

平成 30 年 6 月 15 日現在

機関番号：13601

研究種目：基盤研究(C) (一般)

研究期間：2014～2017

課題番号：26420353

研究課題名(和文) 情報源符号化における遅延と符号化レートの解析

研究課題名(英文) Analysis of delay and coding rate in source coding

研究代表者

西新 幹彦 (NISHIARA, Mikihiko)

信州大学・学術研究院工学系・准教授

研究者番号：90333492

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：本研究ではA*アルゴリズムを応用して最適な準瞬時可変長符号を探索する方法を確立した。最適な符号を見つけるためには最適な二つの符号木を見つける必要があるが、各符号木は情報源シンボルに順番に符号語を割り当てることによって符号を構成していく。割り当てる符号語は単に符号シンボルの列だけでなく語頭条件を満たすのか否かという付加情報を加えることによって特徴付けられる。このアルゴリズムでは、探索の過程で最初に発見した符号が最適符号であることが保証される。すなわち、さらに探索を続けても2番目以降に見つかる符号は最初に見つけた符号の性能を超えられないことを理論的に証明した。

研究成果の概要(英文)：In this research, we established a search method for an optimal almost instantaneous variable length codes based on the A* algorithm. We must find two optimal code trees to construct an optimal code. Each codeword is characterized not only by the sequence of code symbols but also by one-bit information whether the prefix-free condition is satisfied or not. This algorithm guarantees that the first code encountered through the search process is optimal. That is, we mathematically proved that the second and later codes are not better than the first code.

研究分野：情報理論

キーワード：情報理論 符号化遅延

1. 研究開始当初の背景

高度な情報処理が求められる現在の社会では、情報伝達の正確さとともにその早さも社会的に重要な役割をもっている。情報伝達の「速さ」も重要な指針の一つであるが本研究で主に扱うのは「早さ」である。早さとはレスポンスのことであり、社会ではスループットよりもレスポンスが重要な場面がある。一例として、緊急地震速報がある。地上デジタル放送はアナログ放送より約 2 秒遅れており、緊急地震速報も遅れていることが当初問題となった(「緊急地震速報の送信、地デジの遅れ解消へ」, 読売新聞, 2008 年 9 月 12 日)。この問題は緊急信号の処理を映像信号と分けることにより改善された。一般に、情報伝達における符号化効率と符号化に伴う遅延はトレードオフの関係にある。しかし、トレードオフの数理科学的解析や即時性を重視する通信のための符号化法には多くの課題が残されている。

情報源の符号化にブロック符号を用いる場合、そのブロック長はその符号化遅延を表している。したがってブロック長を短くすることが遅延の抑制にもつながる。しかし、ブロック符号を含む一般の概念として逐次符号がある。逐次符号は入力系列の先頭部分から符号系列の先頭部分を生成できることから、ブロック符号よりも符号化遅延を小さくできると認識されている。よってブロック符号の議論だけでは遅延に対する考察としては不十分である。

2. 研究の目的

情報源系列を固定した長さの文字列に分節し、各文字列を符号化するのが固定長ブロック符号である。一般に、ブロック長が長いほど符号化効率がよい。しかし、ブロック長が大きいと、復号器では長い符号語を受け取ってからでないとブロックを復号できないので、復号遅延が大きくなる。すなわち、符号化の効率と復号遅延はトレードオフの関係にある。このバランスの制御方法の確立に向けて、本研究では復号遅延の発生するメカニズムを多角的に捉え、情報源シンボルへの符号語の割り当て方と遅延の関係を調べることを目的とした。

3. 研究の方法

本研究では遅延発生メカニズムを多角的に捉えるため、いくつかの異なる仮定の下で符号化の問題を考えた。

1 番目の問題設定では、次のような符号化の過程を想定した。情報源から出力されるシンボルの時間間隔は一定で、到着したシンボルは符号器に入れられる。符号器は入力されたシンボルの内容ばかりでなく送信機の状態をも監視しながらどの時点でどのような

符号語を出力するのかが決める。符号語はその長さに比例する時間で送信される。符号語が送信されている間にもシンボルは到着し、新たな符号語が生成される。送信機が動作中は新たな符号語は送信バッファに入れられる。送信機は送信バッファが空になるとアイドル状態になる。このような状況の下で、分節木に基づく可変長符号を用いた時の遅延について実験的に検証した。具体的には、入力系列の分節は語頭条件を満たすものとし、分節の任意性を排除した。これは一意な分節により符号化レートを下げる効果がある。語頭条件を満たす分節は分節木を用いて表すことができるが、分節木の葉には符号語を割り当てなければならない。分節が一意であることから使用する符号語の集合として語頭符号を用いた。さらに分節木の兄弟ノードに割り当てられた符号語の語頭が共通している場合、その語頭を前倒して送信できることを活用し、分節木上へ符号シンボルの再配置を行った。兄弟ノードが共通してもつ語頭がなるべく長い方が、送信を前倒しできるシンボルが多いことになる。このことから語頭符号の中でもアルファベットと符号語の順序が同一となる順序保存符号を用いて検証実験を行い、ハフマン符号との比較を行った。

2 番目の問題設定では、符号器として算術符号を用いる状況を想定した。算術符号は逐次符号の一つであるが、理論的な動作原理として符号化の過程で実数値の算術演算を行う。しかし実際の符号器を設計するには、符号器と復号器で完全に同期が取れるように有限精度の演算を行う。この演算精度は遅延にも影響を与える。従来研究では、算術符号を用いた符号器と復号器の間に送信器はなく、どんなに長い符号語も瞬時に復号器に届くと想定されていた。そのもとで算術符号の精度を変えることによって遅延が変化することが示されている。本研究ではシンボルの到着の様子と遅延の関係を調べるため、シンボルの到着がポアソン到着する場合を実験的に検証した。

3 番目の問題設定では、演算精度の低い算術符号と準瞬時可変長符号の関係について調べた。準瞬時可変長符号は復号遅延の上界を指定したもとで設計できる逐次符号である。遅延の制御方法に関する知見を得るため、両者の符号化方法の対応関係を比較した。

4. 研究成果

1 番目の問題設定において、順序保存符号は送信を前倒しできるシンボルが多いことから、遅延を小さくする効果があると予想された。検証実験の結果、情報源の分布に対し、一定範囲の到着レートにおいて順序保存符号の遅延がハフマン符号の遅延よりも小さくなることを確認した。したがって、符号の順序保存が符号の復号遅延に影響を与えていることが確認できた。これにより、遅延を

制御するためには順序保存性も考えなくてはならないという知見が得られた。

2 番目の問題設定に対しては、様々な条件下で実験を行った結果、符号器と復号器が直接接続されている場合、平均遅延を特徴づけているのはシンボルの到着レートであり、到着間隔の分布そのものには影響されないことが分かった。また符号器と復号器の間に送信器と通信路を設置したモデルについても演算精度と遅延の関係を様々な条件下で実験的に検証した。このモデルでは、平均遅延を符号化遅延とバッファ遅延に分けて考えることができる。得られた結果から、符号化遅延は到着間隔の分布にはよらない一方、バッファ遅延は到着間隔の分布に依存することが分かった。いずれの場合も算術符号の演算精度を下げると遅延は小さくなったが、バッファ遅延については他と比べてその程度が小さかった。これは送信器の送信間隔と符号シンボルがバッファに入る間隔の違いに原因があることを示唆しており、遅延に関する重要な知見である。

3 番目の問題設定に対しては、いくつかの情報源シンボルの分布に対して最適な準瞬時可変長符号の構成例が知られていることから、その分布に対して算術符号を構成した。二つの符号は入力シンボルに応じて状態遷移するオートマトンとみなすことができ、さらに詳しくは各状態に符号木が対応する。演算精度の低い算術符号から導かれたオートマトンの各状態の符号を見ると、語頭条件を満たさないということが確認された。2 状態の場合において、符号木の中間ノードまたは葉に対応している符号語が出力されるとき、それぞれ状態 1 または状態 2 に遷移することも確認された。それぞれの符号は語頭条件を満たさないが、もともと算術符号であるということから、誤りなく復号できることは保証されている。また、状態 2 の符号器には特定のパスが存在しないことも確認された。これは算術符号における区間と出力系列の対応から説明することができる。また、算術符号の復号遅延の特徴付けが情報源系列の区間と符号語系列の区間の包含関係という概念を用いて表現できるという見通しが得られた。特に、復号遅延がゼロであることは情報源系列の区間と符号語系列の区間が等しいことによって特徴付けられることが分かった。

算術符号と準瞬時可変長符号を比較すると共通点が多く、算術符号が準瞬時可変長符号と本質的に同じ符号木によって表せる場合があることが分かった。一方、算術符号の構成方法には多様性があり、場合によっては符号木の数(状態数)を増やすこともできる。このことによって符号の性能が準瞬時可変長符号を超えられるかどうかははっきりしない。準瞬時可変長符号と算術符号が本質的に同等であるのか、どういう条件において同等となるのかについて調べるためには、両者そ

れぞれに最適な符号を探索することが必要である。このことから、最適な準瞬時可変長符号を探索する問題に取り組んだ。

準瞬時可変長符号は複数の符号木を用いることによってハフマン符号よりも小さい符号化レートを達成できる符号として知られており、最適な準瞬時可変長符号は整数計画法を用いることで見つけられることが知られている。一方、符号語が互いに語頭・語尾にならないリバーシブル可変長符号の最適符号を探索する方法として A* アルゴリズムを応用することが提案されている。本研究では A* アルゴリズムを別の形で応用して最適な準瞬時可変長符号を探索する方法を確立した。最適な符号を見つけるためには最適な二つの符号木を見つける必要があるが、各符号木は情報源シンボルに順番に符号語を割り当てることによって符号を構成していく。割り当てる符号語は単に符号シンボルの列だけでなく語頭条件を満たすのか否かという付加情報を加えることによって特徴付けられる。このアルゴリズムでは、探索の過程で最初に発見した符号が最適符号であることが保証される。すなわち、さらに探索を続けても 2 番目以降に見つかる符号は最初に見つけた符号の性能を超えられないことを理論的に証明した。

5 . 主な発表論文等

〔雑誌論文〕(計 1 件)

- (1) Mikihiro Nishiara, Ryo Hidai, Decoding Error of Sudoku for Erasure Channels, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.E100-A, no.12, pp.2641-2646, Dec. 2017, 査読有 .

〔学会発表〕(計 14 件)

- (1) 山川誠史、西新幹彦、最適な 2 元 AIFV 符号の幅優先探索アルゴリズム、電子情報通信学会情報理論研究会、2017 年 .
- (2) 打田尚大、西新幹彦、順序保存性のない算術符号の最適性に関する基礎的考察、電子情報通信学会信越支部大会、2017 年 .
- (3) 山川誠史、西新幹彦、最適な 2 元 AIFV 符号の探索アルゴリズム、電子情報通信学会信越支部大会、2017 年 .
- (4) 西新幹彦、着払いコスト制約付き通信路符号化について、第 10 回シャノン理論ワークショップ、2017 年 .
- (5) 西新幹彦、微小な復号誤りを許容する着払いコスト制約付き通信路符号化について、第 40 回情報理論とその応用シンポジウム、2017 年 .
- (6) Mikihiro Nishiara, Ryo Hidai, Decoding Error of Sudoku for Erasure Channels, International Symposium on

Information Theory and Its Applications (ISITA2016), 2016, 査読有.

- (7) 山川誠史、西新幹彦、算術符号と AIFV 符号の関係について、電子情報通信学会信越支部大会、2016 年.
- (8) 山川誠史、西新幹彦、演算精度の低い算術符号の状態遷移について、第 39 回情報理論とその応用シンポジウム、2016 年.
- (9) 西新幹彦、原寛貴、最大歪み測度を用いたレート・歪み問題に関する一定理、第 9 回シャノン理論ワークショップ、2015 年.
- (10) Daobin Liu, 西新幹彦、状態数より少ない符号器を用いた情報源符号化について、電子情報通信学会信越支部大会、2015 年.
- (11) 飯島一貴、西新幹彦、パケット間隔で情報を送る通信路に対する最尤復号のための距離関数、電子情報通信学会信越支部大会、2015 年.
- (12) 西新幹彦、渡部健人、一人のエージェントを雇った CEO 問題におけるレート・歪み関数、第 38 回情報理論とその応用シンポジウム、2015 年.
- (13) 飯島一貴、西新幹彦、パケット間隔で情報を送るシステムの無記憶離散通信路としての通信路容量、電子情報通信学会信州大学学生ランチ研究発表会、2015 年.
- (14) 荻野真志、西新幹彦、遅延特性改善のための分節木上の符号語の配置に関する考察、電子情報通信学会信越支部大会、2014 年.

6. 研究組織

(1) 研究代表者

西新幹彦 (NISHIARA, Mikihiko)
信州大学・学術研究院工学系・准教授
研究者番号 : 90333492