

Lecture 28

Classifiers:

Biomarker Development and Radiomics

MP574: Applications

Sean B. Fain, PhD (sfain@wisc.edu)

Diego Hernando, PhD (dhernando@wisc.edu)

ITK/VTK Applications: Andrew Hahn, PhD (adhahn@wisc.edu)

Learning Objectives

- Introduce common texture measures and shape-measures used to generate features
- Introduce generalized approach to biomarker development - “radiomics”
 - Strategy is to extend the data-driven approach to expand the dimensionality of the data
 - Multiple image contrast and quantitative measures
 - Almost totally data-driven after curation by expert readers

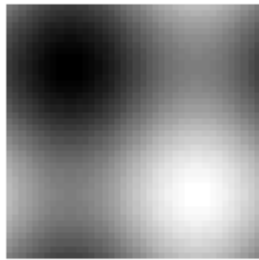
Features and Patterns

- Patterns are ordered sequences of measurements (e.g. shape and texture measures. Could also be functional perfusion etc...)
- Example pattern matrix:

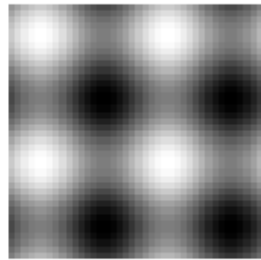
| <i>object</i> | <i>feature1</i> | \dots | <i>feature n</i> | |
|---------------|-----------------|---------|------------------|---------------|
| 1 | i | $i+1$ | i_n | Pattern x_1 |
| \vdots | . | . | . | Pattern x_2 |
| n | . | . | . | Pattern x_n |

Characterization

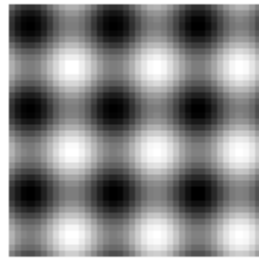
How can we characterize these images?



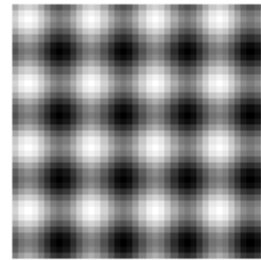
CV = 0.05



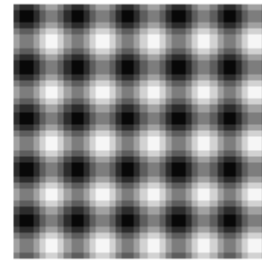
CV = 0.05



CV = 0.05



CV = 0.05



CV = 0.05

CV (Coefficient of Variation) = STD/mean

Textural features

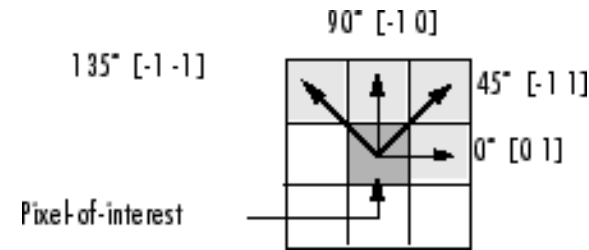
- **First order** texture measures are statistics calculated from the original image values, like variance, and do not consider pixel neighbor relationships.
- **Second order** measures consider the relationship between groups of two (usually neighboring) pixels in the original image.
 - Gray Level Co-occurrence Matrix (GLCM)
 - Run Length Matrix (RLM)
- **Higher order** textures (considering the relationships among three or more pixels) are theoretically possible but not commonly implemented due to calculation time and interpretation difficulty. There has been some recent development of a more efficient way to calculate third-order textures.
 - Neighborhood Gray-Tone-Difference Matrix (NGTDM)

Gray Level Co-occurrence Matrix

A gray level co-occurrence matrix (GLCM) contains information about the positions of pixels having similar gray level values. The GLCM is defined by:

$$P_{ij}(d, \theta) = n_{ij}$$

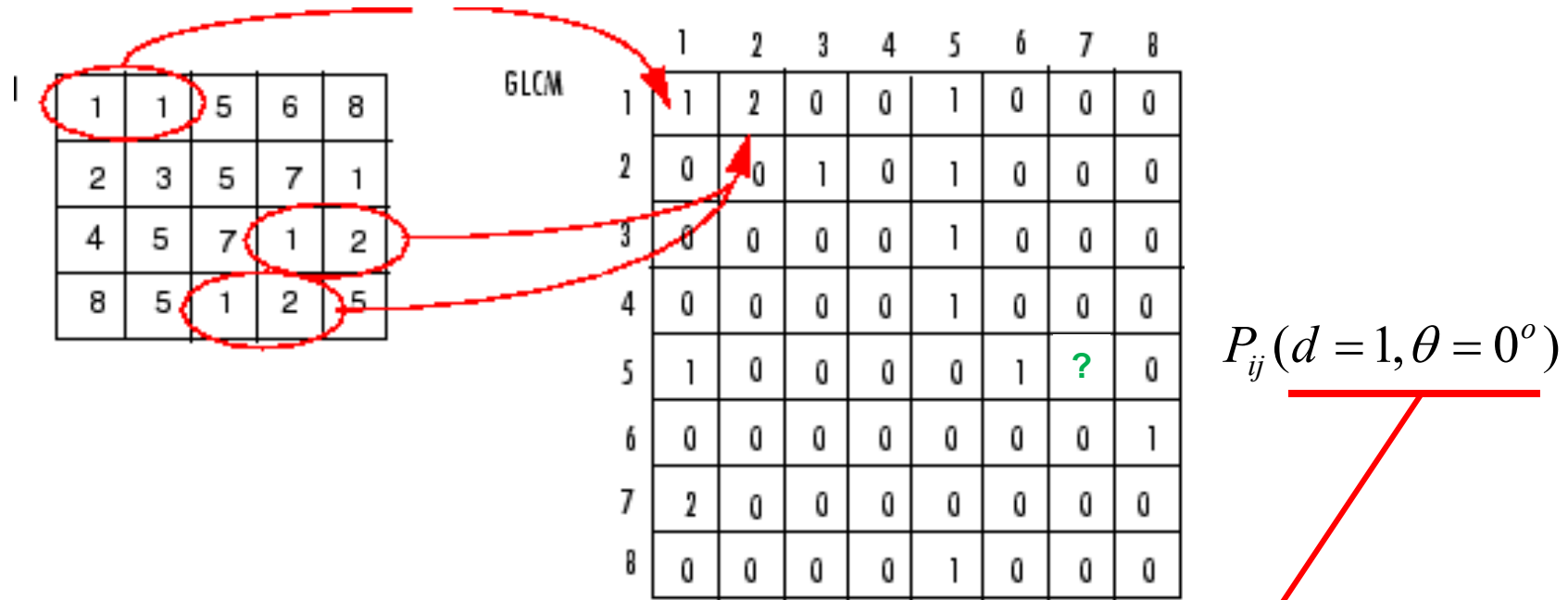
[alternatively, $P(i, j|d, \theta)$]



where n_{ij} is the number of occurrences of the pixel values (i, j) lying at distance d in the direction θ .

The co-occurrence matrix P_{ij} has dimension $n \times n$, where n is the number of gray levels in the image.

Gray Level Co-occurrence Matrix



$$P_{ij}(d=1, \theta=0^\circ)$$

```
GLCM = graycomatrix(I, 'GrayLimits', [1 8], 'Offset', [0 1]);
```

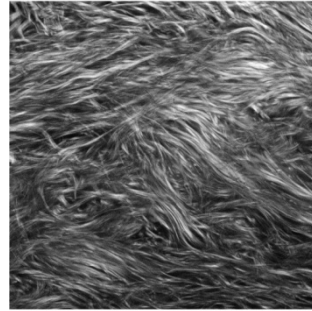
Statistics derived from GLCM

- Energy: $f_1 = \sum_i \sum_j p_{ij}^2$
- Entropy: $f_2 = -\sum_i \sum_j p_{ij} \ln p_{ij}$
- Homogeneity: $f_3 = \sum_i \sum_j \frac{p_{ij}}{1 + (i - j)^2}$
- Contrast: $f_4 = \sum_i \sum_j p_{ij} (i - j)^2$
- Dissimilarity: $f_5 = \sum_i \sum_j p_{ij} |i - j|$
- Correlation: $f_6 = \sum_i \sum_j \frac{(i - \mu_i)(j - \mu_j)}{\delta_i \delta_j} \quad \left\{ \begin{array}{l} \mu_i = \sum_i \sum_j i p_{ij} \\ \delta_i^2 = \sum_i \sum_j i^2 p_{ij} - \mu_i^2 \end{array} \right.$
- ...

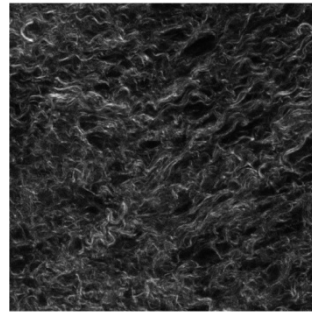
Typically d=1, averaged over all four directions in 2D

SHG images of breast tissues

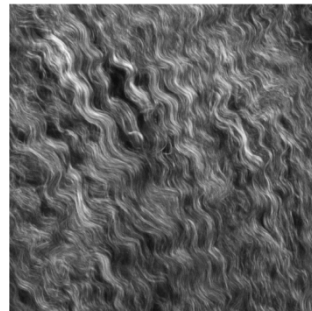
Normal



High risk



Cancer

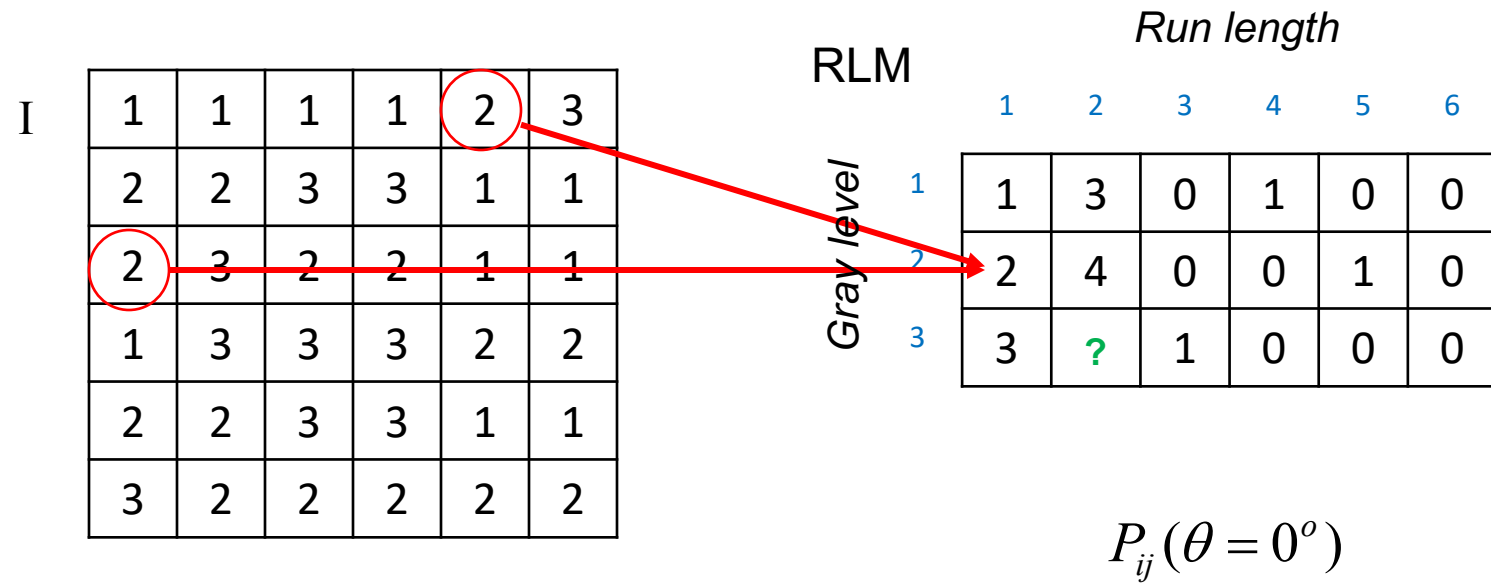


Textures

| | Normal | High Risk | Cancer |
|---------------------|--------|-----------|--------|
| autocorrelation | 13.55 | 5.39 | 13.22 |
| contrast | 0.59 | 0.35 | 0.47 |
| correlation | 0.89 | 0.76 | 0.86 |
| cluster prominence | 81.94 | 18.18 | 42.16 |
| cluster shade | 7.32 | 2.34 | 3.88 |
| dissimilarity | 0.48 | 0.32 | 0.42 |
| energy | 0.14 | 0.30 | 0.19 |
| entropy | 2.58 | 1.82 | 2.29 |
| homogeneity | 0.86 | 0.88 | 0.88 |
| maximum probability | 0.27 | 0.51 | 0.34 |
| sum of squares | 13.59 | 5.45 | 13.24 |
| sum average | 7.03 | 4.47 | 7.05 |
| sum variance | 29.64 | 10.47 | 30.34 |
| sum entropy | 2.13 | 1.56 | 1.93 |
| difference variance | 0.59 | 0.35 | 0.47 |
| difference entropy | 0.84 | 0.67 | 0.76 |

Texture analysis can be performed after wavelet or curvelet transformation

Run Length Matrix

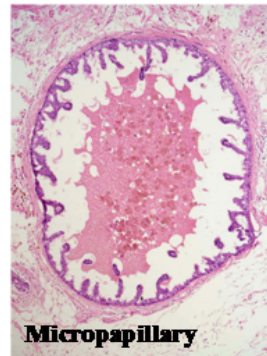


Statistics derived from RLM

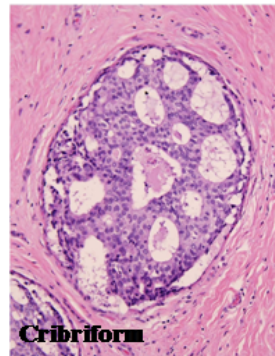
- Short Run Emphasis: $SRE = \sum_i \sum_j \frac{p_{ij}^2}{j^2}$
- Long Run Emphasis: $LRE = \sum_i \sum_j p_{ij} j^2$
- Low Gray-Level Run Emphasis: $LGRE = \sum_i \sum_j \frac{p_{ij}}{i^2}$
- High Gray-Level Run Emphasis : $HGRE = \sum_i \sum_j p_{ij} i^2$
- Short Run Low Gray-Level Emphasis: $SRLGE = \sum_i \sum_j \frac{p_{ij}}{i^2 j^2}$
- Short Run High Gray-Level Emphasis: $SRHGE = \sum_i \sum_j \frac{p_{ij} i^2}{j^2}$
- ...

Morphological classification of human breast cancer

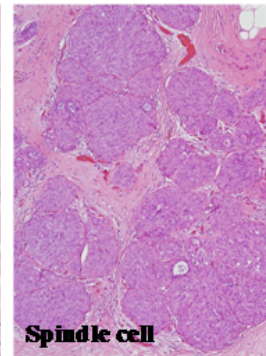
| # | Morphological classification | # | Morphological classification |
|----|------------------------------|----|------------------------------|
| 1 | Inflammatory | 11 | Clear cell |
| 2 | Pregnancy-associated | 12 | Medullary |
| 3 | Comedo | 13 | Secretory |
| 4 | Micropapillary | 14 | Signet ring cell |
| 5 | Papillary | 15 | Mucinous |
| 6 | Cribriform | 16 | Tubular |
| 7 | Solid | 17 | Lobular neoplasia |
| 8 | Clinging | 18 | Mixed cell types |
| 9 | Spindle cell | 19 | Apocrine |
| 10 | Neuroendocrine | 20 | Malignant myoepithelioma |



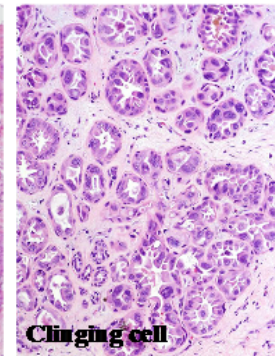
Micropapillary



Cribriform

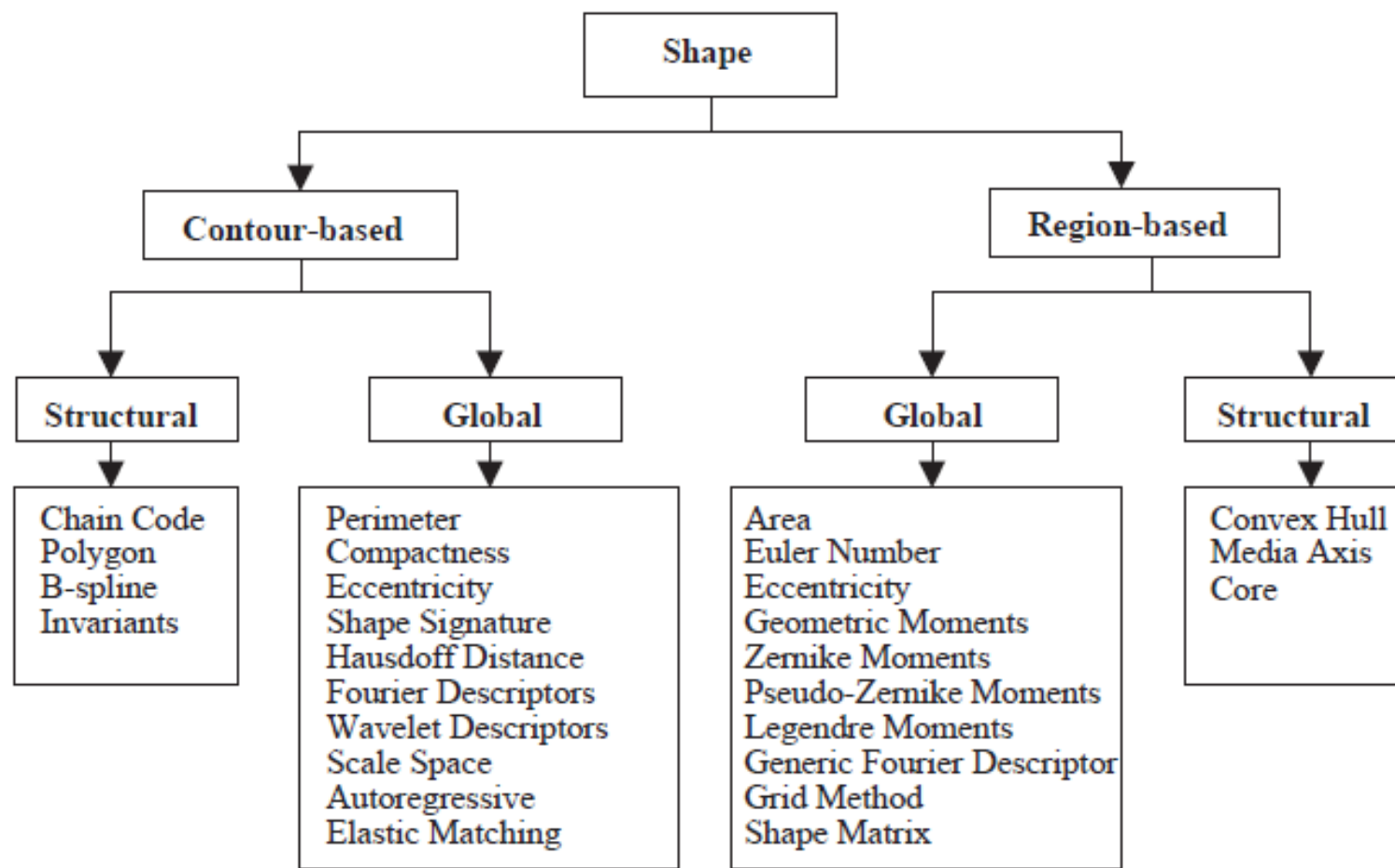


Spindle cell



Clinging cell

Shape Descriptors



Chain Code

A chain code is a lossless compression algorithm for monochrome images

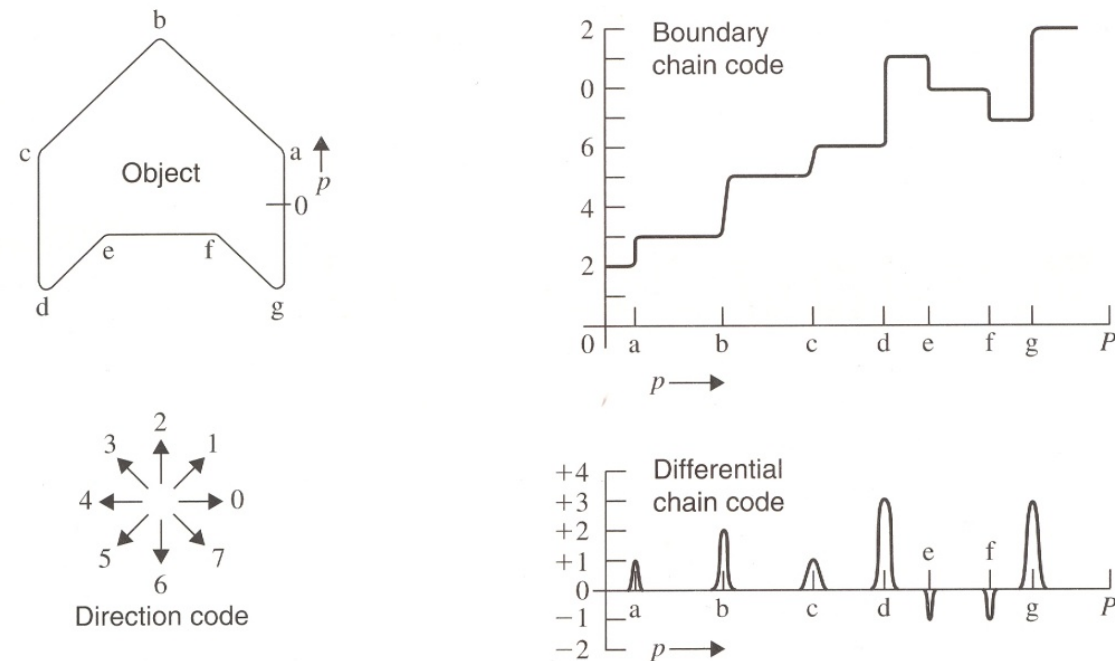
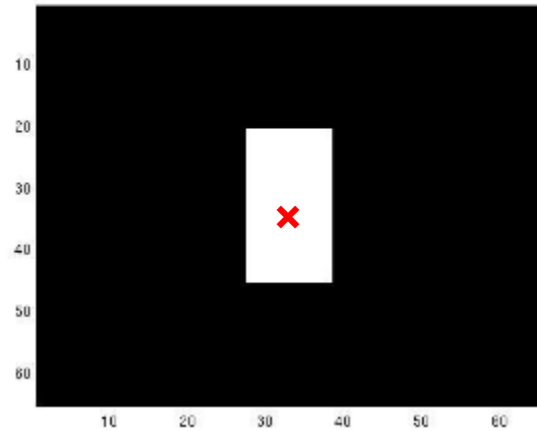


Figure 19-6 The chain code and its derivative

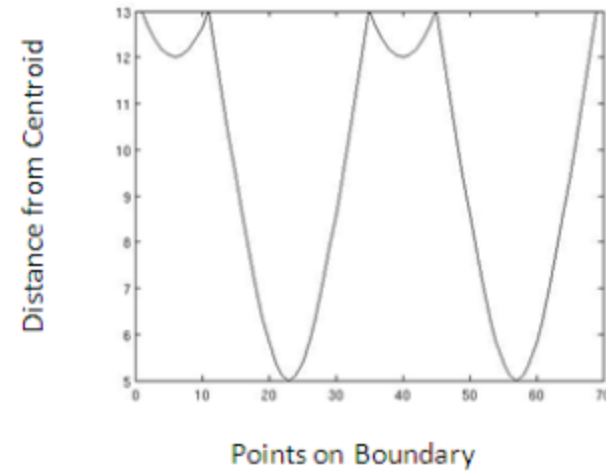
Shape Signature

The 1D functional representation of the 2D shape boundary is called **shape signature**.

Original Rectangle



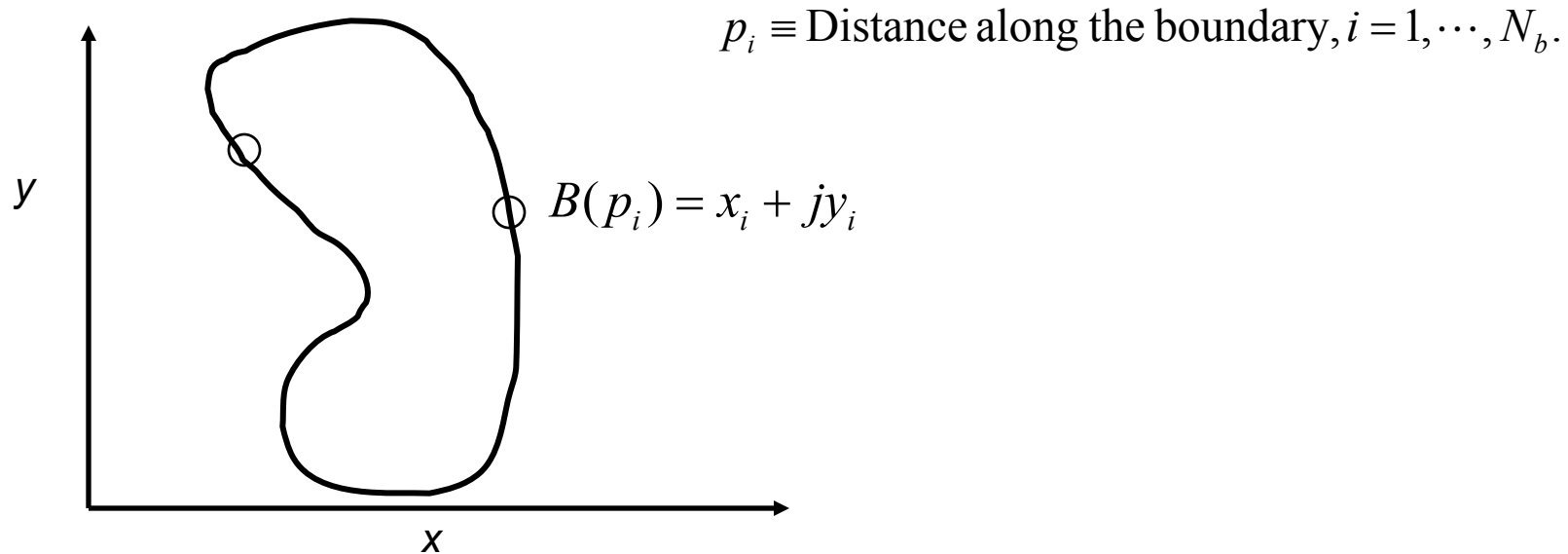
Shape Signature of Original Rectangle



Translation invariant

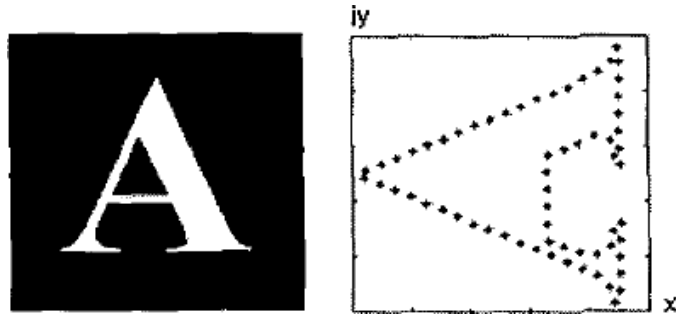
Fourier Descriptors

Fourier descriptors are **invariant** to translation, rotation, scaling, and change of starting point



Fourier transform of $B(p_i)$ results in a periodic function with coefficients characteristic of object shape.

Example of Fourier Shape Descriptor

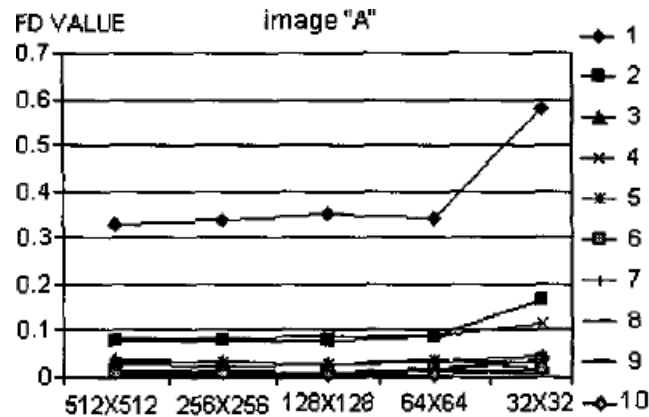


$$s(n) = x(n) + iy(n), \quad n = 0, 1, 2, \dots, N - 1$$

$$z(k) = \frac{1}{N} \sum_{n=0}^{N-1} s(n) e^{-i2\pi kn/N}, \quad k = 0, 1, \dots, K - 1$$

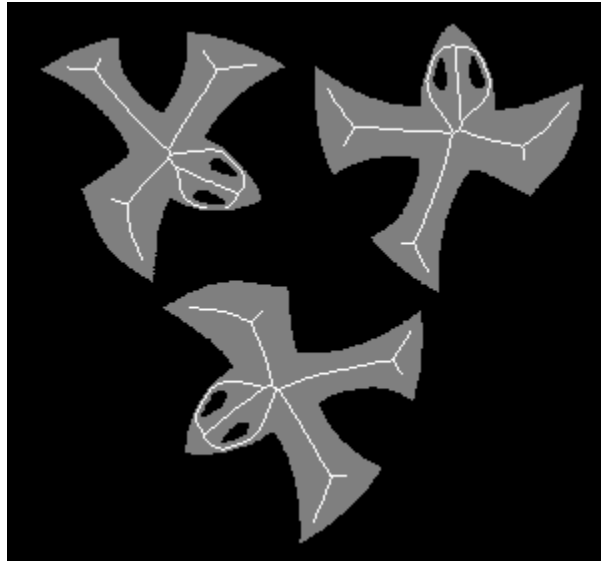
Can be made “invariant” to rigid transformations by normalizing to low frequency terms:

$$c(k - 2) = \frac{|z(k)|}{|z(1)|}, \quad k = 2, 3, \dots, K - 1$$



Skeleton Descriptors

Skeleton is an important topological descriptor of a 2D binary object



```
BW2 = bwmorph(BW,'skel',Inf);
```

- Average length of the skeleton of the object
- Ratio of number of points on the skeleton to number of points inside the object
- Ratio of number of points on the skeleton to number of points on the boundary of the object
- Ratio of number of branch points to number of points on the skeleton

Moments

Recall that the set of moments of a bounded function $f(x, y)$

$$M_{jk} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^j y^k f(x, y) dx dy$$

$$\mu_{jk} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^j (y - \bar{y})^k f(x, y) dx dy$$

$$M_{00} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy$$

Central Moments

$$\bar{x} = \frac{M_{10}}{M_{00}}, \quad \bar{y} = \frac{M_{01}}{M_{00}}$$

Center of gravity of the object

Normalized Moments for Multi-scale Object Recognition



- Central moments are computed using the centroid of the image object
 - Equivalent to an image whose center has been shifted to coincide with the object centroid
 - Implies invariant to translations
- Consider scaling by:
- $\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, S_x = S_y = \alpha$

- Consider moments under scaling:

$$f(x', y') = f(\alpha x, \alpha y); x' = \alpha x, y' = \alpha y$$

$$\mu'_{pq} = \iint (\alpha x - \alpha \bar{x})^p (\alpha y - \alpha \bar{y})^q f(x', y') dx' dy'$$

$$\begin{aligned} \mu'_{pq} &= \alpha^{p+q+2} \mu_{pq} \\ \mu'_{00} &= \alpha^2 \mu_{00} \end{aligned}$$

- Normalized central moments:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \text{ where } \gamma = (p + q + 2)/2; p + q = 2, 3, \dots$$

Invariant Moment Descriptors

A set of 7 **invariant moments** (IM) which are invariant to rotation, scaling and translation are given by Hu

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

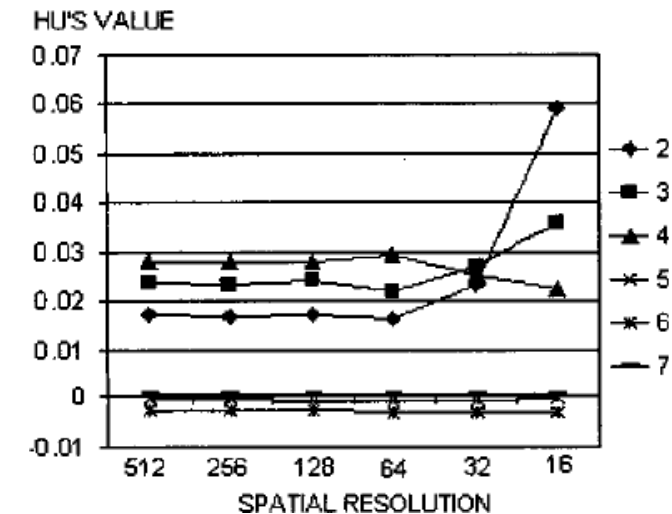
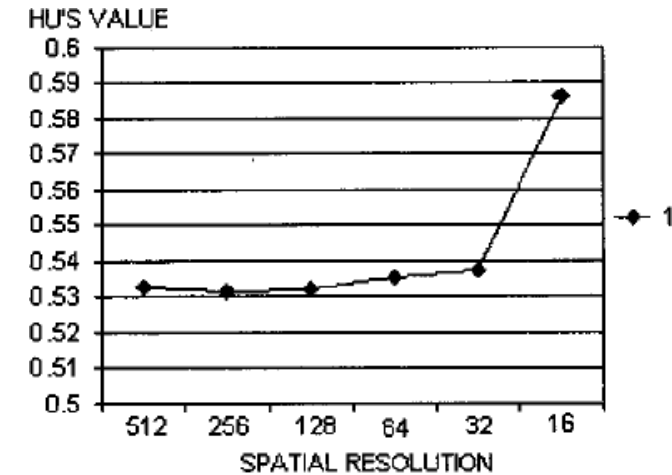
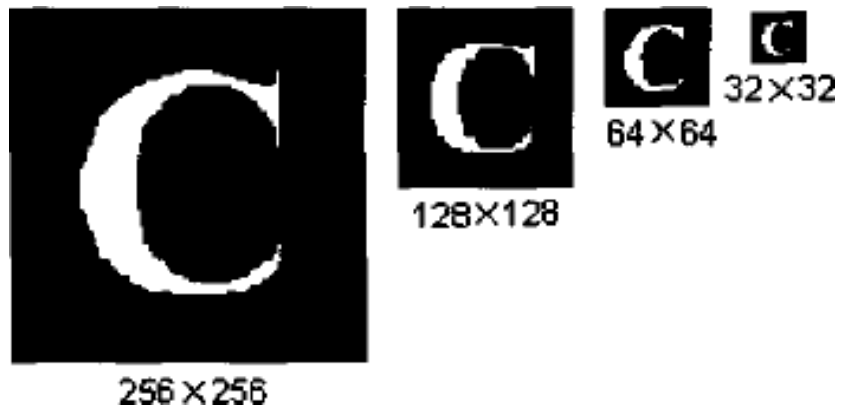
$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ + 4\eta_{11}^2(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

where $\eta_{pq} = \mu_{pq} / \mu_{00}^\gamma$ and $\gamma = 1 + (p + q)/2$ for $p + q = 2, 3, \dots$

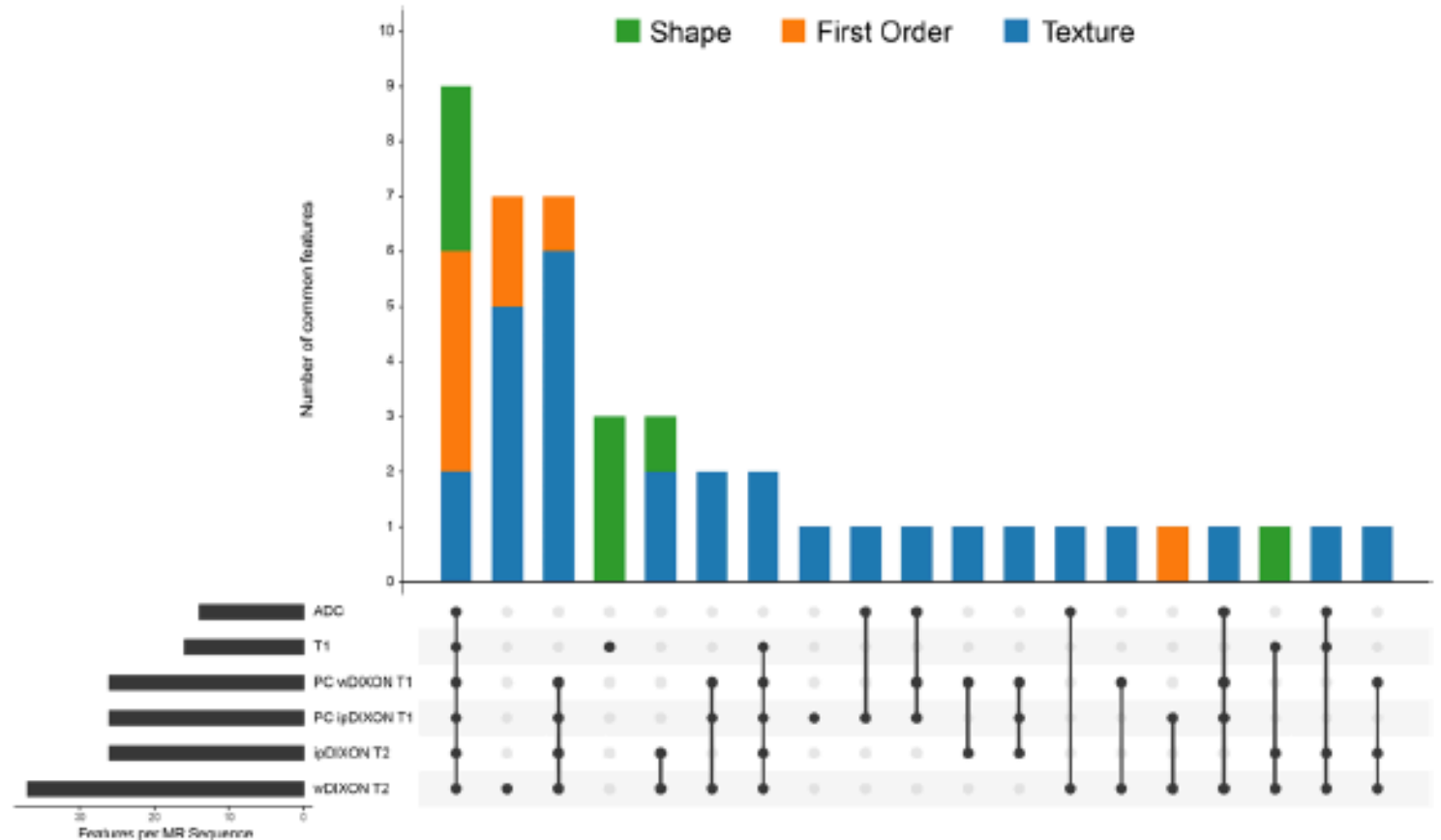
- High degree of information redundancy.
- Higher-order moments are very sensitive to noise

Example of Fourier Shape Descriptors vs. Moments for Multi-scale Object Recognition



Biomarker Development Process

1. Develop curated dataset
 - Expert readers manually delineate disease features
 - Preferably multi-parametric datasets (MRI contrasts)
2. Identify features
 - Large number of candidate features
 - Use statistical criteria to identify reproducible features (ICC, CCC)
3. Identify non-redundant (un-correlated) features clusters
 - Step-wise regression models
 - Spearman correlation matrix



Intra-Class and Cross-Correlation Coefficients

- Used to identify thresholds for feature and cluster selection
- Intra-class correlation coefficient:

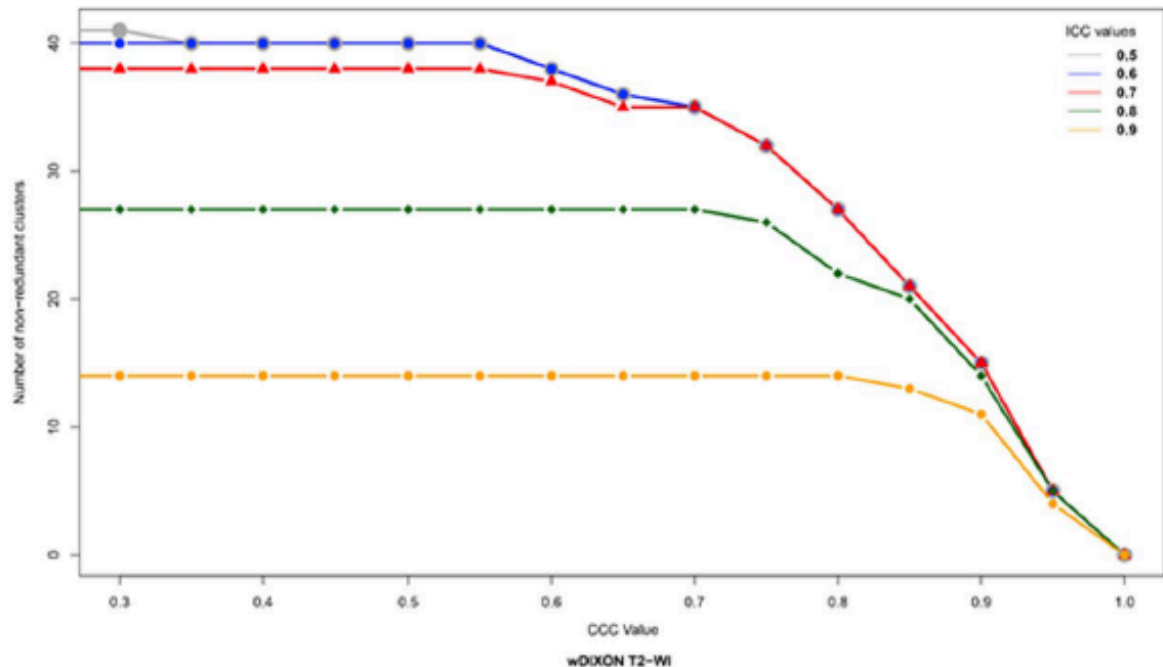
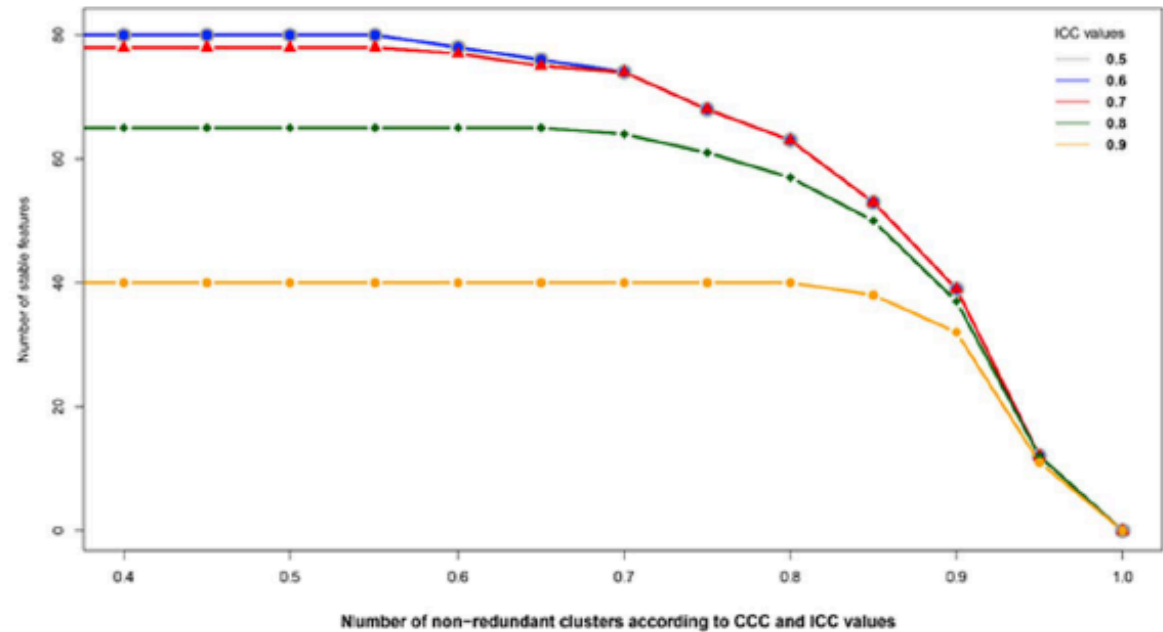
$$ICC = \frac{Cov(x_1, x_2)}{\sigma_{x_1}^2 + \sigma_{x_2}^2}$$

- Proportion of total variance that is "between groups."
- Reproducibility

- Cross-correlation coefficient:

$$\rho = \frac{Cov(x_1, x_2)}{\sigma_{x_1} \sigma_{x_2}}$$

- Pearson correlation coefficient
- Redundancy



Feature Selection and Hierarchical Clustering

Deciding on the variables that provide the best prediction

1. Best model: search for all possible models and take the one with the highest R^2 (lowest Mallows' C_p)
2. Stepwise regression: useful when the number of predictor variables is large.
 - a. Backward elimination: start with all variables, test them one by one and remove any that are not significant.
 - b. Forward selection: start with no variable in the model and try out one by one and include them if they are statistically significant.

`stepwise()`, `stepwisefit()`

Summary

- Biomarker Development and Radiomics (Data-Mining)
- Strengths
 - Expands dimensionality of the data set to improve diagnostic power
 - Agnostic, data-driven reduced subjective bias imposed by more deterministic approaches
- Limitations
 - Features drawn from the data are typically highly correlated
 - Statistical clustering approaches to improve conditioning and optimize classification
 - Requires large amounts of high quality training data to improve accuracy