# Optimal Approximation of Multidimensional Discrete Signals with Applications

*W.-S. Lu*

Dept. of Electrical and Computer Engineering
University of Victoria
Victoria, B.C., Canada, V8W 3P6
email: wslu@ece.uvic.ca

## Abstract

This paper presents an analytical solution to a general approximation problem for multidimensional (m-D) discrete signals. The solution is optimal in the Frobenius-norm sense and includes the well known singular value decomposition (SVD) based approximation as a special case with $m = 2$. Computational issues involved in the algorithm implementation as well as potential applications of the algorithm to image and video coding are discussed.

## 1. Introduction

This paper is concerned with the problem of approximate representations of multidimensional (m-D) discrete signals with $m \geq 2$ by a linear combination of a selected set of basis functions, which is often referred to as a dictionary, in a certain optimal sense. In particular, we are interested in the special cases of $m = 2$ and $m = 3$ for the purpose of transform coding of still images and video signals. Signal representation for m-D discrete signals have been investigated in [1]–[4] in a matching pursuit or basis pursuit framework and in [5][6] using an iterative least squares technique. This paper describes an analytical solution to a general approximation problem for m-D discrete signals. The problem is formulated in an inner-production space of dimension $\gamma$ with $\gamma = \prod_{i=1}^{m} K_i$ where $K_i$ denotes the number of entries in the $i$th dimension. This inner-product space setting provides a straightforward relation to the m-D signal space, facilitates the definitions of inner-product related concepts in the m-D signal space, allows the use of basis signals which may not be of rank-one, and more importantly, leads to a least-squares solution to the approximation problem. It is shown that the solution includes the well know singular value decomposition (SVD) based approximation as a special case for $m = 2$ when the basis signals are generated from the SVD of the signal, which are of rank-one and orthonormal. Some computational issues involved in the algorithm's implementation as well as potential applications of the algorithm to image and video coding are discussed.

## 2. Notation and Preliminaries

Let $s(k_1, \ldots, k_m)$ with $1 \leq k_i \leq K_i$ be the given $m$-D signal of finite support, and $\mathcal{B} = \{b_i(k_1, \ldots, k_m),\ 1 \leq i \leq N\}$ be the set of basis signals. For $\mathcal{B}$ to be a set of basis functions, it must satisfy (a) the rank of each $b_i \in \mathcal{B}$ is greater than or equal to one; and (b) $\{b_i,\ 1 \leq i \leq N\}$ are linearly independent. Two terms need to be defined here: a set of m-D signals $b_1, \ldots, b_K$ are said to be linearly independent if

$$\sum_{i=1}^{K} k_i b_i = 0$$

implies $k_i = 0$ for $1 \leq i \leq K$; and $b(k_1, \ldots, k_m)$ is said to be a rank-one signal if

$$b = u_1 \bullet u_2 \bullet \cdots \bullet u_m \tag{1a}$$

where $u_i \in R^{K_i \times 1}$, and $\bullet$ denotes the outer product operation meaning that the $(k_1, \ldots, k_m)th$ entry of $b$ is simply the product of $u_1(k_1)$, $u_2(k_2)$, $\ldots$, $u_m(k_m)$, i.e.,

$$b(k_1, \ldots, k_m) = \prod_{j=1}^{m} u_j(k_j) \tag{1b}$$

The rank of a m-D signal is greater than one if it is a nonzero signal and it cannot be expressed as in (1a). A better way to define a m-D signal with rank greater than one will be described shortly. Note that the basis signal set $\mathcal{B}$ may contain signals whose rank is strictly greater than one.

It is quite natural to define the inner product of two m-D signals as

$$\ll a,\ b \gg = \sum_{k_1=1}^{K_1} \cdots \sum_{k_m=1}^{K_m} a(k_1, \ldots, k_m) b(k_1, \ldots, k_m) \tag{2}$$

The m-D signal space equipped with the inner product (2) admits important geometrical concepts. For example, $a$ is said to be orthogonal to $b$ if $\ll a,\ b \gg = 0$; and the norm of signal $b$ is defined by

$$||b||_F = \left[ \sum_{k_1=1}^{K_1} \cdots \sum_{k_m=1}^{K_m} b^2(k_1,\ \ldots,\ k_m) \right]^{1/2} \tag{3}$$

where $||\bullet||_F$ indicates the norm definition here is a generalization of the Frobenius norm for matrices. As will be seen in the next section, the inner-product space induced from (2) can be connected to a normal vector space of higher dimension.

# 3. Problem Formulation and an Analytical Solution

## 3.1. Problem Formulation

For a given positive integer $m$, there is an one-to-one correspondence between the m-D discrete signal space $\mathcal{S} = \{s(k_1,\ \ldots,\ k_m),\ 1 \le k_i \le K_i\}$ and the vector space $\mathcal{V}$ of dimension $\gamma$ with $\gamma = \prod_{i=1}^{m} K_i$. The mapping that facilitates the correspondence is

$$s \in \mathcal{S} \longleftrightarrow s_v \in R^{\gamma \times 1} \tag{4a}$$

where $s(k_1,\ \ldots,\ k_m) = s_v(i)$ with

$$i = (k_m - 1) \prod_{i=1}^{m-1} K_i + (k_{m-1} - 1) \cdot \prod_{i=1}^{m-2} K_i + \cdots + (k_2 - 1)K_1 + k_1 \tag{4b}$$

With $m = 2$, for example, (4b) becomes $i = (k_2 - 1)K_1 + k_1$ so the vector $s_v$ corresponding to a $K_1$ by $K_2$ matrix $s(k_1,\ k_2)$ is generated by stacking the column vectors of the matrix in an orderly fashion. In general, it is straightforward to verify that for $a,\ b \in \mathcal{S}$.

$$\ll a,\ b \gg = \langle a_v,\ b_v \rangle$$

where $\langle \bullet, \bullet \rangle$ denotes the usual inner product in Euclidean space $R^{\gamma \times 1}$, and $||b||_F = ||b_v||_2$ where $||\bullet||_2$ denotes the usual Euclidean norm. It is also easy to verify that m-D signals are linearly independent in $\mathcal{S}$ if and only if they are linearly independent in $\mathcal{V}$; and an m-D signal has rank $r$ if one needs at least $r$ linearly independent signals each of which corresponds to a rank one m-D signal in $\mathcal{S}$ to linearly represent it. In what follows we state the approximation problem in space $\mathcal{S}$, then reformulate it in space $\mathcal{V}$.

Given a set of basis functions $\mathcal{B}$, our goal is to find weights $w_i$, $1 \le i \le N$, such that

$$J = ||s(k_1,\ \ldots,\ k_m) - a(k_1,\ \ldots,\ k_m)||_F^2 \tag{5a}$$

is minimized where the approximation $a$ is given by

$$a = \sum_{i=1}^{N} w_i b_i, \quad b_i \in \mathcal{B} \tag{5b}$$

In space $\mathcal{V}$, we denote the corresponding signals of $s$, $a$ and $b_i$ by $s_v$, $a_v$ and $b_i^{(v)}$, respectively, and rewrite (5) as

$$J = \left|\left| s_v - \sum_{i=1}^{N} w_i b_i^{(v)} \right|\right|_2^2 \tag{6}$$

The weights $w_i^*$ $(1 \le i \le N)$ that minimize $J$ in (6) are used to construct

$$a^* = \sum_{i=1}^{N} w_i^* b_i \tag{7}$$

which is for a given $\mathcal{B}$ an optimal approximation of m-D signal $s$ in the Frobenius norm sense. In the rest of paper $a^*$ in (7) is referred to as the Frobenius (F)-optimal approximation of $s$.

## 3.2. An Analytical Solution

The approximation error $J$ in (6) can be written as

$$\begin{aligned} J &= \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \langle b_i^{(v)},\ b_j^{(v)} \rangle \\ &\quad -2 \sum_{i=1}^{N} w_i \langle s_v,\ b_i^{(v)} \rangle + ||s_v||_2^2 \\ &= w^T Q w - 2w^T p + ||s_v||_2^2 \end{aligned} \tag{8}$$

where $Q = (q_{ij})$ with $q_{ij} = \langle b_i^{(v)},\ b_j^{(v)} \rangle$, $p = (p_i)$ with $p_i = \langle s_v,\ b_i^{(v)} \rangle$, and $w = (w_i)$. Since $\{b_i^{(v)}\}$ are linearly independent, $Q$ is symmetric and positive definite. Hence $J$ is globally convex with respect to $w$, and its minimum is achieved at

$$w^* = Q^{-1} p \tag{9}$$

# 4. Special Cases and Computation Complexity

## 4.1. Special Cases

(a) An important case is when $m = 2$. The signals in question are then discretized images or alike. A well known method to approximate a given matrix $s$ of rank $r$ is to use the SVD of $s$:

$$s = \sum_{i=1}^{r} \sigma_i u_i v_i^T \tag{10}$$

where $\sigma_1 \geq \cdots \geq \sigma_r > 0$, $\{u_i,\ i = 1,\ \ldots,\ r\}$ and $\{v_i,\ i = 1,\ \ldots,\ r\}$are two orthonormal vector sets, i.e., $\langle u_i,\ u_j\rangle = \langle v_i,\ v_j\rangle = \delta_{ij}$. It is known [7] that for each $K \leq r$, the optimal rank-$K$ approximation of $s$ in the Euclidean and Frobenius norm sense is given by

$$a_K = \sum_{i=1}^{K} \sigma_i u_i v_i^T \tag{11}$$

If we denote $b_i = u_i v_i^T$ and let $\mathcal{B} = \{b_i,\ i = 1,\ \ldots,\ K\}$, then we can write the optimal approximation as

$$a_K = \sum_{i=1}^{K} \sigma_i b_i$$

Note that the basis signals in $\mathcal{B}$ are all rank one signals and are orthonormal. Thus $Q = I$ and (9) becomes

$$w^* = p$$

i.e.,

$$\begin{aligned} w_i^* &= p_i = \langle s_v,\ b_i^{(v)}\rangle = \ll s,\ b_i \gg \\ &= \sum_l \sum_k s_{lk} u_{il} v_{ik} = u_i^T s v_i \end{aligned}$$

On the other hand, it follows from (10) that

$$\sigma_i = u_i^T s v_i$$

Hence $w_i^* = \sigma_i$ meaning that (9) covers the SVD-based approximation as a special case when $m = 2$ and $\mathcal{B} = \{u_i v_i^T,\ i = 1,\ \ldots,\ K\}$.

(*b*) Consider now the case when all the basis signals in $\mathcal{B} = \{b_i,\ 1 \leq i \leq N\}$ have rank one and are linearly independent, but not necessarily from the SVD of the signal to be approximated. The dimension $m$ may exceed 2. In this case we write each rank-one basis function in $\mathcal{B}$ as

$$b_i(k_1, \ldots, k_m) = \prod_{j=1}^{m} u_j^{(i)}(k_j)$$

and compute

$$\begin{aligned} J &= \|s - a\|_F^2 \\ &= \sum_{i=1}^{N} \sum_{l=1}^{N} q_{il} w_i w_l - 2 \sum_{i=1}^{N} p_i w_i + \|s\|_F^2 \\ &= w^T Q w - 2 w^T p + \|s\|_F^2 \end{aligned}$$

where

$$\begin{aligned} w &= [w_1 \ \cdots \ w_N]^T, \\ Q &= \{q_{il}\}, \quad p = [p_1 \ \cdots \ p_N]^T, \\ q_{il} &= \prod_{j=1}^{m} \langle u_j^{(i)},\ u_j^{(l)}\rangle \qquad (12) \\ p_i &= \sum_{k_1=1}^{N} \cdots \sum_{k_m=1}^{N} \left[ s(k_1, \ldots, k_m) \prod_{j=1}^{m} u_j^{(i)}(k_j) \right] \end{aligned}$$

(12) indicates that matrix $Q$ is the Schur (i.e., entrywise) product of $Q_1, Q_2, \ldots, Q_m$ where each $Q_i$ is given by

$$Q_i = \begin{bmatrix} \langle u_i^{(1)}, u_i^{(1)}\rangle & \langle u_i^{(1)}, u_i^{(2)}\rangle & \cdots & \langle u_i^{(1)}, u_i^{(N)}\rangle \\ \langle u_i^{(2)}, u_i^{(1)}\rangle & \langle u_i^{(2)}, u_i^{(2)}\rangle & \cdots & \langle u_i^{(2)}, u_i^{(N)}\rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle u_i^{(N)}, u_i^{(1)}\rangle & \langle u_i^{(N)}, u_i^{(2)}\rangle & \cdots & \langle u_i^{(N)}, u_i^{(N)}\rangle \end{bmatrix} \tag{13}$$

which is symmetric and positive definite if and only if the set of basis vectors $\mathcal{U}_i \equiv \{u_i^{(1)},\ \cdots\ u_i^{(N)}\}$ is linearly independent. Using the Schur product theorem [8, Ch. 5], it can be shown that $Q$ is also positive definite if and only if every $\mathcal{U}_i$ is linearly independent. Under such circumstances $J$ is strictly convex with respect to $w$ and the optimal weight $w^*$ that achieves the global minimum of $J$ is given by

$$w^* = Q^{-1} p \tag{14}$$

and the F-optimal approximation is given by

$$a = \sum_{i=1}^{N} w_i^* b_i \tag{15}$$

An interesting case, again with $m = 2$, that involves no SVD of the signal blocks is the two-dimensional (2-D) discrete cosine transform (DCT), which can be expressed in the form [9]

$$\theta(k, l) = \frac{2}{N} \alpha(k) \alpha(l) \cdot \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} x(m, n) \cos \frac{\pi k(2m+1)}{2N} \cos \frac{\pi l(2n+1)}{2N}$$

where $k, l = 0,\ 1, \ldots,\ N-1$, and $\alpha(0) = 1/\sqrt{2}$, $\alpha(j) = 1$ for $j \neq 0$. For a fixed pair $k$ and $l$, the transformation kernel can be expressed as an $N$ by $N$ matrix with the $(i, j)$th entry being $\cos \frac{\pi k(2i-1)}{2N} \cos \frac{\pi l(2j-1)}{2N}$. With $k$ and $l$ varying from 0 to $N-1$, these matrices form a set of basis functions $\mathcal{B} = \{b_l,\ l = 1,\ \ldots,\ N^2\}$. Assuming each basis function $b_l$ has been normalized such that $\|b_l^{(v)}\|_2 = 1$ where $b_l^{(v)}$ denotes the vector in space $\mathcal{V}$ that corresponds to $b_l$, then we can show that the matrix $B = [b_l^{(v)} \cdots b_{N^2}^{(v)}]$ is orthogonal, hence this 2-D DCT based basis function set leads (14) to $w^* = p$.

(c) If $\mathcal{B}$ is overcomplete, then $Q$ will no longer be positive definite and the approximation problem at hand fits into a matching pursuit framework if one decides to search a locally optimal solution or into a basis pursuit scenario when a globally optimal solution is sought. By working in inner-product space $\mathcal{V}$ in which weights $w_i^*$ $(1 \le i \le N)$ that minimize $J$ in (6) are sought, a matching pursuit algorithm for best approximating a m-D signal can be established [3].

### 4.2. Computation Complexity

There are several good features of the solution expressed in either (9) or (14), and we would like to describe two of them here. First, the matrix $Q$ is *independent* of the input signal $s$, therefore $Q^{-1}$ can be evaluated off line as long as the basis functions have been chosen. Second, if for some reasons one decides to add new basis functions to $\mathcal{B}$, the inverse matrix $Q^{-1}$ can be fully used to quickly evaluate the inverse of the augmented $Q$ matrix, denoted by $\tilde{Q}$, by using the well-known formula for inverting block matrices. Specifically let $\tilde{Q}$ be the $Q$ matrix after new basis functions are included in $\mathcal{B}$. From (8) or (13) it follows that $\tilde{Q}$ can be expressed as

$$\tilde{Q} = \left[\begin{array}{c|c} Q & A_{12} \\ \hline A_{12}^T & A_{22} \end{array}\right]$$

where $A_{12}$ and $A_{22}$ are the only blocks involving new basis functions.

The inverse of $\tilde{Q}$ is given by

$$\tilde{Q}^{-1} = \begin{bmatrix} Q^{-1} + Q^{-1}A_{12}\hat{A}_{22}^{-1}A_{12}^TQ^{-1} & -Q^{-1}A_{12}\hat{A}_{22}^{-1} \\ -\hat{A}_{22}^{-1}A_{12}^TQ^{-1} & \hat{A}_{22}^{-1} \end{bmatrix}$$

where

$$\hat{A}_{22} = A_{22} - A_{12}^TQ^{-1}A_{12}$$

Note that if $Q^{-1}$ is already available, then evaluating $\tilde{Q}^{-1}$ using the above formula is computationally inexpensive as long as the number of new basis functions is not large.

## 5. An Illustrative Example

As an example we consider approximating still images using the 2-D DCT based basis signals with $N = 8$. The set $\mathcal{B}$ in this case contains 64 $8 \times 8$ signals and, as was mentioned in the previous section, the associated B is a $64 \times 64$ orthogonal matrix.

The image in question is partitioned into $8 \times 8$ blocks, and each block is approximated by a linear combination of several basis signals. In general, two simple steps are needed to carry out the approximation: first we compute $w^*$ using all 64 basis signals, and keep only $K$ basis functions that correspond to the largest $w_i^*$'s in magnitude where $K$ is the number of basis functions allowed to use in the approximation; second, these basis functions are used to form a new $\mathcal{B}$ to compute the new $w^*$. With the 2-D DCT based $\mathcal{B}$, however, the fact that B is orthogonal immediately implies that only the first step is needed and the weights $w_i^*$ do not have to be recalculated. Figures 1 and 2 show the approximation of two sample $256 \times 256$ images, namely the building and the fruits, with 256 gray scales using 2, 3, and 4 basis signals. The approximations can be evaluated using the signal-to-noise ratio (SNR) defined by

$$\mathrm{SNR} = 20 \log_{10} \left[ \frac{||s||_F}{||s-a||_F} \right]$$

and the evaluation results are listed in Table 1.

**Table 1: Evaluation Results**

| Number of Basis Functions | SNR for "building" | SNR for "fruits" |
|---|---|---|
| 2 | 10.6944 dB | 15.1105 dB |
| 3 | 12.6670 dB | 17.4176 dB |
| 4 | 14.0938 dB | 19.1333 dB |

It is observed that the 2-D DCT basis functions yield better approximations for image "fruits" than for image "building" in which edges are the dominant component of the image. One way to improve the approximation accuracy without increasing the number of basis functions used is to augment $\mathcal{B}$ by including a set of edge signals [10]. Of course $\mathcal{B}$ in this case becomes overcomplete and the algorithm proposed here needs modifications. An easy approach to handling the situation is that one can use the algorithm here as applied to the set of DCT basis functions to select, say $n_1$, best basis functions, next the same algorithm is applied to the set of edge type basis functions to select, say $n_2$, best basis functions. Together these $(n_1 + n_2)$ basis functions are used to form an improved $\mathcal{B}$ to which the algorithm is once again applied. Technical details and simulation results using such modified method will be reported elsewhere.

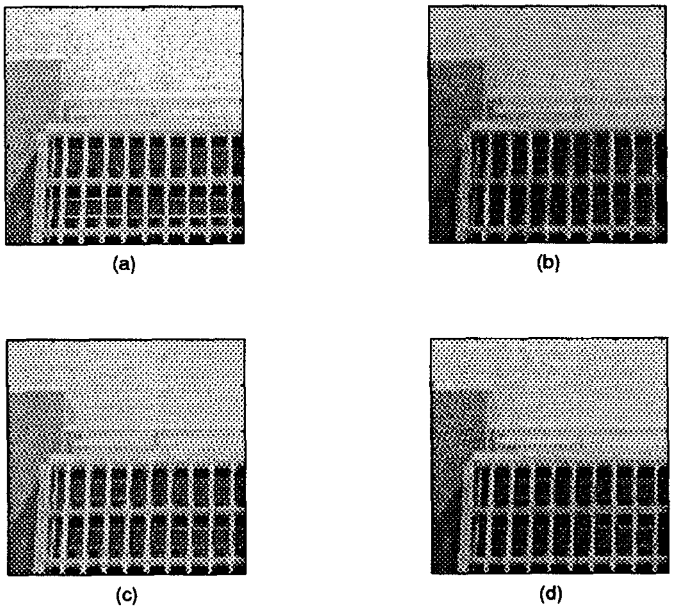


**Figure 1: (a) original "building" image; approximations using (b) 2 basis signals; (c) 3 basis signals; and (d) 4 basis signals.**

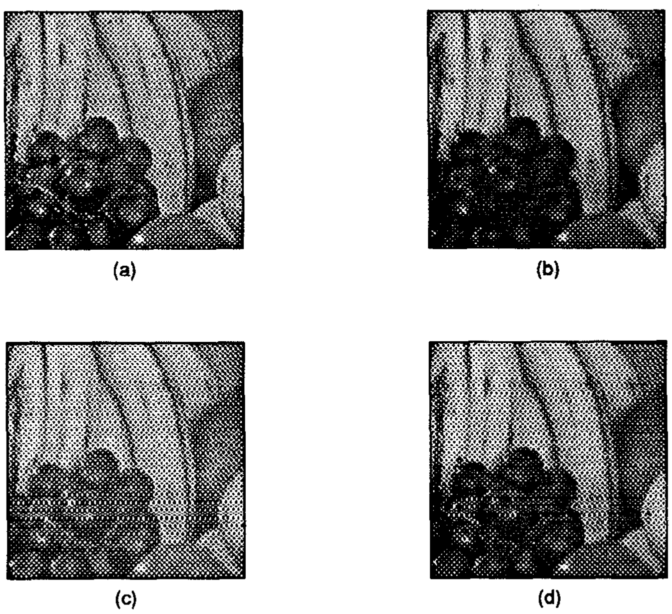


**Figure 2: (a) original "fruits" image; approximations using (b) 2 basis signals; (c) 3 basis signals; and (d) 4 basis signals.**

### Acknowledgment

The author is grateful to Micronet, Networks of Centres of Excellence Program and to the Natural Sciences and Engineering Research Council of Canada, for supporting this work.

## References

[1] S. Chen and D. Donoho, "Basis pursuit," in *Proc. 28th Asilomar Conference*, pp. 41-44, Nov. 1994.

[2] Y. C. Pati, R. Rezallfar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition." in *Proc. Asilomar Conf.*, pp. 40-44, Nov. 1993.

[3] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing,* vol. 41, pp. 3397-3414, Dec. 1993.

[4] M. Vetterli and T. Kalker, "Matching pursuit for compression and application to motion compensated video coding," in *Proc. ICIP*, 1995.

[5] A. P. Berg and W. B. Mikhael, "An efficient structure and algorithm for the mixed transform representation of signals," in *Proc. 29th Asilomar Conf.*, Monterey, CA., Nov. 1995.

[6] A. P. Berg and W. B. Mikhael, "An iterative formulation for the least squares approximation of 2-D signals," in *Proc. ISCAS'96*, vol. 2, pp. 360-363, Atlanta, May 1996.

[7] G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, 1973.

[8] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, New York: Cambridge University Press, 1994.

[9] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall, 1984.

[10] A. P. Berg and W. B. Mikhael, "Fidelity enhancement of transform based image coding using nonorthogonal basis images," in *Proc. ISCAS'96*, vol. 2, pp. 437-440, May 1996.