Limits of Error Correction Coding in Video Watermarking

M. Ambroze[a], M. Tomlinson[a], C. Serdean[b] and G. Wade[c]

[a] Department of Communications and Electronic Engineering, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK

[b] Electronics Department, University of Kent at Medway, Chatham ME5 9UK, UK

[c] School of Electrical Engineering and Computer Science, University of Newcastle, NSW2308, Australia

ABSTRACT

The paper discusses the limits of error correction coding for spread spectrum-based video watermarking. The error correction code has as input the watermark data bits and as output the values which will be scaled and used to modify the video pixels (transform coefficients). The data rate of the watermark can increase only at the expense of increasing code rate. Theoretically, the scheme is seen as a communication channel with Gaussian additive noise interference. Shannon's (ideal) spherical codes are used as the error correcting code to calculate the minimum signal to noise ratio (SNR) necessary for a coding scheme with a given block length to achieve a given error probability. This limit is different from Shannon's asymptotic limit, which is valid for infinite block lengths and zero error probability. In practice, in order to verify the Gaussian channel assumption, the error correction code is a concatenation of codes, of which the innermost is a repetition code. Several practical codes of different length and rates, such as turbo codes and BCH codes are investigated and their performance compared to that of the ideal code of the same size. The compromise block length/code rate is investigated for several marking schemes and attacks.

Keywords: Video watermarking, Turbo coding, Shannon limit

1. INTRODUCTION

Video watermarking can be seen as a communication channel and error correction codes can be used to improve the error rate.[1,2] A thorough presentation of coding schemes is given in,[2] although their utility is slightly hidden in the mathematical detail. The error correction codes are chosen from a large collection, including convolutional codes with soft decision decoding,[3] block codes, such as BCH codes with hard decision decoding[4] (which inevitably lose $1-2$dB of their maximum likelihood performance due to hard decision) and, recently, turbo codes[5,6] and related schemes, due to their large coding gain. The problem of applying coding after a spread spectrum scheme (equivalent with concatenation with a repetition code) or directly on the video samples (in spatial or transform domain) is approached.[2,7] The main advantage of the latter method is the block size, the disadvantage being the non-Gaussian statistics of the noise, which in this case is the video itself. What is the best method? What is the best code? What is the limit of coding for the video watermarking channel? It is known that coding cannot give unlimited gains: there is an ultimate limit to what can be achieved with error correction, and that is the Shannon limit.[8] This limit depends on the block error rate that is needed, the block length and the rate of the code considered. In this paper, we see the video watermark as a hidden communications channel. We use filtering, interleaving, and spread spectrum to justify an AWGN assumption. A single error correction code is used to protect the watermark. If the information block length is increased, its rate increases such that the code length is a constant dependent on the available number of pixels. We calculate the minimum sample-level signal to noise ratio SNR_0 necessary to obtain a probability of block error of $P_w = 10^{-8}$ for different block lengths (and their corresponding rates) using the Shannon limit for finite block lengths.[9] We then consider several coding schemes (turbo codes, BCH codes, and short binary codes) and compare their performance with the limit. The SNR that is available in the channel is also determined for several simple schemes, such as the direct insertion scheme in the spatial domain and the scheme using a highpass filter to vary the amplitude of the watermark. The reduction in SNR_0 due to MPEG2 compression attacks is estimated. Based on the measured SNR_0, the capacity of the watermarking channel is determined for the given channel model.

2. OUTLINE OF THE PAPER

The paper is structured as follows: Section 3 presents Shannon's limit and discusses the available coding gain as a function of block length, code rate and block error probability. Section 4 describes the watermarking channel model and its parameters. The assumption of Gaussian noise is justified and the method to obtain such a channel is described. In Section 5 several practical error correction codes are presented, such as turbo codes and short codes, and their performance is compared to Shannon's limit. In Section 6 the available sample-level SNR_0 for several watermarking schemes is determined, with no attacks and MPEG2 compression attacks. The variation of SNR_0 with the MPEG2 bit rate is determined. Based on the available SNR_0 and the minimum SNR curves from Shannon limit and the error correction codes, the channel capacity is determined. Practical codes are chosen to approach this limit and their choice is justified by closeness to Shannon limit. A short discussion of "dirty paper" codes which reject the host signal interference is presented in Section 7. Section 8 presents the conclusions.

3. SHANNON'S LIMIT

Consider a code with coded block length n, information block length k and code rate $r = k/n$. There are several levels at which the Shannon limit is considered:

Infinite block length, zero rate. The ultimate, asymptotic level, which says that a code of infinite length and zero rate can achieve infinitely small error probability on an AWGN channel provided that the bit energy to noise ratio, $E_b/N_o > -1.6$dB. This is rather approximate for real codes, which have finite rate and length.

Infinite block length, nonzero rate. The limit for infinite block length and non-zero code rate r is $E_b/N_o = \frac{2^{2r}-1}{2r}$.

Finite block length, nonzero rate. The Shannon limit given the code rate r, the coded block length n and the probability of error P_w.

The Shannon limit for finite block length is calculated using the sphere packing bound,[9] which was originally derived by Shannon.[8] The probability of block error P_w is lower bounded by the probability of error of (ideal) spherical codes:

$$P_w \geq Q_n(\theta_{S_n}, A) \tag{1}$$

The computation proceeds by determining the solid angle θ_{S_n} from

$$\Omega_n(\theta_{S_n}) = \int_0^{\theta_{S_n}} \frac{n-1}{n} \frac{\mathbf{\Gamma}\left(\frac{n}{2}+1\right)}{\mathbf{\Gamma}\left(\frac{n+1}{2}\right)\sqrt{\pi}} \sin(\phi)^{n-2} d\phi = \frac{1}{2^{rn}} \tag{2}$$

which is then used in

$$Q_n(\theta_{S_n}, A) = \int_{\theta_{S_n}}^{\pi} \frac{(n-1)\sin\phi^{n-2}}{2^{n/2}\sqrt{\pi}\Gamma\left(\frac{n+1}{2}\right)} \int_0^{\infty} s^{n-1} e^{-(s^2+nA^2-2s\sqrt{n}A\cos\phi)/2} ds d\phi \tag{3}$$

where $A = \sqrt{2rE_b/N_o}$. From Equations (3) and (1), a minimum value of E_b/N_o at which a probability P_w of block error can be attained is calculated, and we will denote it as $sh(n, k, P_w)$ in the remainder of the paper.

A simple approximation for large blocks, is[9]:

$$k(\Delta SNR_{dB})^2 \geq (27 \pm 2) P_{w,dB}^{-1} \tag{4}$$

where ΔSNR_{dB} is the difference between the limit corresponding to a block length k and the asymptotic Shannon limit for the given code rate. The closeness to Shannon limit of practical codes depends on the P_w

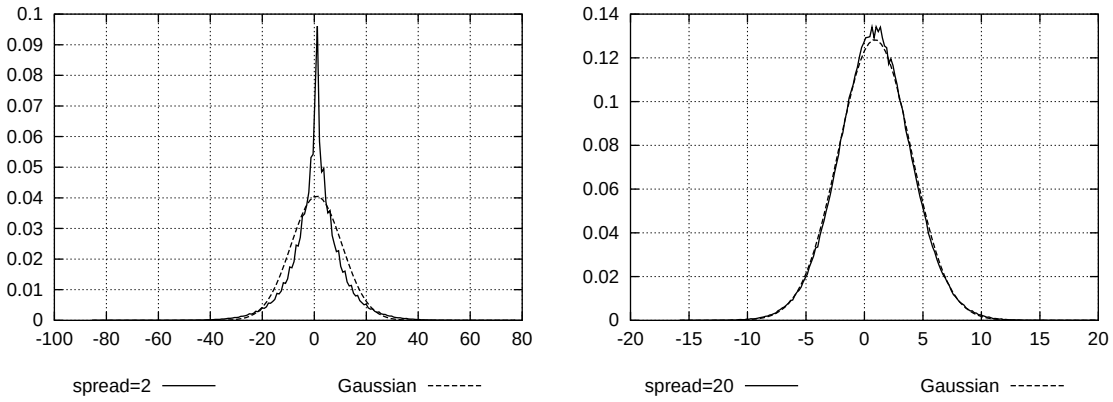

Figure 1. Gaussian distribution of video interference due to spreading.

considered. As P_w is reduced, it becomes more difficult to obtain codes which are close to Shannon's limit. The gain that can be obtained by coding increases with P_w, as shown in table 1. The gain that could be obtained by a real, finite length code can be estimated more precisely by using equation 4. Thus, for $k = 1000$ and $P_w = 10^{-8}$ we have $P^{-1}_{w,dB} = -10 \log_{10} P_w = 80$ and thus $\Delta SNR_{dB} \geq \sqrt{2} \approx 1.4dB$. This means that, for a block size $k = 1000$ and a $P_w = 10^{-8}$, a code cannot possibly get closer than about 1.4dB from the corresponding asymptotic Shannon limit, which for low rate codes can be approximated to -1.6dB. Since the SNR necessary to reach $P_w = 10^{-8}$ for uncoded is $SNR_u = 12$dB, the gain achievable by coding with a block length $k = 1000$ is $SNR_u + 1.6 - 1.4 = 12.2$dB. The gains for other values of P_w are calculated in a similar way and presented in table 1.

P_w	10^{-4}	10^{-8}	10^{-16}	10^{-24}
$k = 1000$	9	12.2	14.88	16.26
$k = \infty$	10	13.6	16.88	18.76

Table 1. Possible coding gains [dB] for low rate codes for a given P_w. Equation 4 is used for $k = 1000$, and a value of -1.6dB is used for $k = \infty$.

4. THE WATERMARKING CHANNEL

The basic signal to noise ratio is given by the ratio of the energy of the watermark and the energy of the host signal. A problem in investigating the SNR is the fact that the distribution of the interference values is not Gaussian. As shown in Fig. 1, a spreading factor higher than 20 produces a distribution which is close to Gaussian. By employing an interleaver (scrambler) in order to break the correlation of the image coefficients, a SNR value which is close to the theoretical value based on the energy of the two signals is obtained. This means that even in case of MPEG2 compression attacks (see section 6), the SNR can be calculated for a larger spreading factor and then extrapolated to the basic SNR_0 (the sample-level watermark/video energy ratio). As long as the code rates considered are low enough (less than 1/20), this value can then be used to determine the performance of the scheme.

A requirement of practical schemes is that the watermark is contained in a short "video segment". This requirement constrains the watermark rate and determines the minimum SNR_0 at which a given P_w can be obtained. If we consider a video segment containing 25 frames (1s in PAL), the maximum number of coefficients is the number of pixels n. A code with a given block length k and rate $r = k/n$ will have an energy per bit to noise ratio given by

$$2rE_b/N_o = SNR_0 \tag{5}$$

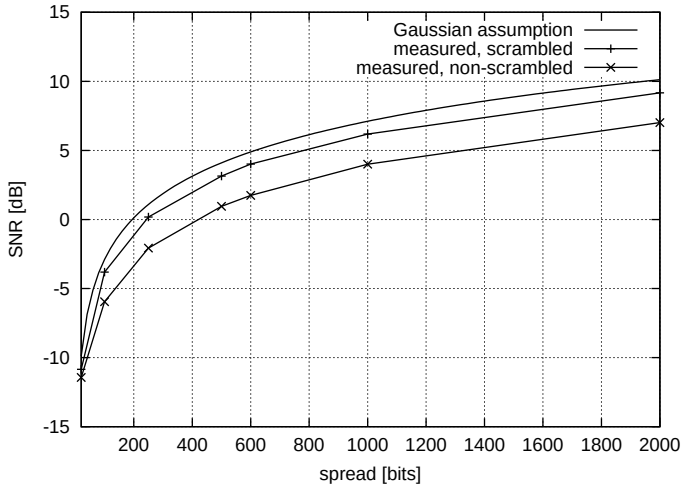

Figure 2. Example SNR dependence on block length.

The code will work at a given P_w as long as

$$E_b/N_o \geq E_b/N_o(code, P_w) \tag{6}$$

where $E_b/N_o(code, P_w)$ denotes the bit energy to noise ratio for which a given code has a probability of block error P_w. By combining equations 5 and 6, we have

$$SNR_{0,dB} \geq (E_b/N_o)_{dB}(code, P_w) + 10\log_{10}(2r) = (E_b/N_o)_{dB}(code, P_w) + 10\log_{10} 2 + 10\log_{10} k - 10\log_{10} n \tag{7}$$

The minimum value of $SNR_{0,dB}$ can be obtained by replacing $(E_b/N_o)_{dB}(code, P_w)$ with Shannon's limit for a given block size k, rate $r = k/n$ and probability of block error P_w:

$$(E_b/N_o)_{dB}(code, P_w) \geq sh(n, k, P_w) \tag{8}$$

In this paper we consider n fixed, corresponding to 1s in PAL, and determine the maximum block length k that can be used to obtain a probability of error P_w. In this case, the block length k is numerically equal to the watermark data rate in bits/s. Since the video has $n = 25 \cdot 720 \cdot 576 = 10368000$ pixels, the equations above become:

$$SNR_{0,dB} \geq -67dB + (E_n/N_o)_{dB}(code, P_w) + 10\log_{10} k \tag{9}$$

Equation 9 is plotted in Fig. 3 for ideal spherical codes (label "Shannon") and for $(E_n/N_o)_{dB}(code, P_w)$ replaced with the asymptotic value -1.6dB (label "Asymptote"). The block error probability is $P_w = 10^{-8}$. Note that the figure only plots the case of low rate codes, for which the asymptotic value is very close to -1.6dB and the Gaussian noise assumption holds for the watermarking scheme. Fig. 4 is a more detailed version of Fig. 3 for short/medium block lengths. It can be observed that for short block lengths there is a significant difference between the asymptotic and finite block length Shannon limits. As the block length increases, the limit for finite block length approaches the asymptotic limit. The difference for a block length of $k = 5000$ is less than 1dB. This difference is dependent on P_w and increases with decreasing P_w.

5. SHORT AND LONG CODES

As shown in Fig. 3, if the coded block length n is kept constant, the coding gain due to the increase of information block length k occurs at the expense of increased code rate. It is apparent that the coding gain reduces but does not compensate the loss due to increase in code rate r. In this situation, the most powerful scheme is the scheme that embeds a single bit. If the available SNR_0 is high enough, more than one bit can be

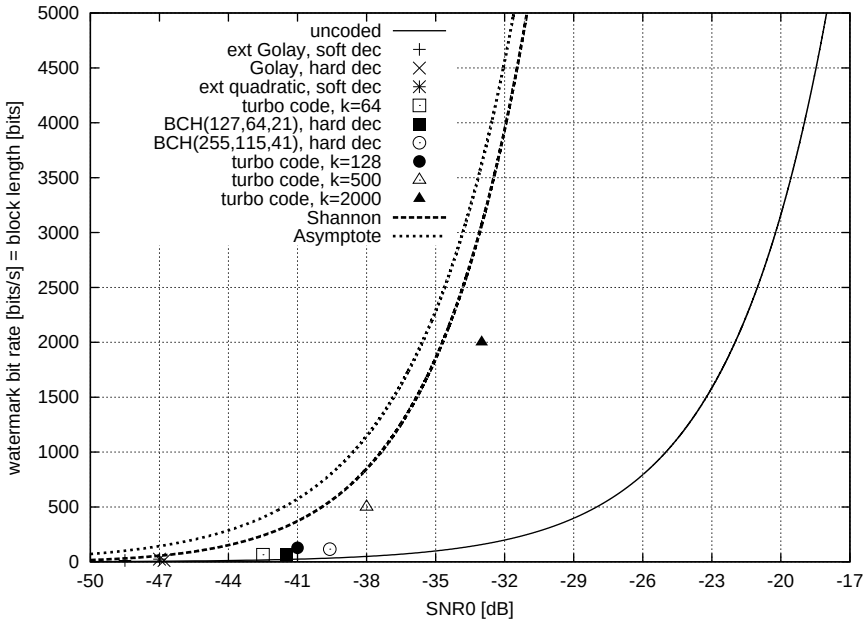

Figure 3. Shannon limit SNR_0 ranges for $P_w = 10^{-8}$.

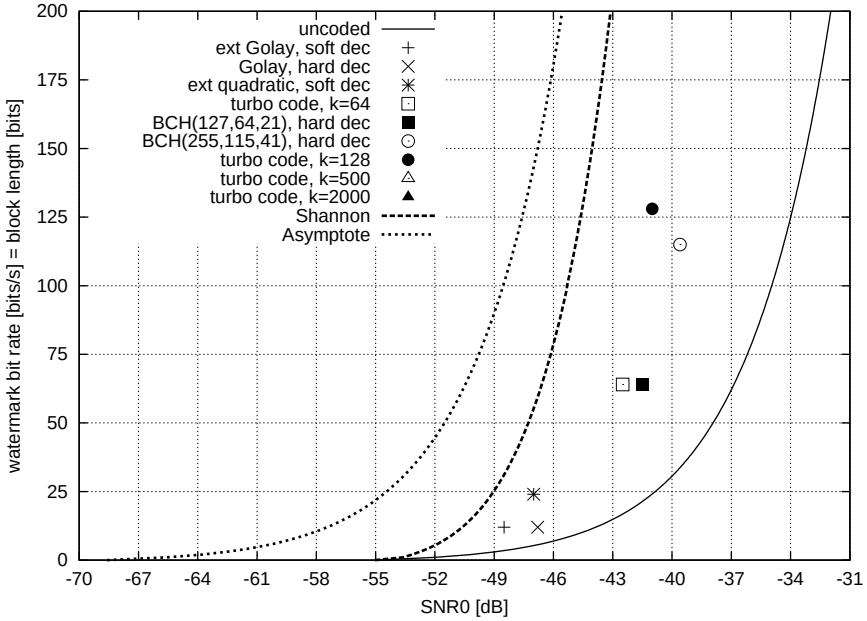

Figure 4. Shannon limit SNR_0 ranges for $P_w = 10^{-8}$, and small block lengths.

embedded and the use of coding can maximise the watermark data rate. In the case of attacks that drastically reduce SNR_0, small block lengths k might have to be used.

From a practical coding point of view, Fig. 3 has two regions: long (information) block lengths, for which the Shannon limit can be closely approached using concatenated codes and iterative decoding, and short/medium block lengths, for which it is more difficult to find codes that approach the (finite block length) limit.

The practical codes shown in Fig. 3 and 4 are mainly BCH codes (with hard decision decoding) and turbo codes. A single block is embedded in one video segment. Since the rate of the code does not match the considered rate, they are serially concatenated with repetition inner codes. The concatenation means that the code does not fully utilise the rate of the scheme, which is reflected in the relatively large distance from the limit, even in the case of perfect codes, such as the Golay code, and soft decision decoding. The situation can be improved by using BCH codes with lower rates and soft decision decoding.[1] The "uncoded" graph presented in the figures represents the performance a block of k uncoded bits, with an inner (rate $r = k/n$) repetition code.

At a block length $k = 64$, the BCH$(127, 64, 21)$ ($n = 127, k = 64, d_{min} \geq 21$) code with hard decision decoding achieves $P_w = 10^{-8}$ for an $E_b/N_o = 7.6$dB. Using equation 6, we obtain an SNR_0 value of $-67 + 7.6 + 10\log_{10} 64 = -41.34$dB, which is about 6dB away from the sphere packing bound, as shown in Fig. 4. At the same SNR_0, a perfect spherical code can accommodate more than 200 bits/s. Short turbo codes improve on the performance of the BCH code by just 1dB. This is due to the reduced minimum Hamming distance of the turbo code for this block length. A significant improvement could be obtained in this region by reducing the rate of the component codes for the turbo codes and using multiple or serial concatenations.[10] A BCH$(255, 115, 41)$ code with hard decision decoding is about 5dB away from the optimal code at a block length $k = 115$. The situation is better for long block lengths, where turbo codes are about 2dB away from optimal for both $k = 500$ and $k = 2000$. Note that this performance is dependent on P_w and it can get better if P_w is increased and worse if P_w is decreased.

6. SNR FOR SIMPLE WATERMARKING SCHEMES

The value of SNR_0 can be determined in several ways:

- As the ratio of of the watermark energy and the video energy. This can be used to estimate the SNR_0 in the absence of the MPEG2 attacks.
- By determining the distribution of the correlation peaks for a given spreading rate that verifies the Gaussian assumption. The corresponding SNR can then be determined as $SNR = \mu^2/\sigma^2$, where μ is the average of the distribution and σ is the standard deviation. The SNR_0 can then be extrapolated as $SNR_0 = SNR/s$, where s is the spread value. This can be used to estimate SNR_0 even in the presence of MPEG2 attacks.

SNR_0 values have been measured for two simple marking scheme: video independent spatial marking and video dependent spatial marking using a highpass filter.[6] In the former case, the spread watermark is directly added to the pixel values. In the latter case, the spread watermark is first modulated by the output of a highpass filter on the image (which makes it video dependent) and then added to the pixel values. In each case, a constant watermark amplification factor α has been set such that the watermark is (almost) invisible. The results are presented in Fig. 5 in the case of no attacks. It can be observed that the SNR for the video dependent scheme is not always better than that of the video dependent scheme. This is due to two possible factors:

- The video dependent scheme has artifacts at high frequencies which are more visible and thus the scaling factor had to be reduced.
- The scaling factor for the video independent scheme is too high in certain regions, such as plain regions where it should be set to zero.

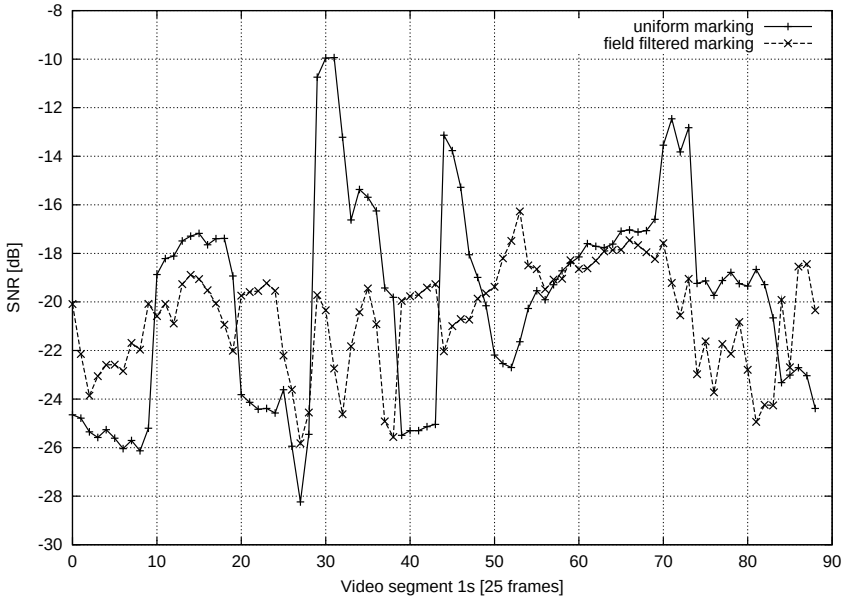

Figure 5. SNR_0 values for different video segments, no MPEG2 compression

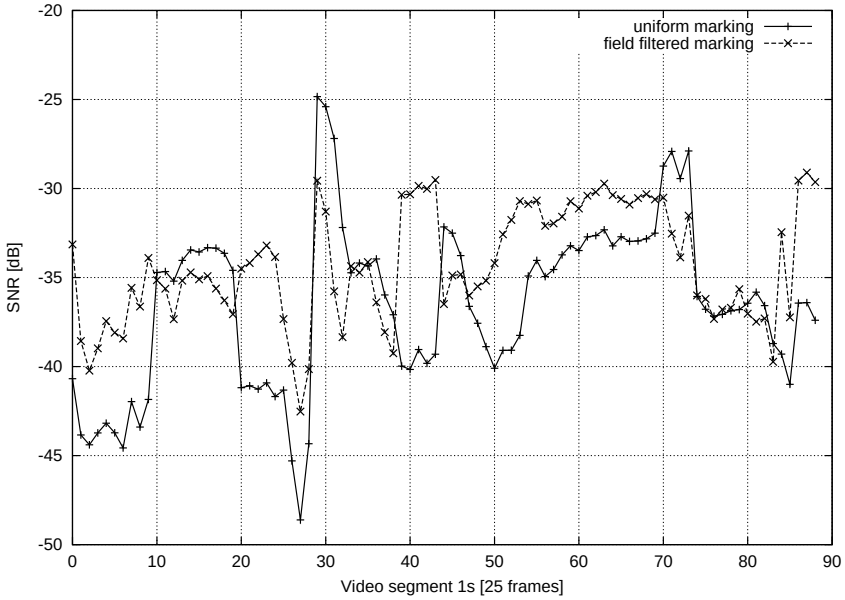

Figure 6. SNR_0 values for different video segments, 6Mbit/s MPEG2 compression

The advantage of video dependent schemes is apparent when the video is compressed. This is shown in Fig. 6 for MPEG2 compression at 6Mbit/s and Fig. 7 for MPEG2 compression at 3Mbit/s. As the compression increases, it becomes noticeable that the video dependent scheme is more resilient to this attack. By comparing Fig. 5, 6 and 7 it can be observed that the effect of MPEG2 compression on SNR is a reduction of -12dB for a rate of 6Mbit/s MPEG2 and -18dB for a rate 3Mbit/s MPEG2 as compared to the uncompressed situation.

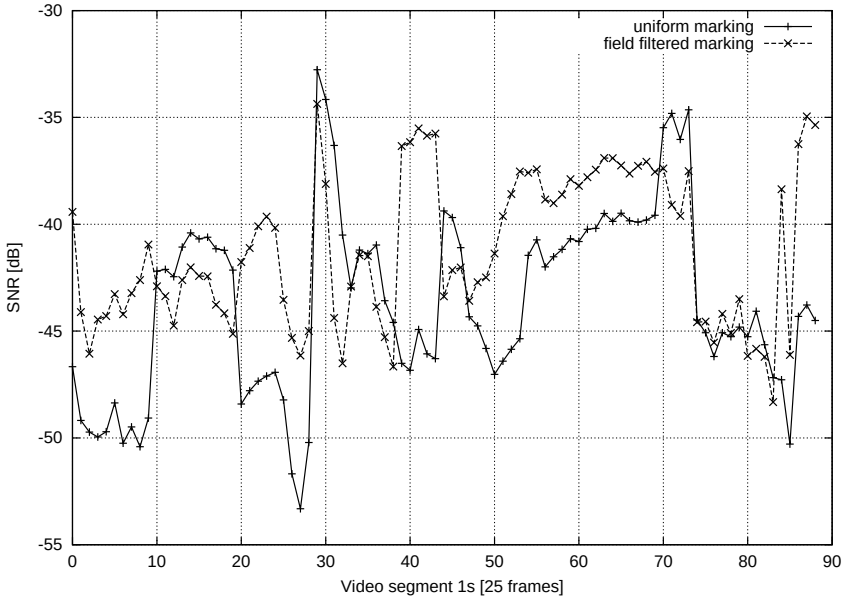

Figure 7. SNR_0 values for different video segments, 3Mbit/s MPEG2 compression

It can be observed that a typical SNR_0 for these schemes in the absence of the attacks is around -22dB which allows for a watermark data rate of about 3000 bits/s without coding. Coding can improve this rate to values above 5kbit/s. In the case of an 6Mbit/s MPEG2 attack, this SNR_0 value is reduced by a value of about -12dB, giving a $SNR_0 = -34$dB. In this case, the uncoded situation allows for a data rate of just 200 bits/s, but coding can easily improve this situation, and this is the case when long turbo codes are effective. Turbo coding can provide a data rate in the range of 500-2000 bits/s. In the case of a 3Mbit/s MPEG2 attack, the situation becomes critical, with a SNR_0 around -42dB. In this case, low rate, short codes have to be used.

The embedded energy can be increased by using better embedding schemes, such as DCT domain embedding using Just Noticeable Difference (JND) measures, or wavelet domain embedding. A DCT based scheme can produce a $SNR_0 = -17$dB with no attacks, which from Fig. 3 allows for an uncoded data rate higher than 5kbit/s. A wavelet based scheme can have a $SNR_0 = -13$dB in the absence of attacks, which also leads to a very high uncoded data rate. When the video is compressed at 3Mbit/s, the signal to noise ratio drops to about $SNR_0 = -39$dB for the JND scheme and about $SNR_0 = -30$dB for the wavelet-based system. This signifies that the JND scheme can work at up to 500bit/s with 3Mbit/s compression, whereas the wavelet-based scheme can achieve 200bit/s with no coding (Fig. 4) and more than 5kbit/s with coding (Fig. 3).

7. DIRTY PAPER CODES

The channel presented so far in this paper is not the best channel model for watermarking schemes. It has been shown that the watermarking channel is better modeled as a channel with side information.[11] The knowledge of the host signal can be used at the encoder to design the codewords in such a way that the host interference is eliminated.[12] This has resulted in the concept of "dirty paper codes", codes which eliminate the interference due to the host signal. This means that the error correction coding has to cope only with the distortion due to attacks, such as the MPEG2 attacks. Although this scheme is clearly better than the spread spectrum scheme, the comparison has to take into account several aspects, of which the most important is the adaptability of the the dirty paper code to HVS marking schemes. This dictates the amount of energy the watermark has, and thus the SNR when the video is compressed using MPEG2. We assume that the scheme is a non-HVS scheme and thus its energy is similar to that of the video independent scheme presented in section 6. Practical

implementations of dirty paper codes in the form of dither modulation are presented in.[11, 13, 14] This shows that the scheme has an impressive performance under MPEG2 compression, even at 3Mbit/s. The SNR is in the range of -18dB to -17dB, which means that without coding it could embed as much as 5kbit/s. The use of coding can allow for very large blocks, which can be calculated using the asymptotic formula 4 or just the asymptotic Shannon limit. It is known that turbo codes can approach this limit for very large blocks.

8. CONCLUSIONS

The paper investigates the performance of additive spread spectrum schemes with error correction coding. The watermarking scheme is modeled as a communications channel and a SNR value is calculated for different schemes and watermark bit rates. Using the SNR value and optimal spherical codes introduced by Shannon, the maximum bit rate of such a scheme for a given probability of error P_w is calculated. This illuminates the trade-off block length/code rate/watermark data rate and identifies the coding scheme to be used depending on the available watermark energy to interference energy ratio. Note that the limit presented here is not an ultimate limit for watermarking. It depends on modeling the channel as a communications channel and using spreading and scrambling to obtain a Gaussian interference. Schemes that use the correlation of the image rather than trying to break it could improve system performance. Also, a large part of the interference energy comes from the host signal. Host interference rejecting methods such as quantisation index modulation thus have a great advantage over additive spread spectrum. Classical error correction coding can also be used with these schemes to largely increase their data rate.

Acknowledgments

This work has been carried out under grant GR/R14606/01 awarded by the Engineering and Physical Sciences Research Council (EPSRC).

REFERENCES

1. S. Baudry, J. Delaigle, B. Sankur, B. Macq, and H. Maitre, "Analyses of error correction strategies for typical communication channels in watermarking," *Signal Processing* **81**, pp. 1239–1250, June 2001.
2. F. Perez-Gonzalez, J. Hernandez, and F. Balado, "Approaching the capacity limit in image watermarking: a perspective on coding techniques for data hiding applications," *Signal Processing* **81**, pp. 1215–1238, June 2001.
3. J. Hernandez, J. Delaigle, and B. Macq, "Improving data hiding by using convolutional codes and soft-decision decoding," in *SPIE Security and Watermarking of Multimedia Contents*, **3791**, (Santa Clara, USA), Jan. 2000.
4. J. Hernandez, F. Perez-Gonzalez, and J. Rodriguez, "The impact of channel coding on the performance of spatial watermarking for copyright protection," in *ICASSP98*, **5**, (Seattle, USA), May 1998.
5. C. Berrou, P. Thitimajshima, and A. Glavieux, "Near Shannon limit error correcting coding and decoding: turbo codes," in *Proc. IEEE International Conference on Communications*, pp. 1064–1070, (Geneva, Switzerland), May 1993.
6. A. Ambroze, G. Wade, M. Tomlinson, J. Stander, and M. Borda, "Turbo code protection of a video watermarking channel," *IEE Vision, Image and Signal Processing* **148**, pp. 54–58, Feb. 2001.
7. F. Balado and F. Perez-Gonzalez, "Coding at the sample level for data hiding: turbo and concatenated codes," in *SPIE Security and Watermarking of Multimedia Contents*, **4314**, (Santa Clara, USA), Jan. 2001.
8. C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal* **27**, pp. 379–423, July 1948.
9. S. Dolinar, D. Divsalar, and F. Pollara, "Code performance as a function of block size," *JPL TDA Progress Report* **42-133**, pp. 1–23, May 1998.
10. M. Ambroze, *On turbo codes and other concatenated schemes in communication systems.* PhD thesis, University of Plymouth, Plymouth, UK, Nov. 2000.

11. B. Chen and G. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory* **47**, pp. 1423–1443, May 2001.
12. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory* **29**, pp. 439–441, May 1983.
13. B. Chen and G. Wornell, "Dither modulation: a new approach to digital watermarking and information embedding," in *SPIE Security and Watermarking of Multimedia Contents*, **3657**, pp. 342–353, 1999.
14. B. Chen and G. Wornell, "Preprocessed and postprocessed quantization index modulation methods for digital watermarking," in *SPIE Security and Watermarking of Multimedia Contents*, **3791**, pp. 48–59, 2000.