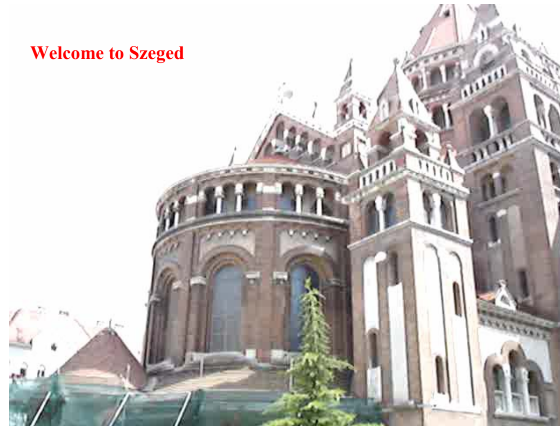
SSIP 2005 Szeged July 1st

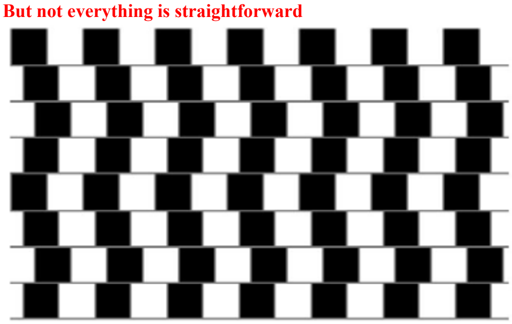
Methods for the analysis of a series of image in time.

Andrew Todd-Pokropek
University College London,
A.Todd@ucl.ac.uk

Welcome to Szeged



But not everything is straightforward



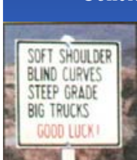
Are the horizontal lines parallel or do they slope?

Acknowledgements

- Image Science group UCL (CS and Medphys)
- IRC (UCL, KCL, Oxford, Manchester, IC)
- Yale (J. Duncan)
- and many others

Outline

- Segmentation, feature detection, NRR
- Examples of temporal series
- Reductions of dimensionality
  - PCA based methods
  - ICA based methods
- Change detection (Kalman filtering)
- Conclusions and the future



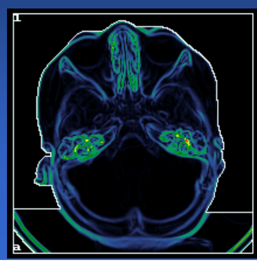
A road map

What is segmentation?

Engineers v. Scientists (Pure and Applied)

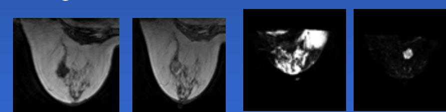
Segmentation

- Classic 'edge detection' methods
  - Gradient (Sobel etc), zero crossings of Laplacian
  - Canny
  - Marr Hildreth
- Phase congruency
- Model based
  - Medial axis
  - Active shape
- Clustering
  - Split merge
  - K-Means
  - Affinity
- etc



Three research strands

- Non-rigid registration
  - change detection
  - voxel-based morphometry
  - segmentation



Pre-contrast Post-contrast Subtract Subtract NRR

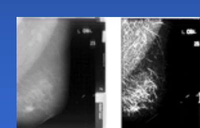
2nd strand

- Non-rigid registration
- Shape and appearance models
  - segmentation
  - normal variation and pathology

Object and template Canny edged image Find Correspondence

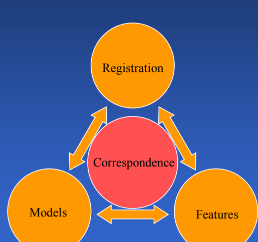
3rd strand

- Non-rigid registration
- Shape and appearance models
- Feature detection
  - 'interesting' structure
  - abnormal structure



Mammogram Linear features

A Unified View

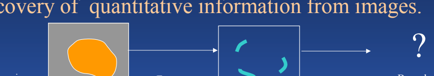


Underlying unity not currently exploited

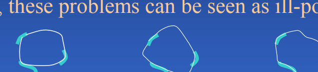
- Models ⇔ Registration
  - NRR to build models
  - models to constrain NRR
- Registration ⇔ Features
  - features to improve NRR
  - NRR defines corresponding features
- Features ⇔ Models
  - features to enrich models
  - models to locate features

Models in Image Analysis

- Lack of image quality and/or features often limit the recovery of quantitative information from images.



- Overall, these problems can be seen as ill-posed



- Models can help constrain solutions in plausible ways:

The Deformable contour model

- Deformable contour model (or “snake”) can be represented by a set of controls points developed through the solution of energy minimization using variational calculus
- This model requires initial control points which roughly delineate the volume of interest on several slices
- New Control points on each slice are generated from cubic spline interpolation to obtain continuity and smoothness

Contd. Deformable contour model

The total energy of snake can be represented by

$$E_{total} = \int_0^1 E_{snake}[v(i)]\,di = \alpha E_{int}[v(i)] + \beta E_{image}[v(i)]$$

The internal energy is

$$\alpha E_{int}[v(i)] = \alpha_1 |v_i - v_{i-1}|^2 + \alpha_2 |v_{i-1} - 2v_i + v_{i+1}|^2$$

The external energy is

$$\beta E_{image} = \beta_1 E_{int\ en}[v(i)] + \beta_2 E_{grad}[v(i)]$$

In this process, modified greedy optimisation technique has been used

Snakes

- Balloons
- Shrink wrapping
- Gsnakes
- Tsnakes
- 2-D to 3-D

Deformable model

- GVF (Gradient Vector Flow) [5] image forces

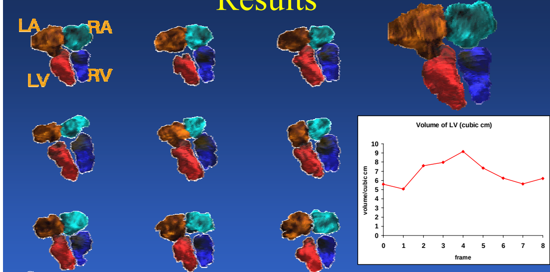
$$\underline{F}^i = \underline{F}^i_{tension} + \underline{F}^i_{bending} + \underline{F}^i_{GVF}$$

$$\underline{F}^i_{GVF} = \underbrace{\mu(|\nabla f|)\nabla^2 \underline{u}}_{\text{Diffusion of edge map in absence of local edges}} - \underbrace{|\nabla f|^2(\underline{u} - \nabla f)}_{\text{Local edges}}$$

- Creation of GVF field
  - Gaussian intensity distribution within blood pool yields initial classification of boundary.
  - GVF field created from anisotropic diffusion of edge boundaries.

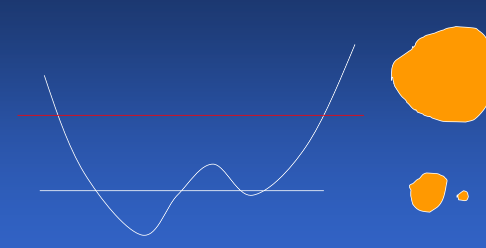
[5] Prince and Xu, 1997

Results



Comments
- Partial volume effect causes fusion of intra chamber structures with myocardium
- RMS projected errors to manual tracings are below 2.5mm (2 datasets).

Level Sets



Relationship to deformable contours

- Ignores the tension component
- $\varphi(X(s,t),t) = 0$ the level set at zero

Differentiate:

- $\delta\varphi/\delta t + \text{grad } \varphi . \delta X/\delta t = 0$

Fuzzy Classification and Fuzzy Connectedness

- Segmentation and classification
- Classification can lead to segmentation and vice-versa.
- Classification refers to the labelling of pixels in an image that may result in the segmentation of objects or regions.
- The grey level intensity value is the most common feature. Texture is an alternative.
- Pixels with similar feature vectors form clusters in the feature space that can be separated by lines or curves.
- In reality, partitioned regions do overlap at the border and the classes are not separable which brings fuzzy Clustering
- Fuzzy membership functions has been assigned a pixel to classes with any value between 0 and 1.
- Any pixel can be assigned to more than one class simultaneously where the membership value of a pixel i to each class k is

$$\mu_{ik} \in [0,1] \ \ and \ \sum_k \mu_{ik} = 1$$

Fuzzy Clustering works as follows:

1. Initialise c cluster centres
2. Begin iteration

   i. Calculate distance function

   $$d_{ik} = \|x_i - m_k\|$$

   ii. Assign a fuzzy membership value to each pixel x_i for each cluster

   $$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{ij}} \right)^{2(m-1)}}$$

   iii. Re-calculate cluster centres

   $$m_k = \frac{\sum_{i=1}^{n} (\mu_{ik})^m x_i}{\sum_{i=1}^{n} (\mu_{ik})^m}$$

   iv. At each iteration, recalculate the membership value

3. Stop iteration when appropriate stopping criterion is satisfied.

Contd. Fuzzy clustering and fuzzy connectedness

- The overall objective function by this classification process is given by

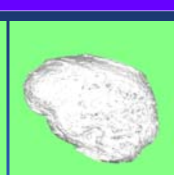
$$F_m = \sum_{i=1}^{n} \sum_{k=1}^{c} (\mu_{ik})^m (\|x_i - v_k\|)^2$$

- Using spatial information, image elements that constitute a region can be called as "accumulated voxels". These voxels can be determined by the similarity of image elements and of intensity-based features associated with image elements as well as by their spatial connectivity.
- Fuzzy connected object is that object in an image where every pixel is spatially adjacent, homogenous in pixel intensities and their fuzzy membership values are high.
- An image element will be considered to belong to that object whose strength of connectedness is highest.

Examples of WM, GM and CSF which are segmented by applying relative fuzzy connectedness are shown below:

(a) Segmented volume WM

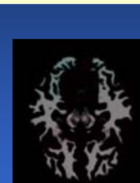
(b) Segmented volume GM



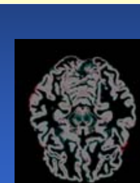
(c) segmented volume CSF

Comparison between segmented matter with simulated segmented matter

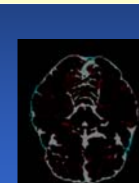
To create a colour overlay model to make an objective comparison by merging segmented WM with simulated WM, segmented GM with simulated GM and segmented CSF with simulated CSF



(a) Matched WM



(b) Matched GM



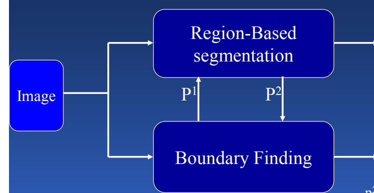
(c) matched CSF

Model based fitting

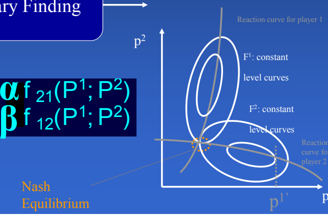
- Shape Priors: Cootes and Taylor (IPMI93) ; Grenander/ Miller (atlases/templates- 1991) ; Vemuri, et al. (MedIA97); Leventon, Grimson, et al. (CVPR00)
- Integrated Methods: Region grow w/ edges- Pavlidis and Liow (PAMI91); Zhu and Yuille (ICCV95) ; Ahuja (PAMI96) ; level sets - Tek and Kimia (ICCV95)
- Segmenting Cortical Gray Matter: Macdonald and Evans (SPIE95) ; Davatzikos and Prince (TMI95); Davatizikos and Bryan (TMI96); Teo and Sapiro (TMI97) ; Xu and Prince (MICCAI98) ;
- Multiple Objects/Level Sets and Priors: Tsai, Wells, Grimson, Willsky (IPMI03); Leventon, Grimson,Faugeras (CVPR00);

Integrated Segmentation via Game Theory

(Chakraborty & Duncan, PAMI 99)



$$F^1(P^1;P^2) = f_1(P^1) + \alpha f_{21}(P^1;P^2)$$
$$F^2(P^1;P^2) = f_2(P^2) + \beta f_{12}(P^1;P^2)$$

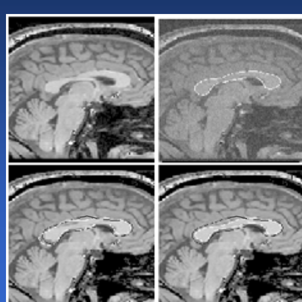


Corpus Callosum Result

Sagittal MR (1mm³ gradient echo)

Black = initial contour

White = gradient-based boundary finding



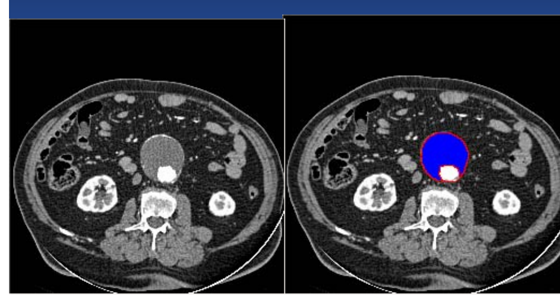
Expert-traced contour

Black = initial contour

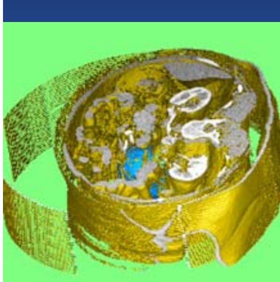
White = game-theory result

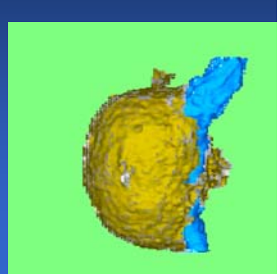
Ack. Duncan

Watersheds

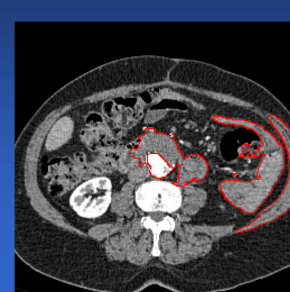


3D task



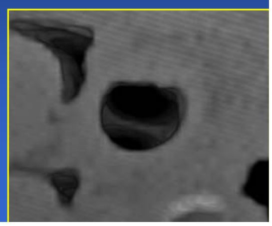


Leakage



Colon segmentation



Ack Sorantin


Key question: how much
does the tumour invade
The colonic wall.

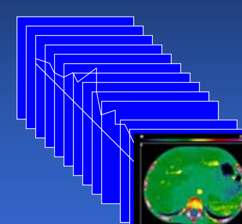
CT Lung lesions

Aims and objectives

- Automating analysis of multiple slices
- Isolating lung field
- Identifying structures
- Eliminating blood vessels and airways
- Classification of nodules on 3-D
- Determination of extent in 3-D
- Tracking in time and finding correspondences
- Problem is false positives

Handling of temporal data

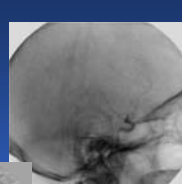
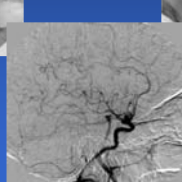
Special class of 3-D data processing
- Looking for change
- Looking for (derived) function

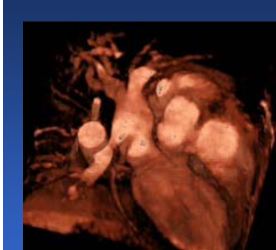


Consider set of time curves for every pixel
- Dual curve/image data set

Weather Satellites

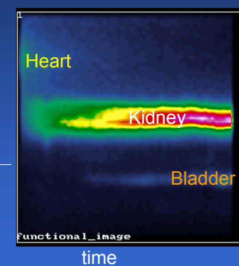
Looking for change


Before

After

Subtraction



Data compression/ projection

- Removing redundancy
- Reducing dimensionality
- Projection against
  - time (summation)
  - y (vertical axis)
  - oblique
- Constraints are required (a priori information)



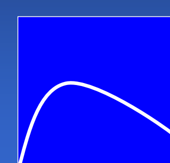
Function fitting

- For cyclical function
  - $A[i,j] = \Sigma_k \; C[i,j,k] \; \cos(\omega k)$
  - $B[i,j] = \Sigma_k \; C[i,j,k] \; \sin(\omega k)$

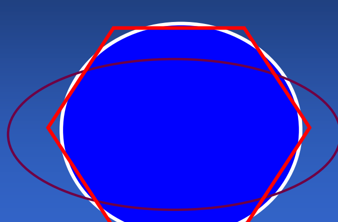
  - $AMP[i,j] = sqrt (A[i,j]^2 + B[i,j]^2)$
  - $PHASE[i,j] = tan_{-1} (B[i,j] / A[i,j])$
- First Fourier component

Functional Images

- Image of a derived function
  - Rate of increase/ decrease
  - Time to max
  - Variance image
- Example Kalman filtering
  - Estimating current values
  - And statistical model

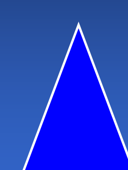


Fourier Contours



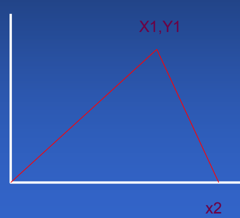
Active Shape Model A simple example

- A set of triangles
- Characterised by 6 parameters
  - {x1,y1,x2,y2,x3,y3}
- Simpler description possible



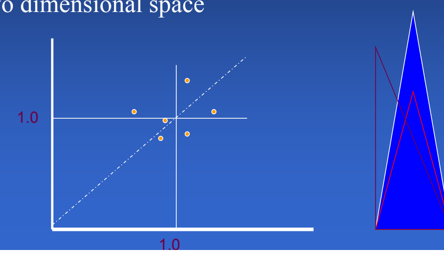
Triangle space

- Fix the origin
- Fix the x axis
- 3 parameters
- Also lengths of 3 sides



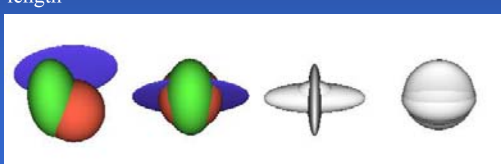
Normalise scale

- Take ratios of lengths of 1 side v. other two
- Two dimensional space



Where to place the nodes

- Regular
- Maximum radius of curvature
- Minimal description length

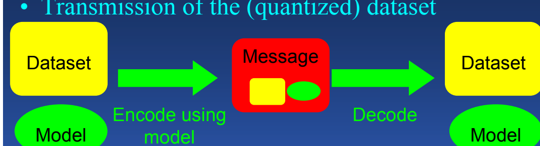


Minimum description length

- Energy of the model (PCA)
- Energy of the description (points)
- Minimise total
- The two energies have to be in comparable units
- Add a 'lamda'
- Minimise that also

What is MDL?

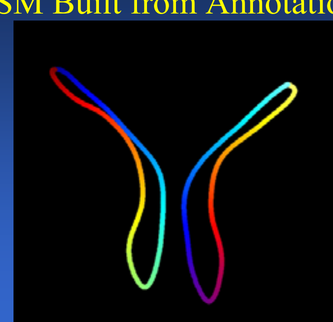
- Transmission of the (quantized) dataset



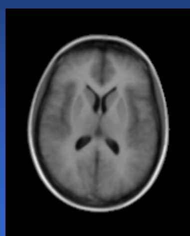
- Must have exact reconstruction of dataset.
- Optimal model ≡ shortest total message length.

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{model parameters}} + \mathcal{L}_{\text{data encoded using model}} + \mathcal{L}_{\text{residual/unmodelled bits}}$$

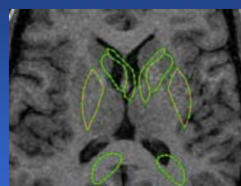
SSM Built from Annotation



Example from Cootes and Taylor

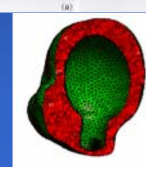


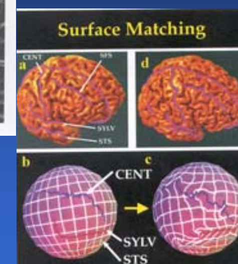
1st Mode of active appearance model



ASM model fitting

3D deformations

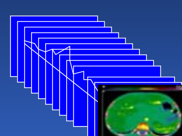


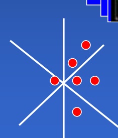


Ack. Thompson

Model based approach

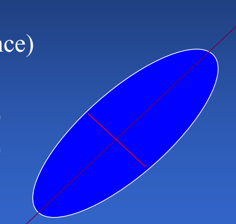
- $Y = X + \epsilon$
  - Y is data matrix (m,n)
  - X is model
  - ϵ is error
- Decompose
  - $X = F\ G'$
    - F (m,k)
    - G (k,n)





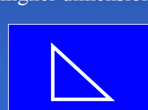
Principal Component Analysis

- F are Principal Axes
- G are Weights (variance)
- C (covariance matrix)
  - $C = (Y - y_m)'\ (Y - y_m)$



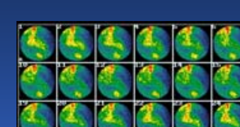
Oblique rotation

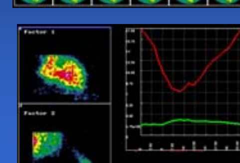
- PCA solution is orthogonal
- Make linear combinations
  - Oblique rotation
  - To satisfy constraint (positivity)
  - For example is higher dimensional space



Factor analysis: an example

- Decomposition into principal component
- Oblique rotation (based on constraints)
- Display of images (eigenfuctions) and curves (eigenvalues)
- Segmentation, model fitting and quantitation



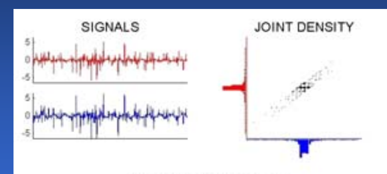


Independent Component Analysis

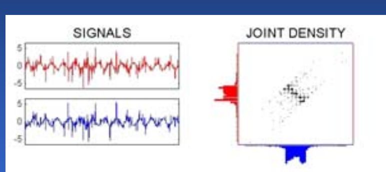
- Cocktail party problem
- Assume signals strictly independent
- Prewhitening

- Components not ordered.

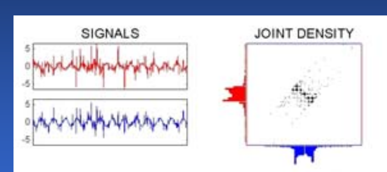
Example



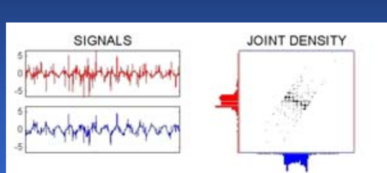
Input signals and density



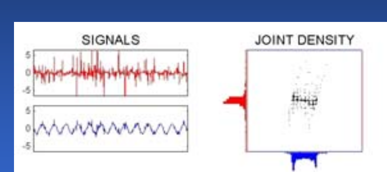
Whitened signals and density



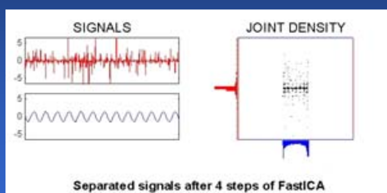
Separated signals after 1 step of FastICA

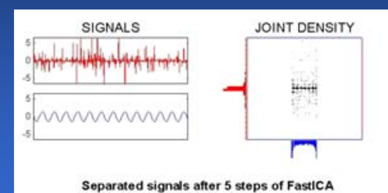


Separated signals after 2 steps of FastICA

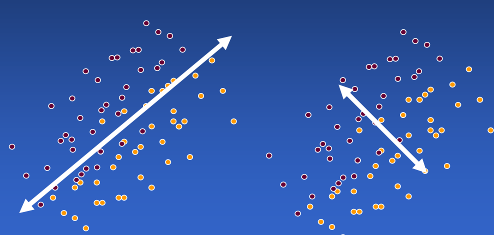


Separated signals after 3 steps of FastICA





PCA v. ICA

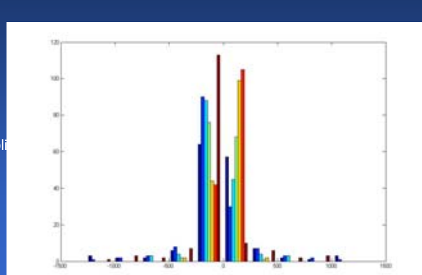


How ICA works

- Originally from signal processing
- Three algorithms:
  1. Fast ICA
  2. InfoMax
  3. JADE
- To obtain 'vectors'

Ordering ICA

Ampli



Colour represents component

Bayesian Analysis

Bayes' theorem

$$\Pr(P \mid E) = \frac{\Pr(\mathrm{E} \mid \mathrm{p})\Pr(\mathrm{P})}{\Pr(E)}$$

- Using the prior knowledge of shape to bias the boundary finding

$$\Pr(P_{map} \mid E) = \max \frac{\Pr(E \mid P)\Pr(P)}{\Pr(E)}$$

E is image object, P variables in template P_{map} is desired result

After serial inference

$$\Pr(P) = \prod_{i=1}^{N} \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{(p_i - m_i)^2}{2\sigma_i^2}}$$

$$E = P + noise$$

$$\Pr(E \mid p) = \prod_{A} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{(E(x,y) - P(x,y))^2}{2\sigma_n^2}}$$

$$M(P) = \sum_{i=a}^{t} [-\frac{(P_i - m_i)^2}{2\sigma_i^2}] + \frac{1}{\sigma_n^2} \sum_{1}^{N} E(x(p,n), y(p,n))$$

m is mean, σ is SD, for N points (x,y) over A for whole image

Neighbour-Constrained Segmentation

(Yang, Staib, Duncan, IPMI03)

Observation:

- Neighbouring structures often have a consistent image location and shape ;
- Relative positions or shapes among neighbors can be modeled based on statistical information from a training set.

→

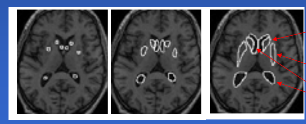
•Maximum A Posterior (MAP) framework:

Assume image I has M objects of interest: $S_1, S_2, \ldots S_M$

$$\hat{S}_i = \underset{S_i}{\arg\max}\, p(S_1, S_2, \ldots, S_M / I)$$

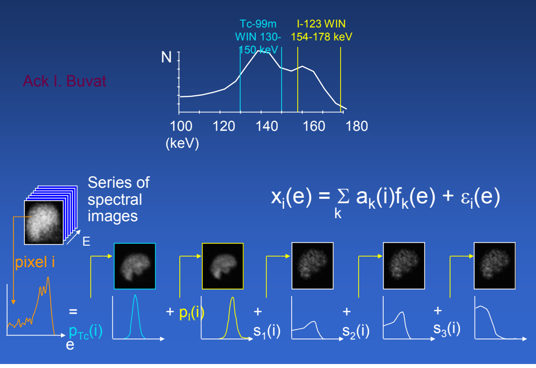
$$= \underset{S_i}{\arg\max}\, p(I / S_1, S_2, \ldots, S_M)\, p(S_1, S_2, \ldots, S_M) \quad i = 1, 2, \ldots, M$$

image gray level info — neighbour (shape + distance) prior info



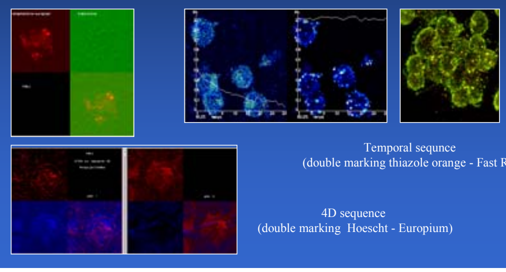
Detection of 8 sub-cortical structures using neighbor priors

Spectral Analysis



Confocal microscopy

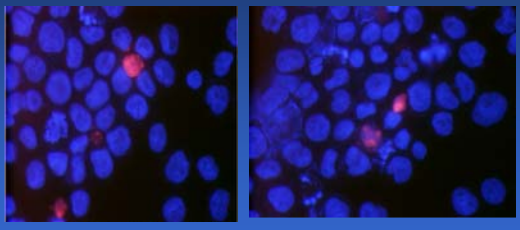
- Factor analysis of confocal image sequences of human papillomavirus DNA revealed with Fast Red in cervical tissue sections stained with TOTO-iodide. ANALYT QUANT CYTOL HISTOL, 2000, 22(2): 168-174, O.; 01-05, KAHN E et al : Confocal image characterization of human papillomavirus DNA sequences revealed with Europium in HeLa cell nuclei stained with Hoechst 33342. ANALYT QUANT CYTOL HISTOL, 2001, 23(2):101-108 O.



Temporal sequnce
(double marking thiazole orange - Fast Red

4D sequence
(double marking Hoescht - Europium)

Labelled leucemic cells



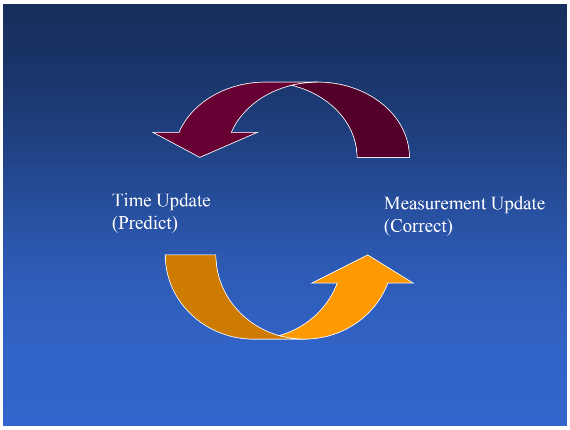
Kalman Filtering

- Problem
  - To estimate the state of $x \in \Re^n$
  - where
$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$$
  - With a measurement $z \in \Re^n$ that is
$$z_k = Hx_k + v_k$$

Random variables w_k and v_k are process and measurement noise

Q is noise covariance and R is measurement noise covariance

A relates previous step to current step, B is optional, H relats to changes in measurements



Update equations

- Filter time update

$$\hat{x}_k^- = A\hat{x}_{k-1}^- + Bu_{k-1}$$ Project the state ahead

$$P_k^- = AP_{k-1}A^T + Q$$ Project the error covariance

- Filter measurement update

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$ Compute the Kalman gaina

$$\hat{x}_k^- = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$ Update estimates with measurements

$$P_k = (I - K_k H)P_k^-$$ Update error covariance

^ indicates a postiori estimate ⁻ indicates a priori estimate

General Conclusions

- Ensure it is a good problem
- Acquire high quality data (as far as possible)
- Validate
- Evaluate
- Adapt