Bayesian image modeling and phase transition in generalized sparse Markov random fields and loopy belief propagation

Kazuyuki Tanaka
GSIS, Tohoku University, Sendai, Japan
http://www.smapip.is.tohoku.ac.jp/~kazu/

Collaborators
Muneki Yasuda (GSIS, Tohoku University, Japan)
Sun Kataoka (GSIS, Tohoku University, Japan)
D. M. Titterington (Department of Statistics, University of Glasgow, UK)
Outline

1. Pairwise Markov Random Fields for Natural Images
2. Supervised Learning of Pairwise Markov Random Fields by Loopy Belief Propagation
3. Bayesian Image Modeling by Generalized Sparse Prior
4. Noise Reductions by Generalized Sparse Prior
5. Phase Transitions of Generalized Sparse Prior
6. Concluding Remarks
Probabilistic Model and Belief Propagation

Bayes Formulas

Probabilistic Models

Bayesian Networks
Markov Random Fields

Belief Propagation
=Bethe Approximation

\[ M_{i \leftarrow j}(f_i) = \text{Constant} \times \sum_{f_j} U(f_i, f_j) \prod_{k \in \partial j} M_{j \leftarrow k}(f_j) \]

\[ \partial j \equiv \{k | \{j,k\} \in E\} \]

\[ V: \text{Set of all the nodes (vertices) in graph } G \]
\[ E: \text{Set of all the links (edges) in graph } G \]

\[ P(f) \propto \prod_{\{i,j\} \in E} U(f_i, f_j) \]

G = (V, E)
In Bayesian image modeling, natural images are often assumed to be generated by according to the following pairwise Markov random fields:

\[
\Pr\{F = f\} = P(f) \propto \prod_{\{i, j\} \in E} U(|f_i - f_j|) = \exp \left( - \sum_{\{i, j\} \in E} \ln(U(|f_i - f_j|)) \right)
\]

Hamiltonian of Classical Spin System

\[
f_i \in \{0, 1, \cdots, q - 1\}
\]

\[
V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \text{ Set of All the pixels}
\]

\[
E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{5, 6\}, \{6, 7\}, \{7, 8\}, \{9, 10\}, \{10, 11\}, \{11, 12\}, \{1, 5\}, \{2, 6\}, \{3, 7\}, \{4, 8\}, \{5, 9\}, \{6, 10\}, \{7, 11\}, \{8, 12\}\}
\]

Set of All the Nearest Neighbour Pairs of Pixels
Conventional Pairwise Markov Random Fields

$$\Pr\{F = f\} = P(f) \propto \prod_{\{i,j\} \in E} U(\|f_i - f_j\|) = \exp\left( - \sum_{\{i,j\} \in E} \left( - \ln(U(\|f_i - f_j\|)) \right) \right)$$

**Gaussian Prior**

$$- \ln(U(x)) \propto x^2$$

**Huber Prior**

$$- \ln(U(x)) \propto \begin{cases} |x|^2 & (x < c) \\ 2c|x| - c^2 & (x > c) \end{cases}$$

**Saturated Quadratic Prior**

$$- \ln(U(x)) \propto \frac{x^2}{x^2 + c^2}$$

Convex Functions near \(x=0\)
Supervised Learning of Pairwise Markov Random Fields

Supervised Learning Scheme by Loopy Belief Propagation in Pairwise Markov random fields:

\[ P(f) \propto \exp \left( - \sum_{\{i,j\} \in E} \left( - \ln U(f_i - f_j) \right) \right) \]

30 Standard Images

Supervised Learning
LBP in Inverse Problem for Ising Model

\[ P(s) \propto \exp \left( h \sum_{i \in V} s_i + J \sum_{\{i, j\} \in E} s_i s_j \right) \quad (V, E): \text{Square Grid Graph} \]

LBP for Ising Model on Square Grid Graph

\[ h = -3 \text{arctanh} m + 2 \sum_{s=\pm 1} \text{arctanh} \left( \frac{m + cs}{1 + ms} \right) \]

\[ J = \frac{1}{4} \sum_{s=\pm 1} \sum_{s' = \pm 1} s s' \ln (1 + ms + ms' + css') \]

Sample Average from the given \(|V|\)-dimentional data \(s\)

\[
\begin{align*}
 m & \leftarrow \frac{1}{|V|} \sum_{i \in V} s_i \\
 c & \leftarrow \frac{1}{|E|} \sum_{\{i, j\} \in E} s_i s_j \\
\end{align*}
\]

\[ \Rightarrow \begin{pmatrix} h \\ J \end{pmatrix} \]

Inverse Problem

September, 2013
University of Roma La Sapienza, Italy
Supervised Learning Scheme by Loopy Belief Propagation in Pairwise Markov random fields:

\[ P(f) \propto \prod_{\{i,j\} \in E} U(|f_i - f_j|), \quad \ln U(|f_i - f_j|) = \frac{1}{2} \sum_{m=0}^{q-1} \sum_{n=0}^{q-1} K_{m,n} \Phi_m(f_i) \Phi_n(f_j) \]

\[ -3 \sum_{f_i=0}^{q-1} \Phi_m(f_i) \ln P(f_i) \]

\[ + 4 \sum_{f_i=0}^{q-1} \sum_{f_j=0}^{q-1} \Phi_m(f_i) \Phi_0(f_j) P(f_i, f_j) \]

\[ K_{m,0} = K_{0,m} = -3 \sum_{f_i=0}^{q-1} \Phi_m(f_i) \ln P(f_i) \]

\[ K_{m,n} = K_{n,m} = \sum_{f_i=0}^{q-1} \sum_{f_j=0}^{q-1} \Phi_m(f_i) \Phi_0(f_j) P(f_i, f_j) \]


Histogram from Supervised Data
Supervised Learning of Pairwise Markov Random Fields by Loopy Belief Propagation

Supervised Learning Scheme by Loopy Belief Propagation in Pairwise Markov random fields:

\[ P(f) \propto \exp \left( - \sum_{\{i,j\} \in E} \ln U(|f_i - f_j|) \right) \sim \exp \left( -2.35 \times \sum_{\{i,j\} \in E} |f_i - f_j|^{0.32} \right) \]

\[ q = 256 \]

\[ f_i \in \{0,1,2,\cdots,255\} \]
Supervised Learning of Pairwise Markov Random Fields by Loopy Belief Propagation

Supervised Learning Scheme by Loopy Belief Propagation in Pairwise Markov random fields:

\[ P(f) \propto \exp\left( - \sum_{\{i, j\} \in E} \left( - \ln U \left( |f_i - f_j| \right) \right) \right) \sim \exp\left( - 4.39 \times \sum_{\{i, j\} \in E} |f_i - f_j|^{0.41} \right) \]

\[ f_i \in \{0,1,2,\ldots,15\} \]
Bayesian Image Modeling by Generalized Sparse Prior

Assumption: Prior Probability is given as the following Gibbs distribution with the interaction $\alpha$ between every nearest neighbour pair of pixels:

$$
\Pr\{ F = f \mid p, \alpha \} = P(f\mid p, \alpha) = \frac{1}{Z(p, \alpha)} \exp\left( -\frac{1}{2} \alpha \sum_{i,j \in E} |f_i - f_j|^p \right)
$$

$i$-th pixel

$\{1,2\}, \{3,4\}, \{5,6\}, \{7,8\}, \{9,10\}, \{11,12\}$

State Variable of Light Intensity at $i$-pixel

$f_i$ \in \{0,1,\cdots,q-1\}$

$p=0$: $q$-state Potts model

$p=2$: Discrete Gaussian Graphical Model
Prior in Bayesian Image Modeling

\[
\Pr\{F = f \mid p, \alpha\} = P(f \mid p, \alpha) = \frac{1}{Z(p, \alpha)} \exp\left( -\frac{1}{2} \alpha \sum_{\{i,j\} \in E} |f_i - f_j|^p \right)
\]

\[
\alpha^* = \arg \max_{\alpha} \ln \Pr\{F = f^* \mid p, \alpha\}
\]

\[
u(p, \alpha) \equiv \frac{1}{|E|} \sum_{\{i,j\} \in E} \sum_f |f_i - f_j|^p P(f \mid p, \alpha)
\]

\[
u^* = \frac{1}{|E|} \sum_{\{i,j\} \in E} |f_i^* - f_j^*|^p
\]

In the region of \(0 < p < 0.3504\ldots\), the first order phase transition appears and the solution \(\alpha^*\) does not exist.
Bayesian Image Modeling by Generalized Sparse Prior: Conditional Maximization of Entropy

\[ \Pr\{F = f \mid p, u\} = \arg \max_{P(f)} \left\{ -\sum_{z} P(z) \ln P(z) \right\} \left\{ \sum_{z} \sum_{\{i, j\} \in E} |z_i - z_j|^p P(z) = u \mid E \right\} \]

\[ \Pr\{\tilde{F} = \tilde{f} \mid p, u\} = P\left(f \mid p, \alpha(p, u)\right) = \frac{1}{Z(p, \alpha(p, u))} \exp\left\{ -\frac{1}{2} \alpha(p, u) \sum_{\{i, j\} \in E} |f_i - f_j|^p \right\} \]

Prior Analysis by LBP and Conditional Maximization of Entropy in Bayesian Image Modeling

Prior Probability

\[
\Pr\{F = f \mid p, u\} = P(f \mid p, \alpha(p, u)) = \frac{1}{Z(p, \alpha(p, u))} \exp \left( -\frac{1}{2} \alpha(p, u) \sum_{\{i, j\} \in E} |f_i - f_j|^p \right)
\]

Repeat until \( \alpha(p, u) = A \) converges

\[
M_{j \rightarrow i} \leftarrow \sum_{f_j} \left( \prod_{k \in \delta \setminus i} M_{k \rightarrow j} \right) \exp \left( -\frac{1}{2} A |f_i - f_j|^p \right) \quad (\forall \{i, j\} \in E)
\]

Compute marginals \( P_{(i,j)}(f_i, f_j \mid p, C) \) in LBP by messages \( M \)

\[
\alpha(p, u) = A \leftarrow A \times \frac{1}{u|E|} \sum_{(i,j) \in E} \sum_f |f_i - f_j|^p P_{(i,j)}(f_i, f_j \mid p, A)
\]

September, 2013
University of Roma La Sapienza, Italy
Prior Analysis by LBP

\[ u(p, \alpha(p,u^*)) = u^* \]

\[ u^* = \frac{1}{|E|} \sum_{\{i,j\} \in E} |f_i^* - f_j^*|^p \]

\[ \alpha^* = \arg \max_{\alpha} \ln \Pr\{F = f^* \mid p, \alpha\} \]

\[ = -\frac{1}{2} \alpha \sum_{\{i,j\} \in E} |f_i^* - f_j^*|^p - \ln Z(p, \alpha) \]

\( q = 16 \)
\( p = 0.2 \)
Prior Analysis by LBP

\[ u(p, \alpha(p, u^*)) = u^* \]

\[ u^* = \frac{1}{|E|} \sum_{\{i, j\} \in E} |f^*_i - f^*_j|^p \]

Critical Point by Linear Response in LBP

\[ \alpha^* = \arg \max_{\alpha} \ln \Pr\{F = f^* | p, \alpha\} \]

\[ \ln \Pr\{F = f^* | p, \alpha\} = -\frac{1}{2} \alpha \sum_{\{i, j\} \in E} |f^*_i - f^*_j|^p - \ln Z(p, \alpha) \]
Prior Analysis by LBP

\[ u(p, \alpha(p, u^*)) = u^* \]

\[ u^* = \frac{1}{|E|} \sum_{(i,j) \in E} |f_i^* - f_j^*|^p \]

Critical Point by Linear Response in LBP

\[ \alpha^* = \arg \max_{\alpha} \ln \Pr\{F = f^* | p, \alpha\} \]

\[
\ln \Pr\{F = f^* | p, \alpha\} = -\frac{1}{2} \alpha \sum_{(i,j) \in E} |f_i^* - f_j^*|^p - \ln Z(p, \alpha)
\]
**Prior Analysis by LBP**

\[ u(p, \alpha(p, u^*)) = u^* \]

\[ u^* = \frac{1}{|E|} \sum_{(i,j) \in E} |f_i^* - f_j^*|^p \]

**Critical Point by Linear Response in LBP**

\[ \alpha^* = \arg \max_{\alpha} \ln \Pr\{F = f^* \mid p, \alpha\} \]

\[ \ln \Pr\{F = f^* \mid p, \alpha\} = -\frac{1}{2} \alpha \sum_{(i,j) \in E} |f_i^* - f_j^*|^p - \ln Z(p, \alpha) \]

\( q=16 \)
\( p=0 \)
Prior Analysis by LBP and Conditional Maximization of Entropy in Generalized Sparse Prior

Log-Likelihood for $p$ when the original image $f^*$ is given

$$ \ln \Pr \{ F = f^* \mid p, u^* \} = -\frac{1}{2} \alpha(p, u^*) \sum_{\{i, j\} \in E} |f_i^* - f_j^*|^p $$

Free Energy of Prior

$$ -\ln Z(p, \alpha(p, u^*)) $$

$$ p^* = \arg \max_p \ln \Pr \{ F = f^* \mid p, u^* \} $$

$$ u^* = \frac{1}{|E|} \sum_{\{i, j\} \in E} |f_i^* - f_j^*|^p $$


$q = 16$
Degradation Process in Bayesian Image Modeling

**Assumption:** Degraded image is generated from the original image by Additive White Gaussian Noise.

\[
\begin{align*}
\Pr\{\text{Original Image} \mid \text{Degraded Image}\} & \propto \Pr\{\text{Degraded Image} \mid \text{Original Image}\} \\
& \times \Pr\{\text{Original Image}\}
\end{align*}
\]

\[
\Pr\{G = g \mid F = f, \sigma\} \propto \prod_{i \in V} \exp\left(-\frac{1}{2\sigma^2} (g_i - f_i)^2\right)
\]

\(\sigma > 0\)

- : Pixel of Original Image
  \(f = (f_1, f_2, \ldots, f_{12})^T\)

- : Pixel of Degraded Image
  \(g = (g_1, g_2, \ldots, g_{12})^T\)

September, 2013

University of Roma La Sapienza, Italy
Posterior Probability and Conditional Maximization of Entropy in Bayesian Image Modeling

Posterior Probability

\[
\Pr\{F = f \mid G = g, p, \sigma, u\} = \arg \max_{P(f)} \left\{ - \sum_{z} P(z) \ln P(z) \right\}
\]

\[
\sum_{z} \sum_{\{i, j\} \in E} |z_i - z_j|^p P(z) = u \mid E
\]

\[
\sum_{z} \sum_{i \in V} (z_i - g_i)^2 P(z) = \sigma^2 \mid V
\]

\[
P(f \mid g, p, \sigma, \alpha) \propto \exp \left( - \frac{1}{2\sigma^2} \sum_{i \in V} (f_i - g_i)^2 - \frac{1}{2\alpha} \sum_{\{i, j\} \in E} |f_i - f_j|^p \right)
\]

Lagrange Multipliers

\[
\alpha \propto \Pr\{\text{Degraded Image} \mid \text{Original Image}\} \times \Pr\{\text{Original Image}\}
\]

\[
\Pr\{\text{Original Image} \mid \text{Degraded Image}\}
\]
Noise Reductions by Generalized Sparse Prior

Deterministic Equations of $\alpha$ and $\sigma$

\[
\frac{1}{|E|} \sum_{\{i, j\} \in E} \sum_{z_i} \sum_{z_j} |z_i - z_j|^p P_{ij}(z_i, z_j | g, p, \sigma, \alpha) = \frac{1}{|E|} \sum_{\{i, j\} \in E} \sum_{z_i} \sum_{z_j} |z_i - z_j|^p P_{ij}(z_i, z_j | p, \alpha)
\]

\[
\frac{1}{|V|} \sum_{i \in V} \sum_{z_i} (z_i - g_i)^2 P_i(z_i | g, p, \sigma, \alpha) = \sigma^2
\]

Repeat until $\alpha(t)$ and $\sigma(t)$ converge

[Step 1] Solve

\[
\frac{1}{|E|} \sum_{\{i, j\} \in E} \sum_{z_i} \sum_{z_j} |z_i - z_j|^p P_{ij}(z_i, z_j | p, \alpha(t)) = u(t) \quad \text{w.r.t.} \alpha(t)
\]

[Step 2] Compute $u(t + 1) \leftarrow \frac{1}{|E|} \sum_{\{i, j\} \in E} \sum_{z_i} \sum_{z_j} |z_i - z_j|^p P_{ij}(z_i, z_j | g, p, \sigma(t), \alpha(t))$

and $\sigma(t + 1) \leftarrow \sqrt{\frac{1}{|V|} \sum_{i \in V} \sum_{z_i} (z_i - g_i)^2 P_i(z_i | g, p, \sigma(t), \alpha(t))}$

[Step 3] $\hat{f}_i(g, p) \leftarrow \arg \max_{z_i} P_i(z_i | g, p, \sigma(t), \alpha(t)) (\forall i \in V)$
Noise Reductions by Generalized Sparse Prior

Repeat until $\alpha(t)$ converge

$$\alpha(t) \leftarrow \alpha(t) \times \sqrt{\frac{1}{u(t)} \times \frac{1}{|E|} \sum_{\{i,j\} \in E} |z_i - z_j|^p P_{ij}(z_i, z_j | p, \alpha(t))}$$

Repeat until $\alpha(t)$ and $\sigma(t)$ converge

[Step 1] Solve $\frac{1}{|E|} \sum_{\{i,j\} \in E} |z_i - z_j|^p P_{ij}(z_i, z_j | p, \alpha(t)) = u(t)$ w.r.t. $\alpha(t)$

[Step 2] Compute $u(t+1) \leftarrow \frac{1}{|E|} \sum_{\{i,j\} \in E} |z_i - z_j|^p P_{ij}(z_i, z_j | g, p, \sigma(t), \alpha(t))$

and $\sigma(t+1) \leftarrow \sqrt{\frac{1}{|V|} \sum_{i \in V} (z_i - g_i)^2 P_i(z_i | g, p, \sigma(t), \alpha(t))}$

[Step 3] $\hat{f}_i(g, p) \leftarrow \arg \max_{z_i} P_i(z_i | g, p, \sigma(t), \alpha(t))(\forall i \in V)$

September, 2013
University of Roma La Sapienza, Italy
Noise Reduction Procedures Based on LBP
and Conditional Maximization of Entropy

Input \( p \) and data \( g \)

Repeat until \( u(t), \sigma(t) \)

Marginals and Free Energy in LBP

\[
\ln \Pr\{G = g \mid p, \hat{\sigma}, \hat{u}\} = \ln Z(g, p, \hat{\sigma}^2, \alpha(p, \hat{u})) - \ln Z(p, \alpha(p, \hat{u})) - \frac{1}{2} |V| \ln(2\pi \hat{\sigma})
\]

\[
\hat{p} = \arg \max_p \Pr\{G = g \mid p, \hat{\sigma}, \hat{u}\}
\]

Repeat until \( \alpha(t) \) and Messages converges

\[
\alpha(t) \leftarrow \alpha(t) \times \sqrt{\frac{1}{u(t)} \times \frac{1}{|E|} \sum_{\{i,j\} \in E} \sum_{f_i} \sum_{f_j} |f_i - f_j|^p P_{ij}(f_i, f_j \mid p, \alpha(t))}
\]

Compute marginals \( P_{ij}(f_i, f_j \mid g, p, \sigma(t), \alpha(t)) \) and \( P_i(f_i \mid g, p, \sigma(t), \alpha(t)) \) by messages

\[
u(t + 1) \leftarrow \frac{1}{|E|} \sum_{\{i,j\} \in E} \sum_{f_i} \sum_{f_j} |f_i - f_j|^p P_{ij}(f_i, f_j \mid g, p, \sigma(t), \alpha(t))
\]

\[
\sigma(t) \leftarrow \frac{1}{\sqrt{|V|}} \sum_{f_i} (f_i - g_i)^2 P_i(f_i \mid g, p, \sigma(t), \alpha(t))
\]

\[
\hat{\sigma} \leftarrow \sigma(t), \hat{u} \leftarrow u(t), \alpha(p, \hat{u}) \leftarrow \alpha(t)
\]

\[
\hat{f}_i(g, p) \leftarrow \arg \max P_i(f_i \mid g, p, \hat{\sigma}, \hat{\alpha})
\]
Noise Reductions by Generalized Sparse Priors and Loopy Belief Propagation

Original Image

Degraded Image

Restored Image

Noise Reductions by Generalized Sparse Priors and Loopy Belief Propagation

Original Image  Degraded Image  Restored Image

$p=0.2$  $p=0.5$  $p=1$


September, 2013

University of Roma La Sapienza, Italy
Phase Transition in Sparse Prior for even number $q$

$q=16$

$p=0.2$

$$P(f|p,\alpha) \propto \exp\left(-\frac{1}{2}\alpha \sum_{\{i,j\} \in E} |f_i - f_j|^p\right)$$
Phase Transition in Sparse Prior for odd number $q$

$q=15$
$p=0.2$

$$P(f|p, \alpha) \propto \exp \left( -\frac{1}{2} \alpha \sum_{\{i, j\} \in E} |f_i - f_j|^p \right)$$
Summary

- Formulation of Bayesian image modeling for image processing by means of generalized sparse priors and loopy belief propagation are proposed.
- Our formulation is based on the conditional maximization of entropy with some constraints.
- In our sparse priors, although the first order phase transitions often appear, our algorithm works well also in such cases.
- In the cases of even and odd $q$, our generalized sparse priors are expected to show different kinds of phase transitions of each other.
References


