International Journal Of Engineering Research And Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 15, Issue 1 (January 2019), PP.22-30

Visualization of Large Dataset Using Approximate Commute Time Embedding

Hee Il Hahn
Department of Information and Communications Engineering Hankuk University of Foreign Studies, Yongin Kyonggi-Do, Korea

ABSTRACT: *This paper presents visualization of the data set lying in a higher dimensional space using commute time embedding. However, commute time embedding involves computing spectral decomposition of the graph Laplacian matrix, which requires computational burden proportional to O(n³), which might not be suitable for large scale dataset. Recently, many methods have been proposed to reduce the computational burden. These methods, which usually involves sampling the affinity matrix of the graph, might be suffered from the distortion induced by computing spectral decomposition of the normalized graph Laplacian from the sampled affinity matrix. This paper proposes how to reduce the distortion by preserving the properties the normalized graph Laplacian matrix should be symmetric and positive semidefinite even after its approximation by sampling process. The performance of the proposed algorithm is analyzed by checking the embedding results on a patch graph in both geometric and topological ways.*
Keywords: *Approximate commute time embedding,Normalized graph Laplacian,Persistent homology.*

Date of Submission: 25-02-2019 Date of acceptance: 18-03-2019

I. INTRODUCTION

Over last decades, there have been several different embedding algorithms developed for dimensionality reduction in manifold ways. They make it possible to visualize the dataset lying in a higher dimensional space by embedding them in a two or three dimensional space. This can be a meaningful clue for a desired output in the area of pattern recognition or machine learning, because it helps to imagine what they look like. When dataset lies on a linear subspace, PCA(principal component analysis) is most useful and optimal for embedding as well as dimensionality reduction in terms of maintaining maximum variance of dataset. However, when dataset lies on a nonlinear space, PCA introduces severe error. The manifold learning algorithms replace PCA on a nonlinear space. Although there are lots of manifold learning algorithms, commute time embedding is known to be most suitable for dimensionality reduction.

Basically, manifold learning algorithms usually use kernel-based methods to represent the data of nonlinear structure, because it is more clear in the feature space rather than in the original input space. In kernel-based methods, kernel matrix is defined when input data are mapped to the feature space using a map F : X ® F . Kernel matrix $G \in \mathbf{R}^{n \times n}$, whose element $G_{ij} = k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$, can be represented in terms of an inner product between feature vectors. For example, ISOMAP [14], which is known to be the first manifold learning algorithm, computes the matrix whose elements are the geodesic distance between pairs of nodes. On the contrary, Laplacian eigenmap [2] and commute time embedding [11] employ graph Laplacian matrices as kernels. This scheme needs to compute the spectral decomposition of the kernel matrix, whose computational burden increases in the order of O(n³), where n is the size of dataset. It is hardly applicable when n is sufficiently large.

Recent studies try to solve this problem by computing an approximate commute time without any spectral decomposition. it is approximated by using random projection based on Johnson-Lindenstrauss theorem [1] and Spielman-Teng algorithm [12]. It is widely known that this method can be applied to large scale spectral clustering. However, it is not appropriate to the visualization of dataset because it only approximates commute time between pairs of points rather than embedding coordinate of each point.

In this work, a different approach is proposed to solve this problem by accelerating the computational time of spectral decomposition. Nyström method, proposed by Williams, et al. [15], is adopted in [7] to compute the kernel eigenfunctions approximately from the affinity matrix composed of randomly chosen samples. In spite of lots of researches on approximation of kernel matrices, there have been rarely known about approximation of the normalized graph Laplacian. The properties that a normalized graph Laplacian matrix is symmetric and positive semidefinite (SPSD) should be preserved even after its approximation. However, the method proposed by Choromanska, et al. [6], which approximates a normalized graph Laplacian based on the Nyström method, does not satisfy it any more. They try to create a submatrix by randomly sampling the columns of graph Laplacian and normalizing it directly. However, it induces a severe error because the submatrix is no longer SPSD.

This paper proposes a new method to approximate a normalized graph Laplacian, which reduces the approximation error significantly. The performance of the proposed method is analyzed by comparing the eigenvalues of the normalized graph Laplacian with the true ones as well as their embedding geometries, which are generated using the patch graph constructed using the overlapped samples of data. The patch graph, which is proposed by Taylor, et al. [13], effectively organizes the patches extracted from images or waveforms according to the graph-based metrics. Their recent studies on the patch graph and its embedding give convincing ideas of analyzing signals from the geometric point of view. It is known that commute time embedding results of periodic or quasi-periodic waveforms are represented as closed curves on the low dimensional Euclidean space, while those of aperiodic signals have the shape of open curves [10]. Based on this property, persistent homology is employed to determine the topological structures of the embedding geometries.

The outline of the paper is as follows: The next section reviews construction of patch graphs and its commute time embedding. In section 3, the proposed method to approximate a normalized graph Laplacian is described in detail, from which approximate commute time embedding is constructed. we present in section 4 experiments related to approximation of a normalized graph Laplacian and investigate its topological performance when it is applied to the commute time embedding. Finally in section 5, we conclude with directions for future research.

II. BACKGROUND

2.1Construction ofa Patch Graph

Suppose that maximally overlapped patches ofsize p samples are extracted around each time sample in the following way:

$$\hat{\bar{\mathbf{x}}}_n = \frac{\bar{\mathbf{x}}_n}{\|\bar{\mathbf{x}}_n\|} \in S^{p-2}, \quad n = 1, 2, \cdots, N \quad (1)$$

Where $\bar{\mathbf{x}}_n$ is the mean-centered version of $\mathbf{x}_n = \left(x[n], x[n+1], \cdots, x[n+p-1]\right)^T \in \mathbf{R}^p$ and S^{p-2} represents $p-2$ sphere. The patch $\hat{\bar{\mathbf{x}}}_n$ is obtained by normalizing with the magnitude, so that it may not be sensitive to changes in the local energy of the signal. In this work, a patch $\hat{\bar{\mathbf{x}}}_n$ is regarded as a vector on the $p-2$ dimensional sphere embedded on the $p-1$dimensional Euclidean space. Thus, the signal is reformatted as a patch set, with which the graph of patches is constructed.The weight along the edge connecting the nodes v_i with v_j, which are associated with $\hat{\bar{\mathbf{x}}}_i$ and $\hat{\bar{\mathbf{x}}}_j$, respectively, is defined as follows:

$$w(i,j) = \begin{cases} e^{-\left\|\hat{\bar{\mathbf{x}}}_i - \hat{\bar{\mathbf{x}}}_j\right\|^2 / 2\sigma^2} & v_i, v_j : \text{connected} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Given a set of patches $\hat{\bar{\mathbf{x}}}_1, \cdots, \hat{\bar{\mathbf{x}}}_N$ and some measures of similarity between any pair of data $w(i,j)$, a graph can be constructed, where two vertices v_i and v_j are connected with a weight $w(i,j)$ if v_i is among the k-nearest neighbors of v_j or v_j is among the k-nearest neighbors of v_i [3]. Computing similarities between pairs of patches allows us to map the patches at the ambient space into some geometry at the embedding subspace [15].Given the distance matrix M, whose u,v entries are the distance between pairs of nodes v_u and v_v,it is possible to embed all the nodes into the Euclidean space by computing a gram matrix associated with them.

2.2Commute Time Embedding

Given an affinity matrix $W \in \mathbf{R}^{N \times N}$, the graph Laplacian matrix is defined as $L = D - W$, where the degree matrix D is a diagonal matrix with entries $\sum_{v=1}^{N} w(u,v)$. It is assumed that a patch graph is connected and undirected.Let $L = U \Lambda U^T$ be the spectral decomposition of L, where U is the matrixcontaining all eigenvectors as columns and Λ the diagonal matrix with the eigenvalues $\lambda_1, \cdots, \lambda_N$. Denote by $L^\dagger$ the Moore-Penrose pseudo-inverse of L, $L^\dagger = U \Lambda^\dagger U^T$, where $\Lambda^\dagger = diag\left(\lambda_1^\dagger, \cdots, \lambda_N^\dagger\right)$ is defined as

$$\lambda_i^\dagger = \begin{cases} 1/\lambda_i & \lambda_i \neq 0 \\ 0 & otherwise. \end{cases} \quad (3)$$

Then we have

$$c(i,j) = vol \cdot (e_i - e_j)^T L^\dagger (e_i - e_j) \quad (4)$$

where $vol = \sum_{u,v} w(u,v)$ and $e_i = [0,\cdots,0,1,0,\cdots,0]^T$ is defined as the i^{th} column of the identity matrix I. Hence, $\sqrt{c(i,j)}$ can be considered a Mahalanobis distance with a weighting matrix $vol \cdot L^\dagger$ and $c(i,j)$ of the above equation can be rephrased as

$$\begin{aligned} c(i,j) &= vol \cdot (e_i - e_j)^T U \Lambda^\dagger U^T (e_i - e_j) \\ &= (z_i - z_j)^T (z_i - z_j). \end{aligned} \quad (5)$$

Then, the commute time between v_i and v_j is given as $c(i,j) = (z_i - z_j)^T (z_i - z_j)$, where

$$z_i = \sqrt{vol} \cdot \left(u_{i,2} \big/ \sqrt{\lambda_2}, \cdots, u_{i,N} \big/ \sqrt{\lambda_N} \right) \quad (6)$$

L can be normalized as $L_{sym} = D^{-1/2} L D^{-1/2}$, which is called the normalized graph Laplacian. If L_{sym} is used instead of L, $LU = U\Lambda$ becomes $L_{sym} V = V\Lambda'$ where $\Lambda' = \Lambda D^{-1/2}$ and $V = D^{-1/2} U$. Thus, z_i can be rephrased as

$$z_i = \sqrt{vol} \left(v_{i,2} \big/ \sqrt{\lambda_2' d_i}, \cdots, v_{i,N} \big/ \sqrt{\lambda_N' d_i} \right) \quad (7)$$

This allows us to interpret $\sqrt{c(i,j)}$ as the Euclidean distance between two nodes z_i and z_j on the embedding subspace. For the dimensionality reduction, it is not needed to use all the components in the embedding defined by the above equation. Instead, only the first q components corresponding to the lower eigenvectors is used in the following way [13, 15]:

$$z_i' = \sqrt{vol} \left(\frac{v_{i,2}}{\sqrt{\lambda_2' d_i}}, \cdots, \frac{v_{i,q+1}}{\sqrt{\lambda_{q+1}' d_i}} \right), \quad (8)$$

where λ_k' 'sare assumed to be sorted in the following way:

$$0 \le \lambda_1' \le \lambda_2' \le \cdots \lambda_N' < 2 \quad (9)$$

2.3Persistent Homology

Homology makes it possible to analyze the topological properties of objects given as a simplicial complex [9]. Specifically, the Betti numbers β_n, the ranks of the n^{th} dimensional homology groups, are special types of topological invariants capturing the topological properties of many geometrical constructions. In fact, β_0 gives the number of connected components of a topological space, β_1 measures the number of one dimensional topological holes, and β_2 counts the number of two dimensional topological voids.

The most obvious way to convert a collection of points in a metric space into a global object is to use the point cloud as the vertices of a simplicial complex whose edges are determined by proximity vertices within some specified distance ε. Given an embedding map $f: M \to \mathbf{R}^m$, we can determine the topological invariants of $Y = f(X) \subset \mathbf{R}^m$, from which simplicial complex K is constructed. A simplicial complex is a set K in which any two simplices are either disjoint or they intersect in a common face, which is a simplex of smaller dimension.We need to compute a nested subsequence of complexes called a filtration of a complex $0 = K^0 \subseteq K^1 \subseteq \cdots \subseteq K^m = K$ .It is assumed that $K^n = K$ for $n \ge m$.

Persistent homology computes the homology groups along a filtration of complex and measures their topological importance by persistence intervals of their nontrivial homology classes [18].In this paper, persistent homology is represented using the persistence barcode, which displays graphically the set of persistence intervals as a collection of horizontal line segments.

III. DERIVING APPROXIMATE COMMUTE TIME EMBEDDING

3.1Review of Nyström Method

Symmetric, positive semidefinite kernel matrix G of size $n \times n$ can be decomposed as $G = U\Lambda U^T$, where Λ is a diagonal matrix whose diagonal elements are eigenvalues of G and columns of U are orthogonal eigenvectors of G corresponding to the eigenvalues. Let C be the $n \times c$ matrix obtained by randomly sampling c columns of G uniformly without replacement and $\hat{G}$ be the $c \times c$ matrix consisting of the intersection of these c columns with the corresponding c rows of G. Assume that the rows and columns of G are rearranged in the following way:

$$G_r = \begin{bmatrix} \hat{G} & A^T \\ A & B \end{bmatrix}, \qquad \mathrm{C} = \begin{bmatrix} \hat{G} \\ A \end{bmatrix} \quad (10)$$

Nyström method [7] approximates G as $\tilde{G} = C\hat{G}_k^{\dagger}C^T$ where $\hat{G}_k$ is adjusted so that its rank k be $1 \le k \le c$ and $\hat{G}_k^{\dagger}$ is the Moore Penrose inverse of $\hat{G}_k$. It has been shown that $\tilde{G}$ converges to G as c increases and the spectral decomposition of $\tilde{G}$ is represented as $\tilde{G} = \tilde{U}\tilde{\Lambda}\tilde{U}^T$. Here $\tilde{U}$ and $\tilde{\Lambda}$ are defined as follows:

$$\tilde{U} = \sqrt{\frac{c}{n}} C U_{\hat{G}} \Lambda_{\hat{G}}^{\dagger} \qquad \tilde{\Lambda} = \frac{n}{c}\Lambda_{\hat{G}} \quad (11)$$

Where $\hat{G} = U_{\hat{G}}\Lambda_{\hat{G}}U_{\hat{G}}^T$. It means that the eigenvalues Λ and the corresponding eigenvectors U of G can be approximated to $\tilde{\Lambda}$ and $\tilde{U}$, respectively.

3.2Approximating Spectral Decomposition of the Normalized Graph Laplacian

This section explains how to approximate spectral decomposition $L_{sym} = D^{-1/2}LD^{-1/2}$, given an affinity matrix W. Let V be the $n \times c$ matrix obtained by randomly sampling c columns of an W uniformly without replacement and P be the $c \times c$ matrix obtained by extracting the rows of the same indices as the sampled c columns of W. The rows and columns of W are rearranged as follows:

$$W_r = \begin{bmatrix} P & Q^T \\ Q & H \end{bmatrix}, \qquad \mathrm{V} = \begin{bmatrix} P \\ Q \end{bmatrix} \quad (12)$$

Choromanska, et al [6] approximate L_{sym} by computing $\tilde{L}_{sym} \leftarrow \hat{I} - \sqrt{c/n}D^{-1/2} \cdot V \cdot \Delta^{-1/2}$. Here, $\hat{I}$ is obtained by sampling the columns of the identity matrix I where Ω represents the index set $\{i_1, i_2, \cdots, i_c\}$ sampled uniformly, which is denoted as $\hat{I} \leftarrow I(\ :\ ,\ \Omega\ )$. And the degree matrices $D \in \mathbf{R}^{n \times n}$, $\Delta \in \mathbf{R}^{c \times c}$ are defined as follows:

$$D(i,i) = \sum_{j=1}^{c} V(i,i), \qquad \Delta(i,i) = \sum_{j=1}^{n} V(j,i). \quad (13)$$

They approximate the eigenvalues and eigenvectors of L_{sym} using the spectral decomposition of $\hat{L}_{sym} \leftarrow \tilde{L}_{sym}(\ \Omega,\ :\ )$. The eigenvalues and eigenvectors of $\hat{L}_{sym}$ deviate from those of L_{sym} and do not even satisfy the properties that the least eigenvalue be zero, since it is no longer SPSD.

In this work, a different method to approximate the spectral decomposition of L_{sym} is proposed, which preserves the properties that L_{sym} be SPSD even though it is approximated. The procedure to approximate the spectral decomposition of L_{sym} usingNyström method, is described in detail as follows.

Since W_r is SPSD, P is also SPSD. Firstly, P is normalized using its degree matrix D_P to get $\hat{P} \leftarrow D_P^{-1/2}PD_P^{-1/2}$, where $D_P(i,i) = \sum_j P(i,j)$. However, Q should be normalized differently, as given in $\hat{Q} \leftarrow D_Q^{-1/2}Q\Delta_Q^{-1/2}$, where $D_Q(i,i) = \sum_{j=1}^{c} Q(i,j)$ and $\Delta_Q(i,i) = \frac{c}{n-c}\sum_{j=1}^{c} P(i,j)$, since Q is neither symmetric nor a square matrix. Finally, the eigenvalues and eigenvectors of L_{sym} are approximated as

$\tilde{\Lambda} = 1 - \frac{n}{c}\Lambda_{\hat{P}}$, $\tilde{U} = \sqrt{\frac{c}{n}}\hat{V}U_{\hat{P}}\tilde{\Lambda}^{\dagger}$, respectively,using the spectral decomposition of $\hat{P} = U_{\hat{P}}\Lambda_{\hat{P}}U_{\hat{P}}^{T}$. In other words, L_{sym} can be approximated as $\tilde{L}_{sym} = \tilde{U}\tilde{\Lambda}\tilde{U}^{T}$, which is due to the fact that W and L_{sym} have the same eigenvectors and $\Lambda = I - \Lambda_W$, where Λ_W and Λ are the eigenvalues of W and L_{sym} , respectively.Given the approximate eigenvalues and eigenvectors, commute time embedding can be implemented using (7) or (8).

IV. EXPERIMENTS

Patch sets associated with signals are constructed, where the size of patch is decided experimentally with $p = 25$ samples. When the commute time embedding is performed on the patch set composed of N patches, each vertex is mapped into an $N-1$ dimensional vector, as shown in (7), which causes severe burden and makes it hard to get a feel for what the data look like. In this work, dimensionality reduction is employed so that the data can be embedded on three dimensional space, as depicted in (8),where $q = 3$, because it is possible to visualize them on the embedding subspace if the patch set can be represented on two or three dimensional space.

4.1 Characteristics of Commute Time Embedding

Commute time embedding of some sinusoid is depicted in Fig. 2. The sinusoidal segment is composed of 700 samples, from which 676 patches are extracted from the sinusoidal signal, so that they may be maximally overlapped.

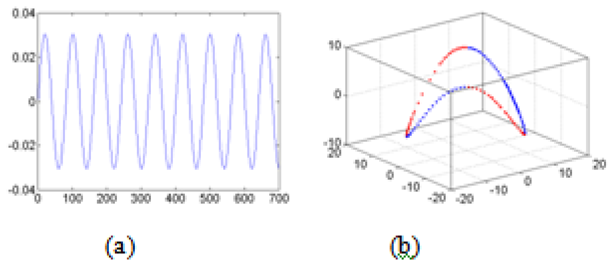


Figure 1. Commute time embedding of a sinusoid signal from which the patch graph is constructed. (a) Sinusoidal waveform (b) Its commute time embedding

In this figure, patches of lower variance are encoded with blue color, while patches of higher variance with red color. Throughout this paper, lower variance means it is less than the median of the distribution of the variances over all patches in the patch set, while higher variance is larger than the median, following the idea of Tayler, et al. [15]. That is, patches of lower variance correspond to the segments that contain smooth or low frequency components and higher variance patches are associated with high frequencies. Commute time embedding shows that each patch is mapped densely to generate a smooth curve inherent to the characteristics of the signal, as depicted in Fig. 1. Indeed, the parametrization of the set of patches according to the spectral graph based metrics can concentrate statistically different patches, which would otherwise be scattered in the space of patch. This concentration property can be exploited so that the set of patches can generate on the embedding subspace a geometric structure intrinsic to the signal, from which patches are extracted.It is known that commute time embedding results of periodic or quasi-periodic waveforms are represented as closed curves on the low dimensional Euclidean space, while those of aperiodic signals have the shape of open curves. Thus, commute time embedding can be used as a tool for geometrization of dataset, where some periodic phenomena are transformed into the geometries composed of topological circles or higher dimensional holes in some space.

4.2The Proposed Approximate Commute Time Embedding

Fig. 2 shows approximate commute time embedding results of the patch graph same as that of Fig. 1(b), generated using Choromanska's method and the proposed method.

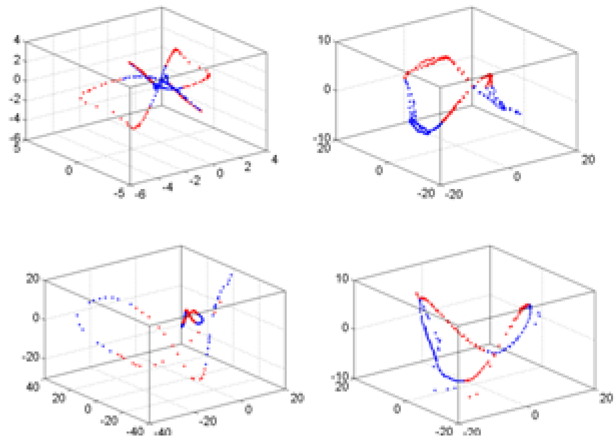


Figure 2. Approximate commute time embedding. The left column: Choromanska's method, the right column: our proposed method, Sampling rates: The first row (59.2%), the second row (88.7%).

As mentioned in the above section, severe distortion occurs on the embedding results of Choromanska's method, compared with those of our proposed method. Irregular scales at the embedding geometries, which are caused by the errors in approximating the eigenvalues of L_{sym}, degrade the overall performance. Note that the phase can be reversed because the embedding map consists of eigenvectors, as given in (8). Table 1 shows the five smallest eigenvalues of the L_{sym} for the patch graph of the sinusoid. According to (9), the smallest eigenvalue of L_{sym} for the connected graph should be zero and the largest one be less than 2. Choromanska's method, given in Table 2, does not satisfy the conditions. However, Table 3 shows that the approximate eigenvalues computed using our proposed method approach the true values as the number of samples increases as well as satisfy the condition given in (9).

Table 1. The five smallest eigenvalues of L_{sym}

	λ_1'	λ_2'	λ_3'	λ_4'	λ_5'
values	0	0.01	0.012	0.040	0.045

Table 2. The five smallest eigenvalues of L_{sym}, obtained using the Choromanska's method

#samples	λ_1'	λ_2'	λ_3'	λ_4'	λ_5'
400	0.637	0.650	0.711	0.729	0.745
600	0.119	0.131	0.136	0.166	0.175

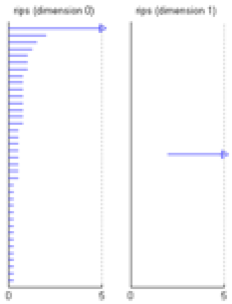
Table 3. The five smallest eigenvalues of L_{sym}, obtained using the proposed method

#samples	λ_1'	λ_2'	λ_3'	λ_4'	λ_5'
400	0.0	0.009	0.011	0.038	0.041
600	0.0	0.010	0.011	0.040	0.046

4.3Topological Analysis

As mentioned in the previous section, commute time embedding results of periodic or quasi-periodic waveforms are represented as closed curves on the low dimensional Euclidean space. In this work, persistent homology is employed to determine the topological invariants of the simplicial complexes constructed by randomly sampling the commute time embedding of the patch graph [18].

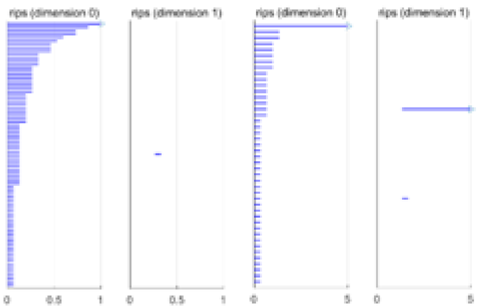
The embedding result of the sinusoid, depicted in Fig. 1, has the geometric structure composed of asingle topological circle.



(a) (b)

Figure 3.The persistence barcodes corresponding to the geometries given in Fig. 1(b). (a) The 0-dimensional barcode (b) The 1-dimensional barcode

Fig. 3shows the persistence barcodes corresponding to the Rips complexes within the range of $0 \le \varepsilon \le 5$, constructed from the geometries of a sinusoid, as given in Fig. 1(b). The computation of Betti numbers β_0 and β_1 is carried out by counting the long bars at the 0- and 1-dimensional persistence barcodes, respectively.It means that we can deduce the topological invariants $\beta_0 = 1$ and $\beta_1 = 1$ of the embedding geometry, given in Fig. 1(b), which can be interpreted as a connected component composed of a single circle. The persistence barcodes depicted in Fig. 4 represent the 0- and 1-dimensional persistent homology groups of the embedding geometries given in Fig.2(a) and (c). Fig. 4(a) shows a single short-lived 1-dimensional homology class, although one short-lived and one long-lived 1-dimensional homology classes are viewed in Fig. 4(b), where short bars are regarded as a topological noise.It means that the embedding geometries generated with the Choromanska's method get topologically correct only when sufficiently enough number of samples should be used to approximate the normalized graph Laplacian.



(a) (b)

Figure 4.The 0- and 1-dimensional persistence barcodes corresponding to the geometries given in (a) Fig. 2(a), (b) Fig. 2(c).

However, the geometries generated with the proposed method, as shown on the right column of Fig. 2, have almost the same topological structures as the original ones given in Fig. 1(b). It can be easily found by checking Fig. 5, which shows the 0- and 1-dimensional persistent barcodes corresponding to the geometries of Fig. 2(b) and Fig.2(d).

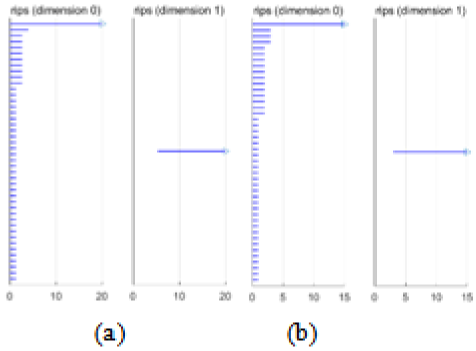


Figure 5.The 0- and 1-dimensional persistence barcodes corresponding to the geometries given in (a) Fig. 2(b), (b) Fig. 2(d).

It is clear that the homology groups of the embedding geometry corresponding to the sinusoid is determined to be:

$$\begin{aligned} H_0(K) &= \mathbf{Z} \\ H_1(K) &= \mathbf{Z}, \end{aligned} \tag{14}$$

which can be easily obtained by counting the long-lived bars on the persistence barcodes calculated from the approximate geometries generated with the proposed method, independently of the sampling rate.

V. CONCLUSION

This work has explored approximation of commute time embedding to reduce the computational burden for large dataset. Our proposed method is based on Nyström sampling method to compute approximate commute time embedding. The strength of our method is that it preserves the properties that the normalized graph Laplacian matrix is SPSD even though it is approximated via sampling process. In addition, the embedding geometries generated by our method preserve the topological properties of the original objects. Thus, our method can be applied efficiently to dimensionality reduction, which is very effective for visualization, as well as spectral clustering or pattern classification of large dataset.
As a future research, we would like to explore its application to pattern classification or manifold visualization in a geometric and topological way.

Acknowledgement
This work was supported by Hankuk University of Foreign Studies Research Fund.

REFERENCES

[1]. D. Achlioptas and P. Niyogi, “Database-friendly random projections: Johnson-Lindenstrauss with binary coins,” Journal of Computer and System Sciences 66, pp. 671-687, 2003.
[2]. M. Belkin and P. Niyogi, “Laplacian eigenmaps for dimensionality reduction and data representation,” Neural Computation15(6), pp.1373-1396, 2003.
[3]. M. Brito, E. Chavez, A. Quiroz, and J. Yukich, “Connectivity of the mutual k-nearest-neighbor graph in clustering and outlier detection,” Statistics and Probability Letter, 1997.
[4]. G. Carlson, “Topology and data,” Bulletin of the American Mathematical Society 46, no. 2, pp.255-308, 2009.
[5]. F. Chung, Spectral graph theory, American Mathematical Society, 1997.
[6]. A. Choromanska, T. Jebara, H. Kim, M. Mohan, and C. Monteleoni, “Fast Spectral clustering via the Nyström method,” ALT2013, LNAI 8139, pp.367-381, 2013.
[7]. P. Drineas and M.W. Mahoney, “On the Nyström method for approximating a Gram matrix for improved kernel-based learning,” Journal of Machine Learning Research 6, pp. 2153-2175, 2005.
[8]. H. Edelsbrunner, D. Letscher and A. Zomorodian, “Topological persistence and simplification,” Discrete Computational Geometry 28, pp. 511-533, 2002.
[9]. C. Fowlkes, S. Belongie, F. Chung, and J. Malik, “Spectral grouping using the Nyström method,” IEEE Trans. Pattern Anal. Mach. Intell., No. 26, pp.214-225, Jan. 2004.
[10]. A. Hatcher, Algebraic Topology. Cambridge University Press, 2002.
[11]. H. Hahn, “Application of persistent homology to topological analysis of waveform signals,” Journal of Mathematics, Statistics and Operations Research, Vol. 3, No. 2, pp. 77-83, 2016.
[12]. N. L. D. Khoa and S. Chawla, “Large scale spectral clustering using approximate commute time embedding,” arXiv:1111.4541v2 [cs.LG] 29 Feb. 2012.

[13]. H. Qiu and E.R. Hancock, "Clustering and embedding using commute times," IEEE Trans. PAMI, Vol. 29, No. 11, Nov., 2007.
[14]. D. A. Spieman and S. Teng, "Nearly linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear system," arXiv:cs/0607105v5 [cs.NA], 2012.
[15]. K.M. Taylor, "The geometry of signal and image patch-sets," PhD Thesis, University of Colorado, Boulder, Dept. of Applied Mathematics, June 2011.
[16]. J.B. Tenenbaum, V. deSilva, and J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction," Science, Vol. 290, pp.2319-2323, 2000.
[17]. C. K. I. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines." In Annual Advances in Neural Information Processing Systems 13: Proceeding of the 2000 Conference, pp. 682-688, 2001.
[18]. A. Zomorodian and G. Carlsson, "Computing persistent homology," Discrete Computational Geometry 33, pp. 249-274, 2005.

Hee Il Hahn" Visualization of Large Dataset Using Approximate Commute Time Embedding"International Journal Of Engineering Research And Development , vol. 15, no. 1, 2019, pp 22-30