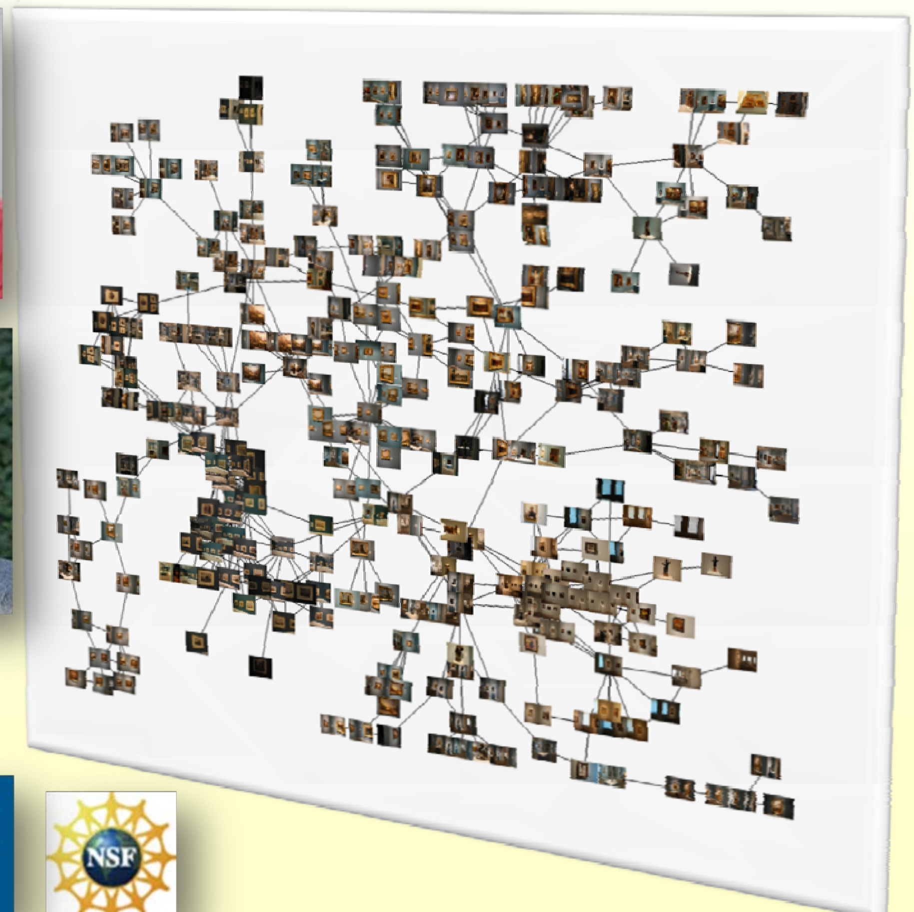
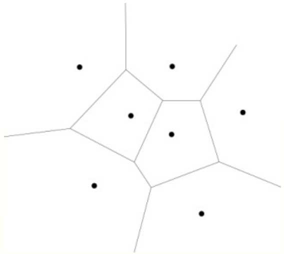
Structure Discovery in Sampled Spaces

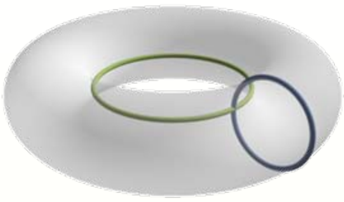
Leonidas Guibas
Computer Science

Gunnar Carlsson
Mathematics



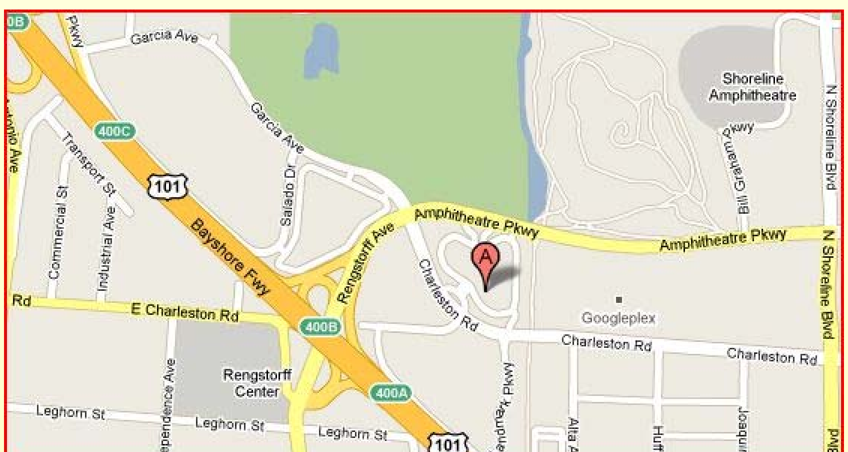


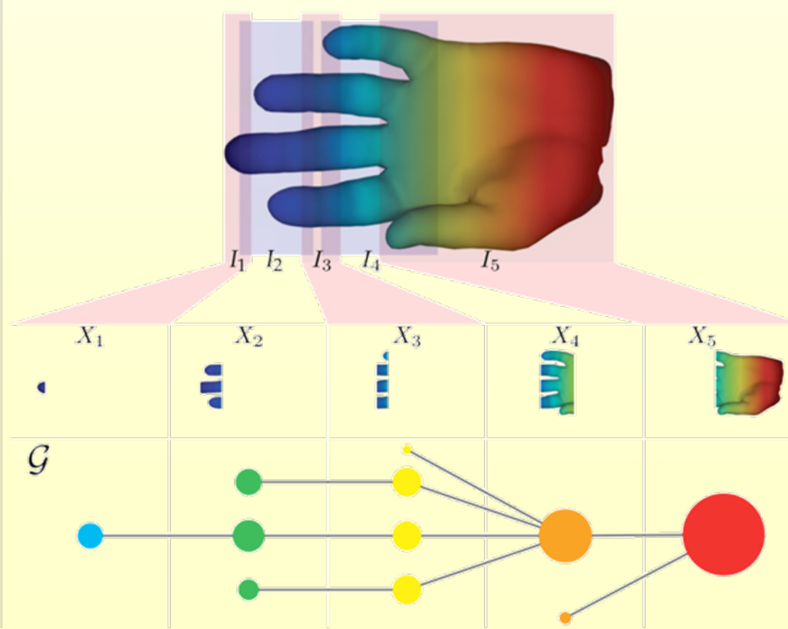
Project Goals



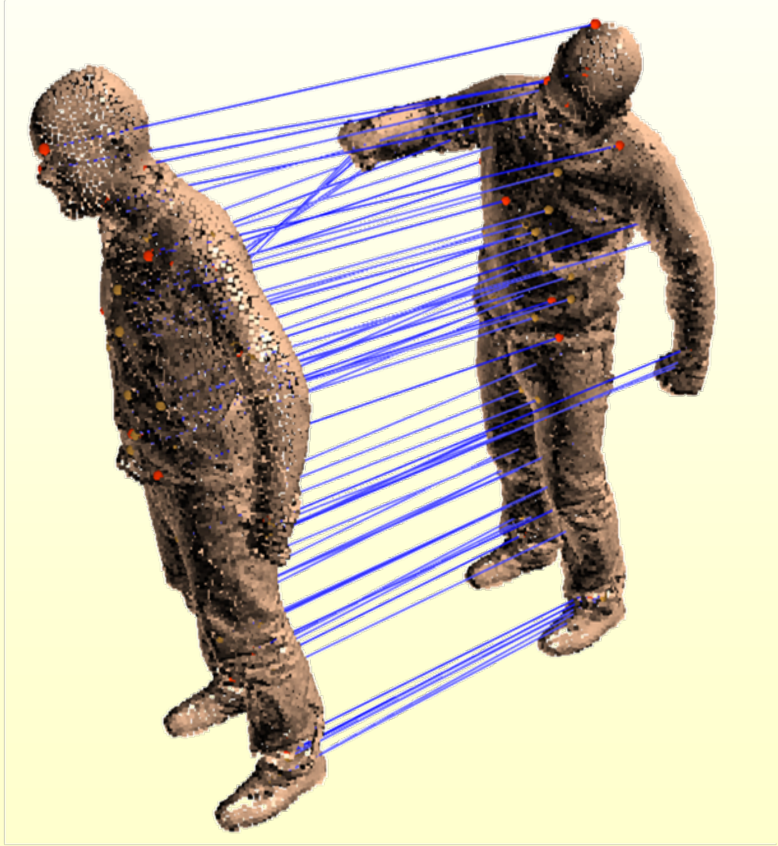
- Bring tools from Computational Geometry and Topology to the analysis and visualization of massive, distributed data sets
- Perform global structure discovery on such data
  - Produce meaningful topological and geometric maps over the data
  - Extract structural similarities or structure preserving correspondences within and across data sets
- Exploit this discovered structure in enabling visual exploration and human interaction with the data

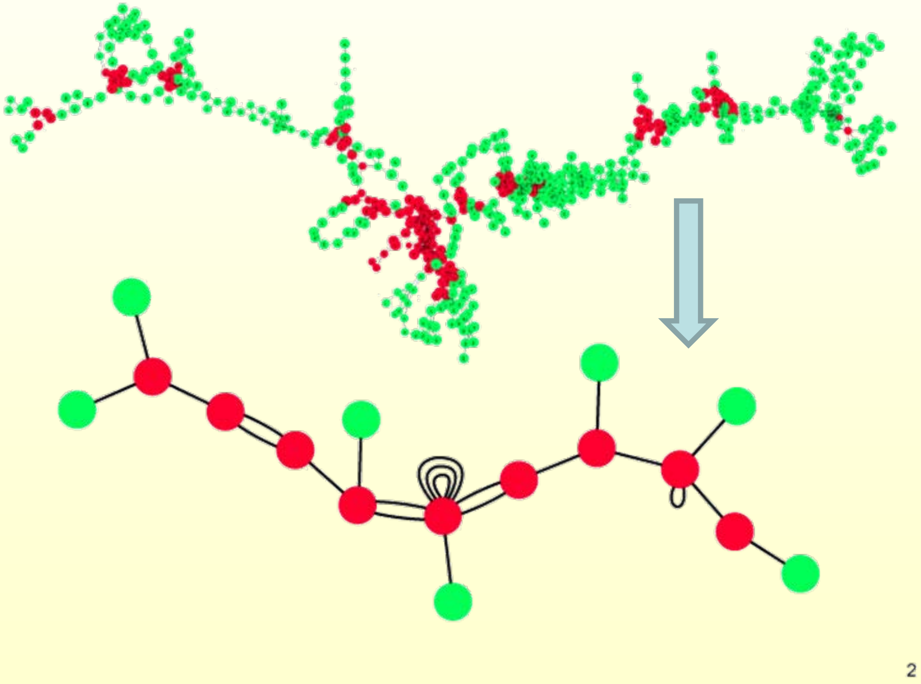
Understand Data via Maps





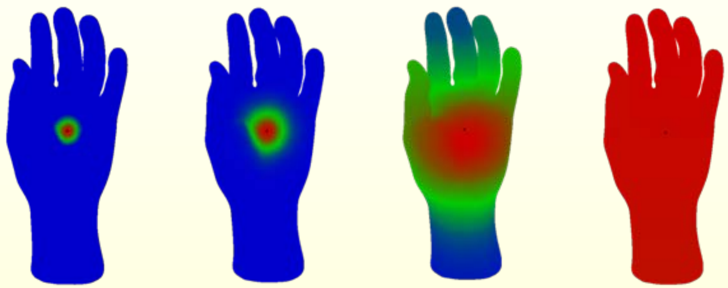
The Problem of Correspondences





Some Tools

Heat Diffusion on Manifolds



$$\frac{\partial u}{\partial t} = \Delta u$$

Δ : Laplace-Beltrami Operator (div grad)

Persistent Homology

$H_k = Z_k / B_k$

Persistence diagrams (barcodes)

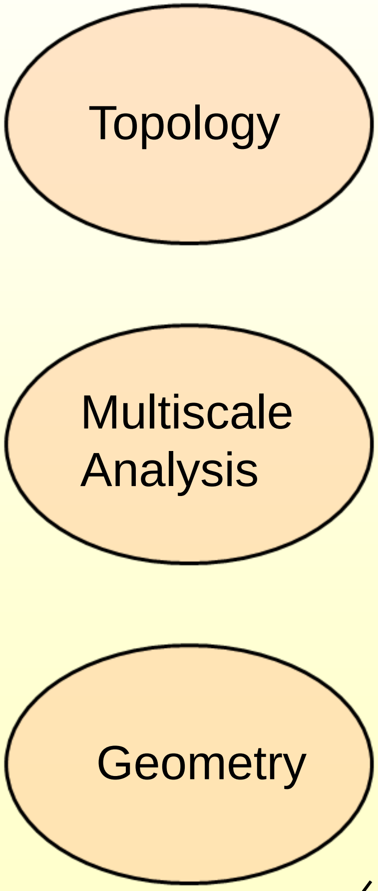
β_0: # components

β_1: # tunnels or loops

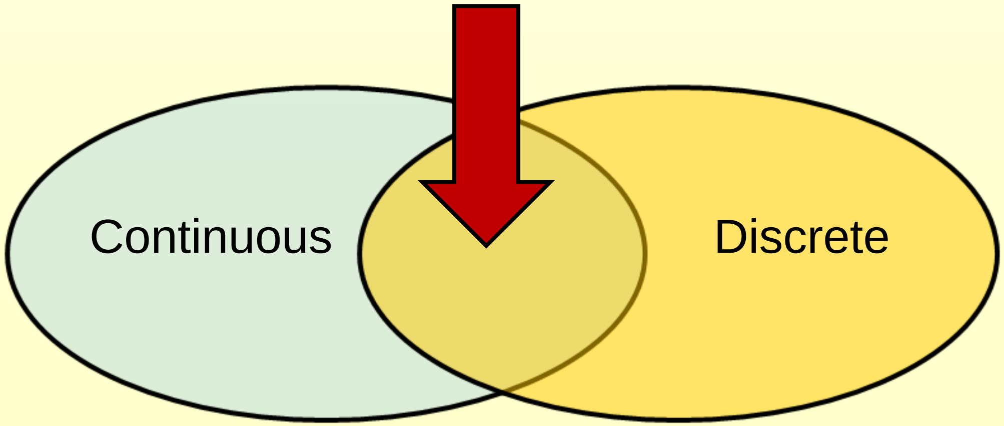
β_2: # voids

Three Quick Vignettes

I. Isometric Descriptors and Shape Correspondences
II. Circular Coordinates for Data Sets
III. Interlinked Image Collections

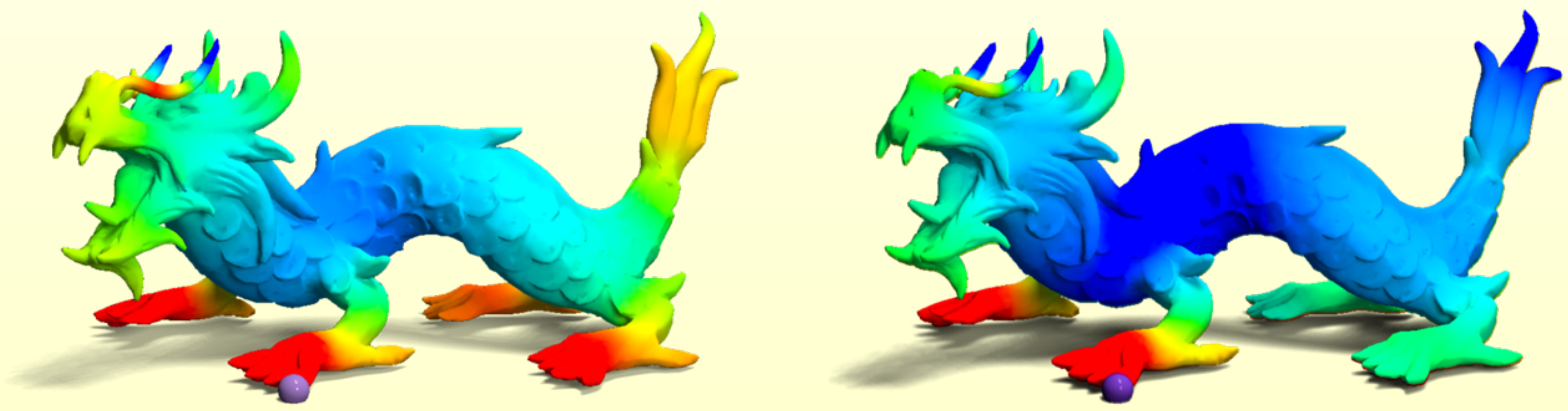


Structure Discovery
in Geometric Data



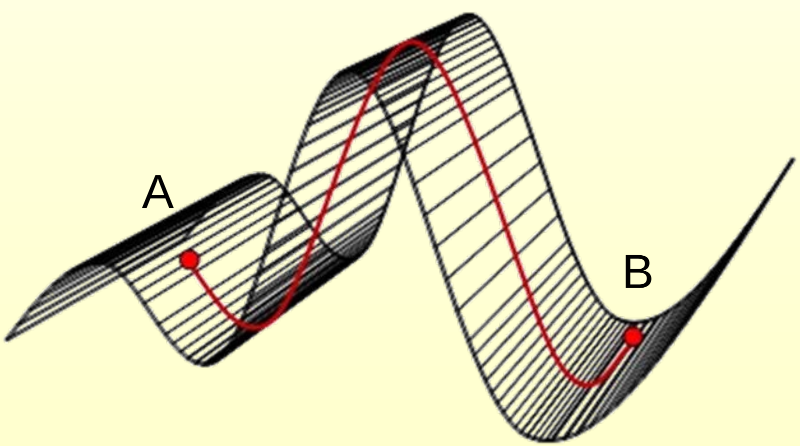
I. Isometric Descriptors and Shape Correspondences

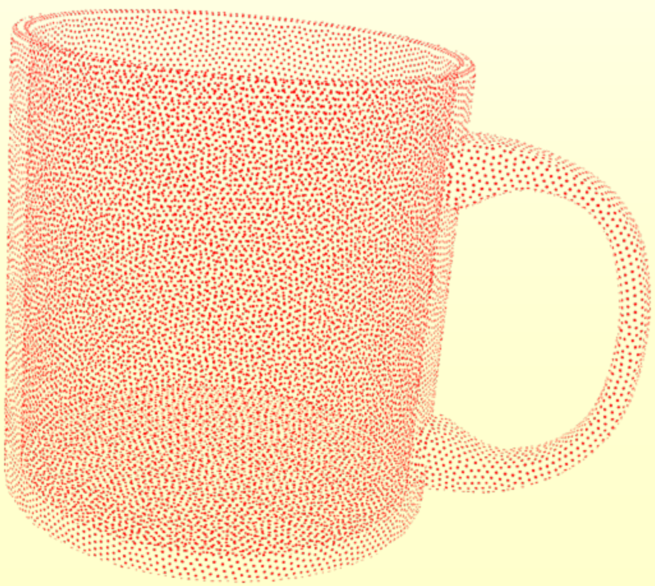
[Ovsjanikov, Sun, G., SGP'08,
Sun, Ovsjanikov, G., SGP'09]



Extrinsic vs. Intrinsic

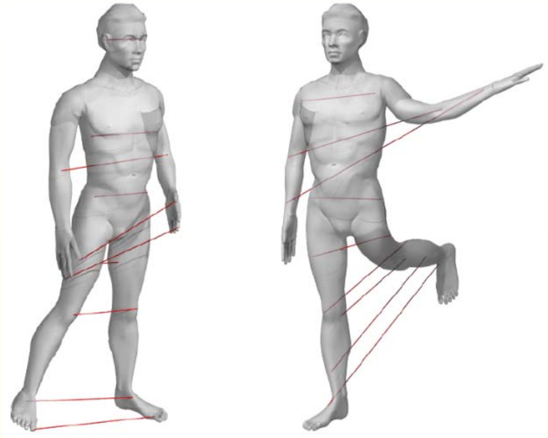
- Most multi-scale methods of geometric analysis, e.g. wavelets, require explicit parametrizations of the geometry, e.g. coordinate functions
- What if we have only metric, or distance information?
- And what if the distances are intrinsic, not extrinsic?





Extrinsic vs. Intrinsic Symmetries

Extrinsic Symmetry



Intrinsic Symmetry

• Invariance under translation, rotation, reflection and scaling (Isometries of the ambient space)

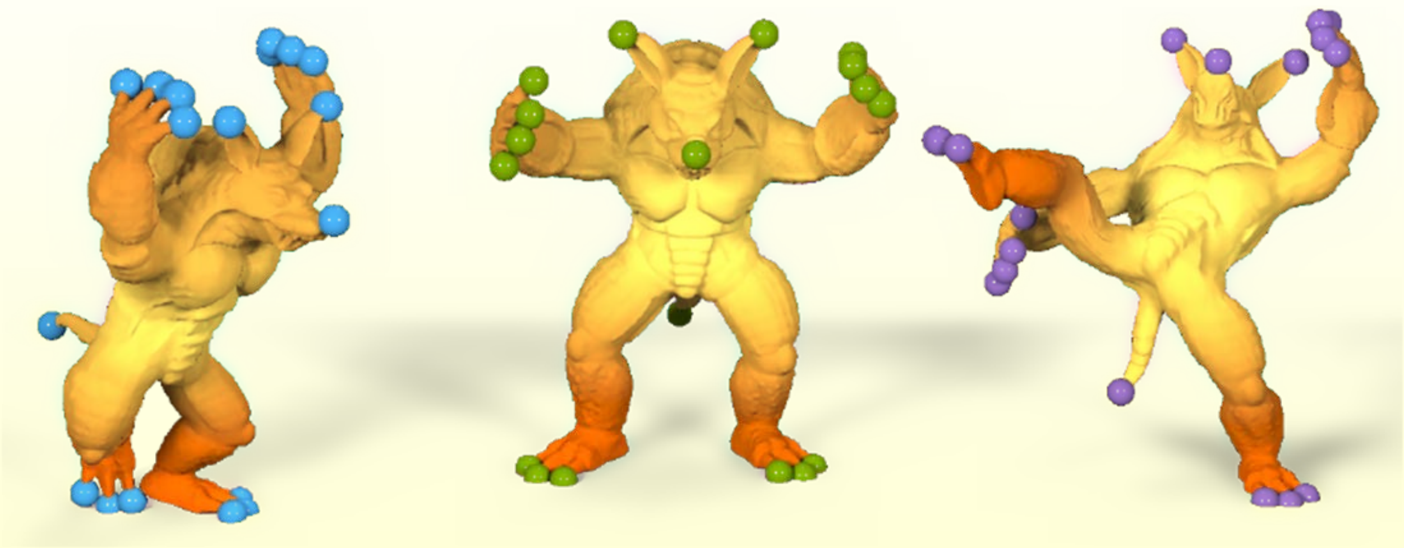
• Break under isometric deformations of the shape

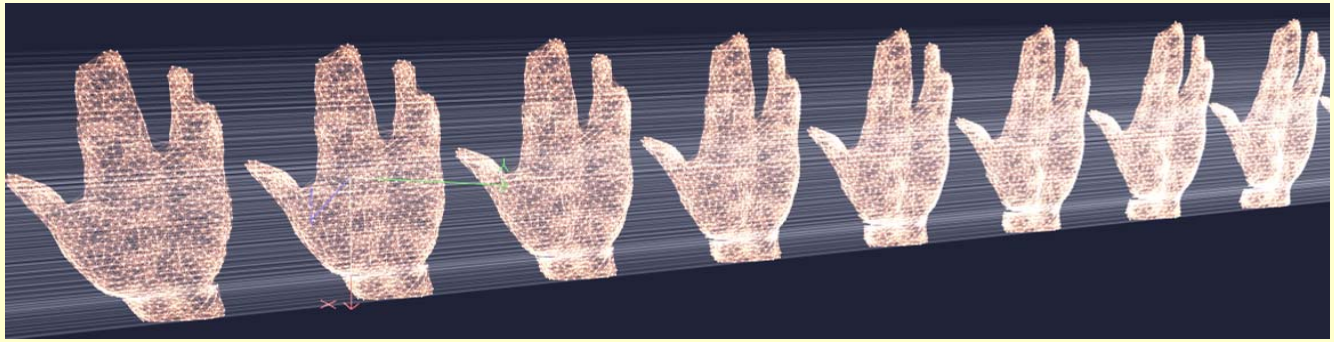
• Invariance of geodesic distances under self-mappings. For a homeomorphism $T : O \rightarrow O$

$$g(\mathbf{p}, \mathbf{q}) = g(T(\mathbf{p}), T(\mathbf{q})) \ \forall \ \mathbf{p}, \mathbf{q} \in O$$

• Persist under isometric deformations

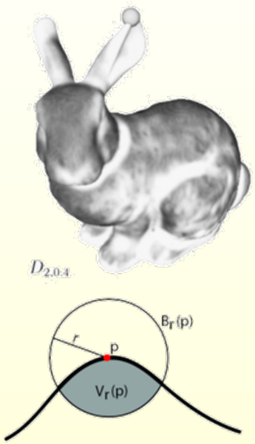
Correspondences are Often Based on Descriptors



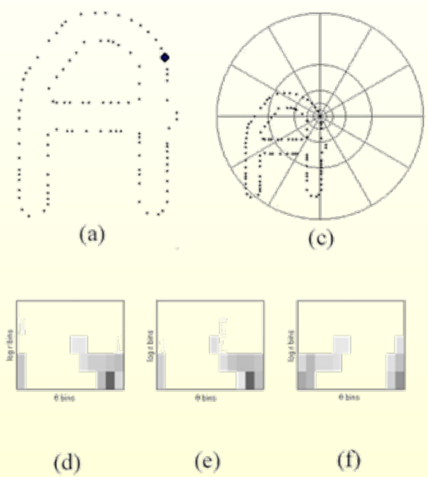


Shape Descriptors

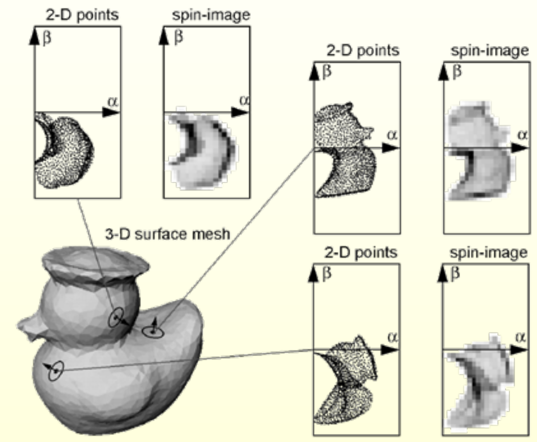
- For shapes, there are many descriptors invariant to rigid motions:



Integral Invariants:
Manay et al. ’04
Pottmann et al. ‘09



Shape Contexts:
Belongie et al. ’00
Frome et al. ‘04

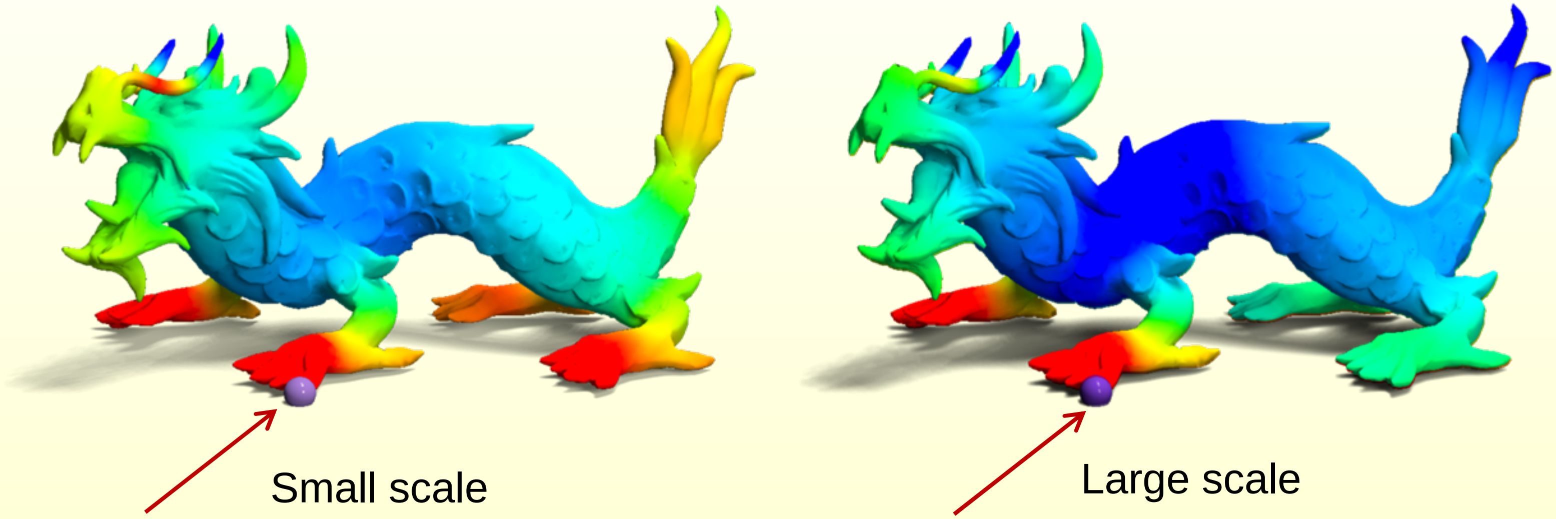


Spin Images:
Johnson, Hebert ‘99

- Many tradeoffs among different descriptors …
- But what about intrinsic descriptors? Heat kernel signatures

The Issue of Scale

- Given a point (●) on a shape, find other points with “similar” neighborhoods



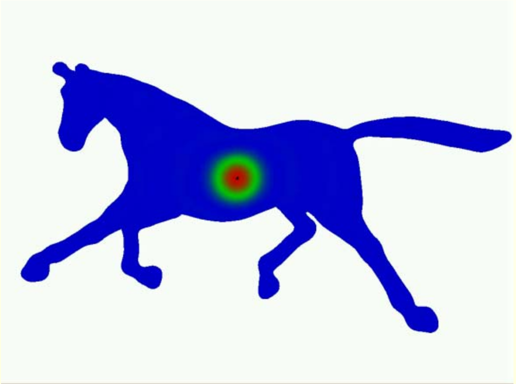
- Inherently multiscale question: on a manifold, locally all points are the same. Need a meaningful way to compare point neighborhoods at different scales
- At what scale do neighborhoods become unique?

Background

- Heat equation on a Riemannian manifold:

  If $u(x,t)$ is the amount of heat at point x at time t, then

$$\frac{\partial u}{\partial t} = \Delta u$$

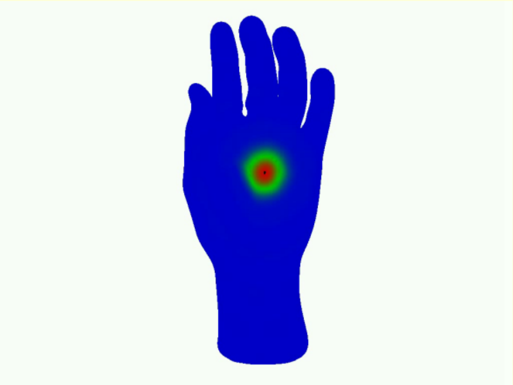


Δ : Laplace-Beltrami Operator (div grad)

- Given an initial distribution $f(x)$. After time t:

$$f(x,t) = e^{-t\Delta} f$$

H_t heat operator

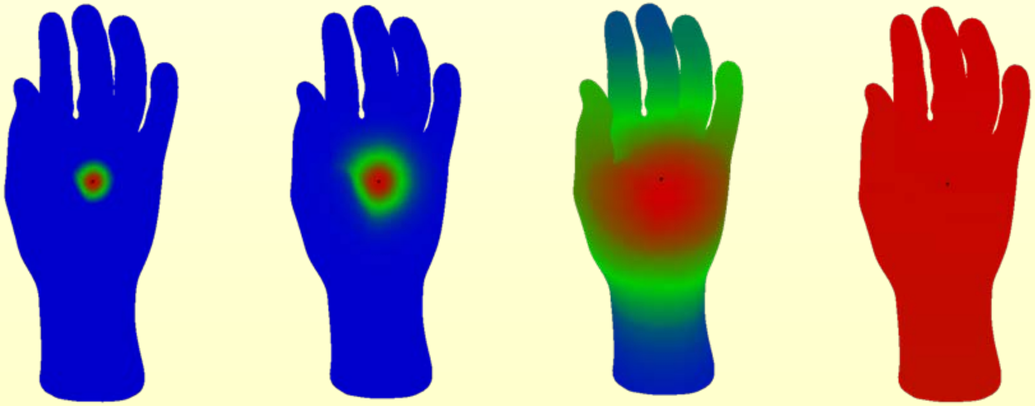


Background

- Heat kernel $k_t(x,y)$:

$$f(x,t) = \int_{\mathcal{M}} k_t(x,y) f(y) dy$$

$k_t(x,y)$: amount of heat transferred from x to y in time t. How well x and y are connected at scale t -- an integral over all paths from x to y



$t = 0.001$ $t = 0.02$ $t = 3$ $t = 40$

Heat Kernel Properties

Basic Properties

- $k_t(x, y) = k_t(y, x)$
- $k_{t+s}(x, y) = \int_M k_t(x, z) k_s(z, y) dz$
- $k_t(x, y) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \phi_i(x) \phi_i(y)$

LB eigenvalues and eigenvectors

Heat Kernel Properties

- Invariant under isometric deformations

  If $T : X \to Y$ is an isometry, then:

$$k_t(x, y) = k_t(T(x), T(y))$$

- Conversely: it characterizes the shape up to isometry.

  If $k_t(x, y) = k_t\,(T(x), T(y))\ \forall\ x, y, t$ then

  $\tilde{M}$ is an isometry

This is because:

$$\lim_{t \downarrow 0} (t \log k_t(x, y)) = -\frac{1}{4} d^2_{\mathcal{M}}(x, y)\ \forall\ x, y$$

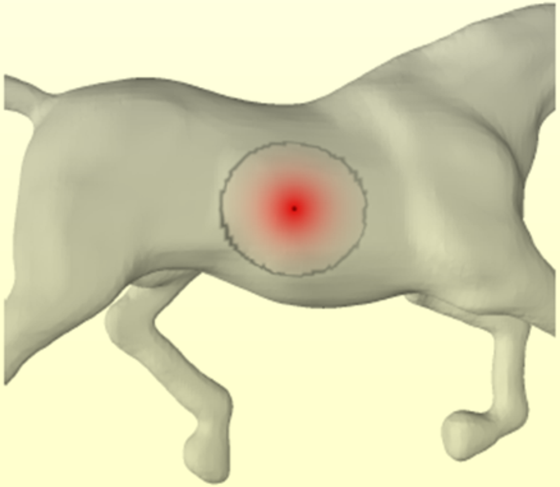
where $d_{\mathcal{M}}(\cdot, \cdot)$ is the geodesic distance

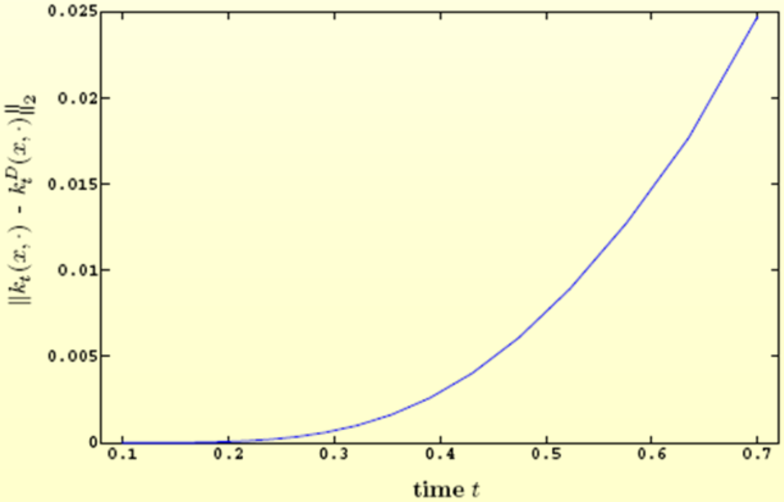
Heat Kernel Properties

- Multiscale:

  For a fixed x, as t increases, heat diffuses to larger and larger neighborhoods

  Therefore, $k_t(x, \cdot)$ is determined by (reflects the properties of) a neighborhood that grows with t



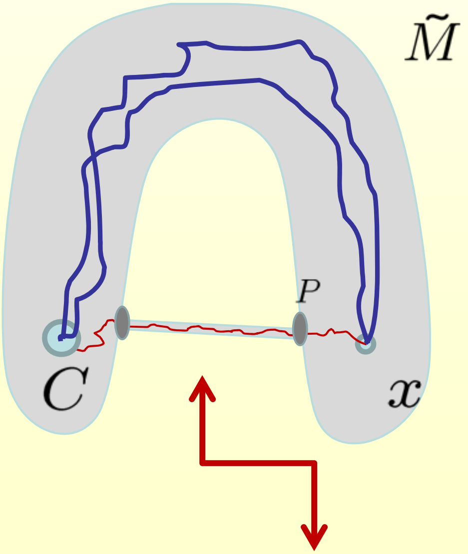


truncation effect

Heat Kernel Properties

- Robustness:

  $k_t(x,\cdot)$ is the probability density function of BM, a weighted average over all paths, which is generally not very sensitive to small perturbations



$$k_t^{\tilde{M}}(x,C) = \mathbb{P}(\tilde{W}_x^t \in C)$$

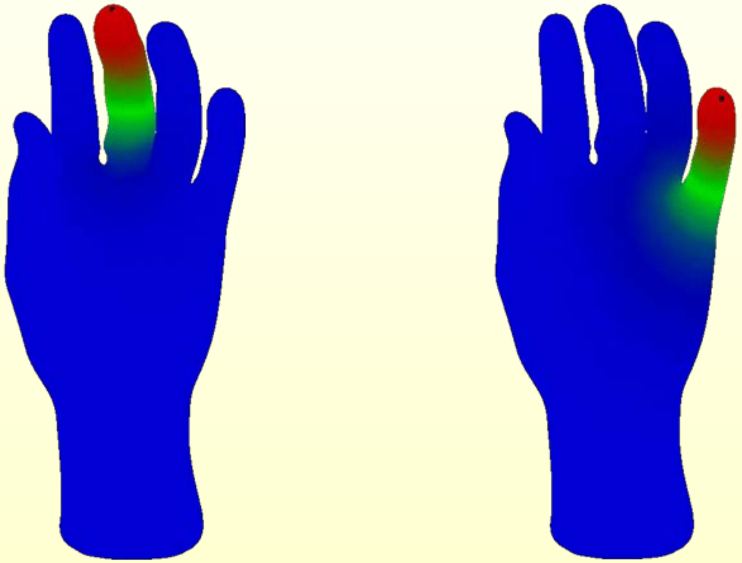
Only paths through the modified area P will change

Defining a Signature

- Let $k_t(x, \cdot)$ be the signature of x at scale t

  The heat kernel has all the properties we want

  Except easy comparison …



- $k_t(x, \cdot)$ is a function on the entire manifold
- Nontrivial to align the domains of such functions across different shapes, or even for different points of the same shape

Defining a Signature

- Let $k_t(x,\cdot)$ be the signature of x at scale t

  The heat kernel has all the properties we want.

  Except easy comparison …

- We define the Heat Kernel Signature (HKS), by restricting to the diagonal of the kernel:

$$\mathrm{HKS}(x) = \{k_t(x,x), t \in \mathbb{R}^+\}$$

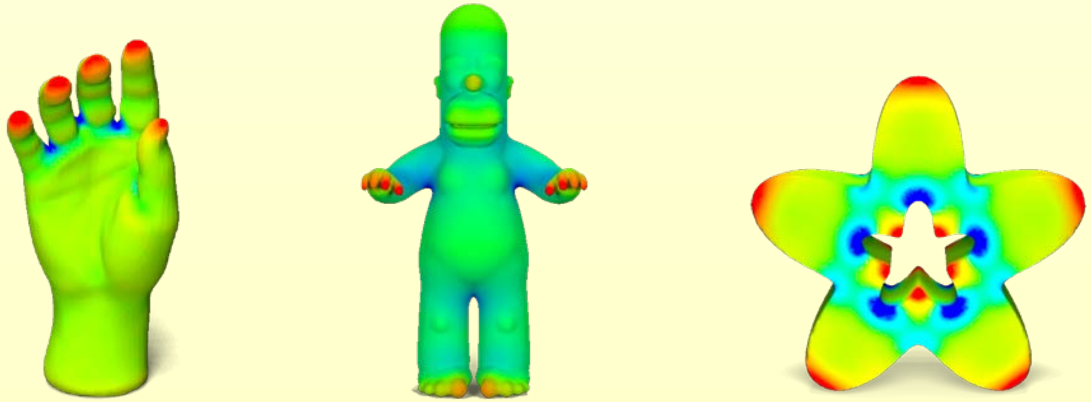
- Now HKSs of any two points can be easily compared, since they are defined on a common domain (time)

Defining a Signature

- Since HKS is a restriction of the heat kernel, it is:
  - Robust
  - Multiscale

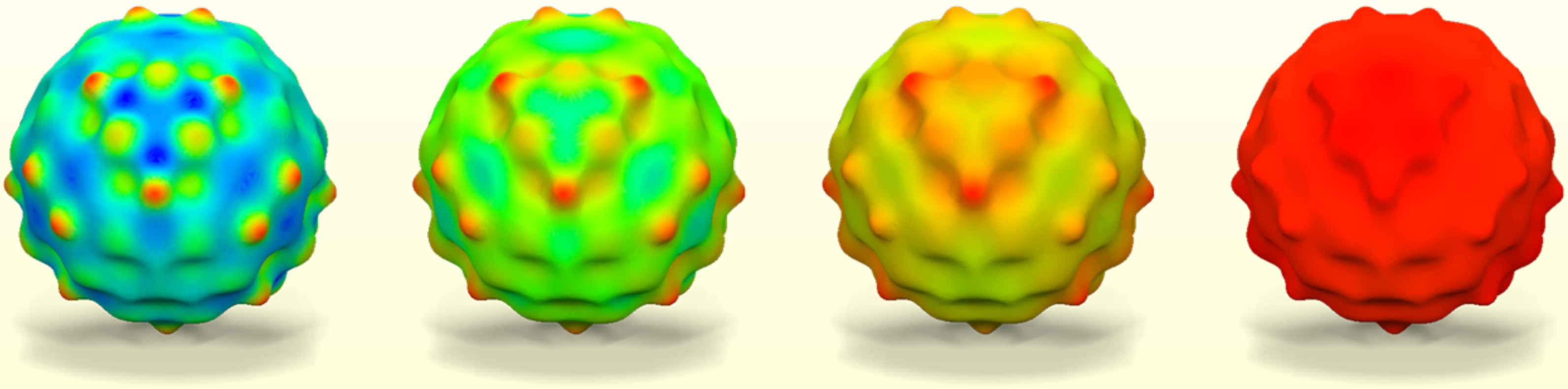
- Question 1: How informative is it?
  - Related to Gaussian curvature for small t :

$$k_t(x,x) = \frac{1}{4\pi t}\sum_{i=0}^{\infty} a_i t^i \qquad a_0 = 1, a_1 = \frac{1}{6}K$$



Defining a Signature

- HKS can be interpreted as a multiscale, robust, intrinsic curvature:



$t = 0.004$ $t = 0.008$ $t = 0.02$ $t = 2$

- HKS computational aspects omitted in this talk

Informative Theorem

- The set of all HKSs on a shape almost always defines it up to isometry

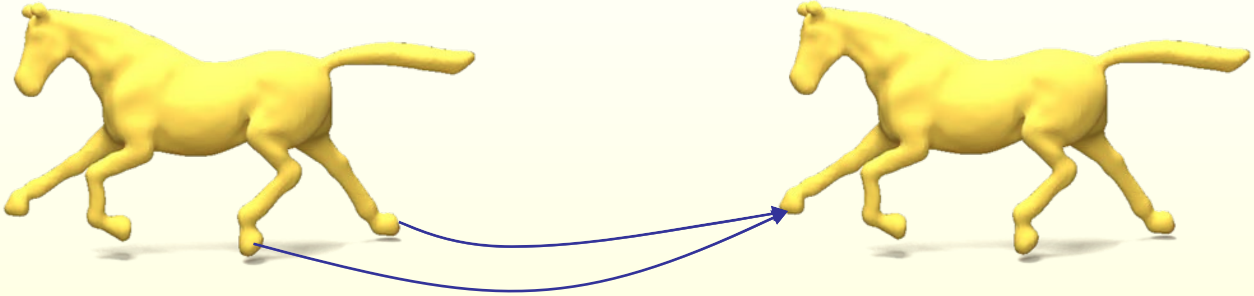
- Theorem: If X and Y are two compact manifolds, such that Δ_X and Δ_Y have only non-repeating eigenvalues, then a homeomorphism $T : X \to Y$ is an isometry if and only if, for all $\boldsymbol{x}$

$$\mathsf{HKS}(x) = \mathsf{HKS}(T(x))$$

- The set of all HKSs characterizes the intrinsic structure of the manifold!

Applications of HKS

- Multi-scale matching, structure discovery



- Feature extraction

Multiscale Matching

- Two heuristics for making HKSs comparisons practical:
  - For a fixed point x, sample HKS on a logarithmic scale at times t_i
  - For a fixed time t scale each HKS, by the sum over all points of *M*

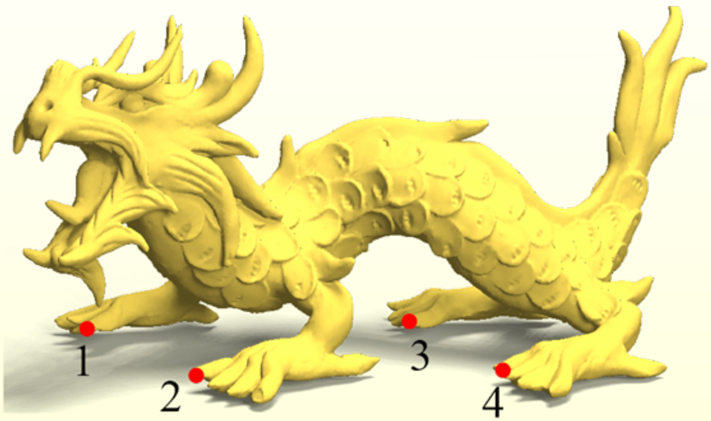
$$\mathrm{HKS}(x) = \left\{ \frac{k_{t_i}(x,x)}{\sum_j e^{-t_i \lambda_j}}, i \in 1, 2, .., 100 \right\}$$

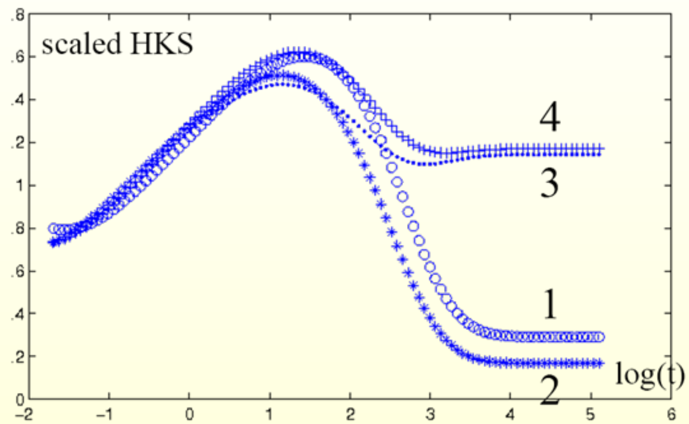
$$t_i = \alpha^i t_0$$

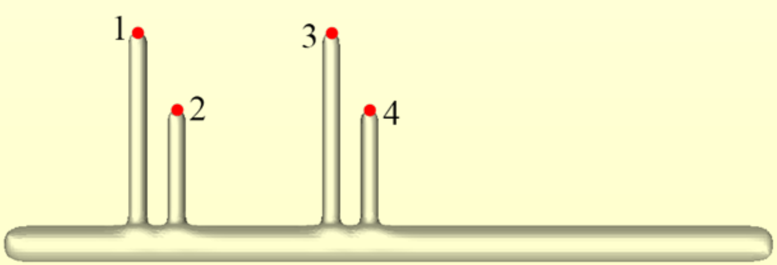
- Compare using L2 norm of these HKS vectors

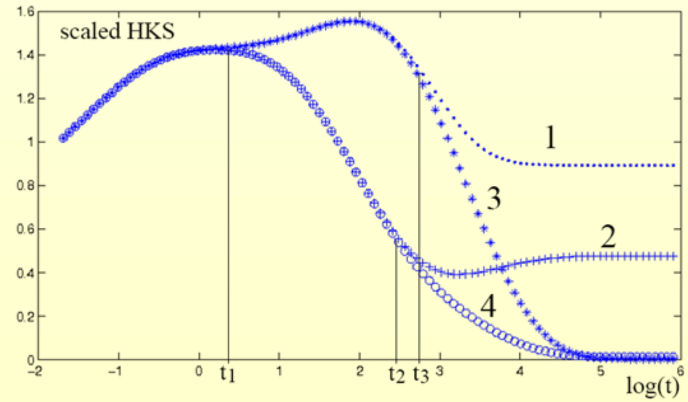
Multiscale Matching

- Comparing points through their HKS signatures:









Multiscale Matching

- Finding similar points – robustly:

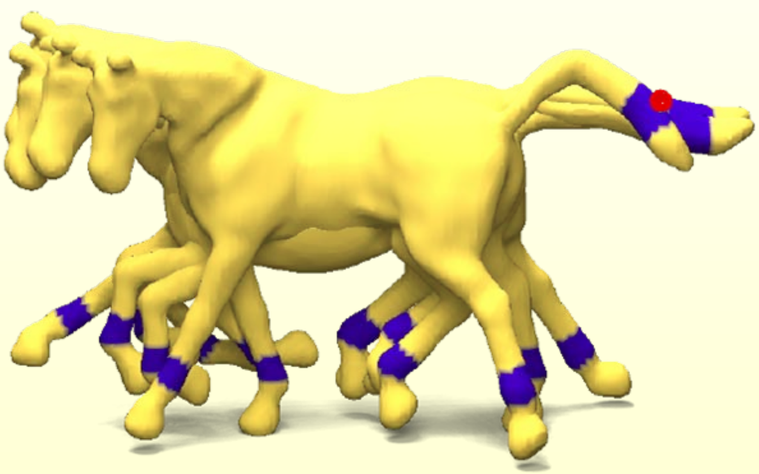
Medium scale

Full scale

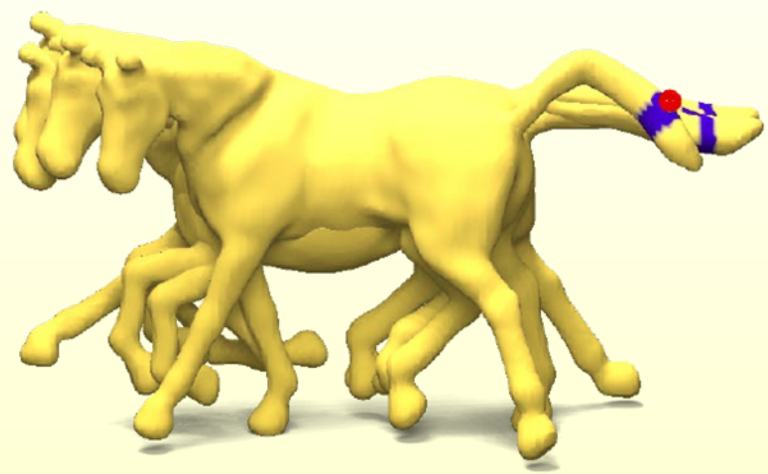
Armadillo

Multiscale Matching

- Finding similar points across multiple shapes:



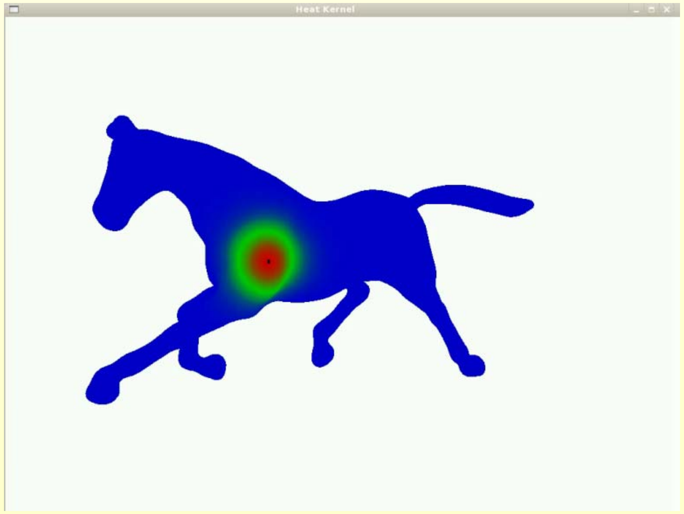
Medium scale

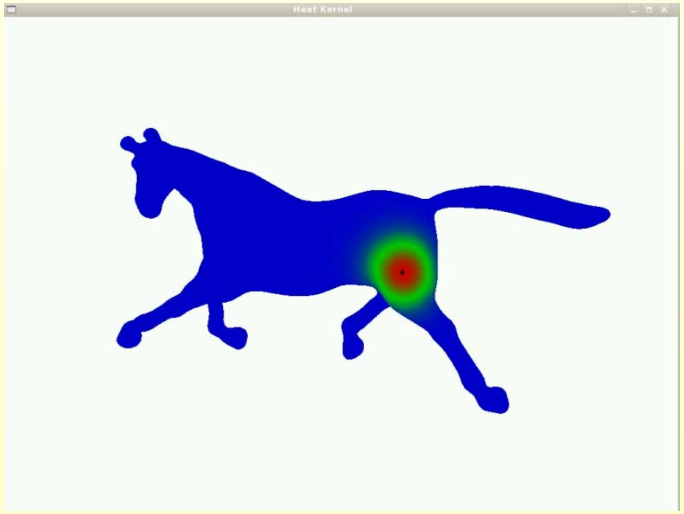


Full scale

Feature Detection

- Persistent feature detection:
  - Intuition: heat diffuses slower at points with high curvature. Heat will tend to concentrate in “hot spots” – extremities of the surface
  - Approach: track the local maximum of the heat kernel for increasing t

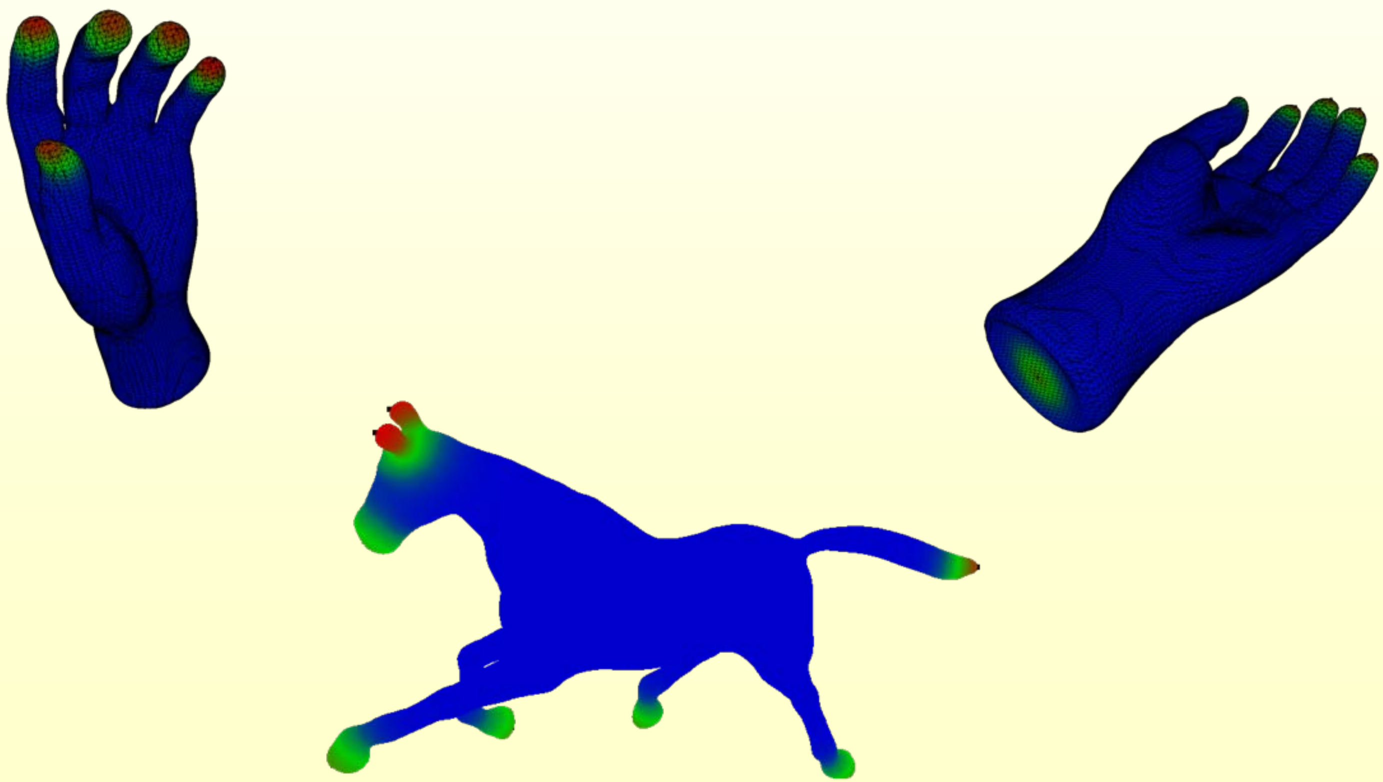




Feature Detection

- Persistent feature detection:
  - Find points that are long term maxima of their heat kernels:

$$k_t(x, \cdot)$$



Feature Detection

- Persistent feature detection:
  - Find points that are long term maxima of their heat kernels:

$$k_t(x, \cdot)$$

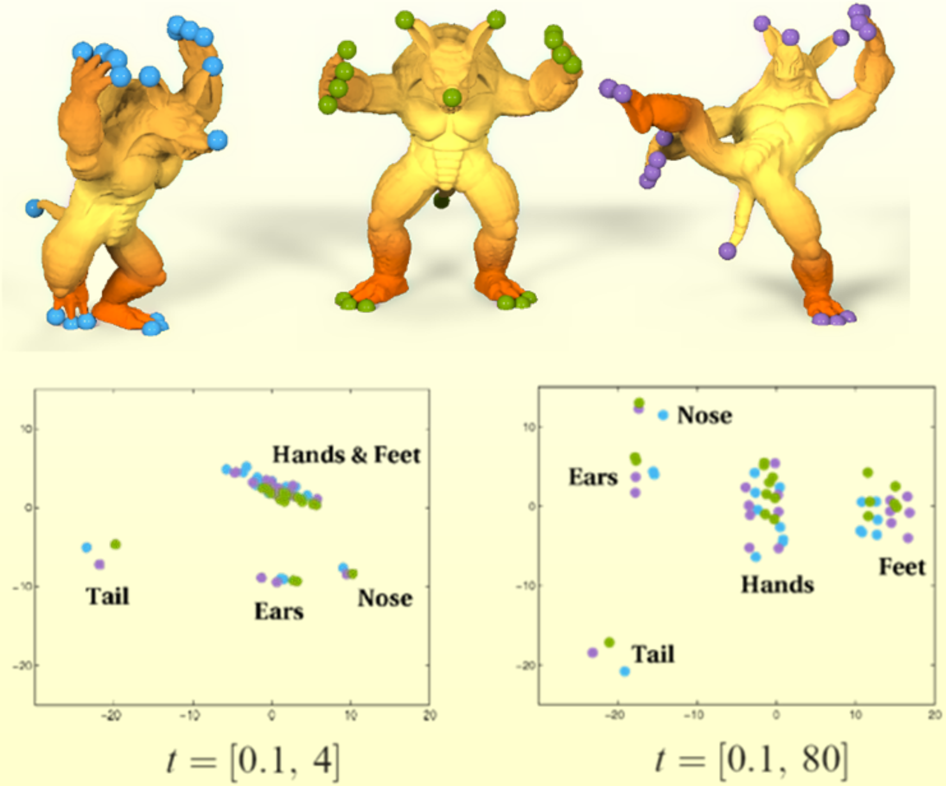
  - This may be expensive since the heat kernel at every point is a function over the whole shape. However, long term behavior at nearby points is similar due to mixing
  - Approximation: find points that are local maxima of

$$k_t(x, x)$$

    for large enough t

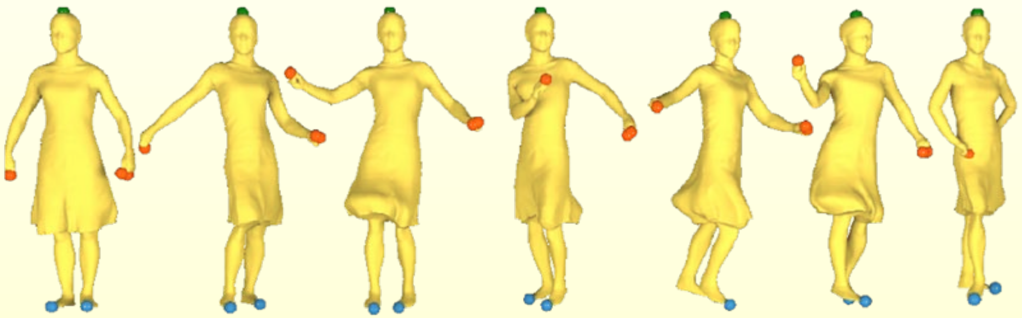
Shared Structure

- 2D MDS embedding of feature points on three shapes according to distances of their HKS

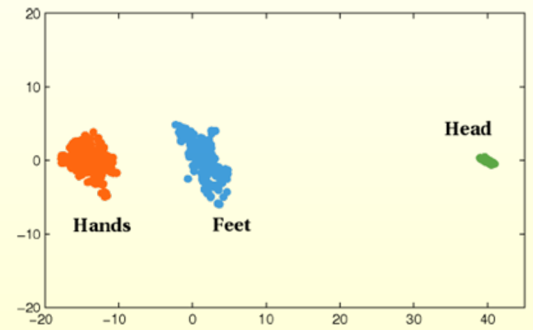


Shared Structure

- 2D MDS embedding of feature points on 175 shapes according to distances of their HKS.



Feature points found on a few poses of the dancer model by Vlasic *et al.*

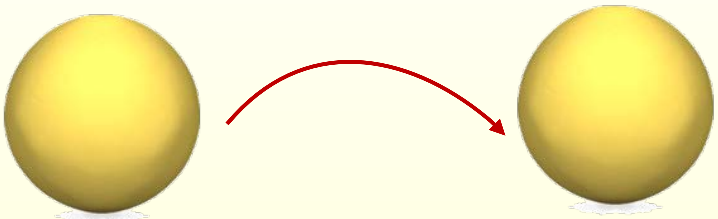


MDS of features from all 175 poses using a full range of scales

Partial and approximate intrinsic symmetries can be detected this way

Informative Theorem

- How general is the theorem?
  - If there are repeated eigenvalues, it does not hold:



On the sphere, $\mathrm{HKS}(x) = \mathrm{HKS}(y)\ \forall\ x, y$ but there are non-isometric maps between spheres.

- Do not know if an "approximate" version of the theorem is true, but suspect so

Intrinsic Measures of Shape Similarity

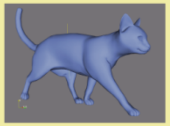
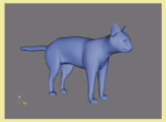
- Gromov-Hausdorff distance: a second order optimization over correspondences

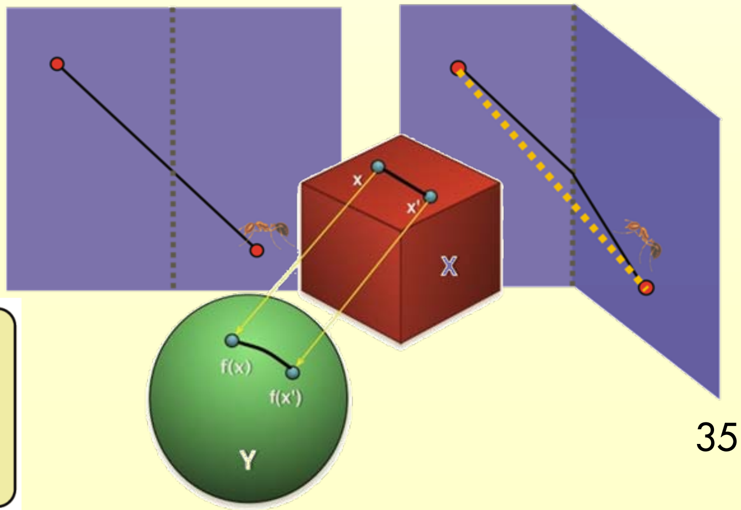
$$\Gamma_{X,Y}(x,y,x',y') := \left| d_X(x,x') - d_Y(y,y') \right|$$

intrinsic distance distortion

$$d_{\mathcal{GH}}(X,Y) = \frac{1}{2} \inf_{R} \|\Gamma_{X,Y}\|_{L^\infty(R \times R)}$$

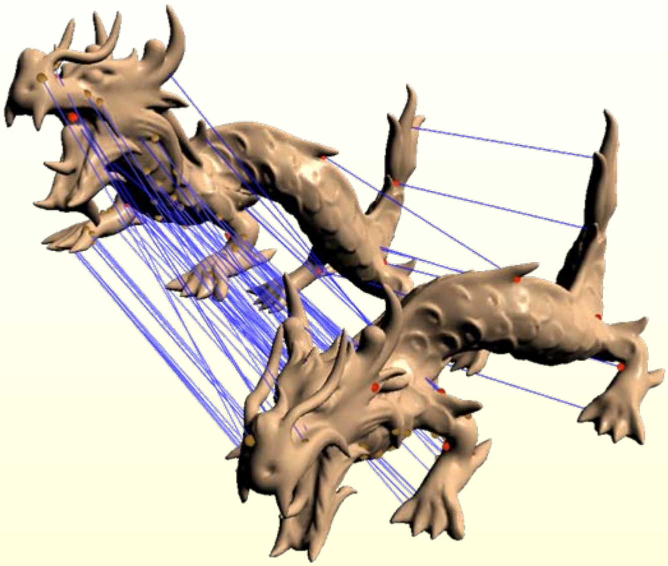
evaluated via intrinsic distances

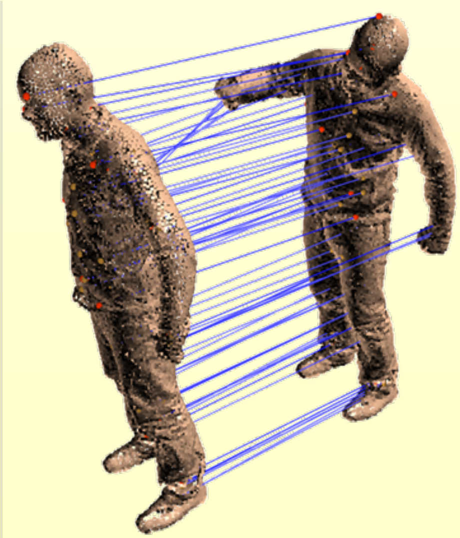
dist( **,**  **) = 0 ?**



Are There Perfect Signatures?

- To optimally align two shapes, is it sufficient to optimally align their point signatures, or certain features derived from these signatures?
- Optimal alignment can be defined in terms of certain intrinsic but hard-to-compute shape distances, such as Gromov-Hausdorff
- If this is so, then we only have a first-order optimization problem to solve …
- Of course this can fail if there are symmetries …





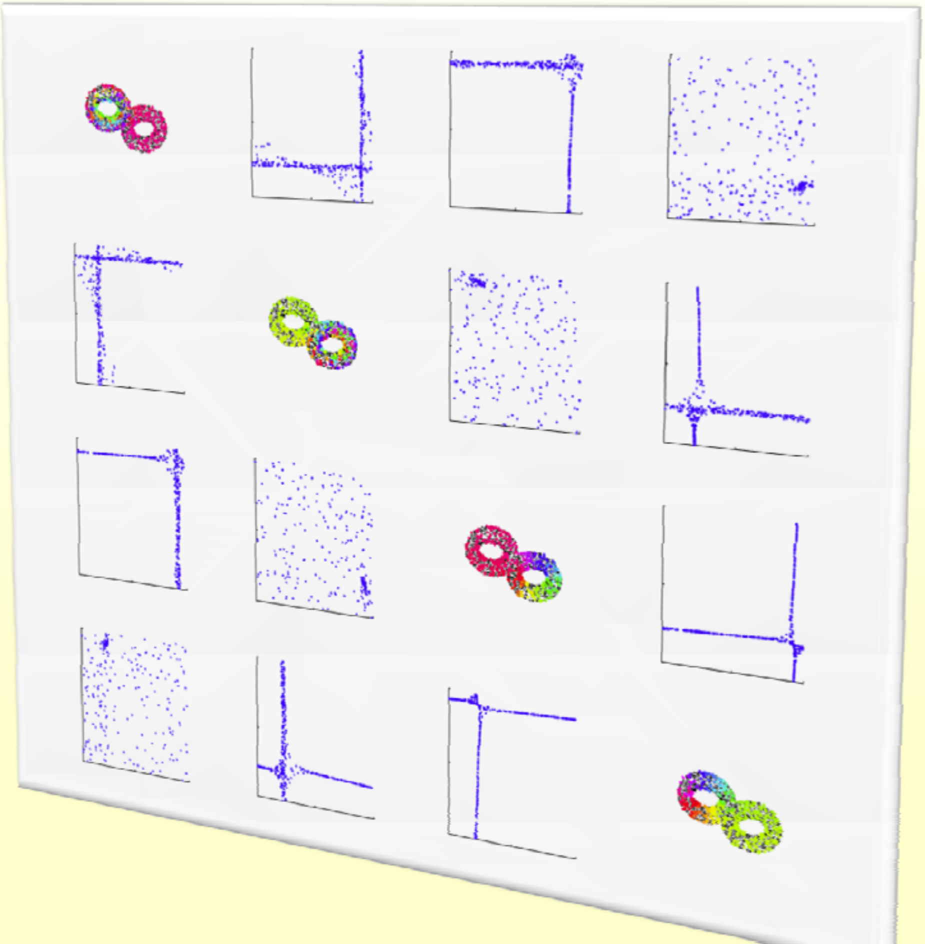
[data sets: Stanford 3D Scanning Repository / Carsten Stoll]

Key Points and Issues

- Heat kernel signatures (HKS) provide a powerful tool for describing shape neighborhoods. They are
  - Robust
  - Multiscale
  - Informative. Related to curvature and geodesics
  - Easily computable
- They can be used to
  - Provide point signature for multiscale matching
  - Extract shape features
  - Discover intrinsic symmetries
  - Study a formal spectral metric between shapes

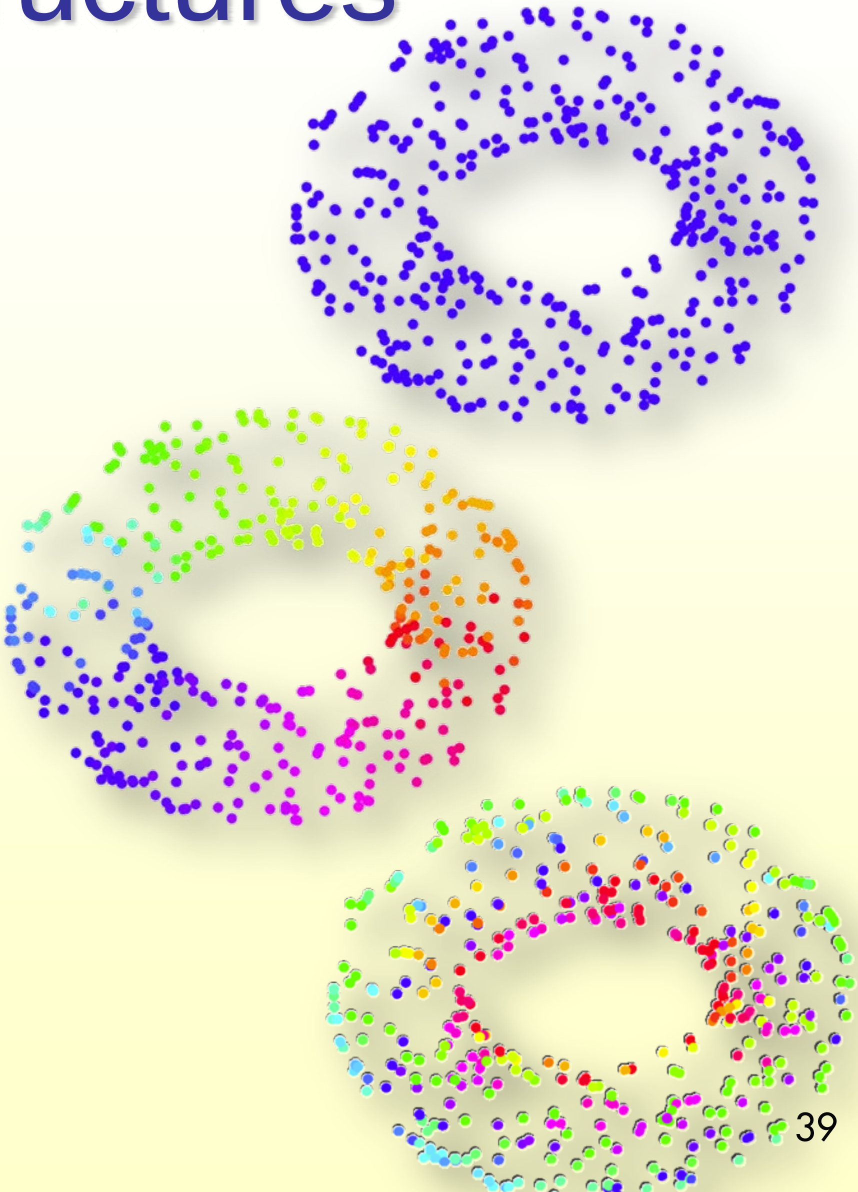
II. Circular Coordinates for Data Sets

[de Silva, Morozov, Vejdemo-Johansson, SoCG'09]



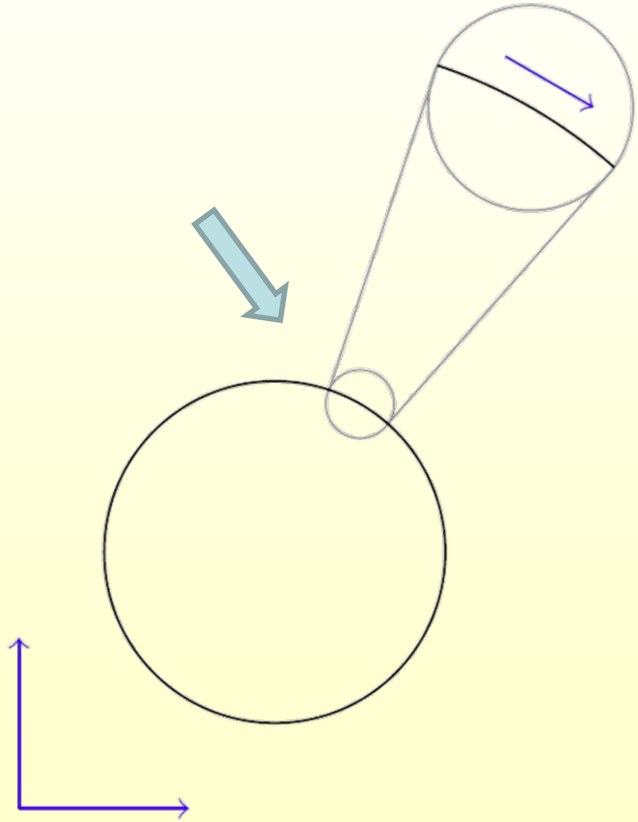
Circular Structures

- Circular structures are often present in data
- Classically
  - Linear coordinatization: find linear transformations from *X* to R^d
  - Principal component analysis, projection pursuit
- Recently
  - Non-linear methods: drop the expectation of linearity for the transformation
  - MDS, kernel methods, locally linear methods



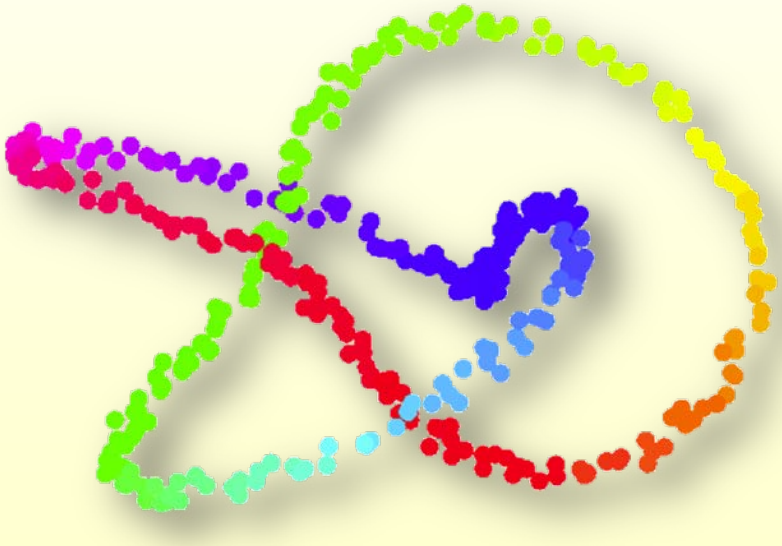
Problematic Cases

- Some shapes take up *too many* coordinates
- Circle - locally 1-dimensional, globally needs 2 coordinates
- Torus - locally 2-dimensional, globally needs 3, or even 4 coordinates



How Can We Fix This?

- Circle-valued coordinates
  - Use $S^1 = [0, 1]/(0 \sim 1)$ as an additional coordinate space
  - Fixes the circle
  - Fixes the torus
  - Occurs naturally:
    - Phase coordinates for waves
    - Angle coordinates for directions



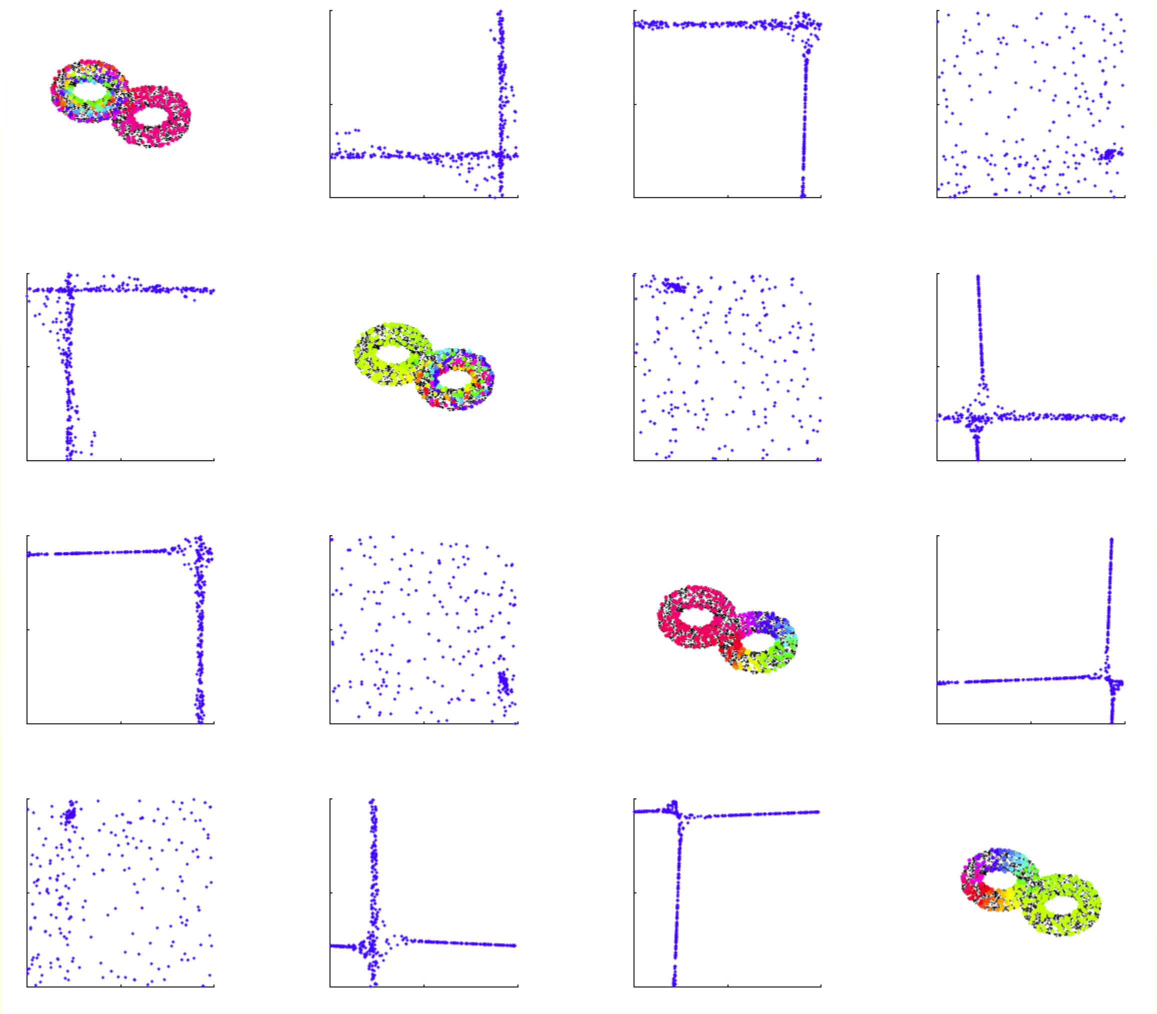
Approach

- Exploit canonical isomorphism

$$H^1(X;Z) \cong [X, S^1]$$

- Use persistent cohomology to pick out features
  - Compute over Z_p, for several p
- Use least-squares smoothing to generate *nice* circle-valued functions from cocycles
- Cohomology is calculated with variant on the persistence algorithm: coboundaries are computed and matched for consecutive simplices

Double Torus Correlation Plots

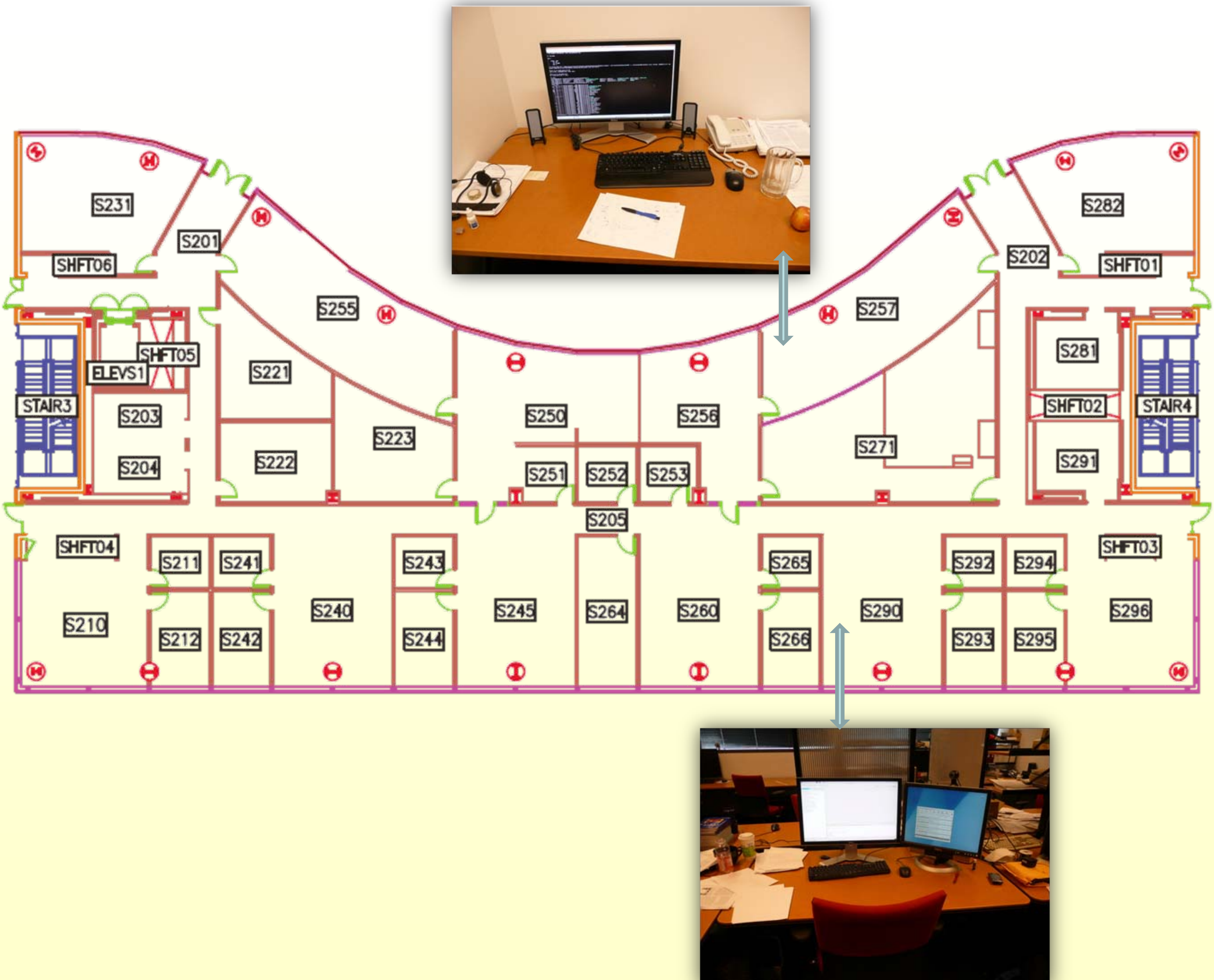


Key Points and Issues

- Circular structures are very common in real data
- Linear structures can also be discovered this way, by appropriate identification of endpoints
- The need for such parametrizations arises in many other problems
- The cohomology persistence algorithm is very lightweight and fast (faster than regular persistence)

III. Interlinked Image Collections

[Heath, Gelfand, Ovsjanikov, Aanjenaya, G., ‘09]


S231
S201
SHFT06
S255
S257
S202
S282
SHFT01
SHFT05
ELEVS1
S221
STAIR3
S203
S204
S222
S223
S250
S256
S271
S281
SHFT02
STAIR4
S291
S251
S252
S253
S205
SHFT04
S211
S241
S243
S265
S292
S294
SHFT03
S210
S240
S245
S264
S260
S290
S296
S212
S242
S244
S266
S293
S295

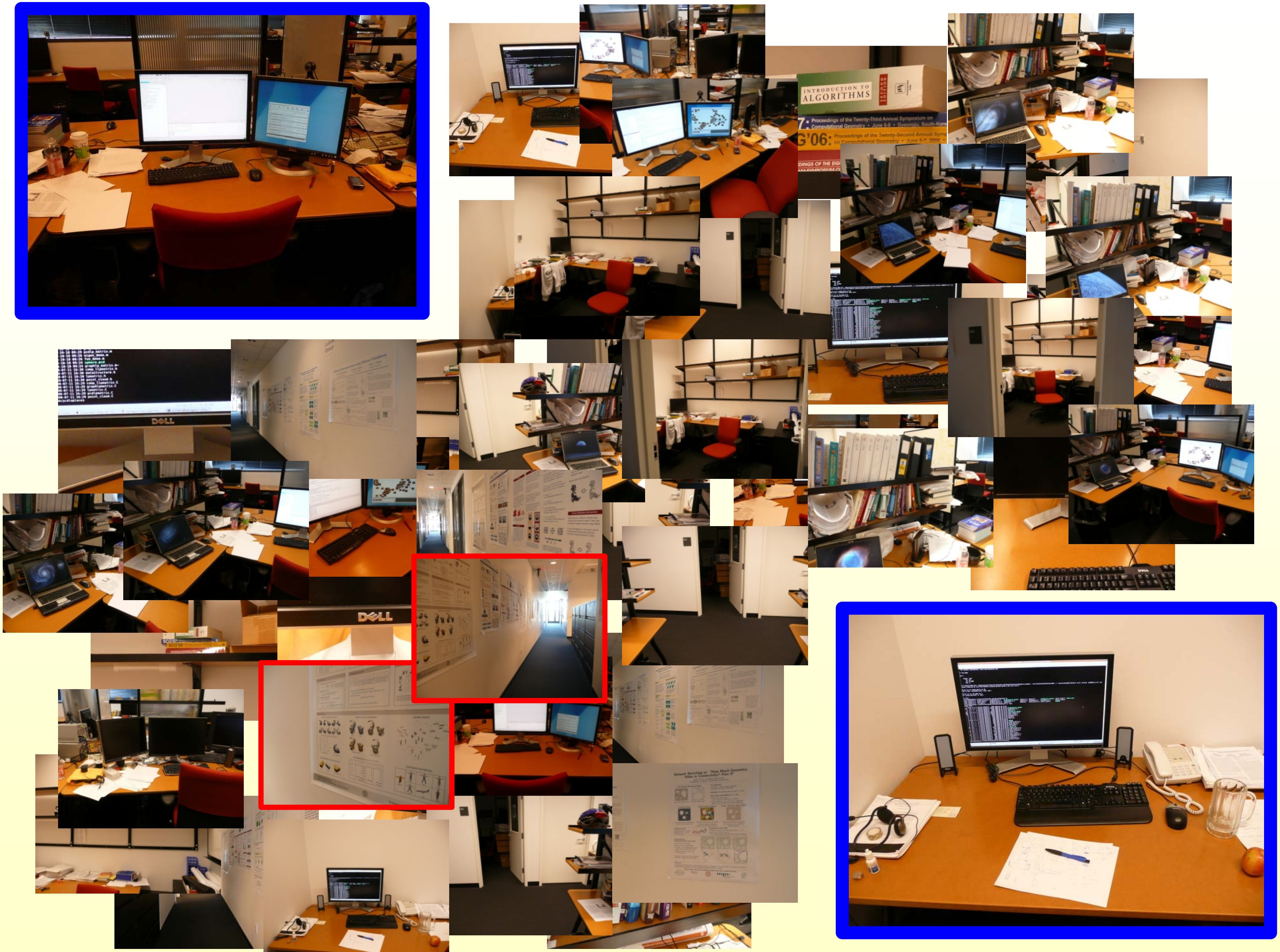
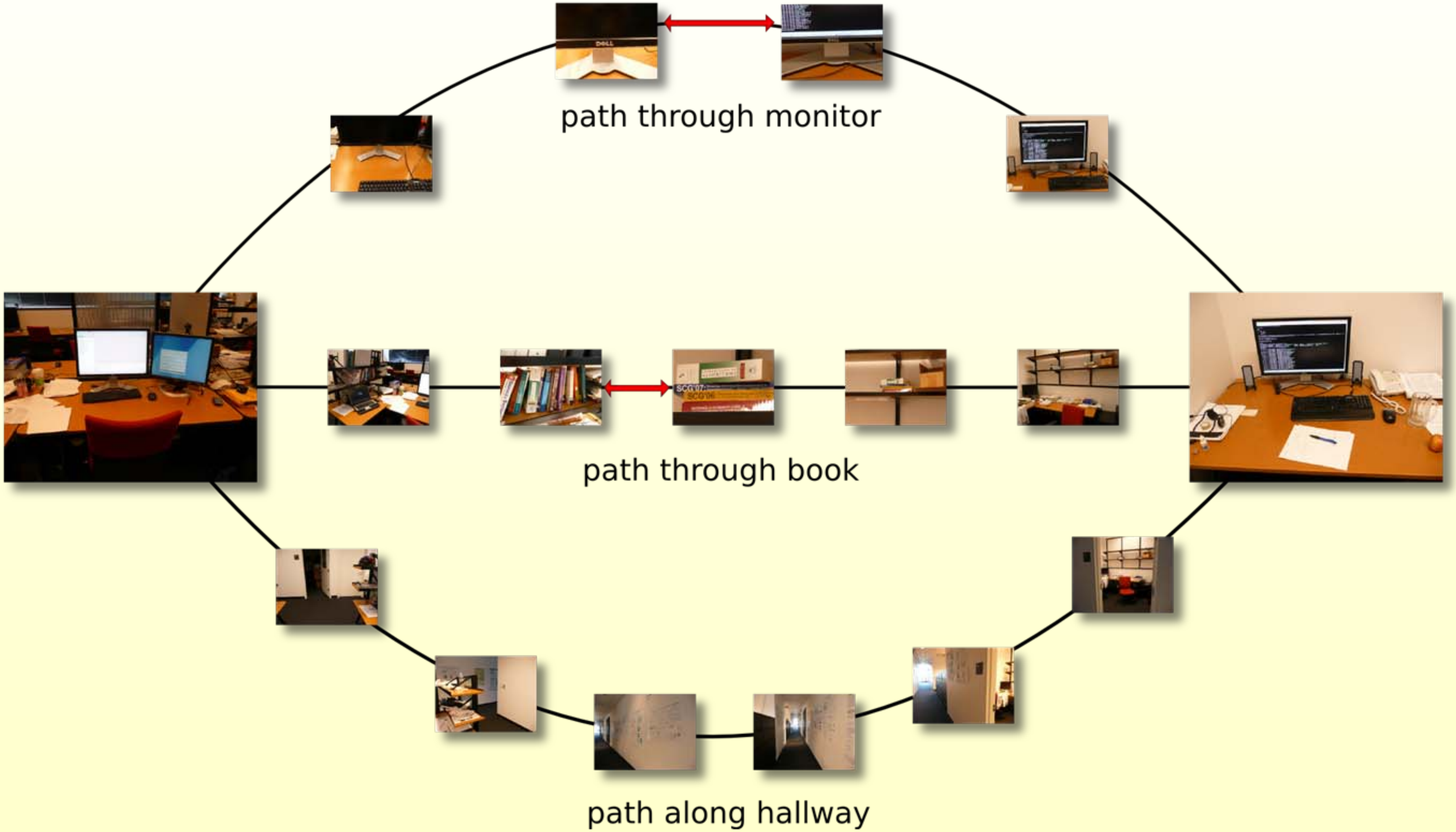

INTRODUCTION TO ALGORITHMS
DELL

Image Match Links

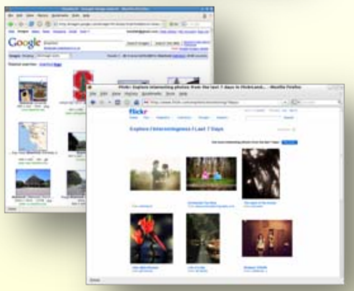
Paths Through Image Collections

Homotopy Classes



Large Scale Image Acquisition

- Acquiring, storing, and sharing large image collections is becoming easier and easier
  - Ubiquitous cell phone cameras
  - Inexpensive storage
  - Wireless networking
- Photo sharing sites (e.g., Flickr, Picasa)
- Systematic commercial acquisition projects (e.g., Google Streets)
- Camera sensor networks



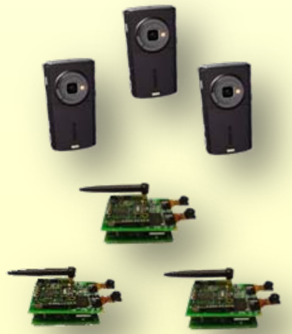
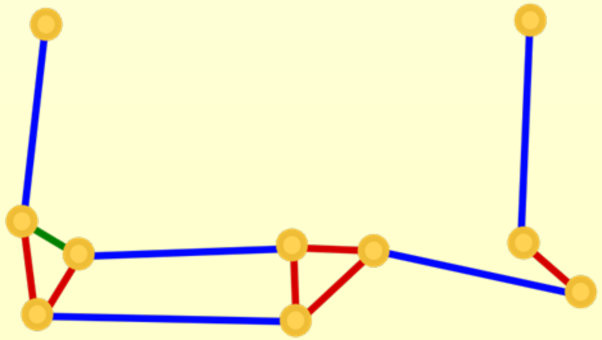


Image Webs

- The idea of Image Webs is to interlink images through a variety of link types, based on both content and image metadata (GPS, time)
- *The same way that the WWW of documents has proved useful, the hope is that interlinked webs of signals will also be valuable for propagating, extracting, and filtering information – and the web types two can cross-link and cross-fertilize*



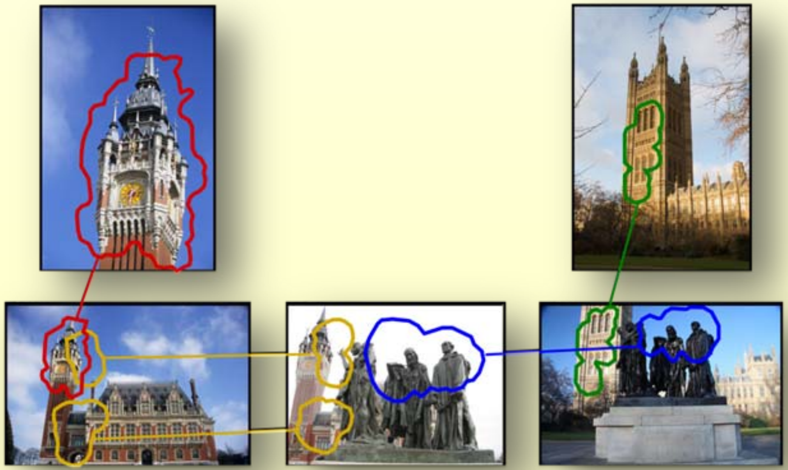
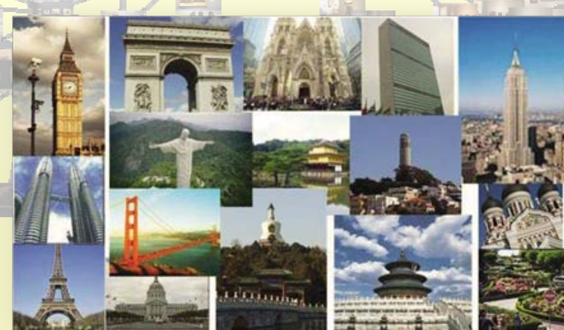


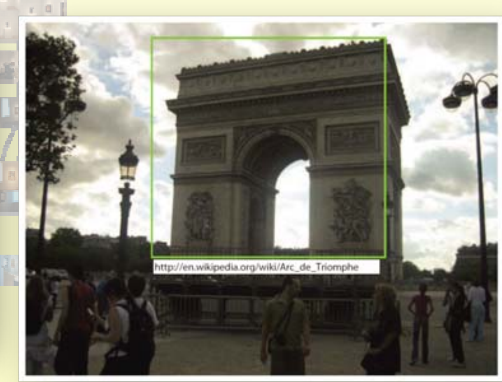
Image Webs Agenda

- Understand the local and global structure of image webs, aiming at a softer, more topological understanding
- Develop efficient construction algorithms
- Explore applications (image browsing, annotation transfer, social networks, etc.)

[Snavely et. al., Siggraph ’06]



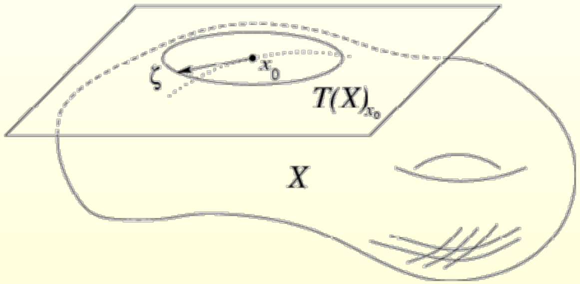
[Zheng et. al., CVPR 2009]



[Gammeter et al., ICCV 2009]

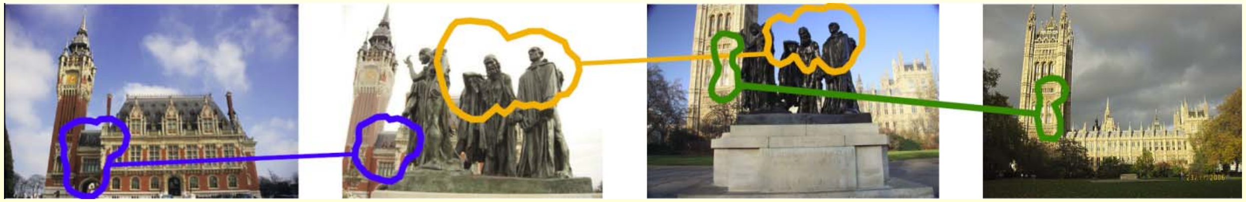
The Space of All Images

- If we frieze time, the local structure of the space of images is well understood: it that of a low dimensional manifold – the manifold of views
- This is also the local structure of an image web based on match links

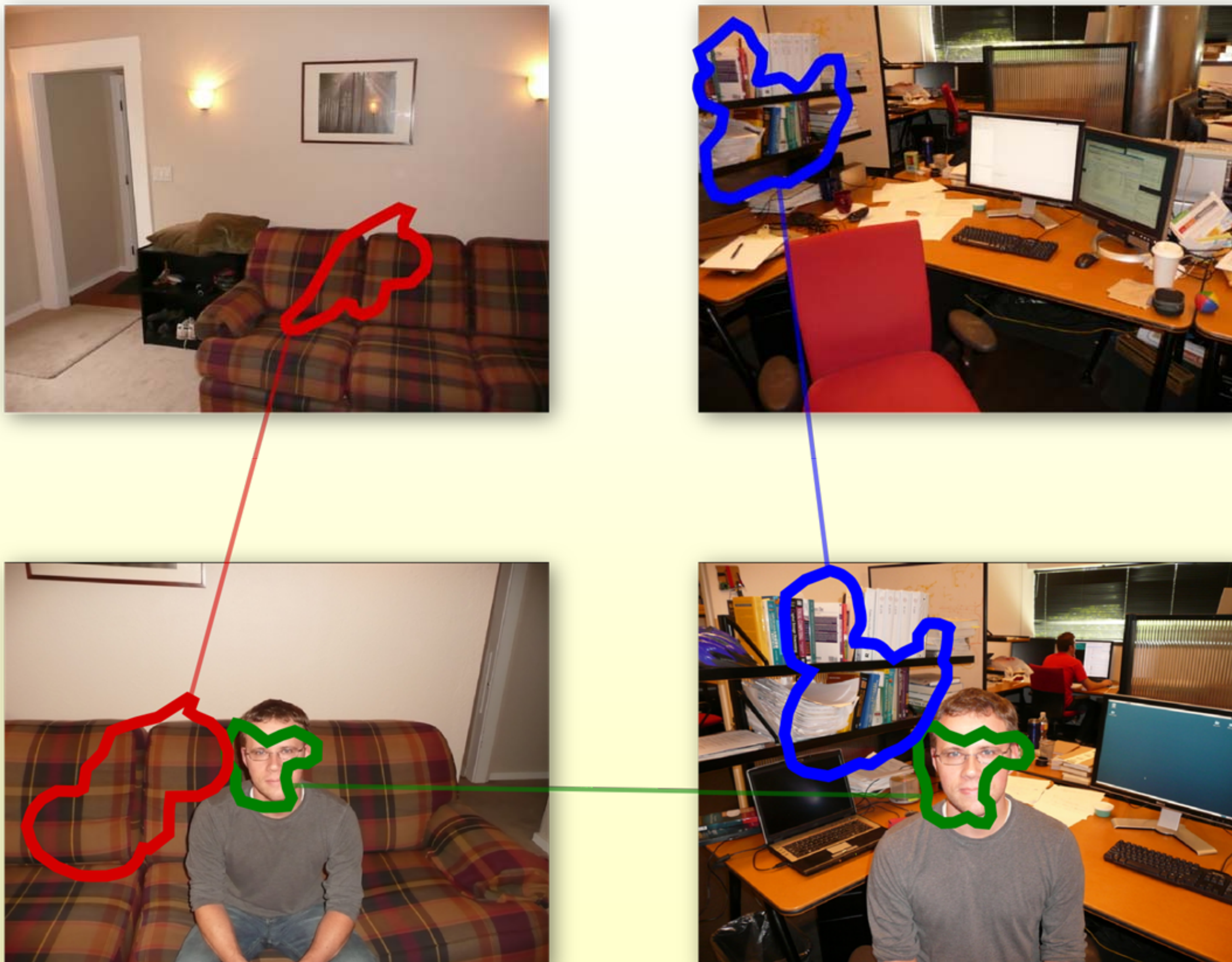


- But at larger scales the structure is more complex
  - because of moving objects
  - because of repeated similar objects
- For us this is exactly the structure that is of interest

Non-Local Links



Proximity Through Mobility: Home to Office



Proximity Through Mobility on the Stanford Campus

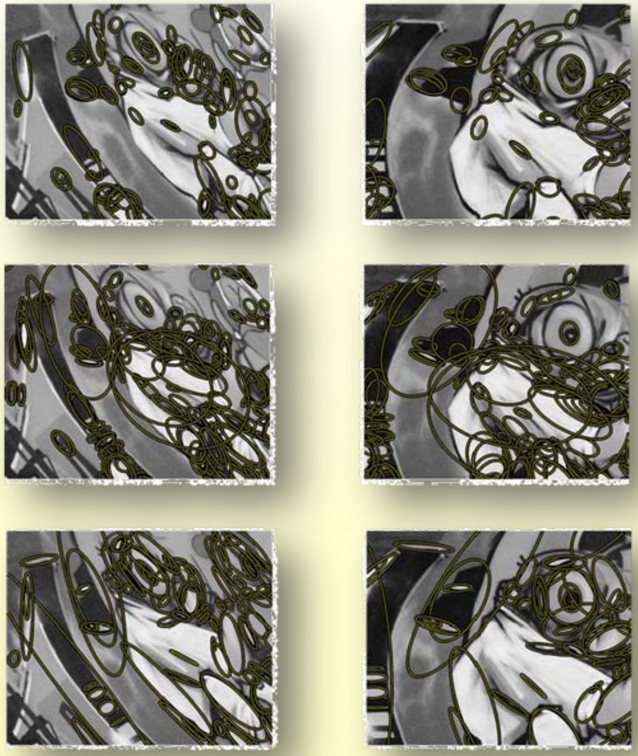
Getting Down to It: Building Image Webs

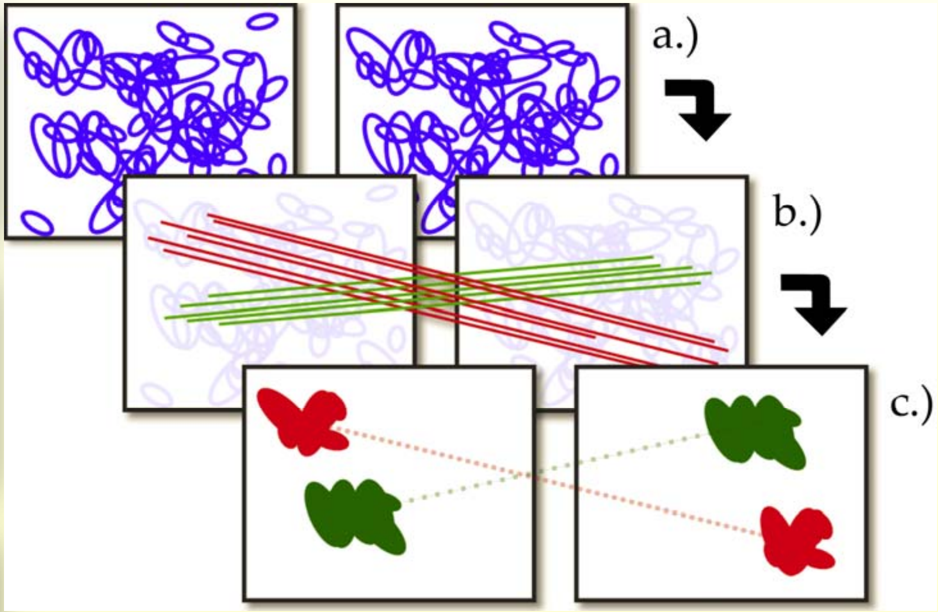
- Feature Extraction: interest points, associated with a region and summarized by a descriptor

Harris-Affine

Hessian-Affine

Maximally Stable Extremal Regions

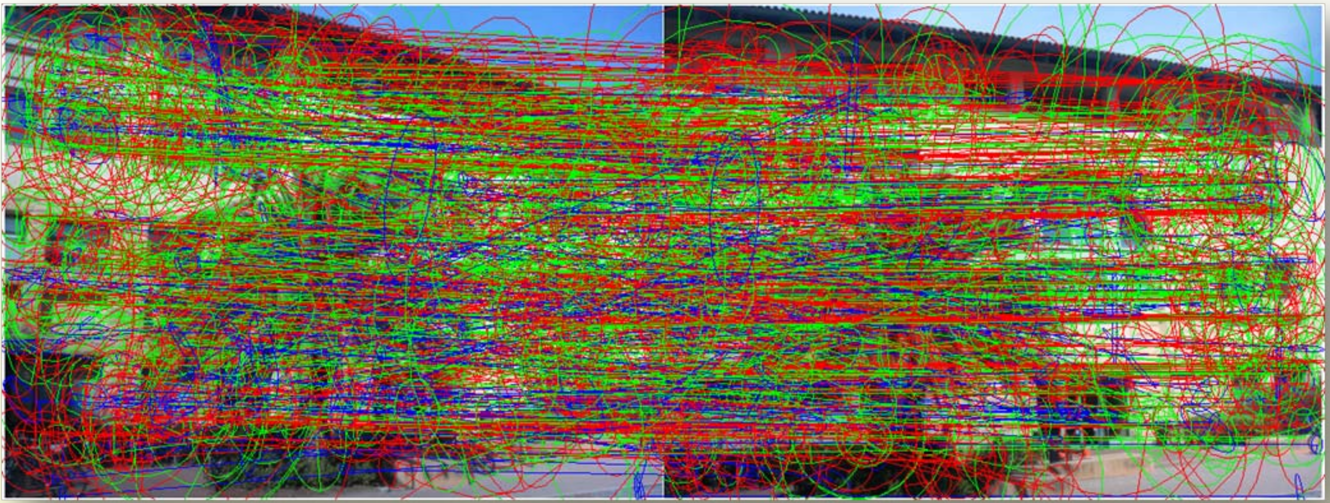




Geometric verification

K. Mikolajczyk, et al. A comparison of affine region detectors. IJCV, 2005

Getting Rid of False Feature Matches

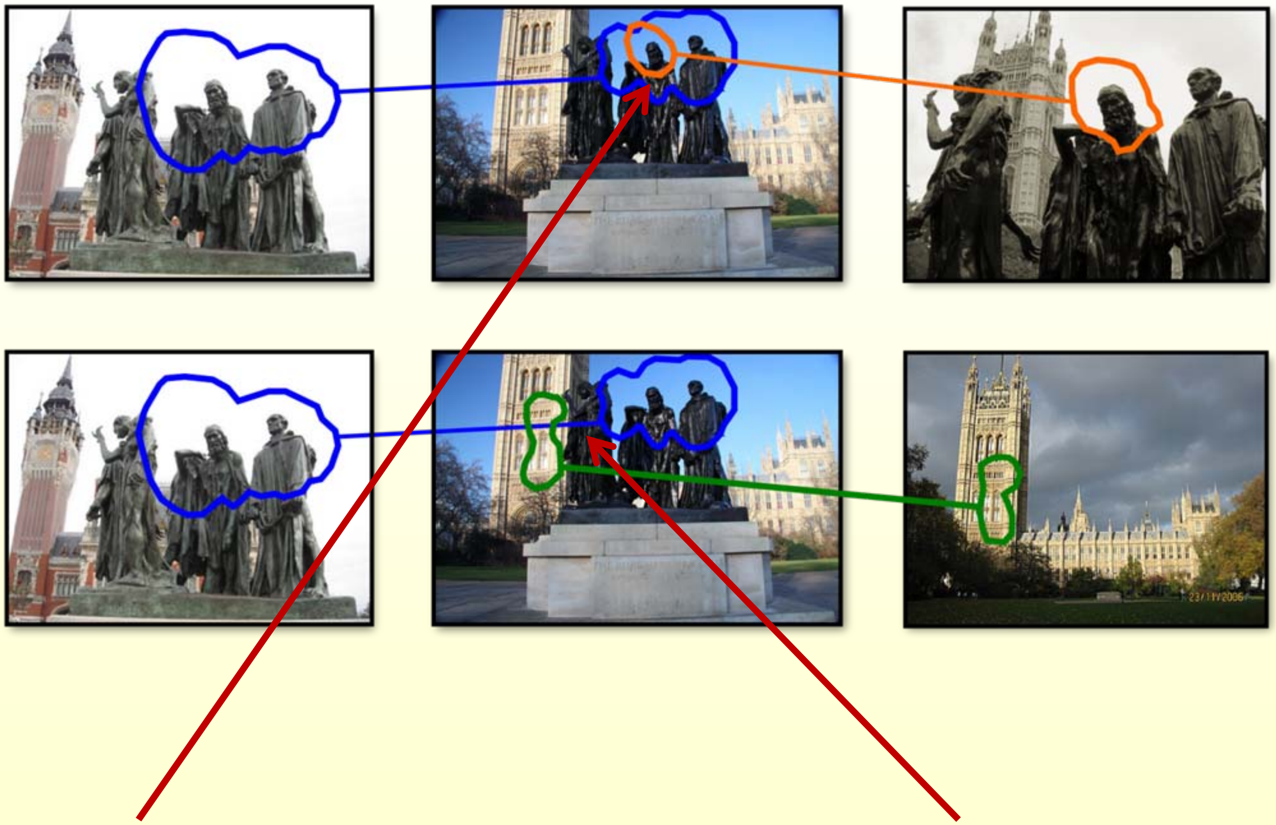


raw matches

after geometric verification

Symmetries and Repetitions: Link Aliasing

Overlap and Pivot Links



Overlap link

Pivot link

Basic element of a Web is a pair (patch, image)

Links and Their Decorations

Link decoration:

- Match (M)-links (quality of match, transform attributes)
- Overlap (O)-links (degree of overlap)
- Pivot (P)-links (patch distance, visual attributes)

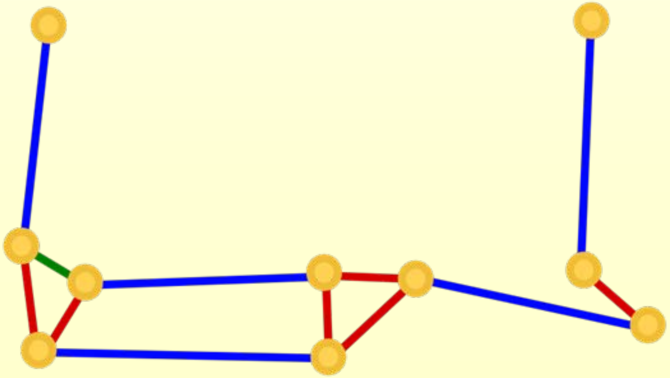


Image Webs Pipeline

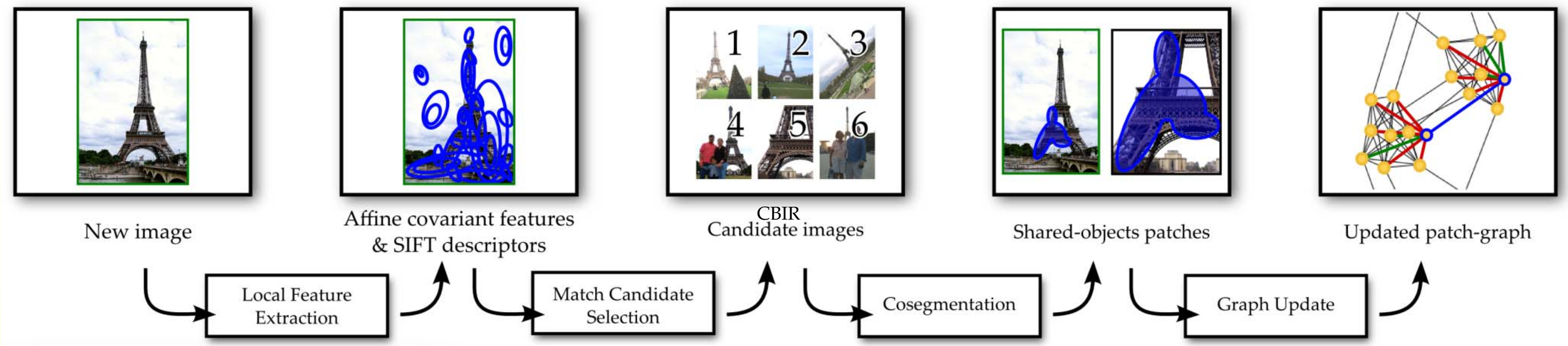


Image webs are complexes (graphs) on image patches, not images

Feature Matching

↓

Geometric Verification

↓

Co-Segmentation

Gaining Efficiency: Pruning Pairs by CBIR Filtering

- Content-Based Image Retrieval (CBIR) via “Bag of Words” models:
  - cluster and quantize descriptors into vocabulary trees
  - use document information retrieval type indices

[Fei Fei, Fergus, Torralba]

- Used to retrieve “visually similar” images – in our case possible Web neighbors for which match links exist

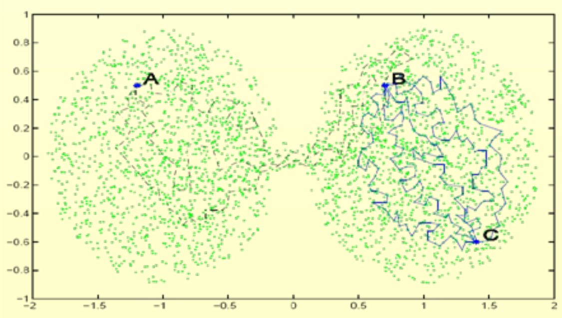
Computation Times (w. a Cluster)

- Image matching steps (VGA image size)
  - Feature extraction (~ 4 sec per image)
  - CBIR indexing (~ 30 sec per image)
  - Cosegmentation operation (~ 1.5 sec per image pair)
- Image Web construction times*
  - Car (70 images ~ 1 minute)
  - Art museum (1200 images ~ 52 minutes)
  - Stanford campus (4200 images ~ 3 hours)

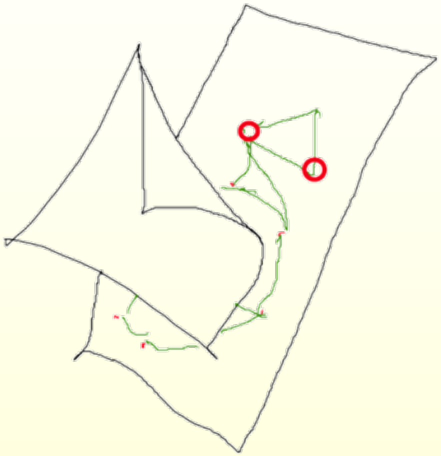
*just cosegmentation stage using up to 500 compute nodes

Scaling Up Web Construction

- We want to build Image Webs with millions of images -- and understand how they are connected
- We cannot afford to try cosegmentation on all image pairs
- CBIR is a useful filter, but …
- Vital connectivity information may reside in sparser areas of the Web



Getting an Unknown Graph to Reveal Itself …



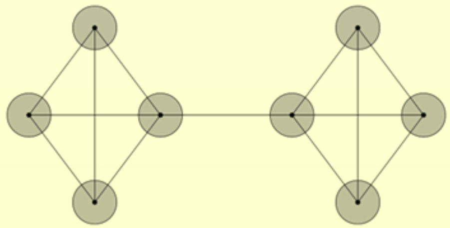
- Testing for the presence of links is expensive
- Which images pairs should we try to connect?
- We seek a sparser graph which captures the connectivity of the unknown Web
  - On the one hand, the CBIR filter favors image pairs where links are likely to exist
  - But how can we tell is a particular link improves connectivity?
  - What should be our ultimate measure of Web utility?
- Spectral graph theory and harmonic analysis to the rescue

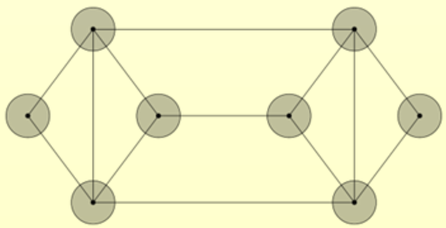
Algebraic Connectivity Measures

- Connectivity of a graph based on heat diffusion notions
  - Second smallest eigenvalue of the graph Laplacian

$$L_{i,j} = \begin{cases} d(i) & \text{if } i = j \\ -1 & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases}$$

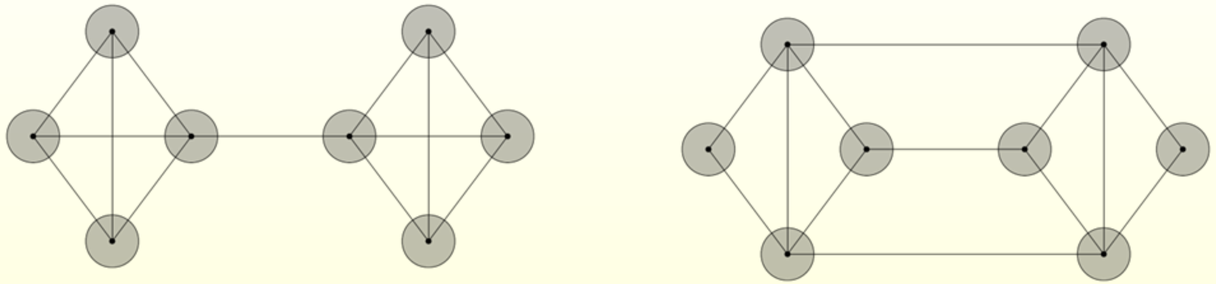
  - Smallest eigenvalue of L is always 0 – and has a constant eigenvector
  - Multiplicity of 0: number of connected components





Algebraic Connectivity

- Connectivity Measure: Second smallest eigenvalue of the graph Laplacian



$\lambda_2 = 0.3542$ $\lambda_2 = 1.0968$

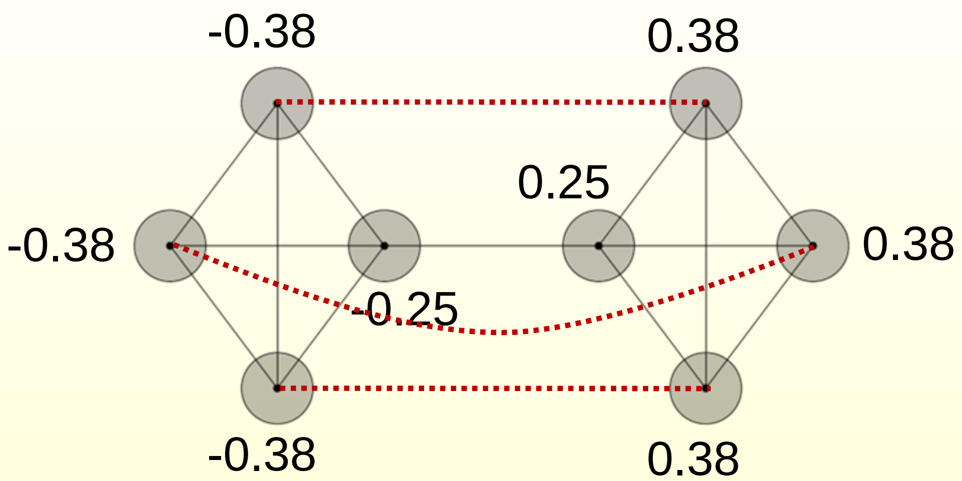
- Related to the diameter *D* of a graph with n nodes, random walk convergence, diffusion distances, and many other measures of graph connectivity
- The eigenvector corresponding to λ_2 is the Fiedler vector, and is often used to partition the graph

Building a "Good" Graph

- Objective:
  - Build a "well connected" graph in minimal time
- Difficulty:
  - Given a graph, finding the k extra edges which maximally increase algebraic connectivity is NP-hard
- Use a greedy strategy:
  - For every potential new connection, test its EdgeRank *R* – how much it will increase connectivity

Building a “Good” Graph

- Use a strategy from graph cuts

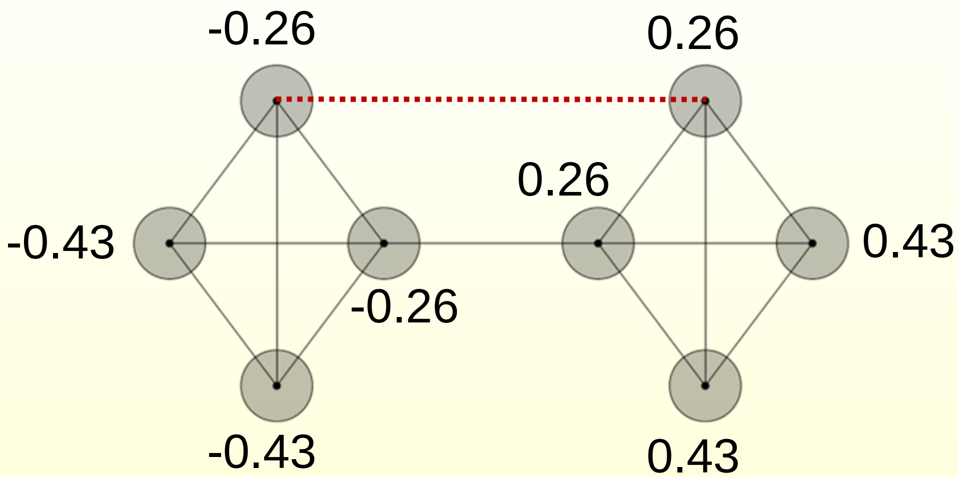


- Assign to each node its value in the Fiedler vector
- Add an edge (i, j) to maximize connectivity score:

$$R(i,j) = \max_{i \nsim j} |\phi_2(i) - \phi_2(j)|$$

Building a “Good” Graph

- Practical considerations



- Update the Fiedler vector after each new edge
- Can use the old estimate as a guess
- Use a *power iteration* to update the Fiedler vector

Building a “Good” Graph

- Power Iteration

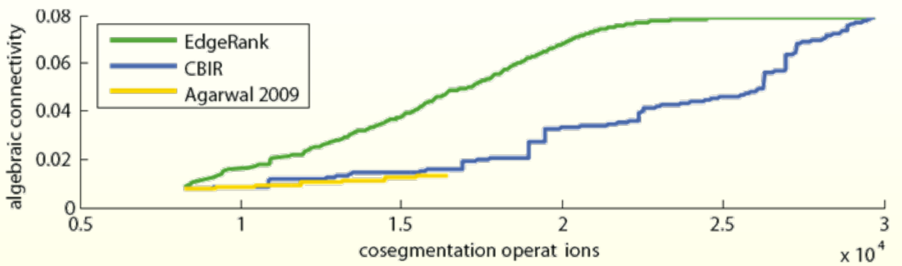
$$u_{i+1} = (2nI - L)u_i$$

$$u_{i+2}(j) = u_{i+1}(j) - \frac{1}{n}\sum_{k=1}^{n} u_{i+1}(k) \ \forall \ j$$

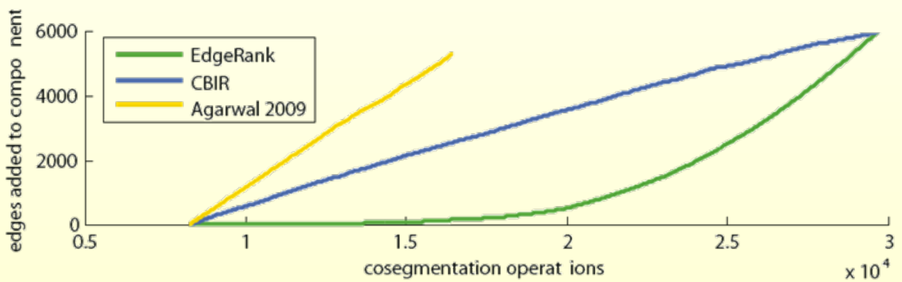
$$u_{i+3} = \frac{1}{||u_{i+2}||}u_{i+2}$$

  - Converges to the Fiedler vector
  - Convergence is fast if have a good estimate. We don’t expect the Fiedler vector to change drastically
  - Small overhead: only 1 vector in memory

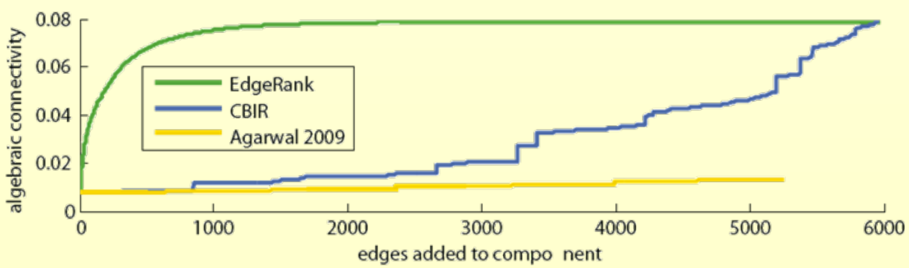
Results on Real Data Sets



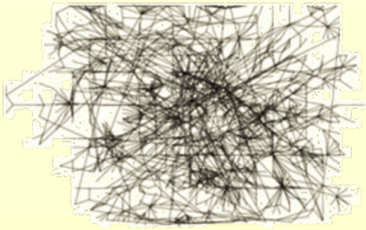
(a) Connectivity / Construction Time



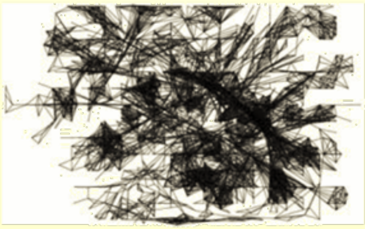
(b) Edges / Construction Time



(c) Connectivity / Edges



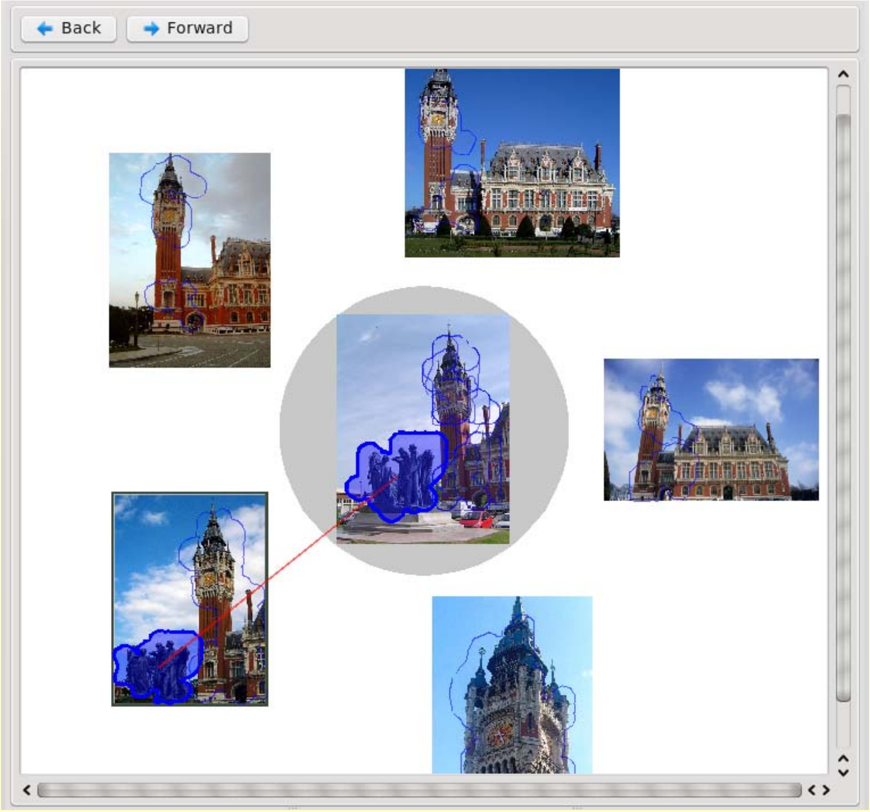
(a) Edge Rank



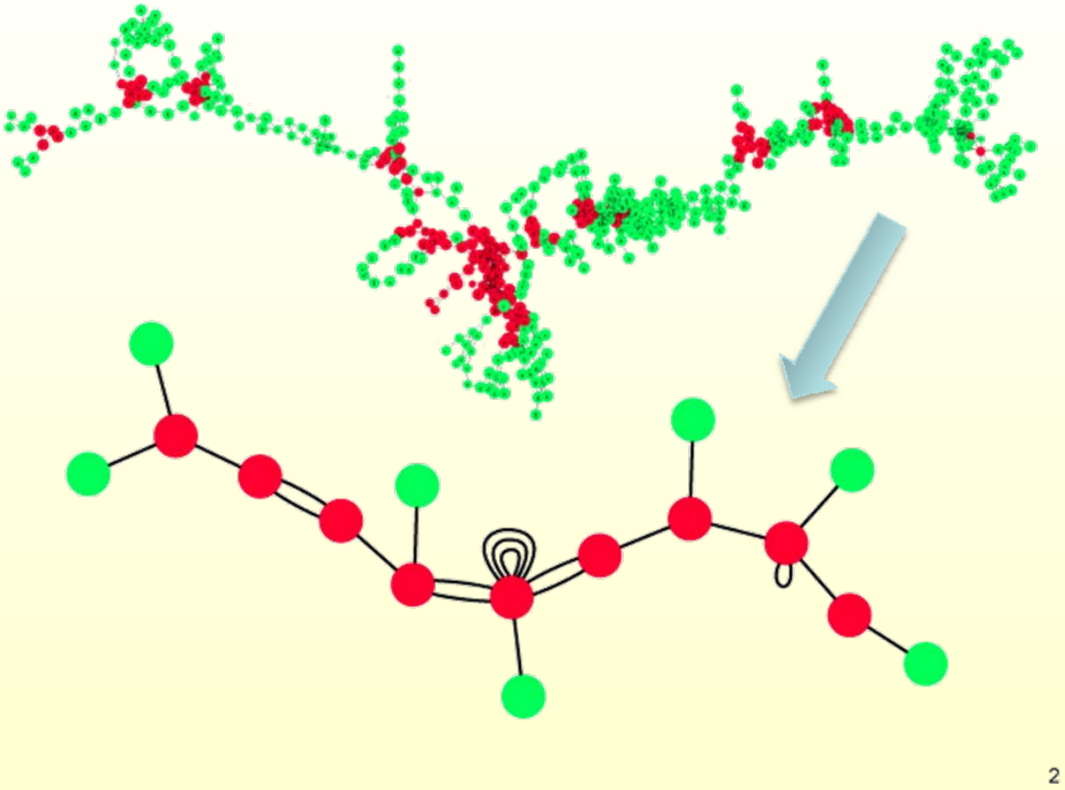
(b) Query Expansion

Applications: An Image Webs Browser

- How can we navigate through large Image Webs effectively?
- How do we mitigate the effects of wrong links?
- How do we extract "persistent" global structure



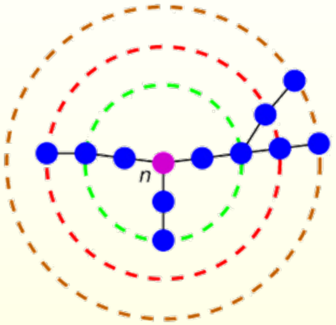
Computing a `Summary Graph'

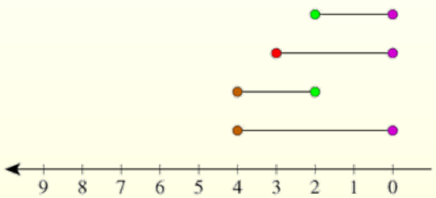


A global map makes navigation easy

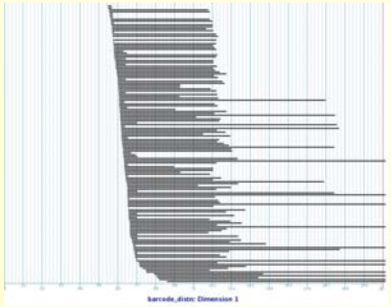
Persistent Local Homology

- Image Webs are often stratified spaces because of the acquisition process – understanding the strata structure helps



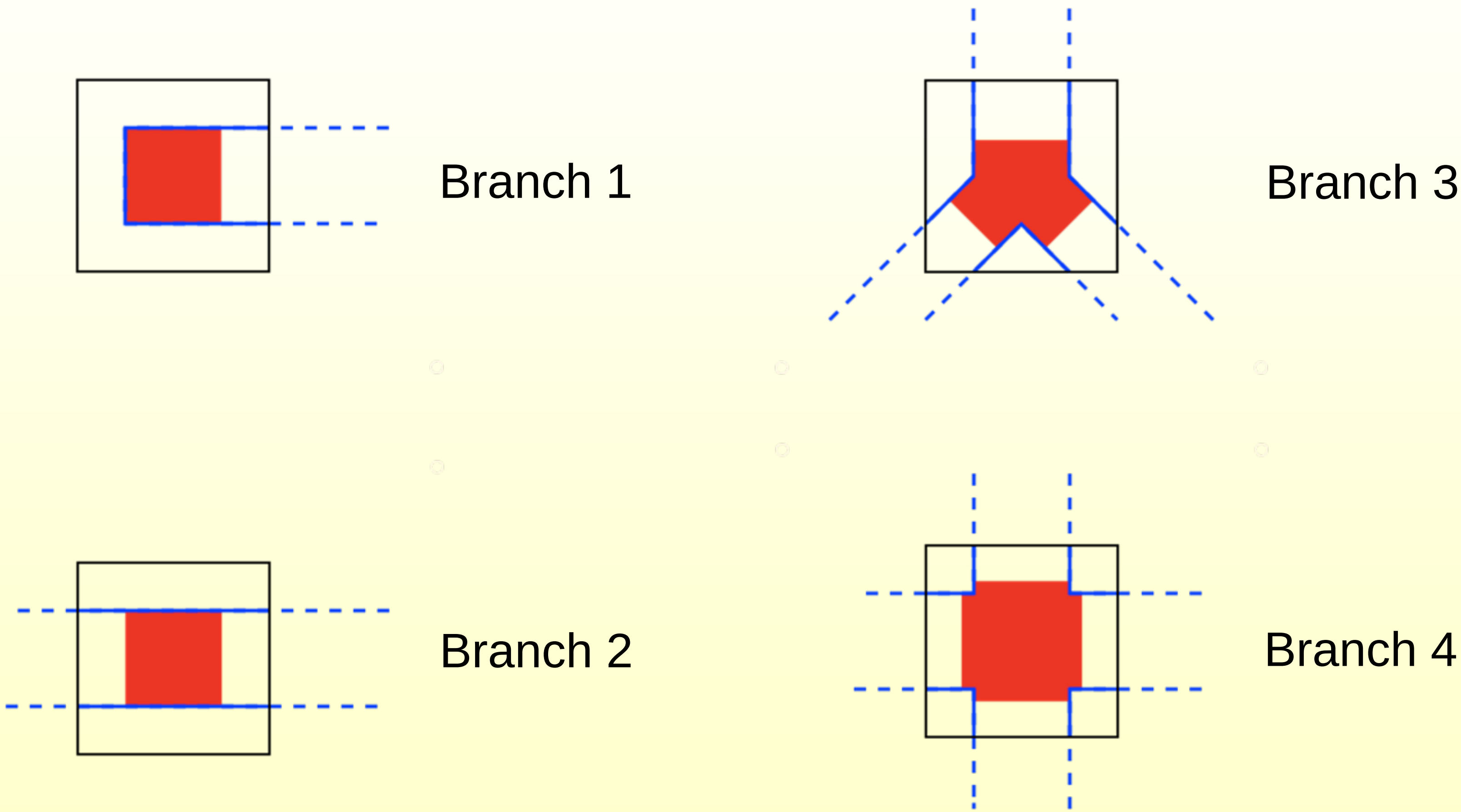


- Use some algebraic topology: image webs as combinatorial complexes
- Rips-Vietoris complex on images, based on distances coming from the links (affine maps)
- Exploit filtered complexes and persistence ideas

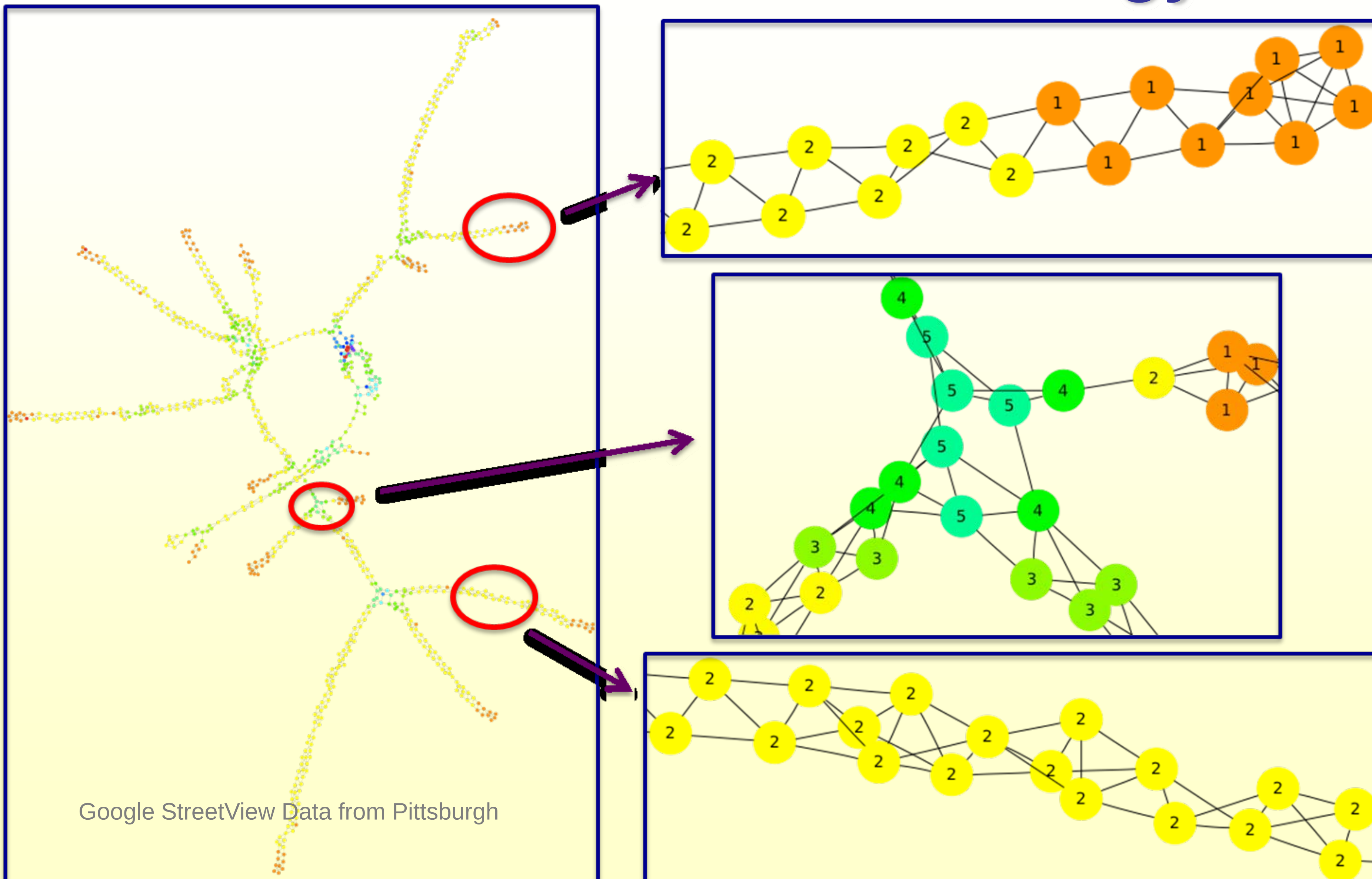


Persistent Local Homology

Different types of nodes in an Image Web:

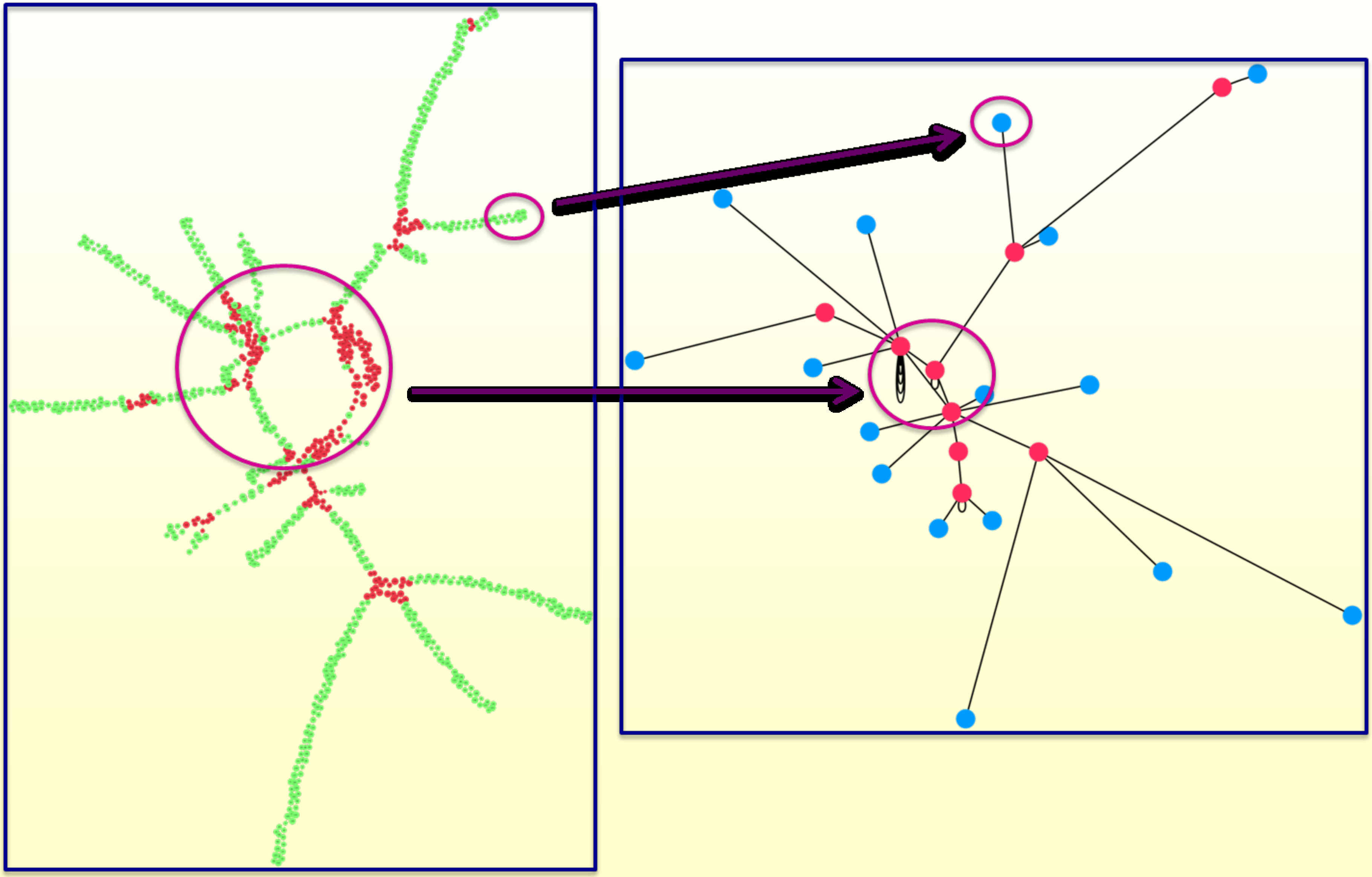


Persistent Local Homology

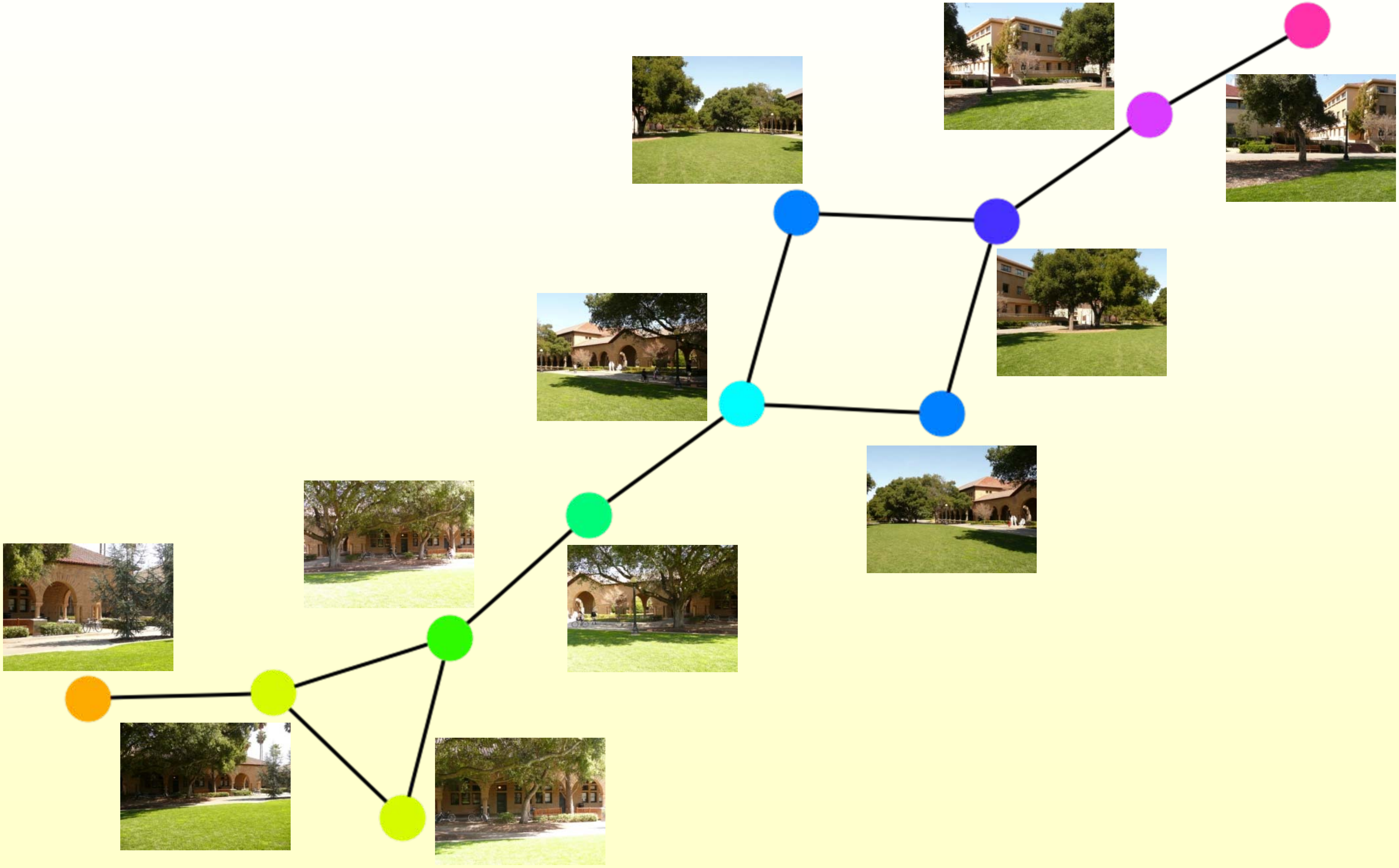


Google StreetView Data from Pittsburgh

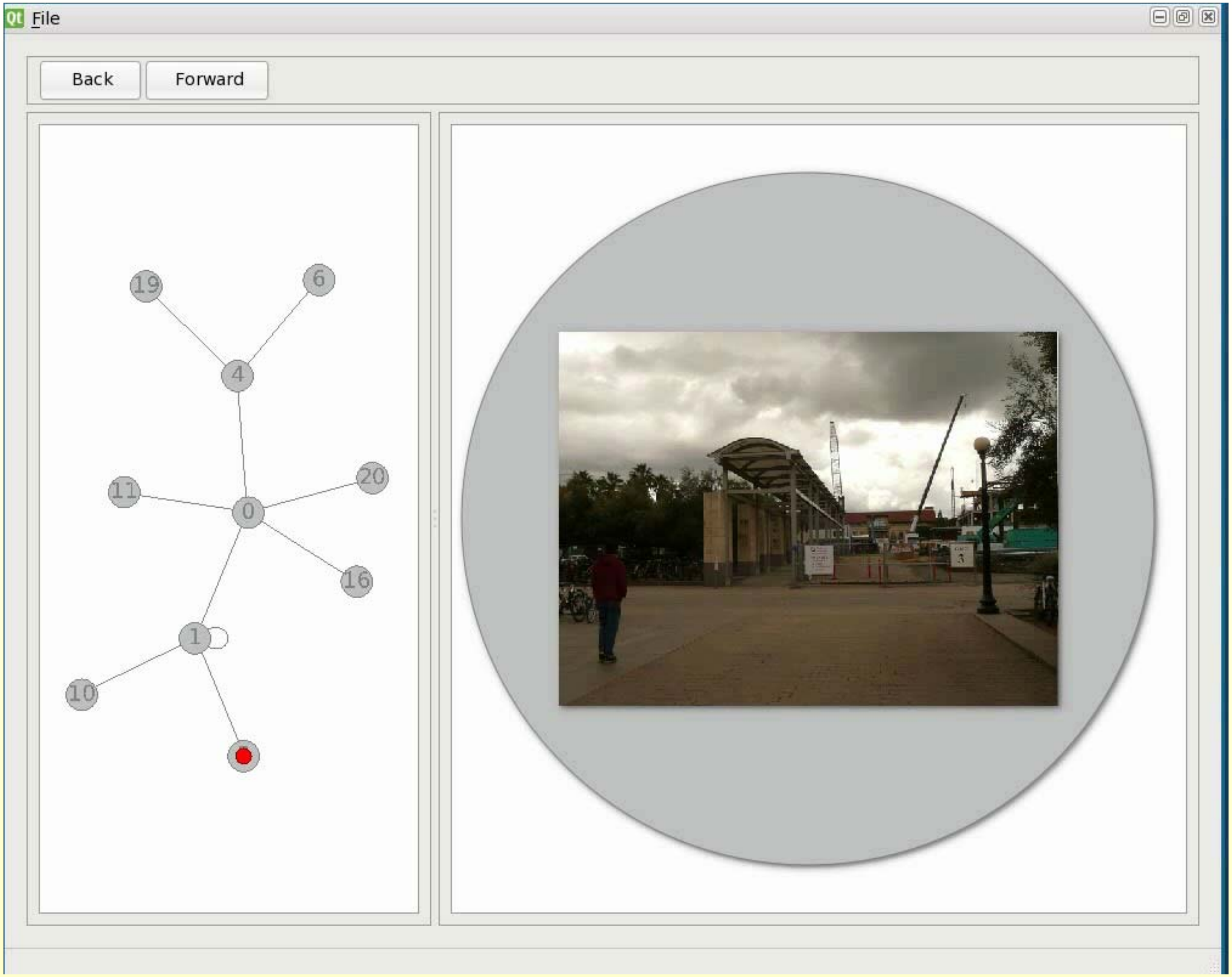
Summarizing Image Webs



Parametrizing Edges/Loops

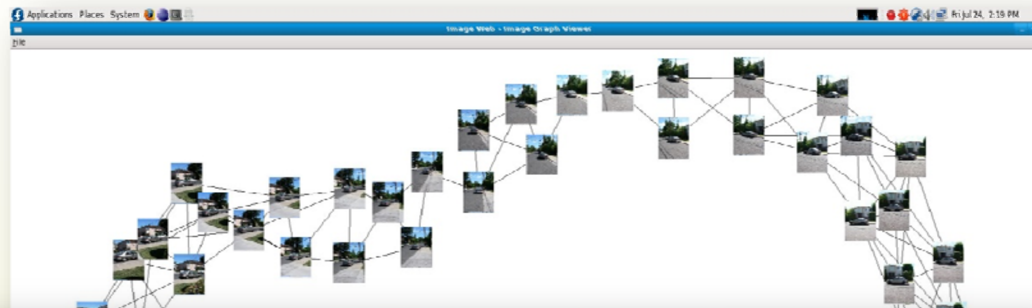


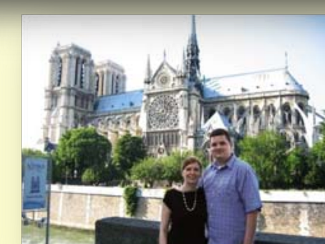
Web Navigation: Video 1



Other Applications

- Object models as subwebs: focus and context
- Annotation transfer
- Linking people through their images
- Mobile webs: photo-guided navigation, collaborative exploration





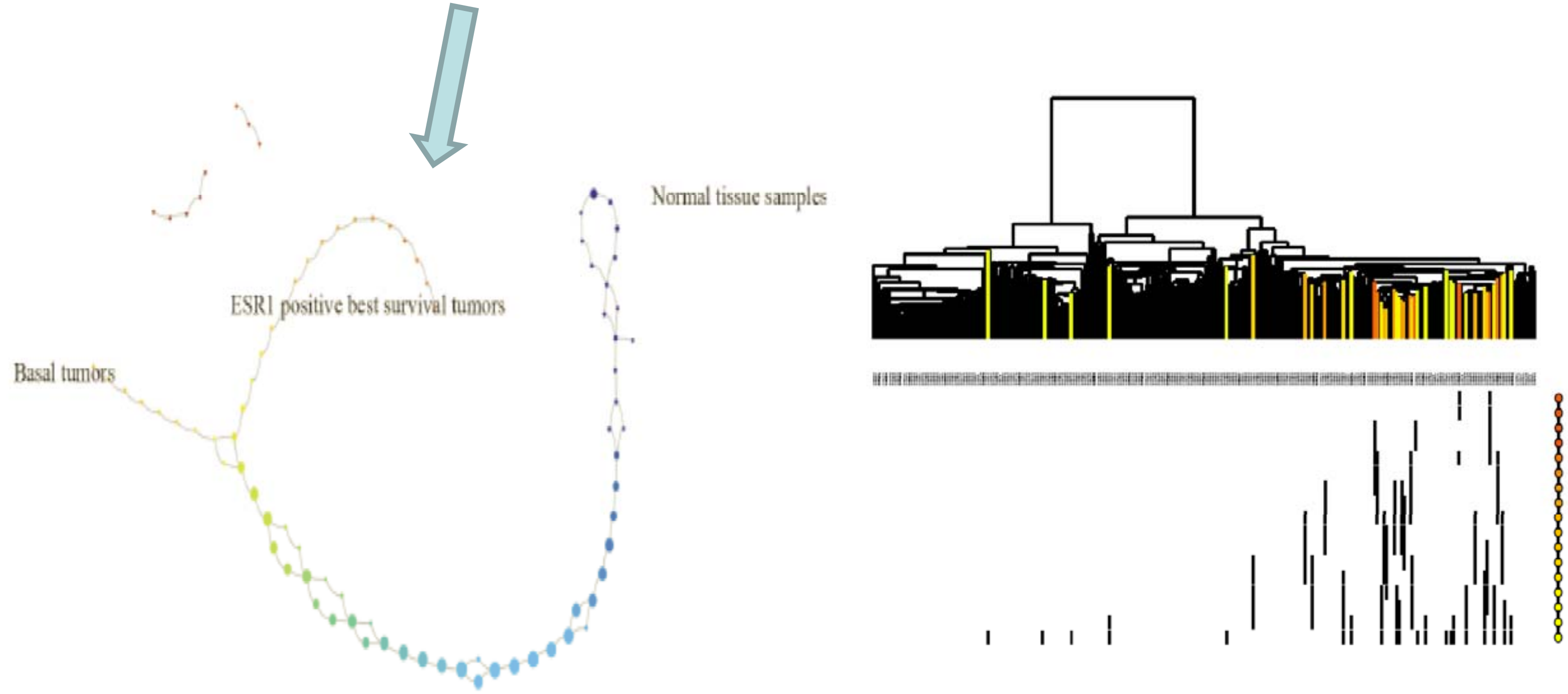
Key Points and Issues

- Interlinked images and other signals contain a wealth of information not apparent in any one image or signal alone
- Such signal webs form networks of maps; maps can be used to carry to transport information and arrive at a global understanding of both the sensed environment and the acquisition process

- The information is in paths induced by the maps

Mapper Application: Breast Cancer Study

This flare consists entirely of patients which survive. This is a new piece of the taxonomy of breast cancer, not identified before, and which cannot be recognized by clustering.



[M. Nicolau, G.C.]

Acknowledgements

- Collaborators:
  - Current and past students: Mridul Aanjenaya, Natasha Gelfand, Kyle Heath, Monica, Nikolau, Maks Ovsjanikov
  - Current and past postdocs: Mikael Vejdemo-Johansson, Dmitriy Morozov, Jian Sun
  - Senior: Vin de Silva
- Sponsors:

FODAVA

National Science Foundation
WHERE DISCOVERIES BEGIN

Homeland Security

The Information is in the Maps

We understand data by studying maps or self-maps among the data, and networks of such maps