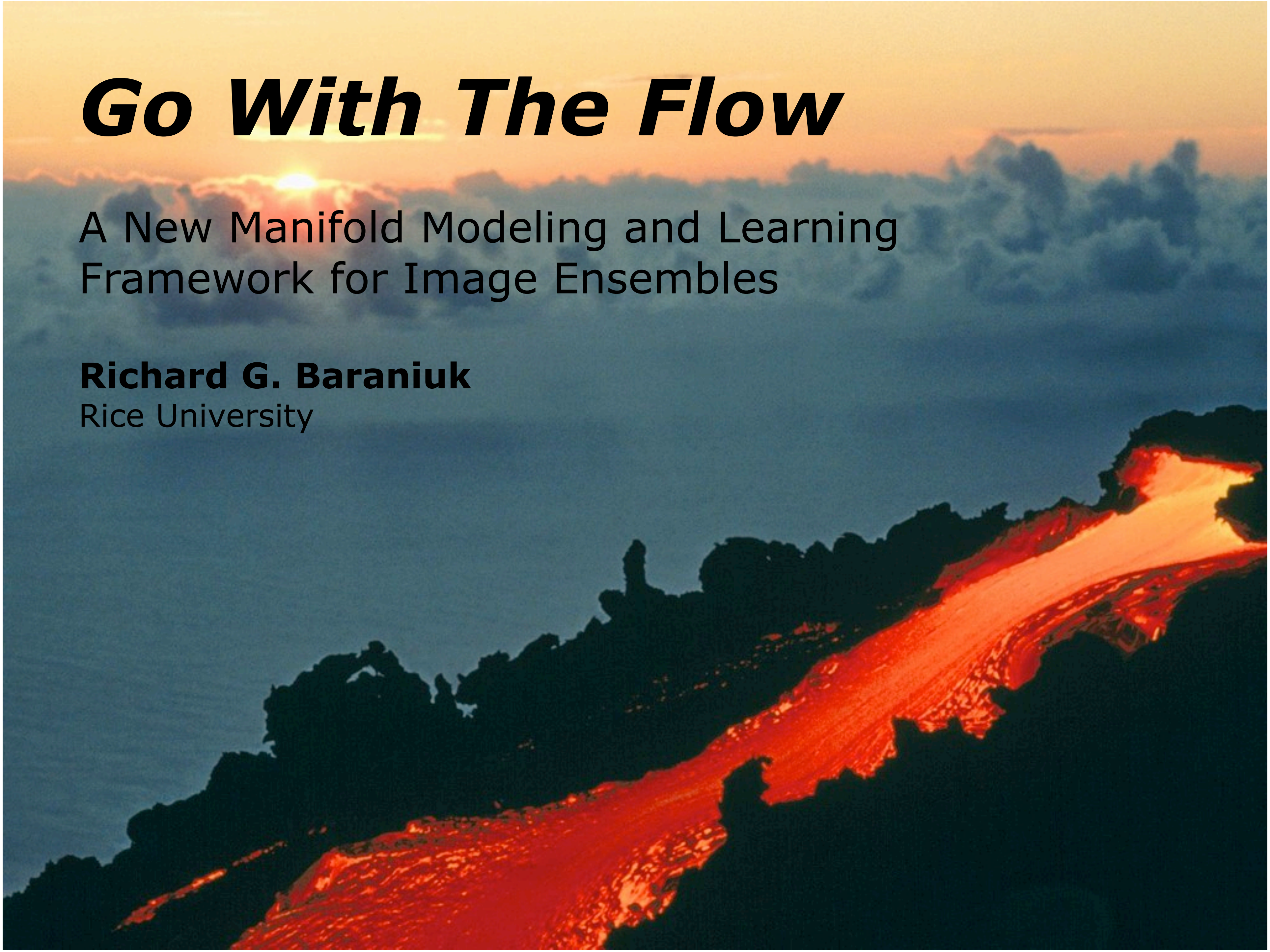
Go With The Flow

A New Manifold Modeling and Learning Framework for Image Ensembles

Richard G. Baraniuk
Rice University



Go With The Flow

A New Manifold Modeling and Learning Framework for Image Ensembles

Richard G. Baraniuk
Rice University

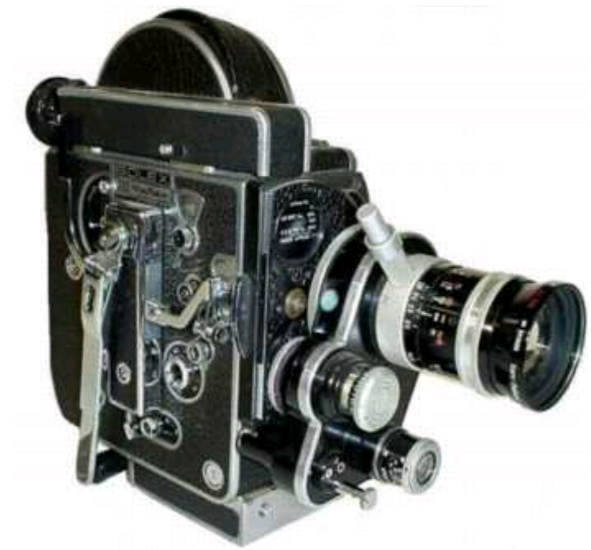
PRAIRIE SWINE CENTRE
Research Profits Everyone

Aswin Sankaranarayanan

Chinmay Hegde

Sriram Nagraj

Digital Sensing Revolution



Pressure is on Digital Sensors

- Success of digital data acquisition is placing increasing pressure on signal/image processing hardware and software to support

higher resolution / denser sampling

» ADCs, cameras, imaging systems, microarrays, …

x

large numbers of sensors

» image data bases, camera arrays, distributed wireless sensor networks, …

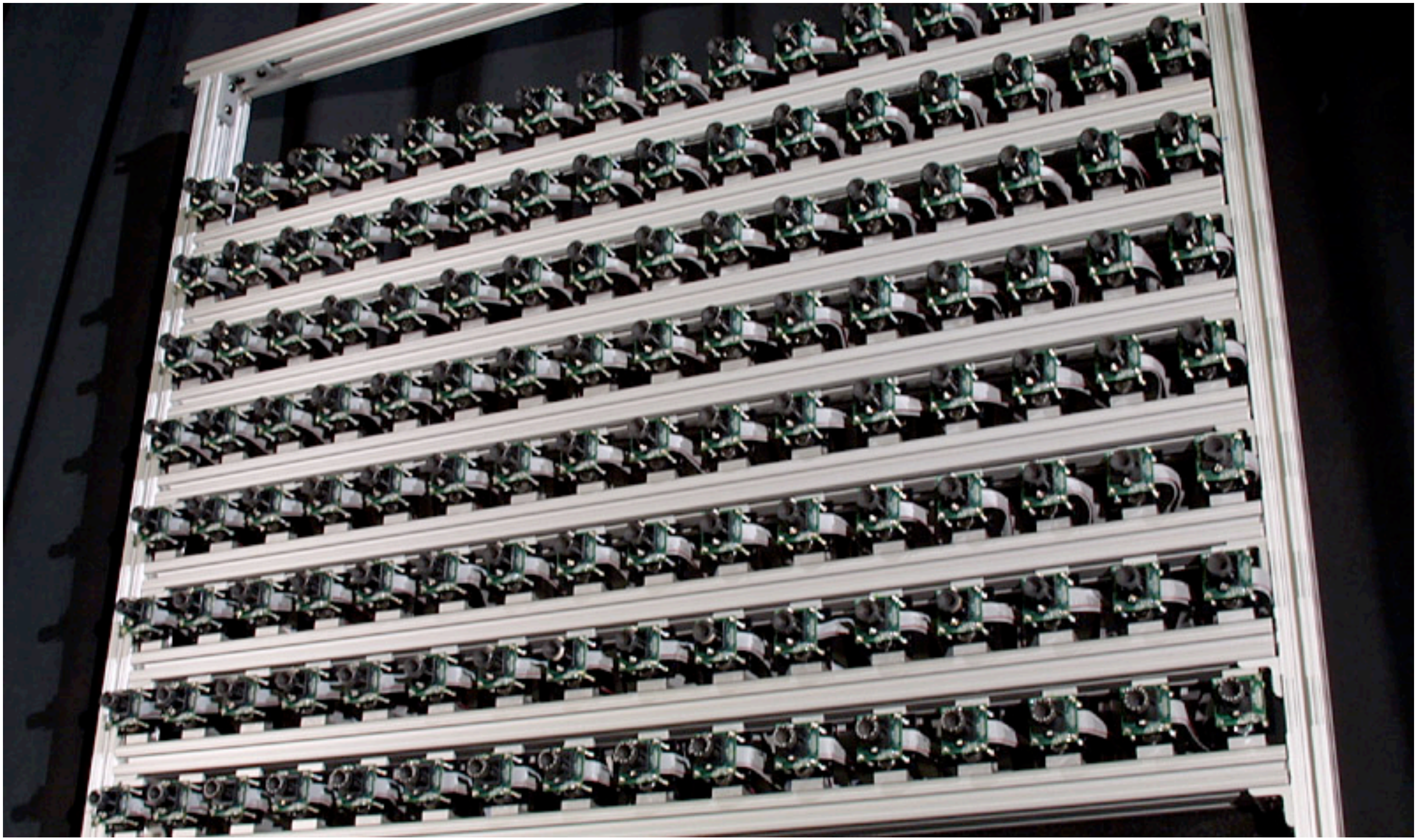
x increasing numbers of modalities

» acoustic, RF, visual, IR, UV

= *deluge of sensor data*

➢ how to efficiently **fuse**, **process**, **communicate**?

Sensor Data Deluge



Sensor Data Deluge

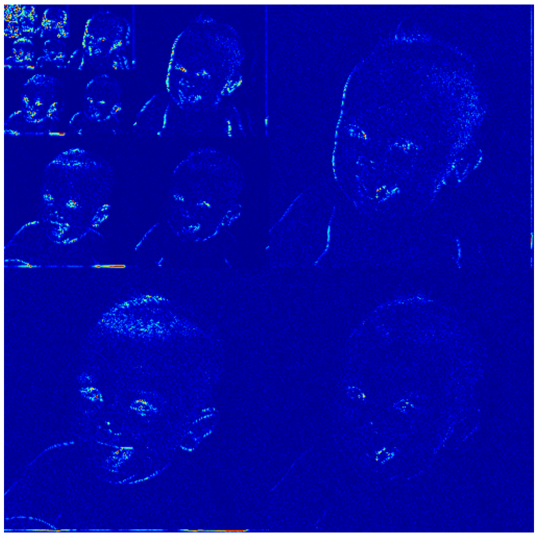
Sensor Data Deluge

Concise Models

- Efficient processing / compression requires concise representation

- **Sparsity** of an individual image

N pixels



$K \ll N$ large wavelet coefficients

(blue = 0)

Concise Models

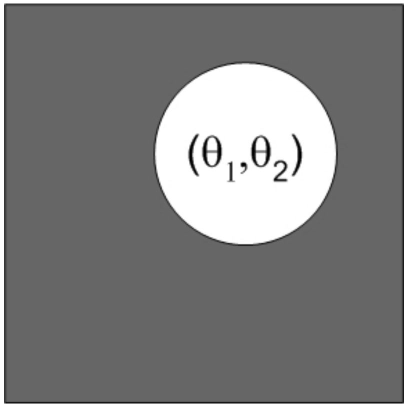
- Efficient processing / compression requires concise representation
- Our interest in this talk: **Collections** of images

Concise Models

- Efficient processing / compression requires concise representation

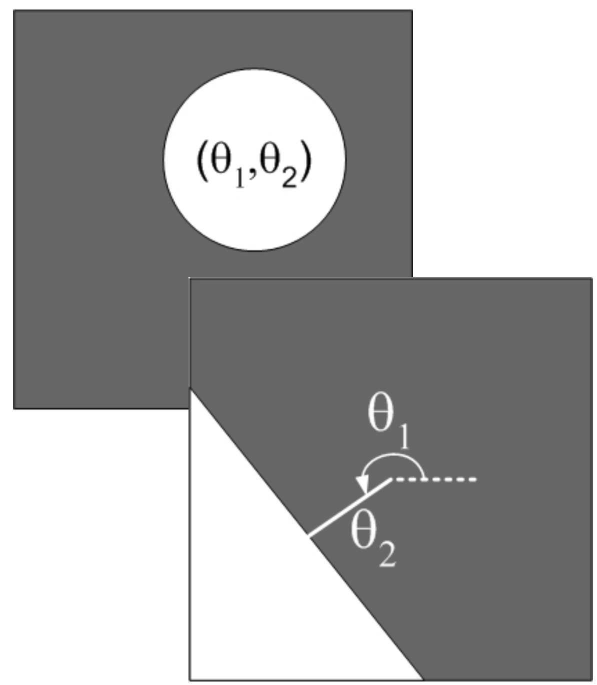
- Our interest in this talk: **Collections** of images **parameterized** by $\theta \in \Theta$

  - **translations of an object**
    θ: x-offset and y-offset



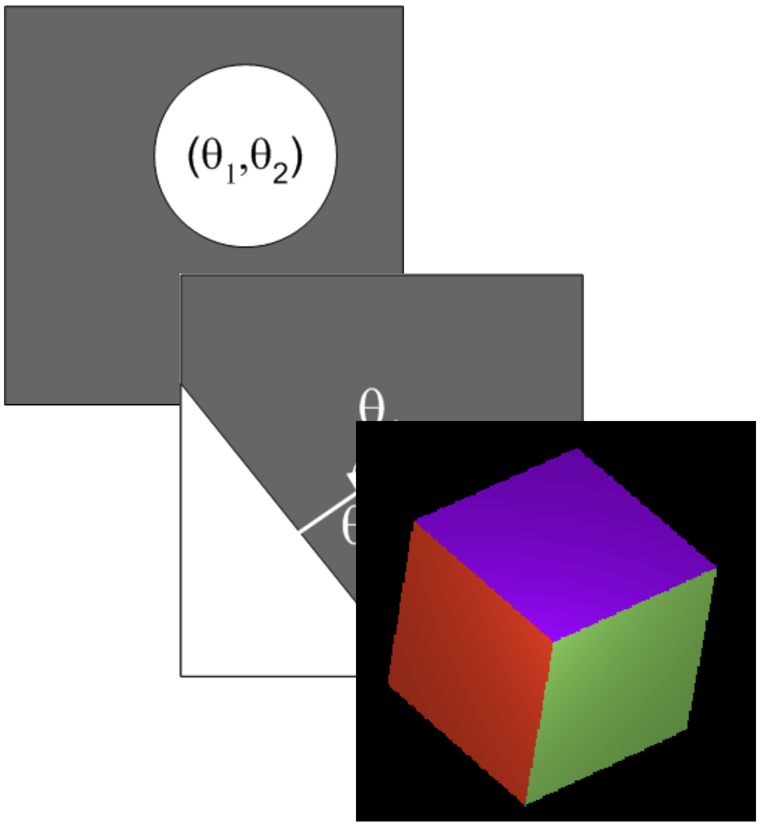
Concise Models

- Efficient processing / compression requires concise representation
- Our interest in this talk: **Collections** of images parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset
  - **wedgelets**
    θ: orientation and offset



Concise Models

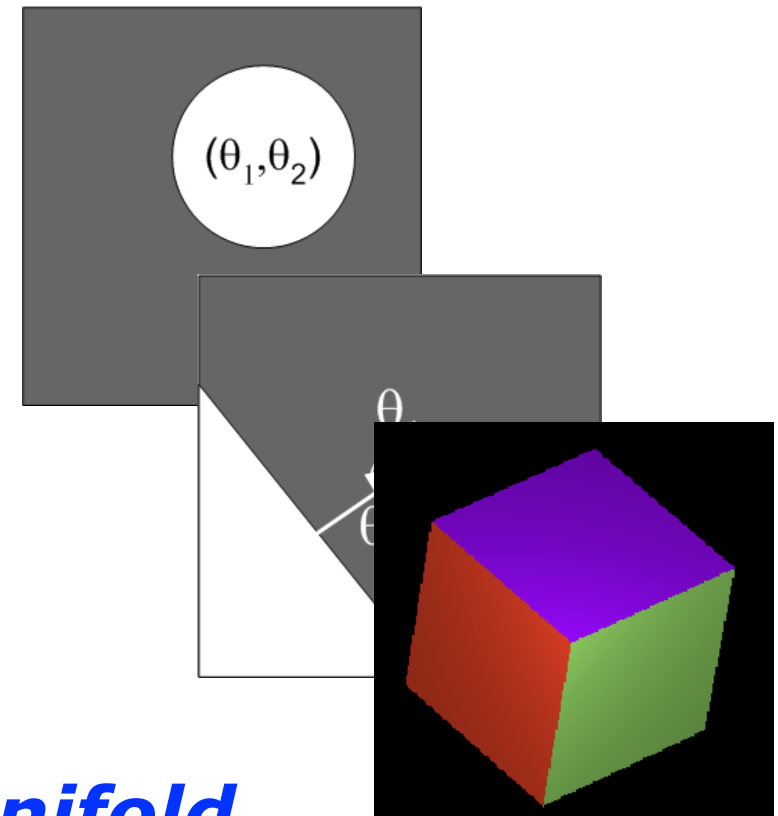
- Our interest in this talk: **Collections** of images parameterized by $\theta \in \Theta$
  - translations of an object
    θ: x-offset and y-offset
  - wedgelets
    θ: orientation and offset
  - **rotations of a 3D object**
    θ: pitch, roll, yaw



Concise Models

- Our interest in this talk: **Collections** of images parameterized by $\theta \in \Theta$

  - translations of an object
    θ: x-offset and y-offset
  - wedgelets
    θ: orientation and offset
  - rotations of a 3D object
    θ: pitch, roll, yaw

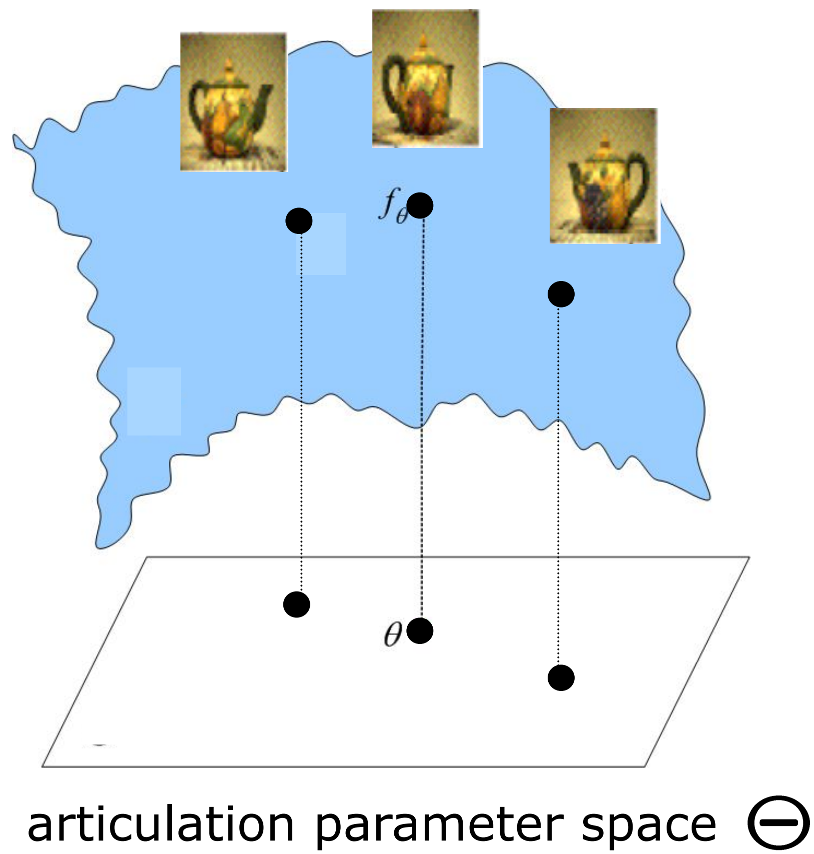


$$\mathcal{M} = \{I_\theta \ : \ \theta \in \Theta\}$$

- **Image articulation *manifold***

Image Articulation Manifold

- In practice: *N*-pixel images: $I \in \mathbf{R}^N$
- In theory: $I \in L^2([0,1] \times [0,1])$
- *K*-dimensional articulation space
- Then $\mathcal{M} = \{I_\theta \,:\, \theta \in \Theta\}$ is a *K*-dimensional manifold in the ambient space
- Concise model when *K*<<*N*



Smooth IAMs

- In practice: *N*-pixel images: $I \in \mathbf{R}^N$
- In theory: $I \in L^2([0,1] \times [0,1])$

- If the images are smooth then so is $\mathcal{M}$

- **Isometry**: locally, image distance $\propto$ parameter space distance

- **Locally linear tangent spaces** are close approximation

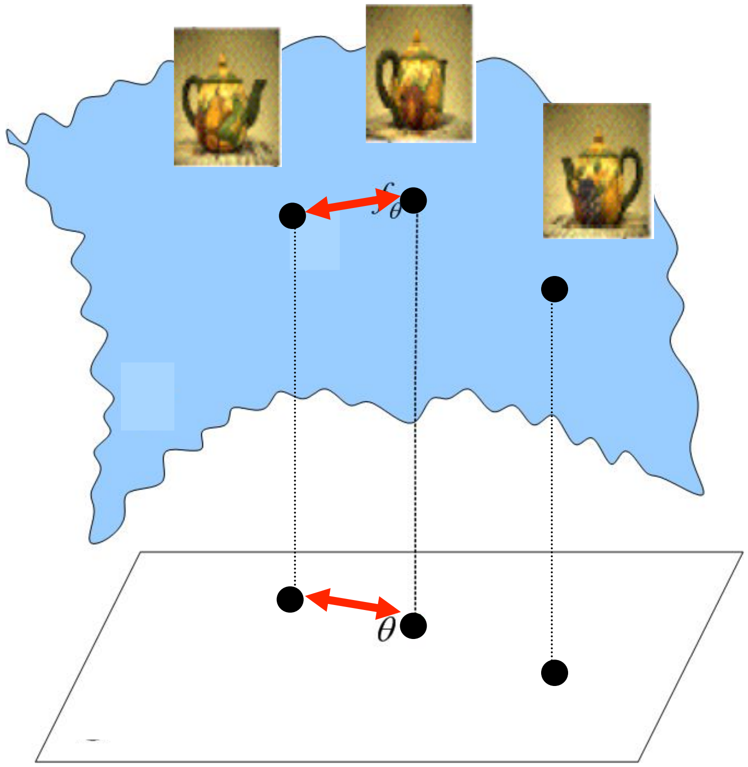


Image Articulation Manifold

- K=1
  rotation

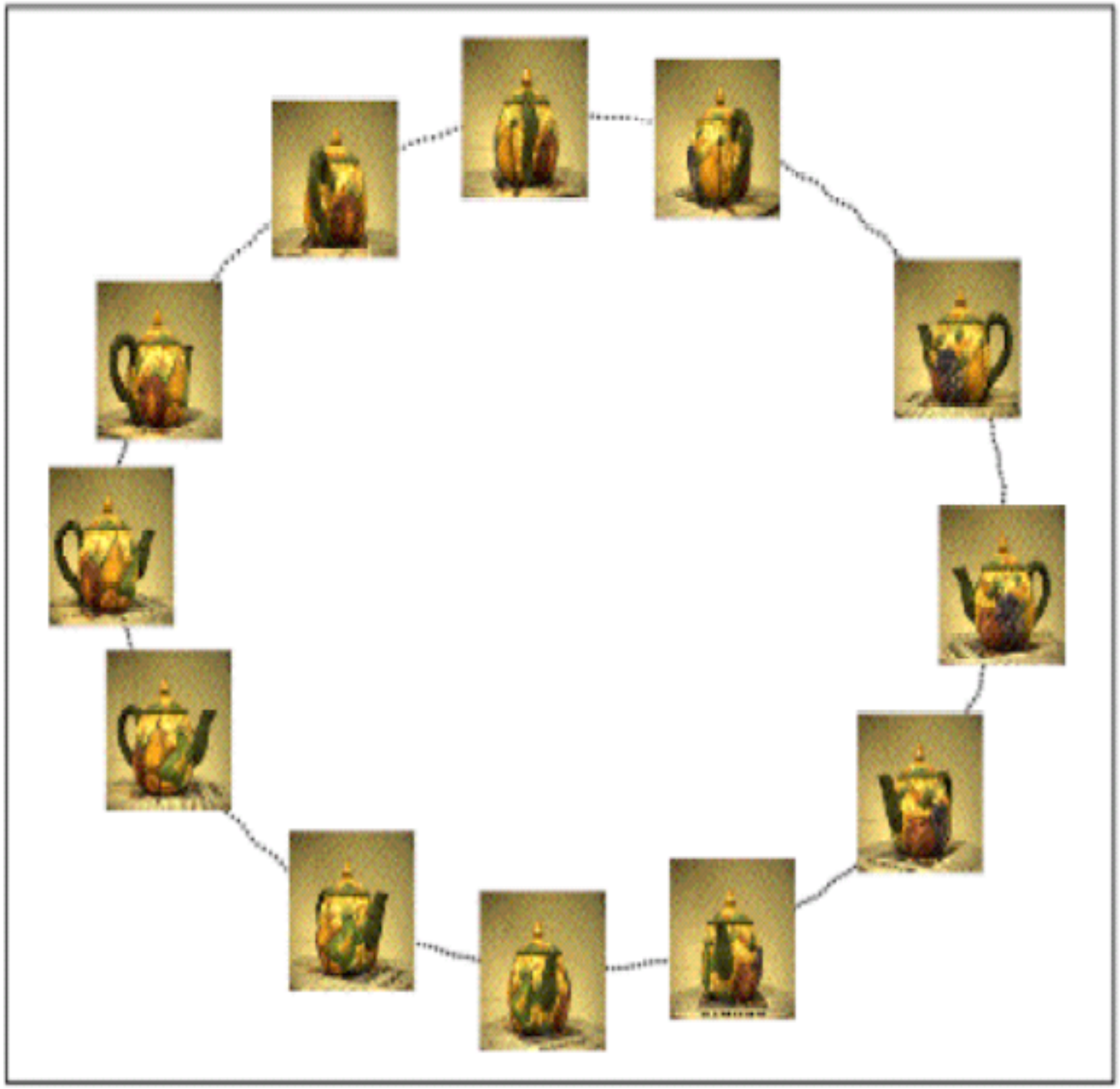
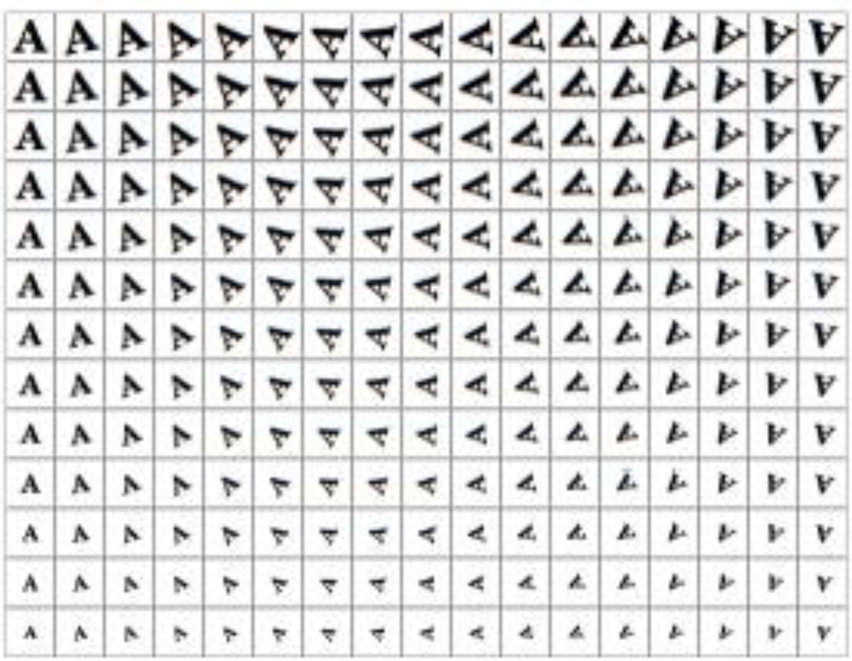


Image Articulation Manifold

- *K*=2
  rotation and scale



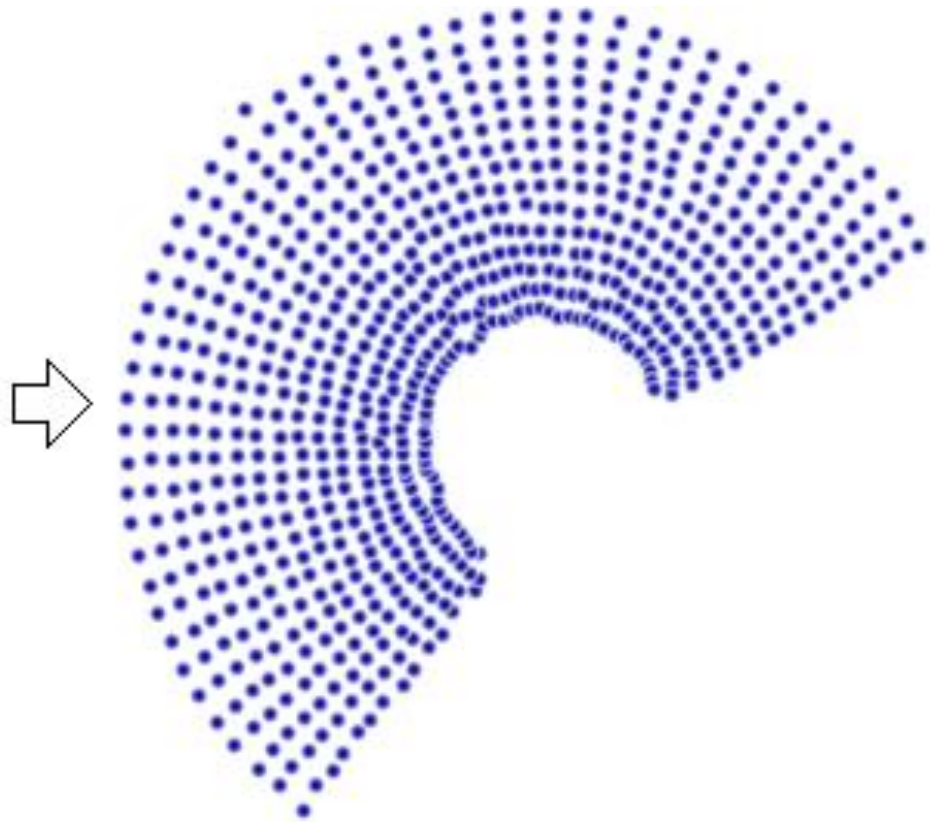
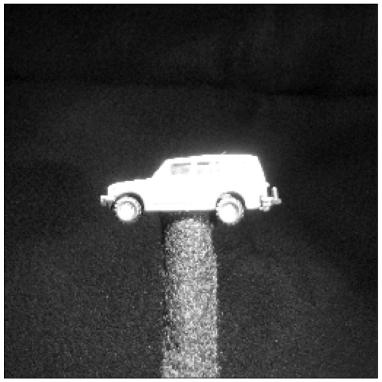


Image Articulation Manifold

- **IAM viewpoint unifying** for a large number of image inference problems
  - detection, classification, estimation, interpolation, …

  involving
  - imaging nuisance parameters
  - multiple sensors/viewpoints

- Example: target classification with unknown imaging parameters



Classification Geometry

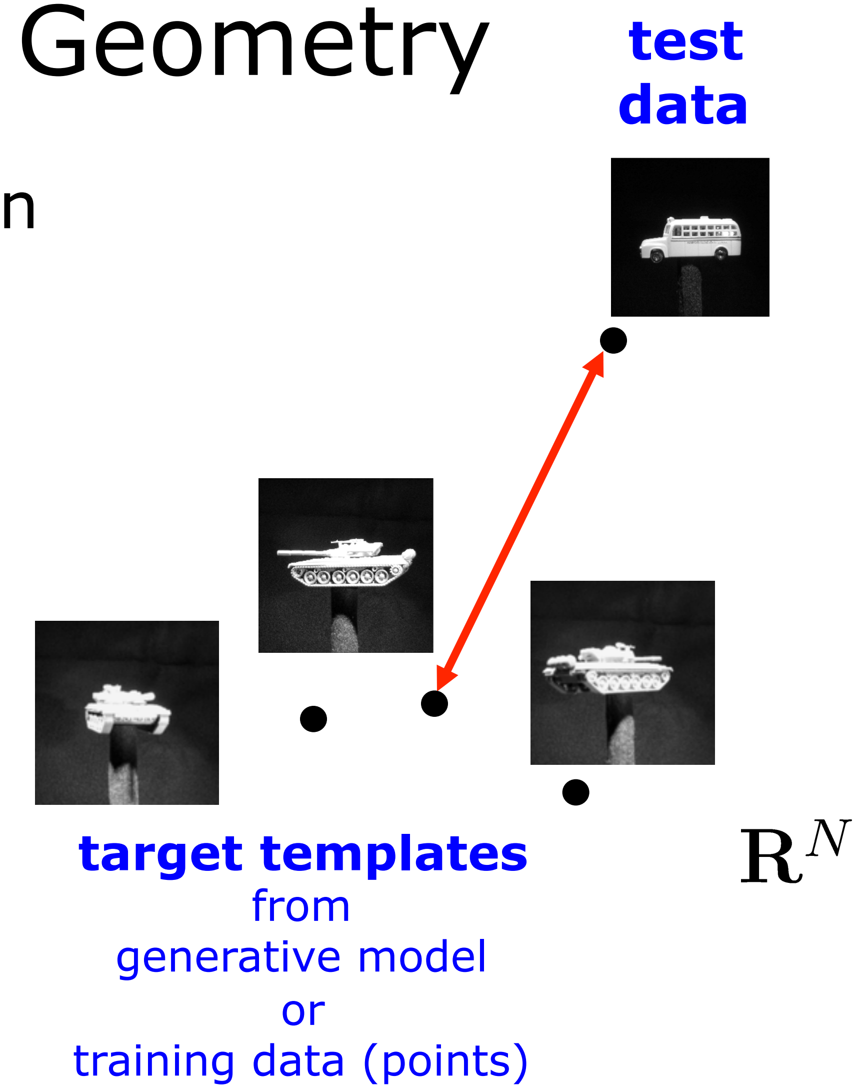
- Classification with K unknown articulation parameters

- Images are points in $\mathbf{R}^N$

- White Gaussian noise

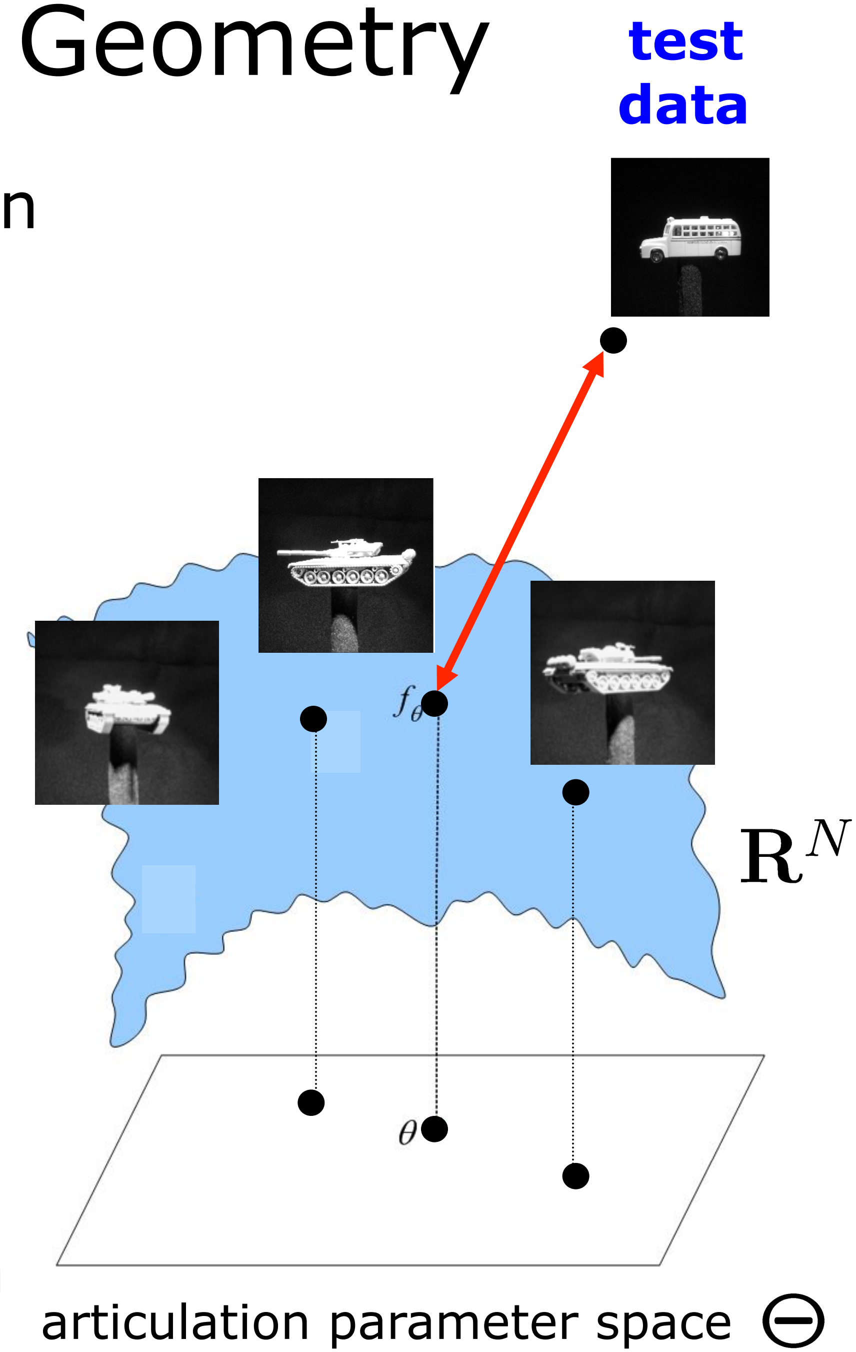
- **Classify** by finding closest target template to data for each class
  - distance or inner product

- "Matched filter"



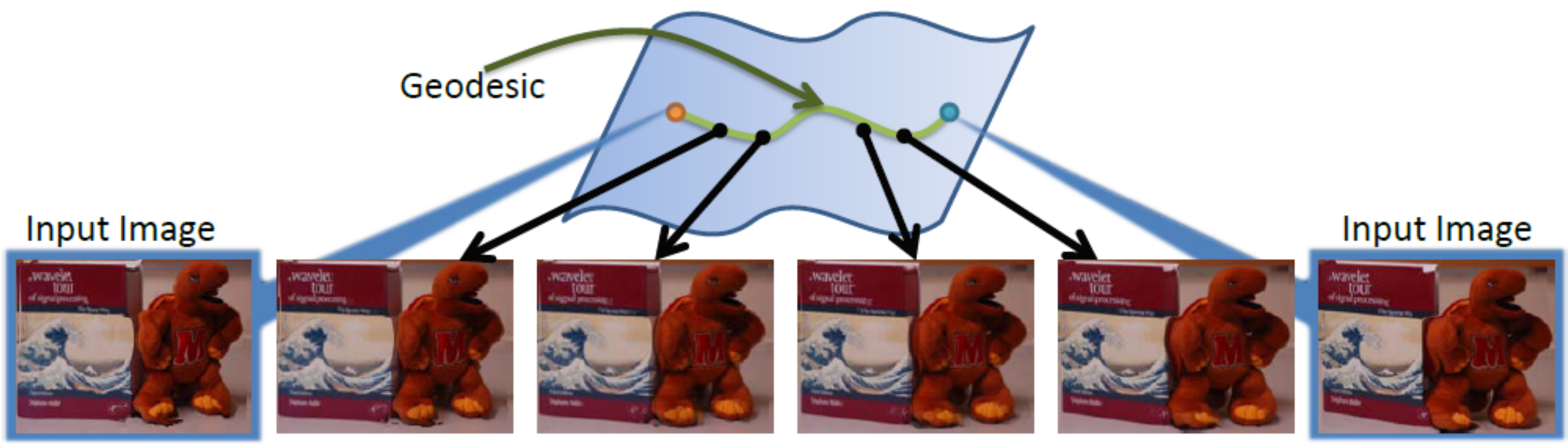
Classification Geometry

- Classification with *K* unknown articulation parameters
- Images are points in $\mathbf{R}^N$
- Classify by finding closest target template to data
- As template articulation parameter changes, points map out a *K*-dim **nonlinear manifold**
- Matched filter classification = **closest manifold search**



Synthesis / Interpolation

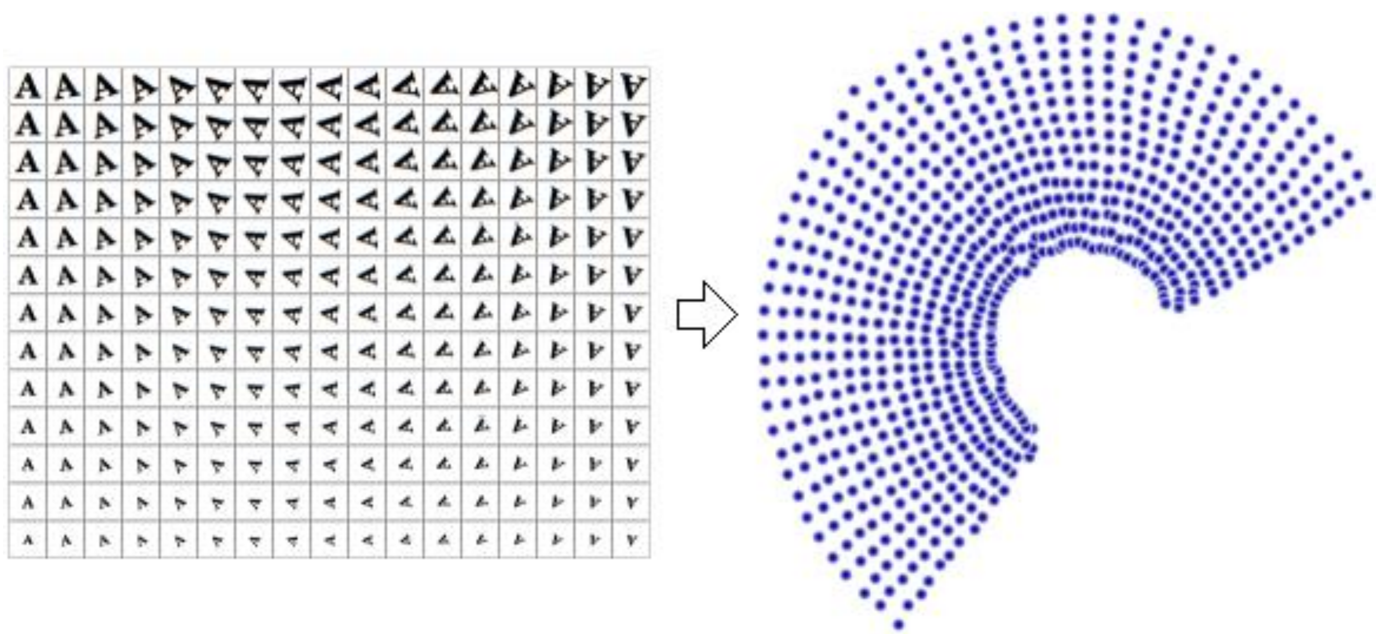
- Can sample points along manifold to **interpolate** or **synthesize** images



Manifold Learning

- Exploit fact that locally
  image distance $\propto$ parameter space distance
  to learn parameter space given a collection of images

- Numerous algorithms: ISOMAP, LLE, LE, HE, …

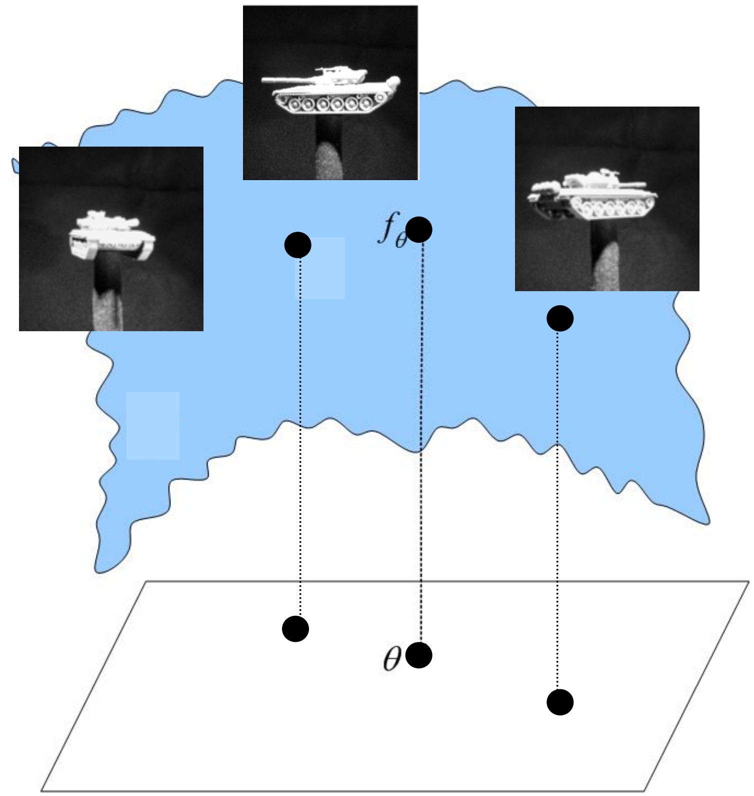


Theory/Practice Disconnect

- Practical image manifolds are **not smooth**!

- If images have sharp edges, then manifold is everywhere **non-differentiable**

[Donoho Grimes,2003]



articulation parameter space Θ

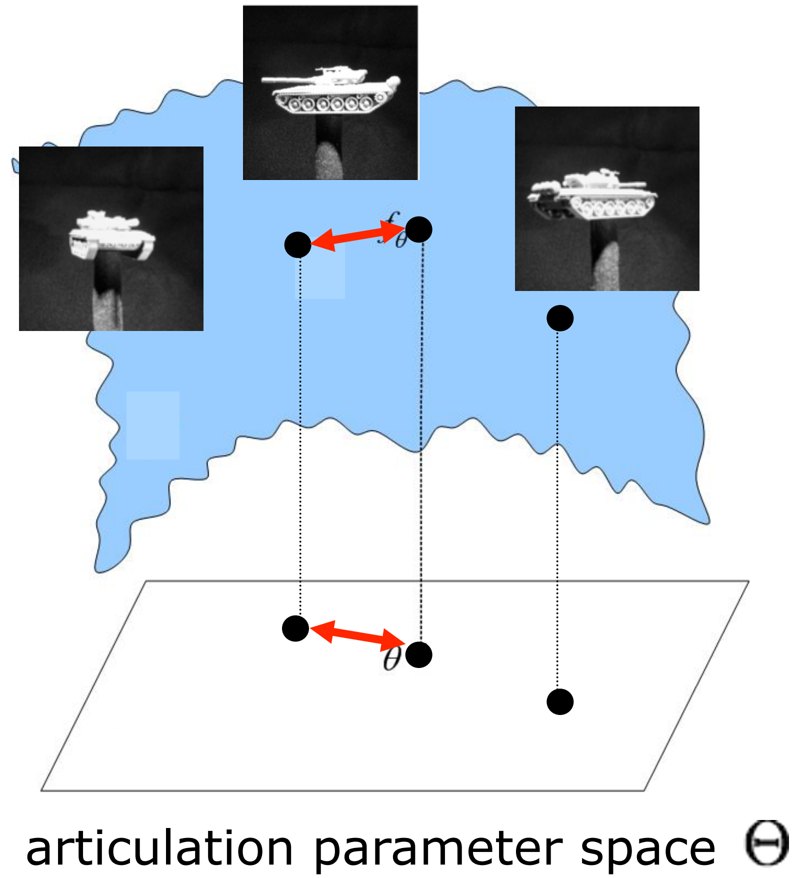
Theory/Practice Disconnect – 1

- **Lack of isometry**

- Local image distance on manifold should be proportional to articulation distance in parameter space

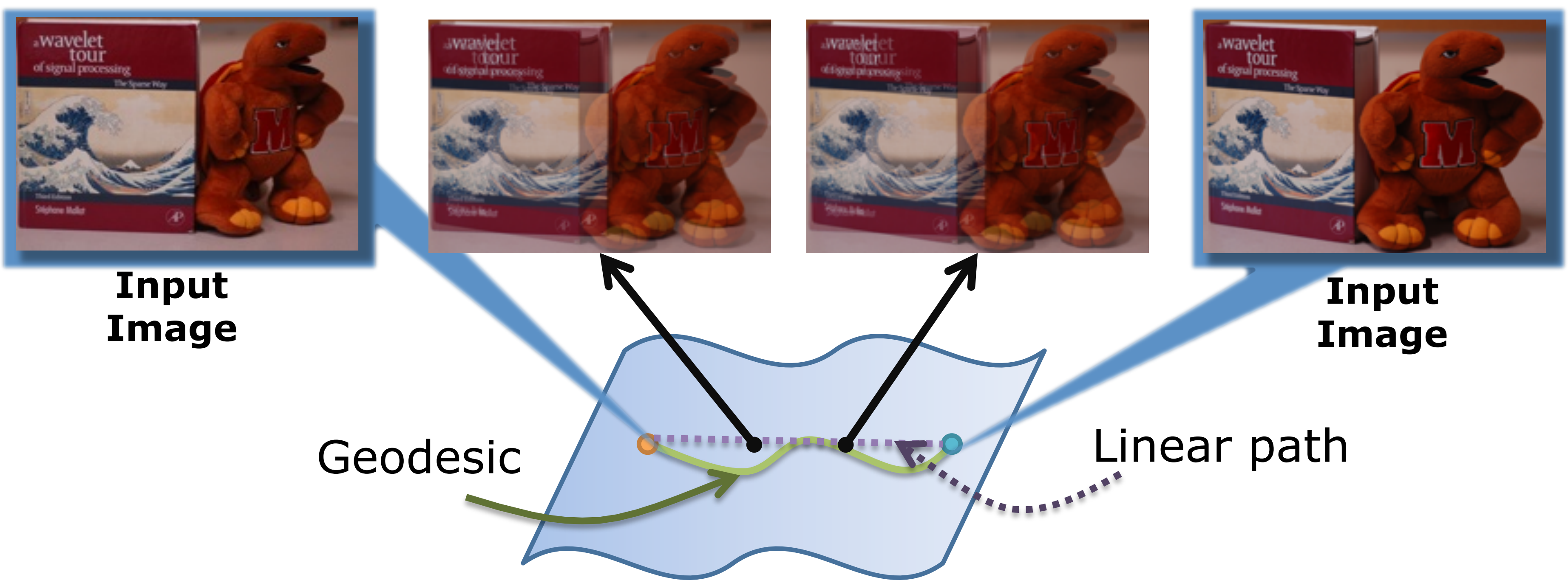
- But true only in toy examples

- Result: poor performance in classification, estimation, tracking, learning, …



Theory/Practice Disconnect – 2

- **Lack of local linearity**
- Local image neighborhoods assumed to form a **linear tangent subspace** on manifold
- But true only for extremely small neighborhoods
- Result: **cross-fading** when synthesizing images that should lie on manifold



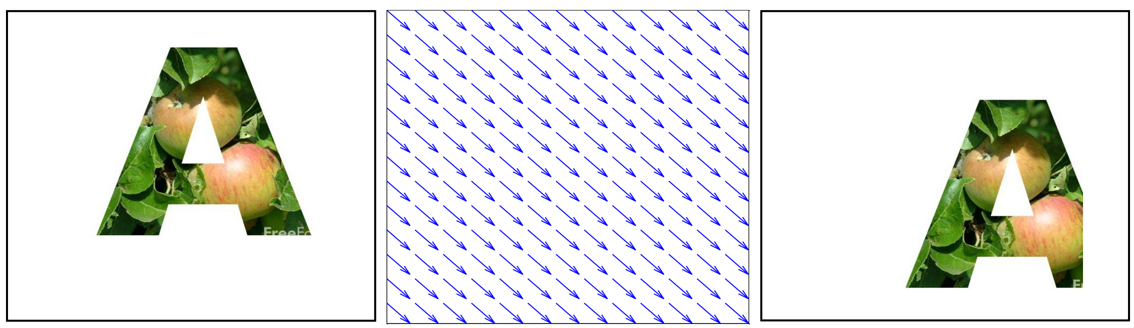
Insight: Leverage Progress in CV

- Computer vision (CV) community has developed powerful tools for image registration
  - **optical flow** for computing dense correspondences between images
  - huge progress over last 5 years

Optical Flow

- **Brightness constancy:** Given two images I_1 and I_2, we seek a displacement vector field $f(x, y) = [u(x, y), v(x, y)]$ such that

$$I_2(x,y) = I_1(x + u(x,y), y + v(x,y))$$



- **Linearized brightness constancy**

$$I_2(x,y) = I_1(x,y) + (\nabla_X I_1)u(x,y) + (\nabla_Y I_1)v(x,y)$$

Optical Flow History

$$I_2(x,y) = I_1(x + u(x,y), y + v(x,y))$$

$$I_2(x,y) = I_1(x,y) + (\nabla_X I_1)u(x,y) + (\nabla_Y I_1)v(x,y)$$

- Dark ages (<1985)
  - special cases of LBC by solving an under-determined set of linear equations
- Horn and Schunk (1985)
  - LBC solved via smoothness prior on the flow
- Brox et al (2005)
  - shows that linearization of brightness constancy a horrible assumption
  - develops optimization framework to handle BC directly
- Brox et al (2010), Black et al (2010), Liu et al (2010)
  - practical systems with reliable code

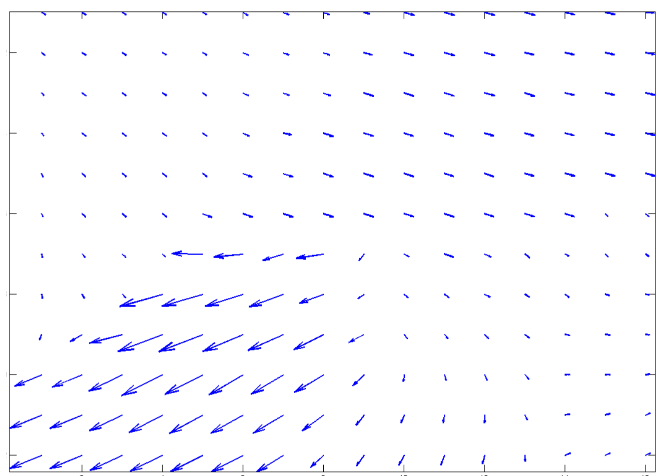
Optical Flow

$$I_2(x,y) = I_1(x+u(x,y), y+v(x,y))$$

$$I_2(x,y) = I_1(x,y) + (\nabla_X I_1)u(x,y) + (\nabla_Y I_1)v(x,y)$$

two-image sequence

optical flow

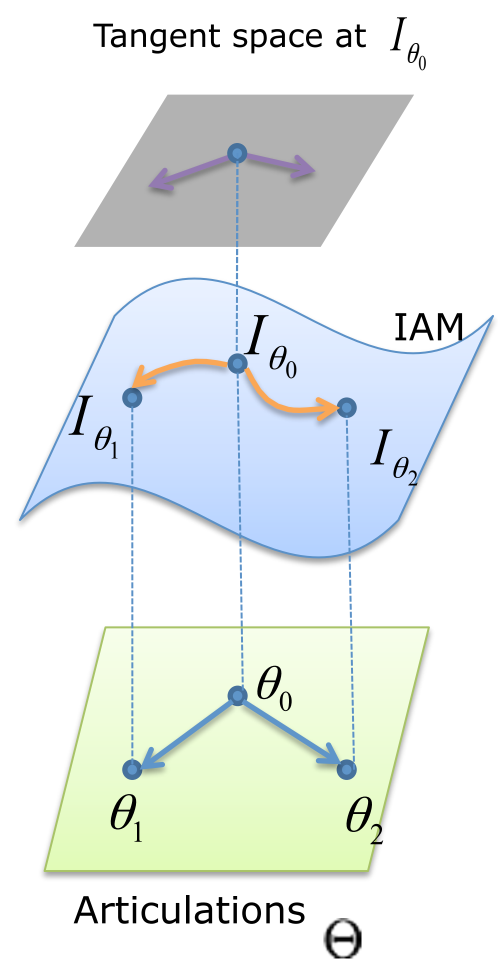


2^{nd} image predicted from 1^{st} via OF

(Figures from Ce Liu's optical flow page and ASIFT results page)

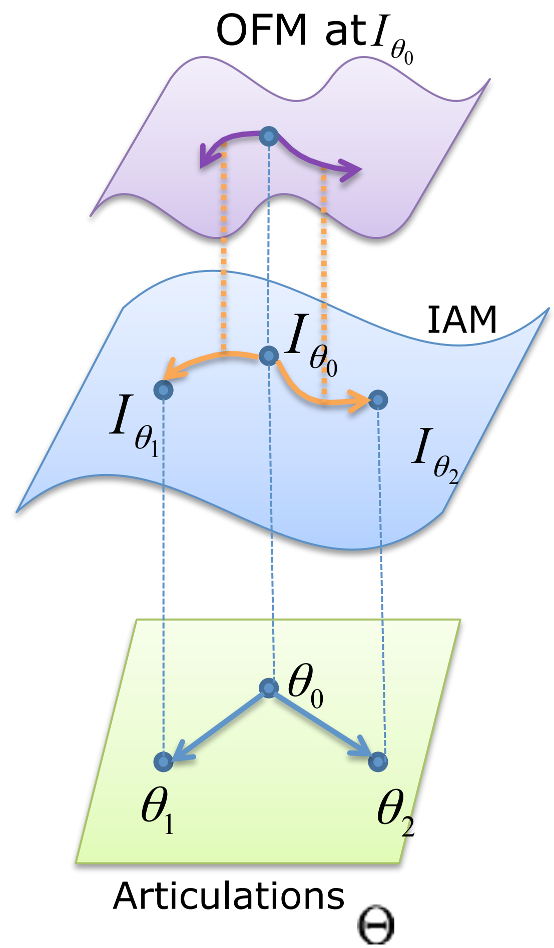
Image Articulation Manifold

- Consider a reference image I_{θ_0} and a *K*-dimensional articulation
- **Linear tangent space** at I_{θ_0} is *K*-dimensional
- Tangent space provides a mechanism to propagate along manifold
- **Problem:** Since manifold is non-differentiable, tangent approximation is poor

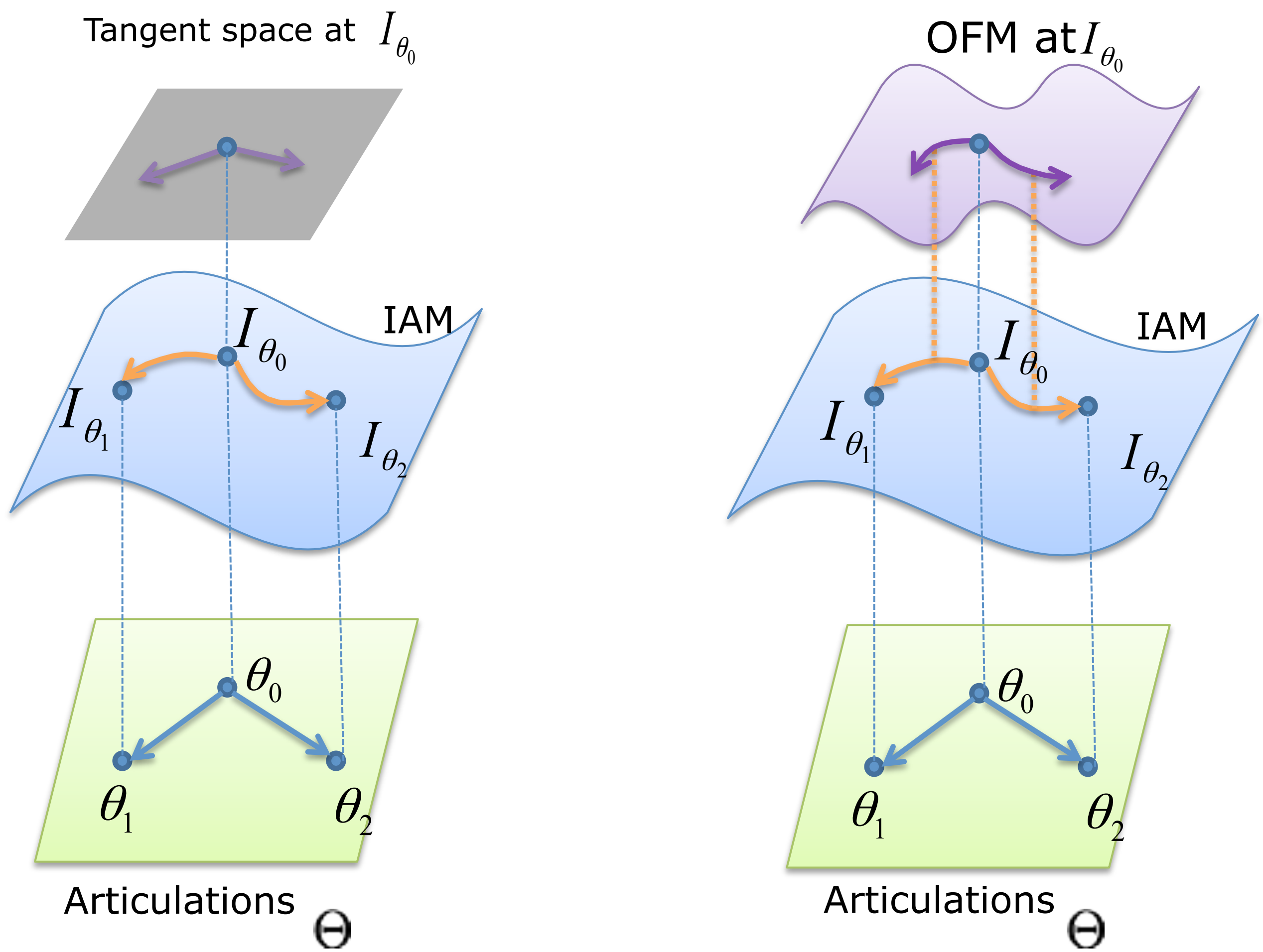


Optical Flow Manifold

- Consider a reference image I_{θ_0} and a K-dimensional articulation
- Collect optical flows from I_{θ_0} to all images reachable by a K-dimensional articulation
- Provides a mechanism to propagate along manifold
- **Theorem:** Collection of OFs is a **smooth, *K*-dimensional** manifold (even if IAM is not smooth)
  [N,S,H,B,2010]



OFMs as Nonlinear Tangent Spaces



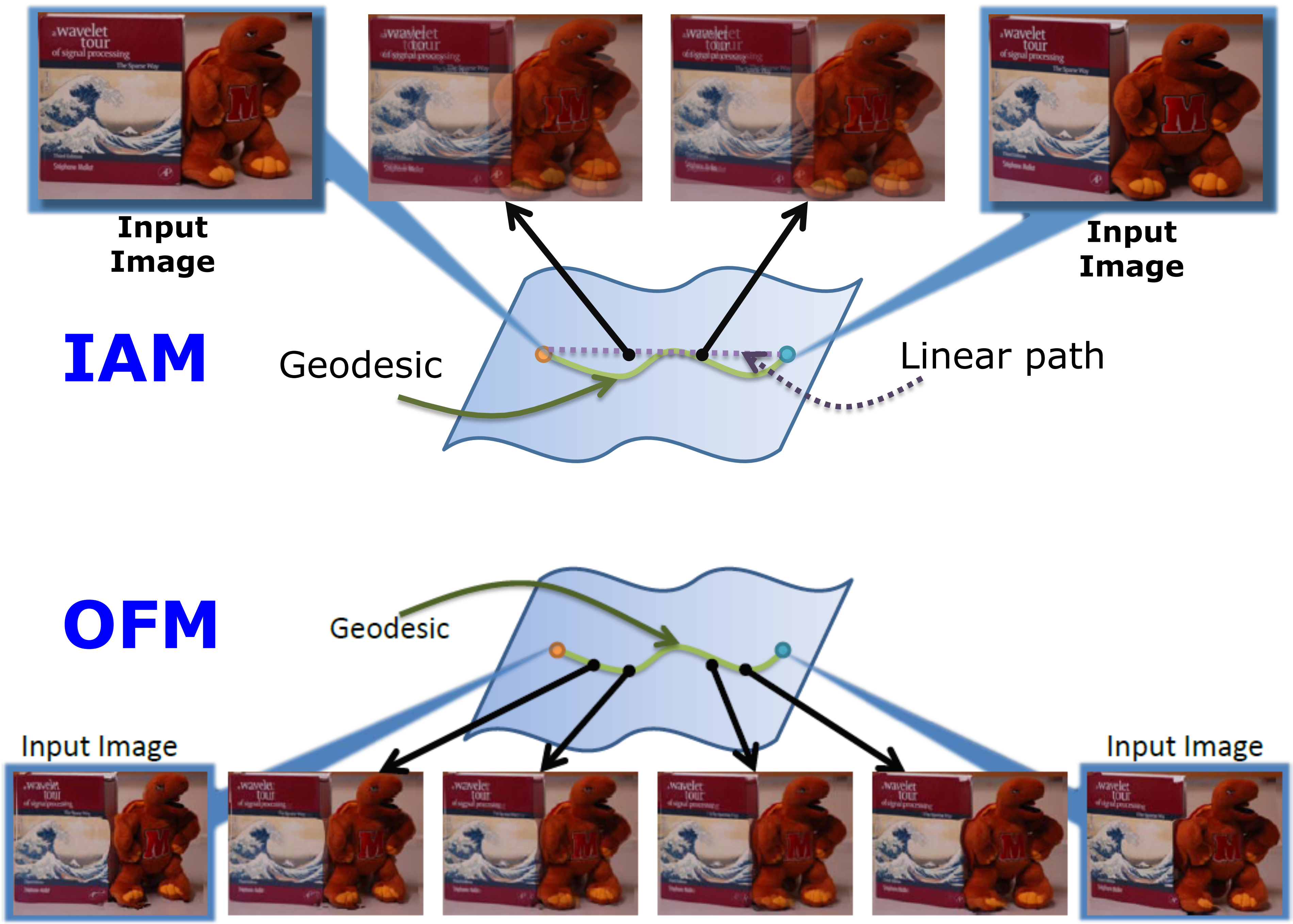
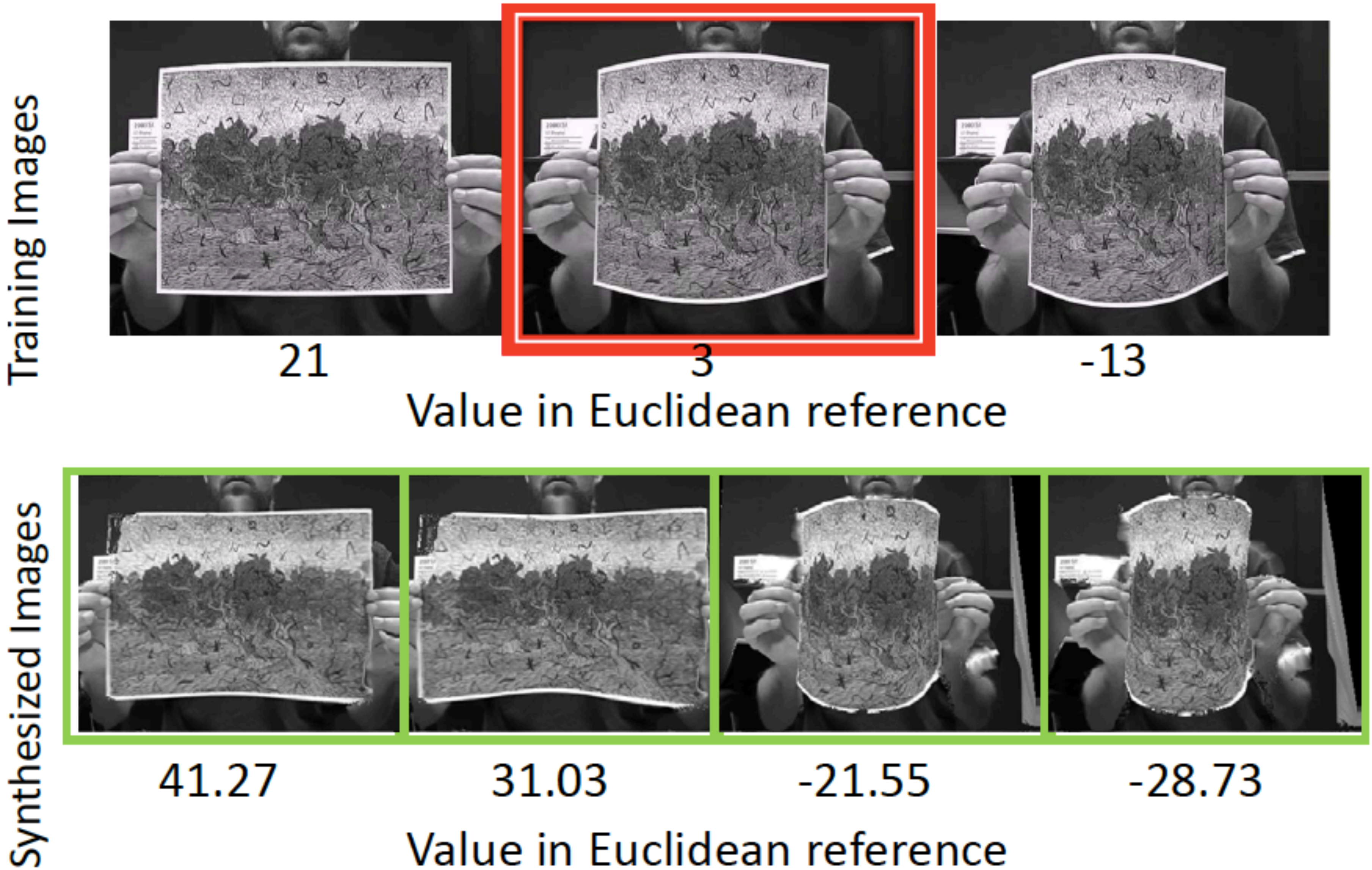
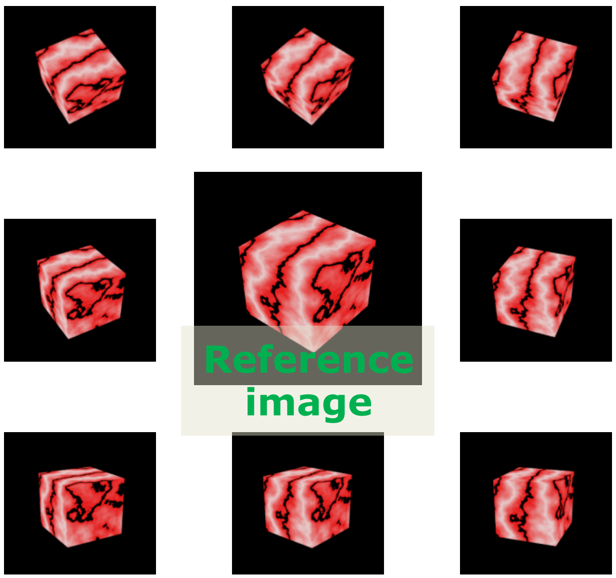

Input Image
Input Image
IAM
Geodesic
Linear path
OFM
Geodesic
Input Image
Input Image

Image Synthesis

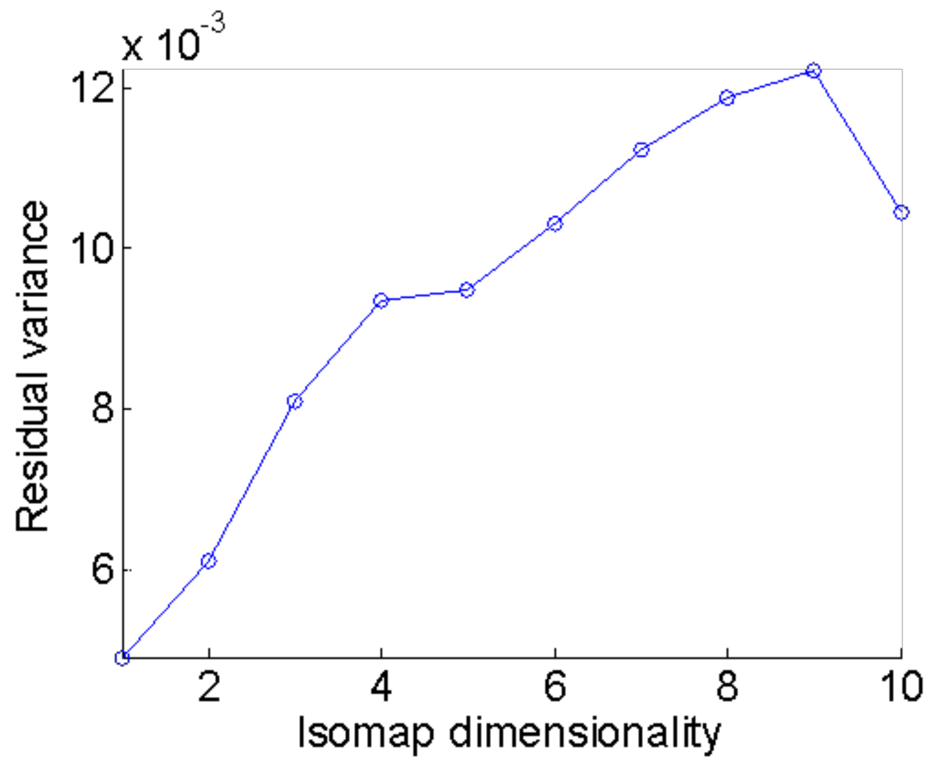


Manifold Learning via ISOMAP

2D rotations

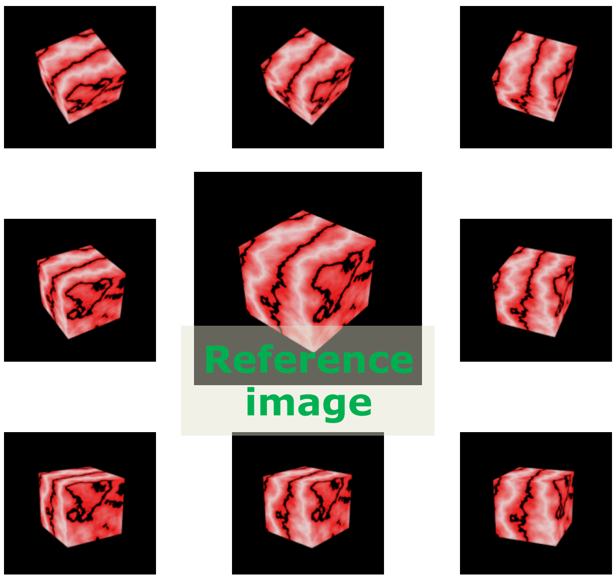


ISOMAP embedding error
for **IAM**

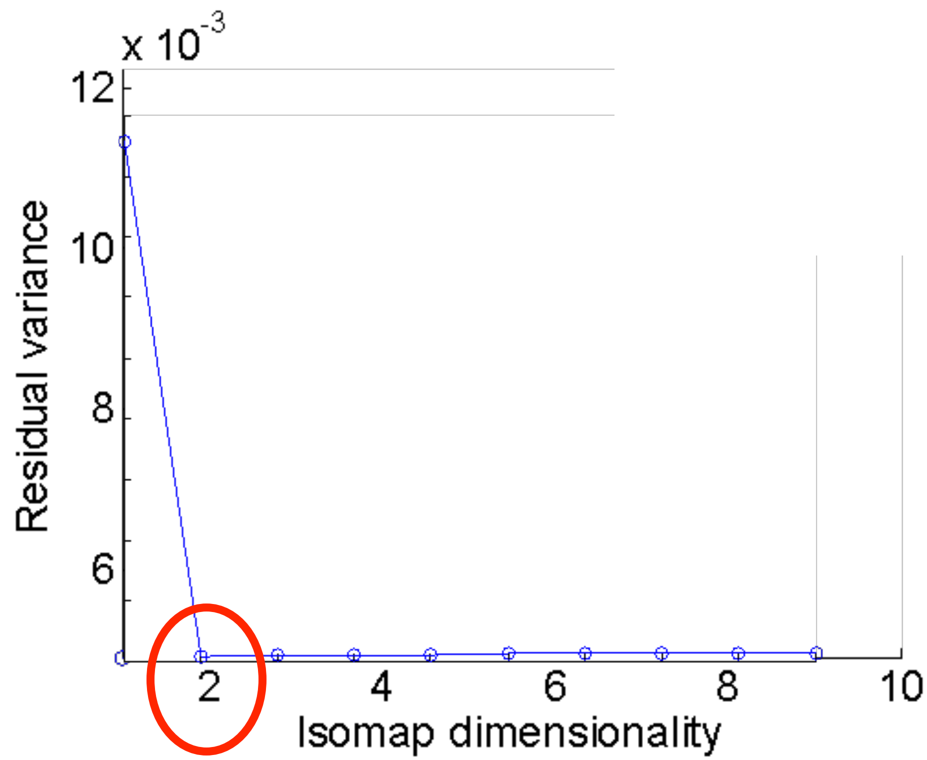


Manifold Learning via ISOMAP

2D rotations

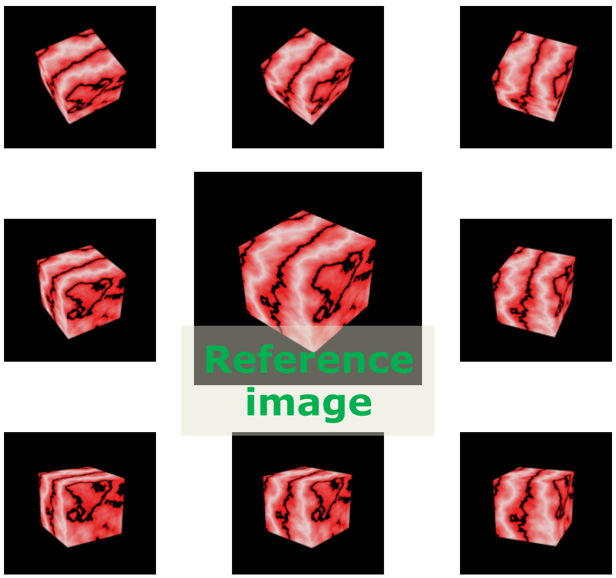


ISOMAP embedding error for **OFM**

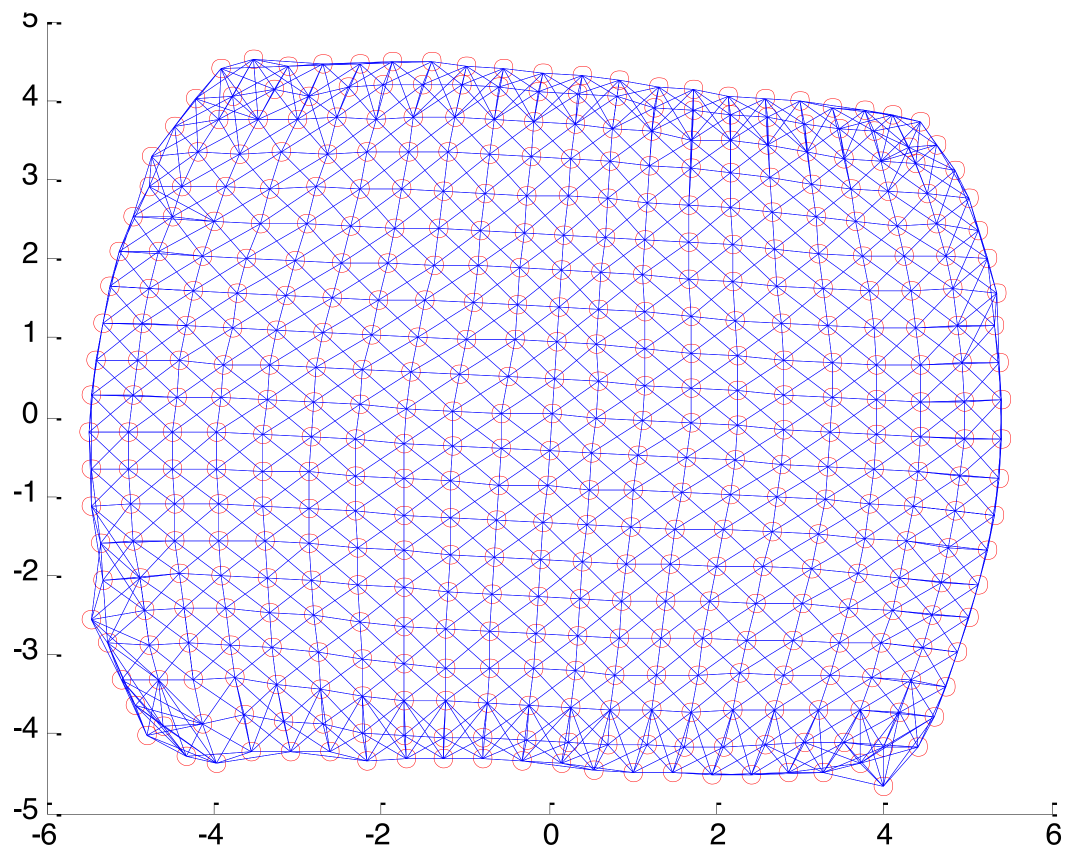


Manifold Learning via ISOMAP

2D rotations



Embedding of **OFM**



Manifold Charting

- **Goal:** build a **generative model** for an entire IAM/OFM based on a small number of base images

- Algorithm:
  - choose a reference image randomly
  - find all images that can be generated from this image by OF
  - compute Karcher (geodesic) mean of these images
  - repeat on the remaining images until no images remain

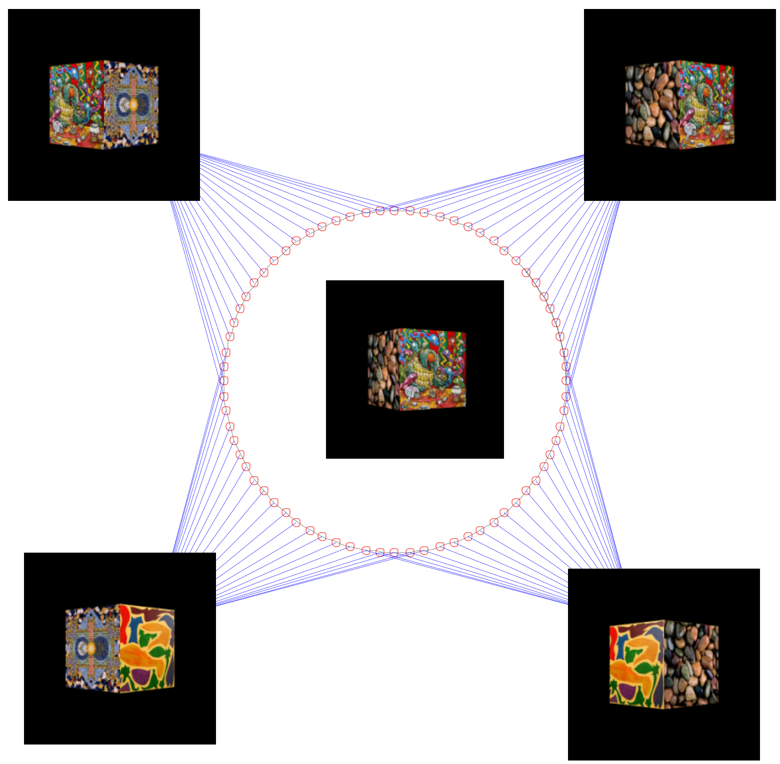
- **Exact representation** when no occlusions

Manifold Charting

- **Goal:** build a **generative model** for an entire IAM/OFM based on a small number of base images

- Ex: cube rotating about axis

- All images of the cube can be representing using 4 reference images + their respective OFMs

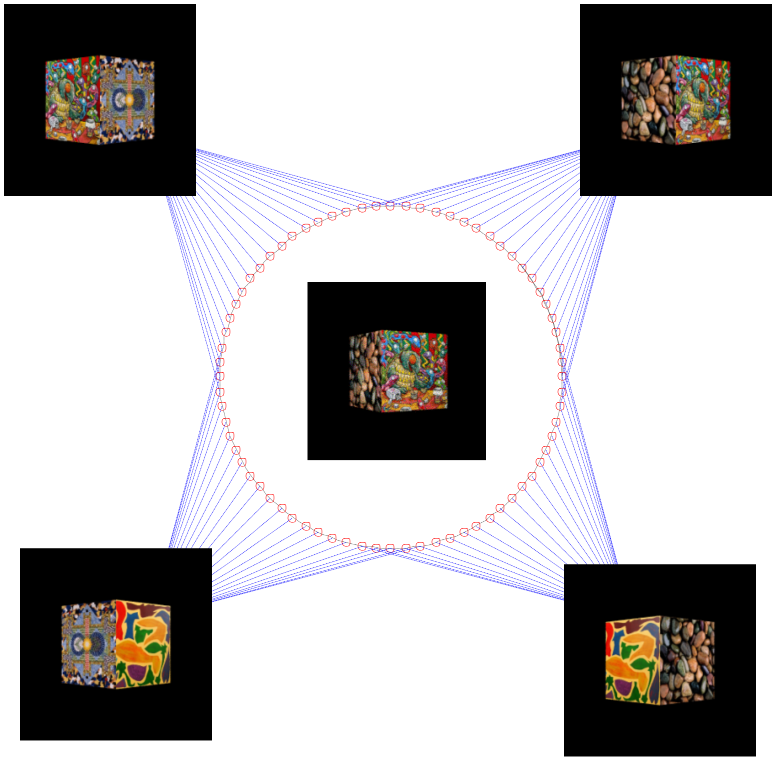


Manifold Charting for Classification

- **Optimal selection** of target templates for classification

- Dramatically reduced number of target templates (**compression**)

- **Optimal "next-view"** selection for adaptive sensing applications



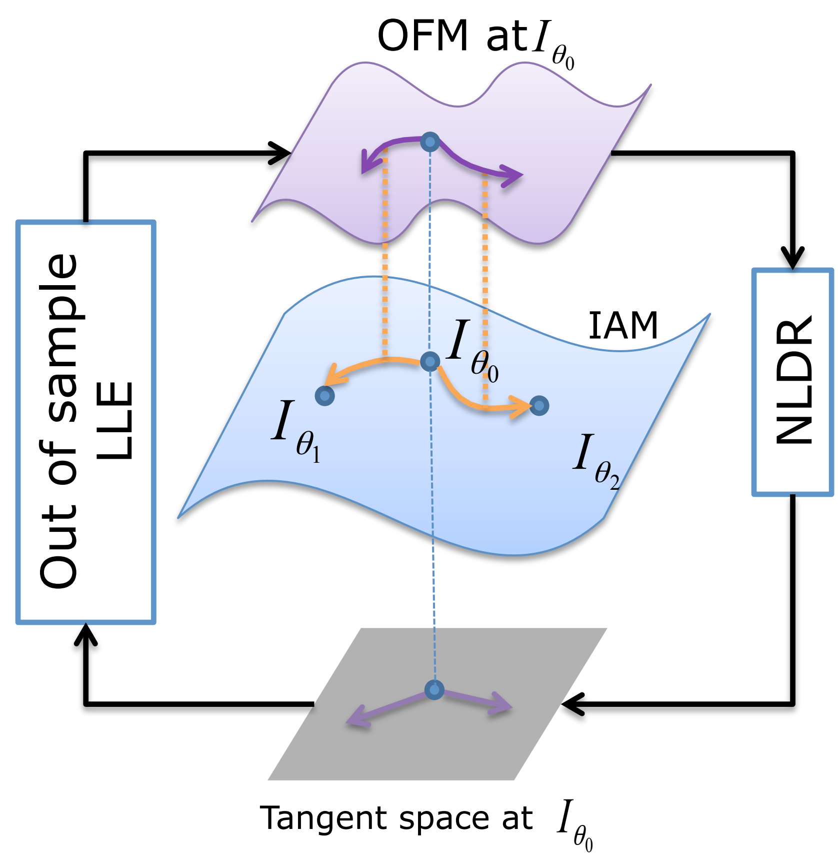
Summary

- Image articulation manifolds (IAMs) are a useful unifying construct for many image processing problems involving image collections and multiple sensors/viewpoints
- But practical IAMs are non-differentiable
  - IAM-based algorithms have not lived up to their promise
- Optical flow manifolds (OFMs)
  - Smooth even when IAM is not
  - OFM ~ nonlinear tangent space
  - Support accurate image synthesis, learning, charting, …

Not in Today's Talk

- Log and Exp maps between "image space" and "parameter space" become simple to calculate because OF varies so smoothly from image to image

- Enables simple and explicit strategies for
  - geodesic computation
  - Karcher means and
  - variances (for statistical models on manifolds)
  - geometric clustering, dendrograms (data organization)
  - image synthesis …



Not in Today's Talk

- For a large class of articulations, the resulting OFM is a **Lie group**
  - affine transformation (translations, videos from aircraft)
  - perspective transformation (scene at infinity, planar scenes)
  - diffeomorphisms (unstructured deformations)

- Lie groups have additional structure!
  - **Analytic generators** when the Lie group has an associated Lie Algebra
  - Ex: Affine groups [Olhausen et al, 2009]

Open Questions

- Our approach was specific to **image manifolds**
- Do there exist mollifying “nonlinear tangent spaces” for other kinds of non-smooth data manifolds?

Open Questions

- **Theorem:**
  $M = O(K \log N)$
  random measurements
  stably embed a *K*-dim manifold
  whp [B, Wakin, *FOCM* '08]

- Q: Is there an analogous result for OFMs?

dsp.rice.edu