10th VIPS Advanced School on Computer Vision and Pattern Recognition

Dissimilarity-based Representation for Pattern Recognition, Final words

Robert P.W. Duin, Delft University of Technology

Pattern Recognition Lab
Delft University of Technology
The Netherlands

//rduin.nl

TUDelft

Learning about the world

Human knowledge grows
in the debate between

-those who see the patterns, and

-those who know the universal laws

TUDelft

Popper versus Kuhn

Karl Popper:
We generate a **conjecture** and
try to **refute** it by an observation.

Thomas Kuhn:
Theories and counter examples can coincide
for a long time until by a **paradigm shift**
a new theory is accepted.

TUDelft

PR History

William of Occam (1284 – 1347)

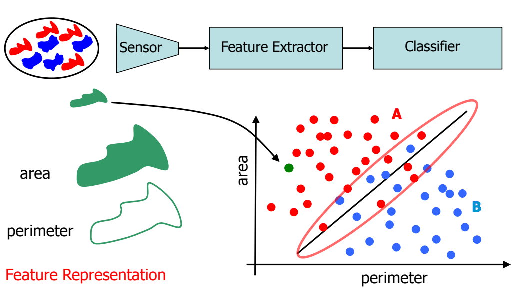
Rev. Thomas Bayes (1702 – 1761)

Fisher's Linear Discriminant

C.R. Rao, 1948,
The utilization of multiple measurements
in problems of biological classification.
Rao's paradox (Curse of dimensionality)

TUDelft

Pattern Recognition System

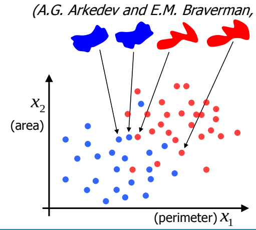


TUDelft

Compactness

Representations of real world similar objects are close.
There is no ground for any generalization (induction) on representations
that do not obey this demand.

(A.G. Arkedev and E.M. Braverman, Computers and Pattern Recognition, 1966.)

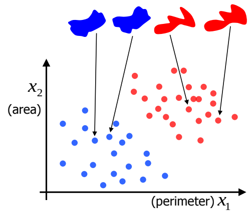


The compactness hypothesis is not sufficient for perfect classification as dissimilar objects may be close.
→ class overlap
→ probabilities

Reduce the set of PR problems to compact problems

TUDelft

True Representations



Similar objects are close
and
Dissimilar objects are distant.

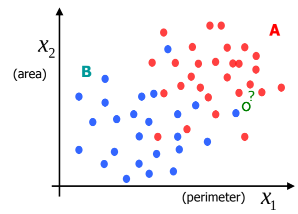
→ no probabilities needed, domains are sufficient!

Distances and Densities

? to be classified as

B – because it is most close to an object B

A – because the local density of A is larger.



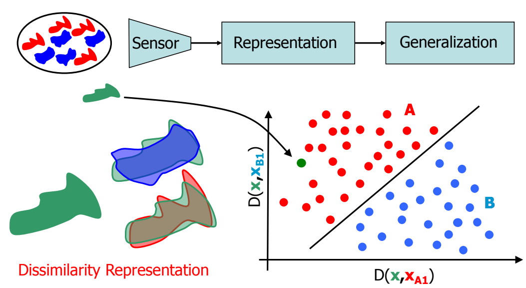
The right question

Fabio Roli (Barcelona, 2000):
What is the core business of pattern recognition?

Finding **representations** of real world problems such that given **generalizations** work.

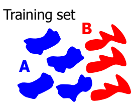
Duin, Roli, de Ridder,
A note on core research issues for statistical pattern recognition, PRL, 2002

Dissimilarities → True Representation



Alternatives for the Nearest Neighbor Rule

Training set



Dissimilarities d_{ij} between all training objects

$$D_T = \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} & d_{27} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} & d_{37} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} & d_{47} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} & d_{57} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} & d_{67} \\ d_{71} & d_{72} & d_{73} & d_{74} & d_{75} & d_{76} & d_{77} \end{pmatrix}$$

Unlabeled object **x** to be classified

$$d_x = (d_{x1}\ d_{x2}\ d_{x3}\ d_{x4}\ d_{x5}\ d_{x6}\ d_{x7})$$

1. Dissimilarity Space
2. Embedding

Pekalska, The dissimilarity representation for PR. World Scientific, 2005.

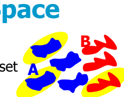
Alternative 1: Dissimilarity Space

Dissimilarities

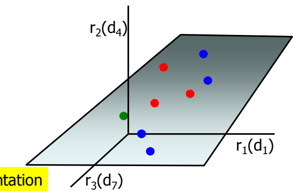
$$D_T = \begin{array}{c} \begin{array}{ccccccc} r_1 & & & r_2 & & & r_3 \end{array} \\ \begin{pmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ d_{21} & d_{22} & d_{23} & d_{24} & d_{25} & d_{26} & d_{27} \\ d_{31} & d_{32} & d_{33} & d_{34} & d_{35} & d_{36} & d_{37} \\ d_{41} & d_{42} & d_{43} & d_{44} & d_{45} & d_{46} & d_{47} \\ d_{51} & d_{52} & d_{53} & d_{54} & d_{55} & d_{56} & d_{57} \\ d_{61} & d_{62} & d_{63} & d_{64} & d_{65} & d_{66} & d_{67} \\ d_{71} & d_{72} & d_{73} & d_{74} & d_{75} & d_{76} & d_{77} \end{pmatrix} \end{array}$$

$$d_x = (d_{x1}\ d_{x2}\ d_{x3}\ d_{x4}\ d_{x5}\ d_{x6}\ d_{x7})$$

Given labeled training set

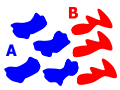


Unlabeled object to be classified



Selection of 3 objects for representation

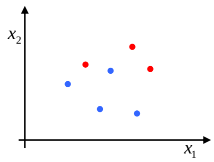
Alternative 2: Embedding



Training set

→ Dissimilarity matrix D → X

Is there a feature space for which Dist(X,X) = D ?



Position points in a vector space such that their Euclidean distances → D

TUDelft

(Pseudo-)Euclidean Embedding

$m \times m$ D is a given, imperfect dissimilarity matrix of training objects.

Construct inner-product matrix: $B = -\frac{1}{2} J D^{(2)} J \quad J = I - \frac{1}{m}\mathbf{11}$

Eigenvalue Decomposition , $B = Q \Lambda Q^T$

Select k eigenvectors: $X = Q_k \Lambda_k^{\frac{1}{2}}$ (problem: $\Lambda_k < 0$)

Let $\Im_k$ be a k x k diag. matrix, $\Im_k(i,i)$ = sign($\Lambda_k(i,i)$)

$\Lambda_k(i,i) < 0 \rightarrow$ Pseudo-Euclidean

$n \times m$ D_z is the dissimilarity matrix between new objects and the training set.

The inner-product matrix: $B_z = -\frac{1}{2}(D_z^{(2)} J - \frac{1}{n}\mathbf{11}^T D^{(2)} J)$

The embedded objects: $Z = B_z Q_k |\Lambda_k|^{-\frac{1}{2}} \Im_k$

TUDelft

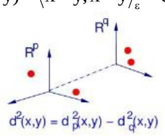
PES: Pseudo-Euclidean Space (Krein Space)

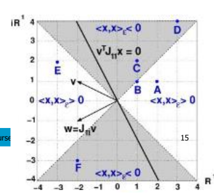
If D is non-Euclidean, B has p positive and q negative eigenvalues.

A pseudo-Euclidean space $\mathcal{E}$ with signature (p,q), k =p+q, is a non-degenerate inner product space $\Re_k = \Re_p \oplus \Re_q$ such that:

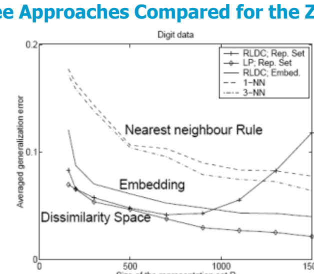
$$\langle x, y \rangle_\varepsilon = x^T \Im_{pq} y = \sum_{i=1}^{p} x_i y_i - \sum_{j=p+1}^{q} x_j y_j \qquad \Im_{pq} = \begin{bmatrix} I_{p \times p} & 0 \\ 0 & -I_{q \times q} \end{bmatrix}$$

$$d_\varepsilon^2(x,y) = \langle x-y, x-y \rangle_\varepsilon = d_p^2(x,y) - d_q^2(x,y)$$





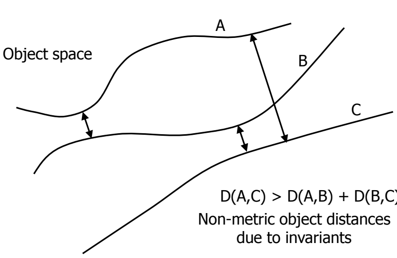
Three Approaches Compared for the Zongker Data



Dissimilarity Space equivalent to Embedding better than Nearest Neighbour Rule

TUDelft

Intrinsicly Non-Euclidean Dissimilarity Measures Invariants

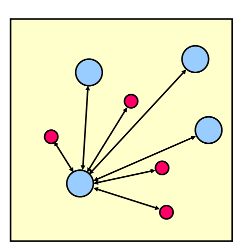


D(A,C) > D(A,B) + D(B,C)

Non-metric object distances due to invariants

TUDelft

Artificial Example ,Ball Distances

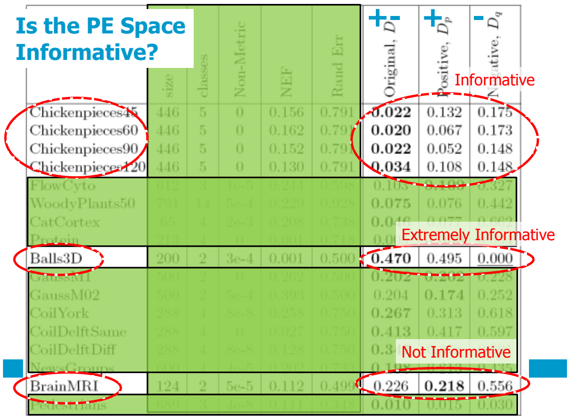


- Generate sets of balls (classes) uniformly, in a (hyper)cube; not intersecting.
- Balls of the same class have the same size.
- Compute all distances between the ball surfaces.

-> Dissimilarity matrix D

TUDelft

Is the PE Space Informative?



	size	classes	Non-Metric	NEF	Rand Err	Original, D	Positive, D_p	Negative, D_q
Chickenpieces45	446	5	0	0.156	0.791	**0.022**	0.132	0.175
Chickenpieces60	446	5	0	0.162	0.791	**0.020**	0.067	0.173
Chickenpieces90	446	5	0	0.152	0.791	**0.022**	0.052	0.148
Chickenpieces120	446	5	0	0.130	0.791	**0.034**	0.108	0.148
FlowCyto	612	3	1e-5	0.244	0.598	0.103	[illegible]	0.327
WoodyPlants50	791	14	5e-4	0.229	0.928	**0.075**	0.076	0.442
CatCortex	65	4	2e-3	0.208	0.738	**0.046**	0.077	0.662
Protein	213	4	0	0.001	0.718	[illegible]	[illegible]	[illegible]
Balls3D	200	2	3e-4	0.001	0.500	**0.470**	0.495	0.000
GaussM1	500	2	0	0.262	0.500	**0.202**	**0.202**	0.228
GaussM02	500	2	5e-4	0.393	0.500	0.204	**0.174**	0.252
CoilYork	288	4	8e-8	0.258	0.750	**0.267**	0.313	0.618
CoilDelftSame	288	4	0	0.027	0.750	**0.413**	0.417	0.597
CoilDelftDiff	288	4	8e-8	0.128	0.750	[illegible]	[illegible]	[illegible]
NewsGroups	600	4	4e-5	0.202	0.733	[illegible]	[illegible]	0.435
BrainMRI	124	2	5e-5	0.112	0.499	0.226	**0.218**	0.556
Pedestrians	689	3	4e-8	0.111	0.348	**0.010**	0.015	0.030

+- + -

Informative

Extremely Informative

Not Informative

Bridging structural and statistical PR

Horst Bunke (2002): *The dissimilarity representation may be a good approach for using statistical tools In structural problems*

→ Objects are not points, they have a size

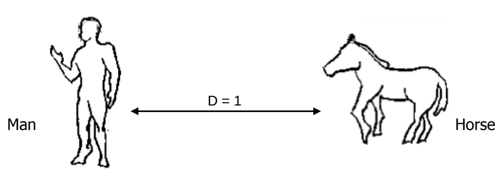
Riesen, Bunke, Graph Classification Based on Vector Space Embedding, IJPRAI, 2009

....

Pekalska, Duin, Gunter, Bunke, On not making dissimilarities Euclidean, SSSPR 2006

....

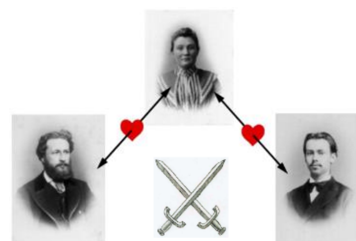
TUDelft



David W. Jacobs, Daphna Weinshall and Yoram Gdalyahu, Classification with Nonmetric Distances: Image Retrieval and Class Representation, *IEEE Trans. Pattern Anal. Mach. Intell*, 22(6), pp. 583-600, 2000.

TUDelft

Intrinsicly Non-Euclidean Dissimilarity Measures



Non-Euclidean human relations

TUDelft

Conclusions

- Real world objects are not points
- Objects have a size
- Relations are non-Euclidean
- Non-Euclidean generalization procedures are needed

TUDelft