COMPARATIVE STUDY OF TWO INFORMED EMBEDDING STRATEGIES FOR AUDIO SPREAD-SPECTRUM DATA HIDING SYSTEMS

C. Baras[1], *N. Moreau*[1], *and P. Dymarski*[2]

[1] GET - ENST, TSI Department,
46 rue Barrault, 75013 Paris, FRANCE
phone: (33)(0) 1 45 81 76 40 , fax: (33)(0) 1 45 81 37 94
[2] Technical University of Warsaw,
15/19 Nowowiejska, 00-665 Warsaw, POLAND
email: baras@tsi.enst.fr, nicolas.moreau@enst.fr, dymarski@tele.pw.edu.pl

ABSTRACT

In the particular application field of broadcasting, audio data hiding systems should ensure an inaudible, reliable and robust transmission for various channel perturbations. In this paper, we present two informed embedding strategies adapted to a closed-loop data hiding scheme. Both strategies aim at maximizing system robustness to additive channel perturbation and at limiting local perceptual distortion. In the first one, system robustness is based on the input signals of the correlator employed in reception process, whereas in the second one, system robustness is related to the transmission error probability. Experimental results on real audio signals are presented to compare the efficiency of the strategies. In terms of transmission reliability, the second strategy is comparable to the first one until 300 bps and is slightly more robust than the first one for higher transmission rate, but requires a higher computational cost than the first one.

1. INTRODUCTION

Audio data hiding research was developed with the growing use of audio signals under digital format. Data hiding is a generic term which groups processes used to embed some binary information into an audio signal without any perceptual degradation. Embedded information brings an added value to the audio signal, which can be interesting for many applications [1]: it can be related to content description for indexing, to labelling for monitoring, to advertising for commercial broadcasting, or even to signature for copyright protection.

Spread-Spectrum (SS) data hiding systems are designed as a communication channel, embedding a binary message in an audio signal. Systems should offer an embedding strategy which conciliates perceptual distortion and information detection constraints. They should also be robust to classical distortions (further referred to as channel perturbations) applied to audio signals. These distortions described in [2] are: filtering, format change, noise addition and dynamics change. In this context, perceptual distortion and transmission reliability with respect to transmission rate and channel perturbations are major issues and define system performances.

State-Of-The-Art in data hiding points out the efficiency of informed embedding strategies. These strategies take into account the *a priori* knowledge of the audio signal during the embedding process to choose adapted watermarks that ease the detection of the embedded information and still respect the inaudibility constraint. Several strategies for SS data hiding with aim at maximizing system robustness to additive channel perturbation have already been proposed. One of them, proposed by Miller *et al.* in [4], uses an iterative embedding algorithm that builds the watermark by adding perceptually shaped components until the robust detection of the information is ensured. Another, envisaged by Malvar *et al.* in [5], adapts the watermark power in order to minimize the error probability without altering the average perceptual distortion. These two strategies succeed in removing a part of the interference of the audio signal on the detection process, yielding high transmission reliability, but they are deficient in perceptual distortion control since they do not limit the local perceptual distortion.

In this paper, we propose and compare two informed embedding strategies, adapted to an informed data hiding system that limits local perceptual distortion. This distortion is controlled by a psychoacoustical model, that limits for each embedded information the choice of the watermarking signal to a finite set, defining the inaudibility region. Both embedding strategies intend to maximize system robustness to additive channel perturbations. They exploit a local copy of the receiver at the embedder to estimate signals taking part in the detection process and choose the appropriated watermark with respect to their own criterion:

- the first strategy, similar to Miller's one and already presented in [3], defines system robustness directly with the input signals of the correlator, employed in decision process.
- the second strategy, as Malvar's one, relies on the probability of making an error during the detection process but is novel in the sense that it ensures a fixed error probability in presence of a channel noise with maximized power.

The outline of the paper is the following. Principles of the proposed data hiding system and its closed-loop design are described in section 2. The two informed embedding strategies are then presented: section 3 expounds their characteristics and section 4 details their implementations. Experimental results are given in section 5 to evaluate and compare the influence of the two informed strategies on system performances.

2. INFORMED DATA HIDING SYSTEM PRINCIPLES

Figure 1 illustrates the proposed closed-loop data hiding system.

2.1 Pre-requisites

The system is based on a non-informed data hiding scheme, designed as a communication channel.

At the embedder, the source encoding process maps the binary message to be embedded into a sequence of L symbols $\{k_l\}_{l=1..L}$, chosen in the set $\{1,...,M\}$. The modulation interface uses an embedding codebook $\underline{S} = \{\underline{s}_k\}_{k=1..M}$ containing M SS waveforms with length N and unit power. Each symbol k_l is mapped into the k_l-th waveform of $\underline{S}$ so that the modulated signal on the l-th symbol interval $[(l-1)N...lN-1]$ is : $\underline{v} = \underline{s}_{k_l}$. To satisfy the inaudibility constraint, the watermarking signal $\underline{t}$ is constructed by filtering $\underline{v}$ with a psychoacoustic shaping filter $H(f)$. $H(f)$ is computed on each symbol interval by a psychoacoustic study of the audio signal $\underline{x}$. The watermarked audio signal $\underline{y}$ is finally obtained by adding the watermarking signal $\underline{t}$ to the audio signal $\underline{x}$.

The receiver scheme is designed supposing that there is no symbol interference. The signal $\hat{\underline{y}}$ resulting from channel perturbations applied on $\underline{y}$ is filtered by the whitening filter $G(f)$ of $\underline{x}$, yielding $\hat{\underline{v}}$. If the channel is free from perturbation, $\hat{\underline{v}}$ can be modelled as

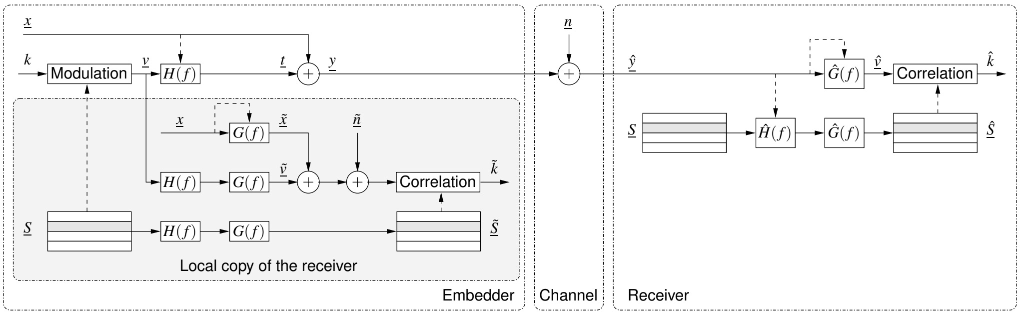

Figure 1: Closed-loop data hiding system.

the received signal of an Additive White Gaussian Noise (AWGN) channel, where the noise is the whitened audio signal. Now, the transmitted watermarking signal has only M possible forms that define the reception codebook $\hat{\underline{S}} = \{\hat{\underline{s}}_m\}_{m=1..M}$. Each of them are computed by filtering the embedding codebook waveforms by $H(f)$ and $G(f)$. Consequently, the chosen receiver is a correlation demodulator. It selects the reception codebook waveform which correlation with the received signal is the highest. Since $\underline{x}$ is not available during the reception process, $H(f)$ and $G(f)$ are approximated by two filters $\hat{H}(f)$ and $\hat{G}(f)$. $\hat{H}(f)$ and $\hat{G}(f)$ are respectively the perceptual shaping filter and the whitening filter of $\hat{\underline{y}}$.

2.2 Using the local copy of the receiver

A local copy of the receiver scheme is introduced at the embedder to take the *a priori* knowledge of the audio signal into account. It allows us to estimate the signals taking part in the reception process. They are: the whitened audio signal $\tilde{\underline{x}}$, the filtered modulated signal $\tilde{\underline{v}}$ and the estimated reception codebook $\tilde{\underline{S}} = \{\tilde{\underline{s}}_m\}_{m=1..M}$. The inaudibility contraint and conditions of a correct detection can now be stated.

Suppose that the symbol k_l has to be embedded during the l-th symbol interval. The inaudibility constraint is ensured by the perceptual shaping filter $H(f)$. Its design only imposes to choose a modulated signal $\underline{v}$ that satisfies the following inequality:

$$\sigma_v^2 = \frac{1}{N}\underline{v}^t\underline{v} \leq 1. \tag{1}$$

Moreover, given signals estimation at the local copy of the receiver, k_l is detected with no error if the following $M-1$ inequalities are satisfied at the input of the correlator:

$$\forall m \neq k_l, (\tilde{\underline{x}}+\tilde{\underline{v}}+\tilde{\underline{n}})^t\tilde{\underline{s}}_{k_l} > (\tilde{\underline{x}}+\tilde{\underline{v}}+\tilde{\underline{n}})^t\tilde{\underline{s}}_m, \tag{2}$$

where $\tilde{\underline{n}}$ is some channel noise, which is not supposed to be known at the embedding process.

At this point, the informed embedding strategy must establish how to choose the adapted watermarking signal that conciliates the inaudibility contraint (1) and correct detection conditions (2) for a channel noise with a maximum variance.

3. INFORMED EMBEDDING STRATEGIES

Given the previous correct detection conditions (2), the embedding strategy can be related to two criteria. The first one deals with maximizing a robustness parameter, that is introduced directly in equation (2) to characterize system robustness to channel perturbation. The second one relies on error probability supposing that $\tilde{\underline{n}}$ can be modelled by a white Gaussian random process. Further, the embedding strategy based on the former will be denoted by RPS (for Robustness Parameter based Strategy) and the one based on the latter will be denoted by EPS (for Error Probability based Strategy).

3.1 Strategy RPS

Since the channel noise is unknown, we substitute $\tilde{\underline{n}}$ in (2) by a robustness parameter σ_n^2, that characterizes system robustness to additive perturbations, as in Miller's strategy [4]. Robust detection contraints (2) become:

$$\forall m \neq k_l, (\tilde{\underline{x}}+\tilde{\underline{v}})^t(\tilde{\underline{s}}_{k_l} - \tilde{\underline{s}}_m) \geq \sigma_n^2. \tag{3}$$

In this context, maximizing system robustness amounts to finding $\tilde{\underline{v}}$ satisfying (3) with a maximum robustness parameter σ_n^2. This is finally related to the following equation:

$$\tilde{\underline{v}} = \arg\max_{\tilde{\underline{u}}} J_1(\tilde{\underline{u}}), J_1(\tilde{\underline{u}}) = \min_{m\neq k_l}(\tilde{\underline{x}}+\tilde{\underline{u}})^t(\tilde{\underline{s}}_{k_l} - \tilde{\underline{s}}_m) \tag{4}$$

Here, system robustness is directly linked to the inputs of the correlator.

3.2 Strategy EPS

By defining the $M-1$ normalized vectors $\tilde{\underline{s}}_{k_l m} = \frac{\tilde{\mathbf{s}}_{k_l} - \tilde{\mathbf{s}}_m}{||\tilde{\mathbf{s}}_{k_l} - \tilde{\mathbf{s}}_m||}$, detection constraints (2) can be rewritten as follows:

$$\forall m \neq k_l, (\tilde{\underline{x}}+\tilde{\underline{v}})^t\tilde{\underline{s}}_{k_l m} > -\tilde{\underline{n}}^t\tilde{\underline{s}}_{k_l m}. \tag{5}$$

The probability of erroneous decision is:

$$P_e = 1 - prob(\forall m \neq k_l, (\tilde{\underline{x}}+\tilde{\underline{v}})^t\tilde{\underline{s}}_{k_l m} > -\tilde{\underline{n}}^t\tilde{\underline{s}}_{k_l m}).$$

We suppose that $\tilde{\underline{n}}$ is a set of $M-1$ zero-mean Gaussian random variables with variance σ_n^2 and distribution $N(0,\sigma_n^2)$. Ideally, if the set $\{\tilde{\underline{s}}_{k_l m}\}_{m\neq k_l}$ is orthogonal, the set $\{n_{k_l m} = -\tilde{\underline{n}}^t\tilde{\underline{s}}_{k_l m}\}_{m\neq k_l}$ could be viewed as $M-1$ random variables, statistically independent [6]. However, the variables $\{n_{k_l m}\}$ are not independent, because the set of vectors $\{\tilde{\underline{s}}_{k_l m}\}_{m\neq k_l}$ is not orthogonal. In this case, the set $\{n_{k_l m} = \tilde{\underline{n}}^t\tilde{\underline{s}}_{k_l m}\}_{m\neq k_l}$ can still be modelled as $M-1$ random variables, yet statistically dependent. Nevertheless, the error probability P_e may be approximated using a formula adequate for independent variables:

$$P_e \approx P_e' = 1 - \prod_{m\neq k_l} prob(c_{k_l m}(\tilde{\underline{v}}) > n_{k_l m}),$$

with $c_{k_l m}(\tilde{\underline{v}}) = (\tilde{\underline{x}}+\tilde{\underline{v}})^t\tilde{\underline{s}}_{k_l m}$. It may be noted, that, if the codebook $\underline{S}$ contains orthogonal vectors, then $P_e \leq P_e'$. Since the $n_{k_l m}$ have the same distribution $N(0,\sigma_n^2)$, P_e' can be evaluated as follows:

$$P_e' = 1 - \prod_{m\neq k_l}\left(1 - Q\left(\frac{c_{k_l m}(\tilde{\underline{v}})}{\sigma_n}\right)\right), \tag{6}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{t^2}{2}) dt$.

We wish to find $\underline{\tilde{v}}$ that minimizes P'_e and maximizes σ_n^2 conjointly. Unfortunately, this problem has one degree of freedom, since for any given $\underline{\tilde{v}}$, P'_e increases with σ_n^2. In this strategy, since the channel noise is unknown, we decide to set P'_e to a fixed value p.

This time, maximizing system robustness deals with finding $\underline{\tilde{v}}$ that ensures $P'_e = p$ with a maximized noise variance σ_n^2.

4. IMPLEMENTATION

We now intend at finding in practice the watermarking signal, defining $\underline{v}$ and $\underline{\tilde{v}}$, which satisfies the inaudibility constraint (1) and the robust detection contraints, depending on the chosen embedding strategy. For both strategies, the received signal is expanded over the reception codebook waveforms. Hence, $\underline{\tilde{v}}$ can be chosen in the signal space defined by $\underline{\tilde{S}}$. Due to filtering linearity, $\underline{v}$ belongs to the signal space defined by $\underline{S}$. Therefore, $\underline{v}$ can be searched as a linear combination of the embedding codebook waveforms: $\underline{v} = \sum_{m=1}^{M} \alpha_m \underline{s}_m$. Using a vector representation, $\underline{v} = \underline{S}\,\underline{\alpha}$, which results in $\underline{\tilde{v}} = \underline{\tilde{S}}\,\underline{\alpha}$. Consequently, the choice of the adapted watermarking signal depends on the choice of the codebook and the evaluation of the coefficients $\underline{\alpha}$.

4.1 Choice of the codebook and a signal to be detected at the receiver

As in Costa's model [7], we structure the embedding codebook $\underline{S}$ as a set of M orthogonal sub-codebooks $\{\underline{S}_m\}_{m=1..M}$. Each subcodebook $\underline{S}_m = \{\underline{s}_m^q\}_{q=1..Q}$ contains Q biorthogonal waveforms, all able to transmit symbol m.

When the symbol k_l is transmitted, the correlator selects the waveform $\underline{\tilde{s}}_k^{opt}$ of $\underline{S}_{k_l}$, whose correlation with the received signal is the highest. Therefore to ease the detection, the modulated signal should be roughly correlated with $\underline{s}_{k_l}^{opt}$. Now, equations (3) and (5) show that the higher the correlation between $\tilde{\mathbf{x}}$ and $s_{k_l}^{opt}$, the easier the detection. Thus, $s_{k_l}^{opt}$ is chosen for both strategies so that:

$$opt = \arg \max_{q=1..Q} J_2(q), J_2(q) = \underline{\tilde{x}}^t \underline{\tilde{s}}_{k_l}^q. \tag{7}$$

4.2 Evaluation of the coefficients α

Given the codebook $\underline{S}$ chosen previously and the waveform $\underline{s}_{k_l}^{opt}$ chosen to detect the symbol k_l, we intend to find the coefficients $\underline{\alpha}$ which permit the inaudible and robust transmission of k_l. Their evaluation is related to two optimisation problems under constraints, depending on the chosen embedding strategy.

4.2.1 Strategy RPS

The RPS strategy consists in choosing $\underline{\alpha}$ that satisfies (1) and (4), that is:

$$\begin{cases} \underline{\alpha} = \arg\max_{\underline{\lambda}} J_3(\underline{\lambda}), \\ J_3(\underline{\lambda}) = \min\limits_{m=1..M, m\neq k_l, q=1..Q} (\underline{\tilde{x}} + \underline{\tilde{S}}\underline{\lambda})^t (\underline{\tilde{s}}_{k_l}^{opt} - \underline{\tilde{s}}_m^q), \\ \frac{1}{N} \underline{\lambda}^t \underline{S}^t \underline{S}\,\underline{\lambda} \le 1. \end{cases} \tag{8}$$

The coefficients α are obtained using a sub-optimal iterative algorithm with a step parameter ρ, inspired from [4], that proceeds as follows:

1. $\underline{\alpha}$ is initially null.
2. $\underline{v} = \underline{S\alpha}$, $\tilde{\mathbf{v}} = \underline{\tilde{S}\alpha}$ and σ_v^2 (as defined by (1)) are computed.
3. If $\sigma_v^2 < 1$, the waveform $\underline{\tilde{s}}_m^q$ with $m \neq k_l$ which minimizes $(\underline{\tilde{x}} + \underline{\tilde{S}\alpha})^t (\underline{\tilde{s}}_{k_l}^{opt} - \underline{\tilde{s}}_m^p)$ is selected and $\underline{\alpha}$ is modified as follows: $\alpha_{k_l}^{opt} = \alpha_{k_l}^{opt} + \rho$, $\alpha_m^q = \alpha_m^q - \rho$, modifying the direction of the watermark from the "most menacing" to the "desirable" signal.
4. Steps 2 and 3 are repeated until $\sigma_v^2 \ge 1$. The coefficients $\underline{\alpha}$ are finally sightly modified to ensure $\sigma_v^2 = 1$. They become: $\underline{\alpha}/||\underline{\alpha}||$.

4.2.2 Strategy EPS

The EPS strategy consists in choosing $\underline{\alpha}$ that satisfies the inaudibility constraint (1) and ensures a fixed value p of P'_e given by (6) with a maximum noise variance σ_n^2. This problem can be solved with an iterative algorithm. It deals with increasing progressively σ_n^2 until obtaining its maximum value. It processes as follows:

1. Given a certain value of σ_n^2, we aim at finding $\underline{\alpha}$ that ensures $P'_e = p$. Nevertheless this problem can have zero, one or several solutions depending on the signals configuration and the chosen value of σ_n^2. Therefore, we prefer choosing $\underline{\alpha}$ that minimize P'_e. It is related to the following optimization problem:

$$\begin{cases} \underline{\alpha} = \arg\min_{\underline{\lambda}} J_4(\underline{\lambda}), \\ J_4(\underline{\lambda}) = 1 - \prod\limits_{m=1..M, m\neq k_l, q=1..Q} \left(1 - Q\left(\frac{c_{k_l mq}(\underline{\lambda})}{\sigma_n}\right)\right), \\ c_{k_l mq}(\underline{\lambda}) = (\underline{\tilde{x}} + \underline{\tilde{S}}\,\underline{\lambda})^t \frac{\underline{\tilde{s}}_{k_l}^{opt} - \underline{\tilde{s}}_m^q}{||\underline{\tilde{s}}_{k_l}^{opt} - \underline{\tilde{s}}_m^q||}, \\ \frac{1}{N} \underline{\lambda}^t \underline{S}^t \underline{S}\,\underline{\lambda} \le 1. \end{cases} \tag{9}$$

   This problem is solved using a sequential quadratic programming method, proposed by Matlab©'s optimization toolbox. It yields a unique solution $\underline{\alpha}$ and the corresponding mimimum value of $P'_e = J_4(\underline{\alpha})$.
2. Since P'_e increases with σ_n^2, σ_n^2 is modified with regard to the previous value of $P'_e = J_4(\underline{\alpha})$. If $P'_e > p$, σ_n^2 is decreased, and if $P'_e < p$, σ_n^2 is increased, following a dichotomous progression.
3. Step 1 and 2 are repeated until P'_e approximates the expected value p or σ_n^2 becomes null.

At the end, in the case where $P'_e \approx p$, the coefficients $\underline{\alpha}$ that ensure a fixed error probability with a maximized noise variance are found. In the case where σ_n^2 is null, no watermarking signal permits a robust transmission of symbol k_l as defined by the EPS strategy. This last case happens in practice in almost 1% of the total number of signals configuration. EPS is then replaced with RPS.

5. EXPERIMENTAL RESULTS

5.1 Test plan

System performances are evaluated through three criteria: (1) the perceptual quality, (2) BERs with respect to transmission rate R for various channel perturbations and (3) the computational cost. We have used a set of 20 audio signals, sampled at $F_e = 44.1$ kHz and watermarked with L binary digits to process L/R seconds of signal. We have decided to transmit $L = 100000$ binary digits to achieve a compromise between accuracy of BER (lower than 10^{-3}) and processing time. Thus BERs lower than 10^{-3} are not reliable.

A various range of perturbations has been considered. We use the automated evaluation tool (and its default parameters) proposed in [2] from which we only select the perturbations adapted to broadcasting application field. We also consider MPEG compression, performed by an MPEG 1 Layer 3 digital encoder and white noise adding with various SNR.

Systems are implemented using Matlab©version 6.1. The machine used for simulations has the following characteristics: Pentium 4, 1.80 GHz, 512MB RAM.

Performances of the two embedding strategies RPS and EPS are studied and compared to those of the Non-Informed Strategy, presented in section 2.1 and further denoted by NIS. The used codebook is structured in $M = 4$ orthogonal sub-codebooks, each containing $P = 8$ biorthogonal waveforms, with a cut-off frequency of 6 kHz. The expected error probability in EPS is fixed at $p = 10^{-3}$.

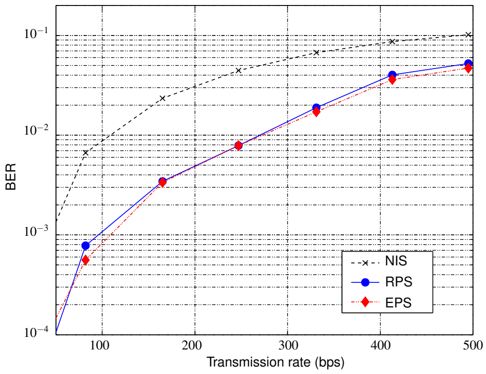

Figure 2: BER vs transmission rate when the channel is free from perturbation.

5.2 Results

Listening tests were performed to evaluate the perceptual distortion introduced by the watermark. These tests are inspired from the UIT-R BS 1116 recommandation and were performed on a set of five listeners. Test results confirm for both informed strategies that the watermark is "perceptible but not irritating" as defined on the perceptual grade .

BERs were measured for different binary transmission rate $R = \frac{\log_2(M)F_e}{N}$ and perturbations. Figure 2 presents the BERs of the three embedding strategies when the channel is free from perturbation. It confirms the efficiency of informed embedding strategies since the BERs obtained with RPS and EPS are divided by almost 10 compared with BERs with NIS when $R < 200$ bps. Transmission reliability with the two informed embedding strategies are quite similar. Indeed, when $R < 300$ bps, BERs with RPS and EPS are the same given the reliability of BERs measures. When $R > 300$ bps, EPS is slightly more efficient than RPS. Error correction code could be introduced to take benefits from this slight improvement. It could yield better BERs at low transmission rates.

Robustness of the two informed embedding strategies to various channel perturbations has also been evaluated for $R = 82$ bps. Figure 3 presents the obtained results. It shows that the two embedding strategies offer the same robustness to channel perturbations. The most severe perturbations are MPEG compression at 64 kbps and echo adding. With these perturbations, BERs are multiplied by 2 compared with those obtained for a channel free from perturbation.

Finally, computational costs of the embedding and the reception processes were measured as the ratio between the simulation time (in seconds) and the duration of the processed signals (in seconds) when $R = 82$ bps. The obtained ratios are detailed in table 1. A real-time embedding process can be achieved with NIS but not with RPS and EPS due to the use of iterative optimization algorithms. Moreover EPS computational cost is multiplied by 5 compared to RPS one due to the complexity of the EPS optimization algorithm. Computational costs of the reception process are quite similar for the three strategies even though EPS and RPS are slighly more costly than NIS. Indeed the computational cost is related to the filtering of each codebook waveform and the number of codebook waveforms used with EPS and RPS is 8 times greater than that of NIS.

To sum up, the RPS strategy is much more efficient than the EPS strategy to achieve a robust transmission with a 10^{-3} reliability, since RPS yields the same BERs as EPS and its computational cost is lower than EPS's one.

6. SUMMARY AND CONCLUSIONS

In this paper, we present and compare two informed embedding strategies adapted to an audio data hiding system designed for broadcast application. These strategies can be summed up as follows: a local copy of the receiver scheme at the embedder is exploited to estimate signals taking part at the receiver stage. Then the watermarking signal is specifically chosen to limit the local perceptual distortion and to reach a reliable transmission of the embedded information with maximized robustness to additive channel perturbations. In the first strategy, robustness is expressed with the input of the correlator used in the detection process. In the second one, robustness relies on the probability of erroneous decision. Experimental results show that the first strategy is much more efficient than the second one when transmission rates are lower than 300 bps, since it yields the same reliability of transmission at lower computational cost than the second one. But the second strategy seems to be promising since it yields a sightly better transmission reliability for transmission rate higher than 300 bps. Error correction code should be introduced to improve transmission reliability and confirm the efficiency of the second strategy at high transmission rates.

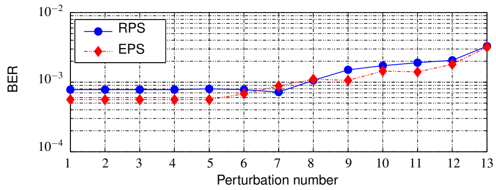

Figure 3: BERs with respect to channel perturbations when $R = 82$ bps. Perturbations are numbered as follows: 1: any, 2: resampling, 3: white noise adding with SNR=50 dB, 4: white noise adding with SNR=60 dB, 5: compressor, 6: white noise adding with SNR=40 dB, 7: MPEG at 96 kbps, 8: low-pass filtering, 9: high-pass filtering, 10: loudness change, 11: requantization, 12: MPEG at 64 kps and 13: echo adding.

Computational Cost	NIS	RPS	EPS
Embedding process	0.5	10.5	51.5
Reception process	2.5	3.8	3.8

Table 1: Computational cost of three embedding strategies.

REFERENCES

[1] L. Gomes, P. Cano, E. Gomez, M. Bonnet and E. Battle, "Audio Watermarking and Fingerprinting: For Which Applications?," in *Journal of New Music Research*, vol. 32, pp. 65-81, 2003.

[2] M. Steinebach, F. Petitcolas, F. Raynal, J. Dittmann, C. Fontaine, S. Seibel, N. Fates and L.C. Ferri, "StirMark benchmark: audio watermarking attacks", in *Proc. Int. Conf. on Information Technology: Coding and Computing*, pp. 49-54, 2001.

[3] C. Baras, N. Moreau and P. Dymarski, "An audio watermarking scheme based on an embedding strategy with maximized robustness to perturbations," in *Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 17, pp. 357-360, 2004.

[4] M. Miller, G. Doerr and I. Cox, "Applying informed coding and embedding to design a robust, high capacity watermark," in *IEEE Trans. on Image Procesing*, vol. 13, pp. 792-807, 2004.

[5] H. Malvar and D. Florencio , "Improved spread spectrum: a new modulation technique for robust watermarking," in *IEEE Trans. on Image Procesing*, vol. 13, pp. 792-807, 2004.

[6] J.G. Proakis, *Digital communications*, McGraw-Hill, 4th edition, 2001.

[7] M. Costa, " Writing on dirty paper," *IEEE Trans. on Information Theory*, vol. 29, pp. 439-441, 1983.