1.3. 符号化におけるロバスト計算

山本博資　　小川朋宏

情報理工学系研究科 数理情報学専攻

概 要

情報の符号化技術は，情報を「効率良く，高品質で，安全に」伝送および記録するために必要不可欠な技術である．これらの要求を満たすために，それぞれ「データ圧縮符号化，誤り訂正符号化，暗号化」技術があるが，本研究ではデータの特性や通信路の状態，不正者からの攻撃方法，あるいは計算機の計算精度によらず，上記の目的をロバストに達成できる符号化技術の開発を目的としている．

1　はじめに

情報化社会を支える必要不可欠な要素として，情報を伝送するための通信路と情報を記録するための記録媒体がある．これらは，情報技術の進歩により，光通信や光磁気ディスクなどが実用化され，大容量で高品質な通信や記録が可能となって来ている．しかし，その一方で，携帯電話の利用を始めとするユビキタスなネットワーク社会では品質の悪い公開の無線通信路を通して情報を伝送し，より小さな記録媒体に大量の情報を記録することが要求されている．このように，性能の悪い通信路や記録媒体を，高品質で安全に効率良く使用できるようにするロバストな符号化技術が今後もますます重要度を増てくるものと思われる．

符号化技術は大きく，次の３つに分類される．

(A) データ圧縮符号化： データ系列を，より短いビット長で表現できるように符号化する．

(B) 誤り訂正符号化： 通信路の雑音あるいは記録媒体のゴミや埃で，正しく受信あるいは読み出しができない場合に，誤りを検出し自動的に訂正する．

(C) 暗号化： 盗聴や改ざんなどの攻撃から，情報を守る．

これらの符号化において，次のような要因により符号化の性能が悪化する．

(a) データの確率構造や通信路の確率構造が未知である．あるいは時間的に変動する．

(b) 符号化, 復号化において正確な計算が困難であり，近似計算を必要とする．

(c) 不正者によるさまざまな攻撃により，情報の安全性が劣化する．

本研究では，このような性能悪化要因にも関わらず，ロバストに「効率良く，高品質で，安全に」符号化が行える技術の開発を目指す．

平成 14 年度は，データ圧縮, 誤り訂正, 暗号の各項目に対して，解決すべき問題点を明らかにすると共に，その幾つかの問題点の解決法を与えた．平成 15 年度〜17 年度は，これらの問題点およびその解決法について理論的に詳細な検討を行うと共に，符号の実装などによる性能評価を行う．最終年度の平成 18 年度には，それらの結果をまとめて，理論の体系化を行う予定である．

次節以後で，各符号化に対する本年度の成果を報告する．

2　データ圧縮符号化

データ圧縮において，対象とするデータの確率構造によらず同一の符号で効率よく圧縮を行える

符号をユニバーサル符号という．ユニバーサル符号は,「ユニバーサルモデル化＋算術符号」を用いる方法と「辞書法, ソート法, 文法法」など，符号化や復号にデータの確率構造を陽に用いない符号化法に分類できる．

前者では，算術符号化における近似計算が，符号化の性能に大きく影響し，特に連続値データやアルファベットサイズが非常に大きい場合は，性能が大きく悪化し易い．このような問題は，音声波形や地震波形などの時間波形のディジタルデータを符号化する場合に特に問題となるが，本研究では，そのような場合に対してもロバストな符号化を行える圧縮方式の検討を行っている．

まず，音声データの圧縮に対しては，差分符号化の残差分布がラプラス分布で近似できることを利用し，有限精度で効率よく符号化できる方法を提案した [1]．これは，正整数のユニバーサル表現である Elias の γ 符号を利用することにより，加算無限個のアルファベットを有限精度で符号化できるように工夫したものであり，従来の Golomb 符号を用いる場合に比べて，一般に大きな改善が得られる．また，地震波に対しては，その従来符号化法を調査し，その問題点の改善を検討している．

ユニバーサル符号の文法法は，従来の符号化法とは異なり，データ系列を生成できる文法規則に符号化する方式であり，多くの研究者の注目を集めている．その中で MPM(Multiple Pattern Matching) 符号は，データ長 n に対して $O(1/\log n)$ の最悪冗長度を達成できるため，特に注目されている．しかし，MPM 符号は長い系列を一度に符号化するためオフライン的な符号である．また，漸近的には $O(1/\log n)$ の最悪冗長度を達成できるものの，実ファイルに対する圧縮性能が悪い欠点を持つ．本研究では，オフライン符号化ではなく，オンラインで高速に符号化が可能な逐次 MPM 符号の提案とその性能評価を行っている [2]．また，漸近的に最悪冗長度 $O(1/\log n)$ を達成しながら，実ファイルに対してもロバストに性能のよい符号化が可能となる改良方式についても検討を行っている．

なお，[1] と [2] は，それぞれ修士 1 年の番伸宏君と石井邦憲君との共同研究である．これらの問題に対して，平成 15 年度以降もより詳しい検討を行う予定である．

3　誤り訂正符号化

最近，誤り訂正符号として，ターボ符号 (Turbo 符号) と低密度パリティ検査符号 (LDPC 符号) が多くの研究者の注目を集めている．これらは共に，Shannon 限界に近い誤り訂正能力を有しており，復号法に Belief Propagation などの近似計算アルゴリズムが用いられている．

ターボ符号は，通常，通信路の SNR 特性が分かっているものとして，復号が行われている．しかし，多くの通信路では，SNR が未知であったり，SNR が時間と共に変動したりする．そのような場合には，SNR を精度よく推定すると共に，SNR の変動によらずロバストに許容復号誤り率を達成できる符号化方式の開発が要求される．SNR の大きな変動に関わらず，ロバストに低誤り率を達成するには，ターボ符号に ARQ（自動再送要求）を組み合わせたターボ符号 ARQ 方式が望ましい．これは，ターボ復号を行っても誤りが残ってしまった場合に，再送によりその誤りを訂正する通信方式である．

従来のターボ符号 ARQ 方式では，ターボ復号に誤りが残っているか否かを，内符号を用いて検出していたため，繰り返し復号を最後まで行わないと，再送要求の判断ができないという欠点があった．そのため，復号計算量が多いというターボ符号の欠点が，さらに悪化されていた．これに対して，本研究ではそのような欠点のない新しいターボ符号 ARQ 方式を提案した [3]．その提案方式では，受信系列の SNR の推定値が，正しく復号できる可能性のない場合は，最初から再送を要求し，無駄な復号計算を省いている．また，繰り返し復号の各段階で求まる復号系列に対する受信系列の SNR を用いることにより，繰り返し復号系列が，正しい送信系列に近付いているかどうかを判断できることを示し，その判断基準を用いて再送要求を行う方式を採用している．さらに，数理実験に

より，従来方式に比べて，低い復号誤り率を少ない復号回数で達成できることを示した．また，スループットの低下は従来方式に比べて若干生じるが，内符号による誤り検査を併用することにより，スループットの低下もなくせることを示した．

今後，この方式に関してより詳細な検討を行う予定である．

LDPC 符号は，誤り訂正符号として高性能であることが広く知られているが，その誤り特性のロバスト性と，誤り訂正符号とデータ圧縮符号の双対性を利用して，有歪み圧縮への応用を検討している．

なお，[3] の研究は，修士２年の山村聡君との共同研究である．

4　情報セキュリティ

通信路や記録媒体への敵からの攻撃に対して情報を安全に守るために，数多くの暗号技術が開発されている．しかしそれらの多くは，敵に知られていない秘密鍵に安全性の根拠を置いているため，秘密鍵に関して何らかの情報が敵に知られたとき，新しい鍵を再度共有し直す必要が生じる．そのような場合に，敵がどのような情報を得ていても，その情報量がある値以下であれば，元の秘密情報と公開通信路上の情報交換で，敵が全く知ることのできない新しい秘密情報を作り出すことができる．その手法を「プライバシー増幅」という．

従来のプライバシー増幅は，正規の送信者かどうかを判断する認証部と新しい秘密を生成する秘密生成部とを同じ長さで使用していたため，最適な方式ではなかった．本研究では，攻撃者が知っている情報量に応じて，最適な分割法が存在することを示した．その結果として，敵の攻撃の種類によらず，ロバストかつ効率的にプライバシー増幅を行う方法の検討を行っている．

記録媒体の安全性を保つためには，情報の盗難だけでなく破壊や記録装置の故障にも対処しなければならない．前者の脅威には，秘密情報のコピーを作ることにより対処できるが，秘密情報のコピーを複数作ると，後者に対する脅威が大きくなってしまう．そのような相矛盾する問題を解決し，盗難と破壊の両方の脅威に対してロバストに強い記録方式として秘密分散法がある．(k,n) しきい値秘密分散法では，秘密情報を n 個の分散情報に分散符号化し，そのうちの任意の k 個が集まると秘密情報が復号できるが，任意の $k-1$ 個では全く秘密情報が復号できない符号化法である．(k,n) しきい値法は，全ての分散情報が対等な働きをするが，一般のアクセス構造を持つ秘密分散法では，秘密情報を復元できる分散情報の集合族（アクセス集合族）と復元できない集合族（非アクセス集合族）を自由に設定することにより，より柔軟な秘密保持特性を達成できる．しかし，その自由度の大きさにより，最適な符号化効率が求められておらず，その結果最適な符号化方法を求める手がかりも明確になっていない．

本研究では，最適な一般アクセス構造の符号化効率の理論的限界を明らかにすると共に，その実用的な符号化方法についても検討を行っている [6].

通常の秘密分散法は，符号化および復号化にコンピュータを用いて複雑な計算を行う必要がある．しかし，地震等でコンピュータが使用できないような非常時でもロバストに秘密情報を復元したい場合がある．そのような場合に使用できる秘密分散法として，視覚復号型秘密分散法がある．視覚復号型秘密分散法では，秘密情報を画像として記録し，それを符号化した分散情報を，透明なシートに印刷して記録する．復号は，アクセス集合に属するシートを重ねるだけで秘密情報が得られるが，非アクセス集合に属するシートからは，秘密情報が全く漏れない方式である．

視覚復号型秘密分散法の従来方式では，秘密画像として，単一の白黒画像，白黒濃淡画像，カラー画像，あるいは複数の白黒画像が取り扱われていた．本研究では，複数のカラー画像や濃淡画像を秘密画像として取り扱える，一般的な符号化方法を理論的に与えると共に，具体的に符号化を行うことにより，その有効性を明らかにした [5]

今後，これら秘密分散法に関してより詳細な検討を行っていく予定である．なお，[5][6] は，博士課程２年の岩本貢君との共同研究である．

5　量子符号化

将来の符号化技術として，量子通信路や量子データに対する符号化技術が重要である．本研究では，そのような符号化問題のうち,「量子仮説検定問題における誤り確率の大偏差型評価」を行っている．

量子仮説検定とは，二つの量子状態を量子力学的観測によって統計的にロバストに識別するための手法である．これは量子系の非可換性に起因する不確定性が最もシンプルに現れる問題であり，量子通信路符号化，量子情報源符号化，エンタングルメント純粋化などの様々な符号化定理を論ずる上で重要な基礎となっている．また，量子コンピュータにおいてノイズの混ざった計算結果をロバストに読み出すためにも必要な手法である．

古典的な情報理論においては，情報スペクトル的方法が様々な符号化定理に関する見通しの良い議論を与える．そこでは，[符号化レートや誤り指数の最適値] = [情報スペクトル] = [エントロピーやダイバージェンスなどの情報量] というように二つの等式を介して符号化定理が証明される．一番目の等式は極めて一般的な状況で証明され，この時点で符号化や検定の最適化がなされる．ただし「情報スペクトル」自身は極限を用いて定義されているため，符号化レートの具体的な計算には右辺を用いる必要がある．そのための二番目の等式は，大数の法則や大偏差原理 (Sanov の定理や Cram'er の定理) などの，純粋に確率論的な定理の帰結として得られる．

量子仮説検定においては，第一種誤り確率と第二種誤り確率をともに指数的に減少させていく場合の誤り指数に関して，古典論と同様な情報スペクトル的公式 (一番目の等式) が最近知られている．しかし，二番目の等式の成立に関しては，本研究で与えた大偏差型の評価式以外には知られていない．この評価式は古典系においては最適なものであり，二番目の等式の証明を与えるが，量子系においては準最適としか言えない．これは量子系における大偏差原理の未整備によるものである．したがって，大偏差原理の視点から二番目の等式に関する研究を行うことが重要であると考える．

また，古典系においては仮説検定やパラメータ推定をはじめとする様々な統計的問題に対して，情報幾何的な手法が有効であることが知られている．中でも仮説検定の誤り指数に関する情報幾何的描像は，大偏差原理と情報幾何の関係を示す基本的な例である．量子仮説検定の誤り指数決定問題は，量子系における大偏差原理と情報幾何学のあるべき姿を探る意味でも重要な問題であり，今後より詳細に検討を行う予定である．

なお，量子符号化に関する研究は，主に小川朋宏が担当している．

参考文献

[1] 番伸宏, 山本博資, “幾何分布およびラプラス分布のための算術符号化アルゴリズム,” 第 25 回情報理論とその応用シンポジウム予稿集，pp.143-146, Dec. 2002

[2] 石井邦憲, 山本博資, “ボトムアップアルゴリズムに基づく遂次 MPM 符号,” 第 25 回情報理論とその応用シンポジウム予稿集，pp.147-150, Dec. 2002

[3] 山村聡, 山本博資, “SNR 推定値を用いたターボ符号 ARQ 方式,” 電子情報通信学会技術報告，情報理論研究会，March 2003

[4] 坂井秀行, 山本博資, “木構造を用いたグループ鍵更新方式に対する性能解析,” 電子情報通信学会技術報告，情報セキュリティ研究会，March 2003

[5] M.IWAMOTO and H.YAMAMOTO, “A Construction Method of Visual Secret Sharing Schemes for Plural Secret Images, ” (submitted to IEICE Transactions on Fundamentals)

[6] 岩本貢, 山本博資, “一般アクセス構造に対する非理想的ランプ型秘密分散法,” 第 25 回情報理論とその応用シンポジウム予稿集，pp.227-230, Dec. 2002