Applications of Multi-transforms to 3-D Spatial Signal Representation

Arun Ramaswamy and Wasfy B. Mikhael
ECE Department, Univ. of Central Florida
Orlando, Fl 32816

Abstract

Three dimensional spatial signal representation employing transforms is generally performed by applying a 2-D transform to the 2-D slices of the 3-D data or a 3-D transform for the entire data. Transforms such as DCT, Walsh and Haar are frequently considered because of their ease of computation. This paper develops a technique called the multi-transform technique for the efficient representation of 3-D spatial data. The main feature of the technique is that two or more transforms can be applied to represent the signal thus benefiting from the compression qualities of each of the transforms. The signal to noise ratio (SNR) of the reconstructed signal employing the proposed technique is higher than the SNR of the signal employing a single transform alone. Results are presented for a 3-D MRI signal representation using the 2-D and 3-D multi-transform techniques employing the DCT / Walsh combination. Comparisons with using DCT alone are made for both cases.

1 Introduction

Constrained representation of three dimensional(3-D) spatial signals has been achieved by processing the signal in the transform domain, [1]. One possible approach is to apply 2-D transforms to the individual 2-D slices of the 3-D data. Alternatively, an invertible 3-D transform maps the signal to a set of coefficients, to take advantage of the correlation in three directions. The transforms employed are usually separable and symmetric and are required to convert the statistically correlated input samples to a set of uncorrelated coefficients. The dominant transform coefficients are retained and the remaining ones are discarded. The performance of the various transforms have been extensively compared [2]. Each of these transforms has its own features, which may make it the more efficient for a specific class of signals.

In both the 2D and 3D approaches, it is however possible that certain spectral features of a signal can be compactly represented by a particular transform, while some other features could be represented more efficiently by another transform. A technique to represent 3-D spatial signals employing two or more transforms is proposed in this paper. Both 2-D and 3-D approaches are considered. First, through adaptive techniques, the signal is appropriately resolved into subsignals, such that each subsignal is efficiently represented by the dominant components in a particular transform domain. Then the entire signal is reconstructed by superimposing the subsets of basis functions from different domains. Also, these subsets are, in general, non-orthogonal with respect to each other. This constrained representation employing basis functions of two or more suitable transforms yields a higher Signal to Noise Ratio(SNR) of the reconstructed signal as compared to reconstruction using a single transform. Here, SNR is defined as

$$SNR = \frac{Energy\ in\ the\ original\ signal}{Energy\ in\ the\ unrepresented\ part\ of\ \ signal} \tag{1}$$

Applications of the proposed technique are presented for medical MRI's. Objective comparisons between the proposed approach using multi-transforms and using a single transform, such as DCT, are made for both the 2-D and 3-D implementations.

The remainder of the summary is arranged as follows. The residual error formulation, error minimization and optimization strategy are considered in section II for the 2-D and 3-D approaches. Section III presents the applications and results. Section IV concludes the paper.

2 Formulation

2.1 Implementation

The block diagram for the multi-transform representation of multi-dimensional signals is shown in Fig. 1. The purpose of the system is to resolve the input

1058-6393/95 $4.00

signal into subsignals such that each of these subsignals is compactly represented in a particular transform domain. Consider a p-dimensional signal $[X]$ to be the input to the system. The technique can be briefly summarized as follows. The input is transformed by the first orthogonal transform $tr1$, to produce the transform coefficient matrix $[A]$, which contains the independent coefficients. Next, the dominant coefficients from the $tr1$ domain, *i.e*, the entries of $[A]$, are extracted for representation. To do that, the entries of $[A]$ are appropriately weighted by an matrix $[\alpha]$, to yield the matrix $[C]$. The difference between $[A]$ and $[C]$ is the residual error or the unrepresented part of the signal in the $tr1$ domain. This residual error is a function of the $[\alpha]$ weights and is referred to as $[R(1)]$. It is assumed that the signal corresponding to the residual, $[R(1)]$ can be compactly represented in a second transform domain, $tr2$. To perform this operation efficiently, a conversion operation, $[T^{1,2}]$ is defined which converts the signal corresponding to the residual of the $tr1$ domain directly to the $tr2$ domain. This operation yields the matrix $[P]$. In general, $[T^{i,j}]$ refers to the operator transforming the signal to the $tr2$ domain by operating directly on the $tr1$ domain coefficients. The usefulness of this operator lies in the fact that it is not necessary to perform an inverse transformation to the input signal domain. The dominant elements of $[P]$ are similarly selected by means of another weight matrix $[\beta]$. In a manner similar to $[R(1)]$, the residual error of the second transform domain $[R(2)]$, is computed and transformed to the third transform domain, $tr3$. This is continued for, say, M transforms. These M transforms are, in general, mutually non-orthogonal. The residual of the last transform domain $[R(M)]$, is the final residual error. This error is a nonlinear function of the $[\alpha]$, $[\beta]$.. weights and represents the part of the input signal which is not represented by any of the applied transforms. Next, the energy in the residual error is minimized using the developed adaptive algorithms. As the energy is minimized, it is shown that the weights approach their optimal values, thus resolving the signal into subsignals in the various transform domains. The signal is finally reconstructed by superimposing the weighted coefficients from the various transform domains.

In this contribution, the technique is applied to represent 2-D(slices of the 3-D data) and the 3-D signal as a whole, *i.e* $p = 2, 3$. Specifically, the formulation presented below considers the two transform case, *i.e*, M is set to 2. The DCT is selected as the first transform, $tr1$, and the $tr2$ transform is the WH transform.

2.2 2-D Formulation

The input signal, $[X]$, the transformation and the weight matrices are considered to be of size $(N \times N)$. The DCT of the input signal is given by

$$[A] = [D][X][D]^T \tag{2}$$

where $[A]$ is the $(N \times N)$ 2-D DCT of the input and $[D]$ is the $(N \times N)$ 1-D DCT transformation matrix [1].

The elements of $[A] = \{a_{ij}\}$ are independent. To select the dominant coefficients for representation, $[A]$ is weighted by the matrix, $[\alpha] = \{\alpha_{ij}\}$. This yields the $(N \times N)$ matrix $[C] = \{c_{ij}\}$. Let the weight α_{ij} correspond to a general (ij) DCT transform coefficient a_{ij}. Therefore, the (ij) element of the weighed transform coefficient matrix can be written as

$$c_{ij} = \alpha_{ij} a_{ij} \tag{3}$$

Due to the nature of the basis functions of the DCT transform, $[C]$ is representative of the narrow band portion of the input signal. The residual in the DCT domain $[R(1)] = \{r(1)_{ij}\}$ is the difference between $[A]$ and $[C]$. This residual is representative of the broadband part of the input spectrum.

$$r(1)_{ij} = (1 - \alpha_{ij}) a_{ij} \tag{4}$$

The residual is converted to the $tr2$ domain by means of the $[T^{1,2}]$ operator. This operator is represented by the matrix $[K]$ and is derived in [3]. This conversion yields.

$$[P] = [K][R(1)][K]^T \tag{5}$$

The matrix $[P] = \{p_{ij}\}$ represents the residual error in the Walsh domain. To select the dominant coefficients of $[P]$, the elements of $[P]$ are similarly weighed by another weight matrix $[\beta] = \{\beta_{ij}\}$ to yield the matrix $[G] = \{g_{ij}\}$. Let β_{ij} represent the weight corresponding to the $tr2$ transform coefficient p_{ij}. Thus,

$$g_{ij} = \beta_{ij} p_{ij} \tag{6}$$

The residual in the Walsh domain, referred to as $[R(2)]$, is the difference between $[P]$ and $[G]$, given by (5) and (6) . Therefore, this residual error can be written as

$$r(2)_{ij} = (1 - \beta_{ij}) p_{ij} \tag{7}$$

Substituting for p_{ij} from (5) and $[R(1)]$ from (4), the final residual error in the Walsh domain, $r(2)_{ij}$ can be derived as

$$r(2)_{ij} = (1 - \beta_{ij}) \sum_{u=0}^{N-1} k_{iu} \sum_{v=0}^{N-1} [(1 - \alpha_{uv}) a_{uv} k_{jv}] \tag{8}$$

The energy, $\Phi(\alpha,\beta)$, in the residual error is computed in the $tr2$ domain as

$$\Phi(\alpha,\beta) = [R(2)]^T.[R(2)] \quad (9)$$

where the superscript T denotes the transpose. It is to be minimized to resolve the input signal into the desired subsignals. The steepest descent approach is one possible method for error minimization and is developed in this paper to minimize the error energy with respect to the unknown weights, [4]. The update equations for the unknown weights are as follows

$$\alpha_{ij}(n+1) = \alpha_{ij}(n) - \mu_\alpha \nabla_{\alpha_{ij}}, \quad i,j = 0,1,..N-1; \quad (10)$$

$$\beta_{ij}(n+1) = \beta_{ij}(n) - \mu_\beta \nabla_{\beta_{ij}}, \quad i,j = 0,1,..N-1; \quad (11)$$

where n is the iteration index, μ is a time varying convergence factor and $\nabla_{\alpha_{ij}}$ and $\nabla_{\beta_{ij}}$ are the instantaneous gradients, *i.e*, the partial derivatives of Φ with respect to α_{ij} and β_{ij} respectively.

Rather than using a constant convergence factor, a time varying μ is employed, which is determined before updating each of the non-zero DCT and Walsh weights. Using the Taylor series expansion for the error energy and the update equations, the optimal convergence factor can be derived as

$$\mu_\alpha = \frac{\Phi}{\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}[\nabla_{\alpha_{uv}}]^2} \quad (12)$$

$$\mu_\beta = \frac{\Phi}{\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}[\nabla_{\beta_{uv}}]^2} \quad (13)$$

The error surface is multimodal because the residual error given by (8) is nonlinear in the $[\alpha]$ and $[\beta]$ weights. Also, the transforms employed are mutually non-orthogonal. Consequently, the highest SNR may not result by simply selecting the largest coefficients from each domain at the start of the error minimization. A coefficient that is initially small in a given transform domain can acquire a large final value after minimizing the residual error. Thus, a strategy is also presented in the paper, where care is taken before discarding any coefficient. Finally, the optimal weights and the coefficients are computed in all the transform domains. These weighed coefficients are inverse transformed to the spatial domain from their respective transform domains. The resulting subsignals are then added to yield the reconstructed signal. For example, in the DCT-Walsh case, the reconstructed signal $[X']$ can computed from

$$[X'] = [D]^T[C][D] + [W]^T[G][W] \quad (14)$$

where $[G]$ and $[C]$ are computed using the optimal weights and $[W]$ is the Walsh matrix.

2.3 3-D Formulation

The fomulation presented below, briefly outlines the basic equations and steps of the implementation for the 3-D case. The input signal $[X]$ and the other matrices are assumed to contain $(N \times N \times N)$ elements. The DCT, $[A] = \{a_{ijk}\}$, of the input signal can be written in terms of the separable 1-D transformation matrix, $[D]$.

$$a_{ijk} = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} x_{uvw} d_{kw} d_{jv} d_{iu} \quad (15)$$

The elements are $[A]$ are weighted by a weight matrix, $[\alpha]$ to select the dominant coefficients, such that the weighted coefficient matrix element c_{ijk} is given by

$$c_{ijk} = \alpha_{ijk} a_{ijk} \quad (16)$$

The residual of $tr1$ equals

$$r(1)_{ijk} = (1-\alpha_{ijk}) a_{ijk} \quad (17)$$

Employing the 1-D $[T^{1,2}]$ matrix operator, $[K]$, this residual is converted to the $tr2$ domain, yielding $[P]$.

$$p_{ijk} = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} r(1)_{uvw} k_{kw} k_{jv} k_{iu} \quad (18)$$

The is similarly weighted by another weight matrix, $[\beta]$ to yield the matrix, $[G]$.

$$g_{ijk} = \beta_{ijk} p_{ijk} \quad (19)$$

The residual of $tr2$ is the final residual error and can be shown to be

$$r(2)_{ijk} = (1-\beta_{ijk}) \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} (1-\alpha_{uvw}) a_{uvw} k_{kw} k_{jv} k_{iu} \quad (20)$$

The residual error energy Φ is given by

$$\Phi = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\sum_{w=0}^{N-1} |r(2)_{uvw}|^2 \quad (21)$$

This energy is minimized with respect to the unknown weights in steepest descent manner similar to the 2-D case. Similar update equations are applied for the unknown $\alpha]$ and $[\beta]$ weights,

$$\alpha_{ijk}(n+1) = \alpha_{ijk}(n) - \mu_\alpha \nabla_{\alpha_{ijk}}, \quad i,j,k = 0,1,..N-1; \quad (22)$$

$$\beta_{ijk}(n+1) = \beta_{ijk}(n) - \mu_\beta \nabla_{\beta_{ijk}}, \quad i,j,k = 0,1,..N-1; \quad (23)$$

where n denotes the iteration index, μ is a time varying convergence factor tailored to each individual weight and ∇ is the gradient.

The gradients can be shown to be

$$\nabla_{\alpha_{ijk}} = -2a_{ijk} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} r(2)_{uvw}[(1-\beta_{uvw})k_{wk}k_{vj}k_{ui}] \quad (24)$$

$$\nabla_{\beta_{ijk}} = -2r(2)_{ijk} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} [(1-\alpha_{uvw})a_{uvw}k_{iu}k_{jv}k_{kw}] \quad (25)$$

The convergence factors can be derived as

$$\mu_\alpha = \frac{\Phi}{\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} [\nabla_{\alpha_{uvw}}]^2} \quad (26)$$

$$\mu_\beta = \frac{\Phi}{\sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} [\nabla_{\beta_{uvw}}]^2} \quad (27)$$

An identical implementation strategy and reconstruction scheme as the 2-D signals is applied for the 3-D case too.

3 Simulation Results

3.1 Simulation I

In this simulation the 2-D multi-transform method was successfully applied individually to eight slices of a MRI, each slice is (256×256). The first two sample slices are shown in Fig. 2. Each slice was divided into subimages of size (8×8). Then each individual subimage is sequentially represented employing multi-transforms. The transforms considered are DCT and Walsh. Coefficient reduction factor is the factor by which the total number of coefficients for the entire image is reduced. For a coefficient reduction factor of say, n, the number of retained coefficients for the entire image is N_r

$$N_r = INT(\frac{512 \times 512}{n}) \quad (28)$$

where INT denotes integer. For a given N_r, the coefficients are allocated for each subimage depending on the detail present in the subimage. For example, a subimage with sharp changes in the grey level is allocated more coefficients than a image with a constant grey level. The number of coefficients allocated to the $i-th$ subimage N_i, is chosen as

$$N_i = INT(\frac{E_{X_i}}{\sum_k E_{X_k}} \times N_r) \quad (29)$$

where E_{X_i} is the variance in the i-th subimage. Once, the optimum number of coefficients, N_i, for each subimage is decided, the multi-transforms are applied to represent and code the subimages sequentially. The results are presented in Table 1. The first two sample reconstructed slices are shown in Fig. 3 for crf of 10.

3.2 Simulation II

The proposed 3-D multi-transform approach employing the DCT and Walsh transforms was utilized to represent the 3-D medical data. The simulation was similar to the 2-D approach, with the difference that $(8 \times 8 \times 8)$ subcubes were represented instead of subimages. Coefficients were distributed to each subcube depending on the variance. Both the DCT and the multi-transform approach were compared in terms of the SNR_i. The results are tabulated in Table 2 and the reconstructed data is shown in Fig. 4 for a crf of 10.

4 Conclusions

A technique is developed for the efficient representation of 3-D signals by means of the 2-D and 3-D implementations. The signal is resolved appropriately into subsignals such that each of the subsignals is compactly represented in a particular transform domain. The signal is then efficiently represented by superimposing the dominant coefficients corresponding to each subsignal. The proposed approach has been successfully developed by addressing several key issues. The residual error is properly formulated. Adaptive algorithms are developed to appropriately resolve the signal between the domains by minimizing the residual error. Extensive simulations confirmed the improvement offered by the proposed technique over a single transform. Sample results are presented to illustrate the improvement.

Acknowledgements

This work is supported by: The Enterprise Florida, Contract # 16-22-749; Harris Semiconductors Sector, SPD Engineering Operations;
Director, Science & Technology, EPCOT Center, Walt Disney World, Florida; Sandia National Laboratories, Contract # 16-22-867.

References

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, 1992.

[2] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*, New York: Springer Verlag, 1975.

[3] W. B. Mikhael and A. Ramaswamy, " Residual error formulation and adaptive minimization for representing non-stationary signals using mixed transforms", IEEE Trans. on Circuits and Sys., Vol. 39, No. 7, pp. 489-492, 1992.

[4] B.Widrow and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1985.

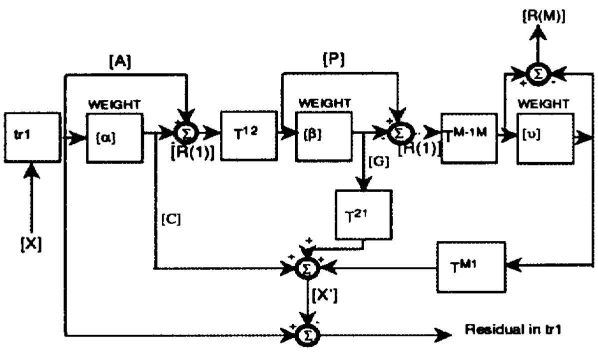
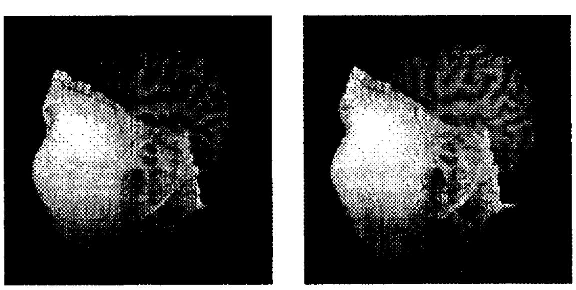
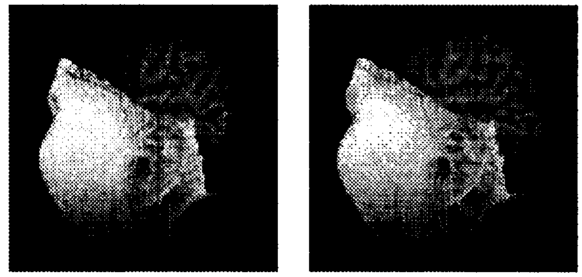


Figure 1: The Multi-transform representation of signals. The adaptive weights help in selecting coefficients from the vaious transform domains. The residual error is minimized adaptively for a compact representation. The scheme is discussed in section II.



(a) (b)

Fig. 2: The first two original frames of the brain MRI .

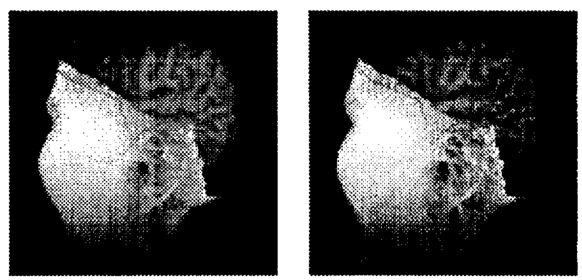


(a) (b)

Fig. 3: The first two reconstructed frames, using the 2-D approach with crf equals 10.

Table. 1: SIGNAL TO NOISE RATIO (dB) FOR CRF OF 10
2-D Approach

Slice	DCT	Multi-transform
1	25.92	27.89
2	26.07	26.90
3	26.32	27.34
4	26.50	28.42
5	26.52	27.54
6	26.71	27.79
7	26.72	27.62
8	26.79	27.98



(a) (b)

Fig. 4: The first two reconstructed frames using the 3-D approach with crf equals 10.

Table. 2: SIGNAL TO NOISE RATIO (dB) FOR CRF OF 10
3-D Approach

Slice	DCT	Multi-transform
1	32.44	33.50
2	32.80	33.01
3	32.90	33.12
4	32.71	33.76
5	32.87	33.94
6	32.66	34.03
7	33.10	33.23
8	32.87	33.91