

平成21年6月18日現在

研究種目：若手研究（B）
 研究期間：2006～2008
 課題番号：18700230
 研究課題名（和文） 低密度パリティ検査符号に基づく有歪圧縮アルゴリズムの符号化ダイナミクスの解析
 研究課題名（英文） Analysis of encoding dynamics for lossy compression based on low-density parity-check codes
 研究代表者
 三村 和史（MIMURA KAZUSHI）
 広島市立大学・大学院情報科学研究科・准教授
 研究者番号：40353297

研究成果の概要：

低密度パリティ検査符号に基づく有歪データ圧縮の符号化について性能評価を行った。復号に用いる疎行列の構成の方法がどのように性能に影響を及ぼすかなどを解析した。ダイナミクスの研究のための理論である経路積分法を拡張した。関連の深い誤り訂正符号である低密度生成行列符号の反復法に基づく復号アルゴリズムの性能評価など、その他の関連研究についても研究を行った。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,200,000	0	1,200,000
2007年度	1,000,000	0	1,000,000
2008年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,200,000	300,000	3,500,000

研究分野：総合領域

科研費の分科・細目：情報学，感性情報学・ソフトコンピューティング

キーワード：有歪圧縮，情報統計力学，経路積分法，レプリカ法

1. 研究開始当初の背景

本研究課題が対象とする有歪圧縮は，主に画像や音声を対象に歪みを許容する代わりに高圧縮を目的とした圧縮方法である。画像，音声，動画などで幅広く利用されている（下図）。例えば，画像では近い画素同士の色が似ているなど，情報の持つ偏りをうまく利用



圧縮前

圧縮後

することによって圧縮される。その一方で，より簡単な問題のように思える無記憶情報源の情報源系列については，まだ有効な圧縮の方法を多くの研究者が模索中といった段階にある。

許容する歪の大きさに対して，どの程度まで情報が圧縮できるかという限界は，理論的に示されている。この理論限界を達成するような，いくつかの有歪圧縮の方法が発見されている。これらの方法では，圧縮符号から元の情報を取り出す復号化をまず定義する。符号化は，圧縮符号の候補をすべて復号してみ，原情報と最も違いが少ない符号を探すことによって行われる。このため符号化に必要な計算コストが大きいという問題がある。

低密度パリティ検査 (LDPC) 符号に基づく

有歪圧縮方法（生成行列に疎行列を用いるため、低密度生成行列（LDGM）符号に基づく有歪圧縮方法とも呼ばれる）については、多項式時間で精度よく符号化が可能な高速符号化アルゴリズムが発見された。

一般には、復号した系列が情報源系列と同じ歪になるような符号語はたくさんある。この符号化アルゴリズムは、確率伝播法に反復履歴を導入することにより、探索する圧縮符号をひとつに絞り込む働きをもつ。すなわち、反復履歴が確率伝播法の動力学の慣性項の役割をはたしているのである。しかしながら、この慣性項は発見的に導入されたものであるため、その設計方法についての指針は明かではない。

効率のよい符号化の方法を調べていくためには、慣性項付き確率伝播法に必要な反復過程のダイナミクスを解析的に評価することが必要である。そこで、経路積分法や密度発展法などを用いてダイナミクスを解析することによって、慣性項付き確率伝播法の理解することを目的とする。また、関連する有歪圧縮法を提案し解析して、符号化の困難さについて検討する。これらによって、この符号化アルゴリズムを効率よく利用できるようになると考えられる。

2. 研究の目的

以下のような段階に問題を細分化して取り組む。

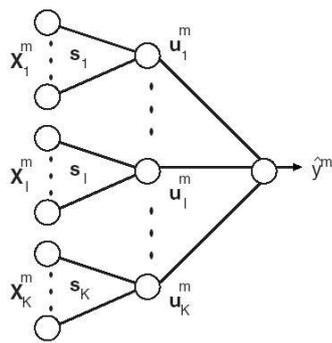
- (1) LDGM 符号に基づく有歪圧縮法の一般化を行なう。イレギュラーな生成行列を用いた場合の性能がどのようになるかを評価する。
 - (2) 性能が理論な限界を達成するような、他の有歪圧縮法を探索する。具体的には、多層パーセプトロンを復号に用いたアルゴリズムについて検討する。多層パーセプトロンは、コミティマシン、パリティマシンといったバリエーションが多く、復号器の複雑さが符号化の性能に与える影響について、調べることができるようになると期待できる。
 - (3) パリティマシン、コミティマシンを復号に用いた有歪圧縮法について、慣性項付き確率伝播法による符号アルゴリズムの性能を調べる。
 - (4) 経路積分法による密結合系の同期更新の反復計算のダイナミクスを解析する。経路積分法はダイナミクスを厳密に扱うことができる。数理的に似た構造を持つ
- CDMA マルチユーザ検出問題について、経路積分法を適用し性質を調べる。
- (5) 経路積分法による密結合系の確率伝播法のダイナミクスを解析する。先と同様に、数理的に似た構造を持つ CDMA マルチユーザ検出問題へ確率伝播法を適用したアルゴリズムについて、その性能評価を行なう。
 - (6) 経路積分法による疎結合系の同期更新の反復計算のダイナミクスを解析する。もっとも単純な SK モデルのダイナミクスを解析する。
 - (7) 誤り訂正符号の LDGM 符号について、近似的確率伝播法に基づく復号アルゴリズムを経路積分法によって解析する。
 - (8) 密度発展法による疎結合系の確率伝播法のダイナミクスを解析する。
- ## 3. 研究の方法
- (1) 復号に用いる疎行列について、行重み分布と列重み分布が与えられた場合の性能評価をレプリカ法によって解析する。
 - (2) 復号にコミティマシン及びパリティマシンを用いた有歪圧縮法を導入し、その性能をレプリカ法によって解析する。
 - (3) 慣性項付き確率伝播法による符号化アルゴリズムを導出する。このアルゴリズムを適用して、実際に圧縮を行い数値実験によって性能を調べる。
 - (4) CDMA マルチステージ検出器を用いてユーザビットを検出した結果（ビット誤り率）のダイナミクス（ステージごとのビット誤り率）を経路積分法によって解析する。
 - (5) 確率伝播法に基づく CDMA マルチユーザ復調をした場合の、ユーザビットを検出した結果（ビット誤り率）のダイナミクスを経路積分法によって解析する。
 - (6) 同期的に状態を更新する疎グラフ上で定義される SK モデルのダイナミクス（時間ごとのオーバーラップ）を経路積分法によって解析する。
 - (7) LDGM 符号について、近似的確率伝播法に基づく復号アルゴリズムをまず導入する。得られた復号アルゴリズムのダイナ

ミクス（反復回数ごとのビット誤り率）を経路積分法によって解析する。

- (8) 疎結合系の LDGM 符号に基づく有歪圧縮の符号化ダイナミクスを、密度発展法によって解析する。

4. 研究成果

- (1) これまでの解析では、レギュラー（行重みと列重みが一定）な疎行列を用いる場合が議論されており、これは圧縮率が整数比に限定されるものであった。イレギュラーな疎行列の場合、すなわち行重み分布と列重み分布を導入した場合を考えることで、任意の圧縮率の場合の圧縮を扱えるようになった。また、イレギュラーな疎行列を用いた圧縮をした場合の理論的な性能限界を与えることができる。
- (2) パリティマシンとコミティマシン（構造は下図）を復号に用いたとき、1RSB 仮定の下で性能は理論限界を達成することを示した。また、中間層や出力層を非単調な出力関数を用いると、偏りのある情報源系列に対応することができる。この非単調なパリティマシンとコミティマシンを復調に用いたときも、理論限界を達成することを示した。これらの復号器を用いる場合は、復号したときに同じ歪を与えるような符号語が中間素子の数を K とすると 2^K 個ある。このために、符号語の数と符号化のしやすさの関係についての類推が可能となった。

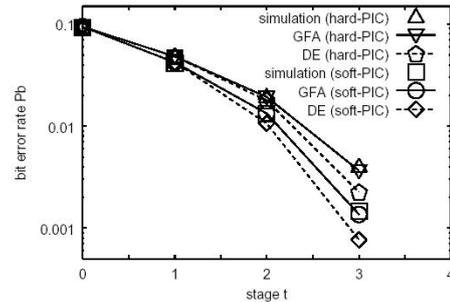


多層パーセプトロンの構造

- (3) 圧縮率を一定とした比較を行なうと、パリティマシン、コミティマシンの両方の場合で、中間素子数が増えるにしたがって歪が大きくなった。この結果は、慣性項付き確率伝搬法によって符号語を探索するときは、最も小さい歪を与える符号語は少ないほうが有利であることを

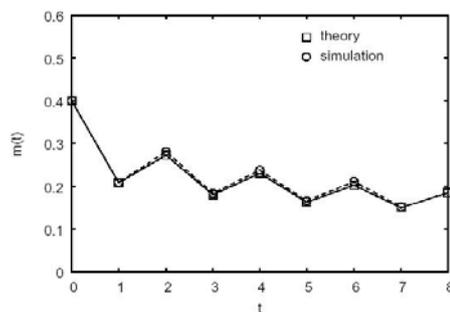
示唆する。

- (4) CDMA マルチステージ検出器について、硬判定ビットを用いた場合と、軟判定ビットを用いた場合を調べた。性能を経路積分法によって評価して、密度発展法と比較した（下図参照。simulation は計算機実験の結果、GFA は経路積分法の解析結果、DE は密度発展法の解析結果を表す）。経路積分法では、過去の反復履歴を理論に取り込むことができるため、反復過程の解析の重要な方法のひとつであることがわかる。



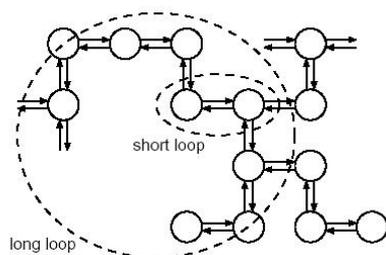
反復回数とビット誤り率の関係

- (5) 確率伝搬法に基づく CDMA マルチユーザ検出器の性能を、経路積分法によって解析した。密度発展法による結果は、2回の反復までは厳密で、その後は少し誤差を含むことを示した。また、経路積分法による解析結果から、十分に反復を繰り返したあとのビット誤り率を評価し、平均場近似の結果と一致することを確認した。ダイナミクスの研究がより明確な仮定の下での、平衡状態の解析としても有効であることがわかった。
- (6) 次数分布を与えることのできる疎なグラフ上で定義された SK モデルを経路積分法で厳密に解析する方法を示した。レギュラーグラフに対して、解析結果を適用して有効性を確認した（下図参照。○は数値実験、□は理論値を表す）。これによって、様々なグラフアンサンブルを厳密に扱うことが可能となった。

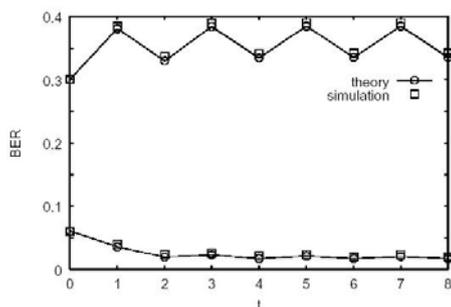


反復回数とオーバーラップの関係

- (7) LDGM 符号の近似的確率伝搬法に基づく復号アルゴリズムを導出した。また、その導出された方法を、(6)の方法によって解析した。この系の疎グラフは、長さが2のループのある鎖状のもので構成される(下図(上)はグラフの例)。このため、状態更新過程の履歴の影響が強い。レギュラーグラフの場合の、典型的なパラメータで数値実験を行なったところ理論とよく一致した(下図(下)参照)。このような、反復履歴の影響(ループの影響)を取り込んだ解析は、有限符号長の近似的評価としても有用である。



グラフの例



反復回数のビット誤り率の関係

- (8) この部分については、現在もひきつづき解析をおこなっている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 5 件)

1. K. Mimura, Typical performance of irregular low-density generator-matrix codes for lossy compression, *Journal of Physics A: Mathematical and Theoretical*, vol. 42, no. 13, 135002, 11 pages (2009).
2. K. Mimura, Parallel dynamics of continuous Hopfield model revisited, *Journal of the Physical Society of Japan*, 78, 3, 033001, 4 pages (2009).
3. F. Cousseau, K. Mimura, T. Omori, M. Okada, Statistical mechanics of lossy

compression for nonmonotonic multilayer perceptrons, *Physical Review E*, 78, 021124 (2008).

4. K. Mimura, M. Okada, Statistical mechanics of lossy compression using multilayer perceptrons, *Physical Review E*, 74, 026108, 11 pages (2006).
5. 三村和史, ニューラルネットワーク理論と符号, *日本神経回路学会誌*, vol. 13, no. 1, 19-27, (2006).

[学会発表] (計 16 件)

1. K. Mimura & A. C. C. Coolen, to be published in *Proceedings of the 2009 IEEE International Symposium on Information Theory (ISIT2009, Seoul, Korea)*, (2009).
2. K. Mimura, *Proceedings of the 2008 IEEE International Symposium on Information Theory and its Applications (ISITA2008, Auckland, New Zealand)*, 595-599 (2008).
3. F. Cousseau, K. Mimura & M. Okada, *Proceedings of the 2008 IEEE International Symposium on Information Theory (ISIT2008, Toronto, Canada)*, 509-513 (2008).
4. K. Mimura & M. Okada, *Proceedings of the 2007 IEEE International Symposium on Information Theory (ISIT2007, Nice, France)*, 2076-2080 (2007).
5. K. Mimura & M. Okada, *Proceedings of the 2006 IEEE International Symposium on Information Theory (ISIT2006, Seattle, US)*, 2373-2377 (2006).

その他, 国内会議 11 件

6. 研究組織

(1) 研究代表者

三村 和史
 広島市立大学大学院
 情報科学研究科・准教授
 研究者番号: 4 0 3 5 3 2 9 7

(2) 研究分担者

なし

(3) 連携研究者

なし