Backward generalizations of PCA for shape representations

Sungkyu Jung
University of Pittsburgh

BIRS–Geometry for Anatomy

Presentation based on Joint work with Dibyendusekhar Goswami, Jörn Schultz, Xiaoxiao Liu, Ritwik Chaudhuri, Steve Marron, Ian Dryden, Steve Pizer

Principal Component Analysis (PCA)

1. Study population of shape representations
2. Exploratory statistics
   - Visualization of data structure
3. Dimension reduction and Estimation of Probability distribution

Generalization of PCA to shape representations

- Shape representations are either spheres, quotient of spheres, or involve position-tuples, directions, and (log) sizes
- PCA suited for this type of manifolds

More details on skeletal models (Pizer)

Models of object interiors *designed* for probability distribution estimation: s-reps

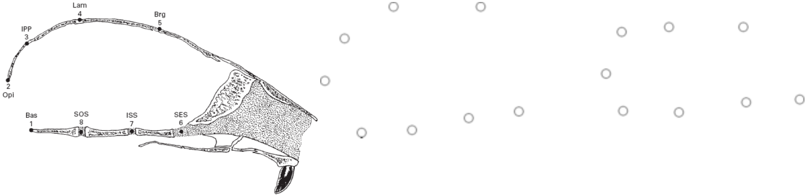
Examples of image/shape analysis

1. Landmark based shapes

Kendall (1984), Bookstein (1986), Dryden and Mardia (1998)

- Work with a set of landmarks on object
- Shape: Invariant under translation, scale, and rotation
- Kendall's Shape Space is a curved manifold ($\mathbf{CP}^{k-2}$)

Example: Shapes of rat skulls

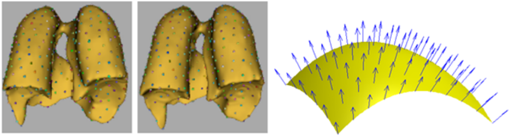


(left) biological landmarks of rat skull, (right) two sets of landmarks

Examples of image/shape analysis

2. Shapes by modern techniques

- Point Distribution Models (PDM):many landmarks automatically determined
- Point and Normal Distribution Models (PNDM): PDM + Normal directions at landmarks
- Warps of an atlas
  - Displacement vector, or
  - Velocity array by t over $[0, 1]$
- Continuous outlines and surfaces (Srivastava, Kurtek)



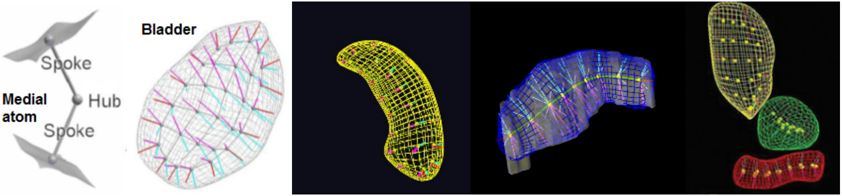
(left) PDM of Lung, (right) illustrative example of Points and Normals

Examples of image/shape analysis

3. Skeletal representations (s-rep)

Siddiqi and Pizer (2008), Pizer et al. (2011)

- Special case: Medial representations (m-rep)
- Capturing interior of objects
- Suitable for statistical analysis
- More details covered in Pizer's talk



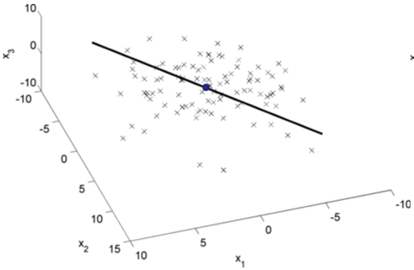
medial atom, slabular m-rep model, slabular s-rep, quasi-tubular s-rep, multi-object

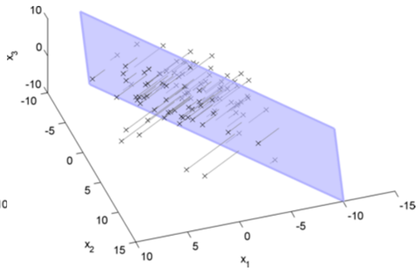
Two equivalent formulations of Euclidean PCA

Forward stepwise view of PCA: center point - line - plane - ...

Backward stepwise view of PCA:

1. Begin with full data space ($\mathbb{R}^d$)
2. Find $d-1$ dim'l affine subspace (best approximates data)
3. Reduce dimension further to $d-2, d-3, \ldots, 0$.





Euclidean case: Forward PCA = Backward PCA

PCA generalizations

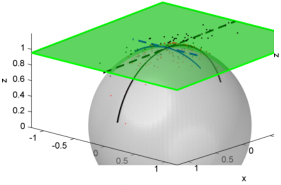
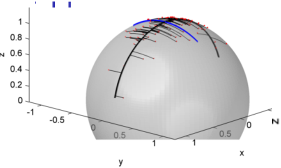
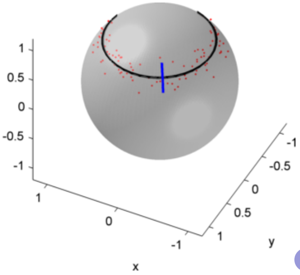
Complications on manifolds

- Orthogonal lines $\rightarrow$ Orthogonal geodesics
- Need to find an appropriate 'mean'

Approaches of manifold PCA

1. Tangent Space Approach [Dryden and Mardia (1998), *Principal Geodesic Analysis* by Fletcher et al. (2004)]
   Forward approximation
2. Direct Geodesic Fitting [*Geodesic PCA* by Huckemann, Ziezold and Munk (2006, 2010)]
   Partially backward approach
3. Small Circle Fitting for S^2 [*Principal Arc Analysis* by Jung, Foskey and Marron '11]
   Backward approach

PCA generalizations

Approaches of manifold PCA

1. Tangent Space Approach [Dryden and Ma
   Geodesic Analysis by Fletcher et al. (2004)]
   Forward approximation
2. Direct Geodesic Fitting [*Geodesic PCA* by Huckemann, Ziezold and Munk (2006, 2010)]
   Partially backward approach
3. Small Circle Fitting for S^2 [*Principal Arc Analysis* by Jung, Foskey and Marron '11]
   Backward approach

De

Analysis of Principal Nested Spheres

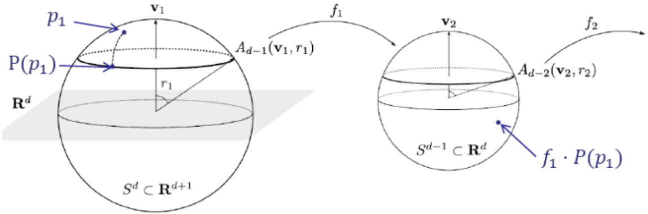
Jung, Dryden and Marron '11

- Generalization of *Principal Arc Analysis* to S^d, $d \geq 2$.
- Decomposition of S^d captures non-geodesic variation in lower dimensional spheres.
- $\mathfrak{A}_k$: k-dimensional Principal Nested Sphere (PNS)

$$\mathfrak{A}_0 \subset \mathfrak{A}_1 \subset \cdots \subset \mathfrak{A}_{d-1} \subset S^d.$$

- Works for Kendall's landmark shape data through the preshape space S^d.
- Fitted in backward stepwise fashion.

Sequence of Principal Nested Spheres



Begin with $\mathbf{x}_1, \ldots, \mathbf{x}_n \in S^d$

1. Fit $\mathfrak{A}_{d-1} \cong (d-1)$-sphere
   - best non-geodesic (d-1) dim'l approximation
2. Fit $\mathfrak{A}_{d-2} \cong (d-2)$-sphere

   $\vdots$
3. Reach $\mathfrak{A}_0$ (PNSmean)
4. Result in $\mathfrak{A}_0 \subset \mathfrak{A}_1 \subset \cdots \subset \mathfrak{A}_{d-1} \subset S^d$

Best fitting subsphere

- Samples: $\mathbf{x}_1, \ldots, \mathbf{x}_n \in S^d$
- Subsphere: $A_{d-1}(\mathbf{v}_1, r_1) \subset S^d$
- Residual ξ of $\mathbf{x}$ from a subsphere A_{d-1}
  - Signed length of the minimal geodesic that joins $\mathbf{x}$ to A_{d-1}.

Subsphere fitting

$\widehat{A}_{d-1} \equiv A_{d-1}(\hat{\mathbf{v}}_1, \hat{r}_1)$ minimizes the sum of squared residuals

$$\sum_{i=1}^{n} \xi_i(\mathbf{v}_1, r_1)^2 = \sum_{i=1}^{n} \{\rho_d(\mathbf{x}_i, \mathbf{v}_1) - r_1\}^2,$$

among all $\mathbf{v}_1 \in S^d$, $r_1 \in (0, \pi/2]$. Detail....

Byproducts of PNS

Euclidean-Type Representation (Principal Scores matrix)

- Stacked residuals from each layer
- Analogue of principal component scores
- Used to visualize the data, and for further analysis

% Variance explained

- Sample variance of residuals (from each layer) over the sum of all variances

Principal Arc

- the *direction* of major variations defined by PNS

A special case: PNG

A *great sphere* is a sphere with radius 1 (or $r = \pi/2$).
Interesting & important special case of PNS:

Principal Nested Great Spheres (PNG)

- Setting $r = \pi/2$ for each subsphere fitting.
- Principal arcs become great circles (i.e. geodesics).
- The principal geodesics, found by PNG, are similar to the geodesic-based PCs.
- Close to the Geodesic PCA (direct geodesic fitting) than the tangent space approach.

Choice between small and great sphere

1. Strictly using small spheres (PNS) —nongeodesic decomp.
2. Adopted tests:
   H_0 : Great Sphere ($r = \pi/2$) vs H_1 : Small sphere ($r < \pi/2$).
3. Strictly using great Spheres (PNG) —geodesic decomp.

Choice between small and great sphere

1. Strictly using small spheres (PNS) —nongeodesic decomp.
2. Adopted tests:
   H_0 : Great Sphere ($r = \pi/2$) vs H_1 : Small sphere ($r < \pi/2$).
3. Strictly using great Spheres (PNG) —geodesic decomp.
4. Soft decision between small and great sphere. —Ongoing work.

Kendall's 2D landmark shape space

Planar shape space Σ_2^k

- A shape is a point in Kendall's shape space with dimension $2k - 2 - 1 - 1$

Preshape space $\mathcal{S}_2^k \cong S^d$

- Preshape is what is left from removing location and scale
- Dimensionality of preshape space is $d = 2k - 2 - 1$

PNS to shape data

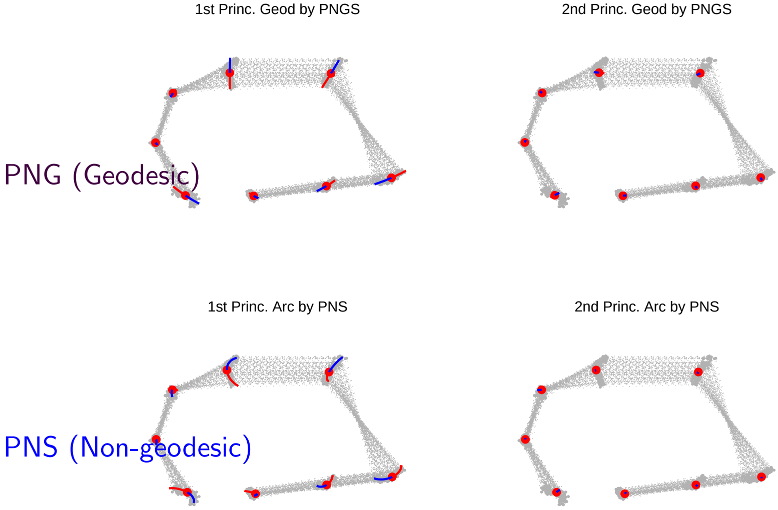
- Desire that $\mathfrak{A}_{d-1}$ of S^d leaves zero residuals.
- Achieved when each shape is aligned to a common base shape (e.g. Procrustes mean)

Example: Shape of Rat Skulls

- Shape data with 8 landmarks on plane in Σ_2^8

- Non-geodesic variation captured by PNS (and not by PNG)
- Scatterplot given by Principal Scores
- Shape changes related to growth of rats.

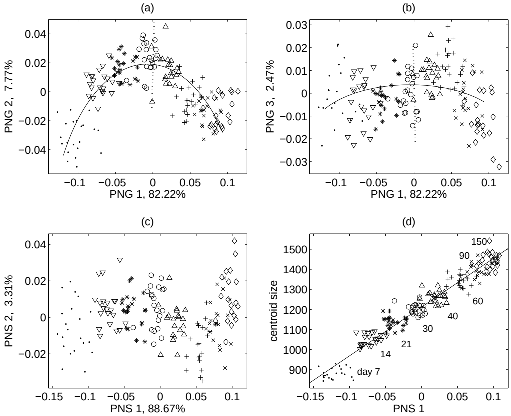
Rat Skulls: Major mode of variation by PNS



Rat Skulls: Scatterplots

PNG →

PNS →

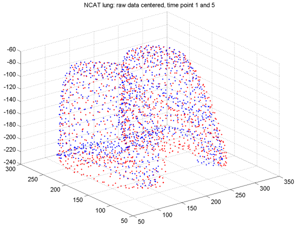


- PNS need 1 mode (PNG need 2 modes) to capture the non-geodesic variation
- Shape change by *growth* of rats explained by PNS 1
- PNS 1 linearly correlated with size of skulls ($R = 0.9705$)

CPNS on Lung Respiratory Motion

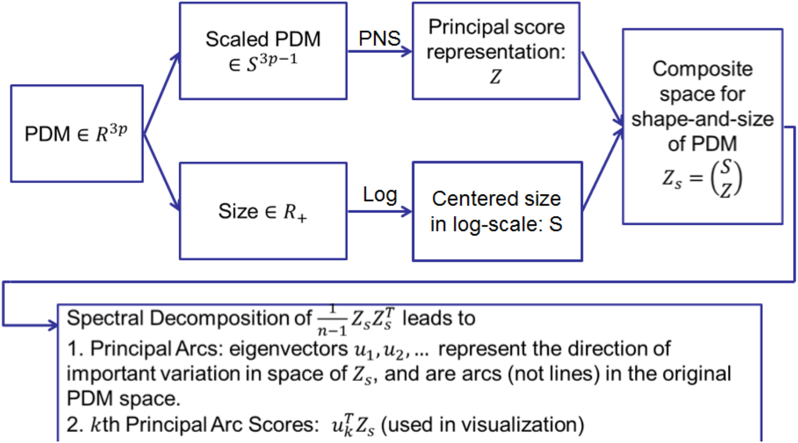
Jung, Liu, Pizer and Marron '10

- PDM represents shape of human lung, pre-aligned with N points
  - Scaled PDM is $\in S^{3N-1}$
  - Size variable is $\in \mathbb{R}^+$
- PDM in $\mathbb{R}^{3N} = ScaledPDM \oplus Size \in S^{3N-1} \otimes \mathbb{R}^+$
  - Thus want composite of PNS (S^{3N-1}) and $\mathbb{R}^1$.



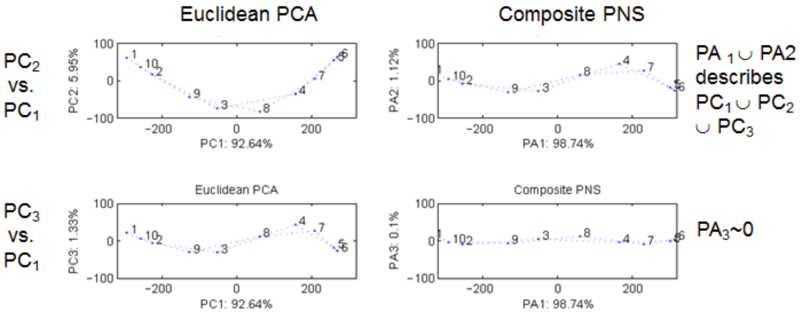
Composite PNS for PDM

Must capture correlations between Euclidean and non-Euclidean features



Respiratory Motion Analysis in the Lung

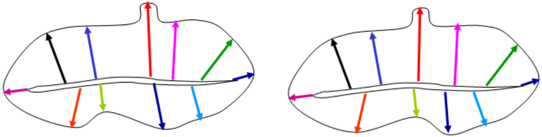
$n = 10, \; N = 10550.$



% of variance	PCA	PGA	KPCA*	Composite PNS
Component 1	92.64%	92.72%	97.94%	**98.74%**
Component 2	5.95%	5.88%	1.89%	1.12%
Component 3	1.33%	1.32%	0.09%	0.10%
			*kernel = $\exp(-10^6 \|\|x-y\|\|^2)$	

S-reps:3D model of object interior

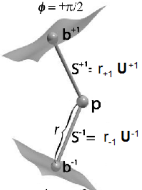
- Interior-filling skeletal model of an object
- Stable topology
  - no branches
  - skeletal locus: fully folded, multi-sided
- Stable geometry
  - as medial as possible
  - correspondence of primitives over population
- Continuous: Folded sheet of non-intersecting spoke vectors
- Types: Slabular and Quasi-tubular
- Discrete: sampled continuous s-reps

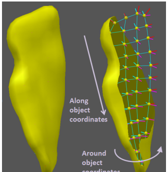


Fitting s-reps to signed distances

By optimization of objective function summing 2 terms

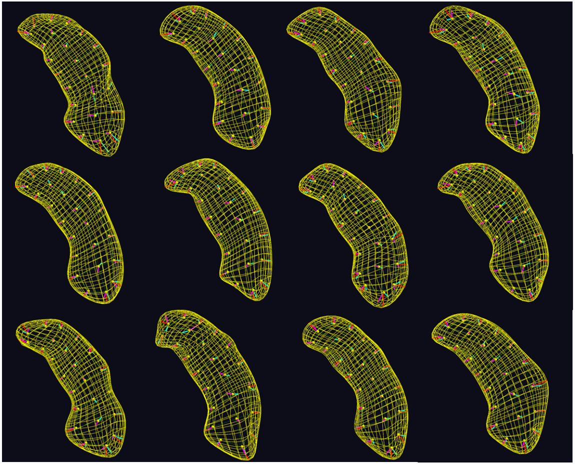
- Geometric properties
  - Spokes do not cross
  - As medial as possible
    - Near orthogonality of spoke directions to Δdistance
    - Near equality of spoke lengths with spokes sharing the same hub
    - Difference of spoke directions nearly normal to skeletal sheet
- Data (distance function) match
  - All spoke ends on boundary
  - End spoke vector triples properly fit into crest of zero level set of distance





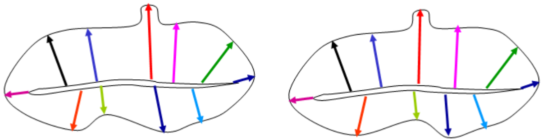
Fitting s-reps: Results

Hippocampi in study of schizophrenia



Correspondence across training s-reps

1. By analogy to shifting points on boundaries in PDMs via entropies (Cates 2007, Oguz 2008)
2. But for spokes:
   - tightest prob. distribution on geometry of spokes tuples
   - uniformity of interior coverage of spokes in each case
3. Retain spoke orthogonality to bdry
4. Results in separated discrete spoke hubs on top & bottom of skeletal sheet



Abstract space of discrete s-reps with n spokes

- Each spoke direction $\in S^2$
- log (spoke length) $\in \mathbb{R}^n$
- After centering and scaling of tuple of $p(u)$ values,
- These points are in $\mathbb{R} \times S^{3(n-1)-1}$ (same as for the PDMs)

The s-rep is a point $\in \mathbb{R}^{n+1} \times (S^2)^n \times S^{2(n-1)-1}$

Composite Principal Nested Sphere is applied

Separately analyze each sphere into Euclideanized variables and,
Composite with Euclidean variable to take correlation into account.

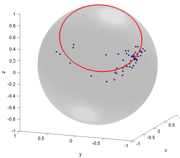
Transformations of S-reps

For global rotation

- Each spoke direction moves on a small circle on S^2;
- the circles share a common axis
- Scaled tuple of spoke tails move on a small circle (1D sphere) on S^{3n-4}

For rotational fold and twists about an axis

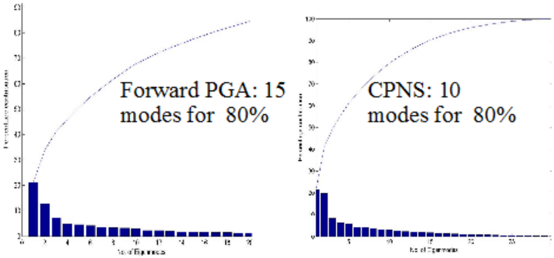
- All spoke directions move on small circles on S^2
- The circles share a common axis



Experimentally, analysis via small sphere motions is useful

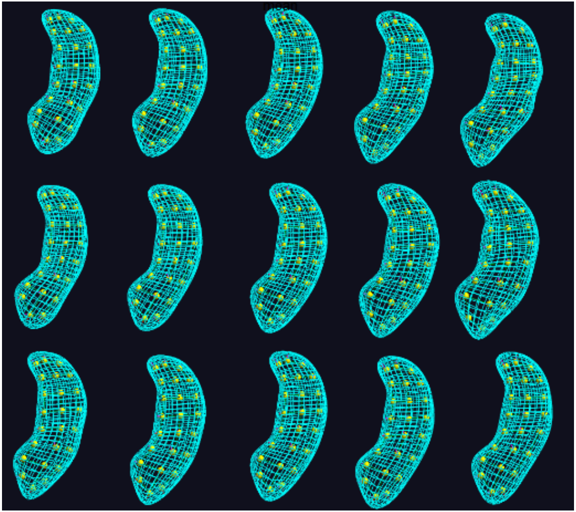
Shape Probabilities via CPNS

- Successive dimension reduction for spherical variables
- Composite scores from each sphere with Euclidean part, then SVD
- Yields fewer eigenmodes to explain variation



Shape Probabilities via CPNS

Modes of variation by principal arcs:



rotation, pinching / elongation, swelling / twisting, swelling in the bottom

Summary

- Backward PCA approaches on spheres and composite space with Euclidean space
- Shown useful for 2D, 3D landmark data (PDM)
- S-reps provide a basis for statistics on objects
- In the size and shape changes of hipposcampi s-reps, composite PNS yields succinct description of data

- Cates, J; PT Fletcher, M Styner, ME Shenton, RT Whitaker (2007). Shape modeling and analysis with entropy-based particle systems. Proc. Information Processing in Medical Imaging, Springer LNCS 4584: 333-345.
- Damon, J (2003). Smoothness and geometry of boundaries associated to skeletal structures I: sufficient conditions for smoothness, Annales Inst. Fourier 53: 1001-1045.
- Damon, J (2008). Swept regions and surfaces: modeling and volumetric properties, Conf. Computational Alg. Geom. 2006, in honor of Andre Galligo, Special Issue of Theor. Comp. Science, 392(1-3): 66-91.
- Fletcher, PT; C Lu, SM Pizer, S Joshi (2004). Principal geodesic analysis for the study of nonlinear statistics of shape, IEEE Transactions on Medical Imaging, IPMI 2003 special issue, 23(8): 995-1005.
- Huckemann, S; H Ziezold (2006). Principal component analysis for Riemannian manifolds with an application to triangular shape spaces. AAPS 38(2): 299-319.
- Huckemann, S; T Hotz, A Munk (2010). Intrinsic shape analysis: geodesic principal component analysis for Riemannian manifolds modulo Lie group actions. Statistica Sinica 20(1):1-100.
- Jung, S; I Dryden, D Marron (2011). Analysis of principal nested spheres. Submitted.

- Oguz, I; J Cates, T Fletcher, R Whitaker, D Cool, S Aylward, M Styner (2008). Entropy-based particle systems and local features for cortical correspondence optimization. Proc ISBI 2008: 637 – 1640.
- K Siddiqi, Pizer, SM (2008). Medial Representations: Mathematics, Algorithms and Applications, Springer.
- Schulz, J; S Jung, S Huckemann (2011): a collection of internal reports submitted to UNC-Gottingen study group on s-rep change under rotational transformations.