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

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

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)
- Tomography (CT, MRI, PET etc.)
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

Projection on the image plane of the 3D scene velocity

Optical center

Image intensity

I_1

I_2

3D motion field

2D motion field
Optical flow example

Stereo reconstruction

**Principle**

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

**Result (3D model)**

Source: [http://lcav.epfl.ch](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 \( u : \mathbb{R}^2 \to [0, 1] \)
  – Discrete – matrix or vector \( u \in \mathbb{R}^{m \times n}, \ u \in \mathbb{R}^{mn} \)
  – Both can be extended to 3D

• Color image = set of 3 or more greyscale images
  – RGB channels are highly correlated \( \rightarrow \) 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): \( \text{max } p(u|z) \)
Maximum likelihood (MLE): \( \text{max } p(z|u) \)
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

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

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

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

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

\[
\min_u - \log p(u|z) \propto - \log p(z|u) - \log p(u)
\]

\[
\ln p(u) = \ln \prod p(u_i) = \sum_i \ln p(u_i)
\]
Image Prior

\[ p(u) \propto \prod_{i} e^{-\lambda \phi(\nabla u_i)} \]

\[ \ln p(u) = -\lambda \sum_{i} \phi(\nabla u_i) + c \]

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

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

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

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

\[- \ln p(u|z) = - \ln p(z|u) - \ln p(u)\]

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

\[\min_u \frac{1}{2\sigma^2} \sum_i (z_i - u_i)^2 + \lambda \sum_i |\nabla u_i|^p\]
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 Daubechies wavelets)
  – Learned by PCA
Synthesis-based sparsity

Bayesian approach applied on transform coefficients:

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

$$\downarrow$$

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

(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

- $l_p$, $0 < p \leq 1$ norms $\|a\|_p^p$

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

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

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

- $l_1$ 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
$\hat{a} = \arg \min_a \|a\|_2^2$ subject to $Da = x$
\[ \hat{a} = \arg \min_a \|a\|_1 \quad \text{subject to} \quad Da = x \]
Deblurring

• Denoising
  \[ z = u + N(0, \sigma^2 I) \]
  \[
  \min_u \frac{1}{2} \| z - u \|^2 + \lambda \| \nabla u \|_{2,1}
  \]

• Deblurring
  \[ z = h \ast u + N(0, \sigma^2 I) = H u + N(0, \sigma^2 I) \]
  \[
  \min_u \frac{1}{2} \| z - H u \|^2 + \lambda \| \nabla u \|_{2,1}
  \]
Super-resolution (with deblurring)

Several possibly shifted blurred images

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

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

\( \text{D}_i \) ... 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

\[
\min_v \frac{1}{2} \int_{\Omega} (\nabla I \cdot v + I_t)^2 dx + \lambda \sum_{i=1}^{2} \| \nabla v_i \|_{2,1}
\]
JPEG compression

8 x 8 blocks

Source image data

FDCT

Determinant encoder

DCT-based encoder

Quantizer

Entropy encoder

Table specifications

Table specifications

Compressed image data

DCT-based decoder

Entropy decoder

Dequantizer

IDCT

Table specifications

Table specifications

Reconstructed image data

C

TQ

C^{-1}

Q^{-1}
Bayesian MAP restoration

MAP – maximum a posteriori probability

\[
\min_u - \log p(z|u) - \log p(u)
\]

\[- \log p(u) = \tau \|W u\|_1\]

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

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)

\[
\min_{u} \| \nabla u \|_{2,1}, \quad s.t. \quad QCu \in (QCz - 0.5, QCz + 0.5)
\]

(Bredies and Holler, 2012)

Or using redundant wavelets

\[
\min_{u} \| Wu \|_{2,1}, \quad s.t. \quad QCu \in (QCz - 0.5, QCz + 0.5)
\]

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

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Segmentation</strong></td>
<td><strong>foreground/background or object number</strong></td>
<td><strong>{0,1}</strong></td>
</tr>
<tr>
<td>Stereo</td>
<td>disparity (inverse depth)</td>
<td>-k..k</td>
</tr>
<tr>
<td>Optical flow</td>
<td>local motion vector</td>
<td>(-k..k) x (-k..k)</td>
</tr>
<tr>
<td>Restoration</td>
<td>intensity</td>
<td>0..255</td>
</tr>
</tbody>
</table>
Segmentation by graph cuts
Graph cuts & Belief propagation

Graph cuts

„Classical local algorithms“

Belief propagation
Markov Random Fields (MRFs)

- Markov Random Field, Gibbs Random Field
  - MRF ↔ 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 $S = \{1, \ldots, m\}$
- $F$ ... set of random variables defined on $S$
- $N$ ... neighborhood system
- $f_i \in \mathcal{L}$ ... (possibly discrete) label
- configuration $f = \{f_1 \ldots f_K\}$,

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

$P(f) > 0$

- Other possible properties – homogeneity, isotropy
Gibbs Random Field

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

Partition function

\[ Z = \sum_{f} e^{-\frac{1}{T} U(f)} \]

Energy function \( U(f) \)

\[ U(f) = \sum_{c \in C} V_c(f) = \sum_{i \in S} V_1(f_i) + \sum_{i \in S} \sum_{i' \in N_i} V_2(f_i, f_{i'}) \]

\( V_c(f) \) ... clique potentials
Hammersley-Clifford theorem

\[ \text{MRF} = \text{GRF} \]

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

\[
P(\mathbf{f}_i | \mathbf{f}_{S-\{i\}}) = P(\mathbf{f}_i | \mathbf{f}_{N_i})
\]

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

\[
P(\mathbf{f}_i | \mathbf{f}_{S-\{i\}}) = \frac{e^{-\sum_{\{c, i \in c\}} V_c(f)}}{\sum_{f'_i} e^{-\sum_{\{c; i \in c\}} V_c(f)}}
\]
MRF = GRF

- MAP-MRF

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

\[
\min_f (-\ln p(f)) = \min_f E(f) + \text{const}
\]

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

Priors on derivatives, usually first derivative

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

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

segmentation, sometimes in stereo

Tikhonov regularization

Discontinuity preserving penalties

TV regularization

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

line process, Mumford-Shah functional

\[ V(f_i, f_j) = \kappa_{ij} \min((f_i - f_j)^2, \text{const}) \]
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:

$$V_1(f_i, d^1, d^2) = \min(\min_{\Delta \in (f_i-\frac{1}{2}, f_i+\frac{1}{2})} |d^1_i - d^2_{i+\Delta}|, \ldots, const)^2$$
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

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

\[ V_1(f_i, d_i) \approx -\ln p(f_i|d_i) \approx -\ln p(d_i|f_i) - \ln p(f_i) \]

\[ V_1(f_i, d_i) \sim \text{probability to be in fg/bg based on a feature space (intensities, texture features etc...) \text{–} modeled for example as a mixture of Gaussians} \]
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)

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

- 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
  \[
  \max_f p(f) = \frac{1}{Z} e^{-E(f)}
  \]
  \[
  \min_f (-\ln p(f)) = \min_f E(f) + \text{const}
  \]

- 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

\[ E(f) = \sum_{i} V_1(f_i) + \sum_{i,j} V_2(f_i, f_j) \]

For \( V_2 \geq 0 \) metric
- \( V_2(a,b) = 0 \iff a = b \)
- \( V_2(a,b) = V_2(b,a) \) (actually not necessary)
- \( V_2(a,b) \leq V_2(a,c) + V_2(c,b) \)

or semimetric (without \( \Delta \)-inequality)

Metric:
\[
\delta(f_i - f_j) = \min(|f_i - f_j|, \text{const}) \quad \text{for any norm } |.|
\]

Semimetric:
\[
\min((f_i - f_j)^2, \text{const})
\]
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

(a) initial labeling
(b) single pixel ICM move
(c) α-β swap move
(d) α-expansion move
\(\alpha\)-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 \(\alpha\)-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 \(V_2(\alpha, \beta)\) (\(\Delta\)-inequality)
- Not worse than 2x optimum
\( \alpha-\beta \) swap algorithm

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

• Arbitrary \textbf{semimetric} \( V_2(\alpha, \beta) \)
  (without \( \Delta \)-inequality)

• No optimality guaranteed
\[ E(f) = \sum_i V_1(f_i) + \sum_{i,j} V_2(f_i, f_j) \]
### α-β Swap Move Graph

The energy function $E(f)$ is given by:

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

**Table:**

<table>
<thead>
<tr>
<th>Term</th>
<th>Expression</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^\alpha_p$</td>
<td>$V_p(\alpha) + \sum_{q \in \mathcal{N}<em>p \cap \mathcal{P}</em>{\alpha\beta}} V(\alpha, f_q)$</td>
<td>$p \in \mathcal{P}_{\alpha\beta}$</td>
</tr>
<tr>
<td>$t^\beta_p$</td>
<td>$V_p(\beta) + \sum_{q \in \mathcal{N}<em>p \cap \mathcal{P}</em>{\alpha\beta}} V(\beta, f_q)$</td>
<td>$p \in \mathcal{P}_{\alpha\beta}$</td>
</tr>
<tr>
<td>$e_{{p,q}}$</td>
<td>$V(\alpha, \beta)$</td>
<td>${p,q} \in \mathcal{N} \cap \mathcal{P}_{\alpha\beta}$</td>
</tr>
</tbody>
</table>
α-β swap move graph

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

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

<table>
<thead>
<tr>
<th>$t^α_p$</th>
<th>$V_p(α) + \sum_{q ∈ N_p \setminus P_{αβ}} V(α, f_q)$</th>
<th>$p ∈ P_{αβ}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^β_p$</td>
<td>$V_p(β) + \sum_{q ∈ N_p \setminus P_{αβ}} V(β, f_q)$</td>
<td>$p ∈ P_{αβ}$</td>
</tr>
<tr>
<td>$e{p,q}$</td>
<td>$V(α, β)$</td>
<td>${p,q} ∈ N$ \hspace{1em} $p, q ∈ P_{αβ}$</td>
</tr>
</tbody>
</table>
Proof step 2: go through 3 types of pairwise configurations. We need binary V to be semi-metric V(α,α)=0
• We know how to transform minimization of $E(f)$ over all possible $\alpha$-$\beta$ swap moves to graph cut problem.
\[ E(f) = \sum_i V_1(f_i) + \sum_{ij} V_2(f_i, f_j) \]
\( \alpha \)-expansion move graph

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

\[
\begin{array}{|c|c|c|}
\hline
\tau^\alpha_p & \infty & p \in \mathcal{P}_\alpha \\
\tau^\alpha_p & V_p(f_p) & p \notin \mathcal{P}_\alpha \\
\tau^\alpha_p & V_p(\alpha) & p \in \mathcal{P} \\
e_{\{p,a\}} & V(f_p, \alpha) & \\
e_{\{a,q\}} & V(\alpha, f_q) & \{p, q\} \in \mathcal{N}, f_p \neq f_q \\
\tau^\alpha_a & V(f_p, f_q) & \\
e_{\{p,q\}} & V(f_p, \alpha) & \{p, q\} \in \mathcal{N}, f_p = f_q \\
\hline
\end{array}
\]
$\alpha$-expansion graph - cuts

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

<table>
<thead>
<tr>
<th>$t_{p}^{\alpha}$</th>
<th>$\infty$</th>
<th>$p \in P_{\alpha}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_{p}^{\bar{\alpha}}$</td>
<td>$V_p(f_p)$</td>
<td>$p \notin P_{\alpha}$</td>
</tr>
<tr>
<td>$t_{p}^{\alpha}$</td>
<td>$V_p(\alpha)$</td>
<td>$p \in P$</td>
</tr>
<tr>
<td>$e_{{p,a}}$</td>
<td>$V(f_p, \alpha)$</td>
<td></td>
</tr>
<tr>
<td>$e_{{a,q}}$</td>
<td>$V(\alpha, f_q)$</td>
<td>${p,q} \in \mathcal{N}, ; f_p \neq f_q$</td>
</tr>
<tr>
<td>$t_{a}^{\bar{\alpha}}$</td>
<td>$V(f_p, f_q)$</td>
<td></td>
</tr>
<tr>
<td>$e_{{p,q}}$</td>
<td>$V(f_p, \alpha)$</td>
<td>${p,q} \in \mathcal{N}, ; f_p = f_q$</td>
</tr>
</tbody>
</table>

$\Delta$ - inequality!
• We know how to transform minimization of $E(f)$ over all possible $\alpha$-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 $\alpha$ and $\beta$
• Maximum complexity $O(n^2 m C_{\text{max}})$, $C_{\text{max}}$ 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 $O(n^2 \sqrt{V} m)$
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
  – $\alpha$-expansion finds a solution $\leq 2\cdot \text{optimum}$
  – In practice both $\alpha$-$\beta$ swap and $\alpha$-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
Convex formulation of multi-label problems

• Continuous counterpart of Ishikawa’s pairwise MRF problem taking huge memory

• “Arbitrary” non-convex data term

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

Functional lifting

\[
\min_u \left( \int_{\Omega} \rho(u(x), x) \, dx + \int_{\Omega} | \nabla u(x)| \, dx \right) \quad u : \Omega \to \Gamma, \ \Gamma = [\gamma_{\min}, \gamma_{\max}]
\]

\[
\phi(x, \gamma) = 1_{\{u(x) > \gamma\}}(x)
\]

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