Geometric Data Analysis for Face Recognition: Classification on the Grassmannians

Jen-Mei Chang

Department of Mathematics
Colorado State University
chang@math.colostate.edu

Workshop on Geometry and Computation, 12/20/2007
Department of Mathematics, Tunghai University

Outline

Outline

Various Forms of Classification

- **Identification/Verification** - Are you who you claim to be?
  - Personal usage, e.g., personal computer access, ATM access, bank account access, cell phone access.
  - Public usage, e.g., company entry access.
- **Detection** - Where are you?
  - Personal usage, e.g., face finding in digital cameras.
  - Public usage, e.g., video surveillance as part of face recognition algorithm.
- **Recognition** - Who are you?
  - Personal usage, .e.g., digital photo sorter, a greeting system on PC.
  - Public usage, e.g., casino and airport security watch list.

Example Biometrics - fingerprint

- Advantages: accuracy, speed, reliable.
- Disadvantages: willingness of people to use it.

pay clock

door lock

palm pilot

cell phone

Example Biometrics - fingerprint

- Advantages: accuracy, speed, reliable.
- Disadvantages: willingness of people to use it.

pay clock

door lock

palm pilot

cell phone

Example Biometrics - fingerprint

- Advantages: accuracy, speed, reliable.
- Disadvantages: willingness of people to use it.

pay clock

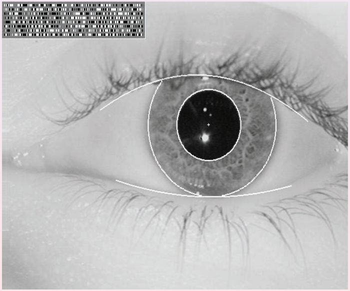
door lock

palm pilot

cell phone

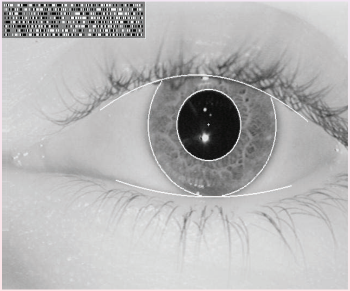
Example Biometrics - iris

- Advantages: uniqueness of our eyes, low likelihood of false positives, speed and ease of use.
- Disadvantages: high-level subject cooperation, storage of data.



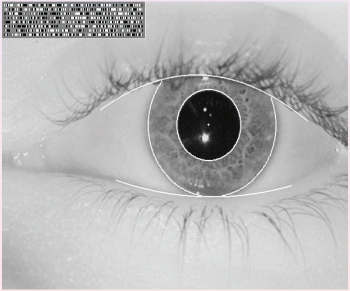
Example Biometrics - iris

- Advantages: uniqueness of our eyes, low likelihood of false positives, speed and ease of use.
- Disadvantages: high-level subject cooperation, storage of data.



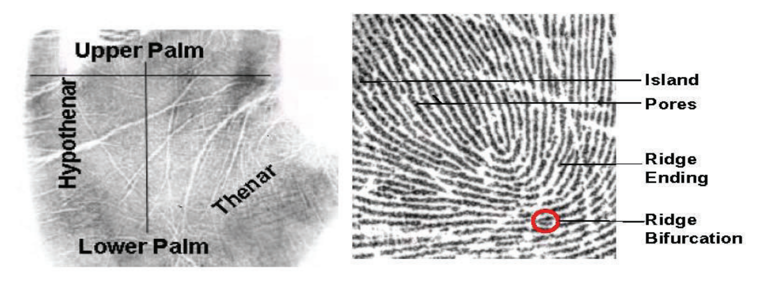
Example Biometrics - iris

- Advantages: uniqueness of our eyes, low likelihood of false positives, speed and ease of use.
- Disadvantages: high-level subject cooperation, storage of data.



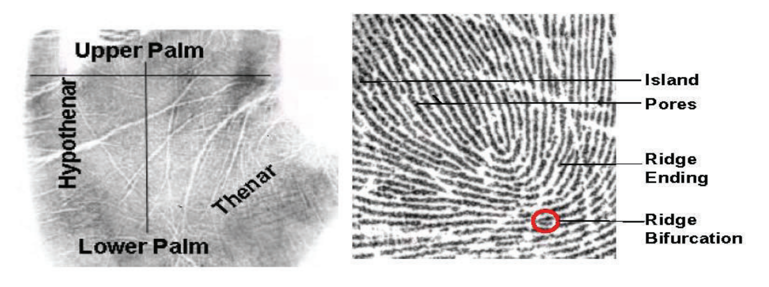
Example Biometrics - palm print

- Advantages: excellent biometric when combined with fingerprint.
- Disadvantages: high error rate, high-level subject cooperation, high precision of hardware.



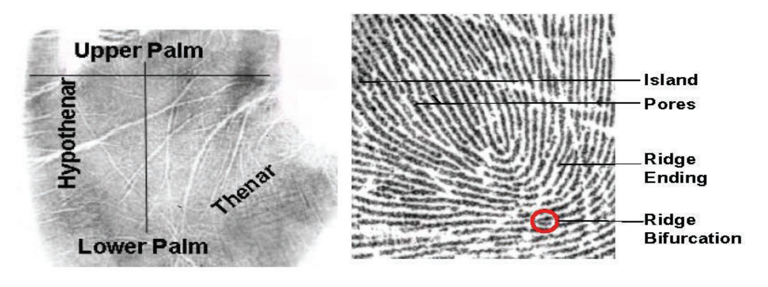
Example Biometrics - palm print

- Advantages: excellent biometric when combined with fingerprint.
- Disadvantages: high error rate, high-level subject cooperation, high precision of hardware.



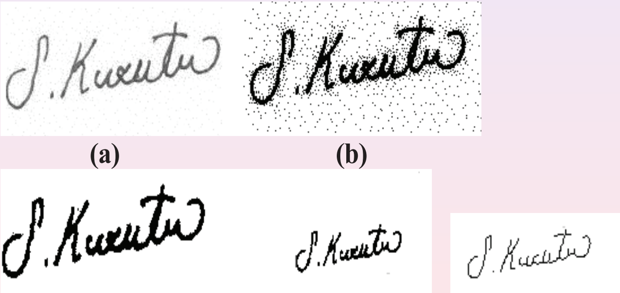
Example Biometrics - palm print

- Advantages: excellent biometric when combined with fingerprint.
- Disadvantages: high error rate, high-level subject cooperation, high precision of hardware.



Example Biometrics - signature

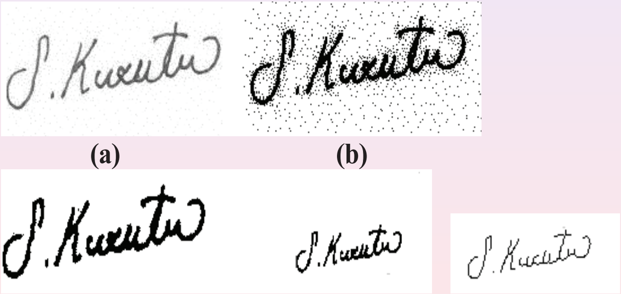
- Advantages: ease of use.
- Disadvantages: sensitive to perturbations, e.g., rotation, scaling, and translation, high false positives.



(a) (b) (c) (d) (e)

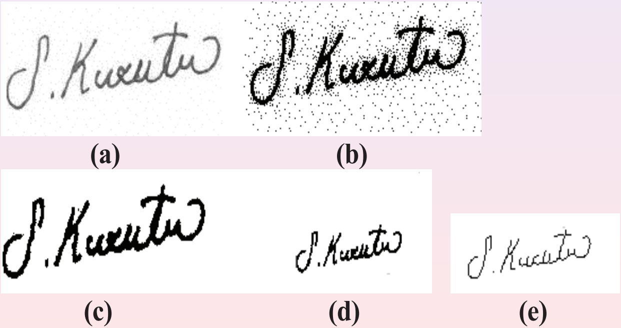
Example Biometrics - signature

- Advantages: ease of use.
- Disadvantages: sensitive to perturbations, e.g., rotation, scaling, and translation, high false positives.


(a) (b) (c) (d) (e)

Example Biometrics - signature

- Advantages: ease of use.
- Disadvantages: sensitive to perturbations, e.g., rotation, scaling, and translation, high false positives.



Example Biometrics - face

- Advantages: non-intrusive, low-level subject cooperation.
- Disadvantages: non-intrusive $\Rightarrow$ violation of privacy, changes in expression, lighting, pose, age, etc. Occlusion.

Original image

Example Biometrics - face

- Advantages: non-intrusive, low-level subject cooperation.
- Disadvantages: non-intrusive $\Rightarrow$ violation of privacy, changes in expression, lighting, pose, age, etc. Occlusion.

Example Biometrics - face

- Advantages: non-intrusive, low-level subject cooperation.
- Disadvantages: non-intrusive ⇒ violation of privacy, changes in expression, lighting, pose, age, etc. Occlusion.

Biometrics Comparison Chart

Biometric	Accuracy	Security Level	User Acceptance
Fingerprint	high	high	medium
Iris	high	high	medium
Signature	low-medium	medium	medium
Voice	low-medium	medium	high
Face	medium-high	medium	medium

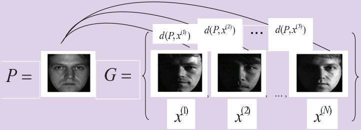
Biometric	Intrusive	Ease of Use	Cost/Hardware
Fingerprint	somewhat	high	special, cheap
Iris	non	medium	special, expansive
Signature	non	high	special, mid-price
Voice	non	high	common, cheap
Face	non	medium	common, cheap

Outline

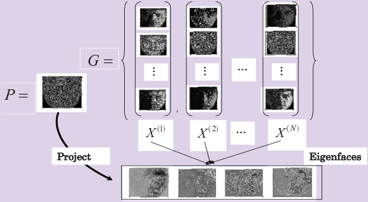
Classification Paradigms

Traditionally

1. single-to-single

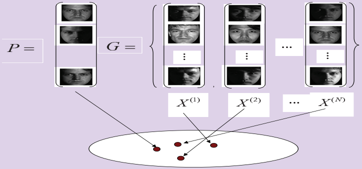


2. single-to-many

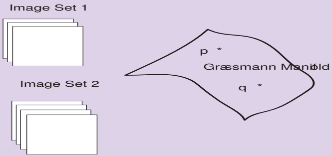


Currently

- many-to-many



- classification on $G(k, n)$



Outline

Definition of the Grassmannians

Definition

The Grassmannian $G(k, n)$ or the Grassmann manifold is the set of k-dimensional subspaces in an n-dimensional vector space K^n for some field K. i.e.,

$$G(k, n) = \{W \subset K^n \mid \dim(W) = k\}.$$

Example

For example, $G(1, 3)$ = set of all lines through the origin in $\mathbb{R}^3$ and $G(2, 3)$ = set of all planes through the origin in $\mathbb{R}^3$.

Definition of the Grassmannians

Definition

The Grassmannian $G(k, n)$ or the Grassmann manifold is the set of k-dimensional subspaces in an n-dimensional vector space K^n for some field K. i.e.,

$$G(k, n) = \{W \subset K^n \mid \dim(W) = k\}.$$

Example

For example, $G(1, 3)$ = set of all lines through the origin in $\mathbb{R}^3$ and $G(2, 3)$ = set of all planes through the origin in $\mathbb{R}^3$.

Definition of the Grassmannians

Definition

$$G(k,n) = \left\{ [p] \mid p \sim q \text{ if and only if } q = Q^T p \text{ for some } Q \in O_k \right\}.$$

i.e., if we vectorize data by columns, then points on G(k,n) are equivalence classes of n-by-k orthogonal matrices, where two matrices are equivalent if their columns span the same k-dimensional linear subspace.

Definition of the Grassmannians

Definition

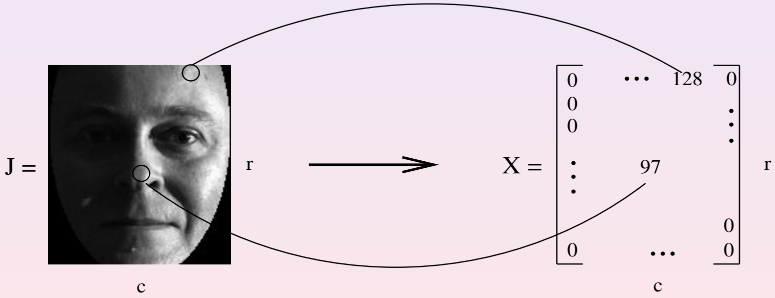
$$G(k,n) = \left\{ [p] \mid p \sim q \text{ if and only if } q = Q^T p \text{ for some } Q \in O_k \right\}.$$

i.e., if we vectorize data by columns, then points on G(k,n) are equivalence classes of *n*-by-*k* orthogonal matrices, where two matrices are equivalent if their columns span the same *k*-dimensional linear subspace.

Outline

Images as Points on the Grassmannians - 1

An r-by-c gray scale digital image corresponds to an r-by-c matrix where each entry enumerates one of the 256 possible gray levels of the corresponding pixel.



Images as Points on the Grassmannians - 2

Now, realize X by its columns and concatenate columns into a single column vector:

$$X = \begin{bmatrix} x_1 & | & x_2 & | & \cdots & | & x_c \end{bmatrix} \in \mathbb{R}^{r \times c} \quad \longrightarrow \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_c \end{bmatrix} \in \mathbb{R}^{rc \times 1}$$

Thus, an image J can be realized as a column vector of length equal to the product of J's resolutions.

Images as Points on the Grassmannians - 3

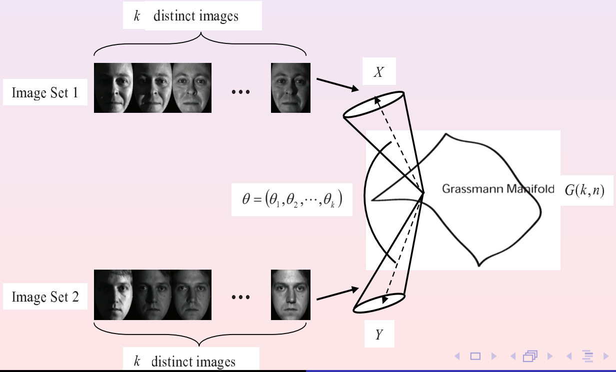
- Now, for a subject i, we collect k distinct images (which corresponds to k column vectors) and concatenate them into a single data matrix $X^{(i)}$ so that

$$X^{(i)} = \left[x_1^{(i)} \;\middle|\; x_2^{(i)} \;\middle|\; \cdots \;\middle|\; x_k^{(i)} \right]$$

  and $rank(X^{(i)}) = k$ with each $x_j^{(i)} \in \mathbb{R}^n$ being an image of resolution n.
- The column space of $X^{(i)}$ then gives a point on the Grassmannian $G(k, n)$.

Images as Points on the Grassmannians - 4

Although the matrix representation of this point on $G(k, n)$ is not unique, $X^{(i)}$ does give a way to represent the equivalence class in $G(k, n)$ on the computer. Thus, A matrix $X \in \mathbb{R}^{n \times k}$ naturally corresponds to a subspace $\mathcal{X} \in G(k, n)$.



Connection

- Thus, a face recognition problem where multiple images are available for the probe and gallery $\rightarrow$ a classification problem on the Grassmannians.
- Like any classification problem, we need a way to measure the *geometrically sound* distance between points in the classification space.
- It turns out that any attempt to construct a unitarily invariant metric on the Grassmann manifold will yield something that can be expressed in terms of principal angles.

Connection

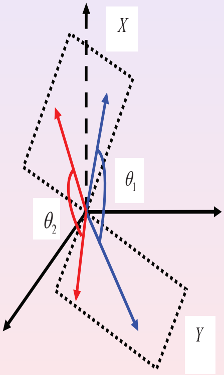
- Thus, a face recognition problem where multiple images are available for the probe and gallery $\rightarrow$ a classification problem on the Grassmannians.
- Like any classification problem, we need a way to measure the *geometrically sound* distance between points in the classification space.
- It turns out that any attempt to construct a unitarily invariant metric on the Grassmann manifold will yield something that can be expressed in terms of principal angles.

Connection

- Thus, a face recognition problem where multiple images are available for the probe and gallery $\longrightarrow$ a classification problem on the Grassmannians.
- Like any classification problem, we need a way to measure the *geometrically sound* distance between points in the classification space.
- It turns out that any attempt to construct a unitarily invariant metric on the Grassmann manifold will yield something that can be expressed in terms of principal angles.

Outline

Principal Angles



Definition (Björck & Golub, 1973)

If X and Y are two vector subspaces of a unitary space $\mathbb{E}^n$ such that $p = \dim(X) \geq \dim(Y) = q \geq 1$, then the *principal angles* $\theta_k \in [0, \frac{\pi}{2}], 1 \leq k \leq q$ between X and Y are defined recursively by

$$\cos(\theta_k) = \max_{u \in X} \max_{v \in Y} \left| u^H v \right| = \left| u_k^H v_k \right|$$

subject to $||u||_2 = ||v||_2 = 1$, $u^H u_i = 0$ and $v^H v_i = 0$ for $i = 1, 2, \ldots, k-1$.

Grassmannian Distances

Let $\theta = (\theta_1, \ldots, \theta_q)$ be the principal angle vector.

Example Grassmannian Distances [Edelman *et al.*, 1999]

- **arc length (geodesic)** $d_g(X, Y) = ||\theta||_2$
- **Fubini-Study** $d_{FS}(X, Y) = \cos^{-1}\left(\prod_{i=1}^{k} \cos\theta_i\right)$
- **chordal (projection F) distance** $d_c(X, Y) = \|\sin\theta\|_2$
- **subspace distance** $d_s(X, Y) = ||\theta||_\infty$

Various Realizations of the Grassmannian

1. First, as a quotient (homogeneous space) of the orthogonal group,

$$G(k,n) = O(n)/O(k) \times O(n-k). \quad (1)$$

2. Next, as a submanifold of projective space,

$$G(k,n) \subset \mathbb{P}(\Lambda^q \mathbb{R}^n) = \mathbb{P}^{\binom{n}{k}-1}(\mathbb{R}) \quad (2)$$

via the Plücker embedding.

3. Finally, as a submanifold of Euclidean space,

$$G(k,n) \subset \mathbb{R}^{(n^2+n-2)/2} \quad (3)$$

via a projection embedding described recently in [Conway *et al.*, 1996].

Various Realizations of the Grassmannian

1. First, as a quotient (homogeneous space) of the orthogonal group,

$$G(k,n) = O(n)/O(k) \times O(n-k). \tag{1}$$

2. Next, as a submanifold of projective space,

$$G(k,n) \subset \mathbb{P}(\Lambda^q \mathbb{R}^n) = \mathbb{P}^{\binom{n}{k}-1}(\mathbb{R}) \tag{2}$$

   via the Plücker embedding.

3. Finally, as a submanifold of Euclidean space,

$$G(k,n) \subset \mathbb{R}^{(n^2+n-2)/2} \tag{3}$$

   via a projection embedding described recently in [Conway *et al.*, 1996].

Various Realizations of the Grassmannian

1. First, as a quotient (homogeneous space) of the orthogonal group,

$$G(k,n) = O(n)/O(k) \times O(n-k). \tag{1}$$

2. Next, as a submanifold of projective space,

$$G(k,n) \subset \mathbb{P}(\Lambda^q \mathbb{R}^n) = \mathbb{P}^{\binom{n}{k}-1}(\mathbb{R}) \tag{2}$$

via the Plücker embedding.

3. Finally, as a submanifold of Euclidean space,

$$G(k,n) \subset \mathbb{R}^{(n^2+n-2)/2} \tag{3}$$

via a projection embedding described recently in [Conway *et al.*, 1996].

The Corresponding Grassmannian Distances

1. The standard invariant Riemannian metric on orthogonal matrices $O(n)$ descends via (1) to a Riemannian metric on the homogeneous space $G(k, n)$. We call the resulting geodesic distance function on the Grassmannian the *arc length* or *geodesic* distance and denote it d_g.
2. If one prefers the realization (2), then the Grassmannian inherits a Riemannian metric from the *Fubini-Study* metric on projective space (see, e.g., [Griffiths & Harris, 1978]).
3. One can restrict the usual Euclidean distance function on $\mathbb{R}^{(n^2+n-2)/2}$ to the Grassmannian via (3) to obtain the *projection F* or *chordal* distance d_c.

The Corresponding Grassmannian Distances

1. The standard invariant Riemannian metric on orthogonal matrices $O(n)$ descends via (1) to a Riemannian metric on the homogeneous space $G(k, n)$. We call the resulting geodesic distance function on the Grassmannian the *arc length* or *geodesic* distance and denote it d_g.
2. If one prefers the realization (2), then the Grassmannian inherits a Riemannian metric from the *Fubini-Study* metric on projective space (see, e.g., [Griffiths & Harris, 1978]).
3. One can restrict the usual Euclidean distance function on $\mathbb{R}^{(n^2+n-2)/2}$ to the Grassmannian via (3) to obtain the *projection F* or *chordal* distance d_c.

The Corresponding Grassmannian Distances

1. The standard invariant Riemannian metric on orthogonal matrices $O(n)$ descends via (1) to a Riemannian metric on the homogeneous space $G(k, n)$. We call the resulting geodesic distance function on the Grassmannian the *arc length* or *geodesic* distance and denote it d_g.
2. If one prefers the realization (2), then the Grassmannian inherits a Riemannian metric from the *Fubini-Study* metric on projective space (see, e.g., [Griffiths & Harris, 1978]).
3. One can restrict the usual Euclidean distance function on $\mathbb{R}^{(n^2+n-2)/2}$ to the Grassmannian via (3) to obtain the *projection F* or *chordal* distance d_c.

Outline

1. Grassmann Separability Measure

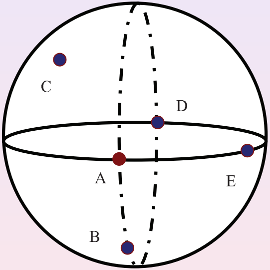
A data set $P = \{C^{(1)}, C^{(2)}, \ldots, C^{(N)}\}$ of N distinct classes having subspace configurations $C^{(i)} = \left\{S_1^{(i)}, S_2^{(i)}\right\}$ for each i is **Grassmann separable** if $\max_{i=j} d_{ij} < \min_{i \neq j} d_{ij}$.

d	$S_1^{(1)}$	$S_1^{(2)}$	$\cdots$	$S_1^{(N)}$
$S_2^{(1)}$	d_{11}	d_{12}	$\cdots$	d_{1N}
$S_2^{(2)}$	d_{21}	d_{22}	$\cdots$	d_{2N}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$S_2^{(N)}$	d_{N1}	d_{N2}	$\cdots$	d_{NN}

FAR at 0% FRR is the ratio of the number of off-diagonal entries that have distances smaller than the maximum of the diagonal entries divided by the total number of off-diagonal entries.

2. Nearest Neighbor Classifier

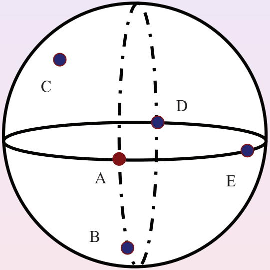
"Distances used in classifications depend on the Geometry of the data"



- What's the nearest neighbor of the point A on this S^2?
- So, if we were to classify A, then it would be assigned the identity of B.

2. Nearest Neighbor Classifier

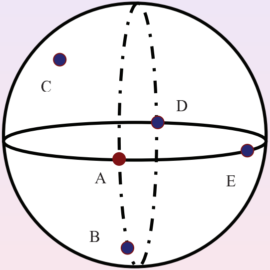
"Distances used in classifications depend on the Geometry of the data"



- What's the nearest neighbor of the point A on this S^2?
- So, if we were to classify A, then it would be assigned the identity of B.

2. Nearest Neighbor Classifier

"Distances used in classifications depend on the Geometry of the data"

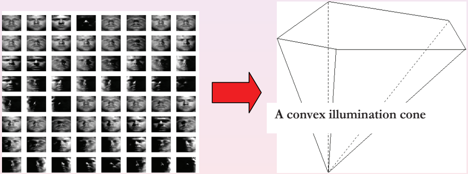


- What's the nearest neighbor of the point A on this S^2?
- So, if we were to classify A, then it would be assigned the identity of B.

Outline

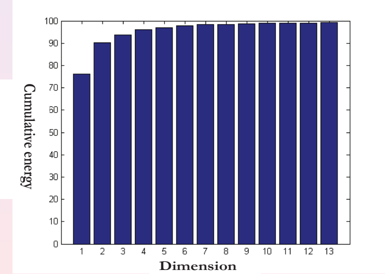
Geometry of Illumination Variations - 1

The set of n-pixel monochrome images of an object of any shape with a general reflectance function, seen under all possible illumination conditions, forms a convex polyhedral cone [Belhumeur & Kriegman, 1998].

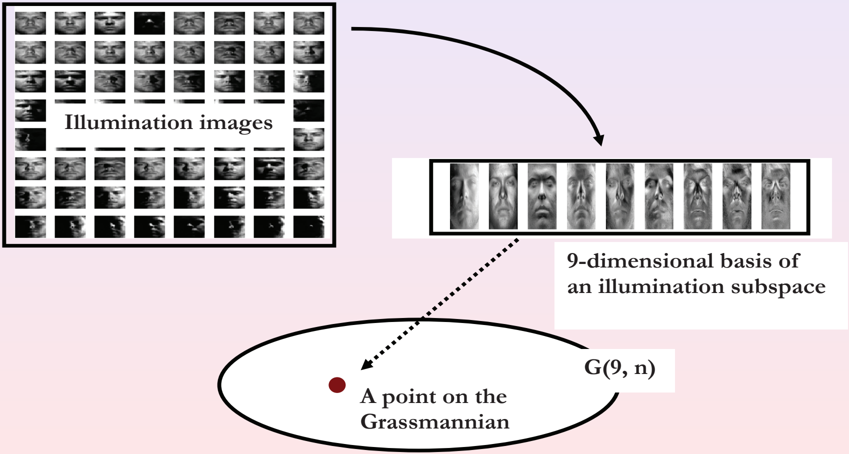


Geometry of Illumination Variations - 2

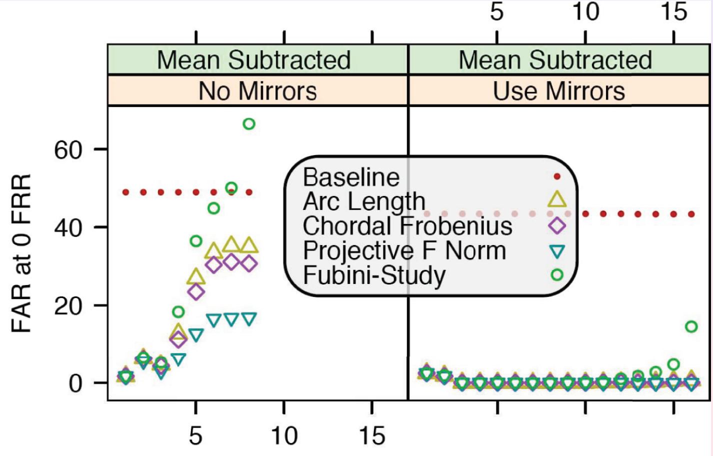
The set of images of a convex, Lambertian object seen under arbitrary distance light sources lies approximately in a 9-dimensional linear subspace with over 99% of the energy [Basri & Jacobs, 2003].



Connection to Classification on the Grassmannians



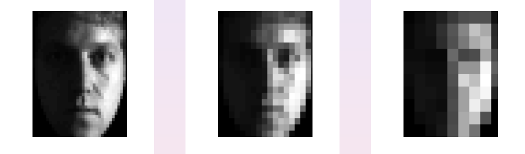
Example Classification Results [Chang *et al.*, 2006b]



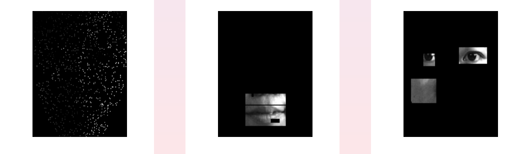
Outline

Three Forms of Projection

The idiosyncratic nature of the illumination spaces persist under all forms of mathematical projections.

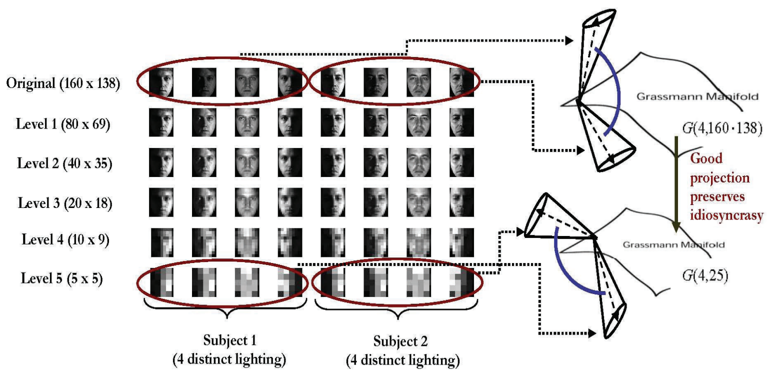


(a) Patch collapsing

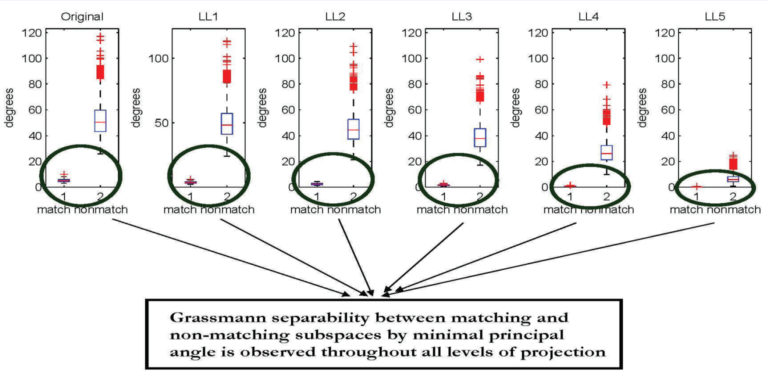


(b) Patch projections

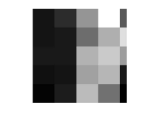
Example Classification Result - Patch Collapsing

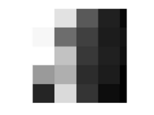


Example Classification Result - Patch Collapsing [Chang *et al.*, 2007a]



Potential Use: 25-pixel Illumination Camera





Large private databases of facial imagery can be stored at a resolution that is sufficiently low to prevent recognition by a human operator yet sufficiently high to enable machine recognition.

Outline

Non-linear Problems

- The method of mapping sets of images to points on the Grassmann manifold is not limited to intrinsically linear data sets.
- Two class problems (e.g. tree or non-tree, cloud or tree, human or non-human, cat or dog) versus multi-class problems (e.g. a group of human subjects, a collection of oranges).

Non-linear Problems

- The method of mapping sets of images to points on the Grassmann manifold is not limited to intrinsically linear data sets.
- Two class problems (e.g. tree or non-tree, cloud or tree, human or non-human, cat or dog) versus multi-class problems (e.g. a group of human subjects, a collection of oranges).

Glasses vs. No-glasses

(a) glasses class

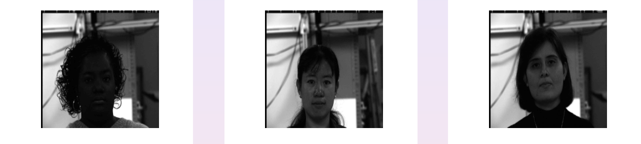
(b) no-glasses class

Empirical Evidences [Chang *et al.*, 2006a]

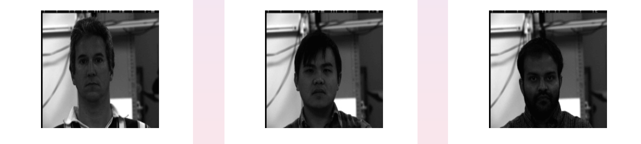
Gallery set size = 10, glasses vs. no-glasses											
	Trial number										
ℓ	1	2	3	4	5	6	7	8	9	10	μ
1	0	0	0	0	0	0	0	0	0	0	0
2	50	38	13	45	35	48	28	38	35	40	37
3	53	55	53	63	48	48	48	53	70	45	53

Table: Misclassification percentage out of 40 testing sets where 20 are of glasses class and 20 are of no-glasses class. Gallery size refers to the number of distinct sets used in representing the glasses and no-glasses classes. The distance function is the ℓ-truncated chordal distance. The experiment is repeated ten times with the mean shown in the last column.

Male vs. Female



(a) female class



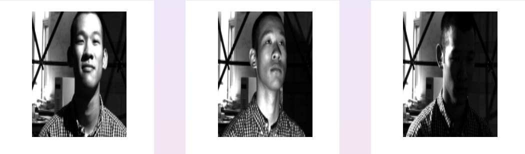
(b) male class

Empirical Evidences [Chang *et al.*, 2006a]

gallery set size = 3, male-female											
	Trial number										
ℓ	1	2	3	4	5	6	7	8	9	10	$\mu \pm \sigma$
1	0	0	0	0	0	0	23	0	10	0	3 ± 8
2	35	0	0	5	13	28	23	5	10	8	13 ± 12
3	30	0	0	0	0	28	28	5	23	0	11 ± 14

Table: Misclassification percentage out of 40 testing sets where 20 are of male class and 20 are of female class. Gallery set size refers to the number of distinct sets used in representing the male and female classes. The distance function is the ℓ-truncated chordal distance. The experiment is repeated ten times where the mean and standard deviation is reported in the last column of the table.

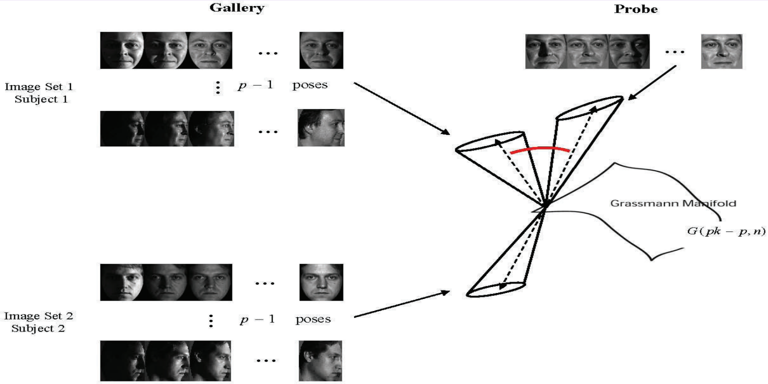
Pose Problem



(a) pose and illumination variations of subject 1

(b) pose and illumination variations of subject 2

Pose Problem - Difficulty



Empirical Evidences [Chang *et al.*, 2007b]

Error Rate (%)	Experiment		
Database	I	II	III
Extended YDB	0	0	6.7
CMU-PIE	0	0	43.2

Table: Average recognition error rate for Experiments I – III with d_c^1 on both Extended YDB and CMU-PIE.

pose	c02	c05	c07	c09	c11	c14	c22
error (%)	13.4	31.3	83.6	73.1	0	1.5	23.9
pose	c25	c27	c29	c31	c34	c37	
error (%)	82.1	22.4	16.4	80.6	76.1	56.7	

Table: Average break-down recognition error rate for each pose in Experiment III using d_c^1 on CMU-PIE.

Outline

Summary

1. When the classification problem is defined as comparing distances of a set of images to a set of images, this corresponds naturally to a classification problem on the Grassmannians.
2. Well-established metrics (in terms of principal angles) on the Grassmann manifold are available for calculating distances between points on the manifold.
3. Example face recognition problems are given to demonstrate the potential use of this paradigm.

Outline

Future Work

1. Consider the statistics on the Grassmann manifold and the use of Karcher mean as a prototype to compress data and accelerates computations.
2. Use perturbation theory for matrices to evaluate the robustness of the Grassmann method.
3. Examine the effect of image resolution to the accuracy of classification algorithms.

Basri, R., & Jacobs, D. 2003.

Lambertian Reflectance and Linear Subspaces.

PAMI, **25**(2), 218–233.

Belhumeur, P., & Kriegman, D. 1998.

What Is the Set of Images of an Object Under All Possible Illumination Conditions.

IJCV, **28**(3), 245–260.

Chang, J.-M., Kirby, M., Kley, H., Beveridge, J.R., Peterson, C., & Draper, B. 2006a.

Examples of Set-to-Set Image Classification.

Pages 102–105 of: Seventh International Conference on Mathematics in Signal Processing Conference Digest.

The Royal Agricultural College, Cirencester: Institute for Mathematics and its Applications.

Chang, J.-M., Beveridge, J.R., Draper, B., Kirby, M., Kley, H., & Peterson, C. 2006b (June).
Illumination Face Spaces Are Idiosyncratic.
Pages 390–396 of: Int'l Conf. on Image Procesing & Computer Vision, vol. 2.

Chang, J.-M., Kirby, M., Kley, H., Peterson, C., Draper, B., & Beveridge, J. R. 2007a.
Recognition of Digital Images of the Human Face at Ultra Low Resolution Via Illumination Spaces.
Pages 733–743 of: Computer Vision – ACCV 2007.
LNCS, vol. 4844.
Springer.

Chang, J.-M., Kirby, M., & Peterson, C. 2007b (September).
Set-to-set face recognition under variations in pose and illumination.

In: 2007 Biometrics Symposium.

- Conway, J., Hardin, R., & Sloane, N. 1996.
  Packing lines, planes, etc.: Packings in Grassmannian Spaces.
  Experimental Mathematics, **5**, 139–159.

- Edelman, A., Arias, T., & Smith, S. 1999.
  The Geometry of Algorithms with Orthogonality Constraints.
  SIAM J. Matrix Anal. Appl., **20**(2), 303–353.

- Griffiths, Phillip, & Harris, Joseph. 1978.
  Principles of Algebraic Geometry.
  Wiley & Sons.