Canonical Correlation Analysis

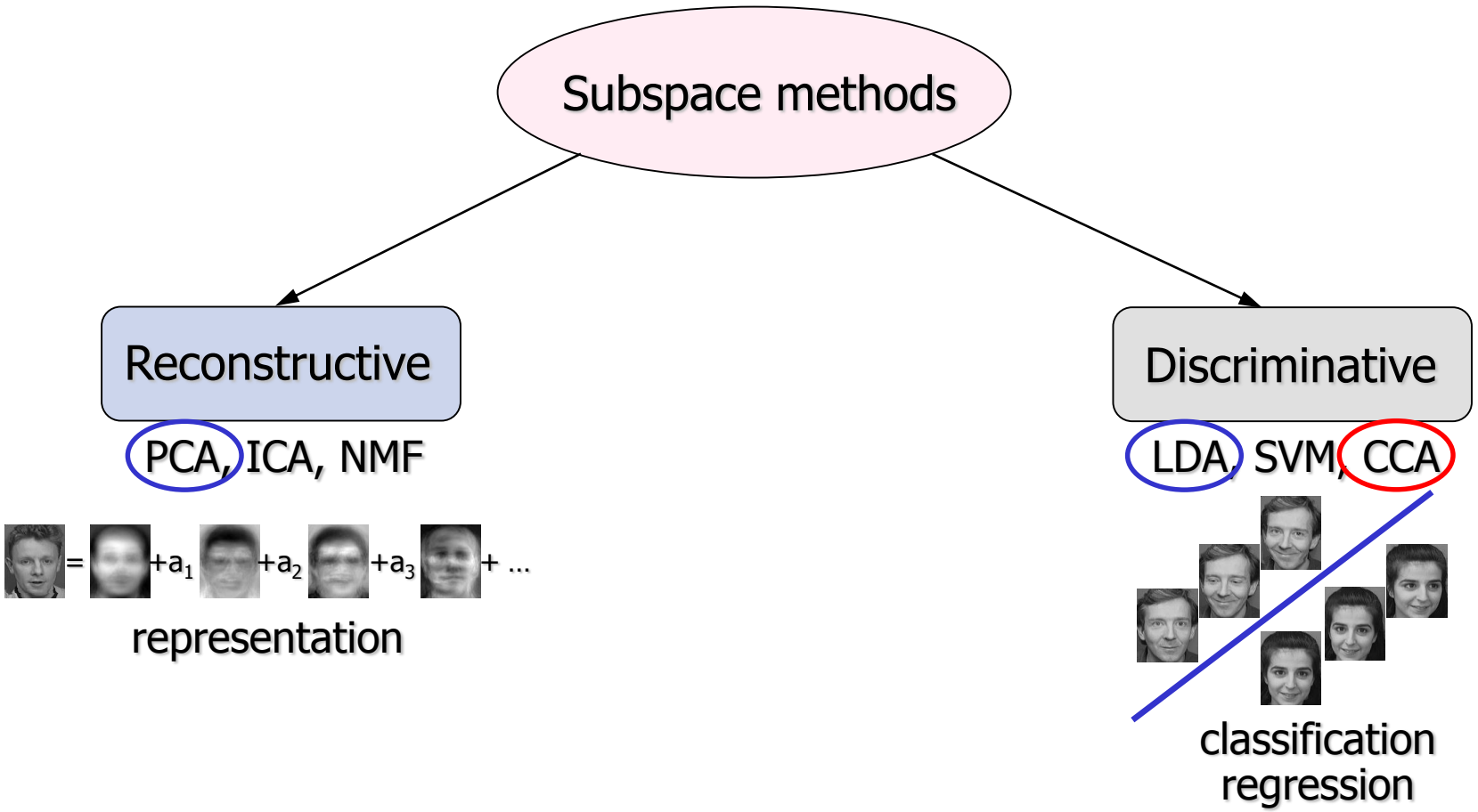
Danijel Skočaj



University of Ljubljana
Faculty of Computer and Information Science

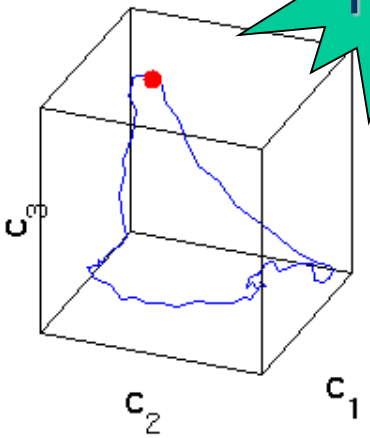
Udine; 5, 12 December 2014

Subspace methods



Object recognition and pose estimation

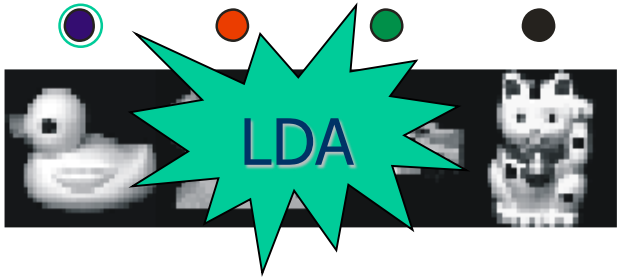




new image:

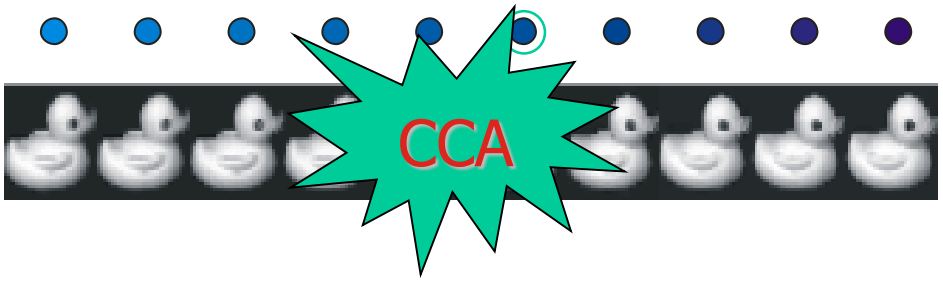


Recognition:

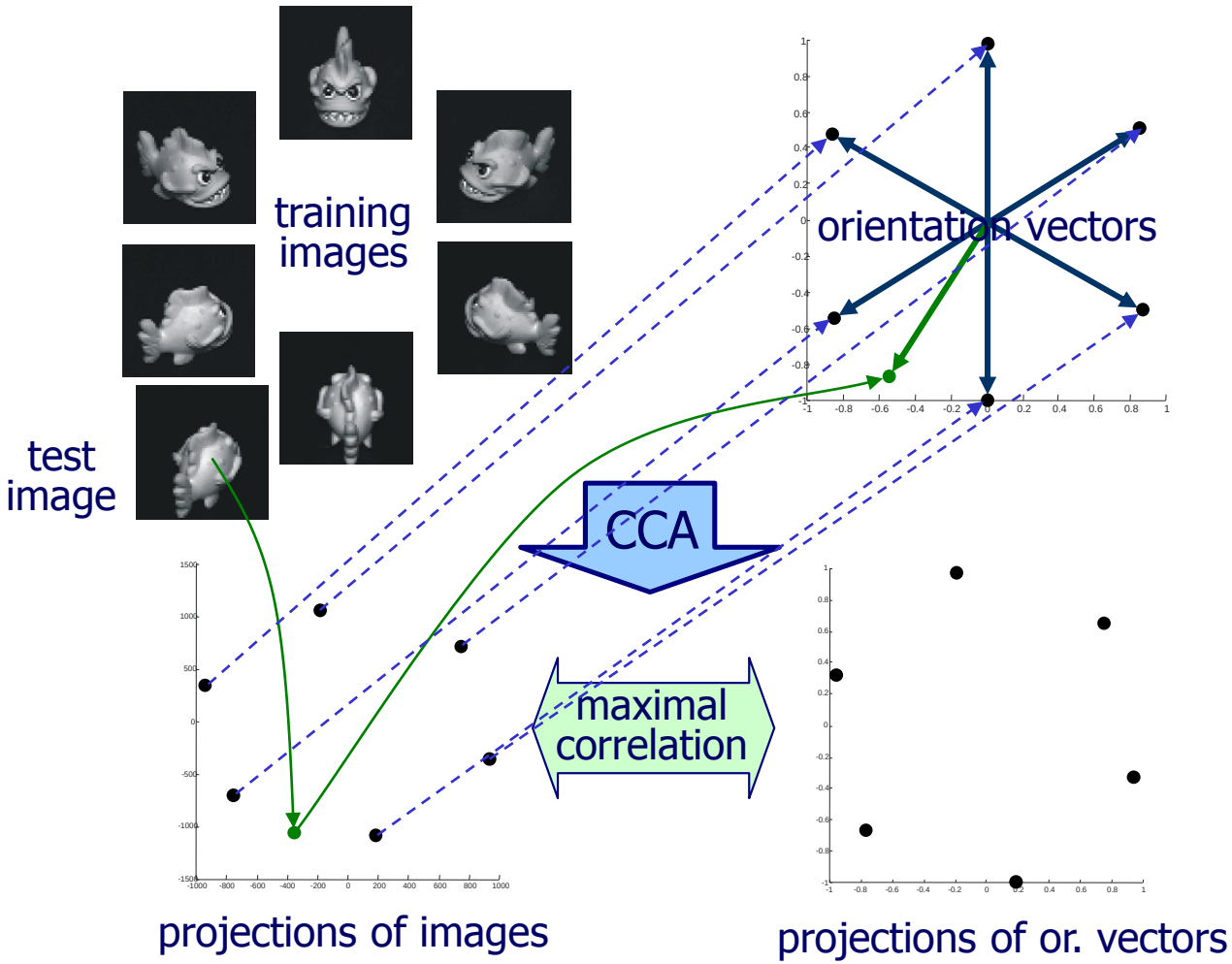


Map the images into
a low-dimensional subspace.

Pose estimation:

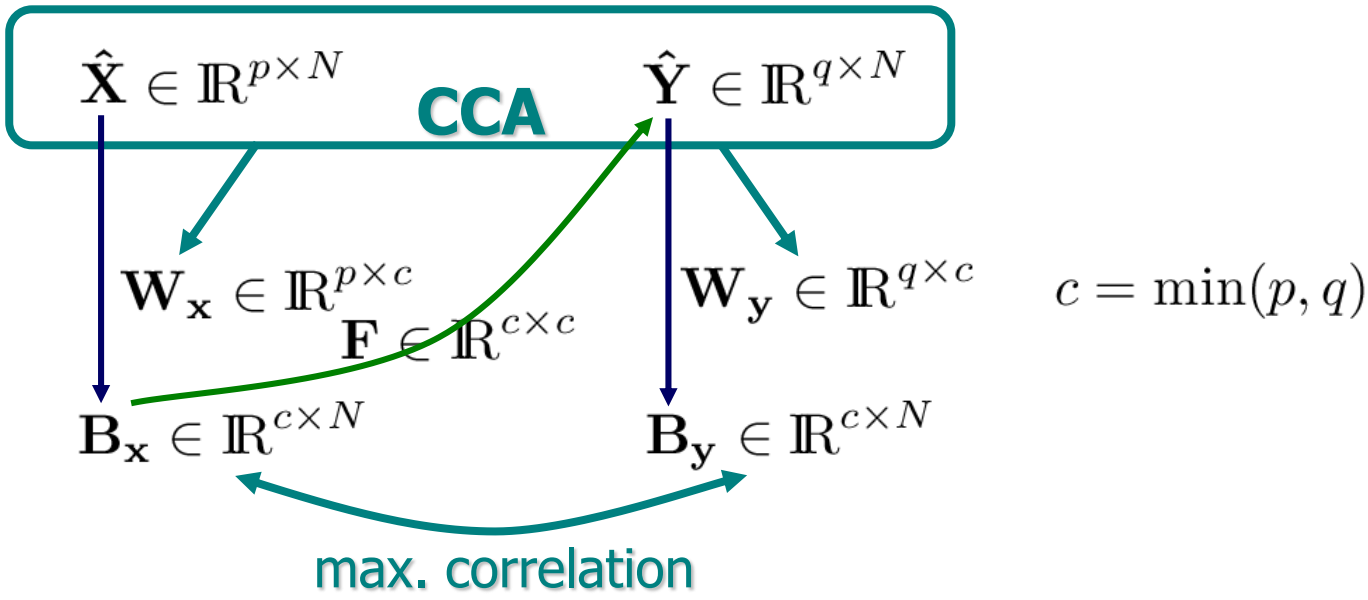


Canonical Correlation Analysis - principle



CCA - principle

CCA finds pairs of directions (*canonical correlation vectors*), that yield maximum correlation between the projections of the input vectors (*canonical correlation coefficients*).



Estimation of CCA vectors

Maximize: $$\rho = \frac{\mathbf{w_x}^\top \mathbf{C_{xy}} \mathbf{w_y}}{\sqrt{\mathbf{w_x}^\top \mathbf{C_{xx}} \mathbf{w_x} \mathbf{w_y}^\top \mathbf{C_{yy}} \mathbf{w_y}}}$$

$$\mathbf{C_{xx}} = \frac{1}{N}\hat{\mathbf{X}}\hat{\mathbf{X}}^\top \in \mathbb{R}^{p\times p} \qquad \mathbf{C_{xy}} = \frac{1}{N}\hat{\mathbf{X}}\hat{\mathbf{Y}}^\top \in \mathbb{R}^{p\times q}$$

$$\mathbf{C_{yy}} = \frac{1}{N}\hat{\mathbf{Y}}\hat{\mathbf{Y}}^\top \in \mathbb{R}^{q\times q} \qquad \mathbf{C_{yx}} = \frac{1}{N}\hat{\mathbf{Y}}\hat{\mathbf{X}}^\top \in \mathbb{R}^{q\times p}$$

Solution:

$$\mathbf{w} = \begin{bmatrix} \mathbf{w_x} \\ \mathbf{w_y} \end{bmatrix} \qquad \mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{C_{xy}} \\ \mathbf{C_{yx}} & \mathbf{0} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{C_{xx}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C_{yy}} \end{bmatrix}$$

Rayleigh quotient: $$r = \frac{\mathbf{w}^\top \mathbf{A} \mathbf{w}}{\mathbf{w}^\top \mathbf{B} \mathbf{w}}$$

Generalized eigenproblem: $$\mathbf{A}\mathbf{w} = \lambda \mathbf{B}\mathbf{w}$$

Dual formulation

If $p \gg N \gg q$:

$$\mathbf{w_x}^* = \hat{\mathbf{X}}\mathbf{w'_x}^* \qquad \mathbf{w}' = \begin{bmatrix} \mathbf{w}'_\mathbf{x} \\ \mathbf{w_y} \end{bmatrix}$$

$$\mathbf{A}' = \begin{bmatrix} \mathbf{0} & \mathbf{C}'_{\mathbf{xx}}\hat{\mathbf{Y}}^\top \\ \hat{\mathbf{Y}}\mathbf{C}'_{\mathbf{xx}} & \mathbf{0} \end{bmatrix} \qquad \mathbf{B}' = \begin{bmatrix} \mathbf{C}'^{\,2}_{\mathbf{xx}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C_{yy}} \end{bmatrix}$$

$$\mathbf{C}'_{\mathbf{xx}} = \hat{\mathbf{X}}^\top\hat{\mathbf{X}} \in \mathbb{R}^{N\times N}$$

Inner-product matrix is smaller!

Generalized eigenvalue problem feasible:

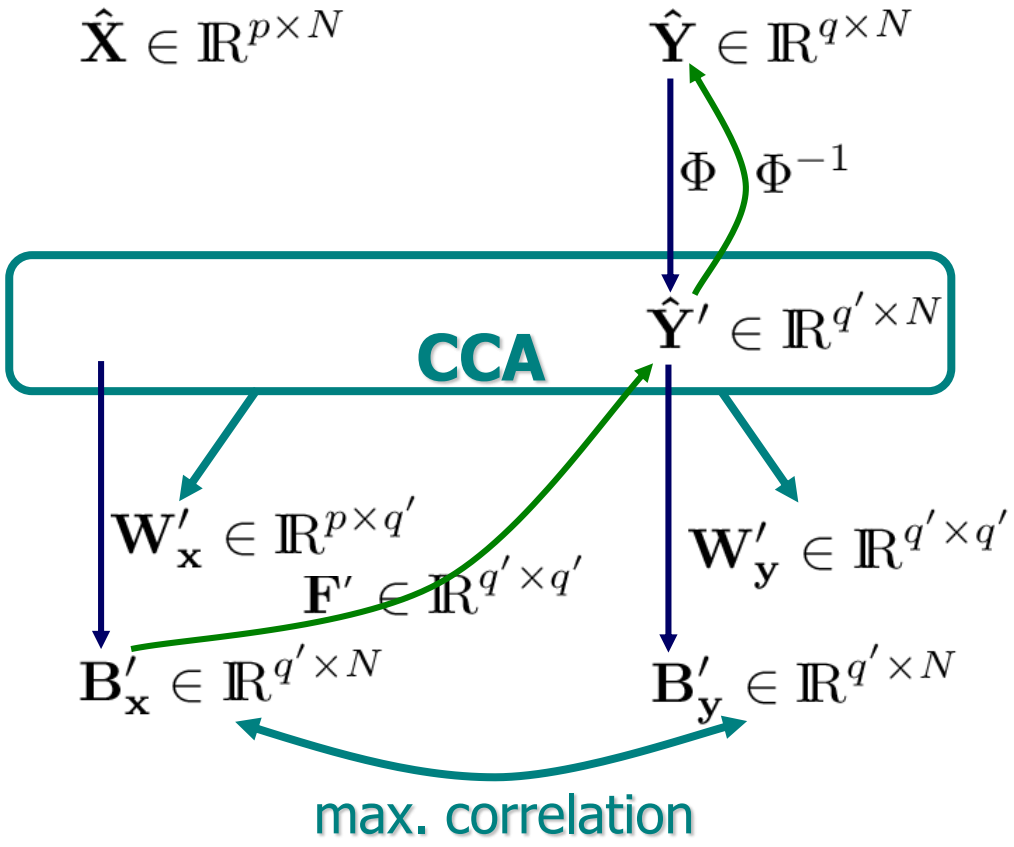
$$\mathbf{A}'\mathbf{w}' = \lambda\mathbf{B}'\mathbf{w}'$$

Algorithm

```
function [Wx,Wy,Xm,Ym]=cca(X,Y)
[p,N]=size(X); [q,N]=size(Y);  k=min([p q N]);
%center data around the mean
Xm=mean(X,2); X=X-repmat(Xm,1,N);
Ym=mean(Y,2); Y=Y-repmat(Ym,1,N);
%construct A and B
if p<=N & q<=N
   Cxx=X*X';  Cyy=Y*Y';  Cxy=X*Y'; Cyx=Y*X';
   A=[zeros(p,p) Cxy; Cyx, zeros(q,q)];
   B=[Cxx zeros(p,q); zeros(q,p) Cyy];
elseif p>N & q>N
   Cxx=X'*X; Cyy=Y'*Y;
   A=[zeros(N,N) Cxx*Cyy; Cyy*Cxx, zeros(N,N)];
   B=[Cxx*Cxx zeros(N,N); zeros(N,N) Cyy*Cyy];
elseif p>=N & q<=N
   Cxx=X'*X;  Cyy=Y*Y';
   A=[zeros(N,N) Cxx*Y'; Y*Cxx, zeros(q,q)];
   B=[Cxx*Cxx zeros(N,q); zeros(q,N) Cyy];
else %p<=N & q>=N
   Cxx=X*X';  Cyy=Y'*Y;
   A=[zeros(p,p) X*Cyy; Cyy*X', zeros(N,N)];
   B=[Cxx zeros(p,N); zeros(N,p) Cyy*Cyy];
end
%generalized eigendecomposition
[Z,d]=eig(A,B);
%sort eigenvalues
d=diag(d);
[ds,sidxs]=sort(real(-d));
rsidxs=sidxs([abs(imag(ds))<10^-10]);
Z=Z(:,rsidxs(1:k));
%map to input spaces
if p<=N & q<=N
   Wx=Z(1:p,:);
   Wy=Z(p+1:end,:);
elseif p>N & q>N
   Wx=X*Z(1:N,:);
   Wy=Y*Z(N+1:end,:);
elseif p>=N & q<=N
   Wx=X*Z(1:N,:);
   Wy=Z(N+1:end,:);
else %p<=N & q>=N
   Wx=Z(1:p,:);
   Wy=Y*Z(p+1:end,:);
end
```

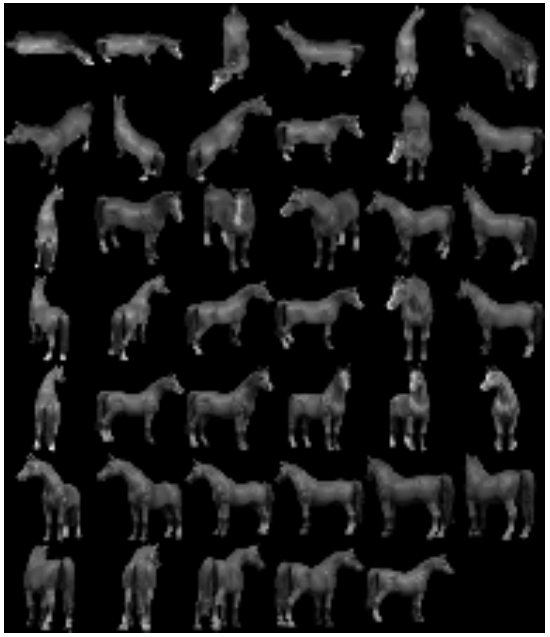
Increasing number of parameters

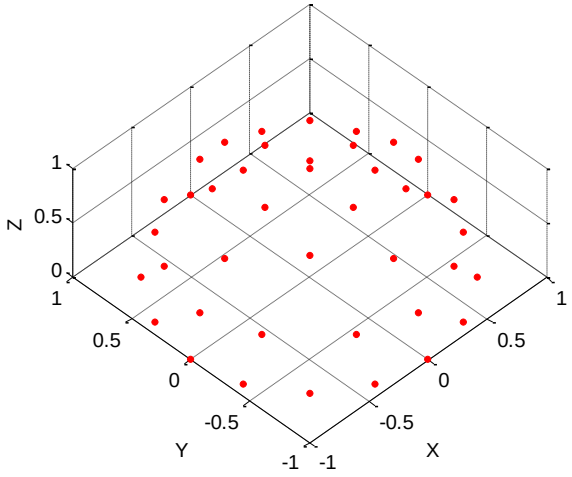
q is small $\Rightarrow$ increase subspace dimension:



Example

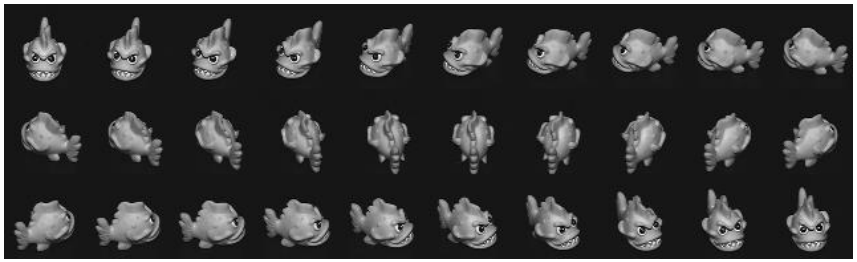
- Estimation of orientation
  - Find three canonical correlation vectors, which yield 3-D coordinates of view-points.

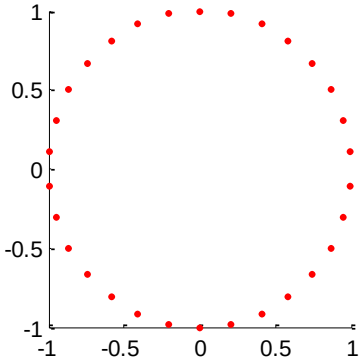


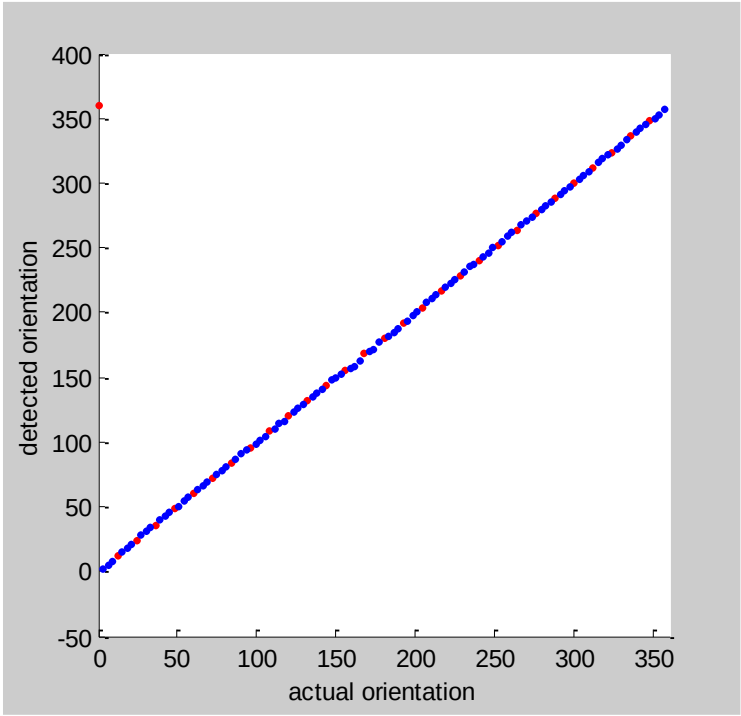


Example

- Estimation of orientation





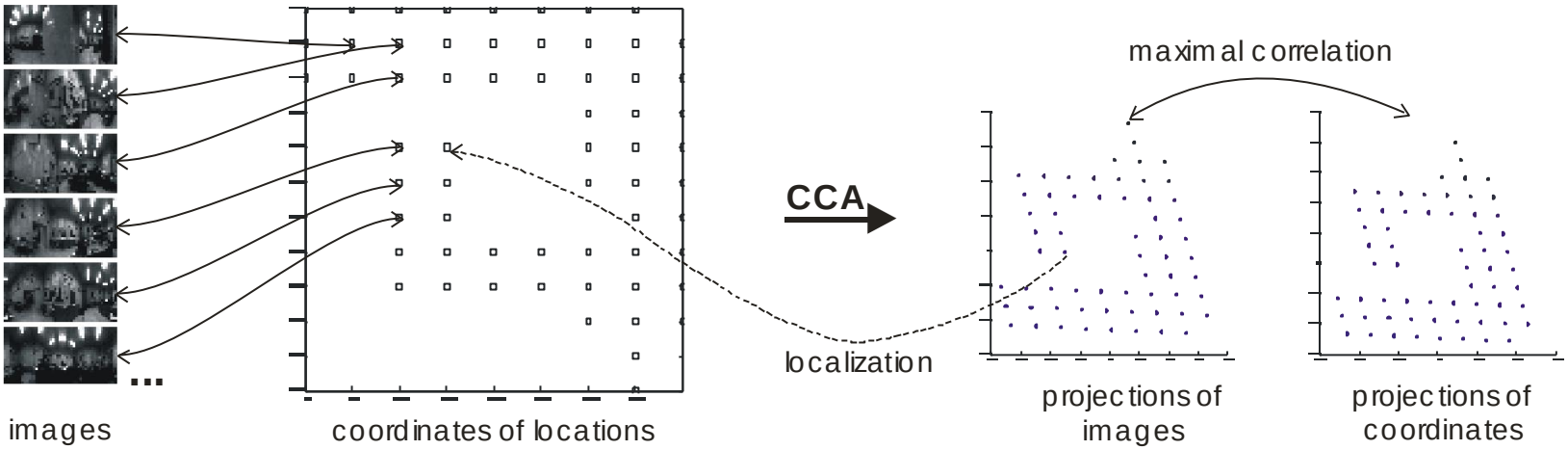


CCA vs. PCA

- Advantages of CCA for regression tasks:
  - We obtain parameters directly from the subspace coefficients using simple linear mapping (without searching for the nearest neighbor).
  - We do not have to keep subspace coefficients of the training images and interpolations.
  - The domain of the results is continuous.
  - It is faster.
- We lose all this advantages if we use KCCA.
- How good (usefull) is CCA in comparison with PCA? $\Rightarrow$ Compare the results.

CCA for mobile robot localization

- CCA is well suited for regression task such as localization of a mobile robot.



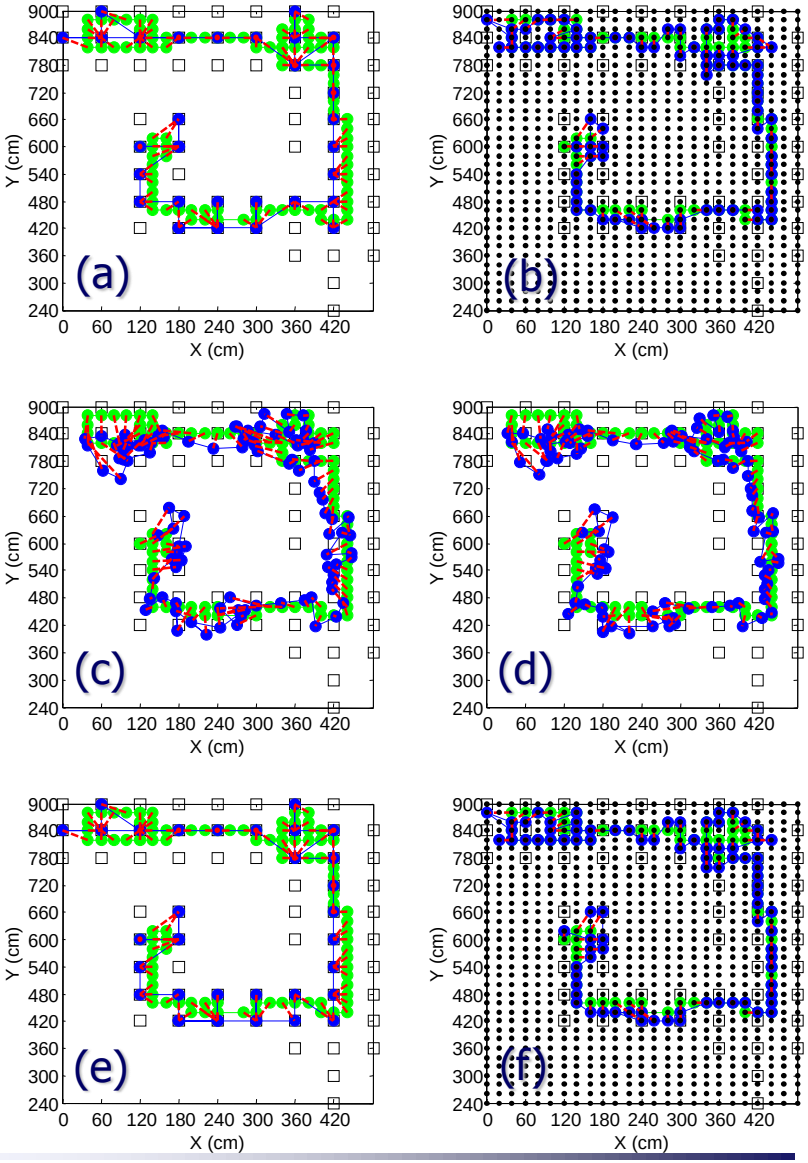
CCA vs. PCA vs. KCCA - 2DOF localization

- CMP panoramic images



- 2DOF (x,y)
- Average error in cm:

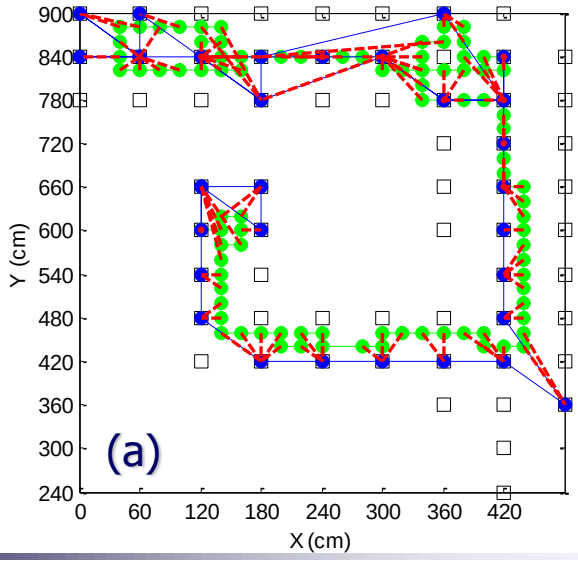
		k	error	secs	elem.
(a)	PCA	10	26.16	0.23	4148
(b)	PCAint	10	13.04	2.77	24500
	PCA	2	77.37	0.21	4148
(c)	CCA	2	34.77	0.02	3906
(d)	CCAsncs	4	27.32	0.03	6520
(e)	KCCA	10	26.65	0.23	15044
(f)	KCCAint	10	13.12	2.81	24500

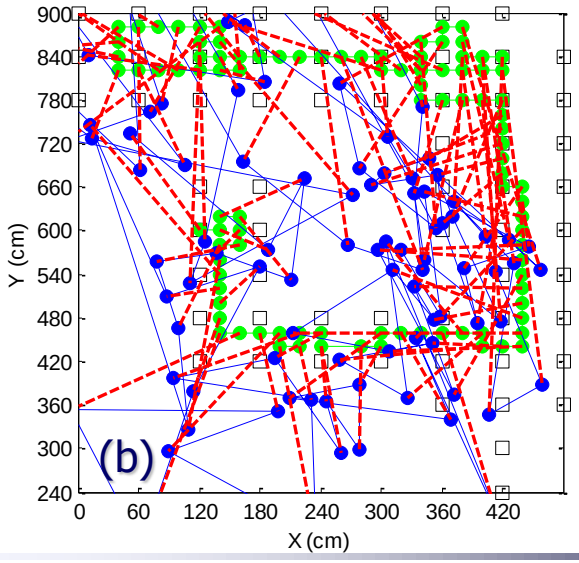


CCA vs. PCA - 3DOF localization

- 3DOF (x,y,θ) – postion and rotation
- Spinning images

		k	error	secs	elem.
(a)	PCA	10	39.46	2.07	22360
	PCA	3	210.63	2.06	8920
(b)	CCA	3	161.35	0.04	5212

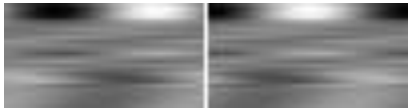
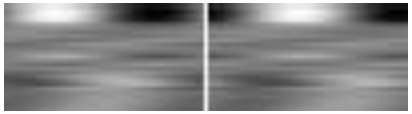
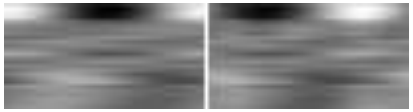
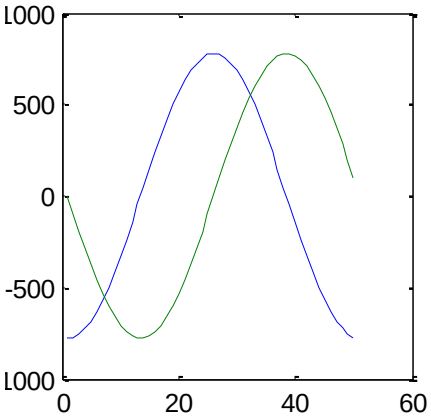
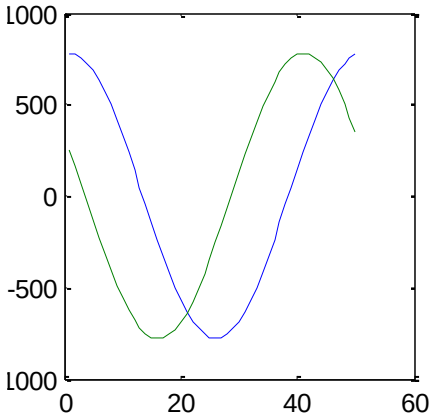
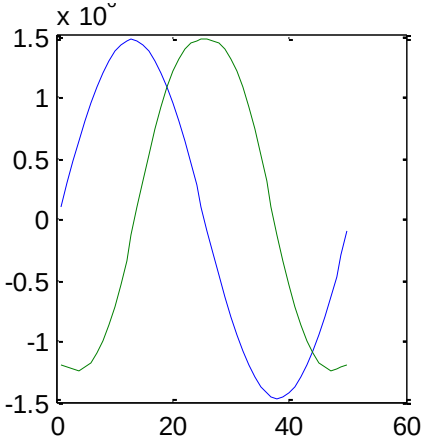




PCA vs. CCA - 1DOF

- Spinning images of one CMP image



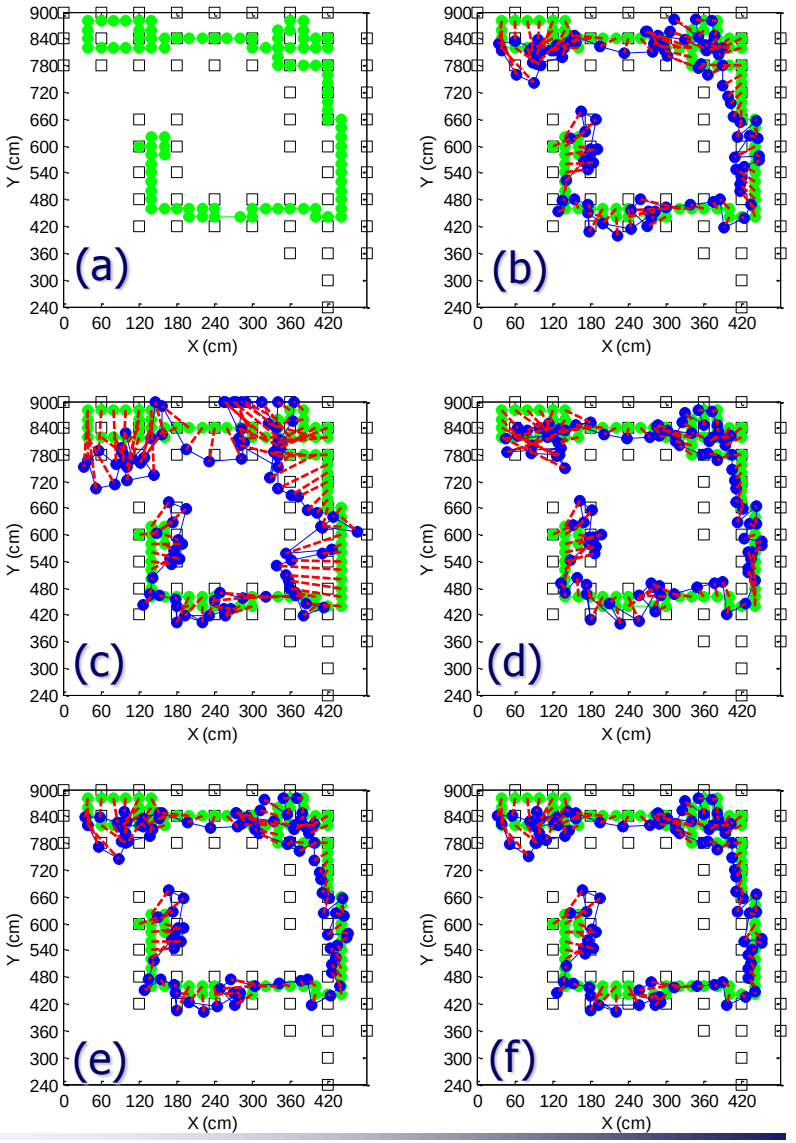
	PCA (k=2)	CCA (k=2) sin,cos	KCCA (k=2)
First two vectors:			
Coeficients:			

- PCA and CCA of spinning images are equivalent

Increasing number of parameters

- 2DOF, CMP lab
- Both parameters mapped using sine and cosine functions.
- Avg. error in cm:

	mapping	k	error	secs	elem.
(a)	ground truth				
(b)	linear	2	34.77	0.02	3906
(c)	sin	2	58.24	0.02	3906
(d)	cos	2	34.33	0.02	3906
(e)	tan	4	30.65	0.02	6520
(f)	sin|cos	4	27.32	0.03	6520



CCA vs. PCA - conclusions

- CCA is significantly faster and simpler than PCA.
- CCA requires less storage requirements.
- 1DOF: PCA and CCA have very similar performance
- 2DOF: PCA ouperforms CCA
  - CCA is better if k=2 (however not very good)
  - PCA is better if k>>2 (which is not possible for CCA)
  - KCCA has very similar performance as PCA for equal k (however we lose advantages of CCA)
- 3DOF: CCA is useless (it is difficult to obtain three parameters from three vectors and coefficients)
- We can improve the results by increasing the number of coefficients by nonlinearly mapping input parameters.
- However, the inverse function has to be known!

Literature

- D. Skočaj, A. Leonardis, *Appearance-Based Localization using CCA*, CVWW 2004
  - http://vicos.fri.uni-lj.si/data/publications/skocajCVWW04.pdf

- T. Melzer, M. Reiter, H. Bischof, *Appearance models based on kernel canonical correlation analysis*, Pattern Recognition, 2003
  - http://www.sciencedirect.com/science/article/pii/S003132030300058X
- Magnus Borga, *Learning Multidimensional Signal Processing*, PhD thesis, Linköping University, 1998
  - http://www.imt.liu.se/mi_old/Publications/pdfs/borga98.pdf