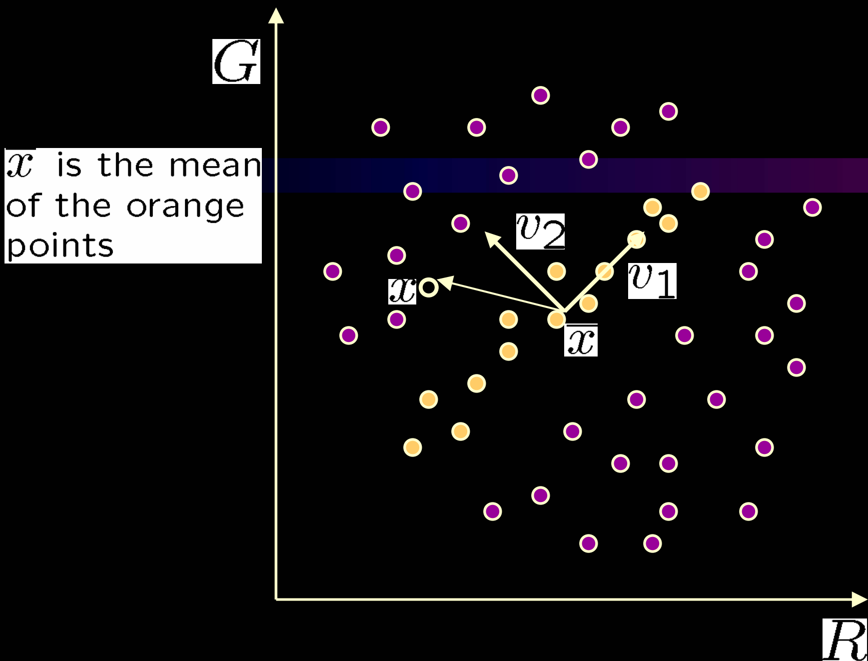
# *Recognition by Appearance*

- Appearance-based recognition is a competing paradigm to features and alignment.
- No features are extracted!
- Images are represented by basis functions (eigenvectors) and their coefficients.
- Matching is performed on this compressed image representation.

# *Eigenvectors and Eigenvalues*




Consider the sum squared distance of a point **x** to all of the orange points:

$$SSD(\mathbf{v}) = \sum_{\text{orange point } \mathbf{x}} \|(\mathbf{x} - \bar{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\|$$

What unit vector **v** minimizes SSD?

$$\mathbf{v}_2 = min_{\mathbf{v}} \; \{SSD(\mathbf{v})\}$$

What unit vector **v** maximizes SSD?

$$\mathbf{v}_1 = max_{\mathbf{v}} \; \{SSD(\mathbf{v})\}$$

$$\begin{aligned} SSD(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x} - \bar{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\| \\ &= \sum_{\mathbf{x}} \mathbf{v}^{\mathbf{T}}(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathbf{T}}\mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}} \left[ \sum_{\mathbf{x}} (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathbf{T}} \right] \mathbf{v} \\ &= \mathbf{v}^{\mathbf{T}}\mathbf{A}\mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}} (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^{\mathbf{T}} \end{aligned}$$

Solution: $\mathbf{v_1}$ is eigenvector of **A** with *largest* eigenvalue
$\mathbf{v_2}$ is eigenvector of **A** with *smallest* eigenvalue

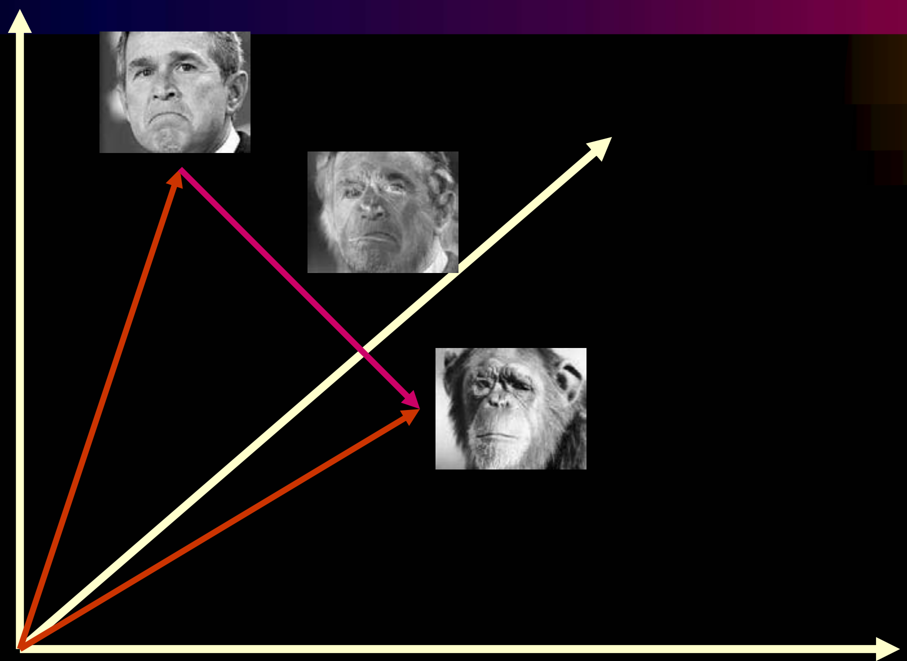
# *Principle component analysis*

- Suppose each data point is N-dimensional
  - Same procedure applies:

$$
\begin{aligned}
SSD(\mathbf{v}) &= \sum_{\mathbf{x}} \|(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}} \cdot \mathbf{v}\| \\
&= \mathbf{v}^{\mathbf{T}}\mathbf{A}\mathbf{v} \quad \text{where } \mathbf{A} = \sum_{\mathbf{x}} (\mathbf{x} - \overline{\mathbf{x}})(\mathbf{x} - \overline{\mathbf{x}})^{\mathbf{T}}
\end{aligned}
$$

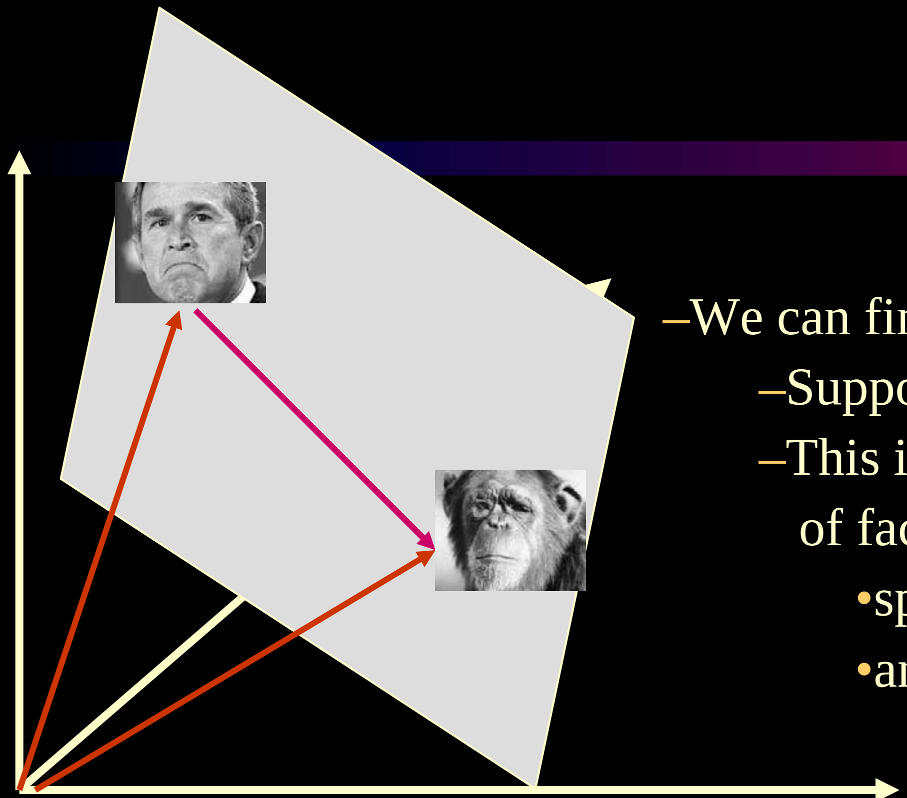
  - The eigenvectors of **A** define a new coordinate system
    - eigenvector with largest eigenvalue captures the most variation among training vectors **x**
    - eigenvector with smallest eigenvalue has least variation
  - We can compress the data by only using the top few eigenvectors

# *The space of faces*



- An image is a point in a high-dimensional space
  - An N x M image is a point in $R^{NM}$
  - We can define vectors in this space

# *Dimensionality reduction*



- –We can find the best subspace using PCA
  - –Suppose it is K dimensional
  - –This is like fitting a "hyper-plane" to the set of faces
    - •spanned by vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_K}$
    - •any face $\mathbf{x} \approx a_1\mathbf{v_1} + a_2\mathbf{v_2} + , \ldots, + a_K\mathbf{v_K}$

The set of faces is a "subspace" of the set of images.

# *Turk and Pentland's Eigenfaces: Training*

- Let F1, F2,…, FM be a set of training face images.
  Let F be their mean and Φi = Fi – F
- Use principal components to compute the eigenvectors
  and eigenvalues of the covariance matrix of the Φi s
- Choose the vector u of most significant M eigenvectors
  to use as the basis.
- Each face is represented as a linear combination of eigenfaces

u = (u1, u2, u3, u4, u5);  F27 = a1*u1 + a2*u2 + … + a5*u5

# *Matching*

unknown face image I → convert to its eigenface representation → $\Omega = (\Omega_1, \Omega_2, \ldots, \Omega_m)$

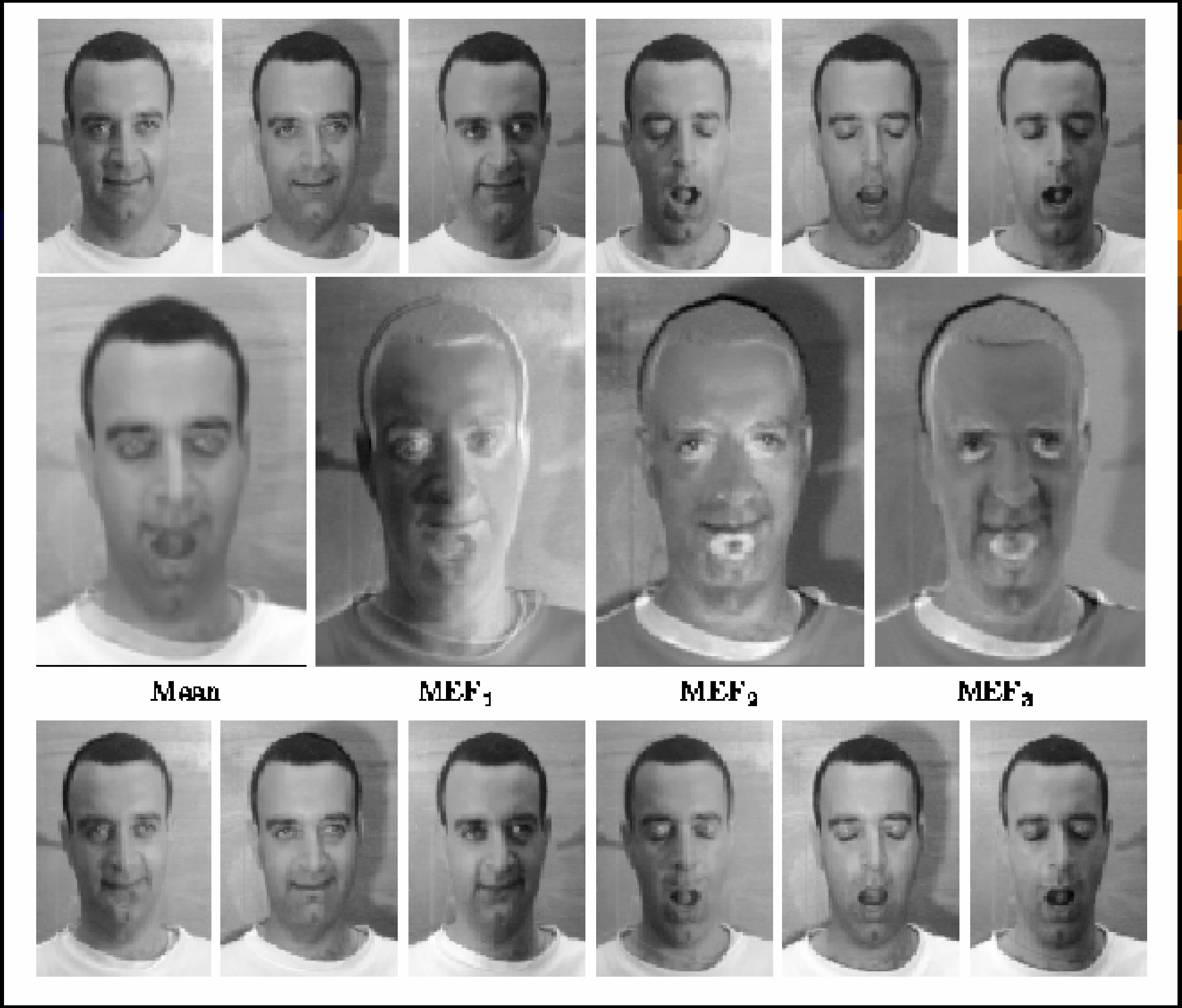
Find the face class k that minimizes

$$\varepsilon_k = \| \Omega - \Omega_k \|$$

training images

mean image

linear approxi-mations




3 eigen-images

# Extension to 3D Objects

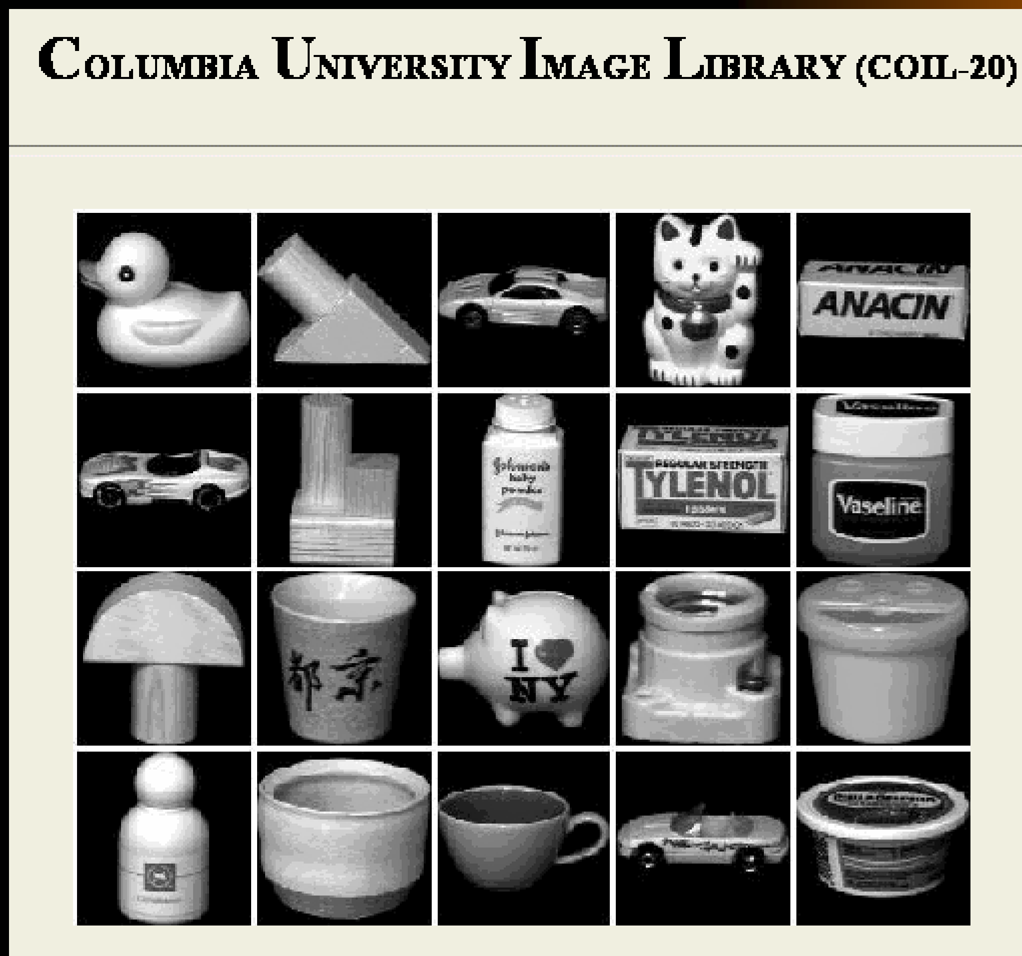
- Murase and Nayar (1994, 1995) extended this idea to 3D objects.

- The training set had multiple views of each object, on a dark background.

- The views included multiple (discrete) rotations of the object on a turntable and also multiple (discrete) illuminations.

- The system could be used first to identify the object and then to determine its (approximate) pose and illumination.

# *Sample Objects*
# *Columbia Object Recognition Database*



# *Significance of this work*

- The extension to 3D objects was an important contribution.

- Instead of using brute force search, the authors observed that

  All the views of a single object, when transformed into the eigenvector space became points on a manifold in that space.

- Using this, they developed fast algorithms to find the closest object manifold to an unknown input image.

- Recognition with pose finding took less than a second.

# Appearance-Based Recognition

- Training images must be representative of the instances of objects to be recognized.
- The object must be well-framed.
- Positions and sizes must be controlled.
- Dimensionality reduction is needed.
- It is not powerful enough to handle general scenes without prior segmentation into relevant objects.

* • Newer systems are using interest operators to identify "parts" and learning objects with these parts.