International Journal of Advances in Electronics Engineering– IJAEE
Volume 6 : Issue 1 *[ISSN : 2278-215X]*
Publication Date: 18 April, 2016

An entropy-based signal restoration method on graphs

Xianwei Zheng, Yuan Yan Tang, Jianjia Pan

Abstract—**This paper proposes an entropy-based best tree decomposition and reconstruction method for signals on graphs. A recently popular method for graph signal decomposition is to down-sample and filter a graph signal by low-pass and high-pass filters, then iterate decomposition on every low-pass component. This method is not well-suited for graph signals with significant high frequency components. We propose a new method to decompose a graph signal according to an entropy based best tree decomposition scheme. This method is adaptive for graph signal decomposition, and it gives more accurate and robust representations for graph signals. The proposed decomposition method is shown to provide an efficient representation well-suited for graph signal compression, or graph-based image compression. The performance of the proposed method is validated by real-world graph signal recovery problems. The new method also achieves better performance than the existing wavelet-like decomposition in graph-based image decomposition and reconstruction.**

Keywords— **Graph Fourier transform, graph signal filter, entropy, graph signal compression.**

I. Introduction

Many type of modern real-world data, for example from social or transportation network, are naturally reside on complex domains, e.g. graphs or networks. The complex underlying structures of these kinds of data disabled the application of the classical signal processing methods. New representation and analysis techniques are required for processing of signals on graphs, which lead to a new research field called **signal processing on graphs**. The core problem in signal processing on graphs is to develop signal processing theory similar to the classical signal processing.

Time-frequency analysis is a very important component in classical signal processing [1]. Analogously, developing vertex-frequency analysis theory on graphs is one of the main topics in graph signal processing theory. Recently, spectral graph theory play a key role in developing a frequency analysis theory on graph signals. The foundation of the theory is to apply graph Laplacian matrix and its eigenvectors to establish the so called graph Fourier transform, which captures the spectral property of the given datasets [2].

Xianwei Zheng
Department of Computer and Information Science, University of Macau
Macau, China

Yuan Yan Tang
Department of Computer and Information Science, University of Macau
Macau, China

Jianjia Pan
Department of Computer and Information Science, University of Macau
Macau, China

In classical signal processing, wavelet transform is ubiquitous in time-frequency signal and image analysis. The significant advantage of the wavelet transform is that it gives time-frequency representations, and present local characteristics of given signals. One of the main topics in graph signal processing is to develop wavelet transform or wavelet-like transform on graphs. Since graphs are irregular domain, local transforms on graphs similar to wavelet transform can be defined in different ways, to name a few, spectral wavelets [3], CKWT [4], Haar-like wavelets [5], diffusion wavelet [6], and separable filterbank wavelets [7].

Classical wavelet transform proceeds filtering and downsampling on each scale or resolution level to establish a multiscale analysis. To establish a wavelet-like transform on graphs, one has to find proper techniques for filtering, downsampling, and multiscale analysis on graphs. Graph filtering can be defined by using graph Fourier transform, similar to classical frequency filtering. Design of downsampling operation on graphs requires more complex techniques.

In [8], Narang and Ortega proposed a downsampling then upsampling method on bipartite graph signals, and then extend it to general graphs by using Harary's algorithm and graph coloring. The corresponding two channel filter banks designed in [8], which are called graph quadrature mirror filterbanks (graph-QMFs), are critically sampled, perfect reconstruction and nearly orthogonal. But the graph-QMFs is in general not compact supported, compact supported filterbanks can directly control the trade-off between localization in the vertex domain and the spectral domain. In [9], graph-QMFs are relaxed by giving up the condition of orthogonality to design a biorthogonal pair of compact supported graph filter banks. Tanaka and Sakiyama extended biorthogonal graph filter banks to the case of M-channel oversampled graph filter banks in [10, 11]. The filter bank design mentioned above lead to a graph decomposition similar to one level classical wavelet transform, but the downsampled graph signal do not inherent any graph structure from the original graph. In [12], Do and Nguyen designed a downsampling method for graph signal by using maximum spanning trees (MST). A graph multiresolution is naturally generated by filtering and downsampling on the maximum spanning tree of a given graph signal.

However, both the graph coloring based and MST-based methods iterating on down-sampling and filtering on the lower-pass components, and such a scheme is not adaptive to different types of graphs signals, especially for signals with significant high frequency components. Then iterating decomposition on the high-frequency components are required for some applications. We first develop a full subband decomposition, i.e., decompose both the high-pass and low-pass components, for image decomposition, but it performs poorer when more decomposition level is

SEEK DIGITAL LIBRARY

implemented. In order to develop an adaptive downsampling method on graph signals or graph based image processing, we propose an entropy based best tree decomposition scheme in this paper. We will propose two types of entropy based best tree decomposition scheme, i.e., time-varying signal entropy best tree and graph signal entropy best tree decomposition. For graph based image processing, an image is considered as a graph signal on an 8-nearest neighbor connected and 4-colored graph, then we use the bipartite graph downsampling and filtering to decompose an image, and since images are time-varying signals, we will use a time entropy best tree to find the adaptive decomposition for an image. Then for graph signals, in order to generate a multi-scale analysis, we apply the MST based downsampling. We will also introduce the definitions of graph entropies for searching of the best tree decomposition.

II. Graph Fourier transform

In this section, we review the definition of graph Fourier transform. Consider a weighted graph $G=(V, E, W)$, where V denotes the set of vertices, E denotes the set of edges, of the graph, respectively. W is the weighted adjacency matrix. If there is an edge $e(i,j)$ connecting nodes i and j, $W_{i,j}$ represents the weight assign to the edge $e(i,j)$, otherwise, $W_{i,j}=0$. Define the degree matrix D associated to the graph G as a diagonal matrix whose i-th diagonal element d_i is equal to the sum of the weights of all the edges incident to vertex i, i.e. $D_{ii}=\sum_j W_{ij}$. Then the graph Laplacian, also called the combinatorial graph Laplacian of G is defined as $L=D-W$. Obviously, L is a real symmetric matrix, and therefore has a complete set of orthonormal basis. Denote these eigenvectors by u_l for $l=0,1,\cdots,N-1$, with associated eigenvalues λ_l, i.e. $Lu_l=\lambda_l u_l$. A normalized Laplacian is defined as $\mathrm{L}=I-D^{-\frac{1}{2}}WD^{-\frac{1}{2}}$. For a connected graph, the spectrum of L, $\sigma(\mathrm{L})\subseteq[0,2]$. In particular, $\max(\sigma(\mathrm{L}))=2$ if the graph is a bipartite graph. In this paper, we will apply the normalized Laplacian, for the applying the structure of the spectrum. The eigenvalue and eigenvectors are denoted by the same notations as mentioned above, and we denote the eigenvector matrix by $U=(u_0,u_1,\cdots,u_{N-1})$. The subspaces spanned by eigenvectors correspond to the same eigenvalue λ is denoted by V_λ, then the corresponding eigenspace projection matrix on V_λ is defined as $P_\lambda:=\sum_{\lambda_l=\lambda} u_l u_l^T$, where u_l^T is the transpose of u_l. Then for any vector $f\in R^N$ defined on the vertices of G, its graph Fourier transform $\hat{f}$ is defined by $\hat{f}(\lambda_l)=\langle f,u_l\rangle=\sum_{n=1}^{N} u_l^*(n)f(n)$. The inverse transform can be derived by: $f(n)=\sum_{l=0}^{N-1}\hat{f}(\lambda_l)u_l(n)$. The Parseval equation holds for the graph Fourier transform, that is, for any $f,g\in R^N$, $\langle f,g\rangle=\langle\hat{f},\hat{g}\rangle$.

Note that graph Fourier transform is consistent with the traditional Fourier transform if a finite discrete periodic signal is defined on a circle graph, since the eigenvector matrix of its Laplacian is the discrete Fourier transform matrix.

III. Graph signal filtering

In classical signal processing, filtering operation on a signal f is defined by computing convolution of f with a filter kernel h, and the output signal $f_{out}=f_{in}*h$ where $*$ is the convolution operator. Taking Fourier transform on both side, we have $\hat{f}_{out}=\hat{f}_{in}\cdot\hat{h}$. Then similarly, graph filtering is also defined via graph Fourier transform, but only on the spectrum domain, that is, $\hat{f}_{out}(\lambda_l)=\hat{f}_{in}(\lambda_l)\cdot\hat{h}(\lambda_l)$. Then taking inverse graph Fourier transform on both sides, we have $f_{out}(\lambda_l)=\sum_l \hat{f}_{in}(\lambda_l)\cdot\hat{h}(\lambda_l)u_l(\lambda_l)$. In a matrix form, the last equation can be written as $f_{out}=\hat{h}(\mathcal{L})f_{in}$, where

$$\hat{h}(\mathrm{L})=U\begin{pmatrix}\hat{h}(\lambda_0) & & & \mathbf{0}\\ & \hat{h}(\lambda_1) & & \\ & & \ddots & \\ \mathbf{0} & & & \hat{h}(\lambda_{N-1})\end{pmatrix}U^*.$$

The critically sampled two-channel filter banks on graphs consist of 4 filter banks: the low-pass and high pass analysis filters H_0,H_1 and the low-pass and high-pass synthesis filters G_0,G_1 [9].

In matrix form, the corresponding transform matrices can be written as, for $i=0,1$:

$$H_i=\hat{h}_i(\mathcal{L})=\sum_{\lambda_\sigma(G)}\hat{h}_i(\lambda)P_\lambda,$$

$$H_i=\hat{g}_i(\mathcal{L})=\sum_{\lambda_\sigma(G)}\hat{g}_i(\lambda)P_\lambda.$$

In order to achieve a critically sampling, that is, the number of analysis filters is equal to the downsampling factor, a downsampling operator on graphs of factor 2 is defined in [8]. Downsampling on graphs is proceeded by choosing signal samples on a subset of nodes $H\subset V$ and discarding the samples on the other nodes H^c. Then an upsampling operator is defined by inserting zeros into the downsampled component on the discarded nodes $H^c\subset V$. This downsampling then upsampling process can be expressed in the following relation: $f_{du}^H=\frac{1}{2}(1+\beta_H(n))f(n)$, where $\beta_H(n)=1$ if $n\in H$, and $\beta_H(n)=-1$ if $n\notin H$. Denote $J_H:=diag(\beta_H(n))$. Then in matrix form, we have $f_{du}^H=\frac{1}{2}(I+J_H)f$. Downsampling on graphs requires to separate the node set V into two subsets, for general graphs, this leads to a technique design of separation. For bipartite graphs, the structure of the graph naturally gives a separation, since by its definition, a bipartite graph is a graph whose nodes can be partition into two disjoint subsets L and H, such that each edge of G connects one node from L to one node from H. The

SEEK DIGITAL LIBRARY

spectrum of a bipartite graph has the following properties [8]:

a) Given a bipartite graph $B = (L, H, E)$ with the Laplacian matrix L and with the partition function β defined as above on L or H, then if λ is a eigenvalue of L with eigenvector u_λ, then $2-\lambda$ is also an eigenvalue of L with eigenvector $u_{2-\lambda} = J_H u_\lambda$.

and the perfect reconstruction condition:

b) Given a bipartite graph $G = (L, H, E)$, and the filter kernels as defined as above. For any $\lambda \in [0,2]$, the perfect reconstruction is equivalent to

$$\hat{h}_0(\lambda)\hat{g}_0(\lambda) + \hat{h}_1(\lambda)\hat{g}_1(\lambda) = 2,$$
$$\hat{h}_0(2-\lambda)\hat{g}_0(\lambda) - \hat{h}_1(2-\lambda)\hat{g}_1(\lambda) = 2.$$

IV. Graph entropy

Similar to classical wavelet transform, the wavelet-like decomposition for graph signals decompose a graph signal into low-frequency and high-frequency components and then iteratively decompose the low-frequency components, i.e., the approximated parts, see Figure 1 for a two-level decomposition of a graph signal and its reconstruction in the setting of two channel graph filter banks.

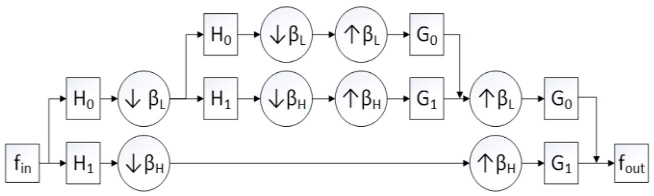


Figure 1: Two channel two level graph filter bank.

However, for graph signals with significant high-frequency components, decomposition on the low-frequency components is not effective for extracting local characteristics in the high-frequency components. We propose an adaptive decomposition for graph signals in this section by using a graph entropy measurement. For that, we will have to introduce a definition of graph entropy. Then using the entropy measurements to establish a best tree decomposition of a given signal. The Shannon entropy of coefficients of a time-varying signal x is defined as $H(x) = -\sum_i x_i^2 log(x_i^2)$, where $log(\cdot)$ is the natural logarithm. A log-energy entropy of signal x is defined as: $E(x) = \sum_i log(x_i^2)$. Now we define the entropy for a graph signal f on graph $G = (V, E, W)$. A graph signal entropy should captures the characteristics of both the signal and the graph it resided. For that, we define $q_i = \sum_j w_{i,j} / \sum_{i,j} w_{i,j}$, which denotes the share of the edge weights related to node i in all the edge weights. Then we define the shannon type graph signal entropy as follows:

$$H_S(f,G) = \sum_{i \in V} p(f_i) log(p(f_i)) \sum_{j \in V} q_j log(q_j),$$

where $p(f_i) = p(f = f_i)$, the probability of $f = f_i$. Then we have, $H_S(f,G) = E(log(p(f)))E(log(q))$, where $E(\cdot)$ is the expectation function.

We can also define a log-energy graph signal entropy:

$$H_{log}(f,G) = \sum_{i \in V} log(p(f_i)) \sum_{j \in V} log(q_j).$$

Note that we substitute the terms f_i^2 by $p(f_i)^2$, and since $p(f_i) > 0$, we skip the power of 2 in the definition. Other graph signal entropies and their detailed properties will be discussed in an extended version of this paper.

In the following, we present the algorithm for graph entropy based best tree decomposition for graph signals via MST-based decomposition, see Algorithm 1.

Algorithm 1 Entropy-based best tree graph signal decomposition Algorithm via MST

Input: Graph signal S; Graph $G = (V, E, W)$, number of decomposition level N; entropy type.
1: Initialization: level $\leftarrow 1$, generate the MST of G: T.
2: **for** $level = 1, 2, \cdots, N$ **do**
3: • **if** $level < N$
4: — for each existing subband I_{level}^k, $1 \leq k \leq 2^{level}$
5: — compute the entropy for each existing subband (parent node) I_{level}^k: $E(I_{level}^k)$.
6: — decompose I_{level}^k into two subbands (children nodes), $L_{level+1}^{2k+1}, H_{level+1}^{2k+2}$, and compute the entropy for each subband $E(L_{level+1}^{2k+1}), E(H_{level+1}^{2k+2})$,
7: — **if** $E(I_{level}^k) < E(L_{level+1}^{2k+1}) + E(H_{level+1}^{2k+2})$
8: — then only retain the parent node and eliminate the children nodes, otherwise, retain both parent and children nodes.
9: — **endif**
10: • **endif**
11: **end for**
Output: The best tree decomposition of I.

Note that, Algorithm 1 can only be applied to MST based decomposition. For graph based image decomposition, sub-images are assigned 8-connected graphs, and the decomposition can be implemented iteratively.

V. Application on image decomposition

In the case of image processing, since images are time-varying signals on regular grids, we will use the classical entropy to find the best tree image decomposition. The entropy we use in this paper is the log-energy entropy, which is defined as $E(x) = \sum_i log(x_i^2)$. In our matlab experiments, we use a variation of log-energy, $E(x) = \sum_i log(p_i)$, where $p_i = p(x = x_i)$. When using the log-energy entropy, we applied a quantization on the signals, that is, $p_i = p(x \mid b_i < x \leq b_{i+1})$, with $b_0 = min(x)$, $b_N = max(x)$, $b_{i+1} - b_i = \frac{1}{N}(b_0 - b_N)$.

SEEK DIGITAL LIBRARY

The best tree decomposition and the pseudo code of the Graph based image best tree decomposition algorithm is shown in Algorithm 2.

Algorithm 2 Graph based image best tree decomposition Algorithm

Input: Data: image I; number of decomposition level N; entropy type.

1: Initialization: level $\leftarrow 1$

2: **for** $level = 1, 2, \cdots, N$ **do**

3: • **if** $level < N$

4: — for each existing subband I^k_{level}, $1 \leq k \leq 4^{level}$

5: — compute the entropy for each existing subband (parent node) I^k_{level}: $E(I^k_{level})$.

6: — decompose I^k_{level} into four subbands (children nodes), $LL^{4k+1}_{level+1}, LH^{4k+2}_{level+1}, HL^{4k+3}_{level+1}, HH^{4k+4}_{level+1}$, and compute the entropy for each subband $E(LL^{4k+1}_{level+1}), E(LH^{4k+2}_{level+1}), E(HL^{4k+3}_{level+1})$, $E(HH^{4k+4}_{level+1})$,

7: — **if** $E(I^k_{level}) < E(LL^{4k+1}_{level+1}) + E(LH^{4k+2}_{level+1}) + E(HL^{4k+3}_{level+1}) + E(HH^{4k+4}_{level+1})$

8: — then only retain the parent node and eliminate the children nodes, otherwise, retain both parent and children nodes.

9: — **endif**

10: • **endif**

11: **end for**

Output: The best tree decomposition of I.

A best tree decomposition by using the graph biorthogonal filter banks graphBior(5,5) from [9] on Lichtenstein image is shown in Figure 2.

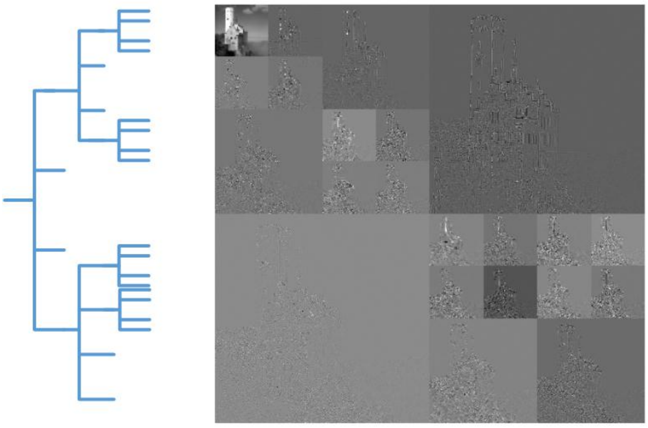


Figure 2. 3 level best tree graph filter bank decomposition for Lichtenstein image.

For comparison, we also present the results of full subband decomposition (i.e., decomposing both high-pass and low-pass components) and wavelet-like decomposition.

Our proposed best tree decomposition performs adaptively in decomposition, and keeps a balance between the full subband decomposition and wavelet-like decomposition, all the experiments used the same filter banks graphBior(5,5). For piecewise smooth images, wavelet-like decomposition gives better result, and the superiority (compare to full subband and best tree decomposition) increases with the decomposition level, see Table I.

TABLE I. RECONSTRUCTION OF IMAGES USING 3% HIGH-PASS COEFFICIENTS FOR BALLET IMAGE: SNR (DB)

level	***Full subband***	***Best tree***	***Wavelet-like***
2	45.09	**46.66**	**46.66**
3	39.37	**41.38**	**41.38**
4	35.60	**39.08**	**39.08**
5	33.13	**38.14**	**38.14**

But for images with significant textures, such as Barbara image, a full subband decomposition outperforms the wavelet-like decomposition, see Table II.

TABLE II. RECONSTRUCTION OF IMAGES USING 3% HIGH-PASS COEFFICIENTS FOR BARBARA IMAGE: SNR (DB)

level	***Full subband***	***Best tree***	***Wavelet-like***
2	**20.76**	**20.76**	19.26
3	**20.17**	20.01	17.53
4	**19.11**	18.93	17.04
5	18.06	**18.20**	16.91

We apply our best tree on the following 6 test images, i.e. Coins, Lena, Ballet, Barbara, Baboon, and Lichtenstein.

Figure 3. Test images.

In the reconstruction, all the low-pass coefficients and 3% of the largest magnitude high-pass are retained. Table III shows the SNR quality measures versus the decomposition levels, for full subband decomposition, best tree decomposition and wavelet-like decomposition. The three schemes are equivalent for 1 level decomposition, for 2-4 level decomposition, the best tree decomposition gives best performances in SNR, while wavelet-like decomposition gives best performance in 5 level decomposition. Decomposing images into low-pass and high-pass components leads to smaller and smaller coefficients as the level increases, and leading to a poorer performance of the full-subband and best tree decomposition.

TABLE III. RECONSTRUCTION OF IMAGES USING 3% HIGH-PASS COEFFICIENTS FOR 6 TEST IMAGE: SNR (DB) (AVERAGE)

level	***Full subband***	***Best tree***	***Wavelet-like***

SEEK DIGITAL LIBRARY

level	***Full subband***	***Best tree***	***Wavelet-like***
2	27.53	**27.62**	27.26
3	25.20	**25.24**	24.83
4	23.57	**23.85**	23.78
5	22.33	22.99	**23.34**

In Table III, we can see that the best tree decomposition does not perform better than wavelet-like decomposition in level 5, simply because the decay of decomposition coefficients.

VI. Conclusion

In this paper, an entropy best tree decomposition for graph signal is proposed. This method is an improvement of the recent developed wavelet-like graph signal decomposition. Wavelet-like graph signal decomposition are not effective for signals with significant high-frequency components. We modified the wavelet-like graph decomposition to a wavelet-packet-like decomposition, which is adaptive to any given graph signals. To achieve that, we define a graph signal entropy, and using it to find a best tree decomposition of a given signal. We applied the algorithm to graph-base image decomposition, our proposed method outperforms the wavelet-like decomposition.

Acknowledgment

This work was supported by the Research Grants of University of Macau MYRG2015-00049-FST, MYRG2015-00050-FST, RDG009/FST-TYY/2012, the Science and Technology Development Fund (FDCT) of Macau 100-2012-A3, 026-2013-A and Macau-China join Project 008-2014-AMJ. This research project was also supported by the National Natural Science Foundation of China 61273244.

References

[1] S. Mallat, A wavelet tour of signal processing: the sparse way. Academic press, 3rd edition, 2008.

[2] 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 Process. Mag., vol. 30, no. 3, pp. 83-98, May 2013.

[3] D. K. Hammond, P. Vandergheynst, R. Gribonval, "wavelets on graphs via spectral graph theory," Appl. Comput. Harmon. Anal., vol. 30, issue 2, pp. 129-150, March 2011.

[4] M. Crovella and E. D. Kolaczyk. "Graph wavelets for spatial traffic analysis," In INFOCOM, 2003.

[5] M. Gavish, B. Nadler, and R. R. Coifman. "Multiscale wavelets on trees, graphs and high dimensional data: Theory and applications to semi-supervised learning," In ICML, pp. 367-374, 2010. 1, 3, 4.5.

[6] R. R. Coifman and M. Maggioni. "Diffusion wavelets," Appl. Comput. Harmon. Anal., vol. 21, issue 1, pp. 53-94, July 2006.

[7] S. K. Narang and A. Ortega. "Multi-dimensional separable critically sampled wavelet filterbanks on arbitrary graphs," In ICASSP, pp. 3501-3504, March 2012.

[8] S. K. Narang, A. Ortega. "Perfect reconstruction two-channel wavelet filter banks for graph structured data," Signal Processing, IEEE Transactions on, vol. 60, no. 6, pp. 2786-2799, June 2012.

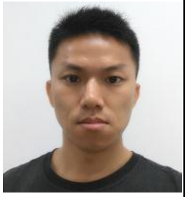
[9] S. K. Narang, A. Ortega. "Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs," Signal Processing, IEEE Transactions on, , vol. 61, no. 19, pp. 4673-4685, October 2013.

[10] Y. Tanaka, A. Sakiyama. "M-Channel Oversampled Graph Filter Banks," Signal Processing, IEEE Transactions on, vol. 62, no. 14, pp. 3578-3590, July 2014.

[11] A. Sakiyama, Y. Tanaka. " Oversampled graph Laplacian matrix for graph filter banks," Signal Processing, IEEE Transactions on, vol. 62, no. 24, pp. 6425-6437, December 2014.

[12] H. Q. Nguyen M. N. Do. "Downsampling of signals on graphs via maximum spanning trees," Signal Processing, IEEE Transactions on, vol. 63, no. 1, pp. 182-191, January 2015.

About Author (s):



Xian Wei Zheng received his B.S. degree and M.S. degree in applied mathematics from Hanshan Normal University and Shantou University, Guangdong, China, in 2009 and 2012, respectively. He is currently a Ph.D. student with Department of Computer and Information Science, University of Macau. His current research interests are graph signal processing, image processing, wavelet analysis and frame theory.

Yuan Yan Tang is a Chair Professor at University of Macau. His current interests include pattern recognition, image processing, and artificial intelligence. He is the Founder and Editor-in-Chief of International Journal on Wavelets, Multiresolution, and Information Processing (IJWMIP). He is the Founder and Chair of pattern recognition committee in IEEE SMC. Dr. Tang is the Founder and General Chair of the series International Conferences on Wavelets Analysis and Pattern Recognition (ICWAPRs). He is the Founder and Chair of the Macau Branch of International Associate of Pattern Recognition (IAPR). Dr. Y. Y. Tang is an IEEE Fellow, and IAPR Fellow.

Jian Jia Pan received his Ph.D. degree in Computer Science from Department of Computer Science at Hong Kong Baptist University, Hong Kong, in Feb. 2014. From Aug. 2014, he works as Post-doctoral Fellow in Department of Computer and Information Science, Faculty of Science and Technology, University of Macau. His research interests fall in the areas of time series analysis, machine learning, pattern recognition, and signal processing.

SEEK DIGITAL LIBRARY