Robust estimation of subspace coefficients

Danijel Skočaj



University of Ljubljana

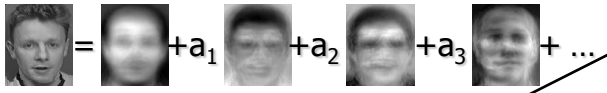
Faculty of Computer and Information Science

Udine; 5, 12 December 2014

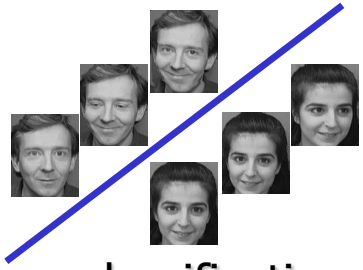
Background

Computer Vision

Subspace methods



representation



classification
regression

Objects often occluded or
appear on different backgrounds



Reconstructive

PCA, ICA, NMF

learning

De la Torre,
Skočaj, ...

recognition

Leonardis, Bischof,
Black, ...

Should have ability to work on
subsets of points in the image

Discriminative

LDA, SVM, CCA

learning

Croux, He,
Hawkins, ...

recognition

Fidler,
Skočaj, ...

Subspace methods

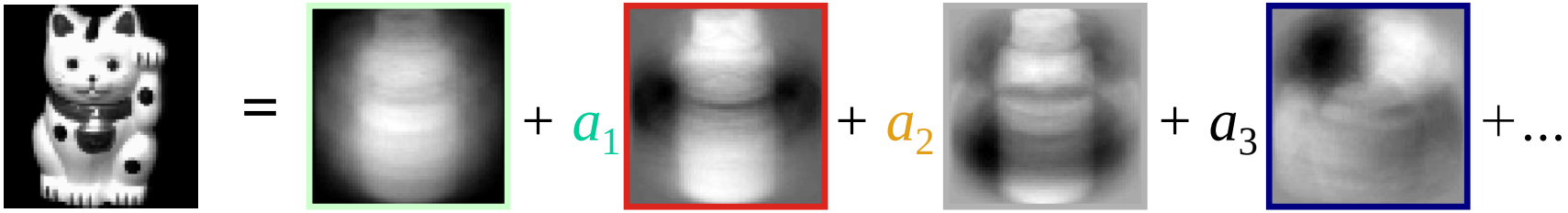
- PCA – Principal Component Analysis
  - Generative, unsupervised method.
  - Minimizes squared reconstruction error.
  - Maximizes variance of projected input vectors.
- LDA – Linear Discriminant Analysis
  - Non-generative, supervised, discriminant method.
  - Maximizes distance of projections between classes.
  - Minimizes distance of projections within a class.
- CCA – Canonical Correlation Analysis
  - Non-generative, supervised method.
  - Maximizes correlation between projections.
  - Enables regression.
- ICA, NMF, TRP, SVM,…

Principal Component Analysis (PCA)

- PCA finds principal vectors $\mathbf{u}_i,\ i = 1 \ldots k$ and vectors of princ. comp. $\mathbf{a}_i \in \mathbb{R}^k, i = 1 \ldots N$, which minimize the squared reconstruction error

$$e = \sum_{i=1}^{M} \sum_{j=1}^{N} \left(\hat{x}_{ij} - \sum_{l=1}^{k} u_{il} a_{lj} \right)^2$$

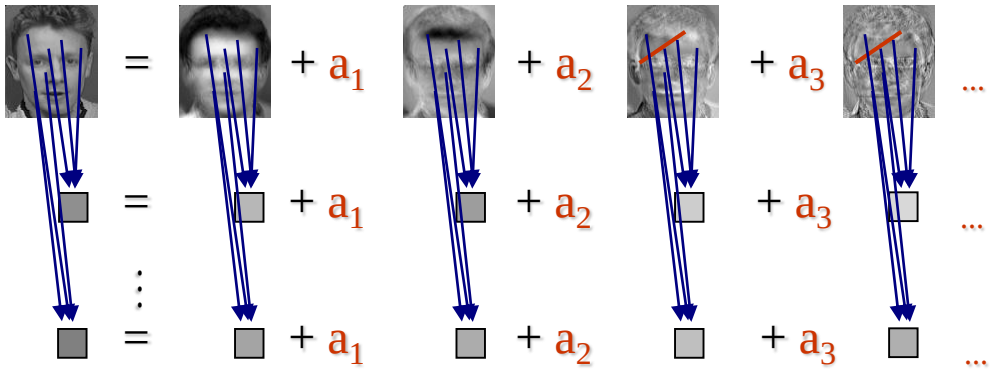
- Projection: $\mathbf{a} = \mathbf{U}^\top (\mathbf{x} - \boldsymbol{\mu})$
- Reconstruction: $\mathbf{x} \approx \mathbf{U}\mathbf{a} + \boldsymbol{\mu}$



Estimation of coefficients

- Instead of projection solve a system of linear equations!

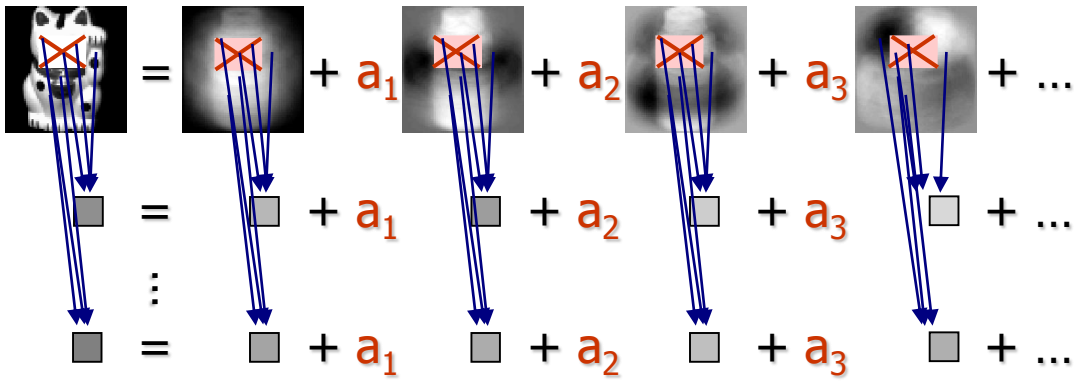
$$x^1 = a_1 u_1^1 + a_2 u_2^1 + \cdots + a_k u_k^1$$
$$x^2 = a_1 u_1^2 + a_2 u_2^2 + \cdots + a_k u_k^2$$
$$\vdots$$
$$x^m = a_1 u_1^m + a_2 u_2^m + \cdots + a_k u_k^m$$



Robust estimation of PCA coefficients

Main idea: Instead of using the standard approach

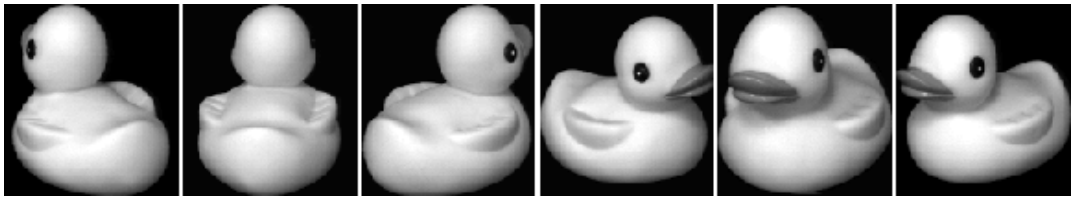
- Select a subset of pixels → linear system of equations
- Find a robust solution of equations.
- Perform multiple hypotheses.



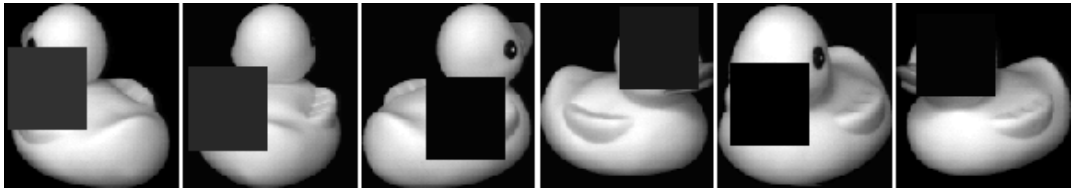
- Hypothesize-and-test paradigm
- Competing hypotheses are subject to a selection procedure based on the MDL principle.

Robust estimation of PC - Example

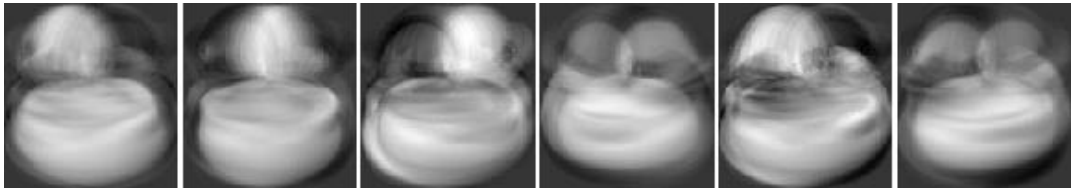
training images



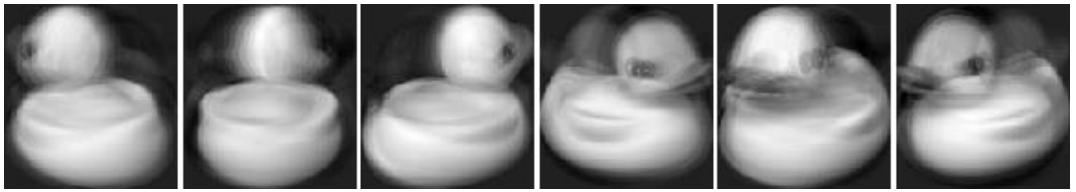
test images



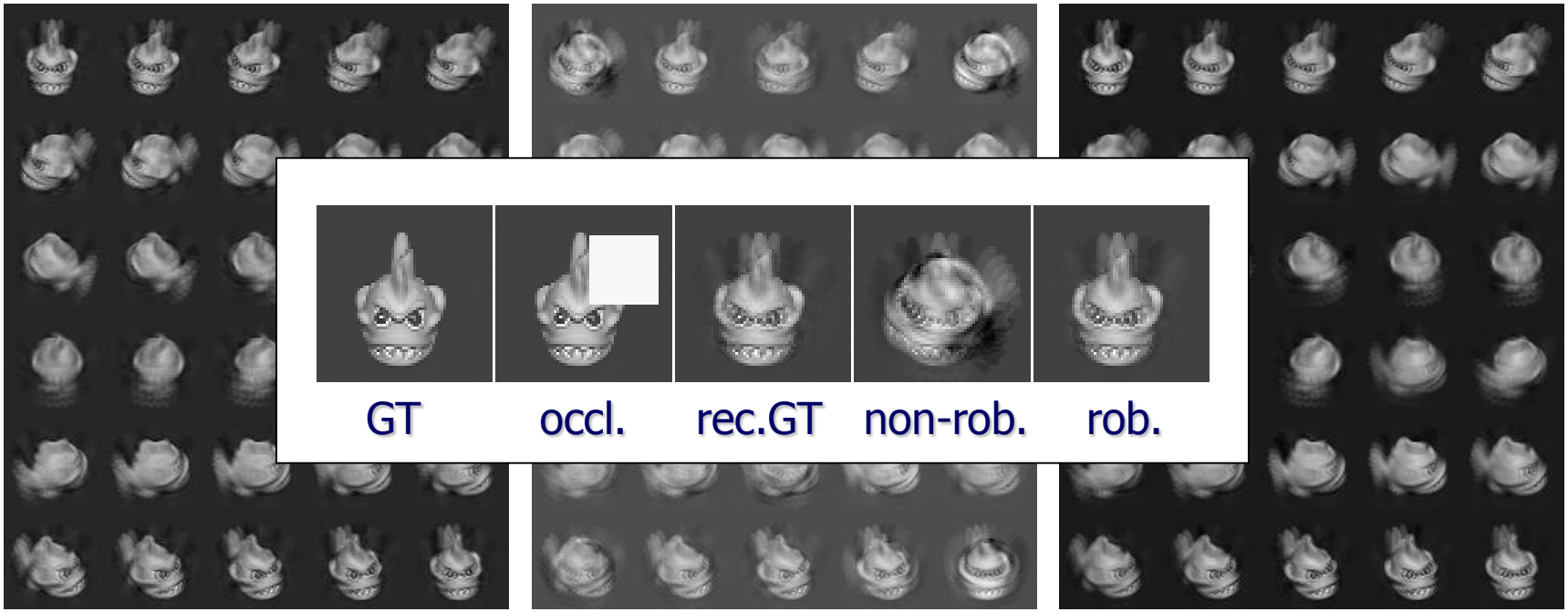
non-robust reconstruction



robust reconstruction



Robust estimation of PC - Example



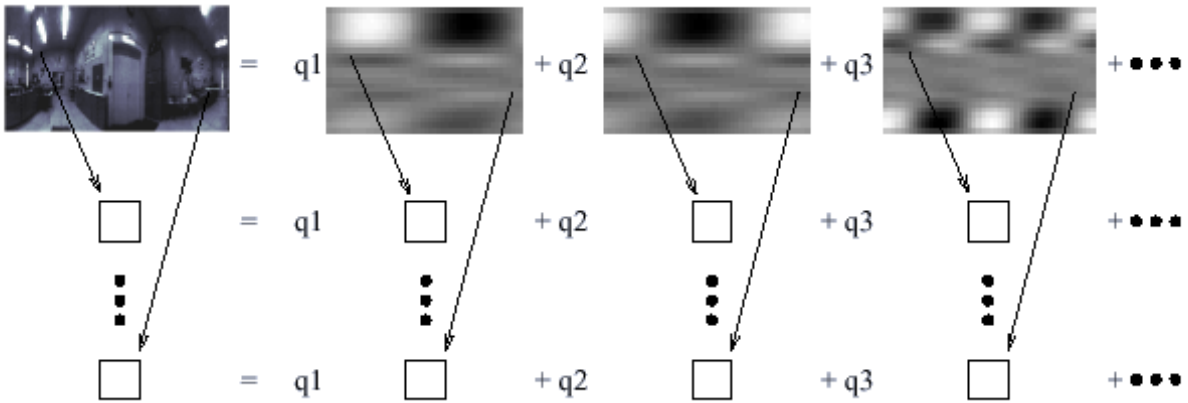
reconstructed non-occluded images

non-robustly reconstructed occluded images

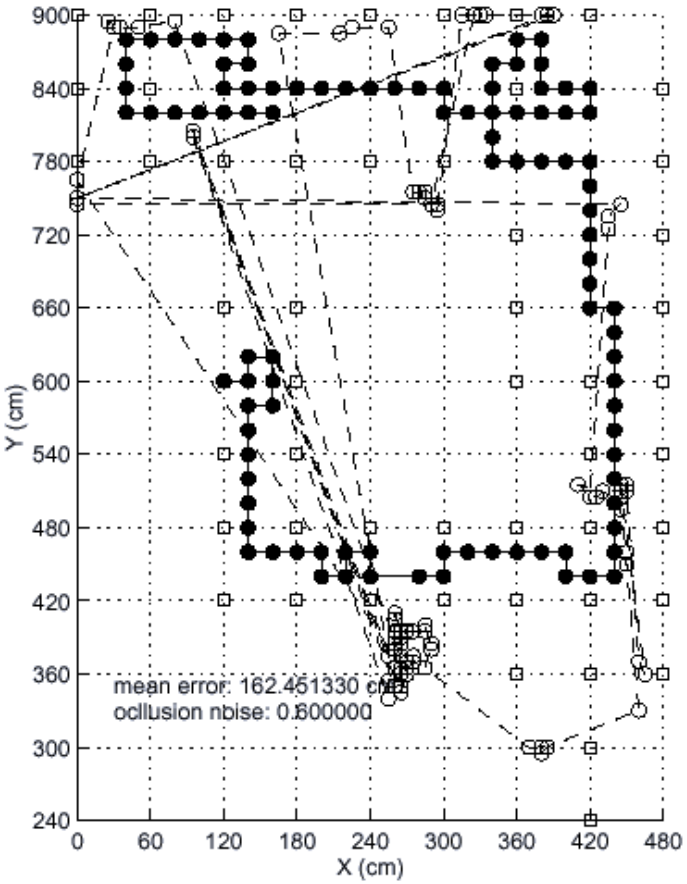
robustly reconstructed occluded images

Robust localisation under occlusions

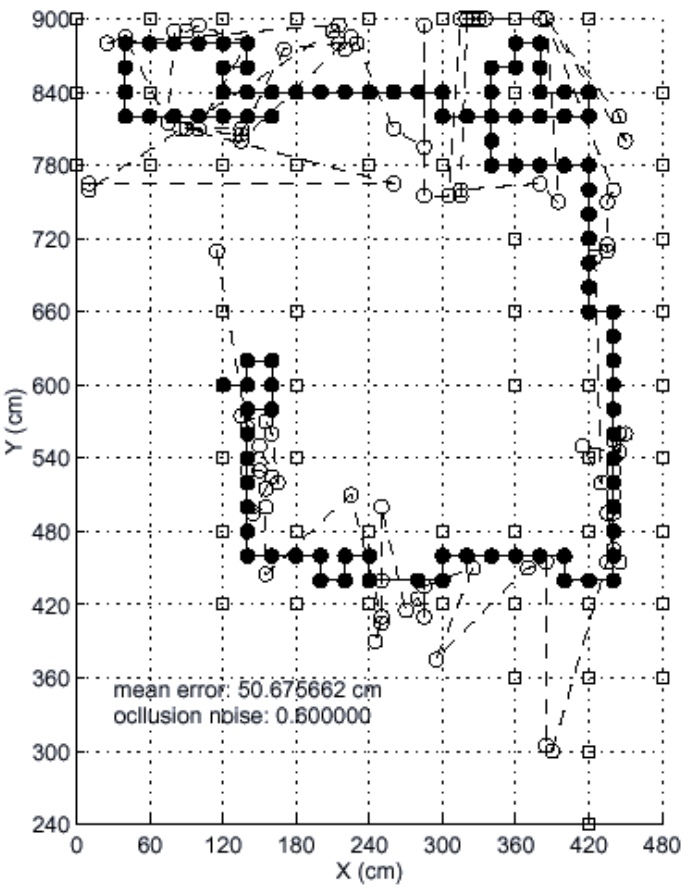




Robust localisation at 60% occlusion



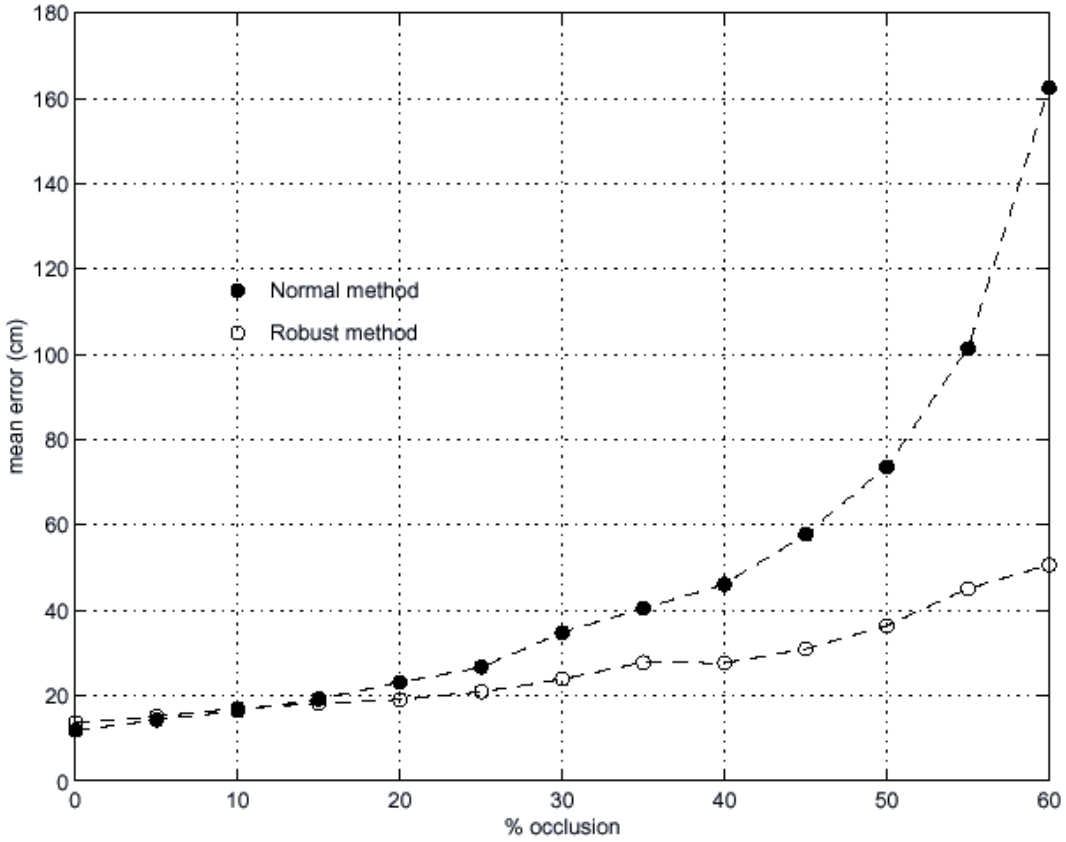
Standard approach



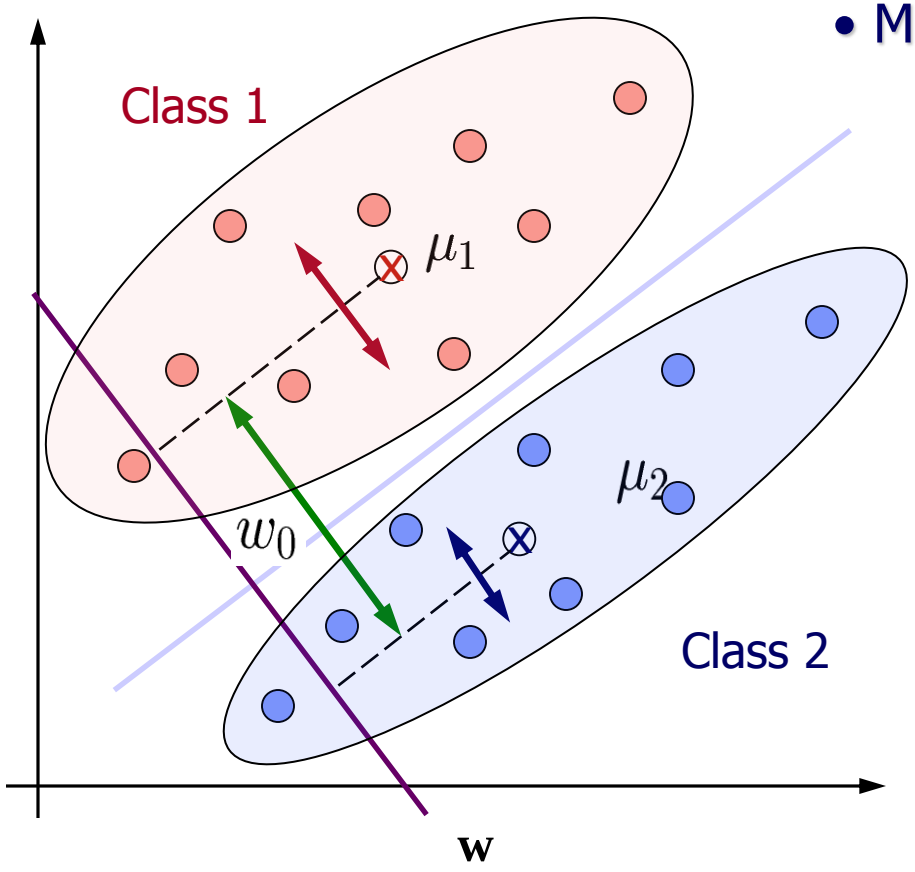
Robust approach

Mean error of localisation

- Mean error of localisation with respect to % of occlusion



Linear Discriminant Analysis (LDA)



- Maximize distance between classes.
- Minimize distance within a class.

Criterion function:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}}$$

S_b … between scatter matrix

S_w … within scatter matrix

w is a solution of a generalized eigenvalue problem:

$$S_b \mathbf{w} = \lambda S_w \mathbf{w}$$

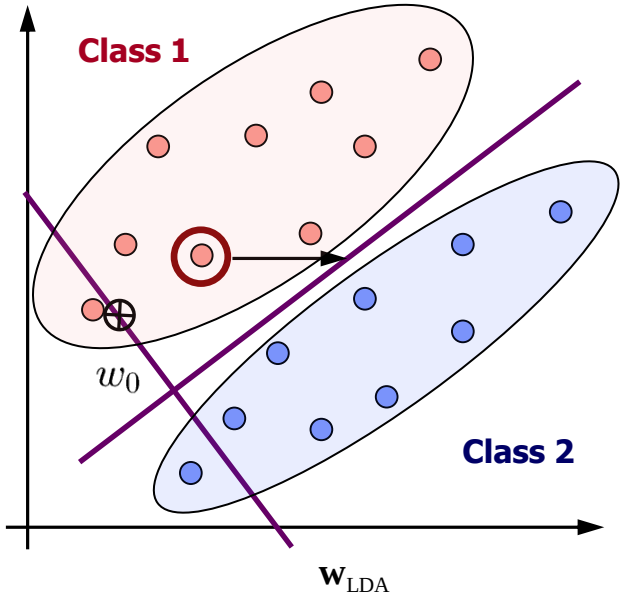
Classification function:

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 \underset{\text{Class 2}}{\overset{\text{Class 1}}{\gtrless}} 0$$

Outliers and occlusions











$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0 \underset{\text{Class 2}}{\overset{\text{Class 1}}{\gtrless}} 0$$

- LDA depends on the inner product of ***w*** and a data vector, thus highly sensitive to outliers and occlusions.

Robustisation

- Reconstructive methods (PCA)
  - Good approximation of data
  - General, task-independent
  - Not very compact, redundant representations
  - ⇨ Enable reconstruction and detection of outliers
- Discriminative methods (LDA)
  - Supervised, task-dependent
  - Compact representations
  - No reconstruction (low dimensional projections)
  - ⇨ Do not enable reconstruction and detection of outliers

⇨ Combine reconstructive model and discriminative classifier

Augmenting the PCA basis

- Embed LDA learning and classification into PCA framework

$$\mathbf{X} \in \mathbf{R}^{m \times n} : PCA(\mathbf{X}) \rightarrow \mathbf{U} = \left[\begin{array}{cc} \mathbf{U}_{\underline{k}} & \mathbf{U}_{\underline{(n-k)}} \end{array} \right] \in \mathbf{R}^{m \times n}$$

$$\mathbf{A} = \left[\begin{array}{c} \mathbf{A}_{\underline{k}} \\ \mathbf{A}_{\underline{(n-k)}} \end{array} \right] \in \mathbf{R}^{n \times n}$$

$$LDA(\mathbf{A}) \rightarrow \mathbf{V} = \left[\begin{array}{c} \mathbf{V}_{\underline{k}} \\ \mathbf{V}_{\underline{(n-k)}} \end{array} \right] \in \mathbf{R}^{n \times (c-1)}$$

- First k << n principal vectors
  - Contain most visual variability
  - No guarantee that it contains most discriminative information

Augmenting the PCA basis

- Augment PCA subspace with c-1 additional basis vectors → keeps discriminative information

$$\widetilde{\mathbf{V}}_{\underline{(n-k)}} = \mathbf{V}_{\underline{(n-k)}}(\mathbf{V}_{\underline{(n-k)}}^{\top}\mathbf{V}_{\underline{(n-k)}})^{-1/2} \in \mathbf{R}^{(n-k)\times(c-1)}$$

$$\widehat{\mathbf{U}} = \left[\begin{array}{cc} \mathbf{U}_{\underline{k}} & \mathbf{U}_{\underline{(n-k)}}\widetilde{\mathbf{V}}_{\underline{(n-k)}} \end{array}\right] \in \mathbf{R}^{m\times(k+c-1)}$$

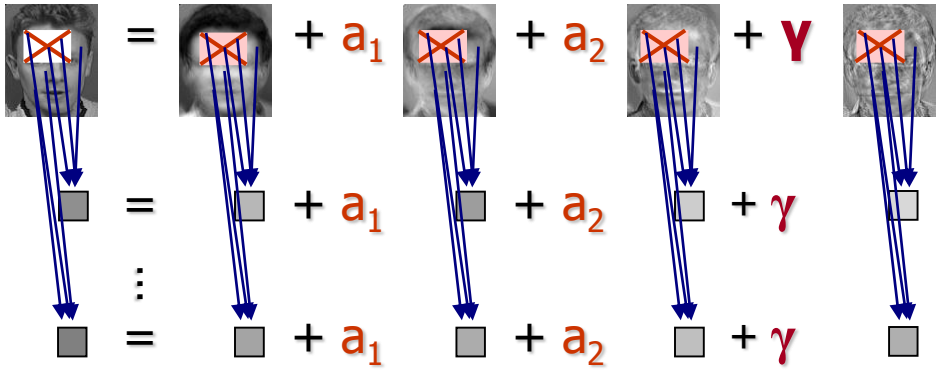
$$\widehat{\mathbf{A}} = \left[\begin{array}{c} \mathbf{A}_{\underline{k}} \\ \widetilde{\mathbf{V}}_{\underline{(n-k)}}^{\top}\mathbf{A}_{\underline{(n-k)}} \end{array}\right] \in \mathbf{R}^{(k+c-1)\times n}$$

$$\widehat{\mathbf{V}} = \left[\begin{array}{c} \mathbf{V}_{\underline{k}} \\ \widetilde{\mathbf{V}}_{\underline{(n-k)}}^{\top}\mathbf{V}_{\underline{(n-k)}} \end{array}\right] \in \mathbf{R}^{(k+c-1)\times(c-1)}$$

- → APCA subspace

Robust estimation of LDA coefficients

- Now we can use robust subsampling-based estimation of coefficients in augmented PCA basis:
  - Select a subset of pixels.
  - Find a robust solution of linear equations.
  - Perform multiple hypotheses.



  - Select the best hypothesis.
- From the robustly obtained coefficients, we can obtain robust LDA coeficients using the standard projection.

LDA classification - Algorithm

- Learning stage:
  - Perform PCA on input images.
  - Apply LDA on full vectors of principal components.
  - Keep k principal vectors and append $c-1$ additional basis vectors to form an augmented PCA basis.
  - Extend the obtained feature vectors with new coefficients.
- Classification stage
  - Calculate the coefficients of a new image in the augmented PCA basis.
  - From the obtained coefficients calculate the LDA coefficients.

The results are equivalent to the results of LDA on the full vectors of principal components!

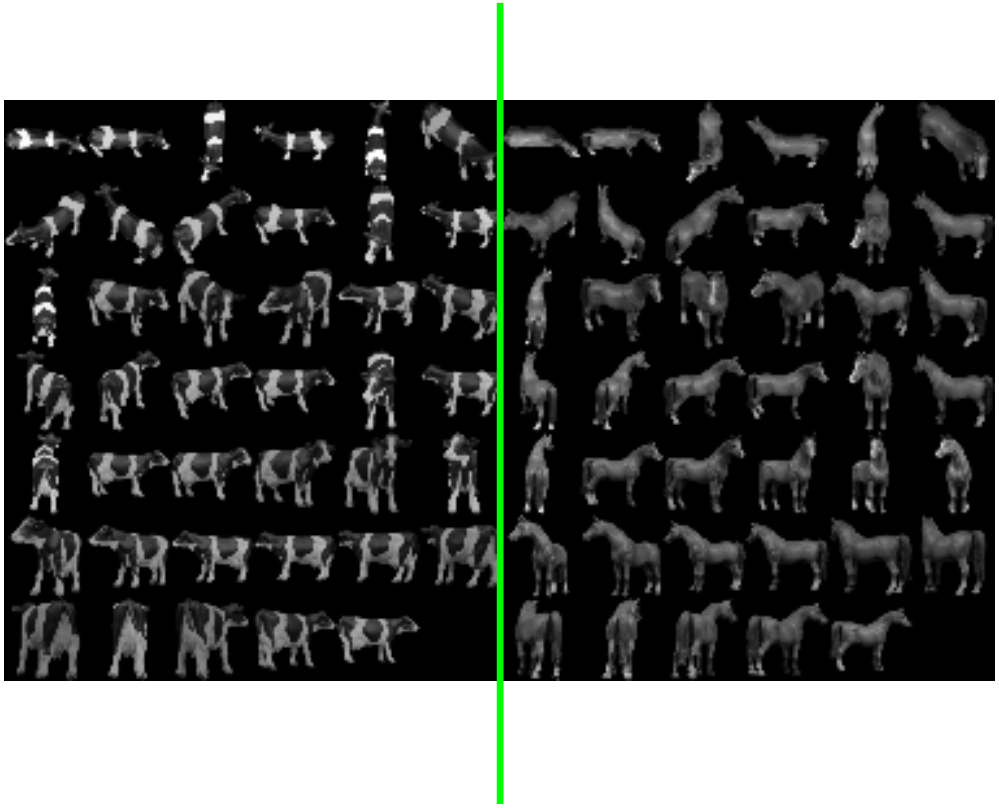
Robust LDA classification - Algorithm

- Learning stage:
  - Perform PCA on input images.
  - Apply LDA on full vectors of principal components.
  - Keep k principal vectors and append $c-1$ additional basis vectors to form an augmented PCA basis.
  - Extend the obtained feature vectors with new coefficients.
- Classification stage
  - Calculate the coefficients of a new image in the augmented PCA basis in a robust manner.
  - From the obtained coefficients calculate the LDA coefficients.

Since the coefficients in the augmented PCA basis are robustly estimated, LDA coefficients are robust as well!

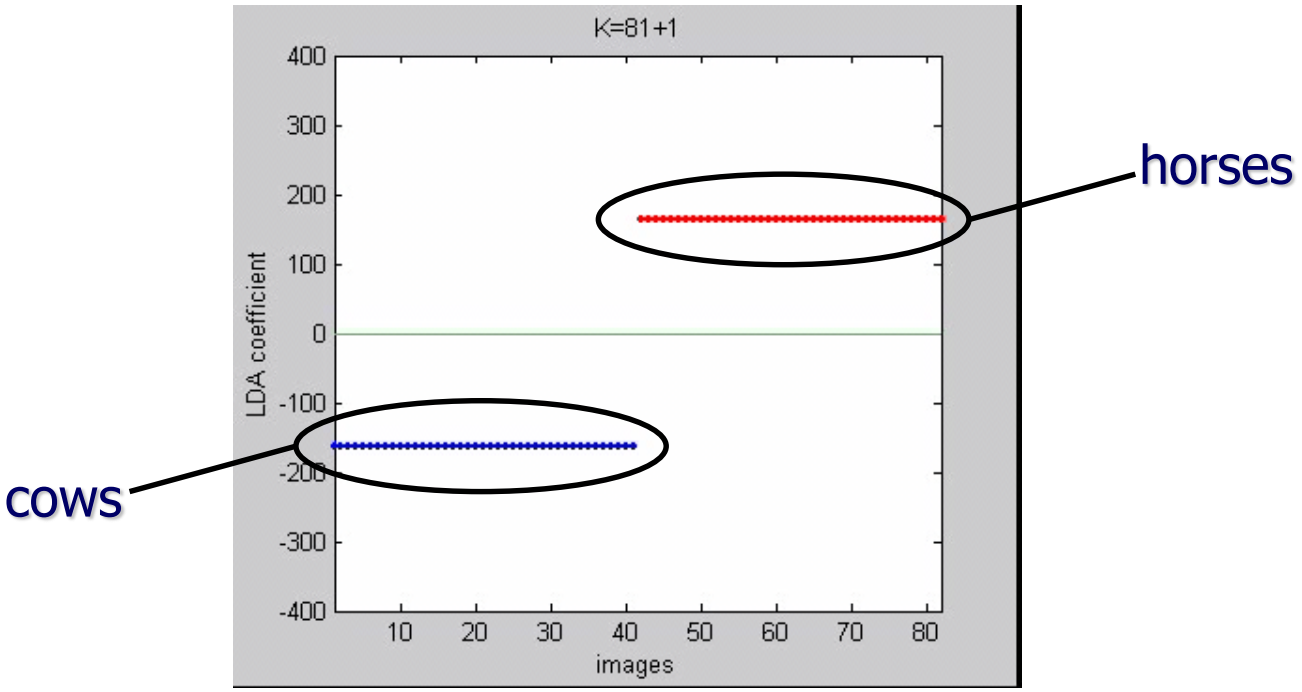
LDA - Example

Find a classification vector, which separates cows and horses.



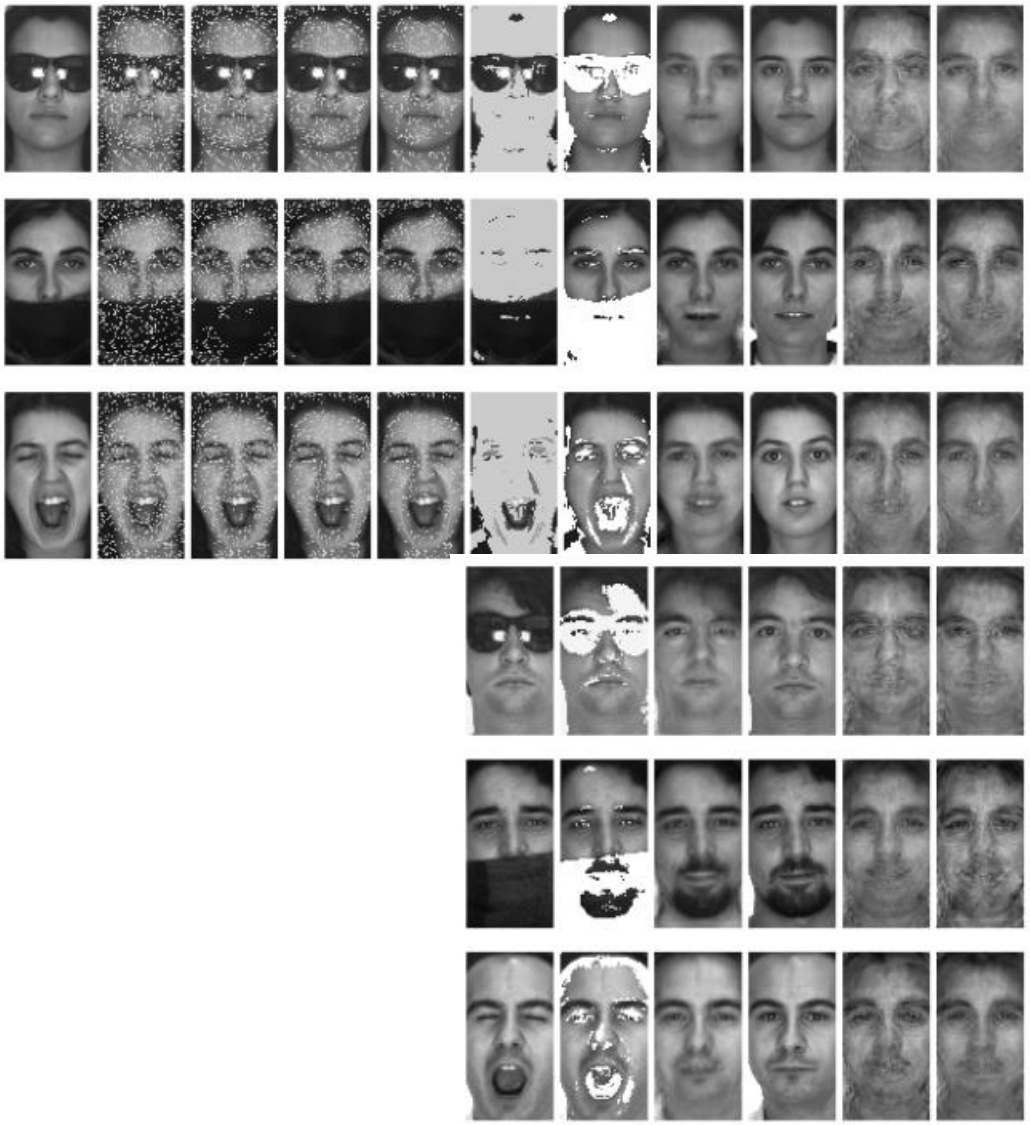
LDA - Example

LDA on K<N-1 principal components –
some discriminative information has been lost!



Estimation of coefficients in augmented PCA basis -
entire discriminative information has been preserved!

Robust face recognition



AR database



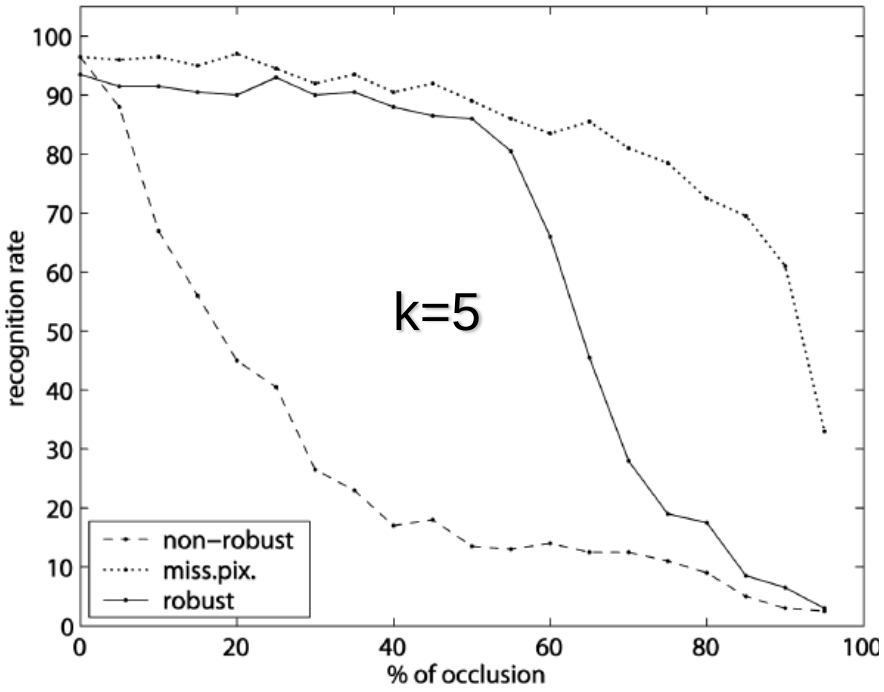
Robust face recognition

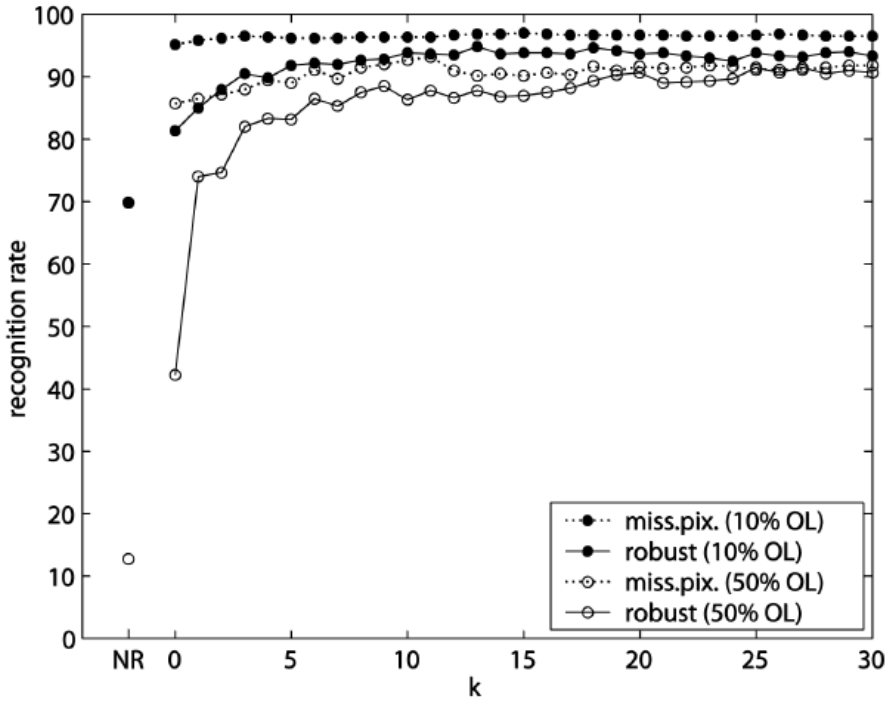




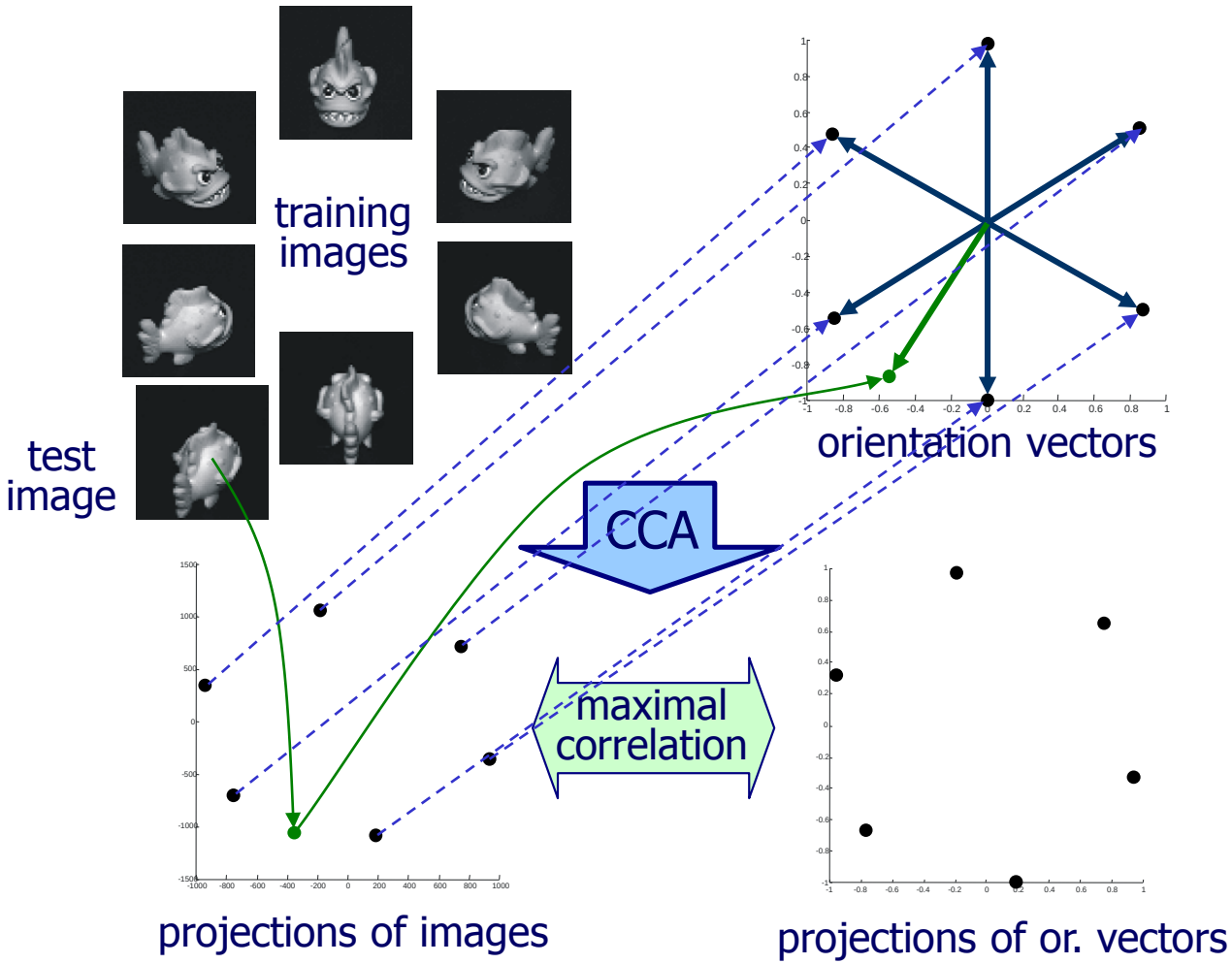


ORL database

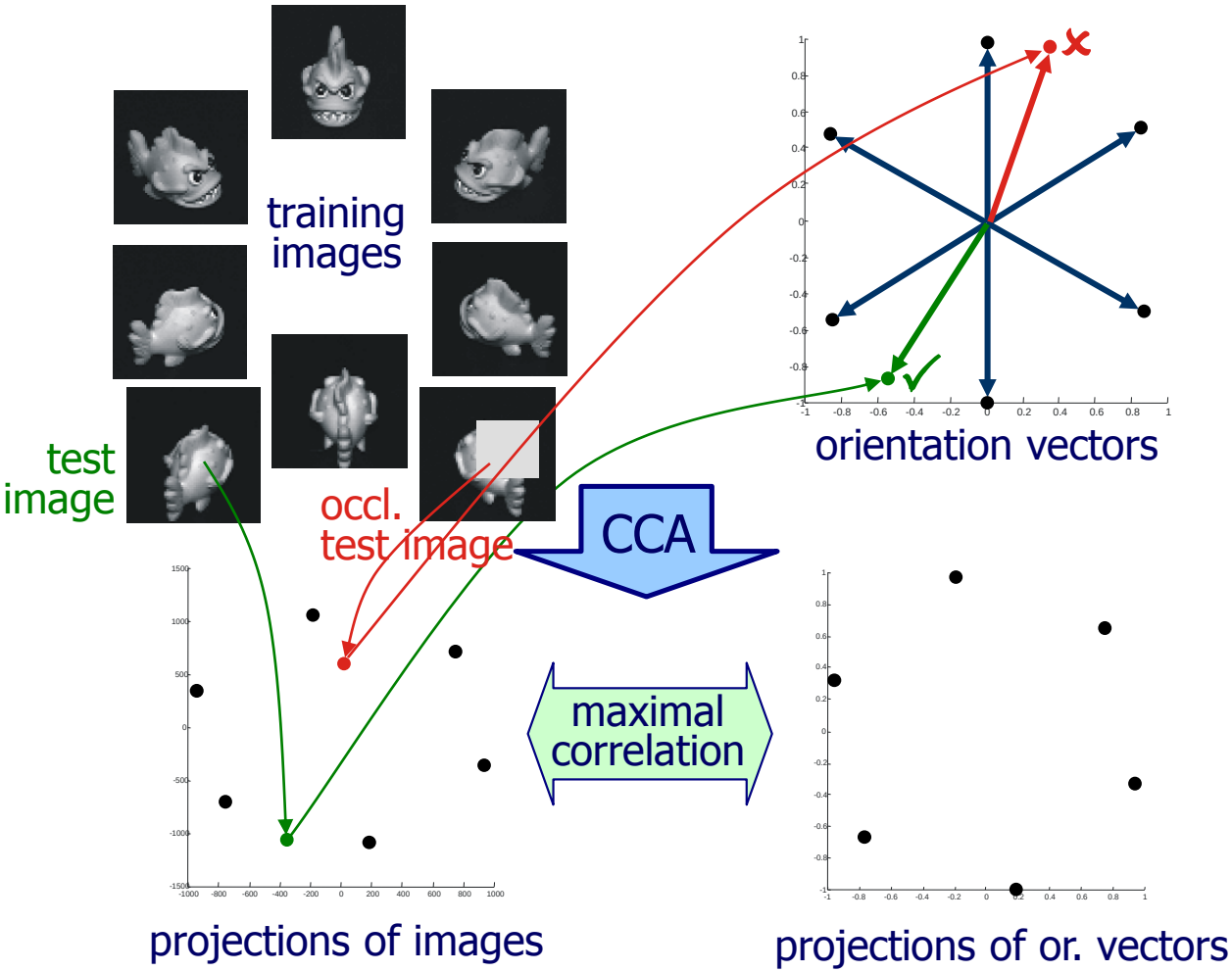




Canonical Correlation Analysis (CCA)



CCA is non-robust

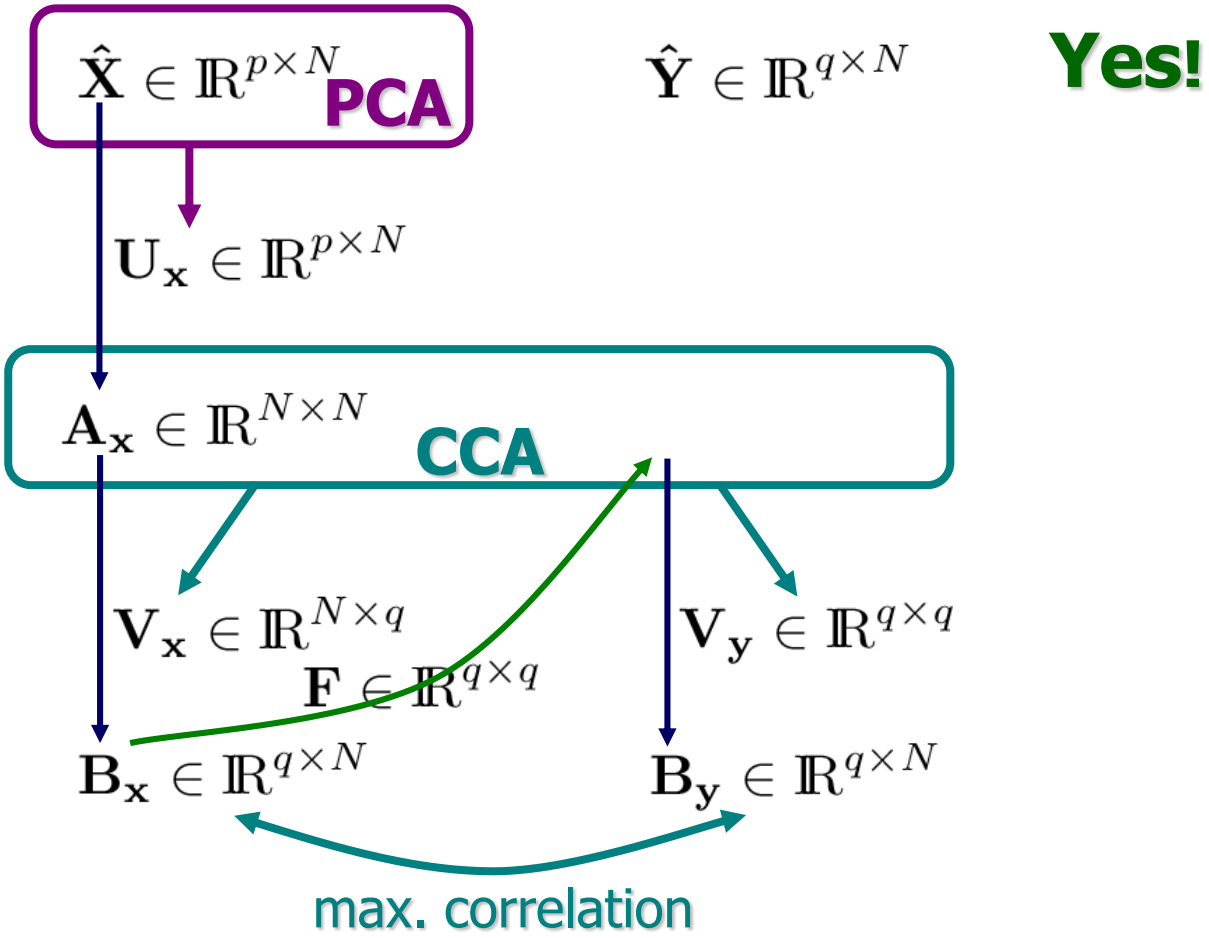


Robustifying CCA

- Robust procedure has to
  - detect outliers
  - estimate coefficients from inliers only.
- CCA yields only a few canonical correlation vectors ⇒ the images can not be reconstructed ⇒ the momentary input can not be compared to the training data ⇒ the outliers can not be detected.
- PCA is reconstructive method
  ⇒ robust approach exists for PCA.
- Embed CCA in PCA space!

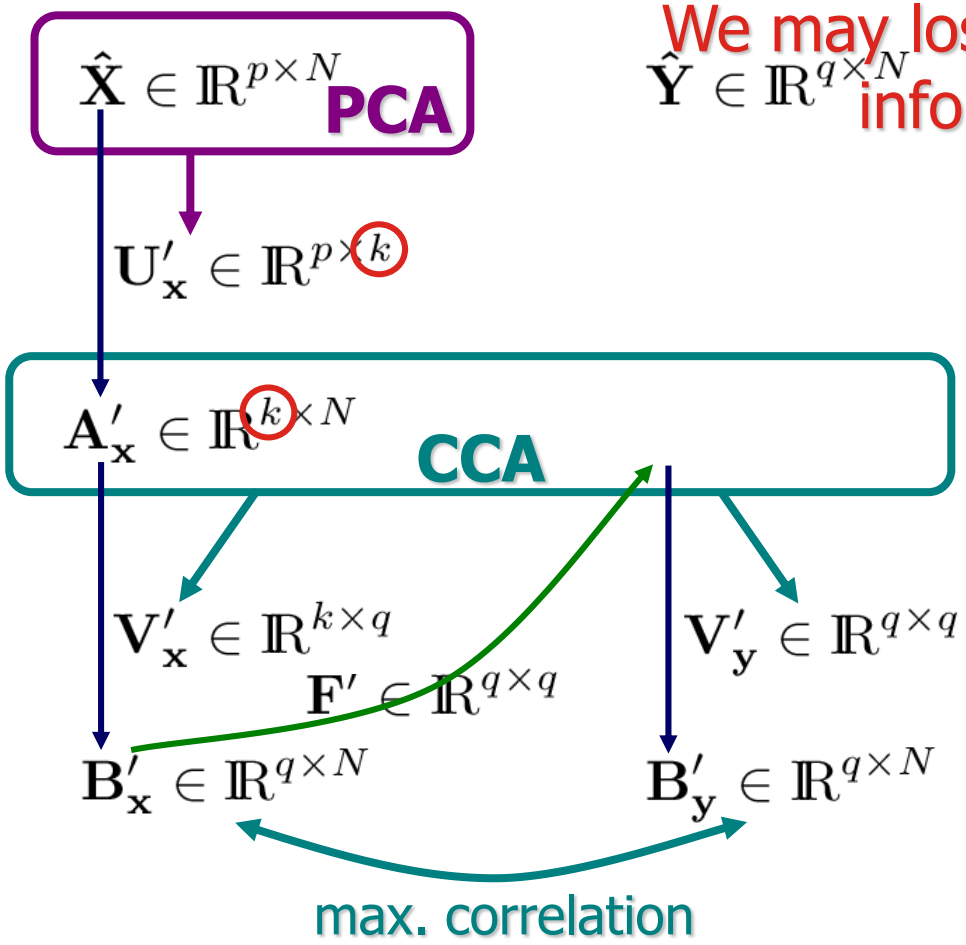
CCA in complete PCA subspace

Can we perform CCA in complete PCA space?



CCA in truncated PCA subspace

Can we perform CCA in truncated PCA space?

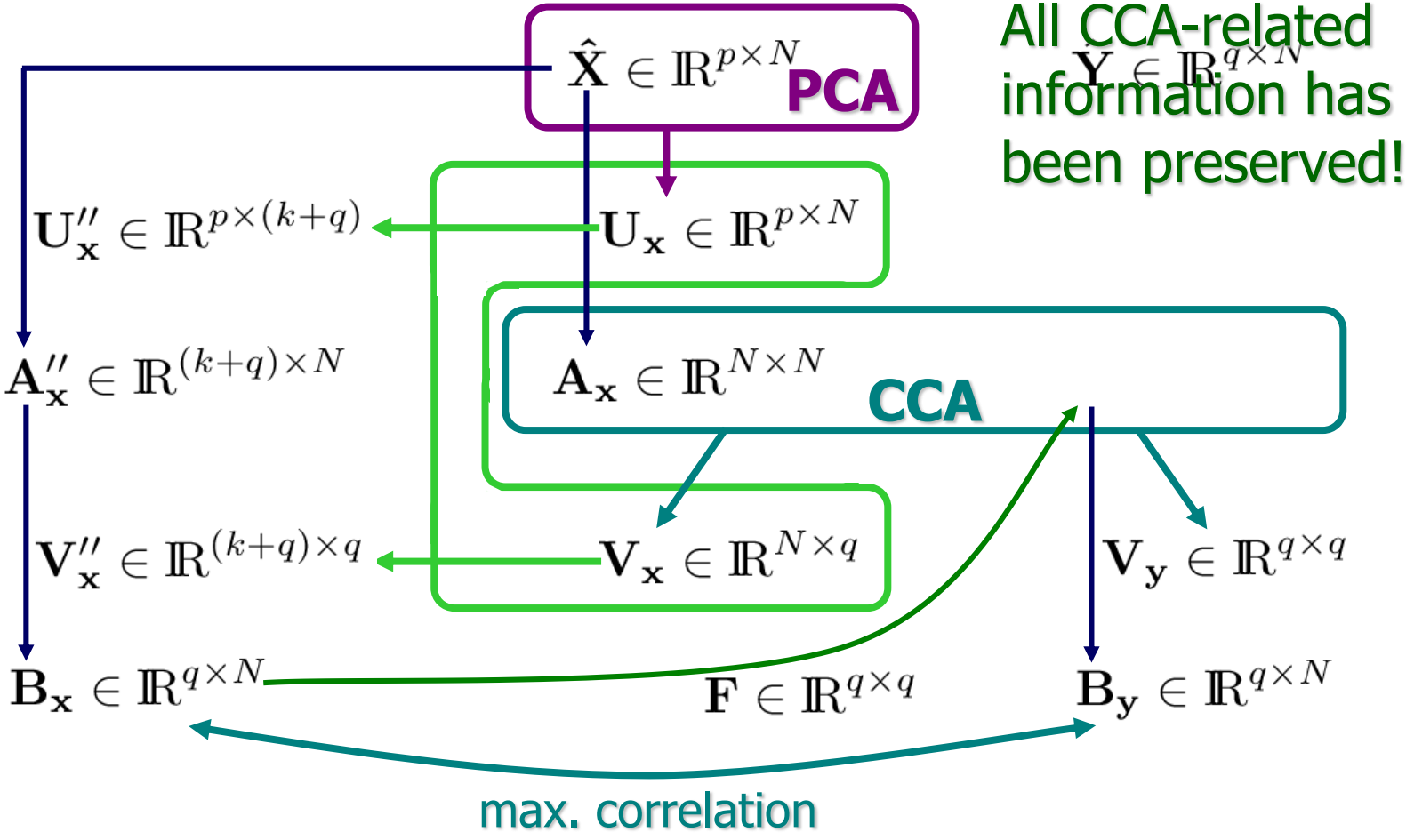


We may lose significant information!

$\hat{\mathbf{Y}} \in \mathbb{R}^{q \times N}$

Augmented principal subspace

Extend the truncated PCA subspace!



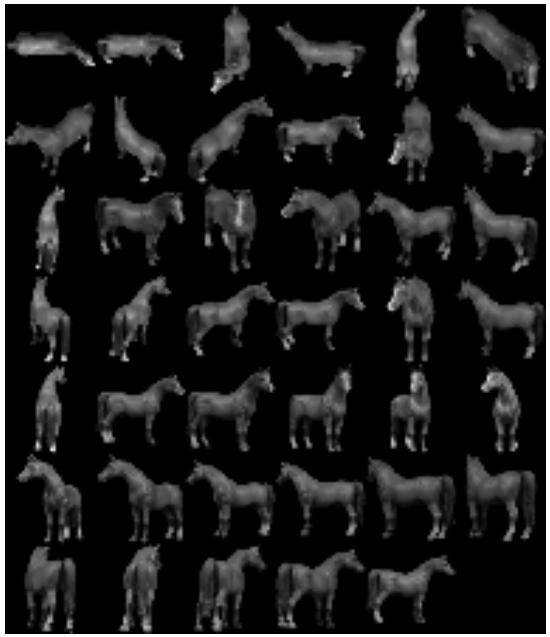
Robust estimation of CCA coefficients

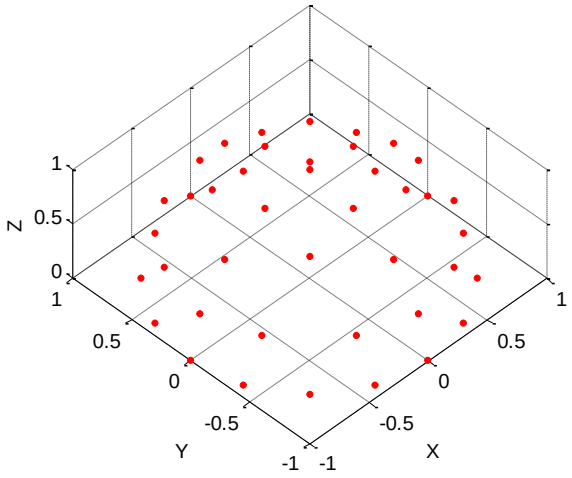
Robust estimation of CCA coefficients:

- In the learning stage:
  - Perform PCA on input images.
  - Perform CCA on full vectors of principal components.
  - Keep K principal vectors and components and augment the PCA basis with c additional basis vectors, which preserve entire regression-relevant information.
- In the regression stage:
  - Estimate coefficients in $(k+c)$-dimensional augmented PCA basis in a robust way.
  - Project the obtained coefficients onto the CCA vectors.

Robust CCA - Example

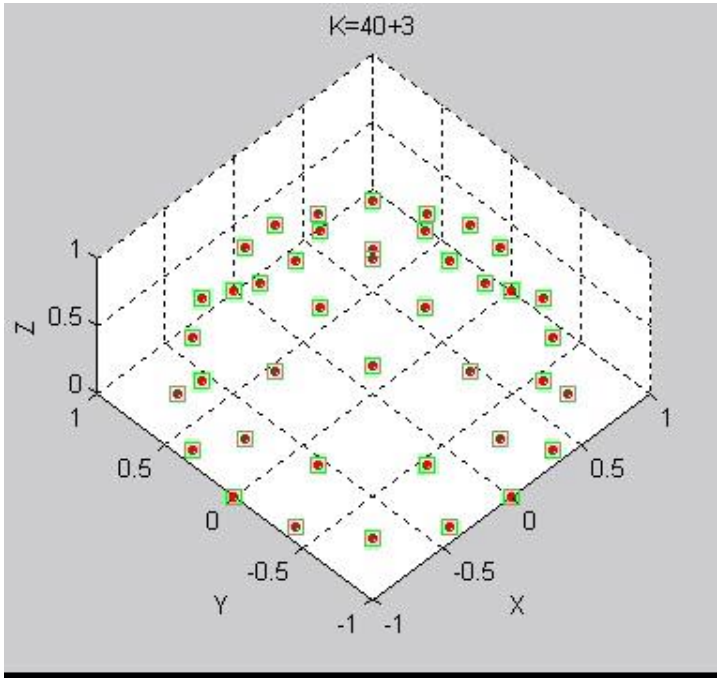
- Find three canonical correlation vectors, which yield 3-D coordinates of view-points.





Robust CCA - Example

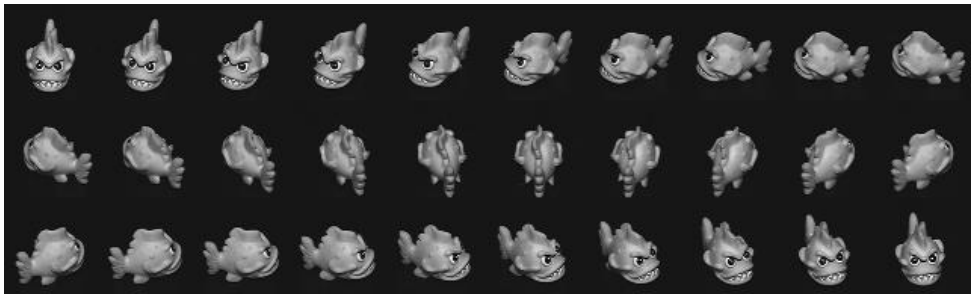
CCA on *K<N-1* principal components:
some regression-relevant information has been discarded!

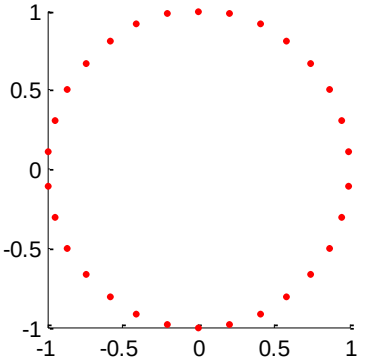


Estimation of coefficients in augmented PCA basis:
entire regression-relevant information has been preserved!

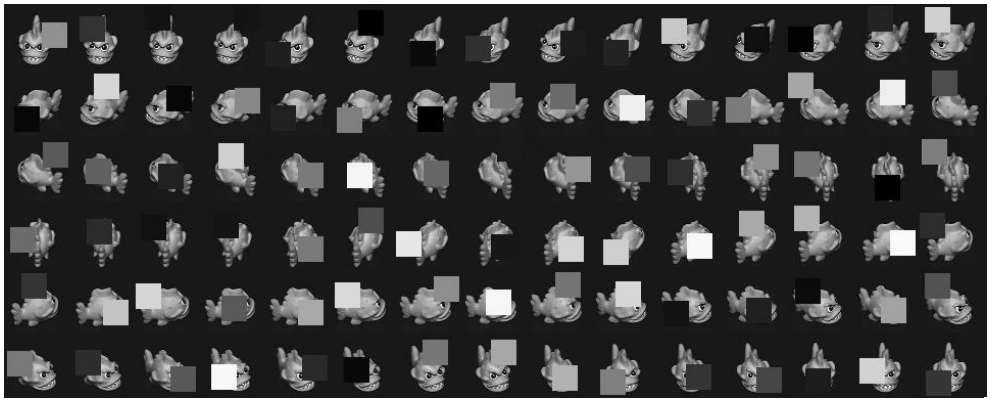
Robust CCA – Experimental results

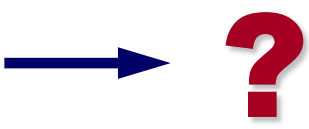
- Estimation of the rotation.
- Training sets:





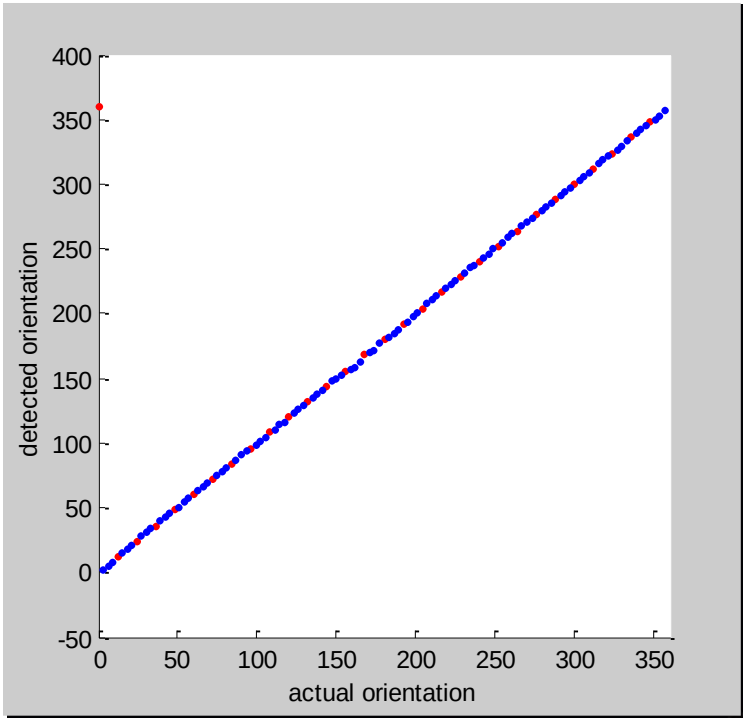
- Occluded test images:





Robust CCA – Experimental results

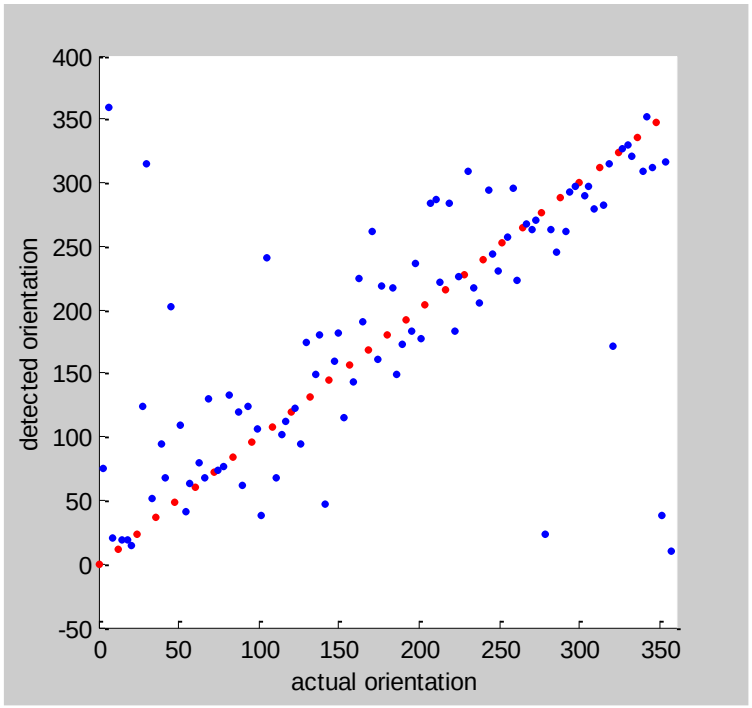
- Non-occluded test images



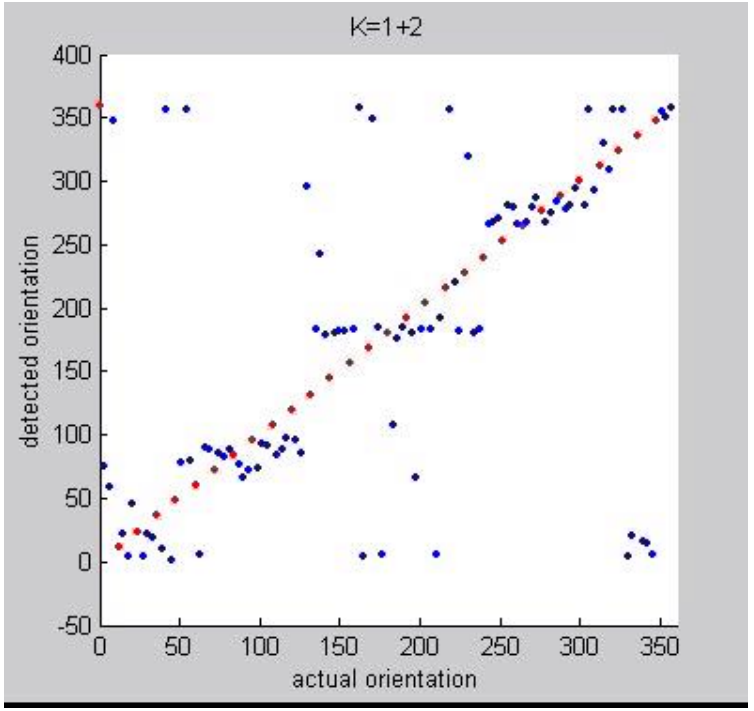
Using eigenvectors of PCA basis

Robust CCA – Experimental results

- Occluded test images



standard



robust

Literature

- A. Leonardis and H. Bischof, Robust Recognition Using Eigenimages, Computer Vision and Image Understanding, 2000
  - http://www.mmp.rwth-aachen.de/teaching/cvws08/additional/leonardis-robusteigen-cviu00.pdf
- S. Fidler, D. Skočaj and A. Leonardis, Combining Reconstructive and Discriminative Subspace Methods for Robust Classification and Regression by Subsampling, IEEE PAMI, 2006
  - http://vicos.fri.uni-lj.si/data/publications/fidlerPAMI06.pdf
- D. Skočaj, A. Leonardis and S. Fidler, Robust estimation of canonical correlation coefficients, OAGM 2004
  - http://vicos.fri.uni-lj.si/data/publications/skocajOAGM04a.pdf