

Some examples of Grassmann and Flag manifolds in data analysis

August 16, 2019

Geometric/Topological Data Analysis?

Geometric/Topological Data analysis is the application of tools from geometry/topology to elucidate structure in large data sets.

This talk will describe a couple of examples utilizing Grassmann and Flag manifolds.

For visualization purposes, the examples will be in the context of digital black and white photographs.

In order to solidify the setting, let's first review a few basic concepts related to digital photographs.

Black and white images as rectangular arrays

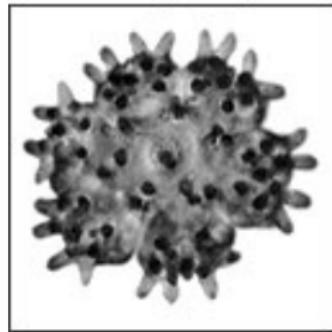
A black and white digital picture is stored as a matrix of numbers.

Each number in the matrix corresponds to a pixel and is a measure of the energy arriving from a corresponding portion of the image.

Frequently, the numbers lie between 0 and 255.

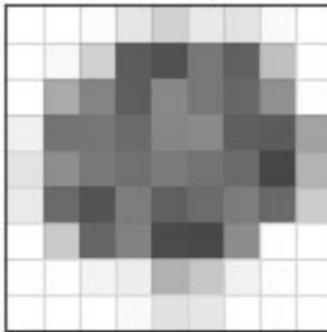
Creation of a Digital Image

Analog Image



(a)

Digital Sampling



(b)

Pixel Quantization

249	244	240	230	209	233	227	251	255
248	245	210	93	81	120	97	193	254
250	170	133	94	137	120	104	145	253
241	116	118	107	134	138	96	92	163
277	142	121	113	124	115	107	71	179
234	106	84	125	97	108	125	106	204
241	202	102	132	75	73	141	246	252
253	252	244	239	178	199	242	250	245
255	249	244	250	226	231	240	251	253

(c)

Figure 1

<http://micro.magnet.fsu.edu/primer/digitalimaging/digitalimagebasics.html>

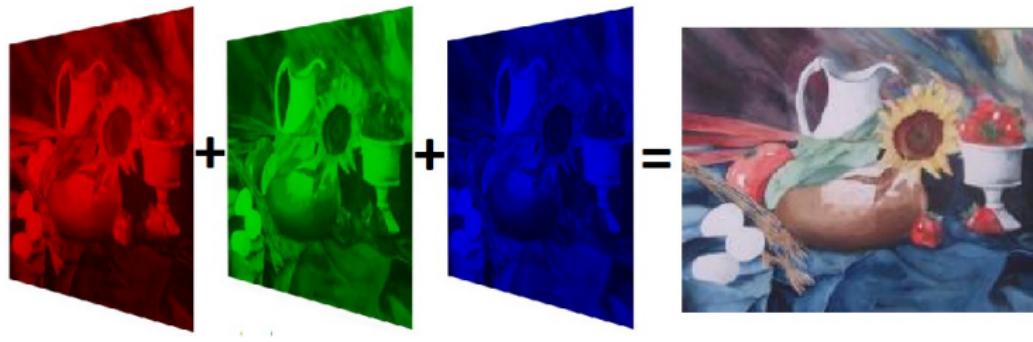
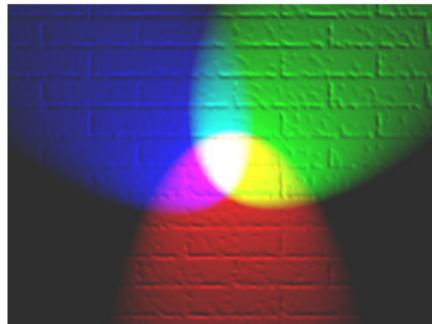
Color images as rectangular hyper-arrays

In a digital camera, a **color** image is typically stored as a three sheeted rectangular array of numbers.

Each sheet in the array corresponds to energy arriving in a particular color band. In many cameras, these sheets are the **red**, **green**, and **blue** contributions.

We can visualize each sheet as a black and white photograph or as a monochromatic image.

By combining basic colors, we can produce a wider range of colors.



Color is really a function of wavelength

We think of color as being composed of three basic colors. We think this way because this is how (most) humans are built. But colors are much richer than what we can see.

Humans are typically tri-chromats. We see three dimensions worth of color.

Some people are bi-chromats (color blind) and see two dimensions worth of color. There are even mono-chromats (who only see in black and white).

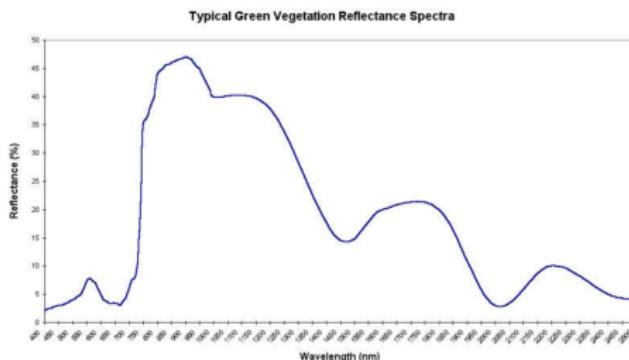
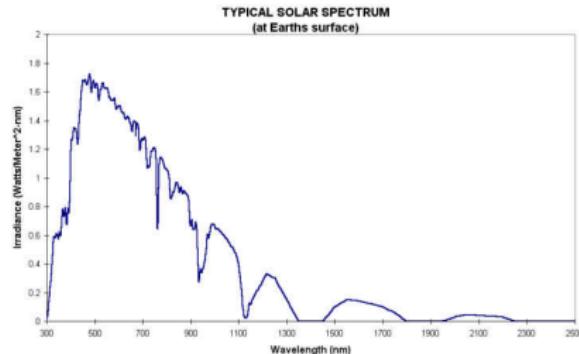
In the other direction, there are some rare people who are tetra-chromats. They see a much richer collection of colors.

The color you see depends on what wavelengths are being reflected by an object.

These reflected wavelengths, in turn, depend on what wavelengths are illuminating an object.

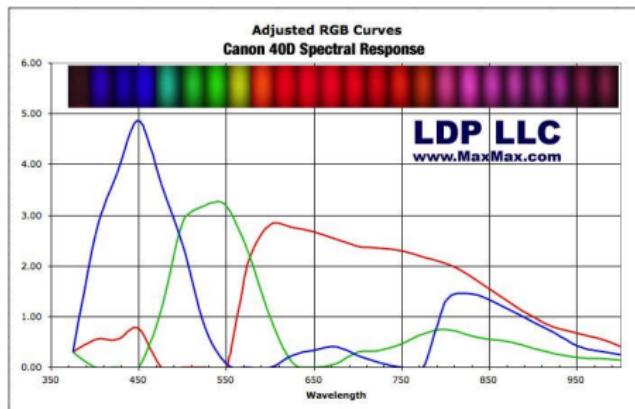
Color forms an infinite dimensional vector space. Humans pick a particular projection onto a three dimensional vector space.

<http://www.cropscan.com/response.html>

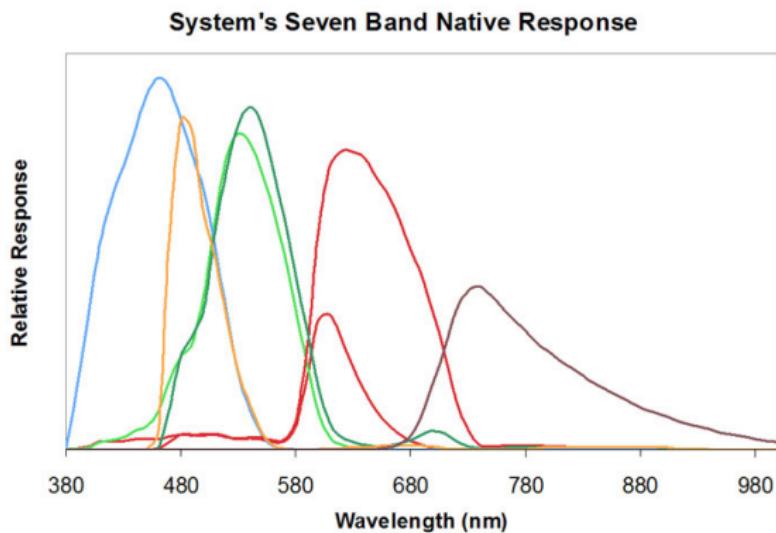


If intensity is considered as a function of wavelength, $I(\lambda)$, then a pixel value (roughly) corresponds to an integral of $I(\lambda)$ multiplied against a frequency response function, $F(\lambda)$.

In other words: Pixel Value = $Constant \cdot \int_{\lambda} I(\lambda) \cdot F(\lambda) d\lambda$.



Some cameras can collect more samples of the spectral curve by integrating against more frequency response curves. Below are some curves from a higher end multi-spectral camera.



<http://www.fluxdata.com/products/high-resolution-3-ccd-multispectral-camera>

Most digital cameras have the ability to collect data from outside the visible spectrum (but can only realize this ability through the use of filters).

It can be surprising when you “shift” data outside the visible spectrum to lie in the visible spectrum.

Below is a flower imaged in the visible spectrum and the same flower imaged in the UV-spectrum.



You are already familiar with this idea of “shifting” non-visible colors into the realm of visible colors if you are a fan of astronomy.

You are already familiar with this idea of “shifting” non-visible colors into the realm of visible colors if you are a fan of astronomy.

Orion Nebula as imaged by the Hubble space telescope



Variations of state

There are several natural variations of state that can be considered, here are a few:

$SO(n), O(n)$

$\mathbb{R}^n, E(n), SE(n)$

Parameter spaces

There are also multiple parameter spaces that are useful for processing large data sets collected under a variation of state:

Grassmann, Veronese, Segre Manifolds

Flag Manifold

Steifel Manifold

Getting data to lie on one of these parameter spaces is often achieved through an application of the Singular Value Decomposition.

This is a method of capturing as much “Energy” as possible in a data set in an efficient manner.

By considering a photograph as a matrix or array, you can apply standard linear algebra techniques such as the **Singular Value Decomposition**:

$$M = U \Sigma V^*$$

where:

U is a unitary matrix,

Σ is a “diagonal” matrix with monotone decreasing non-negative real numbers down the diagonal,

V^* is a unitary matrix.

Columns of $U \leftrightarrow$ eigenvectors of $M M^*$.

Columns of $V \leftrightarrow$ eigenvectors of $M^* M$.

Diagonal entries in $\Sigma \leftrightarrow$ square roots of eigenvalues of $M M^*$

There is a great deal of geometric information that can be extracted from the SVD.

For instance, it provides a way of finding the closest rank r matrix to M .

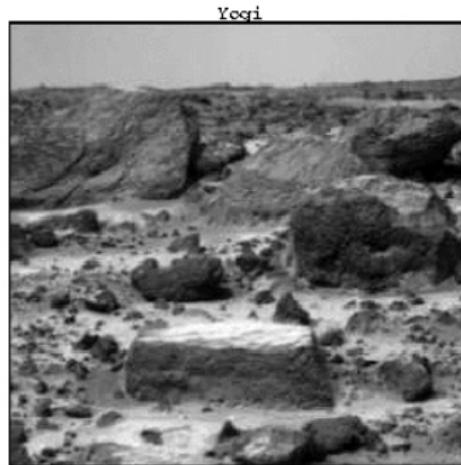
More precisely, it lets you find a rank r matrix M_r that minimizes $\|M - M_r\|$.

In particular,

$$M_r = U \Sigma_r V^*$$

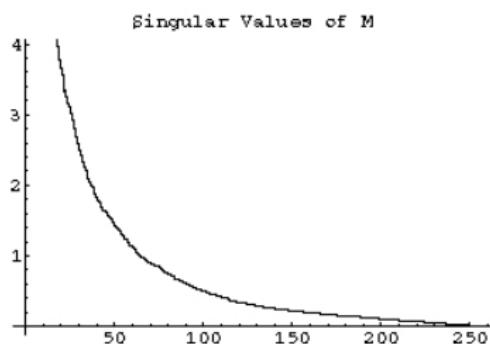
Low Rank Image: Image Compression

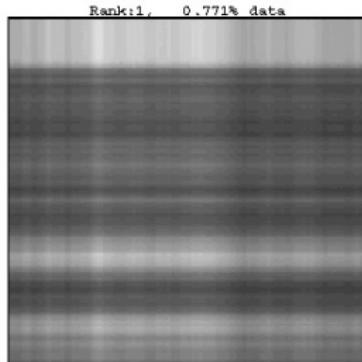
Picture sent back from Mars: Resolution is 256 x 264



<http://www.uwlax.edu/faculty/will/svd/compression/index.html>

Matrix has rank 256, plot of singular values of normalized matrix





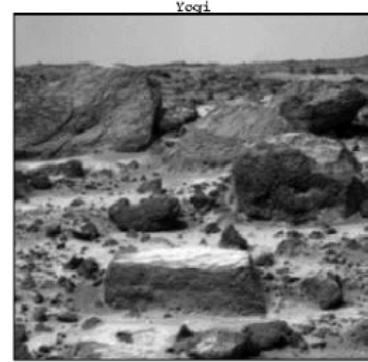
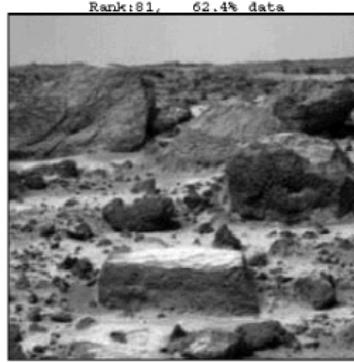
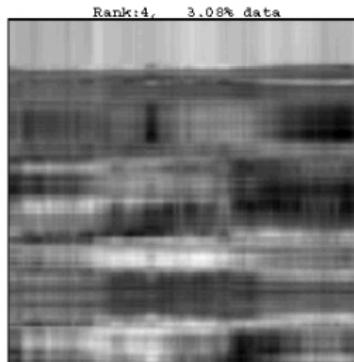
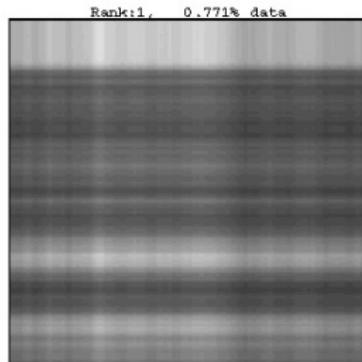
$$= \sigma_1 U_1 V_1^*$$

$$\text{Original picture} = \sigma_1 U_1 V_1^* + \sigma_2 U_2 V_2^* + \cdots + \sigma_k U_k V_k^* + \dots$$

Some background in digital imagery

Grassmann varieties
Experiments on CMU-PIE
A Flag Example

Images as matrices
Images as points
Low dimensional approximation



SVD, Segre Varieties, Secant Varieties

The Segre variety $\mathbb{P}^a \times \mathbb{P}^b$ is essentially the locus of outer products of vectors of length $a + 1$ and $b + 1$.

Thus, the singular value decomposition expresses a picture as a linear combination of points on a Segre Variety.

$$\text{Original picture} = \sigma_1 U_1 V_1^* + \sigma_2 U_2 V_2^* + \cdots + \sigma_k U_k V_k^* + \dots$$

In particular, a rank r approximation of a picture represents the picture as a point on an r -secant variety to a Segre variety.

Digital images as vectors

A black and white digital picture determines a point in $\mathbb{R}^{r \times c}$.

By “flattening” the matrix, we obtain a point in $\mathbb{R}^{rc \times 1}$.

Here we see a 2×3 matrix get flattened to a 6×1 matrix.

$$\begin{bmatrix} 1 & 3 & 17 \\ 2 & 6 & 19 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 2 \\ 3 \\ 6 \\ 17 \\ 19 \end{bmatrix}.$$

Eigenfaces

Consider a matrix, M , whose columns are flattened photographs.

From the SVD: $M = U \Sigma V^*$, the columns of U give a distinguished basis for the column space of M .

If the columns of M have some correlated property then this property is reflected in the columns of U .

For instance, if the columns of M are faces, we get “Eigenfaces”.

Some background in digital imagery

Grassmann varieties

Experiments on CMU-PIE

A Flag Example

Images as matrices

Images as points

Low dimensional approximation



These pictures are an ordered orthonormal basis spanning an approximation of "facespace".

A new face can be written as a linear combination of these pictures plus a "residual" in the null space of the span of the eigenfaces.

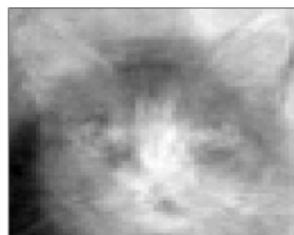
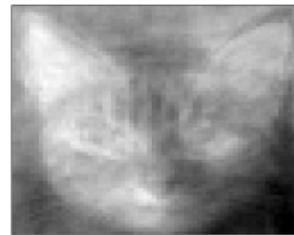
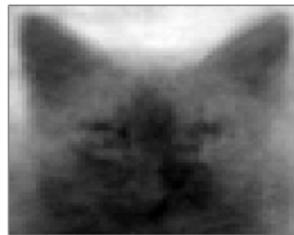
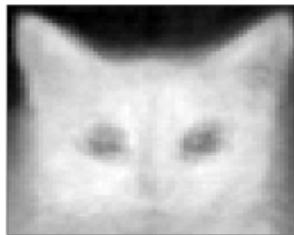
Thus, to a new face we can attach a coordinate tag and measure the "novelty" or length of the residual.

Of course the same process can be carried out with other data sets.

Collection of cats



Eigencats



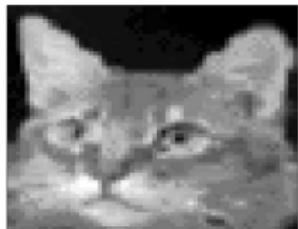
Eigencats and Projection



Novelty= 5.83×10^{-12}

Eigencats and Projection

First Cat



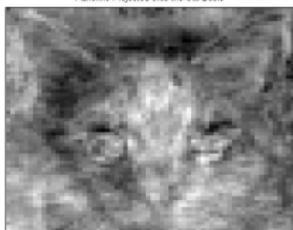
First Cat Projected onto the Cat Basis



$$\text{Novelty} = 5.83 \times 10^{-12}$$



Adrienne Projected onto the Cat Basis



Residual of Adrienne



$$\text{Novelty} = 1.8781 \times 10^3$$

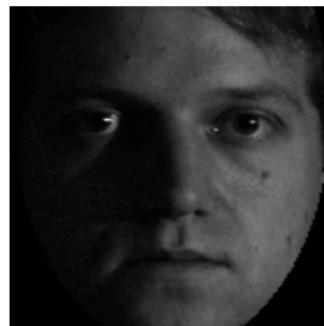
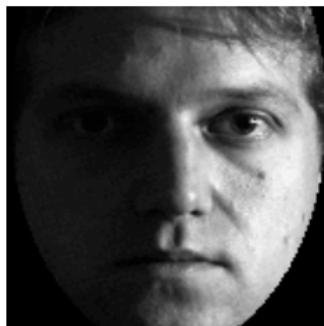
A Grassmann example: Illumination space

We now consider the illumination space of an object.

More precisely, we consider the span of a collection of images of a fixed object under varying illumination conditions.

As a specific example, consider the illumination space of a person.

Images of a person under different illumination conditions



Illumination images form a convex set

If \mathbf{A}, \mathbf{B} are vectors representing two images of the same object under two different illumination conditions, then any convex combination of \mathbf{A} and \mathbf{B} represents the object under a convex combination of these lighting conditions.

Consequence: the set of images of a fixed object collected under varying illumination conditions form a convex set.

i.e. any weighted “average” of two such images is a valid image.

An average illumination image.

The picture in the middle is artificial. It was created by adding the matrices and dividing by 8.



Singular values of an illumination data matrix

Basri and Jacobs (2000) showed that a collection of digital photos of a convex Lambertian object under distant lighting lies close to a 9D linear subspace.

An implication is that the illumination space of many objects is well approximated by a low-dimensional linear space.

Thus, if the columns of a data matrix consist of images of a fixed object taken under a wide variety of illumination conditions, then the singular values of the matrix decay rapidly.

Objects which do not reflect light well are Lambertian, they have the effect of smoothing (rugs and fur are very Lambertian, human faces are somewhat Lambertian).

Objects which do not reflect light well are Lambertian, they have the effect of smoothing (rugs and fur are very Lambertian, human faces are somewhat Lambertian).



Objects which do not reflect light well are Lambertian, they have the effect of smoothing (rugs and fur are very Lambertian, human faces are somewhat Lambertian).



Objects which reflect light very well are non-Lambertian (mirror disco balls and water are non-Lambertian).



Non-Lambertian images - slower decay in singular values

Illumination data matrices of non-Lambertian objects have a slower decay of singular values.

They are less well approximated by a 9-dimensional space; it takes more dimensions to capture their illumination variance.

Let $\text{Gr}(k,n)$ denote the Grassmann variety of k -dimensional subspaces of a fixed n -dimensional vector space.

So, a point on $\text{Gr}(k,n)$ corresponds to a k -dimensional subspace.

Since the singular values of an illumination data matrix decay rapidly, there is a low dimensional linear space that captures the majority of the energy of an illumination data set.

For a given k , the singular value decomposition tells us which k -dimensional linear space best approximates the illumination space:

The column space of the closest rank k matrix will be a point on $\text{Gr}(k,n)$.

Image Set 1

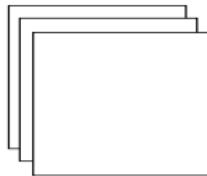
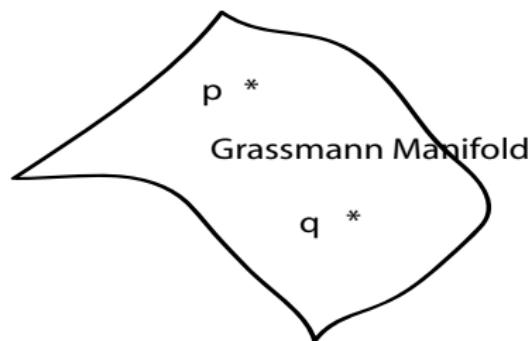


Image Set 2



Principal Angles

Given two k -dimensional spaces U, V , the first principal angle, θ_1 , between the two spaces is the smallest possible angle between vectors $u_1 \in U$ and $v_1 \in V$.

The second principle angle, θ_2 , is the smallest possible angle between vectors $u_2 \in U$ and $v_2 \in V$ subject to the constraint that $u_1 \perp u_2$ and $v_1 \perp v_2$.

Continuing in this way we obtain a vector of principal angles:
 $\Theta(V, W) = (\theta_1, \theta_2, \dots, \theta_k)$.

Metrics on Grassmann varieties

Any orthogonally (unitarily) invariant metric on a Grassmann variety can be expressed as a function of $\Theta(V, W)$.

$\Theta(V, W)$ is readily computed using an SVD-based algorithm.

In particular, if the columns of A, B are orthonormal bases for V, W then the i^{th} singular value of A^*B is $\cos(\theta_i)$.

The CMU-PIE data base is a database (developed at Carnegie-Mellon University) containing images of people with variations in **Illumination, Pose and Expression**.

For the following experiment, we restrict our attention to images in the CMU-PIE database with a variation only in **Illumination**.

Description of experiment

For each of the 67 subjects in the CMU-PIE database:

Randomly select 10 disjoint sets of 10 images.

Their spans produce 10-dimensional estimates of the illumination space for the subject.

This generates a total of 670 matching subspaces and 44,220 non-matching subspaces.

An example of a data point



Let D be a distance measure on a Grassmann variety. When the largest distance between any two matching subspaces is less than the smallest distance between any two non-matching subspaces, the data is called D -Grassmann separable.

Using the first principal angle, we observed a significant gap between matching and non-matching subspaces (approximately 16°) when subspaces are realized as points in $Gr(10, 22080)$.

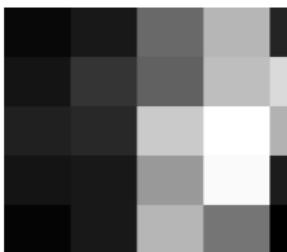
I.e. the data is Grassmann separable by the first principal angle.

Through projection, we can map points on $Gr(10, 22080)$ to $Gr(10, N)$ with $N < 22080$.

A generic projection isn't so far off from an isometry: it roughly preserve lengths and angles between random vectors.

In the next experiment, we describe what happens if we consider a rather special linear map derived from the 2-d Haar wavelet.

Intuitively, we use the Haar wavelet to replace a region of an image with its average pixel value.



Application of the 2D Haar wavelet transform.

The resulting LL image is displayed at each level.

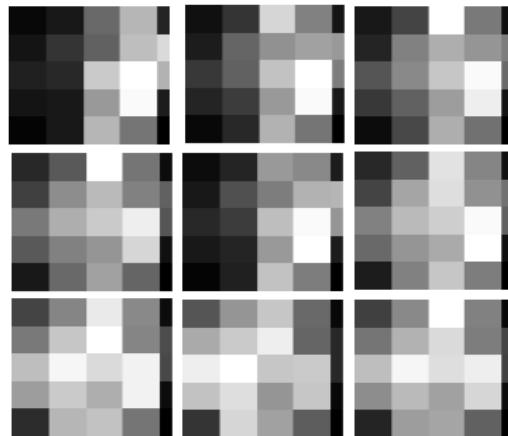
This transform mimics a low resolution camera.

25 pixels

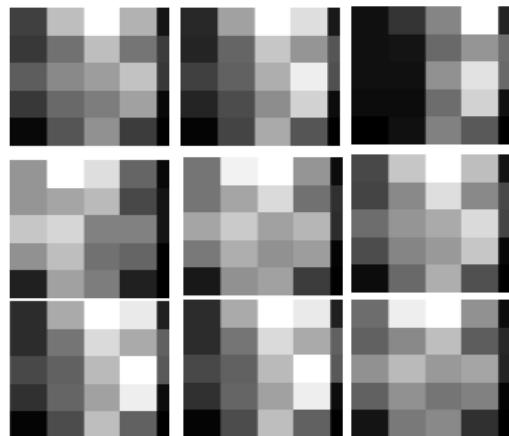
Under drastic resolution reduction, the CMU-PIE data base is still Grassmann separable.

However, the separation gap between matching and non-matching subspaces drops from 16° to 14° , 8° , and 0.17° when subspaces are realized as points in $G(10, 22080)$, $G(10, 360)$, $G(10, 90)$, and $G(10, 25)$, respectively.

Low resolution illumination spaces

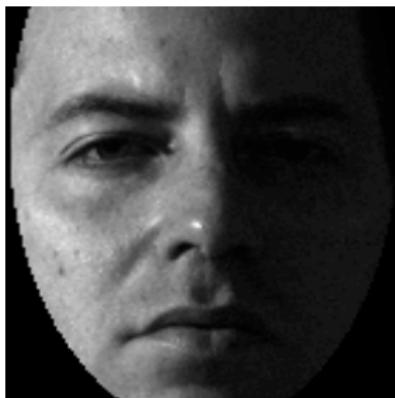


Subject A



Subject B

Sampling of a strip of pixels



Sampling of a random set of pixels



Flags

A Flag is a nested sequence of vector spaces $V_1 \subset V_2 \subset \cdots \subset V_k$.

The signature of the flag is the k -tuple $(\dim(V_1), \dim(V_2), \dots, \dim(V_k))$.

The Flag manifold, $FL(d_1, d_2, \dots, d_k; n)$, is the manifold whose points parametrize the flags of signature (d_1, d_2, \dots, d_k) inside a fixed n -dimensional vector space.

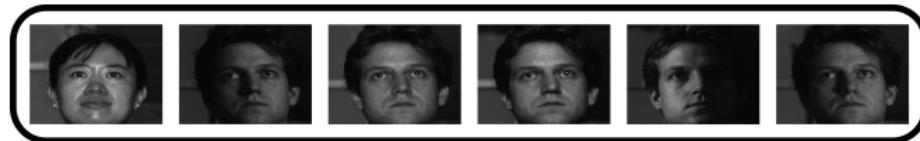
Let W_1, W_2, \dots, W_k be a collection of vector spaces.

Through an optimization problem, one can define a "Flag of best fit" to these vector spaces.

The next example illustrates how to use the flag of best fit to detect a potentially weak (but common) signal in a collection of vector spaces.

An example using the Flag Manifold: the Flag Mean

Goal: Recover an image corresponding to a subject of interest, who has been “weakly” added to a set of illumination subspaces.



Some background in digital imagery
Grassmann varieties
Experiments on CMU-PIE
A Flag Example





Further Problems:

- 1) Given a point $p \in Gr(k, n)$, you can minimize the distance from p to a Schubert variety. Given a collection of points on a Grassmannian, you can find a Schubert variety of best fit.
- 2) Find a collection of points on $Gr(k, n)$ that maximize their minimum distance apart or are an extremum of some other objective function.
- 3) Given a pair of random points on $Gr(k, n)$ find their expected distance apart.

The goal of these problems are automated attribute discovery, subspace packing (coding theory), detection of structure or signal.