International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

A Randomized Approach to Sparse Subspace Clustering using Spectral clustering

Samson Hansen Sackey[1], Samuel Nartey Kofie[1], Abdul Karim Armah[1]
[1]College of IoT Engineering, Hohai University, 213022, China
Email address: samsonsackey@yeah.net, kofano25@yahoo.com, armahabdulkarim1@yahoo.com

Abstract— *The steps taken to segment an in-motion object from its training set is a major feature in a lot of computer vision applications ranging from motion segmentation to image recognition. A random subspace is expressed into sparse representation called Randomized Sparse Subspace Clustering (RSSC), which is capable of intensifying the precision of the subspace cluster on real-life datasets that we tested on significantly. RSSC adopts the assumption that high-dimensional data actually lie on the low-dimensional manifold such that out-of-sample data could be grouped in the neighboring space learned from in-sample data. Experimental results show that RSSC is potential in clustering out-of-sample data.*

Keywords— *Subspace clustering, Sparse subspace, Face grouping, Randomized motion segmentation, Spectral clustering.*

I. INTRODUCTION

Clustering is a very dynamic and principal subject matter when it comes to machine learning, computer vision, signal and image processing, pattern recognition, bioinformatics etc. whose objective is to sort similar patterns into the same cluster by maximizing the inter-cluster dissimilarity and the intra-cluster similarity [1]. Generally, clustering uses a wide variety of techniques for exploratory data analysis with applications ranging from statistics, computer science, and biology to social sciences or psychology [2]. However, clustering is a segmentation phenomenon which does not use spatial information [3]. The sparse representation use vectors lying on a union of subspaces to cluster the data into distinct subspaces which clarifies the fact that each data point in a union of subspaces can constantly correspond with a linear or affine combination of all other points [4]. Sparsity is obtained for other points lying in the same subspace [5, 6]. This tolerates us to construct a similarity matrix, from which the segmentation of the data can be easily obtained using spectral clustering. Our methodology goes higher by posing many advantages over other algorithms.

Subspace clustering is a state-of-the-art problem where one is given points in a high-dimensional ambient space and compare by estimating them to a union of lower-dimensional linear subspaces [7]. Applications of subspace clustering include motion segmentation, face clustering, gene expression analysis, and system identification. Data points with the same identify assembled from the above applications (e.g., face images of a person under varying illumination conditions, feature points of a moving rigid object in a video series) lie on a low dimensional subspace, and the miscellaneous dataset can be modeled by unions of subspaces [8]. In subspace clustering, given the data from a union of subspaces, the objective is to compute the number of subspaces, their dimensions, the segmentation of the data and a basis for each subspace [9].

Consider the set $X \coloneqq X_1 U X_2 U \dots X_L$ containing N data points lying near a union of L subspaces of R^m. At this point, $X_l = \{x_i^{(l)}\}_{i=1}^{n_l}$, where $x_i^{(l)}$ is the ith of the n_l data points belonging to the d_l-dimensional subspace S_l. Assuming X is known, the subspace clustering problem that looks into partitioning the data points in X according to their subspace membership. The values for L, d_l, n_l, and S_l are completely unknown. We expect a subspace clustering algorithm to return a set { $\ddot{X}_1 \ddot{X}_2, \dots, \ddot{X}_{\dot{L}}$ }, where $\dot{L}$ is the estimated number of subspaces and where the sets $\ddot{X}_1$ are disjoint subsets of X such that $\ddot{X}_1 U \ddot{X}_2 U \dots, U \ddot{X}_{\dot{L}} = X$. We declare success if { $\ddot{X}_1, \ddot{X}_2, \dots, \ddot{X}_{\dot{L}}$ } = { $X_1, X_2, \dots, X_L$ }, which implies that $\dot{L} = L$.

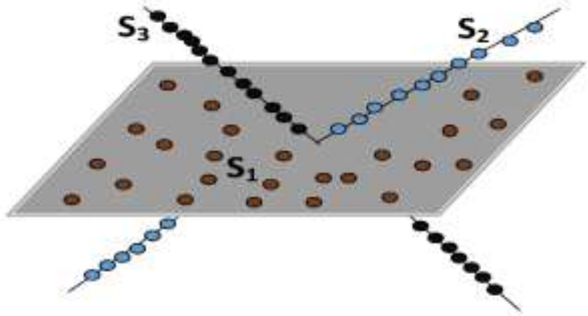

Fig. 1. A set of sample points in R^3 drawn from a union of subspaces: two lines (S_2 *and* S_3) and a plane (S_1)

The main contribution of this paper is to propose an approach that extends sparse subspace clustering (SSC) in a randomly clustered data. The two main requirements of this method are (i) a randomized based sparse subspace clustering was discussed to identify different sets of points that are good representation of a data set and (ii) the proposed technique was used to solve the motion segmentation and face clustering problems.

The paper is organized as follows: In Section 2, we briefly reviewed the related studies on SSC-like methods. In section 3, we present an approach that used SSC to make RSSC are more accessible to large data set. We investigated the application of RSSC to real world data sets in Section 4. Section 5 concludes the paper.

Samson Hansen Sackey, Samuel Nartey Kofie, and Abdul Karim Armah, "A Randomized Approach to Sparse Subspace Clustering using Spectral clustering," *International Research Journal of Advanced Engineering and Science*, Volume 4, Issue 3, pp. 276-280, 2019.

II. Related Study

Clustering is an unsupervised arrangement of arrays into sets (clusters). There are assorted methodologies to clustering data which can be defined with the advantage of the hierarchy below. There are numerous approaches that have been outlined for example algebraic methods, iterative methods, statistical methods and spectral clustering. The main idea behind Generalized Principal Component Analysis (GPCA), which is an algebraic method is to prove that one can fit a union of n subspaces with a set of polynomials of degree n, whose products at a given point provides a vector normal to the subspace containing that point. Factorization based algebraic approaches such as matrix factorization-based algorithms find an initial segmentation by thresholding the entries of a similarity matrix built from the factorization of the data matrix. K-planes and K-subspaces algorithms are generalized by the K-means algorithm from data distributed around cluster centers to data drawn from hyper planes and affine subspaces of any dimensions [15]. Random sample Consensus (RANSAC) is a statistical way of dealing with corrupted points with outliers numerically. It is a kind of fitting pattern to a group of points. The Multi-Stage Learning (MSL) algorithm is a statistical approach that can handle some classes of degenerate motions by refining the solution of SS using the Expectation Maximization algorithm (EM). The Agglomerative Lossy Compression (ALC) algorithm postulates that the data illustrates that of a mixture of degenerate Gaussians [14]. Probabilistic PCA (PPCA) can certainly be drawn-out of a multiplicative model for a union of subspaces by using a mixture of PPCA (MPPCA) model [13]. The local subspace affinity (LSA) algorithm proposed based on a linear projection and spectral clustering [12]. The main difference is that LSA fits a subspace locally around each projected point even though GPCA makes use of the slopes of a polynomial that is generally a fit to the projected data. Spectral Curvature Clustering (SCC) is a spectral clustering based methods that uses multiway similarities that capture the curvature of a collection of points within an affine subspace [31]. Sparse subspace clustering (SSC) is also centered on the knowledge of matching data points as a linear or affine combination of neighboring data points [10, 11]. On top of all, approaches formed on spectral clustering have been shown to achieve very good results for several applications in computer vision.

Liu et al. proposed partitioning framework whereby a data matrix can be transformed into a column blocks and then decomposes the original problem into parallel multivariate Lasso regression sub problems and sample wise operations [16]. In [17], a Laplacian structured representation model was proposed to enhance the representation based clustering techniques by means of introducing local feature similarity earlier information to monitor the encoding procedure, and then develop an efficient Alternating Direction Method of Multipliers (ADMM) algorithm for optimization. A new model to balance the sparsity and rank was presented in [18] which adopts the log-determinant function to control the rank of solutions. Zhan et al. improved spectral clustering by using Nyström sampling method [19]. Subspace clustering was discussed through learning in an adaptive graph affinity matrix [20]. In [21], the integration of self-representation and hypergraph fused and then extended to graph based spectral clustering based hypergraph. Abdolali et al. proposed a method which cannot only allow SSC to scale to large-scale data sets, but that it is also much more robust as it performs significantly better on noisy data and on data with nearby outliers and subspaces, while it is not likely to over segmentation [22]. High clustering accuracy and fast processing speed is more advantageous is subspace clustering [23]. The demonstration of a unique technique to eradicate the special effects of errors from the projection space (representation) was better fairly than from the input space [24, 25]. A similarity matrix was constructed based on locally linear embedding [32]. Bootstrap Algebraic Multigrid was applied to construct a set of vectors associated with graph Laplacian [26]. Tong et al in [27] proposed a new spectral clustering method based on local Principal Component Analysis (PCA) and connected graph decomposition. A unified similarity measure was proposed using SimRank to construct the fuzzy similarity matrix of the recognized graph which can assimilate structural and recognized similarities of nodes into a flexible weighted framework [28]. A hypergraph was proposed to describe the visualization of subspace structures [29]. Zhu et al. propose an iterative optimization method to adapt and adjust to each processes towards the goal of clustering performance, it however, enables the output of good clustering results [30].

III. A Modified Sparse Subspace Clustering

In this section, we initiate our approach i.e. RSSC representation by given enumerating clarifications.

The data matrix is denoted by $X \in R^{n \times N}$, where each column of X is a member of the union of L subspaces i.e. $S_1 \cup S_2 \cup \ldots \cup S_L$. There are N_l data samples in each ι subspace with $N_1 + N_2 + \cdots + N_L = N$. The data matrix X is considered to be a noisy type. We select $X^l \epsilon R^{n \times N_l}$ that is a member of S_l □ R^n with $X = [X^1, X^2, \ldots, X^L]$. We consider $\|\cdot\|$ to represent Euclidean norm (in terms of vectors) and spectral norm (in terms of matrices).

Approach: We solve the following optimization problem.

$$\min_Z \|Z\|_* + \lambda\|Z\|_1 + \beta tr(ZLZ^T)$$

$$s.t. \quad X = XZ, \qquad diag(Z) = 0. \tag{1}$$

Where Z corresponds to a random sparse coefficient matrix, L is an unnormalized laplacian, $\|\cdot\|_*$ represents a convex surrogate of the rank function, λ, β balances the terms in the objective function and $\|\cdot\|_1$ is the l_1 norm.

1. Calculate the coefficients matrix of Z using Equation (1).
2. The spectral clustering technique is then applied on the affinity matrix $|Z| + |Z^T|$ to find the segmentation of the data.

Algorithm 1 (SC)

Given a similarity matrix $S \in R^{n \times n}$, number k of clusters is to be constructed.

i. Construct a similarity graph. Let M be its weighted adjacency matrix.

ii. Compute the normalized Laplacian L_{sym}.

Samson Hansen Sackey, Samuel Nartey Kofie, and Abdul Karim Armah, "A Randomized Approach to Sparse Subspace Clustering using Spectral clustering," I*nternational Research Journal of Advanced Engineering and Science*, Volume 4, Issue 3, pp. 276-280, 2019.

iii. Compute the first k eigenvectors $u_1, ..., u_k$ of L_{sym}.
iv. Let $U \in R^{\,n\,x\,k}$ be the matrix containing the vectors $u_1, ...,$ u_k as columns.
v. From the matrix $T \in R^{n\,x\,k}$ from U by normalizing the rows to norm that is $t_{ij} = u_{ij}/(\sum_k u_{ik}^2)^{\frac{1}{2}}$
vi. For $i = 1, ..., n$, let $y_i \epsilon R^k$ be the vector corresponding to the ith row of T.
vii. Cluster the points (y_i), $i = 1, ..., n$ with the k-means algorithm into clusters $C_1, ..., C_k$.

Therefore, W_l is the similarity matrix of data points in S_l. The matrix for M is defined below.

$$M = \begin{bmatrix} M_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & M_n \end{bmatrix} \Gamma. \quad (2)$$

Algorithm 2 (RSSC)

Input: Data matrix X and preferred number of subspaces k. Under conditions: $\lambda, \gamma, \beta, \rho$.
1. Solve the problem in Equation (1) to obtain Z.
2. Knowing Z, solve the spectral clustering to obtain segmented matrix

Output: Segmented matrix

We solve Z by working out the problem in equation (1) using the Alternating Direction Method of Multipliers s.t. $X = XZ$, $diag(Z) = 0$,

$$Z^{(t+1)} = \underset{Z}{argmin} \frac{1}{\mu^{(t)}} ||Z||_1 + \frac{1}{2} \left|\left|Z \;\; diag(Z)\; U^{(t)}\right|\right|_F^2 \quad (3)$$

$$Z^{(t+1)} = \tilde{Z}^{(t+1)} \qquad diag\left(\tilde{Z}^{(t+1)}\right) \quad (4)$$

$$\tilde{Z}_{ij}^{(t+1)} = S_{\frac{1}{\mu^{(t)}}}(1 + \alpha \Theta_{ij}(U_{ij}^{(t)}) \quad (5)$$

Where $S_\tau(\cdot)$ is the shrinkage thresholding operator, $\mu > 0$ is a parameter, $Y^{(i)}$ is a matrix multiplier and U^t is a threshold matrix entry which depends on Θ_{ij}. In the experiment, $\gamma\ and\ \rho$ are 0.008 and 1.5 respectively.

Given Z, we solve the problem $\underset{Z}{min}\ \beta\ trace(Z^T LZ)\ s.t. Z \epsilon \bar{Z}$ with spectral clustering by singular value decomposition.

Observations: (i) The matrices of the sparse coefficient Z deals effectively with the norm of the data points and the problem formulated can also deal perfectly with noise.
(ii) Fully random model is employed when connection is found between points in different subspaces i.e. for all subspaces are of the same dimension d and there are some $pd + 1$ points in each subspace.

IV. EXPERIMETAL RESULTS

In this section, we determine the effectiveness of the proposed technique by applying the clustering test on some expanded dataset. The experiment is carried out using Matlab 2014a on a PC with an Intel(R) Core(TM) i5-3317U CPU and 6144MB of RAM.

Basically, motion segmentation focuses on classifying repositioning objects in a video sequence. It is the grounds on which all other computer vision actions arises such as metrology, robotics, inspection, video surveillance, video indexing, traffic monitoring, structure from motion, and many other applications.

Suppose that we have tracked N feature points over F frames in a video sequence, $\{x_{ij}\}$, where i =1, · · ·, N and j = 1· · ·, F. Each feature trajectory $y_i \in R^{2f}$ is obtained by stacking the feature points in the video, i.e. $y_i^T = [x_{1i}^T, x_{2i}^T, ..., x_{Fi}^T]$. The trajectories of the overall rigid motion under affine projection spans a $4n$ -dimensional linear subspace.

Our main aim is to set apart feature trajectories according to their motions. However, feature trajectories of n rigid motions lie in a union of n dimensional subspaces of R^{2F}. Therefore, the problem of clustering affine subspace in different motion can be associated to the problem of clustering trajectories.

Fig. 2. Showing the background frame without the moving object by low rank method.

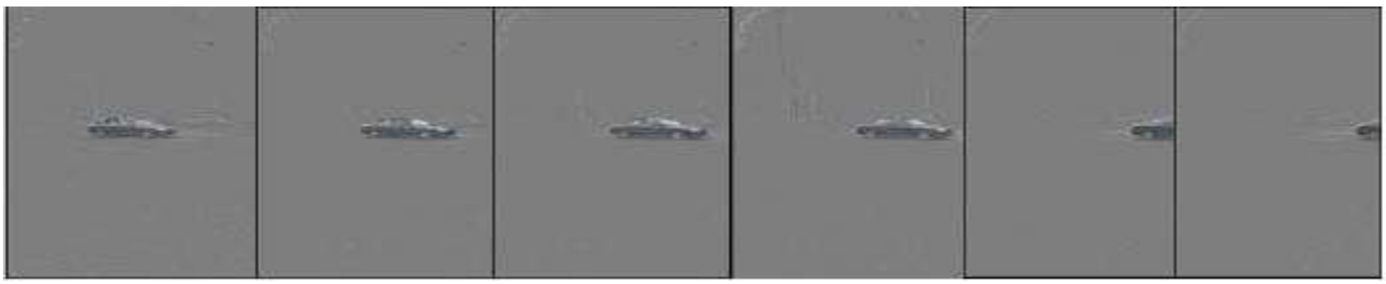

Fig. 3. Given the moving feature trajectory through sparse approach.

Samson Hansen Sackey, Samuel Nartey Kofie, and Abdul Karim Armah, "A Randomized Approach to Sparse Subspace Clustering using Spectral clustering," I*nternational Research Journal of Advanced Engineering and Science*, Volume 4, Issue 3, pp. 276-280, 2019.

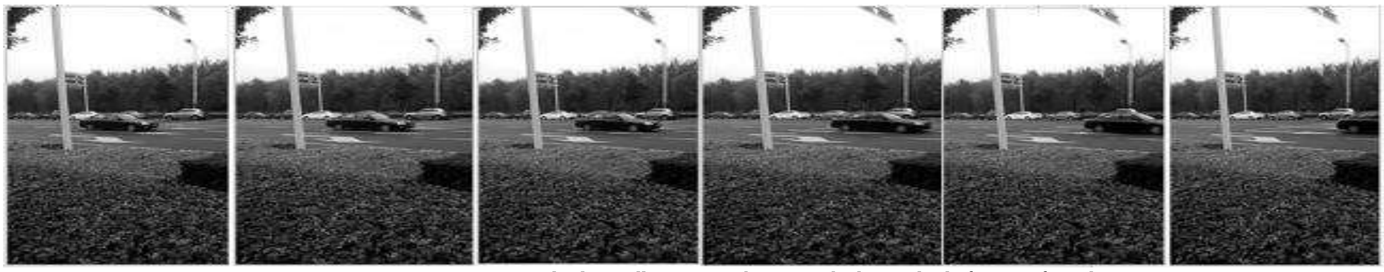

Fig. 4. Motion segmentation: Multiple rigidly moving objects tracked in multiple frames of a video.

The dataset consist of 442 video sequences of 2 or 3 motions. In addition, each video sequence can be well modeled as data points that almost perfectly lie in a union of linear subspaces of dimension of 4. The samples frames are chosen to be equally spaced and spread out as possible. Our algorithm is applied on the motion dataset. However, the dimension is reduced from m dimensions to $4n$ because the rank of the dataset is above bound. Therefore, the purpose is to cluster these trajectories with their respective motion subspaces.

Fig. 5. Separated feature trajectories according to the moving objects.

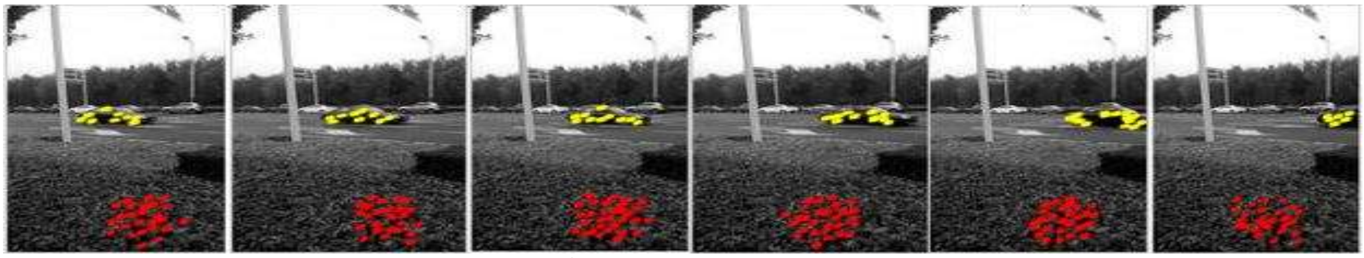

Fig. 6. Separated feature trajectories showing clear distinction.

Considering the problem of clustering images of different faces under variable lighting conditions lie near 9-dimensional linear subspaces. A set of images consisting of different faces are compared to the partitioning of a collection of images with low dimensional subspaces.

A data set consisting of cropped frontal face images of 24 subjects with different facial expressions. The original images have a resolution of 180 x 100 pixels i.e. the images are points in more than 10,000 dimensional ambient space. To reduce the computational cost and the memory requirements of algorithms, we down sample the images to 53 x 76 pixels and treat each 4028D vectorized image as a data point. The results shows that the studied algorithm performs almost better than others.

Fig. 7. Representation of face images of multiple subjects.

In this experiment, we randomly select few images to construct the set. However, the image patches that are used to construct the set come from a selected images, the unpicked images are not included. In processing, grayscale images are used because the dimension of the feature data is reduced as compared to the original color images. If the subspaces corresponding to different persons are extremely close to each other, it will be very hard to cluster.

Fig. 8. Distinguished images belonging to same subjects.

V. Conclusion

Conclusively, we have presented a concise approach to subspace clustering based on sparse representation. RSSC constructs a sparse similarity graph for spectral clustering by using l_1 minimization based coefficients. Since the representation of spectral clustering algorithms has become more and more popular owing to its effectiveness, we studied a simple but effective highly researched spectral clustering method. The experiment was implemented and it showed how effective and efficiency our method is compared to other approaches. As soon as the similarity grid is selected, the algorithm can then run for a number of times. However, spectral clustering cannot detect correctly a cluster in a given data. Nonetheless, it is regarded as one of the strong tool which can yield suitable outcomes if used meticulously.

References

[1] C. D. Larose, D. T. Larose, "Clustering", Chapter in book: Data Science Using Python and R, DOI: 10.1002/9781119526865.ch10, March, 2019.
[2] R. Peng, H. Sun, L. Zanetti, "Partitioning Well-Clustered Graphs: Spectral Clustering Works", SIAM Journal on Computing 46(2):710-743, 2017.
[3] P. Venkataraman, "Natural Clustering of Shaped Clusters", Research, DOI: 10.13140/RG.2.2.10150.98886, February, 2019.
[4] E. Arias-Castro, X. Pu, "A Simple Approach to Sparse Clustering", Computational Statistics & Data Analysis 105, 2016.
[5] P. McCullough, N. G. Polson, "Statistical Sparsity", Biometrika 105(4):797-814, 2018.
[6] L. Shen, B. Suter, E. Tripp, "Structured Sparsity Promoting Functions", Preprint, arXiv:1809.06777v1, 2018.
[7] J. Wang, A. Suzuki, L. Xu, Feng Tian, L. Yang, K. Yamanishi, "Orderly Subspace Clustering", Conference: AAAI 2019,Hawaii, February, 2019.
[8] B. Kelbar, "Subspace Clustering-A survey", Chapter in book: Advances in intelligent Systems and Computing, Springer, Singapore, 2018.
[9] K. Niu, Z. Gao, H. Jiao, X. Qiao, Y. Zhao, "Subspace Clustering for Vector Clusters", Journal of Internet Technology 18(1):87-94DOI: 10.6138/JIT.2017.18.1.20140501, 2017.
[10] E. Elhamifar, R. Vidal, "Clustering disjoint subspaces via sparse representation", In Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, pp. 1926–1929, 2010.
[11] E. Elhamifar, R. Vidal, "Sparse Subspace Clustering: Algorithm, Theory, and Applications", Arxiv preprint arXiv: 1203.1005, 2012.
[12] S. J. Peng, X. Liu, Z. Cui, Z. Xie, D. Chen, "Automatic Motion Capture Data Denoising via Filtered Local Subspace Affinity and Low Rank Approximation", Proceedings of the 2013 International Conference on Computer-Aided Design and Computer Graphics, 2013.
[13] C. Bouveyron, P. Latouche, P. A. Mattei, "Bayesian Variable Selection for Globally Sparse Probabilistic PCA", Electronic Journal of Statistics 12(2), 2016.
[14] T. Zhang, A. Szlam, Y. Wang, G. Lerman, "Hybrid linear modeling via local best-fit flats", In Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1927–1934, 2010.
[15] T. Zhang, A. Szlam, G. Lerman, "Median k-flats for hybrid linear modeling with many outliers", in Workshop on Subspace Methods, 2009.
[16] B. Liu, X. T. Yuan, Y. Yu, Q. Liu, D. N. Metaxas, "Parallel Sparse Subspace Clustering via Joint Sample and Parameter Blockwise Partition", ACM Transactions on Embedded Computing Systems 16(3):1-17DOI: 10.1145/3063316, 2017.
[17] W. Chen, E. Zhang, Z. Zhang, "A Laplacian structured representation model in subspace clustering for enhanced motion segmentation" Neurocomputing DOI:10.1016/j.neucom.12.123, 2016.
[18] Y. Xia, Z. Zhang, "Rank–sparsity balanced representation for subspace clustering" Machine Vision and Applications 29(1), 2018.
[19] Q. Zhan, Y. Mao, "Improved spectral clustering based on Nyström method", IEEE Multimedia Tools and Applications 76(2), 2017.
[20] Z. Wu, M. Yin, Y. Zhou, X. Fang, S. Xie, "Robust Spectral Subspace Clustering Based on Least Square Regression" IEEE Neural Processing Letters, 2017.
[21] Y. Li, S. Zhang, D. Cheng, W. He, G. Wen, Q. Xie, "Spectral clustering based on hypergraph and self-re-presentation" Multimedia Tools and Applications 76(16)DOI: 10.1007/s11042-016-4131-6, 2016.
[22] M. Abdolali, N. Gillis, R. Mohammed, "Scalable and Robust Sparse Subspace Clustering Using Randomized Clustering and Multilayer Graphs", arXiv: 1802.07648v1 [cs.CV] Feb, 2018.
[23] C. Zhao, W. L. Hwang, C.L. Lin, W. Chen, "Greedy Orthogonal Matching Pursuit for Subspace Clustering to Improve Graph Connectivity", Information Sciences 459, 2018.
[24] X. Peng, Z. Yu, Z. Yi, H. Tang, "Constructing the L2-Graph for Robust Subspace Learning and Subspace Clustering", IEEE Transactions on Cybernetics 47(4):1-14, 2016.
[25] A. K. Armah, M. K. Ansong, S. H. Sackey, N. Bulgan, "A Novel Approach Using Convolutional Neural Network to Reconstruct Image resolution", International Journal of Scientific and Research Publications (IJSRP), vol. 9, issue 5, 2019.
[26] P. D'Ambra, L. Cutillo, P. S. Vassilevski, "Bootstrap AMG for spectral clustering: BAMG for spectral clustering", DOI: 10.1002/cmm4.1020, March, 2019.
[27] T. Tong, X. Zhu, T. Du, "Connected graph decomposition for spectral clustering", Multimedia Tools and Applications, DOI: 10.1007/s11042-018-6643-8, 2018.
[28] S. Hechmi, A. Gallas, E. Zagrouba, "Enhanced Similarity Measure for Sparse Subspace Clustering Method", International Work-Conference on Artificial Neural Networks, 2017.
[29] J. Xia, G. Jiang, Y. Zhang, R. Li, W. Chen, "Visual Subspace Clustering Based on Dimension Relevance", Journal of Visual Languages & Computing 41DOI: 10.1016/j.jvlc.2017.05.003, 2017.
[30] X. Zhu, S. Zhang, Y. Li, Y. Fang, "Low-rank Sparse Subspace for Spectral Clustering", IEEE Transactions on Knowledge and Data Engineering PP(99):1-1, 2018.
[31] S. T. Wierzchon, M. A. Klopotek, "Spectral Clustering", doi: 10.1007/978-3-319-69308-8_5, In book: Modern Algorithms of Cluster Analysis, 2018.
[32] P. Shu-Xia, S. Wang Jie, "A spectral Clustering Based on Locally Linear Embedding", DOI: 10.2174/2352096509666160823112400, February, 2017.

Samson Hansen Sackey, Samuel Nartey Kofie, and Abdul Karim Armah, "A Randomized Approach to Sparse Subspace Clustering using Spectral clustering," I*nternational Research Journal of Advanced Engineering and Science*, Volume 4, Issue 3, pp. 276-280, 2019.