Manifold-Based Approaches for Improved Classification

Mark A. Davenport, Chinmay Hegde
Richard G. Baraniuk
Rice University

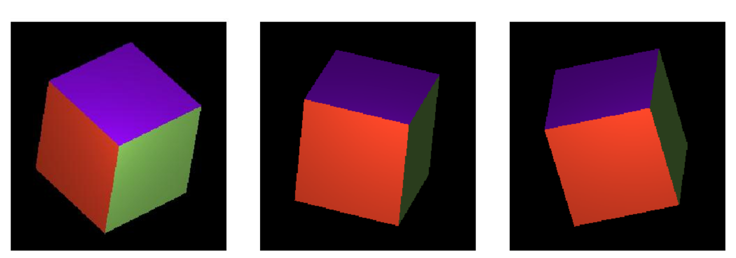
Michael B. Wakin
University of Michigan

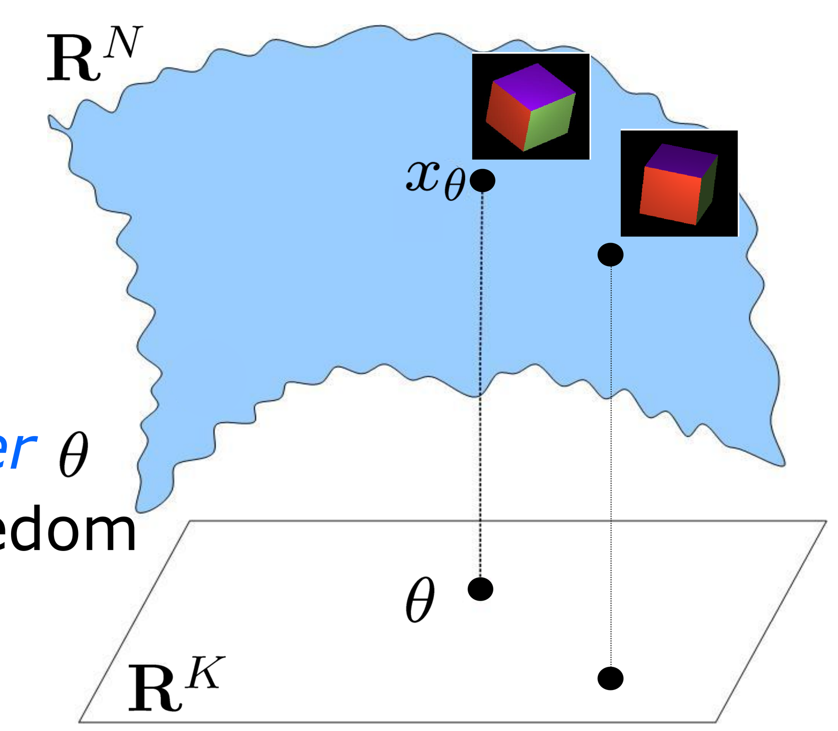
Manifold models for classification

- **Manifold models aid in overcoming the "curse of dimensionality" by providing a low-dimensional model for high-dimensional data**
- **Classification algorithms can be designed to exploit these models**

Manifold models

Many high-dimensional signal ensembles possess intrinsic low-dimensional geometric structure



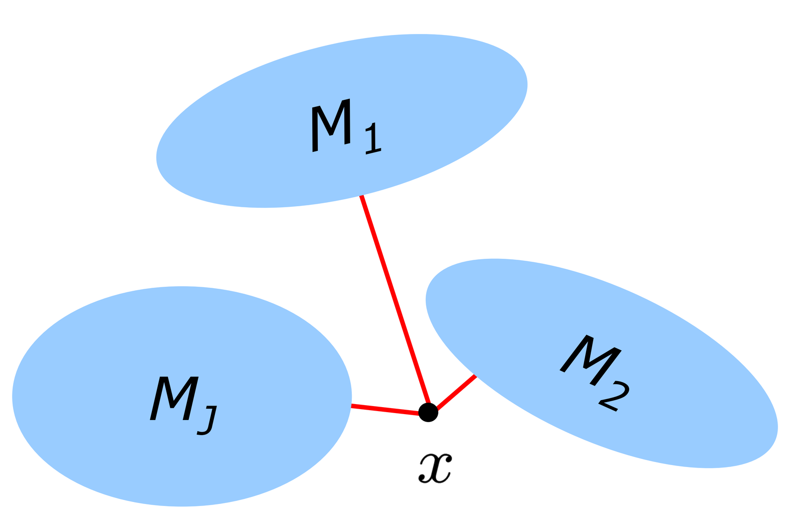


K-dimensional *parameter* θ captures degrees of freedom in signal $x_\theta \in \mathbf{R}^N$

Matched filters and manifolds

The matched filter can be viewed as a "nearest manifold" classifier

$$\mathcal{H}_1: \quad x = \mathcal{T}_{\theta_1} s_1 + n$$
$$\mathcal{H}_2: \quad x = \mathcal{T}_{\theta_2} s_2 + n$$
$$\vdots$$
$$\mathcal{H}_J: \quad x = \mathcal{T}_{\theta_J} s_J + n$$



$$\min_{j, \widehat{\theta}_j} \| x - \mathcal{T}_{\widehat{\theta}_j} s_j \|_2$$

Two stage approach:

- find ML estimate of parameter for each manifold
- classify according to which manifold is closest

Topology-aware classification

- **Number of projections required is linear in the intrinsic dimension K and only logarithmic in the ambient dimension N**
- **Bounds depend on manifold parameters like volume, curvature**

Classification using multiscale manifold navigation

Manifolds generated by images with sharp edges are nowhere differentiable

- parameter estimation for such manifolds becomes unstable
- exploit the *multiscale structure* of such manifolds using Newton's method and nested smoothing kernels
- *model aware* classification

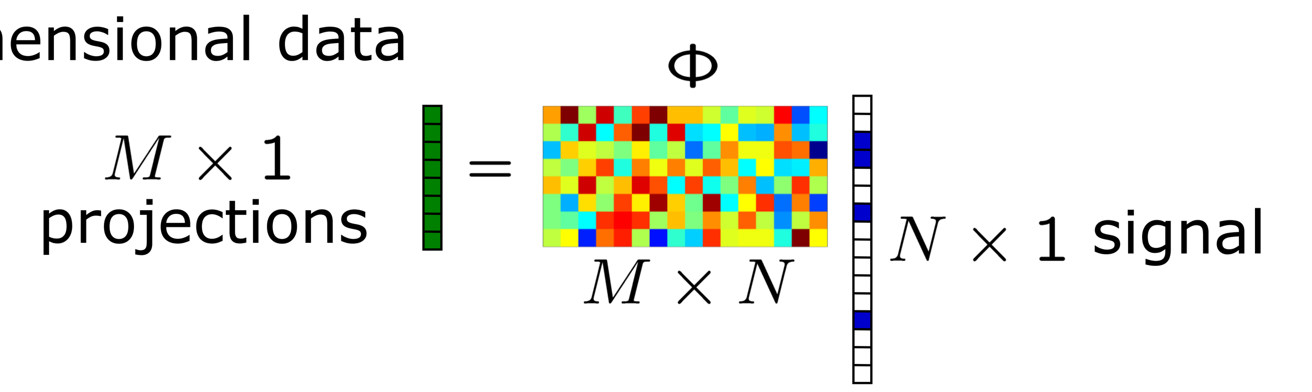
Classification using manifold learning

If we do not explicitly know the manifolds, we must learn the manifolds from training data

- training data is often *coarsely* sampled
- exploit manifold structure in *testing data* in addition to training data
- *model blind* classification

The random projection method

Compute *random linear measurements* of high-dimensional data

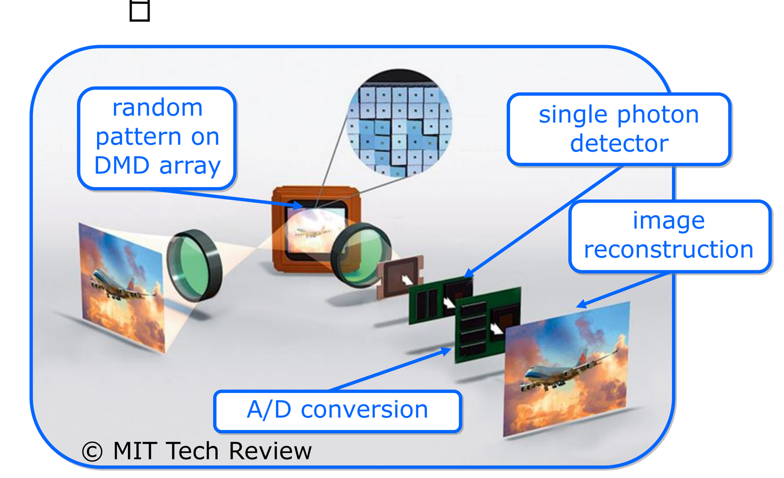


Let $\mathcal{M} \subset \mathbf{R}^N$ be compact, K-dimensional manifold. If Φ is an $M \times N$ random orthoprojector with

$$M = O(K \log(N)/\epsilon^2)$$

then for every pair $x, y \in \mathcal{M}$

$$1 - \epsilon \leq \frac{\|\Phi(x-y)\|}{\|x-y\|} \leq 1 + \epsilon$$



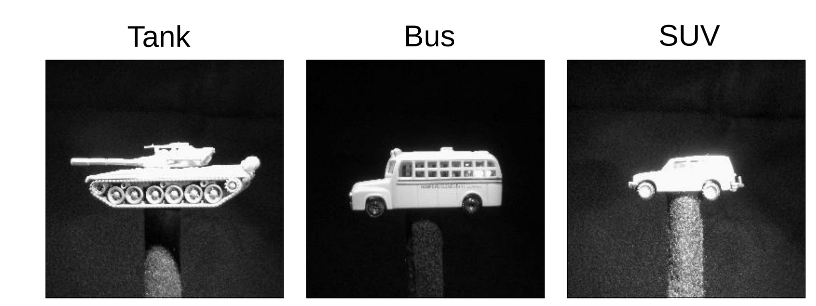
Advantages and applications

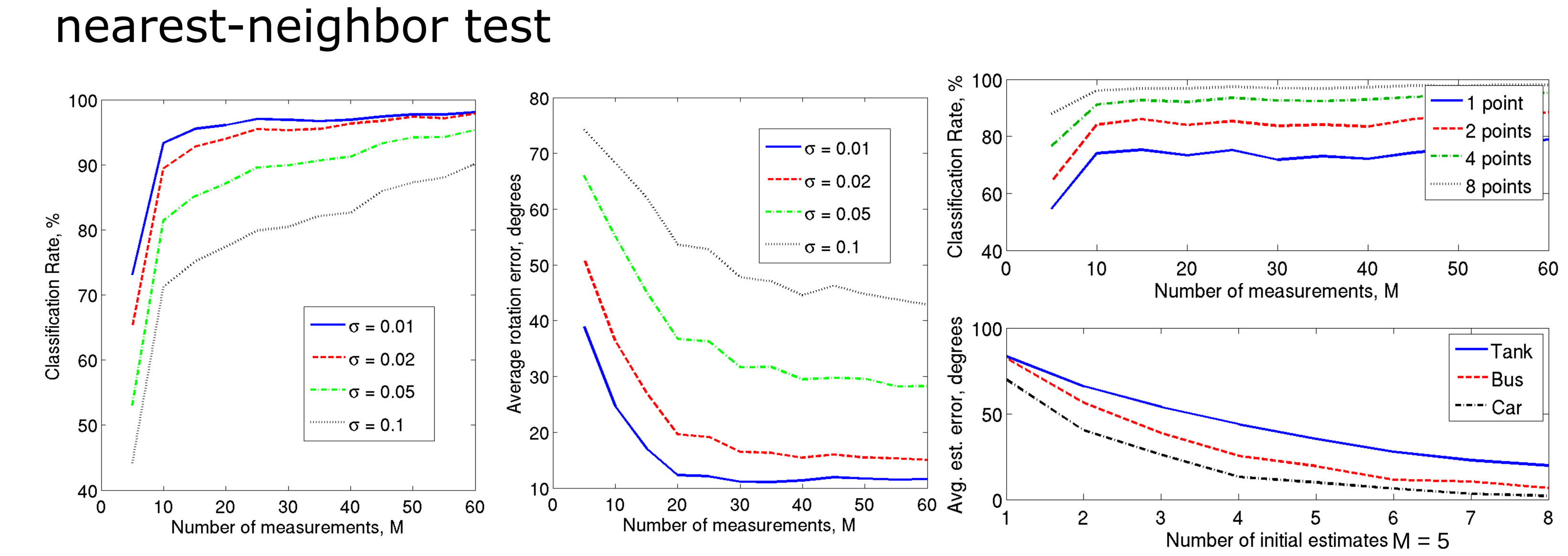
- **Bounds are pessimistic, difficult to compute explicitly**
- **Solution: start small, progressively acquire more measurements**

Experiments

Multiscale manifold navigation

- 3 image classes imaged using single-pixel camera
  - rotations 2º, 4º, …, 360º
  - binary random measurements
  - 5 regularization kernels
- Estimate rotation using multiscale projections
- Identify most likely class using nearest-neighbor test





Classification with manifold learning

- 2 image classes
  - disc or square of fixed radius
  - unknown shift
  - random measurements
- Training data coarsely sampled from the two manifolds
- Append a new batch of data to each of the two training data sets, run ISOMAP, and classify according to which yields lower residual variance