Probabilistic image processing and Bayesian network

Kazuyuki Tanaka

Graduate School of Information Sciences,
Tohoku University
kazu@smapip.is.tohoku.ac.jp
http://www.smapip.is.tohoku.ac.jp/~kazu/

References

- **K. Tanaka: Statistical-mechanical approach to image processing (Topical Review), J. Phys. A, vol.35, pp.R81-R150 (2002).**
- **K. Tanaka, H. Shouno, M. Okada and D. M. Titterington: Accuracy of the Bethe approximation for hyperparameter estimation in probabilistic image processing, J. Phys. A, vol.37, pp.8675-8695 (2004).**

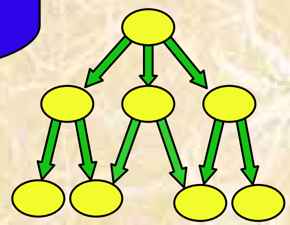
Bayesian Network and Belief Propagation

Bayesian Network

Bayes Formula

Probabilistic Information Processing → Probabilistic Model

Belief Propagation



J. Pearl: Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference (Morgan Kaufmann, 1988).
C. Berrou and A. Glavieux: Near optimum error correcting coding and decoding: Turbo-codes, IEEE Trans. Comm., 44 (1996).

Statistical-Mechanical Approach to Probabilistic Information Processing

Contents

1. Introduction
2. <u>Belief Propagation</u>
3. Bayesian Image Analysis and Gaussian Graphical Model
4. Concluding Remarks

Belief Propagation

- How should we treat the calculation of the summation over 2^N configuration?

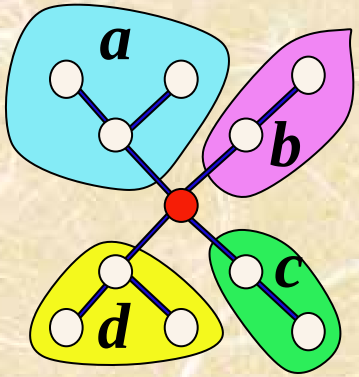
$$\sum_{x_1=0,1}\sum_{x_2=0,1}\cdots\sum_{x_N=0,1} W(x_1, x_2, \cdots, x_N)$$

It is very hard to calculate exactly except some special cases.

- **Formulation for approximate algorithm**
- **Accuracy of the approximate algorithm**

Tractable Model

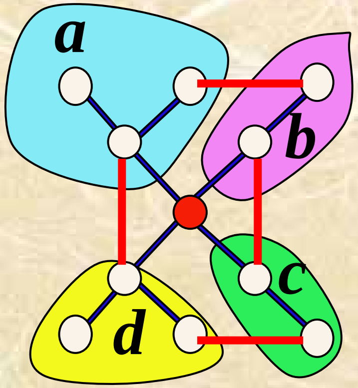
- Probabilistic models with no loop are tractable.



Factorizable

$$\sum_{a}\sum_{b}\sum_{c}\sum_{d} A(\boldsymbol{a},x)B(\boldsymbol{b},x)C(\boldsymbol{c},x)D(\boldsymbol{d},x) = \left(\sum_{\boldsymbol{a}} A(\boldsymbol{a},x)\right)\left(\sum_{\boldsymbol{b}} B(\boldsymbol{b},x)\right)\left(\sum_{\boldsymbol{c}} C(\boldsymbol{c},x)\right)\left(\sum_{\boldsymbol{d}} D(\boldsymbol{d},x)\right)$$

- Probabilistic models with loop are not tractable.



Not Factorizable

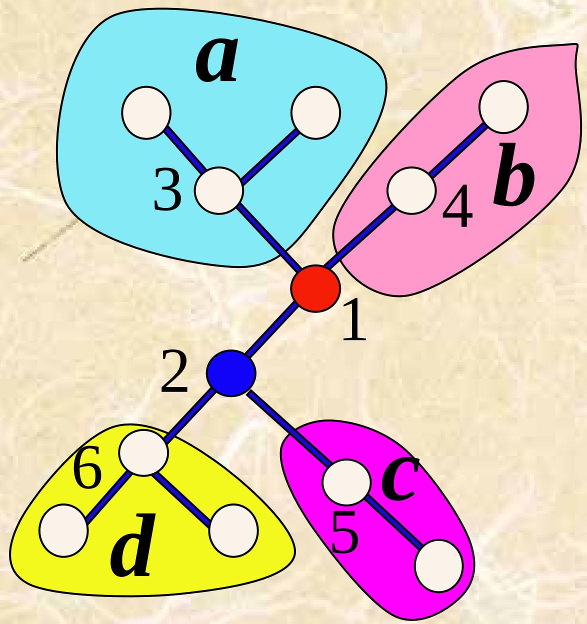
$$\sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\sum_{\boldsymbol{c}}\sum_{\boldsymbol{d}} W(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},x)$$

Probabilistic model on a graph with no loop

$$P(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},x,y)$$
$$\equiv W_A(\boldsymbol{a},x)W_B(\boldsymbol{b},x)W_{12}(x,y)W_C(\boldsymbol{c},y)W_D(\boldsymbol{d},y)$$

$$P_2(y) \equiv \sum_{\boldsymbol{a}}\sum_{\boldsymbol{b}}\sum_{\boldsymbol{c}}\sum_{\boldsymbol{d}}\sum_{x} P(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d},x,y)$$

Marginal probability of the node 2

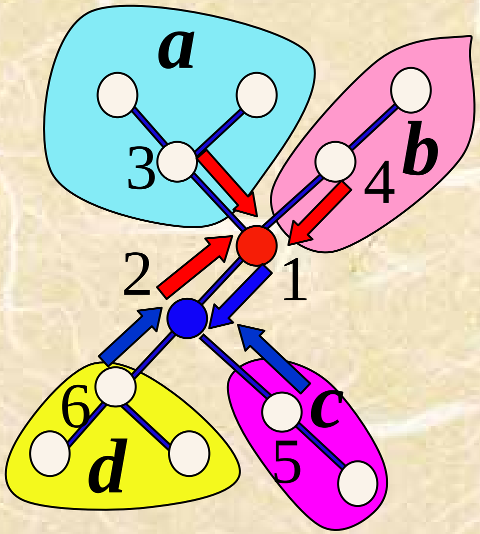


Probabilistic model on a graph with no loop

$$P_2(y) = M_{5\to2}(y) M_{6\to2}(y) M_{1\to2}(y)$$

$$M_{1\to2}(y) = \sum_x W_{12}(x,y) M_{3\to1}(x) M_{4\to1}(x)$$

$$M_{2\to1}(x) = \sum_y W_{12}(x,y) M_{5\to2}(y) M_{6\to2}(y)$$

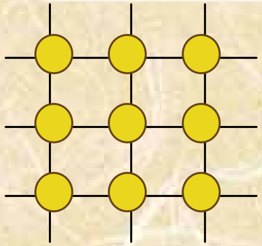


Marginal probability can be expressed in terms of the product of messages from all the neighbouring nodes of node 2.

Message from the node 1 to the node 2 can be expressed in terms of the product of message from all the neighbouring nodes of the node 1 except one from the node 2.

Probabilistic Model on a Graph with Loops

$$P(x_1, x_2, \cdots, x_L) = \frac{1}{Z} \prod_{ij \in N} W_{ij}(x_i, x_j)$$



Ω

$$Z \equiv \sum_{\boldsymbol{x}} \prod_{ij \in N} W_{ij}(x_i, x_j)$$

Marginal Probability

$$P_1(x_1) = \sum_{\boldsymbol{x} \backslash \{x_1\}} P(x_1, x_2, \cdots, x_L)$$

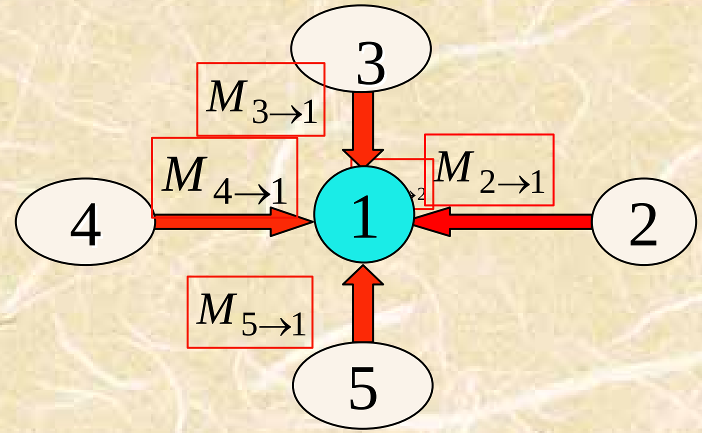
$$P_{12}(x_1, x_2) = \sum_{\boldsymbol{x} \backslash \{x_1, x_2\}} P(x_1, x_2, \cdots, x_L)$$

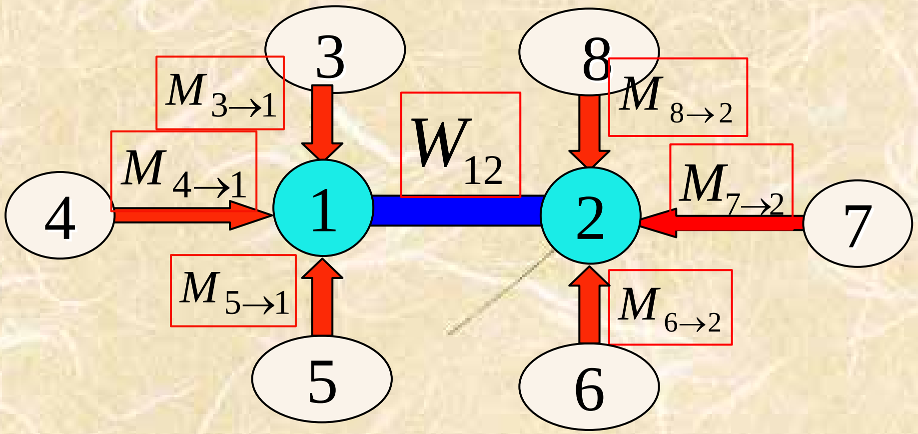
Belief Propagation

$$Q_1(\xi) = \sum_{\varsigma} Q_{12}(\xi, \varsigma)$$

$$M_{1\to2}(\xi) \propto \sum_{\varsigma} W_{12}(\varsigma, \xi) M_{3\to1}(\varsigma) \times M_{4\to1}(\varsigma) M_{5\to1}(\varsigma)$$

Message Update Rule



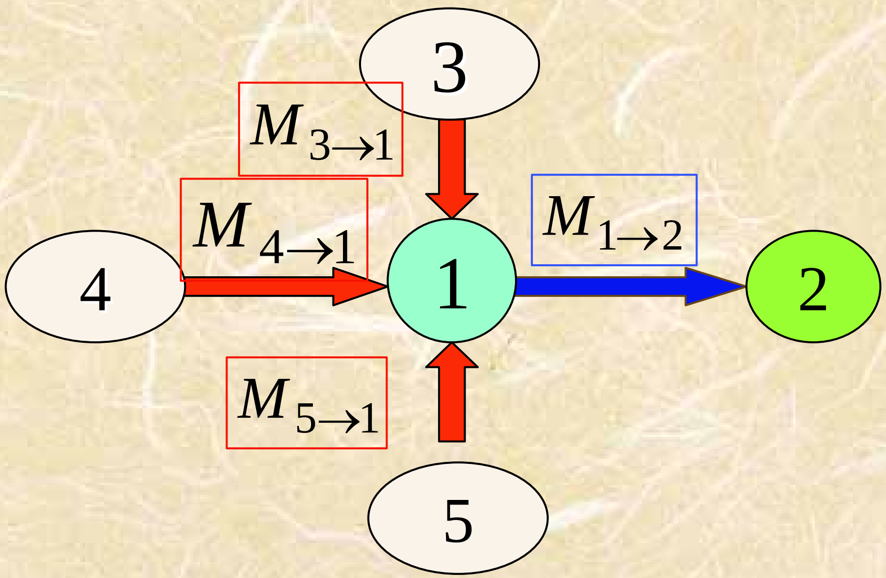


$$Q_1(x_1) = \frac{1}{Z_1} M_{2\to1}(x_1) M_{3\to1}(x_1) \times M_{4\to1}(x_1) M_{5\to1}(x_1)$$

$$Q_{12}(x_1, x_2) = \frac{1}{Z_{12}} M_{3\to1}(x_1) M_{4\to1}(x_1) M_{5\to1}(x_1) \times W_{12}(x_1, x_2) M_{6\to2}(x_2) M_{7\to2}(x_2) M_{8\to2}(x_2)$$

Message Passing Rule of Belief Propagation

$$M_{1\to 2}(\xi) = \frac{\sum_{\varsigma} W_{12}(\varsigma,\xi) M_{3\to 1}(\varsigma) M_{4\to 1}(\varsigma) M_{5\to 1}(\varsigma)}{\sum_{\xi}\sum_{\varsigma} W_{12}(\varsigma,\xi) M_{3\to 1}(\varsigma) M_{4\to 1}(\varsigma) M_{5\to 1}(\varsigma)}$$



Fixed Point Equations for Massage

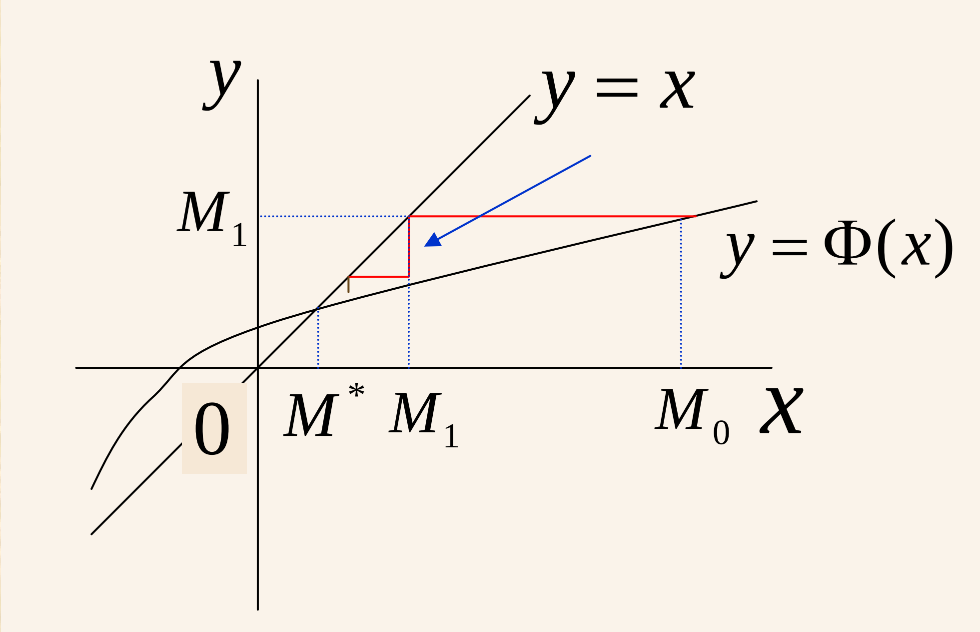
$$\vec{M} = \vec{\Phi}\left(\vec{M}\right)$$

Fixed Point Equation and Iterative Method

Fixed Point Equation $\vec{M}^* = \vec{\Phi}\left(\vec{M}^*\right)$

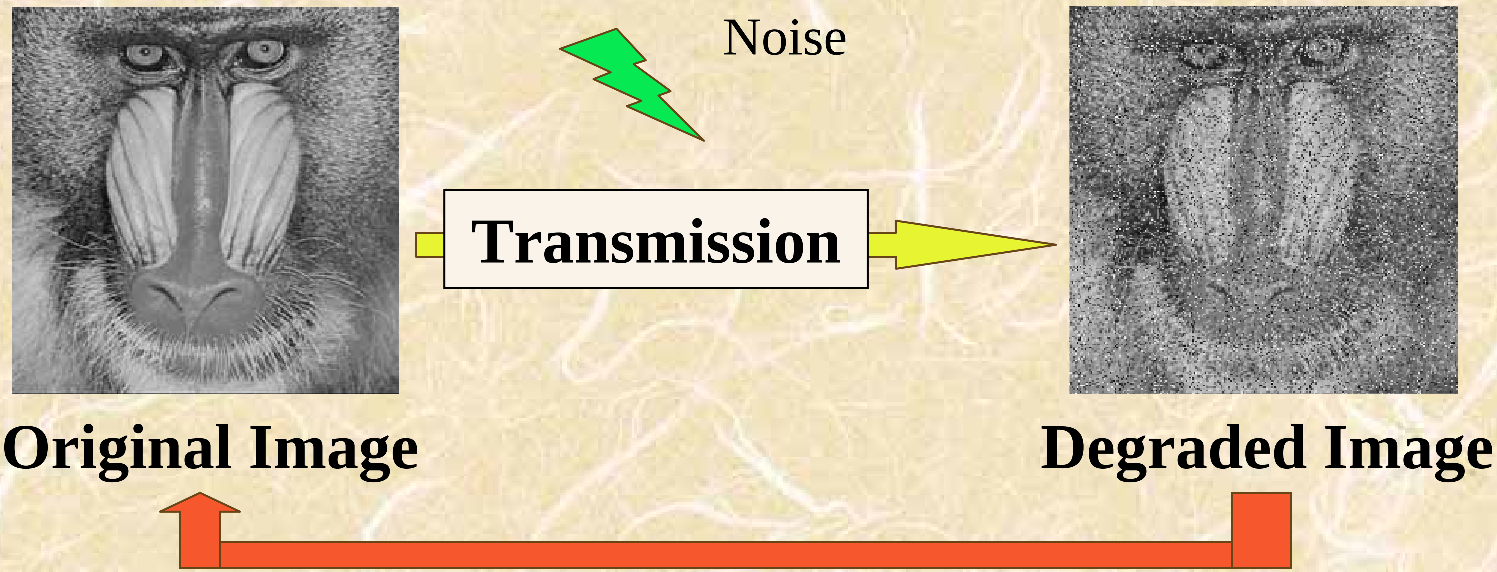
Iterative Method

$$\vec{M}_1 \leftarrow \Phi\left(\vec{M}_0\right)$$
$$\vec{M}_2 \leftarrow \Phi\left(\vec{M}_1\right)$$
$$\vec{M}_3 \leftarrow \Phi\left(\vec{M}_2\right)$$
$$\vdots$$



Contents

Bayesian Image Analysis

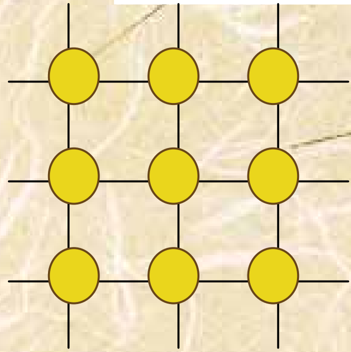


$$\underbrace{\Pr\left\{\text{Original Image}\,\middle|\,\text{Degraded Image}\right\}}_{\text{A Posteriori Probability}}=\frac{\overbrace{\Pr\left\{\text{Degraded Image}\,\middle|\,\text{Original Image}\right\}}^{\text{Degradation Process}}\overbrace{\Pr\left\{\text{Original Image}\right\}}^{\text{A Priori Probability}}}{\underbrace{\Pr\left\{\text{Degraded Image}\right\}}_{\text{Marginal Likelihood}}}$$

Bayesian Image Analysis

Degradation Process

$$f_i, g_i \in (-\infty, +\infty)$$

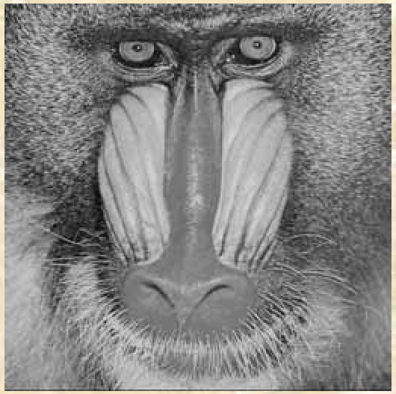


Ω

$$g_i = f_i + n_i$$

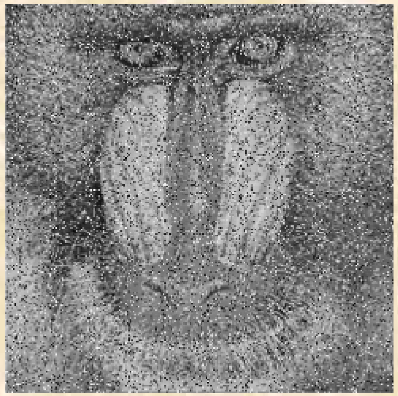
$$n_i \sim N\left(0, \sigma^2\right)$$

$$P(\boldsymbol{g} | \boldsymbol{f}, \sigma) = \prod_{i \in \Omega} \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2} (f_i - g_i)^2 \right)$$



Additive White Gaussian Noise

Transmission



Original Image **Degraded Image**

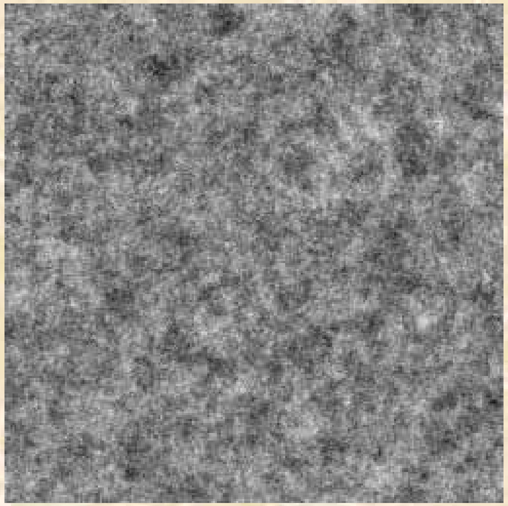
Bayesian Image Analysis

A Priori Probability

$$f_i, g_j \in (-\infty, +\infty)$$

$$P(\boldsymbol{f} \mid \alpha) = \frac{1}{Z_{\mathrm{PR}}(\alpha)} \prod_{ij \in N} \exp\left(-\frac{1}{2} \alpha (f_i - f_j)^2 \right)$$

Generate



Similar?

Standard Images

Bayesian Image Analysis

A Posteriori Probability

$$f_i, g_j \in (-\infty, +\infty)$$

$$P(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma) = \frac{P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)}{P(\boldsymbol{g}|\alpha,\sigma)}$$

$$= \frac{1}{Z_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)}$$

$$\times \exp\left(-\frac{1}{2\sigma^2}\sum_{i\in\Omega}(f_i - g_i)^2 - \frac{1}{2}\alpha\sum_{ij\in N}(f_i - f_j)^2 \right)$$

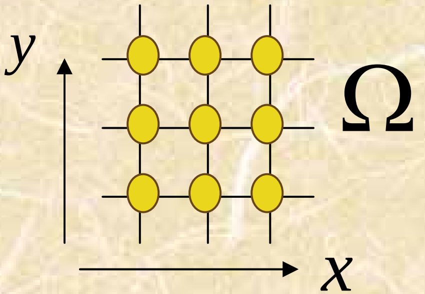
Gaussian Graphical Model

Bayesian Image Analysis

$P(\boldsymbol{f}|\alpha)$ → $\boldsymbol{f}$ — $P(\boldsymbol{g}|\boldsymbol{f},\sigma)$ → $\boldsymbol{g}$

A Priori Probability — Original Image — Degraded Image — Degraded Image

$\boldsymbol{f}=\{f_i|i\in\Omega\}$ $\boldsymbol{g}=\{g_i|i\in\Omega\}$



Pixels

A Posteriori Probability

$$P(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)=\frac{P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)}{P(\boldsymbol{g}|\alpha,\sigma)}$$

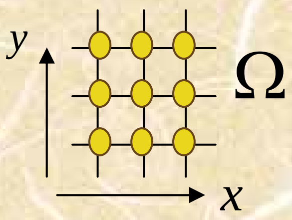
$$\hat{f}_i=\int f_i P(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)d\boldsymbol{f}=\int_{-\infty}^{+\infty} f_i P(f_i|\boldsymbol{g},\alpha,\sigma)df_i$$

Hyperparameter Determination by Maximization of Marginal Likelihood

$$(\hat{\alpha},\hat{\sigma}) = \arg\max_{(\alpha,\sigma)} P(\boldsymbol{g}|\alpha,\sigma)$$

$$\hat{f}_i = \int f_i P(\boldsymbol{f}|\boldsymbol{g},\hat{\alpha},\hat{\sigma})d\boldsymbol{f}$$

$$P(\boldsymbol{g}|\alpha,\sigma) = \int P(\boldsymbol{f},\boldsymbol{g}|\alpha,\sigma)d\boldsymbol{f} = \int P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)d\boldsymbol{f}$$



$$P(\boldsymbol{f},\boldsymbol{g}|\alpha,\sigma) = P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)$$

$$P(\boldsymbol{f}|\alpha) \Rightarrow \boldsymbol{f} \Rightarrow P(\boldsymbol{g}|\boldsymbol{f},\sigma) \Rightarrow \boldsymbol{g}$$

⇓ **Marginalization**

$$P(\boldsymbol{g}|\alpha,\sigma) \Rightarrow \boldsymbol{g}$$

Original Image

$$\boldsymbol{f} = \{f_i | i \in \Omega\}$$

Marginal Likelihood

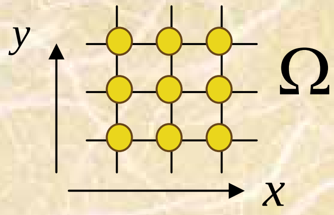
Degraded Image

$$\boldsymbol{g} = \{g_i | i \in \Omega\}$$

Maximization of Marginal Likelihood by EM (Expectation Maximization) Algorithm

Marginal Likelihood

$$P(\boldsymbol{g}|\alpha,\sigma)=\int P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)d\boldsymbol{f}$$



***Q*-Function**

$$Q(\alpha',\sigma'|\alpha,\sigma,\boldsymbol{g})=\int P(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)\ln P(\boldsymbol{f},\boldsymbol{g}|\alpha',\sigma')d\boldsymbol{f}$$

$$\frac{\partial}{\partial\sigma}P(\boldsymbol{g}|\alpha,\sigma)=0,\ \frac{\partial}{\partial\alpha}P(\boldsymbol{g}|\alpha,\sigma)=0$$

Incomplete Data

$$\boldsymbol{g}=\{g_i|i\in\Omega\}$$

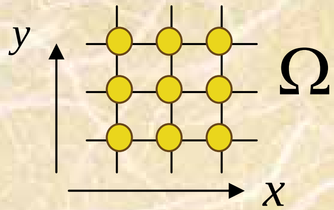
Equivalent

$$\left[\frac{\partial}{\partial\alpha'}Q(\alpha',\sigma'|\alpha,\sigma,\boldsymbol{g})\right]_{\alpha'=\alpha,\sigma'=\sigma}=0,\ \left[\frac{\partial}{\partial\sigma'}Q(\alpha',\sigma'|\alpha,\sigma,\boldsymbol{g})\right]_{\alpha'=\alpha,\sigma'=\sigma}=0$$

Maximization of Marginal Likelihood by EM (Expectation Maximization) Algorithm

Marginal Likelihood

$$P(\boldsymbol{g}|\alpha,\sigma)=\int P(\boldsymbol{g}|\boldsymbol{f},\sigma)P(\boldsymbol{f}|\alpha)d\boldsymbol{f}$$



***Q*-Function**

$$Q(\alpha',\sigma'|\alpha,\sigma,\boldsymbol{g})=\int P(\boldsymbol{f}|\boldsymbol{g},\alpha,\sigma)\ln P(\boldsymbol{f},\boldsymbol{g}|\alpha',\sigma')d\boldsymbol{f}$$

EM Algorithm

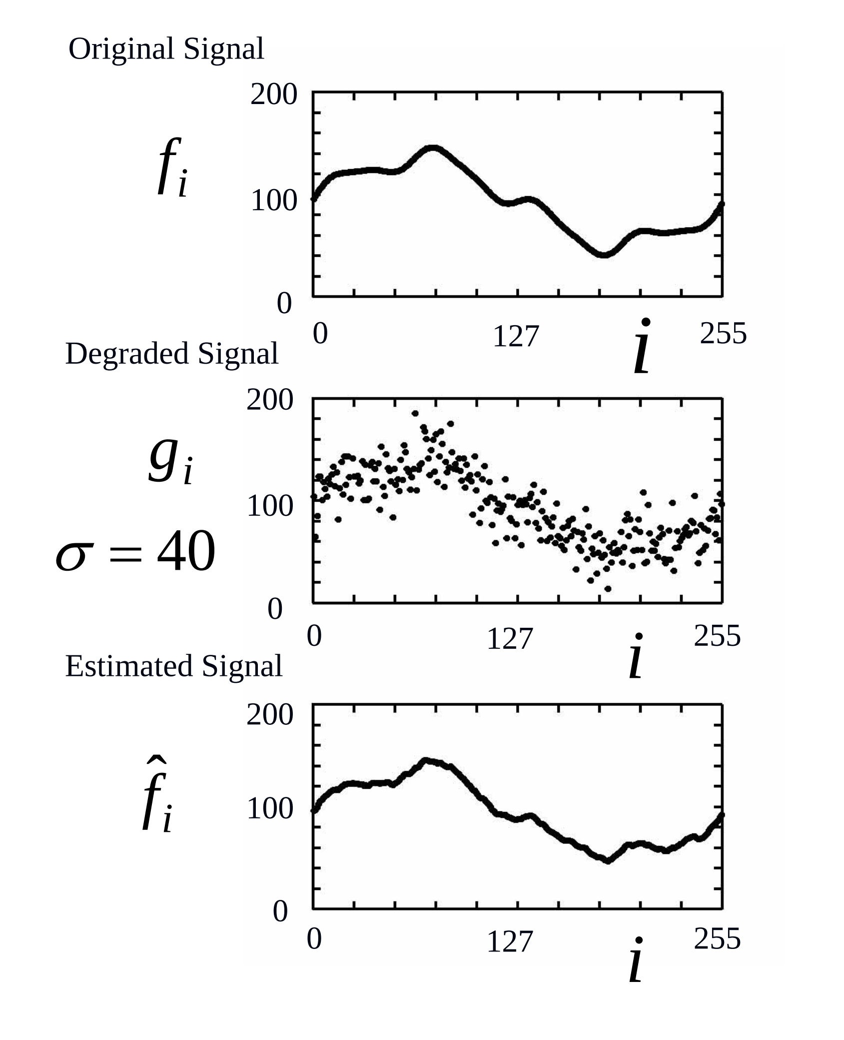
Iterate the following EM-steps until convergence:

$$\text{E-Step}: Q(\alpha',\sigma'|\alpha(t),\sigma(t))\leftarrow\int P(\boldsymbol{f}|\boldsymbol{g},\alpha(t),\sigma(t))\ln P(\boldsymbol{f},\boldsymbol{g}|\alpha',\sigma')d\boldsymbol{f}.$$

$$\text{M-Step}: (\alpha(t+1),\sigma(t+1))\leftarrow\arg\max_{(\alpha',\beta')} Q(\alpha',\sigma'|\alpha(t),\sigma(t)).$$

A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Roy. Stat. Soc. B, 39 (1977).

One-Dimensional Signal



EM Algorithm

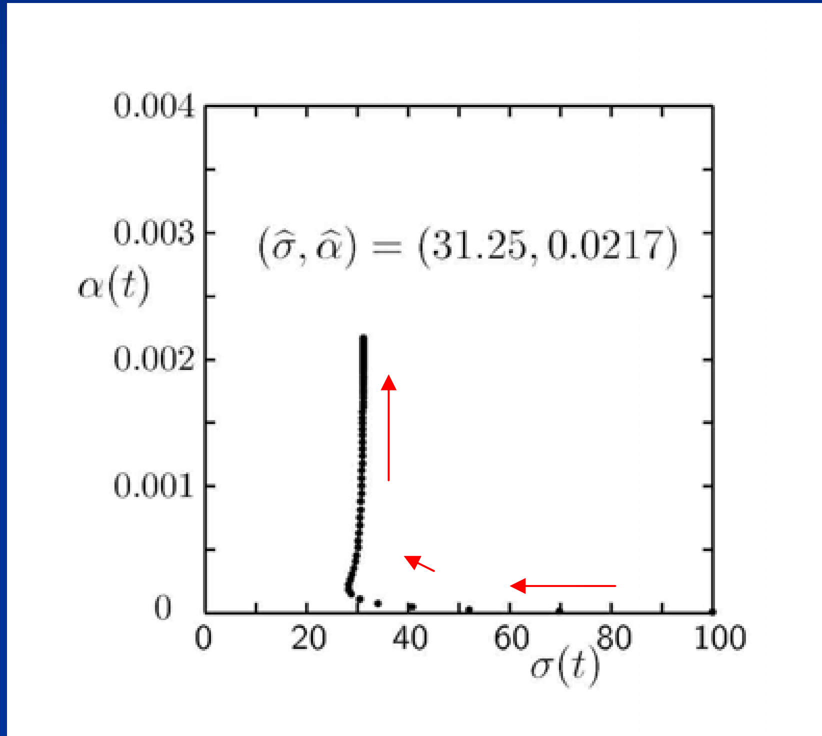
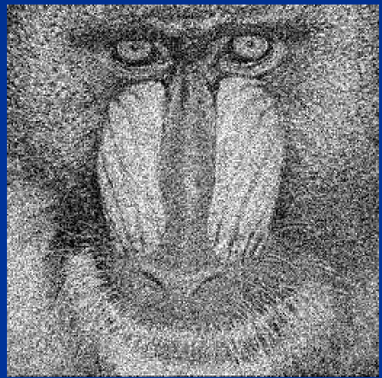


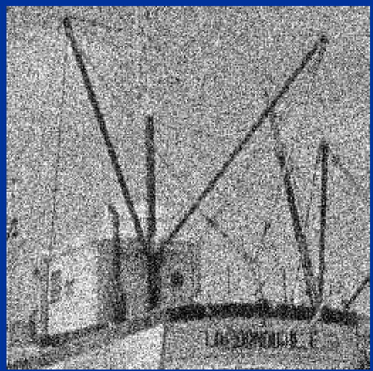
Image Restoration by Gaussian Graphical Model

EM Algorithm with Belief Propagation

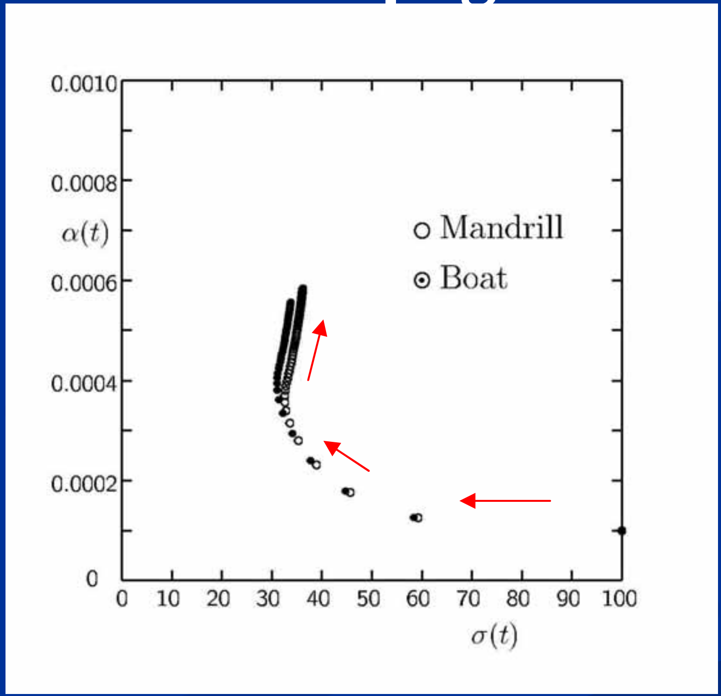
Original Image | **Degraded Image**



MSE: 1512



MSE: 1529



Exact Results of Gaussian Graphical Model

$$P\left(\boldsymbol{f}\middle|\boldsymbol{g},\alpha,\sigma\right) == \frac{1}{Z_{\mathrm{POS}}(\boldsymbol{g},\alpha,\sigma)}\exp\left(-\frac{1}{2\sigma^2}\sum_{i\in\Omega}(f_i-g_i)^2-\frac{1}{2}\alpha\sum_{ij\in N}(f_i-f_j)^2\right)$$

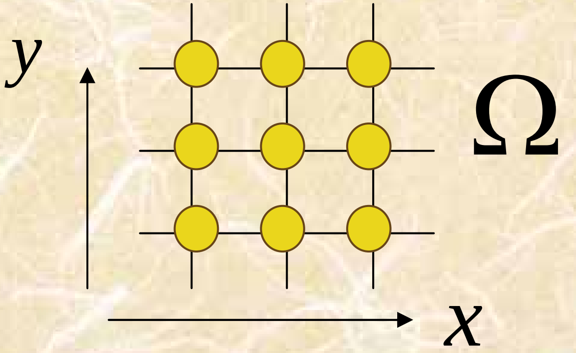
$$= \frac{\exp\left(-\frac{1}{2\sigma^2}\|\boldsymbol{f}-\boldsymbol{g}\|^2-\frac{1}{2}\alpha\boldsymbol{f}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{f}\right)}{\int\exp\left(-\frac{1}{2\sigma^2}\|\boldsymbol{f}-\boldsymbol{g}\|^2-\frac{1}{2}\alpha\boldsymbol{f}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{f}\right)d\boldsymbol{f}}$$

Multi-dimensional Gauss integral formula

$$P\left(\boldsymbol{g}\middle|\alpha,\sigma\right)=\sqrt{\frac{\det(\alpha\boldsymbol{C})}{(2\pi)^{|\Omega|}\det\left(\boldsymbol{I}+\alpha\sigma^2\boldsymbol{C}\right)}}\exp\left(-\frac{1}{2}\alpha\boldsymbol{g}^{\mathrm{T}}\frac{\boldsymbol{C}}{\boldsymbol{I}+\alpha\sigma^2\boldsymbol{C}}\boldsymbol{g}\right)$$

$$(\hat{\alpha},\hat{\sigma})=\arg\max_{(\alpha,\sigma)} P\left(\boldsymbol{g}\middle|\alpha,\sigma\right)$$

$$\hat{\boldsymbol{f}}=\left(\boldsymbol{I}+\alpha\sigma^2\boldsymbol{C}\right)^{-1}\boldsymbol{g}$$



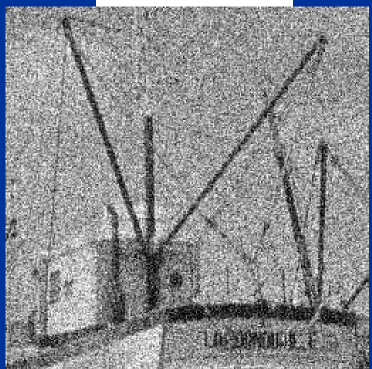
Comparison of Belief Propagation with Exact Results in Gaussian Graphical Model

$$(\hat{\alpha},\hat{\sigma}) = \underset{(\alpha,\sigma)}{\arg\max}\; P(\boldsymbol{g}|\alpha,\sigma)$$

$\sigma = 40$

	MSE	$\hat{\alpha}$	$\hat{\sigma}$	$\ln P(\boldsymbol{g}\|\hat{\alpha},\hat{\sigma})$
Belief Propagation	**327**	**0.000611**	**36.302**	**-5.19201**
Exact	**315**	**0.000759**	**37.919**	**-5.21444**

$\sigma = 40$

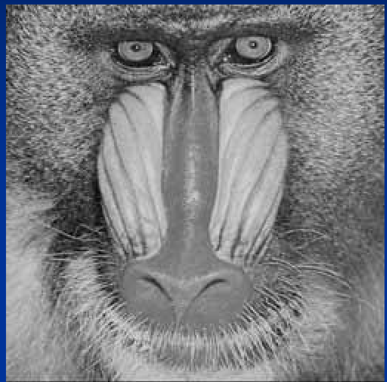


	MSE	$\hat{\alpha}$	$\hat{\sigma}$	$\ln P(\boldsymbol{g}\|\hat{\alpha},\hat{\sigma})$
Belief Propagation	**260**	**0.000574**	**33.998**	**-5.15241**
Exact	**236**	**0.000652**	**34.975**	**-5.17528**

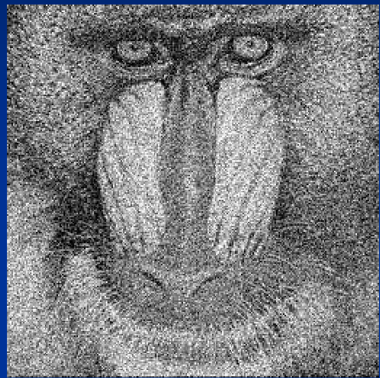
$$\mathrm{MSE} = \frac{1}{|\Omega|}\sum_{i\in\Omega}\left(f_i - \hat{f}_i\right)^2$$

Image Restoration by Gaussian Graphical Model

Original Image

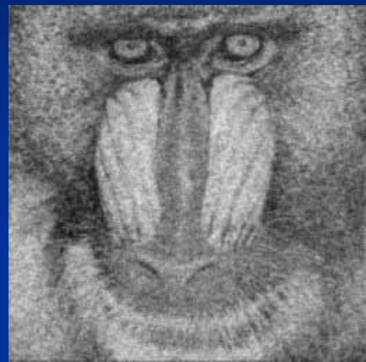


Degraded Image



MSE: 1512

Belief Propagation

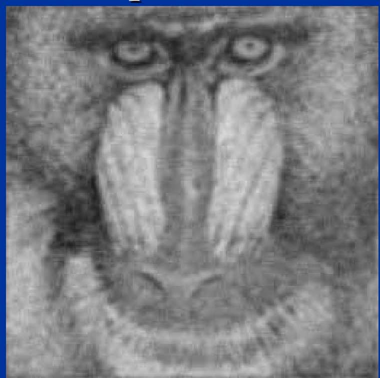


MSE: 325

Exact

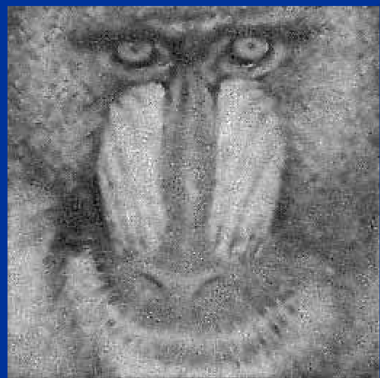
MSE:315

Lowpass Filter



MSE: 411

Wiener Filter



MSE: 545

Median Filter

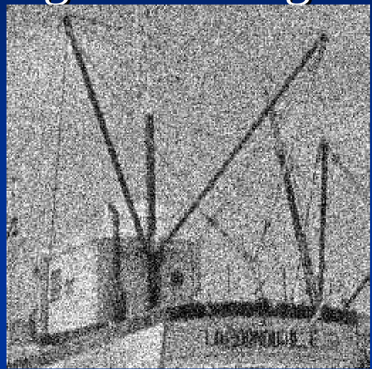
MSE: 447

8 $$\mathrm{MSE} = \frac{1}{|\Omega|}\sum_{i\in\Omega}\left(f_i - \hat{f}_i\right)^2$$

Image Restoration by Gaussian Graphical Model

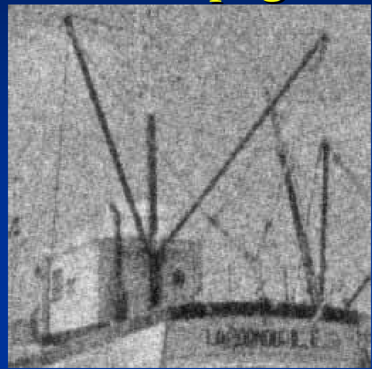
Original Image

Degraded Image



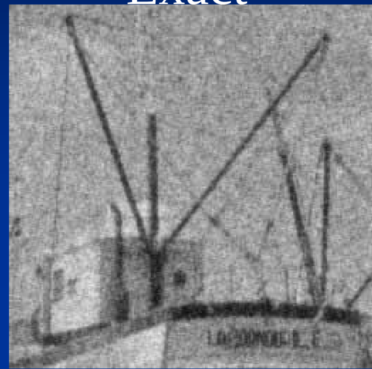
MSE: 1529

Belief Propagation



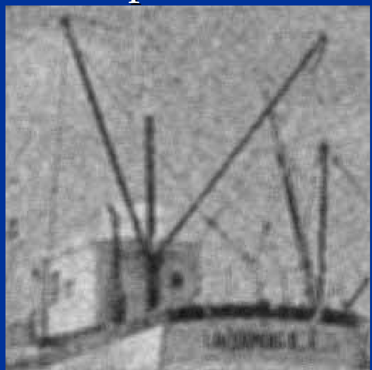
MSE: 260

Exact



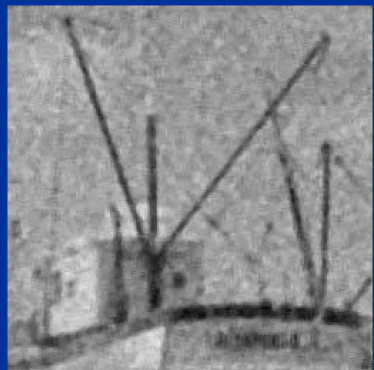
MSE236

Lowpass Filter



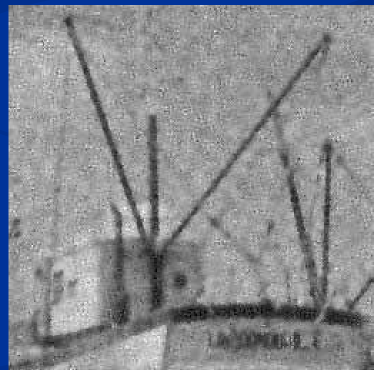
MSE: 224

Wiener Filter



MSE: 372

Median Filter



MSE: 244

$$\mathrm{MSE} = \frac{1}{|\Omega|}\sum_{i\in\Omega}\left(f_i - \hat{f}_i\right)^2$$

Extension of Belief Propagation

- **Generalized Belief Propagation**

  J. S. Yedidia, W. T. Freeman and Y. Weiss: Constructing free-energy approximations and generalized belief propagation algorithms, IEEE Transactions on Information Theory, 51 (2005).

- **Generalized belief propagation is equivalent to the cluster variation method in statistical mechanics**

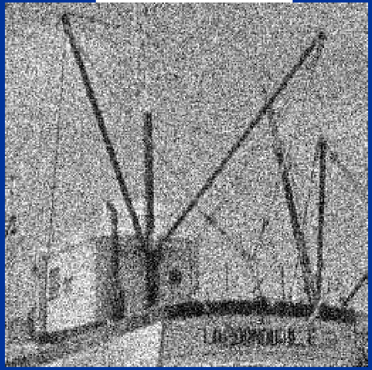
  R. Kikuchi: A theory of cooperative phenomena, Phys. Rev., 81 (1951).

  T. Morita: Cluster variation method of cooperative phenomena and its generalization I, J. Phys. Soc. Jpn, 12 (1957).

Image Restoration by Gaussian Graphical Model

$\sigma = 40$

$\sigma = 40$



$$(\hat{\alpha},\hat{\sigma}) = \arg\max_{(\alpha,\sigma)} P(\boldsymbol{g}|\alpha,\sigma)$$

$$\mathrm{MSE} = \frac{1}{|\Omega|}\sum_{i\in\Omega}(f_i - \hat{f}_i)^2$$

	MSE	$\hat{\alpha}$	$\hat{\sigma}$	$\ln P(\boldsymbol{g}\|\hat{\alpha},\hat{\sigma})$
Belief Propagation	327	0.000611	36.302	-5.19201
Generalized Belief Propagation	315	0.000758	37.909	-5.21172
Exact	315	0.000759	37.919	-5.21444

	MSE	$\hat{\alpha}$	$\hat{\sigma}$	$\ln P(\boldsymbol{g}\|\hat{\alpha},\hat{\sigma})$
Belief Propagation	260	0.000574	33.998	-5.15241
Generalized Belief Propagation	236	0.000652	34.971	-5.17256
Exact	236	0.000652	34.975	-5.17528

Image Restoration by Gaussian Graphical Model and Conventional Filters

$\sigma = 40$

GBP

	MSE			MSE
Belief Propagation	327	**Lowpass Filter**	(3x3)	388
			(5x5)	413
Generalized Belief Propagation	315	**Median Filter**	(3x3)	486
			(5x5)	445
Exact	315	**Wiener Filter**	(3x3)	864
			(5x5)	548

(3x3) Lowpass

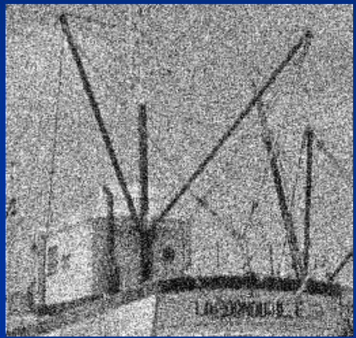
(5x5) Median

(5x5) Wiener

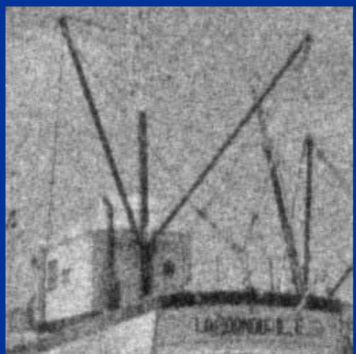
$$\mathrm{MSE} = \frac{1}{|\Omega|}\sum_{(x,y)\in\Omega}\left(f_{x,y} - \hat{f}_{x,y}\right)^2$$

Image Restoration by Gaussian Graphical Model and Conventional Filters

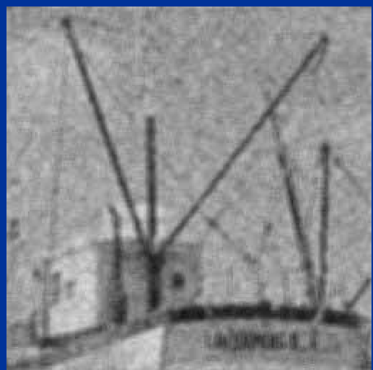
$\sigma = 40$



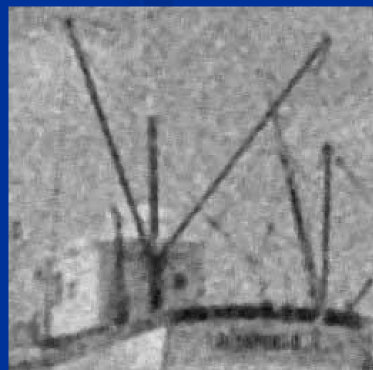
GBP



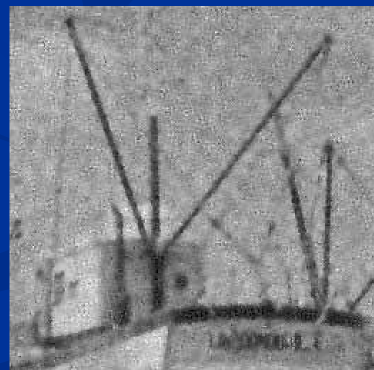
	MSE			MSE
Belief Propagation	260	**Lowpass Filter**	(3x3)	241
			(5x5)	224
Generalized Belief Propagation	236	**Median Filter**	(3x3)	331
			(5x5)	244
Exact	236	**Wiener Filter**	(3x3)	703
			(5x5)	372



(5x5) Lowpass



(5x5) Median



(5x5) Wiener

$$\mathrm{MSE} = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \left(f_{x,y} - \hat{f}_{x,y} \right)^2$$

Contents

Summary

- Formulation of belief propagation
- Accuracy of belief propagation in Bayesian image analysis by means of Gaussian graphical model (Comparison between the belief propagation and exact calculation)