Information Theory and Image/Video Coding

Ming Jiang

School of Mathematical Sciences
Peking University

ming-jiang@pku.edu.cn

March 12, 2012

Outline

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  1. [Geman, 1990].
  2. [Wrinkler, 1995].
  3. [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i. [Geman, 1990].
  ii. [Wrinkler, 1995].
  iii. [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  1. [Geman, 1990].
  2. [Wrinkler, 1995].
  3. [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  - i [Geman, 1990].
  - ii [Wrinkler, 1995].
  - iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Saptial Dependence & Markov Random Fields

- Each pixel value
  - depends only on neighboring pixel values;
  - is independent from far away pixel values.
- Markov random fields provide a flexible mechanism for
  - modeling spatial dependence,
  - image attributes.
- General references for this section are
  i [Geman, 1990].
  ii [Wrinkler, 1995].
  iii [Li, 2009].
- Most of the materials is copy-edited from [Geman, 1990].

Measure on State Space

- Assume that there is a positive measure defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the Borel algebra and κ some Borel measure;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the subset algebra and κ the counting measure.

Measure on State Space

- Assume that there is a positive `measure` defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the `Borel algebra` and κ some `Borel measure`;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the `subset algebra` and κ the `counting measure`.

Measure on State Space

- Assume that there is a positive measure defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the Borel algebra and κ some Borel measure;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the subset algebra and κ the counting measure.

Measure on State Space

- Assume that there is a positive `measure` defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the `Borel algebra` and κ some `Borel measure`;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the `subset algebra` and κ the `counting measure`.

Measure on State Space

- Assume that there is a positive `measure` defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the `Borel algebra` and κ some `Borel measure`;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the `subset algebra` and κ the `counting measure`.

Measure on State Space

- Assume that there is a positive `measure` defined on each state space Λ, respectively, i.e.,
  - $(\Lambda, \mathcal{E})$ is a measurable space with positive measure κ on the σ-algebra $\mathcal{E}$.
- The state space Λ is generally a subset of $\mathbf{R}^q$.
- Two typical cases are
  - if Λ is not of zero measure, $\mathcal{E}$ is the `Borel algebra` and κ some `Borel measure`;
  - if Λ is a finite or countable subset of $\mathbf{R}^q$, $\mathcal{E}$ is the `subset algebra` and κ the `counting measure`.

Random Fields

- A probability measure on $\mathcal{T}$ defines a `random field`:

Definition

Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. The triple $(\Omega, \mathcal{T}, \Pi)$ is called a random field with the site set S and state space Λ if:

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$;
- Π *is a probability measure*
  - *a positive measure such that* $\Pi(\Omega) = 1$.

Random Fields

- A probability measure on $\mathcal{T}$ defines a `random field`:

Definition

Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. The triple $(\Omega, \mathcal{T}, \Pi)$ is called a random field with the site set S and state space Λ if:

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$;
- Π *is a probability measure*
  - *a positive measure such that* $\Pi(\Omega) = 1$.

Random Fields

- A probability measure on $\mathcal{T}$ defines a `random field`:

Definition

Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. The triple $(\Omega, \mathcal{T}, \Pi)$ is called a random field with the site set S and state space Λ if:

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$;
- *Π is a probability measure*
  - *a positive measure such that $\Pi(\Omega) = 1$.*

Random Fields

- A probability measure on $\mathcal{T}$ defines a `random field`:

Definition

Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. The triple $(\Omega, \mathcal{T}, \Pi)$ is called a random field with the site set S and state space Λ if:

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$;
- Π *is a probability measure*
  - *a positive measure such that* $\Pi(\Omega) = 1$.

Random Fields

- A probability measure on $\mathcal{T}$ defines a `random field`:

Definition

Let S be a finite site set and $(\Lambda, \mathcal{E}, \kappa)$ be a state space. The triple $(\Omega, \mathcal{T}, \Pi)$ is called a random field with the site set S and state space Λ if:

- $(\Omega, \mathcal{T}) = (\Lambda, \mathcal{E})^S$;
- *Π is a probability measure*
  - *a positive measure such that $\Pi(\Omega) = 1$.*

Local Characteristics

- The local characteristics refer to the family of univariable, conditional distributions, for $s \in S$ and $x \in \Omega$, and $\lambda \in \Lambda$,

$$\Pi(\lambda | x_{(s)}) \triangleq \Pi_s(x_s | x_{(s)}) \quad (1)$$
$$= \mathbf{Pr}(X_s = x_s | X_r = x_r, r \neq s), \quad (2)$$

where $\lambda = x_s$ and $x_{(s)} = (x_r)_{r \neq s}$.

- Theorem
The distribution Π of the random field $(\Omega, \mathcal{T}, \Pi)$ is determined by its local characteristics.

Local Characteristics

- The `local characteristics` refer to the family of univariable, conditional distributions, for $s \in S$ and $x \in \Omega$, and $\lambda \in \Lambda$,

$$\Pi(\lambda | x_{(s)}) \triangleq \Pi_s(x_s | x_{(s)}) \tag{1}$$
$$= \mathbf{Pr}(X_s = x_s | X_r = x_r, r \neq s), \tag{2}$$

where $\lambda = x_s$ and $x_{(s)} = (x_r)_{r \neq s}$.

- **Theorem**

The distribution Π of the random field $(\Omega, \mathcal{T}, \Pi)$ is determined by its local characteristics.

Local Characteristics

- The `local characteristics` refer to the family of univariable, conditional distributions, for $s \in S$ and $x \in \Omega$, and $\lambda \in \Lambda$,

$$\Pi(\lambda | x_{(s)}) \triangleq \Pi_s(x_s | x_{(s)}) \quad (1)$$

$$= \mathbf{Pr}(X_s = x_s | X_r = x_r, r \neq s), \quad (2)$$

  where $\lambda = x_s$ and $x_{(s)} = (x_r)_{r \neq s}$.

- Theorem

  The distribution Π of the random field $(\Omega, \mathcal{T}, \Pi)$ is determined by its local characteristics.

Proof I

- We will verify that for any $x = (x_i)$ and $y = (y_i)$,

$$\frac{\Pi(x)}{\Pi(y)} = \prod_{i=1}^{N} \frac{\Pi(x_i|x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}{\Pi(y_i|x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}. \tag{3}$$

- Assume (3) holds and that two probability measures Π and μ have the same local characteristics.
- It implies that

$$\frac{\Pi(x)}{\Pi(y)} = \frac{\mu(x)}{\mu(y)}. \tag{4}$$

- It follows that

$$\Pi(x)\mu(y) = \mu(x)\Pi(y), \tag{5}$$

- Summing over $y \in \Omega$ leads to the result $\Pi = \mu$.

Proof I

- We will verify that for any $x = (x_i)$ and $y = (y_i)$,

$$\frac{\Pi(x)}{\Pi(y)} = \prod_{i=1}^{N} \frac{\Pi(x_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}{\Pi(y_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}. \tag{3}$$

- Assume (3) holds and that two probability measures Π and μ have the same local characteristics.
- It implies that

$$\frac{\Pi(x)}{\Pi(y)} = \frac{\mu(x)}{\mu(y)}. \tag{4}$$

- It follows that

$$\Pi(x)\mu(y) = \mu(x)\Pi(y), \tag{5}$$

- Summing over $y \in \Omega$ leads to the result $\Pi = \mu$.

Proof I

- We will verify that for any $x = (x_i)$ and $y = (y_i)$,

$$\frac{\Pi(x)}{\Pi(y)} = \prod_{i=1}^{N} \frac{\Pi(x_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}{\Pi(y_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}. \qquad (3)$$

- Assume (3) holds and that two probability measures Π and μ have the same local characteristics.
- It implies that

$$\frac{\Pi(x)}{\Pi(y)} = \frac{\mu(x)}{\mu(y)}. \qquad (4)$$

- It follows that

$$\Pi(x)\mu(y) = \mu(x)\Pi(y), \qquad (5)$$

- Summing over $y \in \Omega$ leads to the result $\Pi = \mu$.

Proof I

- We will verify that for any $x = (x_i)$ and $y = (y_i)$,

$$\frac{\Pi(x)}{\Pi(y)} = \prod_{i=1}^{N} \frac{\Pi(x_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}{\Pi(y_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}. \tag{3}$$

- Assume (3) holds and that two probability measures Π and μ have the same local characteristics.
- It implies that

$$\frac{\Pi(x)}{\Pi(y)} = \frac{\mu(x)}{\mu(y)}. \tag{4}$$

- It follows that

$$\Pi(x)\mu(y) = \mu(x)\Pi(y), \tag{5}$$

- Summing over $y \in \Omega$ leads to the result $\Pi = \mu$.

Proof I

- We will verify that for any $x = (x_i)$ and $y = (y_i)$,

$$\frac{\Pi(x)}{\Pi(y)} = \prod_{i=1}^{N} \frac{\Pi(x_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}{\Pi(y_i | x_1, \cdots, x_{i-1}, y_{i+1}, \cdots, y_N)}. \tag{3}$$

- Assume (3) holds and that two probability measures Π and μ have the same local characteristics.
- It implies that

$$\frac{\Pi(x)}{\Pi(y)} = \frac{\mu(x)}{\mu(y)}. \tag{4}$$

- It follows that

$$\Pi(x)\mu(y) = \mu(x)\Pi(y), \tag{5}$$

- Summing over $y \in \Omega$ leads to the result $\Pi = \mu$.

Proof: "$\Longleftarrow$" (2)

$$
\begin{aligned}
-V_A(x) &= \sum_{B\subset A} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) \\
&= \sum_{B\subset A, s\notin B, t\notin B} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) + \sum_{B\subset A, s\in B, t\notin B} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) \\
&\quad + \sum_{B\subset A, s\notin B, t\in B} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) + \sum_{B\subset A, s\in B, t\in B} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) \\
&= \sum_{B\subset A-\{s\}-\{t\}} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) + \sum_{s\in B_1\subset A-\{t\}} (-1)^{|A-B_1|} \log \Pi(x_s^{B_1} | x_{(s)}^{B_1}) \\
&\quad + \sum_{t\in B_2\subset A-\{s\}} (-1)^{|A-B_2|} \log \Pi(x_s^{B_2} | x_{(s)}^{B_2}) + \sum_{\{s,t\}\subset B_3\subset A} (-1)^{|A-B_3|} \log \Pi(x_s^{B_3} | x_{(s)}^{B_3}) \\
&= \sum_{B\subset A-\{s\}-\{t\}} (-1)^{|A-B|} \log \Pi(x_s^B | x_{(s)}^B) \\
&\quad + \sum_{B\subset A-\{s\}-\{t\}} (-1)^{|A-B-\{s\}|} \log \Pi(x_s^{B\cup\{s\}} | x_{(s)}^{B\cup\{s\}}) \\
&\quad + \sum_{B\subset A-\{s\}-\{t\}} (-1)^{|A-B-\{t\}|} \log \Pi(x_s^{B\cup\{t\}} | x_{(s)}^{B\cup\{t\}}) \\
&\quad + \sum_{B\subset A-\{s\}-\{t\}} (-1)^{|A-B-\{s\}-\{t\}|} \log \Pi(x_s^{B\cup\{s,t\}} | x_{(s)}^{B\cup\{s,t\}}).
\end{aligned}
$$

Squared-error Loss and MMSE

- Squared-error loss function

$$L(\theta, \hat{\theta}) = \|\theta - \hat{\theta}\|^2 = \sum_s |\theta_s - \hat{\theta}_s|^2. \tag{6}$$

- The Bayesian risk

$$\hat{R} = \int_{\mathfrak{X}} \int_{\Theta} \|\theta - \hat{\theta}(x)\|^2 \mathbf{Pr}(\theta, x) \tag{7}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \left\{ \|\theta\|^2 - 2\left\langle \theta, \hat{\theta}(x) \right\rangle + \|\hat{\theta}(x)\|^2 \right\} \mathbf{Pr}(\theta, x) \tag{8}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \left\{ 2\left\langle \theta_{MMSE}(x), \hat{\theta}(x) \right\rangle - \|\hat{\theta}(x)\|^2 \right\} m(x) \tag{9}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \|\theta_{MMSE}(x)\|^2 m(x) + \int_{\mathfrak{X}} \left\{ \|\theta_{MMSE}(x) - \hat{\theta}(x)\|^2 \right\} m(x) \tag{10}$$

is minimized when $\hat{\theta}(x) = \theta_{MMSE}(x)$.

- References
  - [Wrinkler, 1995].

Squared-error Loss and MMSE

- Squared-error loss function

$$L(\theta, \hat{\theta}) = \|\theta - \hat{\theta}\|^2 = \sum_s |\theta_s - \hat{\theta}_s|^2. \tag{6}$$

- The Bayesian risk

$$\hat{R} = \int_{\mathfrak{X}} \int_{\Theta} \|\theta - \hat{\theta}(x)\|^2 \mathbf{Pr}(\theta, x) \tag{7}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \left\{ \|\theta\|^2 - 2\left\langle \theta, \hat{\theta}(x) \right\rangle + \|\hat{\theta}(x)\|^2 \right\} \mathbf{Pr}(\theta, x) \tag{8}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \left\{ 2\left\langle \theta_{MMSE}(x), \hat{\theta}(x) \right\rangle - \|\hat{\theta}(x)\|^2 \right\} m(x) \tag{9}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \|\theta_{MMSE}(x)\|^2 m(x) + \int_{\mathfrak{X}} \left\{ \|\theta_{MMSE}(x) - \hat{\theta}(x)\|^2 \right\} m(x) \tag{10}$$

is minimized when $\hat{\theta}(x) = \theta_{MMSE}(x)$.

- References
  - [Wrinkler, 1995].

Squared-error Loss and MMSE

- Squared-error loss function

$$L(\theta,\hat{\theta}) = \|\theta-\hat{\theta}\|^2 = \sum_s |\theta_s-\hat{\theta}_s|^2. \qquad (6)$$

- The Bayesian risk

$$\hat{R} = \int_{\mathfrak{X}}\int_{\Theta} \|\theta-\hat{\theta}(x)\|^2 \mathbf{Pr}(\theta,x) \qquad (7)$$

$$= \int_{\mathfrak{X}}\int_{\Theta} \left\{\|\theta\|^2 - 2\left\langle \theta,\hat{\theta}(x)\right\rangle + \|\hat{\theta}(x)\|^2\right\} \mathbf{Pr}(\theta,x) \qquad (8)$$

$$= \int_{\mathfrak{X}}\int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta,x) - \int_{\mathfrak{X}} \left\{2\left\langle \theta_{MMSE}(x),\hat{\theta}(x)\right\rangle - \|\hat{\theta}(x)\|^2\right\} m(x) \qquad (9)$$

$$= \int_{\mathfrak{X}}\int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta,x) - \int_{\mathfrak{X}} \|\theta_{MMSE}(x)\|^2 m(x) + \int_{\mathfrak{X}} \left\{\|\theta_{MMSE}(x)-\hat{\theta}(x)\|^2\right\} m(x) \qquad (10)$$

is minimized when $\hat{\theta}(x) = \theta_{MMSE}(x)$.

- References
  - [Wrinkler, 1995].

Squared-error Loss and MMSE

- Squared-error loss function

$$L(\theta, \hat{\theta}) = \|\theta - \hat{\theta}\|^2 = \sum_s |\theta_s - \hat{\theta}_s|^2. \tag{6}$$

- The Bayesian risk

$$\hat{R} = \int_{\mathfrak{X}} \int_{\Theta} \|\theta - \hat{\theta}(x)\|^2 \mathbf{Pr}(\theta, x) \tag{7}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \left\{ \|\theta\|^2 - 2\left\langle \theta, \hat{\theta}(x) \right\rangle + \|\hat{\theta}(x)\|^2 \right\} \mathbf{Pr}(\theta, x) \tag{8}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \left\{ 2\left\langle \theta_{MMSE}(x), \hat{\theta}(x) \right\rangle - \|\hat{\theta}(x)\|^2 \right\} m(x) \tag{9}$$

$$= \int_{\mathfrak{X}} \int_{\Theta} \|\theta\|^2 \mathbf{Pr}(\theta, x) - \int_{\mathfrak{X}} \|\theta_{MMSE}(x)\|^2 m(x) + \int_{\mathfrak{X}} \left\{ \|\theta_{MMSE}(x) - \hat{\theta}(x)\|^2 \right\} m(x) \tag{10}$$

is minimized when $\hat{\theta}(x) = \theta_{MMSE}(x)$.

- References
  - i [Wrinkler, 1995].

References I

Geman, D. (1990).
Random fields and inverse problems in imaging.
In Ancona, A., Geman, D., Ikeda, N., and Geman, D., editors, *École d'Été de Probabilités de Saint-Flour XVIII i– 1988*, volume 1427 of *Lecture Notes in Mathematics*, pages 113–193. Springer Berlin / Heidelberg.

Li, S. Z. (2009).
Markov Random Field Modeling in Image Analysis.
Springer Publishing Company, Incorporated, 3rd edition.

Wrinkler, G. (1995).
Image Analysis, Random Fields and Dynamic Monte Carlo Methods.
Springer, Berlin–Heideberg.