IMPROVED DIGITAL FM TRANSMISSION WITH LIMITER-DISCRIMINATOR DETECTION

M. LAMNABHI - P. SALEMBIER

Laboratoires d'Electronique et de Physique Appliquée*
3, avenue Descartes, BP. 15 - 94451 Limeil-Brévannes, Cedex - France
Tel. 45.69.96.10

ABSTRACT

In this paper, three techniques for improving the performance of discriminator detection of digital FM are presented : in the first one use is made of detection theory to develop an efficient scheme for detecting the occurence of a click and for correcting its effects in the demodulated signal. The second technique is used to combat the triangular noise. An improvement of the signal to noise ratio in the decoder is achieved by filtering the demodulated signal. The interest of the proposed method lies in the use of partial response encoding technique in the receiver to perform this filtering operation. Finally, the third technique counteracts non-linear distortion by taking advantage of maximum likelihood sequence estimation. It leads to a reduced state Viterbi decoder. At the end of this paper, a specific example is given where these three methods are successfully applied at the same time.

INTRODUCTION

Narrowband digital FM achieves both power and spectral efficiency. Power efficiency is due to the constant envelop property of the transmitted signal which allows the use of amplifiers near their saturation point. Spectral efficiency can be obtained by filtering the baseband signal before modulation. This approach has formed the basis of many studies in the past such as duobinary FM [1], GMSK, GTFM [2] and more generally CPM [3].

Limiter-Discriminator detection of digital FM is often used because of its simplicity and its robustness in the presence of fast fading or frequency shifts (see [1], [2] and [4 - 6]). However the noise at the discriminator output is not white nor Gaussian [8], [9]. At the receiver side, the transmitted signal may be corrupted by three types of disturbances :

1) the "triangular noise" the power spectral density of which is proportional to f^2,
2) an impulse noise called the click noise, for low carrier to noise ratio (CNR),
3) non-linear distortion due to the narrow bandwidth of the receiver IF filter.

In this paper, we present three baseband processing techniques which can be used to improve the receiver performance in the presence of each disturbance. They may be used separately or combined in the receiver. In the first one use is made of detection theory to develop an efficient scheme for detecting the occurence of a click and for correcting its effects in the demodulated signal. The second technique is used to combat the triangular noise. An improvement of the signal to noise ratio (SNR) in the receiver is achieved by filtering the demodulated signal. The interest of the proposed method lies in the use of partial response encoding technique in the receiver to perform this filtering operation. Finally the third technique counteracts non-linear distortion by taking advantage of maximum likelihood sequence estimation. It leads to a reduced state Viterbi decoder. At the end of this paper, a specific example is given where these three methods are successfully applied at the same time.

DIGITAL FM SYSTEM AND TRANSMISSION MODEL

The digital FM system under study is shown in figure 1. The source symbols are binary $\{a_k = \pm 1\}$, equiprobable and independent. The transmitter consists of a correlative level encoder [7], a shaping filter (transfer function $H_{SH}(f)$) used to limit the signal bandwidth, and a frequency modulator. The transmission is corrupted by an additive white Gaussian noise n(t) with the one-sided power spectral density N_0. At the input of the receiver, the signal x(t) can be written as :

$$x(t)= \sqrt{2S} \cos \left[2\pi f_0 t+2\pi f_d \int_{-\infty}^{t} b(\tau)d\tau\right]+ n(t) \quad (1)$$

where S is the signal power, f_0 is the carrier frequency, b(t) is the modulating signal and f_d is the peak frequency devia-

* A member of the Philips Research Organization.

tion. If the bit rate is 1/T, the modulation index is defined as $h = 2f_dT$.

The receiver consists of an IF filter (the transfer function of which is denoted by $H_{IF}(f)$), a limiter, a discriminator, a low-pass filter (transfer function : $H_{LP}(f)$) which limits the bandwidth to that occupied by the modulating signal, and a decoder. At the output of the low-pass filter, the signal z(t) can be written as :

$$z(t) = b'(t) + n_t(t) + n_c(t) \qquad (2)$$

In the above expression b'(t) denotes the digital signal, $n_t(t)$ the triangular noise, and $n_c(t)$ the click noise.

The difference between b'(t) and b(t) represents the non-linear distortion caused by the use of a reduced bandwidth IF filter.

This signal b'(t) is first corrupted by a noise component $n_t(t)$ called the triangular noise. Assuming a linear approximation of the demodulation process (valid for high CNR), the triangular noise results from the filtering process of a white Gaussian noise n(t), by the IF filter and the discriminator. Its power spectral density is then :

$$n_t(f) = \frac{|H_{IF}(f)|^2}{f_d^2} f^2 N_o \qquad (3)$$

It is well known [8-9] that, at a low carrier to noise ratio (CNR ≦ 10 dB) an impulse noise appears. This noise component is characterized by the occurence of spikes which are called "clicks". In the following, we assume that the clicks have a Gaussian temporal shape :

$$cl(t) = \alpha \exp\left(-\frac{t^2}{\beta}\right) \qquad (4)$$

In the next section, we show how clicks can be detected and corrected after demodulation.

CLICK DETECTION AND CORRECTION SYSTEM

Click detection can be performed using the approach suggested in [15]. The main difference here is that use is made of the knowledge of the transmitted signal. Let us assume that the demodulator output is the sum of three components:the digital signal b(t) (or b'(t) if non-linear distortion has to be taken into account), the triangular noise $n_t(t)$ and clicks cl(t) which have to be detected. Thus, the click detection problem involves detecting the presence of a specific signal cl(t) in a noisy signal. The detection noise comprises two components :

$$n_{det}(t) = b(t) + n_t(t).$$

This problem may be viewed as a two-hypotheses problem well-known in detection theory. The two hypotheses are :

$$\begin{cases} y(t) = cl(t) + n_{det}(t) & H_1 = \text{click occurence} \\ y(t) = n_{det}(t) & H_o = \text{no click.} \end{cases}$$

Among the existing criteria, the one which involves maximizing the detection SNR is used here. This maximization can be carried out by using a filter D(f). The detection signal power at the decision time to = 0, is given by :

$$S_{det} = \left[cl(t) * D(t)\Big|_{t_o=0}\right]^2 = \left[\int_{-\infty}^{+\infty} cl(f)\, D(f)\, df\right]^2 \qquad (5)$$

In the above expression, * is the convolution operator. cl(f) and D(f) denote respectively the Fourier transforms of cl(t) and D(t).

The detection noise is defined by

$$N_{det} = \int_{-\infty}^{+\infty} |D(f)|^2\, n_{det}(f)\, df \qquad (6)$$

where $n_{det}(f)$ is the power spectral density of $n_{det}(t)$.

Following Schwartz's inequality, the detection SNR can be upperbounded by :

$$\left(\frac{S}{N}\right)_{det} = \frac{\left[\int_{-\infty}^{+\infty} cl(f)\, D(f)\, df\right]^2}{\int_{-\infty}^{+\infty} |D(f)|^2\, n_{det}(f) df} \leq \int_{-\infty}^{+\infty} \frac{cl^2(f)}{n_{det}(f)} df \qquad (7)$$

the equality is achieved for $D(f) = \dfrac{cl(f)}{n_{det}(f)}$

The detection filter is then defined by :

$$D(f) = \frac{cl(f)}{B(f) + \dfrac{|H_{IF}(f)|^2}{f_d^2} f^2 N_o} \qquad (8)$$

Where B(f) denotes the power spectral density of the digital signal b(t).

Equation (7) shows that the detection SNR upperbound depends on the power spectral density of the digital signal b(t) through $n_{det}(f)$. It is interesting to note that the upperbound can be maximized by reducing the power spectral density of b(t) for low frequencies.

The filter output is then compared to two thresholds (one positive and the other negative). In order to take into account the time duration of the filter impulse response, several processings are made within a decision unit : each time a non-zero value is found after the thresholding operation, a temporal window of fixed length is considered. Inside this window, the highest absolute value defines the click position and polarity. Then the search of a new non-zero value, is carried out outside this window.

Once the occurance of a click is detected, the last step consists of correcting its effect in the demodulated signal. Figure 2 summarizes the entire process. Note that, in order to have as much information as possible, the detection is performed on the signal before the post-demodulation low-pass filter. The signal shape at different steps of the click detection and correction process, is shown in figure 3.

SIGNAL TO NOISE RATIO IMPROVEMENT ON THE RECEIVER SIDE

The aim of this section is to describe a technique which improves the signal-to-triangular-noise ratio before decoding in the receiver. As stated previously, a correlative level encoder is used on the transmitter side. The encoding function may be represented by a polynomial [7] :

$$P(D) = \sum_{i=o}^{N} P_i D^i \quad ; P_o \text{ and } P_n \neq 0 \tag{9}$$

where D denotes the delay operator.
The power of the digital signal before decoding is given by :

$$S = \int_{-\infty}^{+\infty} \sum_{i=o}^{N} \sum_{j=o}^{N} P_i\, P_j \cos\left[2\pi(i-j)fT\right] \times \left|H_{SH}(f) H_{LP}(f)\right|^2 df \tag{10}$$

And the triangular noise power is :

$$N = \frac{No}{f_d^2} \int_{-\infty}^{+\infty} \left|H_{IF}(f)\right|^2 f^2 \left|H_{LP}(f)\right|^2 df \tag{11}$$

This defines the original SNR. When the power spectral densities of the signal and the noise are not proportional, it can be shown that the SNR can be improved by the use of a filtering process after demodulation. Of course, this filter will introduce intersymbol interference (ISI) in the transmitted signal. In order to actually improve the decoding performance, this ISI should be controlled and removed.

To achieve this filtering operation, we propose the use of correlative level "encoding" techniques in the receiver. The motivation for this choice is that the ISI introduced can be easily removed in the decoder. This decoder has then to deal with the cascaded codes composed of the "transmitter encoder" and the "receiver encoder". If the decoder is based on maximum likelihood sequence estimation [11] , this ISI cancellation can generally be carried out with only a small penalty characterized by a decrease of the minimum distance of the code. Thus, the important criterion is the SNR gain minus the loss in the minimum distance (d_{min}).

As the triangular noise power is an increasing function of the frequency, encoding schemes similar to duobinary, which have a low-pass filtering action, seems to be very suitable for improving the SNR. As an example, we examine the case where the "encoding" scheme in the receiver is defined by the following polynomial :

$$P(D) = (1 + D)^R \tag{12}$$

Equation (12) corresponds to R cascaded duobinary encoders in the receiver. In this case, the signal power is :

$$S_R = \int_{-\infty}^{+\infty} \sum_{i=o}^{N} \sum_{j=o}^{N} PiPj \cos\left[2\pi(i-j)fT\right] \cos^{2R}(\pi fT) \left|H_{SH}(f) \cdot H_{LP}(f)\right|^2 df \tag{13}$$

and the noise power becomes :

$$N_R = \frac{No}{fd^2} \int_{-\infty}^{+\infty} \left|H_{IF}(f)\right|^2 f^2 \cos^{2R}(\pi fT) \left|H_{LP}(f)\right|^2 df \tag{14}$$

The gain is defined by

$$G_{dB} = \frac{S_R}{N_R} - \frac{S}{N} - \Delta dmin \tag{15}$$

Figure 4 shows this gain for various codes at the transmitter end (binary, duobinary (1+D), TFM $(1+D)^2$, modified duobinary $(1-D^2)$) and for several cascaded duobinary encoding schemes on the receiver side. For these computations, H_{SH} and H_{LP} are assumed to be ideal rectangular low-pass filters with cut-off frequency 1/2T. And the IF filter has a Gaussian shape defined by :

$$H_{IF}(f) = \exp\left[\frac{-\pi f^2}{2B^2}\right] \tag{16}$$

Where B is chosen so that BT = 1 ; this shape is often used in practice [1], [4], [9]. It can be seen that with a simple duobinary encoding, this gain ranges from 3.1 dB to 4.1 dB (for TFM). In fact, the gain in bit error rate performance achieved after decoding is reduced by the presence, for example, of a non-white noise and error propagation due to the increase of the decoder complexity.

NON-LINEAR DISTORTION CANCELLATION

Non-linear distortion affecting the baseband signal results from the IF band limitation of the FM signal. There are several ways of counteracting this non-linear distortion [12-14]. We are interested here in techniques based on maximum likelihood estimation of the distorted sequences. As it has been seen previously, the transmission system under study includes partial response signaling. We assume now that the decoder at the receiver end is based on the Viterbi algorithm. The aim of this section is to show how the decoding algorithm can be modified in order to take into account not only the correlative level encoding, but also the non-linear distortion. The method described allows the decoder complexity to be increased gradually according to the amount of non linear distortion which has to be cancelled.

From the theoretical point of view, at the decoder input, the received symbol z_k can be expressed as a non-linear function of the source symbols $\{a_{k-i}\}$ $i \geq 0$, corrupted by a noise n_k :

$$z_k = f\,(a_k,\, a_{k-1},\ldots) + n_k \qquad (17)$$

The function f reflects the correlative level encoding, and the channel distortion as well.

The decoder complexity is directly connected to the length of the channel response f. In order to reduce this complexity, the channel response can be approximated by a function f_L which takes into account only L symbols a_k. The choice of this function f_L is difficult, as the decoding process is based on the Euclidean distance between sequences $\{\ldots,\, z_k,\, z_{k+1},\ldots\}$. To avoid this difficulty, another approach will be adopted: the decoder is matched to the channel response f_L and not to the real channel response f. In this case, it can be considered that the decoder receives a sequence with ISI. Our approach consists of minimizing the ISI for each transmitted symbol :

$$\underset{f_L}{\mathrm{MIN}} \left(\sum_{\{a_{L+1},a_{L+2},\ldots\}} \left[f_L(a_k,\ldots,a_{k-2}) - f(a_k,a_{k-1},\ldots)\right]^2 \right) \qquad (18)$$

It is well known that the solution of this minimization problem is given by :

$$z_K^L = f_L(a_k,\ldots,a_{k-1}) = \underset{\Omega}{E}\,(f(a_k,a_{k-1},\ldots)) \qquad (19)$$

In the above expression, E denotes the mathematical expectation with respect to the set $\Omega = \{a_{k-i}/i > L\}$.

Since a Viterbi decoder is used, it is necessary to define the corresponding trellis : the states at time k are $(a_{k-1},\, \ldots,\, a_{k-L})$ and the transitions between states are defined by the set of z_k. In this way, it is possible to gradually increase the receiver complexity to approach the real channel response.

APPLICATION

The aim of this section is to show how the techniques described above can be applied. To this end, a specific example is presented where the three techniques are used at the same time, and the relative improvement given by each technique will be discussed.

On the transmitter side, the selected code has spectral nulls at zero and 1/2T. The spectral null at 1/2T is used to improve the spectral efficiency, whereas the spectral null at zero is used to maximize the click detection SNR. (It has to be pointed out that this last property implies the presence of spectral lines in the modulated spectrum 10 , of course these spectral lines may or may not be a disadvantage depending on the application). The simplest partial response code which has these two properties is defined by the following polynomial $P(D) = 1 - D^2$, and is known as modified duobinary. Its power spectral density has a sine wave shape.

At the receiver end, the transfer function of the IF filter has a Gaussian shape with BT = 1 (see equation (16)). Following (8) the transfer function of the click detection filter is given by :

$$D(f) = \frac{cl(f)}{\mathrm{Sin}^2(2\pi fT) + \alpha \exp\left[\frac{-\pi f^2}{B^2}\right] f^2} \qquad (20)$$

where is a factor which depends on E_b/N_o.

Of course, a linear approximation of the noise power spectral density is used in this formula. If a more accurate estimation is required, classical spectral estimation methods can be used.

It should be noted that, the ideal detection filter has a pole at zero which makes it unrealizable. The practical filter is then an approximation of the ideal one. Figure 5 presents the transfer function of this filter, as well as the detection noise power spectral density. The "simulation results" were computed using a modified periodogramm method.

As discussed in section IV, the SNR before decoding can be improved by using partial response encoding techniques in the receiver. Figure 4 shows that the use of duobinary encoding on the received data provides a gain of 3.1 dB. The decoder complexity increase is moderate. Indeed the Viterbi decoder has to be matched to the following encoding scheme :

$$\underset{\text{Transmitter code}}{(1 - D^2)} \quad \times \quad \underset{\text{Receiver code}}{(1 + D)} \qquad (21)$$

This corresponds to an eight states trellis, and a five level constellation.

To improve the performance in the presence of non-linear distortion, the Viterbi decoder can be matched to both the encoding scheme (equation 21) and to the channel distortion as described in section V. In the following, we show how this can be done without any complexity increase. The Viterbi decoder complexity is proportional to the number of states. An eight states trellis is defined by keeping only four symbols in the channel response. Following equation (19) we have :

$$Z_k^3 = \underset{\Omega}{E} \; f(a_k, a_{k-1}, \ldots) \; ; \Omega = \{a_{k-i} / i > 3\} \quad (22)$$

Therefore, the trellis has the same structure as the one defined by equation (21). The set of symbols corresponding to the transitions in the trellis defines here a 7 level constellation. A comparison between this new trellis and the one defined by equation (21) is presented in figure 6.

The performance of this transmission system have been investigated using computer simulations. The results are presented in figure 7. First, the bit error rate (BER) as a function of the E_b/N_o has been estimated without click correction, duobinary encoding in the receiver or distortion cancellation, and the modulation index which gives the best performance has been found : h = 0.8. Figure 7 shows that, in this case, 11.5 dB of E_b/N_o are required for a BER of 10^{-4}. Then the performance has been investigated either with a click correction or with a duobinary encoding in the receiver. For BER under investigation here (10^{-3} to 10^{-4}), both techniques provide approximately the same gain (between 0.6 and 1 dB each). As a matter of fact, errors are caused as much by clicks as by triangular noise. In order to get a high performance gain, both techniques can be applied jointly. Then figure 7 shows that there is a gain of 2.2 dB at a BER of 10^{-4}. The last curve presents the performance when click correction, duobinary encoding in the receiver and non-linear distortion cancellation are used at the jointly. The gain equals then 2.5 dB (BER = 10^{-4}). The contribution of non-linear distortion cancellation is rather small, but it is performed without any complexity increase in the decoder.

Finally, it is interesting to compare this last scheme with binary FM. It is well known that in the case of discrimator detection and a Gaussian shaped IF filter with BT = 1, binary FM with a modulation index h = 0.7 gives very good performance [1]. Comparison is made in figure 8 where it can be seen that binary FM requires 1.7 dB more of E_b/N_o to achieve the same performance.

CONCLUSION

In this paper, three techniques for improving the performance of discriminator detection of digital FM, have been presented:

- The first one consists of click detection and correction. The detection process makes use of a filter which maximizes the detection SNR. This filter and the system performance depend on the power spectral density of the digital signal.
- The second one improves the signal-to-triangular-noise ratio at the input of the decoder. It is based on the use of a correlative level encoder in the receiver. The decoder has then to be matched to the cascaded code (product of the transmitter and receiver codes).
- The last technique improves the performance of the decoder in the presence of non-linear distortion caused by the band limitation of the FM signal. The technique presented involves modifying the Viterbi decoder so that it takes into account not only the encoding scheme but also the non-linear distortion. It has been shown how this modification can be made using a reduced-state Viterbi decoder. Using this approach, the decoder complexity can be gradually varied in order to cancel more and more non-linear distortion.

The gain provided by those techniques depends on the particular application. As an example, we presented a transmission scheme (with IF bandwidth BT = 1) where these three techniques are applied at the same time. This gives a gain of 2.5 dB in SNR for a BER of 10^{-4}. With a modulation index h = 0.8, this improved scheme saves up to 1.7 dB in SNR compared to binary FM with h = 0.7.

Acknowledgement

The authors would like to express their sincere thanks to O. GUERIF who helped them to investigate the BER performance by computer simulation.

REFERENCES

[1] T.T. Tjhung, K.K. Yeo, P.H. Wittke, "Effects of pulse shaping and soft decisions on the performance of digital FM with discriminator detection", IEEE Trans. Commun., vol. COM-34, pp. 1116-1122, Nov. 86.

[2] K.S. Chung, "Generalized Tamed Frequency modulation and its application for mobile radio communication", IEEE J. Select. Areas Commun., vol. SAC-2, pp. 487-497, July 84.

[3] T. Aulin, C.E.W. Sundberg, "Continuous phase modulation", IEEE Trans. Commun., vol. COM-29, pp. 196-225, March 81.

[4] T.T. Tjhung, K.K. Yeo, P.H. Wittke, "Error performance analysis fornarrow-band duobinary FM with discriminator detection", IEEE Trans.Commun., vol. COM-33, pp. 399-408, May 85.

[5] M.K. Simon, C.C. Wang, "Differential versus limiter-discriminator detection of narrow-band FM", IEEE Trans. Commun., vol. COM-31, pp. 1227-1234, Nov. 83.

[6] A. Svensson, C.E.W. Sundberg, "On error probability for several types of non coherent detection of CPM", GLOBECOM'84, pp. 1-6, Atlanta, Dec. 1984

[7] P. Kabal, S. Pasupathy, "Partial response signaling", IEEE Trans. commun., vol. COM-23, pp. 921-934, Sept. 75.

[8] S.O. Rice, "Noise in FM receivers",Time series analysis, M. Rosenblatt, ed., New York, Wiley 1983.

[9] R.F. Pawula, "On the theory of error rates for narrow-band digital FM", IEEE Trans. Commun., vol. COM-29, pp. 1634-1643, Nov. 81.

[10] G.S. Deshpande, P.H. Wittke, "Correlative encoded digital FM", IEEE Trans. Commun., vol. COM-29, pp. 156-161, Feb. 81.

[11] G.D. Forney, "Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference", IEEE Trans. on Inform. Theory, vol. IT-18, pp. 363-378, May 72.

[12] W.V. Etten, L. Kurz, "Maximum likelihood receivers for data sequences transmitted over non linear channels", A.E.U. Vol. 34, pp. 216-223, 1988

[13] S. BENEDETTO, E. BIGLIERI, "Non linear equalization of digital satellite channels", IEEE J. Select. Areas Commun., vol. SAC-1, pp. 57-62, Jan. 83.

[14] E. Biglieri, S. Barberis, M. Catena, "Analysis and compensation of non linearities in digital transmission systems", IEEE J. Select. Areas Commun., vol. SAC-6, pp. 42-51, Jan. 88.

[15] M.J. Malone, "FM threshold extension without feedback", Proc. IEEE, pp. 200-201, Feb. 68.

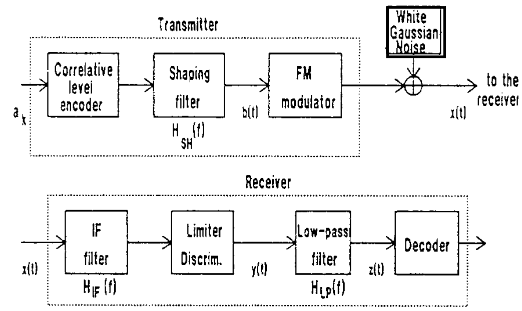

Figure 1 : Digital FM system

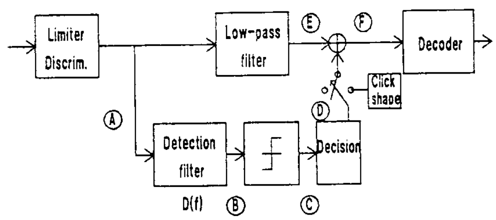

Figure 2 : Click detection and correction system

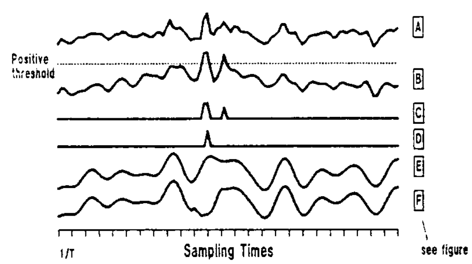

Figure 3 : Click detection and correction process

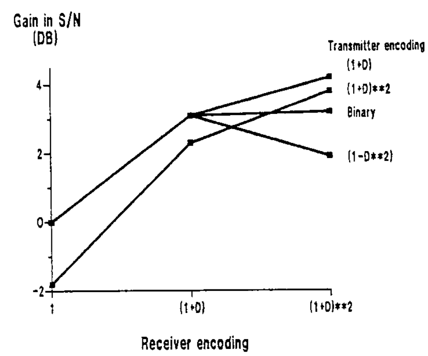

Figure 4 : Signal to Noise Ratio gain in the receiver

36.5.6.

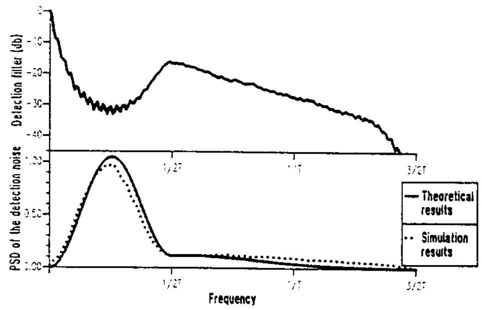

Figure 5 : Click detection filter

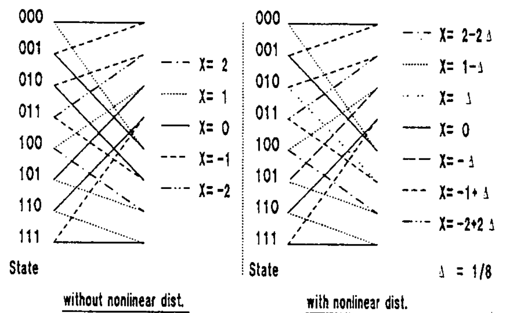

Figure 6 : Modification of the Viterbi decoder to cancel nonlinear distortion

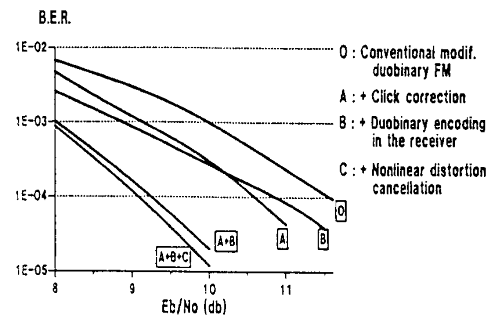

Figure 7 : Performances for modified duobinary FM

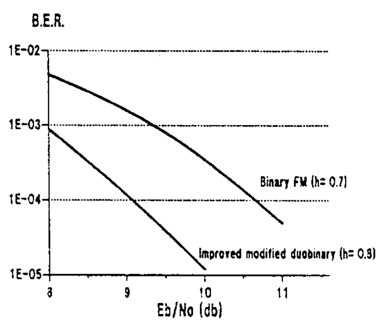

Figure 8 : Performances of improved modified duobinary FM and binary FM