Abstract

The field of computer vision has recently witnessed remarkable progress, due mainly to visual data availability and machine learning advances. Modeling the visual data is challenging due to several factors, such as loss of information while projecting 3D world to 2D plain, high dimensionality of the visual data, and existence of nuisance parameters such as occlusion, clutter, illumination and noise. In this dissertation, we focus on modeling the inter and intra image manifold variability.

The dissertation shows that modeling the image manifold helps to achieve recognition invariance and perform robust regression within the manifold. It leverages the power of Homeomorphic Manifold Analysis (HMA) framework to utilize the known topological information about data manifolds. HMA builds mappings from a conceptual space to the feature space. These mappings are based on topological homeomorphism between points in the two spaces.

The dissertation extends this framework, applied to several applications such as human motion analysis and object recognition in conjunction with pose estimation. We propose Manifold-KPLS, a discriminative nonlinear model for recognition of motion sequences, applied to visual speech recognition. To tackle recognition and pose estimation from single test image, we propose a bi-nonlinear generative framework. The dissertation uses iterative inference techniques to find the optimal category and viewpoint that match a given test image. To speed up the inference, the dissertation proposes a feedforward model, which is more efficient and more accurate for solving the same problem. On the other hand, the dissertation leverages the manifold analysis to propose quantitative measurements for building a CNN variant for simultaneously solving object recognition and pose estimation.