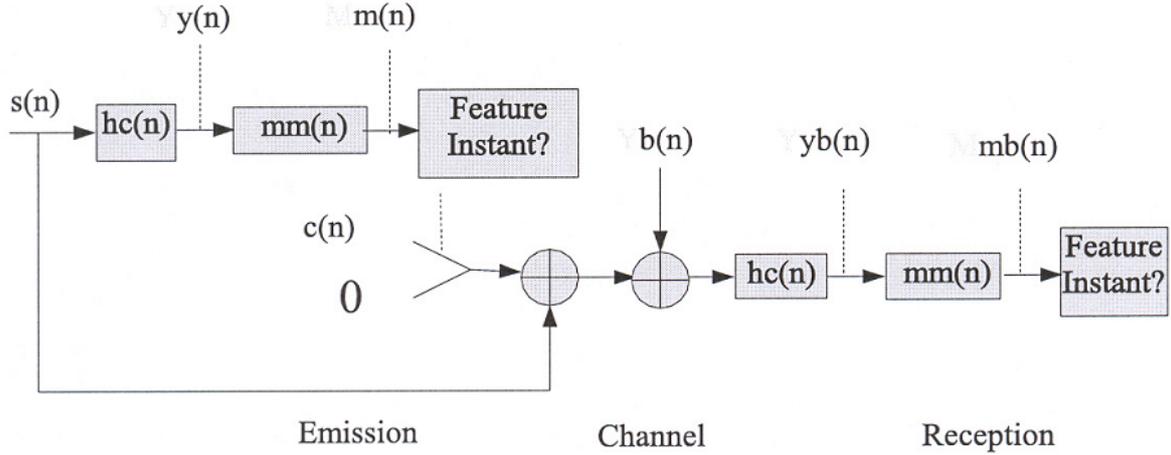


Figure 5. Proposed Informed Synchronization scheme.

The channel noise can be modeled from two points of view:

- Additive White Gaussian Noise (AWGN): A A/D D/A conversion can be done in order to transmit the signal through an analog channel. In this case, AWGN will model the introduced perturbation;
- Compression with lost: the most usual compression format is MP3. In this case the channel perturbation can be studied as a low pass filter followed by an additive compression noise (caused by a re-quantification system).

The detection stage, is analog to the transmission stage. The receptor has to apply the feature detector over the received signal and decode the message. For taking into account the eventual presence of noise, the decision criteria has to be defined. In the following sections a detailed description of the reception stage and detection criteria will be presented.

4.1. Emission Stage

4.1.1. Feature Time Points Definition

Following the ideas exposed in 3.2.1., it is possible to define FTP's and to implement the corresponding detector. FTP's have to match several conditions. They have to be easy to detect, robust to the channel and compression noise and two successive FTP's have to be separated (in the time axis) by a minimal distance. Two different filters have been chosen to derminate FTP's and to fulfill these conditions. They are $hc(n)$ and $mm(n)$.

The first one of the two filters ($hc(n)$) must be chosen agreeing the following criterions:

- To bring out some characteristic of the signal;
- To be robust to channel and compression noise.

As a synchronization charateristic, it could be natural to consider the maximums over the host signal. In analogous way, the maximums over a pass band or low pass filtered version of $s(n)$ could be used as a synchronization criteria (future works will be developed in this line). However, the most of them lack of a suitable noise behavior. In order to have a filter which behavior allows a gain in the SNR, the S.S. theory provides the necessary tools. In fact, the use of a filter matched to a pseudo random code will optimize the SNR and will yield a robust behavior face to the channel noise. Then, the filter $hc(n)$ is defined as

$$h_c(n) = c^*(n - M) \quad (5)$$

Where $c(n-M)$ is the pseudo random code with length M . If the maximums of the filtered signal are taken to realize synchronization, an enhanced behavior concerning the noise will e provided by the system.

Nevertheless, the criteria has to match a supplementary condition. Since each FTP in $s(t)$ will denote the origin of a message code and since overlapping of those message codes is not suitable, the FTP's must have a minimal distance separation equal to the codes length (this is also important when improving robustness by synchronism codes addition as is explained in deep in section 4.1.2.). Although choosing the maximums of $s(n)$ filtered by $h_c(n)$ makes the system robust, this choice does not answer to the minimum distance exigencies. This is why a second filter ($mm(n)$) is proposed. The idea is to analyze the signal over limited duration segments (equal to the code length) and to define the local maximum of each segment as a FTP. This choice will guarantee the minimal required separation. This yields to a morpho-mathematical filter defined by the following equation:

$$m(n) = nm[y(n)] = \max[y(n), y(n-1), y(n-2), \dots, y(n-(N-1))] \quad (6)$$

This filter determines the local maximum over a N samples window in signal $y(n)$. The window size N is chosen to be equal to the length code M and thereby to avoid overlapping between codes. The resulting signal $m(n)$ is a stair-shape signal. The maximums of the doubled filtered signal $m(n)$ are useful for synchronization proposals and can be visually noted as "flat" segments. However, these flat segments are spread over several samples whereas the synchronization instants must be defined over just one sample. In order to define the synchronism instant on just one sample, a particular instant N_0 must be associated to this flat segment of signal $m(n)$. N_0 is therefore defined as the first sample in the flat segment. Figure 6 depicts the process. In (a), the intercorrelation between a realization of $s(n)$ and $c(n-M)$ is illustrated. The circles show how when using the maximums of this signal as synchronization points often generates points that are too close in time. In (b), the local maximum has been taken and the stair-shape signal is visualized. In this case, the circles denote two flat segments corresponding to the local maximums of this signal that are easily detectable. The vertical line shows how these segments are associated to singular sample points defining in this way N_0 . Note that two consecutive instants N_0 are always separated by a distance at least greater than M allowing the insertion of consecutive codes without generating overlapping between them.

This instants (N_0) are determined by the application of a suite of time invariant causal filters ($hc(n)$ and $mm(n)$). This characteristic allows to relate N_0 to an instant over $s(n)$. In fact, if the delay produced by $hc(n)$ is N_c and N_m that one introduced by $mm(n)$, the instant associated to N_0 (called Feature Time Point) over $s(n)$ is:

$$FTP = N_0 - (N_c + N_m) \quad (7)$$

It is important to note that the chosen filter $h_c(n)$ is exactly the same as used in S.S. systems. This brings a SNR gain face to the channel noise. However, $s(n)$ is not treated in the same way then the channel noise. To use $s(n)$ as a known signal has permitted the definition of the FTP's as the maximum of the correlation between $s(n)$ and $c(n)$. From an informed point of view, this filtering system allows the selection of the signal vectors having a maximum of similarity compared with the pseudo random code in a determined period of time. In the next section, the way to increase or control the system robustness by codes addition is explained.

4.1.2. Controlled Robustness Informed Synchronization

The filter $hc(n)$ has the suitable properties face to an additive noise. However,, the interest is focused in controlling the robustness of the system (as enunciated in section 3.2.2.). Once the analyze of the host signal is done and the feature instants are determined, the idea is to reinforce the maximum found in equation 6 by adding a synchronism code equal to the filter impulse response of $hc(n)$ (this is illustrated in figure 5 by a switching system controlled by the FTP's detector). Such an addition yields to interesting perceptual properties because a code will be added in a zone of the host signal that present high correlations and consequently are similar. The synchronism code position is chosen to respect the causality constraints of the system. If the origin of the code is place in the instant FTP the impulsion yield by the correlation will be placed in N_0 reinforcing the maximum found in equation 6.

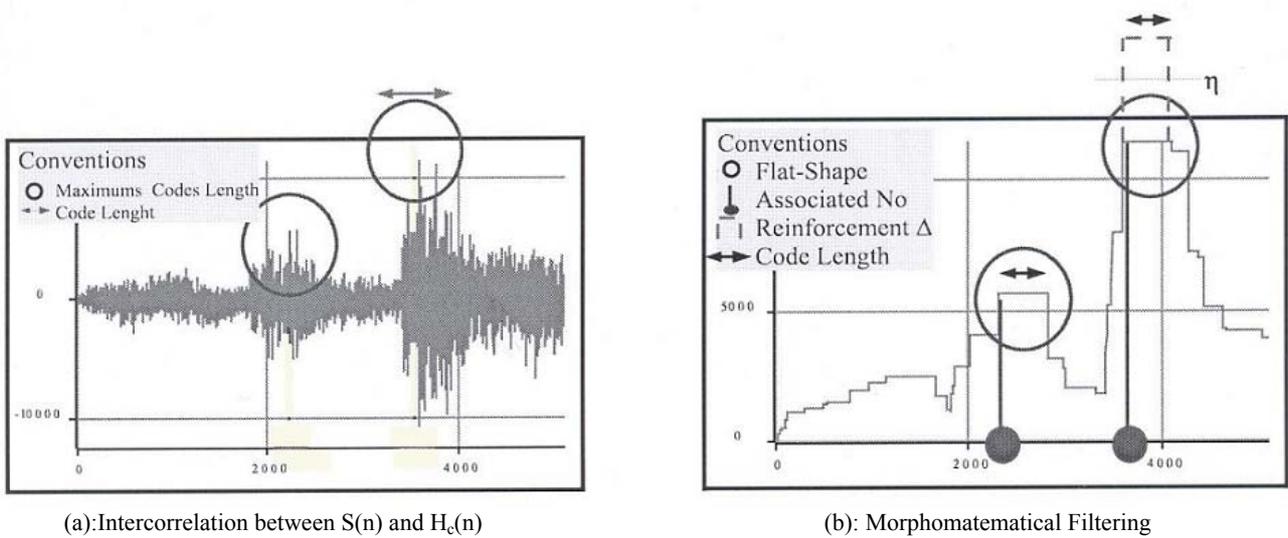


Figure 6. Filtering effect for $h_c(n)$ and $mm(n)$

There is only two possibilities for each instant in $s(n)$. From a communications point of view, it is possible to associate one hypothesis H_0 and H_1 as follows:

- $H_0 : x(n) = s(n)$ if n is not FTP,
- $H_1 : x(n) = s(n) + Ac(n-FTP)$ if n is FTP,

When applied to the filter $h_c(n)$, it implies that:

- $H_0 : y(n) = s(n) \otimes h_c(n)$,
- $H_1 : y(n) = s(n) \otimes h_c(n) + Ac(n-p) \otimes h_c(n)$

When the morpho-mathematical filter is applied it follows:

- $H_0 : m(n) = mm[y(n)=s(n) \otimes h_c(n)]$,
- $H_1 : m(n) = mm [y(n)=s(n) \otimes h_c(n) + Ac(n-p) \otimes h_c(n)]$

We note Δ the term $Ac(n-p) \otimes h_c(n)$. Its importance lies in the fact that the amplitude on the peak present in a feature instant is directly controllable by adding a code weighted by the constant A . Furthermore, the value Δ will determine the robustness degree of the system (The effect of this operation is illustrated in figure 6(b) by the pointed lines).

It is also important to note the choice of the filter $h_c(n)$ combined with the use of $s(n)$ is equivalent to an informed synchronization scheme in which the maximums of the autocorrelation of the synchronism codes are placed on the local maximums of the inter-correlation between $c(n)$ and $s(n)$. From an informed point of view, a vector addition is realized in the FTP's to bring them more easily detectable.

4.2. Reception Stage

The same filters $h_c(n)$ and $mm(n)$ are used in this stage. However, the noise introduced by the channel must be taken into account to define the detection criteria. This criteria is presented in the following subsection.

4.2.1. Detection Criteria

Although the FTP's have been defined as the maximum of the correlation between the host signal and the code, the channel noise influence in the reception stage makes this criterion insufficient. The presence of noise could generate peaks (even if they are often weak they are still maximums) that prevent detection. To face this problem, a comparison with a threshold must be defined.

The first task for the detector is to look for a flat shape segment in $m_b(n)$. Once the flat-shape segment has been identified the system must apply a criterion to determine if this point corresponds to H_1 . If the criterion is not

accomplished, the system will decide H_0 . To define the criterion, a threshold is defined from the two established hypothesis. Based on the fact that the difference between H_0 and H_1 for a specified instant n is Δ , it is known that the minimal difference between a reinforced and a non reinforced point is Δ . The idea is to compare this difference (to do this the analyzed sample and its closest neighborhood are taken) to a predefined threshold. The criterion is:

$$\text{if } m_b(n) - m_b(n-1) < \eta \Rightarrow H_0 \quad \text{decided} \quad (8)$$

$$\text{if } m_b(n) - m_b(n-1) > \eta \Rightarrow H_1 \quad \text{decided} \quad (9)$$

It means that it compares the first sample of $mb(N_0)$ in a found flat-shape with its immediately precedent neighborhood. To determine if this point corresponds to a FTP the threshold is applied allowing the system to decide between hypothesis H_0 and H_1 .

It is important to note that the threshold depends of the codes amplitude A . For the correct behavior of the system, this value has to be fixed and known in reception.

5. RESULTS

The tests have been realized over a digitally recorded sample of an audio broadcasting program sampled to 44.1 kHz (CD quality). The total number of samples is 5120000. The audio file includes speech, music, different kinds of isolated and mixed sounds. The idea is to compare the behavior of the proposed system and a classical S.S. synchronization system over an analogical AWGN channel. In both cases the power of the synchronism has been fixed to the same value. It means that they have the same length N (fixed to 512 samples) and the same amplitude A . This power has been fixed following an audibility criterion: it is the maximum possible power that does not include an audible distortion. Although a young listener has been used to establish this value, the results remains subjective. To deeply evaluate the subjective distortion, more tests more tests should be practiced. For the S.S. system, the K value has been defined (for the average system) as: $K=100$. For simulating an analogical transmission, a delay non multiple of the sampling frequency is used. This is equivalent to realize a A/D D/A conversion. Then, the number of errors has been computed for different SNR's including an ideal noiseless channel. In a second test, the behavior of the system concerning the mp3 compression is analyzed. The same audio file then in the first test have been taken. For each used bit rate, the file has been compressed and immediately decompressed. The detection are realized over the decompressed signals. Five bit rates have been used for this test, varying from 96 to 224 Kilobits par second in steps of 32 kbps. The results are depicted in figure 7.

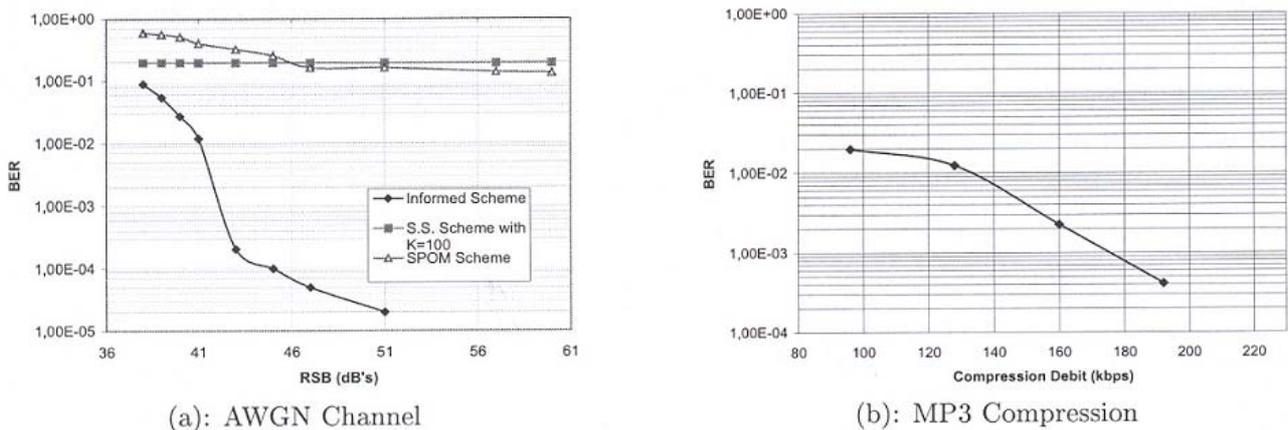


Figure 7. Comparative Results with the same power embedding

The results verifies the fact that a S.S. system treats the host signal as a noise. From an ideal noiseless channel, the power of the host signal perturbs the detection generation an important number of errors. For the informed scheme in the same conditions, the send and received signal are equal and any perturbing noise is added. The consequence is a perfect (zero errors) detection. By the other hand, when SNR is decreased, the S.S. system presents little variations in the number of errors compared with the informed scheme. This can be explained by the fact that the channel noise contribution to the total perturbing signal is little comparatively speaking. In fact, the channel noise is treated jointly with the host signal and its effect is minimized by the average filter. It is also important to note that on one hand, the S.S. and the SPOM (cf. section 2) filter presents improving results compared with the classical S.S. system for RSB under approximately 46 dB's concerning BER. For this test the SPOM filter has not an average filter applied and by consequence it does not generate important delays. However, complementing the SPOM system with an average system yields important improvements. In (b) the results about MP3 compression are depicted. In these tests, the system did not generate errors for debits over 224 kbps (compression rate approximately equal to 3:1).

6. CONCLUDING REMARKS AND FUTURE WORKS

Synchronization in data hiding systems used for data communications is a challenging problem. An informed spread spectrum system presents an innovating solution. When using a pseudo random code to determine the synchronization instants and analyzing the host signal not as a noise, this solution allows high robustness and controlled audio distortion. In this work, just one pseudo random code has been used. However, a codebook containing several codes of many kinds can be used to accomplish this task, future work will be developed in this direction.

REFERENCES

1. A. H. T. L. Boney and K.N. Hamdy, "Digital watermarks for audio", in Eusipco, pp. 473-480, Trieste, 1996.
2. J. L. A, LoboGuerrero P. Bas, "Iterative infirmed audio data hiding scheme using optimal filter", in ICCT chine, pp. 1408-1412, 2003
4. H. S, *Digital communications*, pp. 445-471. John Wiley and Sons, 1988.
5. T. K. Jaap Haitsma, Michel van der Veen and F. Bruekers, "Audio watermarking for monitoring and copy protection", in ACM Multimedia 2000.
6. A. M. I. J. Cox, M. L. Miller, "Watermarking as communications with side information", in Proceeding of IEEE, 87 (7), PP.:1127-1141.
7. M. Costa, "Writing on dirty paper", in IEEE transactions on information Theory, pp. 29:439-411, 1983.
8. I. C. M. L. Miller and J. Bloom, "Informed embedding : Exploiting image and detector information during watermarking insertion", in IIEE International Conference on Image Processing.
9. G. Le Guelvouit, S. Pateux and C. Guillemot, " Perceptual watermarking of non i.i.d signals based on wide spread spectrum using side information", in ICIP02, pp. III: 477-480.
10. J. M. C. P. Bas and B. Macq, "Geometrically invariant watermarking using feature points", in IEEE transactions on image Processing, pp. vol 11, 9 pp. 1014-1028, 2002.
11. J. B. I. J. Cox, L. Matthew, *Digital Watermarking*, pp. 117-153. Morgan Kauffman, 2002.