Estimation of Manifolds from Point Clouds: Building Models from Data

Abstract:

A common observation in data-driven applications is that high dimensional data has a low intrinsic dimension, at least locally. Thus, when one wishes to work with data that is not governed by a clear set of equations, but still wishes to perform statistical or other scientific analysis, an optional model is the assumption of an underlying manifold from which the data was sampled. This manifold, however, is not given explicitly but we can obtain samples of it (i.e., the individual data points). In this talk, we will consider the mathematical problem of estimating manifolds from a finite set of samples, possibly recorded with noise. Using a Manifold Moving Least-Squares approach we provide an approximant manifold with high approximation order in both the Hausdorff sense as well as in the Riemannian metric (i.e., a nearly isometry). In the case of bounded noise, we present an algorithm that is guaranteed to converge in probability when the number of samples tends to infinity. The motivation for this work is based on the analysis of the evolution of shapes through time (e.g., handwriting or primates' teeth) and we will show how this framework can be utilized to answer scientific questions in paleontology and archaeology.