Invited Talk

Department of Computer Science and Engineering

Indian Institute of Technology Kanpur

Venue: RM 101

Time: 11:30 am -12:30 pm, March 20, 2019

Geometric Component Analysis and its Applications to Data Analysis*

Amir Averbuch_

School of Computer Science, Tel Aviv University

Department of Mathematics, Toronto University, Canada

Email: amir@math.tau.ac.il

Dimensionality reduction methods are designed to overcome the `curse of dimensionality' phenomenon that makes the analysis of high dimensional big data difficult. Many of these methods are based on principal component analysis, which is statistically driven, and do not directly address the geometry of the data. Thus, machine learning tasks, such as classification and anomaly detection, may be affected for the bad. This work provides a dictionary-based framework for geometrically driven data analysis, both for linear and diffusion (non-linear) geometries, that includes dimensionality reduction, out-of-sample extension and anomaly detection. The geometry of the data is preserved up to a user-defined distortion rate. In addition, a subset of landmark data points, known as dictionary, is identified by the presented algorithm. The performance of the method is demonstrated on both synthetic and real-world datasets. It achieves good results for unsupervised learning such as cyber tasks.

***Work done with** Amit Bermanis and Moshe Salhov