This is a shortened version of the tutorial given at the ECCV'2002, Copenhagen, and ICPR'2002, Quebec City.

Subspace Methods for Visual Learning and Recognition

Aleš Leonardis

Faculty of Computer and Information Science
University of Ljubljana
Slovenia

Ales.Leonardis@fri.uni-lj.si

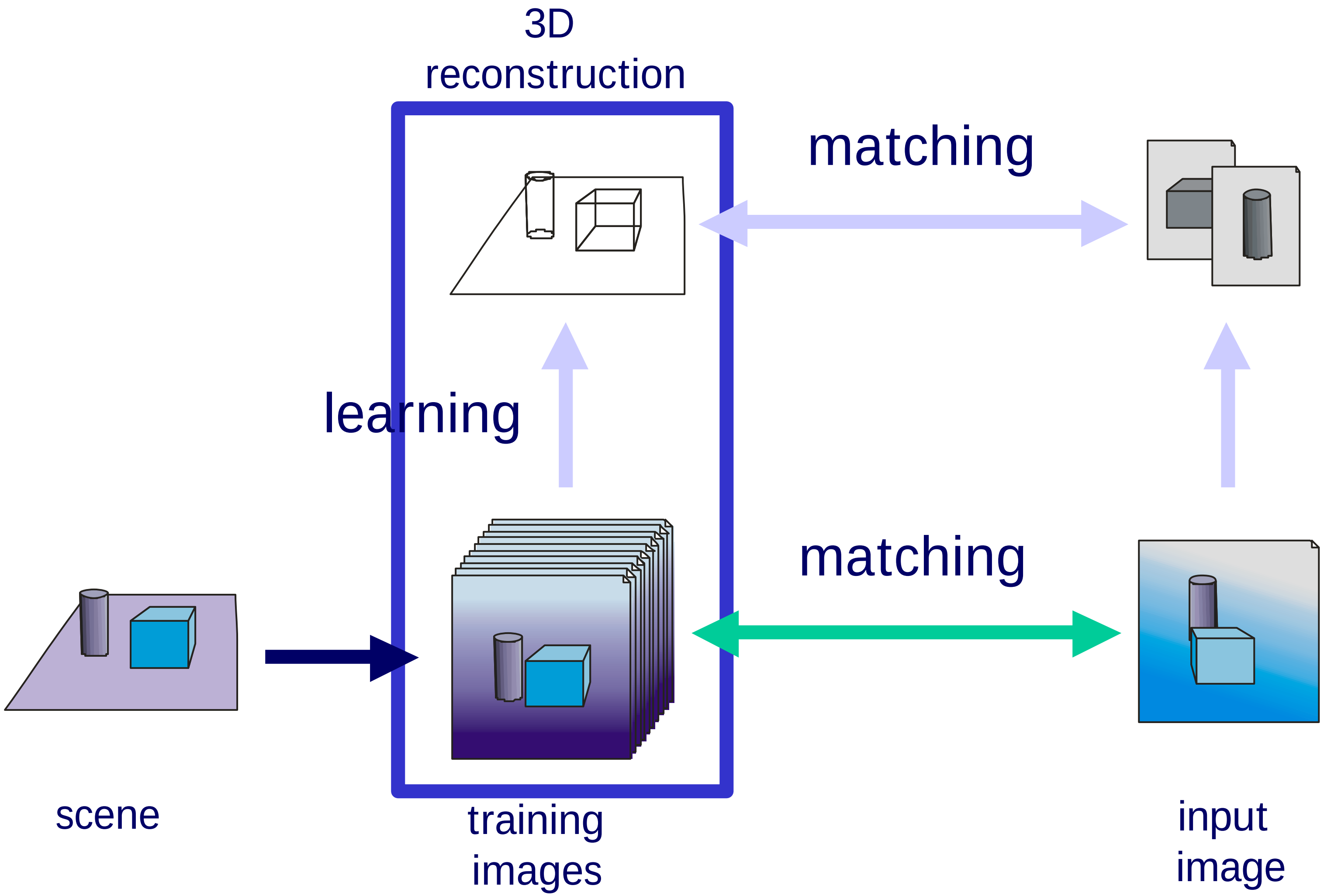
Outline Part 1

- Motivation
- Appearance based learning and recognition
- Subspace methods for visual object recognition
- Principal Components Analysis (PCA)
- Linear Discriminant Analysis (LDA)
- Canonical Correlation Analysis (CCA)
- Independent Component Analysis (ICA)
- Non-negative Matrix Factorization (NMF)
- Kernel methods for non-linear subspaces

Outline Part 2

- Robot localization
- Robust representations and recognition
- Robust PCA recognition
- Scale invariant recognition using PCA
- Illumination insensitive recognition
- Representations for panoramic images
- Incremental building of eigenspaces
- Multiple eigenspaces for efficient representation
- Robust building of eigenspaces
- Research issues

Learning and recognition



Appearance-based approaches

Attention in the appearance-based approaches

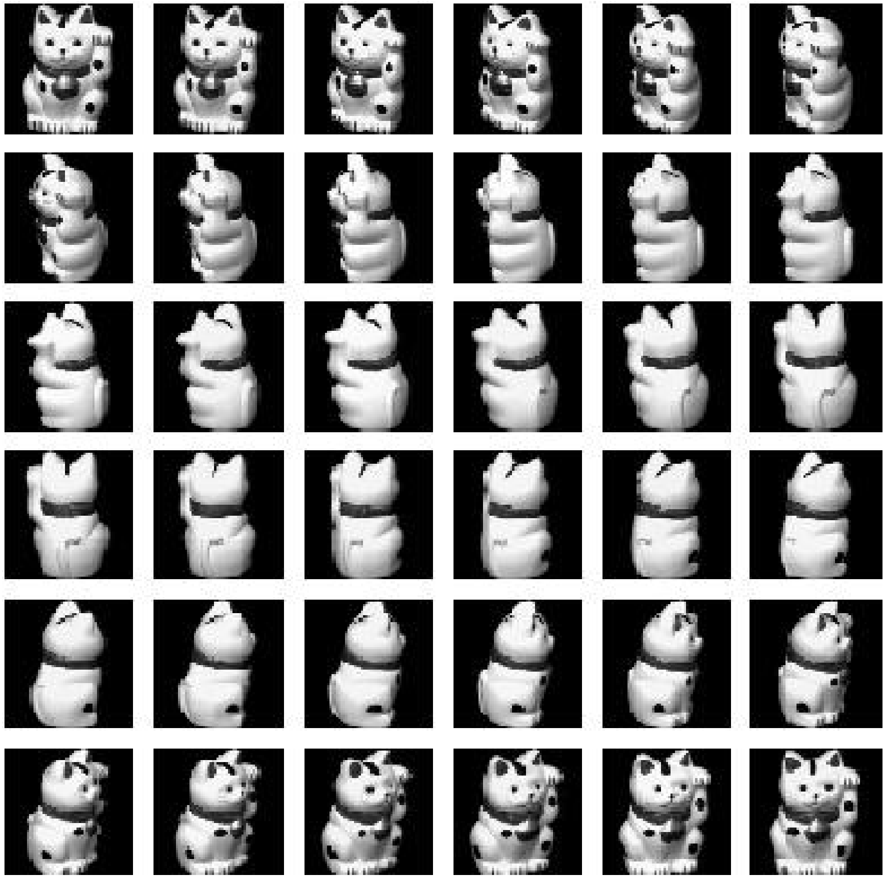
Encompass combined effects of:

- shape,
- reflectance properties,
- pose in the scene,
- illumination conditions.

Acquired through an automatic learning phase.

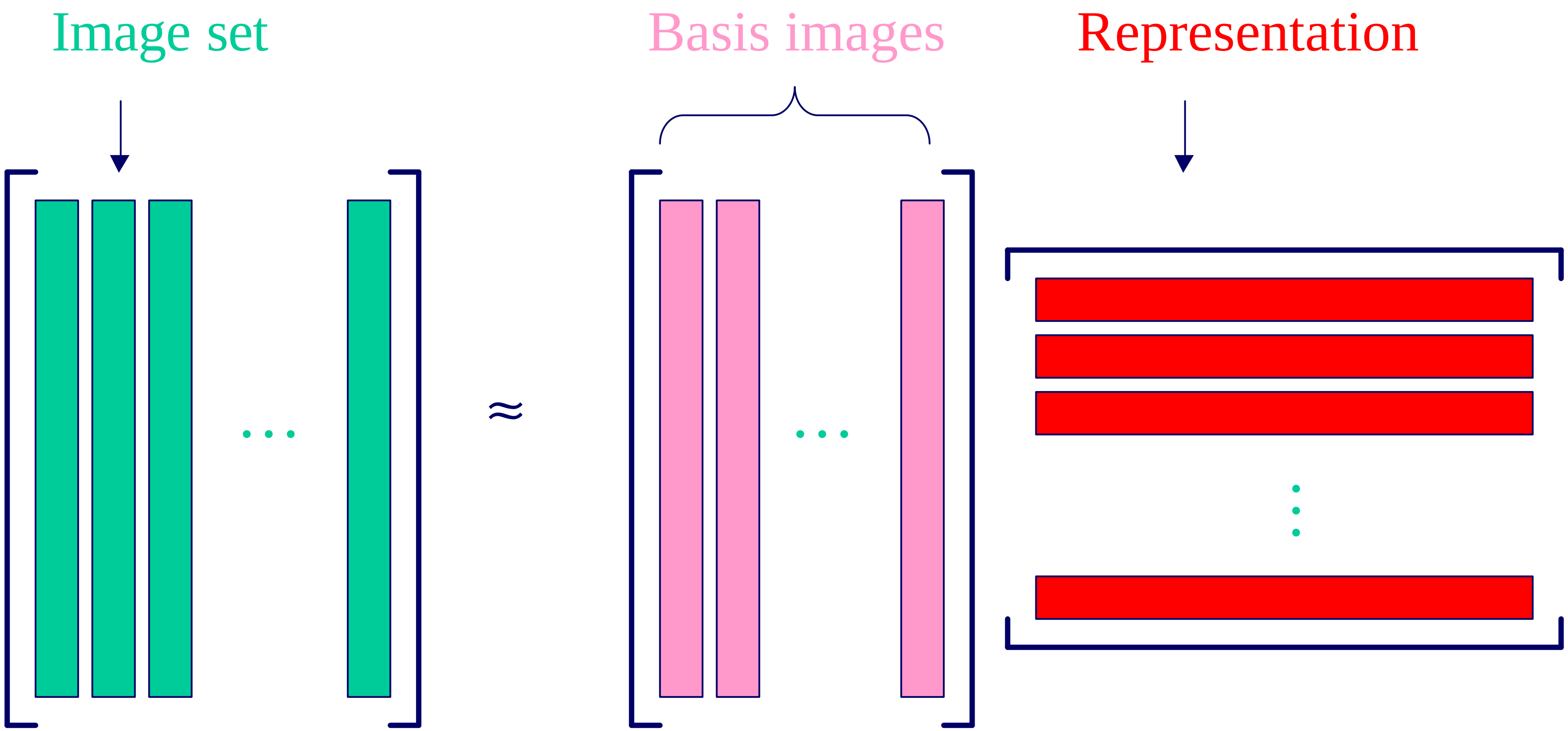
Appearance-based approaches

Objects are represented by a large number of views:



Subspace Methods

- Images are represented as points in the N-dimensional vector space
- Set of images populate only a small fraction of the space
- Characterize subspace spanned by images



Subspace Methods

Properties of the representation:

- Optimal Reconstruction $\Rightarrow$ PCA
- Optimal Separation $\Rightarrow$ LDA
- Optimal Correlation $\Rightarrow$ CCA
- Independent Factors $\Rightarrow$ ICA
- Non-negative Factors $\Rightarrow$ NMF
- Non-linear Extension $\Rightarrow$ Kernel Methods

Eigenspace representation

- **Image set (normalised, zero-mean)**

$$X = [\ \mathbf{x}_0 \quad \mathbf{x}_1 \quad \ldots \quad \mathbf{x}_{n-1}\ ];\ X \in \mathbb{R}^{m\times n}$$

- **We are looking for orthonormal basis functions:**

$$U = [\ \mathbf{u}_0 \quad \mathbf{u}_1 \quad \ldots \quad \mathbf{u}_k\ ];\ k << n$$

- **Individual image is a linear combination of basis functions**

$$\mathbf{x}_i \approx \tilde{\mathbf{x}}_i = \sum_{j=0}^{p} q_j(\mathbf{x}_i)\mathbf{u}_j$$

$$\| \mathbf{x} - \mathbf{y} \|^2 \approx \| \sum_{j=1}^{k} q_j(\mathbf{x})\mathbf{u}_j - \sum_{j=1}^{k} q_j(\mathbf{y})\mathbf{u}_j \|^2 =$$

$$\| \sum_{j=1}^{k} (q_j(\mathbf{x}) - q_j(\mathbf{y}))\mathbf{u}_j \|^2 = \| q_j(\mathbf{x}) - q_j(\mathbf{y}) \|^2$$

Best basis functions n?

- **Optimisation problem**

$$\sum_{i=0}^{n-1} ||\mathbf{x}_i - \sum_{j=0}^{k} q_j(\mathbf{x}_i)\mathbf{u}_j||^2 \rightarrow \min$$

- **Taking the *k* eigenvectors with the largest eigenvalues of**

$$C = XX^T = \left[\begin{array}{cccc} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{n-1} \end{array} \right] \left[\begin{array}{c} \mathbf{x}_0^\top \\ \mathbf{x}_1^\top \\ \dots \\ \mathbf{x}_{n-1}^\top \end{array} \right]$$

- **PCA or Karhunen-Loéve Transform (KLT)**

$$C\mathbf{u}_i = \lambda_i \mathbf{u}_i$$

Efficient eigenspace computation

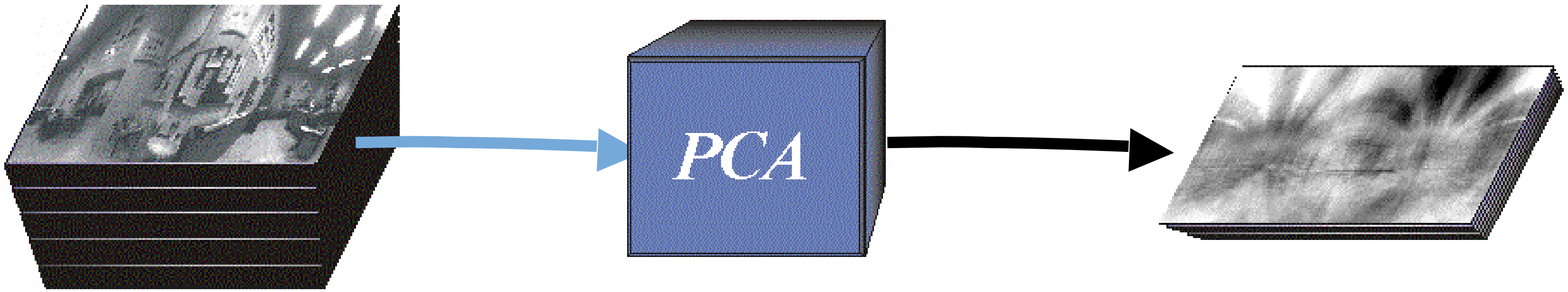
- $n << m$
- Compute the eigenvectors u'_i, $i = 0,\ldots,n-1$, of the inner product matrix

$$Q = X^\top X = \begin{bmatrix} \mathbf{x}_0^\top \\ \mathbf{x}_1^\top \\ \cdots \\ \mathbf{x}_{n-1}^\top \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \ldots & \mathbf{x}_{n-1} \end{bmatrix}; \; Q \in \mathbb{R}^{n\times n}$$

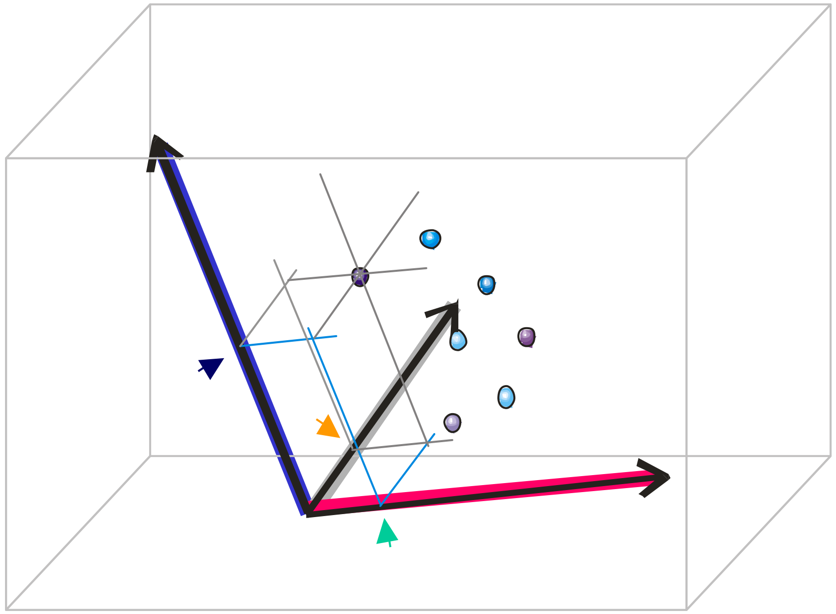
- The eigenvectors of $XX^\top$ can be obtained by using $XX^\top X\nu'_i = \lambda'_i X\nu'_i$:

$$\mathbf{u}_i = \frac{1}{\sqrt{\lambda'_i}} X\mathbf{u}'_i$$

Principal Component Analysis



Principal Component Analysis



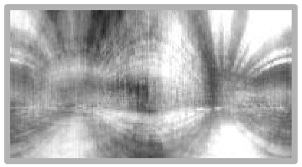
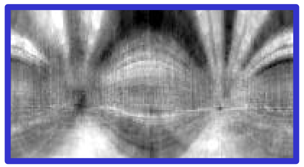
 $= q_1 \cdot$  $+ q_2 \cdot$  $+ q_3 \cdot$  $+ \ldots$

Image representation with PCA

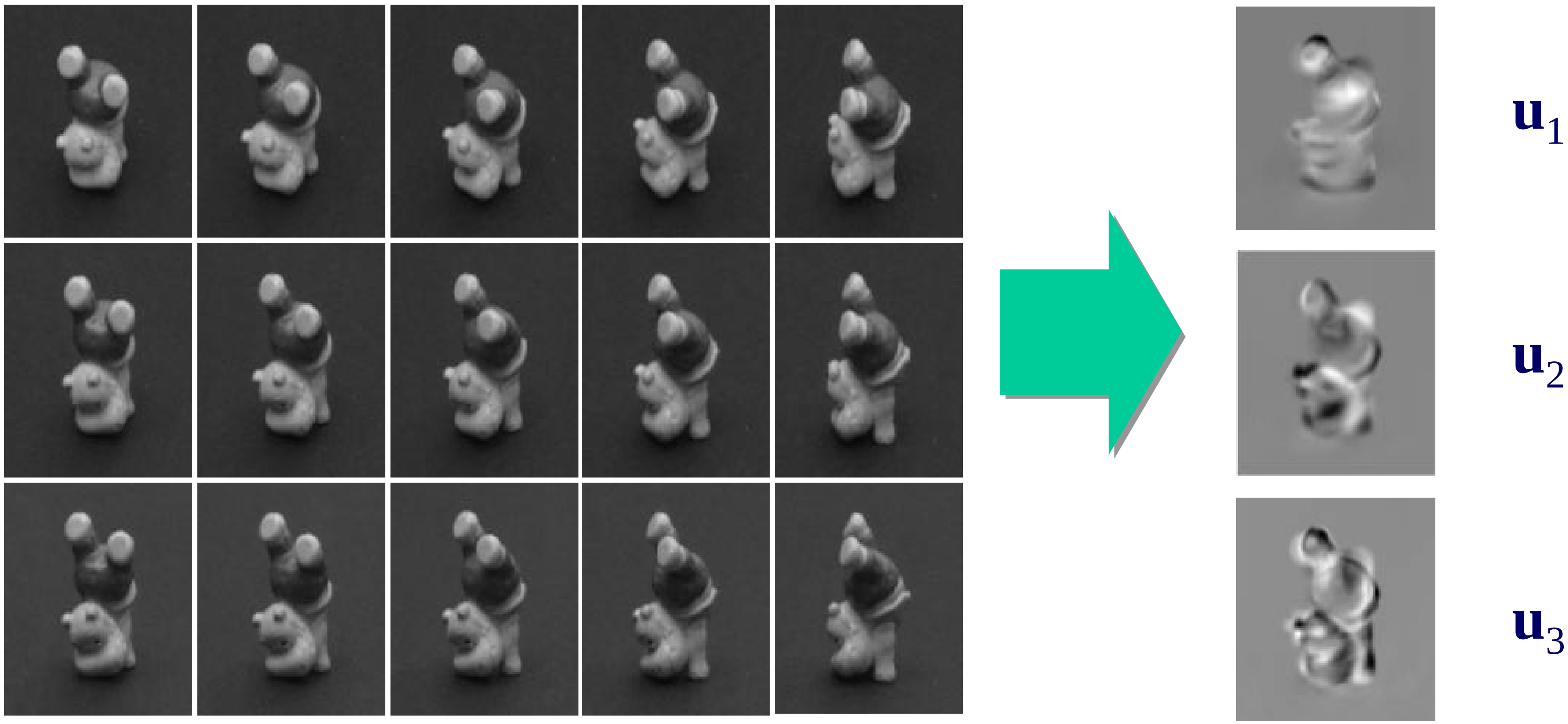
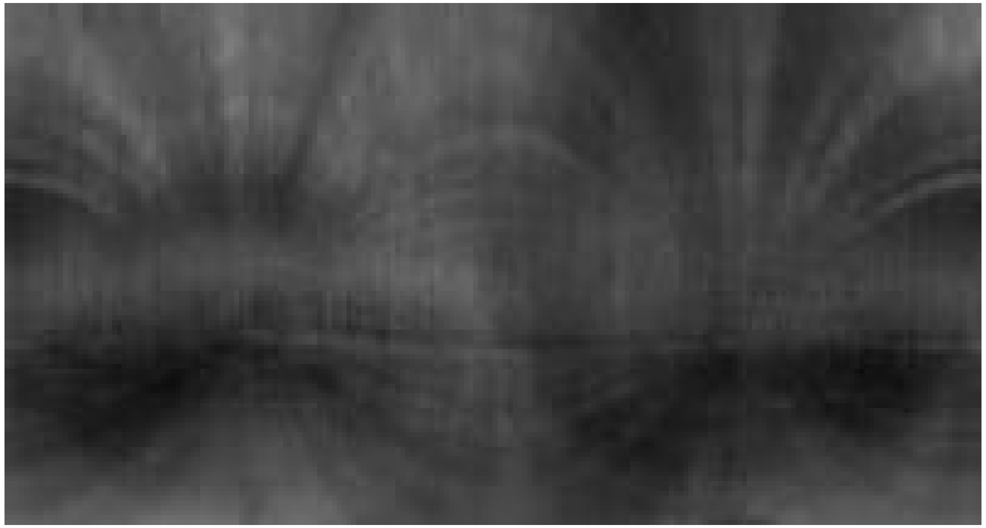


Image presentation with PCA



Properties PCA

- It can be shown that the mean square error between x_i and its reconstruction using only m principle eigenvectors is given by the expression :

$$\boxed{\sum_{j=1}^{N} l_j - \sum_{j=1}^{m} l_j = \sum_{j=m+1}^{N} l_j}$$

- PCA minimizes reconstruction error

- PCA maximizes variance of projection

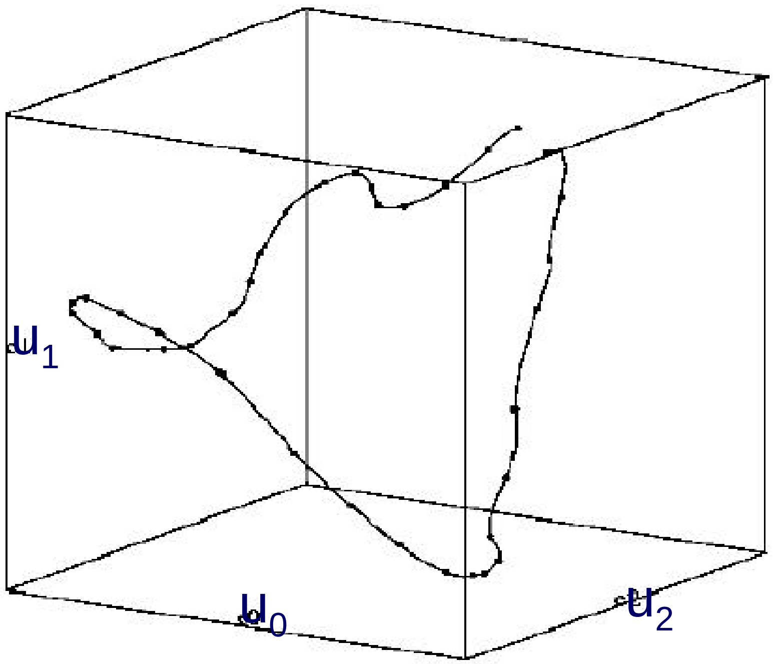
- Finds a more “natural” coordinate system for the sample data.

PCA for visual recognition and pose estimation

Objects are represented as coordinates in an n-dimensional eigenspace.

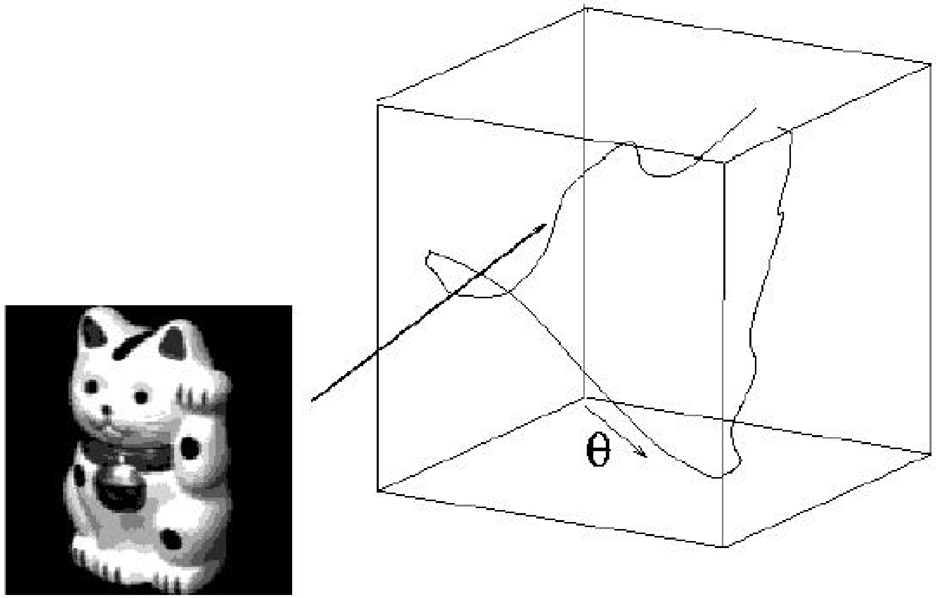
An example:

3-D space with points representing individual objects or a manifold representing **parametric eigenspace** (e.g., orientation, pose, illumination).



PCA for visual recognition and pose estimation

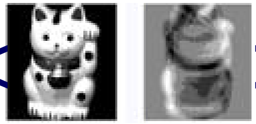
- Calculate coefficients
- Search for the nearest point (individual or on the curve)
- Point determines object and/or pose

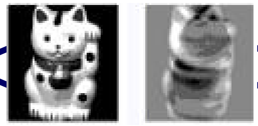
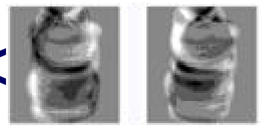
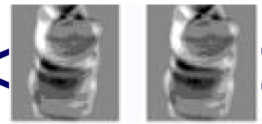


Calculation of coefficients

To recover q_i the image is projected onto the eigenspace

$$q_i(\mathbf{x}) = <\mathbf{x},\mathbf{u_i}> = \sum_{j=1}^{n-1} x_j u_{ij} \qquad 1 \le i \le k$$

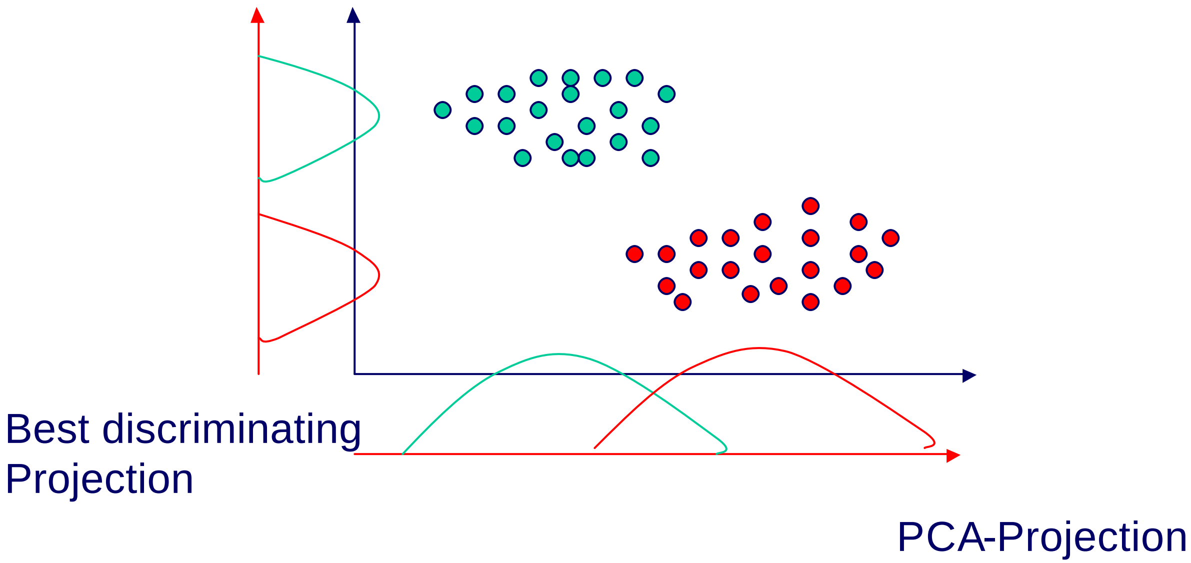
$<$  $> = q_1 <$  $> + q_2 <$  $> + \ldots = q_1$

$<$  $> = q_1 <$  $> + q_2 <$  $> + \ldots = q_2$

- Complete image **x** is required to calculate q_i.
- Corresponds to Least-Squares Solution

Linear Discriminant Analysis (LDA)

- **PCA minimizes projection error**



- **PCA is „unsupervised“ no information on classes is used**
- **Discriminating information might be lost**

LDA

- **Linear Discriminance Analysis (LDA)**
  - Maximize distance between classes
  - Minimize distance within a class

⇨ Fisher Linear Discriminance

$$r(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S_B} \mathbf{w}}{\mathbf{w}^T \mathbf{S_W} \mathbf{w}}$$

LDA: Problem formulation

- n Sample images: $\{x_1, \ldots, x_n\}$
- c classes: $\{c_1, \ldots, c_c\}$
- Average of each class: $$m_i = \frac{1}{n_i} \sum_{x_k \in c_i} x_k$$
- Total average: $$m = \frac{1}{n} \sum_{k=1}^{N} x_k$$

LDA: Practice

- **Scatter of class i:**

$$\mathbf{S}_i = \sum_{x_k \in c_i} (\mathbf{x}_k - m_i)(\mathbf{x}_k - m_i)^T$$

- **Within class scatter:**

$$\mathbf{S}_W = \sum_{i=1}^{c} \mathbf{S}_i$$

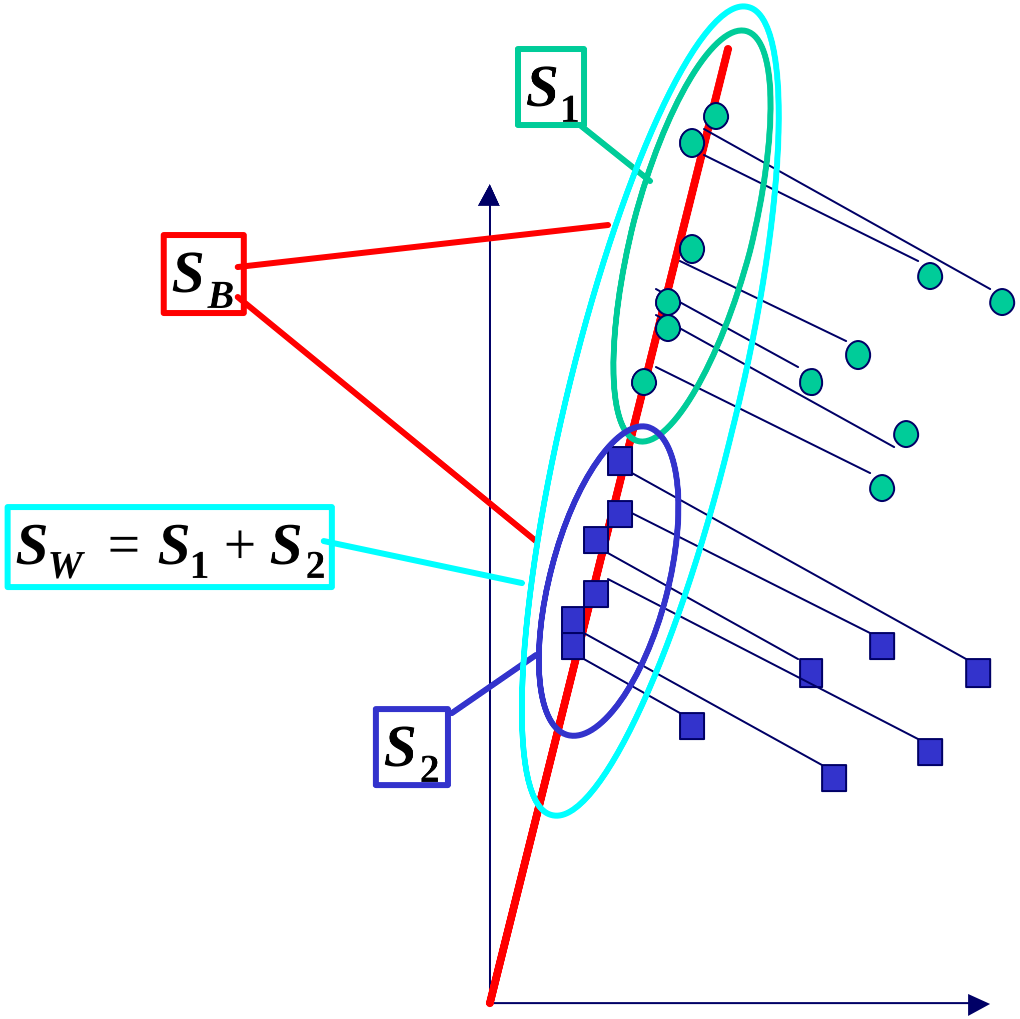
- **Between class scatter:**

$$\mathbf{S}_B = \sum_{i=1}^{c} |c_i| (m_i - m)(m_i - m)^T$$

- **Total scatter:**

$$\mathbf{S}_T = \mathbf{S}_W + \mathbf{S}_B$$

LDA



Good separation

LDA

- **Maximization of**

$$r(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S_B} \mathbf{w}}{\mathbf{w}^T \mathbf{S_W} \mathbf{w}}$$

- **is given by solution of generalized eigenvalue problem**

$$\mathbf{S_B}\mathbf{w} = l\,\mathbf{S_w}\mathbf{w}$$

- **For the c-class case we obtain (at most) c-1 projections as the largest eigenvalues of**

$$\mathbf{S_B}\mathbf{w}_i = l\,\mathbf{S_w}\mathbf{w}_i$$

LDA

- **Example Fisherface of recognition Glasses/NoGlasses (Belhumeur et.al. 1997)**

Canonical Correlation Analysis (CCA)

- **Also „supervised" method but motivated by regression tasks, e.g. pose estimation.**

- **Canonical Correlation Analysis relates two sets of observations by determining pairs of directions that yield maximum correlation between these sets.**

- **Find a pair of directions (canonical factors)** $w_x \in \Re^p$, $w_y \in \Re^q$, **so that the correlation of the projections c =** $w_x^T x$ **and d =** $w_y^T y$ **becomes maximal.**

What is CCA?

Canonical Correlation $0 \leq r \leq 1$

$$r = \frac{E[cd]}{\sqrt{E[c^2]E[d^2]}} =$$

$$\frac{E[\mathbf{w}_x^T \mathbf{x}\mathbf{y}^T \mathbf{w}_y]}{\sqrt{E[\mathbf{w}_x^T \mathbf{x}\mathbf{x}^T \mathbf{w}_x]E[\mathbf{w}_y^T \mathbf{y}\mathbf{y}^T \mathbf{w}_y]}} =$$

Between Set Covariance Matrix

$$\frac{\mathbf{w}_x^T \mathbf{C}_{xy} \mathbf{w}_y}{\sqrt{\mathbf{w}_x^T \mathbf{C}_{xx} \mathbf{w}_x \mathbf{w}_y^T \mathbf{C}_{yy} \mathbf{w}_y}}$$

What is CCA?

- Finding solutions

$$\mathbf{w} = \begin{pmatrix} \mathbf{w}_x \\ \mathbf{w}_y \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} 0 & \mathbf{C}_{xy} \\ \mathbf{C}_{yx} & 0 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} \mathbf{C}_{xx} & 0 \\ 0 & \mathbf{C}_{yy} \end{pmatrix}$$

Rayleigh Quotient

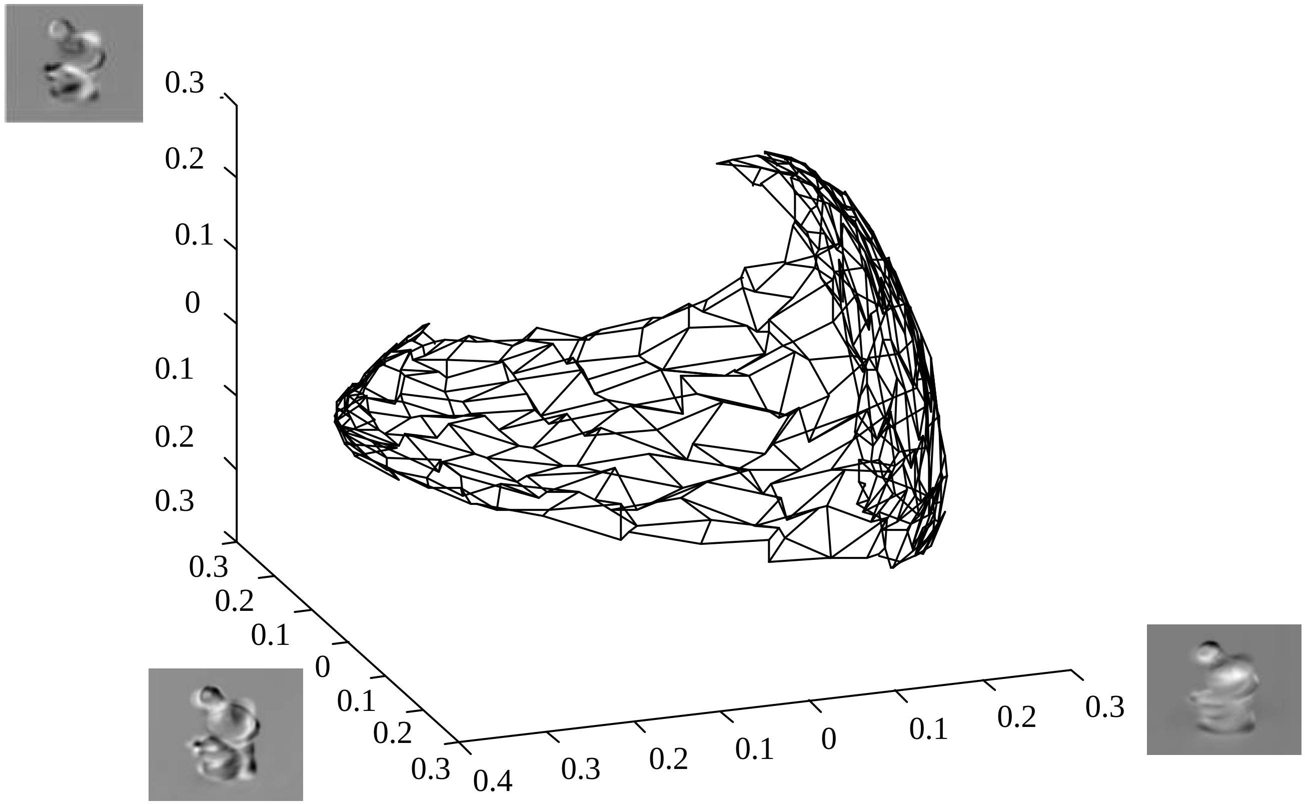
$$r = \frac{\mathbf{w}^T \mathbf{A} \mathbf{w}}{\mathbf{w}^T \mathbf{B} \mathbf{w}}$$

Generalized Eigenproblem

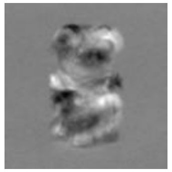
$$\mathbf{A}\mathbf{w} = m\mathbf{B}\mathbf{w}$$

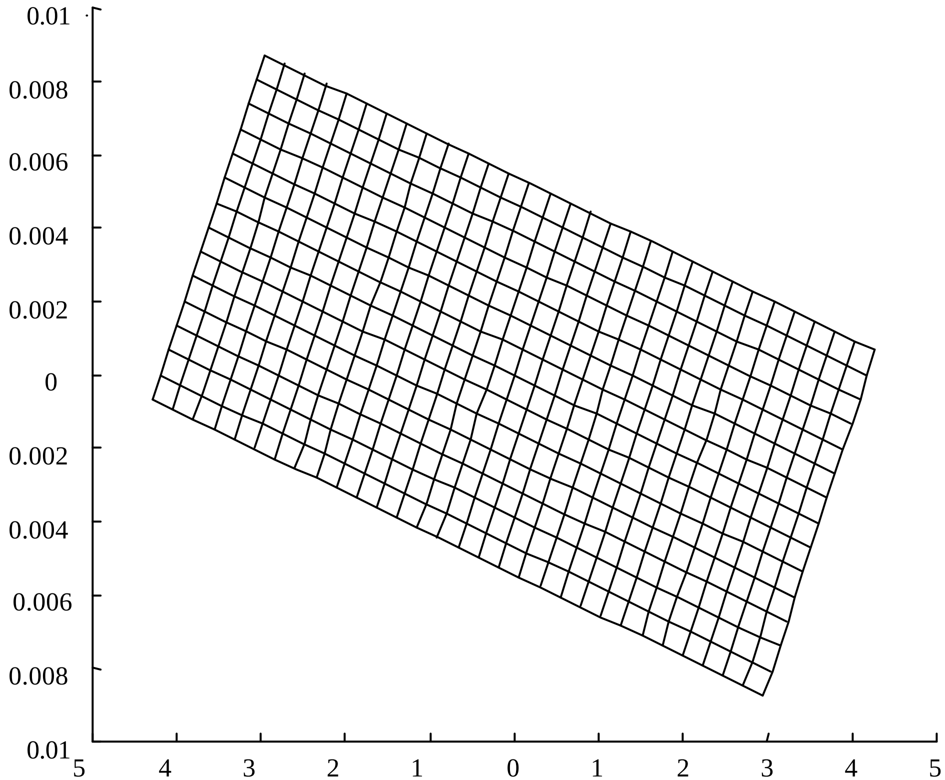
CCA Example

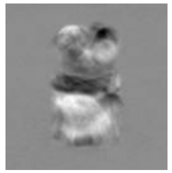
Parametric eigenspace obtained by PCA for 2DoF in pose



CCA representation
(projections of training images onto $\mathbf{w}_{x1}$, $\mathbf{w}_{x2}$)







Independent Component Analysis (ICA)

- ICA is a powerful technique from signal processing (Blind Source Separation)
- Can be seen as an extension of PCA
- PCA takes into account only statistics up to 2nd order
- ICA finds components that are statistically independent (or as independent as possible)

Independent Component Analysis (ICA)

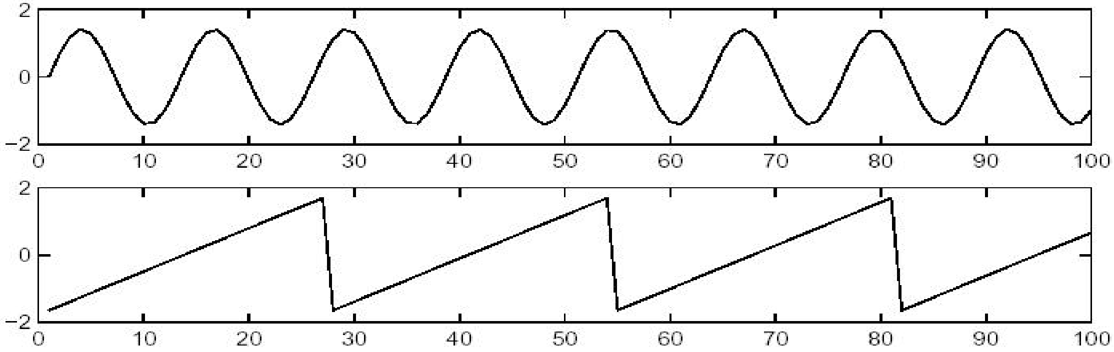
- m scalar variables $X=(x_1 \ldots x_m)^T$
- They are assumed to be obtained as linear mixtures of n sources $S=(s_1 \ldots s_n)^T$

$$\mathbf{X} = \mathbf{A}\mathbf{S}$$

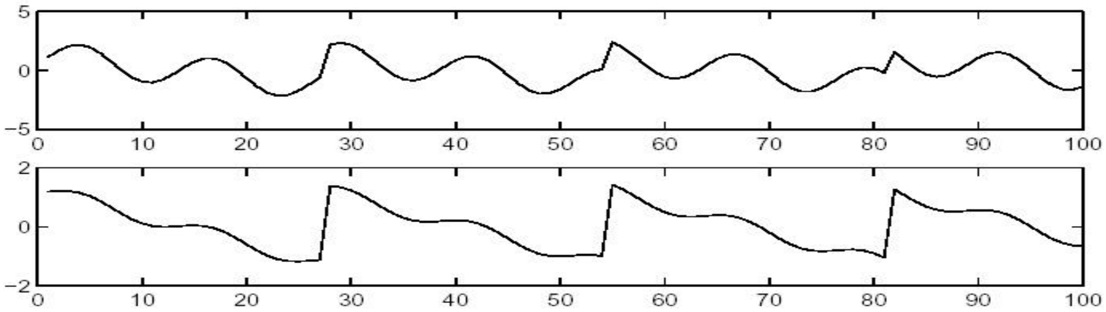
- Task: Given X find A, S (under the assumption that S are independent)

ICA Example

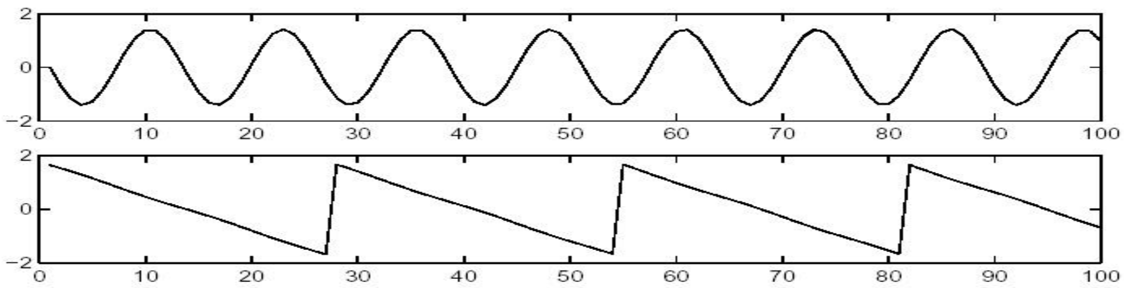
Original Sources



Mixtures

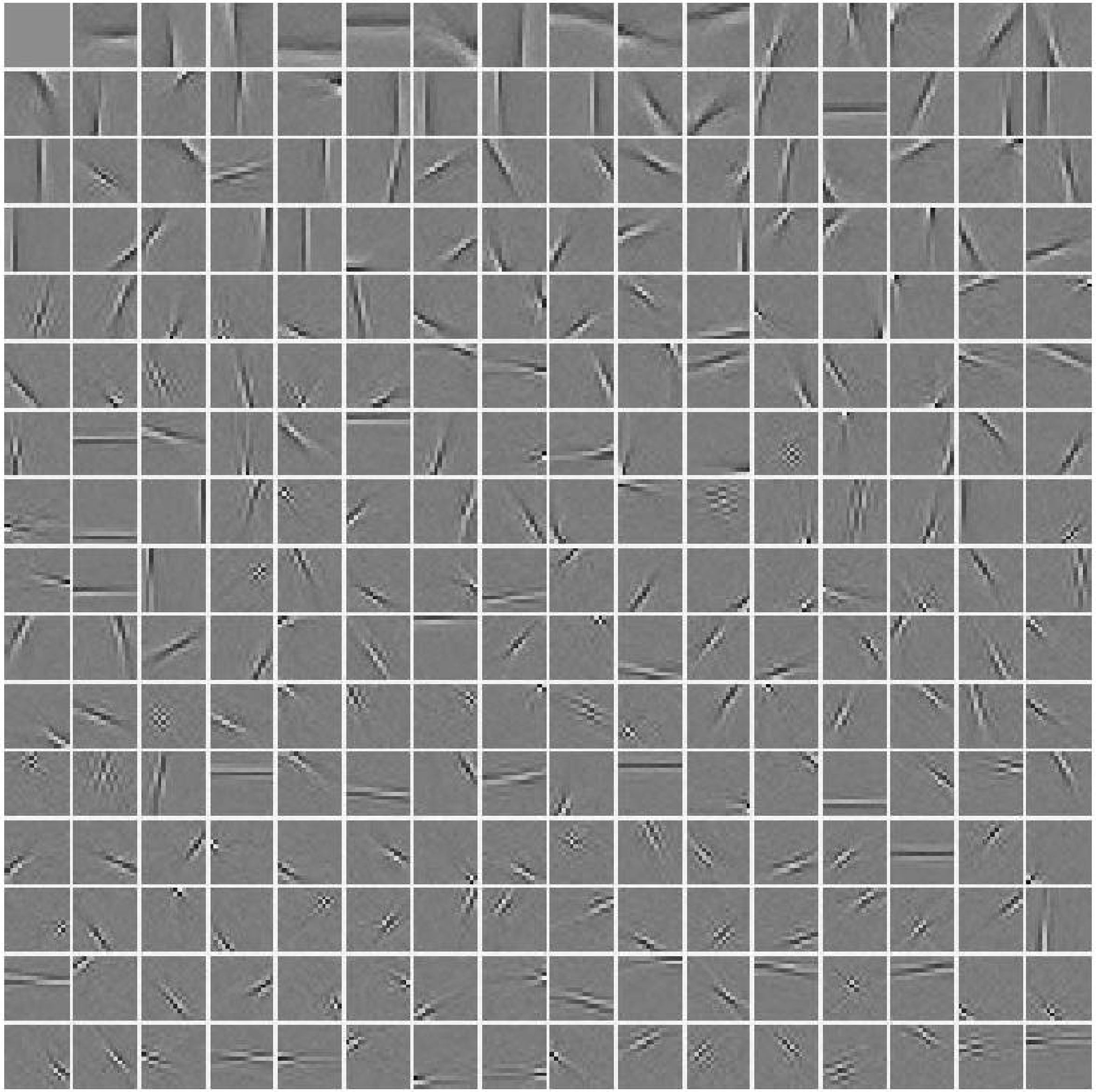


Recovered Sources



ICA Example

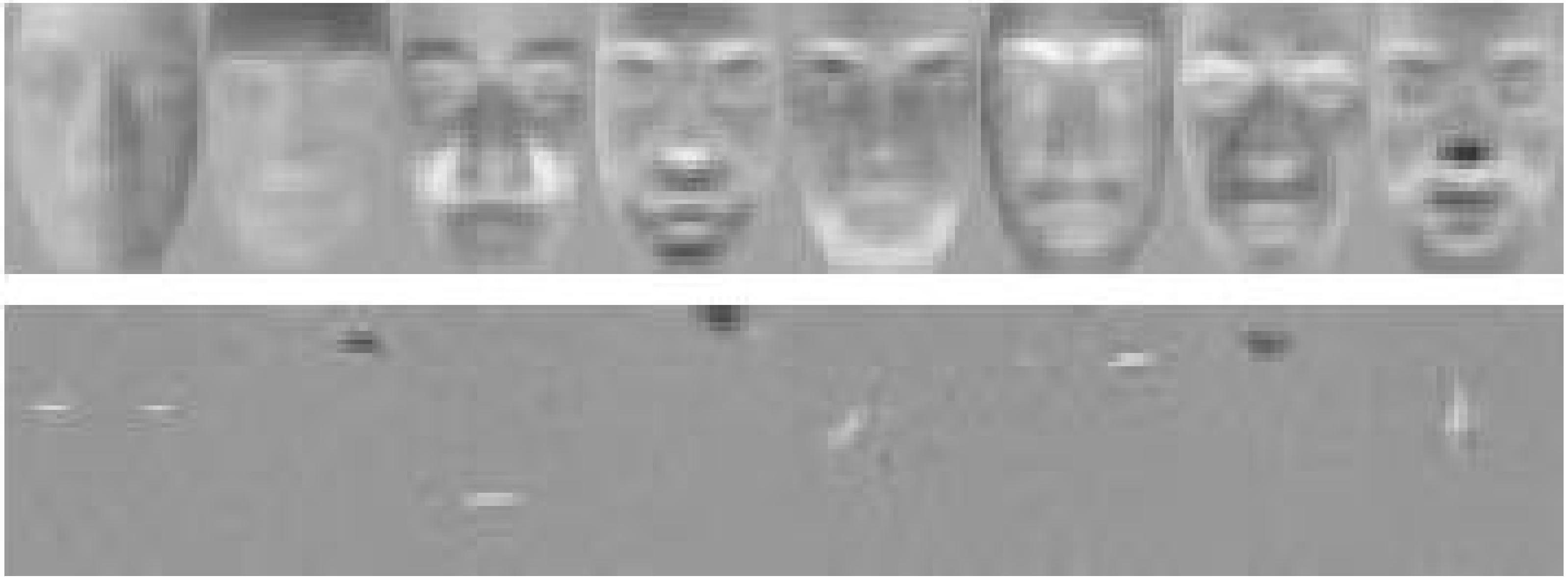
ICA basis obtained from 16x16 patches of natural images (Bell&Sejnowski 96)



Face Recognition using ICA

- **PCA vs. ICA on Ferret DB (Baek et.al. 02)**

PCA



ICA

Non-Negative Matrix Factorization (NMF)

- How can we obtain part-based representation?
- Local representation where parts are added
- E.g. learn from a set of faces the parts a face consists of, i.e. eyes, nose, mouth, etc.
- Non-Negative Matrix Factorization (Lee & Seung 1999) lead to part based representation

Matrix Factorization - Constraints

V* » *WH

- **PCA: *W* are orthonormal basis vectors**

$$W = [\vec{w}_1, \vec{w}_2, \Lambda\ , \vec{w}_n], \qquad \vec{w}_i \cdot \vec{w}_j = d_{ij}$$

- **VQ : *H* are unity vectors**

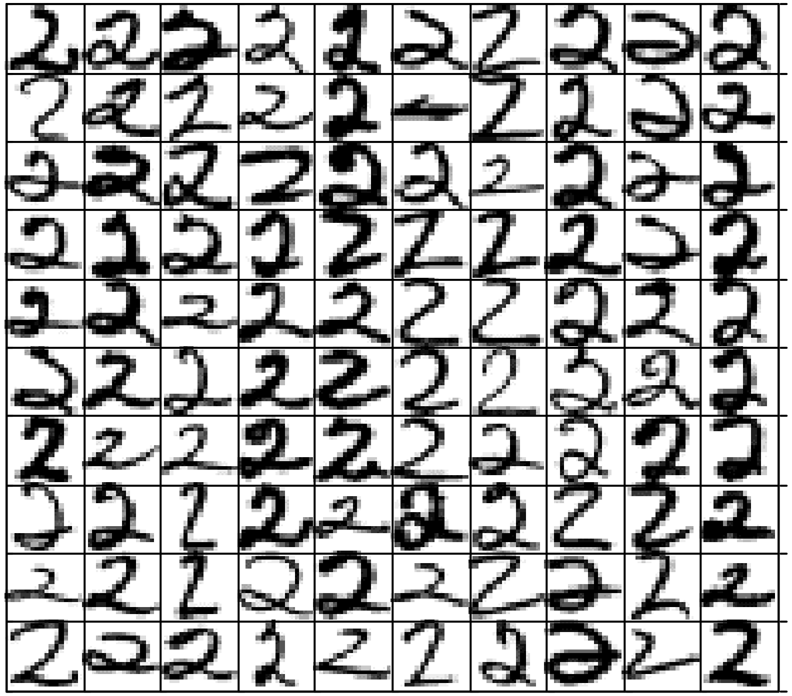
$$H = [\vec{h}_1, \vec{h}_2, \Lambda\ , \vec{h}_n], \qquad \vec{h}_j^{\,T} = [0,0,1,0,\Lambda\ ,0]$$

- **NMF: *V,W,H* are non-negative**

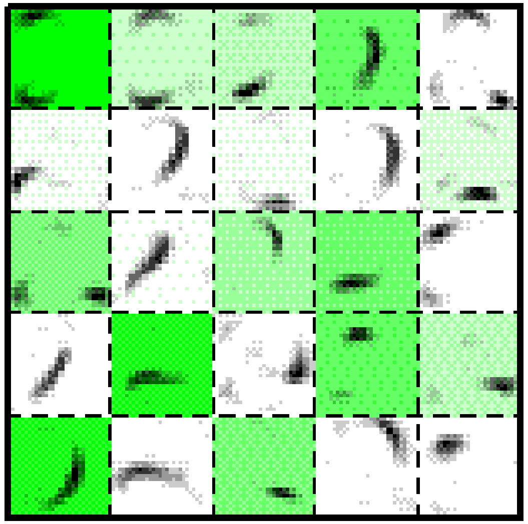
$$V_{ij}, W_{ij}, H_{ij} \geq 0 \qquad \forall i, j$$

Learning

Find basis images from the training set

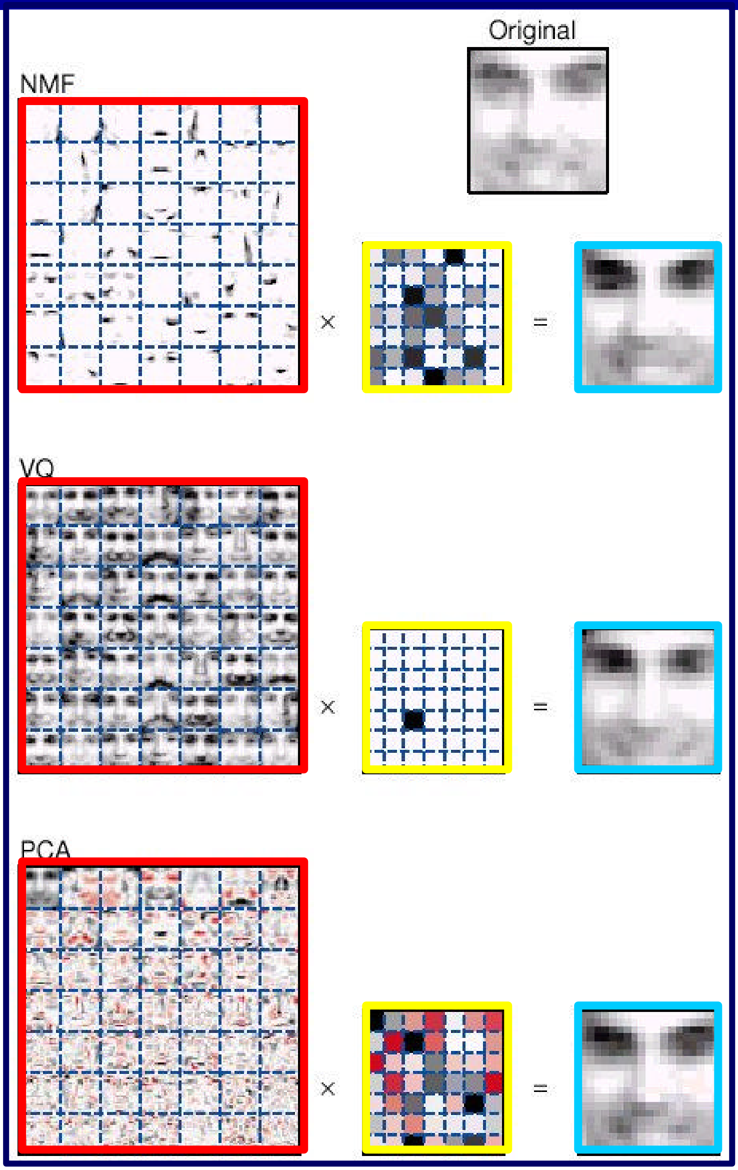


Training images



Basis images

Face features



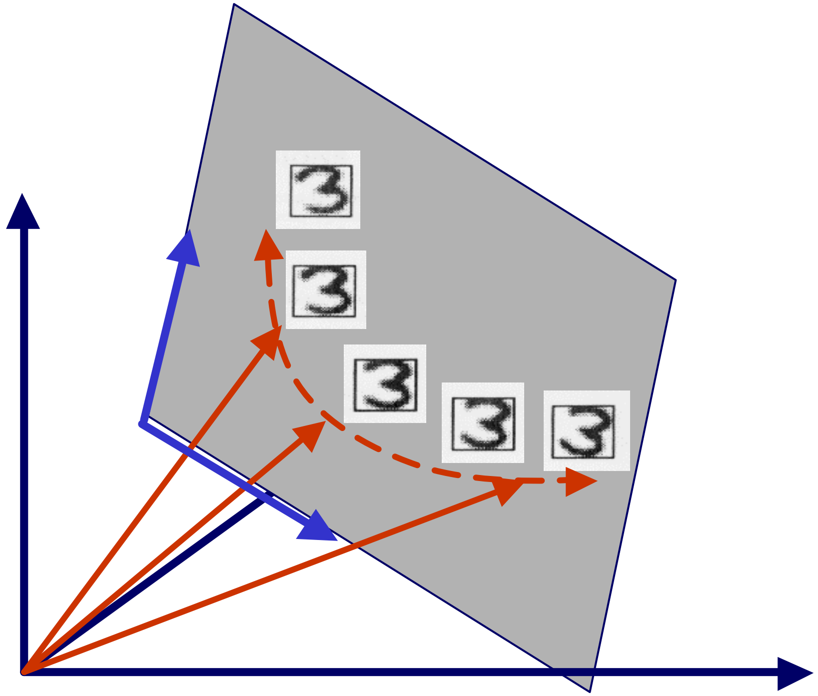
Basis images

Encoding (Coefficients)

Reconstructed image

Kernel Methods

- All presented methods are linear



- Can we generalize to non-linear methods in a computational efficient manner?

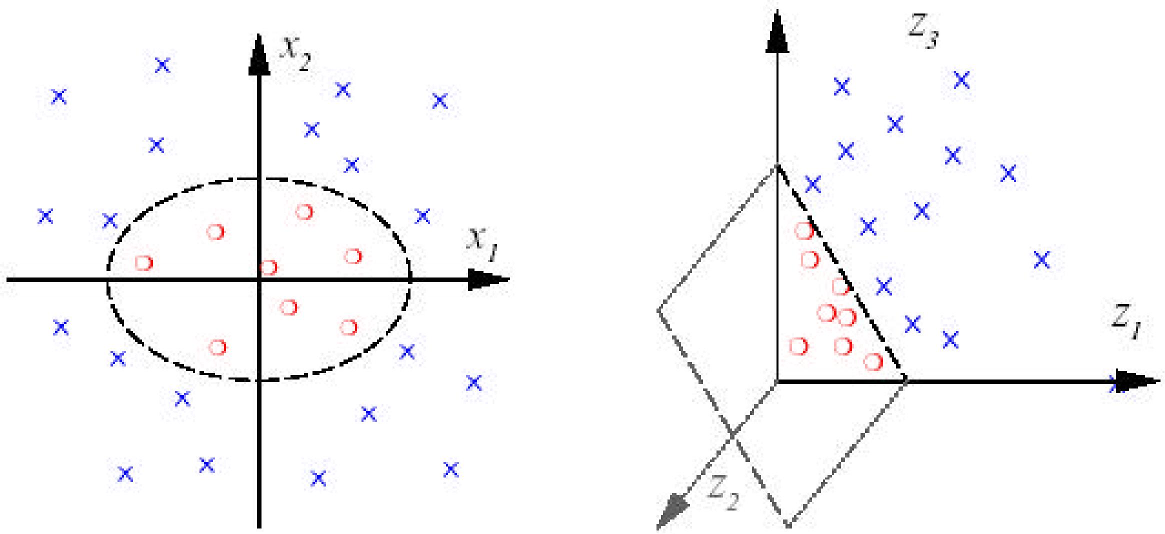
Kernel Methods

- **Kernel Methods are powerful methods (introduced with Support Vector Machines) to generalize linear methods**

BASIC IDEA:

1. **Non-linear mapping of data in high dimensional space**
2. **Perform linear method in high-dimensional space**

   ➡ **Non-linear method in original space**



Outline Part 2

- Robot localization
- Robust representations and recognition
- Robust recognition using PCA
- Scale invariant recognition using PCA
- Illumination insensitive recognition
- Representations for panoramic images
- Incremental building of eigenspaces
- Multiple eigenspaces for efficient representation
- Robust building of eigenspaces
- Research issues

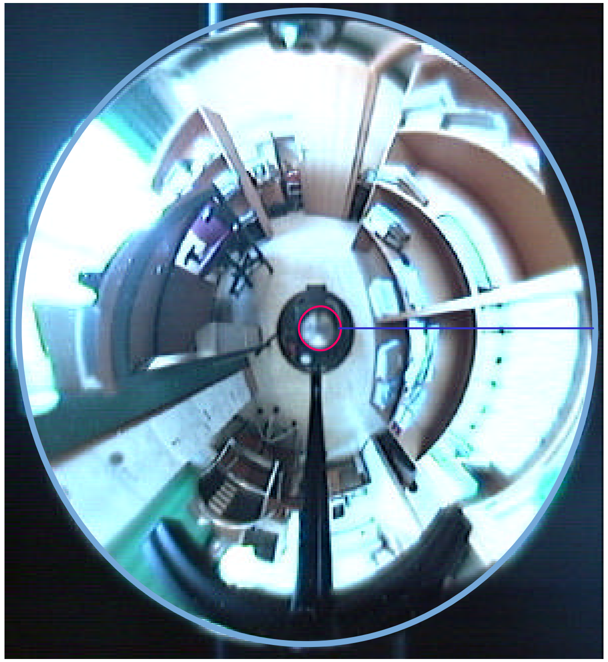
Appearance-based approaches

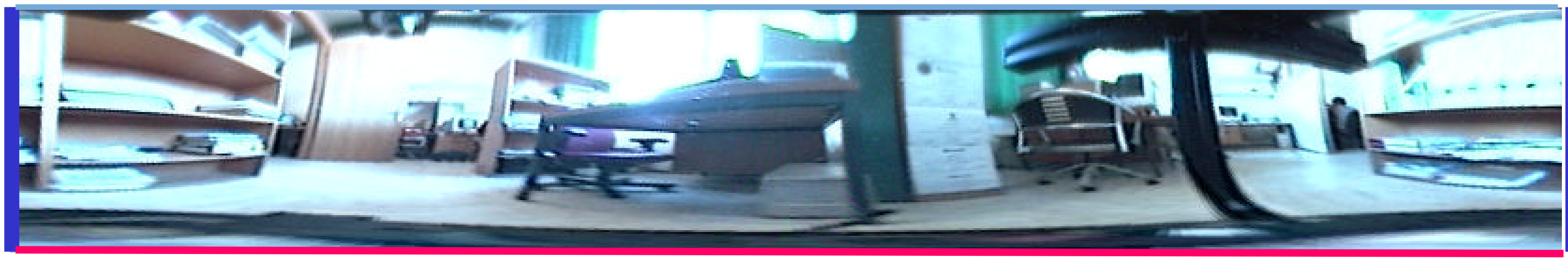
A variety of successful applications:

- Human face recognition e.g. [Turk & Pentland]
- Visual inspection e.g. [Yoshimura & Kanade]
- Visual positioning and tracking of robot manipulators, e.g. [Nayar & Murase]
- Tracking e.g., [Black & Jepson]
- Illumination planning e.g., [Murase & Nayar]
- Image spotting e.g., [Murase & Nayar]
- Mobile robot localization e.g., [Jogan & Leonardis]
- Background modeling e.g., [Oliver, Rosario & Pentland]

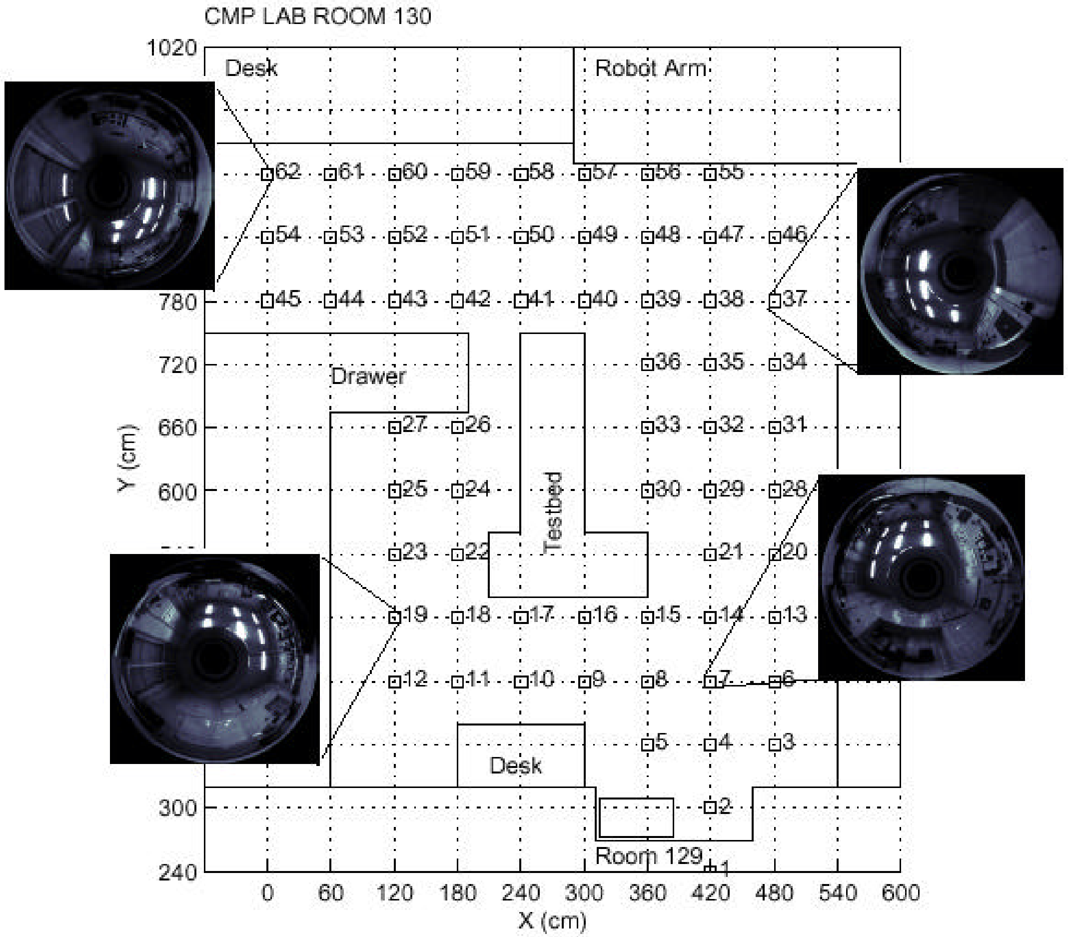
Mobile Robot

Panoramic image





Environment map



- ♦ **environments are represented by a large number of views**
- ♦ **localisation = recognition**

Compression with PCA

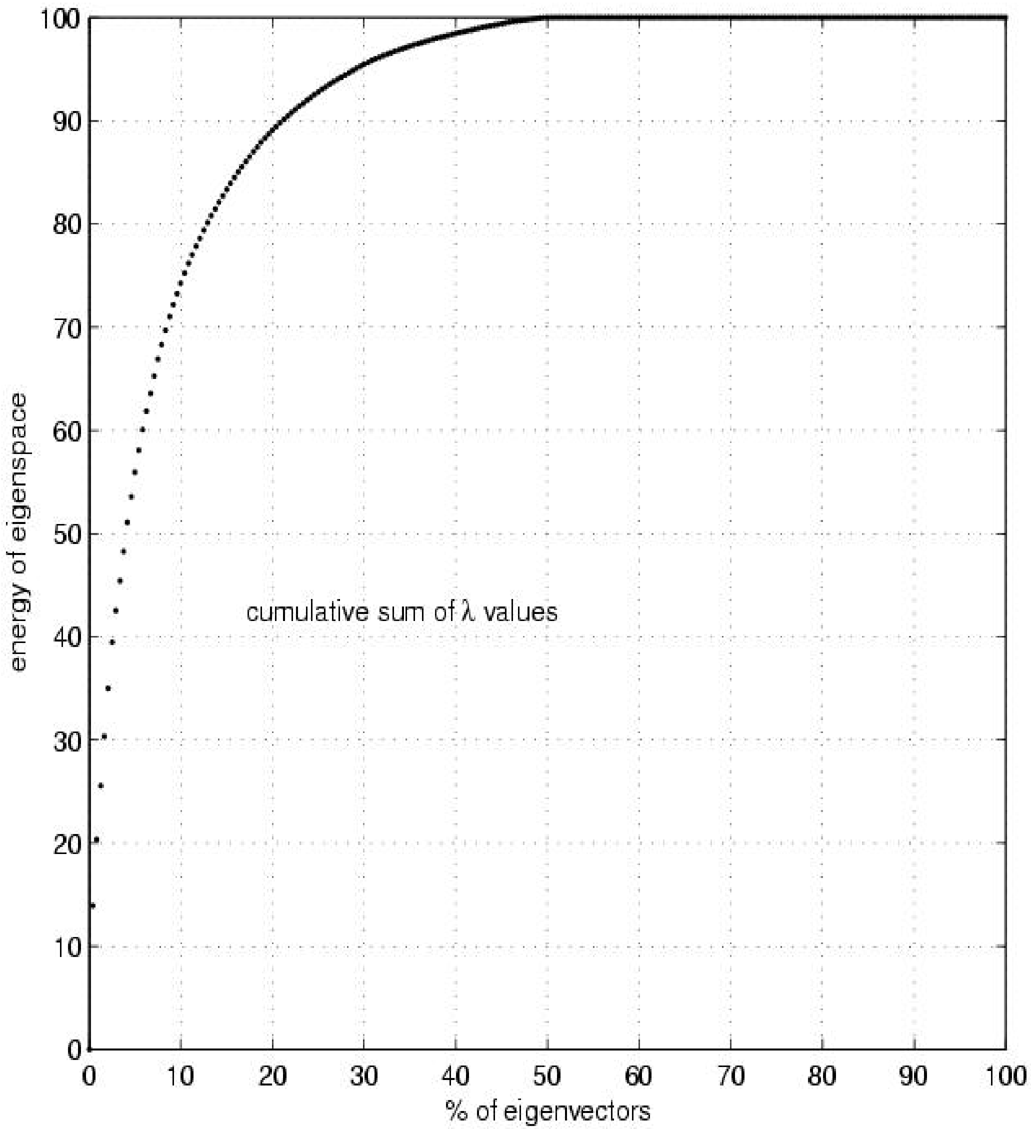
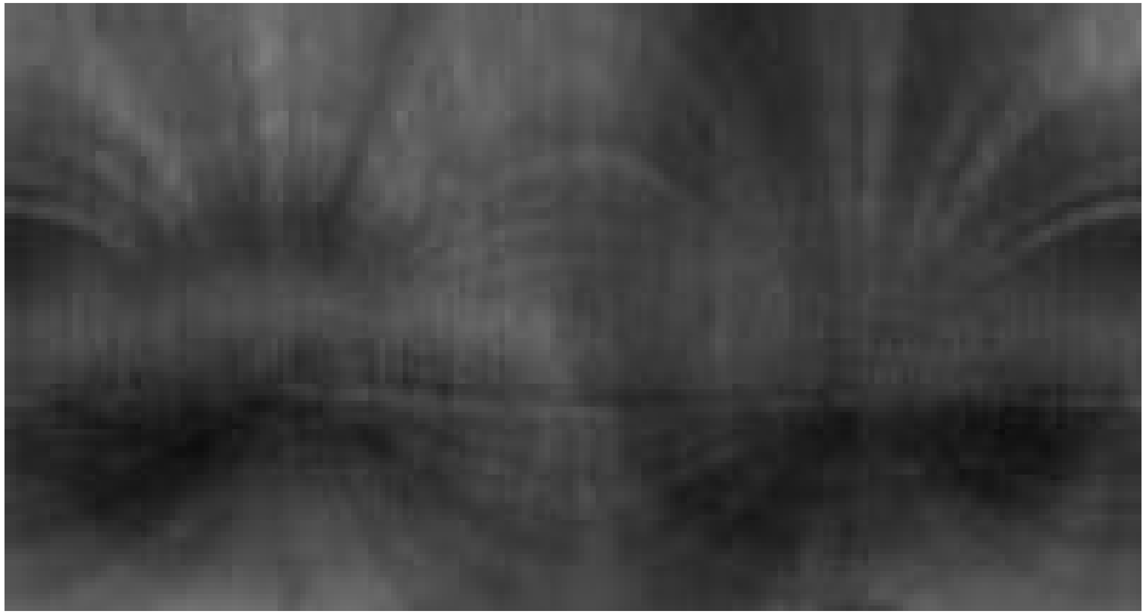
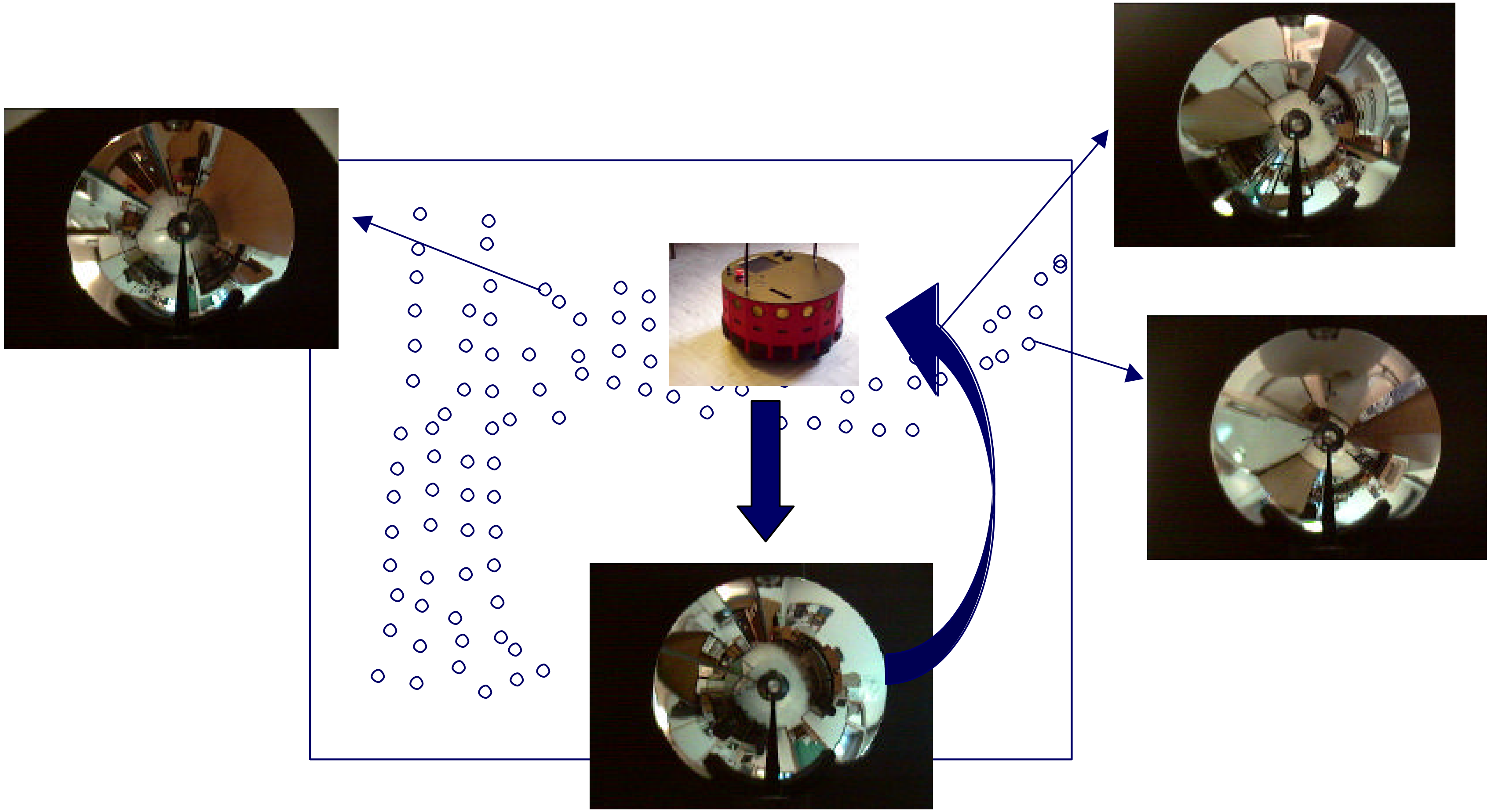


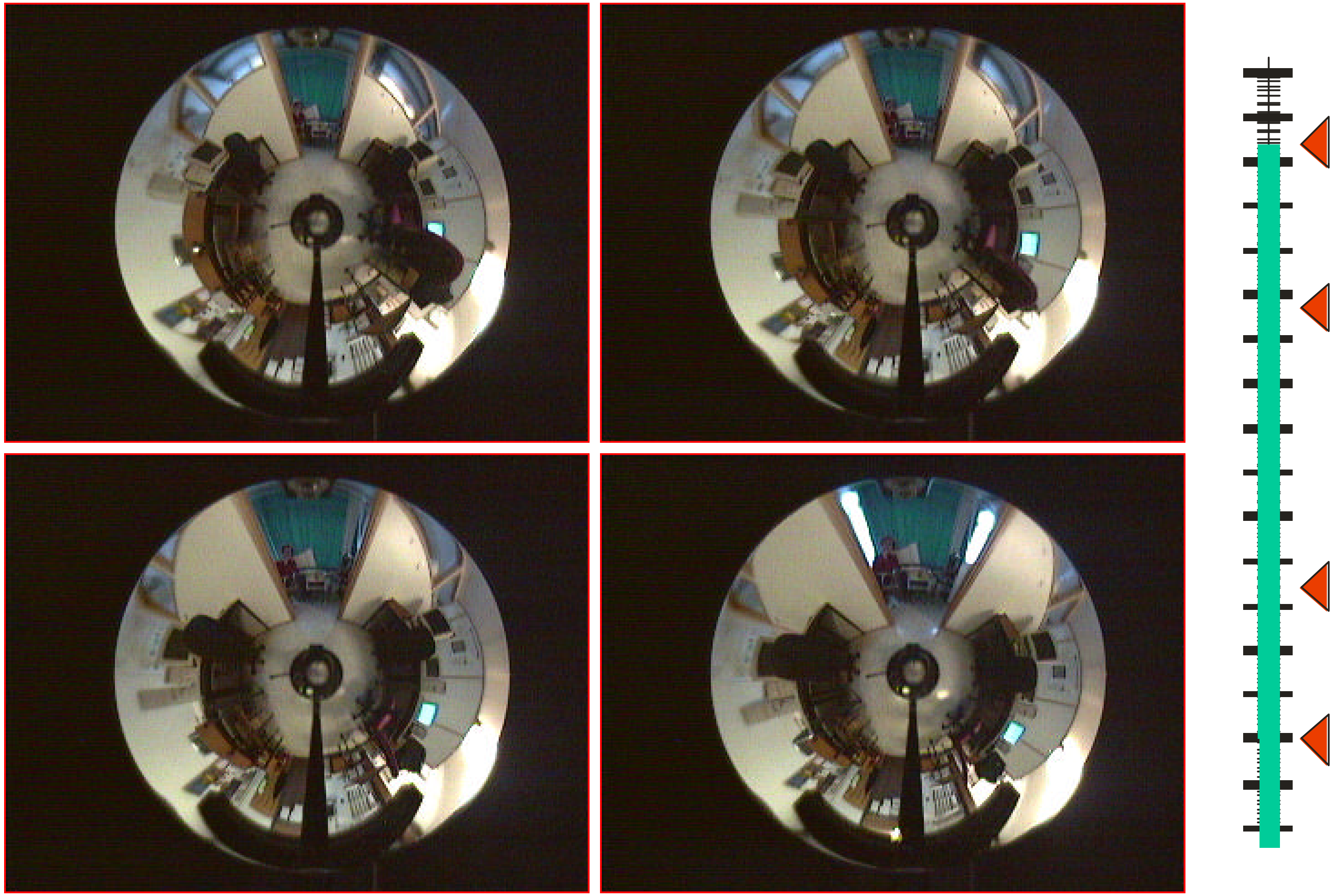
Image representation with PCA



Localisation



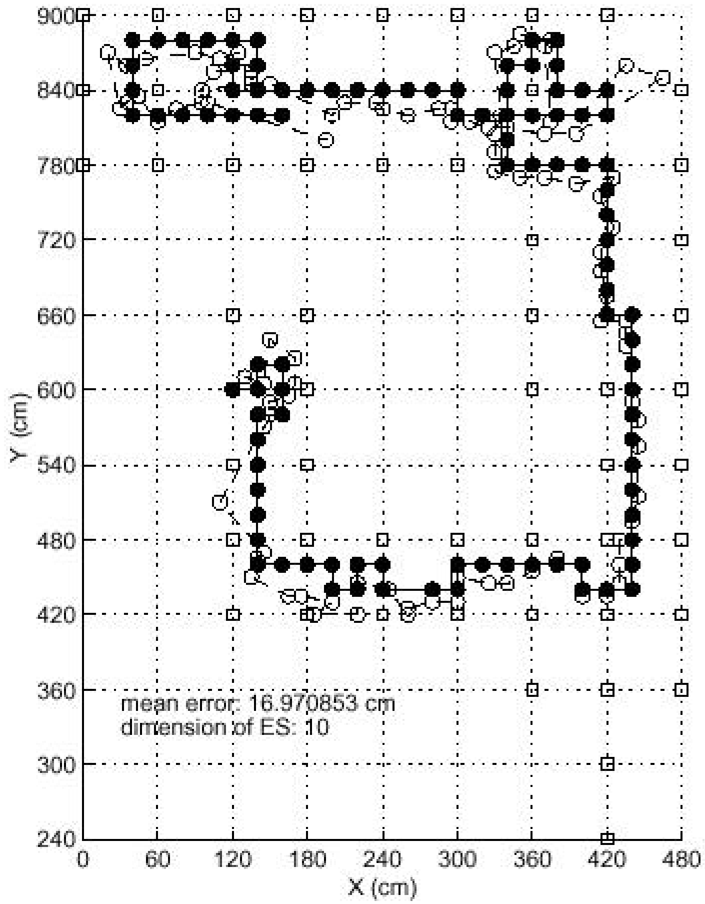
Distance vs. similarity



Robot localisation

- Interpolated hyper-surface represents the memorized environment.
- The parameters to be retrieved are related to position and orientation.
- Parameters of an input image are obtained by scalar product.

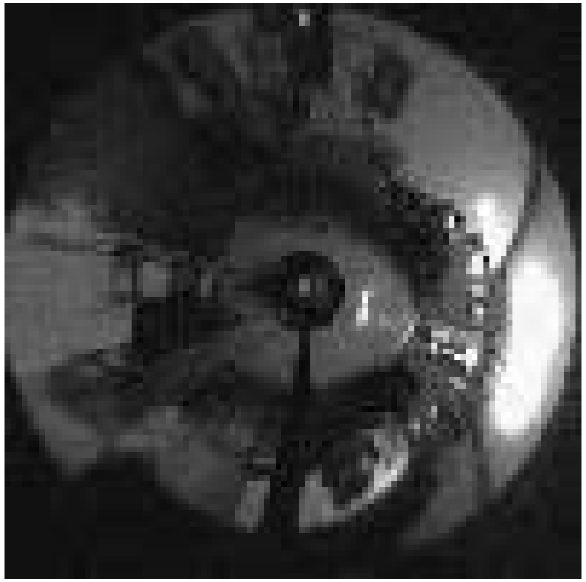
Localisation



Enhancing recognition and representations

- **Occlusions, varying background, outliers**
  - Robust recognition using PCA
- **Scale variance**
  - Multiresolution coefficient estimation
  - Scale invariant recognition using PCA
- **Illumination variations**
  - Illumination insensitive recognition
- **Rotated panoramic images**
  - Spinning eigenimages
- **Incremental building of eigenspaces**
- **Multiple eigenspaces for efficient representations**
- **Robust building of eigenspaces**

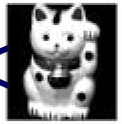
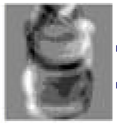
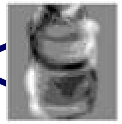
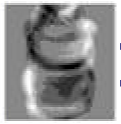
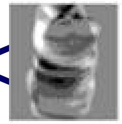
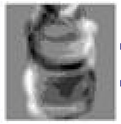
Occlusions

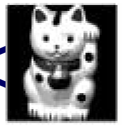
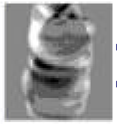
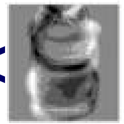
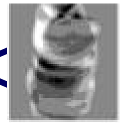
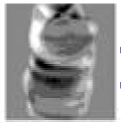


Calculation of coefficients

To recover q_i the image is projected onto the eigenspace

$$q_i(\mathbf{x}) = <\mathbf{x},\mathbf{u_i}> = \sum_{j=1}^{n-1} x_j u_{ij} \qquad 1 \le i \le k$$

$<$   $> = q_1 <$   $> + q_2 <$   $> + \ldots = q_1$

$<$   $> = q_1 <$   $> + q_2 <$   $> + \ldots = q_2$

- Complete image **x** is required to calculate q_i.
- Corresponds to Least-Squares Solution

Non-robustness

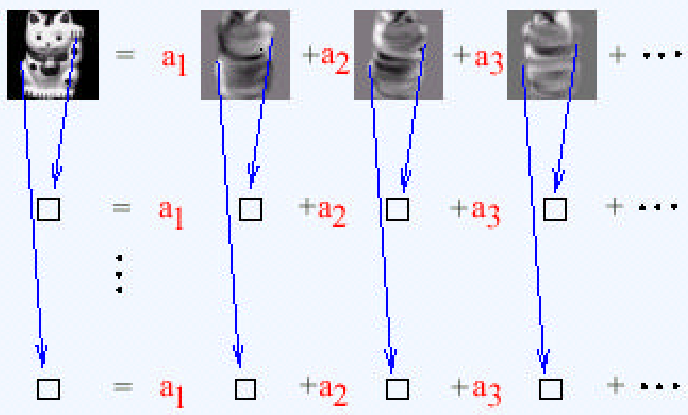
Drawbacks: Prone to errors caused by

- occlusions (outliers)
- cluttered background

Original | Occluded | Reconstruction

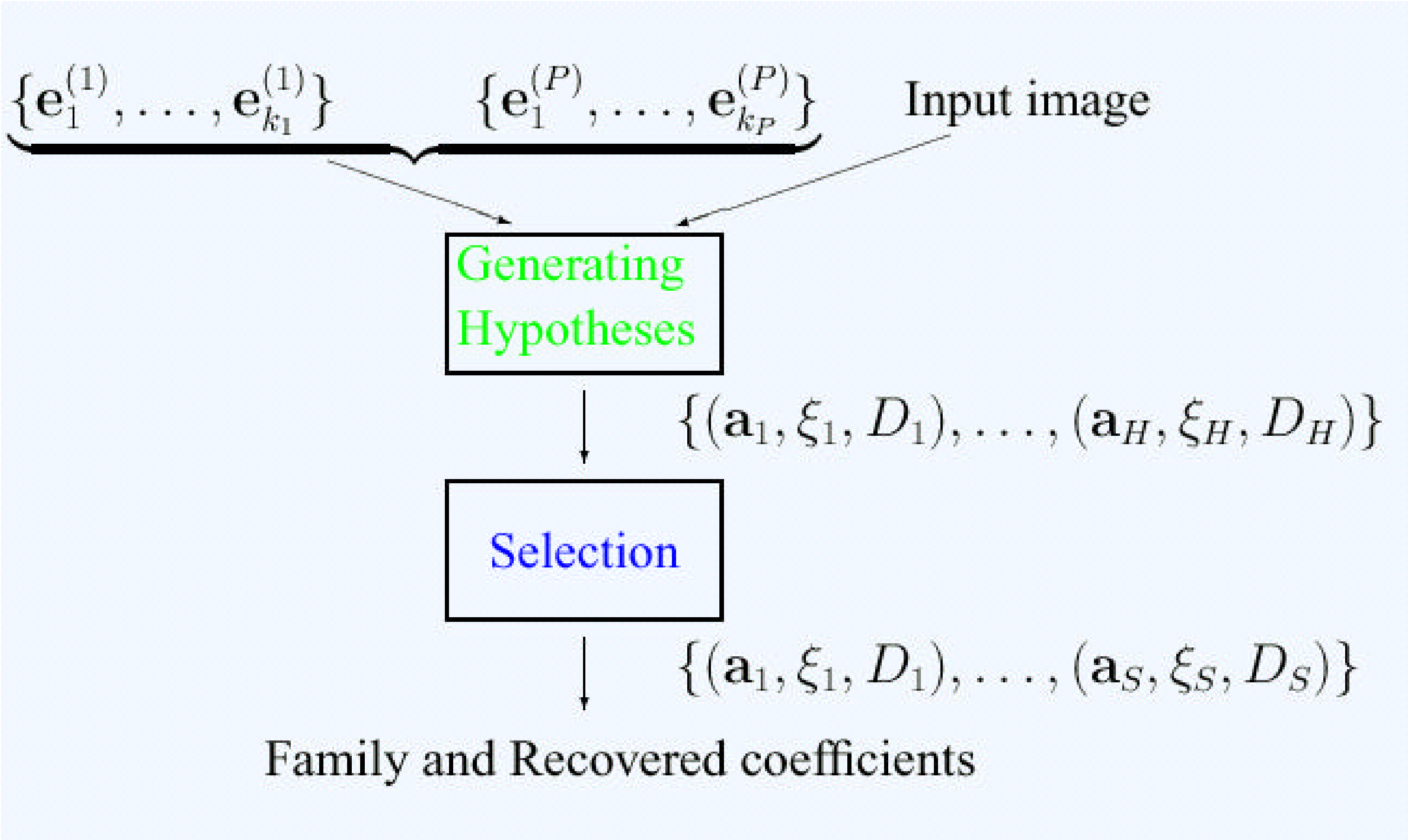
Robust method

- **Major idea:** Instead of using the standard approach we:
  - **subset of data points** $\longrightarrow$ linear system of equations
  - **Robust** solution of this system of equations
  - Perform **multiple hypotheses**



- Hypothesize-and-test paradigm
- Competing hypotheses are subject to a **selection** procedure based on the MDL principle.

Robust algorithm



Selection

Three cases:

1. **One object**: Select best match (c_{ii})
2. Multiple **non-overlapping** objects: Select local maximum (c_{ii})
3. Multiple **overlapping** objects: MDL-criterion:

The objective function:

$$F(\mathbf{h}) = \mathbf{h}^T \mathbf{C} \mathbf{h}$$

$\mathbf{h}^T = [h_1, h_2, \ldots, h_R]$ — set of hypotheses

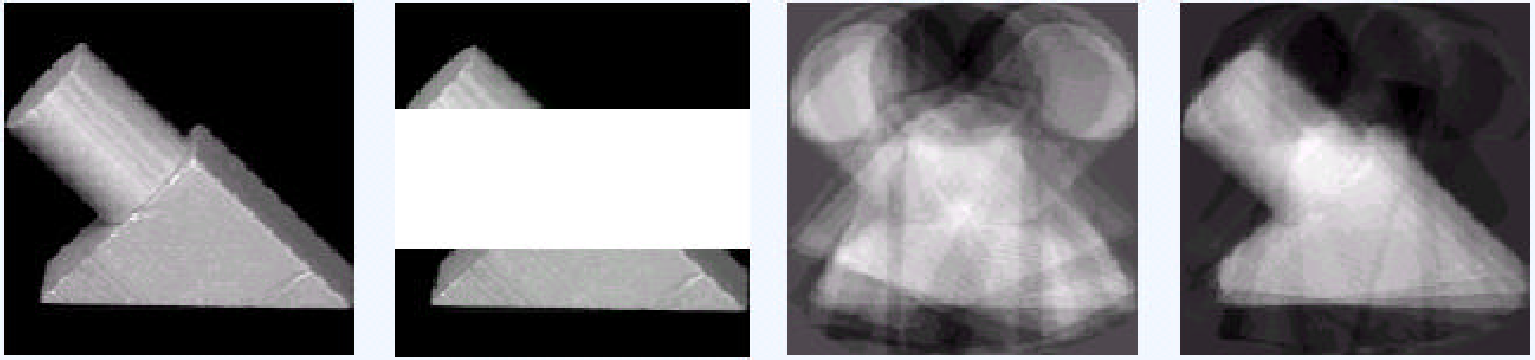
Diagonal terms of $\mathbf{C}$ express the cost-benefit value for hypothesis i

$$c_{ii} = \mathrm{K}_1 |D_i| - \mathrm{K}_2 ||\vec{\xi}_i||_{D_i} - \mathrm{K}_3 N_i$$

Off-diagonal terms handle overlapping hypotheses

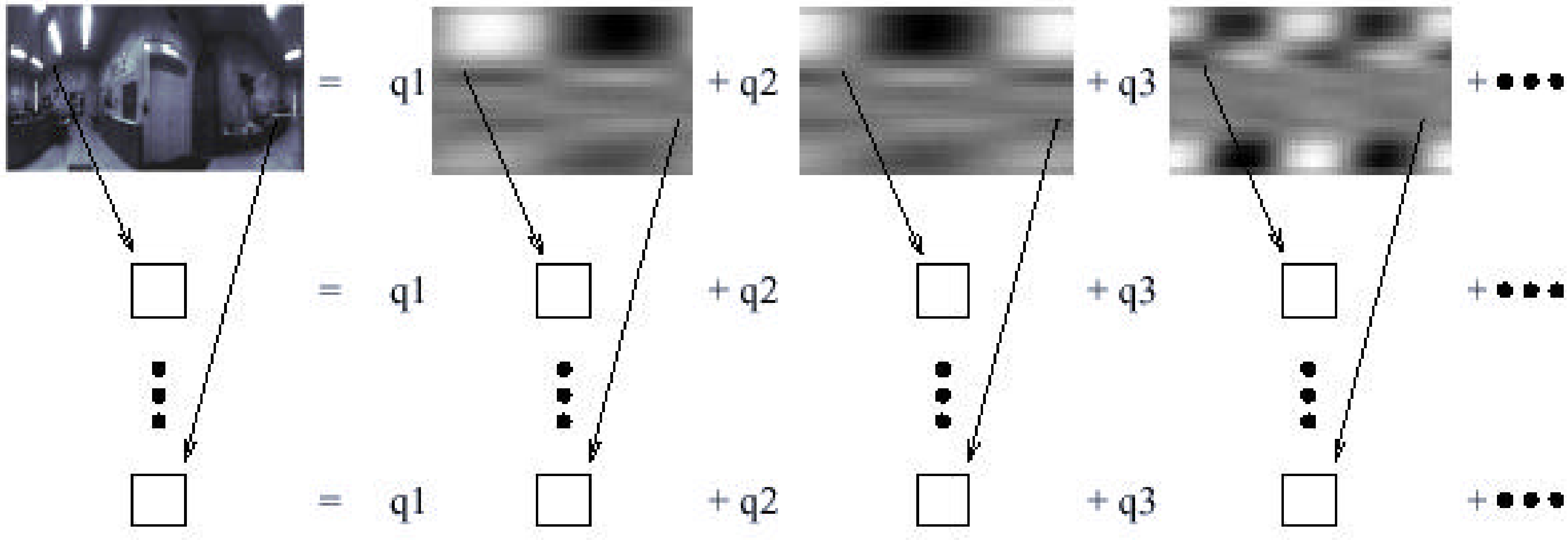
$$c_{ij} = \frac{-\mathrm{K}_1 |D_i \cap D_j| + \mathrm{K}_2 \xi_{ij}}{2}$$

Robust recovery of coefficients

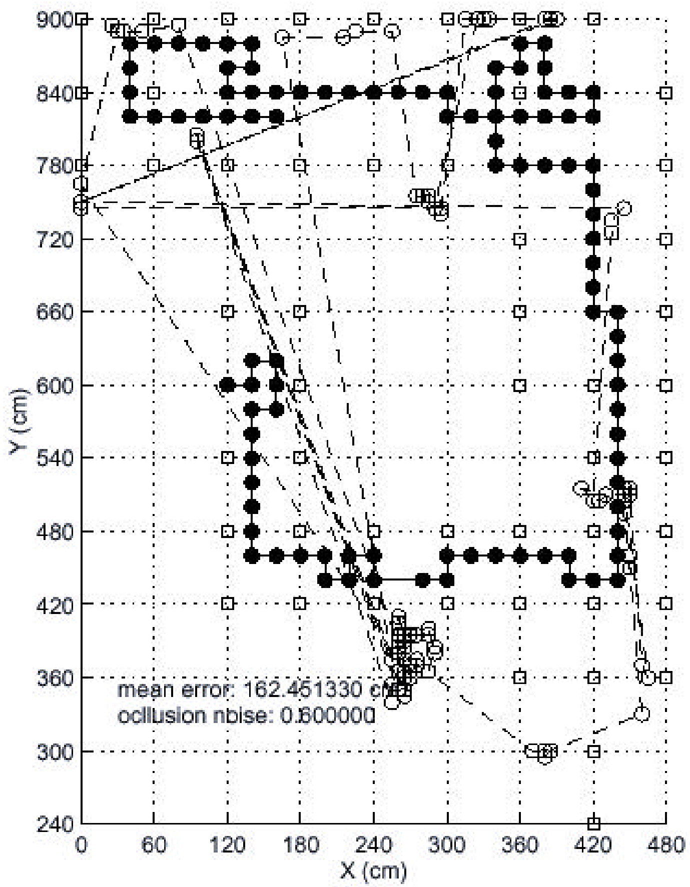


Original	Occluded	Standard	Robust

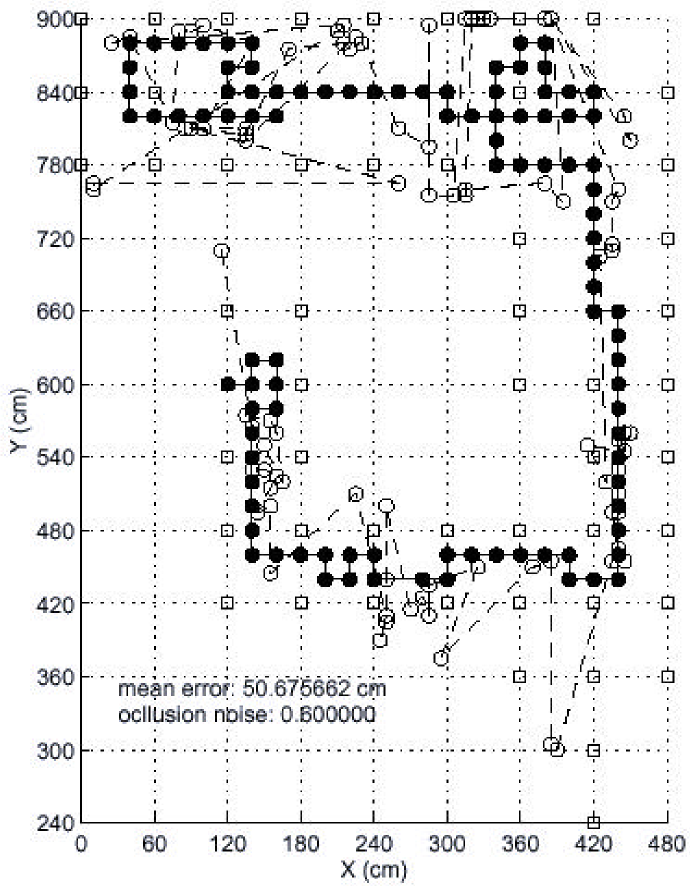
Robust localisation under occlusions



Robust localisation at 60% occlusion



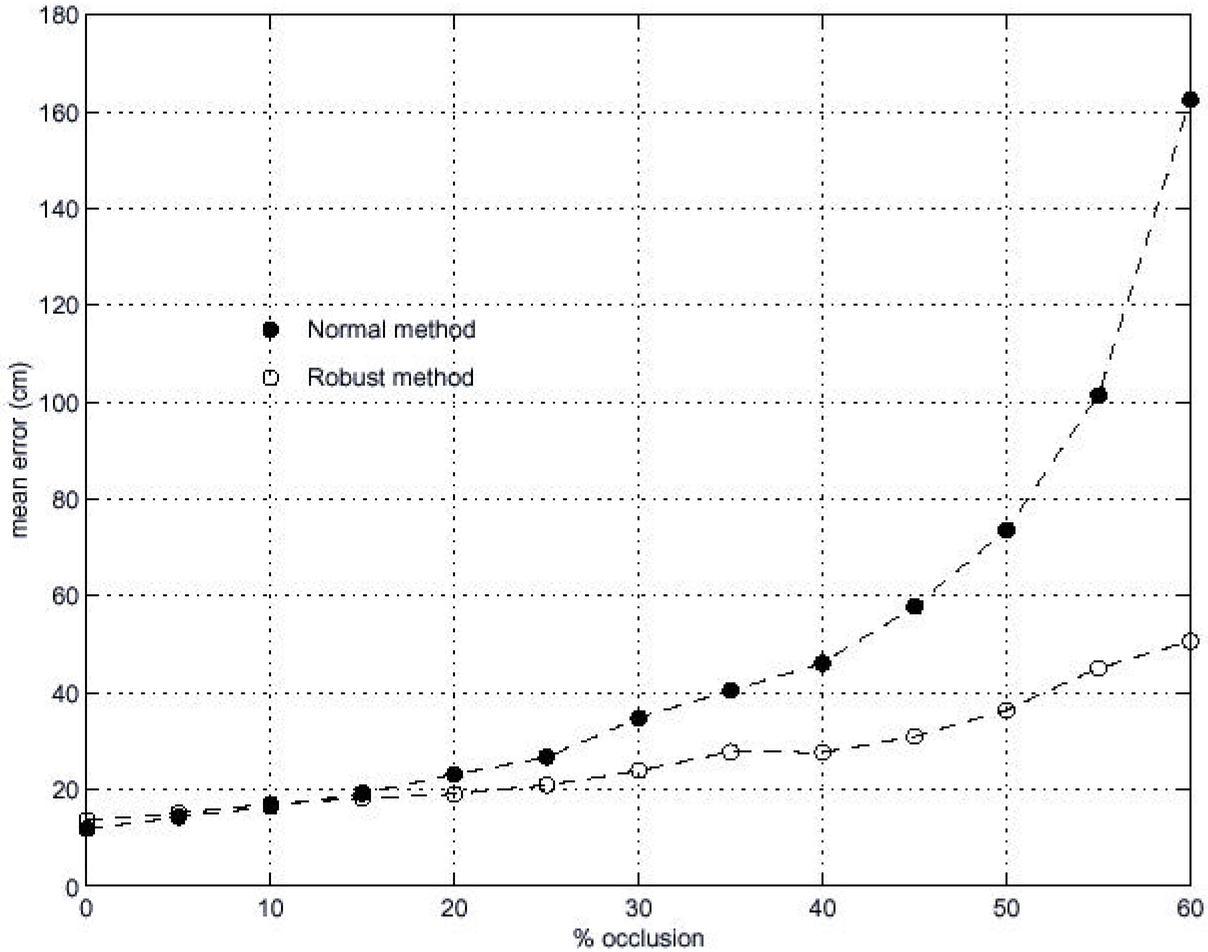
Standard approach



Robust approach

Mean error of localisation

- **Mean error of localisation with respect to % of occlusion**



Illumination insensitive recognition

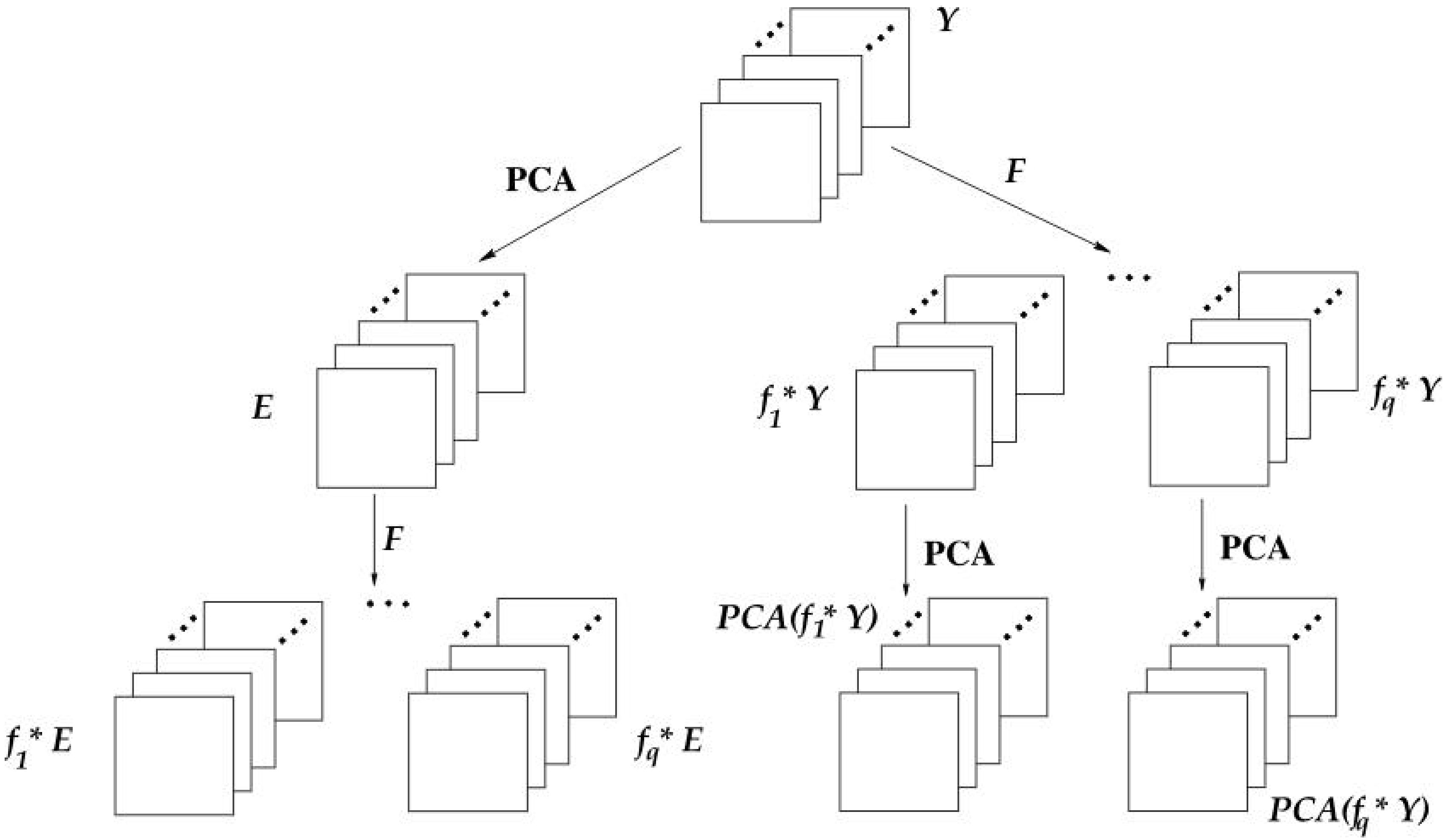
- Recognition of objects under varying illumination
  - **global illumination changes**
  - **highlights**
  - **shadows**
- Dramatic effects of illumination on objects appearance
- Training set under a single ambient illumination

Illumination insensitive recognition

Our Approach

- *Global* eigenspace representation
- *Local* gradient based filters
- Efficient combination of global and local representations
- Robust coefficient recovery in eigenspaces

Eigenspaces and filtering



Filtered eigenspaces

$$y_{r_i} = \sum_{j=1}^{n} q_j e_{j r_i} \quad 1 \leq i \leq k$$

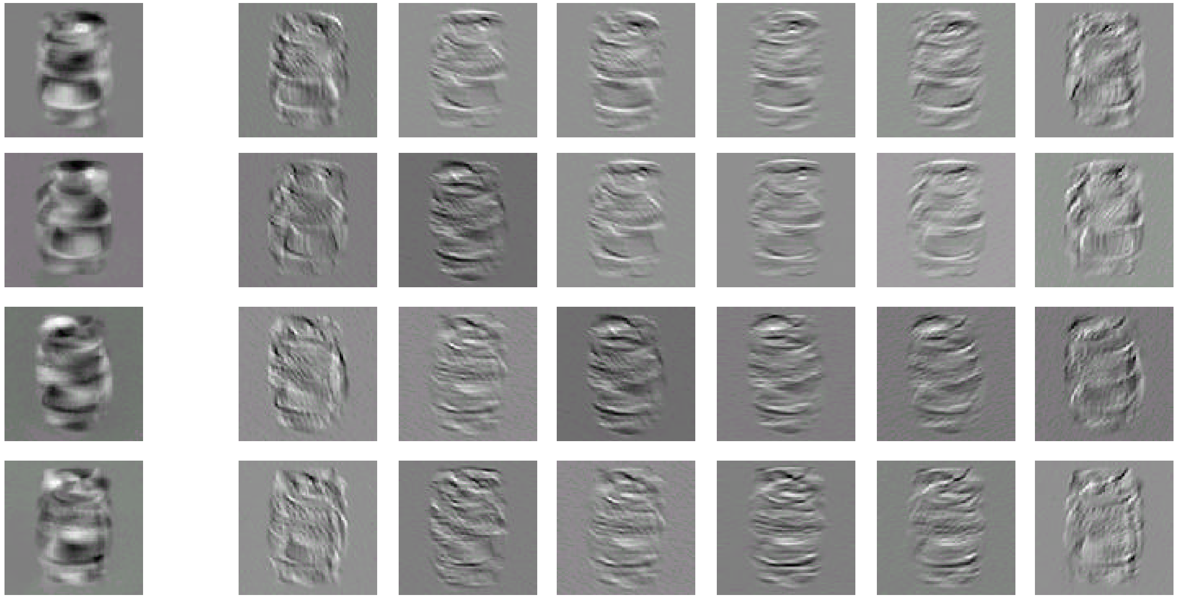
$$(f * x)(r) = \sum_{i=1}^{p} q_j (f * e_i)(r)$$

Global illumination

↓

Gradient-based filters

Steerable filters [Simoncelli]

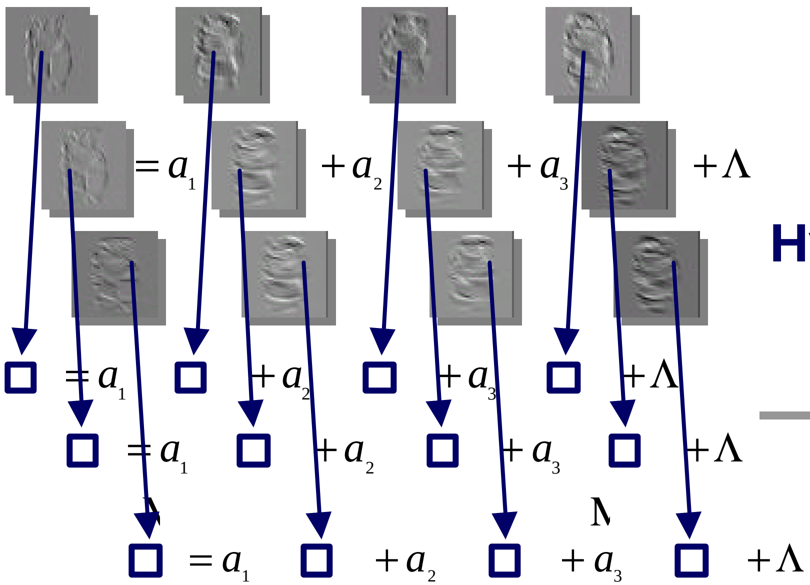


Robust coefficient recovery

Highlights and shadows

Robust coefficient recovery

Robust solution of linear equations

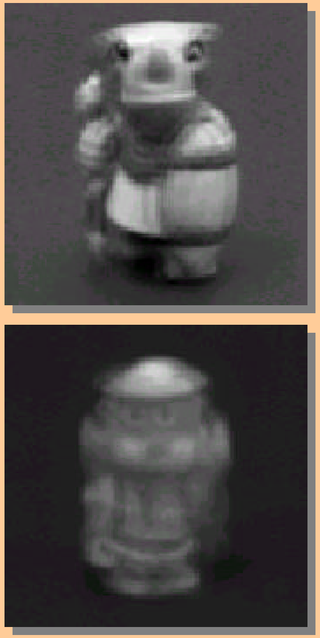


Hypothesize
&
Select

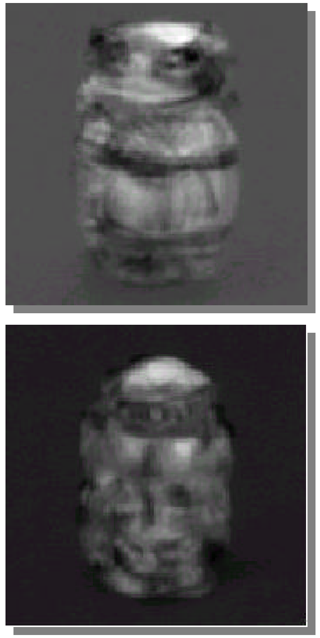
Experimental results

Test images

Our approach



Standard method



→ Demo

Experimental results

Robust filtered method - all eigenvectors used

obj.	1	2	3	4	5	%	ang.
1	360	0	0	0	0	100.0	5.25
2	0	308	16	0	0	95.1	10.55
3	0	0	504	0	0	100.0	1.05
4	19	4	3	332	2	92.2	3.37
5	15	2	17	0	578	94.4	3.34
avg.						96.4	4.19

Standard method - all eigenvectors used

obj.	1	2	3	4	5	%	ang.
1	141	0	14	26	179	39.2	10.50
2	0	254	62	5	3	78.4	18.90
3	0	4	317	0	183	62.9	3.47
4	23	6	38	249	44	69.2	7.11
5	3	1	51	0	557	91.0	6.82
avg.						70.3	8.53

Research issues

- Comparative studies (e.g., LDA versus PCA, PCA versus ICA)
- Robust learning of other representations (e.g. LDA, CCA)
- Integration of robust learning with modular eigenspaces
- Local versus Global subspace represenations
- Combination of subspace representations in a hierarchical framework

Further readings

- Recognizing objects by their appearance using eigenimages (SOFSEM 2000, LNCS 1963)
- Robust recognition using eigenimages (CVIU 2000, Special Issue on Robust Methods in CV)
- Illumination insensitive eigenspaces (ICCV 2001)
- Mobile robot localization under varying illumination (ICPR 2002)
- Eigenspace of spinning images (OMNI 2000, ICPR 2000, ICAR 2001)
- Incremental building of eigenspaces (ICRA 2002, ICPR 2002, CogVis 2002)
- Multiple eigenspaces (Pattern Recognition 2002)
- Robust building of eigenspaces (ECCV 2002)

- Special issue of Pattern Recognition on Kernel and Subspace Methods in Computer Vision (Guest Editors A. Leonardis and H. Bischof), to appear in 2003.