

# 焼きなまし法を用いた次元縮小射影 Simple-Map の中心点探索

今村安伸\* 篠原武\*\*

(九州工業大学 \*情報工学府先端情報工学専攻 \*\*情報工学研究院)

## 1 はじめに

次元縮小射影 Simple-Map[1]では、いくつかの中心点との距離を新しい座標として射影を行う。良い中心点組を用いれば、少ない距離欠損で低次元への射影が可能となる。Simple-Map の中心点探索は、これまで様々に研究されてきた。乱択や局所探索の様な一般的な手法から、二値量子化法[2,3]の様な発見的探索法まで様々ある。また、過去に焼きなまし法[4,5]での中心点探索の研究も行われたが、良い結果は得られていなかった。今回、焼きなまし法へいくつかの改良を施すことにより、同研究室内のベストレコードを更新したので、これについて発表を行う。

## 2 原理

### 2.1 Simple-Map

距離空間においては、距離関数を  $d$  とすると、任意の点  $p, q, r$  に対して、以下の三角不等式が成立つ。

$$d(q, r) \geq |d(p, q) - d(p, r)|$$

ここで、 $p$  との距離はスカラー値となるので、1 次元への射影とみなせる。中心点を複数用いれば、中心点数に等しい次元数への射影が定義でき、射影後のノルムは  $L_2$  となる。

### 2.2 射影関数の評価関数

Simple-Map では、射影後の距離は元の距離を超えないが、縮むことはある。この縮みが小さいほど高性能と言える。そのため、Simple-Map の評価関数として、元の距離に対する射影後の距離の割合(距離保存率)を用いる。実際には、データベース件数  $N$  に対し  $N(N-1)/2$  組の中から  $N'$  組を無作為に抽出し、評価に用いる。

### 2.3 遷移候補の一括列挙

前述の評価関数では、1 つの中心点のみを変化させるとき、動かす中心点以外で最も縮みの小さな距離はまとめ計算可能である。また、そのうちの 1 つの次元のみを変化させるとき、 $L_1$  や  $L_2$  においては、次元ごとの差の絶対値や二乗値の総和を、変化させる次元以外についてまとめ計算可能である。これらを利用すると、1 つの中心点の 1 つの次元に関しては、複数の遷移候補値について評価値計算をまとめて行うことができ、遷移候補の一括列挙に適している。

### 2.4 ギブスサンプラーと焼きなまし法

ギブスサンプラー[6]はマルコフ連鎖モンテカルロ法の手法である。そもそも、焼きなまし法自体がマルコフ連鎖モンテカルロ法の 1 手法であるメトロポリス法[7]からの派生である。一括列挙に対しては、ギブスサンプラーの様な遷移確率を利用することで、通常の焼きなまし法の様に動作する。

### 2.5 採択確率式の変換

焼きなまし法における温度  $T$  において評価値差  $\Delta$  の候補が選択される条件は、一様乱数  $0 \leq \omega \leq 1$  に対して、 $\omega < e^{\Delta/T}$  である。これを変形すると、 $0 < \Delta - T \ln \omega$  である。よって、評価値が下がるときでも、その差が  $-T \ln \omega$  より小さいときは選択されると考えれば、 $-T \ln \omega$  をボーナスとする山登り法と等価である。複数の遷移候補を一括列挙した場合であっても、各々のボーナスを求めればよい。

### 2.6 再標本化と部分評価関数

本研究では  $N'$  組を評価に用いるが、各組の距離保存

率は部分評価関数と見なせる。ここで、 $N'$  組から  $N''$  組を無作為に再標本することを考える。 $N''$  組の再標本のスコア和は正規分布しており、近傍との評価差分も同様に正規分布することが分かる。また、ボーナスも片側正規分布なので、 $N''$  を徐々に  $N'$  に近付ける再標本による評価関数を用いた山登り法は、焼きなまし法と等価となる。

## 3 実験

動画のコマ送り画像から抽出した  $N = 700$  万件の 64 次元の特徴データに対して  $L_1$  距離を用いた、 $N' = 5000$  とし、当研究室での従来ベストレコードだった閾値を局所探索する二値量子化法[3]の実行時間である 1386 秒に収まる様に、焼きなまし法の遷移回数を 40 万回として、実験を行った。

	実行時間	スコア
従来ベストレコード	1386 秒	57.32
提案手法, 初期 $N'' = 320$	1268 秒	57.59
提案手法, 初期 $N'' = 164$	907 秒	57.58
提案手法, 初期 $N'' = 86$	640 秒	57.57

## 4 まとめ

遷移時に評価関数の再標本を用いた山登り法が、焼きなまし法と同じ特性を持つことを示した。また、そこまでの過程において、ギブスサンプラーが焼きなまし法においても有効であること、そして焼きなまし法の採択確率がボーナスと見なせることに言及した。また、実際に実験を行い、当研究室内のベストレコードを更新した。過去に行われた通常の焼きなまし法ではベストレコード更新に至っていなかったため、その原因究明が課題として残るが、今後、Simple-Map の応用射影である座標分割法[8]への適用など、適用範囲を広げていきたい。

## 参考文献

- [1] Takeshi Shinohara, Hiroki Ishizaka : On Dimension Reduction Mapping for Approximate Retrieval of Multi-dimensional Data, Progress in Discovery Science 2001, pp. 224-231, (2002).
- [2] 中島正八 : データベース内オブジェクトの離散化を利用した次元縮小射影 Simple-Map の中心点探索に関する研究, 九州工業大学卒業論文, (2014).
- [3] NGUYEN THI BICH HAU : 量子化を利用した次元縮小射影 Simple-Map の中心点探索に関する研究, 九州工業大学卒業論文, (2015).
- [4] Kirkpatrick, S.; Gelatt Jr, C. D.; Vecchi, M. P. (1983). "Optimization by Simulated Annealing". Science 220 (4598): 671-680.
- [5] 小川史宏 : 次元縮小射影 Simple-Map の中心点探索における焼きなましと局所探索の有効性に関する研究, 九州工業大学卒業論文, (2014).
- [6] Geman, S.; Geman, D. (1984). "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images". IEEE Transactions on Pattern Analysis and Machine Intelligence. 6 (6): 721-741.
- [7] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller: "Equation of State Calculations by Fast Computing Machines". J. Chem. Phys. 21, 1087 (1953).
- [8] 今村安伸 : 空間索引のための射影法 一座標分割と成分分解について一, 九州工業大学大学院修士論文, (2008).