

The Discrete Modal Transform and its Application to Lossy Image Compression

Michail Krinidis[†], Nikos Nikolaidis[†] and Ioannis Pitas[†]

[†]*Aristotle University of Thessaloniki*

Department of Informatics

Box 451, 54124 Thessaloniki, Greece

Address for correspondence :

Professor Ioannis Pitas

Aristotle University of Thessaloniki

54124 Thessaloniki

GREECE

Tel. ++ 30 231 099 63 04

Fax ++ 30 231 099 63 04

email: pitas@zeus.csd.auth.gr

Abstract

This paper introduces the Discrete Modal Transform, a 1D and 2D discrete, non-separable transform for signal processing, which, in the mathematical sense, is a generalization of the well known Discrete Cosine Transform (DCT). A 3D deformable surface model is used to represent the image intensity and the introduced discrete transform is a by-product of the explicit surface deformation governing equations. The properties of the proposed transform are similar to those of the DCT. To illustrate these properties, the proposed transform is applied to lossy image compression and the obtained results are compared to those of a DCT-based compression scheme.

Experimental results show that DMT, which includes an embedded compression ratio selection mechanism, has excellent energy compaction properties and achieves comparable compression results to DCT at low compression ratios, while being in general better than DCT at high compression ratios.

Key words: Signal transforms, Discrete Cosine Transform, 3D deformable models, intensity surface, lossy compression, image decomposition, signal analysis.

1 Introduction

Intensive research has been carried out in the last decades on signal transforms and this topic continues to be of great interest both with respect to its theoretical aspects, as well as in terms of applications in the field of signal processing. A one dimensional ($1D$) signal can be decomposed into a set of transform coefficients by applying a one dimensional transform that employs a set of basis functions. Similarly, a two dimensional ($2D$) signal, e.g. an image, can be decomposed in terms of basis matrices, usually called basis images. Image transforms have been used in several applications, such as image enhancement [1], restoration [2], description [3], filtering [4], data compression [5], [6], feature extraction [7],[8] etc. Various transforms, such as the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT) [9] and the Wavelet Transform [10], have been proposed and applied in digital signal and image processing literature. Among the applications mentioned above, compression is perhaps the most prominent one.

The overall idea behind transform-based compression is to use a transform that decorrelates the input signal and packs its total energy into a small number of coefficients. There is no globally optimal transform for that purpose [6],

but the Karhunen-Loève Transform (KLT) can be considered to be optimal under certain circumstances, i.e. for a Gaussian source at any bit rate and bit allocation strategy [11]. However, the computation of the KLT is expensive and time consuming, since there is no special structure and it is signal-dependent. Effros et al. [12] demonstrate both the failures and the successes of the KLT. Other transforms, such as DCT [13]-[15], DFT [16], Hadamard Transform [17], Slant Transform [18] are computationally faster than the KLT, while exhibiting slightly worse performance than that of the KLT, in terms of energy compaction and decorrelation. Among them, DCT is the most widely used for image compression. Indeed, DCT is incorporated in various image/video compression standards, such as JPEG [13], MPEG 1/2 [19], H.261 [20], and H.263 [21].

In this paper, we propose a new transform that was motivated by the technique presented in [22]-[24] for analyzing non-rigid object motion, with application to medical images. Nastar and Ayache [22] approximated the dynamic object surface deformations using a physics-based deformable model. Based on the same principle, we assume that a $2D$ signal (image intensity) can be represented by a surface in $3D$, namely the intensity surface. The basic idea is to warp a physics-based deformable surface model onto the intensity surface of a target $2D$ image. We, then, utilize an intermediate step of the deformation procedure, which is proven to be a discrete $2D$ transform that can decompose the image into a class of basis images. In a similar manner, by exploiting a deformable open curve (chain), one can form a $1D$ discrete transform in order to represent a $1D$ signal by a set of $1D$ basis functions.

The proposed transform, named Discrete Modal Transform (DMT), is shown to be a generalization of DCT, although it is non-separable and non-orthonormal

and thus, some of the properties of the well known discrete separable transforms do not apply. The fact that DCT can be derived starting from a deformable model that tries to approximate the intensity surface of an image is a significant outcome of this study. A very important characteristic of DMT is that it can be viewed as DCT scaled by a novel analytically derived quantization table that incorporates a physically meaningful parameter for controlling its energy compaction/compression level. The proposed transform is applied to lossy data compression and the results indicate that it can be a useful tool in image compression. Compared to DCT (combined with the JPEG quantization tables), the proposed transform can achieve comparable image quality for low compression ratios and surpasses it in terms of compressed image quality for high compression ratios. Moreover, it has very good energy packing and decorrelation properties.

The remainder of the paper is organized as follows. The physics-based deformable surface model [22] is reviewed in Section 2 whereas in Section 3, the proposed 1D and 2D discrete transform is introduced. The properties of the proposed transform are described in Section 4. In Section 5 the DMT energy packing and decorrelation efficiency are examined. Section 6 presents a JPEG-like lossy image compression scheme that utilizes the proposed transform and provides comparisons with DCT-based compression. Conclusions are drawn in Section 7.

2 3D Physics-Based Deformable Surface Modeling

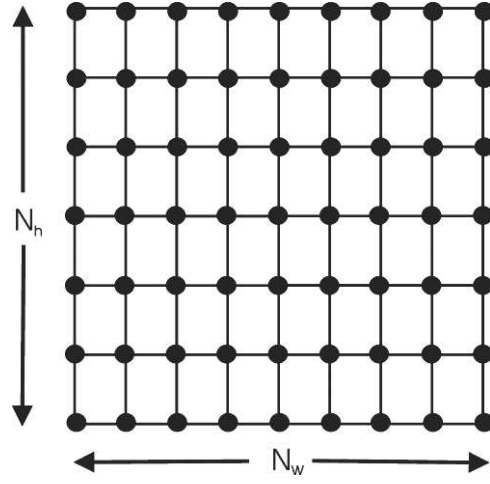
In this section, a physics based deformable surface model, introduced in [22] and [25], will be reviewed in order to make the paper self-contained. There are

two formulations of the model, one for a $3D$ deformable surface model and one for a $2D$ deformable curve model. We choose to analyze the $3D$ model since it is more complicated in general and, at the same time, the assumptions and the deformation procedure are the same for both the $2D$ and the $3D$ model. The $2D$ physics-based deformable curve model is described in brief, at the end of this section, due to its similarity to the $3D$ case.

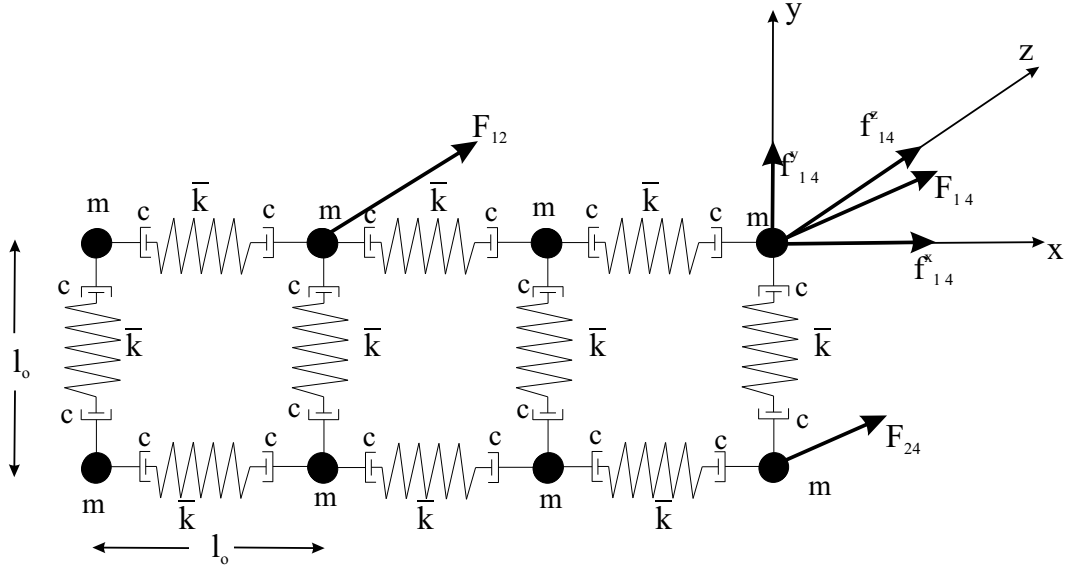
Image intensity $I(x, y)$ can be assumed to define a surface over the image domain that will be subsequently called *intensity surface* $(x, y, I(x, y))$ in the so-called XYI space [26] (Figure 2). A $3D$ physics-based deformable surface model can be used to approximate the intensity surface. The deformable surface model consists of a uniform mesh of $N = N_h \times N_w$ nodes, as illustrated in Figure 1a. In this Section, we assume that N_h and N_w are equal to the image height and width respectively, i.e., that each image pixel corresponds to one mesh node. The node coordinates of the model under examination are stacked in a N -dimensional column vector whose elements are 3-dimensional row vectors:

$$\mathbf{v}^{(t)} = [\mathbf{x}_{11}^{(t)}, \dots, \mathbf{x}_{1N_w}^{(t)}, \mathbf{x}_{21}^{(t)}, \dots, \mathbf{x}_{j'j'}^{(t)}, \dots, \mathbf{x}_{N_h N_w}^{(t)}]^T = [\mathbf{x}_1^{(t)}, \dots, \mathbf{x}_i^{(t)}, \dots, \mathbf{x}_N^{(t)}]^T, \quad (1)$$

where $\mathbf{x}_i^{(t)} = [x_i^{(t)}, y_i^{(t)}, z_i^{(t)}] = [x_i^{(t)}, y_i^{(t)}, I_{Model}^{(t)}(x_i, y_i)]$ and t denotes the t -th deformation time instance. For the problem at hand, each node of the deformable surface model can move only along the z -axis, i.e., the intensity axis, thus $x_i^{(t)} = x_i$ and $y_i^{(t)} = y_i$. Each model node is assumed to have a mass m and is connected to its four neighbors with perfect identical springs of stiffness \bar{k} having natural length l_0 and damping coefficient c (see Figure 1b). Constants \bar{k} and m describe the physical characteristics of the deformable model and determine its behavior. Under the effect of internal and external forces, the



(a)



(b)

Fig. 1. *(a)* Quadrilateral surface (mesh) model, *(b)* Example of a 3D surface model comprised of 8 nodes of mass m connected with identical springs of stiffness \bar{k} and damping coefficient c . Three forces act on three model nodes and result in model deformation.

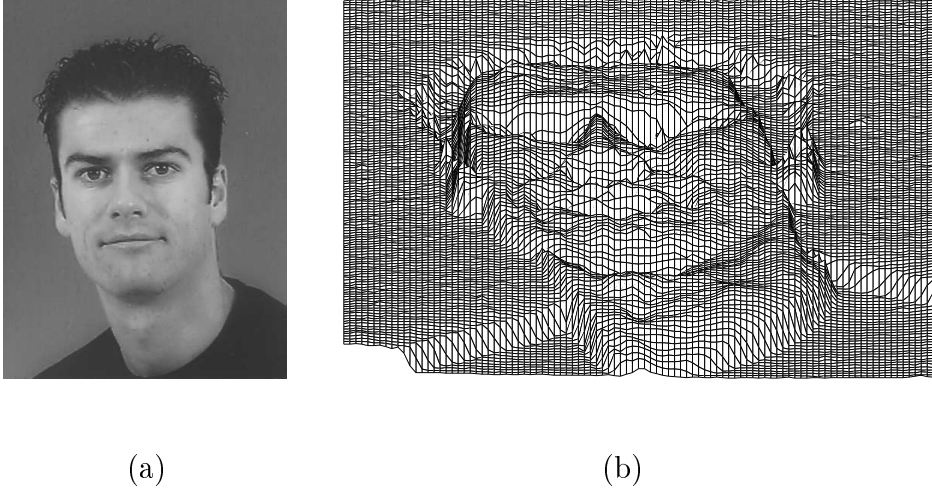


Fig. 2. *(a) Facial image, (b) intensity surface representation of the image.*

mass-spring system deforms to a $3D$ mesh representation of the image intensity surface. Only elastic deformations are considered, i.e., it is assumed that the model recovers its original configuration as soon as all applied forces causing the deformation are removed. When \bar{k} increases and/or m decreases, the ratio $\frac{\bar{k}}{m}$ increases and the surface model tends to behave as a rigid one. This means, in practice, that the intensity surface model can hardly deform. In this case, forces applied to nodes affect entire node neighborhoods (if not the entire grid), depending on the exact value of the ratio $\frac{\bar{k}}{m}$. On the other hand, when m increases and/or \bar{k} decreases, the ratio $\frac{\bar{k}}{m}$ decreases and the intensity surface model tends to be a fully deformable one, meaning that each force essentially affects only the node (mass), where it is applied to.

The model under study is a physics-based system governed by the fundamental

dynamics equation:

$$\mathbf{f}_{el}(\mathbf{x}_i^{(t)}) + \mathbf{f}_d(\mathbf{x}_i^{(t)}) + \mathbf{f}_{ext}(\mathbf{x}_i^{(t)}) = m_i \ddot{\mathbf{x}}_i^{(t)} \quad , \quad i = 1, 2, \dots, N, \quad (2)$$

where $\ddot{\mathbf{x}}_i^{(t)}$ is the acceleration of the i -th node. The external force $\mathbf{f}_{ext}(\cdot)$ applied to each node, results from the attraction of the model by the image intensity (often being proportional to the Euclidean distance between the node coordinates and the corresponding pixel of the image, whose representation in the XYI space is $(x_i, y_i, I(x_i, y_i))$ [27], [28]). The elastic force, $\mathbf{f}_{el}(\cdot)$ on the i -th node is defined as:

$$\mathbf{f}_{el}(\mathbf{x}_i^{(t)}) = -\bar{k} \left(\sum_{j \in \mathcal{T}(i)} \mathbf{b}_{ij}^{(t)} - l_0 \sum_{j \in \mathcal{T}(i)} \frac{\mathbf{b}_{ij}^{(t)}}{\|\mathbf{b}_{ij}^{(t)}\|} \right), \quad (3)$$

where $\mathcal{T}(i)$ denotes the set of the four connected neighbors of node i and $\mathbf{b}_{ij} = \mathbf{x}_i^{(t)} - \mathbf{x}_j^{(t)}$, $j \in \mathcal{T}(i)$ is the vector difference of two nodes. The damping force $\mathbf{f}_d(\cdot)$ is proportional to node velocity $\dot{\mathbf{x}}_i^{(t)}$:

$$\mathbf{f}_d(\mathbf{x}_i^{(t)}) = -c \dot{\mathbf{x}}_i^{(t)}, \quad (4)$$

where c is the damping constant. The above governing equation (2) applies to all N model nodes, leading to a nonlinear system of coupled differential equations, since the displacement of a node depends on the displacement of its neighbors that affects the term $\mathbf{f}_{el}(\cdot)$.

In order to solve this system of coupled differential equations, one can set the natural length of the springs to zero, $l_0 = 0$ and add a constant equilibrium force $\mathbf{f}_{eq} = -\mathbf{f}_{el}$ to the left hand side of (2) [22]. By doing so, the natural state of the model is its initial configuration (Figure 1a). This assumption has the main advantage that the model can be considered within the framework of linear elasticity, i.e., equation (2) is transformed to a set of linear differential

equations with node displacements decoupled in each coordinate, regardless of the magnitude of the displacements.

The deformable surface model in its entirety is ruled by the following Lagrangian dynamics matrix equation [29]:

$$\mathbf{M}\ddot{\mathbf{u}}^{(t)} + \mathbf{C}\dot{\mathbf{u}}^{(t)} + \mathbf{K}\mathbf{u}^{(t)} = \mathbf{f}^{(t)}, \quad (5)$$

where $\mathbf{u}^{(t)} = [\mathbf{u}_1^{(t)} \mathbf{u}_2^{(t)} \dots \mathbf{u}_N^{(t)}]^T$ is the N -dimensional nodal displacements vector $\mathbf{u}^{(t)} = \mathbf{v}^{(t)} - \mathbf{v}^{(t_0)}$, whose elements $\mathbf{u}_i^{(t)}$ are 3-dimensional row vectors. \mathbf{M} , \mathbf{C} , and \mathbf{K} are, respectively, the $N \times N$ mass, damping, and stiffness matrices of the model whose formulation is explained analytically in [25] and $\mathbf{f}^{(t)} = [\mathbf{f}_1^{(t)} \mathbf{f}_2^{(t)} \dots \mathbf{f}_N^{(t)}]^T$ is the N -dimensional vector whose elements are all the external force vectors applied to the model. Each element $\mathbf{f}_i^{(t)}$ of the external forces vector $\mathbf{f}^{(t)}$ can be seen as a virtual spring of natural length zero and stiffness g , joining each node of the surface model to the corresponding pixel in the image (actually the corresponding point on the image intensity surface). Equation (5) is a finite element formulation of the model deformation process.

Instead of finding directly the equilibrium solution of (5), one can use modal analysis [25] and transform it by a basis change [30] to the so-called modal space:

$$\mathbf{u}^{(t)} = \mathbf{\Psi}\tilde{\mathbf{u}}^{(t)}, \quad (6)$$

where $\mathbf{\Psi}$ is a square nonsingular transformation matrix of dimension $N \times N$ and order N to be determined and $\tilde{\mathbf{u}}^{(t)} = [\tilde{\mathbf{u}}_1^{(t)} \tilde{\mathbf{u}}_2^{(t)} \dots \tilde{\mathbf{u}}_N^{(t)}]^T$ is referred to as the *generalized displacement* vector. One effective way of choosing $\mathbf{\Psi}$ is to set it equal to a matrix $\mathbf{\Phi} = [\phi_1, \dots, \phi_N]$, whose columns ϕ_i are the eigenvectors

of the generalized eigenproblem:

$$\mathbf{K}\boldsymbol{\phi}_i = \omega_i^2 \mathbf{M}\boldsymbol{\phi}_i, \quad i = 1, \dots, N, \quad (7)$$

$$\mathbf{u}^{(t)} = \boldsymbol{\Phi} \tilde{\mathbf{u}}^{(t)} = \sum_{i=1}^N \boldsymbol{\phi}_i \tilde{\mathbf{u}}_i^{(t)}, \quad (8)$$

where \mathbf{K} and \mathbf{M} are the stiffness and mass matrices of the model. Equation (8) is referred to as the *modal superposition equation*. The i -th eigenvector $\boldsymbol{\phi}_i$, i.e., the i -th column of $\boldsymbol{\Phi}$ is also called the i -th *vibration mode*. $\tilde{\mathbf{u}}_i^{(t)}$ is the i -th element of $\tilde{\mathbf{u}}^{(t)}$ and ω_i is the corresponding eigenvalue (also called *vibration frequency*). If the matrix $\tilde{\mathbf{C}} = \boldsymbol{\Phi}^T \mathbf{C} \boldsymbol{\Phi}$ is diagonal (called standard Rayleigh hypothesis in [22]), then, the governing matrix-form equation (5) is decoupled into $3N$ scalar equations (N equations for each one of the x, y, z axes) in the modal space:

$$\ddot{\tilde{u}}_{i,x}^{(t)} + \tilde{c}_i \dot{\tilde{u}}_{i,x}^{(t)} + \omega_i^2 \tilde{u}_{i,x}^{(t)} = \tilde{f}_{i,x}^{(t)} \quad (9)$$

$$\ddot{\tilde{u}}_{i,y}^{(t)} + \tilde{c}_i \dot{\tilde{u}}_{i,y}^{(t)} + \omega_i^2 \tilde{u}_{i,y}^{(t)} = \tilde{f}_{i,y}^{(t)}, \quad (10)$$

$$\ddot{\tilde{u}}_{i,z}^{(t)} + \tilde{c}_i \dot{\tilde{u}}_{i,z}^{(t)} + \omega_i^2 \tilde{u}_{i,z}^{(t)} = \tilde{f}_{i,z}^{(t)}, \quad (11)$$

where $i = 1, \dots, N$, $\tilde{\mathbf{u}}_i^{(t)} = [\tilde{u}_{i,x}^{(t)}, \tilde{u}_{i,y}^{(t)}, \tilde{u}_{i,z}^{(t)}]$, \tilde{c}_i is the i -th diagonal element of $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{f}}_i^{(t)} = [\tilde{f}_{i,x}^{(t)}, \tilde{f}_{i,y}^{(t)}, \tilde{f}_{i,z}^{(t)}]$ the i -th component of the transformed external force vector $\tilde{\mathbf{f}}^{(t)}$, where $\tilde{\mathbf{f}}^{(t)} = \boldsymbol{\Phi}^T \mathbf{f}^{(t)}$, $\mathbf{f}^{(t)}$ being the external force vector. Solving these equations at iteration τ leads to $\tilde{\mathbf{u}}_i^{(t)}$ and therefore to $\tilde{\mathbf{u}}^{(t)}$. The displacement vector $\mathbf{u}^{(t)}$ of the model nodes is then obtained by the modal superposition equation (8).

A significant advantage of the formulations described so far, is that the vibration modes (eigenvectors) $\boldsymbol{\phi}_i = [\phi_i(1), \dots, \phi_i(N)]$ and the frequencies (eigenvalues) ω_i of a plane topology (see Figure 1a) have an explicit formulation [22]

and they do not have to be computed using eigen-decomposition techniques:

$$\omega^2(j, j') = \omega_{j \ N_w + j'}^2 = \frac{4\bar{k}}{m} \left[\sin^2 \left(\frac{\pi j}{2N_h} \right) + \sin^2 \left(\frac{\pi j'}{2N_w} \right) \right], \quad (12)$$

$$\phi_{n, n'}(j, j') = \phi_{j \ N_w + j'}(n \ N_w + n') = \cos \frac{\pi j(2n-1)}{N_h} \cos \frac{\pi j'(2n'-1)}{N_w}, \quad (13)$$

where $j = 0, 1, \dots, N_h - 1$, $j' = 0, 1, \dots, N_w - 1$, $n = 1, 2, \dots, N_h$ and $n' = 1, 2, \dots, N_w$.

In our case, where the initial and the final (desirable) deformable surface states, i.e. the initial planar model configuration and the image intensity surface, are known, it is assumed that a constant force load \mathbf{f} is applied to the surface model. The components of the forces in \mathbf{f} along the x and y axes are taken to be equal to zero, i.e. $f_{i,x} = f_{i,y} = 0$. The components of these forces along the z (intensity) axis are taken to be equal to the Euclidean distance between the point $(x, y, I(x, y))$ of the intensity surface and the corresponding model node position in its initial configuration $(x, y, 0)$, i.e., equal to the intensity $I(x, y)$ of pixel (x, y) :

$$f_{(x-1)N_w+y,z} = f(x, y) = I(x, y), \quad x = 1, \dots, N_h, \quad y = 1, \dots, N_w \quad (14)$$

where $f_{(x-1)N_w+y,z}$ is the component along the z axis of the $\mathbf{f}_{(x-1)N_w+y}$ element of vector \mathbf{f} . Under such a condition, the model deforms only along the z axis, affected by constant forces. Thus, equation (5) reduces to the following equilibrium governing equation that corresponds to the static problem:

$$\mathbf{K}\mathbf{u} = \mathbf{f}, \quad (15)$$

or in the modal space:

$$\tilde{\mathbf{K}}\tilde{\mathbf{u}} = \tilde{\mathbf{f}}, \quad (16)$$

where $\tilde{\mathbf{K}} = \Phi^T \mathbf{K} \Phi$ and $\tilde{\mathbf{f}} = \Phi^T \mathbf{f}$, \mathbf{f} being the external force vector.

In the new basis, equation (16) is simplified to $3N$ scalar equations:

$$\begin{aligned}\omega_i^2 \tilde{u}_{i,x} &= \tilde{f}_{i,x}, \\ \omega_i^2 \tilde{u}_{i,y} &= \tilde{f}_{i,y} = 0, \\ \omega_i^2 \tilde{u}_{i,z} &= \tilde{f}_{i,z} = 0,\end{aligned}\tag{19}$$

where $i = 1, \dots, N$. Thus, instead of computing the displacement vector \mathbf{u} from (15), one can compute ω_i^2 and ϕ_i from (12) and (13) respectively, $\tilde{\mathbf{u}}$ from (19) and finally compute \mathbf{u} using (8).

Using (8) and (19) it can be found that the deformations u_{xy} along the intensity axis of the deformable model node that corresponds to pixel (x, y) based on modal analysis for a plane topology and external forces whose z component is given by (14) whereas the other two components are zero, can be described by:

$$u_{xy} = \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} \frac{\sum_{n=1}^{N_h} \sum_{n'=1}^{N_w} I(n, n') \phi_{n,n'}(i, j)}{(1 + \omega^2(i, j)) \sum_{n=1}^{N_h} \sum_{n'=1}^{N_w} \phi_{n,n'}^2(i, j)} \phi_{x,y}(i, j), \tag{20}$$

where u_{xy} is the z component of the element $\mathbf{u}_{(x-1)N_w+y}$ of vector \mathbf{u} (the other two components being zero) and $I(n, n')$ is the image intensity of pixel (n, n') . The above equation presents the final outcome of the aforementioned methodology of deformable surface models adaptation based on Modal Analysis. One can see that deformations are directly related to the eigenvalues $\omega^2(i, j)$ and the eigenvectors $\phi^2(i, j)$ of the model, calculated by (12) and (13) respectively.

Similar to $2D$ signals, one can obtain a representation of a $1D$ signal using $2D$ physics-based deformable modeling. A $1D$ discrete time signal $s(x)$ (with $x = 1, 2, \dots, N$) can be considered as a $2D$ curve $(x, s(x))$. Modeling such a $1D$ signal can be achieved by using an open chain topology of N virtual masses (Figure 3) to approximate the signal. The assumptions for the deformable

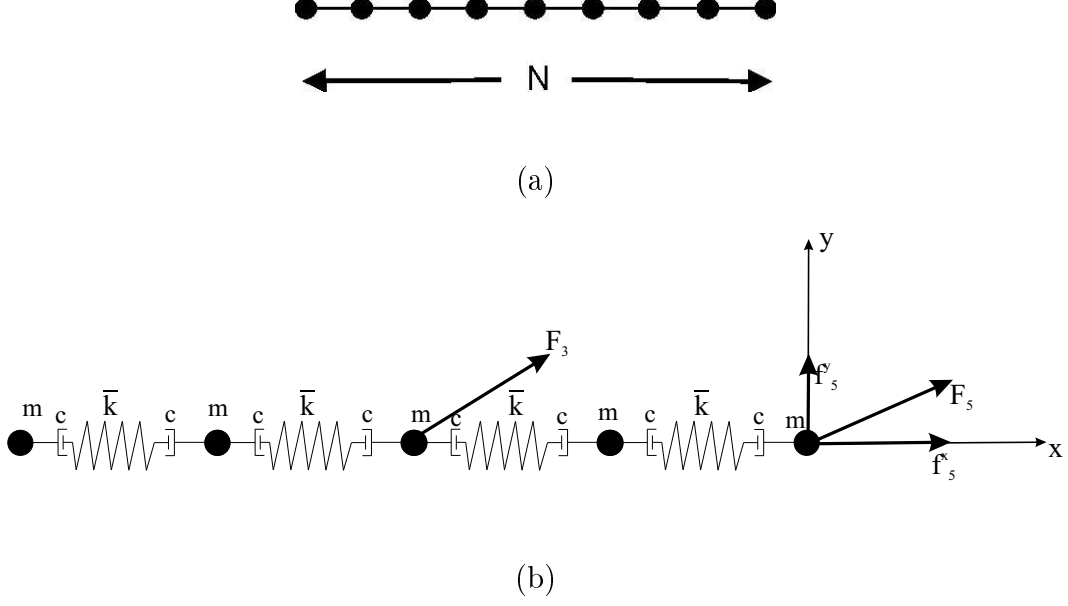


Fig. 3. **(a)** 2D open curve model, **(b)** Example of a 2D curve model comprised of 5 nodes of mass m connected with identical springs of stiffness \bar{k} and damping coefficient c . Two forces act on two model nodes and result in model deformation.

curve model are the same as those of the deformable surface model. The forces acting on the nodes are considered to have only one non-zero component along the $s(x)$ axis. The magnitude of this component is equal to the corresponding signal value $s(x)$. The vibration modes (eigenvectors) ϕ_i and the frequencies (eigenvalues) ω_i of open curves have the following explicit formulation [22]:

$$\omega^2(i) = \frac{4\bar{k}}{m} \sin^2 \left(\frac{\pi i}{2N} \right), \quad (21)$$

$$\phi_n(i) = \cos \frac{\pi i(2n-1)}{2N}, \quad (22)$$

where $\phi_n(i)$ is the n -th element of eigenvector $\phi(i)$, $i \in \{0, 1, \dots, N-1\}$ and $n \in \{1, 2, \dots, N\}$.

Through an analysis similar to that presented for the surface model, it can be

found that the deformations of the node of the 2D deformable curve model that corresponds to sample $s(x)$ of the signal for an open curve (chain topology), can be described by:

$$u_x = \sum_{i=0}^N \frac{\sum_{n=1}^N s(n) \phi_n(i)}{(1 + \omega^2(i)) \sum_{n=1}^N \phi_n^2(i)} \phi_x(i). \quad (23)$$

3 The Discrete Modal Transform (DMT)

In this section, the 1D and 2D Discrete Modal Transform will be introduced. The proposed transform is an intermediate result of the deformation procedure described in Section 2.

In the case of 2D signals, i.e. images, the 3D physics based deformable surface model described in Section 2 is used to approximate the image surface. As already mentioned, the height and width (in nodes) of the deformable model is the height and width (in pixels) of the image and it is assumed, that the surface model can be deformed only along the z -axis.

Thus, the deformations \mathbf{u} of the 3D deformable surface model applied on an image I are given by equation (20) which can be rewritten as:

$$u_{xy} = \sum_{k=0}^{N_h-1} \sum_{l=0}^{N_w-1} \mathcal{F}(k, l) \frac{\phi_{x,y}(k, l)}{\sqrt{\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \phi_{i,j}^2(k, l)}}, \quad x = 1, \dots, N_h, \quad y = 1, \dots, N_w, \quad (24)$$

where

$$\mathcal{F}(k, l) = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} I(i, j) \phi_{i,j}(k, l)}{(1 + \omega^2(k, l)) \sqrt{\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \phi_{i,j}^2(k, l)}}. \quad (25)$$

Equation (24) is applied to each node of the deformable surface model independently.

The term of the deformation equation described by equation (25) is a matrix

of dimensions $N_h \times N_w$. Since (25) involves the image under study, $\mathcal{F}(k, l)$ can be considered to represent the coefficients of a 2D image transform. These coefficients can be used to decompose an image into basis images which have interesting properties.

The normalization factor $\left(\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \phi_{i,j}^2(k, l)\right)$ in the denominator of (25) can be further simplified [31]:

$$\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} \phi_{i,j}^2(k, l) = a(k)a(l), \quad (26)$$

where

$$a(k) = \begin{cases} N, & k = 0 \\ \frac{N}{2}, & k \in \{0, 1, \dots, N-1\}. \end{cases} \quad (27)$$

Thus, using (12), (13), (25), (26) and (27), $\mathcal{F}(k, l)$ can be rewritten as:

$$\mathcal{F}(k, l) = \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(i, j) v_{k,l}(i, j), \quad (28)$$

$$v_{k,l}(i, j) = \frac{\cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w}}{\left[1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h}\right) + \sin^2 \left(\frac{\pi l}{2N_w}\right)\right]\right] \sqrt{a(k)a(l)}}, \quad (29)$$

where $k = 0, 1, \dots, N_h - 1$, $l = 0, 1, \dots, N_w - 1$, $\lambda = 4\frac{\bar{k}}{m}$. Equations (28) and (29) define the proposed 2D Discrete Modal Transform, $\mathcal{F}(k, l)$ being the transform coefficients. From the discussion on Section 2 it is obvious that the parameter λ of the transform controls the elasticity of the corresponding deformable model.

2D DMT is directly related to DCT, which is defined as [9]:

$$\mathcal{C}(k, l) = b(k)b(l) \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(j, j) \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w}, \quad (30)$$

where

$$b(k) = \begin{cases} \sqrt{\frac{1}{N}}, & k = 0 \\ \sqrt{\frac{2}{N}}, & k \neq 0 \end{cases}, \quad k \in \{0, 1, \dots, N-1\}. \quad (31)$$

Table 1

The basis images of the 2D DMT for a block size of dimensions $N_h = 3$, $N_w = 3$ and $\lambda = 1$.

$$\begin{array}{ccc} W_{11} & W_{12} & W_{13} \\ \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \end{bmatrix} & \begin{bmatrix} 0.3266 & 0 & -0.3266 \\ 0.3266 & 0 & -0.3266 \\ 0.3266 & 0 & -0.3266 \end{bmatrix} & \begin{bmatrix} 0.1347 & -0.2694 & 0.1347 \\ 0.1347 & -0.2694 & 0.1347 \\ 0.1347 & -0.2694 & 0.1347 \end{bmatrix} \\ \\ W_{21} & W_{22} & W_{23} \\ \begin{bmatrix} 0.3266 & 0.3266 & 0.3266 \\ 0 & 0 & 0 \\ -0.3266 & -0.3266 & -0.3266 \end{bmatrix} & \begin{bmatrix} 0.3333 & 0 & -0.3333 \\ 0 & 0 & 0 \\ -0.3333 & 0 & 0.3333 \end{bmatrix} & \begin{bmatrix} 0.1443 & -0.2887 & 0.1443 \\ 0 & 0 & 0 \\ -0.1443 & 0.2887 & -0.1443 \end{bmatrix} \\ \\ W_{31} & W_{32} & W_{33} \\ \begin{bmatrix} 0.1347 & 0.1347 & 0.1347 \\ -0.2694 & -0.2694 & -0.2694 \\ 0.1347 & 0.1347 & 0.1347 \end{bmatrix} & \begin{bmatrix} 0.1443 & 0 & -0.1443 \\ -0.2887 & 0 & 0.2887 \\ 0.1443 & 0 & -0.1443 \end{bmatrix} & \begin{bmatrix} 0.0667 & -0.1333 & 0.0667 \\ -0.1333 & 0.2667 & -0.1333 \\ 0.0667 & -0.1333 & 0.0667 \end{bmatrix} \end{array}$$

By comparing (27), (28), (29) to (30), (31) it is easy to conclude that DMT coefficients \mathcal{F} are related with DCT coefficients as follows:

$$\mathcal{F}(k, l) = \frac{1}{\left[1 + \lambda \left(\sin^2\left(\frac{\pi k}{2N_h}\right) + \sin^2\left(\frac{\pi l}{2N_w}\right)\right)\right]} \mathcal{C}(k, l) = \frac{\mathcal{C}(k, l)}{Z(k, l)}. \quad (32)$$

One can easily notice that when $\lambda = 0$, $\mathcal{F}(k, l) = \mathcal{C}(k, l)$. Thus, we can claim, that the proposed transformation is a generalization of the DCT, which is also implied by the fact that they have similar properties, as will be shown in Section 4. Essentially, DMT coefficients are scaled versions of DCT coefficients, the scaling factor being the denominator $Z(k, l)$ in (4). A plot of the scaling

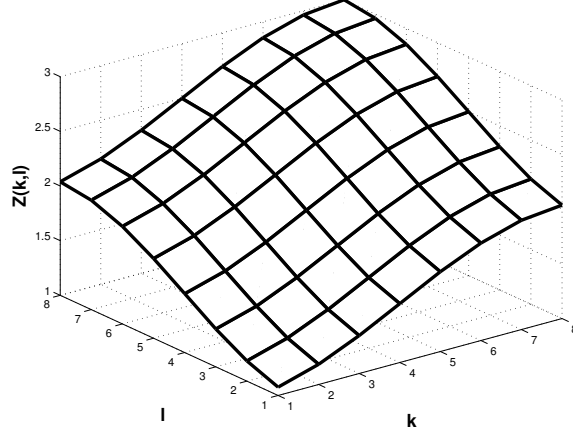


Fig. 4. The denominator of equation (32) for $N_h = N_w = 8$ and $\lambda = 1$.

factors for $N_h = N_w = 8$ and $\lambda = 1$ can be seen in Figure 5. The scaling factor involves the parameter λ , that as will be shown in sections 5 and 6, controls the energy compaction properties of the transform. $Z(k, l)$ plays a role similar to that of quantization matrix used in DCT based coding in JPEG standard. Thus, DMT can be seen as a DCT combined with a new, analytically computed quantization matrix.

The inverse DMT transform is expressed as:

$$I(i, j) = \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \mathcal{F}(k, l) w_{k,l}(i, j), \quad (33)$$

where $i = 0, 1, \dots, N_h - 1$, $j = 0, 1, \dots, N_w - 1$ and $w_{k,l}(i, j)$ are given by:

$$w_{k,l}(i, j) = \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} \frac{1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right]}{\sqrt{a(k)a(l)}}. \quad (34)$$

$w_{k,l}(i, j)$ constitute the basis images of the 2D Discrete Modal Transform. The proof for the inverse 2D DMT is included in the Appendix A.1. The basis images of the proposed transform can be seen in Table 1 for block size of dimensions $N_h = 3$, $N_w = 3$ and $\lambda = 1$. In Table 2, the corresponding basis images of block size 3×3 are illustrated for the DCT. Obviously, the basis

Table 2

The basis images of 2D DCT for a block size of dimensions $N_h = 3$ and $N_w = 3$.

$$\begin{array}{ccc}
W_{11} & W_{12} & W_{13} \\
\begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \\ 0.3333 & 0.3333 & 0.3333 \end{bmatrix} & \begin{bmatrix} 0.4082 & 0 & -0.4082 \\ 0.4082 & 0 & -0.4082 \\ 0.4082 & 0 & -0.4082 \end{bmatrix} & \begin{bmatrix} 0.2357 & -0.4714 & 0.2357 \\ 0.2357 & -0.4714 & 0.2357 \\ 0.2357 & -0.4714 & 0.2357 \end{bmatrix} \\
\\
W_{21} & W_{22} & W_{23} \\
\begin{bmatrix} 0.4082 & 0.4082 & 0.4082 \\ 0 & 0 & 0 \\ -0.4082 & -0.4082 & -0.4082 \end{bmatrix} & \begin{bmatrix} 0.5 & 0 & -0.5 \\ 0 & 0 & 0 \\ -0.5 & 0 & 0.5 \end{bmatrix} & \begin{bmatrix} 0.2887 & -0.5774 & 0.2887 \\ 0 & 0 & 0 \\ -0.2887 & 0.5774 & -0.2887 \end{bmatrix} \\
\\
W_{31} & W_{32} & W_{33} \\
\begin{bmatrix} 0.2357 & 0.2357 & 0.2357 \\ -0.4714 & -0.4714 & -0.4714 \\ 0.2357 & 0.2357 & 0.2357 \end{bmatrix} & \begin{bmatrix} 0.2887 & 0 & -0.2887 \\ -0.5774 & 0 & 0.5774 \\ 0.2887 & 0 & -0.2887 \end{bmatrix} & \begin{bmatrix} 0.1667 & -0.3333 & 0.1667 \\ -0.3333 & 0.6667 & -0.3333 \\ 0.1667 & -0.3333 & 0.1667 \end{bmatrix}
\end{array}$$

images of the two transforms have a similar structure.

In the case of 1D signals, the 2D physics based deformable curve model is used to represent the signal. The length of the model is the signal size and it is assumed that the curve model can be deformed only along the y -axis. Using a methodology similar to that presented above for the 2D case, the 1D Discrete Modal Transform pair for a signal $s(i)$ can be introduced:

$$\mathcal{F}(k) = \sum_{i=0}^{N-1} s(i) \frac{\cos \frac{\pi k(2i+1)}{2N}}{\left[1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N}\right)\right]\right] \sqrt{a(k)}}, \quad (35)$$

$$s(i) = \sum_{k=0}^{N-1} \left[\mathcal{F}(k) \cos \frac{\pi k(2i+1)}{2N} \frac{1 + \lambda \sin^2 \left(\frac{\pi k}{2N}\right)}{\sqrt{a(k)}} \right]. \quad (36)$$

Basis vectors for the proposed 1D Modal Transform ($\lambda = 1$) and the 1D DCT can be seen in Figure 5 for block size 8.

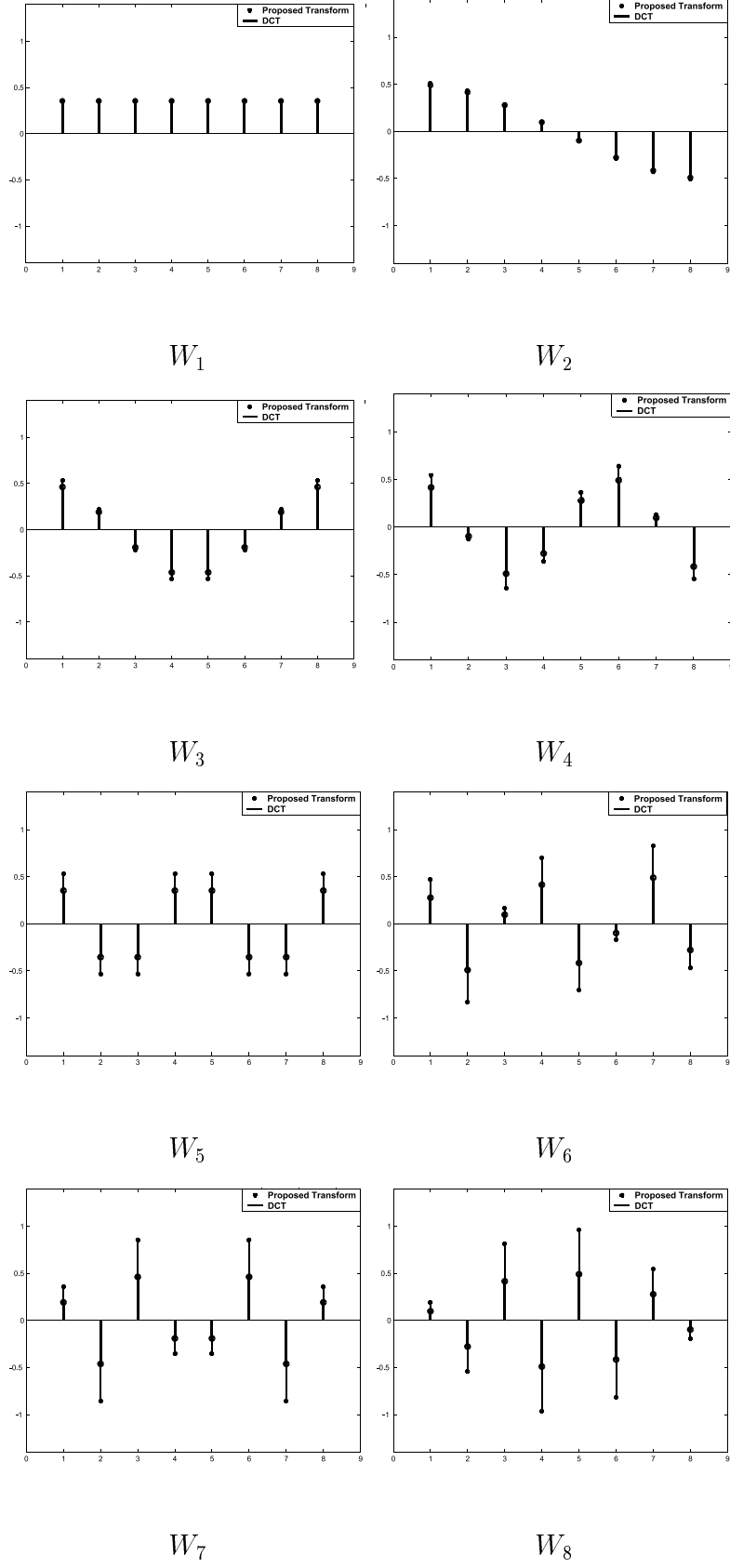


Fig. 5. The basis vectors of block size 8 of the 1D DMT ($\lambda = 1$) and the DCT.

4 Properties of the DMT

In this Section, the properties of the introduced $2D$ transform will be presented and their similarities with those of the DCT will be highlighted. The properties of the $1D$ realization of the transform, are similar to the $2D$ case and thus, will not be presented in this Section.

The DMT is a linear and real transform. Moreover, the DMT and inverse DMT pair form an orthonormal transform pair, i.e.

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} v_{k,l}(i,j) w_{k',l'}(i,j) = \delta(k - k', l - l'), \quad (37)$$

where δ is the unit impulse function. The proof for the orthonormality and the orthogonality of the DMT-inverse DMT pair is included in the Appendix A.2. However, unlike DCT, DMT basis images $w_{k,l}(i,j)$ are orthogonal but not orthonormal (see Appendix A.2), i.e.

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} w_{k,l}(i,j) w_{k',l'}(i,j) = \begin{cases} g \neq 1, & k = k' \text{ and } l = l' \\ 0, & \text{otherwise} \end{cases}. \quad (38)$$

Thus, DMT is not an orthonormal (unitary) transform and the energy is not preserved in the transform domain. Furthermore, unlike DCT, DMT is non-separable, as can be seen from (32).

Since DMT is a non-separable image transform, its computational complexity for a $N \times N$ image is of the order of $O(N^4)$. However, since $2D$ DMT can be calculated using DCT, through (32), its complexity can be further reduced to $O(N^2 \log_2 N)$ if fast DCT implementations are used for its calculation [32].

The basis functions of DMT correspond to well known image processing oper-



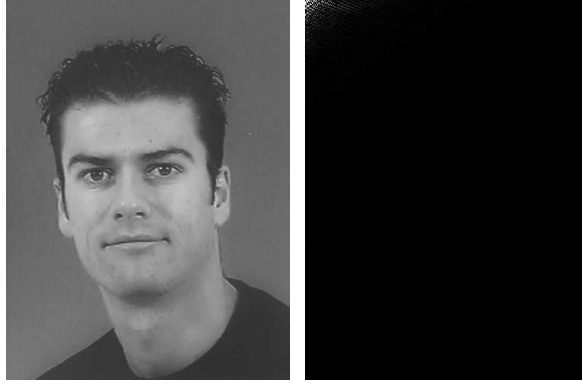
(a)

(b)

Fig. 6. *(a) An outdoor image, (b) DMT coefficients for $\lambda = 1$.*

ators, typically local line and edge detectors, multiplied by a coefficient that depends on λ . For the simplest case, where $N_h = N_w = 3$ (see Table 1), basis images \mathbf{W}_{12} and \mathbf{W}_{21} are the Prewitt operators [33] which detect edges in both vertical and horizontal directions. Additionally, basis images \mathbf{W}_{13} and \mathbf{W}_{31} are vertical and horizontal line detection masks [33]. Moreover, \mathbf{W}_{23} and \mathbf{W}_{32} are edge detectors [34] and \mathbf{W}_{33} is the Laplacian line detection mask. The masks for other values of N_h and N_w also correspond to local line and edge detection operators.

The DMT has excellent energy compaction properties, similar to the ones of the DCT. Figures 6 and 7 illustrate the energy compaction in the frequency domain for two images using the DMT (with $\lambda = 1$). It is obvious, that the energy for both images is packed into the low frequency region, i.e., the top left region. This property will be further investigated in the following section.



(a)

(b)

Fig. 7. *(a)* A human face, *(b)* DMT coefficients for $\lambda = 1$.

5 DMT Energy Packing and Decorrelation Efficiency

In this section, the energy packing ability and the decorrelation efficiency of DMT are examined using test images as well as a stochastic image model.

In the first set of experiments, the DMT was applied to various images and the percentage of the energy compaction for the 3% of the transform coefficients lying at the low-frequency region on the top left quadrant of the frequency domain (Figure 8), was calculated. In other words, the fraction of the total transform domain energy residing into the coefficients of the white area in Figure 8, was calculated. Table 3 shows the energy compaction in the region shown in Figure 8 for various images, for various values of the DMT parameter λ . The block size that was used to compute the transformed images was the actual image size. The results show that the proposed transform compacts the energy into a few coefficients in the low frequency region. More specifically, as λ increases DMT packs more energy in the the low frequency region. Thus,



Fig. 8. *The percentage of the energy compaction is calculated for the white area shown above. This area corresponds to 3% of the frequency domain.*

DMT includes an inherent way of controlling its energy compaction properties, through the parameter λ . This is an important characteristic, especially since λ is directly connected to the deformable model through which the DMT is derived, being the parameter that controls the model elasticity. A large value of λ results in a rigid deformable model that can hardly deform, thus resulting in a coarse, smooth approximation of the target image intensity surface that eliminates the high frequency components of the image. Since DMT coefficients form the image-dependent part of the model deformation (see equation (24)), the energy is concentrated on the low-frequency region. On the other hand, a small value of λ leads to a highly elastic deformable model that deforms to produce a close approximation of the image intensity surface, retaining most of the high frequency components. This in turn results in the energy being more spread over the DMT coefficients domain.

In the following set of experiments, the performance of the proposed transform was evaluated by using two metrics of transform efficiency, i.e. the decorrelation efficiency and energy packing ability, similar to the ones used in [35], [36].

Table 3

Percentage of total energy residing in the 3% low-frequency region (white area of Figure 8) for DMT and different values of λ .

Test images	DMT		
	$\lambda = 1$	$\lambda = 10$	$\lambda = 20$
Lenna	92.85	98.26	100
Baboon	83.33	99.01	100
outdoor image	94.25	100	100
indoor image	82.54	100	100
facial image	94.52	100	100
studio image	74.13	90.43	96.42
average values	86.94	97.95	99.41

These metrics were calculated by assuming a covariance model for the input image and evaluating the covariance function of the transform coefficients. The input image was modeled as a zero mean unit variance wide sense stationary 2D Markov random field with an exponential isotropic non-separable covariance function with interelement correlation coefficient ρ , which is a frequently used model in image processing [32]. The covariance matrix for an $M \times N$ image $I(i, j)$ that adheres to this model is given by:

$$Cov(i, j, i', j') = E[I(i, j)I(i', j')] = \rho^{\sqrt{(i-i')^2 + (j-j')^2}}, \quad (39)$$

where $E[\cdot]$ denotes the expectation operator.

Given an image transform with basis images $v_{kl}(i, j)$ with $k, i \in \{0, 1, \dots, N-1\}$, $l, j \in \{0, 1, \dots, M-1\}$, the covariance matrix $Cov_T(k, l, k', l')$ of the transform coefficients $\mathcal{F}(k, l)$ can be calculated as follows:

$$\begin{aligned}
Cov_T(k, l, k', l') &= E[\mathcal{F}(k, l)\mathcal{F}(k', l')] \\
&= E \left[\left[\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} I(i, j) v_{kl}(i, j) \right] \left[\sum_{i'=0}^{N-1} \sum_{j'=0}^{M-1} I(i', j') v_{k'l'}(i', j') \right] \right] \\
&= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{i'=0}^{N-1} \sum_{j'=0}^{M-1} E[I(i, j)I(i', j')] v_{kl}(i, j) v_{k'l'}(i', j') \\
&= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{i'=0}^{N-1} \sum_{j'=0}^{M-1} Cov(i, j, i', j') v_{kl}(i, j) v_{k'l'}(i', j'). \quad (40)
\end{aligned}$$

$$(41)$$

The metrics of decorrelation and energy packing ability for 1D signals and transforms in [35], [36], were extended to the 2D case. Hence, the decorrelation efficiency of a 2D transform can be expressed as:

$$DE = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |Cov_T(i, j, i, j)|}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{i'=0}^{N-1} \sum_{j'=0}^{M-1} |Cov_T(i, j, i', j')|}. \quad (42)$$

An efficient transform should results in transform coefficients that are as much uncorrelated as possible, i.e. in the ideal case $E[\mathcal{F}(k, l)\mathcal{F}(k', l')] = Cov_T(k, l, k', l')$ should be zero for $k \neq k'$ or $l \neq l'$. In this case $DE = 1$. In general, large values of DE (i.e. values close to 1) indicate good decorrelation properties.

The energy packing ability is calculated by:

$$EPA(\eta) = \frac{\sum_{i=0}^{\eta} \sum_{j=0}^{\eta} |Cov_T(i, j, i, j)|}{\sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |Cov_T(i, j, i, j)|}, \quad (43)$$

where η is the amount of the coefficients retained. Energy packing ability measures, in the transform domain, the ratio of energy that is contained in the $\eta \times \eta$ “low-frequency” coefficients to the energy of all coefficients. High EPA values (i.e. values close to 1) indicate high energy packing. It should be noted that for a non-orthonormal transform, like DMT, the total energy in the transform domain is different from that in the spatial domain. However,

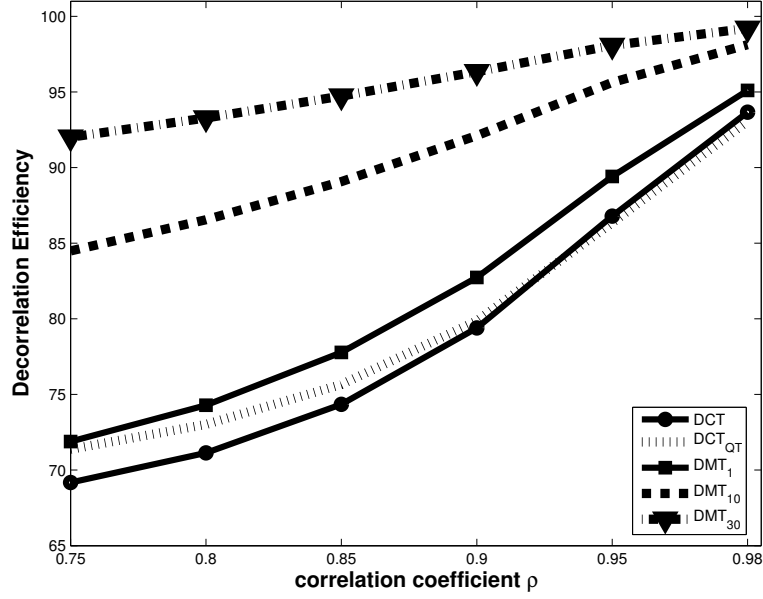


Fig. 9. Decorrelation efficiency (DE) versus ρ . Curves are provided for (a) DMT, $\lambda = 1$ (DMT_1), (b) DMT, $\lambda = 10$ (DMT_{10}), (c) DMT, $\lambda = 30$ (DMT_{30}), (d) DCT, (DCT), (e) DCT with quantization table (DCT_{QT}).

since EPA is expressed as an energy ratio (in the transform domain), it is a valid efficiency metric in this case too.

Plots of the decorrelation efficiency of DMT for values of parameter λ equal to 1 (DMT_1), 10 (DMT_{10}) and 30 (DMT_{30}) when the interelement correlation coefficient ρ (eq. (39)) ranges from 0.75 to 0.98 (i.e. for high correlated 2D signals, a typical situation in images) are plotted in Figure 9. The DE versus ρ values for DCT (denoted as DCT) and DCT scaled with the elements of the quantization table (see Table 4) found in the Annex of the JPEG standard (denoted as DCT_{QT}) are also presented in Figure 9. A transform size of 8×8 , i.e. $N = M = 8$ was used for both DCT and DMT. Curves of the energy packing ability (EPA) versus the interelement correlation coefficient ρ for the same algorithms and $\eta = 2$ and 3 (i.e. for 4 and 9 lower frequency coefficients) are presented in Figures 10 and 11 respectively.

Table 4

DCT Quantization Table: Luminance.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

These Figures show that DMT has very good energy packing and decorrelation properties that get better as λ increases. It should be noted however, that the improved performance of DMT does not stem from the transform *per se*, since DMT is essentially a DCT followed by a division of its coefficients with appropriate, coefficient-dependent, scaling factors. As λ increases, these scaling factors tend to favour the low frequency coefficients, thus increasing the EPA of the algorithm. Hence, safer conclusions regarding the performance of the algorithm with respect to DCT can be obtained within an image compression framework, as will be detailed in the next section. Note also that, as is obvious from (40), (42) and (43), the application of a scaling factor Q on the DCT quantization matrix, does not change the shape of energy packing and decorrelation curves of DCT (i.e. curves DCT_{QT} in Figures 9-11).

6 Application of DMT to Image Compression

Although the main contribution of this paper is the introduction of the DMT, an application of the DMT to lossy grayscale image compression is presented in this Section. The proposed transform is compared with the DCT in terms

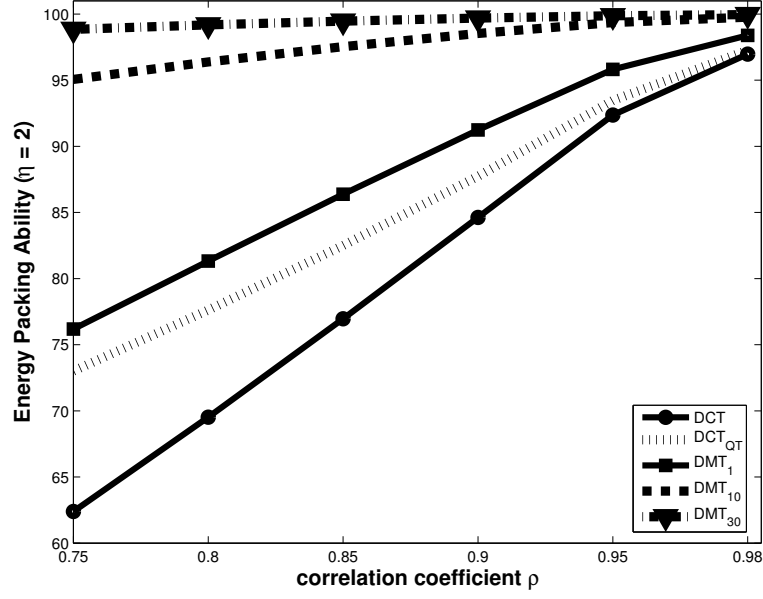


Fig. 10. Energy packing ability $EPA(\eta)$ versus ρ for $\eta = 2$. Curves are provided for (a) DMT, $\lambda = 1$ (DMT₁), (b) DMT, $\lambda = 10$ (DMT₁₀), (c) DMT, $\lambda = 30$ (DMT₃₀), (d) DCT, (DCT), (e) DCT with quantization table (DCT_{QT}).

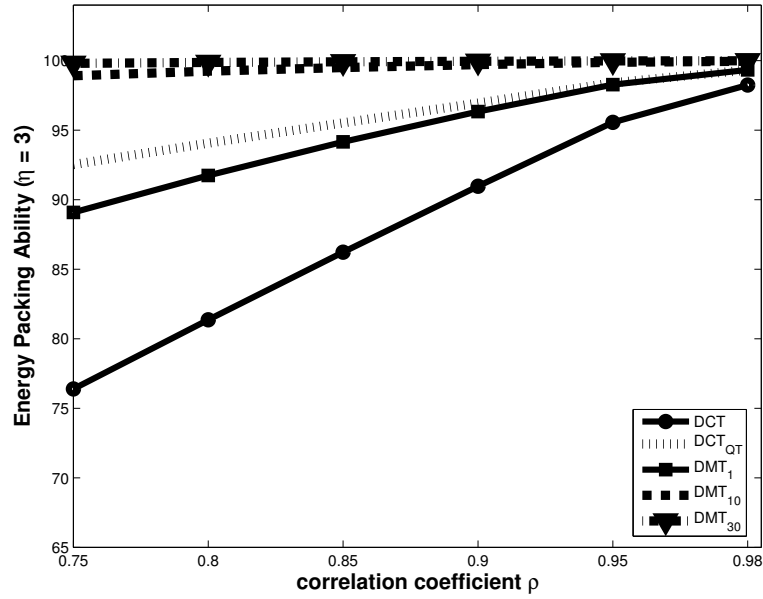
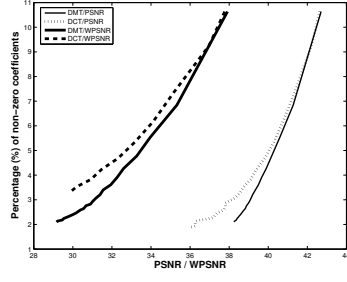


Fig. 11. Energy packing ability $EPA(\eta)$ versus ρ for $\eta = 3$. Curves are provided for (a) DMT, $\lambda = 1$ (DMT₁), (b) DMT, $\lambda = 10$ (DMT₁₀), (c) DMT, $\lambda = 30$ (DMT₃₀), (d) DCT, (DCT), (e) DCT with quantization table (DCT_{QT}).

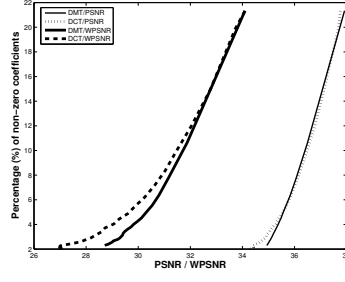
of level of compression and compressed image quality. We chose to use DCT, because of the similarities between the proposed transform and DCT. Furthermore, DCT is the most widely used transform in many image compression schemes, e.g. in JPEG.

The first set of experiments dealt with the evaluation of the quality of the compressed images when DMT and DCT are applied to the target images. To perform image compression, the same procedure is used for both transforms. Both $2D$ transforms are applied to an image, by using 8×8 blocks. In order to discard high-frequency details and achieve compression, the DCT output must be quantized. In this experiment, the quantization table found in the Annex of the JPEG standard [15] (Table 4), multiplied by a scaling coefficient Q to enable variable compression, was used for all image blocks. No quantization table was used for DMT, since the coefficient λ along with the scaling factor in the denominator of (32) essentially act as a quantization table. Once the frequency coefficients of DCT were divided by the values of the quantization table, both the DMT and DCT outputs were rounded to their closest integers. The frequency components that either have a small coefficient or a large divisor in the quantization table will likely round to zero. By sorting the frequency components (e.g. zig-zag scanning), one will typically end up with a run of zeros at the end of the coefficient vector, which can be discarded for compression purposes. In order to acquire the compressed image, first, the output for DCT is multiplied with the quantization table and then, the inverse transforms are applied to the frequency vectors, for both transforms.

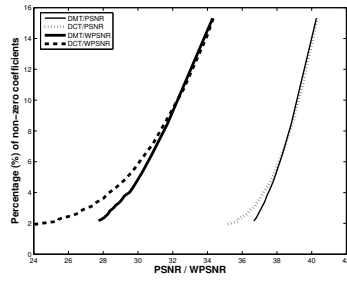
The metrics used for comparing the results in terms of image quality were the peak signal-to-noise ratio (PSNR) between the original and the compressed image, the weighted PSNR (WPSNR) [37] and the total perceptual error (TPE)



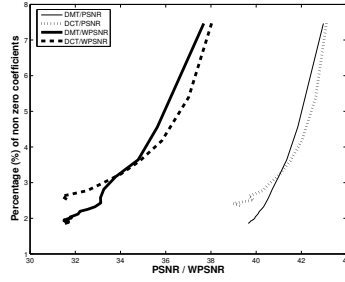
(a)



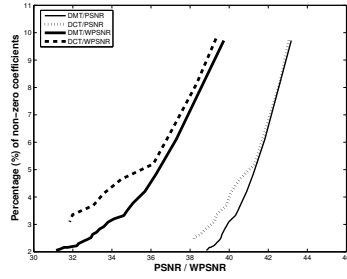
(b)



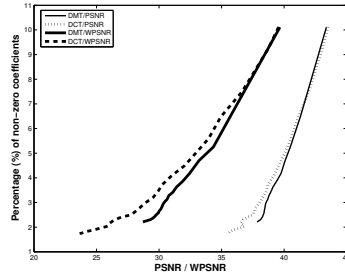
(c)



(d)

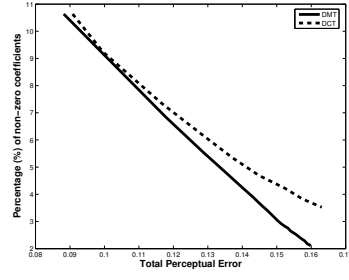


(e)

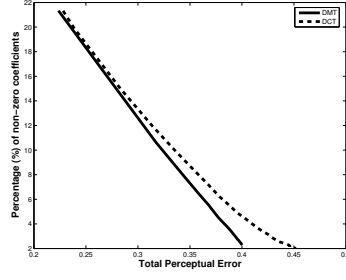


(f)

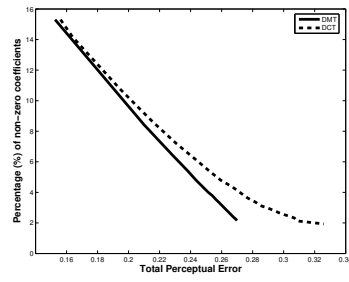
Fig. 12. *PSNR and WPSNR between the original and the compressed image versus the percentage of non-zero coefficients for different values of λ and Q for DMT and DCT respectively, for different test images: (a) Lenna, (b) Mandrill, (c) an outdoor image, (d) an indoor image, (e) a facial image and (f) a studio image.*



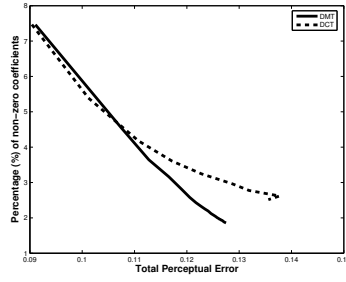
(a)



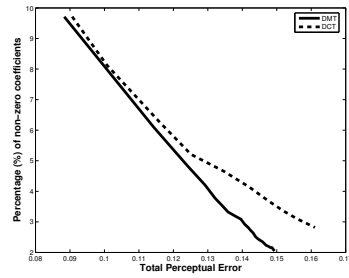
(b)



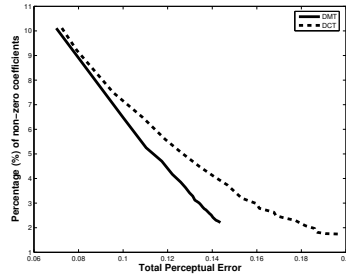
(c)



(d)



(e)



(f)

Fig. 13. The total perceptual error of the Watson metric between the original and the compressed image versus the percentage of non-zero coefficients for different values of λ and Q for DMT and DCT respectively, for different test images: (a) Lenna, (b) Mandrill, (c) an outdoor image, (d) an indoor image, (e) a facial image and (f) a studio image.

of the Watson metric [38]. The formulas used for PSNR and WPSNR are given below:

$$PSNR = 20 \log_{10} \frac{255}{\sqrt{\frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} (I_{original}(i,j) - I_{compressed}(i,j))^2}{N_h N_w}}}, \quad (44)$$

$$WPSNR = 20 \log_{10} \frac{\max(I_{original}(i,j))}{\sqrt{\frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_w} (NVF_{ij} \cdot (I_{original}(i,j) - I_{compressed}(i,j)))^2}{N_h N_w}}}, \quad (45)$$

$$NVF_{ij} = \frac{1}{1 + \sigma_{ij}^2}, \quad (46)$$

where NVF_{ij} is the noise visibility function and σ_{ij}^2 is the local variance of an image I in a window centered on the pixel with coordinates (i, j) . Even though the weighted PSNR and the Watson metric are more sophisticated and efficient than PSNR in representing image quality as perceived by humans, they are not the best way to compare the two algorithms, more so because the JPEG quantization table for DCT was not designed in order to maximize the objective performance but rather the subjective performance of the algorithm. However they have been preferred over subjective quality testing since the latter is a very cumbersome procedure.

In order to compare the compression results of the two transforms various images of different sizes and content, namely facial images, studio images, images depicting humans, indoor and outdoor scenes were used. By varying the coefficients λ and Q for DMT and DCT respectively, different levels of compression were achieved. In this experimental setup, compression was measured as the percentage of non-zero frequency coefficients in the compressed images. Plots of image quality (PSNR, WPSNR and TPE of Watson metric) versus the number of non-zero frequency coefficients for some of the test images can be seen in Figures 12, 13.

Results prove that the DMT can achieve approximately the same or better image quality at the same levels of compression. In general, in this simple compression setup, the DMT achieves better image quality than DCT for high compression levels (small percentage of non-zero coefficients). For example, when only 3% of the coefficients are non-zero DMT achieves a WPSNR improvement over DCT that is equal to 1.6db for an indoor image and 1.8db for a facial image (Figure 2a). The improvements in terms of the TPE of the Watson metric are 0.016 and 0.019 respectively. In low compression levels the two algorithms perform almost the same. Some of the compressed test images are shown in Figure 14. Similar results were obtained in all images used in our experiments.

The second set of experiments dealt with the evaluation of the compression/image quality performance of DMT and DCT when applied for the compression of images in a JPEG-like setup. In this case, the following procedure is applied to both transforms. Both transforms are applied to an image, by using 8×8 blocks. Then, the quantization table found in the Annex of the JPEG standard [15] (Table 4), multiplied by a scaling coefficient Q to enable variable compression, was used for all DCT-transformed image blocks. No quantization table was used for DMT, since as already mentioned, the scaling factor in the denominator of (32) essentially acts as a quantization table. Both the DMT and DCT outputs were rounded to their closest integers. The non zero coefficients were scanned in a zig-zag order and subjected to entropy coding.

A set of images different than the one used in the previous experiment has been used in this case. By varying the coefficients λ and Q for DMT and DCT respectively, different levels of compression were achieved. In this experimental setup, compression was measured in terms of the compression ratio, i.e. as

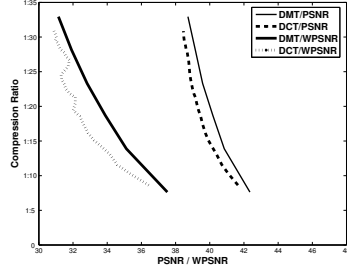


Fig. 14. *Application of DMT and DCT to lossy image compression. Factors Q , λ have been selected so that the two algorithms achieve approximately the same compression for each image. (a), (b), (c): Original images. Compressed images using DMT: (d) $\lambda = 250$, percentage of non zero coefficients = 6% and PSNR = 42.80, (e) $\lambda = 250$, percentage of non zero coefficients = 10% and PSNR = 36.59, (f): $\lambda = 25$, percentage of non zero coefficients = 14.5% and PSNR = 40.27. Compressed images using DCT: (g) $Q = 2$, percentage of non zero coefficients = 6% and PSNR = 42.77, (h) $Q = 2$, percentage of non zero coefficients = 10% and PSNR = 36.45, (i) $Q = 2$, percentage of non zero coefficients = 14.5% and PSNR = 40.11.*

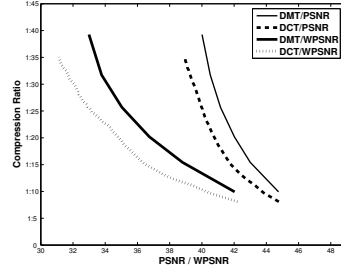
the number of bits in the original image divided by the number of bits in the compressed image, after quantization and entropy coding. Plots of image quality (PSNR, WPSNR and TPE of Watson metric) versus the compression ratio for a subset of the test images used in these experiments can be seen in Figures 15, 16. One can see, that DMT achieves in almost all cases, better image quality than DCT. For example, the WPSNR improvements over DCT are 1.1db for a compression ratio of 1 : 20 in Figure 15b and 0.6db for 1 : 15 compression ratio in Figure 15d. The corresponding improvements in terms of the TPE of the Watson metric are 0.009 and 0.008 respectively. Similar results were obtained in all images used in our experiments.

In the last set of experiments, instead of using standard quantization tables that are based on general experiments for DCT, a technique based on bit-rate control was exploited. More specifically, an optimal bit allocation procedure, reviewed in [39], based on the statistical properties of an image was applied. The image was partitioned in 8×8 non-overlapping blocks and both transforms were applied to each block. The DCT output was scaled by the quantization coefficients that were derived from the statistical properties of the image using the bit allocation procedure in [39] and the selected bit rate. No quantization coefficients were applied to DMT due to its inherent compression control provided by the denominator of (32). Then, both DMT and DCT coefficients were rounded to their closer integer and entropy coding [39] was performed. During the entropy coding, DCT coefficients were ordered by using the zig-zag scan, whereas the DMT coefficients were sorted in ascending order based on the denominator (32):

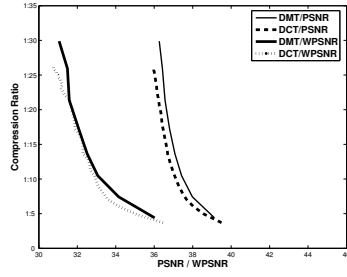
$$Z(k, l) = 1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right], \quad (47)$$



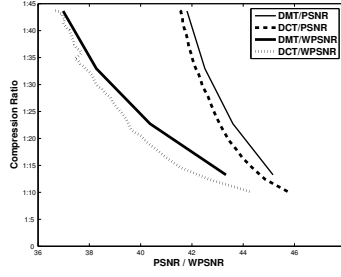
(a)



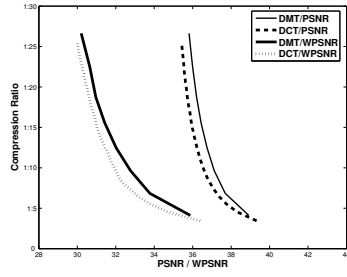
(b)



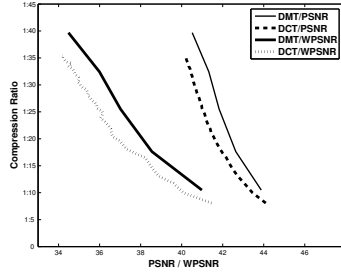
(c)



(d)

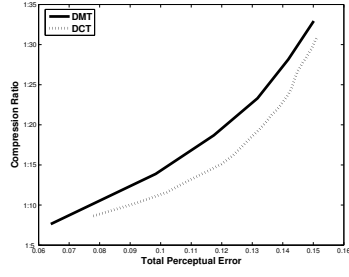


(e)

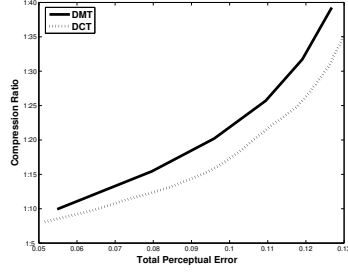


(f)

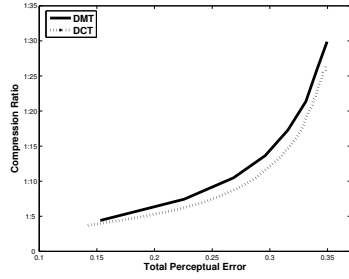
Fig. 15. Compression ratio-distortion curves for both DMT and DCT, for various test images: (a) a lake image, (b) a house image, (c) an animal image, (d) a child image, (e) a flower image and (f) a portrait image. The distortion is measured in terms of PSNR and WPSNR.



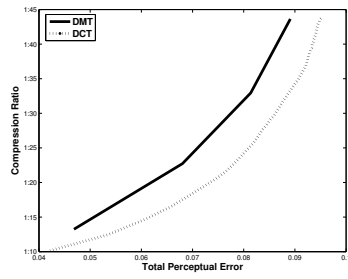
(a)



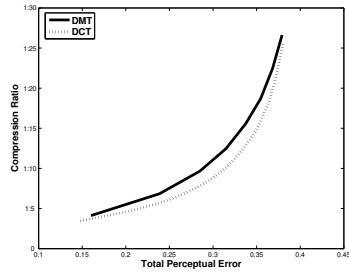
(b)



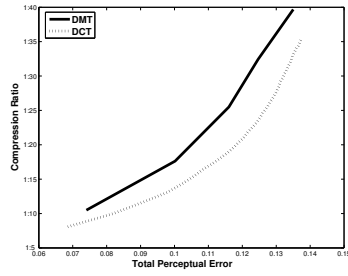
(c)



(d)



(e)



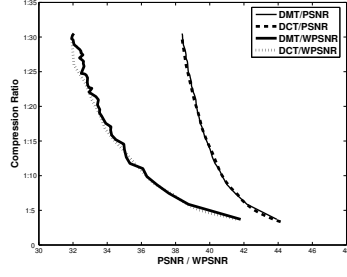
(f)

Fig. 16. *Compression ratio-distortion curves for both DMT and DCT for various test images: (a) a lake image, (b) a house image, (c) an animal image, (d) a child image, (e) a flower image and (f) a portrait image. The distortion is measured in terms of the total perceptual error of the Watson metric.*

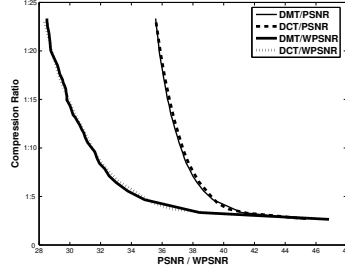
where N_h, N_w are the dimensions of the $2D$ transform. The DMT coefficients are divided by this term, thus, sorting based on this term means that the coefficients which are divided with small values of Z have larger information and must be kept and the coefficients which are divided with large values of Z may be discarded. In Figures 17, 18, the compression ratio-distortion curves for a subset of the test images used in our experiments are depicted. In the case of DMT, compression was controlled by varying λ . Similar to the previous experiment, compression ratio was measured as the number of bits in the original image divided by the number of bits in the compressed image, after quantization and entropy coding, whereas distortion was measured using PSNR and WPSNR (Figure 17) as well as the TPE of the Watson metric (Figure 18). One can see, that DMT achieves in most cases, better image quality than DCT for high compression ratios. For example, the WPSNR improvements over DCT are 0.6db for a compression ratio of 1 : 15 in Figure 17a, 1.1db for a compression ratio of 1 : 35 for the facial image (Figure 2a) and 0.5db for 1 : 35 compression ratio for Figure 17d. The corresponding improvements in terms of the TPE of the Watson metric are 0.004, 0.007 and 0.006 respectively. At lower compression ratios, the two transforms have almost the same performance. Similar results were obtained in all images used in our experiments.

7 Conclusions

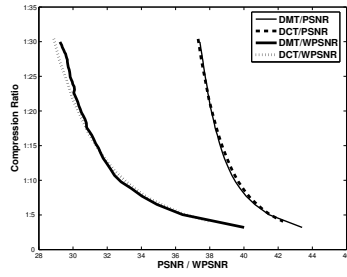
Novel $1D$ and $2D$ discrete, non-separable, signal transforms and their inverse formula were introduced in this paper. The proposed $1D$ and $2D$ transforms are an intermediate result of the deformation procedure of a $2D$ and $3D$



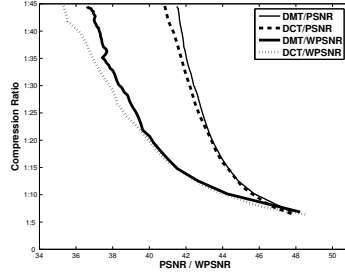
(a)



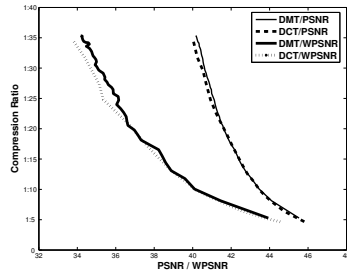
(b)



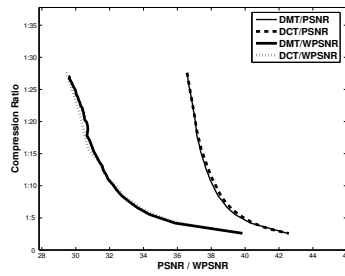
(c)



(d)

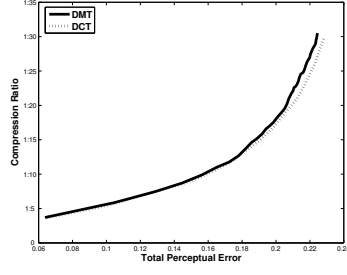


(e)

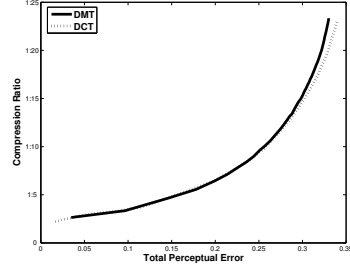


(f)

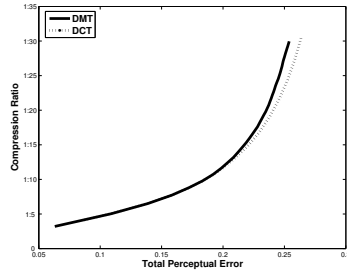
Fig. 17. *Compression ratio-distortion curves for both DMT and DCT, for various test images depicting: (a) a garden, (b) a basket, (c) the sea, (d) a woman, (e) a human face and (f) a forest. The distortion is measured in terms of PSNR and WPSNR.*



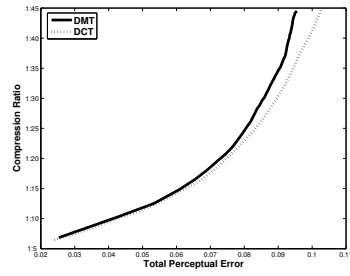
(a)



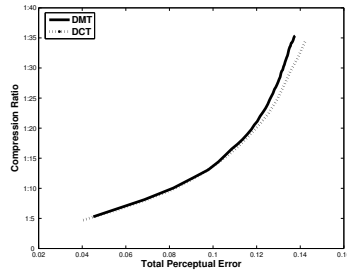
(b)



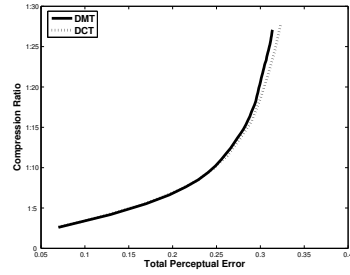
(c)



(d)



(e)



(f)

Fig. 18. *Compression ratio-distortion curves for both DMT and DCT for various test images depicting: (a) a garden, (b) a basket, (c) the sea, (d) a woman, (e) a human face and (f) a forest. The distortion is measured in terms of the total perceptual error of the Watson metric.*

physics-based deformable model respectively. The proposed transform is a DCT-like transform that includes the DCT as a special case and thus exhibits similar properties to DCT. The fact that DCT can be derived using as starting point a deformable model that tries to approximate the intensity surface of an image, is *per se*, an important outcome of this study. Essentially, the coefficients (or equivalently, the basis images) of the proposed transform are scaled versions of the DCT coefficients (or basis images) and so the transform can be also viewed as a means of introducing a new quantization table for DCT. Thus, the proposed transform includes an inherent and physically meaningful compression level selection mechanism which allows handling of the compression ratio without any particular quantization procedure. We applied the proposed transform to lossy image compression and compared it with DCT, since the latter is widely used in image compression. The results show that the proposed transform can achieve very good energy packing and coefficient decorrelation, comparable image quality to DCT at low levels of compression and in most cases, better image quality for high compression levels.

Future work includes performing subjective comparisons between DCT and DMT, using a panel of viewers as well as trying to provide a reasoning also why the quantization matrix included in DMT provides better results than the one included in the JPEG standard. Such an investigation will aim at explaining the fact in terms of the structure and properties of the human visual system.

8 Acknowledgment

This work has been conducted in conjunction with the "SIMILAR" European Network of Excellence on Multimodal Interfaces of the IST Programme of the European Union (<http://www.similar.cc>).

A Appendix

A.1 Inverse DMT Transform

The inverse DMT transform is expressed as:

$$I(i, j) = \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \mathcal{F}(k, l) w_{k,l}(i, j), \quad (\text{A.1})$$

where $\mathcal{F}(k, l)$ denotes the DMT coefficients, N_h and N_w are the dimensions of the transform, $i = 0, 1, \dots, N_h - 1$, $j = 0, 1, \dots, N_w - 1$ and the basis images $w_{k,l}(i, j)$ are given by:

$$w_{k,l}(i, j) = \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} \frac{1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right]}{\sqrt{a(k)a(l)}}. \quad (\text{A.2})$$

In order to prove that (A.1) is indeed the formula of the inverse DMT, we substitute $F(k, l)$ from (28) on the right side of (A.1):

$$\sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(i, j) \frac{\cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w}}{\left[1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right] \right] \sqrt{a(k)a(l)}} w_{k,l}(i, j), \quad (\text{A.3})$$

which can be rewritten as:

$$\begin{aligned}
& \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \frac{1}{a(k)a(l)} \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(i, j) \frac{1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right]}{1 + \lambda \left[\sin^2 \left(\frac{\pi k}{2N_h} \right) + \sin^2 \left(\frac{\pi l}{2N_w} \right) \right]} \\
& \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} = \\
& \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \frac{1}{a(k)a(l)} \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(i, j) \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} \\
& \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} = \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \frac{1}{\sqrt{a(k)a(l)}} \\
& \left[\frac{1}{\sqrt{a(k)a(l)}} \sum_{i=0}^{N_h-1} \sum_{j=0}^{N_w-1} I(i, j) \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} \right] \quad (A.4)
\end{aligned}$$

$$\cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w}. \quad (A.5)$$

The term in brackets in (A.4) is the 2D DCT transform $\mathcal{C}(k, l)$ of image I .

Thus (A.4) can be rewritten as:

$$\begin{aligned}
& \sum_{k=1}^{N_h} \sum_{l=1}^{N_w} \frac{1}{\sqrt{a(k)a(l)}} \mathcal{C}(k, l) \cos \frac{\pi k(2i+1)}{2N_h} \cos \frac{\pi l(2j+1)}{2N_w} = \\
& IDCT(i, j) = I(i, j), \quad (A.6)
\end{aligned}$$

where $IDCT(i, j)$ is the inverse DCT transform.

A.2 Orthonormality of the DMT-inverse DMT pair

A 2D transform pair is orthonormal when the following property is satisfied [32]:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} v_{k,l}(i, j) w_{k',l'}(i, j) = \delta(k - k', l - l'), \quad (B.1)$$

where N is the size of the basis images w , v are the matrices of the transform, $k, l, k', l' = 0, 1, \dots, N-1$ and δ is the unit impulse function:

$$\delta(k, l) = \begin{cases} 1, & k = l = 0 \\ 0, & k \text{ or } l \neq 0 \end{cases}. \quad (\text{B.2})$$

In our case, the transformation matrices v and the basis images w are given by equations (29) and (34) respectively (repeated here for readers' convenience):

$$v_{k,l}(i, j) = \frac{1}{\sqrt{a(k)}} \frac{1}{\sqrt{a(l)}} \cos \frac{\pi k(2i+1)}{2N} \cos \frac{\pi l(2j+1)}{2N} \frac{1}{1 + \lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})}, \quad (\text{B.3})$$

$$w_{k',l'}(i, j) = \frac{1}{\sqrt{a(k')}} \frac{1}{\sqrt{a(l')}} \cos \frac{\pi k'(2i+1)}{2N} \cos \frac{\pi l'(2j+1)}{2N} \left(1 + \lambda \left(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N} \right) \right), \quad (\text{B.4})$$

where $a(\cdot)$ is given by equation (27), and λ is a constant.

Thus, in order to prove that DMT is orthonormal one should prove the following equation:

$$\begin{aligned} B &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{\sqrt{a(k)}} \frac{1}{\sqrt{a(l)}} \cos \frac{\pi k(2i+1)}{2N} \cos \frac{\pi l(2j+1)}{2N} \\ &\quad \frac{1}{1 + \lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})} \frac{1}{\sqrt{a(k')}} \frac{1}{\sqrt{a(l')}} \\ &\quad \cos \frac{\pi k'(2i+1)}{2N} \cos \frac{\pi l'(2j+1)}{2N} \left(1 + \lambda \left(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N} \right) \right) \\ &= \delta(k - k', l - l'). \end{aligned} \quad (\text{B.5})$$

Equation (B.5) can be further analyzed to:

$$B = \frac{1}{\sqrt{a(k) a(k') a(l) a(l')}} \frac{1 + \lambda(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N})}{1 + \lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})} B_1 B_2, \quad (\text{B.6})$$

where B_1 is given by:

$$B_1 = \sum_{i=0}^{N-1} \cos \frac{\pi k(2i+1)}{2N} \cos \frac{\pi k'(2i+1)}{2N}, \quad (\text{B.7})$$

and B_2 is equal to:

$$B_2 = \sum_{j=0}^{N-1} \cos \frac{\pi l(2j+1)}{2N} \cos \frac{\pi l'(2j+1)}{2N}. \quad (\text{B.8})$$

Due to the similarity of B_1 and B_2 one can focus only on the first one.

By a change of the summation index i one can transform B_1 into:

$$\sum_{i=1}^N \cos \frac{\pi k(2i-1)}{2N} \cos \frac{\pi k'(2i-1)}{2N}. \quad (\text{B.9})$$

which can be rewritten to:

$$B_1 = \sum_{i=1}^N \frac{1}{2} \left(\cos \frac{\pi(k-k')(2i-1)}{2N} + \cos \frac{\pi(k+k')(2i-1)}{2N} \right). \quad (\text{B.10})$$

The proof will be separated into 5 distinct cases depending on the values of k , k' and l , l' .

1. $k \neq k'$ or $l \neq l'$

It is known that $0 \leq k < N$ and $0 \leq k' < N$, thus $-N < k-k' < N$ and $-\frac{1}{2} < \frac{k-k'}{2N} < \frac{1}{2}$. Since in this case we have assumed that $k \neq k'$, $\frac{k-k'}{2N} \neq 0$ and thus this term cannot obtain an integer value. As a consequence $\sin\left(\frac{k-k'}{2N}\pi\right) \neq 0$. Moreover, $0 < k+k' < 2N$ and thus using as similar reasoning $\sin\left(\frac{k+k'}{2N}\pi\right) \neq 0$.

The summation $\sum_{i=1}^N \cos(2i-1)x$ can be expressed as follows [40]:

$$\sum_{i=1}^N \cos(2i-1)x = \frac{1}{2} \sin(2Nx) \frac{1}{\sin x}. \quad (\text{B.11})$$

By substituting x with $\frac{\pi(k-k')}{2N}$ in the previous expression, the term $\sum_{i=1}^N \cos \frac{\pi(k-k')(2i-1)}{2N}$ in equation (B.10) can be simplified to:

$$\sum_{i=1}^N \cos \frac{\pi(k-k')(2i-1)}{2N} = \frac{1}{2} \sin \left(\frac{2N(k-k')\pi}{2N} \right) \frac{1}{\sin \frac{(k-k')\pi}{2N}}, \quad (\text{B.12})$$

the term in the denominator being non zero as was shown above. By substituting x with $\frac{\pi(k+k')}{2N}$ in (B.11), the second term of the sum in (B.10) is simplified as follows:

$$\sum_{i=1}^N \cos \frac{\pi(k+k')(2i-1)}{2N} = \frac{1}{2} \sin \left(\frac{2N(k+k')\pi}{2N} \right) \frac{1}{\sin \frac{(k+k')\pi}{2N}}, \quad (\text{B.13})$$

the term in the denominator being again non zero.

Thus, $B_1 = 0$, since $\sin((k-k')\pi)$ and $\sin((k+k')\pi)$ are always equal to zero. In the same way, when $l \neq l'$ $B_2 = 0$ and therefore $B = 0$.

2. $\mathbf{k} = \mathbf{k}' \neq \mathbf{0}$ and $\mathbf{l} = \mathbf{l}' \neq \mathbf{0}$

In this case B_1 is given by the equation:

$$B_1 = \sum_{i=1}^N \frac{1}{2} \left(1 + \cos \frac{\pi k(2i-1)}{N} \right), \quad (\text{B.14})$$

and as we have already seen in the previous case, the second term of the summation is equal to zero. Consequently, $B_1 = \frac{N}{2}$. In the same way, when $l = l' \neq 0$ the factor $B_2 = \frac{N}{2}$. Since k, k', l and l' are all not equal to zero, $a(k) = a(k') = a(l) = a(l') = \frac{N}{2}$. Moreover, the fraction $\frac{1+\lambda(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N})}{1+\lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})}$ in (B.6) is equal to 1 since $k = k'$ and $l = l'$, thus, $B = 1$.

3. $\mathbf{k} = \mathbf{k}' = \mathbf{l} = \mathbf{l}' = \mathbf{0}$

In this case, it is obvious that

$$B_1 = \sum_{i=1}^N \frac{1}{2} (1 + 1) = N. \quad (\text{B.15})$$

In the same way, when $l = l' = 0$, $B_2 = N$. Since $k = k' = l = l' = 0$,

$a(k) = a(k') = a(l) = a(l') = N$ and $\frac{1+\lambda(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N})}{1+\lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})} = 1$. Thus, $B = 1$.

4. $\mathbf{k} = \mathbf{k}' \neq \mathbf{0}$ and $\mathbf{l} = \mathbf{l}' = \mathbf{0}$

From the cases 3 and 4 respectively, one can see that $B_1 = \frac{N}{2}$ and $B_2 = N$. Moreover, $a(k) = a(k') = \frac{N}{2}$, $a(l) = a(l') = N$ and $\frac{1+\lambda(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N})}{1+\lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})} = 1$, thus, $B = 1$.

5. $\mathbf{k} = \mathbf{k}' = \mathbf{0}$ and $\mathbf{l} = \mathbf{l}' \neq \mathbf{0}$

According to the previous case, $B_1 = N$, $B_2 = \frac{N}{2}$, $a(k) = a(k') = N$, $a(l) = a(l') = \frac{N}{2}$, $\frac{1+\lambda(\sin^2 \frac{\pi k'}{2N} + \sin^2 \frac{\pi l'}{2N})}{1+\lambda(\sin^2 \frac{\pi k}{2N} + \sin^2 \frac{\pi l}{2N})} = 1$ and therefore $B = 1$.

Concluding, we can say that DMT is an orthonormal transform pair since equation (B.5) is proven to be equal to:

$$B = \begin{cases} 1, & k = k' \text{ and } l = l' \\ 0, & \text{otherwise} \end{cases} = \delta(k - k', l - l'). \quad (\text{B.16})$$

The proof of the 1D case is trivial and it is omitted.

Moreover, the basis images of DMT are orthogonal and not orthonormal, i.e. the following equation is satisfied:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} w_{k,l}(i, j) w_{k',l'}(i, j) = \begin{cases} g \neq 1, & k = k' \text{ and } l = l' \\ 0, & \text{otherwise} \end{cases}, \quad (\text{B.17})$$

where $w_{k,l}$ form the basis images of DMT. The proof of the afore-mentioned equation is similar to the orthonormality of the DMT pair (B.1) and it is omitted.

References

- [1] S. S. Agaian, K. Panetta, and A. M. Grigoryan. Transform-based image enhancement algorithms with performance measure. *IEEE Transactions on Image Processing*, 10(3):367–382, Mar. 2001.
- [2] M. S. Brown and W. B. Seales. Image restoration of arbitrarily warped documents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1295–1306, Oct. 2004.
- [3] C. Wufan, Z. Jie, S. Yuhua, C. Jianjun, and L. Xianqing. Description of medical images in characteristic subspace and vector quantization coding based on wavelet transformation. In *IEEE International Conference on Image Processing*, volume 1, pages 911–917, Seattle, United States, 16-19 Sept. 1996.
- [4] B. Chiptrasert and K. R. Rao. Discrete Cosine Transform filtering. *Signal Processing*, 19(3):233–245, Mar. 1990.
- [5] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies. Data compression and harmonic analysis. *IEEE Transactions on Information Theory*, 44(6):2435–2476, Oct. 1998.
- [6] H. Feng and M. Effros. On the rate-distortion performance and computational efficiency of the Karhunen-Loeve Transform for lossy data compression. *IEEE Transactions on Image Processing*, 11(2):113–122, Feb. 2002.
- [7] D. Nandy and J. Ben-Arie. Generalized feature extraction using expansion matching. *IEEE Transactions on Image Processing*, 8(1):22–32, Jan. 1999.
- [8] J. T. Chien and C. C. Wu. Discriminant waveletfaces and nearest feature classifiers for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(12):1644–1649, Dec. 2002.

- [9] N. Ahmed, T. Natarajan, and K. Rao. Discrete cosine transform. *IEEE Computers*, C-23:90–93, Jan. 1974.
- [10] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechiess. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205–220, Apr. 1992.
- [11] V. K. Goyal. Theoretical foundations of transform coding. *IEEE Signal Processing Magazine*, 18(5):9–21, Sep. 2001.
- [12] M. Effros, H. Feng, and K. Zeger. Suboptimality of the Karhunen Loeve Transform for transform coding. *IEEE Transactions on Information Theory*, 50(8):1605–1619, Aug. 2004.
- [13] G. K. Wallace. The jpeg still picture compression standard. *Communications of the ACM*, 34(4):30–44, Apr. 1991.
- [14] J. K. Li, J. Li, and C. C. J. Kuo. Layered DCT still image compression. *CirSysVideo*, 7(2):440–443, Apr. 1997.
- [15] W. B. Pennebaker and J. L. Mitchell. *JPEG: Still Image Data Compression Standard*. Van Nostrand Reinhold, New York, 1993.
- [16] Y. Yusong, S. Guangda, W. Chunmei, and S. Qingyun. Invertible integer fft applied on lossless image compression. In *IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, volume 2, pages 1219–1223, 8-13 Oct. 2003.
- [17] I. Valova and Y. Kosugi. Hadamard-based image decomposition and compression. *IEEE Transactions on Information Technology in Biomedicine*, 4:306–319, Dec. 2000.
- [18] Z.-X. Hou, N.-N. Xu, H. Chen, and X.-L. Li. Fast slant Transform with sequency increment and its application in image compression. In *Proceedings of*

2004 *International Conference on Machine Learning and Cybernetics*, volume 7, pages 4085–4089, 26-29 Aug. 2004.

- [19] Didier Le Gall. Mpeg: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):46–58, Apr. 1991.
- [20] Video Codec for Audiovisual Services at $p \times 64$ kbit/s ITU-T Recommendation H.261. *Version 1: ITU-T, ITU-T Recommendation H.261*, 1990.
- [21] ITU Telecom. Standardization Sector of ITU. Video coding for low bitrate communication. *ITU-T Recommendation H.263*, Mar. 1996.
- [22] C. Nastar and N. Ayache. Frequency-based nonrigid motion analysis: Application to four dimensional medical images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1069–1079, 1996.
- [23] C. Nikou, G. Bueno, F. Heitz, and J.P. Armspach. A joint physics-based statistical deformable model for multimodal brain image analysis. *IEEE Transactions on Medical Imaging*, 20(10):1026–1037, 2001.
- [24] S. Krinidis, C. Nikou, and I. Pitas. Reconstruction of serially acquired slices using physics-based modelling. *IEEE Transactions on Information Technology in Biomedicine*, 7(4):394–403, December 2003.
- [25] A. Pentland and S. Sclaroff. Closed-form solutions for physically-based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, Jul 1991.
- [26] B. Moghaddam, C. Nastar, and A. Pentland. A bayesian similarity measure for direct image matching. In *International Conference on Pattern Recognition (ICPR 1996)*, pages 350–358, Vienna, Austria, August 1996.
- [27] G. Borgefors. On digital distance transforms in three dimensions. *Computer Vision and Image Understanding*, 64(3):368–376, 1996.

- [28] P.-E. Danielsson. Euclidean distance transform. *Computer Graphics and Image Processing*, 14:227–28, 1980.
- [29] K. J. Bathe. *Finite Element Procedure*. Prentice Hall, Englewood Cliffs, New Jersey, 1996.
- [30] A. Pentland and B. Horowitz. Recovery of non-rigid motion and structure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):730–742, July 1991.
- [31] M. R. Spiegel. *Mathematical Formulas*. McGraw-Hill, New York, 1996.
- [32] A. K. Jain. *Fundamentals of Digital Image Processing*. Prentice Hall, Englewood Cliffs, New Jersey, 1989.
- [33] R. Gonzalez and R. Woods. *Digital Image Processing*. Addison-Wesley Publishing Company, 1992.
- [34] W. Frei and C.C. Chen. Fast boundary detection: A generalization and a new algorithm. *IEEE Transactions on Computers*, C-26(10):988–998, 1977.
- [35] W. K. Cham and R. J. Clarke. Application of the principle of dyadic symmetry to the generation of orthogonal transforms. *IEE Proceedings, Pt. F*, 133(3):264–270, June 1986.
- [36] R. J. Clarke. *Digital Compression of Still Images and Video*. Academic Press, London, 1995.
- [37] A. Netravali and B. Haskell. *Digital Pictures Representation and Compression*. Plenum Press, New York, 1998.
- [38] A. Mayache, T. Eude, and H. Cherifi. A comparison of image quality models and metrics based on human visual sensitivity. In *Proceedings of the International Conference on Image Processing (ICIP 1998)*, volume 3, pages 409–413, Chicago, IL, USA, October 1998.

- [39] V. Bhaskaran and K. Konstantinides. *Image and Video Compression Standards*. Kluwer Academic Publishers, London, 1995.
- [40] I. S. Gradshteyn, I. M. Ryzhik, and A. Jeffrey. *Table of Integrals, Series and Products, Fifth Edition*. Academic Press, London, 1994.