

TOWARDS A ROBUST DIGITAL WATERMARK

R.G.van Schyndel(*), *A.Z.Tirkel*(+), *C.F.Osborne*(*)

(*) Department of Physics, Monash University, Clayton, 3168, Australia.

(+) Scientific Technology, P.O.Box 3018, E. Brighton, 3187, Australia.

Abstract

This paper discusses the feasibility of coding a robust, undetectable, digital water mark on a standard 512*512 intensity image with an 8 bit gray scale. The watermark is capable of carrying such information as authentication or authorisation codes, or a legend essential for image interpretation. This capability is envisaged to find application in image tagging, copyright enforcement, counterfeit protection, and controlled access. The method chosen is based on linear addition of the water mark to the image data. Originally, the authors made use of one dimensional encoding using m-sequences [1],[2], a process which, whilst showing promise, had considerable shortcomings. This paper presents an analysis of the one dimensional scheme and some constructions to extend this work to two dimensions.

1 Background

There exist two basic classes of electronic water marks: fragile and robust. Interest in both types has increased in recent times because of the explosion in digital communications and the rapidity and ease of transmission of electronic material which is subject to copyright. The authors have been concerned with the construction of the robust type, i.e. one which is resilient to some image distortions such as pixel or bit tampering, cropping, translation, rotation and shear. At this stage, our watermark possesses limited immunity against the first three distortions, but the intention is to improve its performance in the future. This should be contrasted with a novel technique involving a fragile watermark as described in [3], where, by deliberate design, *any* distortions render the watermark non-recoverable and this becomes proof of tampering. Both methods use LSB manipulation. Walton [3], also introduces an ingenious and effective palette manipulation technique to increase the watermark effectiveness by involving the complete RGB image components. A totally different technique and its variations is reviewed in [4]. Its major advantage is its compatibility with the JPEG format, whilst its principal disadvantage is that the watermark recovery requires the presence of the unencoded image. In this respect it differs from the other techniques. Our

technique involves a linear addition of the watermark pattern, followed by a correlative recovery.

Correlation can be defined as: cyclic or extended, global or character specific.

Correlation functions can be decomposed into: even and odd, or periodic and aperiodic. At this stage we are confined to binary characters only. Our watermarks are chosen from two-dimensional array patterns based on m-sequences or extended m-sequences. An m-sequence basis is chosen because of their balance (zero mean), random appearance, resilience to filtering, cropping and individual bit errors, optimal autocorrelation properties and constrained cross-correlation. The water mark can be encoded by the choice of m-sequences and their phases.

2 Method

Our encoding method uses LSB addition for embedding the water mark [1], [2]. In many imaging systems the LSB is corrupted by hardware imperfections or quantisation noise and hence its manipulation is invisible, or of limited significance.

The linear addition process is difficult to crack and makes it possible to embed multiple watermarks on the same image [2]. The decoding process makes use of the unique and optimal auto-correlation of m-sequence arrays to recover the watermark and suppress the image content. Since the correlation process involves averaging over long strings of binary digits, it is relatively immune to individual pixel errors, such as may occur in image transmission. The correlation process requires the examination of the complete bit pattern and must therefore be performed off-line, unless some form of dedicated, real-time, parallel processing is involved.

The decoding process is not completely error free, due to partial correlation of the image data with the encoding sequence. In our previous work, we overcame this by filtering and dynamic range compression [2]. These artificial steps would be undesirable in a practical system. A typical 128*128 (unfiltered) image encoded with a one dimensional watermark is shown in Fig.1(Top left). The message is encoded on a line by line basis, using the ASCII character to select a sequence phase shift. There are numerous message repeats. The decoder output

Fig.1(centre left) shows distinct message correlation peaks (white). Note that there are significant sidelobes due to image crosscorrelation effects. The top half of Fig.1. shows encoded images that have been progressively high-pass filtered, removing 10, 60 and 100 of the spatial frequency components from the total of 128. The watermark peaks survive all these filtering processes, demonstrating the robustness of the technique. The image content in the original and the decoded version is rendered negligible after the second or third of the filters. It is also clear that the filtering introduces progressively more severe ringing in the decoded output. This can result in ambiguities. We are presently investigating the feasibility of rectifying this shortcoming by the introduction of a matched filter in the decoding process. Fig.2. shows similar effects on the watermark alone (encoded on a null image). Contrast enhancement was employed to render the watermark visible.

3. Watermark requirements

An ideal watermark would possess:

- (i) High in-phase autocorrelation peak for rows and columns [All]
- (ii) Low out-of-phase autocorrelation for rows and columns [Costas Arrays]
- (iii) Low cross-correlation between rows and between columns & between rows and columns [Perfect Maps, Hadamard Matrix, Legendre Arrays]
- (iv) Low cross-correlation with image content [Folded M-sequence]
- (v) Array diversity [Gold, Extended Gold Arrays]
- (vi) Balance [All except Costas and some Gold]

The first two criteria are required for unambiguous watermark registration, the third is necessary to avoid scrambling, the fourth minimises image related artefacts, whilst the fifth is concerned with the information capacity of the watermark. The sixth criterion maximises the significance of the correlation operation: in the binary case, the minority symbol determines the correlation score.

Constructions can be optimised for each of these requirements. However, a global optimisation requires compromise. We have examined all the criteria in detail, with the exception of (iv). Presently, we are examining methods of watermark design to minimise crosscorrelation with the image.

3.1 Image crosstalk suppression (ideal)

Clearly, it is possible to analyse the image content, by DCT or Walsh Transform and deduce a low crosscorrelation watermark by remapping any pattern, satisfying all criteria above except (iv). Similar effects could be assured by a random or adaptive search for a mapping to minimise the

crosscorrelation with the image. However, such a procedure is impractical because there is no guarantee of uniqueness and hence the computation of the inverse mapping at the decoder.

3.2 Crosstalk suppression (practical)

There are at least three approaches which do not suffer from the above problem.

- (1) Use longer m-sequences.
- (2) Use high pass filtering.
- (3) Use a "random" mapping.

The first is obvious. The other methods rely on the low overlap of the spatial frequency content of the image and watermark. In most cases (except random or fractal images), the image exhibits a (peaked) spatial frequency content constrained to low frequencies. By contrast, the m-sequence content is almost perfectly white. Therefore, as demonstrated by Fig.1., high pass filtering can reduce image related artefacts, without significantly degrading the peak.

3.3 Analysis

(3) requires an appreciation of the significant moments of the cross-correlation. Since the mean is subtracted from the image in the decoding stage, the correlation is that between two zero-mean functions and hence is itself zero. The variance, however is not so easily constrained. It is the main source of high cross-correlation peaks. Intuitively, this variance can be calculated by resolving each function into an orthogonal basis and applying random-phase statistics to each component. (A one dimensional analogy of this analysis is presented in the Appendix). The cross-correlation can therefore be expressed as a restricted summation over the m

overlapping components. In the case of a "white" image, the total 2^n-1 components would contribute. Hence, assuming laws of large numbers, the ratio of variances is:

$$\frac{S_w}{S_{mv}} \approx \sqrt{\frac{m}{2^n-1}} \quad \mathbf{1}$$

A more complete analysis of these phenomena is presented in the Appendix.

For example, a linear image of 512 pixels can be expressed as a summation of 32 DCT components. The improvement offered by "whitening" is approximately a factor of 4.

It is not necessary to modify the image in order to implement this "whitening" process. It can just as easily be performed by embedding the m-sequence on the image with "random" pixel offsets, or by performing an orderly interleaving operation. The random offsets can be obtained from the m-sequence vector ($n*1$) for one dimensional or ($n*m$) for two dimensional patterns. We are presently investigating and comparing the performance of this technique against that of high-pass filtering.

4 Two-dimensional m-sequence based arrays

M-Sequences can be formed from starting vectors by a Fibonacci recursion relation. They are of maximal length (2^n-1) for a vector of length n . The autocorrelation function of an m-sequence is two valued: 2^n-1 (in phase), -1 (out of phase).

4.1 Extension of one dimensional arrays

A two-dimensional construction can be performed using a row by row phase shift. The effect on columns is that of decimation. Unique phase shifts as determined from Galois Field theory lead to the formation of columns, which are themselves m-sequences. The resulting array is an unbalanced Hadamard Matrix. Alternatively, a long sequence can be folded diagonally into an array format [5]. In this manner, the desirable one-dimensional autocorrelation property can be extended to two dimensions. The encoding and decoding performance of the Hadamard technique suffers from the image related effects because the correlations are performed on the (short and thus interference prone) row or column basis. The folded m-sequence is more immune to these effects, owing to its increased length. However, its information storage capacity is inferior. Watermarks encoded by both methods are presented, compared and analysed in the paper.

4.2 Intrinsic 2D constructions

We have also studied other fundamentally two-dimensional constructions. Costas Arrays are optimal in that their out-of-phase autocorrelation is minimum for shifts in either or both dimensions [6]. (Uniformly low sidelobe point-spread-function). They have been successfully deployed in radar and sonar, where time delays and frequency shifts (Doppler) can occur simultaneously. However, they are highly unbalanced and therefore prone to image related artefacts. Perfect Maps are constructions, where every $m*n$ basis vector occurs once in a large pattern or map and hence can be used for automatic location. (An m-sequence is a one dimensional example of this category). The construction algorithm for Perfect Maps of large dimensions, commensurate with our image sizes is complicated [7]. However, some perfect maps are also Hadamard Matrices. We have examined examples of these, but still found them to be inadequate at rejecting image related artefacts. Legendre sequences and modified Legendre sequences, which are based on a quadratic residue (modulo n) and are similar to m-sequences of non-maximal length are also being studied. They are expected to improve on m-sequences for short lengths only. Extended m-sequences are attractive because they are commensurate with the image size (2^n). Whenever the extension by adding a zero to the m-sequence is performed to the longest run length of zeros, the resulting sequence still exhibits a strong in-phase autocorrelation peak of 2^n . This peak is surrounded by n zero values on either side, making it easy to recognise by filtering techniques. However, this is at the expense of numerous sidelobes at other phase shifts. The effect of these is being investigated. Gold Codes are linear additions of a preferred pair of m-sequences in with a prescribed relative phase shift. Alternatively, they can be viewed as sequences generated by a non-maximal feedback configuration shift register constructed to implement a product of the individual m-sequence generating polynomials. The family of codes generated by all the relative phase shifts and the original parent m-sequences in 2^n+1 , of which approximately half are balanced. The auto and cross correlations are constrained to approximately $2^{n/2}$. These linear codes can be folded into array format, just as m-sequences. They offer greater information storage capacity because of their great diversity and constrained correlations. Gold codes can also be extended to length 2^n in a similar manner to m-sequences.

5 Conclusion

This paper presents a method of encoding and the recovery of a two-dimensional digital water mark on

test images. The performance of the recovery process is analysed and improvements suggested.

6 References

- [1] A.Z.Tirkel, G.A.Rankin, R.M.van Schyndel, W.J.Ho, N.R.A.Mee, C.F.Osborne. Electronic Water Mark. DICTA-93 Macquarie University, Sydney, December 1993. p.666-672.
- [2] R.G. van Schyndel, A.Z.Tirkel, N.R.A.Mee, C.F.Osborne. A Digital Watermark. First IEEE Image Processing Conference, Houston TX, November 15-17, 1994, vol II, p.86-90.
- [3] S.Walton. Image Authentication for a Slippery New Age. Dr.Dobb's Journal, April 1995. p.18-26, 82-87.
- [4] F.M.Boland, J.K.K. Ó Rouanaidh and C.Dautzenberg. Watermarking Digital Images for Copyright Protection.
- [5] F.J. MacWilliams and N.J.A.Sloane. Pseudo-random Sequences and Arrays. Proc.IEEE, vol 64, 1715-1729, Dec.1976.
- [6] S.W.Golomb and H.Taylor. Two-Dimensional Synchronization Patterns for Minimum Ambiguity. IEEE Trans. on Information Theory, vol IT-28, no.4, p.600-604, July 1982.
- [7] T.Etzion. Construction for Perfect Maps and Pseudorandom Arrays. IEEE Trans. on Information Theory, vol 34, no 5, p.1308-1317. September 1988.

APPENDIX

Consider, for the sake of simplicity, a 512 pixel, one dimensional image $I(x)$ and an encoding m-sequence $M(x)$. This analysis can be extended to two dimensions. Both functions can be expressed as Fourier sums:

$$I(x) = \sum_{n=1}^{2^9-1} I_n \cos(\omega_n x + f_n) \quad (\text{A1})$$

and since the m-sequence distribution is white,

$$M(x) = \sum_{n=1}^{2^9-1} m \cos(\omega_n x + f_n) \quad (\text{A2})$$

The encoding process can be described by:

$$W(x) = I(x) + cM(x) \quad (\text{A3})$$

where, in our case, c corresponds to LSB scaling

The decoding process involves the following:

1. Remove mean.
2. Perform correlation with a reference m-sequence in particular phase.
3. Repeat for all m-sequence phases.
4. Compute global correlation maximum and record m-sequence phase.

The correlation can be expressed as:

$$Y(f_m) = Y_m c + \sum_{i=0}^{2^9-1} \sum_{n=1}^{2^9-1} I_n m \cos(\omega_n x_i + f_n) \quad (\text{A4})$$

where Y_m is the two-valued autocorrelation function of the m-sequence (2^9-1 for $m=n$, -1 otherwise). The first term contains information in n or the choice of sequence, whilst the summation term represents the interference (cross correlation with the image content). The value of m is determined by locating the maximum in $Y(\phi_m)$. In order to avoid false positive identification, missed detection and ambiguities, the image crosscorrelation peaks must be smaller than that of the m-sequence. This cross-correlation has zero mean, but its variance can be estimated as:

$$S_Y = \langle Y(f_m) \rangle \approx \left[\sum_{i=0}^{2^9-1} \sum_{n=1}^{2^9-1} \frac{1}{2} I_n^2 \right]^{1/2} \quad (\text{A5})$$

Where the random phase approximation and the laws of large numbers have been implied. A value of $\sigma_Y < (2^9-1)/6$ will guarantee correct detection unconditionally for images with Gaussian statistics.

Most physical images contain only a few significant spectral components (L), around 0 frequency. Assuming a rectangular distribution,

$$S_Y \approx S_I \frac{2^{\frac{9}{2}}}{L^2} \quad (\text{A6})$$

where σ_I is the image variance, whose maximum is approximately $256/6 \approx 43$ for a 256 gray scale gaussian image.

There are two obvious methods of minimising σ_Y . High pass filtering the encoded image before performing the correlation will render the second term in (A4) negligible. However, the filter cut-in frequency is image-dependent. Also, this method introduces a degradation in the m-sequence autocorrelation peak (peak erosion) and an increase in sidelobe levels. There are also effects due to image content beyond the filter cut-in frequency (image leakage). Nevertheless, this method is capable of increasing the peak-to-sidelobe ratio by a factor of 2.

The second method of relative "whitening" described in the main text does not suffer from the above deficiencies, but does affect the spatial properties of the m-sequence. It is not clear if both techniques can be used in cascade, nor if they are commutative.