HYPERGRAPH-BASED IMAGE PROCESSING

Songyang Zhang[1] *Shuguang Cui*[2] *Zhi Ding*[1]

[1]University of California, Davis
[2]Chinese University of Hong Kong, Shenzhen

ABSTRACT

Learning and processing of signals over hypergraph models have gained substantial traction owing to the ability of hypergraphs in characterizing multilateral interactions. In this work, we explore hypergraph spectral analysis and provide alternative definitions of frequency domain operations that are practically useful in image processing. We analyze hypergraph spectral properties and present several application examples, including compression, edge detection and segmentation. Successful experiment results demonstrate the effectiveness and the future prospect of the proposed hypergraph frequency operations in image processing.

Index Terms**—** Image processing, hypergraph signal processing, convolution, stationary process

1. INTRODUCTION

Graph signal processing (GSP) is a graph-theoretic tool to implement signal processing and data analytic tasks based on graph models [1]. A dataset of N data points can be modeled as a graph of N vertices, whose internal relationships are captured by edges. Graph Fourier space and the corresponding spectrum-based methods can be exploited to process signals, including images, in the GSP framework [2]. With the development of graph frequency operations and wavelet analysis [2, 3, 5], graph neural networks (GNN) have been proposed to explore the underlying data structure and complex data analysis [6, 7]. Although the GSP and GNN defined over normal graphs have achieved notable successes, they are constrained by graph edge dimensions since each edge in a normal graph only connects two nodes. Thus, a graph edge can only model pairwise relationship among nodes but cannot describe the often informative multilateral relationships in practical applications. One simple example is the network in online social communities called folksonomies, in which trilateral interactions occur among users, resources, and annotations [8]. As a result, a more general model capable of modeling high-dimensional interactions can provide added value to problems in image processing and other high-dimensional multimedia, such as point clouds and videos.

This material is based on works supported by the National Science Foundation Grant No. 1824553.

To develop a more general model for complex data structure, we venture into the realm of high-dimensional graphs called hypergraphs. Hypergraphs have found successes by generalizing normal graphs in image processing, such as classification [9] and retrieval [10]. Moreover, hypergraph convolutional networks (HGCN) [11, 12] have been extended from GNN to implement learning tasks over images. Different from GNN developed from graph convolution, HGCN lacks the basic support of hypergraph spectral analysis and fundamental hypergraph frequency operations. Fortunately, recent development of hypergraph signal processing (HGSP) provides theoretical foundation for hypergraph spectral analysis. For example, the authors of [13] proposed an analytical tool based on the simplicial complex. Moreover, the authors of [14] developed a more general HGSP framework to analyze the multilateral relationship in multimedia datasets, such as images and point clouds [15, 16]. However, we note that existing hypergraph tools are incomplete as the definitions of hypergraph spectral operations useful in image processing, such as convolution and translation, are still missing.

In this work, we investigate the use of hypergraph spectral analysis and introduce HGSP as tools in image processing. Our contributions can be summarized as follows:

- We provide alternative definitions of hypergraph spectral operations within HGSP [14] for data analysis.
- We present guidelines for applying HGSP in image processing and introduce several applications, such as edge detection, compression, and video segmentation.

When presenting test results, we compare our proposed methods against benchmarks from traditional graph and learning methods. Our experiments demonstrate the effectiveness of HGSP and hypergraph spectral analysis in image processing. Furthermore, we expect the proposed spectral operations to play important roles in the development of HGCN.

2. PRELIMINARIES

In this section, we will overview fundamentals of the general hypergraph signal processing (HGSP) [14].

Within the HGSP framework, a hypergraph with N nodes and longest hyperedge connecting M nodes, is represented by an Mth-order N-dimensional adjacency tensor

$\mathbf{A} \in \mathbb{R}^{\overbrace{N\times N\times\cdots\times N}^{\text{M times}}}$ which can be also decomposed via orthogonal CP decomposition [14], i.e.,

$$\mathbf{A} = (a_{i_1 i_2 \cdots i_M}) \approx \sum_{r=1}^{N} \lambda_r \cdot \underbrace{\mathbf{f}_r \circ ... \circ \mathbf{f}_r}_{\text{M times}}, \tag{1}$$

where $\circ$ is the tensor outer product [14]. Note that the hyperedges with fewer nodes than M are normalized with weights as shown in [14, 17]. In the orthogonal CP decomposition, $\mathbf{f}_r$'s are orthonormal basis called spectrum components and λ_r are frequency coefficients related to the hypergraph frequency. All the spectrum components $\{\mathbf{f}_1, \cdots, \mathbf{f}_N\}$ form the hypergraph spectral space.

Similar to GSP, hypergraph signals are attributes of nodes. Intuitively, the original signal is defined as $\mathbf{s} = [s_1 \quad s_2 \quad ... \quad s_N]^{\mathrm{T}} \in \mathbb{R}^N$. Since the adjacency tensor $\mathbf{A}$ describes the signals in the high-dimensional interactions, we need to define a specific form of the hypergraph signal to work with the representing tensor, i.e.,

$$\mathbf{s}^{[M-1]} = \underbrace{\mathbf{s} \circ ... \circ \mathbf{s}}_{\text{M-1 times}}. \tag{2}$$

The resulted hypergraph signal is an $(M-1)$th-order N-dimensional tensor. Hypergraph shifting is defined as the contraction of the adjacency tensor and the hypergraph signal:

$$\mathbf{s}' = \mathbf{A}\mathbf{s}^{[M-1]}. \tag{3}$$

The insight to interpret hypergraph shifting over the adjacency tensor is shown in [14].

Based on the definition of hypergraph spectrum and hypergraph signals, the hypergraph Fourier transform (HGFT) is defined as

$$\hat{\mathbf{s}} = \mathcal{F}_C(\mathbf{s}) = [(\mathbf{f}_1^{\mathrm{T}}\mathbf{s})^{M-1} \cdots (\mathbf{f}_N^{\mathrm{T}}\mathbf{s})^{M-1}]^{\mathrm{T}}. \tag{4}$$

Because of the page limitation, we shall refrain from presenting the details on many fundamental aspects of HGSP. Certain basic ideas, such as the implementation of HGFT, HGSP filter design, bandlimited signals, and hypergraph frequency analysis as well as their interpretations can be found in [14].

3. HYPERGRAPH FREQUENCY OPERATIONS

In this section, we introduce important operations for the hypergraph frequency analysis useful in image processing.

3.1. Convolution

Convolution is an essential operation in traditional image processing. In DSP and GSP [4, 20], Fourier transform of convolution between two signals is equal to the product between their respective Fourier transforms. Similarly, we generalize the hypergraph convolution denoted by $\diamond$ as

$$\mathbf{x} \diamond \mathbf{y} = \mathcal{F}_C^{-1}(\mathcal{F}_C(\mathbf{x}) * \mathcal{F}_C(\mathbf{y})), \tag{5}$$

where $\mathcal{F}_C$ is the HGFT, $\mathcal{F}_C^{-1}$ is the iHGFT, and $*$ denotes Hadamard product [14]. This definition generalizes the property that the convolution in the vertex domain is equivalent to the product in the hypergraph spectral domain.

3.2. Translation

The classic translation in DSP can be written as the convolution between the signal and an impulse function centered at a certain point. With the definition of hypergraph convolution, we define the hypergraph translation of an original signal $\mathbf{x} \in \mathbb{R}^N$ similar to that in GSP [20] as

$$\mathcal{T}_n \mathbf{x} = \sqrt{N}\mathbf{x} \diamond \Delta_n, \tag{6}$$

where the nth element of the Kronecker $\Delta_n \in \mathbb{R}^N$ is 1 and other elements are 0. Similar to translation in GSP, hypergraph translation is not like the time shift of signal in DSP. Instead, it represents a hypergraph convolution kernel localizing the information near the centered node $\mathbf{v}_n$ [5], which helps capture topological information among pixels in images.

3.3. Sampling and Interpolation

Sampling is an important operation in image processing, which selects a subset of individual data points to estimate the characteristics of the whole population. Similar to sampling signals in time and GSP [21], the HGSP sampling theory can be developed to sample signals over the vertex domain. Since the reduction of tensor order may break the structure of hypergraph and cannot always guarantee perfect recovery, we adopt the dimension reduction of each order. The sampling of a hypergraph signal is defined as follows:

Suppose that Q is the dimension of each sampled order. The sampling operation of a hypergraph signal $\mathbf{s}^{[\mathbf{M-1}]} \in \mathbb{R}^{\overbrace{N\times N\times \ldots\times N}^{M-1 \text{ times}}}$ is defined as

$$\mathbf{s}_{\mathbf{Q}}^{[\mathbf{M-1}]} = \mathbf{s}^{[\mathbf{M-1}]} \times_1 \mathbf{U} \times_2 \mathbf{U} \cdots \times_{M-1} \mathbf{U}, \tag{7}$$

where $\times_n$ denotes the n-mode product [14], the sampling operator is $\mathbf{U} \in \mathbb{R}^{Q\times N}$, and the sampled signal is $\mathbf{s}_{\mathbf{Q}}^{[\mathbf{M-1}]} \in \mathbb{R}^{\overbrace{Q\times Q\times \ldots\times Q}^{M-1 \text{ times}}}$. The interpolation operation is defined by

$$\mathbf{s}^{[\mathbf{M-1}]} = \mathbf{s}_{\mathbf{Q}}^{[\mathbf{M-1}]} \times_1 \mathbf{T} \times_2 \mathbf{T} \cdots \times_{M-1} \mathbf{T}, \tag{8}$$

where the interpolation operator is $\mathbf{T} \in \mathbb{R}^{N\times Q}$.

With the definition of HGSP sampling operations, we have the following theorem.

Theorem 1 *Define the sampling operator* $\mathbf{U} \in \mathbb{R}^{Q\times N}$ *according to* $U_{ji} = \delta[i - q_j]$ *where* $1 \leq q_i \leq N,\ i = 1, \ldots, Q$. *By choosing* $Q \geq K$ *and the interpolation operator* $\mathbf{T} = \mathcal{F}_{[K]}^{\mathrm{T}}\mathbf{Z} \in \mathbb{R}^{N\times Q}$ *with* $\mathbf{ZU}\mathcal{F}_{[K]}^{\mathrm{T}} = \mathbf{I_K}$ *and* $\mathcal{F}_{[K]}^{\mathrm{T}} = [\mathbf{f}_1, \cdots, \mathbf{f}_K]$, *perfect recovery can be achieved, i.e.,* $\mathbf{s} = \mathbf{TUs}$ *for all* K*-bandlimited original signal* $\mathbf{s}$ *and the corresponding hypergraph signal* $\mathbf{s}^{[M-1]}$.

The proof of *Theorem 1* is given in [14]. This theorem shows that perfect recovery is possible for HGSP by choosing suitable sampling and interpolation, which can be further applied in applications, such as image compression.

3.4. Hypergraph Stationary Process

Stationarity is an important property in analyzing observations of random variables, which can be applied in high-dimensional data structure with multiple observations, such as videos or point clouds. In [18], a graph stationary process is defined for graph-based observed samples. Within the HGSP framework of [14], we extend the graph stationarity and define a hypergraph stationary process based on the supporting matrix to process 3D point clouds [15]. Here, we revisit a useful property of hypergraph stationarity for image processing as follows by leaving its proof in [15].

Theorem 2 *A stochastic signal* $\mathbf{x}$ *is weak-sense stationary if and only if it has zero-mean and its covariance matrix has the same eigenvectors as the hypergraph spectrum basis, i.e.,*

$$\mathbb{E}[\mathbf{x}] = \mathbf{0} \tag{9}$$

$$\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{V}\Sigma_{\mathbf{x}}\mathbf{V}^H, \tag{10}$$

where $\mathbf{V} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_N]$ *is the hypergraph spectrum.*

This property implies that we can estimate hypergraph spectral components from the eigenspace of the covariance matrix of observations, which can help estimate the hypergraph spectrum for multimedia datasets. We shall give an application example of this stationarity property in Section 4.3.

4. APPLICATION EXAMPLES

4.1. Image Compression

Efficient compression of signals is important in data analysis and signal processing. Projecting signals onto a suitable orthonormal basis is very common in compression. Within the proposed HGSP framework, we can represent N signals in the original domain with C frequency coefficients in the hypergraph spectrum domain.

More specifically, according to *Theorem 1*, we can losslessly compress a K-bandlimited signal of N signal points with K spectrum coefficients. To test the performance of our HGSP sampling, we compare the results of image compression with those from GSP-based compression method [19].

Fig. 1. Test Set of Images.

Table 1. Compression Ratio of Different Methods

image	ct	lenna	mri	AVG
IANH-HGSP	**1.41**	**1.57**	**1.53**	**1.50**
(α, β)-GSP	1.07	1.07	1.11	1.08
4 connected-GSP	1.07	1.04	1.05	1.05

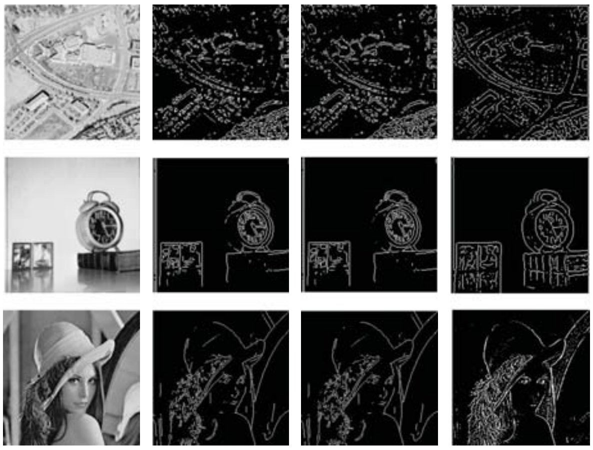
We tested three size-256×256 photo images, shown in Fig. 1. We used the Image Adaptive Neighborhood Hypergraph (IANH) model [22] to construct the hypergraph structure. To lower complexity, we chose three closest neighbors in each hyperedge to construct a third-order adjacency tensor. When applying GSP-based method of [19], we represented images as graphs with 1) the 4-connected neighbor model [23], and 2) the distance-based model in which an edge exists only if the spatial distance is below α and the pixel distance is below β. We use the compression ratio CR= N/C to measure compression efficiency. The results summarized in Table 1 illustrate higher compression ratio (efficiency) achieved by HGSP-based compression over the GSP-based compression. We believe HGSP could play an important role in both lossless and lossy data compression in the future works.

4.2. Edge Detection

Convolution-based method is widely used in the edge detection of images in DSP. Similarly, hypergraph convolution-based operations can also play a role in image edge detection. Here, we introduce a hypergraph translation-based method for edge detection. To implement HGSP-based method, we first design a 9×9 mask and use it as a hypergraph with 81 nodes by the IANH model. Next, we implement hypergraph-based translation at the center of the mask. The summation of the translated result becomes the new value of the center node. Hovering the mask over the whole image, this process can be understood as blurring the hypergraph signals with the information of center node. We can design a threshold for the difference between the translated values and the original pixels to detect edges. We compared the proposed method with the Sobel and Prewitt methods [24] on three SIPI image datasets (http://sipi.usc.edu/database/) as shown in Fig. 2. From the result, we can see that HGSP-based method detects the details and the edges more explicitly. With deeper understanding of hypergraph convolution-based kernels, HGSP shows clear promise in edge detection.

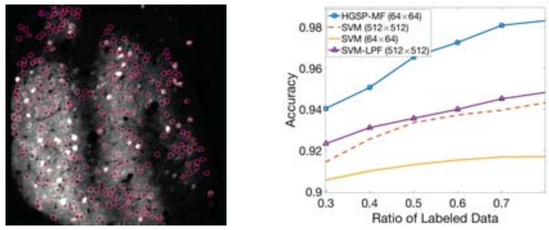
4.3. Hypergraph Convolution Filter for Segmentation

Deep-learning methods, like graph and hypergraph convolutional networks [7, 11, 12] have achieved significant success



(a) Original (b) Sobel (c) Prewitt (d) HGSP

Fig. 2. Results of Edge Detection.



(a) Detection of neurons (b) Comparison with SVM.

(c) Ground Truth of Segmentation (d) HGSP-based Segmentation with MF. (e) SVM-based Segmentation with LPF.

in data analysis. However, these learning methods do not provide analysis and interpolation in the hypergraph spectrum domain. To explore hypergraph convolution operations, we propose a convolution-based filter for semi-supervised learning in video segmentation. Define $\mathbf{V} = [\mathbf{f}_1 \quad \cdots \mathbf{f}_N]$. Given the definition of (5), a single-step convolution filter with parameter $\mathbf{y}$ on signal $\mathbf{x}$ is defined as

$$F_{\mathbf{y}}(\mathbf{x}) = \mathbf{x} \diamond \mathbf{y} = \mathbf{V} diag(\mathbf{f}_i^T \mathbf{y}) \mathbf{V}^T \mathbf{x}. \tag{11}$$

We now explain the use of convolution filter to segment. We consider a dataset (http://neurofinder.codeneuro.org/) in a biomedical application. In such datasets, a series of time-varying images are provided to record the activities of neuron shown as Fig. 3(a). The task here is to segment neurons from the background if we know part of labels in the ground truth shown as Fig. 3(c). Such a video with $K = 3024$ frames and $N = 512^2$ pixels in each frame can be modeled as a hypergraph with N nodes and each node has K observations.

Assume that $\mathbf{x}_i \in \mathbb{R}^N, i = 1, \cdots K$ are the pixels in the ith frame. With the property of hypergraph stationary process in *Theorem 2*, we can easily estimate the hypergraph spectrum $\mathbf{V}$ for the time-varying images by applying decomposition on the normalized covariance matrix from K observations $\mathbf{x}_i$'s. Suppose that ℓ nodes have labels and $N-\ell$ do not. We can estimate the filter parameters $\mathbf{y}$ by minimizing the error between the given labels $\mathbf{L} \in \mathbb{R}^\ell$ and the ℓ corresponding filtered signals $F_{\mathbf{y}}(\mathbf{x})_\ell \in \mathbb{R}^\ell$, i.e.,

$$\min_{\mathbf{y}} ||\mathbf{L} - F_{\mathbf{y}}(\tilde{\mathbf{x}})_l||_2^2, \tag{12}$$

where $\tilde{\mathbf{x}} = \frac{1}{K}\sum \mathbf{x}_i$ is the pixel average. More specifically, this optimization problem has a solution at

$$\mathbf{y} = \mathbf{W}^{-1}\mathbf{L}, \tag{13}$$

where each row of $\mathbf{W} \in \mathbb{R}^{l \times N}$ is the corresponding row of $\mathbf{V} diag(\mathbf{f}_i^T \tilde{\mathbf{x}}) \mathbf{V}^T$ with the same index as the labeled data.

Table 2. Comparison of Computation Time (in Seconds)

HGSP(64×64)	SVM (64×64)	SVM (512×512)
2413	403	82813

From the estimated parameters $\mathbf{y}$, we obtain the filtered signals for all the nodes in $F_{\mathbf{y}}(\tilde{\mathbf{x}})$ and apply a threshold to determine its label (i.e., $\{\pm 1\}$) in the binary classification. To tackle possible noise amplification, we further apply a 4×4 median filter (MF) after the convolution filter.

In the experiment, we cut the original 512×512 images into 64×64 non-overlapping blocks to lower complexity. One result with $\ell = 40\% \cdot N$ labeled data achieves 95.06% accuracy, as shown in Fig. 3(d) and shows much clearer segmentation compared to the SVM result after a 4×4 averaging-kernel lowpass filter (LPF) [25] in Fig. 3(e). Fig. 3(b) further compares our method to SVM with and without LPF in terms of estimated label accuracy. From the comparison, we can see that the performances of SVM improves with larger processing blocks at the cost of computation complexity. LPF can improve the performance of SVM results. Still, the HGSP-based method achieves the best performance. We measured the computation time for these methods in Table 2. It is clear that HGSP-based methods achieves superior accuracy with moderate computation cost.

5. CONCLUSION

In this work, we introduce several fundamental definitions and properties of hypergraph frequency operations for image processing. We further present several application examples based on HGSP operations to illustrate the practical utility of HGSP in multimedia signal processing. Our goal is to develop new signal processing methods and analytical hypergraph tools for effective image processing. Future applications of HGSP beyond the proposed image processing problems include those in virtual reality, point clouds, and audio/video processing. Moreover, HGSP is also expected to be a useful tool in the development and analysis of HGCN.

6. REFERENCES

[1] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808-828, Apr. 2018.

[2] A. Sandryhaila, and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644-1656, Apr. 2013.

[3] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129-150, Mar. 2011.

[4] J. Shi, and J. M. Moura, "Graph signal processing: modulation, convolution, and sampling," arXiv:1912.06762.

[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, Apr. 2013.

[6] B. Xu, H. Shen, Q. Cao, Y. Qiu, and X. Cheng, "Graph wavelet neural network," arXiv:1904.07785.

[7] T. N. Kipf, and M. Welling, "Semi-supervised classification with graph convolutional networks," arXiv:1609.02907.

[8] K. Ahn, K. Lee, and C. Suh, "Hypergraph spectral clustering in the weighted stochastic block model," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 5, pp. 959-974, May 2018.

[9] J. Yu, D. Tao, and M. Wang, "Adaptive hypergraph learning and its application in image classification," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3262-3272, Mar. 2012.

[10] Y. Huang, Q. Liu, S. Zhang, and D. N. Metaxas, "Image retrieval via probabilistic hypergraph ranking," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, Jun. 2010, pp. 3376-3383.

[11] S. Bai, F. Zhang, and P. H. Torr, "Hypergraph convolution and hypergraph attention," arXiv:1901.08150.

[12] N. Yadati, M. Nimishakavi, P. Yadav, V. Nitin, A. Louis, and P. Talukdar, "HyperGCN: A new method For training graph convolutional networks on hypergraphs," in *NeurIPS*, Vancouver, Canada, Dec. 2019, pp. 1509-1520.

[13] S. Barbarossa, and M. Tsitsvero, "An introduction to hypergraph signal processing," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, Mar. 2016, pp. 6425-6429.

[14] S. Zhang, Z. Ding, and S. Cui, "Introducing Hypergraph Signal Processing: Theoretical Foundation and Practical Applications," *IEEE Internet of Things Journal*, vol. 7, pp. 639 - 660, Jan. 2020.

[15] S. Zhang, S. Cui, and Z. Ding, "Hypergraph spectral analysis and processing in 3d point cloud," arXiv:2001.02384.

[16] S. Zhang, S. Cui, and Z. Ding, "Point cloud segmentation based on hypergraph spectral clustering," arXiv:2001.07797.

[17] A. Banerjee, C. Arnab, and M. Bibhash, "Spectra of general hypergraphs," *Linear Algebra and its Applications*, vol. 518, pp. 14-30, Dec. 2016.

[18] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 5911-5926, Aug. 2017.

[19] A. Sandryhaila, and J. M. F. Moura, "Discrete signal processing on graphs: frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042-3054, Apr. 2014.

[20] D. I. Shuman, B. Ricaud, and P. Vandergheynst, "A windowed graph Fourier transform," in *2012 IEEE Statistical Signal Processing Workshop*, Ann Arbor, MI, USA, Aug. 2012, pp. 133-136.

[21] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: sampling theory," *IEEE transactions on signal processing*, vol. 63, no. 24, pp. 6510-6523, Dec. 2015.

[22] A. Bretto, and L. Gillibert, "Hypergraph-based image representation," in *International Workshop on Graph-Based Representations in Pattern Recognition*, Berlin, Heidelberg, Apr. 2005, pp. 1-11.

[23] A. Morar, F. Moldoveanu, and E. Gröller, "Image segmentation based on active contours without edges," in *Proc. of 2012 IEEE 8th International Conference on Intelligent Computer Communication and Processing*, Cluj-Napoca, Romania, Sept. 2012, pp. 213-220.

[24] A. K. Cherri, and M. A. Karim, "Optical symbolic substitution: edge detection using Prewitt, Sobel, and Roberts operators," *Applied Optics*, vol. 28, no. 21, pp. 4644-4648, Nov. 1989.

[25] R. Lionnie, I. K. Timotius, and I. Setyawan, "Performance comparison of several pre-processing methods in a hand gesture recognition system based on nearest neighbor for different background conditions," *Journal of ICT Research and Applications*, vol. 6, no. 3, pp. 183-194, 2012.