Defense Announcement

Data Fusion for Multi-source Data Classification

Yuhang Zhang

Degree: **PhD, Electrical Engineering**

Date: 11/16/2015 Time: 11:00 am

Location: ECE Conference Room, N328-D

Committee Chair: Dr. Saurabh Prasad, ECE

Committee Members: Dr. Jose L. Contreras-Vidal,
Dr. Badri Roysam, ECE
Dr. Demetrio Labate, Math
Dr. Melba Crawford, Purdue

Multi-source data, either from different sensors or disparate features extracted from the same sensor, are often valuable for data analysis due to their potential for providing complementary information. Effective fusion of information from such multi-source data is critical to enhanced and robust interpretation about the underlying classification problem. Nevertheless, multi-source data also bring unique challenges for data processing, e.g., high-dimensional features, lack of compact representation, and insufficient quantity of labeled data. To make the most use of multi-source data and to address the above challenges, in this research, we develop and employ data fusion algorithms for diverse practical applications.

We develop a mixture-of-kernels approach for data fusion, and demonstrate its efficacy at fusion of multi-source data in the kernel space. In the proposed approach, each source of data is represented by a dedicated kernel --- one can then learn a classifier (or an "optimal" feature subspace) by optimizing the kernel parameters for maximum discriminative potential. We analyze this property and apply it to infer the relative importance of different sources of information in a brain machine interface (BMI) application. Additionally, to save the labor of labeling a large quantity of samples in real world remote sensing applications, an ensemble based multiple kernel active learning framework is proposed to effectively select important unlabeled samples from multi-source data for classification. We also propose a multi-source feature extraction method based on a composite kernel mapping, to project the multi-source data to a lower dimensional subspace for effective feature fusion.

Finally, to effectively represent multi-source data in a compact and robust manner, we propose a joint sparse representation model with adaptive locality weights for classification. By adapting the penalty on individual atoms in the dictionary, we show that one can achieve better signal representation and reduce estimation errors. Further, we also develop a kernel variant of the proposed fusion framework, which is conceptually consistent and aligned with the mixture-of-kernels approach developed previously. The proposed algorithms are validated on multiple datasets in two active research areas ---remote sensing and brain machine interface, and have shown significant improvements on the overall classification performance.