Families of Matrices with Good Auto and Cross-Correlation

T.E. Hall
Department of Mathematics and Statistics
Monash University
P.O. Box 28M Victoria 3800
Australia
Email: tom.hall@sci.monash.edu.au

C.F. Osborne
Department of Physics
Monash University
P.O. Box 28M Victoria 3800
Australia
Email: charles.osborne@sci.monash.edu.au

A.Z. Tirkel
Department of Mathematics and Statistics
Monash University
P.O. Box 28M Victoria 3800
Australia
Email: atirkel@bigpond.net.au

ABSTRACT. We construct a family of $p-1$ square $p \times p$ matrices (p is any prime) whose periodic cross-correlation values are uniformly $-p$, 0, $+p$ between all pairs of the matrices in the family. For every one of the matrices in the family, all the off-peak autocorrelation values are $-p$ and 0, while the single peak value is $p(p-1)$. For $p = 127$ (where the values $-p$, 0, $+p$ are below 1% of the size p^2 of the matrices) utilization of this construction has resulted in the superimposed embedding of twelve of the matrices (as watermarks) in the standard image "Lenna" and their subsequent retrieval without recourse to the unmarked image.

ARS COMBINATORIA 61(2001), pp. 187-196

1 Introduction and background

The theme of this paper is the production of families of matrices suitable for watermarking of digital images, an area introduced by our group in 1993 [18]. Digital watermarking requires families of matrices with a high single autocorrelation peak, together with good balance, autocorrelation and cross-correlation. Since image format varies, ranging from 8 bit greyscale to sophisticated colour schemes, the watermark alphabet of symbols should be compatible. Traditionally, sequences have been studied over an alphabet composed of roots of unity. A special case is that of the binary alphabet, where only two roots of unity are involved, -1 and +1. This is ideally suited to watermarking of greyscale images, although it is possible to embed and recover watermarks composed of complex roots of unity in colour and greyscale images, as shown by van Schyndel *et al* [20]. Our construction of arrays and theorems are valid for any alphabet whatsoever, in particular for an alphabet of complex roots of unity.

Sequences, over finite alphabets, with balance and two-valued autocorrelation have found applications in communications and radar: good expositions are given by Golomb [7] and Schroeder [17]. More recently, applications have been made in many areas of instrumentation and physical measurement: Cathignol *et al* [4], Eysholdt *et al* [5].

Matrices (arrays) with window properties, namely perfect maps, have been applied to structured light for spatial indexing of medical images, as described by Ozturk *et al* [16]. In vision research Bearse and Sutter [1] combine arrays in the spatial domain and sequences in temporal domain to develop the multifocal retinogram. Costas arrays have been used in radar and sonar to remove ambiguities [8].

Perfect arrays (arrays for which all the off-peak periodic autocorrelation values are zero, while the single peak is of order the size of the array) are presented and analysed by Lüke [12] and Lüke *et al* [13] and by Bömer and Antweiler [2]. Perfect arrays and uniformly redundant arrays [9] have found application in coded aperture astronomy (see Kopilovich [11]).

Of the above array constructions, only Costas arrays involve families of arrays with good cross-correlation (Maric [15]). However, their autocorrelation peak is low, because of the sparse nature of such arrays. Another method of array construction involves the folding of sequences along the diagonal of a non-square matrix, as illustrated by MacWilliams and Sloane [14]. The sequences must be of composite length in order to fill the matrix in a single pass.

The family of sequences with optimal auto and cross-correlation as determined by the Welch bound is the small Kasami set (and its generalization to No/Kumar sequences). Arrays generated by this method are restricted to the format $(p^n + 1) \times (p^n - 1)$ over p roots of unity (p prime). Examples

are shown by Green *et al* [10].

In this paper, we provide a construction of new families of matrices with a high single autocorrelation peak, and with balance, good autocorrelation and cross-correlation. The family size and the auto and cross-correlation of the family members are the same as that of the folded small Kasami set. However, the arrays are available in square format $p \times p$ for any prime p and the alphabet is flexible.

2 Summary

In Section 3, Preliminaries, we give the definitions for sequences, and then for matrices, of autocorrelation and cross-correlation (all correlations in this paper are even periodic). Correlation is the most important measure of families of matrices for our purposes in their application as watermarks on digital images.

Section 4 gives the basic construction, for each prime p and for each pseudonoise sequence as a seed column, of a new family containing $(p-1)$ square $p \times p$ matrices with good cross-correlation and good autocorrelation. As in the method of construction of some perfect maps [16], our method starts with a suitable seed column and generates the further columns by a sequence of cyclic shifts. Suitable seed columns include all sequences with good balance and autocorrelation. A theoretical treatment of these sequences is given by Everett [6]. When the seed column has two-valued autocorrelation, our constructed matrices have three-valued auto and cross-correlation. The alphabet of the seed column becomes the alphabet of the array. Each matrix is obtained from the seed column by using a sequence of cyclic shifts of that seed to obtain the successive columns of the matrix.

3 Preliminaries

Take any column $\mathbf{c} = (c_i)$ of p symbols $c_0, c_1, \ldots, c_{p-1}$ (these symbols could themselves be rows, matrices or higher dimensional arrays of symbols). For each integer k, we define the k shift (or rotation) $\mathbf{c}^k$ of $\mathbf{c}$ to be the column

$$\mathbf{c}^k = (c_i^k), \text{ where } c_i^k = c_{i+k}$$

and $i + k$ is calculated modulo p. Thus the entries of $\mathbf{c}^k$ in order are $(c_k, c_{k+1}, \ldots, c_{k+p-1})$ where the subscripts are calculated modulo p. We say the column $\mathbf{c}$ has period p if $\mathbf{c} = \mathbf{c}^k$ happens only when k is a multiple of p.

Shifts (or rotations) of matrices are defined similarly, as follows: for any $p \times q$ matrix $\mathbf{A} = (a_{ij})$, $i = 0, 1, \ldots, p-1$, $j = 0, 1, \ldots, q-1$, and any integers k, l, the shifted matrix of $\mathbf{A}$, $\mathbf{A}^{(k,l)} = (a_{i,j}^{(k,l)})$ is defined by:

$$a_{i,j}^{(k,l)} = a_{i+k,j+l}$$

where $i + k$ is calculated modulo p and $j + l$ is calculated modulo q. Informally, we can say that there has been a horizontal rotation to the left by l columns and a vertical rotation upwards by k rows (in either order). Likewise, we say that $\mathbf{A}$ has period (p, q) if $\mathbf{A} = \mathbf{A}^{(k,l)}$ only if k is a multiple of p and l is a multiple of q.

For two columns $\mathbf{a} = (a_i)$, $\mathbf{b} = (b_i)$ of the same length p, of symbols, the dot (or scalar) product $\mathbf{a} \cdot \mathbf{b}$ is defined by

$$\mathbf{a} \bullet \mathbf{b} = \sum_{i=0}^{p-1} a_i \cdot b_i$$

where $a_i \cdot b_i$ is some scalar (number) previously defined in terms of the symbols a_i, b_i (this allows a_i, b_i to be real numbers, complex numbers, sequences or arrays). The (even periodic) cross-correlation sequence of $\mathbf{a}$ and $\mathbf{b}$ is the sequence of numbers

$$\mathbf{c}(k) = \mathbf{a} \bullet \mathbf{b}^k, \quad k = 0, 1, 2, \ldots, p-1$$

When $\mathbf{a} = \mathbf{b}$ we call the sequence $\mathbf{c}(k)$ the autocorrelation sequence of $\mathbf{a}$.

Example 3.1: For the binary Legendre sequence $(0, 1, 1, -1, 1, -1, -1)$, the autocorrelation sequence $c(k)$ is $(6, -1, -1, -1, -1, -1, -1)$.

For two $p \times q$ matrices $\mathbf{A} = (a_{ij})$, $\mathbf{B} = (b_{ij})$ their dot (or scalar) product is

$$\mathbf{A} \bullet \mathbf{B} = \sum_{i=0}^{p-1} \sum_{j=0}^{q-1} a_{ij} \cdot b_{ij}$$

where $a_{ij} \cdot b_{ij}$ is some scalar (number) previously defined in terms of the symbols a_{ij}, b_{ij}.

The number $\mathbf{A} \cdot \mathbf{B}$ is one of the numbers forming the (even periodic) cross-correlation between $\mathbf{A}$ and $\mathbf{B}$: the complete list of numbers is the function (or $p \times q$ matrix)

$$\mathbf{c}(k, l) = \mathbf{A} \bullet \mathbf{B}^{(k,l)}, \quad k = 0, 1, \ldots, p-1, \quad l = 0, 1, \ldots, q-1$$

When $\mathbf{A} = \mathbf{B}$ we call $c(k, l)$ the autocorrelation of $\mathbf{A}$. Cross and autocorrelation of matrices were introduced by Calabro and Wolf [3].

We call a sequence or matrix balanced if all non-zero alphabet symbols occur equally often.

4 $p \times p$ Construction

Construction 4.1: For a seed column $\mathbf{c}$ of length p and for $m = 1, 2, \ldots, p-1$ we construct a $p \times p$ matrix $\mathbf{A}_m$ with columns

$$c_0 = c, c_1 = c_0, c_2 = c_1^m, c_3 = c_2^{2m}, \ldots, c_{j+1} = c_j^{mj} \quad j = 0, 1, \ldots, p-1.$$

We write $\mathbf{A}_m = (c_0, c_1, \ldots, c_{p-1})$. We call mj the (differential) shift from c_j to c_{j+1}. The (accumulated) shift from c_0 to c_j is

$$0 + m + 2m + \cdots + (j-1)m = \frac{mj(j-1)}{2}$$

calculated modulo p.

Theorem 4.2. (Cross-correlation between $\mathbf{A}_m$ and $\mathbf{A}_{m'}$, $m \neq m'$)

(a) *If $p \geq 3$ is prime, $m \neq m'$, $m, m' \in \{1, 2, \ldots, p-1\}$, then $\mathbf{A}_m = (\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{p-1})$ and any shift $\mathbf{A}_{m'}^{(k,l)} = (\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{p-1})$ of $\mathbf{A}_{m'}$ agree on at most two columns, that is, $\mathbf{c}_j = \mathbf{b}_j$ for at most two values of j in $\{0, 1, \ldots, p-1\}$.*

(b) *If the seed column $\mathbf{c}$ is a (binary) Legendre sequence of length p, with autocorrelation values $p-1, -1, -1, \ldots, -1$ then the cross-correlation values of each distinct pair of matrices $\mathbf{A}_m$ and $\mathbf{A}_{m'}$ say of the family $\mathcal{A}_p = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{p-1}\}$ are $-p$, 0, $+p$ (when 0, 1, 2 columns agree respectively).*

Proof: (a) Column Agreements.
We find an equation (about accumulated shifts) equivalent to $\mathbf{c}_j = \mathbf{b}_j$. We know $\mathbf{c}_j = \mathbf{c}^{mj(j-1)/2}$. Likewise, column $j + l$ of $\mathbf{A}_{m'}$ is $\mathbf{c}^{m'(j+l)(j+l-1)/2}$. Thus the j column of $\mathbf{A}_{m'}^{(k,l)}$ is $\mathbf{b}_j = \mathbf{c}^n$ where $n = m'(j+l)(j+l-1)/2 + k$.

Since $\mathbf{c}$ has period p, we have that $\mathbf{c}_j = \mathbf{b}_j$ if and only if

$$\frac{mj(j-1)}{2} = \frac{m'(j+l)(j+l+1)}{2} + k \pmod{p} \qquad (1)$$

Multiply Equation (1) by 2 and collect j^2 and j terms:

$$(m'-m)j^2 + (m + m'(2l-1))j + (m'l(l-1) + 2k) = 0 \pmod{p} \quad (2)$$

Now $m' - m \neq 0 \pmod{p}$ so the above quadratic equation is not of the form $0 = 0$, and so has at most two solutions for j in $\{0, 1, 2, \ldots, p-1\}$, which completes the proof (recall that a polynomial equation $a_0 + a_0 x + \cdots + a_0 x^n = 0$, with coefficients a_i from a field F, has at most n solutions in F).

(b) Correlation Calculation
When 0 columns agree, each pair of columns contributes -1, giving a total of $-p$. When 1 pair of columns agree, they contribute $p-1$, and the other $p-1$ pairs of columns contribute -1 each, giving a total of $(p-1) + (p-1)(-1) = 0$. When 2 pairs of columns agree, those two pairs contribute $2(p-1)$ and the remaining $p-2$ pairs contribute -1 each, giving a total of $2(p-1) - (p-2) = p$.

Theorem 4.3. (Autocorrelation of $\mathbf{A}_m$)

(a) *If $p \geq 3$ is prime, and the column* $\mathbf{c}$ *has period* p*, then for* $m = 1, 2, \ldots, p-1$*, the matrix* $\mathbf{A}_m = (\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{p-1})$ *and each non-equal shift* $\mathbf{A}_m^{(k,l)} = (\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{p-1})$ *agree on at most one column, that is,* $\mathbf{c}_j = \mathbf{b}_j$ *for at most one value of* j *in* $\{0, 1, \ldots, p-1\}$.

(b) *If again the seed column* $\mathbf{c}$ *is a (binary) Legendre sequence of length* p*, with autocorrelation values* $(p-1), -1, -1, \ldots, -1$ *then the autocorrelation values of each matrix* $\mathbf{A}_m$ *in the family* $\mathcal{A}_p = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{p-1}\}$ *are* $-p$, 0, $p(p-1)$ *(when* 0, 1 *or all* p *columns agree respectively).*

Proof: (a) Column Agreements
We obtain an equation (about shifts) which is equivalent to $\mathbf{c}_j = \mathbf{b}_j$. We know that $\mathbf{c}_j = \mathbf{c}^{mj(j-1)/2}$. Likewise, $\mathbf{c}^{m(j+l)(j+l-1)/2}$ is the j-column of $\mathbf{A}_m^{(0,l)}$.

The j-column of $\mathbf{A}_m^{(k,l)}$ is thus $\mathbf{b}_j = c_{j+l}^k = c^n$ where $n = \frac{m(j+l)(j+l-1)}{2} + k$.

Since c has period p, we have that $\mathbf{c}_j = \mathbf{b}_j$ if and only if

$$\frac{mj(j-1)}{2} = \frac{m(j+l)(j+l-1)}{2} + k \pmod{p} \tag{3}$$

Now

$$\begin{aligned}(j+l)(j+l-1) &= j(j-1) + jl + l(j-1) + l^2 \\ &= j(j-1) + l(2j-1+l)\end{aligned}$$

so Equation (3) becomes

$$\frac{ml(2j-1+l)}{2} + k = 0 \pmod{p} \tag{4}$$

We note that the excluded case $k = 0 = l \pmod{p}$ would make Equation (4) have the form $0 = 0$, with solutions $i = 0, 1, 2, \ldots, p-1$ (meaning $\mathbf{A} = \mathbf{A}^{(k,l)}$)

If $k \neq 0 = l \pmod{p}$, there are no solutions of (4) for j, and if $l \neq 0 \pmod{p}$ then the unique solution in $\{0, 1, \ldots, p-1\}$ is

$$j_0 = \frac{-\frac{2k}{ml} - l + 1}{2} \pmod{p}$$

calculated in the field $\mathbf{Z}_p$ (where 2, m and l have multiplicative inverses). This completes the proof.

(b) Correlation Calculation
This is proved similarly to part (b) of Theorem 4.2.

Remark 4.4: As well as Legendre sequences, suitable binary seed columns for the construction include m-sequences for Mersenne primes, Hall sequences, biquadratic residue sequences, octic residue sequences etc. [6], [7].

The seed sequences need not be confined to a binary alphabet. Any pseudonoise sequence over roots of unity would also be suitable for the construction.

Examples 4.5: Two small examples will make the construction clear. We take for the seed column $\mathbf{c}$ the sequence of Example 3.1, namely the Legendre sequence for $p = 7$, namely $0\,1\,1\,-1\,1\,-1\,-1$. From the family $\mathcal{A}_7 = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_6\}$, we give the matrices $\mathbf{A}_1$, with a relative shift sequence $1\,2\,3\,4\,5\,6\,0$ and $\mathbf{A}_2$ with a relative shift sequence $2\,4\,6\,1\,3\,5\,0$. We have rotated the shift sequences one position to obtain symmetry between left and right in the matrices.

0	1	-1	-1	-1	1	0
1	1	1	0	1	1	1
1	-1	-1	1	-1	-1	1
-1	1	-1	1	-1	1	-1
1	-1	0	-1	0	-1	1
-1	-1	1	1	1	-1	-1
-1	0	1	-1	1	0	-1

Matrix $\mathbf{A}_1$: $m = 1$

0	1	-1	-1	-1	1	0
1	-1	0	-1	0	-1	1
1	1	1	0	1	1	1
-1	-1	1	1	1	-1	-1
1	-1	-1	1	-1	-1	1
-1	0	1	-1	1	0	-1
-1	1	-1	1	-1	1	-1

Matrix $\mathbf{A}_2$: $m = 2$

The cross-correlation matrix for these two arrays is:

$$\begin{bmatrix} 7 & 7 & 7 & 7 & 7 & 0 & 7 \\ -7 & 7 & 7 & -7 & 0 & -7 & 0 \\ 7 & 0 & 0 & 7 & -7 & -7 & -7 \\ 7 & -7 & -7 & 7 & -7 & 7 & -7 \\ 0 & -7 & -7 & 0 & 7 & -7 & 7 \\ -7 & 7 & 7 & -7 & -7 & 7 & -7 \\ -7 & -7 & -7 & -7 & 7 & 7 & 7 \end{bmatrix}$$

Entries of -7 correspond to no columns matching, those of 0 correspond to one column matching and those of 7 correspond to two columns matching. For any pair of members of the family $\mathcal{A}_7 = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_6\}$, the entries in their cross-correlation matrix are likewise $-7, 0, 7$.

The autocorrelation matrices of $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_6$ are equal and are

$$\begin{bmatrix} 42 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \\ -7 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Each 0 entry corresponds to one column agreeing, and each -7 entry corresponds to no columns agreeing. The pattern shown is general, in the sense that for each prime $p \geq 3$ Equation (4) has (a) one solution for non-zero horizontal shift and any vertical shift (b) no solution for zero horizontal shift and non-zero vertical shift and (c) full agreement for zero shifts in both directions.

5 An Application

In [19] the authors, together with I. Svalbe and R. van Schyndel, report on an embedding of the arrays discussed above as watermarks in a digital image of "Lenna". We took $p = 127 = 2^7 - 1$ and chose one of the eighteen maximal length sequences (m-sequences) of length 127, denoted by c, as a seed column for the construction. For each of $m = 1, 2, \ldots, 126$, as the value of the multiplier, we considered the 127×127 array $\mathbf{A}_m$ constructed with c as a seed column and with a sequence of relative shifts: $0, m, 2m, \ldots, 126m \pmod{127}$. Twelve of these matrices were selected with a random number generator, and all twelve superimposed on the image of Lenna on the same 127×127 pixels. Each of these matrices was cyclically shifted in two dimensions. This was done with many choices of the set of twelve arrays. In all cases unambiguous retrieval of each of the twelve watermarks was accomplished by use of each array as a template, and without recourse to the unmarked image. The process automatically retrieves the horizontal and vertical shifts of each of the twelve matrices, which allows us to store 168 ($= 12 \times 14$) bits of information in the composite watermark. The retrieval of the individual arrays was possible because of the low cross-correlation values for each pair of the 126 available matrices, and the low off-peak autocorrelation values of each matrix (less than 1% of the peak).

References

[1] M.A. Bearse and E.E. Sutter, Imaging localized retinal dysfunction with the multifocal electroretinogram, *Journal of the Optical Society of America* **A13** (1996), 634–640.

[2] L. Bomer and M. Antweiler, Two-dimensional perfect binary arrays with 64 elements, *IEEE Trans. Inform. Theory* **36** (1990), 411–414.

[3] D. Calabro and J.K. Wolf, On the synthesis of two-dimensional arrays with desirable correlation properties, *Information and Control* **11** (1968), 537–560.

[4] D. Cathignol, C. Fourcade and J.Y. Chapelon, Transcutaneous blood flow measurement using pseudo-random noise doppler flowmeter, *IEEE Transactions on Biomedical Engineering* **27** (1980), No.1, 30–36.

[5] U. Eysholdt and C. Screiner, Maximum length sequences - a fast method for measuring brain stem auditory evoked responses, *Proc. IEEE Conference: Frontiers of Engineering in Health Care*, Houston 1981, 306–309.

[6] D. Everett, Periodic digital sequences with pseudonoise properties, *G.E.C. Journal*, **33**, No.3, (1966), 115–126.

[7] S.W. Golomb, *Shift Register Sequences*, Holden Day, San Francisco, 1967.

[8] S.W. Golomb and H. Taylor, Construction and properties of Costas arrays, *Proceedings IEEE*, **72**, No. 9, September 1984, 1143–1163.

[9] S.R. Gottesman and E.E. Fenimore, New family of binary arrays for coded aperture imaging, *Applied Optics*, **28**, No. 20, 15 October 1989, 4344–4352.

[10] D.H. Green and S.K Amarasinghe, Families of sequences and arrays with good periodic correlation properties, *IEE proceedings-E*, **138**, No. 4, 1991, 260–268.

[11] L.E. Kopilovich, On perfect binary arrays, *Electron Lett.* **24** (1988), 566–567

[12] H.D. Luke, Sequences and arrays with perfect periodic correlation, *IEEE Trans. Aerospace and Electronic Systems* **24** (1988), 287–294.

[13] H-D. Luke, L. Bomer and M. Antweiler, Perfect binary arrays, *Signal Processing* 17 (1989), Elsevier Science Publishing, 69–80.

[14] F.J. MacWilliams and N.J.A. Sloane, Pseudo-random sequences and arrays, *Proc. IEEE* 64 (1976), 1715–1729.

[15] S.V. Maric, I. Seskar and E.L. Titlebaum, On Cross-Ambiguity Properties of Welch-Costas Arrays, *IEEE Trans. Aerospace and Electronic Systems* 30 (1994) no. 4, 19–27.

[16] C. Ozturk, J. Nissanov and S. Dubin, Generation of perfect map codes for an active stereo imaging system, 22nd IEEE Annual Northeast Bioengineering Conference, Rutgers University, New Brunswick, N.J., March 14-15, 1996, 76–77.

[17] M.R. Schroeder, *Number Theory in Science and Communication*, Third Edition, Springer (1997) ISBN 3-540-62006-0.

[18] A.Z. TirkeL G.A. Rankin, R.M. van Schyndel, W.J. Ho, N.R.A. Mee, and C.F. Osborne, Electronic water mark, Proceedings of DICTA-93, Macquarie University, Sydney, December 1993, 666–673.

[19] R. van Schyndel, A.Z. Tirkel, I.D. Svalbe, T.E. Hall, C.F. Osborne, Algebraic Construction of a New Class of Quasi-Orthogonal Arrays in Steganography, SPIE Electronic Imaging 1999, San Jose, USA, 354–364.

[20] R. van Schyndel, A.Z. Tirkel, I.D. Svalbe, A Multiplicative Color Watermark, IEEE-EURASIP Workshop on Non-Linear Signal and Imaging Processing, Antalya, Turkey, 1999, 336–340.