

- [7] G. Carayannis *et al.*, "A fast sequential algorithm for least-squares filtering and prediction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, Dec. 1983.
- [8] J. M. Cioffi and T. Kailath, "Fast recursive least squares transversal filters for adaptive processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, 1984.
- [9] L. B. Jackson and S. L. Wood, "Linear prediction in cascade form," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 518–528, Dec. 1978.
- [10] M. Nayeri and K. Jenkins, "Alternate realizations to adaptive IIR filters," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 485–496, Apr. 1989.
- [11] U. Forssen, "Analysis of adaptive FIR filters in cascade form," *IEEE Trans. Circuits Syst. II*, vol. 41, pp. 392–401, June 1994.
- [12] J. M. Raulin *et al.*, "A 60 channel PCM-ADPCM converter," *IEEE Trans. Commun.*, vol. 30, pp. 567–573, Apr. 1982.
- [13] W. J. Song and M. S. Park, "A complementary pair LMS algorithm for adaptive filtering," in *Proc. ICASSP, 1997*, vol. 3, pp. 2261–2264.
- [14] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

DCT/DST and Gauss–Markov Fields: Conditions for Equivalence

José M. F. Moura and Marcelo G. S. Bruno

Abstract—The correspondence addresses the intriguing question of which random models are equivalent to the discrete cosine transform (DCT) and discrete sine transform (DST). Common knowledge states that these transforms are asymptotically equivalent to first-order Gauss causal Markov random processes. We establish that the DCT and the DST are exactly equivalent to homogeneous one-dimensional (1-D) and two-dimensional (2-D) Gauss noncausal Markov random fields defined on finite lattices with appropriate boundary conditions.

I. INTRODUCTION

In this correspondence, we establish the second-order equivalence between the discrete sine transform (DST) and the discrete cosine transform (DCT) and arbitrary order noncausal Gauss–Markov random fields (GMrf's) defined on a finite lattice. We prove this by showing that the DST and the DCT diagonalize the covariance matrix associated with these fields. Following [1], we work with the inverse of the covariance matrix, which is called the potential matrix, that is highly structured; for homogeneous noncausal GMrf's of arbitrary order, it is given by a Toeplitz canonical matrix plus a boundary matrix.

Section II expresses the Toeplitz component of the potential matrix as matrix polynomials that are diagonalizable by either the DST or the DCT plus a perturbation matrix. Section III shows that for a given arbitrary order one-dimensional (1-D) GMrf, particular choices of boundary conditions (bc's) lead to a boundary matrix that cancels the perturbation term in the expansion of the Toeplitz canonical matrix. The final result is then an overall potential matrix that is

Manuscript received September 24, 1996; revised March 6, 1998. This work was supported in part by DARPA under Grant DABT 63-98-1-0004. The work of M. G. S. Bruno was also supported in part by CNPq-Brazil. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jitendra K. Tugnait.

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213-3890 USA.

Publisher Item Identifier S 1053-587X(98)05966-2.

diagonalizable by either the DST or the DCT. Results are extended to two-dimensional (2-D) GMrf's in Section IV using the Kronecker product. The same techniques can be used to show similar results for the other sinusoidal transforms introduced in [2].

II. NOTATION AND PRELIMINARIES

Let $\mathbf{e}_1 = [1 \ 0 \ \cdots \ 0]^T$ and, for $1 \leq i \leq N$, \mathbf{e}_i be a zero vector except for entry i , which is a 1. We define the following $N \times N$ matrices whose entries are zero, except as indicated:

- reflection matrix \mathbf{J} , with counter diagonal of ones;
- forward shift \mathbf{K}_1 , with first upper diagonal of ones;
- backward shift $\mathbf{K}_2 = \mathbf{K}_1^T$, with lower diagonal of ones;
- the powers of the shift operators \mathbf{K}_1^i and \mathbf{K}_2^i , $1 \leq i \leq N-1$, with the i th-upper, or, respectively, i -lower, diagonal of ones;
- $\mathbf{F}_0 = \mathbf{0}$, and symmetric matrices $\mathbf{F}_i = \mathbf{J}(\mathbf{K}_1^{N-i} + \mathbf{K}_2^{N-i})$, $0 \leq i \leq N-1$, with the $(i-1)$ th lower and upper counter diagonals of ones;
- $\mathbf{H} = \mathbf{K}_1 + \mathbf{K}_2$, with the first upper and lower diagonal of ones;
- $\bar{\mathbf{H}}$ is like \mathbf{H} , with, in addition, the entries $(1, 1)$ and (N, N) , which are also ones.

Eigenstructure of \mathbf{H} and $\bar{\mathbf{H}}$: Matrices \mathbf{H} and $\bar{\mathbf{H}}$ are symmetric tridiagonal matrices. Their eigenstructure is well known, e.g., [3]. The eigenvectors of \mathbf{H} are the basis vectors of the DST, and the eigenvectors of $\bar{\mathbf{H}}$ are the basis vectors of the DCT. In other words, their eigenvectors are the rows of the orthogonal transform matrices \mathbf{S} and \mathbf{C} for the DST and DCT, respectively, which are defined as

$$S_{kn} = \left[\sqrt{\frac{2}{N+1}} \sin \frac{\pi(k+1)(n+1)}{N+1} \right]_{0 \leq k, n \leq N-1} \quad (1)$$

$$C_{kn} = \begin{cases} \sqrt{\frac{1}{N}}, & k=0, 0 \leq n \leq N-1 \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)k}{2N}, & 1 \leq k \leq N-1 \\ & 0 \leq n \leq N-1. \end{cases} \quad (2)$$

Toeplitz Matrices: We decompose a banded symmetric $N \times N$ Toeplitz matrix as

$$\forall m \leq N: \quad \mathbf{T} = b_1 \mathbf{I} + b_2 (\mathbf{K}_1 + \mathbf{K}_2) + b_3 (\mathbf{K}_1^2 + \mathbf{K}_2^2) + \cdots + b_m (\mathbf{K}_1^{m-1} + \mathbf{K}_2^{m-1}). \quad (3)$$

Powers $[\mathbf{K}_1^{i-1} + \mathbf{K}_2^{i-1}]$: The following lemma relates $(\mathbf{K}_1^i + \mathbf{K}_2^i)$ to the powers of \mathbf{H}^i and of $\bar{\mathbf{H}}^i$.

Lemma II.1: Let $\mathbf{F}_{-1} = \mathbf{F}_0 = \mathbf{0}$ and \mathbf{F}_i , $i \geq 1$ be defined as before. Then

$$0 \leq i \leq N-1: \quad \mathbf{K}_1^i + \mathbf{K}_2^i = P_i(\mathbf{H}) + \mathbf{F}_{i-1} \quad (4)$$

where $P_0(\mathbf{H}) = 2\mathbf{I}$, and $P_i(\mathbf{H})$, $1 \leq i \leq N-1$ are matrix polynomials in \mathbf{H}

$$i \text{ even: } P_i(\mathbf{H}) = \mathbf{H}^i - i\mathbf{H}^{i-2} - \alpha_1^i \mathbf{H}^{i-4} - \cdots - \alpha_{(i-2)/2}^i \mathbf{I} \quad (5)$$

$$i \text{ odd: } P_i(\mathbf{H}) = \mathbf{H}^i - i\mathbf{H}^{i-2} - \alpha_1^i \mathbf{H}^{i-4} - \cdots - \alpha_{(i-3)/2}^i \mathbf{H}. \quad (6)$$

Similarly

$$0 \leq i \leq N-1: \quad \mathbf{K}_1^i + \mathbf{K}_2^i = Q_i(\bar{\mathbf{H}}) - \mathbf{F}_i \quad (7)$$

where $Q_0(\bar{\mathbf{H}}) = 2\mathbf{I}$ and $Q_i(\bar{\mathbf{H}})$, $1 \leq i \leq N-1$ are matrix polynomials in $\bar{\mathbf{H}}$

$$i \text{ even: } Q_i(\bar{\mathbf{H}}) = \bar{\mathbf{H}}^i - i\bar{\mathbf{H}}^{i-2} - \alpha_1^i \bar{\mathbf{H}}^{i-4} - \cdots - \alpha_{(i-2)/2}^i \mathbf{I} \quad (8)$$

$$i \text{ odd: } Q_i(\bar{\mathbf{H}}) = \bar{\mathbf{H}}^i - i\bar{\mathbf{H}}^{i-2} - \alpha_1^i \bar{\mathbf{H}}^{i-4} - \cdots - \alpha_{(i-3)/2}^i \bar{\mathbf{H}}. \quad (9)$$

TABLE I
RECURSION FOR COEFFICIENTS α_i^j

	α_k^i	
	$\alpha_1^3 = 0$	
i odd, $i \geq 3$	$\alpha_1^{i+1} = \alpha_1^i - (i-1)$	$i \geq 3$
	$\alpha_k^{i+1} = \alpha_k^i - \alpha_{k-1}^{i-1} \quad 2 \leq k \leq \frac{i-3}{2}$	$i \geq 7$
	$\alpha_{(i-1)/2}^{i+1} = -\alpha_{(i-3)/2}^{i-1}$	$i \geq 5$
i even, $i \geq 4$	$\alpha_1^{i+1} = \alpha_1^i - (i-1)$	$i \geq 4$
	$\alpha_k^{i+1} = \alpha_k^i - \alpha_{k-1}^{i-1} \quad 2 \leq k \leq \frac{i-2}{2}$	$i \geq 6$

The coefficients α_i^j are computed according to the recursion in Table I.

The proof of Lemma II.1 is found in a technical report available from the authors.

III. DCT/DST AND 1-D GMRF'S

We establish the second-order equivalence between the DCT (or DST) and 1-D noncausal GMrf's by determining the GMrf's for which the DCT (or the DST) is their Karuhnen-Loève transform (KLT). We do this for arbitrary-order, finite-length, spatially homogeneous, noncausal GMrf's with appropriately chosen bc's.

The KLT of a GMrf diagonalizes its covariance matrix and is determined by the eigenvectors of the GMrf covariance matrix. Finding this eigenstructure is, in general, hard because there is no direct parametrization of the GMrf covariance. Reference [1] shows, however, that for arbitrary-order GMrf's on finite lattices, the inverse of the covariance (the potential matrix) has a well-defined parametrized structure. We work with the potential matrix and use this parametrization to show when the DCT or the DST diagonalize the potential matrix, hence, the covariance matrix of the GMrf.

Consider a zero-mean spatially homogeneous m th-order GMrf $\{x(n)\}$ defined on the finite lattice $1 \leq n \leq N$ and described by the finite difference equation

$$x(n) = -a_1[x(n-1) + x(n+1)] - \dots - a_m[x(n-m) + x(n+m)] + v(n) \quad (10)$$

where $\{v(n)\}$ is correlated noise. Stack $x(n)$ and $v(n)$, $1 \leq n \leq N$ in two column vectors \mathbf{x} and \mathbf{v} , and define the vector of bc's \mathbf{x}_{bc}

$$\mathbf{x}_{bc} = [x(0) \dots x(-(m-1)) \dots x(N+1) \dots x(N+m)]^T. \quad (11)$$

With \mathbf{x}_{bc} , we extend (10) up to the boundary. In matrix form, (10), $1 \leq n \leq N$, becomes

$$\mathbf{A}_c \mathbf{x} = \mathbf{G}_{bc} \mathbf{x}_{bc} + \mathbf{v}. \quad (12)$$

By inspection, the matrices \mathbf{A}_c and \mathbf{G}_{bc} are Toeplitz and Hankel, respectively

$$\mathbf{A}_c = \mathbf{I} + \sum_{i=1}^m a_i [\mathbf{K}_1^i + \mathbf{K}_2^i]$$

and

$$\mathbf{G}_{bc} = - \sum_{i=1}^m a_i \mathbf{F}_i. \quad (13)$$

The structure of \mathbf{G}_{bc} follows by writing (10) for $n = 1, \dots, m$ and $n = N - (m-1), \dots, N$ and collecting the terms in the boundary values $x(0), \dots, x(-(m-1)), \dots, x(N+1), \dots, x(N+m)$.

The bc's of interest are homogeneous, i.e.,

$$\mathbf{G}_{bc} \mathbf{x}_{bc} = -\mathbf{A}_{bc} \mathbf{x}. \quad (14)$$

Equation (14) defines \mathbf{A}_{bc} implicitly. Substituting (14) in (12), we get

$$\mathbf{A} \mathbf{x} = \mathbf{v}, \quad \text{with } \mathbf{A} = \mathbf{A}_c + \mathbf{A}_{bc}. \quad (15)$$

We recall from [1] that

- \mathbf{A} potential matrix;
- \mathbf{A}_c canonical potential matrix;
- \mathbf{A}_{bc} bc matrix;

and that the covariance matrix Σ_x of \mathbf{x} is proportional to \mathbf{A}^{-1} . Hence, the DCT or the DST are the KLT of a GMrf when they diagonalize its potential matrix \mathbf{A} . Let $\mathbf{x} = \{x(n): 1 \leq n \leq N\}$ be an m th-order GMrf as in (10).

Theorem III.1—DST and GMrf: The DST is the KLT of the m th-order GMrf \mathbf{x} with bc's

$$1 \leq k \leq m-1: \quad x(0) = 0, \quad x(-k) = -x(k) \quad (16)$$

$$2 \leq k \leq m: \quad x(N+1) = 0, \\ x(N+k) = -x(N-k+2). \quad (17)$$

Theorem III.2—DCT and GMrf: The DCT is the KLT of the m th-order GMrf \mathbf{x} with bc's

$$0 \leq k \leq m-1: \quad x(-k) = x(k+1) \quad (18)$$

$$1 \leq k \leq m: \quad x(N+k) = x(N-k+1). \quad (19)$$

To prove the theorems, we show that the bc's in the theorems lead to an \mathbf{A}_{bc} that cancels a Hankel perturbation term in the expansion of the canonical potential matrix \mathbf{A}_c (see below). This is done by working with (13) and (14).

Proof of Theorem III.1: Equation (13) expresses \mathbf{A}_c , which is a symmetric Toeplitz matrix, as the sum of powers of the matrices \mathbf{K}_i , $i = 1, 2$. From Lemma II.1, substituting (4) in (13), we rewrite \mathbf{A}_c as

$$\mathbf{A}_c = \sum_{i=0}^m d_i a_i P_i(\mathbf{H}) + \sum_{i=2}^m a_i \mathbf{F}_{i-1} \quad (20)$$

where $a_0 = 1$, $d_0 = 1/2$, $d_i = 1$, $1 \leq i$, $P_0(\mathbf{H}) = 2\mathbf{I}$, and $P_1(\mathbf{H}) = \mathbf{H}$. To obtain a global potential matrix \mathbf{A} whose eigenmatrix is the DST matrix, it suffices to choose a bc matrix \mathbf{A}_{bc} that cancels the second term in (20) of \mathbf{A}_c ; in other words

$$\mathbf{A}_{bc} = - \sum_{i=2}^m a_i \mathbf{F}_{i-1} \quad \text{leading to} \\ \mathbf{A} = \mathbf{A}_c + \mathbf{A}_{bc} = \sum_{i=0}^m d_i a_i P_i(\mathbf{H}). \quad (21)$$

With this \mathbf{A}_{bc} , matrix \mathbf{A} is a polynomial in \mathbf{H} and, hence, diagonalizable by the DST matrix.

Homogeneous bc's: In matrix form, the bc's given by (16) and (17) in Theorem III.1 lead to

$$\mathbf{x}_{bc} = -[\mathbf{e}_2 \dots \mathbf{e}_m \mathbf{0} \dots \mathbf{0} \mathbf{e}_{N-m+1} \dots \mathbf{e}_{N-1}] \mathbf{x} \\ = -\mathbf{M}_s \mathbf{x}. \quad (22)$$

bc Matrix \mathbf{A}_{bc} : Substituting (22) in (14) and using the definitions of \mathbf{G}_{bc} in (13) and of \mathbf{M}_s in (22), we get

$$\mathbf{A}_{bc} = -\mathbf{G}_{bc}(-1)\mathbf{M}_s \\ = \mathbf{G}_{bc} \mathbf{M}_s \\ = - \sum_{i=1}^m a_i \mathbf{F}_i [\mathbf{e}_2 \dots \mathbf{e}_m \mathbf{0} \dots \mathbf{0} \mathbf{e}_{N-m+1} \dots \mathbf{e}_{N-1}]. \quad (23)$$

Multiplying out the matrices in (23) leads to \mathbf{A}_{bc} given by (21), proving the theorem.

Proof of Theorem III.2: From Lemma II.1, the canonical matrix \mathbf{A}_c is

$$\mathbf{A}_c = \sum_{i=0}^m d_i a_i Q_i(\overline{\mathbf{H}}) - \sum_{i=1}^m a_i \mathbf{F}_i \quad (24)$$

where $a_0 = 1$, $d_0 = 1/2$, $d_i = 1$, $1 \leq i$, $Q_0(\overline{\mathbf{H}}) = 2\mathbf{I}$, and $Q_1(\overline{\mathbf{H}}) = \overline{\mathbf{H}}$. The matrix \mathbf{A}_{bc} needed to cancel the second term in (24) is

$$\mathbf{A}_{bc} = \sum_{i=1}^m a_i \mathbf{F}_i. \quad (25)$$

Homogeneous bc's: In matrix form, the bc's given by (18) and (19) in Theorem III.2 lead to

$$\begin{aligned} \mathbf{x}_{bc} &= [\mathbf{e}_1 \cdots \mathbf{e}_m \mathbf{0} \cdots \mathbf{0} \mathbf{e}_{N-m+1} \cdots \mathbf{e}_N] \mathbf{x} \\ &= \mathbf{M}_c \mathbf{x}. \end{aligned} \quad (26)$$

bc Matrix \mathbf{A}_{bc} : Replacing \mathbf{x}_{bc} in (14) and again using \mathbf{G}_{bc} in (13) and \mathbf{M}_c in (26)

$$\begin{aligned} \mathbf{A}_{bc} &= -\mathbf{G}_{bc} \mathbf{M}_c \\ &= - \left(- \sum_{i=1}^m a_i \mathbf{F}_i \right) [\mathbf{e}_1 \cdots \mathbf{e}_m \mathbf{0} \cdots \mathbf{0} \mathbf{e}_{N-m+1} \cdots \mathbf{e}_N]. \end{aligned} \quad (27)$$

Multiplying out leads to the matrix \mathbf{A}_{bc} in (25), proving the theorem.

Remark: We make the following observations.

- The DCT bc's in Theorem III.2 are the asymmetric Neumann bc's [1]. In DCT-based standards, the image is segmented in smaller rectangular blocks, which are then independently DCT transformed. The bc's in Theorem III.2 assume symmetry across boundaries, i.e., they ensure continuity across the blocks. They are particularly suited to processing partitioned blocks when the image is smooth.
- The bc's corresponding to the DST in Theorem III.1 are a shifted version of the asymmetric Neumann bc's with an inversion of sign. This suggests that the DST is well suited when there are edges at the boundaries of the image blocks with the picture intensity changing, say, from bright in one block to average at the edge and then to dark in the next block.
- Perturbation matrices similar to (23) and (27) for the 1-D case are also in [4] and [5].

IV. DCT/DST AND 2-D GMRF'S

We consider GMrf's on square lattices $L = \{(i, j): 1 \leq i, j \leq N\}$. We collect the field values¹ $\mathbf{X} = \{x(i, j): (i, j) \in L\}$ along row i in a column vector \mathbf{x}^i and stack these vectors in a long column vector \mathbf{x} . From [1], it follows that the potential matrix \mathbf{A} of the 2-D GMrf is highly sparse with a block structure, which is compactly captured by Kronecker products. The Kronecker product of two matrices A and B corresponds to the matrix $A \otimes B = [a_{ij} B]$; see [3] for properties.

Consider a 2-D (homogeneous) GMrf on a finite lattice, where $\{\eta_{ij} \equiv \eta, (i, j) \in L\}$ is the corresponding *noncausal* 2-D neighborhood system. The MMSE description of the field is

$$1 \leq i, j \leq N: \quad x(i, j) = \sum_{(i-k, j-l) \in \eta} -a(k, l) x(i-k, j-l) + v(i, j). \quad (28)$$

The coefficients $a(k, l)$ are such that they satisfy the *symmetry* conditions $a(-k, l) = a(k, l)$, $a(k, -l) = a(k, l)$, and $a(-k, -l) = a(k, l)$. We also represent $a(k, l)$ by a_k^l .

¹We also represent the field values at pixel (i, j) as x_{ij} .

Let W_p be the maximum Euclidean distance between sites (i, j) and $(i-k, j-l)$ such that $(i-k, j-l) \in \eta$. Collecting (28) in matrix format, we get the following structure for the canonical component (\mathbf{A}_c) of the potential matrix (see [1, Th. 2])

$$\begin{aligned} \mathbf{A}_c &= \mathbf{I} \otimes \mathbf{D}_0 + \mathbf{K}_1 \otimes \mathbf{D}_1 + \mathbf{K}_2 \otimes \mathbf{D}_1^T + \cdots \\ &\quad + \mathbf{K}_1^m \otimes \mathbf{D}_m + \mathbf{K}_2^m \otimes \mathbf{D}_m^T \end{aligned} \quad (29)$$

where \mathbf{D}_0 is centrosymmetric, Toeplitz, and banded; the off-diagonal blocks \mathbf{D}_i are persymmetric, Toeplitz, and banded; and m is the largest integer such that $m \leq W_p$.

We focus attention on arbitrary-order models with *diagonal symmetry*. Diagonal symmetry forces the off-diagonal Toeplitz blocks \mathbf{D}_i to be symmetric. Hence, the canonical component \mathbf{A}_c is

$$\begin{aligned} \mathbf{A}_c &= \mathbf{I} \otimes \mathbf{D}_0 + (\mathbf{K}_1 + \mathbf{K}_2) \otimes \mathbf{D}_1 + \cdots \\ &\quad + (\mathbf{K}_1^m + \mathbf{K}_2^m) \otimes \mathbf{D}_m \end{aligned} \quad (30)$$

$$= \sum_{j=0}^m d_j [\mathbf{K}_1^j + \mathbf{K}_2^j] \otimes \mathbf{D}_j \quad (31)$$

where $d_0 = 1/2$ and $d_j = 1$, $1 \leq j \leq m$. The \mathbf{D}_j 's have the following structure [1]:

$$0 \leq j \leq m: \quad \mathbf{D}_j = \sum_{k=0}^{p(j)} d_k a_j^k [\mathbf{K}_1^k + \mathbf{K}_2^k]. \quad (32)$$

In (31) and (32), m is the largest integer $m \leq W_p$, $p(0) = m$, and $p(j)$, $1 \leq j \leq m$ is such that

$$p^2(j) + j^2 \leq W_p^2. \quad (33)$$

In order to obtain a potential matrix that is diagonalizable by a Kronecker product of sinusoidal orthogonal transform matrices ($\mathbf{C} \otimes \mathbf{C}$ or $\mathbf{S} \otimes \mathbf{S}$), we must add to the canonical *block Toeplitz* \mathbf{A}_c an appropriate perturbation *block Hankel* matrix of bc's. The desired perturbation matrices are

$$\begin{aligned} (\mathbf{A}_{bc})_{\text{DST}} &= - \sum_{j=2}^m \left\{ \mathbf{F}_{j-1} \otimes \left[\sum_{k=0}^{p(j)} d_k a_j^k P_k(\mathbf{H}) \right] \right\} \\ &\quad - \sum_{j=0}^m \left\{ [d_j P_j(\mathbf{H})] \otimes \left[\sum_{k=2}^{p(j)} a_j^k \mathbf{F}_{k-1} \right] \right\} \\ &\quad - \sum_{j=2}^m \left\{ \mathbf{F}_{j-1} \otimes \left[\sum_{k=2}^{p(j)} a_j^k \mathbf{F}_{k-1} \right] \right\} \end{aligned} \quad (34)$$

$$\begin{aligned} (\mathbf{A}_{bc})_{\text{DCT}} &= \sum_{j=1}^m \left\{ \mathbf{F}_j \otimes \left[\sum_{k=0}^{p(j)} d_k a_j^k Q_k(\overline{\mathbf{H}}) \right] \right\} \\ &\quad + \sum_{j=0}^m \left\{ [d_j Q_j(\overline{\mathbf{H}})] \otimes \left[\sum_{k=1}^{p(j)} a_j^k \mathbf{F}_k \right] \right\} \\ &\quad - \sum_{j=1}^m \left\{ \mathbf{F}_j \otimes \left[\sum_{k=1}^{p(j)} a_j^k \mathbf{F}_k \right] \right\} \end{aligned} \quad (35)$$

where $d_0 = 0.5$, $d_j = 1$, $j \geq 1$, $\mathbf{F}_{-1} = \mathbf{F}_0 = \mathbf{0}$, and the polynomials P_k and Q_k are those in (4) and (7), whose coefficients are obtained using the recursions in Table I.

The bc's that correspond to (34) and (35) are in the next two theorems. Proofs are in a technical report.

Theorem IV.1—DST and GMrf: The DST is the KLT of the m th-order symmetric GMrf \mathbf{X} with the following bc's: For $-(m-1) \leq i$, $j \leq N+m$

$$x_{i0} = x_{i(N+1)} = x_{0j} = x_{(N+1)j} = 0$$

$\forall k, k+1 \leq W_p, k > 0:$

$$x_{i(-k)} = -x_{ik} \text{ and } x_{(-k)j} = -x_{kj}$$

$\forall k, k \leq W_p, k > 1:$

$$x_{i(N+k)} = -x_{i(N-k+2)} \text{ and } x_{(N+k)j} = -x_{(N-k+2)j}. \quad (36)$$

Theorem IV.2—DCT and GMrf: The DCT is the KLT of the m th-order symmetric GMrf \mathbf{X} with the following bc's: For $-(m-1) \leq i$, $j \leq N+m$

$\forall k, k+1 \leq W_p, k \geq 0:$

$$x_{i(-k)} = -x_{i(k+1)} \text{ and } x_{(-k)j} = -x_{(k+1)j}$$

$\forall k, k \leq W_p, k \geq 1:$

$$x_{i(N+k)} = -x_{i(N-k+1)} \text{ and } x_{(N+k)j} = -x_{(N-k+1)j}. \quad (37)$$

V. CONCLUSION

We show that the DCT and the DST are statistically *equivalent* to homogeneous *noncausal* GMrf's by establishing that they are the KLT of 1-D or 2-D arbitrary-order homogeneous GMrf's with appropriate bc's. This is an exact equivalence that contrasts with the common argument of *asymptotic* equivalence of the DCT and the DST with certain first-order *causal* Gauss Markov random processes.

The bc's for the DCT assume symmetry across boundaries. Because this ensures continuity across blocks, the DCT is particularly suited to processing partitioned blocks when an image is smooth. The bc's for the DST correspond to intensity changes, which makes it appropriate when the image goes from dark to bright or bright to dark at the boundaries of the image blocks.

REFERENCES

- [1] J. M. F. Moura and N. Balram, "Recursive structure of noncausal Gauss Markov random fields," *IEEE Trans. Inform. Theory*, vol. 38, pp. 334–354, Mar. 1992.
- [2] A. K. Jain, "A sinusoidal family of unitary transforms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, pp. 356–365, Oct. 1979.
- [3] —, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [4] V. Sánchez, P. García, A. Peinado, J. Segura, and A. Rubio, "Diagonalizing properties of the discrete cosine transforms," *IEEE Trans. Signal Processing*, vol. 43, pp. 2631–2641, Nov. 1995.
- [5] V. Sánchez, A. Peinado, J. Segura, P. García, and A. Rubio, "Generating matrices for the discrete sine transforms," *IEEE Trans. Signal Processing*, vol. 44, pp. 2644–2646, Oct. 1996.

Graph Theoretic Techniques for Pruning Data and Their Applications

Tetsuya Hoya

Abstract—In pattern recognition tasks, we usually do not pay much attention to the arbitrarily chosen training set of a pattern classifier beforehand. This correspondence proposes several methods for pruning data sets based upon graph theory in order to alleviate redundancy in the original data set while retaining the original data structure as far as possible. The proposed methods are applied to the training sets for pattern recognition by a multilayered perceptron neural network (MLP-NN) and the locations of the centroids of a radial basis function neural network (RBF-NN). The advantage of the proposed graph theoretic methods is that they do not require any calculation for the statistical distributions of the clusters. The experimental results in comparison both with the k -means clustering and with the learning vector quantization (LVQ) methods show that the proposed methods give encouraging performance in terms of computation for data classification tasks.

Index Terms—Data-pruning algorithms, graph theoretic techniques, pattern classification.

I. INTRODUCTION

Neural networks have played a significant role in pattern classification studies with the rapid development in new computer architectures as well as with advances in vector quantization networks, Bayesian, or nearest neighbor-based classifiers. In pattern recognition tasks using NN's, normally a large amount of data for training NN's are required to generalize the mapping of the input–output relation. This is time consuming. The need for pruning the training data sets is, hence, desirable because the sets usually are not chosen with much care, and there is, therefore, a large possibility of containing redundant information in the training phase. Moody and Darken [1] used the centroids of RBF-NN's obtained by the k -means clustering algorithm. In an earlier work, the k -nearest neighbor (k -NN)-based algorithm [2] was used for the reduction in the size of RBF-NN's [3]. This method, however, involves the estimation of the *a posteriori* distributions, which is one of common problems in statistical analysis. Graph theory provides a number of ways to solve problems in a variety of disciplines and has also been used for clustering [4]–[6]. In a recent paper [7], a method to the data reduction was proposed, which automatically selects the number of exemplars for each class. In that paper, a minimal/shortest spanning tree (MST/SST) was used to calculate the density of the patterns in the training data set of the k -NN and the learning vector quantization neural network (LVQ-NN) classifiers. In this correspondence, three graph theoretic techniques for data pruning are proposed and applied to the data sets for digit word recognition. It is also proposed that the representative vectors should suitably be chosen to contain "well-balanced" information of the given training data. The proposed methods are designed to consider both hierarchical and nonhierarchical aspects of the data sets and do not require any arduous statistical approximation. In Sections III–V, these pruning methods are developed. The performances of the proposed methods

Manuscript received July 23, 1997; revised January 5, 1998. The associate editor coordinating the review of this paper and approving it for publication was Prof. Yu-Hen Hu.

The author is with the Signal Processing and Digital Systems Section, Department of Electrical and Electronic Engineering, Imperial College of Science, Technology, and Medicine, University of London, London, U.K. (e-mail: t.hoya@ic.ac.uk).

Publisher Item Identifier S 1053-587X(98)05967-4.