Medical image analysis tools

Image generation, processing, segmentation and registration

Gábor Székely
Computer Vision LAboratory
ETH Zürich

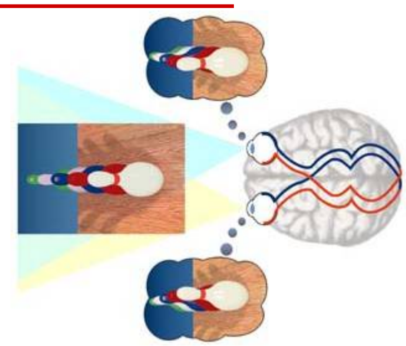
Content

- Introduction, human and machine vision
- Medical image formation
- Basics of signal processing and discrete spaces
- Image reconstruction
- Image processing tools
- Image segmentation
- Image registration

Computer vision

- Superiority of human vision
  - In many respects can hardly be outperformed in near future
  - Scene interpretation and understanding
  - Intuition
  - 3D stereo reconstruction
- Machine vision strength
  - Quantification
  - 3D reconstruction from image slices
  - Dealing with high dimensional and/or vectorial fields
  - Reproducibility
  - stamina

The human visual system

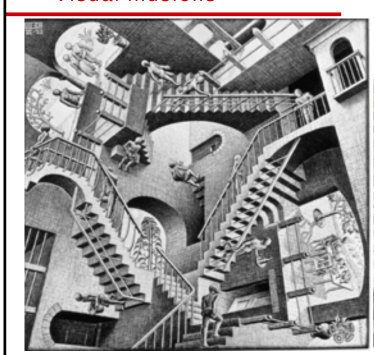


Machine vision systems

Luminance → Structure

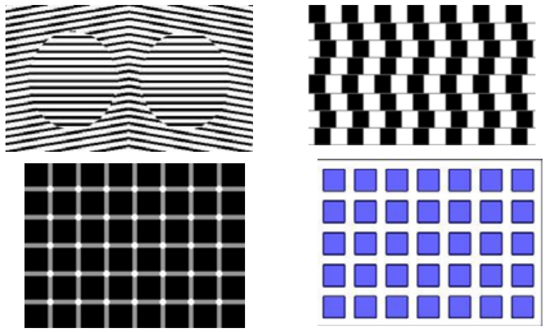
- Goal: identification and delineation of object
  a complex, interconnected process
- Solution approaches: generic properties of objects
  - Homogeneity inside of the object
  - Distinction from the background
- Fundamental limitation
  - Not all information is contained in the image
  - The related problems are mathematically ill posed
  - Prior knowledge is used for solution
  - No dividing line between knowledge and prejudice

Visual illusions

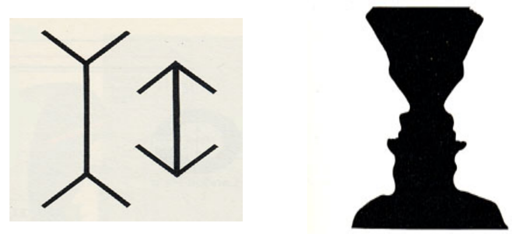


M.C. Escher (1898-1972)

The Paggendorf illusions

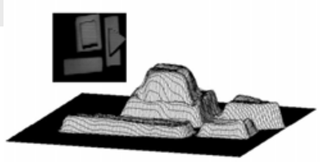


Mueller-Lyer illusion

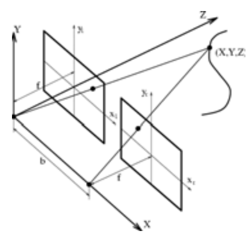


Examples of image formation

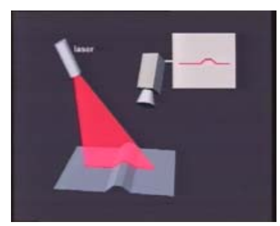
- Human vision: 3D color scene → stereo 2D color image pair
- Remote sensing: 2D/3D scene, multiple physical properties → multi-channel 2D images
- Photographic camera: 3D color scene → 2D color/bw image
- Range images: 3D scene → 2D image with distance to observer



Extraction of 3D information



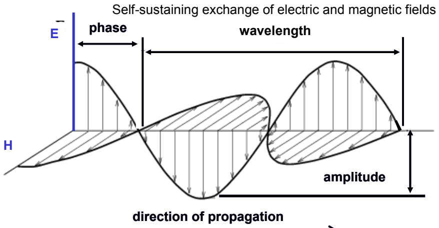
Passiv: human eye



Activ: pattern projection

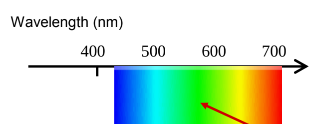
Light as electromagnetic wave

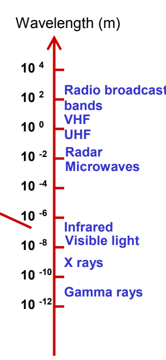
Self-sustaining exchange of electric and magnetic fields



1. wavelength
2. direction
3. amplitude *E*
4. phase
5. direction of polarisation

The electromagnetic spectrum





Interaction of light and matter

phenomenon	example
absorption	blue water
scattering	blue sky, red sunset
reflection	coloured ink
refraction	dispersion by a prism

Dependent on wavelength

+ diffraction

Transmission of EM waves

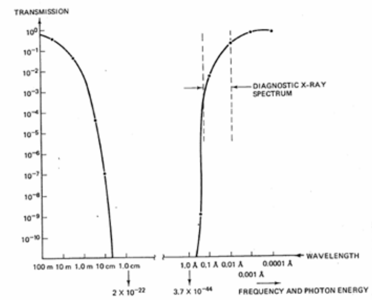
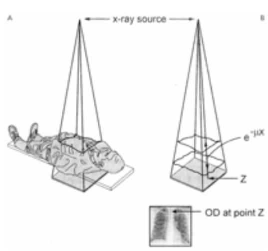


FIG. 1.1 Transmission of EM waves through 25 cm of soft tissue.

Seeing through the human body I



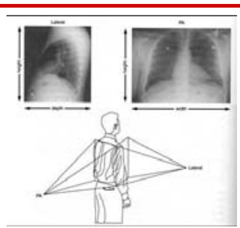
- Using high frequency radiation: X-ray
- Design considerations
  - Sufficiently low wavelength for transparency
  - Sufficiently low frequency to remain selective
    contrast formation
  - Ionizing radiation
- Projective image formation
  loss of depth information

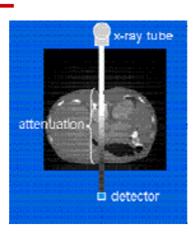
Seeing through the human body II

- Low frequency radiation: ultrasound
- Design considerations
  - Possibly low frequency for maximal penetration
  - Possibly high frequency to avoid diffraction effects
    radiation wavelength at ~1.5MHz in tissue is ~1mm
    wavelength should be significantly higher than the size of objects to be resolved
- Non-ionizing radiation

Frequency	2–15 MHz
Wavelength (in muscle)	0.78–0.1 mm
Penetration	12–1,6 cm
Spatial resolution	3.0–0.4 mm

X-ray imaging



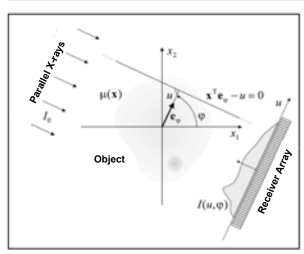


$$I = I_0 e^{-\int_{xi}^{xo} \mu(x)dx}$$

- Contrast based on tissue-dependent attenuation
- Problem: loss of depth

CT imaging

- Goal: depth recovery
- Relying on Radon transform
- Absorption imaging

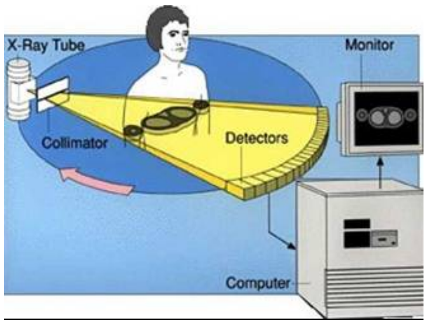


$$r(u,\varphi) = \mathcal{R}\{b(\mathbf{x})\} := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b(\mathbf{x})\,\delta(\mathbf{x}^T\mathbf{e}_\varphi - u)\,d\mathbf{x}$$

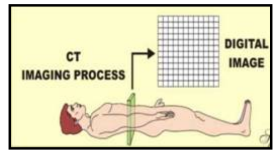
$$\varphi \in [0°, 180°), \quad u \in \mathbb{R}, \quad \mathbf{e}_\varphi = \begin{pmatrix}\cos\varphi \\ \sin\varphi\end{pmatrix}.$$

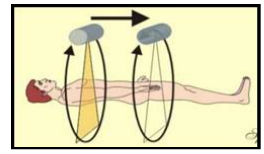
- Two equivalent representations
- b(**x**) can be reproduced from r(u,φ)

Carrying out the Radon Transformation



Multi-slice CT acquisition

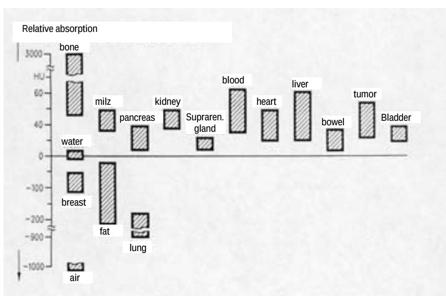




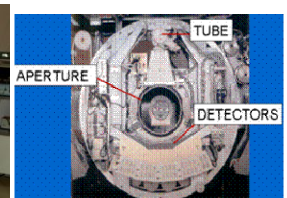
$$CT(x,y) = 1000\left[\frac{\mu(x,y) - \mu_{water}}{\mu_{water}}\right]$$

- We talk about image reconstruction later
- Complete 3D reconstruction possible
- Spiral acquisition
- Intensity: normalized absorption (Hounsfield)

X-ray absorption in tissue

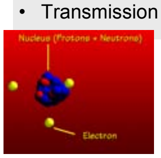


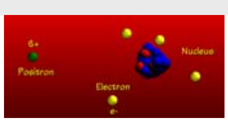
CT practical realization

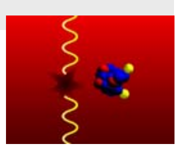


PET Imaging

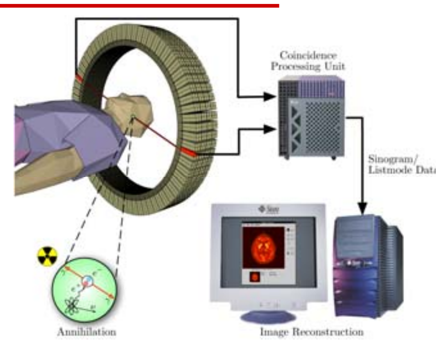
- Intracorporal radiation source
- Physiological imaging: radiolabeled pharmaceuticals
- Positron emitting isotopes
  - Short lifetime: isotope must be locally produced (cyclotrone)
  - detection of γ photons resulting from e^{+} + e^{-} annihilation
- Coincidence imaging to achieve reasonable SNR
- Same tomographic principle as by CT
- Transmission imaging



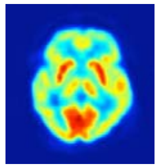


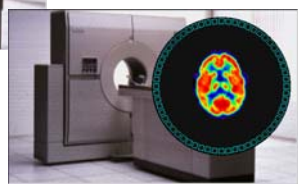


PET imaging device



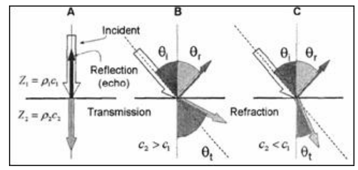
PET realization





Ultrasound imaging

- Acoustic Impedance, Z = r c
  r = density (kg/m3), c = speed of sound (m/sec)
- Large differences in Z cause reflection, small differences allow transmission
- Imaging principle: reflexion

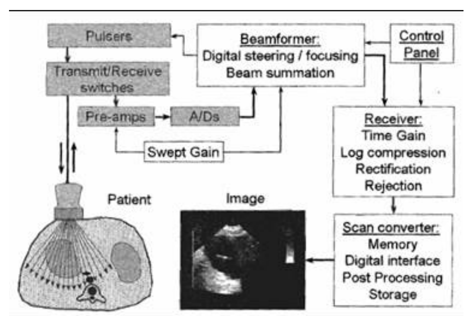


Sound behaviour on tissue boundaries

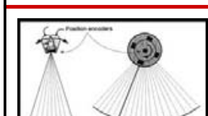
- The intensity reflection coefficient (normal incidence)
  $R = I_r/I_i = ((Z_2 - Z_1)/(Z_2 + Z_1))^2$
- The transmission coefficient $T = 1 - R$
  $T = (4Z_1Z_2)/(Z_1+Z_2)$
- Determined basically by density difference
  very big between air and tissue: air is a problem

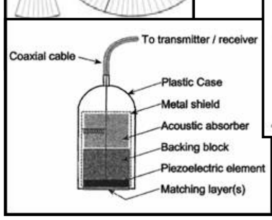
Tissue Interface	Pressure Reflection	Intensity Reflection
Liver–kidney	−0.006	0.00003
Liver–fat	−0.10	0.011
Fat–muscle	0.12	0.015
Muscle–bone	0.64	0.41
Muscle–lung	−0.81	0.65
Muscle–air	−0.99	0.99

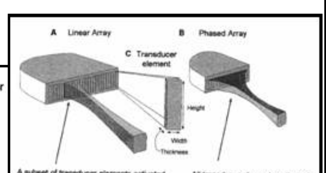
US technical realization



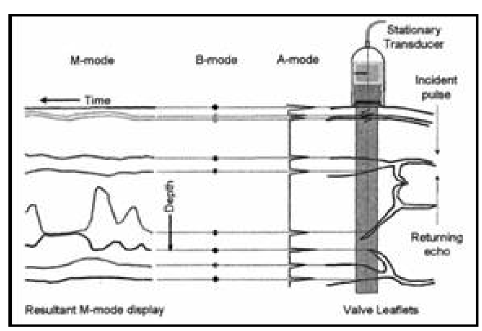
US Transducers



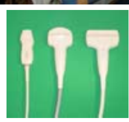


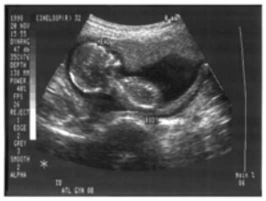


US signal visualization

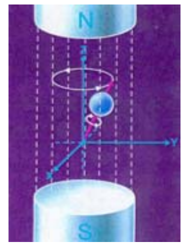


B-mode US imaging





Magnetic resonance imaging

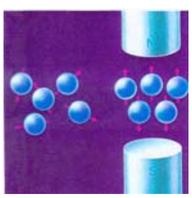


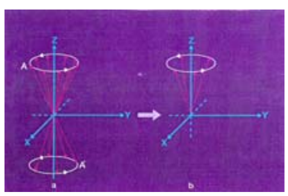
$\omega_{Larmor} = \gamma \cdot B_0$

- Absorption and emission of RF waves
- Emission imaging with external excitation
- Spins (protons) in magnetic field precession
- Two energy states
  - Difference in RF energy range
  - Can be excited by external radiation
  - Population difference is very small low sensitivity (SNR)
  - Relaxation after excitation terminated re-emitted RF energy can be detected
- Resonance frequency determined by local magnetic field

Effects of RF absorption

- Absorption of energy (magnetization in z direction)
- Synchronization of precession (magnetization in xy plane)

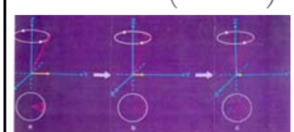


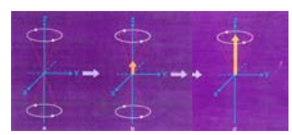


Relaxation effects

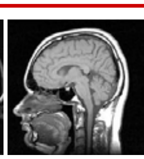
- Longitudinal relaxation: loss of energy (T1)
- Transversal relaxation: loss of coherence (T2)
- Strong dependence on molecular environment
- Highly versatile contrast formation
- No known health risks

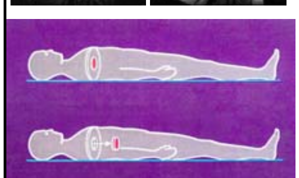
$$M_z(t) = M_0 \cdot \left(1 - c e^{-\frac{t}{T_1}}\right) \qquad M_T(t) = M_T(0) \cdot e^{-\frac{t}{T_2}}$$





Advanced MR imaging techniques





- Signal influenced by many parameters
- Very versatile tissue contrast
  - Perfusion imaging
  - MR angiography
  - fMRI
  - Diffusion imaging
  - Tissue tracking (MR tagging)
  - Diffusion imaging
  - MR elastography

Comparison of achievable resolution

Modality	Δ (mm)	Comments
Screen film radiography	0.08	Limited by focal spot and detector resolution
Digital radiography	0.17	Limited by size of detector elements
Fluoroscopy	0.125	Limited by detector and focal spot
Screen film mammography	0.03	Highest resolution modality in radiology
Digital mammography	0.05–0.10	Limited by size of detector elements
Computed tomography	0.4	About ½-mm pixels
Nuclear medicine planar imaging	7	Spatial resolution degrades substantially with distance from detector
Single photon emission computed tomography	7	Spatial resolution worst toward the center of cross-sectional image slice
Positron emission tomography	5	Better spatial resolution than with the other nuclear imaging modalities
Magnetic resonance imaging	1.0	Resolution can improve at higher magnetic fields
Ultrasound imaging (5 MHz)	0.3	Limited by wavelength of sound

Formal characterization of image analysis systems

f → $\boxed{\mathcal{F}}$ → g $\quad \mathcal{F}(f(t)) \rightarrow g(t)$

- Signal mapping
- Special case : linear shift invariant systems

$$\mathcal{F}(af(t)+bg(t)) = a\mathcal{F}(f(t)) + b\mathcal{F}(g(t))$$

$$\mathcal{F}(f(t)) = g(t) \rightarrow \mathcal{F}(f(t-t_0)) = g(t-t_0)$$

Characterization of LSI systems

- LSI system ≡ convolution
- Fully characterized by the point spread function (PSF) answer to Dirac δ (*h(t)*)
- Response can be calculated by convolution
- Harmonic waves are the eigenfunctions
- Intimate relationship with the Fourier-transform
- Convolution theorem

$$\mathcal{F}(f(t)) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau = f \otimes h(t)$$

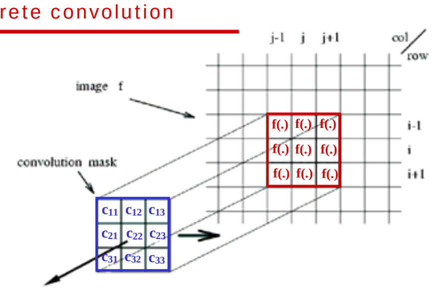
$$F(\nu) = \mathscr{F}(f(x)) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i\nu x}dx \qquad f(x) = \mathscr{F}^{-1}(F(\nu)) = \int_{-\infty}^{\infty} F(\nu)e^{2\pi i\nu x}d\nu$$

$$\mathscr{F}(f(x) \otimes g(x)) = F(\nu)G(\nu)$$

Discretization of signals

- Sampling
  - Fourier pair of discrete signal
  - Problems of artifacts
    - Aliasing
    - Leakage
- Quantization
  - Finite number of intensity levels
  - Contouring artifacts if not sufficiently fine dithering techniques

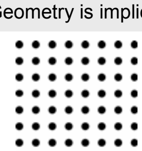
Discrete convolution



o (i,j) = c_{11} f(i-1,j-1) + c_{12} f(i-1,j) + c_{13} f(i-1,j+1) +
c_{21} f(i,j-1) + c_{22} f(i,j) + c_{23} f(i,j+1) +
c_{31} f(i+1,j-1) + c_{32} f(i+1,j) + c_{33} f(i+1,j+1)

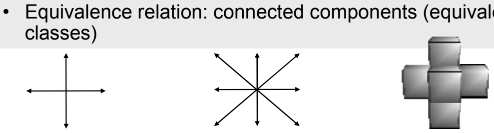
Structure of discrete image spaces

- Two issues: topology and distance closely coupled (otherwise problems)
- Three possible regular rasters in 2D
  - Cartesian
  - Hexagonal
  - Triangular
- Geometry is implicitely involved



Discrete topology

- Defined by neighborhood relation several topologies can be selected for the same space
- 4/8 neighborhoods in 2D, 6/18/26 neighborhoods in 3D
- Induced topology: connectivity
- Equivalence relation: connected components (equivalence classes)

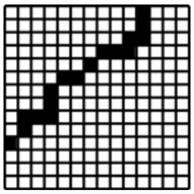


4-Neighbourhood $X \circledR Y$ **8-Neighbourhood** $X \circledR Y$ **6-Neighbourhood**

$$X,Y \in \Sigma, I(X) = I(Y)\, X,Y \in S\; X \sim_{\circledcirc} Y \leftrightarrow \exists\{X_1, X_2, \ldots X_n\} \subset S$$

$$X = X_1, \quad Y = X_n \text{ and } X_i \circledcirc X_{i+1} \quad i = 1,\ldots,n-1$$

Properties of discrete topological spaces



- Considering binary images
- Expected equivalence with continuous spaces

  problems with component connectivity (Jordan theorem)
- Solution: mixed topology (foreground 8, background 4)
- Typical operations
  - Thinning
  - Skeletonization

Distance on discrete spaces

- Distance mapping (metric) $d(x,y)$: $\mathbf{S} \times \mathbf{S} \rightarrow \mathbf{R}^+$
  - Symmetric
  - Transitive
  - Triangle inequality
- Every topology induces a distance (minimal connected path)
- Regular distances (all induced by a neighbourhood)

Topology induced distance

$$d(X,Y) = min\,\{l\{X_1,\dots X_n\} \mid X = X_1 {}^{©} X_2 {}^{©} \dots {}^{©} X_n = Y\}$$

Regular metric

$$\forall P,Q \in \Sigma, P \,{}_{©}\, Q \text{ or } \exists R \in \Sigma,\ R \neq P \text{ and } R \neq Q \text{ such that } d(P,Q) = D(P,R) + D(R,Q)$$

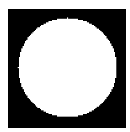
Properties of discrete distance metrics

- Regular distances show unnatural anisotropy
- D_4 circles are diamonds, D_8 circles are squares
- Usage of Euclidean distance necessary
  - Not regular on the discrete raster
  - Conflict with topology
  - Special algorithms needed for its estimation
- Alternative solution: approximation of the Euclidean distance by regular metrics
  - Chamfer metrics
  - Summing of the steps with different weights
  - Characterized by scale weights (<3,4>, <5,7,11>)
  - Neighborhood must be extended for larger environments

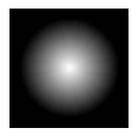
    move of knight on a 5x5 environment

Discrete distance transform

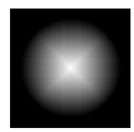
- Calculated usually by Euclidean distance approximation
- Usage
  - Robot path planning
  - Skeletonization
  - Morphological operations



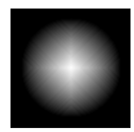
Binary image



DT D_{Euclid}



DT D_8



DT D_4

Mathematical morphology

- Structural processing of binary imaging
- Enforcing consistency within a neighbourhood (structuring element)
- Convolution-like operation but using non-linear (rank order) operators
- Structuring element is usually selected as diamond or square
- Iterative application of basic operators

$$\{I_1, I_2, \dots, I_n\} \xrightarrow{sorting} \{I_1^*, I_2^*, \dots, I_n^*\}$$

Morphological operators

- Erosion, $E(I_c) = I^*_1$

  shrinking, one layer is eroded from the object
- Dilation $D(I_c) = I^*_n$

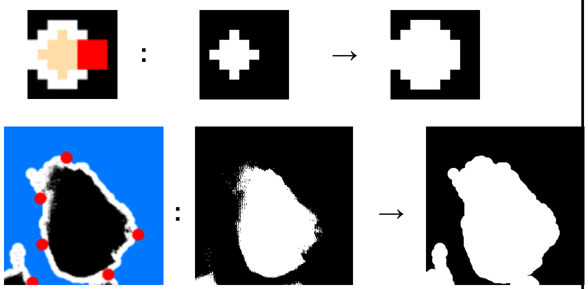
  thickening: one layer is added to the object
- Operator combination
  - Goal: morphological adjustments without size changes
  - Opening: erosion followed by dilation

    breaking narrow bridges

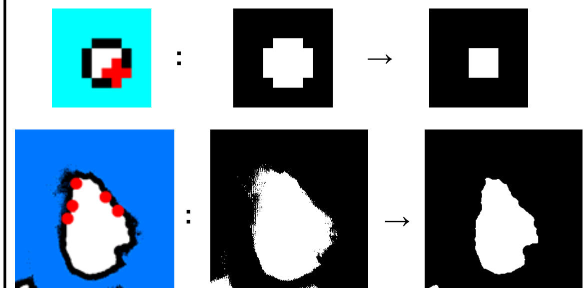
    deleting small components
  - Closing: dilation followed by erosion

    filling small holes

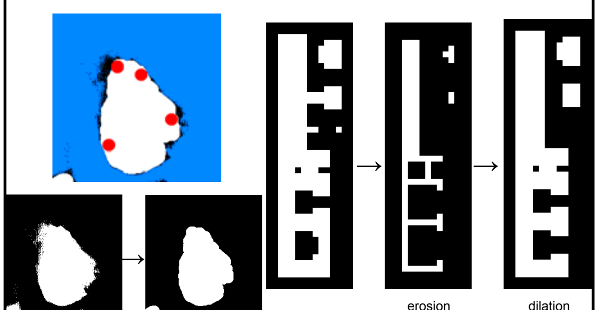
Examples dilation



Examples erosion



Examples opening



Examples closing

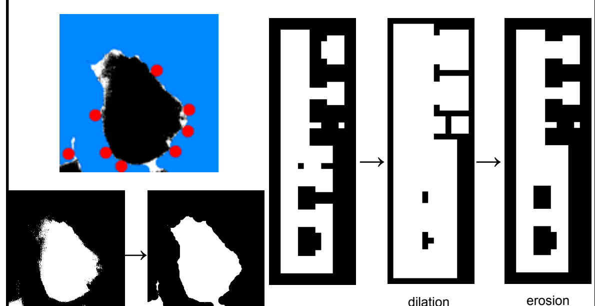
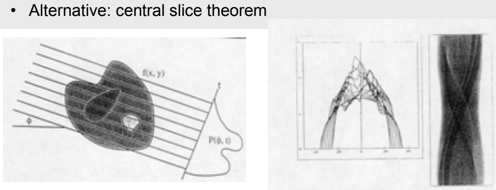
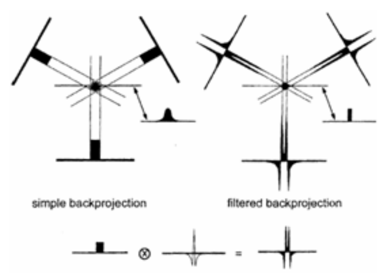


Image reconstruction from projections

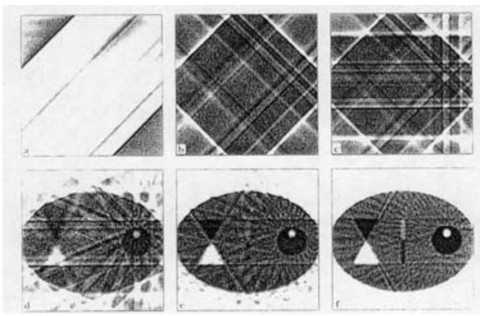
- Radon transform: representation as sinogram
- Inversion techniques
  - Straightforward: backprojection
  - Problem with the transfer function: 1/r filtering necessary
  - Alternative: central slice theorem



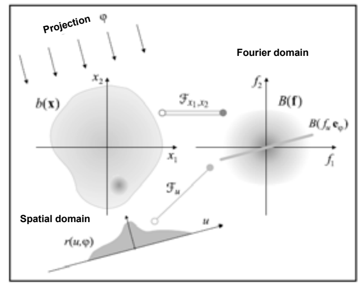
Filtered backprojection



Reconstruction by filtered backprojection



The central slice theorem



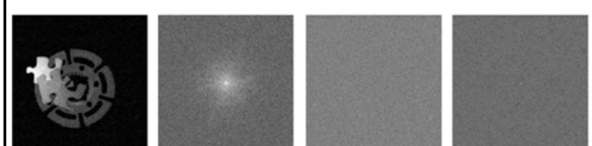
$$\mathcal{F}_u\{r(u,\varphi)\} = B(f_u\mathbf{e}_\varphi) \qquad B(\mathbf{f}) = \mathcal{F}_\mathbf{x}\{b(\mathbf{x})\}$$

Image processing needs

- The acquired images are imperfect distortions disturb later processing
- Major problems
  - Noise
  - Spatial distortion
  - Non-homogeneous intensity (bias)
- Possible approaches
  - Pre-processing: correction before further analysis prototypical example: noise reduction
  - Simultaneous processing and correction prototypical example: bias correction

Noise reduction

- Frequency characteristics of images and noise
- SNR is bad at high frequencies
- Lowpass filter is a straightforward solution
- Blurring edges, losing sharpness



spatial signal frequency spatial noise frequency

Linear noise reduction filters

Box filter
major problems

1	1	1
1	1	1
1	1	1

1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

Binomial filter
filter family (Pascal Δ)

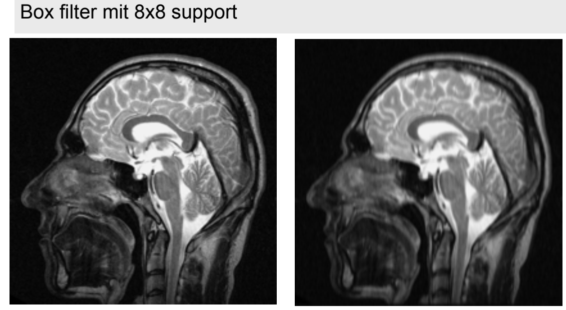
1	2	1
2	4	2
1	2	1

Gaussian filter
„optimal“
Limit of binomial filters
Continuous family (σ)

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}; \quad G(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

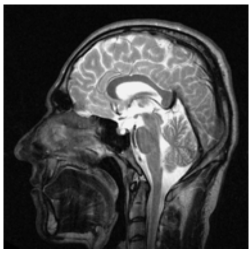
Box filtering results

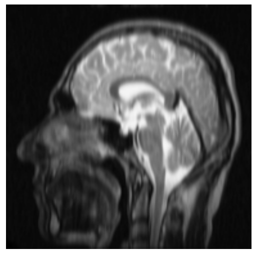
Box filter mit 8x8 support



Gauss filtering results

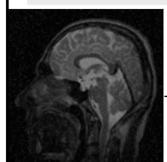
Gauss filter σ = 5 pixel



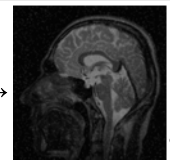


Non-linear noise reduction: the median filter

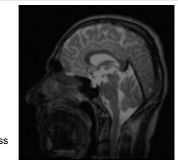
- Noise vs. Sharpness tradeoff
- Linear filters cannot offer a solution: frequencies are handled in a structure-independent fashion (LSI!)
- Traditional solution: median filter
- Yet another rank order operator, $M(I_c) = I^*_{\lceil n/2 \rceil}$ where $\lceil n/2 \rceil$ is the integer round-up operator



→



Gauss



Median

Anisotropic diffusion

- The Gauss function is the Green function of the isotropic diffusion equation
- Perona & Malik: make the diffusion coefficient dependent on image content
- Stop diffusion at the boundaries: strong gradients

$$\frac{\partial f(\vec{x},t)}{\partial t} = div(c(\vec{x},t)\vec{\nabla} f(\vec{x},t))$$

general diffusion

$$\frac{\partial f(\vec{x},t)}{\partial t} = c\Delta f(\vec{x},t)$$

isotropic diffusion (c constant)

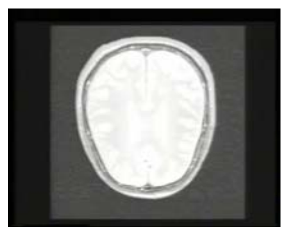
general solution of the Homogeneous diffusion equation Initial condition *f(0,x) = i(x)*

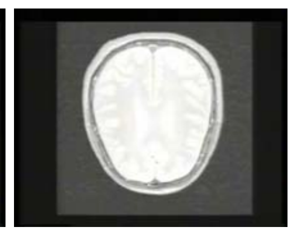
$$o(\vec{x}) = i(\vec{x})^{\otimes} \frac{1}{(2\pi)^{(n/2)}\sqrt{ct}} e^{-\frac{<\vec{x},\vec{x}>}{4ct}}$$

selection of c

$$c(|\vec{\nabla} f|) = e^{-\frac{|\vec{\nabla} f|^2}{2\kappa^2}} \qquad c(|\vec{\nabla} f|) = \frac{1}{1 + \left(\frac{|\vec{\nabla} f|}{\kappa}\right)^2}$$

Comparison between isotropic and anisotropic diffusion





Homogeneity-based image segmentation

$$\boxed{\text{Object}} \xrightarrow{\text{Sensors}} \boxed{\text{Measurement data}} \xrightarrow{\text{feature extraction}} \boxed{\text{features}} \xrightarrow{\text{classificator}} \boxed{\text{Object classes}}$$

- Characterization of the object pixels by selected features
- Looking for areas with similar features
- Prototype: thresholding
- Threshold determination: Bayesian estimates
- Both for scalar and vectorial features

The Bayesian framework

Object classes $\Omega = \{\omega_1, \omega_2, \ldots \omega_s, \omega_r\}$

Features $\vec{v} = (v_1, v_2, \ldots, v_n) \in \boldsymbol{R}^n$

Joint probability density $p(\vec{v}, \omega_i)$

Marginal probability densities $p(\vec{v}) = \sum^{s}_{i} p(\vec{v}, \omega_i)$ $\qquad P(\omega_i) = \int p(\vec{v}, \omega_i) d\vec{v}$

Prior probability

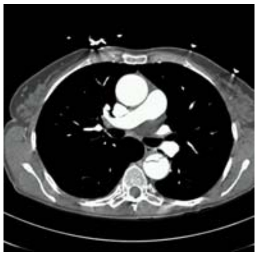
Class specific Probability density $p(\vec{v}|\omega_i) = \frac{p(\vec{v}, \omega_i)}{P(\omega_i)}$

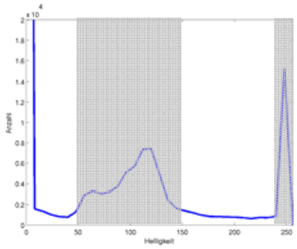
a posteriory probability $P(\omega_i|\vec{v}) = \frac{p(\vec{v}, \omega_i)}{p(\vec{v})}$

Class selection should maximize a posteriori probability
Estimation through the Bayes theorem

$$P(\omega_i|\vec{v}) = \frac{p(\vec{v}|\omega_i)P(\omega_i)}{p(\vec{v})}$$

Thresholding of a CT image

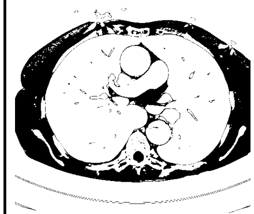


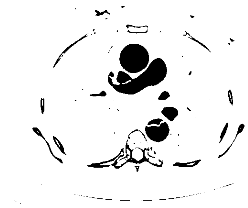


soft tissue

bone
contrast agent

Thresholding results





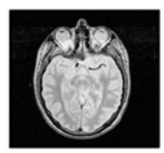
soft tissue

bone
contrast agent

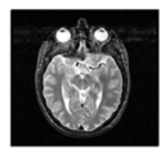
Segmentation of dual echo MRI data

- Brain segmentation from neuroradiological MRI data
- Simultaneous acquisition of early and late echoes
- Identification of main tissue classes
  - Gray matter (GM)
  - White matter (WM)
  - Cerebrospinal fluid (CSF)

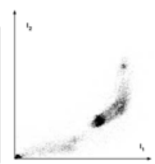
Segmentation process



early echo

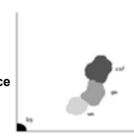


late echo

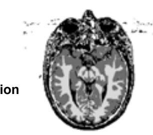


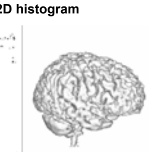
2D histogram

Tessellated
feature space



resulting
segmentation

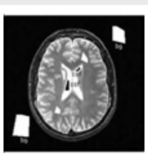




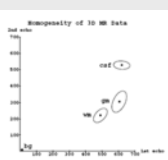
How to learn the priors

- Marginal density of features: histogram
- Prior probability: existing knowledge
- Class specific probabilities: supervised procedures
  - Estimated from examples
  - Parametric: fitting the parameters of the distribution
  - Non-parametric: direct feature space tessellation

Learning
examples



Fitted Gaussian
distributions



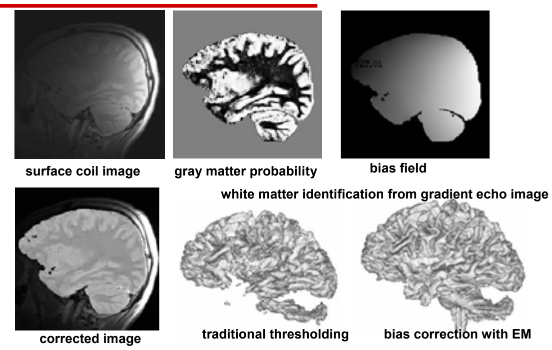
Problems of thresholding

- Works well as long as
  - Tissue properties are homogeneous
  - Tissue classes are reasonably separated
- Noise: previous noise reduction
  - Blurring enlarges class distributions
  - Increasing danger of class overlap
- Problem of inhomogeneity: bias field
- No spatial relations taken into consideration
  - Not well correlated with human image interpretation
  - Enforcement of spatial coherence

Bias field correction

- Simultaneous estimate of a posteriori probabilities and the bias field
- Iterative adjustment of the classification resp. the bias field
  - If the bias field is known, the thresholding can be done after corection
  - If the classification is known, the bias field can be estimated from class homogeneity conditions
  - The expectation maximization (EM) algorithm
  - Both steps rely on Bayesian estimate of a posteriori probabilities

EM results

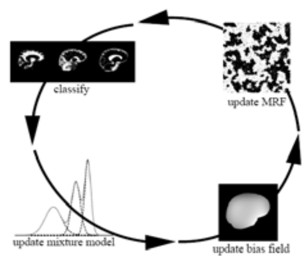


Spatial coherence by post-processing

- Similar behaviour of neighbouring pixels
- Morphological correction after segmentation
- Example: automatic brain extraction from gradient-echo MR images
- Problems
  - Noisy pixels: wrong components and holes
  - Wrong or undesired connectivity (brain to skull, eye nerves)

Thresholded component | erosion | connected component | dilation

Spatial coherence enforcement by Markov random field



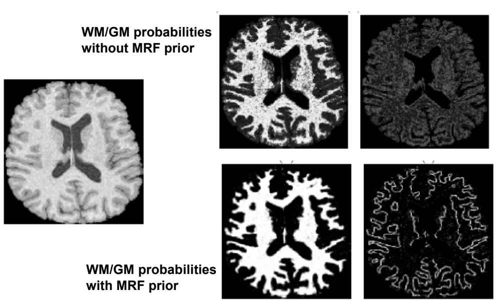
- Modeling label (L) dependency from neighbourhood (Φ) as a Gibbs distribution
- The energy function (U) depends on the direct neighbours: MRF based on Potts (Ising) model
- ξ and *v* describe the cost of transition from label l_j to $l_{j'}$
- Integration into the EM framework: estimating ξ and *v* with or without constraints

$$f(\boldsymbol{L} \mid \Phi) = Z(\Phi)^{-1} \exp[-U(\boldsymbol{L} \mid \Phi)]$$

$$Z(\Phi) = \textstyle\sum_{\boldsymbol{L}} \exp[-U(\boldsymbol{L} \mid \Phi)]$$

$$U(\boldsymbol{L} \mid \Phi) = \frac{1}{2} \sum_j \Big(\sum_{j' \in \mathcal{N}_j^p} \xi_{l_j l_{j'}} + \sum_{j' \in \mathcal{N}_j^o} \nu_{l_j l_{j'}} \Big)$$

Result of MRF/EM segmentation

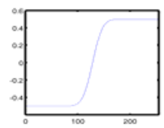


Boundary-based segmentation

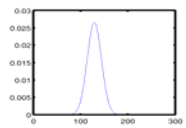
- Identification of object boundaries
- Searching for strong changes in image features
  - Intensity
  - Texture
  - ...
- Edge detection
  - Tools of mathematical analysis
    - edges practically always modelled as 1D structures
    - intensity function orthogonal to boundary
  - Signal processing alternative: high pass filters
  - Practically identical results

Mathematical analysis of the intensity function

- Ideal edge as function $f(t) \leftrightarrow t$



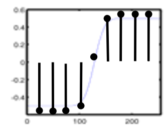
- Change in the function $\frac{df(t)}{dt} \leftrightarrow t$



- Searching for extremal positions

Mathematical analysis on sampled functions

- The image is sampled



- Definition of derivative

$$\frac{df(t)}{dt} = \lim_{\Delta t \to 0} \frac{f(t+\Delta t) - f(t)}{\Delta t}$$

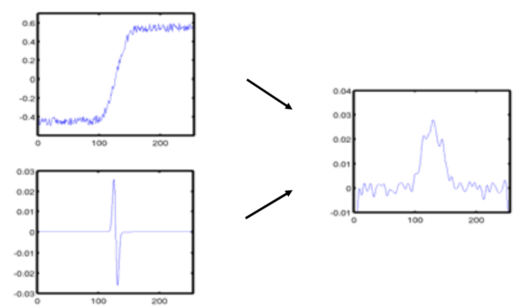
Discrete derivatives

- Finite difference approximation

$$\frac{df(t)}{dt} \approx \frac{f(t_{n+1}) - f(t_n)}{t_{n+1} - t_n}$$

- Problems with intensity differences
  - Noise amplification
  - Noise suppression necessary
  - Combination with low-pass filter: Gauss function
  - Resulting filter: first derivative of Gaussian

1D detection example



How to find edge direction

- Direction of steepest intensity change: gradient
  directional derivative with maximal magnitude
- Gradient direction: orthogonal to edge
- Gradient magnitude: edge strength
- Non-linear, but can be estimated through the partial derivatives *δf/δx* and *δf/δy*

$$\nabla f(x,y) = \left(\frac{\delta f}{\delta x}, \frac{\delta f}{\delta y}\right)$$

$$|\nabla f(x,y)| = \sqrt{\left(\frac{\delta f}{\delta x}\right)^2 + \left(\frac{\delta f}{\delta y}\right)^2}$$

$$\theta = \arctan\left(\frac{\delta f}{\delta y} / \frac{\delta f}{\delta x}\right)$$

Estimate of partial derivatives

- High-pass filter in derivative direction
- Smoothing (low-pass filter) orthogonally
  improving SNR along the edge
- Most common implementations
  - Sobel masks
  - Partial derivatives of the Gauss function (*δG(x,y)/dx, δG(x,y)/δy*) sampled and cut off to a finite mask

Sobel masks

-1	0	1
-2	0	2
-1	0	1

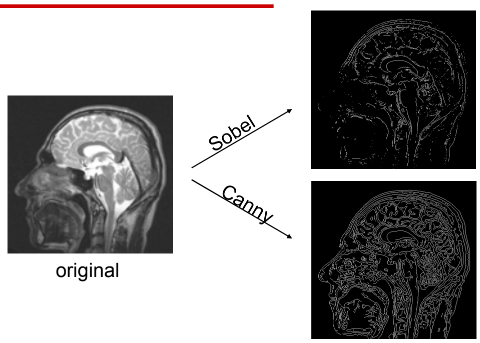
-1	-2	-1
0	0	0
1	2	1

The Canny edge detector

- Does not rely on gradient
  not necessarily compatible with the image raster
- Evaluation of directional derivatives
  - Noise reduction by Gauss-filtering
  - Discrete difference in all 4 directions defined by the 8-neighbourhood
  - Looking for direction channel with maximal magnitude
    value *r(x,y)* is an estimate for gradient magnitude
  - Non maximum suppression
    deleting all responses which are smaller than any neighbour in the selected direction channel
  - Hysteresis thresholding with $t_{high} > t_{low}$
    - If $r(x,y) > t_{high}$ edge point is safe and kept
    - If $r(x,y) < t_{low}$ edge point is due to noise and neglected
    - For $t_{low} < r(x,y) < t_{high}$ edge point is only kept if connected to safe edge pixels

Example of image gradient

Result of edge detection



Edge fragment grouping

- Boundary edges are hardly ever complete
- Gaps must be closed for object delineation
- Edge fragment grouping
  - One of the mysteries of human vision
  - Not much known about the underlying mechanisms
  - Prior knowledge plays a dominant role (Gestalt laws, e.g.)
- Grouping in computer vision
  - Relies also on image-independent assumptions
  - Two characteristic procedures
    - Hough transform
    - Elastically deformable models

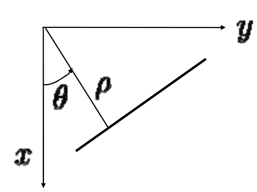
The Hough transform

- Detection of geometric objects with simple parameterization
- Original version identifies straight lines
- Processes binary edge images
- Building up an accumulator for the line parameters
- Every edge point votes for all possible lines which pass through it
- Looking for local maxima in the accumulator
  - Identifies potential lines
  - Intensity = number of votes (i.e. number of points on the line)
- Can be implemented for other structures (circles, ...)
- Advantage: closes arbitrary gaps (can also harm)
- Disadvantage: slow

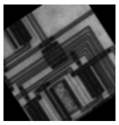
Line parameterization

- Most straightforward: $y = a\,x + b$
  problems with vertical lines (*a* becomes infinite)
- Alternative parameterization

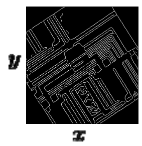
$$\rho = x\cos(\theta) + y\cos(\theta)$$

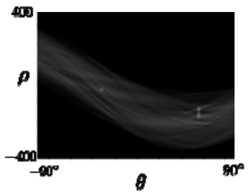


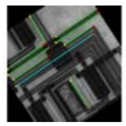
Line detection by Hough transform



$$\rho = x\cos(\theta) + y\cos(\theta)$$







Elastically deformable contour models

- An interactive approach to grouping
- The user provides a rough approximation of the contour
- The model deforms itself in order to adjust to nearby edge segments
- Deformation is based on physical analogy
- Behaviour of a rubber band (snake)
- Basic implementation: two energy terms
  - Potential (image, external) energy (attraction by image features, e.g. the gradient magnitude as a function graph)
  - Elastic (deformation, internal) energy (trying to avoid too much tension and keep the curve smooth)
- Searching for position minimizing the overall energy

Simple snakes formulation

Evolving curve $\vec{v}(s,t) = (x(s,t), y(s,t)) \quad 0 \le s \le 1, \ t \ge 0$

Initialization $\vec{v}(s,0) = (x(s,0), y(s,0))$

Internal energy $E_I(\vec{v}) = -\int_0^1 P(\vec{v}(s,t))ds$

Image potential $P(x,y) = ||\vec{\nabla} I(x,y)|| \approx \sqrt{\Delta_x^2(I \circledcirc G_\sigma) + \Delta_y^2(I \circledcirc G_\sigma)}$

Deformation energy thin plate spline model punishing stretching and bending
$$E_D(\vec{v}) = \frac{1}{2}\int_0^1 \alpha(s)\left|\left|\frac{\partial \vec{v}(s,t)}{\partial s}\right|\right|^2 + \beta(s)\left|\left|\frac{\partial^2 \vec{v}(s,t)}{\partial s^2}\right|\right|^2 ds$$

Total energy to be minimized

$$E(\vec{v}(s,t)) = E_D(\vec{v}(s,t)) + E_I(\vec{v}(s,t)) = \frac{1}{2}\int_0^1 \alpha(s)\left|\left|\frac{\partial \vec{v}(s,t)}{\partial s}\right|\right|^2 + \beta(s)\left|\left|\frac{\partial^2 \vec{v}(s,t)}{\partial s^2}\right|\right|^2 ds - \int_0^1 P(\vec{v}(s,t))ds$$

Energy minimization

Direct minimization techniques

$$E = \frac{1}{2}\int_0^1 \alpha(s)\left|\left|\frac{\partial \vec{v}}{\partial s}\right|\right|^2 + \beta(s)\left|\left|\frac{\partial^2 \vec{v}}{\partial s^2}\right|\right|^2 ds - \int_0^1 P(\vec{v})ds \rightarrow min$$

Variational calculus: Euler-Lagrange differential equation

$$-\frac{\partial}{\partial s}\left(\alpha(s)\frac{\partial \vec{v}}{\partial s}\right) + \frac{\partial^2}{\partial s^2}\left(\beta(s)\frac{\partial^2 \vec{v}}{\partial s^2}\right) = -\vec{\nabla} P(\vec{v})$$

For constant α and β

$$\begin{aligned} -\alpha\frac{\partial^2 x}{\partial s^2} + \beta\frac{\partial^4 x}{\partial s^4} &= -\frac{\partial P}{\partial x} \\ -\alpha\frac{\partial^2 y}{\partial s^2} + \beta\frac{\partial^4 y}{\partial s^4} &= -\frac{\partial P}{\partial y} \end{aligned}$$

After finite difference discretization

$$\vec{v}^{[t]} = \mathbf{K}^{-1} P_{\vec{v}}(\vec{v}^{[t-1]})$$

K is the stiffness matrix (pentadiagonal
Can only be inverted if sufficient boundary conditions given

Extension to full dynamic system

Additional energy term: kinetic energy
$$E_K(\vec{v}) = \frac{1}{2}\int_0^1 \mu(s)\left(\frac{\partial \vec{v}(s,t)}{\partial t}\right)^2 ds$$

Total energy
$$E(\vec{v}) = E_K(\vec{v}) + E_I(\vec{v}) + E_D(\vec{v})$$

Energy dissipation needed for Realistic behaviour Rayleigh functional
$$D(\vec{v}_t) = \frac{1}{2}\int_0^1 \gamma(s)\,|\vec{v}_t|^2\, ds$$

Non-conservative system Minimization by Hamilton principle
$$\int_0^1 E(\vec{v}) + D(\vec{v}_t)ds$$

Euler-Lagrange differential Equation for constant α, β, γ, μ
$$\begin{aligned} \mu\frac{\partial^2 x}{\partial t^2} + \gamma\frac{\partial x}{\partial t} - \alpha\frac{\partial^2 x}{\partial s^2} + \beta\frac{\partial^4 x}{\partial s^4} &= -\frac{\partial P}{\partial x} \\ \mu\frac{\partial^2 y}{\partial t^2} + \gamma\frac{\partial y}{\partial t} - \alpha\frac{\partial^2 y}{\partial s^2} + \beta\frac{\partial^4 y}{\partial s^4} &= -\frac{\partial P}{\partial y} \end{aligned}$$

Iterative solution after finite difference discretization and defining the boundary conditions

$$\vec{v}^{[t]} = [(\mu+\gamma)\mathbf{I} + \mathbf{K}]^{-1}\left(-P_{\vec{v}}\Big|_{\vec{v}^{[t-1]}} + (2\mu+\gamma)\vec{v}^{[t-1]} - \mu\vec{v}^{[t-2]}\right)$$

Snakes example

- Additional interactive forces for guiding the evolution (spring, volcano)
- Extension to 3D (deformable surfaces)

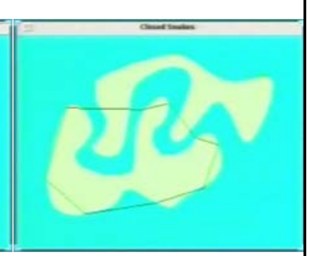


Image registration

- Establishment of a dense mapping between two images
- Classification
  - Degrees of freedom of the applied transformation
    - Rigid
    - Similarity/affine/projective
    - Elastic or non-rigid
  - Targets used (extrinsic – marker based/intrinsic)
  - Correspondence establishment (feature, intensity based)
  - Dimensionality (2D/2D, 3D/3D, 2D/3D)
  - Modalities involved (mono/multi-modal)
  - Origin of the images (data from the same or different patients)
- The crucial point is correspondence establishment
  is there any and if yes, how to find it?

Correspondence establishment

- Feature based
  - Points, curves or surfaces
  - Quality of fit measured by distance, minimizing the sum of squared distances
  - Correspondence between points and 1D/2D features: projection
    Iterative Closest Point (ICP) procedures
  - Straightforward and fast but needs reasonable feature detection and inherits all its errors
- Intensity based
  - Compares image intensities when applying transformations
  - Directly uses image information, no preprocessing needed
  - Similarity measure necessary

Similarity measures for intensity-based registration

Measuring similarity between the reference image (*A*) and the floating image after applying transformation *T* (B^T) over the common domain Ω^T_{AB} containing *N* pixels

Sum of squared differences

$$SSD = \frac{1}{N} \sum_{\vec{s}_A \in \Omega^T_{AB}} \left(A(\vec{s}_A) - B^T(\vec{s}_A)\right)^2$$

Sum of absolute differences
Both rely on identical contrast in *A*, *B*

$$SAD = \frac{1}{N} \sum_{\vec{s}_A \in \Omega^T_{AB}} \left|A(\vec{s}_A) - B^T(\vec{s}_A)\right|$$

Cross correlation: allows linear relationship between image intensities

$$CC = \frac{\sum_{\vec{s}_A \in \Omega^T_{AB}} \left(A(\vec{s}_A) - \overline{A}\right) \cdot \left(B^T(\vec{s}_A) - \overline{B}\right)}{\left[\sum_{\vec{s}_A \in \Omega^T_{AB}} \left(A(\vec{s}_A) - \overline{A}\right)^2 \cdot \sum_{\vec{s}_A \in \Omega^T_{AB}} \left(B^T(\vec{s}_A) - \overline{B}\right)^2\right]^{1/2}}$$

Mutual information as similarity measure

- Based on information theory
- „Find as much as possible of the complexity that is in the separate datasets so that at the same time they explain each other well“

Shannon entropy

$$H = -\sum p(s) \log p(s)$$

Mutual entropy of the images

$$H(A, B^T) = -\sum_{a \in \Omega^T_a} \sum_{b \in \Omega^T_b} p^T_{AB}(a,b) \log p^T_{AB}(a,b)$$

Mutual information

$$I(A, B^T) = H(A) + H(B^T) - H(A, B^T)$$

Marginal entropies

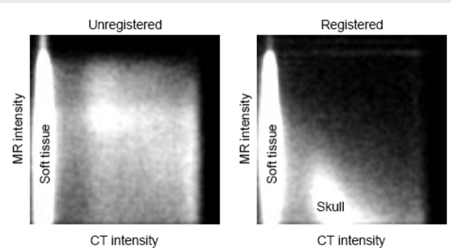
$$H(A) = -\sum_{a \in \Omega^T_a} p^T_A(a) \log p^T_A(a)$$
$$H(B^T) = -\sum_{b \in \Omega^T_b} p^T_B(b) \log p^T_B(b)$$

Mutual information from intensity probability distributions

$$I(A, B^T) = \sum_{a \in \Omega^T_a} \sum_{b \in \Omega^T_b} p^T_{AB}(a,b) \log \frac{p^T_{AB}(a,b)}{p^T_A(a) p^T_B(b)}$$

Estimation of MI for digital images

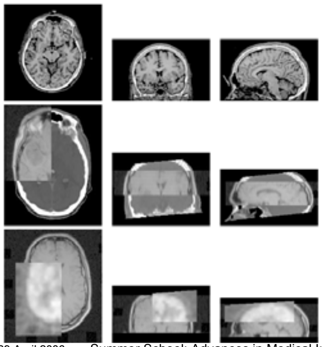
- Relying on 2D histograms
- Its dispersion is increasing with misregistration



Components of the registration process

- Representation of the transformation
  - Critical for non-rigid registration
  - Physical analogies (elastic deformation, viscous fluid)
  - Hierarchical parametric control (spline, subdivision)
- Interpolation methods
  - Calculating image intensities
  - Generating dense deformation field
- Estimation of MI
  Several deficiencies
  - Interpolation artefacts
  - Statistical inconsistency for few samples
- Search for solution
  - Solving the partial differential equation
  - Parametric optimization (sensitivity to local extrema)

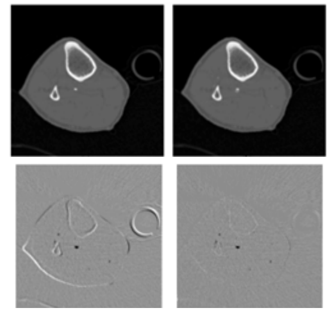
Rigid CT/MR registration results



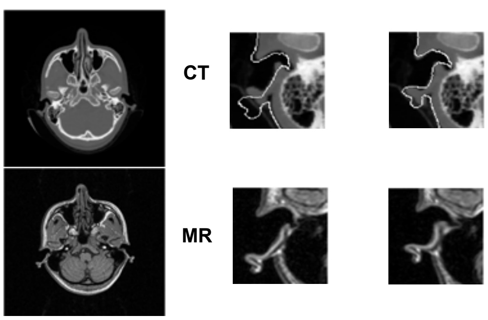
Nonrigid CT/CT registration results

CT/CT angiography image pair vessel detection obstructed by bone misregistration

Difference image after rigid and non-rigid registration



Nonrigid CT/MR registration



Clinical applications of registration

- Intra-patient matching
  - Diagnostic follow-up
  - Time series analysis
  - Information fusion
  - Pre-operative planning
  - Intra-operative navigation
    - Fusion of pre- and intra-operative images
    - Registration of the patient with pre-operative images
- Inter-patient registration
  - Atlas matching: transfer of related information
  - Evaluation of patient collectives (fMRI)

Summary

- A large number of image processing tools are available today
- Several standardization efforts
- Image registration is a broadly used clinical tool
  - Non-rigid matching is in its infancy
  - Feasible inter-patient matching has not even born
- Image segmentation is still unsolved
  - Key is the incorporation of prior knowledge
  - Several promising approaches exist (e.g. statistical shape and appearance modeling)
  - No generic solution can be expected in the near future
  - Hardly any solution in sight for scenes altered by complex pathologies