

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 5月31日現在

機関番号：12608
 研究種目：基盤研究（C）
 研究期間：2010～2012
 課題番号：22560362
 研究課題名（和文） LDPC行列を用いたSlepian-Wolf符号化システムの実現と性能評価
 研究課題名（英文） Practical Slepian-Wolf Coding System by using LDPC matrices
 研究代表者
 植松 友彦（UYEMATSU TOMOHIKO）
 東京工業大学・大学院理工学研究科・教授
 研究者番号：60168656

研究成果の概要（和文）：

Slepian-Wolf システムにおいて LDPC(Low Density Parity Check)行列による実用的な符号について、以下の符号の構成法を得た。(1) 情報源の確率分布が未知でも符号化率が与えられれば構成でき、任意に小さい誤り率を達成する符号。(2) 相関を有する2つの情報源が二元対称通信路を通じて結びつけられている場合、最尤復号に迫る復号性能が線形時間で行える可変符号化率の符号。(3) 2つの符号器間の同期のずれがあっても、符号化や復号が行える符号。

研究成果の概要（英文）：

This research investigates some practical codes for Slepian-Wolf system by using LDPC(Low Density Parity Check) matrices, and obtains the following codes. (1) The code which can be constructed without any statistical property of sources, and can achieve arbitrarily small error probability. (2) The variable rate code which can be decoded with linear complexity of its length. (3) The asynchronous Slepian-Wolf code.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	500,000	150,000	650,000
2011年度	1,400,000	420,000	1,820,000
2012年度	600,000	180,000	780,000
年度			0
年度			0
総計	2,500,000	750,000	3,250,000

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：Slepian-Wolf 符号化 LDPC 符号 情報源符号化

1. 研究開始当初の背景

複数の情報源の符号化を論じる多端子情報理論における最も基本的なモデルとして Slepian-Wolf システムがある。このシステムでは相関を持つ2つの情報源からの一定長の出力列をそれぞれ別々の符号器によって独立に符号化し、復号器は2つの符号器の出

力系列を用いて2つの情報源から出力された系列を復号する。このシステムは、1972年に Slepian と Wolf によって提案され、その性能の解析が行われた。その結果、独立に情報源を符号化しているにもかかわらず、一度に符号化する情報源からの出力列が十分長ければ、符号器が協力して符号化した場合

と同一の圧縮率を任意に小さな誤り率で達成できることが示された。しかしながら、SlepianとWolfはそのような符号の存在を証明しただけであり、実際に最良の圧縮率を任意に小さい誤り率で達成する符号を構成する方法は未知であった。1990年代に入ってから、申請者らを含む幾つかのグループにおいて最良の圧縮率を達成できる符号の構成法が示された。2005年には、二元無記憶情報源に限った場合、そのような符号が、列重みの制限されたLDPC（低密度パリティ検査）行列を用いて符号化を行う線形符号の中に存在することを申請者らのグループが示した。しかしながら、このように符号のクラスがかなり制限されているにもかかわらず、実用的な符号の構成法は未知のままであった。

2. 研究の目的

本研究では、Slepian-WolfシステムにおいてLDPC行列による実用的な符号の存在を明らかにすると共に、そのような符号の設計指針を与えることを目的としている。具体的には、Slepian-Wolfシステムに用いる符号の構成には次のような課題があり、これらの課題を解決する。

(1) 情報源の統計的性質が未知でも符号が構成できること。

これまでのLDPC行列を用いた符号に関する研究では、符号構成の際に情報源の統計的性質を知っている必要があるため、情報源の統計的性質が分からなくても構成できるLDPC行列を用いた符号の存在を明らかにする。

(2) 構成された符号が符号長の線形時間で符号化や復号が可能なこと。

通常、符号長を長くするほど誤り率が小さくなるため、可能な限り符号長を長くして符号化や復号を行いたい。しかしながら、符号長に伴って符号化や復号にかかる計算量が指数関数的に増加すれば、そのような符号は現実的には利用することが出来ない。そのため、たとえ情報源の統計的性質を使わずに符号が構成できたとしても、符号長に伴う線形時間で符号化や復号が可能でなければ実用的な符号とは言えない。他方、LDPC行列を用いた符号の研究は、通信路符号化において活発に議論されており、LDPC行列を用いた通信路符号では線形時間で最尤復号に迫る復号が行えることが知られている。この成果がLDPC行列を用いた本研究においても利用できるかを明らかにする。

(3) 一つのLDPC行列を用いて可変符号化率の符号を構成すること。

LDPC行列を用いた符号では符号の符号化率が変わるごとに異なる行列を用意する必要があり、情報源に応じて符号化率を変化させるようなシステムでは、符号器で用意しておく行列の数が増加し、符号器の構成が複雑

になる。これを解決するには、例えば行列を一つだけ用意して、符号化率に応じて、その行列の一部の行を削除して符号化すれば、符号器はただ一つの行列のみを記憶しておけばよいので、符号器の構成は簡単になる。このような可変符号化率を有する符号のLDPC行列を用いた構成可能性について検討する。

3. 研究の方法

平成22年度は、現実的な符号の構成のために理論面を中心とした研究を行う。具体的には、研究の目的で最初の課題として挙げた、情報源の統計的性質が分からなくても最良の圧縮率を任意に小さい誤り率で達成できる符号がLDPC行列を用いた符号のクラスの中に存在するかどうかを明らかにする。

平成23年度は、LDPC符号を用いた通信路符号化における幾つかの最近の成果がSlepian-Wolf符号化にも応用できないかについて検討を行い、次の2つの課題を検討する。

(1) 符号長に伴う線形時間で符号化や復号の可能性の追求

この課題は符号の実用面において最重要であり、通信路符号化における幾つかの成果の応用可能性について検討する。特に、LDPC行列を用いた通信路符号化においては線形時間で最尤復号に迫る復号性能が得られることが知られており、この成果がSlepian-Wolfシステムに対しても利用できないかについて検討する。

(2) 一つのLDPC行列であらゆる符号化に対応できるような符号の構成

この課題については、AhlswedeとDueckによる符号の階層的構成法が応用できないかについて検討する。彼らは、通信路符号化において、符号語の候補を用意し、その中から符号語を1個ずつ付け加え、付け加える度に得られる符号もまた通信路容量を達成できる良い符号になることを示した。すなわち、彼らの構成法で構成した符号の部分符号もまた通信路容量を達成できる良い符号になっており、符号を一つ用意すれば、その部分符号を用いることで任意の符号化率に対応できることを示した。この成果をSlepian-Wolfシステムに応用できないかについて検討し、符号を階層的に構成することで、どのような符号化率にも対応できる符号を構成する。

平成24年度は、2つの符号器の間で同期が取れない場合の符号化法について検討する。実際にSlepian-Wolfシステムを構成した場合、2つの符号器は互いに独立に符号化を行うので、同期を取ることが常にできるとは限らない。また、情報源における相関にも時刻のずれが生じていることも考えられる。一方、復号器においては、符号器の間で生じた同期

のずれが系列長を超えない範囲で未知であるとする。このような状況下でも符号化や復号が行える符号を構成することは、Slepian-Wolf システムの実現に向けて重要であり、具体的には情報源の統計的性質や同期のずれの大きさを知らずに符号化ならびに復号が行える符号の構成法を検討する。

4. 研究成果

平成 22 年度は、相関を有する定常無記憶情報源 (X, Y) の確率分布 Q を知らなくても 2 つの符号器の符号化率が与えられれば構成でき、かつ情報源 (X, Y) の達成可能領域に符号化率が含まれていれば任意に小さい誤り率を達成できる符号を、LDPC 行列を用いて構成する方法について検討を行い、次の成果を得た。

- (1) 相関を有する定常無記憶情報源 (X, Y) の達成可能領域を

$$R(X, Y) = \{(R1, R2) : R1 \leq H(X|Y), \\ R2 \leq H(Y|X), R1+R2 \leq H(X, Y)\}$$

によって定める。但し、 $(R1, R2)$ は 2 つの符号器の符号化率を表し、レート対と呼ぶ。また、 $H(X|Y)$ と $H(Y|X)$ はそれぞれ条件付きエントロピーと同時エントロピーを表す。予め定められたレート対 $(R1, R2)$ を達成可能領域内に含む有限個の定常無記憶情報源からなる混合情報源に対し、誤り率が符号長のマイナス 3 乗のオーダーよりも速く減衰する Slepian-Wolf 符号が構成できれば、構成した符号は、符号化率を達成可能領域に含む任意の定常無記憶情報源に対しても、符号長を長くしたとき任意に小さい誤り率を達成できることを明らかにした。

- (2) 上記の混合情報源に対し、誤り率について所望の条件を満足する LDPC 行列を用いた Slepian-Wolf 符号の存在を証明し、その結果として、情報源の確率分布を知らずに、符号化や復号が行えるユニバーサルな Slepian-Wolf 符号が存在することを明らかにした。

平成 23 年度は、実用的な符号の構成のための理論について研究を行い、以下の成果を得た。

- (1) 符号長に伴う線形時間での符号化・復号の可能性の追求

最新の通信路符号化における幾つかの成果が Slepian-Wolf 符号化にも応用できないかについて検討を行った。その結果、相関を有する 2 つの情報源 (X, Y) が二元対称通信路を通じて結びつけられている場合、Schonberg らや Stankovic らによる符号構成法を用いれば、通信路符号化法における復号法がそのまま適用できることを明らかにした。特に、LDPC 行列を用いた通信路符号化では線形時間で最尤復号に迫る復号が行える

ことが知られており、この成果が Slepian-Wolf 符号化システムに対しても直接利用できることが分った。また、符号化を線形時間で行う手法についても、LDPC 行列を用いた通信路符号に対する成果が応用できることを確認した。

- (2) 一つの LDPC 行列であらゆる符号化率に対応できるような符号の構成

通信路用の線形符号を一つ用意し、その符号の生成行列を 2 つの符号器に分け与えることで、符号化レートの和 $R1+R2$ が一定という条件下で達成可能領域における任意の点を実現する符号化法が Schonberg らによって既に提案されている。彼らの符号化法は、誤り訂正符号の効率的な復号法が利用可能である点と、レート対 $(R1, R2)$ 毎に符号構成を行わなくて良いという点で実用上重要であるが、一様分布に従う 2 つの二元情報源の相関が二元対称通信路で与えられるときのみ、その理論的な最適性が示される。他方、この符号に対し、符号長を長くしたときに任意に小さい誤り率を達成できるためのレート対の取りうる値の範囲や、符号長を長くしたときに、誤り率が指数関数的に減少する速度を示す誤り指数の解析などの理論解析は行われていなかった。そこで、Schonberg らの成果を非一様な分布に従う二元情報源に対して用いることを検討した。このとき、符号長を長くしたときに任意に小さい誤り率を達成するために符号化レート対 $(R1, R2)$ が満足すべき条件が

$R1 \leq H(Y|X), R2 \leq H(X|Y), R1+R2 \leq H(X, Y)$ であることを明らかにすると共に、その誤り率 Pe が

$$Pe \leq \exp\{-N E(R1, R2, Q)\}$$

を満足することを示した。但し、 N は符号長を表し、 Q は情報源 (X, Y) の確率分布である。ここで、誤り指数 $E(R1, R2, Q)$ は、

$E(R1, R2, Q) = \max \min [Er(Rc, e), Er(Rx, p)]$ によって定まり、 p は情報源 X が記号 0 を出力する確率、 e は 2 つの情報源 X と Y の出力記号が一致しない確率、 $Er(R, q)$ は誤り率 q を有する 2 元対称通信路における Gallager の誤り指数、 \max は $Rx+Rc=R1+R2$ を満足する (Rx, Rc) について取られる。更に、符号化レート対 $(R1, R2)$ の選び方によっては、従来知られていた線形符号の誤り指数よりも大きくなることを示した。

平成 24 年度は、2 つの符号器が独立に符号化を行う際に同期が取れない場合の符号化について検討を行った。その結果、符号器間の同期ずれの大きさ(「遅延時間」と呼ぶ)と符号長との比が、符号長を長くするに従って零に収束する場合、符号長が長くなるにつれて誤り率が零に収束する符号が、遅延時間によらず存在することを示した。次に、無記憶情報源について、遅延時間と符号長との比が

符号長を長くするに従って零に収束する場合、2つの符号器のレート対が情報源によって定まる達成可能領域の内部に属するならば、符号長が長くなるにつれて誤り率が零に収束する符号が、遅延時間ならびに情報源の確率分布によらず存在することを示した。更に、上記の符号は全て、LDPC行列を用いることで構成できることを明らかにした。

参考文献

- [1] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. Inform. Theory, vol.19, pp.471-480, July 1973.
- [2] D. Schonberg, K. Ramchandran, and S. S. Pradhan, "Distributed code constructions for the entire Slepian-Wolf rate region for arbitrarily correlated sources," Proc. of Data Compression Conf. 2004, pp. 292-301, 2004.
- [3] V. Stankovic, A. D. Liveris, Z. Xiong, and C. N. Georghiades, "On code design for the Slepian-Wolf problem and lossless multiterminal networks," IEEE Trans. on Inform. Theory, vol.52, pp.1495-1507, 2006.

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1件)

- ① T. Matsuta, T. Uyematsu, and R. Matsumoto, Universal Slepian-Wolf Source Codes Using Low-Density Parity-Check Matrices, IEICE Trans. Fundamentals, 査読有, E93-A, 2010, 1878-1888.

[学会発表] (計2件)

- ① 松田 哲直, 植松 友彦, 非同期型 Slepian-Wolf符号化システムにおけるユニバーサル符号化, 電子情報通信学会情報理論研究会, 2013年3月7日, 関西学院大学梅田キャンパス(大阪府).
- ② 植松 友彦, 線形符号を用いた Slepian-Wolf符号化の誤り指数, 電子情報通信学会情報理論研究会, 2012年3月1日, 慶応義塾大学日吉キャンパス(神奈川県).

6. 研究組織

(1) 研究代表者

植松 友彦 (UYEMATSU TOMOHIKO)
東京工業大学・大学院理工学研究科・教授
研究者番号：60168656

(2) 研究分担者

なし

(3) 連携研究者

なし