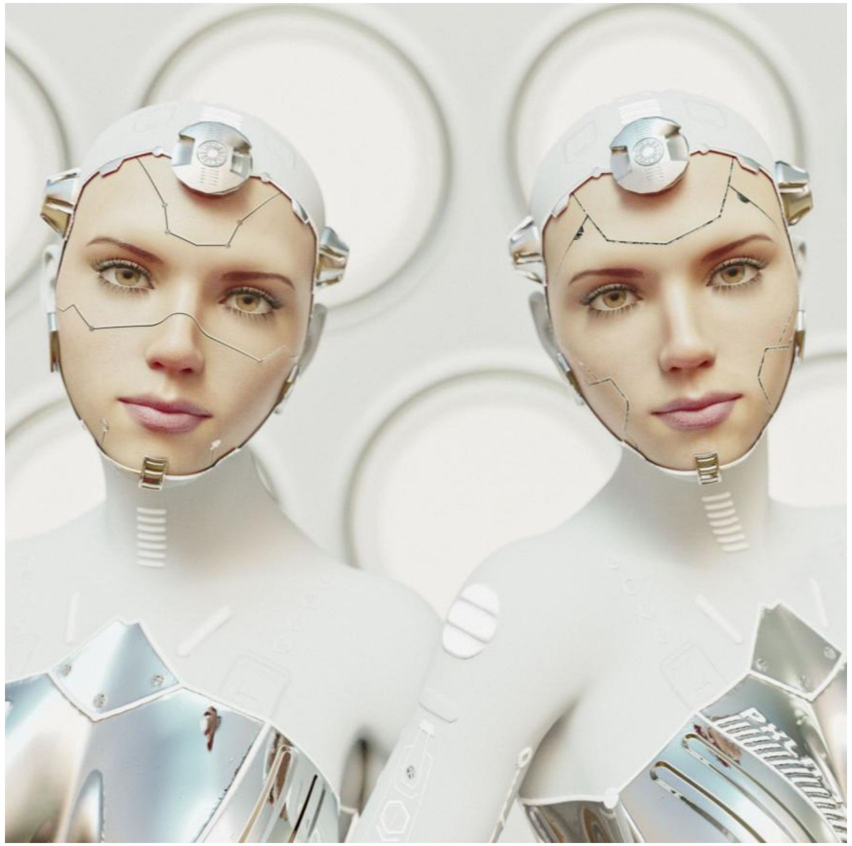
APPEARANCE MATCHING



Appearance matching

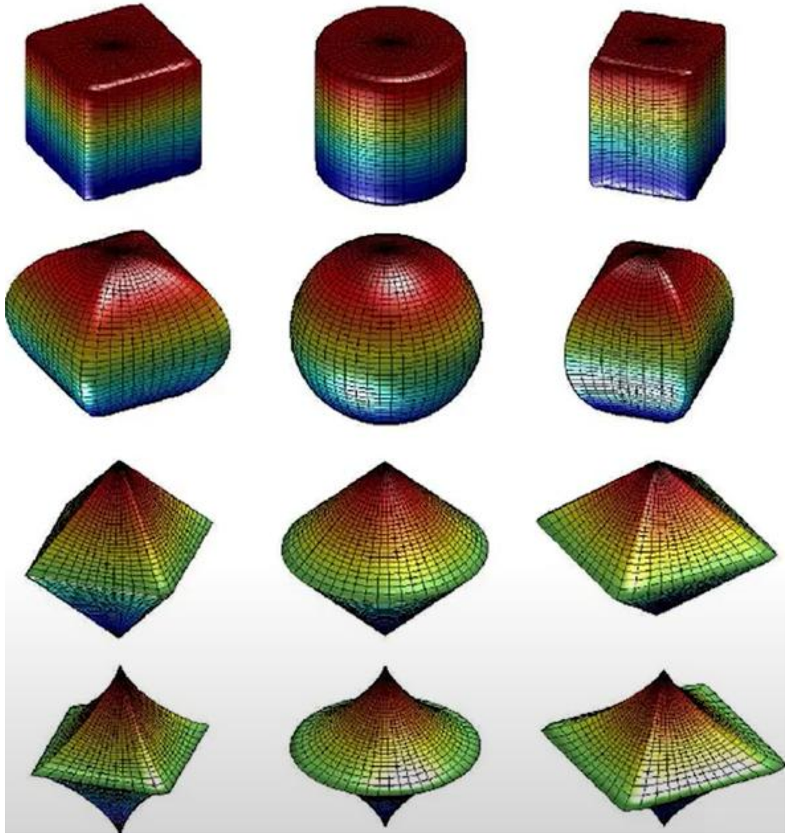
Representation and recognition of 3D object using visual appearance

1. Shape vs Appearance
2. Learning Appearance
3. Principle Component Analysis
4. Parametric Appearance Representation
5. Appearance Matching

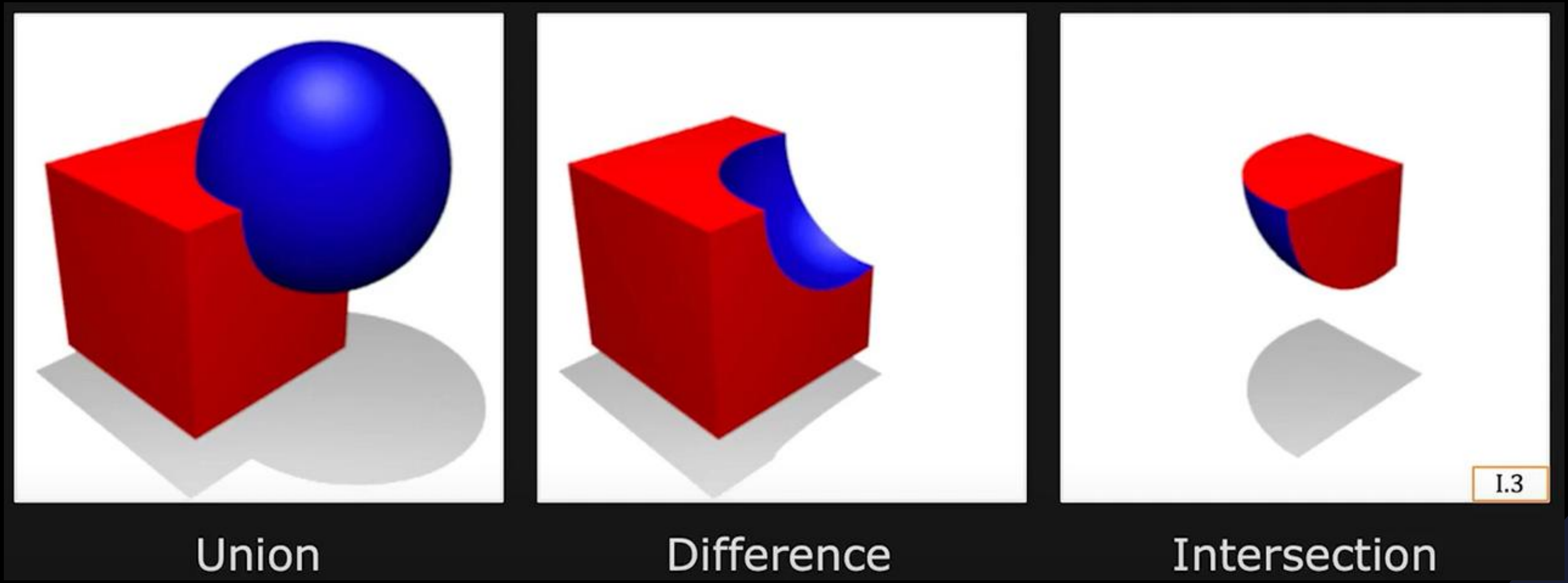
VOXEL REPRESENTATION

ANALYTICAL REPRESENTATION

$$|x|^r + |y|^s + |z|^t = 1$$



Constructive Solid Geometry (CSG)



Issues with 3D Shape Matching

- Requires measuring of 3D shapes
  1. *Creating large database*
  2. *Recognition*

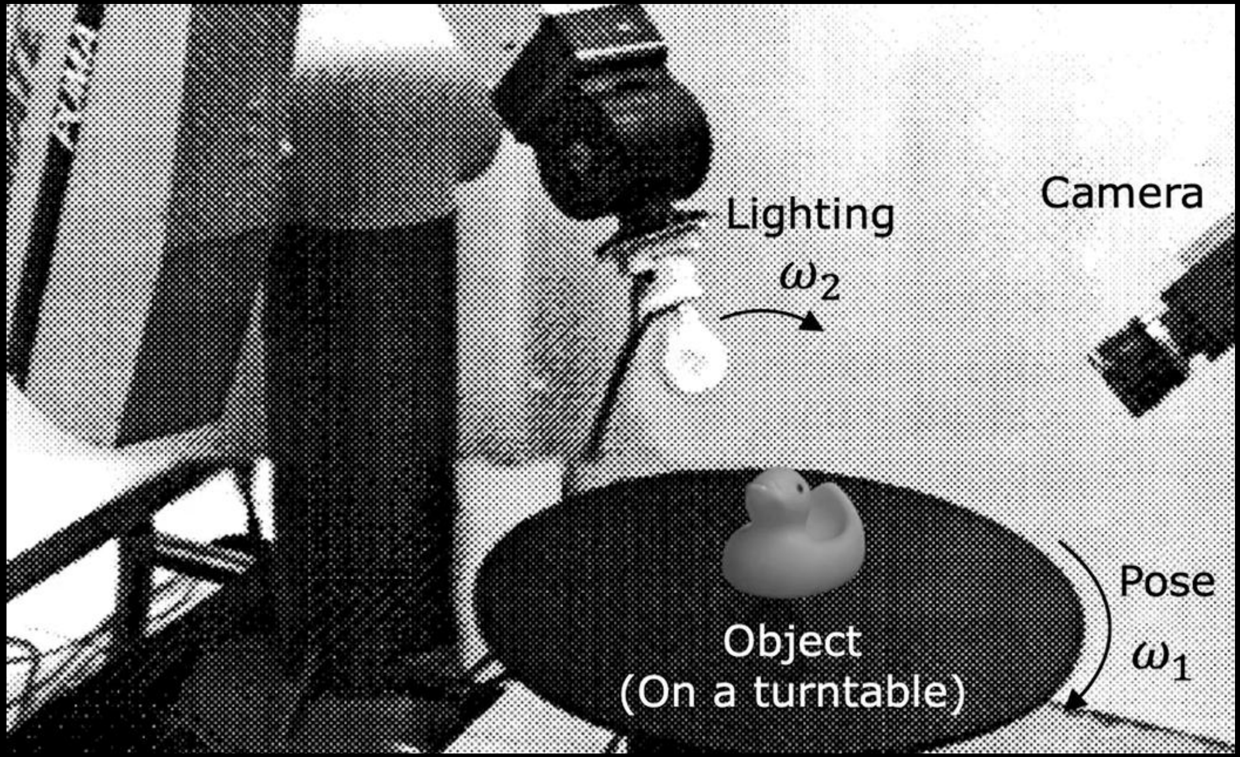
- Computationally expansive

Visual Appearance

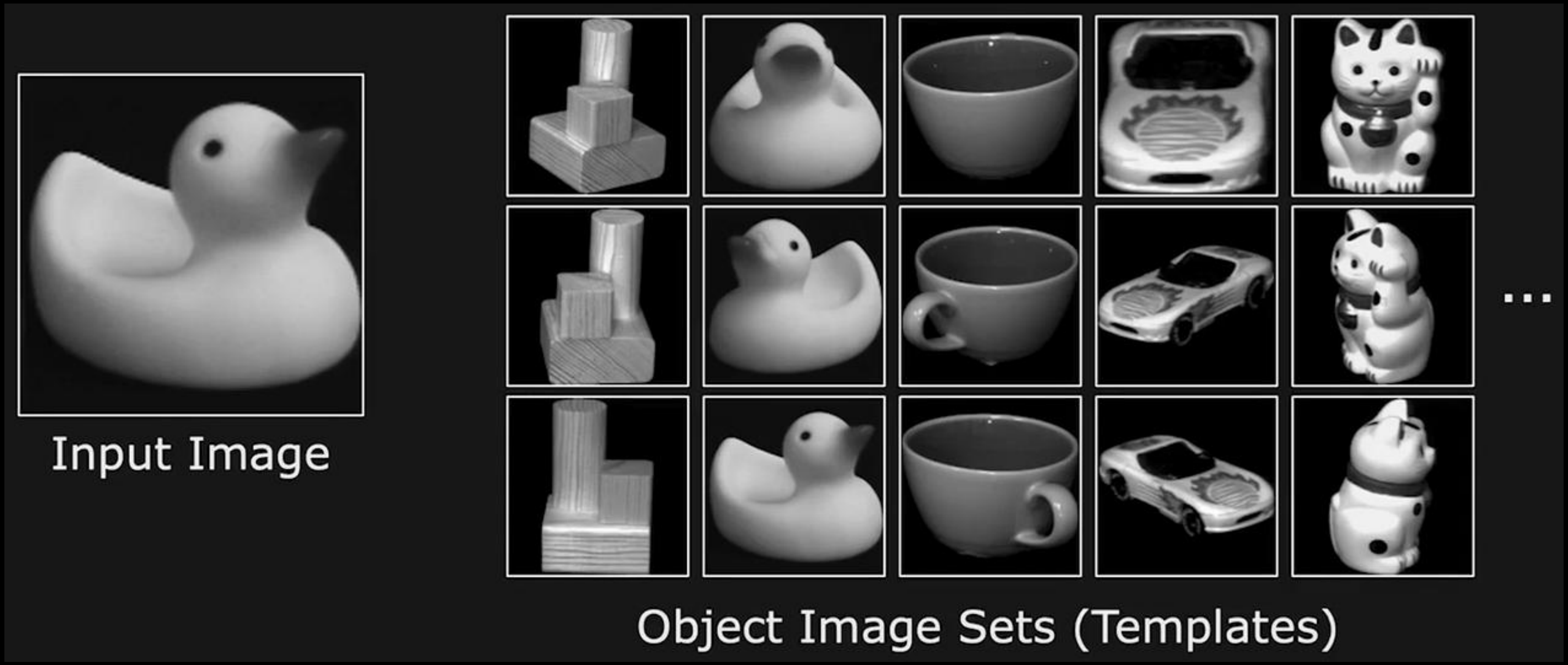
$$\text{Visual Appearance} = \mathcal{F}\left\{\begin{array}{ll}\left.\begin{array}{l}\text{Shape}\\ \text{Reflectance}\end{array}\right\} & \text{Intrinsic Parameters}\\[1em] \left.\begin{array}{l}\text{Pose}\\ \text{Illumination}\end{array}\right\} & \text{Extrinsic Variables}\end{array}\right.$$

Appearance learning by humans

Appearance learning by machines



Template matching (naïve method)



Template matching (naïve method)

- Sum of Absolute difference (SAD)
- Sum of Square difference (SSD)
- Normalized Correlation (NCC)

- Thousands of templates for possibly millions of objects in dataset
- Too costly and time consuming

Object image set

Highly Redundant

Reduce dimensionality

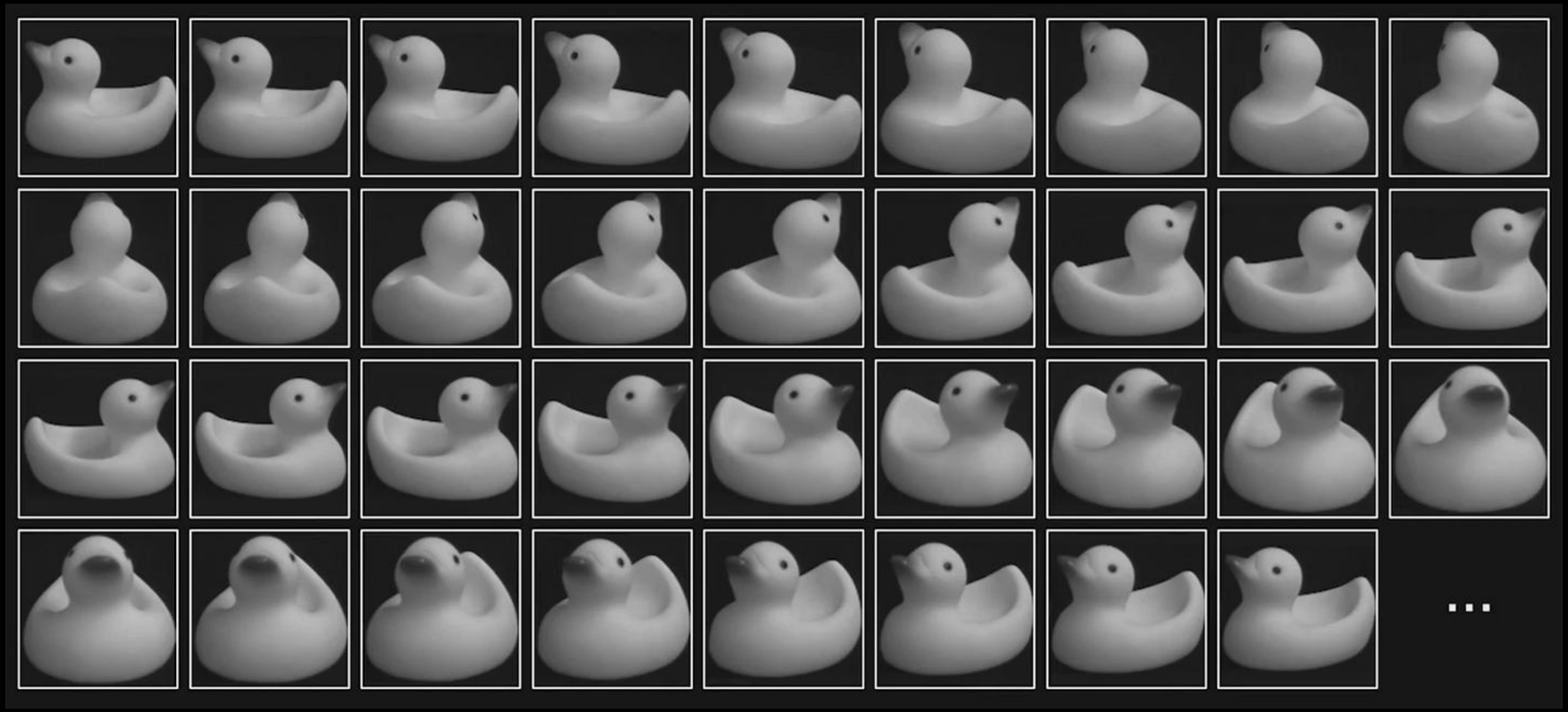


Image representation in Vector

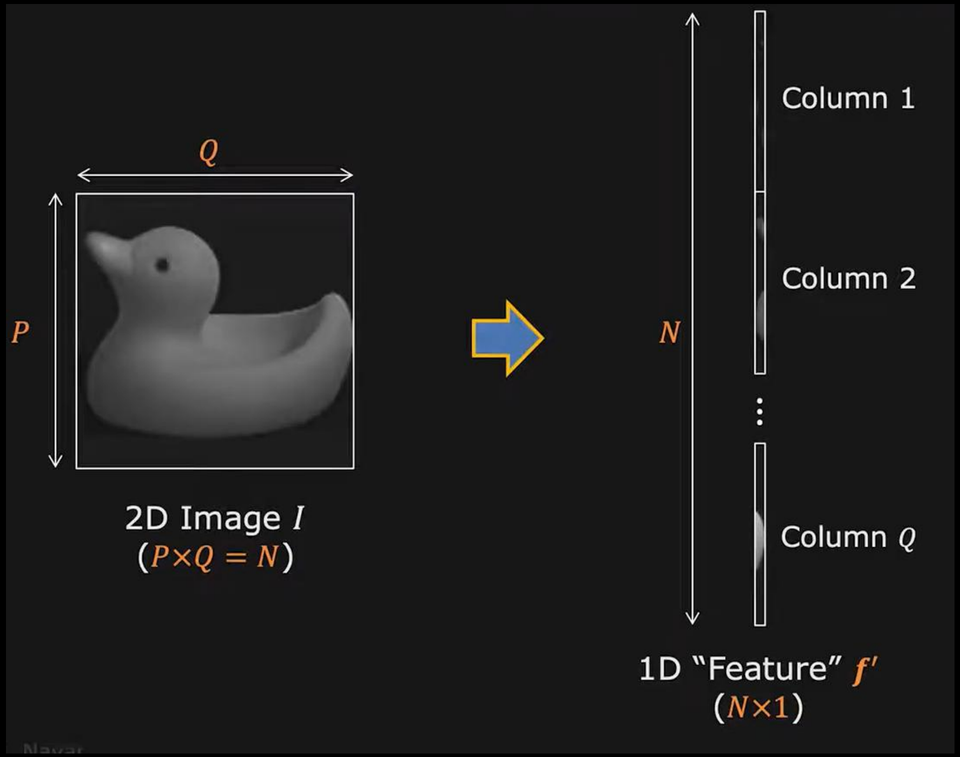
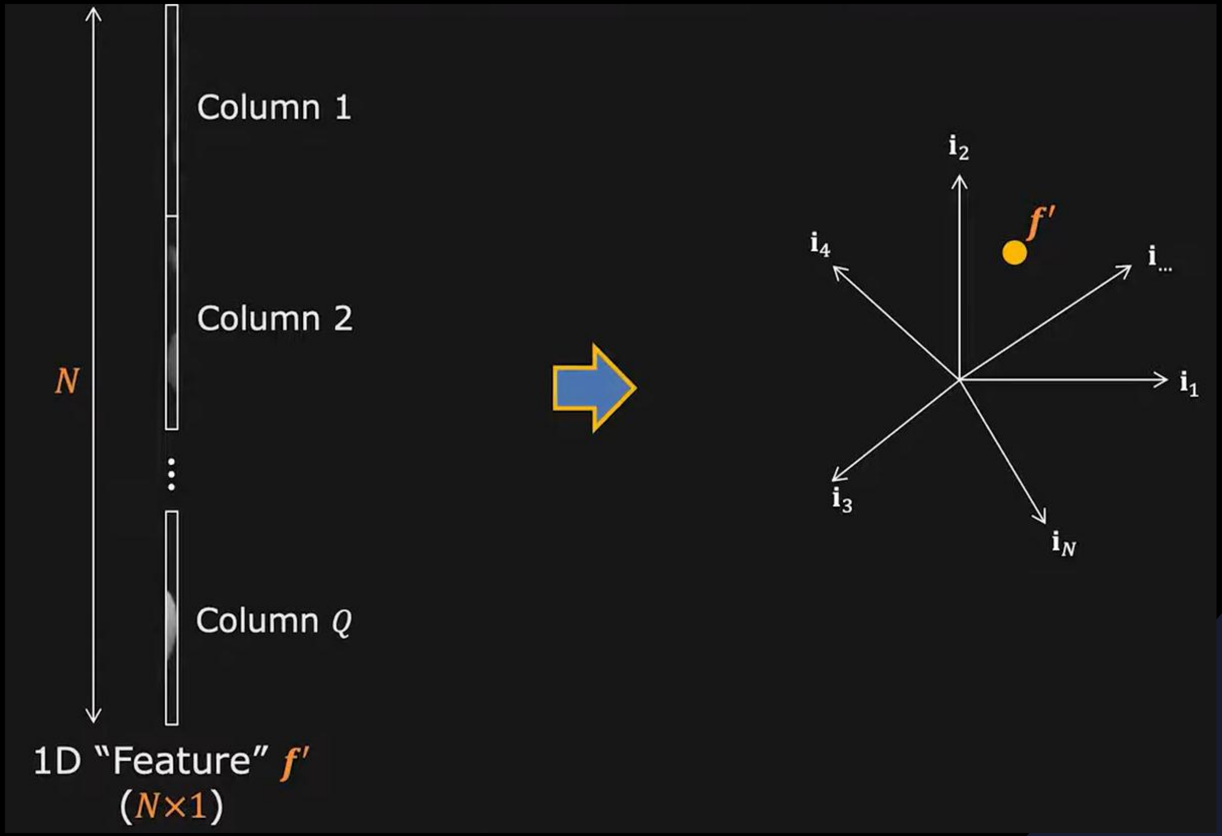


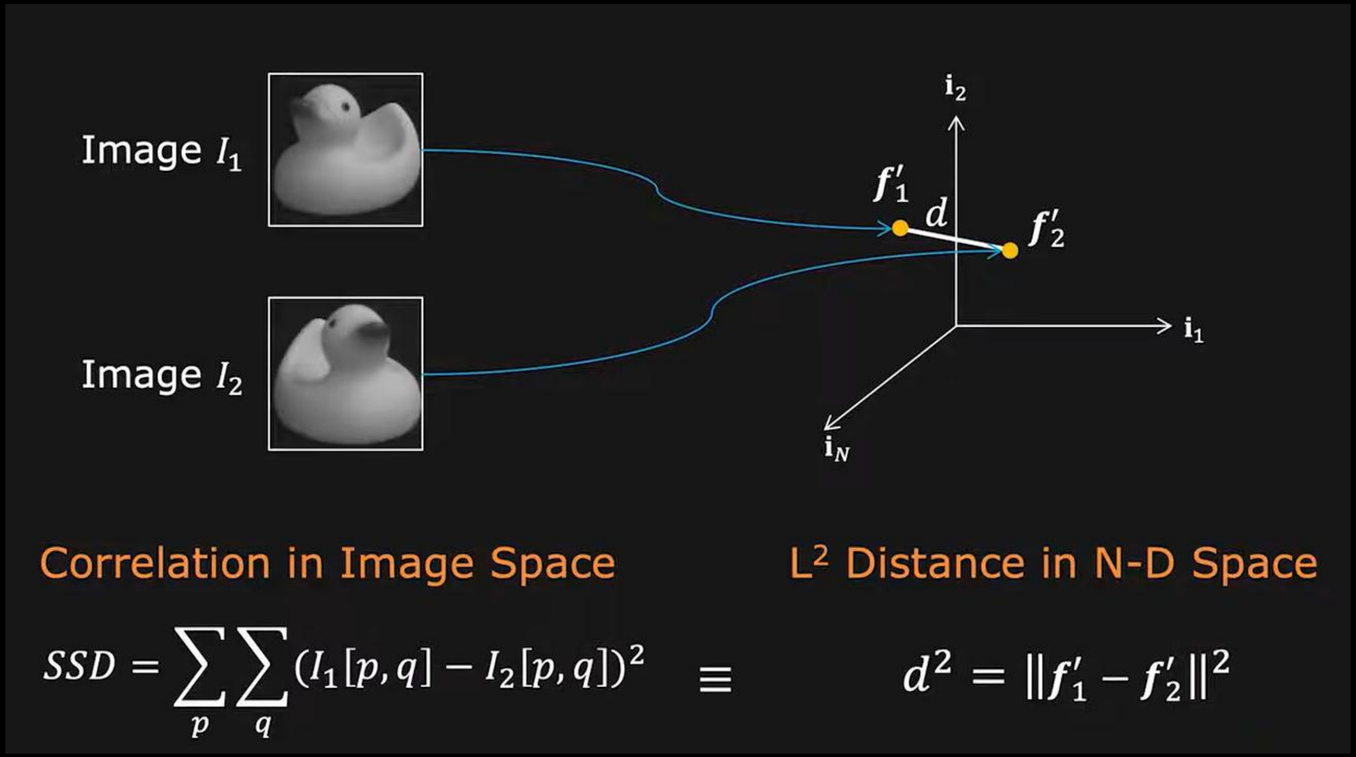
Image in N-D space

- Each dimension represents image intensity at corresponding pixel
- $i_1, i_2, \dots, i_N$ are arranged in Orthonormal basis

$$i_1 = \{1\ 0\ 0\ \dots\ 0\}^T$$



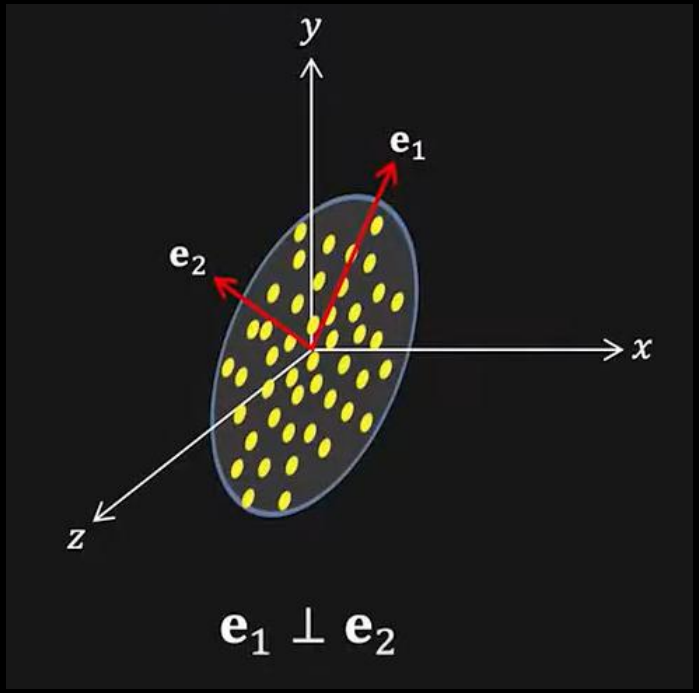
Template Matching (N-D space)



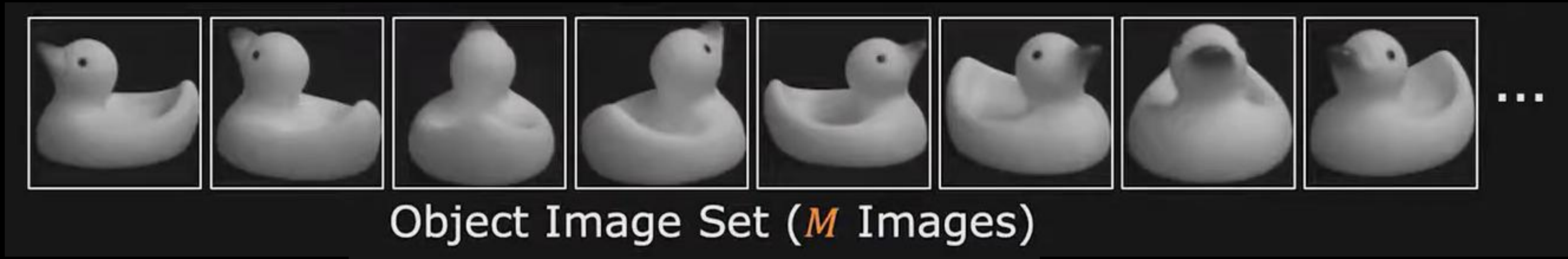
Dimensionality Reduction

- Distribution of points in 3D space, actually lie on a 2D plane.

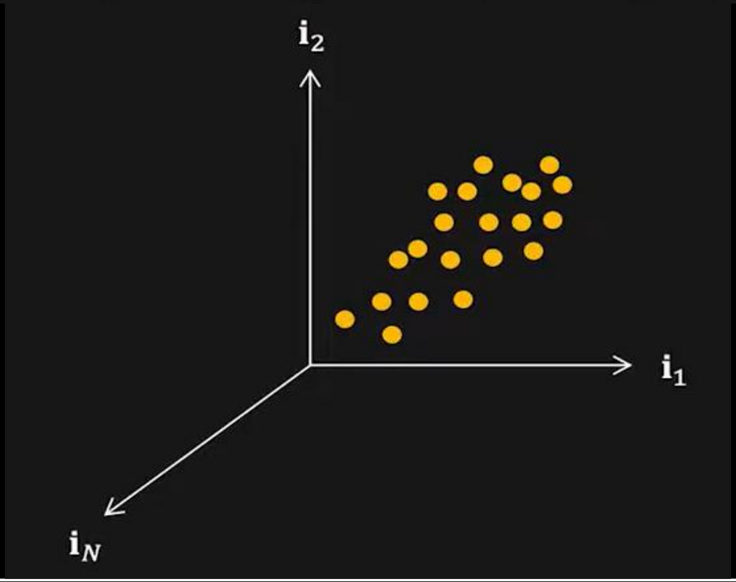
- $\{e_1, e_2\}$ can be utilized (2D) to represent every point rather using 3D representation.



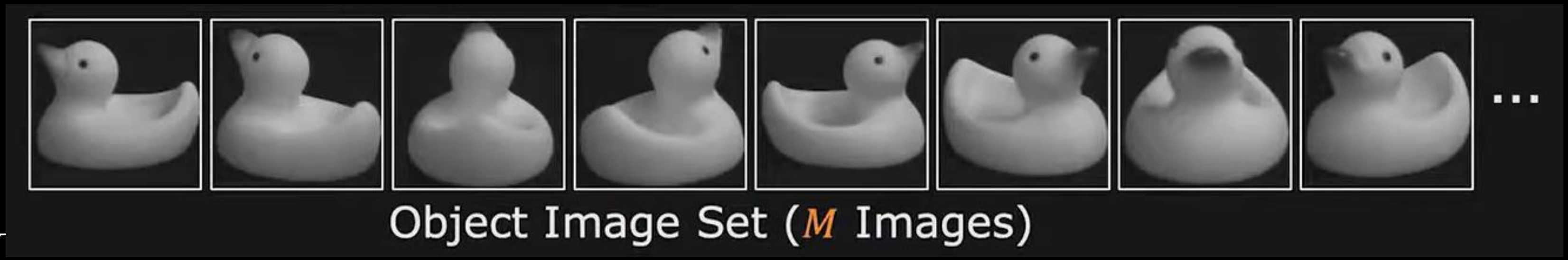
Appearance Distribution

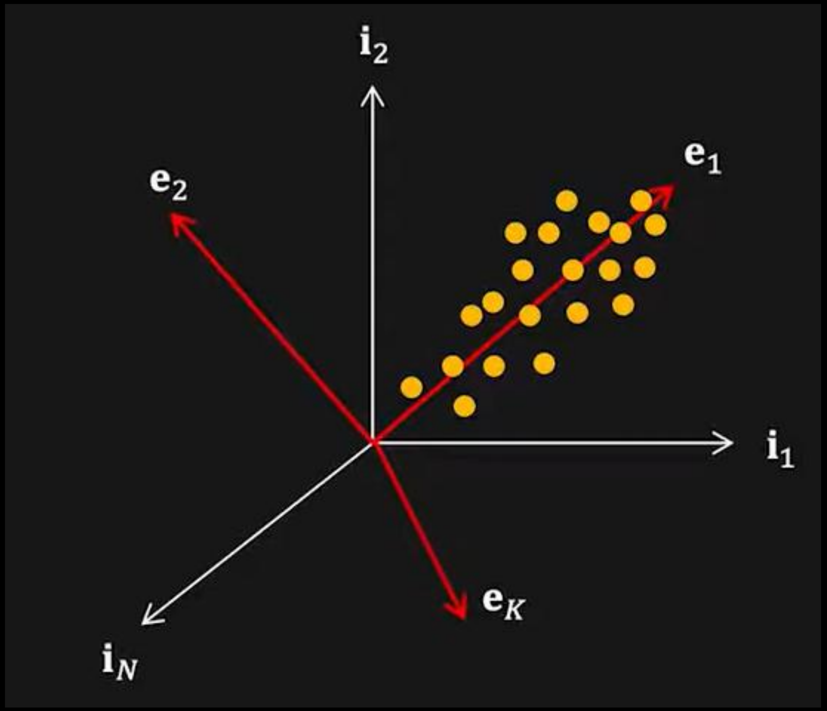


- Distribution is highly structured because of high correlation between every image.
- Can be represented using lower number of dimensions



Dimensionality Reduction





- $\{e_1, e_2, \ldots, e_K\}$ is Orthonormal basis where $K < N$

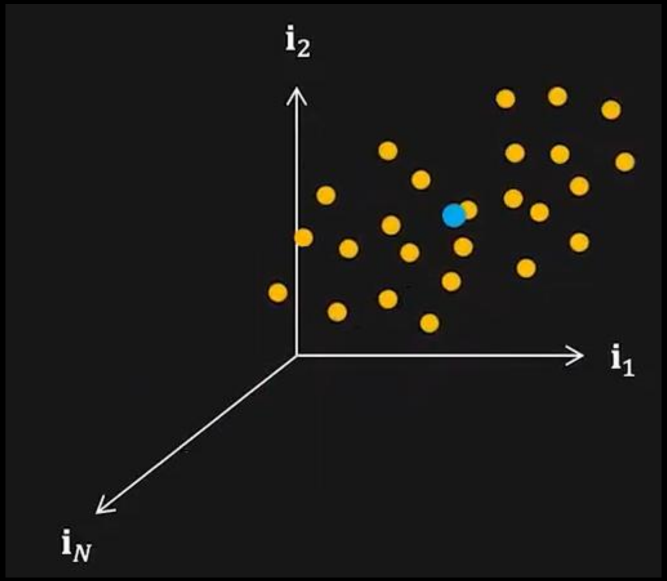
Mean Image

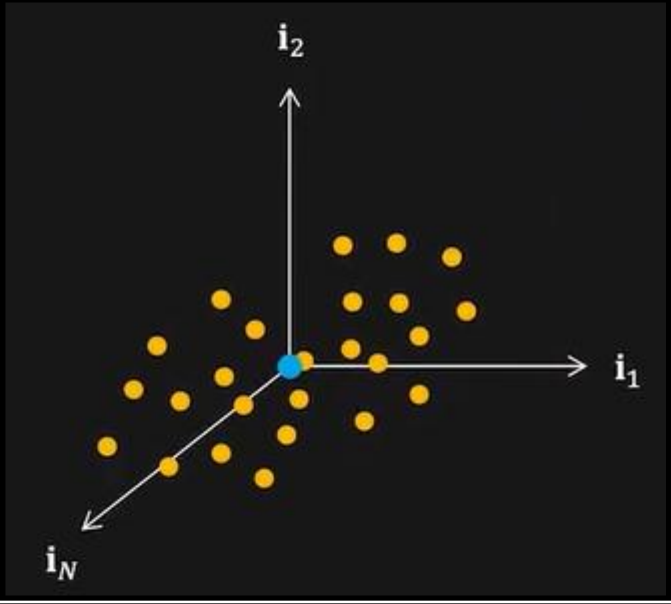
- For M images $\{f'_1, f'_2, \ldots, f'_M\}$ of an object, the mean image is:

$$\mathbf{c} = \frac{1}{M}\sum_{m=1}^{M} f'_m$$

- Subtracting mean from every image for shifting the origin to the centroid of the distribution

$$f_m = f'_m - \mathbf{c}$$



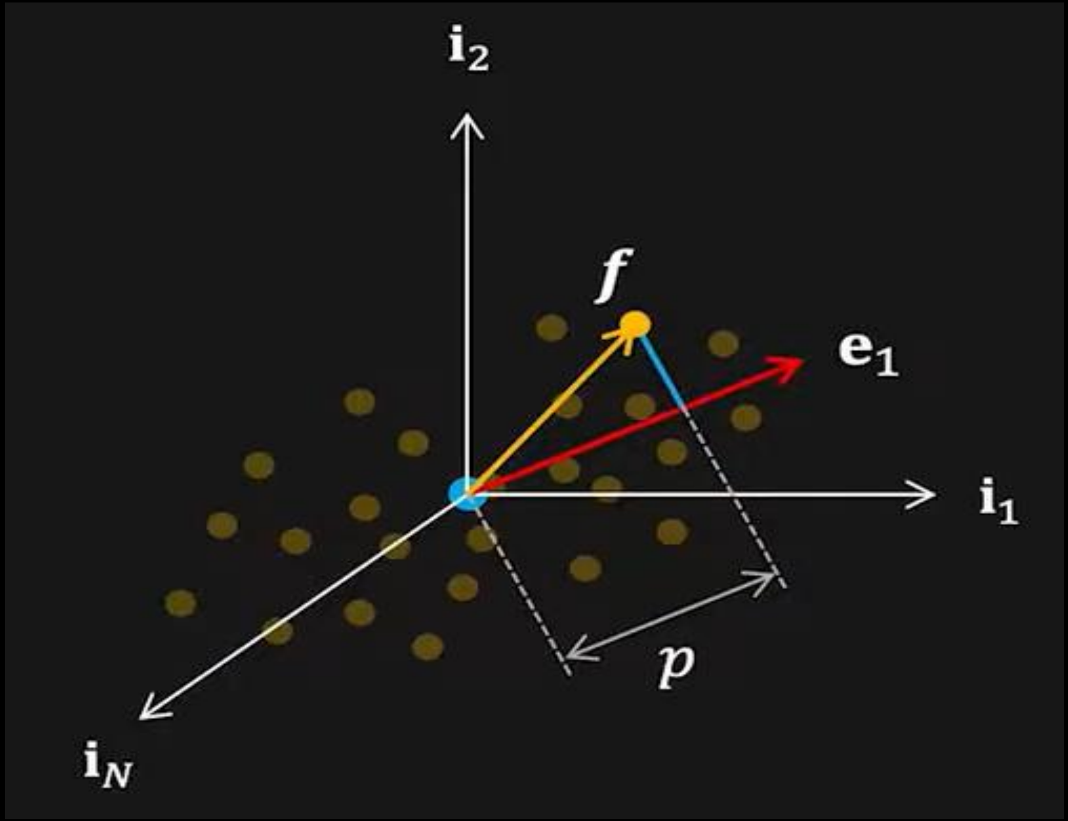


Principal Components

- 1^{st} Principal Component (e_1): Direction of maximum variance in the image set (equivalent to Least Squares Fitting of a Line)

- Image representation: Project the image f onto the principal component e_1.

- $p = e_1 * f$ (Dot Product)
- Image is represented by a single number p

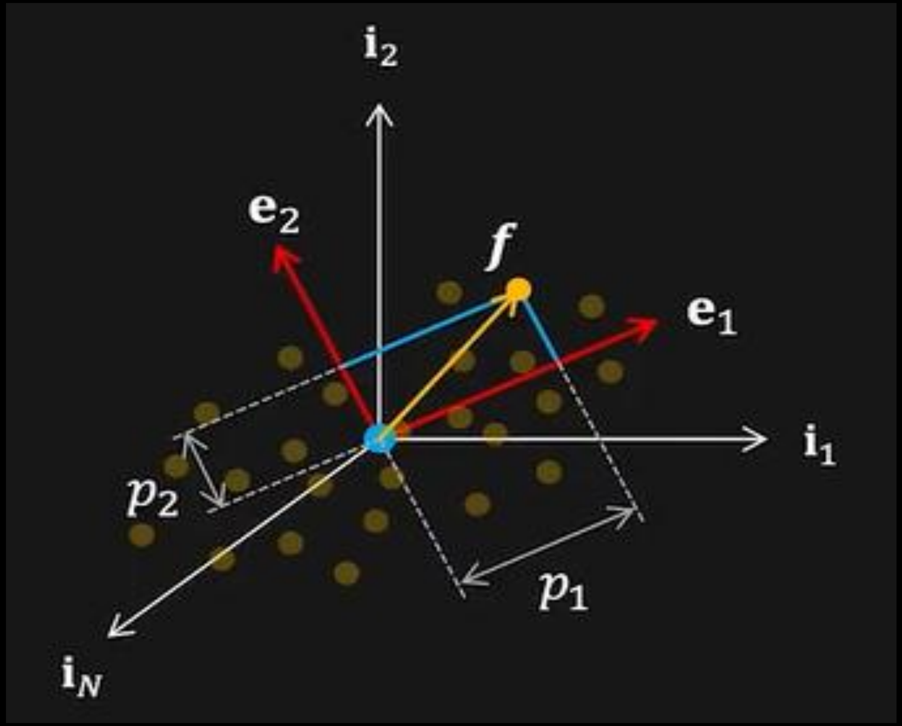


Principal Components

- 2^{nd} Principal Component (e_2): Direction of 2^{nd} maximum variance in the image set such that e_1 and e_2 are perpendicular to each other.

- Image representation: Project the image f onto the principal components $\{e_1, e_2\}$

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = [\mathbf{e}_1 \quad \mathbf{e}_2]^T \boldsymbol{f}$$

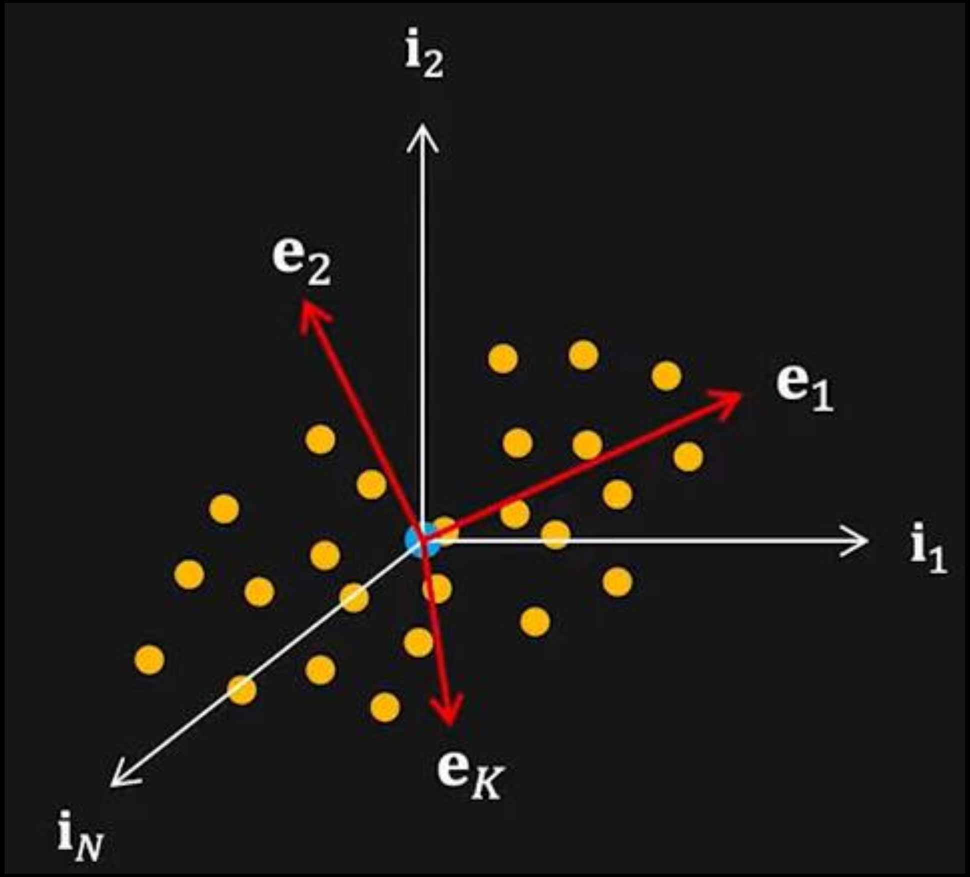


Principal Components

- 2^{nd} Principal Component (e_2): Direction of K^{th} maximum variance in the image set such that e_1 , e_2 , ... , e_K are perpendicular to each other.

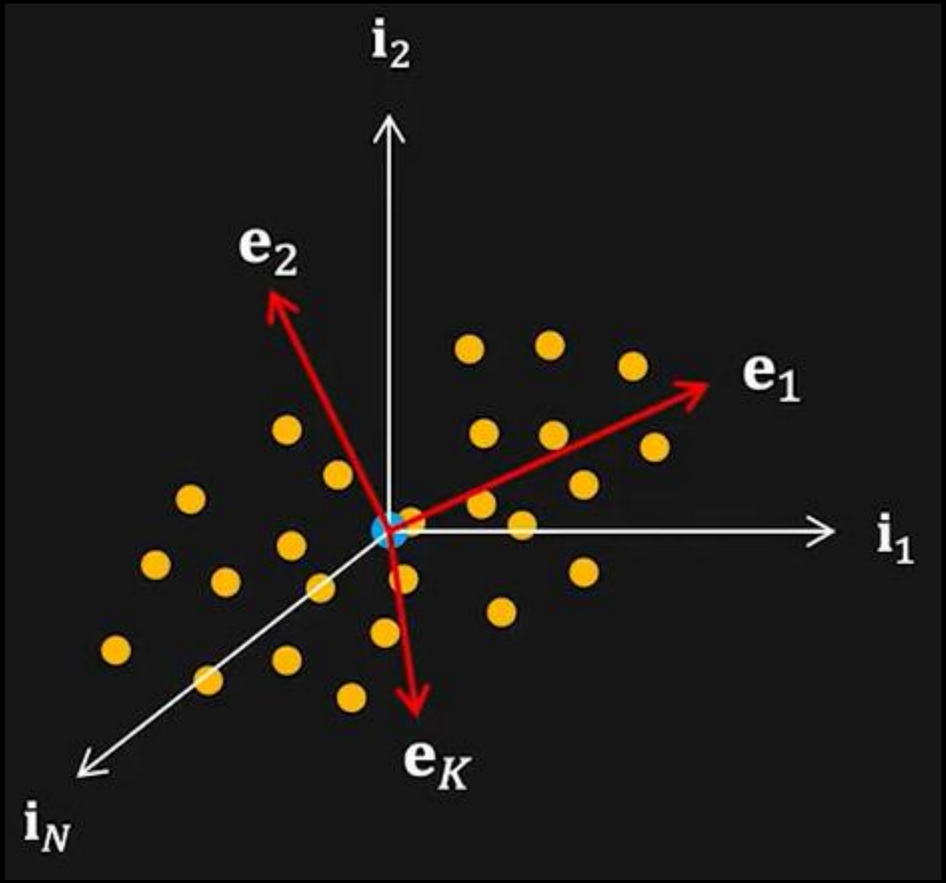
- Image representation: Project the image f onto the principal components $\{e_1 , e_2 , ... , e_K\}$

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{bmatrix} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_K]^T \boldsymbol{f}$$



Forward Projection

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{bmatrix} = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_K]^T \boldsymbol{f}$$



Back Projection

$$f \approx p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + \cdots + p_K\mathbf{e}_K$$

$$f \approx \sum_{k=1}^{K} p_k\mathbf{e}_k$$

- $\{e_1, e_2, \ldots, e_K\}$ are referred as Linear Subspace