COST B21 „Physiological modelling of MR image formation"

Texture feature selection based on clustering quality

Artur Klepaczko, Andrzej Materka

ZAKŁAD ELEKTRONIKI MEDYCZNEJ

Instytut Elektroniki Politechniki Łodzkiej

Cyprus, 1 October 2004

Instytut Elektroniki Politechnika Łódzka

Wydział Elektrotechniki i Elektroniki

the Faculty of Electrical and Electronic Engineering

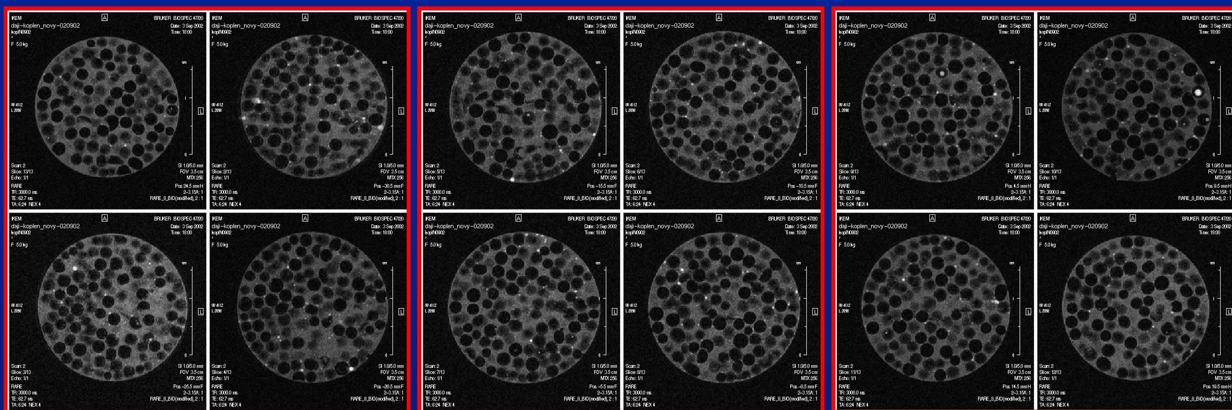
Aim

Development of unsupervised technique for best texture feature selection

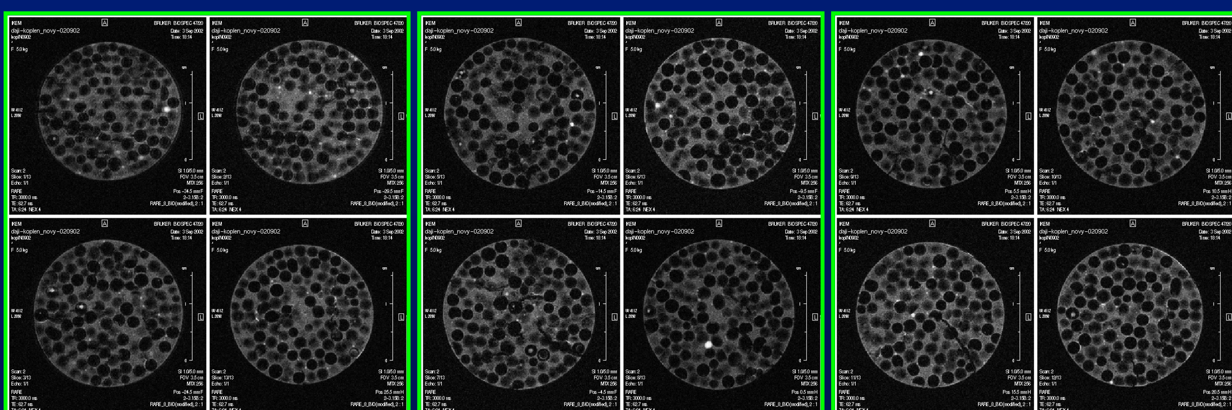
Expected advantages

- Dimensionality reduction
- No need for class labels
- More objective analysis

Motivation



Dan's „class 1" PSAG images



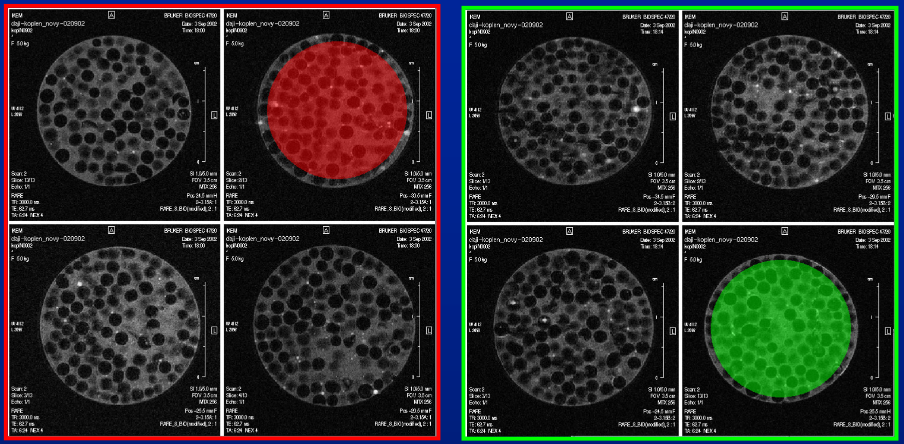
Dan's „class 2" PSAG images

Spheres of the same diameter, the same texture expected, yet MaZda classifies them into two different classes.

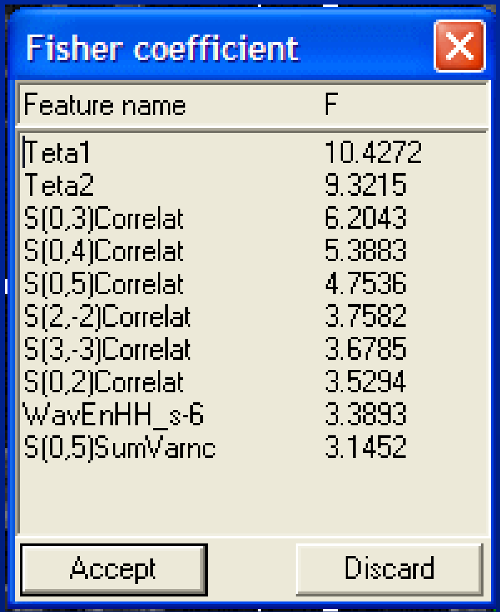
Is MaZda classifier too sensitive?

Call for „invariant" features, stable for objects of the same internal structure.

Supervised classification

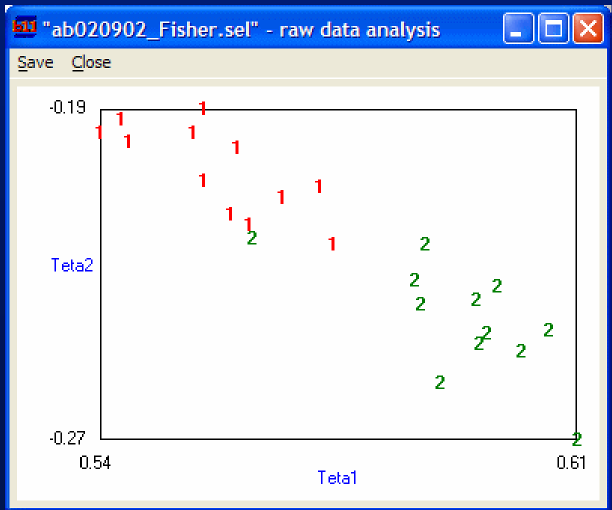


Regions of interest, class 1, class 2



Feature name	F
Teta1	10.4272
Teta2	9.3215
S(0,3)Correlat	6.2043
S(0,4)Correlat	5.3883
S(0,5)Correlat	4.7536
S(2,-2)Correlat	3.7582
S(3,-3)Correlat	3.6785
S(0,2)Correlat	3.5294
WavEnHH_s-6	3.3893
S(0,5)SumVarnc	3.1452

10 MaZda features separating the classes best



Scatter plot, best two feature space

Images randomly split into 2 classes do not demonstrate different textures.

Is partitioning real, or forced by *a priori* information (class labels)?

Clustering

Method: Similarity-Based Clustering (SCM)

- Recent (Yang & Wu, 2004)
- Robust to initial cluster number and sample membership
- Different cluster volumes
- Robust to noise and outliers

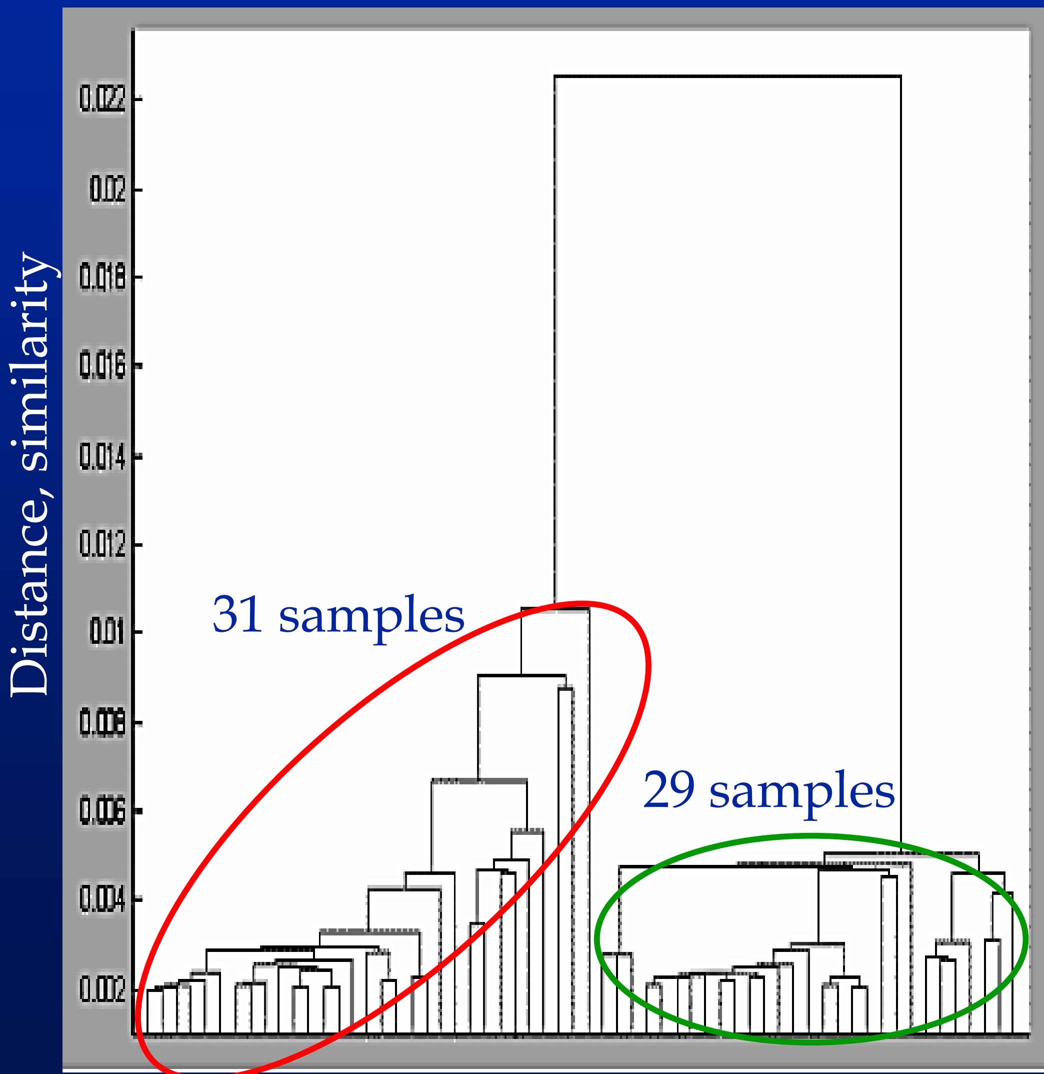
Material

- All 120 images
- Class labels excluded
- 4 features: Theta1 – Theta4

Steps

- Estimation of number of clusters (through peaks in a similarity function)
- Similarity Clustering Algorithm (relocation of points in the feature space)
- Agglomerative Hierarchical Clustering (final data grouping, forming a tree)

Clustering



Distance of the final link is much bigger than distances of the links below.

The partitioning can be done at the top level.

Division based on the tree:
- all „class 1"samples belong to one cluster,
- all „class 2" samples (except for two) belong to the second cluster.

Labelling is justified. There are apparently two subsets of data.

Is partitioning real, or forced by *prior* feature selection?

Novel feature selection method

Concept

Best features should give best quality clusters.

Algorithm

- Select a variety of different subsets of texture parameters
- Perform automatic, unsupervised clustering for each subset
- Evaluate quality of each cluster
- Indicate texture parameters that provide best clusters

Cluster quality measures

Compactness

$$\sum_{i=1}^{c} d_i \cdot \frac{n_i}{n}$$

d_i – cluster diameter
n_i – number of points in custer
n – total number of points
c - number of clusters

Cophenetic correlation

$$cc = \frac{\sum_{i<j}\left(Y_{ij} - y\right)\left(Z_{ij} - z\right)}{\sqrt{\sum_{i<j}\left(Y_{ij} - y\right)^2 \sum_{i<j}\left(Z_{ij} - z\right)^2}}$$

Y_{ij} – distance in feature space
Zij – distance in tree
y, z – mean values

Inconsistency

$$ic_k = \frac{l_k - \mu_k}{\sigma_k}$$

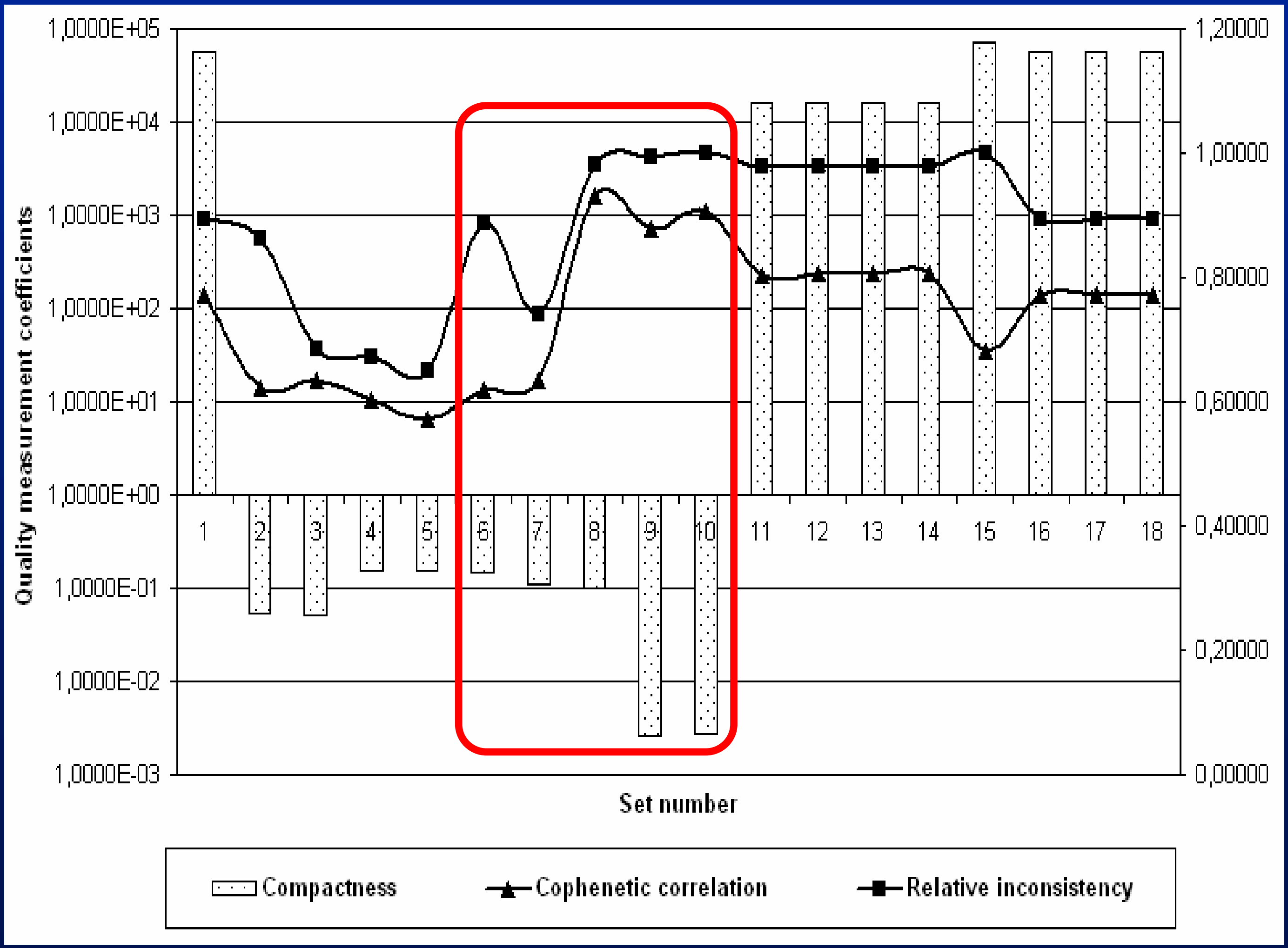
l_k – length of link
μ_k – mean length of links
σ_k – standard deviation of…

Clustering validity

$$cv = cc \cdot \frac{ic_{n-1}}{\max_k ic_k}$$

c_{n-1} – inconstistency of the last link

Clustering quality for 5-feature subsets



Subsets

Theta1: # 6, 7, 8, 9, 10
Theta2: # 7, 8, 9, 10
Theta3: # 8, 9, 10
Theta4: # 9, 10

Other features

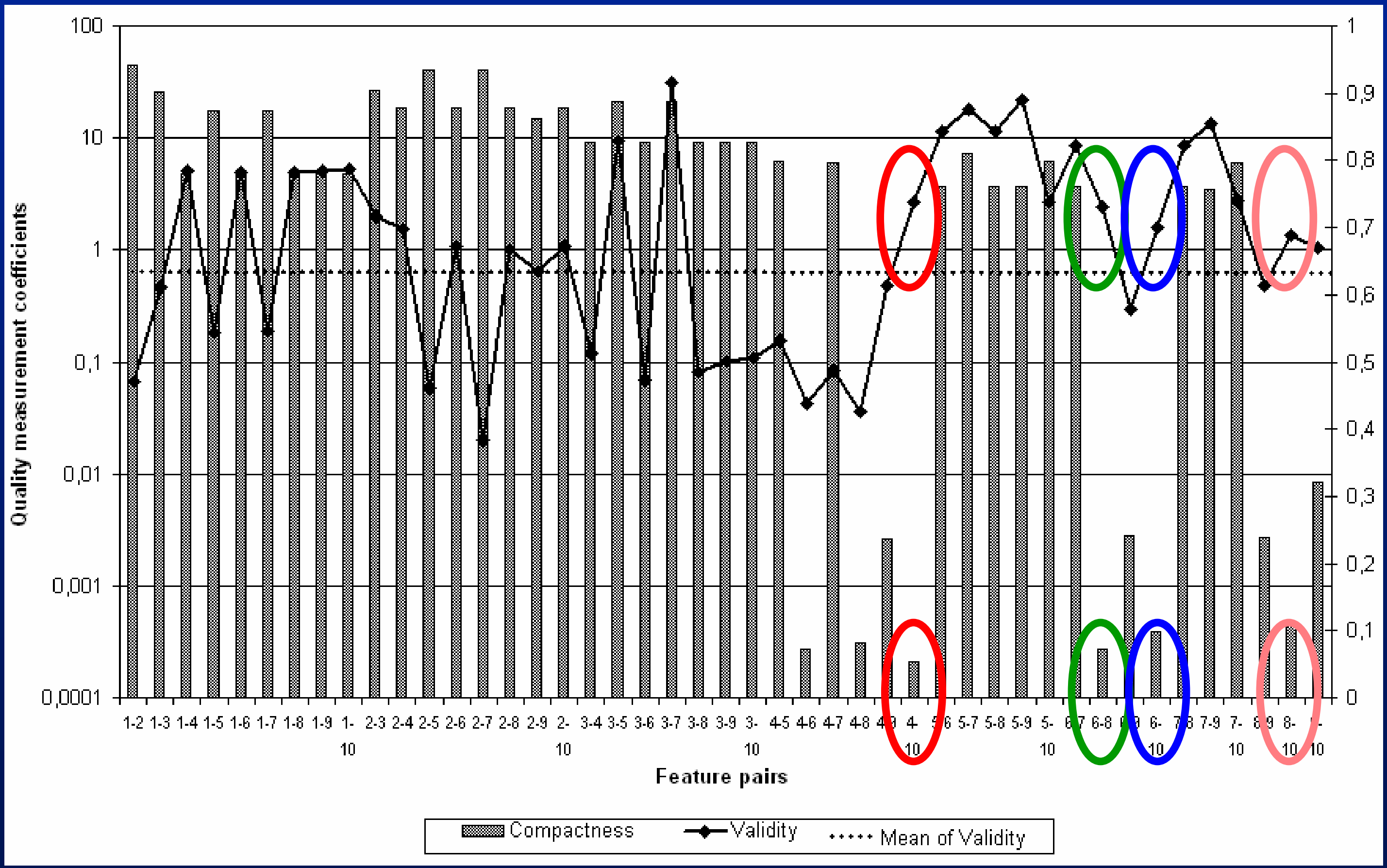
Wavelets
Absolute gradient
Run-length matrix

Considering all three measures and many feature combinations are necessary.

Quality-based feature selection algorithm

- Divide feature vector into N-element subsets (N=10 in our study)
- For every subset create all possible pairs of features (45 for 10-element vector)
- For each pair perform SCM clustering
- Evaluate clusters quality (compute *compactness* and *clustering validity*)
- From every subset choose a winning pair (minimal compactness and validity over the average)
- Eliminate worst features and repeat the algorithm (best features minimize compactness and maximize clustering validity)

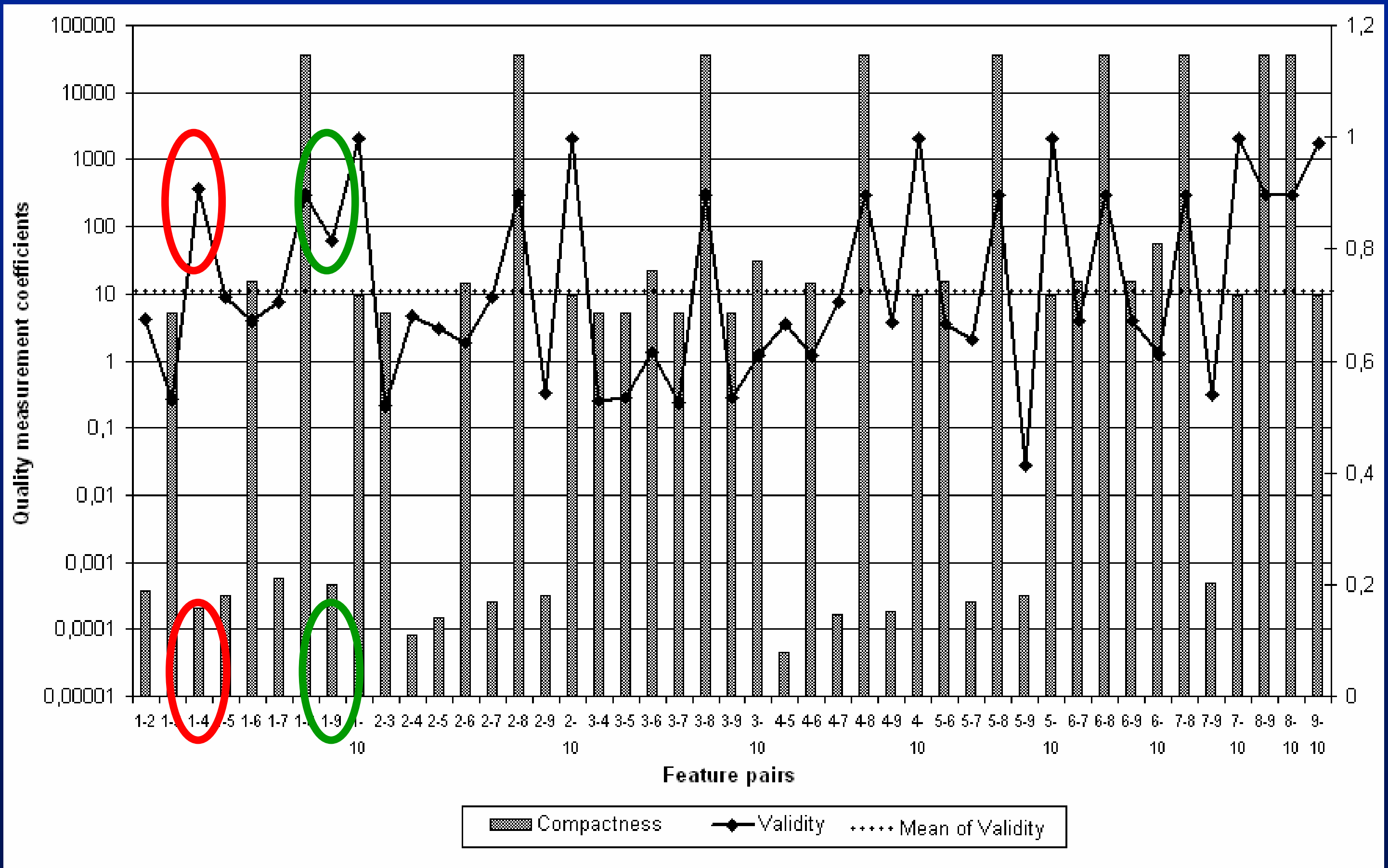
Clustering quality for Theta1 subsets



No	Parameter name
1	S(3,-3)DifVarnc
2	WavEnLH_s-1
3	S(0,3)Contrast
4	S(1,-1)Correlat
5	S(1,1)SumOfSqs
6	S(4,-4)SumEntrp
7	S(0,5)SumOfSqs
8	S(3,3)SumEntrp
9	S(0,2)Entropy
10	Teta1

A Klepaczko, A Materka: Texture Feature Selection Based on Clustering Quality

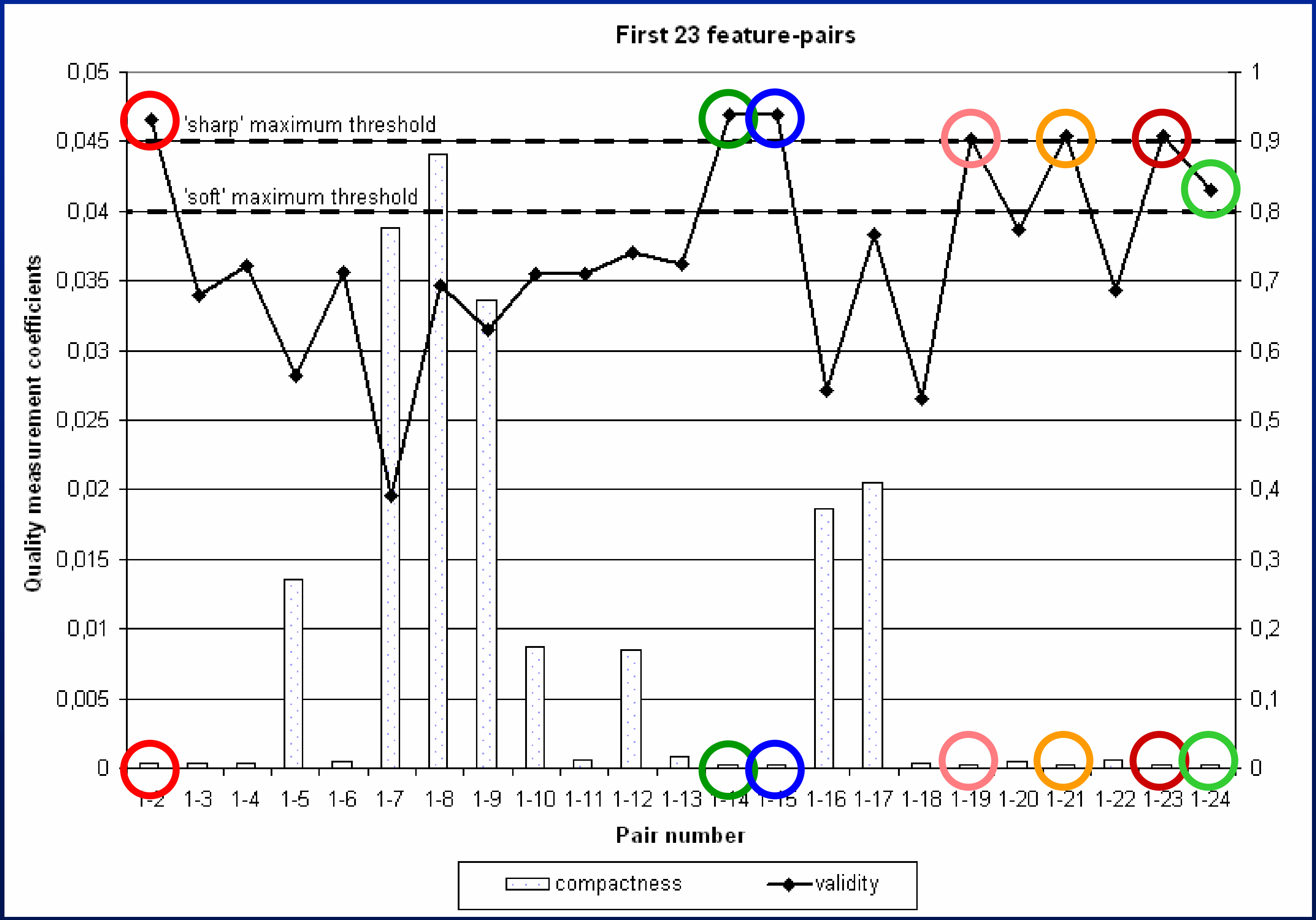
Clustering quality for Theta2 subsets



No	Parameter name
1	Teta2
2	45dgr_Fraction
3	S(2,2)Contrast
4	S(1,-1)AngScMom
5	Horzl_ShrtREmp
6	S(0,5)Contrast
7	S(4,0)SumEntrp
8	WavEnLL_s-4
9	S(3,-3)Correlat
10	S(4,0)DifVarnc

Next stage:
24 best parameters
(276 pairs).

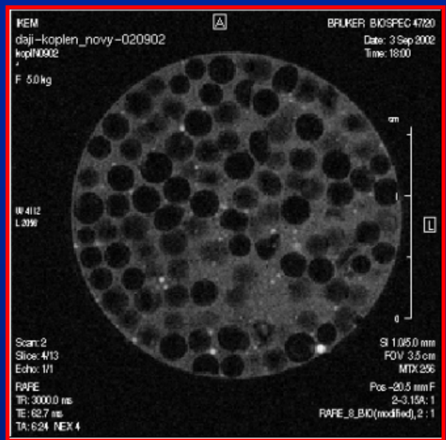
First 23 feature pairs (out of 276 pairs)



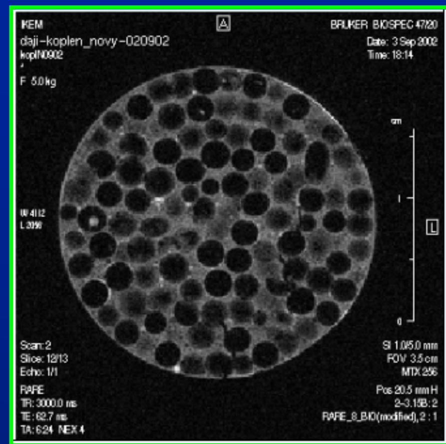
Final-stage features

No	Parameter name
1	Teta1
2	Teta2
3	S(4,0)AngScMom
4	S(3,3)AngScMom
5	S(0,2)AngScMom
6	S(1,-1)AngScMom
7	S(1,1)AngScMom

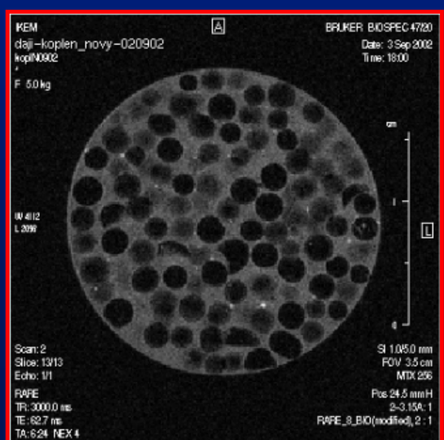
Supervised classification



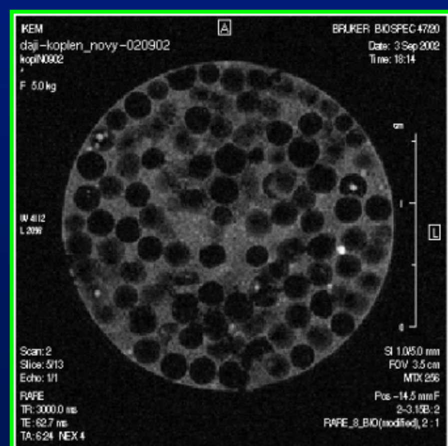
Best class 1 (#4)



Best class 2 (#24)



Missclassified class 1 (#1)

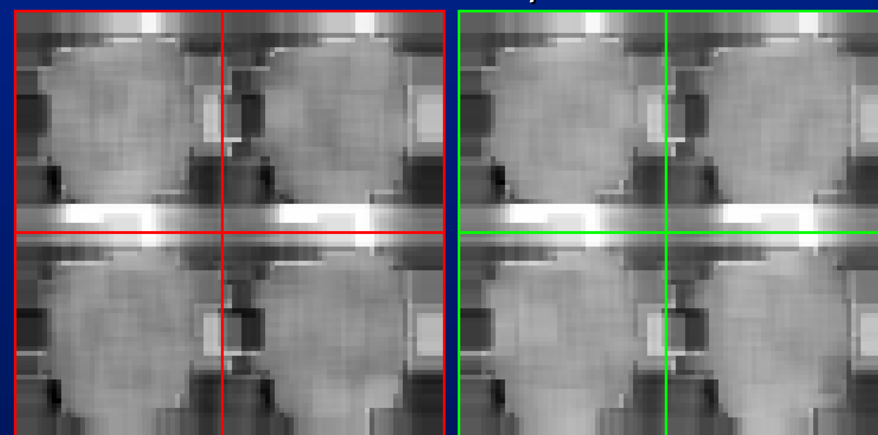


Missclassified class 2 (#17)

Differences in texture for images of „class 1" and „class 2" do exist.

Black spheres are distributed more uniformly in „class 2" images.

Theta1 maps



class 1 *class 2*

There are differences in phantom internal structure that lead to measurable difference in texture.

Conclusion

- Novel unsupervised technique for texture feature selection
- Gives the same result as supervised one (e.g. based on Fisher coefficient)
- Straightforward – simple math

- Objective texture discrimination
- Best features selection
- Dimensionality reduction
- Possible technique for texture homogeneity evaluation of test objects

- Future work: more systematic search through the feature space