IMPLEMENTATION OF AN AUDIO WATERMARKING SYSTEM USING CDMA MODULATION

Said ghnimi, Adnen cherif

Signal Processing Laboratory, Science Faculty of Tunis
2092, University of Tunis El Manar, Tunisia
said.ghnimi@gmail.com ; Adnane.cher@fst.rnu.tn

ABSTRACT

In this paper, we will present an audio watermarking system based on CDMA modulations. After having presented the watermarking structure and its configuration, we will develop the numerical implementation of our interface programmed under Matlab. Validation tests are conducted in a real time acquisition for the verification of the watermarking system performances and its robustness.

Key words : watermarking, modulation, Audio protection, CDMA, information security .

1. INTRODUCTION

With the development of audio and speech transmission with several numerical formats, the protection of the intellectual properties rights became a major problem.
Watermarking was a potential solution to this problem since a few years [1]. It consists in inserting an indelible mark directly in the audio signal; generally it has to satisfy the following constraints: inaudibility, the robustness with the operations of traditional treatment on the signals and resistance to attacks deliberated on third. These studies made it possible to extend watermarking for new applicability: one of them consist in using the audio signal like a transmission channel conveying binary data [2- 3 - 4]. The system of watermarking is presented like a chain of communication at the very particular properties. Useful information is transmitted with a low power in front of a noise, the audio signal, strongly correlated and no stationary.
If this application freed from the concept of " pirate ", it must nevertheless respect the traditional constraints of watermarking especially to satisfy objectives of flow and reliability of transmission.

In fact the couple flow - binary error rate (TEB) defines the concept of performance of a system of watermarking dedicated to the transmission of information. This couple is also added the factor cost in term of computing time, of which the taking into account is necessary during the implementation system of watermarking.
One proposes is to study the contribution of various modulations by spreading out of spectrum on the performances of a system of watermarking to vocation of transmission of information [5]. After description system of watermarking's system used, we will detail the techniques of modulation considered. An analysis of experimental results of these modulations in term of flow, computing TEB and time will enable us to choose the best system's modulation.

2. THE WATERMARKING SYSTEM

The first watermarking system, developed in [6,7], For signals audio frequencies sampled with 44.1kHz, is structured in the shape of a communication's chain, figure 1.

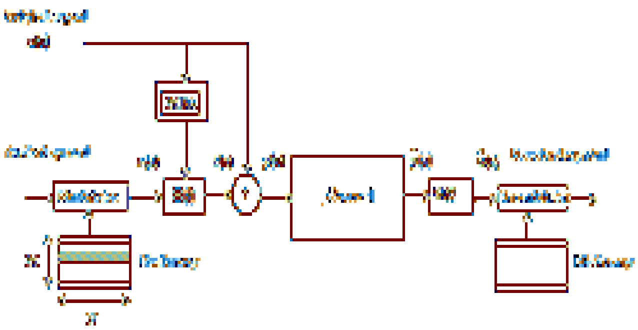

Figure 1. Principle of the watermarking system

The process of watermarking can be seen like a chain of disturbed communication. watermarking is the signal to be transmitted and the audio signal is regarded as a bruit[7], this approach is represented by the figure 1 where one finds the three stages traditional of a chain of communication; the transmitter siege of the insertion of information, the channel, the receiver siege of detection[12].

Information to be inserted or message source can take various forms; text, word, image, signal. Identifying the audio signal host or his author, it is subjected to an operation of conversion which transforms it into a binary sequence taking its values in {0,1} or more generally the message can be converted into a sequence of symbol (S_K) Taking its values in an alphabet finished A.

This message undergoes a stage of modulation to transform to be into a physical signal, the modulated signal; this signal must then be formatted spectral to satisfy the constraint of inaudibility. This working exploits a psychoacoustic model which extracts from the audio signal a threshold of masking translating the frequential limit of inaudibility of a signal in the presence of the audio signal. From this masking threshold is deduced the filter H which formats the modulated signal so that its spectral concentration of power coincides with the threshold of masking of the audio signal. The first stage of its implementation consists in determining the curve of masking [7]. Once the threshold of calculated masking, the system synthesizes the filter of working associated, the calculation of the coefficients of the filter amounts solving a linear system. Indeed, we must determine the filter:

$$H(z) = \frac{b_o}{1 + \sum_{i=1}^{p} a_i z^{-i}} \quad (1)$$

such that the signal filtered y(n)=h(n)*v(n) is the best estimate of the signal $m(n)$ whose spectrum is the curve of masking of $x(n)$ within the meaning of the criterion

$$\min\left(\left\| y(n) - m(n) \right\|\right)^2 \quad (2)$$

For that we thus seek an auto regression process y(n) associated the polynomial

$$A(z) = \sum_{i=1}^{p} a_i z^{-i} \quad (3)$$

checking the relation of recurrence:

$$\sum_{i=0}^{p} a_i y(n-i) = b_0 v(n). \quad (4)$$

v(n) being a white vibration, on has the relation then

$$\forall k \geq 1 : E[y(n-k)v(n)] = 0 \quad (5)$$

If one poses m(n)=y(n)+v(n), and

$$y(n) = -\sum_{i=1}^{p} a_i m(n-i) \quad \text{then } \forall k \geq 1 \quad (6)$$

$$E[m(n-k)(m(n)-y(n))]=0. \quad (7)$$

y(n) is thus indeed the best linear regression on average quadratic of m(n).

The coefficients of the filter are given by the equation of Yule-Walker:

$$\begin{bmatrix} R_m(0) & \cdots & R_m(P-1) & R_m(P) \\ \vdots & & & \\ R_m(P-1) & & \left|R_m^{P-1}\right| & \\ R_m(P) & & & \end{bmatrix} \begin{bmatrix} 1 \\ a_i \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} b_0^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (8)$$

Where them $R_m(k) = E\left[m(n-k)m(n)\right]$ (9) are the coefficients of auto covariance signal $m(n)$ and R_m^{p-1} is the positive matrix of auto covariance of P-1(matrix order of Toeplitz) of the signal $m(n)$ One deduces the system from it:

$$\begin{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_p \end{bmatrix} = (R_m^{p-1})^{-1} \begin{bmatrix} R_m(1) \\ \vdots \\ R_m(P) \end{bmatrix} \\ b_0^2 = R_m(0) + \sum_{i=1}^{p} R_m(i) a_i \end{bmatrix} \quad (10)$$

This system of equations is solved efficiently using the algorithm of Levinson-Durbin, by exploiting the Toeplitz structure of the matrix R_m

After filtering of $v(n)$ by $H(z)$ we obtain the signal of tattooing $t(n)$.This signal is inserted in the signal of music by simple addition to form the tattooed signal of music $y(n)$.

The power of signal of watermarking is thus increased (offering a report/ratio of power between the signal of watermarking and the audio noise more favourable for the detection of information tattooed) while respecting the constraint of inaudibility, two types of receivers can be considered: one by filtering of Wiener [8] is the different one by filtering whitening [9] in both cases the goal of the receiver is to estimate to a certain extent the signal modulated starting from the received signal The recovery of the signal of tattooing $\hat{v}(n)$ starting from the tattooed audio signal $\hat{y}(n)$ exploit a filtering of Wiener. This filter thus carries out the equalization of the channel and the bleaching of the noise, i.e. of the signal of music [11].

The inversion of the channel is carried out by recomputing the filter *H* starting from the tattooed audio signal. So it exploited the model psychoacoustics and must thus be given on windows of analysis identical to those of $H(f)$ therefore of size N_{HG}.

For each window of analysis, the theory developed by Wiener makes it possible to determine an optimal filter $W(f)$, within the meaning of the criterion of minimization of the power of the error in estimation): this filter restores the original signal as well as possible to tattoo $v(n)$ starting from the data of $\hat{y}(n)$ The objective is thus to determine the impulse response $w(n)$ of support W who minimizes the criterion of estimate MSE:

$$MSE = E\left[(v(n)-\hat{v}(n))^2\right] = E\left[(v(n)-\sum_{i\in w} w(i)\hat{y}(n-i))^2\right]$$

$$= E\left[v^2(n) - 2\sum_{i\in w} w(i)E[v(n)\hat{y}(n-i)] + \sum_{i\in w\, j\in w} w(i)w(j)E[\hat{y}(n-i)\hat{y}(n-j)]\right] \quad (11)$$

The minimization of the MSE requires the calculation of the derivative partial and the resolution of the system thus obtained:

$$\forall i \in W, \frac{\partial MSE}{\partial w(i)} = -2E[v(n)\hat{y}(n-i)] + 2\sum_{j\in w} w(j)E[\hat{y}(n-i)\hat{y}(n-j)] = 0 \quad (12)$$

The solution called the equation of Wiener Hopf is given by:

$$\forall i \in W, \sum_{j\in w} w(j)E[\hat{y}(n-i)\hat{y}(n-j)] = E[v(n)\hat{y}(n-i)]. \quad (13)$$

By making the assumption that $y(n)$ was disturbed too much by the operation of compression –decompression i.e. only $\hat{y}(n) \cong y(n)$ one a:

$$\hat{y}(n) = x(n) + t(n) = x(n) + \sum_{i\in H} h(i)v(n-i) \quad (14)$$

$v(n)$ and $x(n)$ being not coloured and $v(n)$ is white. One deduces.

$$E[v(n)\hat{y}(n-i)] = E[v(n)t(n-i)] \quad (15)$$

That is to say the equation:

$$\forall i \in W, \sum_{j\in w} w(j)E[\hat{y}(n-i)\hat{y}(n-j)] = \sum_{j\in H} h(j)E[v(n)v(n-j-i)]. \quad (16)$$

Let us pose $R_{\hat{y}}$ the matrix of covariance of $\hat{y}(n)$ and R_v that of $v(n)$ thus P_w, Q_w, lengths of the parts causal and no causal of the filter of Wiener $W(f)$, and P_h the length of the impulse response $h(n)$ .The equation can then be written in the form:

$$\forall i \in [-Q_w, P_w], \quad \sum_{j=-Q_w}^{P_w} w(j)R_{\hat{y}}(i-j) = \sum_{j=0}^{P_h-1} h(j)R_v(i+j) \quad (17)$$

One deduces the coefficients from the filter of Wiener:

$$\begin{bmatrix} w(-Q_w) \\ \vdots \\ w(0) \\ \vdots \\ w(P_w) \end{bmatrix} = A^{-1}B \begin{bmatrix} h(0) \\ \vdots \\ h(P_h-1) \end{bmatrix} \quad (18)$$

$$A = \begin{bmatrix} R_{\hat{y}}(0) & \cdots & R_{\hat{y}}(-Q_w-P_w) \\ \vdots & & \vdots \\ R_{\hat{y}}(Q_w) & \ddots & R_{\hat{y}}(-P_w) \\ \vdots & & \vdots \\ R_{\hat{y}}(Q_w+P_w) & \cdots & R_{\hat{y}}(0) \end{bmatrix}$$

$$B = \begin{bmatrix} R_v(-Q_W) & \cdots & R_v(-Q_w+P_h+1) \\ \vdots & & \vdots \\ R_v(0) & \ddots & R_v(P_h-1) \\ \vdots & & \vdots \\ R_v(P_w) & \cdots & R_v(P_w+P_h-1) \end{bmatrix}$$

One can then carry out the detection of Cached information using a demodulator by correlation and of a dictionary of reception (identical to that of emission in the case of Wiener).

3. IMPLEMENTATION OF THE MODULATION CDMA

Modulation CDMA (Code Division Multiple Access) is our choice like a modulations by spreading out of the spectrum is usually used in the communication systems [13]. It makes it possible to several " users " to divide the same transmission chain. That Ci can also be seen as the transmission of information broken up on a basis of M = m orthogonal vectors Let us explain principle in details more:
Let us consider a dictionary of M = m orthogonal forms of wave, i.e. of the white vectors of duration N_{bits} and a m-uplet has transmitter (bi, ..., b_m). Each element b_k of the m-uplet is associated in a single way the one of the shapes of wave of the dictionary d_k (n). This form of wave is then balanced by the physical amplitude has k associated with the bit considered b_k the information transmitted during time symbol is the sum of all the contributions of the binary characters of the m-uplet, namely: $V(n) = \sum_{k=1}^{m} a_k d_k(n)$.

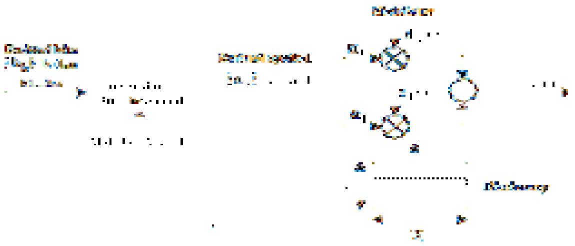

Figure 2 Principle of modulator CDMA

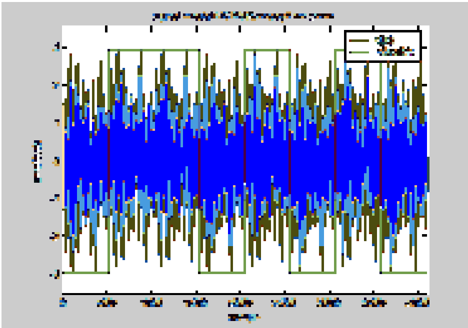

Figure 3 Modulation CDMA: bits sent in series and visualization of the emitted bits

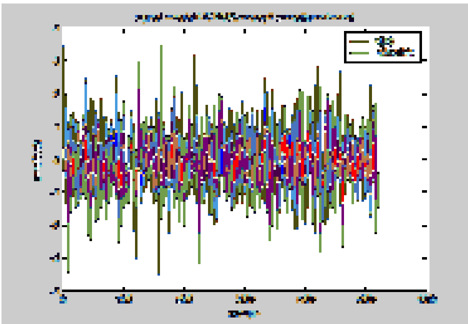

Figure 4 Modulation CDMA: bits sent simultaneously and visualization of the emitted bits

4. EXPERIMENTAL RESULTS

4.1. Experimental protocol

The performances of these modulations by spreading out of spectrum are evaluated by the data information of the EP according to flow Rs=m*Rb This TEB means, are obtained by watermarking of B bits of sample's information of 5 music's signals and 5 signals of word, sampled with 44.1 kHz . The TEB is an effective estimator of transmitted error's probability P for a precision of $\Delta P = \sqrt{\frac{Pe}{\mathrm{B}}}$ and a rate of confidence of 71%. Three noted impose themselves:

– It appears illusory to make function the system of watermarking to flows leading to a TEB higher than 5%, insofar as the system has vocation to transmit information.
– The precision indicates that it would be necessary to transmit 1000 of bits to obtain a relative precision near to 0, very expensive simulation in computing times. The choice was thus made transmit B = 100 bits of information, good compromise between the precision of the results (0.32% for a EP of 1%) and the computing time.
– The tests of watermarking's system showed the strong dependence of the EP to the audio signal to tattoo: for one they, the TEB obtained are often definitely higher (about twice) at the average TEB. The origin of these differences still remains to be clarified. Nevertheless, the variations of the EP according to the flows are identical whatever the signal used.

4. 2. Results and discussions

A. Criterion of computing time

In this part, we studied the various modulations which are used in the chain of watermarking, which helps us in the analysis of modulation's performances.

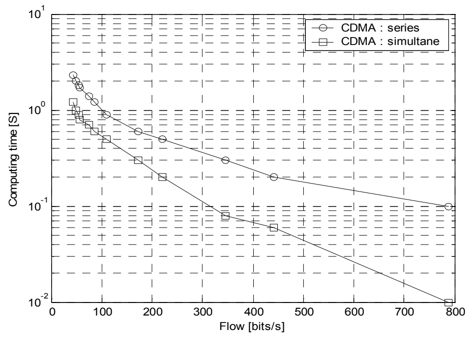

Figure 5 Computing Times of the system of watermarking for a channel without disturbance

Figure 5 represents the execution time of all the system of watermarking for the software Matlab with values of varying flow of 40 A 800(bits/s). The number of the emitted bits is worth 100 (bits).
We check by comparing the two curves represented on this figure that the contribution in term of computing time of watermarking 's system based on simultaneous modulation CDMA very significant compared to that in system of watermarking is based on modulation CDMA in series. this is clear according to the temporal margin between the two curve of figure 6.
In addition notices of it that the computing time decreases with the fur and measurement which the flow increases. This can be explained by the fact why when the flow dominoes, the time of the signal to treat increases (for a

constant number of emitted bits (100(bits)). Indeed the time of the audio signal is given by the following relation:

$$T_{signal\ audio} = (N_{be}/R) \qquad (19)$$

A study of SNR shows that for all the modulations some is the flow or the size of dictionary, average value SNR on windows of size NR is identical. From this point of view, the modulations are thus equivalent.
The study of the cost in term computing time shows that simultaneous modulation CDMA is the best. Thus techniques the most adapted to the system of watermarking [5]. They make it possible to improve the performances of the system for transmission's rates.

B. Evaluation of the transmission reliability

The vocation of the system being to transmit information via an audio signal, it must offer to the traditional chains of communications; a reliability of transmission as high as possible for a rate of transmission as large as possible.
Simulations which we carried out are executed on only one computer without disturbance of the channel. The tattooed signal (received) is thus a simple recopy of the emitted audio signal ($y(n) = \hat{y}(n)$) On is not interested in this case in the phenomenon of synchronization and the problems which can be generated by the analogical passage of the audio signal or by its compression decompression.

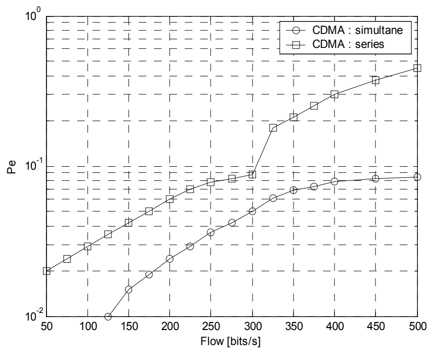

Figure 6 Error probability of CDMA Modulation

The probability of error of the Matlab system is obtained for values of varying flow from 50 to 500 (bits/s) is illustrated on figure 6 for a dictionary of two symbols. The number of emitted bits is worth 100 bits.
According to figure 6 the probability of modulation's error CDMA in series is more significant than that of simultaneous modulation CDMA in particular between 300 and 500(bits/s).
One deduces from it that the technique of simultaneous modulation CDMA is more reliable during the transmission of information on the chain of audio watermarking

C. Evaluation of the capacity of insertion

The capacity of insertion is the maximum quantity of information being able to be transmitted for a quasi null probability of errors. The flow of exig insertion 3rd by the system of watermarking is strongly depend on the application envisaged (of a hundred bits for the animation of the clone in project ARTUS to a thousand in the case of applications for the reduction of noise) [10]. In both cases, it is advisable to choose a system of watermarking allowing a capacity of insertion the largest possible.
The insertion of information representing the signal modulated CDMA sent in series in an original audio signal (xn), we gave a signal watermarking tn to represent in the following figure:

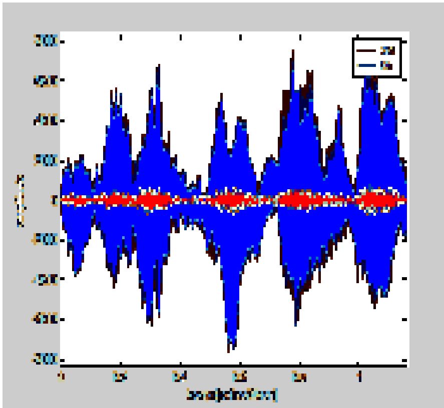

Figure 7 Temporal valuation of signals xn and tn in real times

By introducing the signal modulated CDMA sent simultaneously and the signal of masking led of the original audio signal (xn) into a filter of working, one obtains the signal (tn) visualized below.

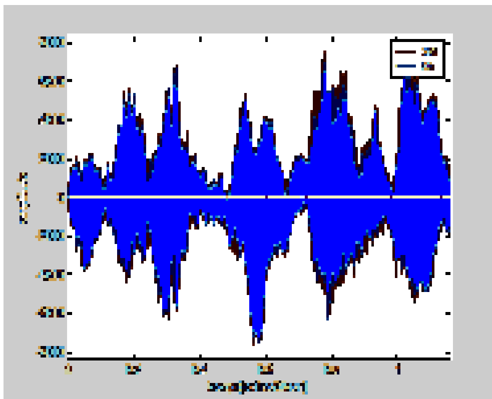

Figure 8. Temporal valuations of signals xn and tn in real times

When the bits are emitted in series it has a widening spectral more significant than for the case or several bits to send simultaneously, each bits occupies a plug of frequency which enables us to avoid the overlapping and the recovery of the bits on the level of detection ensures a better distinction of the bits representing all information.
When we makes the sum of the watermarking signal (tn) and of the original signal (xn), we obtain the signal (yn) as illustrated in figure 9.

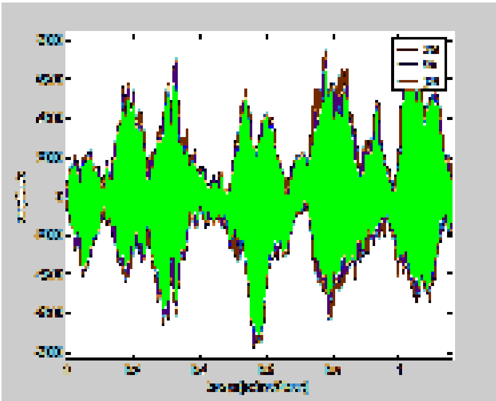

Figure 9 Temporal valuation of signals xn and tn and yn in real times

By making the sum of the signal watermarking (tn) and the original signal (xn), we obtained the watermarked signal (yn) .

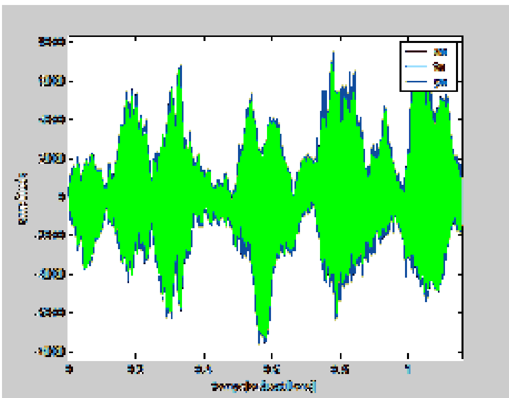

Figure 10 Temporal valuations of signals xn and tn and yn in real times

The modulation by spreading out of spectrum makes it possible to establish the original signal; it seems a white vibration on a wide strip. The observation of the two figures presented shows that modulation CDMA for the case of the bits to send simultaneously made it possible to almost completely eliminate the effect of interference.

5. CONCLUSIONS

In this paper, we succeeded to implement a watermarking audio interface by using a CDMA modulation. The principal contribution developed in this paper is the system adaptation with watermarking specifications such as the capacity of insertion and the reliability of transmission.
We demonstrated that the CDMA simultaneous modulation is most powerful that the serial one insuring transmissions without errors and with low bit rates about 512 bits/s.

REFERENCES

[1] N. KaMatcu, M. Yanada, and H. Taminaga« Authentication. System using digital watermark in Telematics » Symposium on Cryptography and information Security 1989 Japon.
[2] W. Bender, D. Gruhl, N. Morimoto, A. Lu, « Technique for data hiding», IBM Systems journal 35 (3/4), pp. 313-336, 1996.
[3] G. Caronni, « Assuring Ownership rights for digitals images », proc Reliable IT, Systems vis 95 vieweg publishing company 1995.
[4] I. Cax, j. Kilian, F.T. Leighten, and T. Shamam, A. Secure, « Robust watermark for multimedia, information Hiding », First Int. Workshop, proc, 1996.

[5] Cléo Baras, « Etude de la mise en forme de l'information binaire dans un système de tatouage audio » 2002.
[6] L. Boney, A. TewFik, and K. Hamdy, « Digital Watermarks for audio signals », IEEE, Int conf. Multimedia computing systems, Hiroshima June 1996.
[7] R. A. Garcia, « Digital watermarking of audio signals using a psychoacaustique auditory model and spread spectrum » Theory 107 The AES convention New York Septembers 1999.
[8] S. Larbi. M. J. Aïdane et N. Moreau « A new Wiener filtering based detection schema for time domains perceptual audio watermarking » ICASSP, Vol. 5, p. 949-52, 2004.
[9] H. KIM « Stochastic model based audio watermark and whitening filter for improved detection » proceedings of the international conference on Acaustique, speech, signal processing 2000.
[10] P.Bas.«Projet Artus»
http://www.lis.inpg.Fr/pageperso/bas/artus/portARTUS

[11] Brossier (J-M), « Signal et Communication numérique Egalisation et synchronisation », collection traitement du signal 1997.
[12] J. G. proakis, « Digital Communication » Mc Grow-Hill 1995
[13] Alejandro Lobo Guerrero Pérez, « Etude de techniques de tatouage audio pour la transmission de données 2004. »