Statistical Models for Image Processing: Hierarchical Representation of Global and Local Structures of Images

Yuta Nakahara, Center for Data Science, Waseda University

12/12 Waseda University – Academia Sinica Data Science Workshop

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

What is image processing?

- Denoising

http://people.tuebingen.mpg.de/burger/neural_denoising/

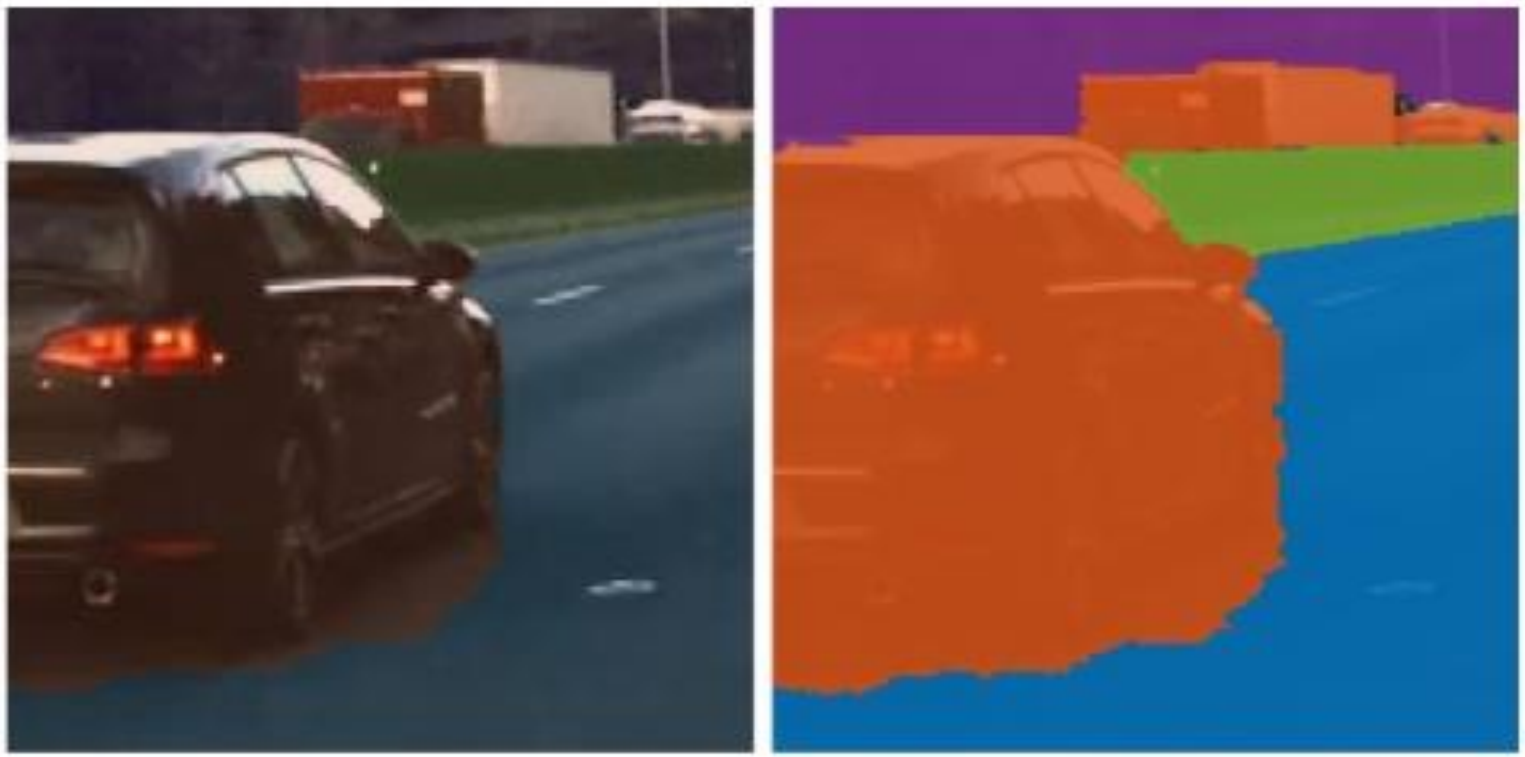
What is image processing?

■Edge detection

http://www.mis.med.akita-u.ac.jp/~kata/image/sobelprew.html

What is image processing?

- Segmentation



https://jp.mathworks.com/discovery/image-segmentation.html

What is image processing?

- Recognition

Cat

Dog

What is image processing?

■Recognition

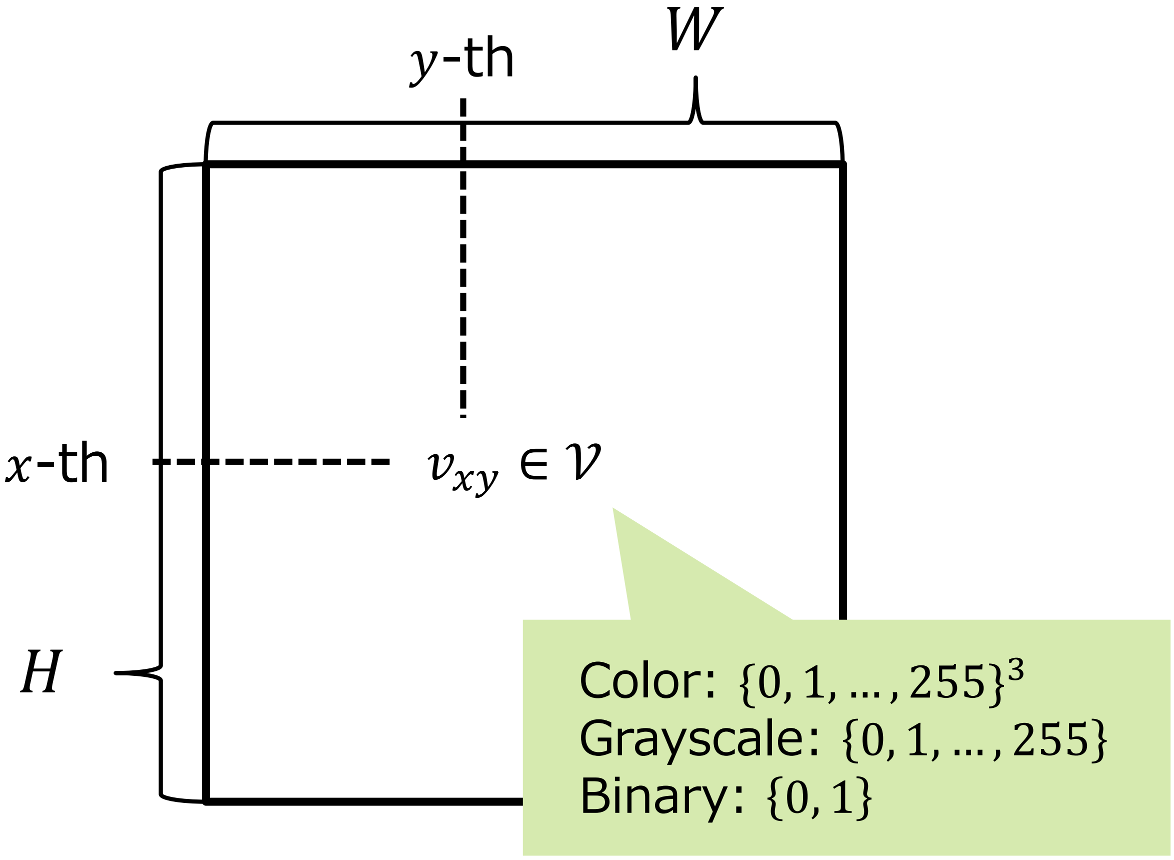
Cat

Dog

Image processing = Processing whose inputs are images

Notation

- Digital images



$$\boldsymbol{v} = [v_{00}, v_{01}, \dots, v_{H-1\,W-1}]$$

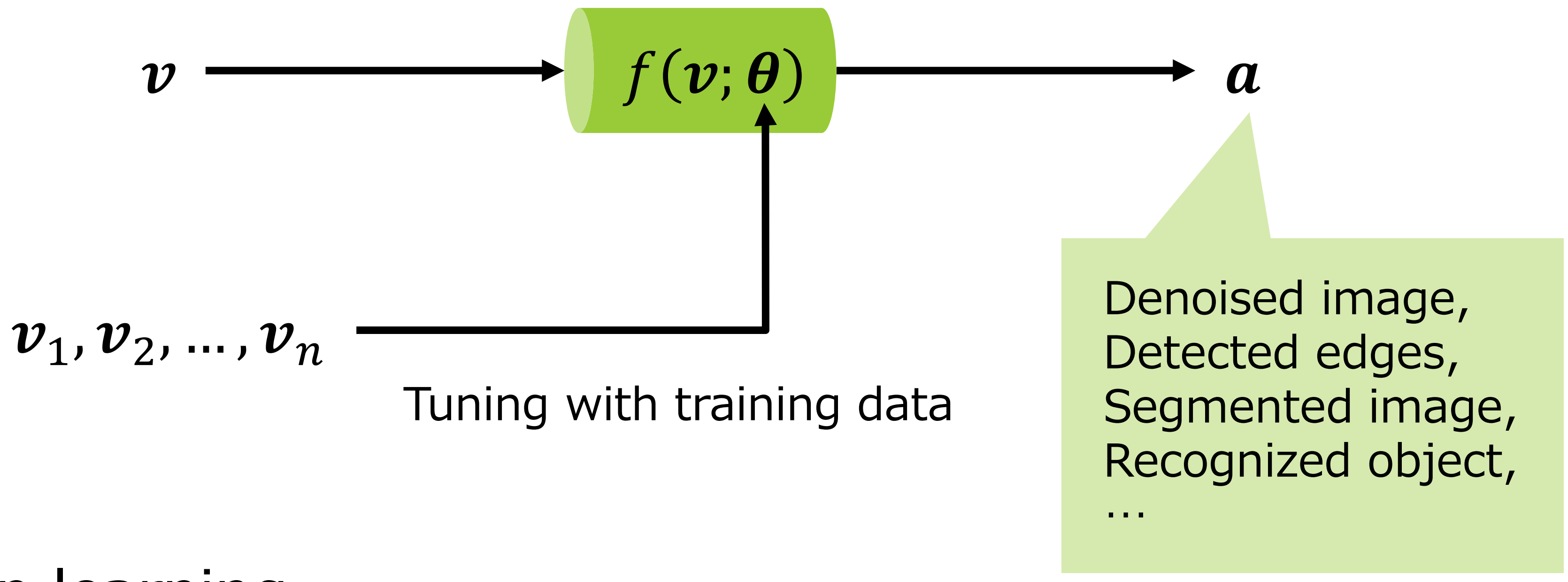
Two types of image processing

- Image processing without statistical models
  - Each pixel value v_{xy} is just an integer or real number.
  - Filtering, Deep learning, ...

- Image processing with statistical models
  - Each pixel value v_{xy} is a realized value of random variable V_{xy}.
  - Ising model, Hidden Markov model, ...
  - This study

Image processing without statistical models

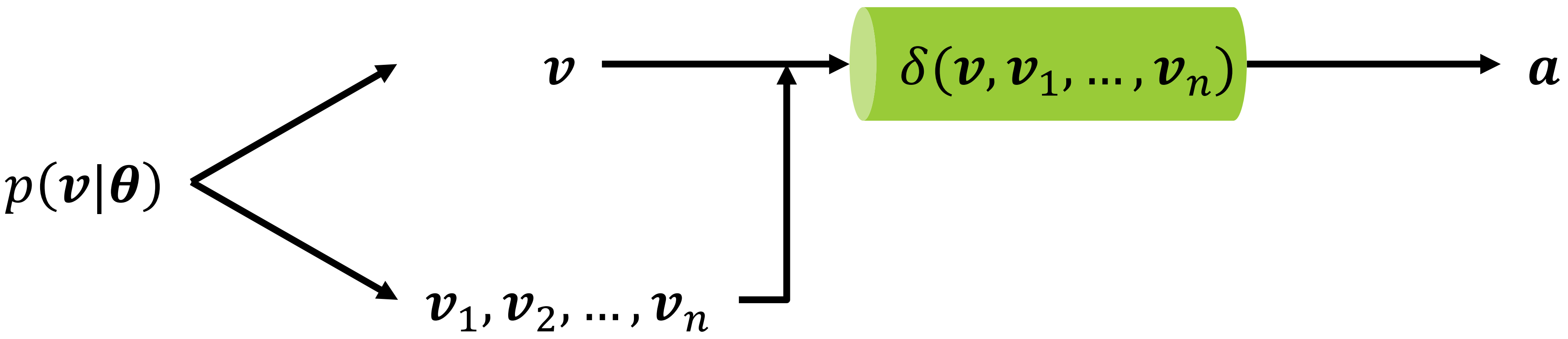
- Mathematical model



- Ex. Filtering, Deep learning, ...
- Practically, we have enough performance.
- Theoretically, it lacks statistical optimality or guarantee.

Image processing with statistical models

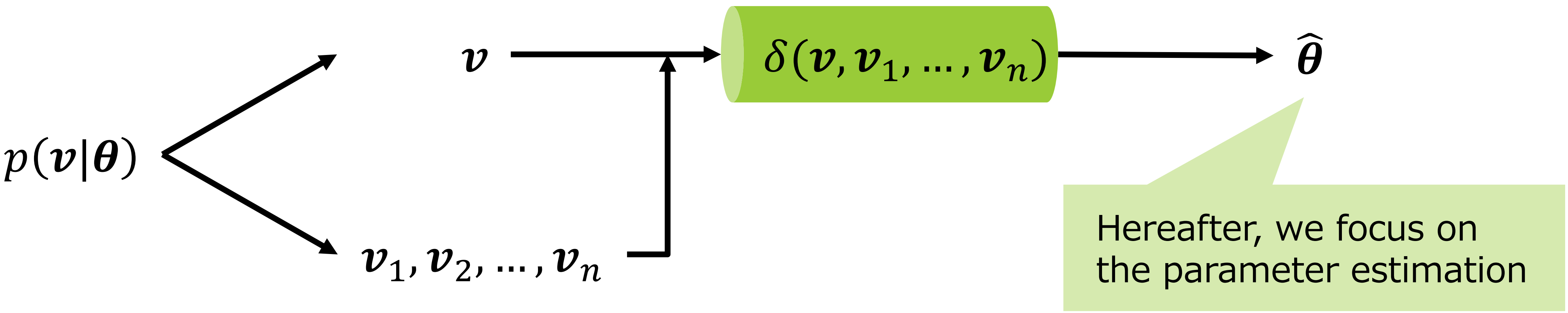
- Mathematical model



- Ex. Ising model, Hidden Markov model, ...
- We can discuss statistical optimality or guarantee of $\delta(\cdot)$ to $p(\boldsymbol{v}|\boldsymbol{\theta})$
- The main concern: Flexibility of model $\longleftrightarrow$ Efficiency of learning

Image processing with statistical models

- Mathematical model



- Ex. Ising model, Hidden Markov model, ...
- We can discuss statistical optimality or guarantee of $\delta(\cdot)$ to $p(\boldsymbol{v}|\boldsymbol{\theta})$
- The main concern: Flexibility of model $\longleftrightarrow$ Efficiency of learning

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

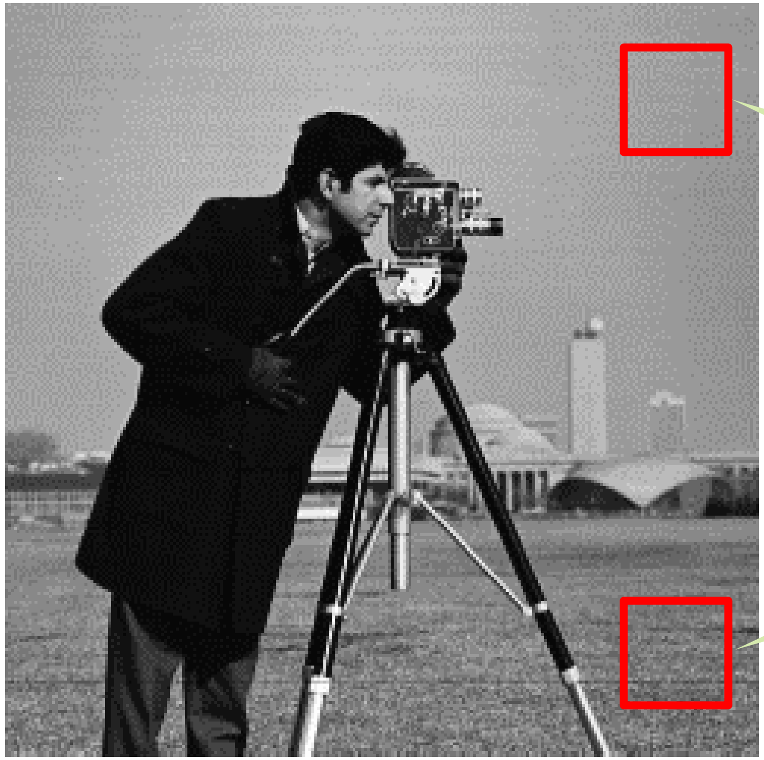
Global and local structure of images

■Local structure

Neighboring pixels have similar values

Global and local structure of images

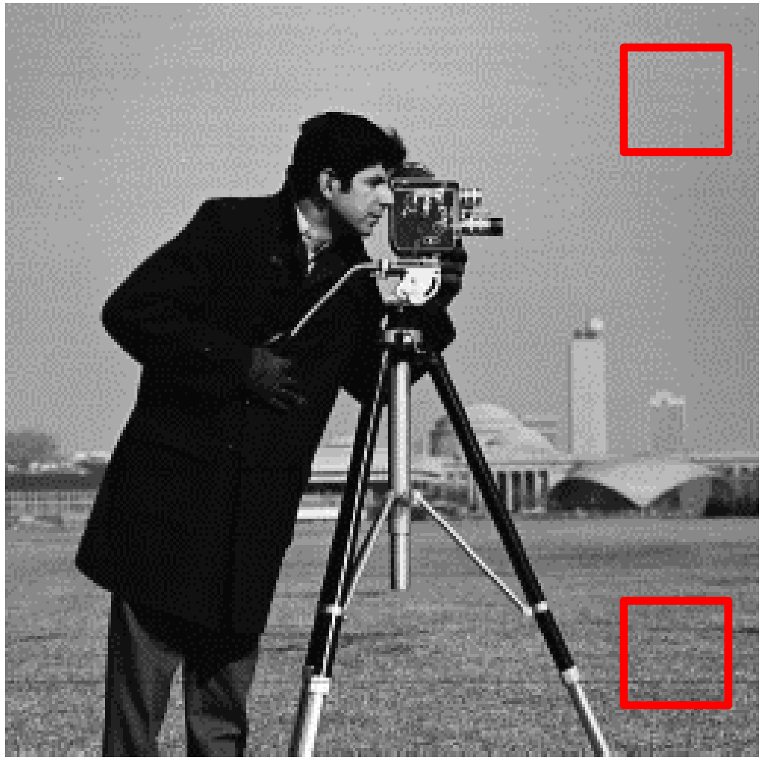
- Global structure



Pixel values have different properties
in regions at a distance from each other

Global and local structure of images

■Hierarchical representation



$m \in \mathcal{M}$: a model for the global structure
$\boldsymbol{\theta}^m \in \boldsymbol{\Theta}^m$: parameters for the local structures

$$m \sim p(m)$$

$$\boldsymbol{\theta}^m \sim p(\boldsymbol{\theta}^m | m)$$

$$\boldsymbol{v} \sim p(\boldsymbol{v} | \boldsymbol{\theta}^m, m)$$

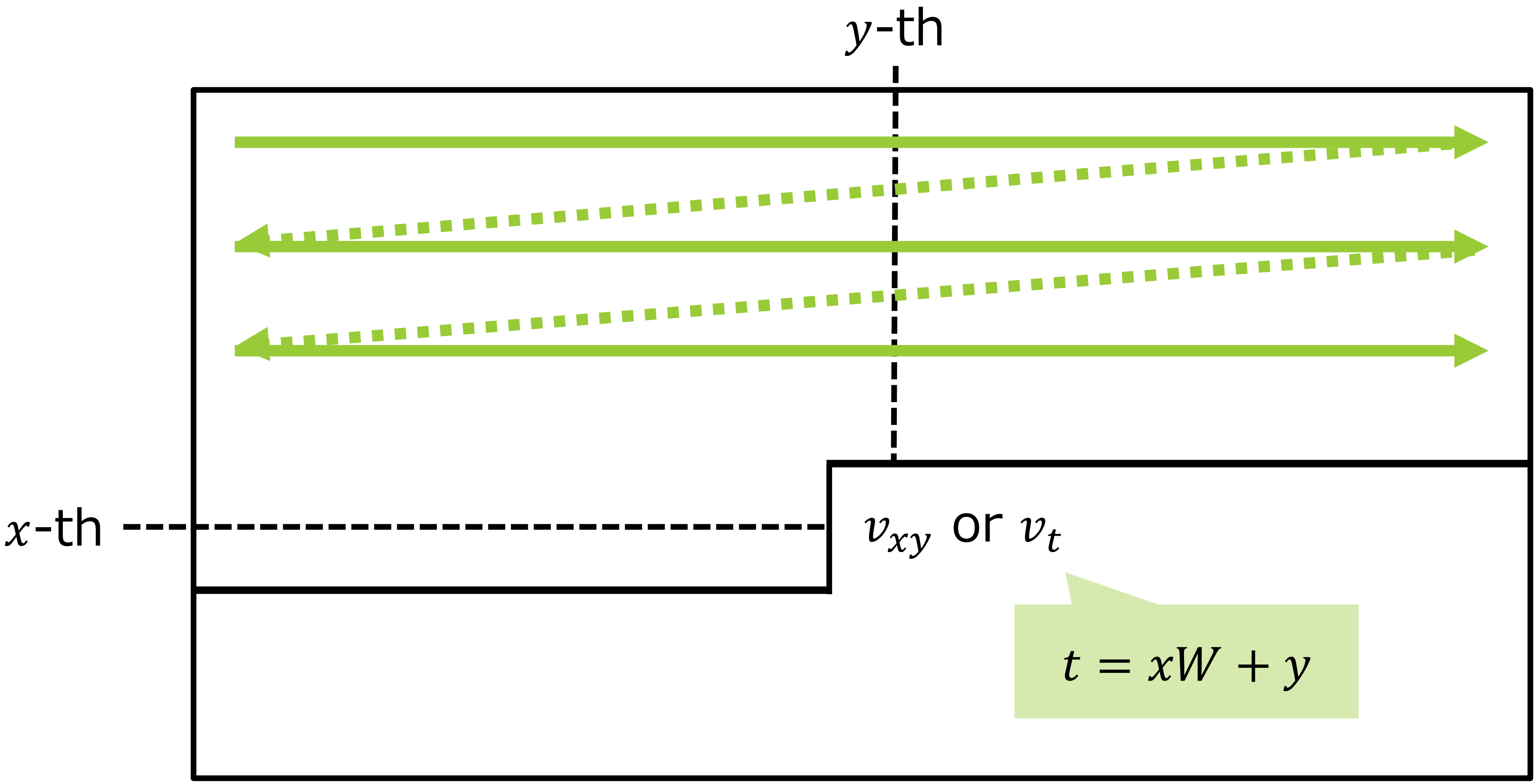
Hierarchical Bayesian statistical model

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

Two-dimensional auto-regressive model

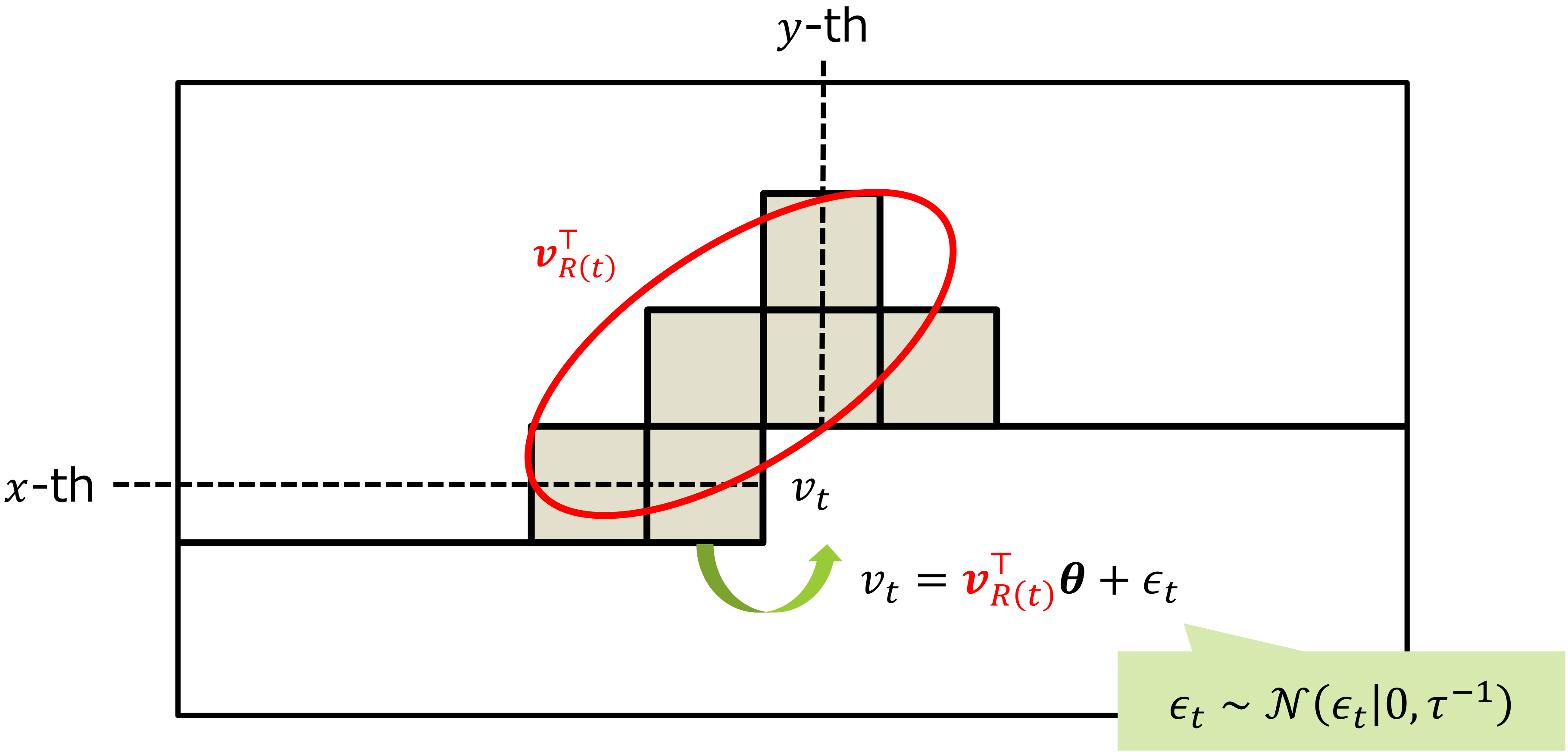
- Generative model [Nakahara et al., 2020]



(More precisely, v_{xy} is normalized and quantized. See our paper.)

Two-dimensional auto-regressive model

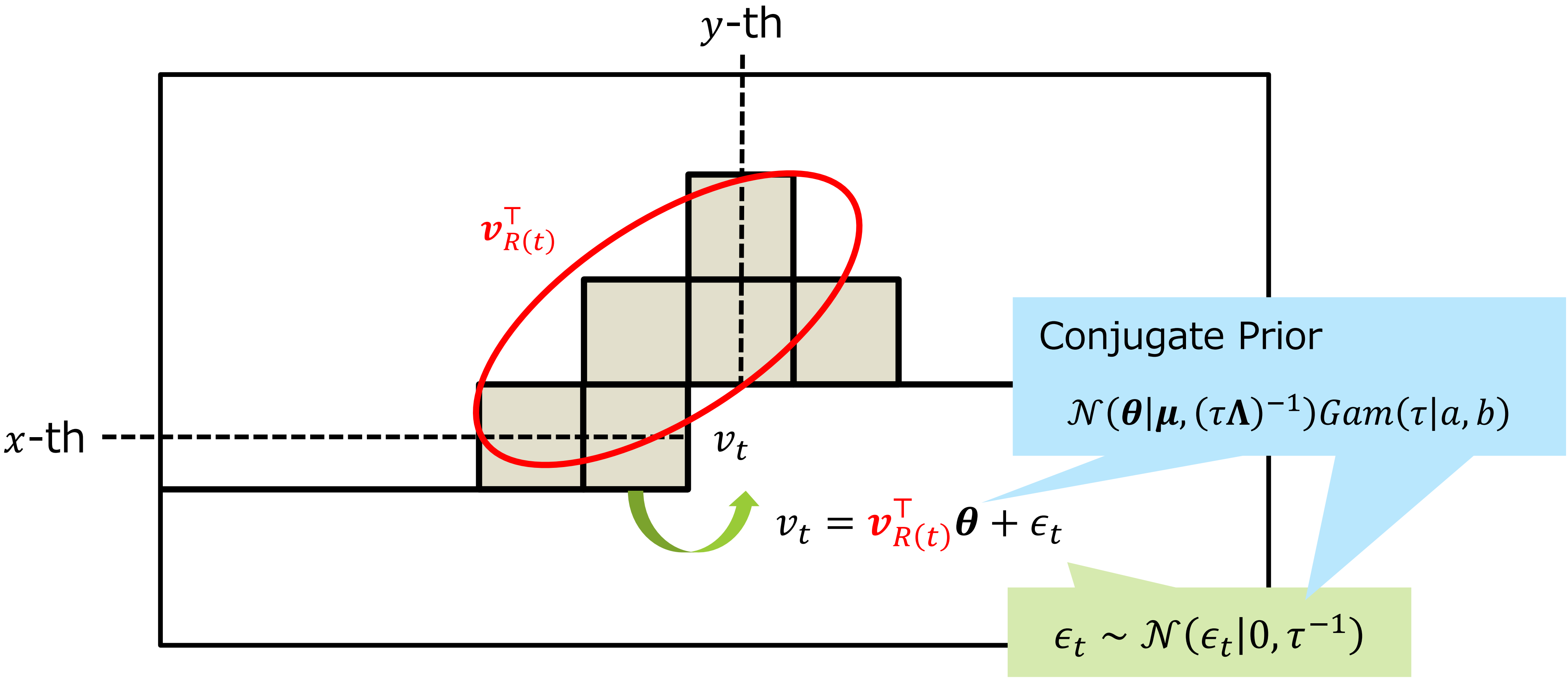
- Generative model [Nakahara et al., 2020]



(More precisely, v_{xy} is normalized and quantized. See our paper.)

Two-dimensional auto-regressive model

- Generative model [Nakahara et al., 2020]



(More precisely, v_{xy} is normalized and quantized. See our paper.)

Parameter estimation

- The prior $p(\boldsymbol{\theta}, \tau)$ is assumed to be normal-gamma distribution.
- The posterior also becomes normal-gamma distribution.

$$p(\boldsymbol{\theta}, \tau | v^t) = \mathcal{N}(\boldsymbol{\theta} | \boldsymbol{\mu}_t, (\tau \boldsymbol{\Lambda}_t)^{-1}) Gam(\tau | a_t, b_t),$$

where

$$\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}_{t-1} + \boldsymbol{v}_{R(t)} \boldsymbol{v}_{R(t)}^\top$$
$$\boldsymbol{\mu}_t = \boldsymbol{\Lambda}_t^{-1} \left(v_t \boldsymbol{v}_{R(t)} + \boldsymbol{\Lambda}_{t-1} \boldsymbol{\mu}_{t-1} \right)$$
$$a_t = a_{t-1} + \frac{1}{2}$$
$$b_t = b_{t-1} + \frac{1}{2} \left(-\boldsymbol{\mu}_t^\top \boldsymbol{\Lambda}_t \boldsymbol{\mu}_t + v_t^2 + \boldsymbol{\mu}_{t-1}^\top \boldsymbol{\Lambda}_{t-1} \boldsymbol{\mu}_{t-1} \right)$$

- For example, we can estimate $\boldsymbol{\theta}$ and τ by

$$\widehat{\boldsymbol{\theta}} = \boldsymbol{\mu}_t, \qquad \hat{\tau} = \frac{a_t}{b_t}$$

Conclusion of the model for local structure

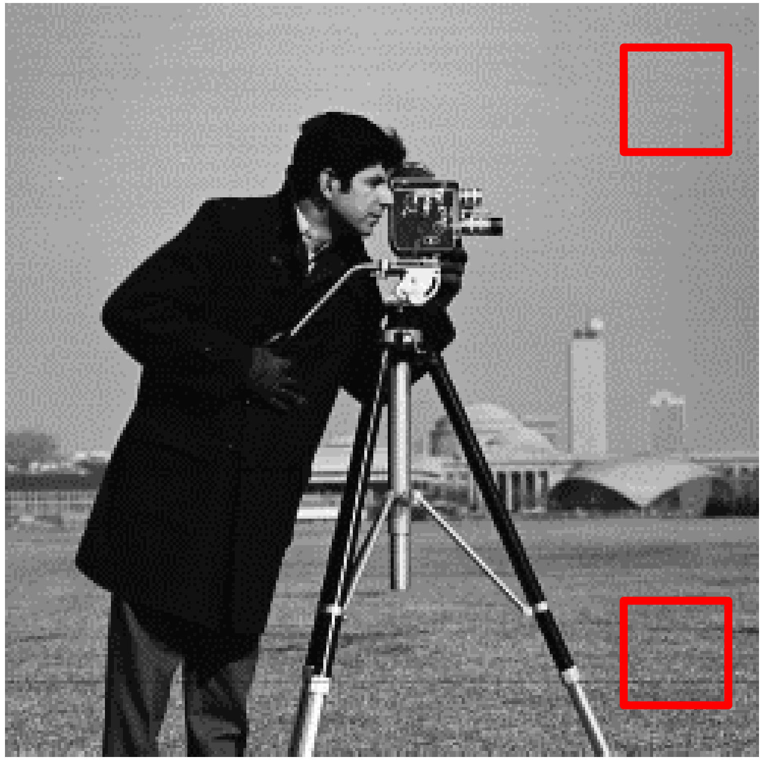
- Pixel values are indexed in raster scan order.
- Explanatory variables are selected from spatially neighboring pixels
- We can estimate the parameters $\boldsymbol{\theta}$ and τ in a similar manner with the usual Bayesian linear regression model.

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

Global and local structure of images

■Hierarchical representation



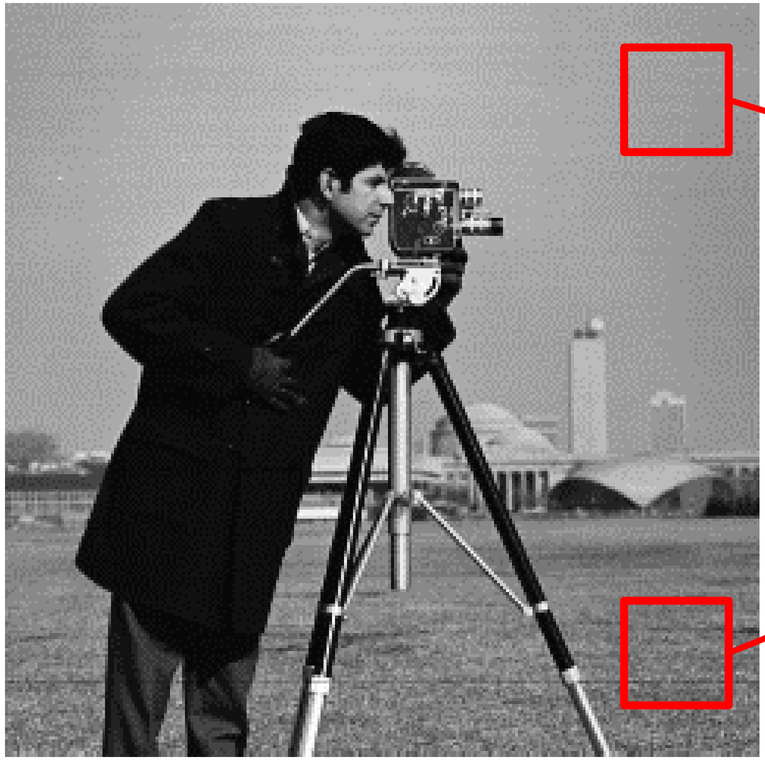
$m \in \mathcal{M}$: a model for the global structure
$\boldsymbol{\theta}^m \in \boldsymbol{\Theta}^m$: parameters for the local structures

$$m \sim p(m)$$
$$\boldsymbol{\theta}^m \sim p(\boldsymbol{\theta}^m | m)$$
$$\boldsymbol{v} \sim p(\boldsymbol{v} | \boldsymbol{\theta}^m, m)$$

Hierarchical Bayesian statistical model

Our idea

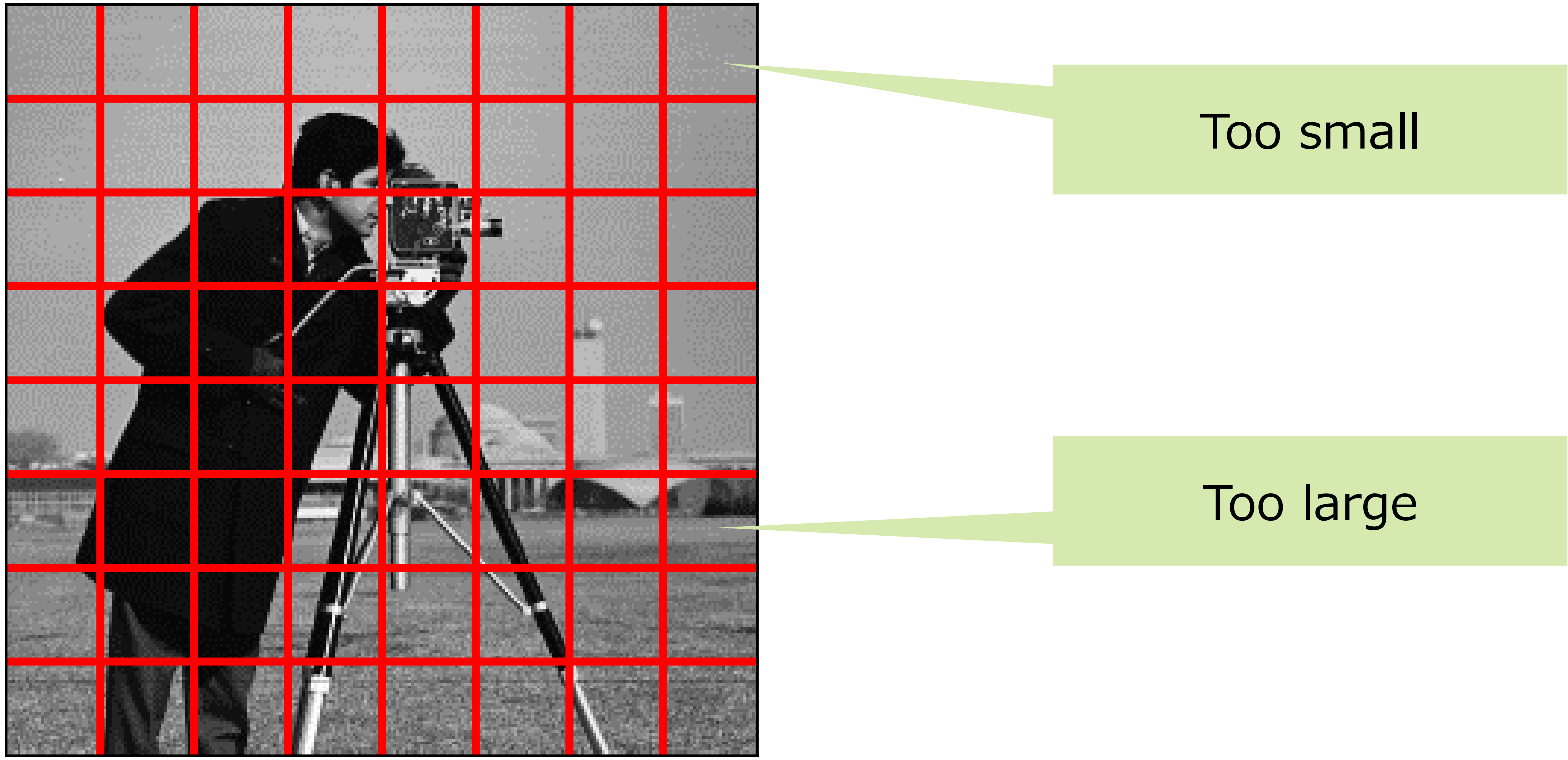
- Each smooth region can be represented by two-dimensional auto-regressive model.



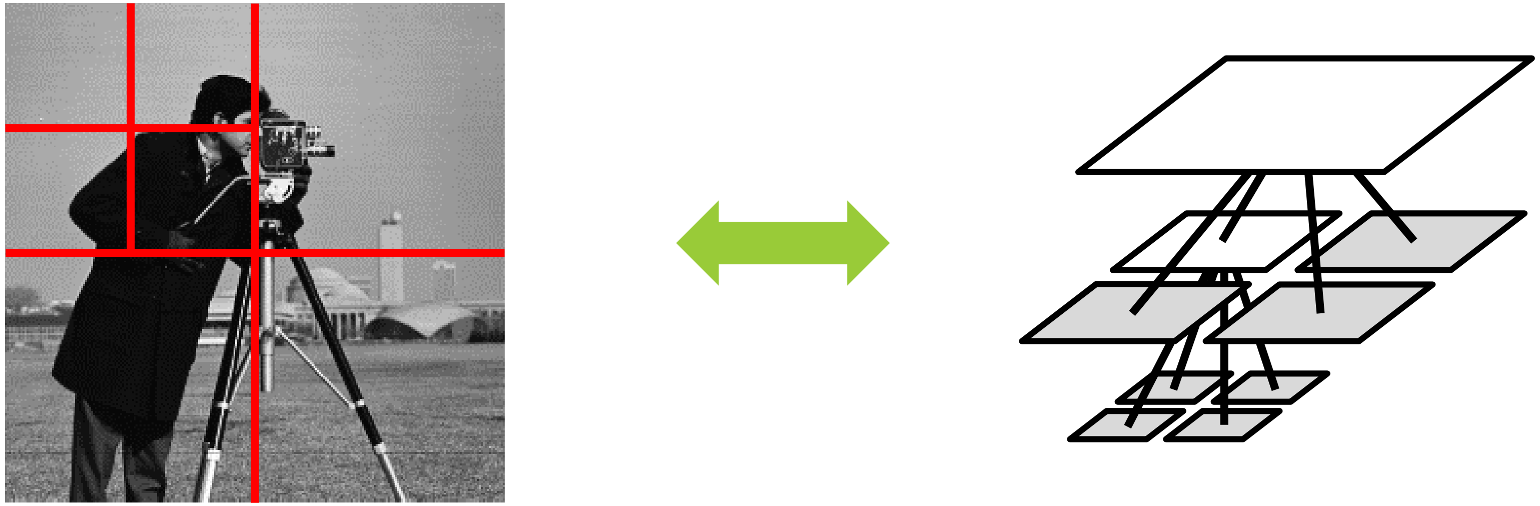
$$p(\boldsymbol{v}|\boldsymbol{\theta},\tau)$$
$$\neq$$
$$p(\boldsymbol{v}|\boldsymbol{\theta}',\tau')$$

Trivial way

- Divide the image into fixed size blocks
- Assume different stochastic models to them



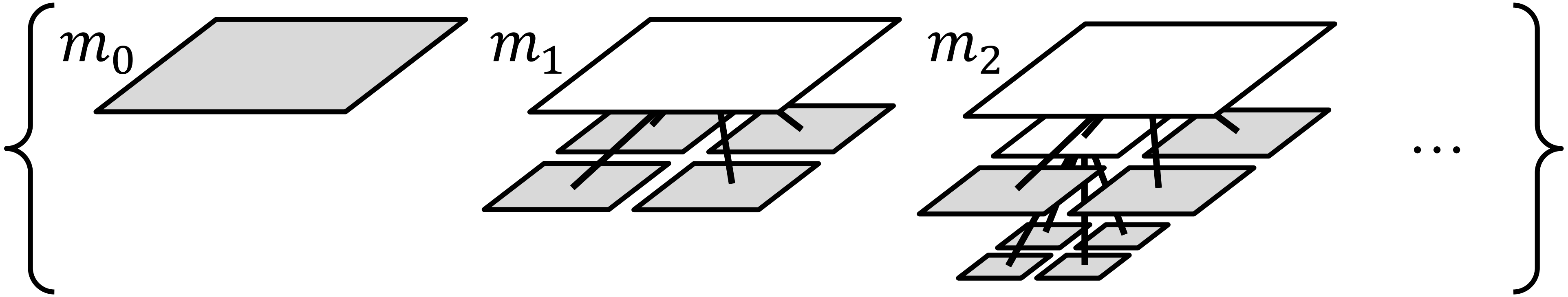
Quadtree

- It effectively represents variable size block segmentation
- 
- Previous studies regard it just as a procedure
- In this study, we regard the quadtree as a part of stochastic generative model of the images.

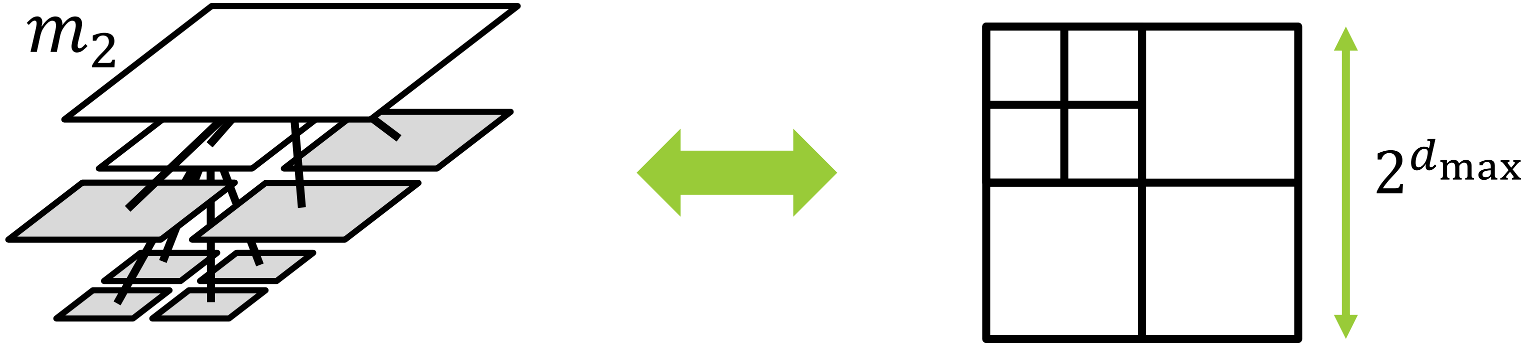
Proposed stochastic model [Nakahara et al., 2020]

- Let both width and height are $2^{d_{\max}}$.
- Consider the set of the quadtrees whose depth $\leq d_{\max}$.

$\mathcal{M} =$

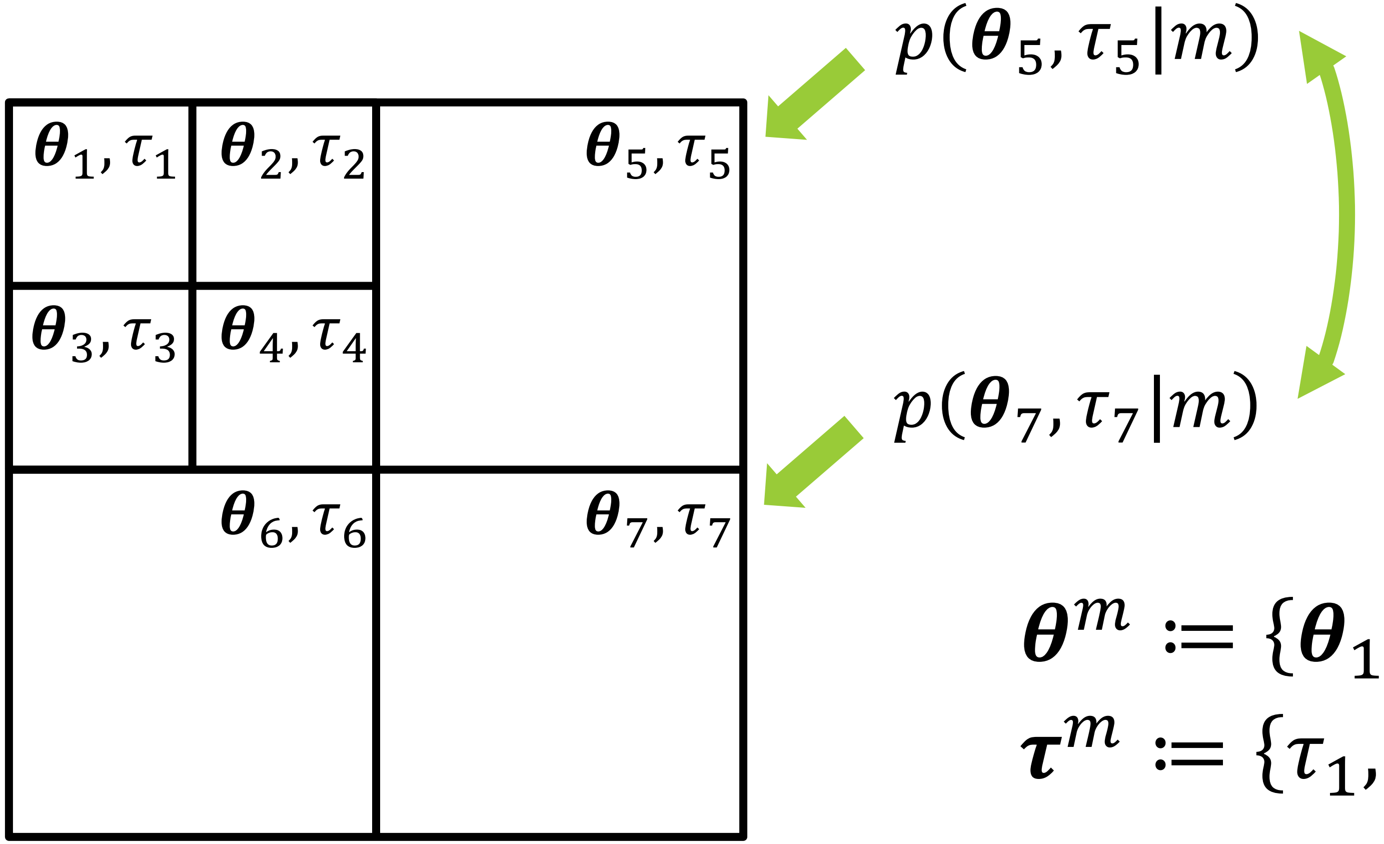


- One of them is chosen with probability $p(m)$.



Proposed stochastic model [Nakahara et al., 2020]

- Parameters $\boldsymbol{\theta}_s$ and τ_s are independently assigned to each block s with probability $p(\boldsymbol{\theta}_s, \tau_s | m)$.

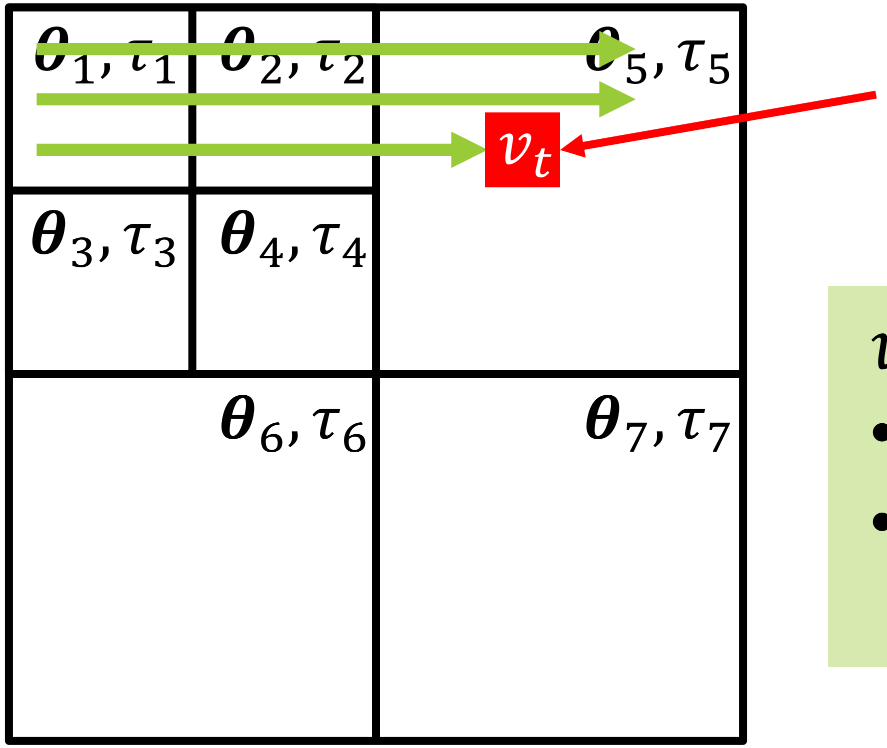


$$\boldsymbol{\theta}^m := \{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_7\}$$

$$\boldsymbol{\tau}^m := \{\tau_1, \tau_2, \dots, \tau_7\}$$

Proposed stochastic model [Nakahara et al., 2020]

- Pixel value v_t at block s is generated in order of the raster scan with probability $p(v_t|v^{t-1},\boldsymbol{\theta}_s,\tau_s,m)$.



v_t depends only on
- The past sequence v^{t-1}
- The parameters $\boldsymbol{\theta}_s$ and τ_s of the block s which contains v_t

Prior distribution

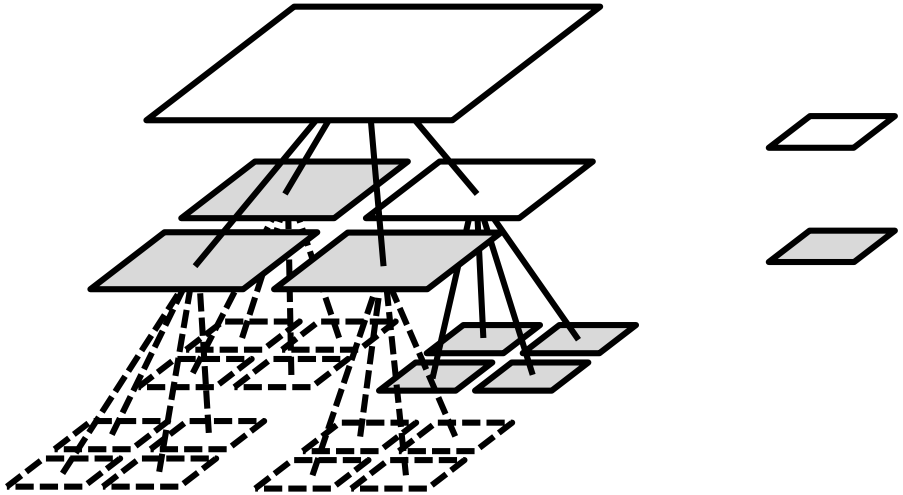
- We estimate m by MAP estimation.

Probability that block s is divided

Prior of models

$$p(m) = \prod_{s \in \mathcal{L}^m} (1 - g_s) \prod_{s \in \mathcal{I}^m} g_s$$

$\mathcal{L}^m$: Leaf node of m
$\mathcal{I}^m$: Inner node of m
$g_s \in [0, 1]$: Hyperparameter of s



: divided with probability g_s

: NOT divided with probability $1 - g_s$

Prior distribution

- We estimate m by MAP estimation.

Probability that block s is divided

Prior of models

$$p(m) = \prod_{s \in \mathcal{L}^m} (1 - g_s) \prod_{s \in \mathcal{I}^m} g_s$$

$\mathcal{L}^m$: Leaf node of m
$\mathcal{I}^m$: Inner node of m
$g_s \in [0, 1]$: Hyperparameter of s

This is known to be conjugate prior of $p(v^t | \boldsymbol{\theta}^m, m)$ [Matsushima et al., 2009]

Posterior of models

$$p(m | v^t) = \prod_{s \in \mathcal{L}^m} (1 - g_{s|t}) \prod_{s \in \mathcal{I}^m} g_{s|t}$$

$\mathcal{L}^m$: Leaf node of m
$\mathcal{I}^m$: Inner node of m
$g_{s|t} \in [0, 1]$: Hyperparameter of s

Hyperparameter updating

- $g_{s|t}$ can be updated as follows. It takes only $O(d_{\max})$.

Prior

(Marginalized) likelihood

$$g_{s|t} = \frac{g_{s|t-1}\, q(v_t|v^{t-1}, s_{\text{child}})}{q(v_t|v^{t-1}, s)}$$

Normalization constant

s and s_{child} are on the path from the root node to the node corresponds to v_t

Predictive distribution of the 2D-AR model

$$q(v_t|v^{t-1}, s) = \begin{cases} St(v_t|\eta_t, \lambda_t, \nu_t), & |s| = 1 \\ (1 - g_{s|t-1}) St(v_t|\eta_t, \lambda_t, \nu_t) + g_{s|t-1} q(v_t|v^{t-1}, s_{\text{child}}), & \text{otherwise} \end{cases}$$

Maximization of posterior

Posterior of models

$$p(m|v^t) = \prod_{s \in \mathcal{L}^m} (1 - g_{s|t}) \prod_{s \in \mathcal{I}^m} g_{s|t}$$

$\mathcal{L}^m$: Leaf node of m
$\mathcal{I}^m$: Inner node of m
$g_{s|t} \in [0, 1]$: Hyperparameter of s

- This posterior can be maximized by dynamic programing.

$$\max_m p(m|v^t) = \phi_t(s_\lambda)$$

$$\phi_t(s) = \begin{cases} 1, & |s| = 1 \\ \max\{1 - g_{s|t}, g_{s|t}\phi_t(s_{\mathrm{child}_1})\phi_t(s_{\mathrm{child}_2})\phi_t(s_{\mathrm{child}_3})\phi_t(s_{\mathrm{child}_4})\}, & \text{otherwise} \end{cases}$$

- $m^{\mathrm{MAP}} = \arg\max p(m|v^t)$ can be found by backtracking after the above maximization with flag variables.

Conclusion of the model for Global structure

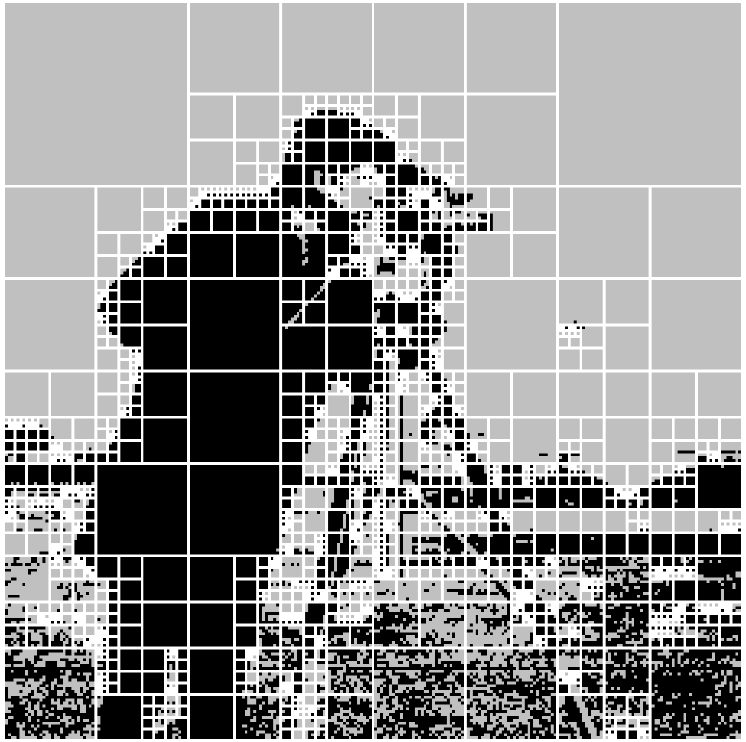
- We interpreted the quadtree from just a procedure into the stochastic model.
- Using the conjugate prior over the quadtrees, posterior distribution can be calculated by updating the hyperparameters. It takes only $O(d_{\max})$.
- Posterior can be maximized by dynamic programing.

Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

Numerical experiments

- Model and parameter estimation for benchmark images



Outline

- Introduction
  - What is image processing?
  - Two types of image processing
- Hierarchical statistical models for images
- A statistical model for local structure of images
- A statistical model for global structure of images
- Numerical experiments
- Conclusion and future works

Conclusion and future works

- Conclusion
  - I introduced an idea to represent the structure of images by hierarchical statistical models
  - The 2D-AR model is for the local structure
  - The model on the quadtree is for the global structure

- Future works
  - Our proposed model can be used for other problems
    - Feature extraction
    - Image generation
    - Image inpainting

Reference

- For local structure
  - Y. Nakahara and T. Matsushima, “Autoregressive Image Generative Models with Normal and t-distributed Noise and the Bayes Codes for Them,” 2020 International Symposium on Information Theory and Its Applications (ISITA), Oct. 2020.
- For global structure
  - Y. Nakahara and T. Matsushima, “A Stochastic Model of Block Segmentation Based on the Quadtree and the Bayes Code for It,” 2020 Data Compression Conference (DCC), pp. 293-302, Mar. 2020.
- For conjugate prior on trees
  - T. Matsushima and S. Hirasawa, “Reducing the space complexity of a Bayes coding algorithm using an expanded context tree," 2009 IEEE International Symposium on Information Theory (ISIT), pp. 719-723, June 2009.