Mathematics in Image Processing

Michal Šorel
Department of Image Processing
Institute of Information Theory and Automation (ÚTIA)
Academy of Sciences of the Czech Republic
http://zoi.utia.cas.cz/

Mathematics in image processing

Mathematics in image processing , CV etc.	My subjective importance
Linear algebra	70%
Numerical mathematics – mainly optimization	60%
Analysis (including convex analysis and variational calculus)	50%
Statistics and probability – basics + machine learning	30%
Graph theory (mainly graph algorithms)	15%
Universal algebra (algebraic geometry, Gröbner bases…)	not much

Probably similar for many engineering fields…

Talk outline

- What is digital image processing? Typical problems and their mathematical formulation.
- Bayesian view of inverse problems in (not only) image restoration, analysis and synthesis based sparsity
- Discrete labeling problems and Markov random fields (MRFs, CRFs)

Image processing and related fields

- Image processing
  - Image restoration (denoising, deblurring, SR)
  - Computational photography (includes restoration)
  - Segmentation
  - Registration
  - Pattern recognition
  - Many applied subfields – image forensics, cultural heritage conservation etc.
- Computer vision – recognition and 3D reconstruction but growing overlap with image processing
- Machine learning
- Compressive sensing (intersects with computational photography)

Image restoration (inverse problems)

– Denoising
– Deblurring (defocus, camera motion, object motion)

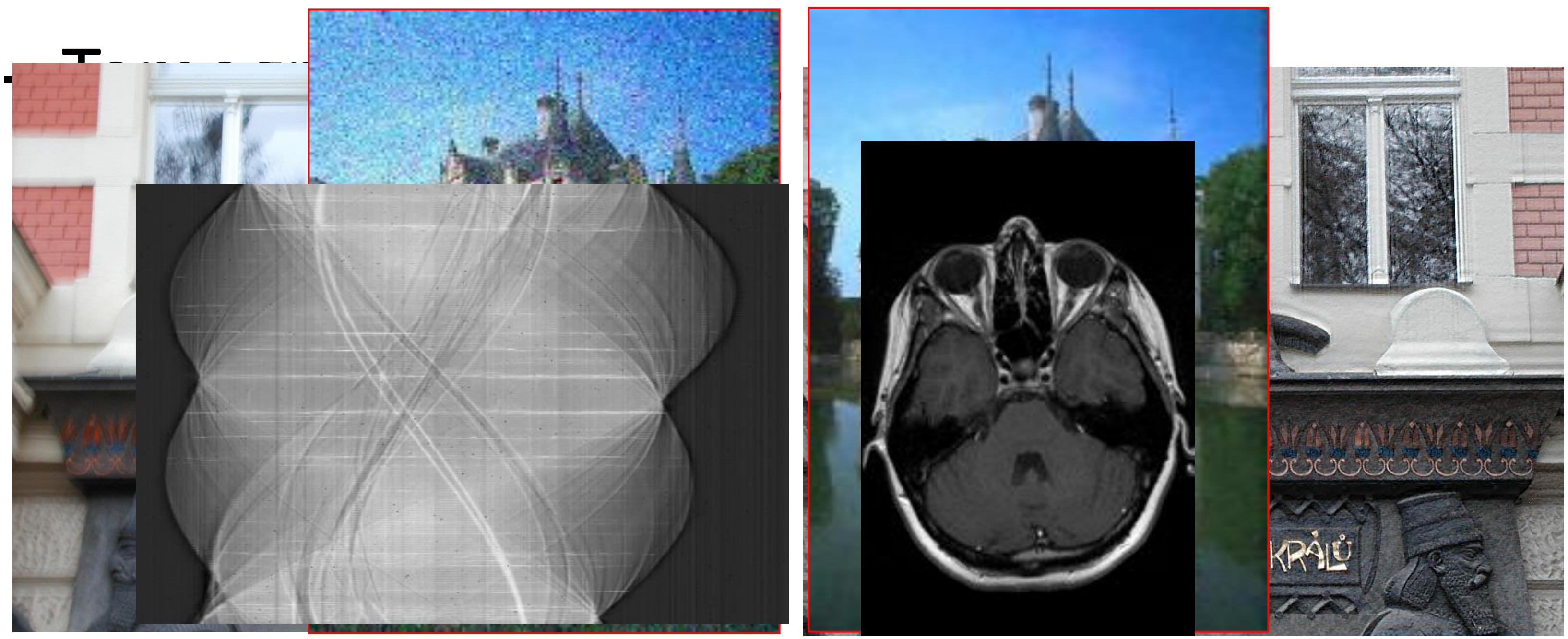


Image segmentation and classification

- Separating objects, categories, foreground/background, cells or organs in biomedical applications etc.

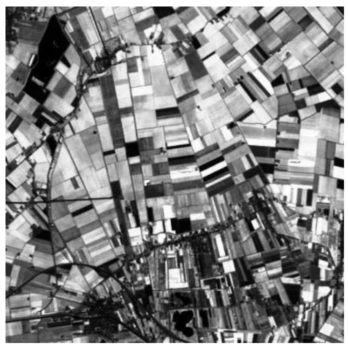
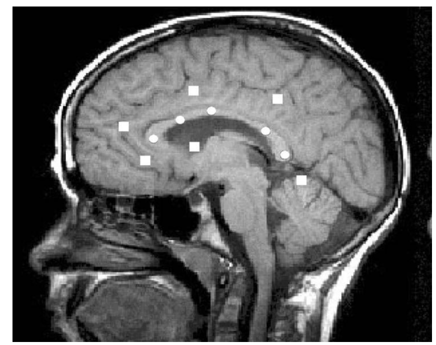
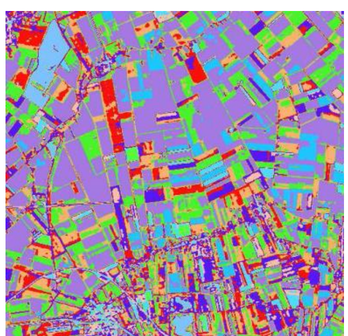
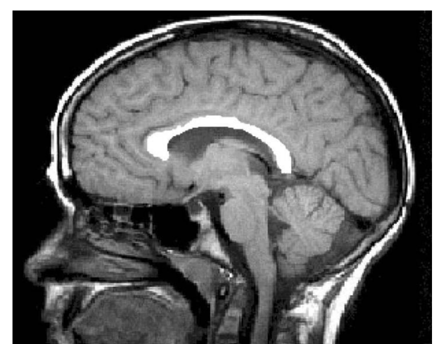
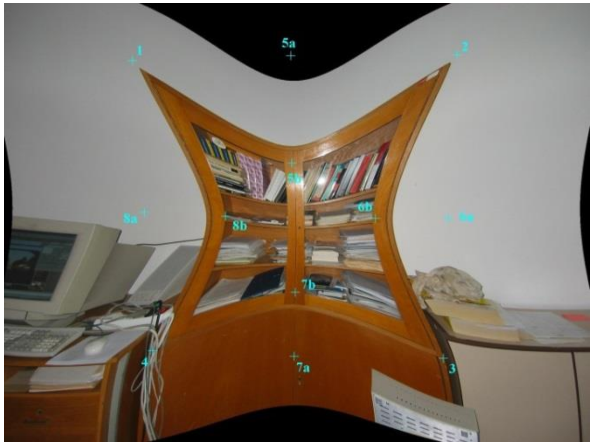
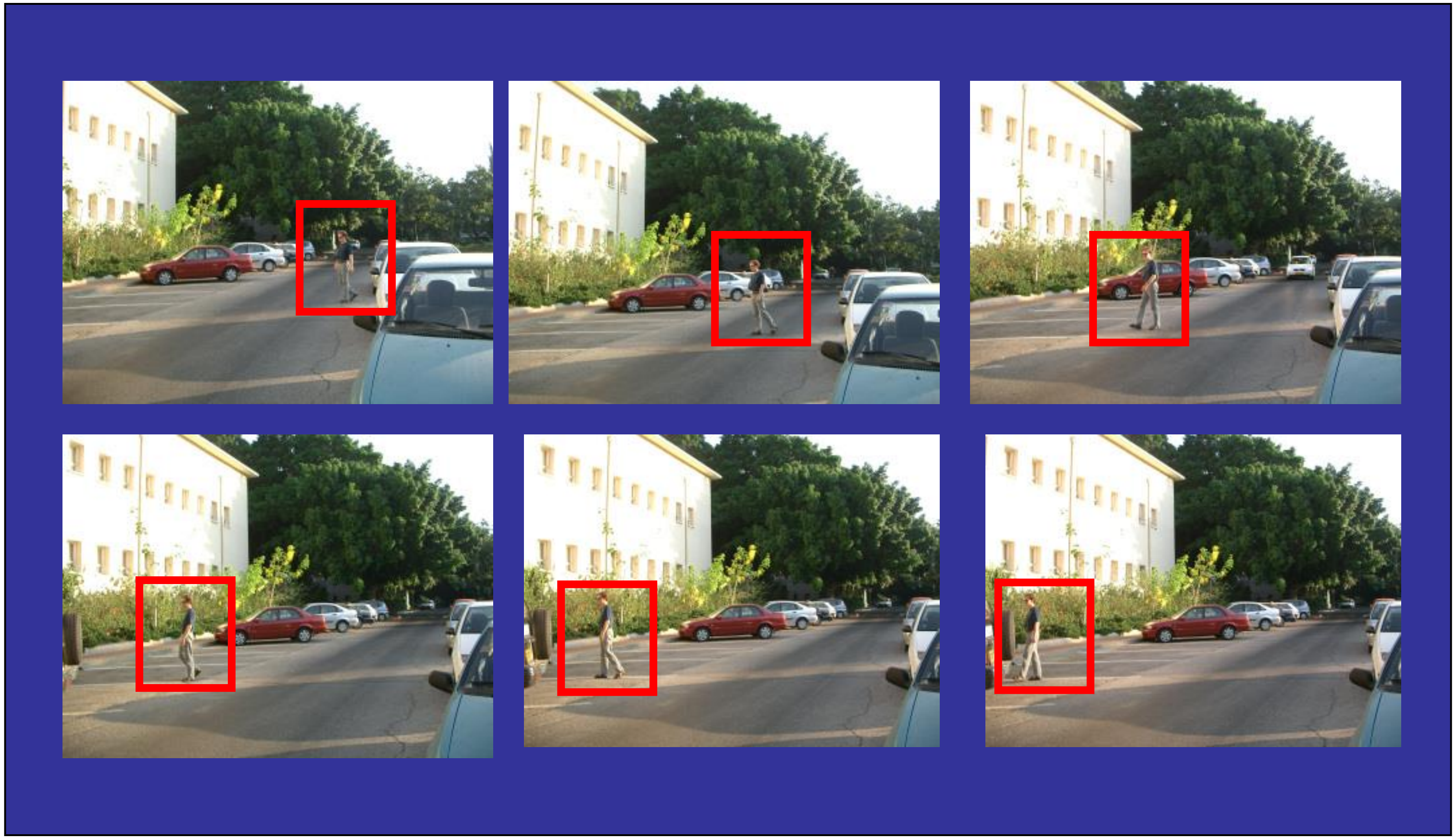





Image Registration

- Transforming different sets of data into one coordinate system
- Transform is constrained to have a specific form (rotation, affine, projective, splines etc.)
- Important general forms – optical flow & stereo



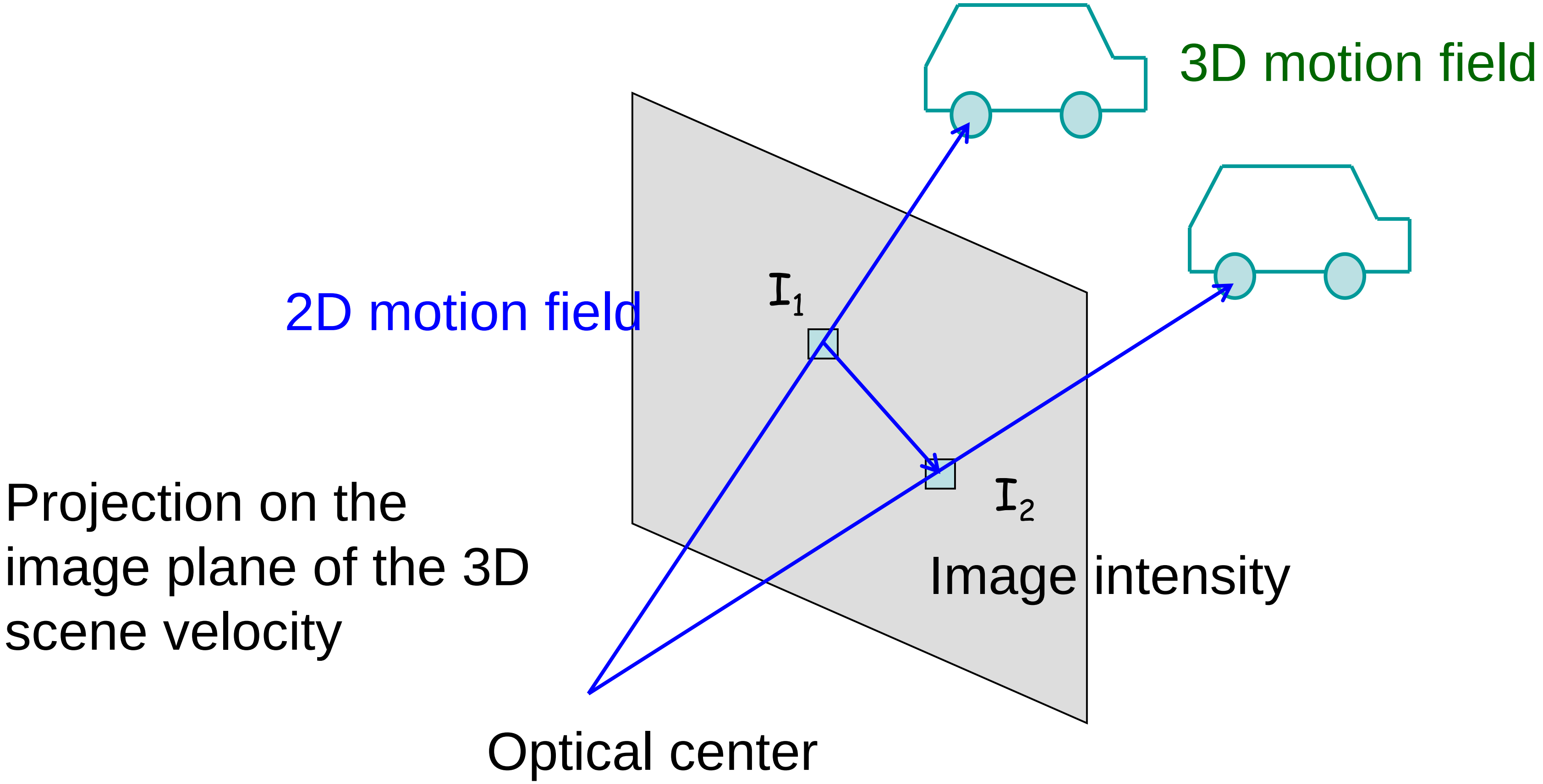
Optical flow



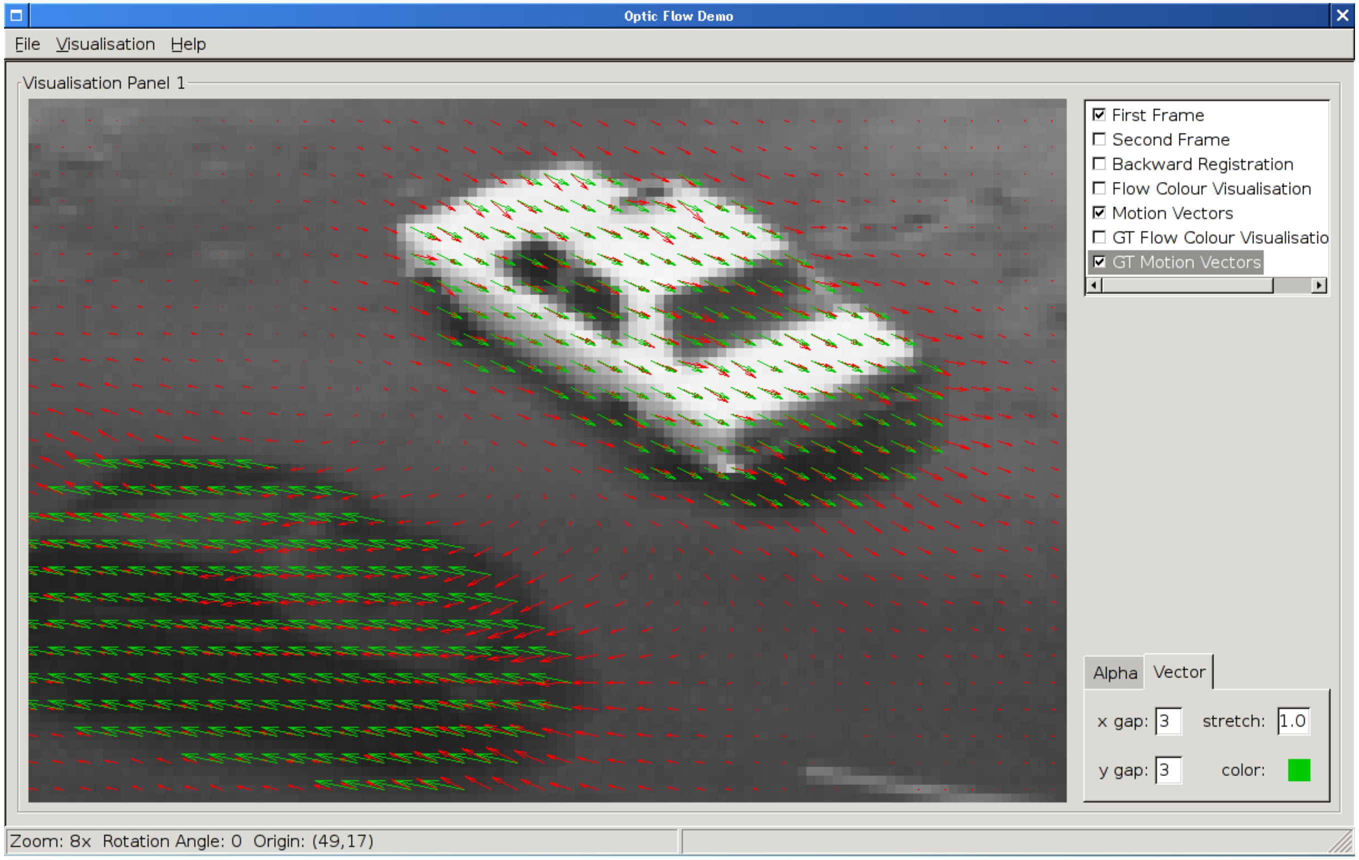
Sequence of images contains information about the scene,

We want to estimate motion – special case of image registration

2D Motion Field = Optical Flow



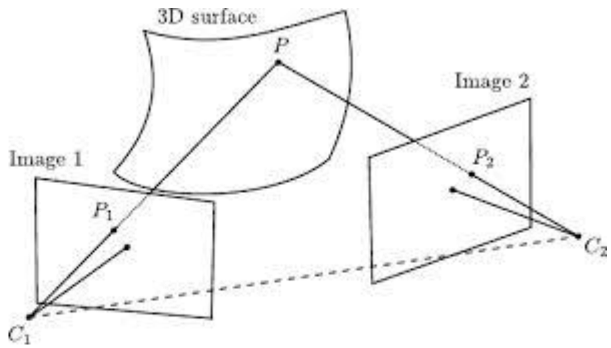
Optical flow example



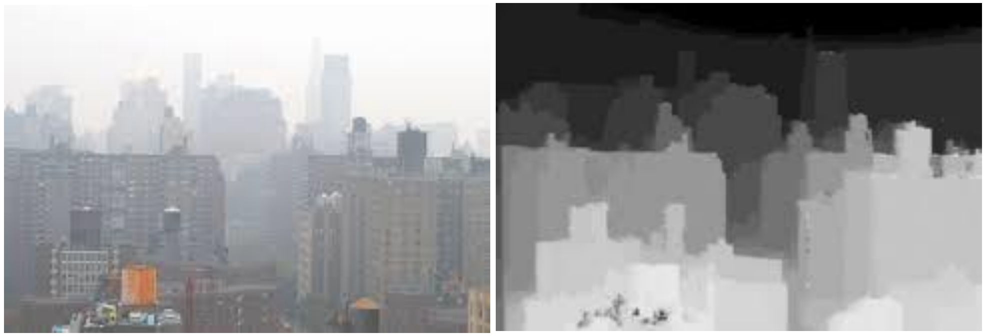
Source: CBIA Brno, http://cbia.fi.muni.cz

Stereo reconstruction

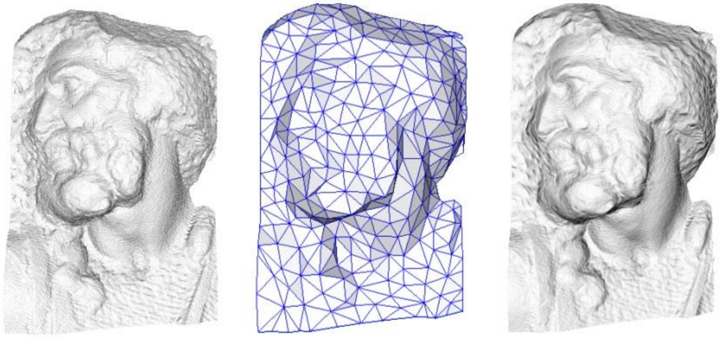
Principle



Result (**depth map** or **disparity map**)



Result (3D model)



original mesh
4M triangles

simplified mesh
500 triangles

simplified mesh
and normal mapping
500 triangles

Source: http://lcav.epfl.ch

Image processing problems

- Image restoration
  - denoising
  - deblurring
  - tomography
- Segmentation and classification
- Image registration
  - optical flow
  - stereo

Mathematical image

- Greyscale image
  - Continuous representation $u : R^2 \rightarrow \langle 0, 1 \rangle$
  - Discrete – matrix or vector $u \in R^{m \times n}, \quad u \in R^{mn}$
  - Both can be extended to 3D
- Color image = set of 3 or more greyscale images
  - RGB channels are highly correlated → many algorithms work with greyscale only

Inverse problems in image restoration

- Denoising
- Linear image degradations
  - Deconvolution and deblurring
  - Super-resolution
  - CT, MRI, PET etc. reconstruction (reconstruction from projections)
- JPEG decompression

Image degradations

- Gaussian noise $z = N(u, \sigma I) = u + N(0, \sigma I)$
- Homogeneous blur = convolution with a kernel h (PSF – Point-spread function)

$$z(x) = \int h(x-s)u(s)ds = h * u = Hu$$

- Spatially-varying blur

$$z(x) = \int h(x-s;s)u(s)ds = Hu$$

Presentation outline

- What is digital image processing? Typical problems and their mathematical formulation.

- Bayesian view of inverse problems in (not only) image restoration, sparsity
- Discrete labeling problems and Markov random fields (MRFs, CRFs)
  – Surprising result: a large family of non-convex MRF problems can be solved exactly in polynomial time/ reformulated as convex optimization problems

Bayesian Paradigm

$$p(u|z) = \frac{p(z|u)\,p(u)}{p(z)}$$

a posteriori distribution
unknown

likelihood
given by our problem

a priori distribution
our prior knowledge

z … observation, u … unknown original image

Maximum a posteriori (MAP): max p(u|z)

Maximum likelihood (MLE): max p(z|u)

MAP corresponds to regularization

$$\max_u p(u|z) \propto p(z|u)p(u)$$

$$\min_u -\log p(u|z) \propto -\log p(z|u) - \log p(u)$$

data term

regularization term

Data term for image denoising

$$\max_u p(u|z) \propto p(z|u)p(u)$$

$$\min_u -\log p(u|z) \propto -\log p(z|u) - \log p(u)$$

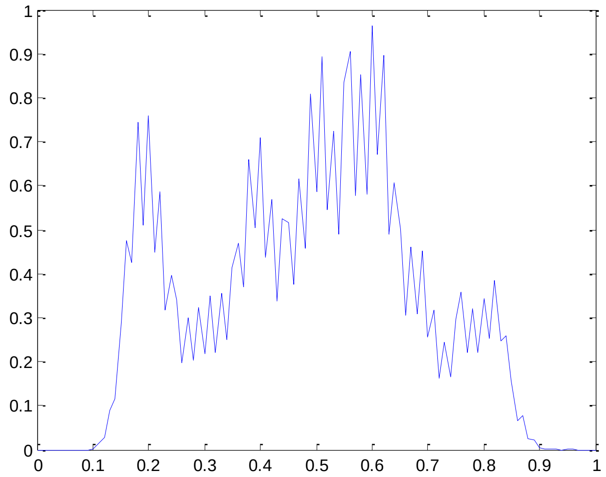
$$p(z|u) = \frac{1}{(2\pi\sigma^2)^{N/2}} \prod_{i=1}^{N} e^{\frac{(z_i-u_i)^2}{2\sigma^2}}$$

$$-\ln p(z|u) = -\ln k \prod_i e^{\frac{(z_i-u_i)^2}{2\sigma^2}} = \boxed{\frac{1}{2\sigma^2}\sum_i (z_i-u_i)^2} + c$$

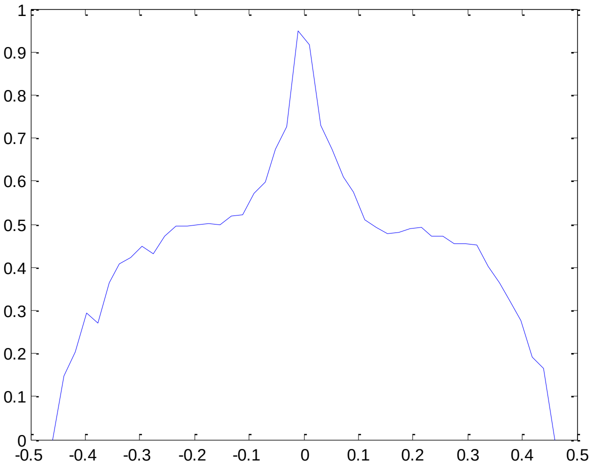
Image Prior

$$\min_u -\log p(u|z) \propto -\log p(z|u) \boxed{-\log p(u)}$$

$$\ln p(\mathbf{u}) = \ln \prod_i p(\mathbf{u}_i) = \sum_i \ln p(\mathbf{u}_i)$$

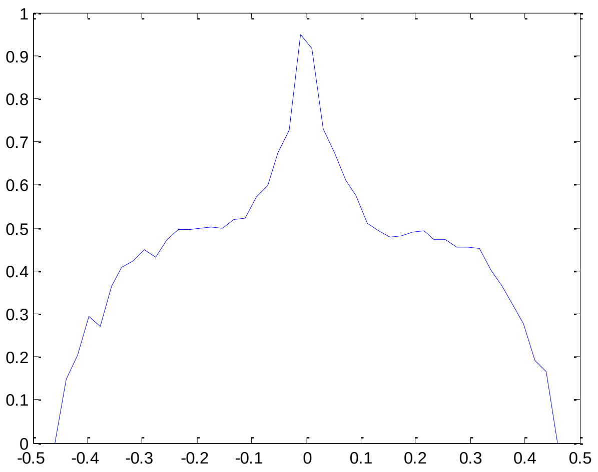


Intensity histogram



Gradient histogram

Image Prior



Gradient histogram

$$p(\mathbf{u}) \propto \prod_i e^{-\lambda \phi(\nabla u_i)}$$

$$\ln p(\mathbf{u}) = -\lambda \sum_i \phi(\nabla u_i) + c$$

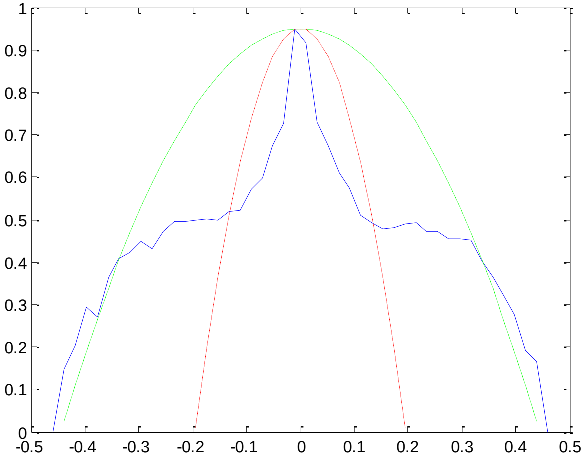
Theory on when we can do this will be given later (CRF)

Tikhonov versus TV Image Prior

$$Q(u) = \lambda \int |\nabla u|^2 = \lambda \|\nabla u\|_2^2$$

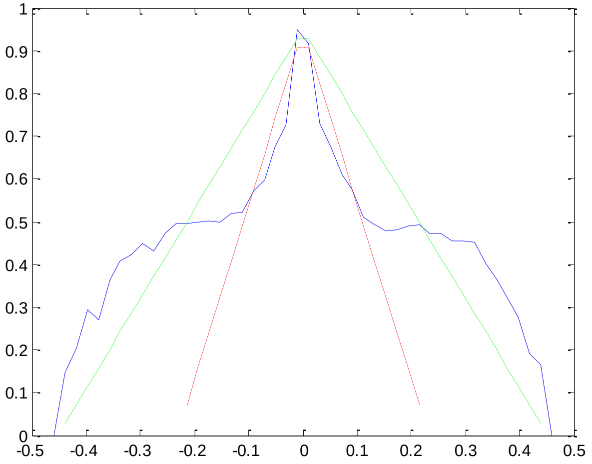
Tikhonov regularization

$$p(\mathbf{u}) \propto \prod_i e^{-\lambda |\nabla u_i|^2} = e^{-\lambda \mathbf{u}^T \mathbf{L} \mathbf{u}}$$



$$Q(u) = \lambda \int |\nabla u| = \lambda \|\nabla u\|_{2,1}$$

TV regularization (isotropic)

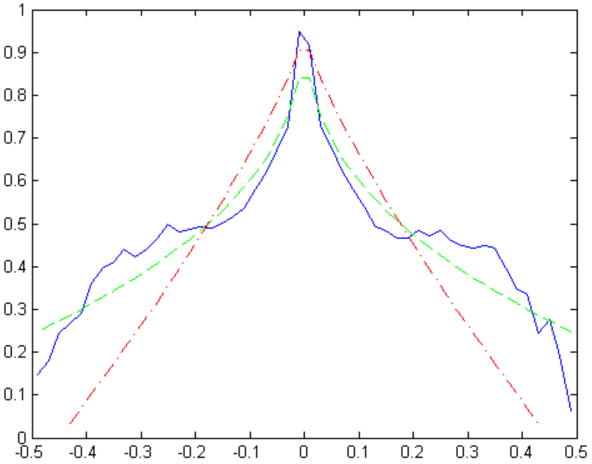


Non-convex Image Prior

$$Q(u) = \lambda \int |\nabla u|^{0.8}$$

$$Q(u) = \lambda \int |\nabla u|^{0.4}$$

Non-convex regularization



Bayesian MAP approach for denoising

$$-\ln p(u|z) = -\ln p(z|u) - \ln p(u)$$

$$\frac{1}{2\sigma^2}\sum_i (z_i - u_i)^2 + \lambda \sum_i \phi(|\nabla u_i|)$$

$$\min_{\mathbf{u}} \frac{1}{2\sigma^2}\sum_i (z_i - u_i)^2 + \lambda \sum_i |\nabla u_i|_p^p$$

Analysis-based **sparsity**

- TV regularization can be extended to other sparse representations

$$\min_u \frac{1}{2}\|z-u\|^2 + \lambda\|\nabla u\|_{2,1}$$
$$\min_u \frac{1}{2}\|z-u\|^2 + \lambda\|Wu\|_1$$

- W often a set of convolutions with highpass filters
  - Wavelets (property of the Daubechie wavelets)
  - Learned by PCA

Synthesis-based sparsity

Bayesian approach applied on transform coefficients:

$$\min_u \frac{1}{2}\|z-u\|^2+\lambda\|Wu\|_1$$

$$\Downarrow$$

$$\min_u \frac{1}{2}\|z-W^T w\|^2+\lambda\|w\|_1$$

(for a Parseval frame W)

PETER G. CASAZZA AND JANET C. TREMAIN: A BRIEF INTRODUCTION TO HILBERT SPACE FRAME THEORY AND ITS APPLICATIONS

Measures of sparsity

- $l_p\,, \quad 0 < p \leq 1$ norms $\|\mathbf{a}\|_p^p$

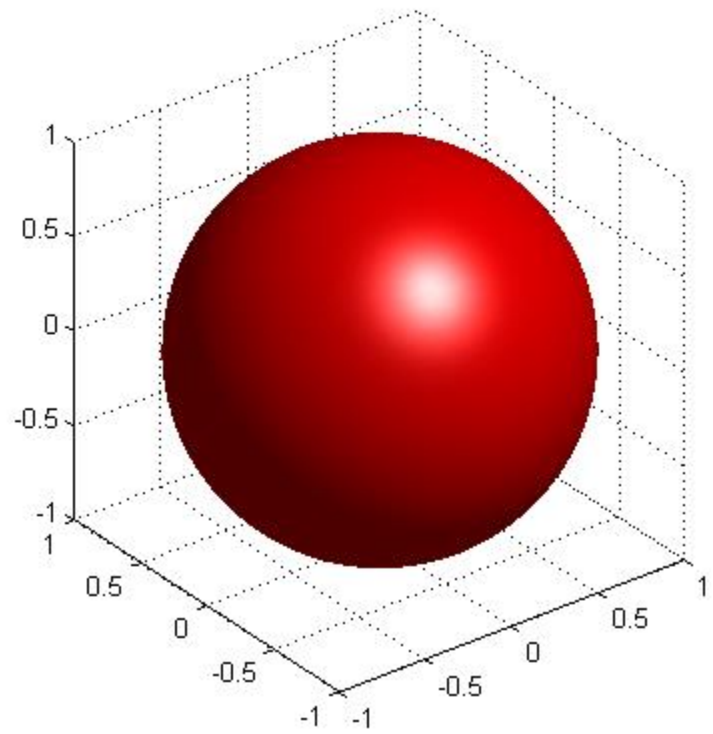
$$\|\mathbf{a}\|_p = \left(\sum_i |a_i|^p\right)^{\frac{1}{p}}$$

- l_0 norm, counts nonzero elements
- many other sparsity measures
  - smooth l_1

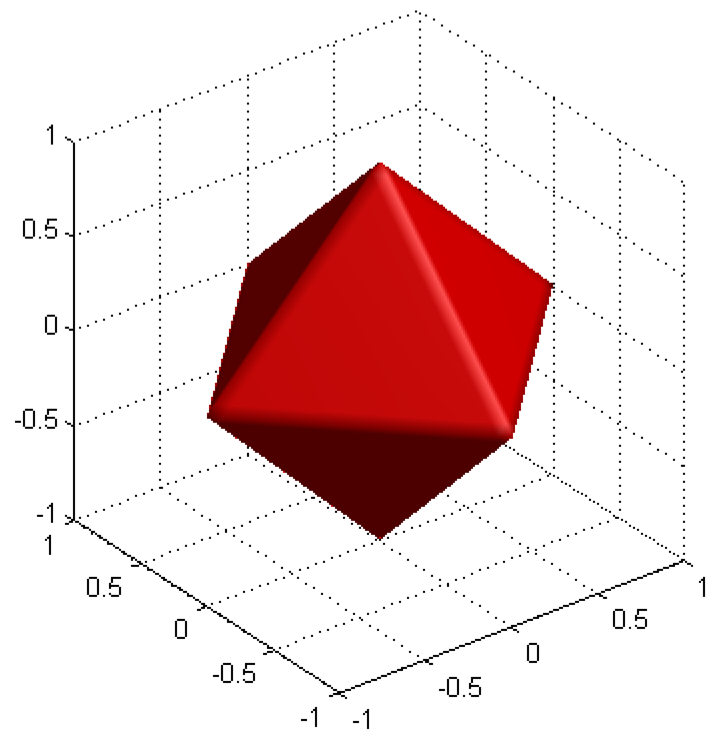
$$\rho(\mathbf{a}) = \|\mathbf{a}\|_1 - \epsilon \log\left(1 + \frac{\|\mathbf{a}\|_1}{\epsilon}\right)$$

- l_1 is the only sparsity enforcing convex p-norm

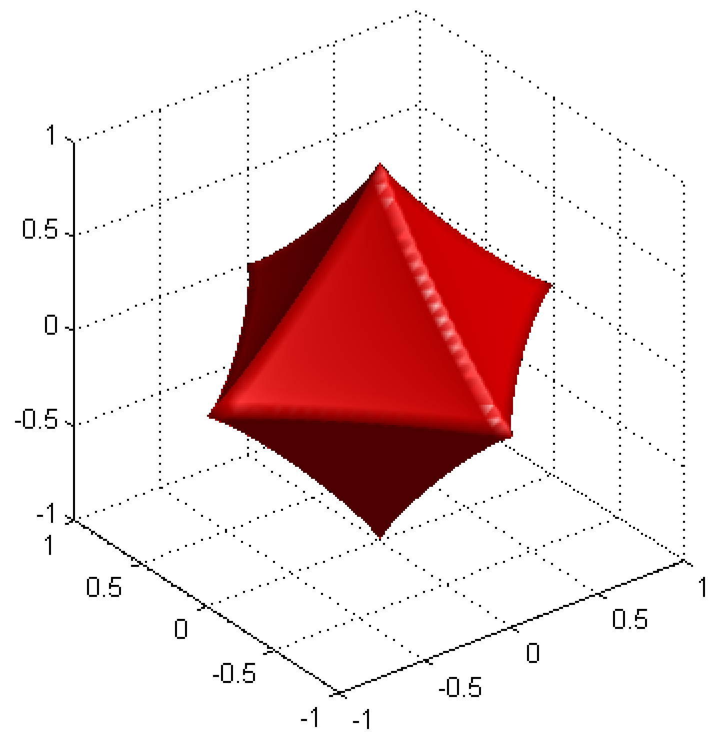
l_2 unit ball



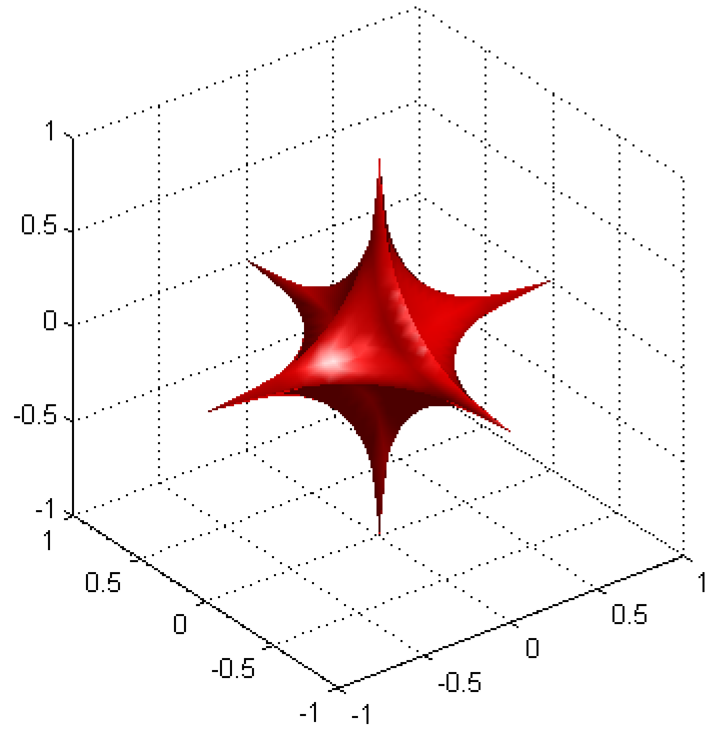
l_1 unit ball



$l_{0.9}$ unit ball

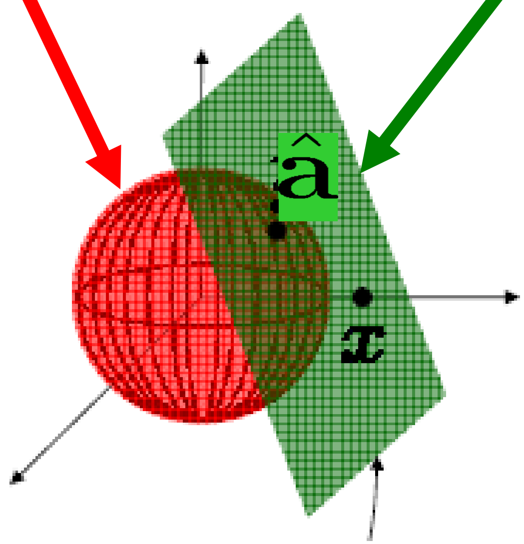


$l_{0.5}$ unit ball



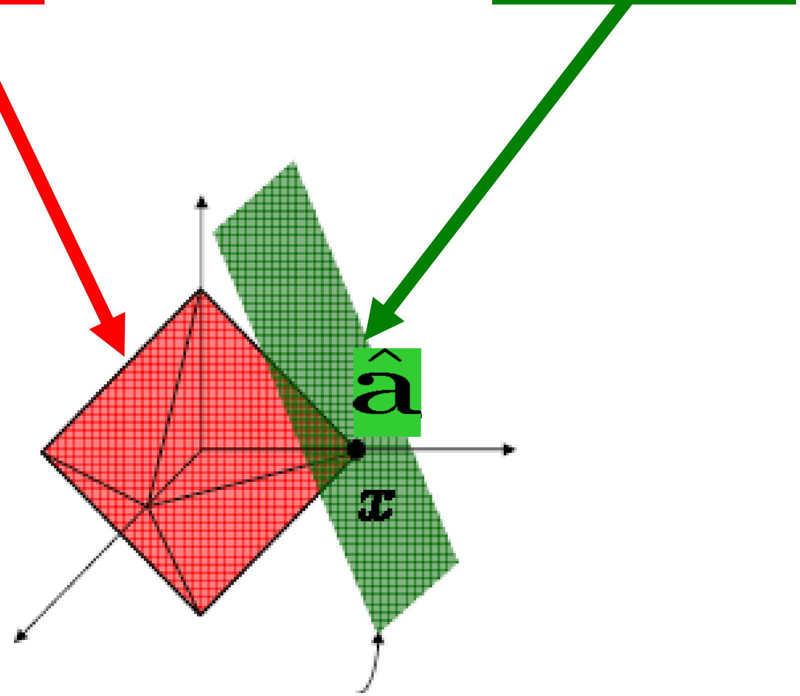
l_2-norm

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}} \|\mathbf{a}\|_2^2 \quad \text{subject to} \quad \mathbf{D}\mathbf{a} = \mathbf{x}$$



l_1-norm

$$\hat{\mathbf{a}} = \arg\min_{\mathbf{a}} \|\mathbf{a}\|_1 \quad \text{subject to} \quad \mathbf{Da} = \mathbf{x}$$



Deblurring

- Denoising $z = u + N(0, \sigma^2 I)$

$$\min_u \frac{1}{2}\|z - u\|^2 + \lambda\|\nabla u\|_{2,1}$$

- Deblurring

$$z = h * u + N(0, \sigma^2 I) = Hu + N(0, \sigma^2 I)$$

$$\min_u \frac{1}{2}\|z - Hu\|^2 + \lambda\|\nabla u\|_{2,1}$$

Super-resolution (with deblurring)

Several possibly shifted blurred images

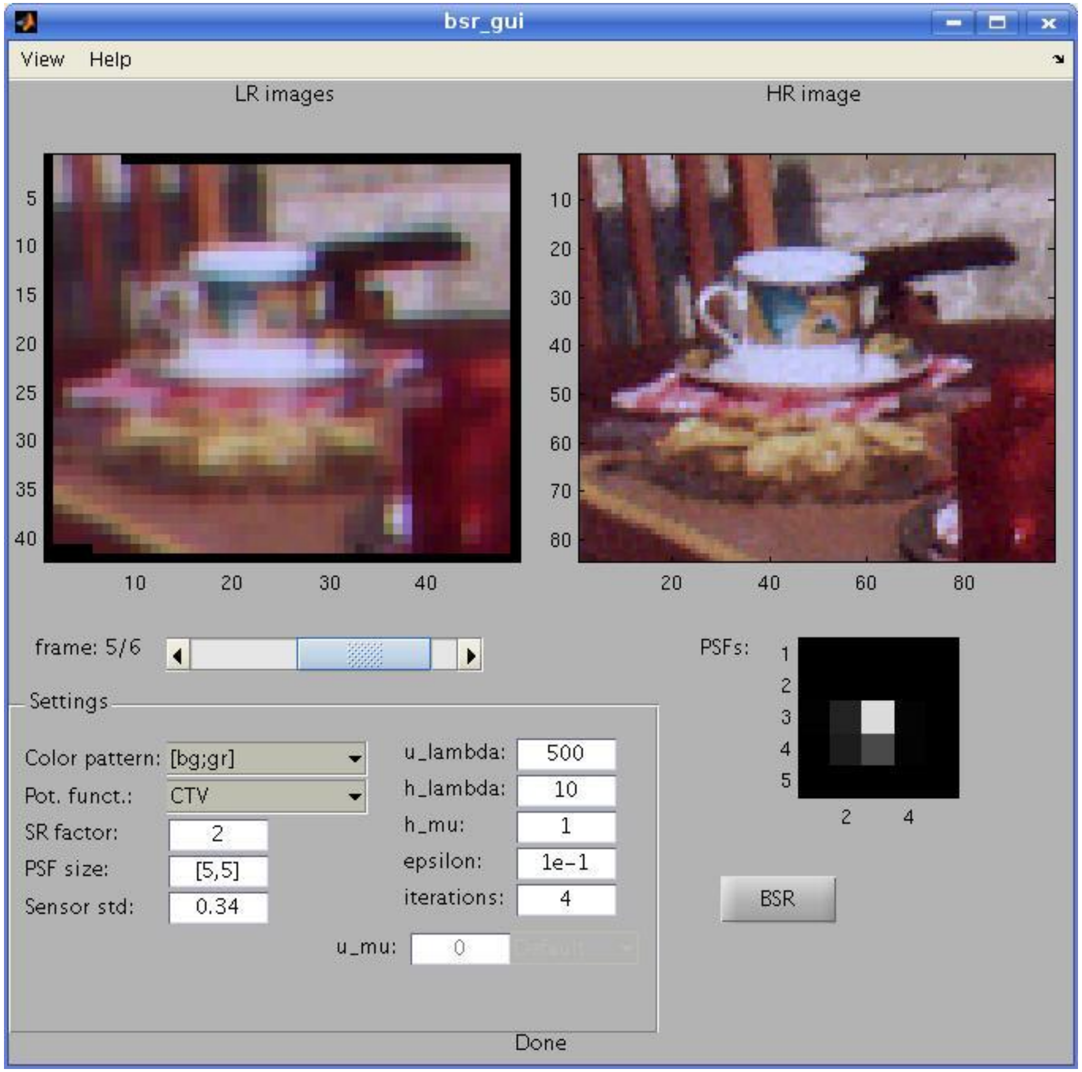
$$z_i = DH_i u + N(0, \sigma^2 I)$$

$$\min_u \frac{1}{2} \sum_i \|z_i - DH_i u\|^2 + \lambda \|\nabla u\|_{2,1}$$

D_i … downsampling operator

Convolutions represent also the shift

Super-resolution

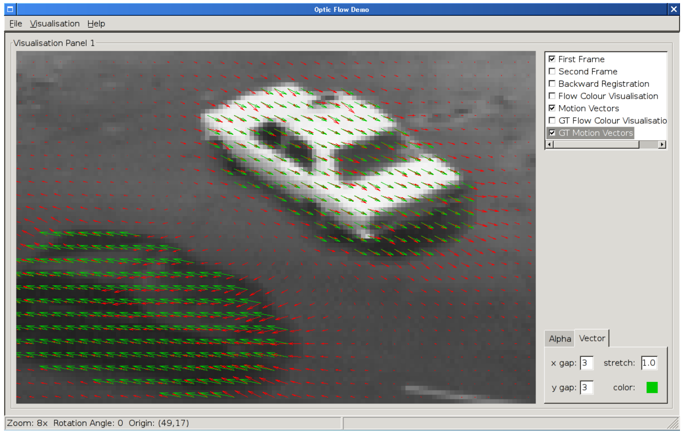


$$\min_u \frac{1}{2} \sum_i \|z_i - DH_i u\|^2 + \lambda \|\nabla u\|_{2,1}$$

http://zoi.utia.cas.cz/bsr-toolbox

Optical flow

- Based on the assumption of constant brightness and Taylor series



$$I(t, x(t), y(t)) = I(t_0, x(t_0), y(t_0))$$

$$\left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}\right) \cdot \nabla I + \frac{\partial I}{\partial t} = 0 \text{ at } t = t_0$$

- Optical flow is the velocity field

$$\mathbf{v}(t_0) = \left(\frac{\partial x}{\partial t}(t_0), \frac{\partial y}{\partial t}(t_0)\right)$$

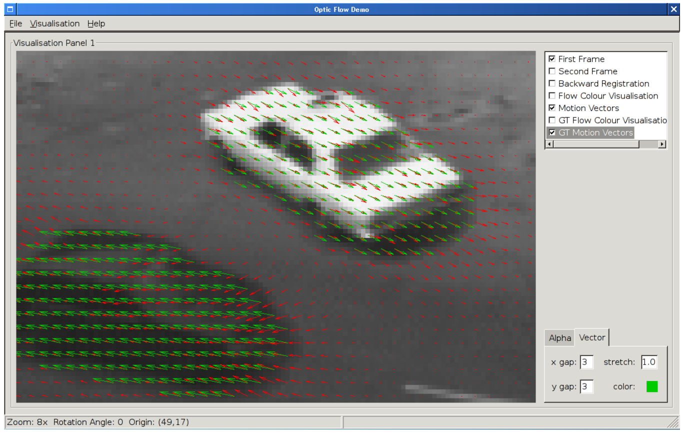
Optical flow

$$\min_{\mathbf{v}} \frac{1}{2} \int_{\Omega} (\nabla I \cdot \mathbf{v} + I_t)^2 dx + \lambda \sum_{i=1}^{2} \|\nabla \mathbf{v}_i\|_{2,1}$$

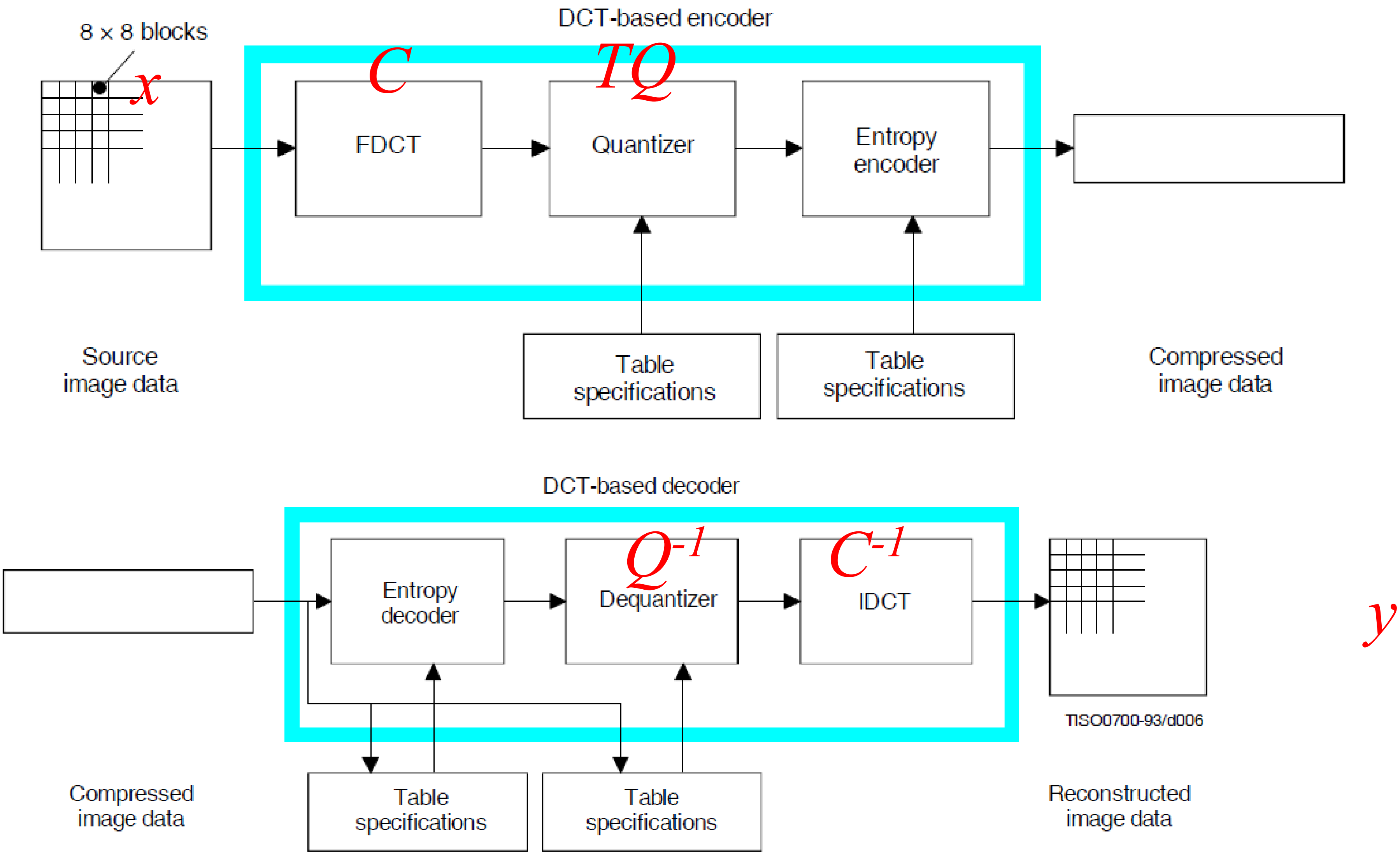
Data term

Weighting parameter

Regularization term



JPEG compression



50
jpg

Bayesian MAP restoration

MAP – maximum a posteriori probability

$$\min_u -\log p(z|u) - \log p(u)$$

$$-\log p(u) = \tau \|Wu\|_1$$

$$-\log p(z|u) = \begin{cases} 0 & QCu \in (QCz - 0.5, QCz + 0.5\rangle \\ \infty & otherwise \end{cases}$$

C … 2D cosine transform (orthogonal 64x64 operator)
Q … diagonal quantization operator (division by entries q_i of the quantization table)

Bayesian JPEG decompression

Using total variation (TV)

$$\min_u \|\nabla u\|_{2,1}, s.t. \quad QCu \in (QCz - 0.5, QCz + 0.5\rangle$$

(Bredies and Holler, 2012)

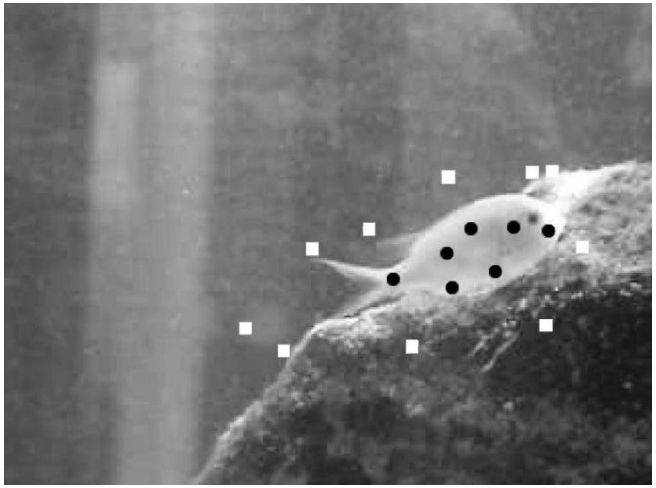
Or using redundant wavelets

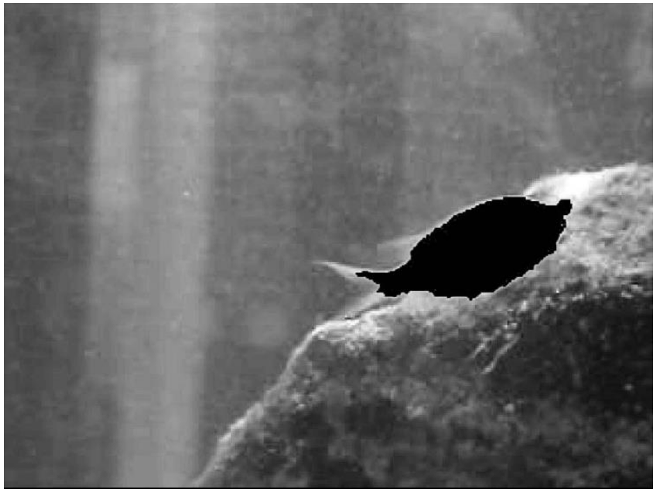
$$\min_u \|Wu\|_{2,1}, s.t. \quad QCu \in (QCz - 0.5, QCz + 0.5\rangle$$

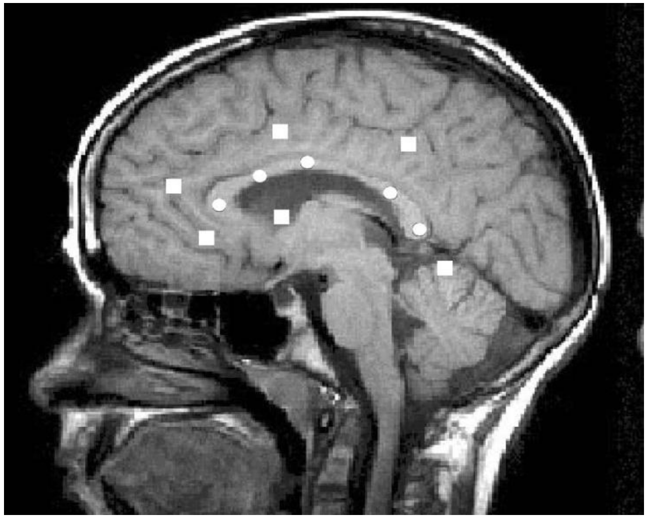
C … 2D cosine transform (orthogonal 64x64 operator)
Q … diagonal quantization operator (division by entries q_i of the quantization table)

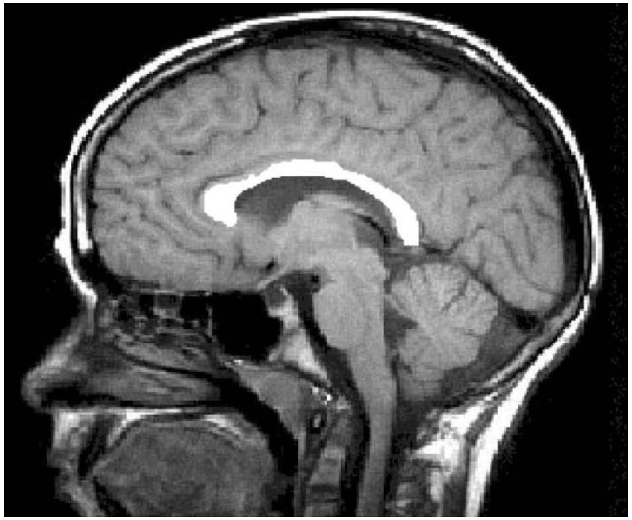
50
jpg

50
est







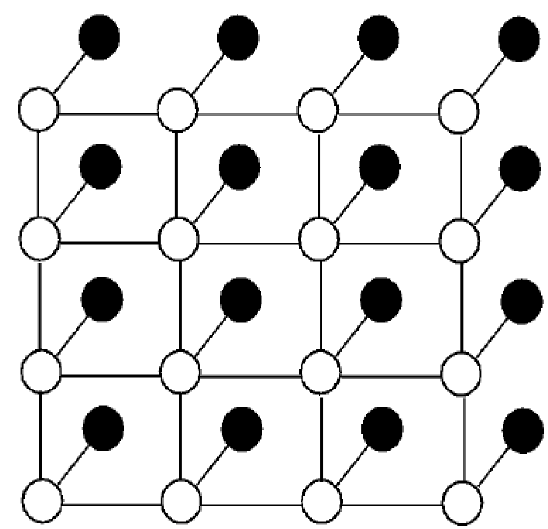


Convex variational problems

- Denoising, deblurring, SR, optical flow, JPEG decompression …
- Solution by convex optimization (interior point, proximal methods)
  N. Parikh, S. Boyd: Proximal Algorithms
- What to do for discrete or non-convex problems such as segmentation and stereo?

Discrete labeling problems

- For each site (pixel) we look for a label (or a vector of labels)
- Labels depend on local image content and a smoothness constraint
- Image restoration, segmentation, stereo, and optical flow are all labeling problems

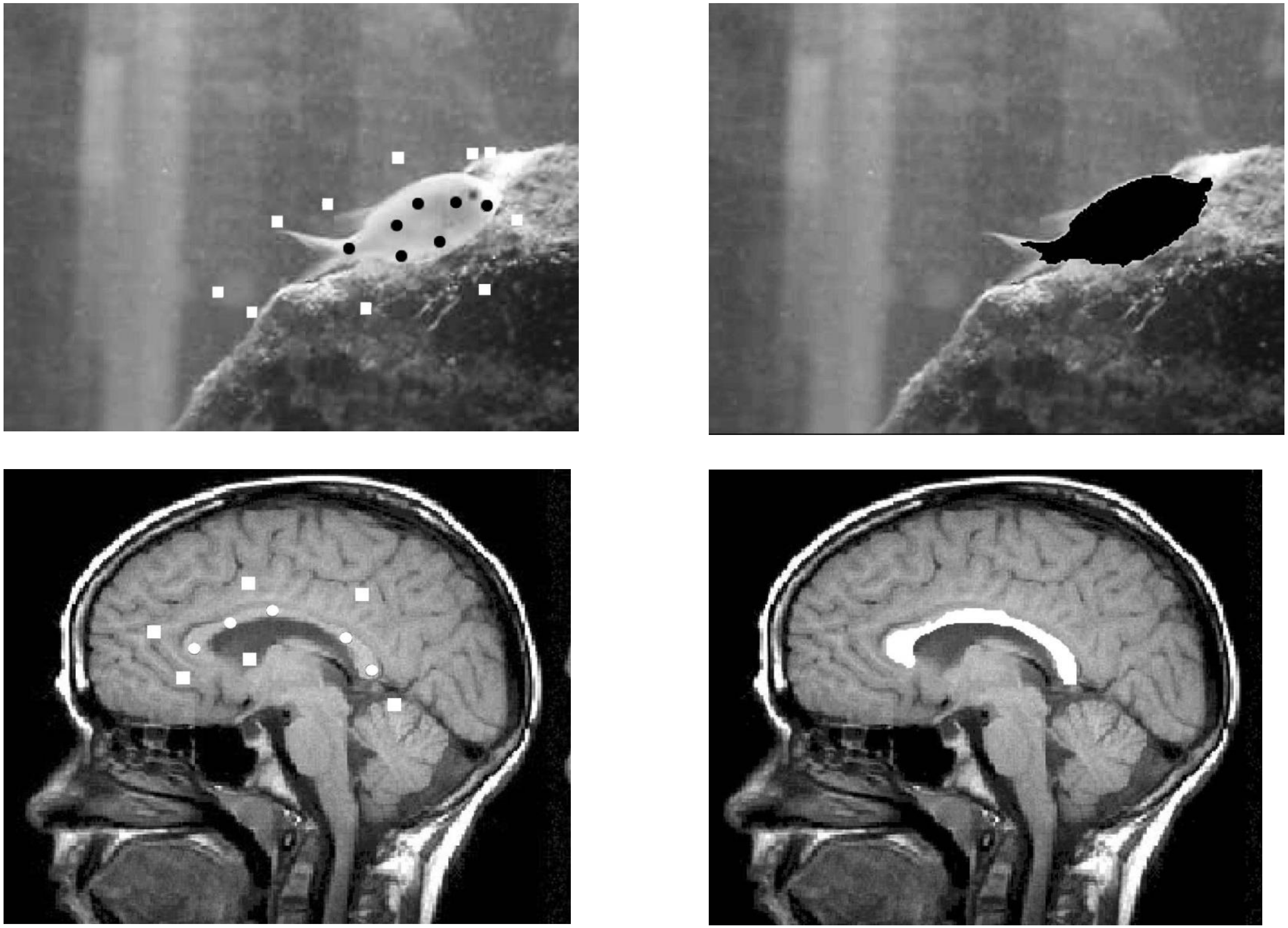


Discrete labeling problems

- For each site (pixel) we look for a label (or a vector of labels)
- Labels depend on local image content and a smoothness constraint

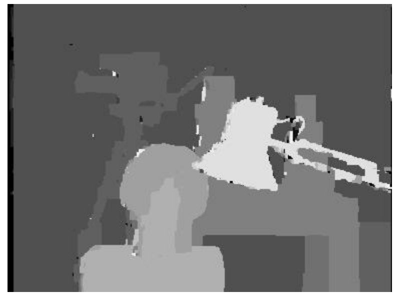
Segmentation	foreground/background or object number	{0,1} {1.. k}
Stereo	disparity (inverse depth)	-k..k
Optical flow	local motion vector	(-k..k) x (-k..k)
Restoration	intensity	0..255

Segmentation by graph cuts



Graph cuts & Belief propagation

Graph cuts



Belief propagation

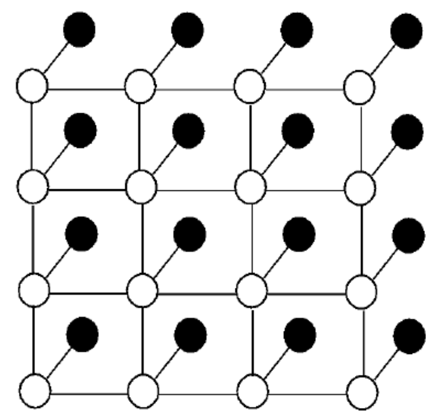
„Classical local algorithms“

Markov Random Fields (MRFs)

- Markov Random Field, Gibbs Random Field
  - MRF $\Leftrightarrow$ GRF (Hammersley-Clifford theorem)
- MRF models including smoothness priors
  - stereo
  - segmentation
  - restoration (denoising, deblurring)
- Discrete optimization on MRFs based on graph cuts

Markov Random Field (MRF)

- sites S = {1, ... , m}
- F ... set of random variables defined on S
- N ... **neighborhood system**
- $f_i \in \mathcal{L}$ ... (possibly discrete) label
- **configuration** f = {f_1 ... f_K},



$$P(f_i | f_{S-\{i\}}) = P(f_i | f_{N_i})$$

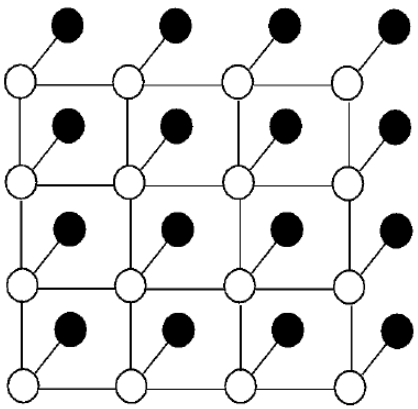
$$P(f) > 0$$

- Other possible properties – homogeneity, isotropy

Gibbs Random Field

$$P(f) = \frac{1}{Z} e^{-\frac{1}{T}U(f)}$$

P(f) > 0 !



Partition function $Z = \sum_{f} e^{-\frac{1}{T}U(f)}$

Energy function U(f)

$$U(f) = \sum_{c \in \mathcal{C}} V_c(f) = \sum_{i \in \mathcal{S}} V_1(f_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{N}_i} V_2(f_i, f_{i'})$$

$V_c(f)$ ... clique potentials

Hammersley-Clifford theorem

MRF = GRF

F is an MRF on S with respect to N

if and only if

F is a Gibbs random field on S with respect to N

MRF ... conditional independence of non-neighbor nodes (variables)

GRF ... global function depending on local “compatability functions”

Hammersley-Clifford theorem - proof

- An MRF is also a GRF – complicated, introduction of canonical potentials needed
- A GRF is a MRF $P(f_i | f_{S-\{i\}}) = P(f_i | f_{N_i})$

$$P(f_i | f_{S-\{i\}}) = \frac{P(f)}{\sum_{f_i \in \mathcal{L}} P(f')} = \frac{e^{-\sum_{c \in \mathcal{C}} V_c(f)}}{\sum_{f'_i} e^{-\sum_{c \in \mathcal{C}} V_c(f)}}$$

$$P(f_i | f_{S-\{i\}}) = \frac{e^{-\sum_{\{c, i \in c\}} V_c(f)}}{\sum_{f'_i} e^{-\sum_{\{c; i \in c\}} V_c(f)}}$$

MRF = GRF

- MAP-MRF

$$\max_f p(f) = \frac{1}{Z} e^{-E(f)}$$

$$\min_f (-\ln p(f)) = \min_f E(f) + const$$

- How to incorporate smoothness?
  – Penalties/potentials similar for most applications

Smoothness prior

Priors on derivatives, usually first derivative

$$V(f_i, f_j) = \kappa_{ij}\ \delta(f_i - f_j)$$

segmentation, sometimes in stereo

$$V(f_i, f_j) = \kappa_{ij}\ (f_i - f_j)^2$$

Tikhonov regularization

Discontinuity preserving penalties

TV regularization

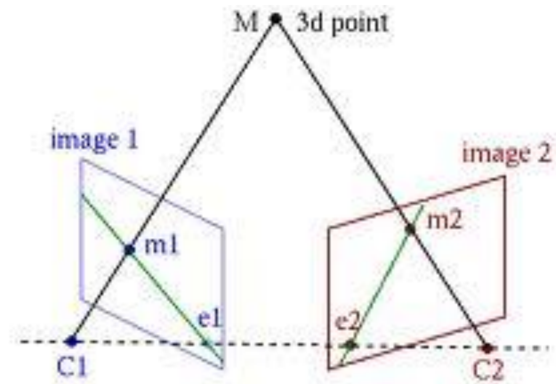
$$V(f_i, f_j) = \kappa_{ij}\ |f_i - f_j|$$

line process, Mumford-Shah functional

$$V(f_i, f_j) = \kappa_{ij} \min((f_i - f_j)^2, const)$$

MAP-MRF for stereo (Boykov & al.)

2 images d^1, d^2 on the input

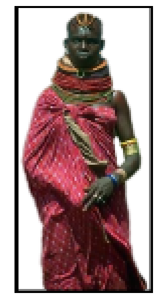


$$E(f) = \sum_i V_1(f_i, d^1, d^2) + \kappa \sum_i \delta(f_{i+1} - f_i)$$

Birchfield-Tomasi matching cost – insensitive to sampling:

$$V_1(f_i, d^1, d^2) = \min\left(\min_{\triangle \in \langle f_i - \frac{1}{2}, f_i + \frac{1}{2} \rangle} |d_i^1 - d_{i+\triangle}^2|, \ldots, const \right)^2$$

MAP-MRF for segmentation



- “ “GrabCut” — Interactive Foreground Extraction using Iterated Graph Cuts”, C. Rother, V. Kolmogorov, A. Blake, SIGGRAPH 2004

MAP-MRF for segmentation

- "Grab cut" example

$$V_2(f_i, f_j) = \kappa_{ij}\, \delta(f_i - f_j) = \gamma e^{-\frac{||d_i - d_j||^2}{2\sigma^2}} \delta(f_i - f_j)$$

$$V_1(f_i, d_i) \cong -\ln p(f_i|d_i) \cong -\ln p(d_i|f_i) - \ln p(f_i)$$

$V_1(f_i,d_i)$ ~ probability to be in fg/bg based on a feature space (intensities, texture features etc...) – modeled for example as a mixture of Gaussians

MAP-MRF for restoration

- Denoising (with anisotropic TV regularization)
  - 2D indexing - only this slide

$$E(f) = \frac{1}{\sigma^2}\sum_{ij}(f_{ij} - d_{ij})^2 + \kappa\sum_{ij}|f_{i+1,j} - f_{ij}| + \kappa\sum_{ij}|f_{i,j+1} - f_{ij}|$$

- Deblurring (with TV regularization)

$$E(f) = \frac{1}{\sigma^2}||f * h - d||^2 + \kappa\sum_{ij}|f_{i+1,j} - f_{ij}| + \kappa\sum_{ij}|f_{i,j+1} - f_{ij}|$$

- Discrete methods not efficient for restoration!

MRFs - Summary

- Common framework for many image processing a CV problems
- Fits well to the Bayesian framework
- MRF = GRF

MAP-MRF using graph cuts

- MAP – Maximum a posteriori probability

$$\max_f p(f) = \frac{1}{Z} e^{-E(f)}$$

$$\min_f (-\ln p(f)) = \min_f E(f) + const$$

- Graph cuts = min-cut ~ max-flow (Ford-Fulkerson theorem)
- Much better than simulated annealing based methods, often very close to global optimum

Graph cuts minimization

$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

For $V_2 \geq 0$ metric

- $V_2(a,b) = 0 \Leftrightarrow a = b$
- $V_2(a,b) = V_2(b,a)$ (actually not necessary)
- $V_2(a,b) \leq V_2(a,c) + V_2(c,b)$

or semimetric (without Δ-inequality)

Metric:

$$\delta(f_i - f_j)$$

$$\min(|f_i - f_j|, const)$$ for any norm |.|

Semimetric:

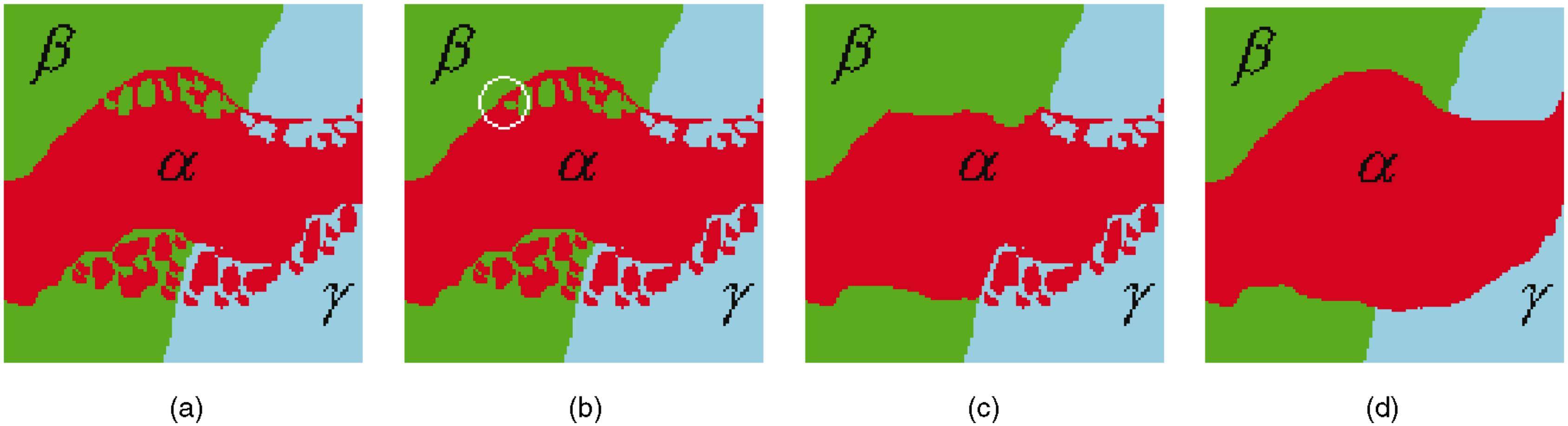
$$\min((f_i - f_j)^2, const)$$

Graph cuts minimization

$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

- General strategy – minimum if no possible decrease of E(f) in one "move"
- Iterated conditional modes (ICM) iteratively minimizes each node (pixel) ➡ easily gets trapped in a local minimum (~ gradient descent)
- Simulated annealing – global moves but without any specific direction ➡ slow
- Graph cuts – use much larger set of "moves" so that the minimum over the whole set can be found in a reasonable (polynomial) time

α-β swap and α-expansion moves



initial labeling

single pixel ICM move

α-β swap move

α-expansion move

α-expansion algorithm

1. Start with an arbitrary labeling f
2. Set success := 0
3. For each label $\alpha \in \mathcal{L}$
    3.1. Find $\hat{f} = \arg\min E(f')$ among f' within one α-expansion of f
    3.2. If $E(\hat{f}) < E(f)$, set f := $\hat{f}$ and success := 1
4. If success = 1 goto 2
5. Return f

- Arbitrary **metric** $V_2(\alpha,\beta)$ (Δ-inequality)
- Not worse than 2x optimum

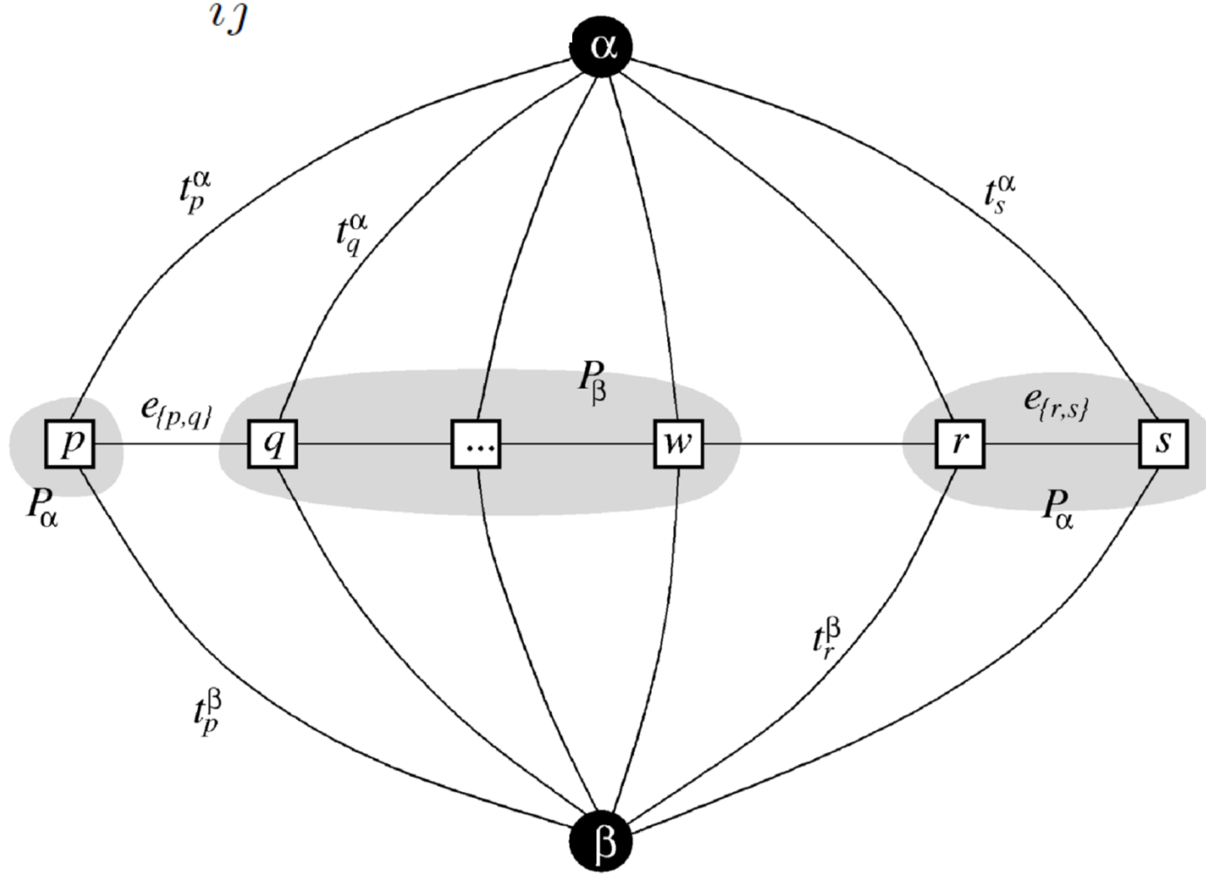
α-β swap algorithm

1. Start with an arbitrary labeling f
2. Set success := 0
3. For each pair of labels $\{\alpha, \beta\} \subset \mathcal{L}$
   3.1. Find $\hat{f} = \arg\min E(f')$ among f' within one α–β swap of f
   3.2. If $E(\hat{f}) < E(f)$, set f := $\hat{f}$ and success := 1
4. If success = 1 goto 2
5. Return f

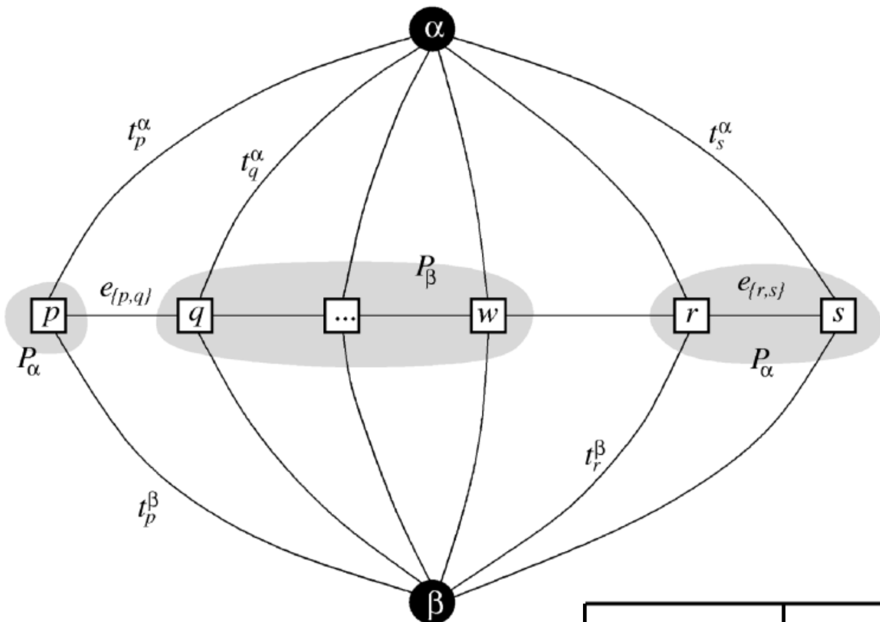
- Arbitrary **semimetric** $V_2(\alpha,\beta)$ (without Δ-inequality)
- No optimality guaranteed

α-β swap move graph

$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$



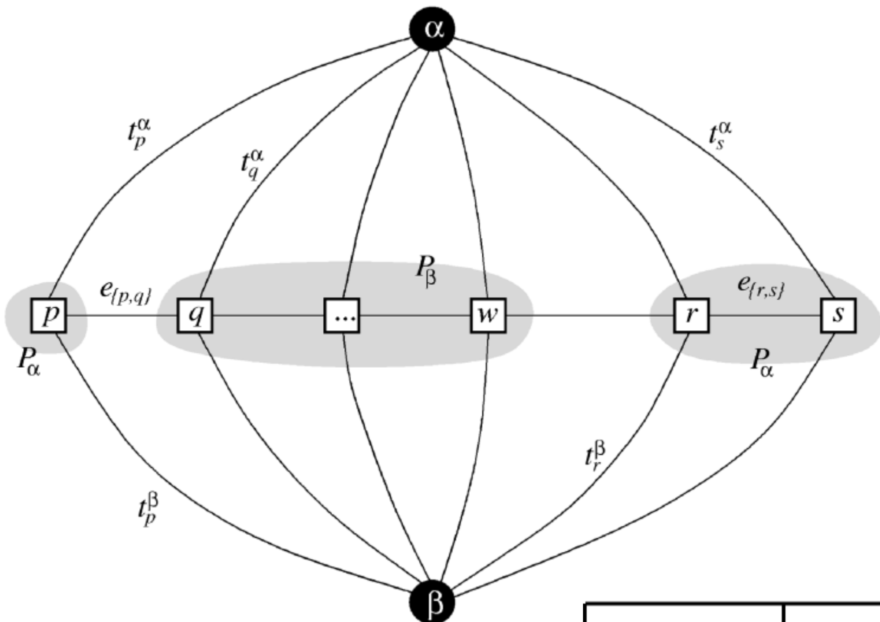
α-β swap move graph



$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

t_p^α	$V_p(\alpha) + \sum_{\substack{q \in \mathcal{N}_p \\ q \notin \mathcal{P}_{\alpha\beta}}} V(\alpha, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
t_p^β	$V_p(\beta) + \sum_{\substack{q \in \mathcal{N}_p \\ q \notin \mathcal{P}_{\alpha\beta}}} V(\beta, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
$e_{\{p,q\}}$	$V(\alpha, \beta)$	$\{p,q\} \in \mathcal{N}$ $p,q \in \mathcal{P}_{\alpha\beta}$

α-β swap move graph

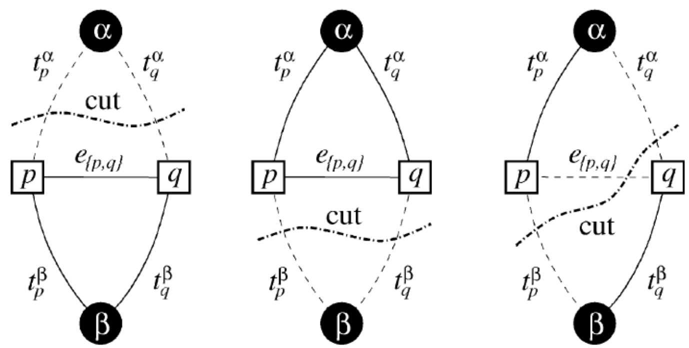


$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

Proof step 1: For each p in the set $P_{\alpha\beta}$, the minimum cut contains exactly one edge t_p

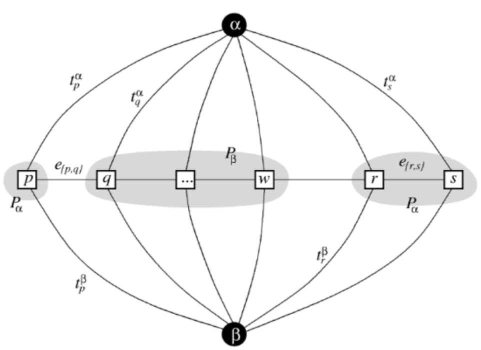
t_p^α	$V_p(\alpha) + \sum_{q \in \mathcal{N}_p, q \notin \mathcal{P}_{\alpha\beta}} V(\alpha, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
t_p^β	$V_p(\beta) + \sum_{q \in \mathcal{N}_p, q \notin \mathcal{P}_{\alpha\beta}} V(\beta, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
$e_{\{p,q\}}$	$V(\alpha, \beta)$	$\{p,q\} \in \mathcal{N}$, $p,q \in \mathcal{P}_{\alpha\beta}$

α-β swap move graph



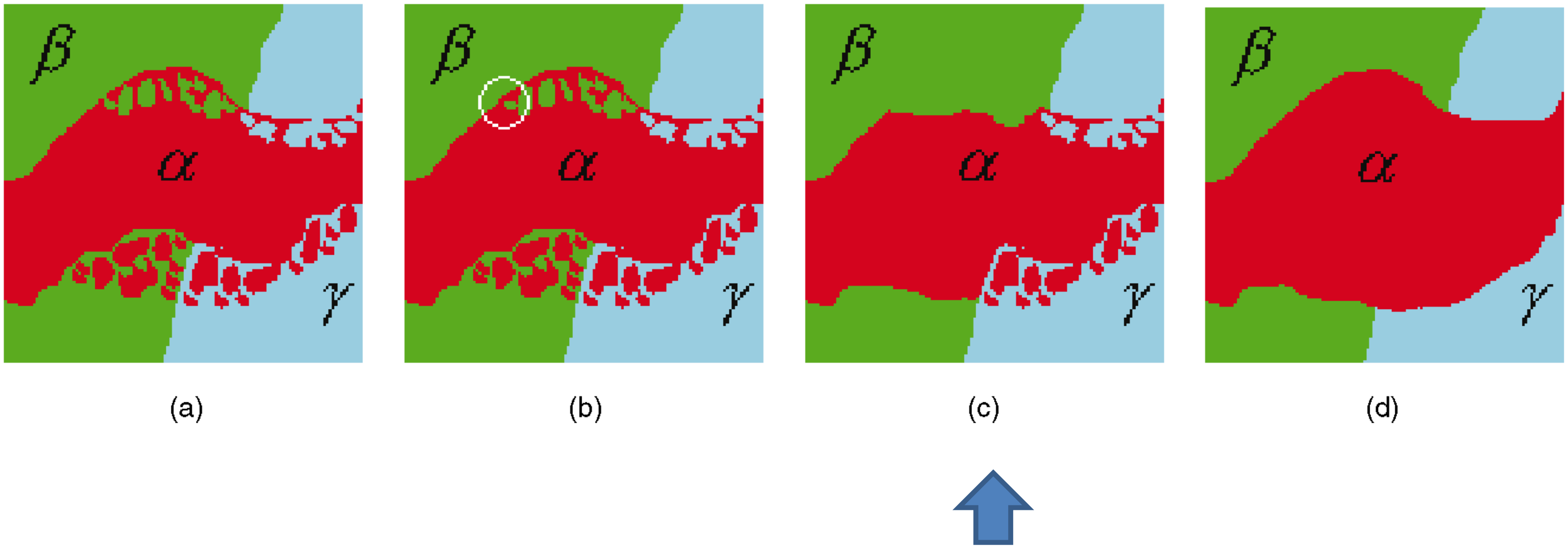
$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

Proof step 2: go through 3 types of pairwise configurations. We need binary V to be semi-metric V(α,α)=0



t_p^α	$V_p(\alpha) + \sum_{q \in \mathcal{N}_p, q \notin \mathcal{P}_{\alpha\beta}} V(\alpha, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
t_p^β	$V_p(\beta) + \sum_{q \in \mathcal{N}_p, q \notin \mathcal{P}_{\alpha\beta}} V(\beta, f_q)$	$p \in \mathcal{P}_{\alpha\beta}$
$e_{\{p,q\}}$	$V(\alpha, \beta)$	$\{p,q\} \in \mathcal{N}$, $p,q \in \mathcal{P}_{\alpha\beta}$

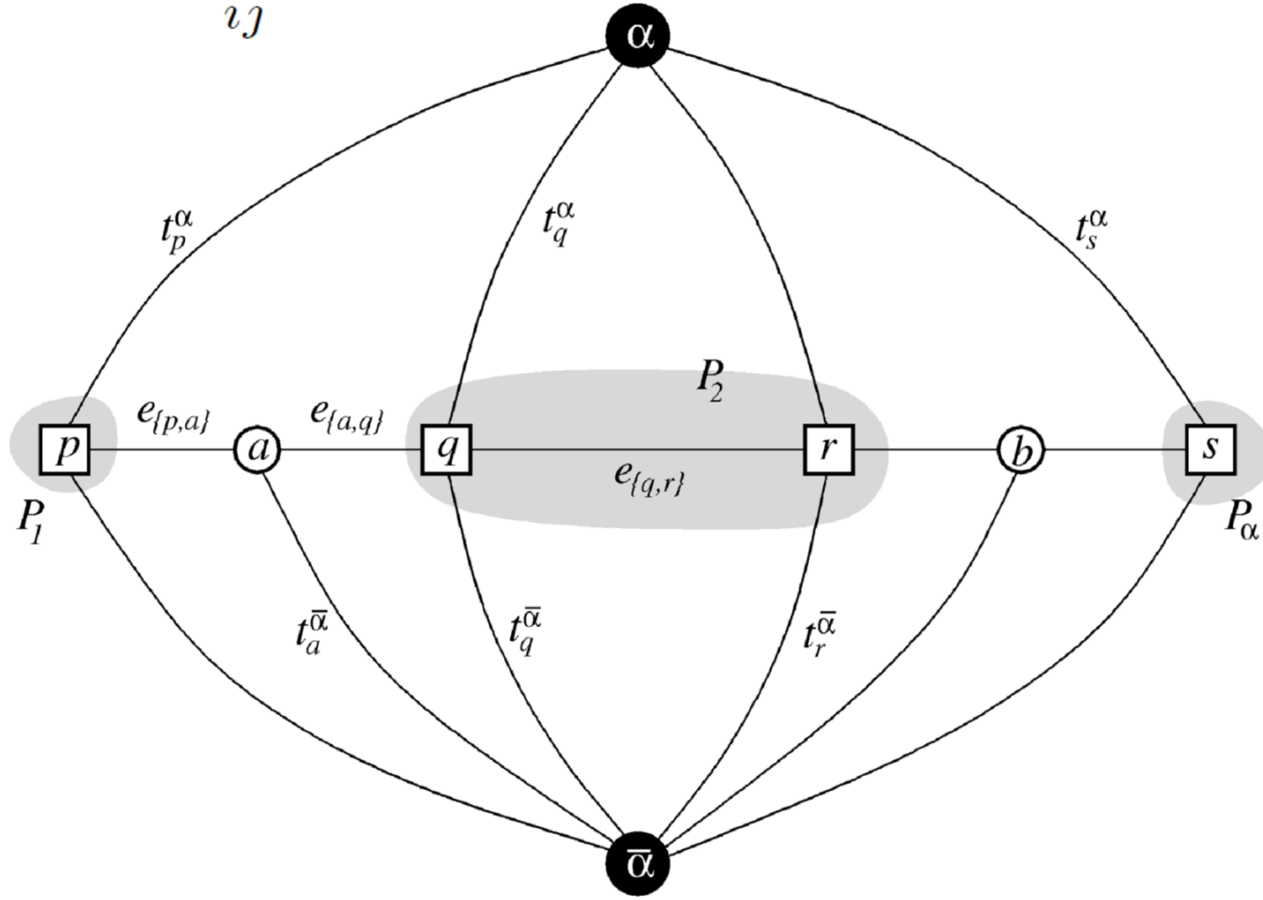
α-β swap - summary



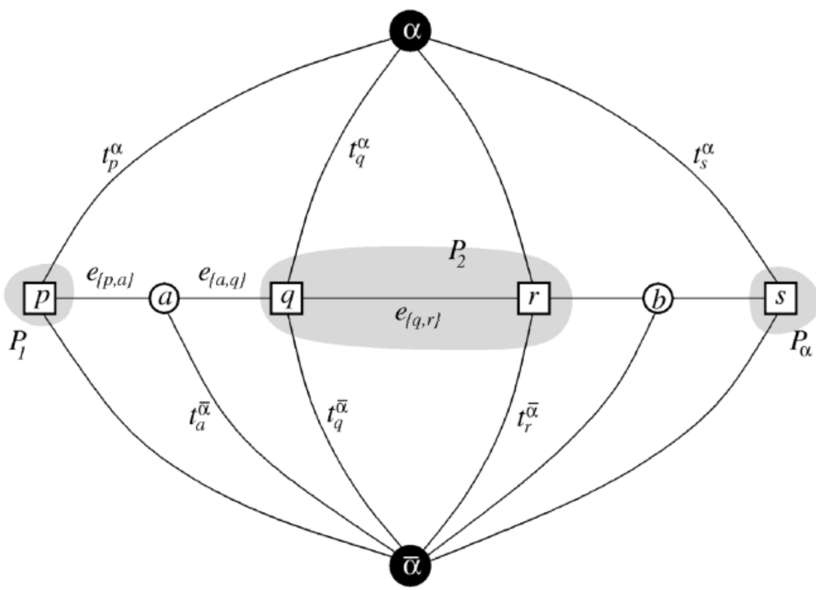
- We know how to transform minimization of E(f) over all possible α-β swap moves to graph cut problem

α-expansion move graph

$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$



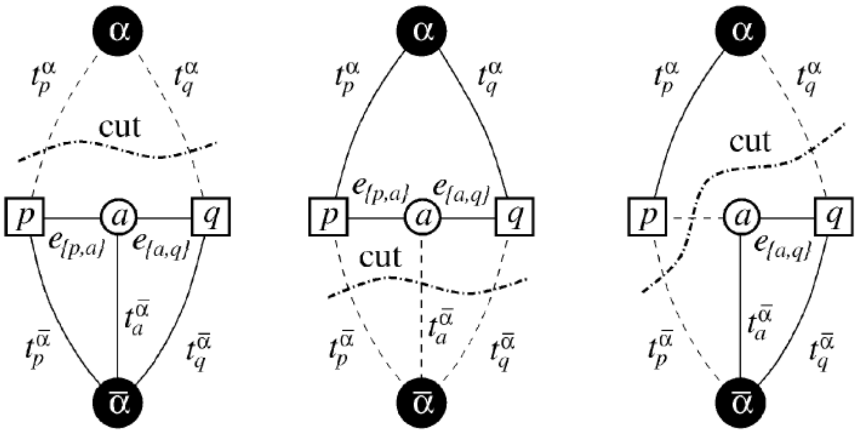
α-expansion move graph

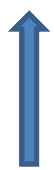


$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

$t_p^{\bar{\alpha}}$	∞	$p \in \mathcal{P}_\alpha$
$t_p^{\bar{\alpha}}$	$V_p(f_p)$	$p \notin \mathcal{P}_\alpha$
t_p^{α}	$V_p(\alpha)$	$p \in \mathcal{P}$
$e_{\{p,a\}}$	$V(f_p, \alpha)$	
$e_{\{a,q\}}$	$V(\alpha, f_q)$	$\{p,q\} \in \mathcal{N},\ f_p \neq f_q$
$t_a^{\bar{\alpha}}$	$V(f_p, f_q)$	
$e_{\{p,q\}}$	$V(f_p, \alpha)$	$\{p,q\} \in \mathcal{N},\ f_p = f_q$

α-expansion graph - cuts



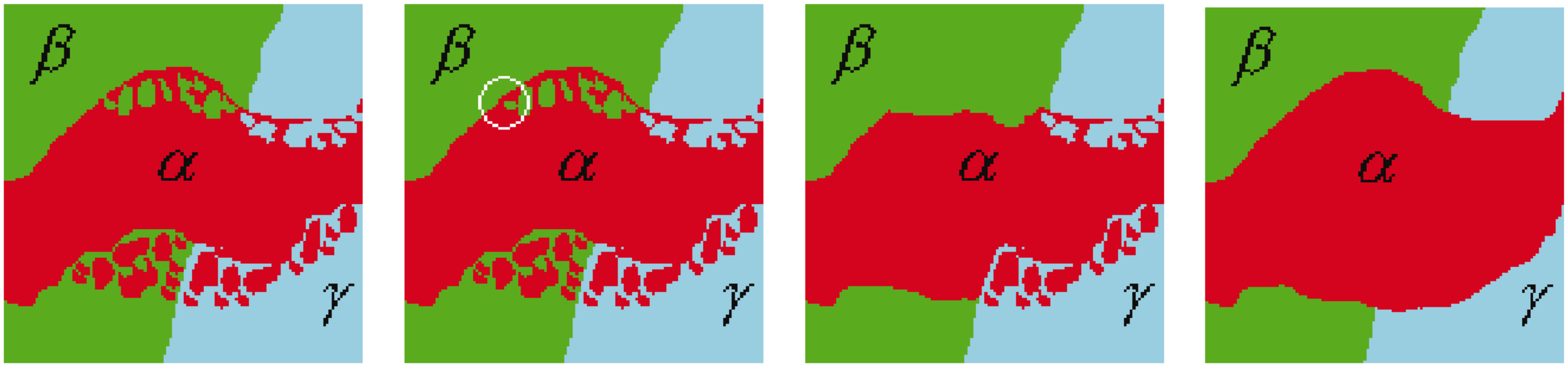


$$E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j)$$

$t_p^{\bar{\alpha}}$	∞	$p \in \mathcal{P}_\alpha$
$t_p^{\bar{\alpha}}$	$V_p(f_p)$	$p \notin \mathcal{P}_\alpha$
t_p^{α}	$V_p(\alpha)$	$p \in \mathcal{P}$
$e_{\{p,a\}}$	$V(f_p, \alpha)$	$\{p,q\} \in \mathcal{N},\ f_p \neq f_q$
$e_{\{a,q\}}$	$V(\alpha, f_q)$	
$t_a^{\bar{\alpha}}$	$V(f_p, f_q)$	
$e_{\{p,q\}}$	$V(f_p, \alpha)$	$\{p,q\} \in \mathcal{N},\ f_p = f_q$

Δ - inequality !

α-expansion - summary



(a) (b) (c) (d)

- We know how to transform minimization of E(f) over all possible α-expansion moves to graph cut problem
- What remains? - how to find the minimum cut

Graph cuts algorithm

- "Augmenting path" type algorithm with simple heuristics
  - Looks for a non-saturated path ~ path in residual graph
  - Simultaneously builds trees from α and β
- Maximum complexity $O(n^2 m C_{max})$, C_{max} cost of the minimum cut
- Actually typically linear with respect to the number of pixels
- On our problems faster than good combinatorial algorithms - Dinic $O(n^2 m)$, Push-relabel $O(n^2 \sqrt{m})$

Graph cuts - summary

- Minimization of E(f) by finding min-cut in a graph in polynomial time

2 label minimization can be done in polynomial (and typically linear) time with respect to the number of pixels

- K>2 labels – NP hard
  - Equivalent to Multiway Cut Problem
  - α-expansion finds a solution $\leq$ 2*optimum
  - In practice both α-β swap and α-expansion algorithms get very close to global minimum

Graph cuts – additional example

- " "GrabCut" — Interactive Foreground Extraction using Iterated Graph Cuts", C. Rother, V. Kolmogorov, A. Blake, SIGGRAPH 2004

Discrete optimization in MRFs - summary

- Conditional independence is strong structural information that can be exploited
- Gives useful approximations for difficult (NP-hard) problems
- For convex problems mostly better to use continuous methods

References

- Graph Cuts
  - “Fast Approximate Energy Minimization via Graph Cuts” - Y. Boykov, O. Veksler, R. Zabih, PAMI 2001 (Augmenting path min-cut algorithm)
  - “An Experimental Comparison of Min-Cut/Max-flow Algorithms for Energy Minimization in Vision” – Y. Boykov, V. Kolmogorov, PAMI 2004 (Graph construction for α-β swap and α-expansion moves)
  - “ “GrabCut” — Interactive Foreground Extraction using Iterated Graph Cuts”, C. Rother, V. Kolmogorov, A. Blake, SIGGRAPH 2004
- Belief propagation
  - “Understanding Belief Propagation and its Generalizations” - J.S. Yedidia, W.T.Freeman, Y.Weiss (Mitsubishi electric research laboratories, Technical report, 2002)

Convex formulation of multi-label problems

- Continuous counterpart of Ishikawa's pairwise MRF problem taking huge memory
- "Arbitrary" non-convex data term

$$\min_u \left(\int_\Omega \rho(u(x), x) dx + \int_\Omega |\nabla u(x)| dx \right)$$

Pock, Schoenemann, Graber, Bischof, Cremers: A **Convex Formulation of Continuous Multi-label Problems** (2008)

Functional lifting

$$\min_u \left(\int_\Omega \rho(u(x), x) dx + \int_\Omega |\nabla u(x)| dx \right) \qquad u : \Omega \to \Gamma,\ \Gamma =< \gamma_{min}, \gamma_{max} >$$

$$\phi(x, \gamma) = \mathbf{1}_{\{u(x) > \gamma\}}(x)$$

Representing u in terms of its level sets

Layer cake formula

$$u(x) = \gamma_{min} + \int_\Gamma \phi(x, \gamma) d\gamma$$

$$\min_{\phi \in D'} \left(\int_\Sigma \rho(x, \gamma) |\partial_\gamma \phi(x, \gamma)| + |\nabla \phi(x, \gamma)| d\Sigma \right)$$

$$D' = \{\phi : \Sigma \to \{0, 1\} \mid \phi(x, \gamma_{min}) = 1,\ \phi(x, \gamma_{max}) = 0\}$$

$$D = \{\phi : \Sigma \to \langle 0, 1 \rangle \mid \phi(x, \gamma_{min}) = 1,\ \phi(x, \gamma_{max}) = 0\}$$

Mathematics in image processing

Many image processing/CV problems can be formulated as optimization problems and solved by variational or discrete algorithms within a common framework

- image restoration (denoising, deblurring, SR, JPEG decompression)
- image segmentation
- optical flow
- stereo