Convert

Michał Strzelecki

COST B11, Brussels, 15-16 December 2000

Convert

Convert provides:

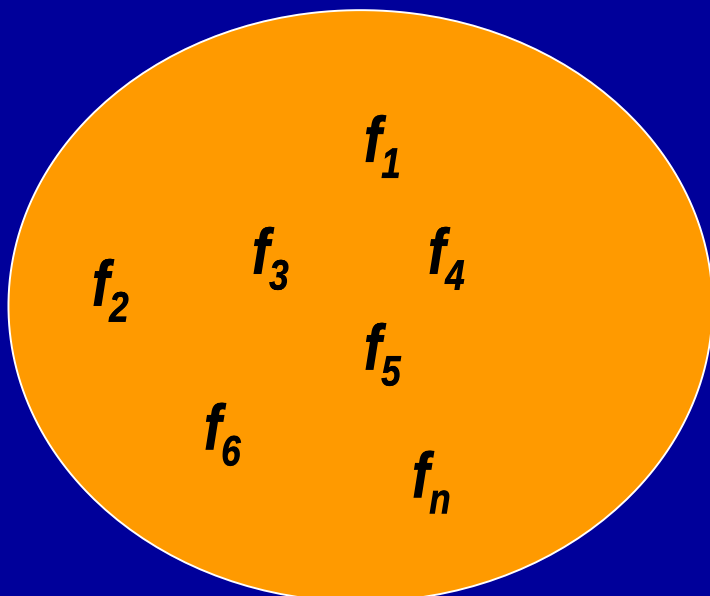
- appropriate file format for b11 (conversion of `*.par` files into `*.sel` file)
- feature reduction

Convert

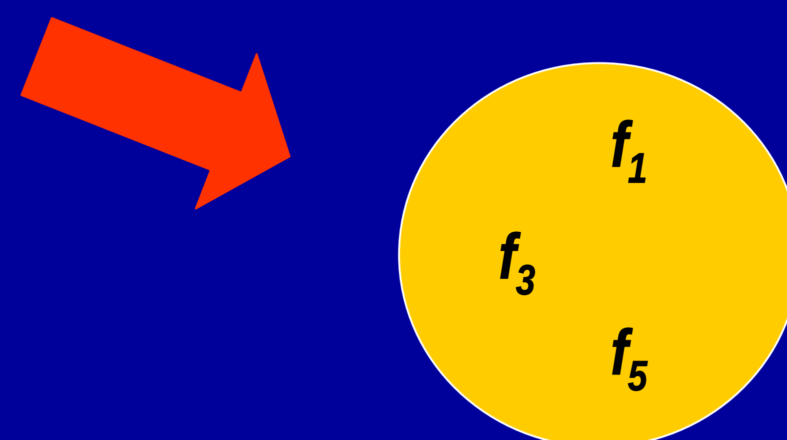
Why feature reduction?

- it is not known *a priori* which features are best for given texture analysis - one has to consider as many features as possible,
- it is very difficult to manage with over 250 features generated by MaZda,
- large number of features requires large number of data samples (which are not available normally).

Feature reduction by <u>selection</u>

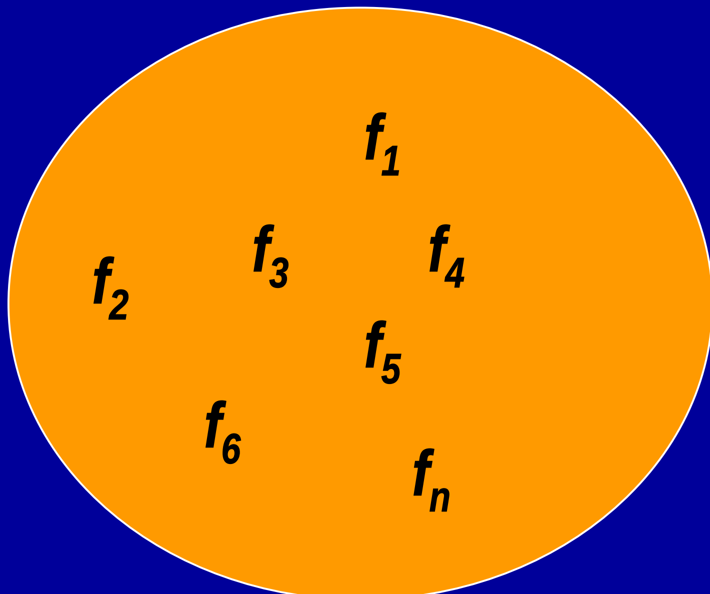


a subset of features is chosen based on given mathematical criterion

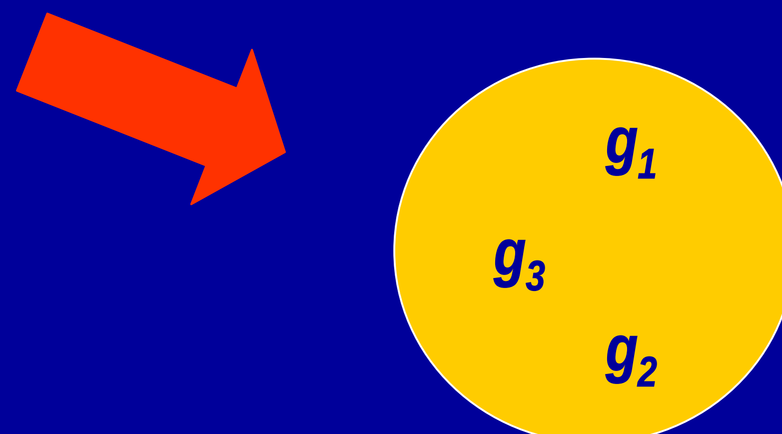


Criteria used in Convert: Fisher, POE, MDM

Feature reduction by extraction (projection)



an original set of features is transformed into different feature set, smaller than original



Transforms used in b11: PCA, LDA, NDA

Convert

Feature selection methods:

- Fisher coefficient
- minimisation of classification error probability
- multidimensional discrimination measure

Fisher coefficient

$$F = \frac{D^2}{V^2} = \frac{\dfrac{1}{1-\sum_{k=1}^{K} P_k^2} \sum_{k=1}^{K}\sum_{j=1}^{K} P_k P_j \left| \boldsymbol{\mu}_k - \boldsymbol{\mu}_j \right|^2}{\sum_{k=1}^{K} P_k V_k^2}$$

P_k - probability of given class k,

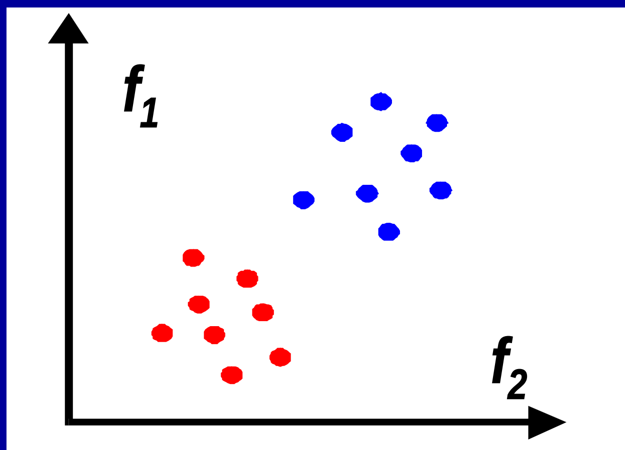
P_k=(# samples of class k)/(total # samples)

V_k , μ_k - variance and mean value for class k

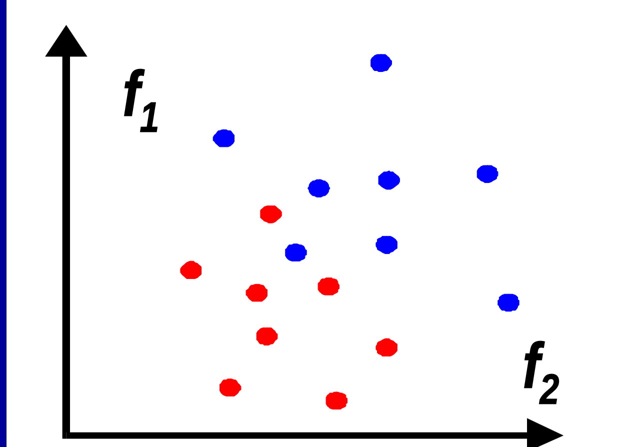
K - number of classes

Interpretation of Fisher coefficient

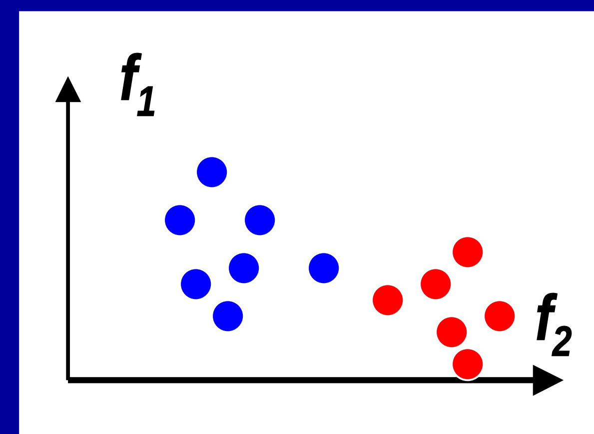
For F>6, perfect class separation can be expected



$F(f_1), F(f_2) \uparrow$

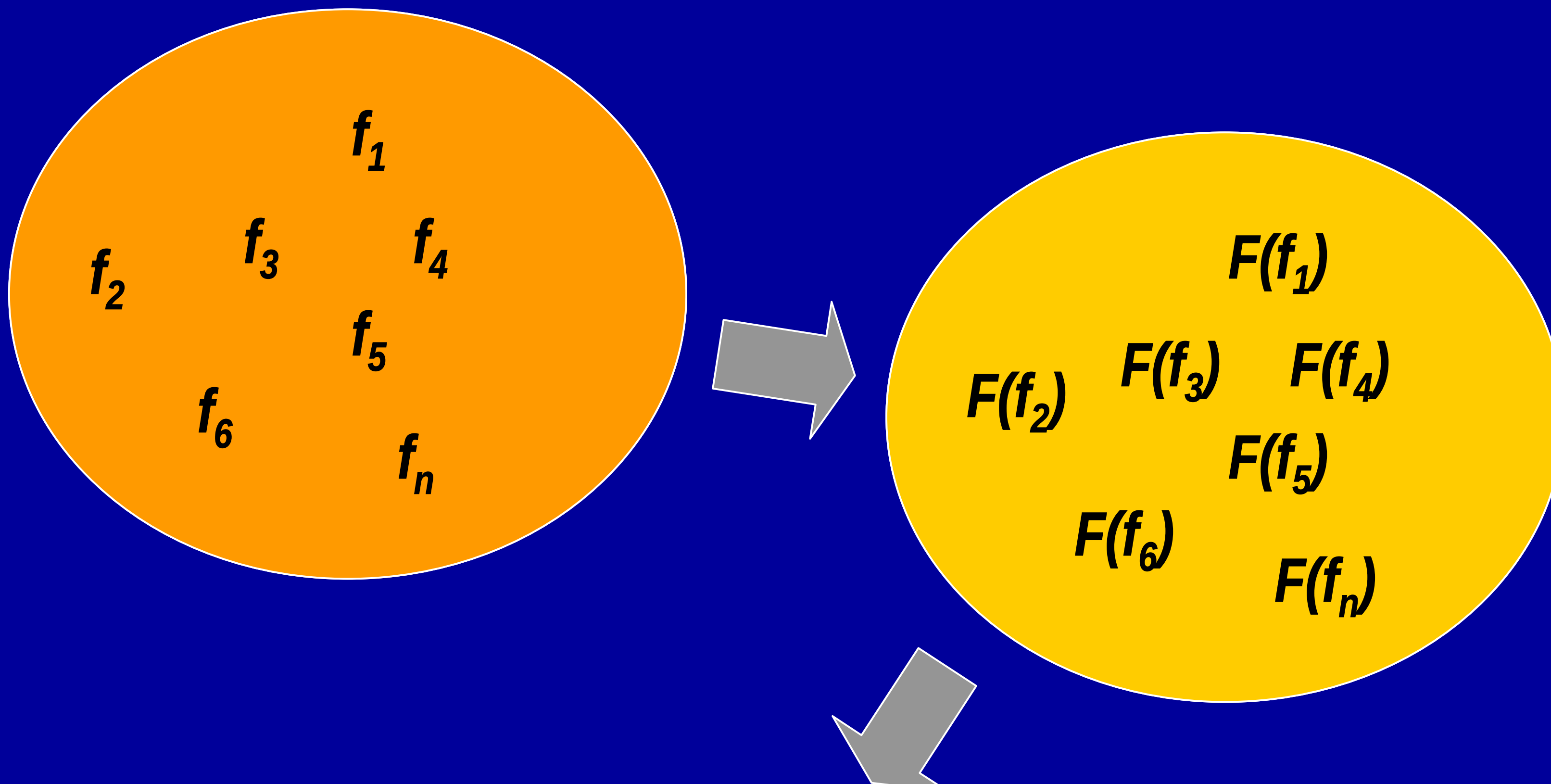


$F(f_1), F(f_2) \downarrow$



$F(f_1) \downarrow F(f_2) \uparrow$

Feature selection using Fisher coefficient

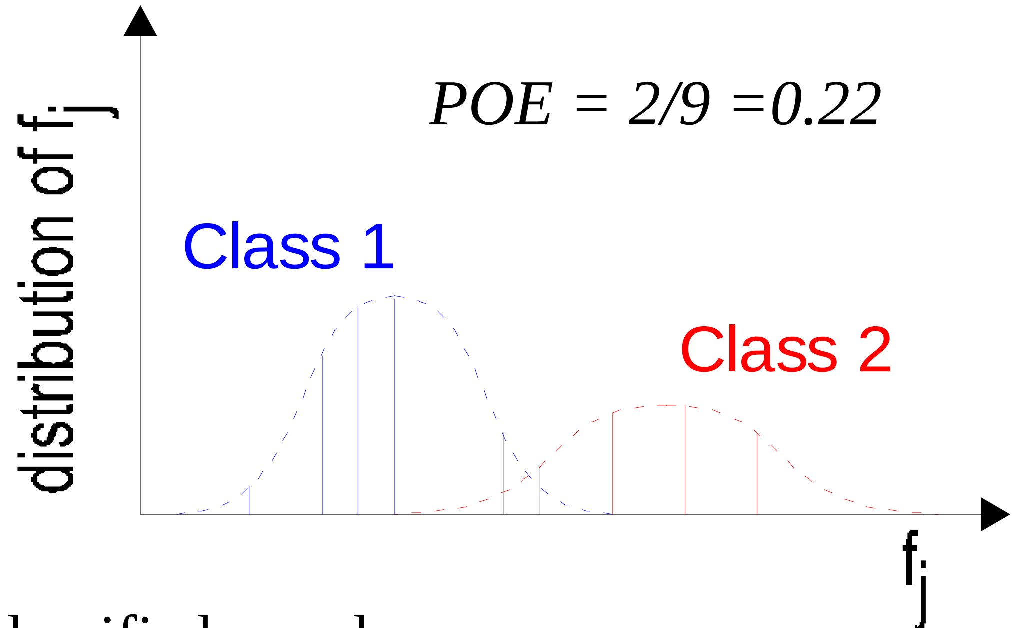


$$\underbrace{F(f_3) > F(f_5) > \dots > F(f_6)} > F(f_1) > \dots > F(f_n)$$

choice of *M* features with highest F coeff. value (*M*=10 in Convert)

Minimisation of classification error probability (POE)

$$f^1 = f_j : \min_j [POE(f_j)]$$



$$POE = \frac{\text{number of not correctly classified samples}}{\text{total number of samples}}$$

$$f^2 = f_j : \min_j [POE(f_j) + \left| CC(f^1, f_j) \right|]$$

$$f^n = f_j : \min_j [POE(f_j) + \frac{1}{n-1} \sum_{k=1}^{n-1} \left| CC(f^k, f_j) \right|]$$

Feature selection using POE method

<u>Step 1</u>

Find a feature with a minimum POE

<u>Step 2</u>

For remaining features, find a feature for which sum of its POE and correlation coefficient of this feature and feature selected in step 1 is minimal

<u>Step 3</u>

For remaining features, find a feature for which sum of its POE and mean correlation coefficient of this feature and features already selected in previous steps is minimal

<u>Step 4</u>

Repeat step 3 until appropriate number of features will be selected

Multidimensional discrimination measure (MDM)

- This measure is similar to Fisher coefficient.
- It is applied to the whole set of features, instead of one as in the case of Fisher coefficient.
- Higher value of this measure signifies better classification ability of given feature set.

Multidimensional discrimination measure (MDM)

For each feature f_i, a difference between MDM of whole feature set and MDM of whole feature set except f_i is calculated:

$$U_i = MDM(f_1, f_2, \ldots, f_{259}) - MDM(f_1, f_2, \ldots, f_{i-1}, f_{i+1}, \ldots, f_{259})$$

then 10 feature with highest U_i is selected.

High value of U_i indicates that feature f_i posses large discriminative power.

Multidimensional discrimination measure (MDM)

Calculation of MDM requires inversion of feature correlation matrix. If most of features are highly correlated, matrix is close to singular and U coefficients might be not accurate.