Please attend the proposal defense of Rexford Aboagye, a Master's Degree candidate in Computer Science and an advisee of Dr. Gary Holness. His proposal defense will be on:

**Date: 15-Nov-2016**
**Time: 11:00 am-12:00 noon**
**Location: Mishoe Science Center North Room 331**

Title: **Balancing Algorithmic and Architectural Considerations for Dimensionality Reduction on GPUs**

Abstract:

The raw input format for data sets derived from natural sources (images, audio, motion) consists of signal and noise . Data sets also have redundant features. An image for example, might have random variation of brightness or color information, considered as noise due to bad sensor or circuitry of a scanner of digital camera. Noise and redundant features increases the dimensionality of data sets. Computational learning theory tells us that construction of feasible learning models is dependent, in part, on the input representation's dimensionality. For practical problems, this generally requires the experimenter to select those principle dimensions that result in the ability to draw discriminations. An Example of such a dimensionality reduction technique is Principal Component Analysis (PCA).

Principal component analysis (PCA) computes the most meaningful representation to re-express a noisy, garbled data set. Principal component analysis (PCA) has long been a robust and efficient technique for dimensionality reduction of data set n X d. It has a wide range of usage in several fields of scientific research e.g., data clustering, machine learning, pattern recognition, image analysis, information retrieval, signal processing, bioinformatics etc. However, as the size of d increases, the computation of PCA becomes very slow or incomputable on traditional general purpose serial devices.

Harnessing the tremendous potential of parallel computing and Graphical Processing Units (GPU), We propose a real-time solution for laying out PCA on GPUs using Open Computing Language (OpenCL). Our technique measures the cause and effect of latency and bandwidth due to data movement and space size in Heterogeneous Computing by remodeling the PCA algorithm to run in parallel.

Given a set of points drawn from a smooth manifold in an abstract feature space, our technique is capable of determining the structure of the surface and laying out the algorithmic processes of computing a parallel mean, transpose and correlation coefficient matrix by balancing the trade-off between algorithmic design and architectural structure.

We demonstrate the performance of our method by reducing the dimensionality of a real world laser induced spectroscopy dataset. The system is tested on a double Tesla C870 GPU card and the code is written in python(pyOpenCL).