SYMMETRY-BASED GRAPH FOURIER TRANSFORMS FOR IMAGE REPRESENTATION

Alessandro Gnutti[⋆]*, Fabrizio Guerrini*[⋆]*, Riccardo Leonardi*[⋆] *and Antonio Ortega*[†]

[⋆] Department of Information Engineering, University of Brescia, Italy
[†] Department of Electrical Engineering, University of Southern California, Los Angeles, USA

ABSTRACT

It is well-known that the application of the Discrete Cosine Transform (DCT) in transform coding schemes is justified by the fact that it belongs to a family of transforms asymptotically equivalent to the Karhunen-Loève Transform (KLT) of a first order Markov process. However, when the pixel-to-pixel correlation is low the DCT does not provide a compression performance comparable with the KLT. In this paper, we propose a set of symmetry-based Graph Fourier Transforms (GFT) whose associated graphs present a totally or partially symmetric grid. We show that this family of transforms well represents both natural images and residual signals outperforming the DCT in terms of energy compaction. We also investigate how to reduce the cardinality of the set of transforms through an analysis that studies the relation between efficient symmetry-based GFTs and the directional modes used in H.265 standard. Experimental results indicate that coding efficiency is high.

Index Terms— Graph Fourier Transform, Discrete Cosine Transform, Karhunen-Loève Transform, Symmetry, H.265.

1. INTRODUCTION

A signal given as a vector in $\mathbb{R}^N$ is implicitly represented as a sequence of length N w.r.t. the standard basis. In transform coding schemes [1], an orthogonal transform is used to change the basis motivated by the fact that transform coding may be more effective in the transform domain than in the original signal space thanks to sparsification. The Karhunen-Loève Transform (KLT) [2] is a particular type of orthogonal transform that depends on the covariance of the source and it allows to achieve the optimal energy compaction. However, since the computation time of the associated eigen-decomposition of the covariance matrix is not negligible and since a fast implementation is in general not possible for the KLT, its practical use is actually limited. Under the first-order stationary Markov condition, the covariance matrix becomes the Toeplitz matrix

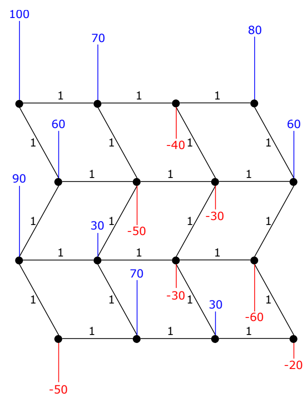
$$\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \rho^2 \\ \vdots & \ddots & \ddots & \ddots & \rho \\ \rho^{N-1} & \cdots & \cdots & \rho & 1 \end{bmatrix}. \tag{1}$$

When the inter-pixel correlation is strong the coefficient ρ is close to 1 and it is proven that the Discrete Cosine Transform (DCT) asymptotically becomes the KLT [3]. This is one of the reasons why the DCT has been so widely used for image/video coding. However, the first-order stationary Markov condition generally occurs for natural image and video signals, but for other kinds of data the sample statistics can be totally different or unknown. For example, even if in video coding scenarios the DCT is applied on residual signals obtained from inter or intra-frame prediction, on the other hand in residual blocks the inter-pixel correlation may be smaller causing performance decrease of the DCT in terms of approximation ability when low amplitude coefficients are dropped. Thus, since the optimality of the DCT holds only under stationarity assumptions regarding the pixel statistical distribution, searching for a unitary transform (or a class of unitary transforms) able to adapt to different inter-pixel correlation configurations is nowadays a reason of interest [4].

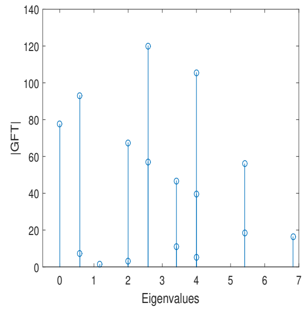
Graph signal processing (GSP) [5] is a novel framework that models inter-sample relation by edge weights in a graph. Signal values lie on the vertices of a weighted graph and the corresponding Graph Fourier Transform (GFT) can be defined through the associated graph Laplacian matrix. This model is very useful since it allows to establish a connection between signal samples and consequently to obtain a transform able to provide for a better energy compaction of the underlying signal [6]. As a matter of fact, focusing on image processing applications, an image block can be represented as a graph where different connectivity patterns lead to different interpretations in the graph transform domain. Some examples can be found in [7], [8] and [9] where optimal GFTs have been studied for inter-predicted video coding, intra-prediction image coding and for piecewise smooth images, respectively. Furthermore, graphs can be also learned from data so that sparse covariance matrices can be designed to approximate the KLT. For example, in [10] graph learning problems are posed as the estimation of graph Laplacian matrices under given structural constraints.

In this work we propose a set of graphs that allows us to characterize symmetries in a 2-D signal. The study of totally or partially symmetric grids is mainly motivated by the fact that symmetries are often present in real-world data [11], [12]. For example, in [13] the authors have worked with graphs estimated from inter-pixel correlation in residual blocks after predictive coding in H.265 standard, showing that the blocks considered exhibit different kinds of symmetry and indicating that there exist potentially useful 2-D GFTs with symmetric grids. Furthermore, when L is a bisymmetric matrix the corresponding GFT allows a fast implementation as in the DCT case [14]; actually, this condition is true for several types of symmetric grids. As such we shall show that the set of symmetry-based GFTs can be compared with the DCT for both natural images and residual data, outperforming the classical transform in both cases.

The rest of the paper is organized as follows. In Section 2 we introduce notation on graphs, focusing on the definition of the GFT. In Section 3 we define a set of symmetry-based GFTs comparing their approximation ability with the DCT and showing a significant performance improvement in terms of energy compaction. In Section 4 we show that the set of effective transforms can be reduced when applied on data corresponding to residuals obtained from intra-



(a) Example of a signal on a 2-D grid graph. Positive values are pictured in blue, negative values in red.



(b) The GFT in the corresponding graph spectral domain. Multiple values correspond to multiple eigenvalues.

Fig. 1: Representation of a signal in the vertex domain and graph spectral domain.

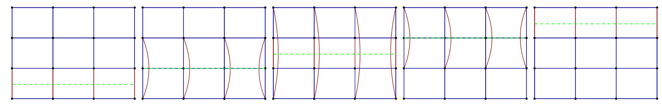
prediction based on the directional modes adopted in H.265. We draw some conclusions in Section 5.

2. PRELIMINARIES

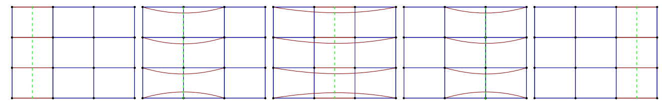
In graph signal processing, signals are defined on an undirected, connected, weighted graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{W}\}$. Such a graph is represented by *a)* a set of vertices $\mathcal{V} = \{1, 2, \ldots, N\}$ to which the signal samples are connected, *b)* a set of edges $\mathcal{E}$ that connect the nodes of the graph that carry the signal values and *c)* a weighted adjacency matrix $\mathbf{W}$ that has the edge weights as entries, i.e., it establishes the relations among the signal samples. In Fig. 1a an example of signal on a graph is shown. Naming $\mathbf{D}$ the diagonal degree matrix of $\mathbf{G}$, we can define the GFT for a given graph through its graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$. The eigen-decomposition of $\mathbf{L}$ is performed, such that $\mathbf{L} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$: the columns of $\mathbf{T}$ represent the basis vectors of the related GFT. Each transform coefficient is associated to an eigenvalue related to a corresponding eigenvector. Note that since the graph Laplacian $\mathbf{L}$ is a real symmetric matrix, it has a complete set of orthonormal eigenvectors, each of them linked to a real, non-negative eigenvalue. Fig. 1b shows the GFT of the signal in Fig. 1a.

3. SYMMETRY-BASED GFT

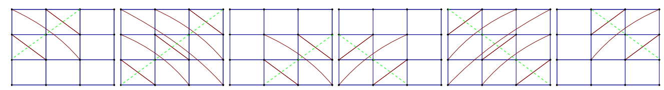
In this work we concentrate our preliminary analysis on 4×4 image blocks. We wish to define a set of graphs whose symmetric grids may well represent symmetries in a 2-D signal. Indeed we assume that symmetric graphs are good at representing signals characterized by the corresponding kind of symmetry. Several symmetry axes could be examined by varying angle and position; due to the discrete nature of a graph, the slopes of the axes that better fit the grid are basically $0°$, $45°$, $90°$ and $135°$, and for each one of them all the significant positions are evaluated. Furthermore, in order to evaluate a complete set of symmetric configurations, all the significant centrosymmetries are taken into account in addition to reflection symmetries. The graph construction based on the mentioned symmetries proceeds as follows (see Fig. 2): for reflection symmetries, nodes in the specular position w.r.t. the symmetry axis are connected, whereas



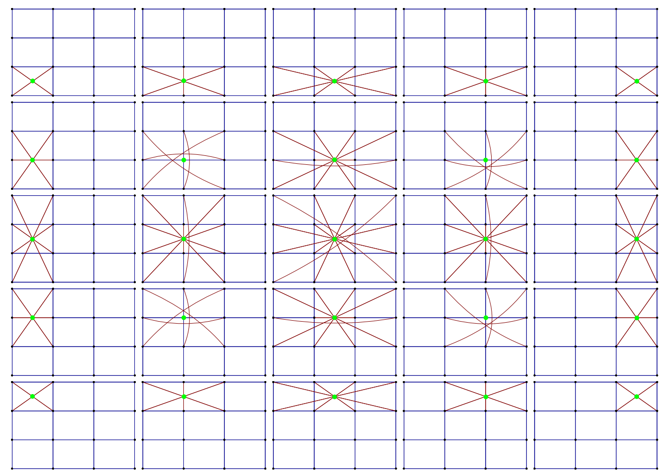
(a) UD symmetric configurations.



(b) LR symmetric configurations.



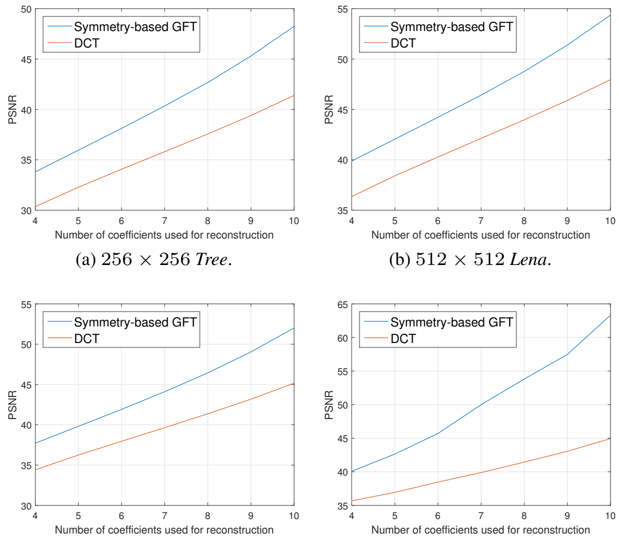
(c) Diagonal and anti-diagonal symmetric configurations.



(d) Centrosymmetric configurations.

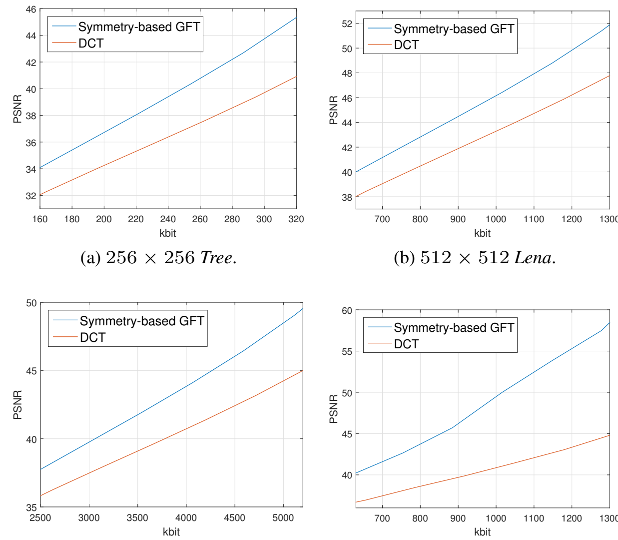
Fig. 2: Set of investigated graphs. Blue indicates edges with weight set to 0.01, red denotes edges with weight set to 1. Green dashed segments and green circles identify the reflective symmetries and centrosymmetries, respectively.

for centrosymmetries nodes are linked so that graphs are invariant under the point reflection. The associated weights are set to 1 (red edges in Fig. 2). For example, in the first graph of Fig. 2a the symmetry axis is drawn as a green dashed line and it is located between the third and fourth row of the graph, therefore the resulting edges connect the node in position $(3, k)$ with the one in position $(4, k)$, for $k = 1, 2, 3, 4$. By contrast, in the first graph of Fig. 2d the center of symmetry is marked as a green circle and it is located between the third and fourth row and first and second column of the graph, and consequently the node in position $(3, 1)$ is connected with the one in position $(4, 2)$ and the node in position $(3, 2)$ is connected with that in position $(4, 1)$. Furthermore, in order to preserve full connectivity for each graph while selecting its specific symmetry, non participating nodes remain connected as in a 2-D grid configuration with the weight set to 0.01 (blue edges in Fig. 2). Indeed, since the eigenvalue $\lambda = 0$ of $\mathbf{L}$ has multiplicity equal to the number of connected components of a graph, and since such an eigenvalue is associated to a constant eigenvector w.r.t. each set of connected nodes, in the case of fully connected graphs the projection of the signal on the associated eigenvector will correspond to its mean. Note that the



(a) 256 × 256 *Tree*. (b) 512 × 512 *Lena*.
(c) 1024 × 1024 *Man*. (d) 512 × 512 *Texture mosaic #3*.

Fig. 3: PSNR vs number of coefficients performance between the DCT and our proposed transforms.



(a) 256 × 256 *Tree*. (b) 512 × 512 *Lena*.
(c) 1024 × 1024 *Man*. (d) 512 × 512 *Texture mosaic #3*.

Fig. 4: PSNR vs variable rate performance between the DCT and the proposed transforms.

	Tree	*Lena*	*Man*	*Texture*
Delta PSNR	2.7	2.6	2.8	6.5
Delta rate	-22%	-20%	-23%	-43%

Table 1: *BD* rate associated to Fig. 4.

ratio between weights identifying a symmetric correspondence and weights used to preserve a 2-D grid topology has been empirically set to 100. In particular, considering a smaller value does not allow to properly characterize the symmetry in signals, whereas a larger value do not basically change the structure of the eigenvectors of the associated Laplacian matrix. As a result the defined set includes graphs characterized by the following symmetries:

- five up-down (UD) symmetries (Fig. 2a);
- five left-right (LR) symmetries (Fig. 2b);
- three diagonal symmetries and three anti-diagonal symmetries (Fig. 2c);
- twenty-five centrosymmetries (Fig. 2d).

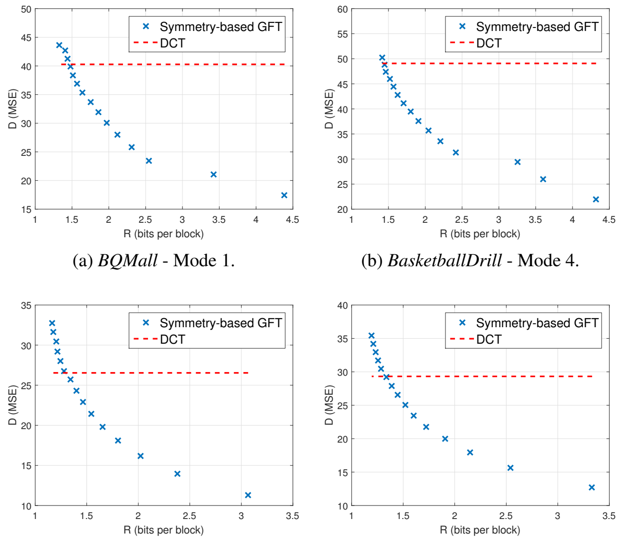
3.1. Representativeness of the symmetry graphs

The symmetry-based GFTs obtained from the graphs in Fig. 2 are tested in order to evaluate their energy compaction ability. The experiments have been carried out on a variety of standard images obtained from USC-SIPI Image Database [15]. We use a "brute-force" approach where for each 4×4 block of the image all the 41 GFTs are tested. After applying a non-linear approximation where the K smallest coefficients (in absolute value) are zeroed, the corresponding 4×4 reconstructed blocks are calculated through the inverse GFTs. A winner is set according to the GFT leading to the smallest mean squared error (MSE). Finally, the peak signal-to-noise ratio (PSNR) is computed comparing the original and the reconstructed image to determine the quality of reconstruction. In Fig. 3 the comparison with the DCT[1] is depicted for images of different size. The results indicate that the set of symmetry-based GFTs clearly outperforms the DCT independently of the type of image (natural or texture) or its size (256×256, 512×512 or 1024×1024), but this does not take into account signaling overhead. In fact, in a coding scenario overhead-information is needed to index the optimal GFT for every block. To introduce this overhead cost, once a different number of transform coefficients is zeroed 8 bits per coefficient are attributed to the surviving coefficients for both the DCT and the GFT; furthermore $\lceil \log_2 41 \rceil = 6$ overhead-bits are added to each image block to obtain final rate of the GFT. Fig. 4 shows the performance between the graph transforms and the DCT applied to the above mentioned images. Even in this context, the set of symmetry-based GFTs continues to exhibit performance significantly better than the DCT as additionally reported in Tab. 1 through the Bjøntegaard's metric (*BD* rate). It allows to compute the average gain in PSNR or the average per cent saving in bit rate between two rate-distortion curves. In the first case a negative (positive) value indicates a decrease (increase) of PSNR for the same bit rate, e.g., in *Lena* the set of symmetry-based GFTs has an average gain of 2.6 dB w.r.t. the DCT. In the second case a negative (positive) value identifies a decrease (increase) of bit rate for the same PSNR, e.g., in *Lena* the set of symmetry-based GFTs requires -20% of bit rate on average w.r.t. the DCT.

4. EXPERIMENTAL RESULTS IN VIDEO CODING

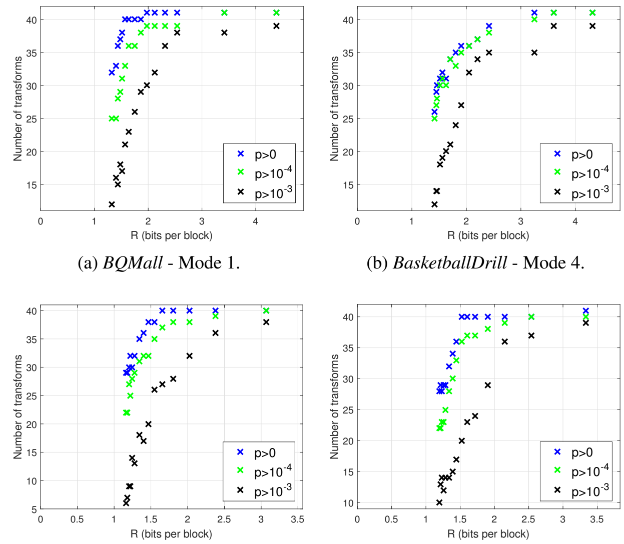
In the previous section we showed the good approximation ability for the proposed set of GFTs. Advantageous performance can still be obtained when an extra-bit cost is added to signal the optimal transform needed for each block. However, removing or decreasing such overhead information would be desirable to reduce both the rate and the complexity of a real system. In this section, the set

[1] For the rest of the paper the acronym DCT is used instead of 2D-DCT.



(a) *BQMall* - Mode 1.
(b) *BasketballDrill* - Mode 4.
(c) *Mobcal* - Mode 9.
(d) *Shields* - Mode 11.

Fig. 5: *R-D* performance for single modes.



(a) *BQMall* - Mode 1.
(b) *BasketballDrill* - Mode 4.
(c) *Mobcal* - Mode 9.
(d) *Shields* - Mode 11.

Fig. 6: Number of transforms used for single modes.

of transforms is applied on 4×4 residual blocks taken by four test video signals: *BQMall*, *BasketballDrill*, *Mobcal* and *Shields*. The residual blocks have been obtained during the intrapicture prediction phase in H.265 standard. The block database is divided in 35 classes, one for each directional mode adopted in H.265, where each class contains the blocks predicted by the corresponding directional orientation. The purpose is to provide an analysis aimed at identifying a connection between directional modes and optimal symmetry-based GFTs when applied on blocks derived by a specific prediction mode. In particular, we ask if and in which terms the size of set of symmetry-based transforms can be reduced based on the directional mode used to generate the residual. The procedure is similar to the method described beforehand, and it is intended to be applied to each class separately. The 4×4 residual block is transformed through the entire set of the symmetry-based GFTs. The signal is reconstructed keeping the 25% of the transform coefficients, i.e., 4 coefficients, and the corresponding MSE is calculated. The most efficient transform based on the smallest associated MSE is selected. At this point of the process the consequent distribution probability of the transforms returns the best scenario in terms of distortion *D*, i.e., the mean of all the MSE values is as small as possible, but the worst one in terms of rate *R* corresponding to the overhead-bits to signal the chosen transform, i.e., it maximizes the number of employed transforms. To minimize *D* for a given *R* all the blocks are newly processed by considering a new cost function, specifically a Lagrange multiplier method formulation is adopted as follows:

$$\begin{gathered} min\{J\} \\ J = D + \gamma R \end{gathered} \tag{2}$$

where *J* denotes the Lagrangian cost function and γ is the so-called Lagrange multiplier. For a fixed γ value the best transform is now selected not only based on distortion (MSE) but on rate calculated with the Huffman coding associated to the probability distribution of using a particular graph configuration. Once all the blocks have been processed, the probability distribution is updated and the procedure is iterated until the convergence is reached returning a rate-distortion point. Consequently an *R-D* curve is produced for different values of γ. Fig. 5 shows the result for the four test video signals: these curves negligibly change for different modes, for this reason only one mode is illustrated for each video. For all video signals, independently of the mode, the performance of our proposed set of transforms significantly outperforms the DCT also for rate values considerably lower than the 6 bits per block overhead cost considered in the preliminary compression scenario described in the previous section. It means that for each mode a reduced subset of transforms is sufficient to achieve notable *R-D* performance. As a matter of fact, in Fig. 6 the number of transforms actually used in relation to each single mode is shown. The blue crosses list the set of transforms used for at least one block, while the green and the black ones consider only the transforms with probability greater than 10^{-4} and 10^{-3} respectively. The results prove that several symmetry-based GFTs can be discarded from the set of transforms when applied on data corresponding to residuals predicted by a specific directional mode. This analysis allows to conclude that a significant overhead information decrease is feasible thanks to the knowledge of the prediction mode from which the residual is generated.

5. CONCLUSIONS

In this paper we have explored various types of GFTs characterized by totally and partially symmetric grids. We have compared this set of transforms with the 2D-DCT on a wide dataset including natural and texture images and residual signals. The experimental tests show that the symmetry-based GFTs outperform the 2D-DCT in approximation ability and even in a preliminary compression scenario. Furthermore the relation between the symmetry-based GFTs and the directional modes employed in H.265 standard has been examined and verified, showing that the cardinality of the set of effective transforms can be reduced. Even if the results presented here are preliminary, the proposed set of symmetry-based GFTs has shown significant performance increase in terms of energy compaction leading to a high efficiency coding.

6. REFERENCES

[1] V. K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, Sep 2001.

[2] C. Maccone, "A simple introduction to the klt (karhunen—loève transform)," in *Deep Space Flight and Communications: Exploiting the Sun as a Gravitational Lens*, pp. 151–179. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[3] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, vol. C-23, no. 1, pp. 90–93, Jan 1974.

[4] Shuyuan Zhu, Siu-Kei Au Yeung, and Bing Zeng, "In search of "better-than-dct" unitary transforms for encoding of residual signals," vol. 17, pp. 961 – 964, 12 2010.

[5] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, May 2013.

[6] Leo J. Grady and Jonathan R. Polimeni, *Discrete Calculus: Applied Analysis on Graphs for Computational Science*, Springer Publishing Company, Incorporated, 1st edition, 2010.

[7] H. E. Egilmez, A. Said, Y. H. Chao, and A. Ortega, "Graph-based transforms for inter predicted video coding," in *2015 IEEE International Conference on Image Processing (ICIP)*, Sept 2015, pp. 3992–3996.

[8] W. Hu, G. Cheung, and A. Ortega, "Intra-prediction and generalized graph fourier transform for image coding," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1913–1917, Nov 2015.

[9] Wei Hu, Gene Cheung, Antonio Ortega, and Oscar C. Au, "Multiresolution graph fourier transform for compression of piecewise smooth images," *IEEE Trans. Image Processing*, vol. 24, no. 1, pp. 419–433, 2015.

[10] H. E. Egilmez, E. Pavez, and A. Ortega, "Graph learning from data under laplacian and structural constraints," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 6, pp. 825–841, Sept 2017.

[11] A. Gnutti, F. Guerrini, and R. Leonardi, "Representation of signals by local symmetry decomposition," in *Proc. of the 23rd European Signal Processing Conference (EUSIPCO '15)*, 2015, pp. 983–987.

[12] Fabrizio Guerrini, Alessandro Gnutti, and Riccardo Leonardi, "Even/odd decomposition made sparse: A fingerprint to hidden patterns," *Signal Processing*, vol. 141, pp. 273 – 286, 2017.

[13] K. S. Lu and A. Ortega, "Fast implementation for symmetric non-separable transforms based on grids," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017, pp. 4109–4113.

[14] K. S. Lu and A. Ortega, "Symmetric line graph transforms for inter predictive video coding," in *2016 Picture Coding Symposium (PCS)*, Dec 2016, pp. 1–5.

[15] "The usc-sipi image database," http://sipi.usc.edu/database/.