Watermarking uncompressed video: an overview

Invited Paper

C. Serdean, A. Ambroze, and G. Wade (Department of Communication and Electronic Engineering, University of Plymouth)

M. Borda (Communications Department, Technical University of Cluj-Napoca, Romania)

I. Nafornita (Communications Department, Politechnica University of Timisoara, Romania)

Abstract

The paper overviews the watermarking of broadcast quality video, based upon the usual transform domain, spread spectrum approach. Aspects covered include perceptual-based marking, the advantages of wavelet based marking, and the performance of sliding correlators. The paper also derives an operational capacity for practical watermark channels, and explores the capacity improvement through the use of channel coding. The capacity is deduced by examining the correlation distribution at retrieval, and can be determined given MPEG-2 compression, geometric attack, visual thresholds, and channel coding. It is found that FEC based on multiple parallel concatenated convolutional codes (3PCCCs) can give over an order improvement in capacity for compressed video, and typically gives 0.5 kbit/s capacity even under a combined compression-geometric attack.

1 Introduction

Hidden data or a watermark is inserted into a video sequence for the purposes of copyright protection and video 'fingerprinting', and it can be performed either on uncompressed video (ITU-R 601) e.g. in studios, or on MPEG compressed video [1]. The former is discussed here, with emphasis upon the achievable channel capacity given typical attack conditions and advanced channel coding. The system under consideration is shown in Fig. 1. It uses the well-known spread spectrum approach, but also protects marking using FEC; N_b coded or uncoded bits are embedded over a sequence of frames, giving N_b normally distributed crosscorrelation peaks.

1.1 Requirements

Apart from the fundamental requirement that the watermark should not be visible under comparative studio viewing conditions, the design of a watermarking system for studios is subject to basic constraints [2], some of which are summarised in Table 1. Throughout this paper, we assume marking is on the luminance component of ITU-R 601 component digital signals in Y, Cr, Cb format (720×576 pixels).

1.	Watermark minimum segment (WMS) : 1 – 5 sec
2.	Bit error and false alarm probability: 10^{-8}
3.	Cryptographically strong, with a watermarking key
4.	Watermark detection: single-ended in that the unmarked host video is not available for retrieval (blind watermarking)
5.	Robustness to :
	• MPEG-2 compression (≥ 2 Mbit/s) • PAL encoding • analogue (VHS) and digital recording • Studio processing: A/D and D/A conversion, sampling rate/aspect ratio/frame rate conversion • Geometric attack: picture shift, cropping and scaling to minimum picture size, unnoticeable rotation • Collusion attack

Table 1 : Basic watermarking requirements for studio signals

A small WMS is required to facilitate studio editing (cut and paste). A minimum payload is 64 watermark bits/WMS, although here we will examine the potential for higher capacity. Attacks can be unintentional. For example, studio mixing can shift the frame 20 pixels, standards conversion can omit frames, and jitter can occur on VHS recording. The collusion attack is a serious problem; it averages multiple versions of the same video sequence (each version having a different fingerprint) with the objective of removing the fingerprint.

The requirements for video are different from those for image watermarking. For example, geometric attacks such as StirMark can perform serious rotation, scaling and translation (RST) on images and have led to the development of RST invariant transforms. For broadcast video this is less of a problem and the basic requirement here is robustness to *unnoticeable* rotation.

2 Watermark embedding

Watermark embedding can be carried out in either the spatial domain or the transform domain. Transform domain marking is generally preferred since it is easy to avoid marking high video frequencies (which tend to be attenuated by compression), and because it is naturally suited for perceptual marking based upon the human visual system (HVS). Also, from an information theoretic argument, transform domain marking can give increased channel capacity compared to spatial domain marking [3].

The FFT coefficients offer the possibility of phase modulation, but in order to obtain a real image/frame, it is necessary to maintain complex conjugate symmetry. Effectively this halves the potential marking capacity. In addition, the phase is quite sensitive to MPEG compression. Many transform domain schemes have been based upon modulation of the discrete cosine transform (DCT) coefficients, as shown in Fig.2. An advantage in developing a DCT-based system is that it is also appropriate for marking the MPEG-2 bitstream. However, recent work

on image watermarking has shown that marking the coefficients of the discrete wavelet transform (DWT) can have significant advantages, and this is discussed in section 5.2.

2.1 SNR in a video watermarking channel

Figure 2 shows the well known spread spectrum watermarking approach based upon the discrete cosine transform (DCT). For chip rate c_r, each data bit u_j is spread as $b_i = u_j,\ j c_r \le i < (j+1)c_r$, and the product $b_i p_i$ is formed, where $\{p_i\}$ is a binary {±1} PN sequence. This sequence spreads u_j over many 8×8 pixel blocks distributed over a number of video frames. Video dependent marking is an essential component of a successful marking scheme [4] if only that it ensures that marking energy is low in low detail areas of a video frame. A simple video dependent watermark is

$$w_i = \alpha b_i p_i |C_i| \tag{1}$$

Here, C_i is a DCT coefficient and α is selected to give a mark below the threshold of visual perception ($\alpha << 1$). If there is no attack or filtering ($\hat{C}_i' = C_i'$), and assuming $u_j = 1$, the normalized crosscorrelation is

$$d_j = \frac{1}{c_r}\sum_{i=jc_r}^{(j+1)c_r-1} p_i C_i \ + \ \frac{\alpha}{c_r}\sum_{i=jc_r}^{(j+1)c_r-1} |C_i|\, p_i^2 \tag{2}$$

We are interested in the distribution of these crosscorrelation peaks since it determines the detected bit error rate (BER). If, for simplicity, we assume C_i is i.i.d. with $C_i \sim N(0,\sigma_c^2)$, then d_j has the approximate distribution

$$d_j \sim N\left(\alpha\mu_{|C|}, \frac{\sigma_c^2}{c_r}\right) \tag{3}$$

where $\mu_{|C|}$ is the mean of $|C_i|$, $\alpha << 1$ and $p_i \in \{\pm 1\}$. The normal distribution follows from the Central Limit theorem since the cross correlator performs a sequence of correlation sums. Clearly, the BER will reduce as c_r increases (as expected) due to reduced distribution variance. The variance of d_j arises mainly from the first term in (2) and so we conclude that the non-zero crosscorrelation of the PN sequence with the DCT coefficients is a source of noise in the channel.

In practice, the distribution will be significantly affected by other factors. For example, it is widely recognised that crosscorrelation can be improved by inserting a 3×3 spatial filter in the video path (Fig.2). This removes low frequency video components prior to crosscorrelation and gives a distribution with larger mean and smaller variance. In practice, compared to filtering, balancing the PN sequence to ensure that it has zero mean gives only a relatively small improvement. Since marking is video dependent, the distribution will also depend upon the choice of sequence. Figure 3 illustrates the point for two standard MPEG video test sequences (ITU-R 601 format); it is apparent that sequence ‘flower garden’ will have a larger BER than sequence ‘mobile’. The underlying normal distribution is shown dotted.

Generalising, for any particular system the distribution mean μ, and variance σ^2 define a SNR of the channel: $SNR = (\mu/\sigma)^2$. The corresponding BER for an uncoded system is simply $BER_u = Q[\mu/\sigma] = Q[\sqrt{SNR_u}]$. For a coded system, μ and σ define a signal to noise ratio SNR_c at the decoder input, and the decoded bit error rate is $BER_c = f(SNR_c)$ where f is a known function for a particular iterative decoder (Fig.6(b)).

2.2 Perceptual-based marking

Rather than use (1), a better approach is to base marking on both the video sequence and the HVS. In this paper we compute a perceptual threshold or *just noticeable difference,* JND_i, for each DCT coefficient [5][6] and watermark as

$$C_i' = C_i + \alpha\, b_i\, p_i\, JND_i \tag{4}$$

where α is a small constant (see Fig.2). Marking is HVS based since JND_i is computed from physcovisual properties of the eye, and it is video dependent since JND_i is also a function of C_i. Human perception is also incorporated by making JND_i a function of the MPEG-2 default quantization matrix elements Q_i [7].

We compute *JND* values using a simplified form of Watson's model [8] and then enhance it to account for lateral inhibition masking [9]. First we compute the frequency sensitivity (modulation transfer function) of the eye as $T_F(i) = Q_i/2$. This greatly simplifies Watson's approach and avoids the need for empirical parameters. The *luminance masking* thresholds T_L and *contrast masking* thresholds T_C for block *k* are then [8]

$$T_L(i,k) = T_F(i) \cdot \left[\frac{C_{0,k}}{\overline{C}_0}\right]^{0.649} \tag{5}$$

$$T_C(i,k) = T_L(i,k) \cdot \max\left[1, \left(\frac{|C_{i,k}|}{T_L(i,k)}\right)^{w}\right] \tag{6}$$

where $\overline{C}_0$ corresponds to the mean DC coefficient over a frame, and $w = 0$ for the DC coefficient and $w = 0.7$ elsewhere. Equation (6) accounts for three aspects of the HVS, and, in a similar way to [9], it could be extended by incorporating lateral inhibition masking:

$$T_{LI}(i,k) = \left\{ \begin{array}{l} T_C(i,k) \quad \text{if } \left(T_C(i,k) > \mu(N_{i,k}) \text{ or } (T_C(i,k) - \sigma(N_{i,k})) < \dfrac{Q_i}{32} \right) \\ T_C(i,k) - \sigma(N_{i,k}) \quad \text{otherwise} \end{array} \right\} \tag{7}$$

where $\sigma(N_{i,k})$ and $\mu(N_{i,k})$ are the standard deviation and the mean for the eight neighbours of $T_C(i,k)$. We obtained the condition in (7) from subjective tests. Using the basic *JND* definition in [8], the values for block *k* are then

$$JND_{i,k} = \frac{Q_i}{2\,T_{LI}(i,k)} \tag{8}$$

Equation (8) ensures that coefficients that are less visible are marked with greater energy. In practice, the theoretical *JND* values have been found to be within a factor 2 or 3 of the actual perceptual threshold, and this is accounted for by the factor α in (4).

3 Channel capacity

If we regard the watermark channel as a communications system with input X (the watermark data) and output Y, the channel capacity is defined as the maximum mutual information:

$$C_{chan} = \max_{p(x)} I(X;Y) = \max_{p(x)} \left[h(X) - h(X \mid Y) \right] = \max_{p(x)} \left[h(Y) - h(Y \mid X) \right] \qquad (9)$$

where the maximum is taken over all possible distributions $p(x)$. Term $h(X \mid Y)$ represents information loss due to channel 'noise', which will be a combination of the host video and signal processing (compression/attack). If the loss is modelled as the addition of an independent Gaussian noise source, $Z \sim N(0, \sigma_z^2)$, i.e. $Y_i = X_i + Z_i$, where Z is a continuous random variable, then (9) reduces to [10]

$$C_{chan} = \frac{1}{2} \log_2 \left(1 + \frac{\sigma_x^2}{\sigma_z^2} \right) \quad \text{bits/symbol} \qquad (10)$$

providing $X \sim N(0, \sigma_x^2)$. In [3], σ_x^2 was estimated from acceptable JPEG performance, and an equivalent Gaussian noise variance was computed for the host and JPEG compression. In [11] the noise was restricted to an AWGN attack, the host noise being virtually eliminated by using it as side information during embedding.

In this paper we invoke *JNDs* to maximise the signal power. From the discussion in section 2.1, the channel in Fig.1 can be modelled as a gain factor cascaded with a Gaussian noise source $Z \sim N(0, \sigma_z^2)$ (the gain and variance depending upon the host video, MPEG compression, and geometric attack). For example, we could estimate a basic operational capacity as follows. Suppose that all N_p pixels ($N_p = 720 \times 576$) in the frame are transformed via DCT blocks, and that the channel noise is simply that of the host video. Assuming C_i is i.i.d. for simplicity, the noise power per video frame is $N_p \overline{C^2}$, where $\overline{C^2}$ is the mean coefficient power for a particular video sequence. If only one data bit is embedded per video frame (corresponding to $c_r = N_p$), and there is no FEC, the SNR is

$$SNR = N_p \frac{\left(\alpha \, \overline{JND} \right)^2}{\overline{C^2}} = N_p \, \overline{SNR} \qquad (11)$$

where $\overline{SNR}$ is a measured mean SNR for the video sequence and $\overline{JND}$ is the mean JND for the sequence. If N_b data bits are embedded into a frame the signal to noise ratio per uncoded data bit reduces to $SNR_u = SNR / N_b$, and the data rate or capacity for an uncoded system of frame rate F_r is

$$D_r = N_b F_r = \frac{k \, N_p \, \overline{SNR} \, F_r}{SNR_u} \quad \text{bits/s} \qquad (12)$$

In (12) we account for the fact that, in general, only a fraction k of the coefficients in each DCT block are marked. Also, in the following work we will define the operational capacity as the maximum value of D_r for which the *BER* does not exceed a tolerable level (typically 10^{-8}).

Equation (12) has been used to estimate the capacity for the uncoded, uncompressed video test sequence 'flower garden' (graph (b) in Fig. 4), assuming the *JND*-based marking in (4). Coefficient α in (4) was set to give marking below the threshold of visibility (visible *JND* marking appearing as fine grain noise). Graph (a) in Fig.4 shows the corresponding simulated capacity, obtained by embedding N_b bits over N_f frames ($D_r = N_b F_r / N_f$) using (4). This gives N_b correlation peaks and the resulting distribution gives $SNR_u = (\mu/\sigma)^2$. The discrepancy between graphs (a) and (b) is attributed to highpass filtering prior to crosscorrelation. Graph (c) shows the simulated result for MPEG-2 compression to 6 Mbit/s. The reduced capacity due to compression is more clearly shown in Fig.5. Using $BER_u = Q\left(\sqrt{SNR_u}\right)$, graph (a) shows that the uncompressed capacity is around 3 kbit/s, and graph (b) gives experimental confirmation of (a) by directly counting data errors. Graph (c) shows that MPEG-2 compression reduces the capacity to the order of 100 bit/s.

4 Use of Turbo codes

Consider a coded system of rate R (Fig.1) and assume that N_b coded bits are embedded over a sequence of N_f video frames. The watermark data rate is now

$$D_r = \left(\frac{N_b R}{N_f}\right) F_r \quad \text{bits/sec} \tag{13}$$

Assuming as before that all N_p pixels in a frame are transformed, the spread spectrum chip rate is

$$c_r = \frac{k N_f N_p}{N_b} = \frac{k N_p R}{D_r} \cdot F_r \tag{14}$$

According to (14), for a fixed D_r, the use of FEC reduces the chip rate by a factor *R*. As indicated in (3), this increases the variance of the channel distribution, resulting in increased *BER*, and the FEC decoder must more than compensate for this increase in order to provide coding gain.

For a channel with a potentially large BER (due to attack) it is essential to use soft decision decoding, and in practice this restricts the choice of FEC to convolutional codes. Viterbi decoding is the usual ML decoder for an AWGN channel, and so has been used to protect watermarked video [12]. Here, we are interested in the recovery of channel capacity that can be achieved by using Turbo codes in an attacked watermarking channel. Turbo codes offer better coding gain and could be used at low rate in this application (with consequent improvement in performance). We consider a code of rate *R* and interleaver size *N*, and treat the FEC as block coding (block length *N*). Note from Fig.1 that coding precedes spread-spectrum so that the input to the iterative decoder will be the output of the spread-spectrum correlator. This scheme ensures that the watermark channel up to the decoder input approximates to a Gaussian channel, and the latter is the usual assumption for Turbo code

systems. An alternative scheme is to place the FEC encoder after the spread-spectrum process. This has the potential to provide large block lengths for the encoder and so improve its performance, but has the disadvantage that the channel at the decoder input is poorly defined and will have a very low SNR.

In this paper a multiple parallel concatenated convolutional code (3PCCC) has been used to protect the watermark channel and the encoder is shown in Fig.6(a). The use of two interleavers (I_1 and I_2) rather than one as in the basic Turbo code reduces the error rate floor and so gives improved performance [13]. Each recursive systematic code (RSC) is an optimum (5,7) code [14], giving an unpunctured code rate $R = 1/4$. Figure 6(b) shows the simulated performance of the overall code for several interleavers.

Figure 4 shows that an uncoded watermark channel at 1 kbit/s corresponds to $SNR_u \approx 4$ dB when 6 Mbit/s MPEG-2 compression is used. Whilst this low SNR is unusable without FEC, Fig.6(b) shows that that iterative decoding should be effective against this sort of attack. Figure 7 illustrates the effectiveness of iterative decoding against a 6 Mbit/s MPEG-2 attack, and also compares the performance of the heuristic marking scheme in (1) (Fig.7(a)) with that of the *JND* scheme in (4) (Fig.7(b)). The BER of the uncoded system is computed as before, whilst that of the coded system is computed as $BER_c = f(SNR_c)$ using Fig.6(b). Without FEC the attack reduces the 10^{-8} capacity to around 300 bits/s, but with FEC the capacity can be over 3 kbit/s. In each case, α was selected to give marking just below the threshold of visibility. Note that *JND* marking is superior to the heuristic scheme, and can give a capacity over 8 kbit/s for uncompressed video (not shown).

4.1 Combined attacks

In practice watermarked video is likely to suffer from a combination of attacks, such as MPEG-2 compression and geometric distortion, and an attack of this nature can defeat many watermarking schemes [15][16]. Figure 8 shows the effect of a combined compression and line cut attack upon the *JND*-based marking scheme in (4). In order to combat the line cut we use a 2-D sliding correlator. This moves the known PN sequence over a small search window relative to the received coefficient block in order to locate the correlation peak. The 10^{-8} capacity is relatively low for an uncoded system (graph (a)), and a possible explanation is that compression reduces the watermark amplitude to below the performance threshold of the sliding correlator. However, graphs (b) and (c) show that performance can be significantly improved through the use of FEC.

5 System improvements

5.1 3-D correlator

Spread spectrum systems are vulnerable to synchronisation error, as can occur in a geometric attack. In Fig.8 we used a 2-D (spatial) correlator to combat a line cut attack, and Fig.9 shows a 3-D (spatial/temporal) correlator for combating temporal attacks, such as a frame cut. All three macro blocks shown are marked with the same data bit, but they are randomly placed within frames (as indicated) in order to improve security (within a macro block the data is at the same spatial location in order to perform correlation, the typical marking depth being four frames).

The use of a sliding correlator has the disadvantage of decreasing the effective chip rate since a local crosscorrelation peak is computed for small blocks (the overall crosscorrelation being the sum of the local correlations). This amounts to a 'correlator loss', and is illustrated in Fig.10. The loss is about 3 dB for the 3-D correlator (Fig.10 graph (b)), and is the same for a frame cut (irrespective of position within the sequence). If a 2-D correlator is used, the loss varies significantly and can be large e.g. 12 dB if the frame cut occurs at frame 30 within a 120 frame sequence (Fig.10 graph (d)).

5.2 Use of the DWT

The human observer tends to process video information by independently processing multiple frequency channels. In a similar way, the multiresolution decomposition of the discrete wavelet transform (DWT) also enables signals in different channels to be processed independently. *This implies that frequency bands can be processed independently without significant perceptible interaction between them, and that the need for a complex perceptual marking algorithm (JNDs) is reduced.* Looked at another way, the multiresolution aspect also enables a watermark to have spatially local and spatially global components.

Figure 11(a) illustrates a simple watermarking scheme based on the DWT [17]. For one level of decomposition, the DWT generates 4 subimages (coefficients) gg, gh, hg, hh. A second level of decomposition is obtained by similarly transforming the approximation coefficients gg. Watermarking is performed by multiplying the *detail* coefficients gh, hg, hh by a (secret) constant, K, its value being chosen to avoid visible artefacts. This procedure tends to mark perceptually significant regions in a video frame e.g. edges, as required for robust watermarking. The watermarked frame is I_w and the watermark itself is W. Assuming no attack for simplicity, Fig. 11(b) shows that the watermark can be extracted blind by multiplying the detail coefficients gh, hg, hh by K^{-1}. In the presence of an attack, $W' \neq W$ and watermark detection can be carried out by crosscorrelating W' with W, and thresholding the result. Clearly, this simple system conveys just 1 bit of information. Also, more secure marking would use a PN sequence to identify coefficients to be marked.

Using the system in Fig.11, it has been shown that DWT marking is more robust to compression compared to DCT marking [17]. More comprehensive work confirms this and shows that significant improvements can also obtained wrt geometric attacks e.g. frame shift, cropping and scaling [18]. The watermarking component with local spatial support tends to be robust to cropping, whilst the component with global support tends to be robust to lowpass filtering.

5.3 Choice of PN sequence

In order to provide a degree of security, each data bit in Fig.2 is spread using a different PN sequence. Each PN sequence is a subset of a large PN sequence, and the key to this sequence is also used to select random locations for the marked blocks. Evenso, the multiplicative congruential algorithm [19] used to generate the large sequence is known to be insecure. In general, the PN sequences used should have the following properties:

- large chip rate e.g. $c_r \approx 10^6$
- large autocorrelation and small crosscorrelation

- cryptographically secure

If we use selected pairs of *m* sequences (*preferred sequences*), it is possible to obtain good autocorrelation and reasonably low cross-correlation. For an *n* stage LFSR the cross-correlation function can then be 3-valued: $\{-1, -t, t-2\}$ where $t = 2^{\lfloor (n+2)/2 \rfloor} + 1$. Better cross-correlation is achieved using Kasami sequences, the 3-values being: $\{-1, -(2^{n/2}+1), 2^{n/2}-1\}$, *n* even. For $n = 6$, preferred pairs would give $\{-1, -17, 15\}$, compared to $\{-1, -9, 7\}$ for Kasami sequences. However, an *m* sequence generated from an *n*-stage linear feedback shift register is not cryptographically secure since it can be deduced from just $2n$ bits of the sequence. A partial but still unsatisfactory solution is to associate some non-linearity with the register. In order to improve security, Kasami sequences could be encrypted using one of the well known algorithms e.g. DES ECB, DES CBC, IDEA ECB, or IDEA CBC [20]. Experiment shows that the resulting crosscorrelation is somewhat inferior to pure Kasami sequences [17].

6 Conclusions

A spread spectrum-based video watermark data channel is conveniently characterised by the Gaussian distribution at the output of the sliding correlator. This distribution defines a SNR for the channel, from which can be deduced an operational channel capacity for a system subject to perceptual marking, combined attack, and channel coding. The Gaussian input to the FEC decoder, and the fact that low code rates can be tolerated, makes iterative decoding particularly appropriate for the protection of a watermarked channel. The computational complexity of such decoding is still relatively small compared to that of the sliding correlator.

As expected, channel capacity increases through the use of perceptual marking (*JNDs*) and FEC, and reduces when watermarked video is subjected to attacks. MPEG-2 compression to 6 Mbit/s reduces the 10^{-8} capacity from over 8 kbit/s (uncompressed video, *JND* marking) to about 300 bits/s, although an order improvement is achieved through FEC. The vulnerability of spread spectrum systems to synchronisation error has been highlighted, although a 3-D correlator has proved effective against temporal attack, such as a frame cut. A combination of MPEG-2 compression and simple geometric attack can severely reduce capacity, although useful improvements can still be made through the use of FEC. Under such an attack, FEC enables the current watermarking scheme to achieve a typical capacity of 500 bits/s, although this is very video dependent. Increased robustness to attack should be achievable by replacing the DCT with the DWT.

References

1. Hartung, F., and Girod, B. 'Digital Watermarking of Raw and Compressed Video', *Proc. SPIE 2952: Digital Compression Technologies and Systems for Video Communication'*, October 1996, 205-213.

2. Union Europeenne de Radio-Television (EBU-UER) Watermarking Working Group, February 2000.

3. Ramkumar, M., and Akansu, A. 'Information Theoretic Bounds for Data Hiding in Compressed Images', IEEE 2nd Workshop on Multimedia Signal Processing, Redondo Beach, CA, 1998.

4. Swanson, M., Zhu, B., Chau, B., and Tewfik, A. 'Object-based transparent video watermarking', *Electronic Proc. IEEE Signal Processing Society*, 1997 Workshop on Multimedia Signal Processing', Princeton, New Jersey, 23-25 June, 1997.

5. Wolfgang, R., Podilchuk, C., and Delp, E. 'Perceptual Watermarks for Digital Images and Video', *Proc. IEEE*, Vol. 87, No.7, July, 1999, 1108-1126.

6. Kim, S., Suthaharan, S., Lee, H., and Rao, K. 'Image watermarking scheme using visual model and BN distribution', *Electronic Letts.*, Vol. 35, No.3, February 1999, 212-213.

7. ISO/IEC 13818-2, 'Information Technology - Generic coding of moving pictures and associated audio information'.

8. Watson, A. 'DCT quantization matrices visually optimised for individual images', Human Vision, Visual Processing and Digital Display IV', *Proc. SPIE 1913-14*, 1993, 1-15.

9. Kim, S. 'Image watermarking scheme using visual model and BN distribution', *Electronics Letts.*, Vol. 35, No.3, 4 Feb. 1999, 212-213.

10. Cover, T., and Thomas, J. 'Elements of Information Theory', Wiley, 1991.

11. Eggers, J., Su, J., and Girod, B. 'A Blind Watermarking Scheme Based on Structured Codebooks', IEE Secure Images and Image Authentication, London, UK, April 2000, 1-6.

12. Hernandez, J., Delaigle, J., and Macq, B. 'Improved data hiding by using convolutional codes and soft-decision decoding', *Proc. SPIE*, 'Security and Watermarking of Multimedia Contents', San Jose, CAL., 24-26 January 2000, 24-47.

13. Divsalar, D. and Pollara, F. 'Multiple Turbo Codes for Deep Space Communications', *TDA Progress Report 42-121*, May 1995, 66-77.

14. Benedetto, S., Garello, R. and Montorsi, G. 'A Search for Good Convolutional Codes to be Used in the Construction of Turbo Codes', *IEEE Trans. on Communications*, Vol.46, No.9, Sept. 1998, 1101-1105.

15. Petitcolas, F., Anderson, R., and Kuhn, M. 'Attacks on Copyright Marking Schemes', Second Workshop on Information Hiding', in Vol. 1525 of *Lecture Notes in Computer Science*, Portland, Oregon, USA, 218-238, 14-17 April, 1998.

16. Hartung, F., Su, J., and Girod, B. 'Spread Spectrum Watermarking: Malicious Attacks and Counterattacks', *Proc. of SPIE*, Vol. 3657: Security and Watermarking of Multimedia Contents, January 1999.

17. 'Preliminary Report on Watermarking', Borda, M. and Deac, V., Communication Dept., Technical University of Cluj-Napoca, Romania, and Nafornita, I. and Isar, A., Politechnica University of Timisoara, Romania, June 2000.

18. Podilchuck, C., and Zeng, W. ‘Image-Adaptive Watermarking Using Visual Models’, IEEE Journal, SAC, Vol.16, No.4, May 1998, 525-538.

19. Press, W. and Teukolski, S. ‘Numerical recipes in C’, Cambridge University Press, 1993.

20. Stallings, W. ‘Network and Internetwork Security Principles and Practice’, Prentice Hall, 1995.

Figure Captions

Fig.1: Channel coding in a watermark channel.

Fig.2: Transform domain spread spectrum watermarking and retrieval.

Fig.3: Channel distributions for two video sequences ($u_j = 1 \ \forall \ j$)

Fig.4: SNR for sequence ‘flower garden’ ; (a) uncompressed, filtered; (b) uncompressed, unfiltered, equation (12); (c) MPEG-2 compressed to 6 Mbit/s, filtered.

Fig.5: BER for sequence ‘flower garden’; (a) uncompressed; (b) uncompressed simulation; (c) MPEG-2 compressed to 6 Mbit/s, filtered.

Fig.6: Rate ¼ 3PCCC FEC: (a) encoder; (b) simulated performance of an iterative decoder

for interleaver sizes of 500 and 2000.

Fig.7: Combating a 6 Mbit/s MPEG-2 attack with iterative decoding ($N = 2000$) for sequence ‘basketball’: (a) heuristic marking, $\alpha = 0.004$; (b) JND marking, $\alpha = 3$.

Fig.8: Combined compression and line cut attack on sequence ‘basketball’: (a) uncoded; (b) coded, $N = 500$; (c) coded, $N = 2000$; (d) Shannon limit.

Fig.9: 3-D sliding correlator.

Fig.10: Correlator loss for sequence ‘flower garden’ (120 frames) : (a) no attack, no sliding; (b) 3-D correlator with/without frame cut attack; (c) 2-D correlator, frame cut attack after 60 frames; (d) 2-D correlator, frame cut attack after 30 frames.

Fig. 11: DWT domain watermarking: (a) embedding; (b) retrieval, assuming no attack

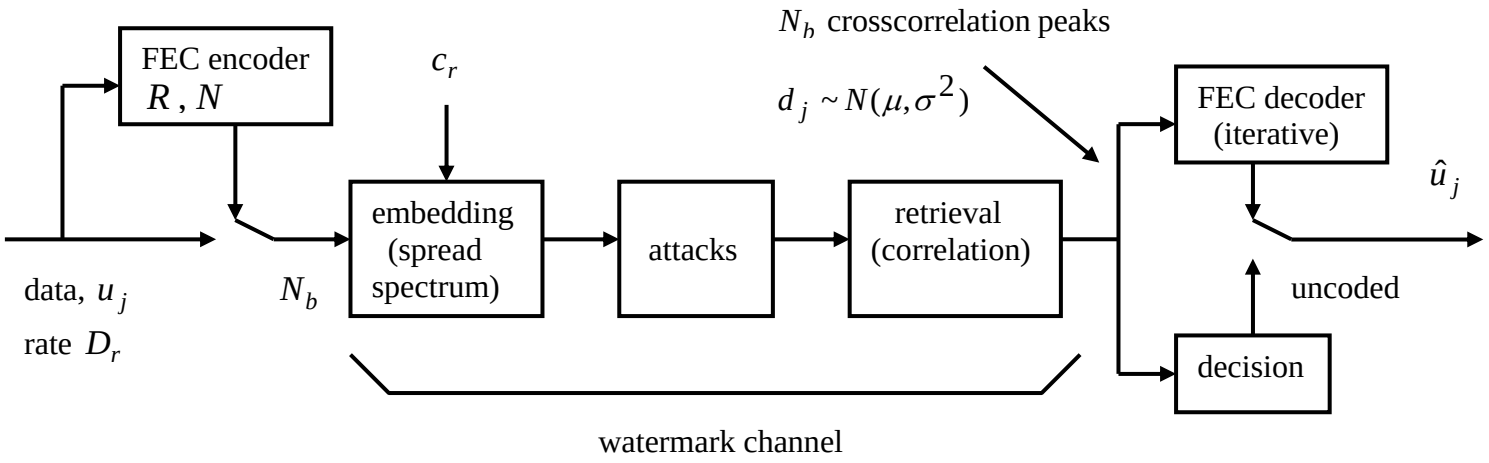

Figure 1 : channel coding in a watermark channel

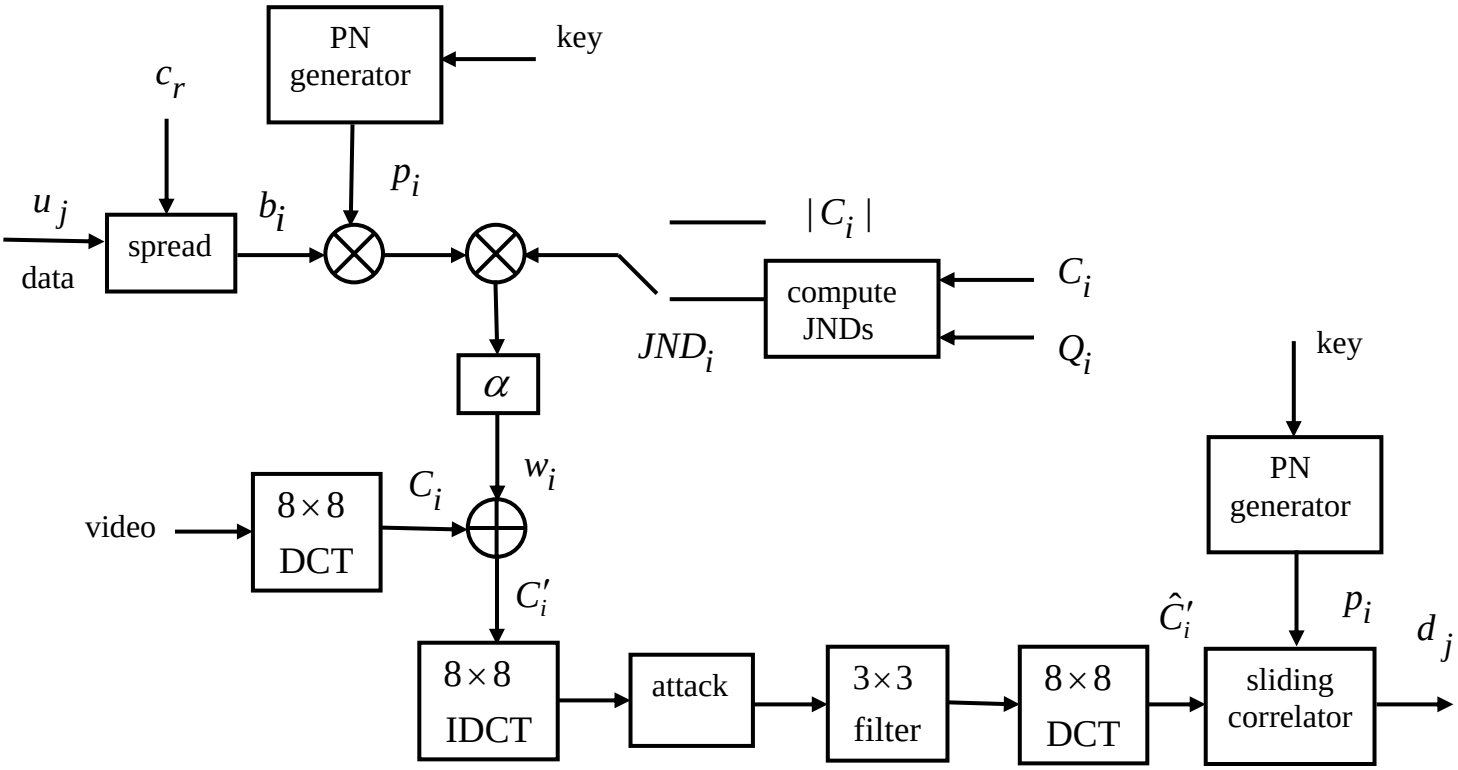

Figure 2 : Transform domain spread spectrum watermarking and retrieval

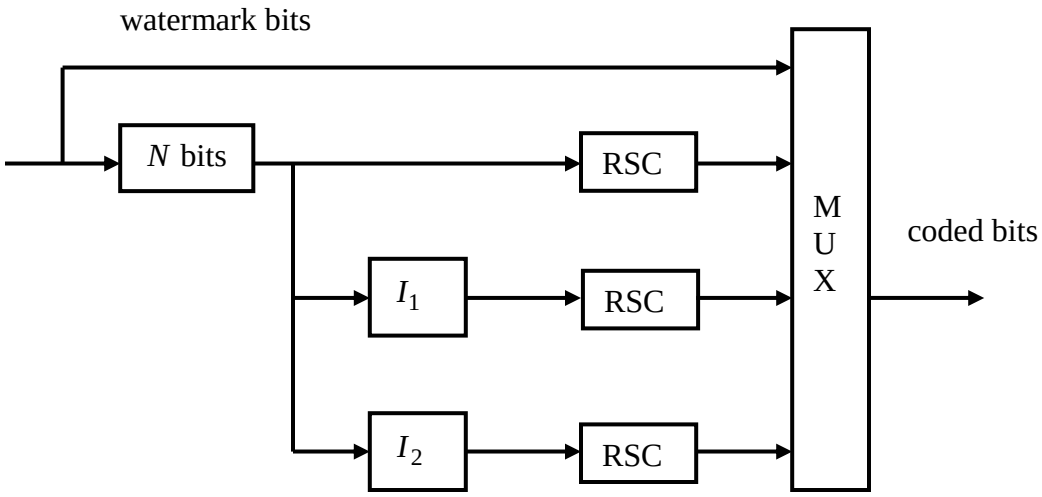

Fig. 6(a) Rate ¼ 3PCCC FEC encoder

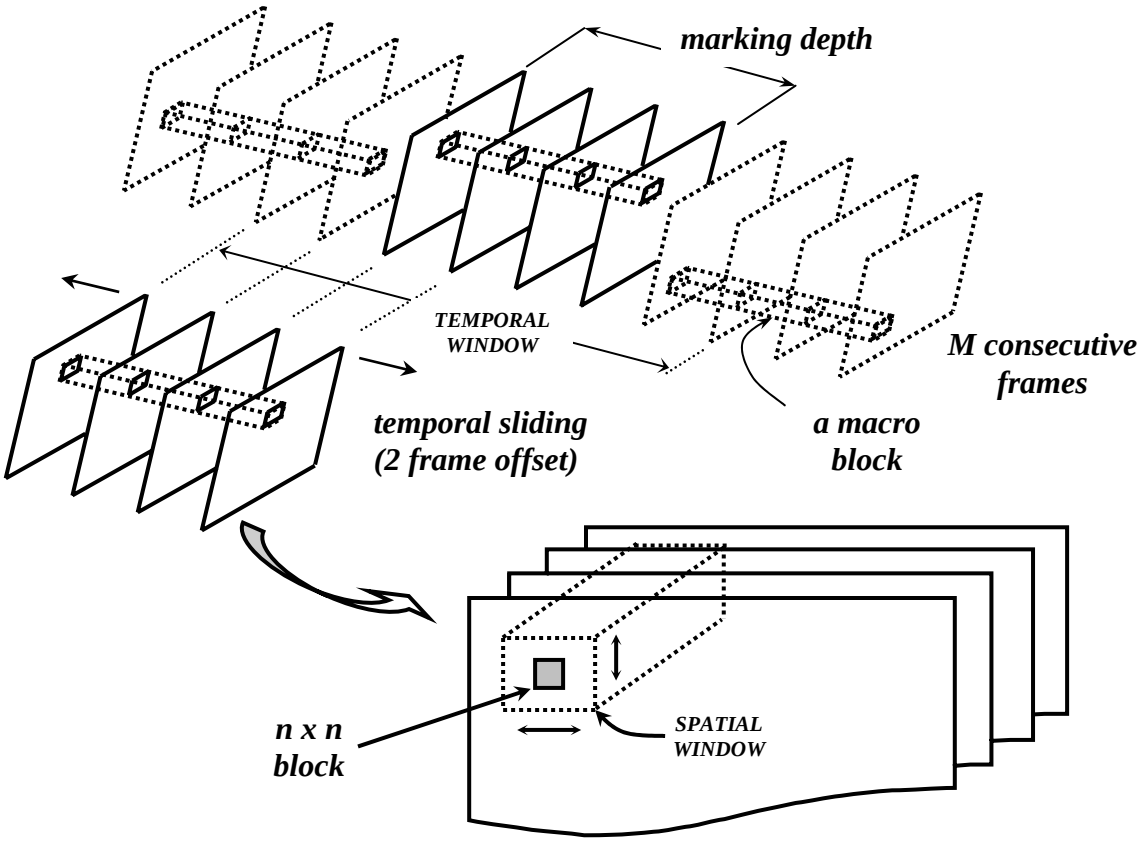

Fig. 9 : 3-D sliding correlator

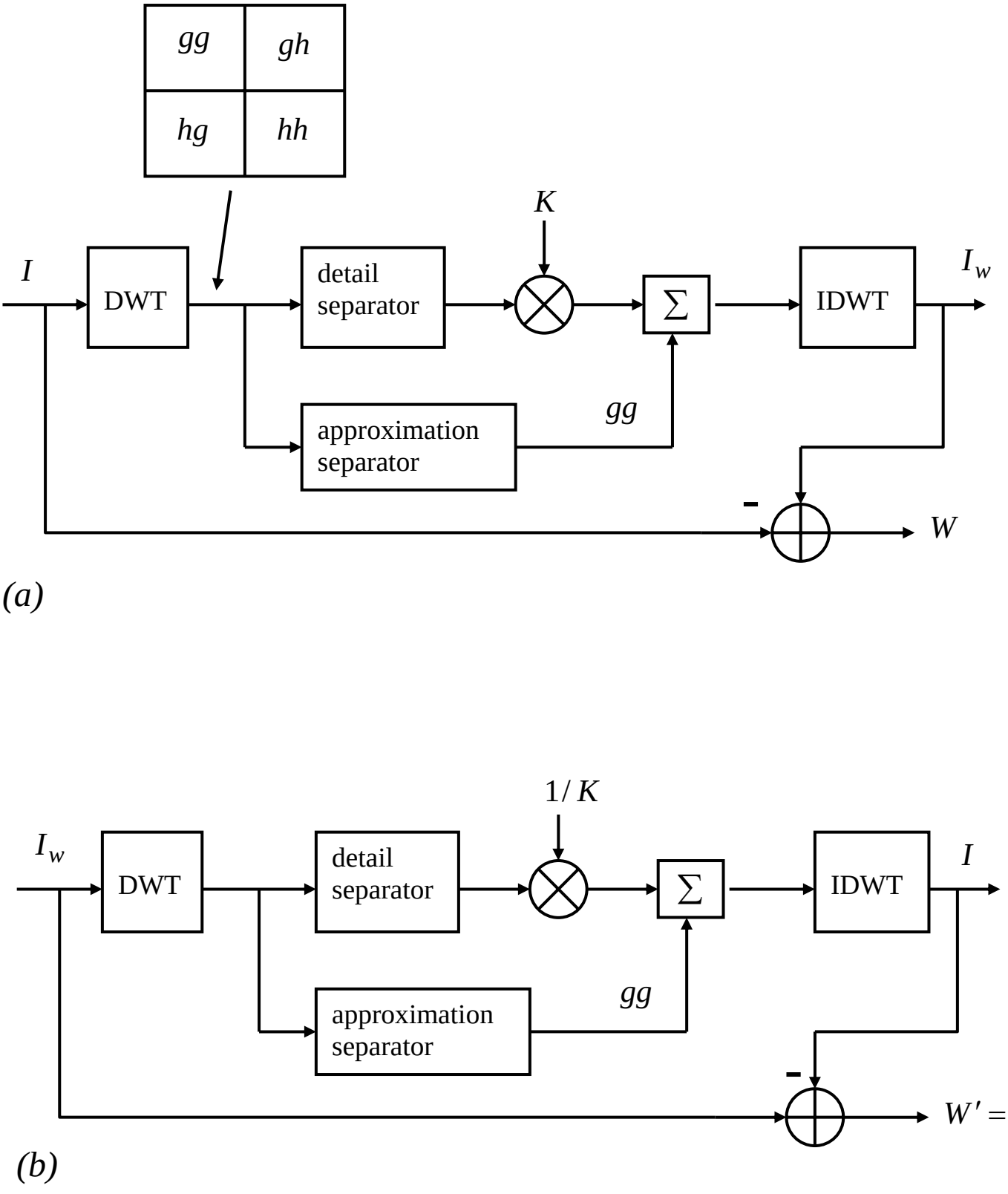

Fig. 11 : DWT domain watermarking

(a) embedding, (b) retrieval, assuming no attack