

情報理論に関する研究

教授　植松 友彦

研究分野：ノンパラメトリックデータに対する情報理論、ネットワーク情報理論

ホームページ: http://www.it.ce.titech.ac.jp/

●研究内容

(1)ノンパラメトリックなデータに対する情報理論の構築

(2)ネットワーク情報理論の研究

(3)相関を有する情報源からの乱数生成

●研究テーマ

１．ノンパラメトリックなデータに対する情報理論の構築

　これまでの情報理論は、データを生成するメカニズムとして確率モデルである情報源や通信路を仮定し、確率過程によってデータが生成・伝達されると考えて理論を構築して来た。しかしながら、ビッグデータの時代になると、収集されたデータの背後に確率モデル（パラメトリックモデル）が必ずしも存在する保証はなく、データ自身に基づいて情報理論を構築する必要がある。

　そこで、本研究では、機械学習やビックデータ解析への応用を目的とし、生成モデルを仮定しないデータの圧縮・識別・分類、ならびに複数のデータ間の相関検出を行う方法を確立する。その結果、データ圧縮やデータ分類の性能限界を明らかにすることで、現在の技術水準が究極の性能に対してどのレベルまで来ているのか、あるいは、より高性能なシステムの設計・開発の指針を示す。

　これまでの研究成果としては、ノンパラメトリックデータに関するデータ圧縮について検討を行い、符号化レート（圧縮率）の限界を明らかにすると共に、具体的な固定長符号化ならびに可変長符号化法を提案し、ブロック長を長くしたときに、符号化レートがこの限界に漸近することを明らかにしている。

生成モデルを仮定せず

データだけに基づいた

圧縮、識別、分類法

確率モデル

ビッグデータ

２．ネットワーク情報理論に関する研究

　相関を有する２つの一般情報源からの出力列をそれぞれ独立に符号化して受信者に送り、受信者は２つの符号器の出力から２つの情報源の出力列を共に復元する符号化問題をSlepian-Wolf符号化問題と呼ぶ。Slepian-Wolf符号化問題において、復号誤り率εを許容したとき、２つの符号器の符号化レートが満足すべき領域をε達成可能領域と呼ぶ。1998年にHanはε達成可能領域を情報スペクトルによって表現することに成功した。これに対して我々は、ε達成可能領域をスムース最大エントロピーに類似した情報量を用いて記述し、このε達成可能領域の内界と外界がスムース最大エントロピーを用いて記述できることを示した。尚、任意に小さい復号誤り率を考えた場合、得られた内界と外界は一致することを明らかにした。

　次に、相関を有する２つの一般情報源からの出力列をそれぞれ独立に符号化して受信者に送り、受信者は２つの符号器の出力から片方の情報源の出力列のみを復元するWyner-Ahlswede-Korner符号化問題を検討し、復号誤り率εを許容したとき、２つの符号器の符号化レートが満足すべき領域（ε達成可能領域）がスムース最大エントロピーとスムース最大ダイバージェンスを用いて記述できることを明らかにした。

　更に、歪みを許容した固定長符号化において、復号器のみが情報源と相関を有する副情報源からの出力を参照できる場合、すなわちWyner-Ziv符号化問題におけるレート歪み関数がスムース最大ダイバージェンスを用いて記述できることを明らかにした。

復号器

相関を有する

情報源

符号器1

符号器2

●教員からのメッセージ

　植松研究室では、使い捨てではない画期的な研究による人類の知的遺産の構築を理念として、情報理論の分野における世界的レベルでの貢献、世界初あるいは世界一の研究成果を目指して研究を行っています。数学的モデルによって情報の本質を明らかにしようとする野心溢れる学生の入門を歓迎します。

●関連する業績、プロジェクトなど

１．T. Uyematsu and T. Matsuta, “Equivalence among Some Information Measures for Individual Sequences,” Proc. SITA 2018, pp.226-231, December 2018.

２. T. Uyematsu and T. Matsuta, “Revisiting the Slepian-Wolf Coding Problem for General Sources: A Direct Approach,” Proc. IEEE Inter. Symp. on Inform. Theory, pp. 1336-1340, June-July 2014.

# ３．T. Uyematsu and T. Matsuta, “Source coding with side information at the decoder revisited,” Proc. IEEE Inter. Symp. on Inform. Theory, pp. 1570-1575, June 2015.

# ４．T. Uyematsu and T. Matsuta, “Revisiting Wyner-Ziv Source Coding Problem Using Smooth Min and Max Renyi Divergence,” Proc. SITA 2015, pp. 49-54, November 2015.