

Potts モデルを用いた動画画像解析

木下 敬介
ATR 人間情報科学研究所
kino@atr.jp

イリス・フェルミン
アストン大学
i.fermin@aston.ac.uk

あらまし 画像の動き解析は、画像認識技術の根幹であり、侵入者監視や人間の動きや表情解析など、応用範囲も広い。従来は、動きのモデルをあらかじめ与え、それに適合する動きを推定するという手法が主であった。本研究では、各画素を粒子に見立て、磁性体のモデルである Potts モデルとモンテカルロ法を適用し、統計的な手法で動きをクラスタリングする。システムの温度を変化させると相転移が生じるが、超常磁性の状態では複数のクラスタが現れることが知られている。画素間の適切な相互作用と適切な温度を与え、超常磁性の状態をつくり出すことで、動きのクラスタを抽出する。

Motion Analysis based on Potts Model

Keisuke Kinoshtia
ATR Human Information Science Labs.

Iris Fermin
Aston University

Abstract The computation of visual motion is important for tasks such as stereo motion, depth estimation and control of locomotion. The classification of regions into clusters that contain the same motions is meaningful for discrimination of static and moving objects, understanding gestures. We present a new approach for clustering motion from spatio-temporal image information based on Potts model and Monte Carlo simulation. Analogous to spin model in statistical physics, each pixel is assigned its own spin. They are statistically estimated considering interaction among neighbouring pixels. Choosing appropriate temperature, super-para magnetism phase emerges allowing the extraction of clusters that have the same motion.

1 はじめに

映像中の動きを抽出し、分類することは、画像認識技術の中核であり、応用範囲も広い。交通監視、侵入者検出、人物の認識、テレビ映像の切り出しの前処理には必須である。また、映像中の動きから対象物体の 3 次元形状を推定する Structure from Motion と呼ばれる手法にも深くかかわっている。

画像の動き解析の研究の歴史は長く、1980 年代からオプティカルフロー解析として知られ、各画素の明るさ情報から BCC(Brightness Constant Constraint) と呼ばれる拘束条件 [1] を導き出し、動きの推定が行われてきた。また、画像列の前後の画素の相関を計算することで対応点を追跡する方法、まず画像を同一の性質をもつ領域に分割し、分割後の領域の動きを追跡する方法も提案されている [2][3]。いずれも、画像処理の分野では古典的な問題である。また、統計的な動き解析として Geman らの研究は有名である [4]。

本研究では、画像の動きの抽出と分類を磁性体の物理モデルである Potts モデル [5] で解析する方法を提案する。Potts モデルは、磁性体のスピンモデルである

Ising モデルのスピンの状態を 2 値 (+1 と -1) から多値 $(1, \dots, q)$ に拡張したものである。Potts モデルを利用した画像解析の研究として [6] がある。Blatt らは、映像中にちりばめられた点群を磁性体の粒子に見立て、点間の局所的な相互作用の集積により、大局的なクラスタを生成している。Potts モデルでは、温度と局所的な相互作用を変化させることで、クラスタの結果が異なってくる。Blatt らは、温度を連続的に変化させ、システムが、強磁性から超常磁性、そして常磁性の状態に変化していくことを、システムの磁化率を監視しながら検出し、超常磁性の状態が、最もよいクラスタが得られると結論付けている。

一方、Potts モデルを動きのクラスタリングに利用するためには、局所的な相互作用をうまく定義しなくてはならない。BCC だけでは動きを一意に決定できないので、Blatt らが使用した相互作用の定義をそのまま流用するわけにはいかない。本研究では、各画素の BCC と、クラスタの動きベクトルを使って動きのクラスタリングに適した局所的な相互作用を与える。

一般に Potts モデルのような統計的な手法では、計算時間が膨大にかかる。ここでは、モンテカルロ法と

Swendsen-Wang のアルゴリズム [7][8] を使って効率化を図っている。

2 Potts モデル

Potts モデルは、Ising モデルを多値に拡張したもので、磁性体のスピンモデルとして知られている。各粒子には、 q 個の状態 $(1, \dots, q)$ のうちの 하나가割り当てられる。粒子 i に割り当てられたスピンの状態を σ_i と表すと、システムのハミルトニアンは、

$$\mathcal{H} = - \sum_{\langle i, j \rangle} J_{ij} \delta_{\sigma_i \sigma_j} \quad \sigma_i = 1, \dots, q \quad (1)$$

で表される。 $\delta_{\sigma_i \sigma_j}$ は、粒子 i と j の間の相関を、 J_{ij} は両者の相互作用であり、これらの積の総和は、近傍領域 $\langle i, j \rangle$ で定義される。いま、 N 個の粒子がある場合、 $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_N)$ は、このシステムの状態を表しており、ある状態 $\boldsymbol{\sigma}_h$ になる確率は、温度を T として

$$P(\boldsymbol{\sigma}_h) \propto \exp\left(-\frac{\mathcal{H}(\boldsymbol{\sigma}_h)}{zT}\right) \quad (2)$$

と表せる。ただし、 z は正規化するための定数である。

このシステムの温度 T を変化させると、相転移が起きる。 T が小さい場合は強磁性 (ferromagnetism) を示し、 T が大きい場合は常磁性 (paramagnetism) となる。強磁性ではスピンは一方向に収束し、常磁性ではスピンはばらばらになり秩序を持たない。しかし、これらの間に超常磁性 (super-paramagnetism) が現れる場合がある。これは、磁区が小さくなり、磁化が熱的に揺らいでいる状態である。つまり、粒子がクラスタとなり、複数のクラスタが存在するようになる [9]。本研究では、画素を粒子にみたて、超常磁性となる温度を探ることで、画像の動きをクラスタリングする。

Potts モデルでは、粒子間の相互作用 J_{ij} が重要な働きをする。Blatt らの研究では、点 i と点 j の距離を d_{ij} とし、

$$J_{ij} = J_{ji} = \frac{1}{K} \exp\left(-\frac{d_{ij}^2}{2a^2}\right) \quad (3)$$

と置いている (K, a に関しては後述)。つまり、点同士の距離が近ければ近いほど J_{ij} は大きくなる。いわば、 d_{ij} は点 i と点 j の類似度に相当し、この類似度に応じて相互作用が定義される。

(2) を計算する場合、状態数が非常に大きい (N^q) ため、そのまま計算することは不可能である。そこで、モンテカルロ法を使わざるを得ない。中でも、Swendsen-Wang アルゴリズム [7, 8] は、スピンモデルを計算するのに適した手法である。これは、粒子同士の「リンク」を確率的に結合・切断し、結合したリンクをもつ粒子の集合をクラスタとする。各クラスタにランダムにスピンの状態を割り当て、再びリンクの結合と切断の操作を繰り返す。従来のモンテカルロ法のように各粒子別個にラ

ンダムにスピンの状態を割り当てるのではなく、クラスタに属する粒子はまとめて同じスピンの状態が割り当てられるので効率的に計算が可能となる。

3 動画像の拘束条件

画像に動きがある場合、オプティカルフローが生じる。いま、画像の輝度を $I(x, y, t)$ とする。二枚の画像間で、ある画素の移動前後の輝度に変化しないという仮定を置く。 $I(x, y, t)$ を x, y, t で偏微分したものを $\mathbf{f} = (f_x, f_y, f_t)$ 、その画素の動きベクトルを $\mathbf{u} = (u, v)$ とすると、

$$f_x u + f_y v + f_t = 0 \quad (4)$$

という拘束式が得られる。これは BCC と呼ばれ、動きベクトルは、画像上で $\mathbf{f}$ で表される直線上に乗ることを示している。しかし、 u, v を、この拘束式だけから一意に求めることはできない。そのため、[1] では、動きに対して局所的な滑らかさ拘束を付加することで動き場を推定している。

一方、Potts モデルを使って動きのクラスタリングをする場合、(1) にあるように、 J_{ij} を定義する必要がある。これは画素 i の近傍にある画素 j のそれぞれの動きベクトル $\mathbf{u}_i$ と $\mathbf{u}_j$ がどれくらい似ているのかを示す指標でなくてはならない。しかし、(4) からは $\mathbf{u}_i$ と $\mathbf{u}_j$ を決定できないという問題点が生じる。その解決策のひとつとして、

$$d_{ij} = \frac{\|\mathbf{f}_i \times \mathbf{f}_j\|}{\|\mathbf{f}_i\| \|\mathbf{f}_j\|} \quad (5)$$

を使う方法がある [10]。これは、 $\mathbf{u}_i$ と $\mathbf{u}_j$ が同じであるなら、対応する拘束直線は同じであるという仮定に基づいている。

ここでは、クラスタの情報を陽に使用する方法を提案する。いま、画素 i が属している (仮の) クラスタの動きを $\mathbf{U}_i = (U_i, V_i)$ とする。もし、画素 i がこのクラスタに属しているなら、 $\mathbf{U}_i$ は、(4) の拘束を満たすはずである。そこで、画素 i と画素 j の動きの類似度の指標として、

$$\begin{aligned} d_{ij} &= \sqrt{d_i^2 + d_j^2} \\ d_i &= \frac{|f_{xi}U_i + f_{yi}V_i + f_{ti}|}{\sqrt{f_{xi}^2 + f_{yi}^2}} \\ d_j &= \frac{|f_{xj}U_j + f_{yj}V_j + f_{tj}|}{\sqrt{f_{xj}^2 + f_{yj}^2}} \end{aligned} \quad (6)$$

を導入する。 d_i は、クラスタの動きベクトル $\mathbf{U}_i$ と、 $\mathbf{f}_i$ が表す拘束直線との距離を表している。本研究では (6) を採用する。

4 Potts モデルによるクラスタリング

Potts モデルにおいては、システムの挙動は温度 T と相互作用 J_{ij} によって決まる。 J_{ij} に関しては (6) を採用し、温度 T に関しては、低い温度から高い温度まで T をスイープし、適切な温度を探る。一般に、強磁性から常磁性への相転移はキュリー温度 T_c で起こるとされ、 T_c は

$$T_c = \frac{e^{-1/2}}{4 \log(1 + \sqrt{q})} \quad (7)$$

で近似できる。超常磁性もこの近辺で表れると考えられるので、温度 T を $0 < T < 4T_c$ の範囲でスイープする。また、各温度での磁化率 χ は相転移がどのように変化したかの指標となる。

以下、各温度 T における Potts モデルと SW アルゴリズム、および (6) を使った動きのクラスタリング手法の具体的な流れを示す。

1. 画像中の画素を粒子に見立て、 $\sigma_i (i = 1 \dots N)$ にスピンの状態 ($1 \dots q$) をランダムに割り当てる。まだクラスタは生成されていないので、各画素が画素数 1 のクラスタとなる。つまり、クラスタの総数は N である。クラスタの動きベクトルにはランダムな値を格納する。
2. (6) に基づき、 J_{ij} を計算する。 $\hat{K}$ は近傍の画素数の平均であり、 a は、 d_{ij} の平均とする。
3. 近傍の画素とのリンクを次に示す確率に従って切断・結合する。

$$P_{\text{delete}}^{ij} = \exp\left(-\frac{J_{ij}}{T}\right) \quad (8)$$

$$P_{\text{bond}}^{ij} = 1 - P_{\text{delete}}^{ij} \quad (9)$$

ただし、結合は $\sigma_i = \sigma_j$ の場合のみに限定する。

4. 結合されたリンクを有する画素の集合をクラスタとする。
5. 各クラスタで、動きベクトル U を計算する。これは、クラスタに含まれる画素の拘束式 (4) を連立させて解く。クラスタに含まれる画素数が少ない場合や (4) を連立させても解が不定になる場合は、 U にランダムな値を格納する。
6. 各画素にランダムにスピンの状態を割り当てる。ただし、同じクラスタに属する画素には同じスピンの状態を割り当てる。
7. 2. の処理へ。これがモンテカルロ法の 1 ステップに相当する。
8. 2.~7. を規定の回数反復処理した後、各種の統計情報を計算する。まず、 c_{ij} を、画素 i と画素 j が同一

クラスタに属する場合は 1、そうでない場合は 0 とし、すべてのモンテカルロ法のステップの平均 $\langle c_{ij} \rangle$ を計算する。この $\langle c_{ij} \rangle$ を使って、画素 i と画素 j が同一のクラスタに属する確率 $\langle \delta_{ij} \rangle = \frac{(q-1)\langle c_{ij} \rangle + 1}{q}$ を計算する。 $\langle \delta_{ij} \rangle > \theta$ である画素のリンクを結合し、結合を持つ画素の集合をクラスタとする。これが温度 T におけるクラスタリングの結果となる。

9. 磁化率 χ を計算するためには磁化 m が必要である。 m は、各モンテカルロ法のステップで、 $m = \frac{qN_{\text{max}} - N}{N(q-1)}$ として計算する。ただし、 $N_{\text{max}} = \max(N_1, \dots, N_q)$ であり、各スピン値をとる画素数の最大となるものである。 χ は、 m の分散を使って $\chi = \frac{N}{T}(\langle m^2 \rangle - \langle m \rangle^2)$ と計算できる。

温度をスイープしながら以上の操作を行うと、磁化率 χ は、温度が低いとき、つまり強磁性ではほぼ 0 であり、その後 T_c 近辺で急激に上昇する。さらに温度を上昇させると、 χ は下落し、しばらく χ が変化しない範囲が表れる。これが超常磁性の状態である。この状態がクラスタリングに適した温度であるといえる。その後は、急激に 0 に低下し、常磁性の状態となる。

5 実験

図 5 に示すような 2 枚の画像間の動きのクラスタリング実験を行った。元の画像は 480×640 画素であるが、 f を計算するに当たり、まずガウシアンフィルタで画像全体を事前にスムージング処理し、解像度を $1/8$ にしている。 f の各要素の値を図 5 の 2 段目右、3 段目左右にそれぞれ示す。近傍領域は、画素のまわり 8 近傍とし、 $q = 11$ とした。また、モンテカルロ法の試行は 4000 回、クラスタを決定する際のパラメータ θ は 0.5 とした。

図 5 右に χ の温度 T に対する変化を示す。 $T = 0.25$ の近辺で χ の変化が少なくなっており、超常磁性の状態であることがわかる。このときのクラスタリングの結果を図 5 左に示す。頭と胴部がほぼ同じ動きをし、右腕はこれと異なる動きをしている。

当然ながら T が 0 に近い場合は、すべての画素が一つのクラスタに包含され、 T が 0.3 を超えると、クラスタと呼べるものはなくなり、各画素がばらばらの状態になった。

6 まとめ

Potts モデルを使った画像の動きのクラスタリング手法を提案した。このとき重要となるのは、画素間の相互作用 J_{ij} をいかに定義するかである。本研究では、その画素の BCC に対して、その画素が属しているクラスタの動きベクトルが合致することを要求して J_{ij} を決定した。

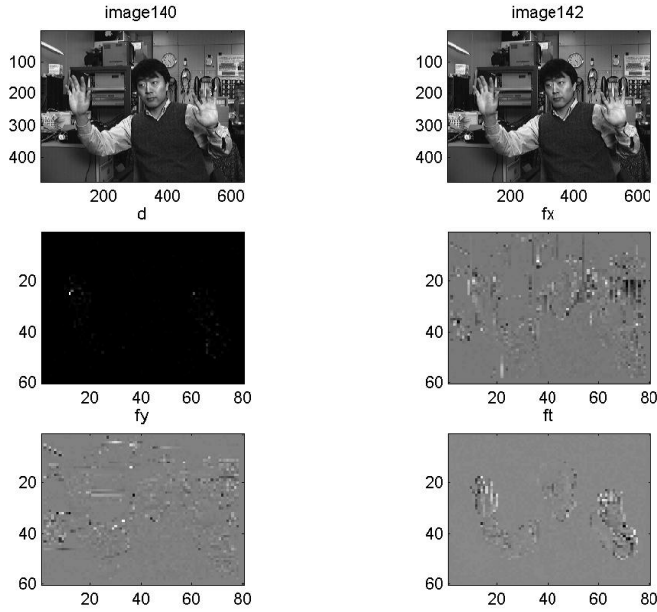


図 1: 実験に使用した画像。上段は動きの前後の画像。
 f_x, f_y, f_z はそれぞれ、 $\mathbf{f} = (f_x, f_y, f_z)$ を表している。

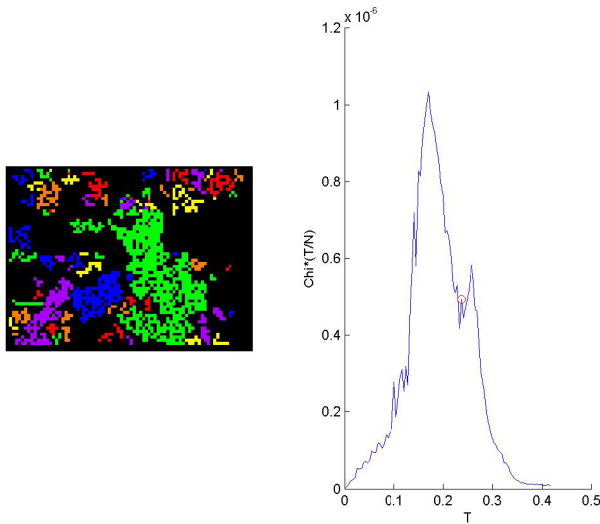


図 2: 磁化率 χ の温度 T に対する変化と $T = 0.25$ のときのクラスタリングの結果。

本研究のような統計的な動きの解析の利点として、局所的な相互作用を定義するだけで、大局的な動きのクラスタリングができるということが挙げられる。従来の研究では、剛体運動や動きの滑らかさ拘束といった、動きのモデルが必要であったが、本研究ではそのような事前知識は最小限にすることができる。一方で、計算時間が膨大に必要であるという欠点がある。モンテカルロ法と SW アルゴリズムを使用して計算時間を大幅に短縮できるが、リアルタイムで処理することは今のところ困難である。また、磁化率 χ を相転移の状態の指標として使用したが、その他の指標（例えばエントロピー）を導入することも考慮するべきであろう。

実験では、単純な動きを取り上げたが、今後、より複雑な動き、例えば、複数の動きが重なり合い、交じり合っている状況や、水の流れのように動きのベクトルが連続的に変化している状況にも適用して有効性を確認することが必要である。

謝辞

本研究は独立行政法人情報通信研究機構の研究委託「人間情報コミュニケーションの研究開発」により実施したものである。

参考文献

- [1] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [2] C. Fermüller, D. Shulman, and Y. Aloimonos. The statistical of optical flow. *Computer Vision and Image Understanding*, 82:1–32, 2001.
- [3] H. Sawhney, Guo Y., and R. Kumar. Independent motion detection in 3d scenes. *IEEE Trans. on Pattern Recognition and Machine Intelligence*, 22(10):1191–1199, 2000.
- [4] Geman S. and Geman D. Sthocastic relaxation, gibbs distributions, and the bayesian restoration of images. *Transaction in Pattern Analysis and Machine Intelligence*, 6(6).
- [5] F.Y. Wu. The potts model. *Review of Modern Physics*, 54(1):235–268, 1984.
- [6] M. Blatt, S. Wiseman, and E. Domany. Clustering data through an analogy to the potts model. In *Neural Information Processing Systems, NIPS*, volume 8, 1995.
- [7] J-S Wang. Cluster monte carlo algorithms and their applications. *Recent Developments in Computer Vision*, S.Z. Li et.al. (eds.), LNCS 1035:307–, 1996.
- [8] R. Swendsen and J-S. Wang. Nonuniversal critical dynamics in monte carlo simulation. *Physics Review Letters*, 58(2):86–88, 1987.
- [9] S. Wiseman, M. Blatt, and Domany E. Superparamagnetic clustering of data. *Physics Review E*, 57:3767–3783, 1998.
- [10] Keisuke Kinoshita and Iris Fermin. Clustering image sequence based on potts model of the motion field. In **情報処理学会研究報告 2003-CVIM-138**, pages 193–200, 2003.