On the Application of Heat Diffusion across a Manifold for Dimensionality Reduction

Piwinski, John (School: BASIS San Antonio Shavano Campus)

Dimensionality reduction techniques are becoming widely used in academia, business, and machine learning to reduce the time required to process a higher dimensional dataset. In this context, a manifold serves as a lower dimensional space which is embedded in a higher dimensional space as a surface. The surface of the manifold often possesses curvature which increases the difficulty of finding the distance across the manifold between two points on the manifold. As a result, many existing dimensionality reduction techniques assume 1) the geodesic distances and the Euclidean distances are equivalent, 2) assume points are uniformly distributed over the manifold, or 3) entirely discard the presence of a manifold. These simplifications lead to distortions in the outputted lower dimensional rendition of the data as they often failed to properly preserve the distances between points on the manifold and the curvature of the manifold. A new dimensionality reduction technique, Heat Diffusion dimensionality Reduction (HDR), was designed to simulate heat diffusion across the manifold to better approximate the distances between higher dimensional points on the manifold. Time of execution and quality of low dimensional outputs for reduction of a three dimensional Mobius strip into two dimensions were compared between HDR, Isomap, Principal Component Analysis (PCA), and t-Distributed Stochastic Neighborhood Embedding (t-SNE). Although HDR required more time to reduce data than PCA and Isomap, HDR performed 6.3 seconds faster than t-SNE (high perplexity) and best preserved the curvature of the manifold compared to all the other techniques.

Awards Won:

Third Award of $1,000

National Security Agency Research Directorate : Second Place Award "Mathematics"$750

American Mathematical Society: Certificate of Honorable Mention