



Title	マルコフ確率場を用いた動画画像処理における移動速度場と確率モデルの同時ベイズ推定
Author(s)	稲垣, 佑哉
Citation	北海道大学. 修士(情報科学)
Issue Date	2010-03-25
Doc URL	<a href="http://hdl.handle.net/2115/43999">http://hdl.handle.net/2115/43999</a>
Type	theses (master)
File Information	inagaki2009.pdf



[Instructions for use](#)

マルコフ確率場を用いた動画像処理における移動速度場と  
確率モデルの同時ベイズ推定

稲垣 佑哉

北海道大学 大学院情報科学研究科 複雑系工学講座 混沌系工学研究室

平成 21 年度 修士論文

# 目次

<b>1</b>	<b>はじめに</b>	<b>3</b>
<b>2</b>	<b>マルコフ確率場と平均場近似</b>	<b>4</b>
2.1	移動速度ベクトル場推定の確率モデル	4
2.1.1	バイズの公式と事後確率	4
2.2	ライン場とセグメンテーション場	5
2.2.1	ライン場とその定義	5
2.2.2	セグメンテーション場とその定義	5
2.2.3	バイズの公式の修正	6
2.3	尤度関数：順過程とその確率モデル	6
2.4	事前確率：未知確率場の生成モデル	7
2.5	エネルギー関数とその構成要素の物理的意味	7
2.6	平均場方程式とその導出	8
2.6.1	MAP 推定と MPM 推定	8
2.6.2	自由度に関する和の指数関数的爆発と平均場近似	9
2.7	ハイパーパラメータ：確率モデルのマクロな特徴付け	11
<b>3</b>	<b>予備実験：平均場の発散とその回避</b>	<b>11</b>
3.1	正則化項のスケーリング係数	11
3.2	統計的性能評価に関する指標	12
3.3	計算機実験の結果	13
3.3.1	自乗誤差/ビット誤り率のスケーリング係数依存性	13
3.3.2	エネルギーの時間発展	13
3.3.3	ノイズの影響	14
<b>4</b>	<b>周辺尤度最大化規準とハイパーパラメータ推定</b>	<b>14</b>
4.1	カルバック・ライブラ情報量の非負性と周辺尤度の最適性	16
4.2	周辺尤度の平均場近似	16
<b>5</b>	<b>確率場とハイパーパラメータの同時推定</b>	<b>17</b>
5.1	ボルツマン・マシン学習方程式	17
5.1.1	ボルツマン・マシン学習方程式の平均場近似と MCMC 法によるハイブリッド解法	18
5.1.2	マルコフ連鎖モンテカルロ法による学習方程式のサンプリング計算	20
5.2	実験結果 1：学習方程式の平均場近似と MCMC 法によるハイブリッド解法	21
5.3	実験結果 2：MCMC 法による学習方程式のサンプリング計算	23
<b>6</b>	<b>おわりに</b>	<b>24</b>

# 1 はじめに

画像工学における物体移動速度ベクトルの推定の技術は、異なる2フレーム間の物体の変位の推定、動画の圧縮などに応用される。この物体移動速度ベクトルの推定は、その解が一意でないため不良設定問題に属する。例えば、画素値が単色である領域においては、単なる画素値のマッチングでは一意に解が決定できない。また、画像中のノイズや歪みによって解が不安定になることもある。

この不良設定問題を解く既存方法として *Smoothness constrains* (連続性による制約) [1] や *Discontinuity constrains* (不連続性による制約) [2] など、物体移動速度ベクトルに各種制約条件を導入する手法 — 正則化理論 — が広く用いられている。そのような標準正則化理論の一つとして位置づけられるマルコフ確率場によるアプローチにおいて「移動物体の速度ベクトル推定」は最小エネルギーを探索する組み合わせ最適化問題とみなすことができる。最小エネルギーの主な探索手法として、**シミュレーテッド・アニーリング法 (SA 焼きなまし法)** や**反復条件付き最大化法**が知られている。しかし、SAは大域的最適解にたどり着くことが証明されているが、膨大な計算時間を要する。一方、反復条件付き最大化法は高速で解を見つけることができるが、その解が必ずしも高精度とはいえない。

これらの問題を解決するために、Zhang ら [3] によって提案された**マルコフ確率場**に基づく平均場解法がある。この手法は以下にあげる2つの大きな特徴をもつ。

- 移動速度ベクトルに連続性/非連続性の制約を導入。
- 画像復元/分割やセンサーフュージョンにおいて有用である**平均場近似**の導入。

これら2つの特徴によって、Zhang らの手法は高速/高精度な解の探索が可能とされている。

しかし、マルコフ確率場モデルに対し、平均場近似に基づく反復解法を構成した場合、正則化項の不適切な規格化に起因する「平均場の発散」が生じるといった問題が生じる。また、エネルギー関数に確率モデルをマクロに特徴つける**ハイパーパラメータ**が複数組み込まれているが、それらの値が何ら理論的根拠もなく「アド・ホック」に設定されているといった問題がある。画像は、物体や背景の模様や、ノイズなどによって各々特性が異なるものである。そのような特性の異なる画像に対して一意に固定されたパラメータでそれぞれ最適な解を得ることは難しい。

そこで本研究では、正則化項に新たなハイパーパラメータとなるスケーリング係数を導入し、発散回避を試みる。さらに、アド・ホックに設定されているハイパーパラメータの値を各画像に対して適切に推定することで、より高精度になることを目標とする。特に、周辺尤度に関するハイパーパラメータの勾配で構成される**ボルツマン・マシン学習方程式**を平均場近似、および、**マルコフ連鎖モンテカルロ法**を用いて数値的に処理することで、ハイパーパラメータと確率場を同時に推定する方式を提案し、その性能を「精度」「計算時間」の観点から詳細に評価する。

本稿の構成は以下の通りである。第2節では我々の移動速度推定をマルコフ確率場でモデル化し、事後確率最大化推定、および周辺化事後確率推定を平均場近似のもとで定式化する。続く第3節では、正則化項の不適切な規格化に起因する平均場変数の発散が生じることを示し、この「発散」を回避するため、正則化項にスケーリング係数を導入し、これを平均自乗誤差最小化規準のもとで最適化する。得られる最適値に対し、エネルギー、自乗誤差、ビット誤り率の時間発展を平均場方程式の反復に対して調べ、同時にノイズの影響についても議論する。第4節では第3節で求めたスケーリング係数値を周辺尤度最大化規準の観点から再考する。第5節では周辺尤度最大化規準を用いてボルツマン学習方程式を導出し、平均場近似、マルコフ連鎖モンテカルロ法、オイラー法などを用い

てハイパーパラメータの適切な値を確率場と同時に推定する。それら結果についての評価を多角的に行う。最後の第6節はまとめである。

## 2 マルコフ確率場と平均場近似

この節では本論文で扱う動画像を時刻ラベル付きの静止画像の集合として定義する。また、異なるフレーム間での物体の「移動方向」とその「幅」を速度ベクトルとして導入し、静止画像の各画素ごとにその速度ベクトルを割り当てた場合の速度ベクトルの集合を「速度場」として定義する。2フレーム分の静止画像があたられた場合の移動速度場の推定は「不良設定」であるため、そのままでは解くことができない。そこで、ここでは、Zhang *et al* [3] に倣い、ベイズ推定の枠組みを適用し、事前確率として拘束条件を導入することで移動速度ベクトルをベイズ推定するための枠組みを定式化する。

### 2.1 移動速度ベクトル場推定の確率モデル

本論文で我々は動画処理を行うが、動画を静止画像が時間軸上に連続して並ぶものとして定義する。従って、2次元正方格子上で定義される静止画像  $\mathbf{x} = \{x_i, i \in S\}$  ( $S$  は一枚の静止画像に含まれる画素集合。画像の「番地」を与える添字  $i$  は正方格子上の座標  $(x, y)$  と一対一に対応しているものとする) にもう1次元分、時刻を表す添字を与える。具体的にはその時刻を  $\tau$  で区別し、時刻  $\tau$  での画像画素の集合を  $\mathbf{x}^\tau = \{x_i^\tau, i \in S\}$  で表す。

ところで、動画の連続する2フレーム間で各画素の位置を比較すると、このフレーム間で移動が観測される物体を構成する画素はその番地変えているはずであるから、各画素ごとに時刻  $\tau-1$  と  $\tau$  で指定される2フレーム間での「移動幅」と「向き」を定義しなければならない。そこで、ここでは時刻  $\tau$ 、画素  $i$  における変位の速度ベクトル  $d_i^\tau$  とその集合を  $\mathbf{d}^\tau = \{d_i^\tau, i \in S\}$  で導入する。つまり、各画素  $i$  に対し、連続する2フレーム内で以下のような関係が成り立つ。

$$x_i^\tau = x_{i-d_i^\tau}^{\tau-1} \quad (1)$$

ここで、画素の番地を表す添字  $i$  は2次元座標  $(x, y)$  と一対一に対応していることに連動し、 $d_i^\tau$  も2次元速度場  $(v_x^\tau(i), v_y^\tau(i))$  と一対一に対応しているものとする。このように、 $d_i^\tau$  は2次元ベクトルであり、本研究ではこのベクトルの成分のもつ値の最大値を5とする。

#### 2.1.1 ベイズの公式と事後確率

上記のように動画を構成する各フレームでの静止画像、および、連続する2フレーム間での物体移動を各画素ごとに定義される移動速度ベクトルの集合 — 移動速度場 — で表すことにすると、静止画像が2フレーム分与えられた場合、その2フレーム間での物体の移動速度場を未知変数とし、これを統計的に推定する問題として定式化することができる。本論文では、この未知変数の推定を**移動速度場の推定**と呼ぶ。

このとき、動画中の連続する2フレーム分の画像、 $\mathbf{x}^\tau$  と  $\mathbf{x}^{\tau-1}$  が与えられたもとでの移動速度ベクトル  $\mathbf{d}^\tau$  を推定する  $r$  問題に対し、確率論的方法を用いれば、事後確率  $P(\mathbf{d}^\tau | \mathbf{x}^\tau, \mathbf{x}^{\tau-1})$  は**ベイズの公式**から

$$P(\mathbf{d}^\tau | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) = \frac{P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau | \mathbf{x}^{\tau-1})}{\sum_{\mathbf{d}^\tau} P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau | \mathbf{x}^{\tau-1})} = \frac{P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau | \mathbf{x}^{\tau-1})}{P(\mathbf{x}^\tau | \mathbf{x}^{\tau-1})} \quad (2)$$

が得られるが,  $P(\mathbf{x}^{\tau+1}|\mathbf{x}^{\tau+1})$  は速度場  $\mathbf{d}^{\tau}$  に依存しない定数であることから,  $P(\mathbf{x}^{\tau}|\mathbf{d}^{\tau}, \mathbf{x}^{\tau+1})P(\mathbf{d}^{\tau}|\mathbf{x}^{\tau})$  を最大化する  $\mathbf{d}^{\tau}$  が事後確率最大推定値 (Maximum A Posteriori: MAP) を与える.

## 2.2 ライン場とセグメンテーション場

移動物体の速度ベクトル場推定に対し, 単純な画素値のマッチングでは解が一意に定まらない. この「不良設定問題」を解消するため, ここではライン場 (Line fields) とセグメンテーション場 (Segmentation fields) とそれぞれ呼ばれる確率場を導入する.

### 2.2.1 ライン場とその定義

ライン場とは速度ベクトル間の連続性を保障するものであり, 「隣接している画素の速度ベクトル同士は似ている」という予測に基づいており,  $\mathbf{l} = \{l(i, j) | l(i, j) \equiv (h_i, v_i, h_j, v_j) \in S\}$ , と表す.  $h_i$  は水平方向,  $v_i$  は垂直方向に隣接している画素との速度ベクトルの連続性を表す. 具体的には以下のように定義する<sup>1</sup>.

$$h_i^{\tau} = \begin{cases} 0 & (\text{水平方向に隣接している画素の間で速度ベクトル } \mathbf{d}^{\tau} \text{ が不連続}) \\ 1 & (\text{水平方向に隣接している画素の間で速度ベクトル } \mathbf{d}^{\tau} \text{ が連続}) \end{cases} \quad (3)$$

$$v_i^{\tau} = \begin{cases} 0 & (\text{垂直方向に隣接している画素の間で速度ベクトル } \mathbf{d}^{\tau} \text{ が不連続}) \\ 1 & (\text{垂直方向に隣接している画素の間で速度ベクトル } \mathbf{d}^{\tau} \text{ が連続}) \end{cases} \quad (4)$$

となる (図 1 参照).

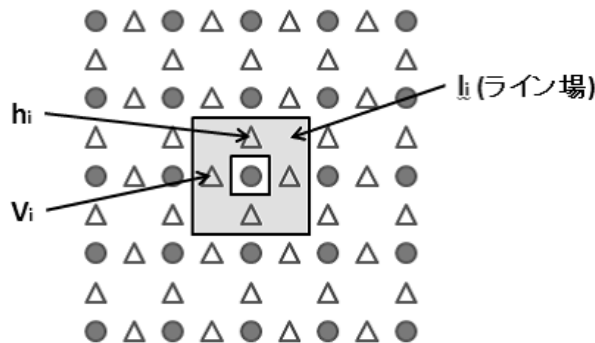


図 1: ライン場の具体的なイメージ.

### 2.2.2 セグメンテーション場とその定義

セグメンテーション場とは, 速度ベクトル場を「予測可能領域」と「予測不可能領域」の2つの集合に分割モデル化するものである. ここで, 予測不可能領域とは移動物体に隠れていた部分であり, 一単位時間手前の時刻の画像において一致する画素が存在しない部分である. そのため,

<sup>1</sup>具体的には  $j$  が  $i$  の水平方向に 1 画素分大きい場合の  $l(i, j)$  を  $l_x(i, i+1)$  と書くことにすれば,  $l_x(i, i+1) = v_i, l_x(i, i-1) = v_j, l_y(i, i+1) = h_i, l_y(i, i-1) = h_j$  で定義する.

その部分に対応する速度ベクトルを推定することはできない。セグメンテーション場を導入することで、推定可能な推定すべき速度ベクトルを把握することができる。セグメンテーション場は  $s = \{s_i | s_i \in \{0, 1\}\}$  と表すことができる。具体的には

$$s_i^\tau = \begin{cases} 0 & \text{(画素 } i \text{ が予測可能な場合)} \\ 1 & \text{(画素 } i \text{ が予測不可能な場合)} \end{cases} \quad (5)$$

として定義する (図 2 参照)。

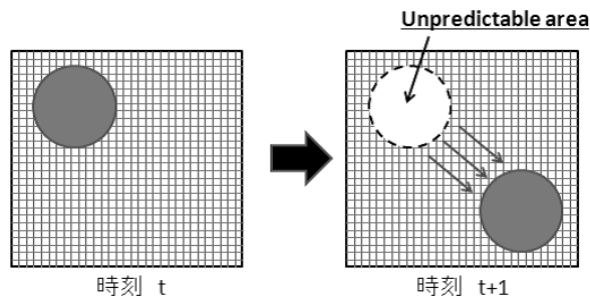


図 2: セグメンテーション場の具体的なイメージ。

### 2.2.3 ベイズの公式の修正

ライン場とセグメンテーション場、これら 2 つの確率場を導入することで、適切な速度ベクトルの推定が可能となる。このとき、ベイズの公式 (2) 式は以下のように修正される。

$$\begin{aligned} P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) &= \frac{P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^{\tau-1})}{\sum_{\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau} P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^{\tau-1})} \\ &= \frac{P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau-1}) P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^{\tau-1})}{P(\mathbf{x}^\tau | \mathbf{x}^{\tau-1})} \end{aligned} \quad (6)$$

ここで、 $P(\mathbf{x}^\tau | \mathbf{x}^{\tau+1})$  は未知変数  $\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau$  に依存しない定数であることから、事後確率最大推定 (MAP 推定) は  $P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau+1}) P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^\tau)$  を最大化する速度場  $\mathbf{d}^\tau$  を求めることになる。ここで、 $P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau+1})$  は尤度、 $P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^\tau)$  は事前確率を表し、物体移動や動画像の性質を考慮した上で適切に確率モデル化しなければならない。以下でそれぞれの確率モデルについてみていくことにする。

## 2.3 尤度関数：順過程とその確率モデル

尤度  $P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau-1})$  は時刻  $\tau-1$  での静止画像 (フレーム)  $\mathbf{x}^{\tau-1}$ 、及び、速度場  $\mathbf{d}^\tau$ 、ライン場  $\mathbf{l}^\tau$ 、セグメンテーション場  $\mathbf{s}^\tau$  が与えられた条件下での次時刻  $\tau$  での静止画像 (フレーム)  $\mathbf{d}^\tau$  の生成確率を表すものであり、事後確率によって速度場を推定する問題を「逆問題」、事後確率を「逆モデル」とすれば、この尤度はその意味で「順モデル」とみなすことができる。この尤度をここでは次のように仮定する。

$$P(\mathbf{x}^\tau | \mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau, \mathbf{x}^{\tau-1}) \propto \exp \left[ -E^{(1)}(\mathbf{x}^\tau | \mathbf{s}, \mathbf{d}, \mathbf{l}, \mathbf{x}^{\tau-1}) \right] \quad (7)$$

ここで、上記指数関数の肩に負符号をつけた量を「エネルギー関数」とみなすと、 $E^{(1)}(\mathbf{x}^\tau | \mathbf{s}, \mathbf{d}, \mathbf{l}, \mathbf{x}^{\tau-1})$  は次式で与えられる。

$$E^{(1)}(\mathbf{x}^\tau | \mathbf{s}, \mathbf{d}, \mathbf{l}, \mathbf{x}^{\tau-1}) = \frac{1}{2\sigma^2} \sum_i (1 - s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} \quad (8)$$

ここに、 $\mathbf{N}(i)$  は画素  $i$  の隣接画素集合を表す。本研究で対象とする 2 次元正方格子の場合にその濃度は  $|\mathbf{N}(i)| = 4$  となる。また、上記エネルギー関数に現れる  $\sigma, \alpha_l$  はハイパーパラメータであり、確率モデルをマクロに特徴つける変数である。このようなパラメトリックな確率分布を扱う本研究においては、「確率モデルの推定」とはこれらハイパーパラメータの推定を意味するものであることに注意されたい。

## 2.4 事前確率：未知確率場の生成モデル

事前確率  $P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^\tau)$  は、速度ベクトル場、ライン場、および、セグメンテーション場の生成モデルであり、これら未知確率場に確率的な制約条件を課す。この事前確率が

$$P(\mathbf{d}^\tau, \mathbf{s}^\tau, \mathbf{l}^\tau | \mathbf{x}^\tau) \propto \exp \left[ -E^{(2)}(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau) \right] \quad (9)$$

と指数関数的に振る舞うと仮定すると、上記指数関数の肩に負符号をつけた量であるエネルギー関数は以下のような式で書くことができる。

$$\begin{aligned} E^{(2)}(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau) &= \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - d_j|^2})(1 - l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(s_i - s_j)) + T_s \sum_i s_i \end{aligned} \quad (10)$$

ここで、 $\lambda_d, \lambda_s, \lambda_l, \beta_d$ 、および、 $T_s$  は事前確率をマクロに特徴つけるハイパーパラメータである。

## 2.5 エネルギー関数とその構成要素の物理的意味

事後確率  $P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1})$ 、すなわち、2 枚のフレームが与えられた場合の所望の確率変数がとる確率は、先述の尤度  $P(\mathbf{x}^\tau | \mathbf{d}, \mathbf{s}, \mathbf{l}, \mathbf{x}^{\tau-1})$ 、および、事前確率  $P(\mathbf{d}, \mathbf{s}, \mathbf{l} | \mathbf{x}^{\tau-1})$  の積： $P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \propto P(\mathbf{x}^\tau | \mathbf{d}, \mathbf{s}, \mathbf{l}, \mathbf{x}^{\tau-1})P(\mathbf{d}, \mathbf{s}, \mathbf{l} | \mathbf{x}^{\tau-1})$  で与えられる。すなわち

$$\begin{aligned} P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) &\propto \exp \left[ -E^{(1)}(\mathbf{x}^\tau | \mathbf{s}, \mathbf{d}, \mathbf{l}, \mathbf{x}^{\tau-1}) - E^{(2)}(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau) \right] \\ &\equiv \exp \left[ -E(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \right] \end{aligned} \quad (11)$$

であるから、上記事後確率の対数  $-\log P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1})$  で与えられる系全体のエネルギー関数は

$$\begin{aligned} E(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) &= \frac{1}{2\sigma^2} \sum_i (1 - s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - d_j|^2})(1 - l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(s_i - s_j)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} \\ &+ T_s \sum_i s_i \end{aligned} \quad (12)$$



で定義される。

この右辺第1,4項が尤度からの寄与, 第2,3,5項が事前分布からの寄与である。(12)式第1項は画像の任意の一点*i*の移動範囲*d<sub>i</sub>*が連続する2フレーム間で極端に大きくならない条件を確率モデル化したものである。また, 第2項は移動速度間の連続性を保障する項であり,  $\beta_d \rightarrow 0$ の極限で

$$\lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2e^{-\beta_d |d_i - d_j|^2}) \simeq 2\lambda_d \beta_d \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j)) |d_i - d_j|^2 + (\mathbf{d} \text{ に依らない定数項}) \quad (13)$$

となり, これは「強磁性多値イジング模型」のエネルギー関数と一致する。第3項はセグメンテーション場とライン場の相関を, 第4項は同一フレーム内の画素間距離とライン場間の相関, 第5項はセグメンテーション場の個数( $T_s$ が「化学ポテンシャル」)の制御をそれぞれ確率モデル化した事前分布の一部である。従って, ベイズ推定の立場からは $\{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}\}$ が外から与えられる「入力データ」, ベイズ推定すべき確率変数は未知の確率場 $\{\mathbf{s}, \mathbf{d}, \mathbf{l}\}$ である。

## 2.6 平均場方程式とその導出

以上で事後確率, および, MAP推定を行う際のエネルギー関数が構築できた。ここで, 未知の確率場 $\{\mathbf{s}, \mathbf{d}, \mathbf{l}\}$ のそれぞれの自由度を $n_s, n_d, n_l$ とすると, エネルギー関数の最小化問題は超高次元空間の最適化問題となり, その解空間の大きさは

$$n_s^N \times n_d^N \times n_l^N = \exp[N \log(n_s n_d n_l)] \quad (14)$$

のように指数オーダーにまで及ぶ。本研究での計算機実験では具体的に $s_i, l_i = \{0, 1\}$ なので,  $n_s = n_l = 2$ ,  $d_i = \{0, 1, \dots, d_{max}-1\}$ ,  $d_{max} = 5$ なので,  $n_d = 5$ , すなわち,  $20^N$ もの和を計算しなければならない。 $N = 30 \times 30 = 900$ であることを考慮すると, この手の計算はもはや動画画像処理において現実的でないことがわかる。

### 2.6.1 MAP推定とMPM推定

この最適化問題をシミュレーテッド・アニーリングや遺伝的アルゴリズムなどの進化的計算で解くことも可能であるが, ここではMAP解に対応する各確率場の推定値が $\beta$ なる制御変数を導入することで, 注目する変数, 例えば,  $s_i$ に関し

$$\langle s_i \rangle \equiv \lim_{\beta \rightarrow \infty} \frac{\sum_{\{\mathbf{s}, \mathbf{d}, \mathbf{l}\}} s_i \exp[-\beta E(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1})]}{\sum_{\{\mathbf{s}, \mathbf{d}, \mathbf{l}\}} \exp[-\beta E(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1})]} \quad (15)$$

と書けることに注意する<sup>2</sup>。ここで, 上記のゼロ温度極限 $\beta \rightarrow \infty$ をとることなく,  $\beta = 1$ で固定して求まる上記の解(15)を**周辺化事後確率最大推定 (Maximizer of Posterior Marginal: MPM)**と呼び, 各画素ごとに同じエネルギー関数を持つ「アンサンブル (集合)」を用意し, それぞれのアンサンブルで注目する画素における確率場 (例えば,  $s_i$ ) 対し, 各離散値 ( $s_i$  と言えば,  $s_i = 0, \dots, n_s - 1$  の  $n_s$  個の離散値) に「投票」を行い, 最も票を集めた離散値を該当する確率場の値とするものである。事前確率を「一様分布」とした場合, MAP推定は「最尤推定」に, MPM推定は「確率場ごとの最尤推定」に相当する。

<sup>2</sup>これらは離散値であるべきなので適切に量子化する。

## 2.6.2 自由度に関する和の指数関数的爆発と平均場近似

しかし、上記の MAP, MPM 推定のいずれの推定値を採用するにせよ、 $\mathbf{s}, \mathbf{d}, \mathbf{l}$  に関する和:

$$\sum_{\{\mathbf{s}, \mathbf{d}, \mathbf{l}\}} (\dots) \equiv \sum_{\mathbf{s}} (\dots) \sum_{\mathbf{d}} (\dots) \sum_{\mathbf{l}} (\dots) = \sum_{s_1=0}^{n_s-1} \dots \sum_{s_N=0}^{n_s-1} \sum_{d_1=0}^{n_d-1} \dots \sum_{d_N=0}^{n_d-1} \sum_{l_1=0}^{n_l-1} \dots \sum_{l_N=0}^{n_l-1} (\dots) \quad (16)$$

の実行ステップは既に述べたように、画像サイズ  $N$  の増加とともに  $e^{N \log(n_s n_d n_l)}$  のように指数関数的に増大し、現実的でなくなるので、ここでは平均化すべき変数を除きそれ以外はその変数に関する (15) 式の意味での「平均値」で置きなおしてエネルギーを「一体問題化」する。すなわち、 $s_i$  に関してこの一体化エネルギーは

$$E \simeq E^0(s_i) \equiv -\frac{s_i}{2\sigma^2} (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + T_s s_i + \lambda_s \sum_{j \in \mathbf{N}(i)} (1 - \langle l(i, j) \rangle) (1 - \delta(s_i - \langle s_j \rangle))$$

となる。以降、(15) 式の意味での平均を  $\langle \dots \rangle$  で表すことに注意されたい。同様に  $d_i$  のみを残すと

$$E \simeq E^0(d_i) \equiv \frac{(1 - \langle s_i \rangle)}{2\sigma^2} (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - \langle l(i, j) \rangle)$$

また、 $l(i, j)$  についてのみ残すと

$$\begin{aligned} E &\simeq E^0(l(i, j)) \\ &\equiv \lambda_d (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - l(i, j)) + \lambda_s (1 - l(i, j)) (1 - 2\delta(\langle s_i \rangle - \langle s_j \rangle)) + \frac{\alpha_l l(i, j)}{(x_i^\tau - x_j^\tau)^2} \end{aligned}$$

が得られる。従って、平均値:  $\langle s_i \rangle, \langle d_i \rangle, \langle l(i, j) \rangle$  等に対し、次の自己無撞着な方程式を得る。

$$\begin{aligned} \langle s_i \rangle &\equiv \frac{\sum_{s_i=0}^1 s_i e^{-\beta E^0(s_i)}}{\sum_{s_i=0}^1 e^{-\beta E^0(s_i)}} \\ &= \frac{\sum_{s_i=0}^1 s_i e^{-\beta \left\{ -\frac{s_i}{2\sigma^2} (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + T_s s_i + \lambda_s \sum_{j \in \mathbf{N}(i)} (1 - \langle l(i, j) \rangle) (1 - \delta(s_i - \langle s_j \rangle)) \right\}}}{\sum_{s_i=0}^1 e^{-\beta \left\{ -\frac{s_i}{2\sigma^2} (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + T_s s_i + \lambda_s \sum_{j \in \mathbf{N}(i)} (1 - \langle l(i, j) \rangle) (1 - \delta(s_i - \langle s_j \rangle)) \right\}}} \end{aligned} \quad (17)$$

$$\begin{aligned} \langle d_i \rangle &\equiv \frac{\sum_{d_i=0}^{d_{\max}} d_i e^{-\beta E^0(d_i)}}{\sum_{d_i=0}^{d_{\max}} e^{-\beta E^0(d_i)}} \\ &= \frac{\sum_{d_i=0}^{d_{\max}} d_i e^{-\beta \left\{ -\frac{(1 - \langle s_i \rangle)}{2\sigma^2} (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - \langle l(i, j) \rangle) \right\}}}{\sum_{d_i=0}^{d_{\max}} e^{-\beta \left\{ -\frac{(1 - \langle s_i \rangle)}{2\sigma^2} (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - \langle l(i, j) \rangle) \right\}}} \end{aligned} \quad (18)$$

$$\begin{aligned} \langle l(i, j) \rangle &\equiv \frac{\sum_{l(i, j)=0}^1 l(i, j) e^{-\beta E^0(l(i, j))}}{\sum_{l(i, j)=0}^1 e^{-\beta E^0(l(i, j))}} \\ &= \frac{\sum_{l(i, j)=0}^1 l(i, j) e^{-\beta \left\{ \lambda_d (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - l(i, j)) + \lambda_s (1 - l(i, j)) (1 - 2\delta(\langle s_i \rangle - \langle s_j \rangle)) + \frac{\alpha_l l(i, j)}{(x_i^\tau - x_j^\tau)^2} \right\}}}{\sum_{l(i, j)=0}^1 e^{-\beta \left\{ \lambda_d (1 - 2e^{-\beta_d |d_i - \langle d_j \rangle|^2}) (1 - l(i, j)) + \lambda_s (1 - l(i, j)) (1 - 2\delta(\langle s_i \rangle - \langle s_j \rangle)) + \frac{\alpha_l l(i, j)}{(x_i^\tau - x_j^\tau)^2} \right\}}} \end{aligned} \quad (19)$$

ここで、上記方程式には確率変数が含まれず、全て確定的変数で閉じていることに注意されたい<sup>3</sup>。従って、上記連立方程式をある種の「非線形写像」とみなし、全ての画素に対して同期的状態更新:

$$\begin{aligned} \langle s_i \rangle^{(t+1)} &= \frac{\sum_{s_i=0}^1 s_i \exp \left\{ -\beta \left\{ -\frac{s_i}{2\sigma^2} (x_i^\tau - x_{i-(d_i)}^{\tau-1})^2 + T_s s_i + \lambda_s \sum_{j \in \mathbf{N}(i)} (1 - \langle l(i,j) \rangle^{(t)}) (1 - \delta(s_i - \langle s_j \rangle^{(t)})) \right\} \right\}}{\sum_{s_i=0}^1 \exp \left\{ -\beta \left\{ -\frac{s_i}{2\sigma^2} (x_i^\tau - x_{i-(d_i)}^{\tau-1})^2 + T_s s_i + \lambda_s \sum_{j \in \mathbf{N}(i)} (1 - \langle l(i,j) \rangle^{(t)}) (1 - \delta(s_i - \langle s_j \rangle^{(t)})) \right\} \right\}} \\ \langle d_i \rangle^{(t+1)} &= \frac{\sum_{d_i=0}^{d_{\max}} d_i e^{-\beta \left\{ \frac{(1 - \langle s_i \rangle^{(t)})}{2\sigma^2} (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{j \in \mathbf{N}(i)} (1 - 2e^{-\beta d_i |d_i - \langle d_j \rangle^{(t)}|^2}) (1 - \langle l(i,j) \rangle^{(t)}) \right\}}}{\sum_{d_i=0}^{d_{\max}} e^{-\beta \left\{ \frac{(1 - \langle s_i \rangle^{(t)})}{2\sigma^2} (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{j \in \mathbf{N}(i)} (1 - 2e^{-\beta d_i |d_i - \langle d_j \rangle^{(t)}|^2}) (1 - \langle l(i,j) \rangle^{(t)}) \right\}}} \\ \langle l(i,j) \rangle^{(t+1)} &= \frac{\sum_{l(i,j)=0}^1 l(i,j) e^{-\beta \left\{ \lambda_d (1 - 2e^{-\beta d_i |d_i - \langle d_j \rangle^{(t)}|^2}) (1 - l(i,j)) + \lambda_s (1 - l(i,j)) (1 - 2\delta(\langle s_i \rangle^{(t)} - \langle s_j \rangle^{(t)})) + \lambda_l \frac{\alpha_l l(i,j)}{(x_i^\tau - x_j^\tau)^2} \right\}}}{\sum_{l(i,j)=0}^1 e^{-\beta \left\{ \lambda_d (1 - 2e^{-\beta d_i |d_i - \langle d_j \rangle^{(t)}|^2}) (1 - l(i,j)) + \lambda_s (1 - l(i,j)) (1 - 2\delta(\langle s_i \rangle^{(t)} - \langle s_j \rangle^{(t)})) + \lambda_l \frac{\alpha_l l(i,j)}{(x_i^\tau - x_j^\tau)^2} \right\}}} \end{aligned}$$

をある精度の範囲内で全ての方程式が満たされるまで反復的に解けばよい。具体的にここでは  $\epsilon_t \equiv N^{-1} \{ |\langle \mathbf{s} \rangle^{(t)} - \langle \mathbf{s} \rangle^{(t-1)}|^2 + |\langle \mathbf{d} \rangle^{(t)} - \langle \mathbf{d} \rangle^{(t-1)}|^2 + |\langle \mathbf{l} \rangle^{(t)} - \langle \mathbf{l} \rangle^{(t-1)}|^2 \}^{1/2} < \epsilon (\ll 1)$  のタイプの収束条件を課す。この反復過程で「逆温度」 $\beta$  を  $\beta \rightarrow \infty$  と制御することで事後確率最大推定 (MAP) 推定が得られる<sup>4</sup>。一方、 $\beta = 1$  で固定し、反復を繰り返すとその解が MPM 解に相当する。いずれの方法に対しても、得られるベクトル  $\langle \mathbf{d} \rangle^{(\infty)}$  が求める移動速度ベクトルに場相当する。

上記のアプローチを**平均場アルゴリズム (Mean-field algorithm)**と呼ぶ。計算量は指数オーダーから  $\sim N$  のオーダーまで低減されていることに注意されたい。しかし、計算量とは別の問題として、各更新式の強い非線形性により、反復式から得られる解の軌道が振動したり、場合によっては「カオス」が生じたり、あるいは次節でみるように平均場自体が発散してしまうという困難が現れる。これらの問題の解消については次節以降で詳しくみていく。

<sup>3</sup> $\langle d_i \rangle$  は水平/鉛直の 2 成分を持つので、その具体的な実装には注意が必要である。ここでは、 $i \mapsto (X, Y)$  なる関係で 2 次元座標系を組むと、それら 2 成分はそれぞれ

$$\begin{aligned} \langle d^x[X][Y] \rangle &= \frac{\sum_{d=0}^{d_{\max}} d^x[X][Y] e^{-\beta f_x}}{\sum_{d=0}^{d_{\max}} e^{-\beta f_x}} \\ \langle d^y[X][Y] \rangle &= \frac{\sum_{d=0}^{d_{\max}} d^y[X][Y] e^{-\beta f_y}}{\sum_{d=0}^{d_{\max}} e^{-\beta f_y}} \end{aligned}$$

ただし、

$$\begin{aligned} f_x &\equiv -\frac{(1 - \langle s[X][Y] \rangle)}{2\sigma^2} (x^\tau[X][Y] - x^{\tau-1}[X - d^x[X][Y]][Y - d^y[X][Y]])^2 \\ &+ \lambda_d \sum_{(K,L) \in \mathbf{N}(X,Y)} (1 - 2e^{\beta d} \{ |d^y[X][Y] - \langle d^x[K][L] \rangle|^2 + |d^x[X][Y] - \langle d^y[K][L] \rangle|^2 \}) (1 - \langle l[X][Y][K][L] \rangle) \\ f_y &\equiv -\frac{(1 - \langle s[X][Y] \rangle)}{2\sigma^2} (x^\tau[X][Y] - x^{\tau-1}[X - d^y[X][Y]][Y - d^x[X][Y]])^2 \\ &+ \lambda_d \sum_{(K,L) \in \mathbf{N}(X,Y)} (1 - 2e^{\beta d} \{ |d^x[X][Y] - \langle d^y[K][L] \rangle|^2 + |d^y[X][Y] - \langle d^x[K][L] \rangle|^2 \}) (1 - \langle l[X][Y][K][L] \rangle) \end{aligned}$$

となる。 $f_x, f_y$  の右辺第 1 項は  $x \leftrightarrow y$  の変換に対して不変ではないので、一般に  $\langle d^x[X][Y] \rangle \neq \langle d^y[X][Y] \rangle$  となることに注意されたい。

<sup>4</sup>具体的には  $\beta \rightarrow \beta + 0.01$  で  $\beta = 1.0$  から 1.5 へ増加させる。1.5 は我々のシステムの実質的「低温極限」である。

## 2.7 ハイパーパラメータ：確率モデルのマクロな特徴付け

ここまでの議論で、「ミクロな」確率場についての決定方程式が得られたが、エネルギー関数に含まれ、確率モデルを「マクロに」特徴つけるハイパーパラメータに対し、適切な値を割り当てなければ確率場についての更新式を組み立てることはできない。そこで、ここでは各ハイパーパラメータの役割について概観しておく。

1.  $\sigma^2$ ：白色正規雑音の分散に関連したパラメータである。つまり、

$$x_i^\tau = x_{i-(d_i)}^{\tau-1} + \eta_i, \quad \eta_i = \mathcal{N}(0, \sigma^2)$$

である。また、標準正則化理論の立場では、正則化項の強さを制御する変数とみなせる。

2.  $\beta_d$ ：隣接するピクセル間の速度ベクトルの違いをどの程度まで許容するかを決定する。 $\beta_d$ を増やせばそれだけ許容範囲は狭くなる。
3.  $\lambda_d, \lambda_s, \alpha_l$ ：正則化パラメータである。例えば、 $\lambda_d$ を大きくすることで速度ベクトル場の連続性の制約を強調する。
4.  $\beta$ ：逆温度。これはハイパーパラメータというより、制御変数とみなす方が適切な場合もある。

先行研究である Zhang ら (1995) の研究においては、これらのハイパーパラメータの値は「アドホック」に設定されていることに注意されたい。

## 3 予備実験：平均場の発散とその回避

上記の平均場アプローチで得られる推定値の精度は確率モデルを特徴つけるハイパーパラメータの選び方に強く依存する。ここではまず Zhang ら [3] によって与えられたこれらに対する典型値である次の値:  $(\beta, \sigma^2, \lambda_d, \beta_d, \alpha_l, T_s, \lambda_s) = (1, 0.2, 2.5, 4, 200, 5, 2)$  を代入し、上記の非線形写像を反復的に解いてみる。具体的に実験で用いた人工画像とその移動後の画像、および、正解画像を図 3 に載せる。それぞれの画像サイズは  $30 \times 30$  である。この図は探索範囲内の各画素間に大きな画素値の差がない連続した 2 枚の単純な人工画像であり、探索範囲内の各画素間に大きな画素値の差がある連続した 2 枚の画像である。Zhang ら [3] は上述のように、各ハイパーパラメータ値を「アド・ホック」に設定しており、適切値を同時推定することはしていない。一方、実画像は物体や背景の模様、ノイズなどの影響により、画像により特性は異なる。そのような特性の異なる画像に対し、各ハイパーパラメータの値をアド・ホックに与えて計算することは適切/実用的ではない。事実、探索範囲内の画素値の差が大きい画像に対し、上記手続きを実行した場合、(12) 式第 1 項が極端に大きくなり、その結果、反復式の各平均場の値が指数関数的に増え発散してしまう。そこで本研究では今までアド・ホックに設定されていたパラメータの適切な値を検討する。

### 3.1 正則化項のスケーリング係数

探索範囲内の各画素間に画素値の大きな差がある連続した 2 枚の画像に対し、平均場が発散してしまう理由はエネルギー関数 (12) 第 1 項である正則化項:  $(1/2\sigma^2) \sum_i (1 - s_i)(x_i^\tau - x_{i-(d_i)}^{\tau-1})^2$  が大きな画素値をはさんで極端に大きくなってしまふことに起因する。従って、ここでは Zhang ら [3] によってアド・ホックに定められたハイパーパラメータ  $\sigma^2$  を  $\sigma^2 \rightarrow \mu\sigma^2$  のようにスケーリングし、スケーリング係数  $\mu$  をハイパーパラメータとし、連続する 2 枚の画像フレーム  $\{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}\}$  に依存したその適切値を探っていく。

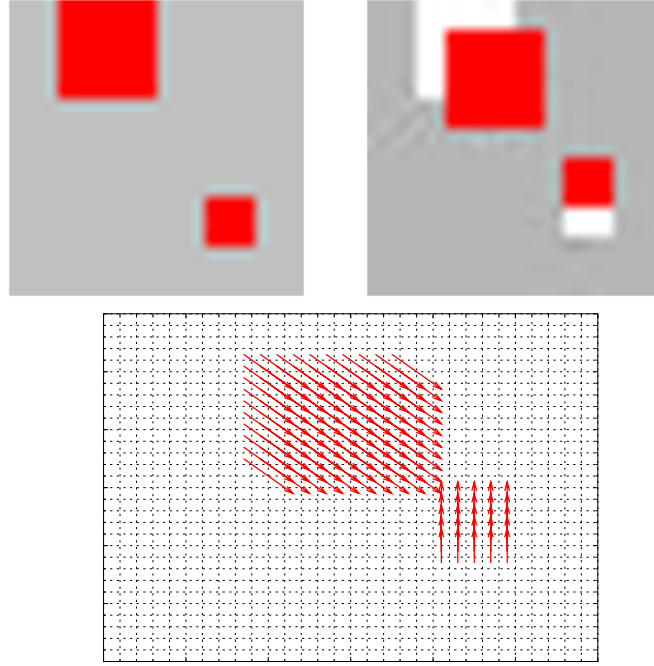


図 3: 実験で用いる典型的人工画像: 移動前画像 (左上), 移動後画像 (右上). 実際には, 背景部分, セグメンテーション部分, 物体部分に 適時階調値を割り当てて用いる. 下図は正解速度ベクトル場.

### 3.2 統計的性能評価に関する指標

移動速度推定の性能を評価する際, 次の 2 種類の平均自乗誤差を用いる.

$$D_1(\mu) \equiv \frac{1}{N_1} \sum_{i=1}^N (1 - s_i) (d_i^{(0)} - d_i)^2$$

$$D_2(\mu) \equiv \frac{1}{N_2} \sum_{i=1}^N s_i (d_i^{(0)} - d_i)^2$$

ここで,  $N_1 \equiv \sum_{i=1}^N (1 - s_i)$ ,  $N_2 \equiv \sum_{i=1}^N s_i$  であり,  $\mathbf{d}^{(0)}$  は人工画像に対し与えた正解速度ベクトル場である. 従って, 平均自乗誤差  $D_1$  は正解速度ベクトル場と推定速度ベクトル場の間の自乗誤差をセグメンテーション場の部分のみ取り出して評価したものであり, もう一方の  $D_2$  は全速度ベクトル場の中でセグメンテーション場を除く部分だけを切り出して比較した場合の自乗誤差である.

また, 本稿では上記自乗誤差とは別に正解ベクトルから食い違ったピクセル数を「ビット誤り率」として評価指標とし, それぞれ

$$\delta_1(\mu) \equiv \frac{1}{N_1} \sum_{i=1}^N (1 - s_i) \delta_{d_i^0, d_i}$$

$$\delta_2(\mu) \equiv \frac{1}{N_2} \sum_{i=1}^N s_i \delta_{d_i^0, d_i}$$

で定義する.  $\delta_{x,y}$  はクロネッカ・デルタである.

### 3.3 計算機実験の結果

ここでは具体的に前節で導入したスケーリング係数を用いた平均場反復式を構成し、その性能を数値計算により議論する。

#### 3.3.1 自乗誤差/ビット誤り率のスケーリング係数依存性

図??に示した人工画像と正解速度ベクトルに対し反復式でMPM推定を構成し、これら2種類の平均自乗誤差  $D_1, D_2$ , およびビット誤り率  $\delta_1, \delta_2$  をスケーリング係数  $\mu$  の関数としてプロットしたものが図4である。この図より、 $D_1$  は  $\mu$  の増加とともにある閾値を境にゼロから増加に転じ、

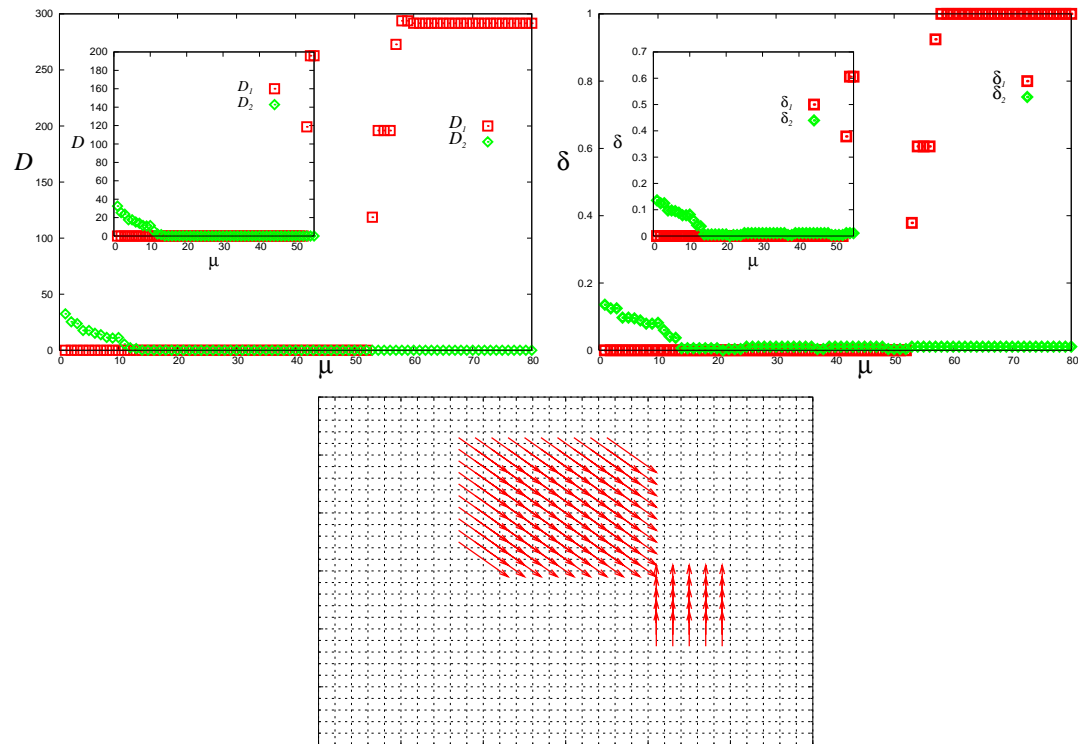


図4: 2種類の平均自乗誤差  $D_1, D_2$ (左上), およびビット誤り率  $\delta_1, \delta_2$  のスケーリング係数  $\mu$  依存性(右上). 下図は最適なスケーリング係数  $\mu_* \simeq 21$  を用いて得られた移動速度ベクトル場の様子. 用いた移動前画像の背景部分が  $Q = 0$ , セグメンテーション部分が  $Q = 40$ , 物体部分が  $Q = 10 \sim 30$  の5種類の画素値からなる.

逆に  $D_2$  は  $\mu$  の増加とともに有限値からゼロへと漸近する.  $\delta_1, \delta_2$  にも同様の振る舞いが見られる. 従って、これらの振る舞いから、2種類の平均自乗誤差、あるいはビット誤り率が同時にゼロになるスケーリング係数  $\mu$  を選ぶべきであるという一つの指針が得られる. この指針で最適なスケーリング係数を求めるとほぼ  $\mu_* \simeq 21$  である. この値を用いた非線形写像の反復で得られる実行結果を図4に載せる.

#### 3.3.2 エネルギーの時間発展

ここでは前節で求めた最適スケーリング係数  $\mu_* \simeq 21$  に対し、平均場の反復によるエネルギー関数のステップ依存性を調べてみる. 具体的には平均場状態更新の各ステップ  $\tau$  で次のエネルギー

関数:

$$E(t; \mu_*) \equiv E(\langle \mathbf{s} \rangle^{(t)}, \langle \mathbf{d} \rangle^{(t)}, \langle \mathbf{l} \rangle^{(t)} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \quad (20)$$

を更新ステップ数  $t$  の関数としてプロットする. その結果を図 5(左) に載せる. また, 同図内に平均場の反復過程で  $\beta^{-1}$  を「徐冷」させた場合のエネルギー時間発展の様子も重ね打ちする. この図

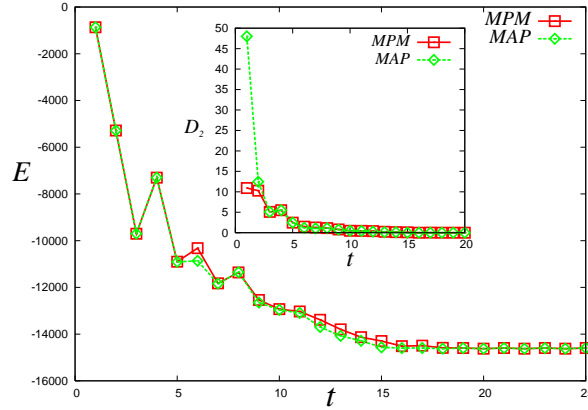


図 5: エネルギー関数の平均場状態更新数  $t$  依存性.  $\beta = 1$  に固定した MPM 解と  $\beta = 1.0$  から  $\beta = 1.5$  に増加させた MAP 解. 組み込み図は自乗誤差  $D_2$  の時間変化.

より, エネルギー最小化規準としては微量ではあるが, MAP 解が MPM 解と比べより低いエネルギー状態を達成していることがわかる.

### 3.3.3 ノイズの影響

ベイズ推定に基づく画像修復や誤り訂正符号では, ノイズの影響が大きい場合には MAP 解よりも MPM 解が平均自乗誤差を小さくすることが知られている [4, 5, 6]. そこで, ここでは物体移動後のフレームに「手ぶれ」などの誤差がノイズとして混入した場合の速度ベクトル場推定の精度を MAP および MPM で比較し, 結果を図 6 に載せる. この図より, ノイズの無い状況下と比べ, MAP 推定値に対する自乗誤差は MPM 解のそれと有為な差はみられない. そこで, 移動前画像の背景部分が  $Q = 0$ , セグメンテーション部分が  $Q = 30$ , 物体部分が  $Q = 10 \sim 18$  の 5 種類の画素値からなる画像を移動前画像の背景部分が  $Q = 0$ , セグメンテーション部分が  $Q = 50$ , 物体部分が  $Q = 10 \sim 40$  の 30 種類の画素値からなる画像変更した場合の結果を図 7 に載せる. この図より, 制御変数  $\beta$  を固定した場合, 自乗誤差, エネルギーともに振動現象を起こすことが見てとれる.

## 4 周辺尤度最大化規準とハイパーパラメータ推定

前節では平均場の発散を回避するため, 正則化項にスケーリング係数を導入し, それを最適化することで発散を回避し, 平均場方程式の振る舞いを安定化させることができた. スケーリング係数は新たなハイパーパラメータとなったが, より根源的で本質的な問題は, Zhang ら (1995) の先行研究ではハイパーパラメータを対象画像に応じてその都度, アド・ホックに (ある意味で「いい加減に」) 決めているため, 平均場方程式の動作が常に不安定となり, この事実が得られる速度場の精度

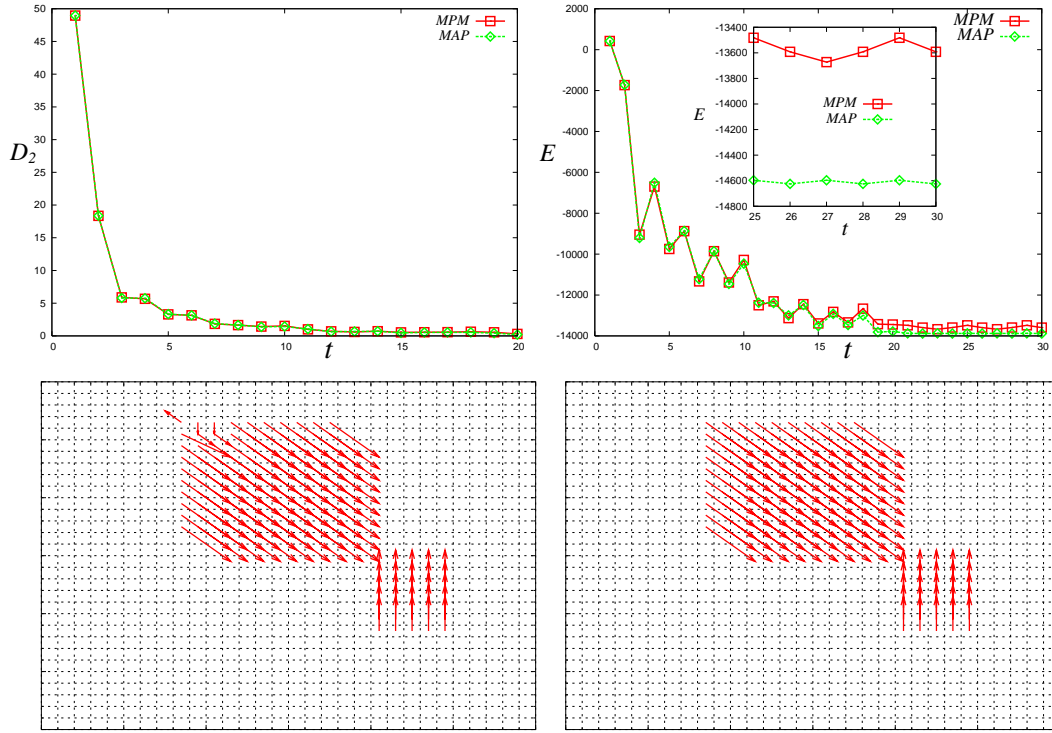


図 6: MAP 解と MPM 解の比較. エネルギー (右上) と自乗誤差 (左上) の時間発展の様子. 用いた移動前画像の背景部分が  $Q = 0$ , セグメンテーション部分が  $Q = 30$ , 物体部分が  $Q = 10 \sim 18$  の 5 種類の画素値からなる. 下図は得られる速度ベクトル (左が MPM 解).

に影響を与えている点である. その意味で何らかの適切な規準に基づき, これらハイパーパラメータをシステマティックに決定し, 正しいベイズ推定を行うことが何よりも重要であると考えられる.

さて, 統計学においては, 観測不可能な未知パラメータ (あるいは「隠れ変数」と言ってもよい)  $\{s, d, l\}$  を含む確率モデルに対し, ハイパーパラメータ:  $\xi \equiv \{\mu, \sigma, \lambda_d, \lambda_s, T_s, \alpha_l, \beta_d\}$  を決める指針としては広く周辺尤度最大化規準が知られている. 周辺尤度とは我々の確率モデルに対し

$$-F_{\xi}(\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}) \equiv \log \sum_{\{s, d, l\}} P(s, d, l | \mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}) \quad (21)$$

で与えられる. つまり, 「隠れ変数」に対してはその自由度で和をとってしまった尤度関数を周辺尤度と呼ぶ. この周辺尤度を入力データ  $\{\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}\}$  に対して平均化したものは真のハイパーパラメータ  $\xi^0$  の値で最大化される. すなわち

$$\left[ -F_{\xi^0}(\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}) \right]_{\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}} \geq \left[ -F_{\xi}(\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}) \right]_{\mathbf{x}^{\tau}, \mathbf{x}^{\tau-1}} \quad (22)$$

が成り立つ [5, 7, 8, 9].



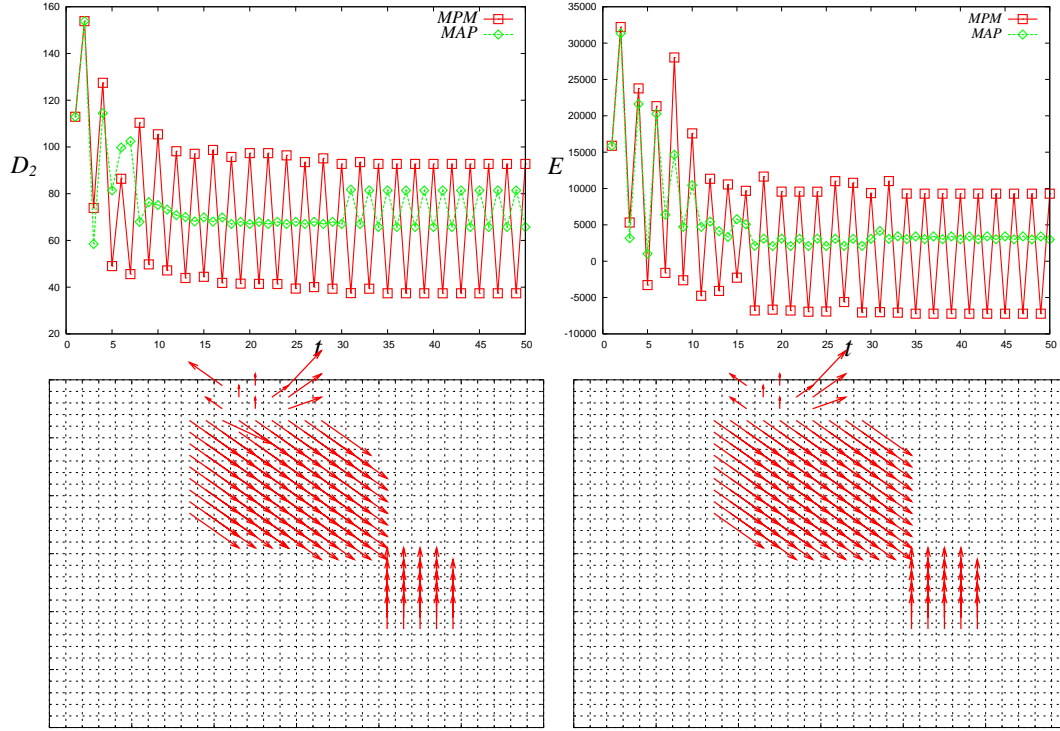


図 7: MAP 解と MPM 解の比較. エネルギー (左上) と自乗誤差 (右上) の時間発展の様子. 用いた移動前画像の背景部分が  $Q = 0$ , セグメンテーション部分が  $Q = 50$ , 物体部分が  $Q = 10 \sim 40$  の 30 種類の画素値からなる. 下図は得られる速度ベクトル (左が MPM 解).

#### 4.1 カルバック・ライブラ情報量の非負性と周辺尤度の最適性

不等式 (22) はカルバック・ライブラ (KL) 情報量の非負性から簡単に証明することができる. つまり, 真の確率モデル  $P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})$  とその推定モデル  $P_{\xi}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})$  との間の KL 情報量は

$$\begin{aligned}
 KL(P_{\xi_0} \| P_{\xi}) &= \sum_{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}} P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \log \left\{ \frac{P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})}{P_{\xi}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})} \right\} \\
 &= \sum_{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}} P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \log P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) - \sum_{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}} P_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \log P_{\xi}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \\
 &= [-F_{\xi_0}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})]_{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}} - [-F_{\xi}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})]_{\mathbf{x}^\tau, \mathbf{x}^{\tau-1}} \geq 0
 \end{aligned} \tag{23}$$

すなわち, 周辺尤度の最適性についての不等式 (22) 式が成り立つ.

#### 4.2 周辺尤度の平均場近似

従って, 我々が前節で導入した指針によって決定された  $\mu_*$  値と  $\xi$  を前節と同一の値に選んだ場合に周辺尤度を最大化する  $\mu$  の値を比較することは意味がある. しかし, この周辺尤度の計算ではやはり (21) 式に現れる確率変数の和の問題が現れ, 厳密計算を遂行することができない. そこで, ここでも次のような平均場近似を行う.

$$F \simeq F_{\xi}^{(mf)}(\mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \equiv \log \sum_{\{\langle \mathbf{s} \rangle^{(\infty)}, \langle \mathbf{d} \rangle^{(\infty)}, \langle \mathbf{l} \rangle^{(\infty)}\}} P(\langle \mathbf{s} \rangle^{(\infty)}, \langle \mathbf{d} \rangle^{(\infty)}, \langle \mathbf{l} \rangle^{(\infty)} | \mathbf{x}^\tau, \mathbf{x}^{\tau-1}) \tag{24}$$

ここで、上記の和は各確率変数の全ての自由度に渡ってとるのではなく、平均場方程式の収束値  $\langle \mathbf{s} \rangle^{(\infty)}, \langle \mathbf{d} \rangle^{(\infty)}, \langle \mathbf{l} \rangle^{(\infty)}$  を代入することにより算出されるので計算量は圧倒的に少なくなる。図 8 に

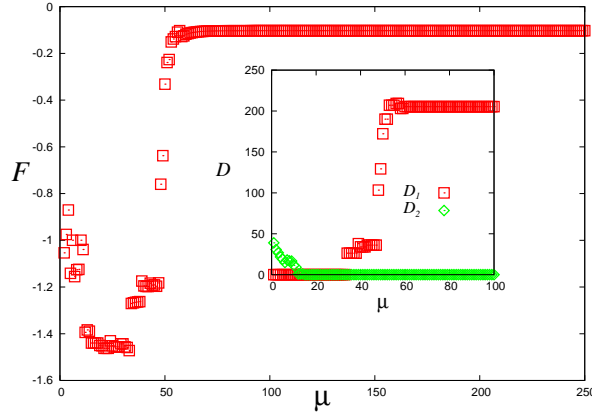


図 8: 平均自乗誤差  $D_1, D_2$  の  $\mu$  依存性 (組み込み図) と周辺尤度に負符号をつけたもの (自由エネルギー  $F^{(mf)}$ ) の  $\mu$  依存性.  $\mu_* \simeq 21$  近傍で最小値をとることがわかる.

負符号をつけた周辺尤度 (自由エネルギー  $F^{(mf)}$ ) の  $\mu$  依存性を載せる. この図より, 周辺尤度を最大化するスケーリング係数の値は  $\mu_* \simeq 21$  であり, これは前節で  $D_1, D_2$  の振る舞いの考察から得られた最適値とほぼ一致することがわかる. 従って, このスケーリング係数  $\mu$  の確定に関しては, 平均場近似のもとで書き換えられた周辺尤度最大化規準が有効であると考えられる. しかし, その他のハイパーパラメータに関し, 上記のような平均場近似で書き換えられた周辺尤度に関する勾配法を構成すると, 各ハイパーパラメータが線形に現れるため, うまく学習方程式を構成することができない. そこで, 次節では各確率場とハイパーパラメータを同時推定する方法について議論する.

## 5 確率場とハイパーパラメータの同時推定

前節では, 平均場近似のもとで書き換えられた周辺尤度最大化規準がスケーリング係数の決定に対しては有効に機能することがわかった. しかし, その他のハイパーパラメータに関し, 同様に平均場近似で書き換えられた周辺尤度に関する勾配法を構成すると, 各ハイパーパラメータが線形に現れるため, うまく学習方程式を構成することができない. そこで, 本節では, 実際に周辺尤度に関するハイパーパラメータの勾配で構成される **ボルツマン・マシン学習方程式** を導き, その方程式に含まれるミクロな確率場をその平均場でおき直す方法, および, **マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo : MCMC)** に基づくサンプリング計算により, ハイパーパラメータの最適値の推定を試みる.

### 5.1 ボルツマン・マシン学習方程式

具体的な学習方程式はハイパーパラメータ  $\xi$  に共役な統計量  $C(\mathbf{s}, \mathbf{d}, \mathbf{l})$  に対し, ハイパーパラメータの最適値を求めるために最急降下法による以下のような決定方程式で与えられる.

$$\frac{d\xi}{dt} = -\frac{\partial F_\xi(\mathbf{x}^\tau, \mathbf{x}^{\tau-1})}{\partial \xi} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} C(\mathbf{s}, \mathbf{d}, \mathbf{l}) P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau+1})}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} P(\mathbf{s}, \mathbf{d}, \mathbf{l} | \mathbf{x}^\tau, \mathbf{x}^{\tau+1})} \quad (25)$$

この式を各ハイパーパラメータについて書き下すと以下のようなになる。

$$\frac{dB}{dt} = \frac{\partial L}{\partial B} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_i (1-s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}}, \quad B \equiv \frac{1}{2\mu\sigma^2} \quad (26)$$

$$\frac{d\lambda_d}{dt} = \frac{\partial L}{\partial \lambda_d} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1-2e^{-\beta_d |d_i-d_j|^2})(1-l(i,j)) \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}} \quad (27)$$

$$\frac{d\lambda_s}{dt} = \frac{\partial L}{\partial \lambda_s} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1-l(i,j))(1-2\delta(s_i-s_j)) \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}} \quad (28)$$

$$\frac{d\alpha_l}{dt} = \frac{\partial L}{\partial \alpha_l} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}} \quad (29)$$

$$\frac{d\beta_d}{dt} = \frac{\partial L}{\partial \beta_d} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1-l(i,j)) |d_i-d_j|^2 e^{-\beta_d |d_i-d_j|} \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}} \quad (30)$$

$$\frac{dT_l}{dt} = \frac{\partial L}{\partial T_s} = -\frac{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} \left\{ \sum_i s_i \right\} e^{-U}}{\sum_{\mathbf{s}, \mathbf{d}, \mathbf{l}} e^{-U}} \quad (31)$$

最適値を求めるために各々の式に対して右辺の計算が必要になる。しかし、画像サイズが大きくなるにつれ、この種の計算時間が膨大となり、動画像処理の要求する時間内での処理は難しくなる。

### 5.1.1 ボルツマン・マシン学習方程式の平均場近似と MCMC 法によるハイブリッド解法

そこで、ここでは学習方程式右辺を平均場近似を用いて書き直す。最も簡単には、変数の中で (26) 式から (31) 式右辺の共役統計量の部分に現れる確率場以外はその平均値で置き換えることにする。例えば、ハイパーパラメータ  $B$  に関しては

$$\begin{aligned} \frac{dB}{dt} &= \frac{\partial L}{\partial B} = -\frac{\sum_{s_i, d_i} \left\{ \sum_i (1-s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 \right\} e^{-\langle U \rangle_{s_i, d_i}}}{\sum_{s_i, d_i} e^{-\langle U \rangle_{s_i, d_i}}} \quad (32) \\ \langle U \rangle_{s_i, d_i} &\equiv B \sum_i (1-s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1-2e^{-\beta_d |d_i-d_j|^2})(1-l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1-l(i,j))(1-2\delta(s_i-s_j)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i s_i \end{aligned} \quad (33)$$

となる。このように他のハイパーパラメータの決定方程式に対しても同様に平均場で近似する。多少複雑になるが、以下にそれらの結果を列記すると

$$\begin{aligned} \frac{d\lambda_d}{dt} &= -\frac{\sum_{d_i, d_j, l_{ij}} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1-2e^{-\beta_d |d_i-d_j|^2})(1-l(i,j)) \right\} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}}{\sum_{d_i, d_j, l_{ij}} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}} \quad (34) \\ \langle U \rangle_{d_i, d_j, l_{ij}} &\equiv B \sum_i (1-s_i)(x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1-2e^{-\beta_d |d_i-d_j|^2})(1-l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1-l(i,j))(1-2\delta(s_i-s_j)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i s_i \end{aligned} \quad (35)$$

$$\frac{d\lambda_s}{dt} = -\frac{\sum_{l_{ij}, s_i, s_j} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(s_i - s_j)) \right\} e^{-\langle U \rangle_{l_{ij}, s_i, s_j}}}{\sum_{l_{ij}, s_i, s_j} e^{-\langle U \rangle_{l_{ij}, s_i, s_j}}} \quad (36)$$

$$\begin{aligned} \langle U \rangle_{s_i, s_j, l_{ij}} &\equiv B \sum_i (1 - s_i) (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |\langle d_i \rangle - \langle d_j \rangle|^2}) (1 - l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(s_i - s_j)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i s_i \end{aligned} \quad (37)$$

$$\frac{d\alpha_l}{dt} = -\frac{\sum_{l_{ij}} \left\{ \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} \right\} e^{-\langle U \rangle_{l_{ij}}}}{\sum_{l_{ij}} e^{-\langle U \rangle_{l_{ij}}}} \quad (38)$$

$$\begin{aligned} \langle U \rangle_{l_{ij}} &\equiv B \sum_i (1 - \langle s_i \rangle) (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |\langle d_i \rangle - \langle d_j \rangle|^2}) (1 - l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(\langle s_i \rangle - \langle s_j \rangle)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i \langle s_i \rangle \end{aligned} \quad (39)$$

$$\frac{d\beta_d}{dt} = -\frac{\sum_{d_i, d_j, l_{ij}} \left\{ \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j)) |d_i - d_j|^2 e^{-\beta_d |d_i - d_j|} \right\} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}}{\sum_{d_i, d_j, l_{ij}} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}} \quad (40)$$

$$\begin{aligned} \langle U \rangle_{d_i, d_j, l_{ij}} &\equiv B \sum_i (1 - \langle s_i \rangle) (x_i^\tau - x_{i-d_i}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - d_j|^2}) (1 - l(i,j)) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - l(i,j))(1 - 2\delta(\langle s_i \rangle - \langle s_j \rangle)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{l(i,j)}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i \langle s_i \rangle \end{aligned} \quad (41)$$

$$\frac{dT_s}{dt} = -\frac{\sum_{s_i} \left\{ \sum_i s_i \right\} e^{-\langle U \rangle_{s_i}}}{\sum_{s_i} e^{-\langle U \rangle_{s_i}}} \quad (42)$$

$$\begin{aligned} \langle U \rangle_{s_i} &\equiv B \sum_i (1 - s_i) (x_i^\tau - x_{i-\langle d_i \rangle}^{\tau-1})^2 + \lambda_d \sum_{i,j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |\langle d_i \rangle - \langle d_j \rangle|^2}) (1 - \langle l(i,j) \rangle) \\ &+ \lambda_s \sum_{i,j \in \mathbf{N}(i)} (1 - \langle l(i,j) \rangle)(1 - 2\delta(s_i - \langle s_j \rangle)) + \alpha_l \sum_{i,j \in \mathbf{N}(i)} \frac{\langle l(i,j) \rangle}{(x_i^\tau - x_j^\tau)^2} + T_s \sum_i s_i \end{aligned} \quad (43)$$

が得られる。これら平均場近似を行った決定方程式を用いることによって、計算時間を大幅に短縮することができる。しかし、既に述べたように、上式右辺の  $\{\dots\}$  に現れる各ハイパーパラメータに共役な統計量の部分も平均場近似で書き換えてしまうと、右辺にはハイパーパラメータが線形に現れるため、適切な値に収束させることができない。そこで、ここではその困難を解消するため、「部分的に」MCMC法を用いる。具体的には、平均場近似で残した自由に動かさうパラメータのみに関するエネルギー（例えば  $\langle U \rangle_{s_i, d_i}$  であれば、 $s_i, d_i$  のみをMCMCの状態変数とする）に対し、次節で説明するMCMC法を適用し、右辺の  $\{\dots\}$  に現れる各ハイパーパラメータに共役な統計量の

期待値をサンプリングで計算する。従って、この方法は**平均場近似と MCMC 法のハイブリッド解法**と呼ぶことができる。

さて、時間微分に関しては適時、数値差分化を行う。本研究ではオイラー法(オイラー差分)を用いるため、

$$\begin{aligned}
B(t + \Delta t) &= B(t) + \Delta t \left\{ \frac{\sum_{d_i, s_i} \left\{ \sum_i (1 - s_i) (x_i^T - x_{i-d_i}^{T-1})^2 \right\} e^{-\langle U \rangle_{s_i, d_i}}}{\sum_{d_i, s_i} e^{-\langle U \rangle_{s_i, d_i}}} \right\} \\
\lambda_d(t + \Delta t) &= \lambda_d(t) + \Delta t \left\{ \frac{\sum_{d_i, d_j, l_{ij}} \left\{ \sum_{i, j \in \mathbf{N}(i)} (1 - 2e^{-\beta_d |d_i - d_j|^2}) (1 - l(i, j)) \right\} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}}{\sum_{d_i, d_j, l_{ij}} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}} \right\} \\
\lambda_s(t + \Delta t) &= \lambda_s(t) + \Delta t \left\{ \frac{\sum_{l_{ij}, s_i, s_j} \left\{ \sum_{i, j \in \mathbf{N}(i)} (1 - l(i, j)) (1 - 2\delta(s_i - s_j)) \right\} e^{-\langle U \rangle_{l_{ij}, s_i, s_j}}}{\sum_{l_{ij}, s_i, s_j} e^{-\langle U \rangle_{l_{ij}, s_i, s_j}}} \right\} \\
\alpha_l(t + \Delta t) &= \alpha_l(t) + \Delta t \left\{ \frac{\sum_{l_{ij}} \left\{ \sum_{i, j \in \mathbf{N}(i)} \frac{l(i, j)}{(x_i^T - x_j^T)^2} \right\} e^{-\langle U \rangle_{l_{ij}}}}{\sum_{l_{ij}} e^{-\langle U \rangle_{l_{ij}}}} \right\} \\
\beta_d(t + \Delta t) &= \beta_d(t) + \Delta t \left\{ \frac{\sum_{d_i, d_j, l_{ij}} \left\{ \sum_{i, j \in \mathbf{N}(i)} (1 - l(i, j)) |d_i - d_j|^2 e^{-\beta_d |d_i - d_j|} \right\} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}}{\sum_{d_i, d_j, l_{ij}} e^{-\langle U \rangle_{d_i, d_j, l_{ij}}}} \right\} \\
T_s(t + \Delta t) &= T_s(t) + \Delta t \left\{ \frac{\sum_{s_i} \left\{ \sum_i s_i \right\} e^{-\langle U \rangle_{s_i}}}{\sum_{s_i} e^{-\langle U \rangle_{s_i}}} \right\}
\end{aligned}$$

として数値処理を行う。つまり、(33)式から(43)式を連立させオイラー法によって更新していく。本研究では時間刻み幅を  $\Delta t = 0.001$  で固定した。具体的な手順としては、まずマイクロな変数(速度ベクトル場, ライン場, セグメンテーション場)を各平均場方程式(17)(18)(19)を用いて求める。次にその求めたマイクロな変数の値を式(33)~式(43)の各決定方程式に代入し、オイラー法を使い、かつ、共役統計量の期待値を部分的に MCMC 法で実行することで、マクロな変数(各ハイパーパラメータ)を求める。そして、マクロな変数の値を再び各平均場方程式に代入し、マイクロな変数を更新する。また、マイクロな変数をパラメータ決定方程式に代入しマクロな変数を更新する。このような手続き繰り返し更新していく。また、各ハイパーパラメータの初期値は Zhang の手法で設定されていた値とする。この手続きで得られる各ハイパーパラメータの収束点が最適解に近い値となるはずである。

### 5.1.2 マルコフ連鎖モンテカルロ法による学習方程式のサンプリング計算

前節では、ハイパーパラメータの最適値の推定方法について説明した。決定方程式を平均場近似を用いて書き直し、部分的に MCMC 法と組み合わせることで、計算時間を大幅に短縮することができると思われるが、計算時間を短縮することで、精度が犠牲になる恐れがある。一般に計算時間と精度の間にはトレード・オフの関係があるからである。そこで、ここでは精度を改善する手だてとして、先述の決定方程式右辺の計算の全てを MCMC 法を利用して実行する。モンテカルロ法とは人工的に事後分布を計算機上に再現し、そこからのサンプリングを何度も行うことにより、所望の期待値を近似的に求める計算手法である<sup>5</sup>。解析的に解くことができない問題に対しても、原理的には十分な回数のサンプリングを行うことにより、精度良く近似解を求めることができる。具体的には、以下の手順で実行される。

<sup>5</sup>本研究で用いるモンテカルロ法は「ギブス・サンブラ」とも呼ばれる。

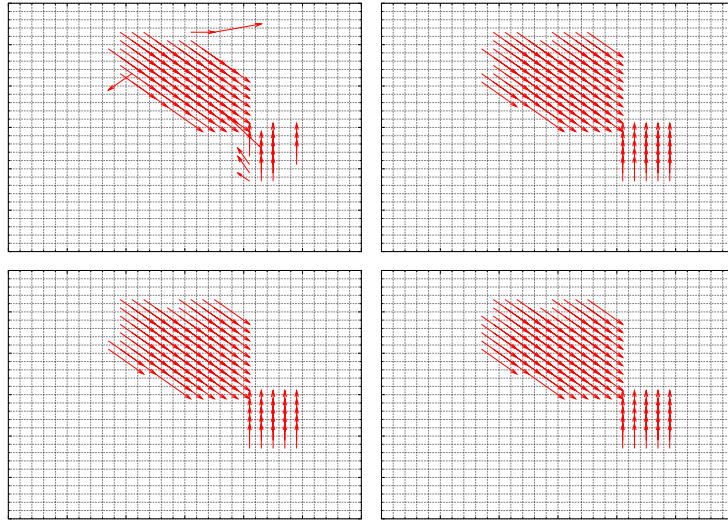


図 9: ハイパーパラメータ推定を導入した手法の実行結果 (ベクトル) の推移.  $t = 0$ (左上),  $t = 10$ (右上),  $t = 20$ (左下),  $t = 30$ (右下),  $t = 0$ (左上) が Zhang ら (1995) の手法に相当する.

- (1) 画像中のある画素 (位置) における, ある確率場に着目する.  
(例: 画素  $i$  におけるセグメンテーション場, つまり  $s_i$  に着目する.)
- (2) 着目した確率場の値を変更する.  
(例:  $s_i = 1$  であるならば,  $s_i = 0$  に変更する.)
- (3) 変数の値を変更することにより, 変更前のエネルギー  $E$  より変更後のエネルギー  $E'$  のほうが小さくなれば, その状態受け入れる. つまり, 状態を更新する. ただし, 変更後のエネルギーのほうが大きくなった場合でも  $\exp[(E - E')/T]$  の確率で状態更新を行う.
- (4) 上記 (1) から (3) の手続きを全ての画素, 確率場 (速度ベクトル場, ライン場, および, セグメンテーション場) に対して行う.

上記 (1)-(4) の手続きを行った後の状態に対し, 各パラメータの決定方程式 (33) から (43) における共役統計量の部分をサンプリング計算する. 求まる値が各決定方程式右辺の近似値となる.

## 5.2 実験結果 1: 学習方程式の平均場近似と MCMC 法によるハイブリッド解法

ここでは Zhang ら (1995) の手法と学習方程式の平均場近似に基づくパラメータ推定を導入した提案手法との実験結果を比較する. まず, 図 9, および, 図 10 に Zhang の手法による実行結果 (ベクトル) と提案手法の実行結果 (各オイラー・ステップ  $t$  での結果) の一例を示す. これらの図の違いは初期条件のみである. それぞれの実験に用いた画像は図 3 であり, 画素値は乱数により決定している. Zhang らの手法の結果は  $t = 0$  の結果に相当する. 図 9, および, 図 10 より, ハイパーパラメータ推定を導入した手法の方が Zhang の手法よりオイラー差分の繰り返し回数  $t$  が増えるにつれて速度ベクトル推定の精度が良くなっていることがわかる. 次に, ハイパーパラメータ推定を導入した提案手法の各オイラー・ステップ  $t$  における速度ベクトル誤差の推移を図 11 にプロット

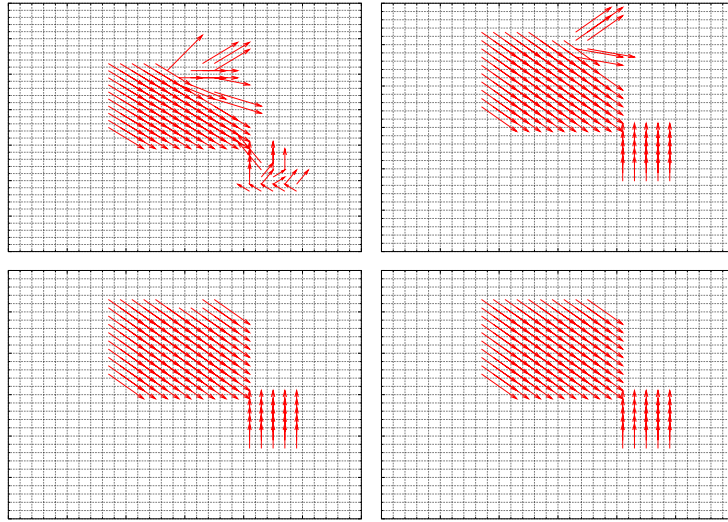


図 10: ハイパーパラメータ推定を導入した手法の実行結果 (ベクトル) の推移.  $t = 0$ (左上),  $t = 1$ (右上),  $t = 2$ (左下),  $t = 3$ (右下).  $t = 0$ (左上) が Zhang らの手法 (1995) に相当する.

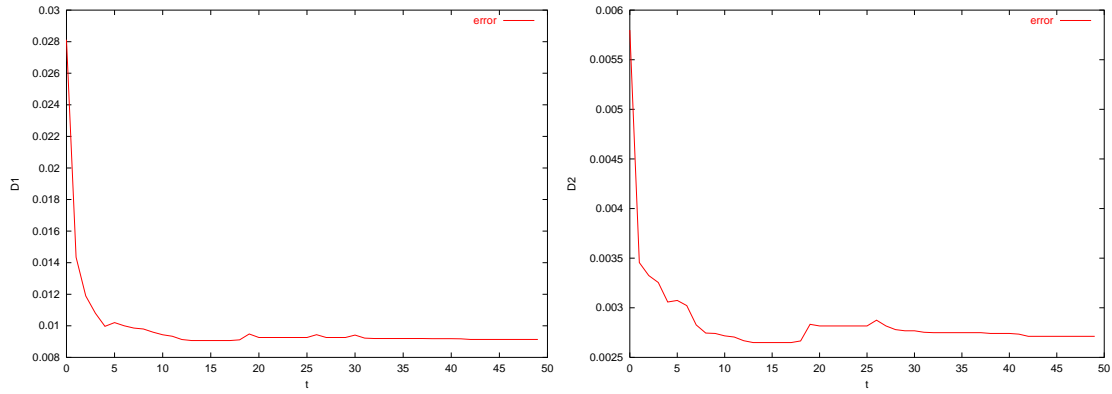


図 11: ハイパーパラメータ推定を導入した手法の各オイラー・ステップ  $t$  における速度ベクトルの誤差の推移. 左が  $D_1$ , 右が  $D_2$

する. ここでは以下の 2 種類の誤差を採用した. まずは

$$D_1 = \frac{1}{N} \sum_i (1 - \cos \theta_i) \quad (44)$$

であり, 角度  $\theta_i$  は正解の速度ベクトル場を  $\mathbf{a} = \{\vec{a}_1, \dots, \vec{a}_N\}$ , 実行結果の速度ベクトル場を  $\mathbf{b} = \{\vec{b}_1, \dots, \vec{b}_N\}$  とすると

$$\cos \theta_i = \frac{\vec{a}_i \cdot \vec{b}_i}{|\vec{a}_i| |\vec{b}_i|} \quad (45)$$

で定義される. この  $D_1$  は全体的なベクトルの向き of 誤差を評価している. 次に

$$D_2 = \frac{1}{N} \sum_i \left( 1 - \frac{|\vec{b}_i|}{|\vec{a}_i|} \right) \quad (46)$$

は全体的なベクトルの大きさの誤差を評価している. これら 2 種類の誤差によって速度ベクトル推定の結果を評価する. また, 誤差は 20 種類の画像に対する平均値としている. この図より, パラメータ推定を導入した提案手法は  $t = 10$  程まではステップ  $t$  が増えるにつれて誤差が小さくなり精度が良くなっていることがわかる.  $t = 10$  以降は誤差はほぼ変わらずに収束している. 以上より, 速度ベクトル, 誤差の結果から, Zhang らの手法 (1995) の結果と比べて, パラメータ推定を導入した提案手法が良好な結果を示すことがみてとれる.

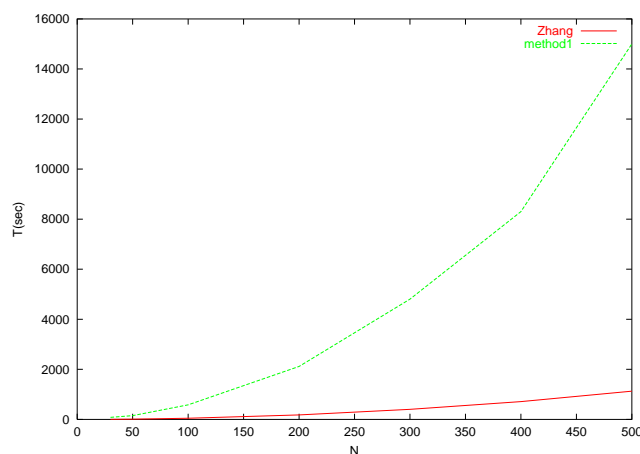


図 12: Zhang の手法と提案手法の画像サイズに対する計算時間の比較. 横軸は画像の一辺の長さ, 縦軸は計算時間 (秒)

次に, Zhang らの手法とパラメータ推定を導入した提案手法の各画像サイズに対する計算時間の比較を図 12 に示す. 具体的には Zhang らの手法は  $t = 1$  ステップにおいて平均場方程式の反復が 50 回, 提案手法は 1 ステップにおいてミクロな変数 (速度ベクトル場, ライン場, および, セグメンテーション場) に関しては平均場方程式の反復が 50 回, マクロな変数 (ハイパーパラメータ) の更新に関しては, 学習方程式に現れるミクロな変数の決定に対して部分的に用いる MCMC 法のステップを  $MCS = 100$  として実行し, 両手法が  $t = 50$  ステップに要した計算時間を比較する. 図 12 より, パラメータ推定を導入した場合の計算時間は画像サイズが大きくなるにつれ非常に膨大となっており, Zhang らの手法と比べても計算時間が大きいことが分かる. この結果は  $t = 50$  ステップに要した時間に対しての比較結果であり, 図 11 の結果より, 提案手法は  $t = 10$  においてほぼ収束することから, 実際の計算時間は図 12 の約  $1/5$  まで短縮されることが予想できるが, 画像サイズが大きいものに対しては, この計算時間が現実的ものとは考えづらい. 以上の考察より, 提案手法は Zhang の手法より精度面では格段に良くなるが, 計算時間は増加すると結論づけられる.

### 5.3 実験結果 2 : MCMC 法による学習方程式のサンプリング計算

実験結果 1 ではミクロな変数, つまり, 速度ベクトル場, ライン場, および, セグメンテーション場は平均場方程式 (17)(18)(19) を解くことによって計算し, マクロな変数, つまり各ハイパーパラメータに関しては平均場方程式の解を代入することによってその右辺が計算されたボルツマン・マシン学習方程式をオイラー差分で計算する (部分的に MCMC 法を利用), というように, それぞれを別々に算出した. そこで, ここではミクロな変数を平均場方程式で解くのではなく, 全てを MCMC 法によって計算を行うことで, ハイパーパラメータと確率場を同時に推定することを考える. つまり, 学習方程式の右辺計算においては各確率場 (速度ベクトル場, ライン場, および, セグ



メンテーション場)の値を変更し、エネルギーの大きさを比較することで状態を更新していた — メトロポリス法 — わけだが、各ステップでの各変数の状態を加算していき (サンプリング), 最終的に  $MCS$  (モンテカルロ・ステップ数) で割るという手順で各確率場の値が算出されることになる。つまり、ミクロな変数は以下の式で計算できる。

$$\langle s_i \rangle = \frac{\sum_{k=1}^{MCS} s_i^{(k)}}{MCS} \quad (47)$$

$$\langle d_i \rangle = \frac{\sum_{k=1}^{MCS} d_i^{(k)}}{MCS} \quad (48)$$

$$\langle l_{ij} \rangle = \frac{\sum_{k=1}^{MCS} l_{ij}^{(k)}}{MCS} \quad (49)$$

$s_i^{(k)}$ ,  $d_i^{(k)}$ ,  $l_{ij}^{(k)}$  は、各モンテカルロ・ステップ  $k$  における各変数の値である。この手順では、ミクロな変数とマクロな変数を同時に解いていることとなり、計算時間の短縮が期待できる。実際に、計算を行い先程の実験結果と比較してみる。ただし、モンテカルロ法においてはライン場、速度ベクトル場、セグメンテーション場の初期値によっては局所解に陥る可能性がある。そこで、本研究では適当な初期値を設定するために、モンテカルロ法を実行する前にミクロな変数について一度平均場方程式を解き、その解を初期値として与えることとする。ある程度適当な状態を初期値を設定することで、局所解に陥る可能性は減少する。実際に、初期値による実行結果 (ベクトル) の違いを図 13 に示す。適当な初期値を与えない場合、速度ベクトルを精度良く推定することができなかった。

ここで、実験結果 1 での手法を「提案手法 1」、実験結果 2 での提案手法を「提案手法 2」と名づけ、この両者を比較する。この図より、オイラーステップ数が少ない領域では**提案手法 1**の精度が良いが、オイラー・ステップ数が増加するにつれ、**提案手法 2**の精度が勝っていくことがわかる。続いて、**提案手法 1**と**提案手法 2**の計算時間について比較してみる。具体的には**提案手法 1**に関してはミクロな変数に対して単位オイラー・ステップあたり 50 回の平均場方程式の反復、マクロな変数に対しては単位オイラー・ステップあたり  $MCS = 100$  の MCMC 法を用いる、一方、**提案手法 2**に関してはマクロな変数に対して単位オイラー・ステップあたり、100MCS とし、それぞれ 50 オイラー・ステップに要した計算時間を比較する。両手法の計算時間を示したのが図 15 である。計算時間については**提案手法 2**の方が高速で計算できている。これらの結果から単純に比較すると、オイラー・ステップ数が少ない時点では**提案手法 1**が高精度、オイラー・ステップ数が多い時点では**提案手法 2**が高精度、計算時間に関しては**提案手法 2**が高速という結果となっている。**提案手法 1**は**提案手法 2**よりも計算時間がかかるが、オイラー・ステップ数が小さい時点では**提案手法 2**よりも精度が良く、さらに、**提案手法 2**よりも速く収束する。一方、**提案手法 1**は**提案手法 2**よりもより少ない計算時間で済むが、サンプリングに際するモンテカルロステップが増加するにつれ徐々に精度が良くなる、つまり、**提案手法 1**と比較して誤差の減少率がゆるやかであるため、最終的に高精度な結果を得るためには、より多くのモンテカルロ・ステップを要する。つまり、**提案手法 1**よりも計算時間が掛かる可能性がある。画像サイズが小さい場合は、両手法の計算時間に大きな差は見られないため**提案手法 1**が有効であると考えられる。

## 6 おわりに

本稿では動画像の連続する 2 フレーム間の物体の移動速度ベクトル場を推定するためのマルコフ確率場を用いた動画像の確率モデルに対し、周辺尤度最大化規準のもとで確率場とハイパーパラ

メータを同時に推定する方法を提案した。提案手法では周辺尤度を最大化する際に用いられる学習方程式に含まれる期待値を平均場近似, および, マルコフ連鎖モンテカルロ法を用いて近似計算し, 学習方程式の収束点としてハイパーパラメータと確率場が同時に確定できる。得られる結果はいずれも先行研究の精度を超えるものであった。また, この結果はそれまで何ら根拠なく「アド・ホック」に決められ, 先行研究で用いられていたハイパーパラメータ値が周辺尤度最大化規準のもとで正当化されるものではなく, その意味において統計学的にみて間違っており, 我々の手法で得られるパラメータこそが統計学的に正しく, かつ, 最適なものであることを明確に示すものである。

## 謝辞

指導教員である井上純一先生に感謝いたします。また, 小野哲雄先生には有益な助言を頂きました。また, 自由で快適な環境で研究することができたのは研究室の皆さんのおかげです。ここに感謝いたします。

## 参考文献

- [1] B.K. Horn and B.G. Schunck, *Artif. Intell.*, **17**, pp. 185-203 (1981).
- [2] J. Konard and E. Dubois, *Proc. Conf. Vision Interface*, pp. 51-60 (1989).
- [3] J. Zhang and G.G. Hanauer, *IEEE Transaction on Image Processing*, **40**, No.1, pp. 19-33 (1995).
- [4] 西森秀稔, 「スピングラス理論と情報統計力学」新物理学選書, 岩波書店 (1999).
- [5] 田中和之, 「確率モデルによる画像処理入門」森北出版 (2006).
- [6] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer (2006).
- [7] J. Inoue and K. Tanaka, *Physical Review E*, **65**, 016125 (2002).
- [8] K. Tanaka and J. Inoue, *IEICE Trans. on Information and Systems* **E85-D**, No. 3, pp. 546-557 (2002).
- [9] K. Tanaka, J. Inoue and D. M. Titterington, *Journal of Physics A : Mathematical and General*, **36** pp. 11023-11035 (2003).
- [10] K. Tabushi and J. Inoue, *Proceedings of the IEEE Neural Networks for Signal Processing*, **XI**, pp.133-142 (2001), J. Inoue and K. Tabushi, *International Journal of Modern Physics B*, **17**, No. 29, pp. 5525-5539 (2003).

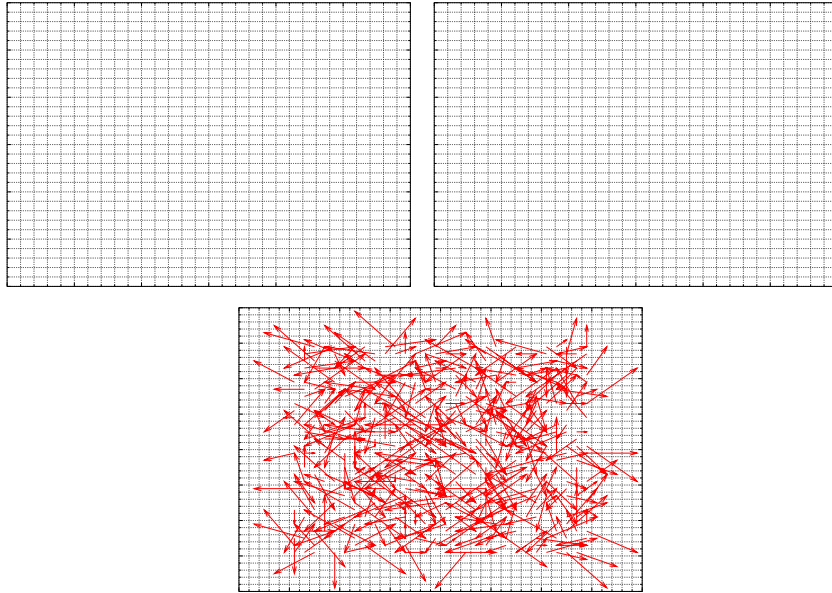


図 13: ミクロな変数の初期値依存性すべての初期値が 0(左), すべての初期値が 1(中), 初期値がランダム (右) の場合のそれぞれの実行結果 (ベクトル)

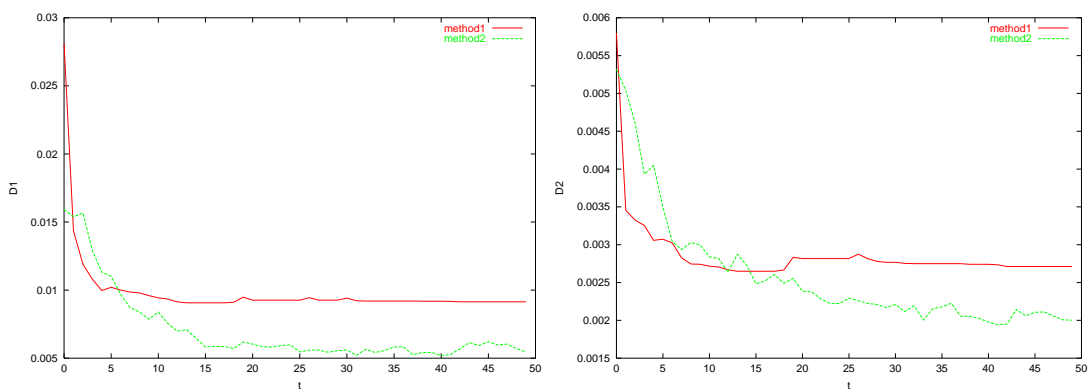


図 14: 提案手法 1(method1) と提案手法 2(method2) の各オイラー・ステップにおける速度ベクトルの誤差の推移の比較. 左が  $D_1$ , 右が  $D_2$

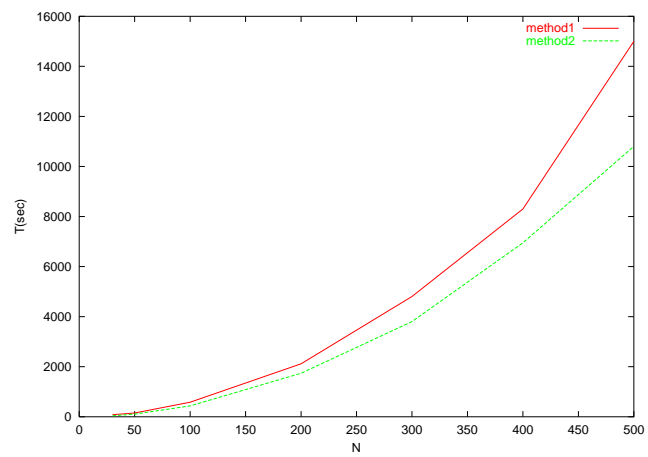


図 15: 提案手法 1 と提案手法 2 の画像サイズに対する計算時間の比較. 横軸は画像の一辺の長さ, 縦軸は計算時間 (秒)