LAPPED TRANSFORMS: A GRAPH-BASED EXTENSION

Keng-Shih Lu and Antonio Ortega

Dept. of Electrical and Computer Engineering, University of Southern California, Los Angeles, USA

ABSTRACT

Lapped transforms are transform coding tools with basis functions that overlap across blocks in order to reduce blocking artifacts. In this work, we take the uniform line graph model interpretation of the discrete cosine transform (DCT) and extend it to lapped transforms. We first extend the conditions of perfect reconstruction and orthogonality to lapped transforms on graphs, where different transforms are allowed for different blocks. Then, with the focus on line graphs, we design a lapped graph Fourier transform (LGFT) that has these properties, with significantly reduced blocking artifact. Experimental results show that the proposed LGFT can achieve improved transform coding gain as compared to other existing transforms.

Index Terms— Lapped transform, graph Fourier transform, blocking artifact, transform coding.

1. INTRODUCTION

The discrete cosine transform (DCT) is extensively used for compression, partially because the DCT is a good approximation to the statistically optimal Karhunen-Loève transform (KLT) when the pixel data follows a high correlation model [1]. One drawback of block transforms, such as the DCT, is that pixels in different blocks are processed independently, so the correlation between pixels on the boundary of two adjacent blocks cannot be captured. This often leads to the so-called blocking artifact: an artificial discontinuity across the block boundary in the reconstructed signal. A well-known approach that can significantly reduce blocking artifacts is based on the design of lapped transforms [2, 3], where a transform is applied to overlapping image regions. Designs of lapped transforms include the lapped orthogonal transform (LOT) [4], and a pre- and post-filtering framework [5], which has been adopted in the Daala codec [6].

Recently, several graph Fourier transforms (GFT), also known as graph-based transforms (GBT), have been proposed for encoding particular types of pixel data. For example, the asymmetric discrete sine transform (ADST) [7] is a GFT associated to a line graph identical to that of the DCT, except for an added self-loop, and has been used for encoding intra-predicted residual blocks. In [8] several graph-based transforms have been designed for piecewise smooth images. Non-separable GFTs for 2D blocks can also be defined based on grids [9] in order to capture diagonal edge orientations. However, despite the convenience that graph-based transforms provide, similar to the block-based DCT, most existing graph-based transforms for pixel data are applied block-wise without overlap, and may lead to blocking artifacts. Thus, an LOT-like graph-based transform that can mitigate blocking artifacts would be of practical interest. To the best of our knowledge, lapped transforms have not been studied in the context of graph signal processing.

In this work, we extend the notion of lapped transform to graph signals by proposing lapped graph Fourier transforms (LGFTs): a family of lapped transforms that have different basis functions for different blocks so as to capture distinct local statistical properties for each block, while having perfect reconstruction and orthogonality properties. The conventional LOT can be seen to provide near optimal transform coding gain for signals with a *uniform line graph model*, while our proposed LGFT is a generalization of the LOT that considers *more general graph-based models*. The LGFT matrix for each block can be obtained from a Kron reduction [10] of the graph, followed by an eigen-decomposition. We particularly focus on line graphs with non-uniform weights (an example is shown in Fig. 1). Such graphs can model pixel data with a smoothness prior (e.g. with prior information of image discontinuity locations), and can be used for encoding piecewise smooth images [8]. Hu et. al. [8] design a *block transform* for data modeled by line graphs with weak edges, e.g., different transforms would be used for blocks $i+1$ and $i+2$ in Fig. 1. Instead, we develop lapped transforms, so that the output for block $i+2$ in Fig. 1 uses information from block $i+1$. Our experimental results show that LGFT can provide a better transform coding gain as compared to several existing transforms. Preliminary experiments on piecewise smooth images are also provided to demonstrate a potential application of this approach.

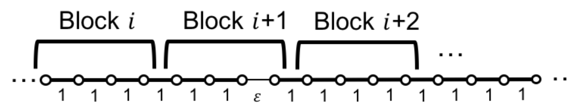


Fig. 1: An example of line graph with non-uniform edge weights.

2. PRELIMINARIES

2.1. Graph Signal Processing

We consider a weighted undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$, where $\mathcal{V}$ is the vertex set and $\mathcal{E}$ is the edge set of the graph. The entry $w_{i,j} \geq 0$ in the weight matrix $\mathbf{W}$ represents the weight of edge $(i,j) \in \mathcal{E}$, and $w_{i,j} = 0$ if $(i,j) \notin \mathcal{E}$. The graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ is the diagonal degree matrix with $d_{i,i} = \sum_{j=1}^{n} w_{i,j}$. Based on the definition of Laplacian matrix, for a given signal $\mathbf{x} \in \mathbb{R}^n$, the *Laplacian quadratic form*

$$\mathbf{x}^\top \mathbf{L} \mathbf{x} = \sum_{(i,j)\in\mathcal{E}} w_{i,j}(x_i - x_j)^2 \tag{1}$$

measures the variation of $\mathbf{x}$ on the graph.

The graph Fourier transform (GFT) is an important tool in graph signal processing [11, 12, 13]. The transform matrix $\mathbf{U}$ is defined as the matrix of eigenvectors of the graph Laplacian: $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$. Based on this definition, the GFT basis functions $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are mutually orthogonal unit norm vectors that correspond to the smallest to the largest variations on the graph:

$$\mathbf{u}_1 = \underset{\|\mathbf{f}\|=1}{\operatorname{argmin}} \quad \mathbf{f}^\top \mathbf{L} \mathbf{f}, \quad \mathbf{u}_k = \underset{\mathbf{f}\perp\mathbf{u}_1,\ldots,\mathbf{u}_{k-1},\|\mathbf{f}\|=1}{\operatorname{argmin}} \quad \mathbf{f}^\top \mathbf{L} \mathbf{f}.$$

In particular, when random signals $\mathbf{x}$ are modeled by a Gaussian Markov random field (GMRF) $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{L}^{\dagger})$ [14] (a Gaussian random vector whose inverse covariance is in the form of a graph Laplacian), then the GFT optimally decorrelates those signals. As an example, the DCT is the GFT of a line graph with uniform edge weights [15], meaning that it is the optimal transform for signals following a uniform line graph model.

2.2. Lapped Orthogonal Transforms

The lapped orthogonal transform (LOT) is a lapped transform that has orthogonal basis functions, and is applied to a length $2M$ region (two blocks of length M). Given the block size M, the LOT matrix is a $2M \times M$ matrix $\mathbf{R} = (\mathbf{E}^{\top},\ \mathbf{F}^{\top})^{\top}$. If we denote $\mathbf{x}$ as the input signal and $\mathbf{y}$ as the transform domain vector, then we can define a transform matrix $\mathbf{T}$ such that $\mathbf{y} = \mathbf{T}^{\top}\mathbf{x}$, and the reconstructed signal $\hat{\mathbf{x}} = \mathbf{T}\mathbf{y}$, where

$$\mathbf{T} = \begin{pmatrix} \ddots & & & \\ & \mathbf{R} & & \\ & & \mathbf{R} & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} \ddots & & & \\ \ddots & \mathbf{E} & & \\ & \mathbf{F} & \mathbf{E} & \\ & & \mathbf{F} & \ddots \\ & & & \ddots \end{pmatrix}.$$

The transform $\mathbf{R}$ is a valid LOT if and only if $\mathbf{T}\mathbf{T}^{\top} = \mathbf{T}^{\top}\mathbf{T} = \mathbf{I}$, meaning that $\hat{\mathbf{x}} = \mathbf{x}$ and columns of $\mathbf{T}$ are orthogonal. The general solution has the form [16]:

$$\mathbf{E} = \mathbf{P}\mathbf{Q}, \quad \mathbf{F} = (\mathbf{I} - \mathbf{P})\mathbf{Q}, \tag{2}$$

where $\mathbf{P}$ can be any symmetric projection matrix, and $\mathbf{Q}$ can be any orthogonal matrix, both with size $M \times M$.[1]

Many lapped transform designs use the DCT as a key component. One example of such a design is:

$$\begin{aligned} \hat{\mathbf{E}} &= \frac{1}{2}\left(\mathbf{U}_e - \mathbf{U}_o,\ (\mathbf{U}_e - \mathbf{U}_o)\mathbf{Z}\right), \\ \hat{\mathbf{F}} &= \frac{1}{2}\left(\mathbf{J}(\mathbf{U}_e - \mathbf{U}_o),\ -\mathbf{J}(\mathbf{U}_e - \mathbf{U}_o)\mathbf{Z}\right), \end{aligned} \tag{3}$$

where $\mathbf{U}_e$ and $\mathbf{U}_o$ are $M \times M/2$ matrices whose columns are the length-M DCT functions even and odd symmetry, respectively. The matrix $\mathbf{Z}$ is a cascade of plane rotations [4] or a product of DST-IV and DCT-II [2], and $\mathbf{J}$ is the order reversal permutation matrix:

$$\mathbf{J} = \begin{pmatrix} & & & 1 \\ & & 1 & \\ & \iddots & & \\ 1 & & & \end{pmatrix}.$$

Based on the fact that DCT approximates the KLT, the design in (3) approximates the optimal solution characterized in [3]. The projection and orthogonal matrices corresponding to (3) are

$$\hat{\mathbf{P}} = \frac{1}{2}(\mathbf{I} - \mathbf{U}_e\mathbf{U}_o^{\top} - \mathbf{U}_o\mathbf{U}_e^{\top}), \quad \hat{\mathbf{Q}} = (\mathbf{U}_e,\ -\mathbf{U}_o\mathbf{Z}). \tag{4}$$

3. LAPPED TRANSFORMS ON GRAPHS

We now propose the lapped graph Fourier transform (LGFT) by first investigating the conditions for perfect reconstruction and orthogonality, and then incorporating the graph variation (1) into the optimality criterion.

[1] In the literature, the general solution is sometimes represented as $\mathbf{E} = \mathbf{Q}\mathbf{P}$, $\mathbf{F} = \mathbf{Q}(\mathbf{I} - \mathbf{P})$. In fact, this form and (2) are interchangeable, and we focus on (2) here as it can be extended to the graph-based design more easily.

3.1. Conditions of Perfect Reconstruction and Orthogonality

We denote a graph signal $\mathbf{x} \in \mathbb{R}^{NM}$ as $\mathbf{x} = (\mathbf{x}_1^{\top}, \dots, \mathbf{x}_N^{\top})^{\top}$, modeled by a GMRF:

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \mathbf{L}^{\dagger} = \mathbf{C} = \begin{pmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & \cdots & \mathbf{C}_{1,N} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} & \cdots & \mathbf{C}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}_{N,1} & \mathbf{C}_{N,2} & \cdots & \mathbf{C}_{N,N} \end{pmatrix}\right),$$

where $(\mathbf{x}_k^{\top}, \mathbf{x}_{k+1}^{\top})^{\top}$ corresponds to the k-th and $(k+1)$-th blocks, and $\mathbf{C}_{p,q}$ is the (p,q)-th $M \times M$ block component of the covariance matrix $\mathbf{C}$. Unlike the conventional LOTs, where a common model is used for all blocks, here we consider different models for different blocks, such as a line graph model with non-uniform weights, where $\mathbf{C}_{k,k}$ are different for different k. Under this assumption, we revisit the conditions of perfect reconstruction and orthogonality. First, we define a lapped transform matrix $\mathbf{R}_k = (\mathbf{E}_k^{\top}, \mathbf{F}_k^{\top})^{\top}$ for the k-th and $(k+1)$-th blocks $(\mathbf{x}_k^{\top}, \mathbf{x}_{k+1}^{\top})^{\top}$, where $\mathbf{R}_k$'s are different in order to capture distinct statistical properties for different blocks. Based on this definition, the overall transform matrix $\mathbf{T}_{\text{LGFT}}$ is

$$\mathbf{T}_{\text{LGFT}} = \begin{pmatrix} \ddots & & & \\ & \mathbf{R}_k & & \\ & & \mathbf{R}_{k+1} & \\ & & & \ddots \end{pmatrix} = \begin{pmatrix} \ddots & & & \\ & \mathbf{E}_k & & \\ & \mathbf{F}_k & \mathbf{E}_{k+1} & \\ & & \mathbf{F}_{k+1} & \\ & & & \ddots \end{pmatrix}.$$

Then, we obtain the output signal $\mathbf{y} = (\mathbf{y}_1^{\top}, \dots, \mathbf{y}_N^{\top})^{\top}$ and the reconstructed signal $\hat{\mathbf{x}} = (\hat{\mathbf{x}}_1^{\top}, \dots, \hat{\mathbf{x}}_N^{\top})^{\top}$ with

$$\begin{aligned} \mathbf{y}_k &= \mathbf{E}_k^{\top}\mathbf{x}_k + \mathbf{F}_k^{\top}\mathbf{x}_{k+1}, \\ \hat{\mathbf{x}}_k &= \mathbf{F}_{k-1}\mathbf{y}_{k-1} + \mathbf{E}_k\mathbf{y}_k \\ &= \left(\mathbf{E}_k\mathbf{E}_k^{\top} + \mathbf{F}_{k-1}\mathbf{F}_{k-1}^{\top}\right)\mathbf{x}_k + \mathbf{F}_{k-1}\mathbf{E}_{k-1}^{\top}\mathbf{x}_{k-1} + \mathbf{E}_k\mathbf{F}_k^{\top}\mathbf{x}_{k+1}. \end{aligned}$$

By comparing $\mathbf{x}_k$ and $\hat{\mathbf{x}}_k$, we obtain the conditions for perfect reconstruction and aliasing cancellation. In addition, the orthogonality constraint, $\mathbf{R}_k^{\top}\mathbf{R}_k = \mathbf{I}$, is equivalent to $\mathbf{E}_k^{\top}\mathbf{E}_k + \mathbf{F}_k^{\top}\mathbf{F}_k = \mathbf{I}$. Thus, for each k, the desired transform would satisfy

$$\mathbf{E}_k\mathbf{E}_k^{\top} + \mathbf{F}_{k-1}\mathbf{F}_{k-1}^{\top} = \mathbf{I}, \qquad \text{(Perfect reconstruction)} \tag{5}$$

$$\mathbf{E}_k\mathbf{F}_k^{\top} = \mathbf{F}_k\mathbf{E}_k^{\top} = \mathbf{0}, \qquad \text{(Aliasing cancellation)} \tag{6}$$

$$\mathbf{E}_k^{\top}\mathbf{E}_k + \mathbf{F}_k^{\top}\mathbf{F}_k = \mathbf{I}. \qquad \text{(Orthogonality)} \tag{7}$$

Due to the highly nonlinear nature of these constraints as well as the large number of degrees of freedom in designing $\mathbf{E}_k$ and $\mathbf{F}_k$, we propose a LGFT construction that generalizes the LOT solution of (2). We select

$$\mathbf{E}_k = \mathbf{P}\mathbf{Q}_k, \quad \mathbf{F}_k = (\mathbf{I} - \mathbf{P})\mathbf{Q}_k, \tag{8}$$

where the projection matrix $\mathbf{P}$ is common for all k so that (5) can be satisfied. Based on (8), one can verify that (5)-(7) are always satisfied as long as $\mathbf{P}$ is a symmetric projection matrix and $\mathbf{Q}_k$ are orthogonal matrices. Thus, (8) is a sufficient condition of perfect reconstruction and orthogonality for LGFT.

3.2. Proposed LGFT Construction

An optimality criterion for the conventional LOT based on the transform coding gain can be defined as [2, 17]:

$$G_{TC} = \frac{\frac{1}{M}\sum_{i=1}^{M}\sigma_{k,i}^2}{\left(\prod_{i=1}^{M}\sigma_{k,i}^2\right)^{1/M}}, \tag{9}$$

where $\sigma_{k,i}^2 = \mathrm{Var}(\mathbf{y}_k(i))$ is the variance of the i-th transform coefficient in block k. With $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$, $\sigma_{k,i}^2$ is the (i,i) entry of

$$\mathbb{E}[\mathbf{y}_k \mathbf{y}_k^\top] = \begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix}^\top \begin{pmatrix} \mathbf{C}_{k,k} & \mathbf{C}_{k,k+1} \\ \mathbf{C}_{k+1,k} & \mathbf{C}_{k+1,k+1} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{F} \end{pmatrix} \tag{10}$$

$$= \mathbf{Q}^\top \underbrace{\begin{pmatrix} \mathbf{P} \\ \mathbf{I}-\mathbf{P} \end{pmatrix}^\top \begin{pmatrix} \mathbf{C}_{k,k} & \mathbf{C}_{k,k+1} \\ \mathbf{C}_{k+1,k} & \mathbf{C}_{k+1,k+1} \end{pmatrix} \begin{pmatrix} \mathbf{P} \\ \mathbf{I}-\mathbf{P} \end{pmatrix}}_{\mathbf{G}_k} \mathbf{Q}. \tag{11}$$

It has been shown [2] that, for a fixed $\mathbf{P}$, the optimal $\mathbf{Q}$ that maximizes G_{TC} is the eigenmatrix of $\mathbf{G}_k$. In the data model considered in conventional LOT, $\mathbf{G}_k = \mathbf{G}$ is common for all k. In (2), $\mathbf{Z}$ is typically chosen as an orthogonal transform with fast implementation such that $\hat{\mathbf{Q}}$ approximates the eigenmatrix of $\mathbf{G}$.

Here, we extend the LOT design problem to LGFT design as follows. Let the signal $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{C} = \mathbf{L}^\dagger)$ be a GMRF, where $\mathbf{L}$ is a graph Laplacian corresponding to graph $\mathcal{G}$ and $\mathbf{C}_{k,k}$ can be different for different k. Note that $(\mathbf{x}_k^\top, \mathbf{x}_{k+1}^\top)^\top$ is also a GMRF with length $2M$:

$$\begin{pmatrix} \mathbf{x}_k \\ \mathbf{x}_{k+1} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \overline{\mathbf{L}}_{\mathcal{S}_k}^\dagger\right), \quad \overline{\mathbf{L}}_{\mathcal{S}_k} := \begin{pmatrix} \mathbf{C}_{k,k} & \mathbf{C}_{k,k+1} \\ \mathbf{C}_{k+1,k} & \mathbf{C}_{k+1,k+1} \end{pmatrix}^\dagger.$$

The matrix $\overline{\mathbf{L}}_{\mathcal{S}_k}$ is a Laplacian obtained by Kron reduction [10] of the graph nodes $\mathcal{V}_{\mathcal{S}_k} = \{(k-1)M+1, \dots, (k+1)M\}$ corresponding to blocks k and $k+1$. In general, the Kron reduction of a subset of graph nodes $\mathcal{V}_{\mathcal{S}} \in \mathcal{V}$ can be expressed as

$$\overline{\mathbf{L}}_{\mathcal{S}} = \mathbf{L}_{\mathcal{S},\mathcal{S}'} \mathbf{L}_{\mathcal{S}',\mathcal{S}'}^{-1} \mathbf{L}_{\mathcal{S}',\mathcal{S}}, \quad \mathcal{V}_{\mathcal{S}'} = \mathcal{V} \backslash \mathcal{V}_{\mathcal{S}}.$$

As in (10), we consider (8) and replace the block components of $\mathbf{C}$ by $\overline{\mathbf{L}}_{\mathcal{S}_k}$. With the assumption that $\overline{\mathbf{L}}_{\mathcal{S}_k}$ can be different for different k, we would like to choose $\mathbf{Q}_k$ that diagonalizes

$$\mathbf{H}_k := \begin{pmatrix} \mathbf{P} \\ \mathbf{I}-\mathbf{P} \end{pmatrix}^\top \overline{\mathbf{L}}_{\mathcal{S}_k} \begin{pmatrix} \mathbf{P} \\ \mathbf{I}-\mathbf{P} \end{pmatrix}. \tag{12}$$

In particular, if $\mathbf{L}$ corresponds to a line graph, then the Kron reduction $\overline{\mathbf{L}}_{\mathcal{S}_k}$ is the Laplacian of the line graph segment of $\mathcal{G}$ with nodes $\mathcal{V}_{\mathcal{S}_k}$. Thus, the LGFT matrix $\mathbf{R}_k$ of a line graph with $n = NM$ can be designed based on the k-th line graph segment with length $2M$.

While the choice of $\mathbf{Q}_k$ is straightforward with a given $\mathbf{P}$, finding $\mathbf{P}$ for global optimal solution in term of transform coding gain, even in conventional LOT design, is a challenging problem [4]. Following the LOT design, we adopt the matrix $\hat{\mathbf{P}}$ in (4) associated to (3). With this choice of $\mathbf{P}$, *this LGFT design can be regarded as a generalization of the LOT*, where $\mathbf{R}_k$ reduces to LOT when $\mathbf{Q}_k = (\mathbf{U}_e, -\mathbf{U}_o\mathbf{Z})$.

To summarize, given a line graph $\mathcal{G}$ with Laplacian $\mathbf{L}$ and block size M, the proposed LGFT can be constructed as follows:

1. Pick $\mathbf{P}$ as in (4).
2. Pick $\mathbf{Q}_k$ for each k as the eigenmatrix of the $\mathbf{H}_k$ in (12).
3. Obtain LGFT components as given in (8).

Similar to the LOT that achieves nearly optimal G_{TC} when $\mathbf{L}$ is associated to a uniform line graph, the proposed LGFT construction, based on a fixed $\mathbf{P} = \hat{\mathbf{P}}$, can approximately optimize G_{TC} when the line graph has non-uniform weights.

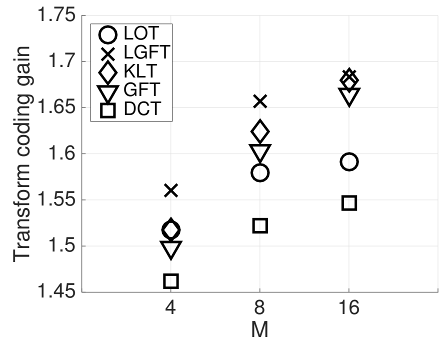


Fig. 2: Transform coding gains for different transforms and block sizes M with signals modeled by a particular line graph. The KLT, GFT, and DCT shown here are defined in a block-based manner.

4. EXPERIMENTAL RESULTS

We evaluate the proposed LGFT on a class of line graphs, where some edges have weights ε that characterize weak correlations, and other edges have weights 1. This line graph model has been used for encoding intra-predicted images [18] and piecewise smooth images [8, 19].

4.1. Transform Coding Gain with a Particular Line Graph

First, we consider a line graph with length $n = 400$ and $\varepsilon = 0.05$, where weak weights are located at edges $(h, h+1)$ with $h \in \{15, 30, 45, 60, \dots, 390\}$. A covariance matrix $\mathbf{C} = (\mathbf{L} + 0.2\mathbf{I})^{-1}$ is used in this experiment to avoid the matrix singularity issue. We compare the transform coding gains with LOT, LGFT, DCT, KLT, and GFT, where the three latter transforms are designed and applied in a non-overlapping block-based manner. The implementation of LOT follows from (3), where $\mathbf{Z}$ is composed of a DCT-II and a DST-IV, as suggested in [3]. For each transform, block sizes of $M = 4$, 8, and 16 are considered.

In Fig. 2 we show the transform size versus the transform coding gain G_{TC}. For this model $\mathcal{N}(\mathbf{0}, \mathbf{C})$, G_{TC} is upper-bounded by 1.736, derived from the length-n KLT of the overall model. We can see that the LGFT yields the highest gain for all block sizes included in this experiment. For some block sizes larger than $M = 16$, the transform coding gains with the block-based KLT and GFT are higher than that of the LGFT since there are fewer block boundaries. In practical coding scenarios, additional information such as the positions of weak edges may be required as bit rate overhead. This has not been considered in the analysis here.

We show the LGFT basis functions $\mathbf{R}_2$ with $M = 8$ in Fig. 3. Note that $\mathbf{R}_2$ is designed for the blocks $(\mathbf{x}_2^\top,\ \mathbf{x}_3^\top)^\top$ with nodes 9 to 24. The associated model to these blocks is the line graph segment with length 16 as shown in Fig. 1, where a weak edge $(15, 16)$ lies between the 7th and 8th nodes. We can observe in most basis functions a discontinuity between entries 7 and 8 that corresponds to the lower correlation, showing that the LGFT basis can capture the local weak correlation in the graph topology through different choices of $\mathbf{Q}_k$.

4.2. LGFT for Image Coding

In the second experiment, we apply an LGFT to image coding. A 480×640 piecewise smooth image from the `Tsukuba` dataset [20] is used for this experiment. For demonstration and comparison purpose, we apply different transforms (LGFT, LOT, DCT, and

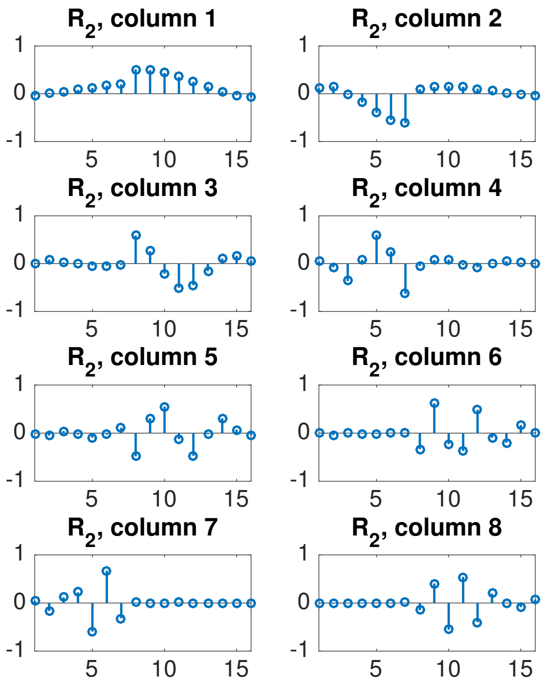


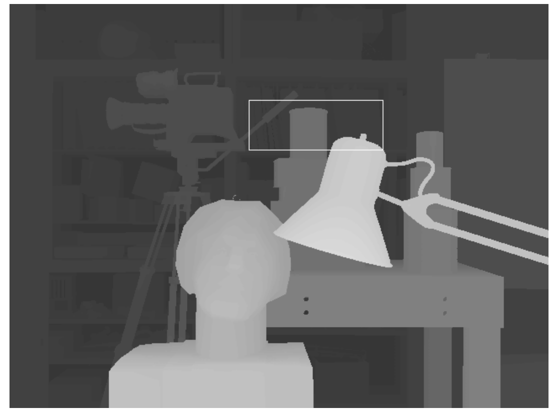
Fig. 3: Basis functions of the LGFT for $k = 2$ with $M = 8$.

QP	PSNR (dB)			
	LGFT	LOT	GFT	DCT
25	48.24	48.12	48.63	48.38
30	44.84	44.74	45.11	44.93
35	41.94	41.83	42.11	41.91
40	37.74	37.73	37.69	37.62
45	33.15	33.15	33.09	33.06

Table 1: Quality comparison of different horizontal transforms on full test image.

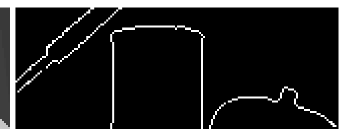
GFT) horizontally, then a common transform (block-based DCT) vertically. We quantize the coefficients and apply inverse horizontal transforms and vertical block-based DCT to reconstruct the signal. To define the line graphs for LGFT and GFT, we apply a Sobel edge detector [21] to obtain an edge map. For each image row we define a line graph based on image discontinuity information in the Sobel edge map: an edge weight of the line graph is $\varepsilon = 0.7$ if any of the two corresponding pixels is an edge point in the Sobel map. We use a block size $M = 8$, and quantization parameters (QP) ranging from 25 to 45, corresponding to quantization factors $2^{(\mathrm{QP}-4)/6}$. Note that, in practical image coding scheme, the edge location can be transmitted as side information to the decoder for unique decodability. In this experiment, we only compare the distortion but not the number of bits required for encoding. This means that we can fairly compare LGFT with GFT, but not with DCT and LOT, which do not require side information.

Table 1 shows the peak-signal-to-noise ratio (PSNR) of different transforms with different QPs. We note that, similar to the GFT that provides a higher PSNR than the DCT, the LGFT gives a higher PSNR than LOT for almost all QPs. This gain results from the fact that the line graph model, which LGFT and GFT are based on, can better capture the discontinuities in the image signal. In Fig. 4 we show the original image, the Sobel edge map, and the reconstructed images using different transforms with QP=30. With this quantization level, while LGFT does not give higher PSNRs than GFT, it yields reduced blocking artifacts (reduced image discontinuities along horizontal direction) as compared to the GFT.

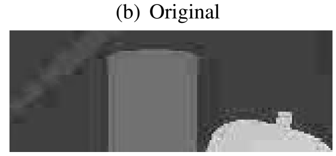


(a) Full Image

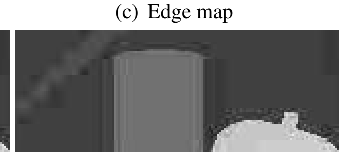
(b) Original



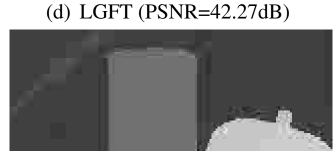
(c) Edge map



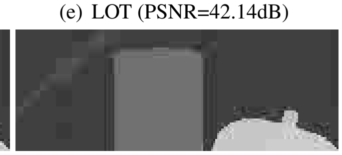
(d) LGFT (PSNR=42.27dB)



(e) LOT (PSNR=42.14dB)



(f) GFT (PSNR=42.57dB)



(g) DCT (PSNR=42.33dB)

Fig. 4: Subjective comparison different transforms with QP=30. (a) The full piecewise smooth image and a 160×60 patch. (b) Original image patch. (c) Sobel edge map. (d)-(f) Recovered image patches with different transforms.

5. CONCLUSION

In this work, we focus on the design of lapped transforms on line graphs with non-uniform weights. In particular, we derive the conditions for perfect reconstruction and orthogonality of the lapped graph Fourier transform (LGFT), which is more general than the lapped orthogonal transform (LOT). Then, we propose a design of LGFT on an arbitrary line graph, where different transform functions are applied to different blocks to adapt to different local statistical properties. Experimental results show that on a nonuniform line graph, the LGFT can achieve a better energy compaction than the block-based graph Fourier transform and a conventional LOT in terms of transform coding gain. The extensions to different lengths of overlap and to graphs with more general topologies, as well as fast implementations of the LGFT, will be explored in future work.

6. REFERENCES

[1] R. J. Clarke, *Transform Coding of Images*, Academic Press, Inc., Orlando, FL, USA, 1985.

[2] H. S. Malvar, *Signal Processing with Lapped Transforms*, Artech House communications library. Artech House, 1992.

[3] H. S. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 6, pp. 969–978, June 1990.

[4] H. S. Malvar and D. H. Staelin, "The LOT: transform coding without blocking effects," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 4, pp. 553–559, April 1989.

[5] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-filtering," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1557–1571, June 2003.

[6] J. Valin, T. B. Terriberry, N. E. Egge, T. Daede, Y. Cho, C. Montgomery, and M. Bebenita, "Daala: Building a next-generation video codec from unconventional technology," in *2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP)*, Sept 2016, pp. 1–6.

[7] J. Han, A. Saxena, V. Melkote, and K. Rose, "Jointly optimized spatial prediction and block transform for video and image coding," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1874–1884, Apr 2012.

[8] W. Hu, G. Cheung, A. Ortega, and O. C. Au, "Multiresolution graph Fourier transform for compression of piecewise smooth images," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 419–433, Jan 2015.

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

[10] F. Dörfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, Jan 2013.

[11] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[12] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013.

[13] A. Ortega, P. Frossard, J. Kovačević, JMF Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.

[14] H. Rue and L. Held, *Gaussian Markov random fields: theory and applications*, CRC Press, 2005.

[15] G. Strang, "The discrete cosine transform," *SIAM Rev.*, vol. 41, no. 1, pp. 135–147, 1999.

[16] G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley-Cambridge Press, 1997.

[17] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice Hall Professional Technical Reference, 1990.

[18] 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.

[19] D. Zhang and J. Liang, "Graph-based transform for 2d piecewise smooth signals with random discontinuity locations," *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1679–1693, April 2017.

[20] M. Peris, S. Martull, A. Maki, Y. Ohkawa, and K. Fukui, "Towards a simulation driven stereo vision system," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, Nov 2012, pp. 1038–1042.

[21] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.