

自己評価報告書

平成 23 年 5 月 20 日現在

機関番号：32706
研究種目：若手研究（B）
研究期間：平成 20 年度～平成 23 年度
課題番号：20760250
研究課題名（和文） 十分統計量に着目した情報源および通信路に対するユニバーサル符号の理論解析と設計
研究課題名（英文） Theoretical Analysis and Design of Universal Codes Using Sufficient Statistics
研究代表者：有村 光晴（MITSUHARU ARIMURA）
湘南工科大学・工学部・講師
研究者番号：80313427

研究分野：工学

科研費の分科・細目：電気電子工学・通信・ネットワーク工学

キーワード：情報理論，データ圧縮，十分統計量

1. 研究計画の概要

本研究の目的は、無歪みデータ圧縮アルゴリズムを対象として、これまで以上に性能の良いデータ圧縮アルゴリズムを開発するための方法論を構築することを目指すことである。

その手段として、十分統計量を用いてアルゴリズムの理論的解析を行なうことで、圧縮アルゴリズムのユニバーサル性について、情報理論的・統計学的に統一的な解釈を与えることを目指す。具体的には、これまでに提案されている各種アルゴリズムを大きく FV 符号(固定長ブロック・可変長符号語符号)と VF 符号(可変長ブロック・固定長符号語符号)の 2 種類のクラスに分類し、それぞれのクラスの符号に対して、ユニバーサル符号と呼ばれている符号を十分統計量を用いて解釈することにより、その圧縮性能のユニバーサル性がどこから来ているのかを明らかにする。

2. 研究の進捗状況

FV 符号と VF 符号の両方について十分統計量を用いた解析を行なう予定であったが、初年度の研究によって、VF 符号の解析そのものがこれまで理論的に余り成されてきていないことが明らかになった。そのため、VF 符号については、ユニバーサル符号に関する解析ではなく、VF 符号そのものの圧縮性能に関する解析を先に行なった。

(1) これまで提案されている VF 符号のアルゴリズムの一つである Tunstall 符号について、これまで符号語長の期待値のみが解析されている。しかし十分統計量を用いた解析を行なう場合には、全体のデ

ータに対する平均値ではなく、個々のデータが明らかになった形で結果の式が出ている必要がある。そこで本研究では、これまでとは異なり Tunstall 符号の符号語長の個別冗長さの評価を経由することで、符号化レートが情報源のエントロピーに確率収束することを証明した。

- (2) Tunstall 符号は有限アルファベットに制限されたアルゴリズムであり、理論的な興味からは無限アルファベットの VF 符号の性能についても興味がある。そこで、可算無限アルファベットの一般情報源に対して、VF 符号で達成できる冗長さについて解析を行なった。
- (3) (2)の研究を行なっている過程で、FF 符号に対する冗長さを一緒に解析することとなったが、この解析により、情報スペクトル理論においてこれまで明らかになっていなかった、情報スペクトルの幅に対して操作的な新しい意味を与える事が出来た。

3. 現在までの達成度

②おおむね順調に進展している。

VF 符号に対する十分統計量を用いたユニバーサル性の解析については、VF 符号そのものの基礎的な解析ができたので、十分統計量を用いたユニバーサル VF 符号の解析全体としては、7 割程度の達成度と考えられる。残りの 3 割程度については、最終年度に進める予定である。一方、FV 符号に対する解析については、最終年度に進める予定である。

4. 今後の研究の推進方策

Tunstall 符号について、最終年度中に、これまでの結果をそのまま拡張する形で、データをブロックに分割して VF 符号化するユニバーサル符号化アルゴリズムを解析する予定である。これにより、十分統計量を用いた VF 符号の理論的な解析がある程度行なえる予定である。

また、FV 符号に対する解析についても、本年度は行なう予定であり、VF 符号に対する成果と合わせて、概ね予定した研究が達成できると考えている。

5. 代表的な研究成果

(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 1 件)

- ① Mitsuharu Arimura, On the Average Coding Rate of the Tunstall Code for Stationary and Memoryless Sources, IEICE Trans. Fundamentals, Vol. E93-A, pp. 1904-1911, 2010. (査読有)

[学会発表] (計 7 件)

- ① 古賀弘樹, 有村光晴, 岩田賢一, 情報スペクトルの幅と固定長符号化の最悪冗長度, 電子情報通信学会技術研究報告, No. IT-2010-83, pp. 93-98, 大阪大学, March 3-4, 2011.
2011 年 3 月 3 日発表 (査読無)
- ② Mitsuharu Arimura and Ken-ichi Iwata, A Converse Coding Theorem for Variable-to-Fixed Length Source Coding of General Sources, Proc. the 33rd Symposium on Information Theory and its Applications (SITA2010), pp. 349-352, Matsushiro, Nagano, Japan, Nov. 30-Dec. 3, 2010.
2010 年 12 月 1 日発表 (査読無)
- ③ Mitsuharu Arimura and Ken-ichi Iwata, The Minimum Achievable Redundancy Rate of Fixed-to-Fixed Length Source Code for General Sources, Proc. 2010 International Symposium on Information Theory and its Applications (ISITA2010), pp. 595-600, Taichung, Taiwan, Oct. 17-20, 2010.
2010 年 10 月 19 日発表 (査読有)
- ④ 有村光晴, 岩田賢一, 一般情報源に対する無歪み FF 符号の最小達成可能冗長度レート, 電子情報通信学会技術研究報告, No. IT2010-21, pp. 57-62, 工学院大学, July 22-23, 2010.

2010 年 7 月 22 日発表 (査読無)

- ⑤ Mitsuharu Arimura and Ken-ichi Iwata, On the Achievable Redundancy Rate of Fixed Length Source Code for General Sources, Proc. 2010 IEEE International Symposium on Information Theory (ISIT2010), pp. 126-130, Austin, TX, USA, June 13-18, 2010.

2010 年 6 月 15 日発表 (査読有)

- ⑥ 有村光晴, 岩田賢一, 一般情報源に対する FF 符号および VF 符号の冗長度レートについて, 電子情報通信学会技術研究報告, No. IT2009-136, pp. 413-418, 信州大学, March 4-5, 2010.

2010 年 3 月 5 日発表 (査読無)

- ⑦ 有村光晴, 定常無記憶情報源に対する Tunstall 符号の個別冗長度, 第 32 回情報理論とその応用シンポジウム (SITA2009) 予稿集, pp. 619-624, 山口県 山口市湯田温泉, Dec. 1-4, 2009.

2009 年 12 月 3 日発表 (査読無)

[図書] (計 1 件)

- ① 白木善尚編, 村松純, 岩田賢一, 有村光晴, 渋谷智治著, 情報理論, オーム社, 2008 年 9 月.
(pp. 89-135 を執筆)