

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 9 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2013～2016

課題番号：25840018

研究課題名(和文) 高度な信号推定技術によるNMR構造解析の高精度化とin-cell NMRへの応用

研究課題名(英文) Fast NMR data acquisition using advanced signal estimation for in-cell NMR

研究代表者

角越 和也 (Sumikoshi, Kazuya)

東京大学・大学院情報学環・特任助教

研究者番号：90431832

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究は、核磁気共鳴(NMR)法を用いた分子構造情報の取得に要する計測を、高度な信号推定手法を用いて高速化し、経時変化の激しい分子についても高精度な構造情報の取得を可能にすることを目的とする。そのために、短時間で取得可能な少数の標本点における測定データから、圧縮センシングと呼ばれる手法を用いてNMRスペクトルの推定を行うプログラムの開発を行った。計測データの特性を考慮したスペクトルの推定を行うことにより、従来法を上回る精度を達成した。

研究成果の概要(英文)：The aim of this project is to develop a fast NMR data acquisition method using an advanced signal estimation technique in order to be able to get structural information even from molecules which deteriorate in a short time. Toward this goal, we developed a software that can estimate NMR spectra from small amount of sampling data, which can be quickly obtained, by using a novel technique called compressed sensing. With the estimation using the deep knowledge of the characteristics of NMR data, we achieved higher performance than conventional methods.

研究分野：生命情報科学

キーワード：NMR 核磁気共鳴 信号処理

1. 研究開始当初の背景

核磁気共鳴 (NMR) 法は、分子の構造情報を原子分解能で得ることのできる実験的手法の一つである。水溶液中で動いている分子を非破壊的に測定できるため、生体内に近い環境下での生体高分子の構造、運動の解析に適するという特徴がある。特に近年では、生体内の環境そのものである、生細胞中のタンパク質を標的にした計測 (in-cell NMR 法) にまで、その可能性が広がっている。

しかし、NMR 法は測定感度が低いこと、また、異なる原子からのシグナル同士が重なり合ってしまうことなどの技術的困難さを抱えているため、生体高分子の測定には、複数回の測定の積算によるシグナルの強調や、多次元 NMR 測定によるシグナルの分離といった方策が採られる。ただし、それには大量の標本点での多数回の計測を要し、測定時間が非常に長くなるという問題点があり、NMR 法の適用範囲を限定する要因となっている。特に、試料内の細胞の寿命により測定時間に制約のある in-cell NMR 法でのタンパク質の測定には深刻な問題となり、通常数日かかる測定を数時間のオーダーに減らすことを可能とする技術が求められている。

2. 研究の目的

そこで本研究では、多次元 NMR 測定に要する計測時間を大幅に削減するため、短時間で取得可能である、少数の標本点のデータから、NMR スペクトルを高精度に推定、再構成する手法を開発すること、また、その手法を用いた推定を行う際に必要となる調整を容易にするプログラムの開発を行うことにより、経時変化の激しい分子についても高精度な構造情報の取得を可能にすることを目的とする。

上記の目的を達成するために行う具体的な項目は、(1) 少数の標本点データから信号を推定する基礎プログラムの開発、(2) 前記基礎プログラムに対し、NMR 計測データの特性を考慮した手法の改良を加えることによる推定性能の向上、(3) 実際の NMR 測定データ、ならびに人工的に作成したデータを用いた性能評価、(4) 対象データの性質に応じて調整を必要とする、推定パラメータの調整作業を助けるための半自動的チューニングプログラムの開発、の 4 つである。

3. 研究の方法

多次元 NMR 測定は、測定における複数のパラメータを少しずつ変化させながらデータの取得を行う。原理的には、所望するスペクトルの情報を正確に得るためには、標本化定理と呼ばれる定理に従った個数のデータを取得する必要がある。その計測時間を減らす一つの策としては、元来必要とされる標本点の一部を間引いた計測を行い、間引かれた点については、対象データが持つとされる性質を用いて値を推定することなどにより補

うことにより、近似的なデータの取得を行う方法がある。そうすることにより、例えば、標本点の個数を $1/N$ にすることができれば、原理的には計測時間を $1/N$ に削減することが可能となる。

NMR 測定データに関しては、スペクトルが疎である、つまり一部を除くと、値がほぼ 0 であるという特徴がある。疎であるという特徴を仮定して推定を行うことができる高性能な手法として、圧縮センシング (Compressed sensing, CS) と呼ばれる最新手法が提案されており、本研究では、推定に圧縮センシングを用いた手法の開発を行う。

圧縮センシングについて簡単に説明すると、間引かれたデータに合致するスペクトルは無限に存在するが、その内、疎であるものほど真のスペクトルに近いと考え、原信号の推定を行う (実際は、(測定値との 2 乗誤差) \times (スペクトルの絶対値の和) の最小化を行う) のがその原理である。推定スペクトルを得る際には、「測定値との 2 乗誤差」と「スペクトルの絶対値の和」の重みを決める、上記パラメータを設定する必要がある。その最適な値は各データの性質に依存しており、高精度の推定のためには、 α の調整が不可欠となる。実際の作業においては、得られたスペクトルを見ながら α の調整を繰り返すことになるが、手間と時間が非常にかかるため、その調整作業を助けるためのツールを作成する。当該ツールでは、所望する条件 (どのエリアにある閾値以上の信号がどのくらいあるか) を入力すると、それに合う α の値の探索を行う。

4. 研究成果

(1) 圧縮センシングによる信号推定を行う基礎プログラムの開発

圧縮センシングを用いた推定を行うには、計測データの空間から、疎に表現される空間への写像を定義する必要がある。基本的には、 N 次元 NMR におけるスペクトルは、 N 次元の複素数の計測データをフーリエ変換した複素数の値として得られる。よって、そのフーリエ変換を写像と定義するのが最も自然である。本研究では、まず、基礎となる信号推定プログラムを作成し、上記フーリエ変換を写像として使い、スペクトルの推定を行った (以下、「推定 1」と表記)。実際の NMR 計測データを用いて実験したところ、データを $1/8$ に間引いても高い再現性を示すスペクトルを得ることができた。図 1 に、得られた推定スペクトルの例を示す。

(2) NMR 計測データの特性を考慮した手法の改良

より大幅に間引かれたデータからも NMR スペクトルの取得を可能にすることを目標とし、NMR 計測データの特性をより深くまで掘り下げて推定に取り込むことにより、さらなる推定性能の改善を試みた。圧縮センシ

グの性能を大きく左右する要素には、1. 測定データの空間から疎な表現となる空間への写像の方法、2. 最適化の目的関数の構成がある。その2つにおいて、NMR 計測データの特性を鑑みた改良を加えることにより、推定性能を向上させることができた。それらについて以下に述べる。

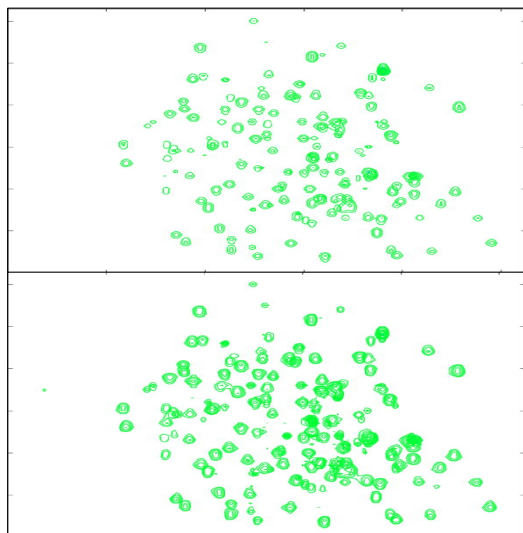


図 1. 1/8 に間引かれたデータを用いた NMR スペクトルの推定結果 (上: 原スペクトル、下: 推定スペクトル)

測定データの空間から疎な表現となる空間への写像に対する改良

圧縮センシングにおいては、写像後の空間における表現がより疎であるほど、推定の性能が高まるという性質がある。つまり、データをより疎に表現可能である空間を見つけることが性能の向上につながる。それに関して考慮する NMR 計測データの特性として、スペクトルの実数部と虚数部の値は互いに独立でないこと、また実数部の方が虚数部より疎な表現になりがちであることを利用し、写像を行う関数の改良を行った。その改良を適用したプログラムを用いて、1/16 に間引かれた計測データからの推定 (以下、「推定 2」と表記) 評価を行ったところ、性能の向上が見られた。性能の指標として、原スペクトルと推定スペクトルの 2 乗誤差の平均の平方根 (以下「RMS」と表記、値が小さいほど性能が高いことを示す) を用いて評価したところ、推定 1 の結果と比較し、約 76.9 % の RMS にまで下げることができた。

最適化の目的関数の構成に関する改良

さらなる性能向上のために考慮可能で、かつ写像の改良として取り込むことのできない、その他の NMR 計測データの特性としては、周波数空間での表現がとり得る値の範囲が挙げられる。取得されたデータが位相補正されていると仮定すると、周波数空間での値の範囲を限ることができ。それから導かれる制約条件を最適化の目的関数に導入すること

により、さらなる改善に成功した。RMS を用いて性能を評価したところ、推定 1 と比較し、約 75.9 % の RMS (推定 2 と比較すると、98.6 %) となり、僅かながらではあるが、本改良による改善が見られた。推定 2 と推定 3 による推定スペクトルの例を図 2 に示す。

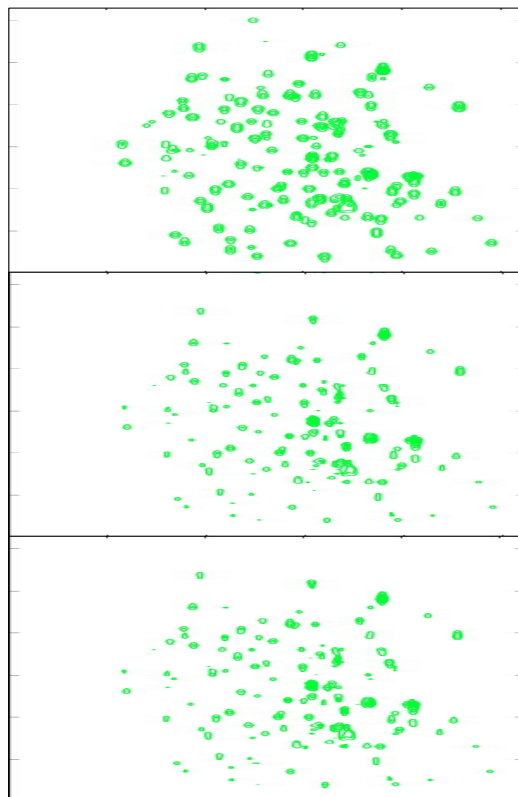


図 2. 1/16 に間引かれたデータを用いた NMR スペクトルの推定結果 (上から、原スペクトル、推定スペクトル (推定 2)、推定スペクトル (推定 3))

また、これらの改良を適用して得られた結果と、Maximum Entropy 法を用いた従来法により得られた結果との比較を、1/16 に間引かれた実際の NMR 計測データを用いて行ったところ、本手法の結果の RMS は、従来法の結果の RMS と比較し、約 96.6 % となり、よりよい性能を達成することができた。

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 2 件)

M. Banno, Y. Komiyama, W. Cao, Y. Oku, K. Ueki, K. Sumikoshi, S. Nakamura, T. Terada, K. Shimizu, Development of a sugar-binding residue prediction system from protein sequences using support vector machine, Journal of Computational Biology and Chemistry, 査読有, Vol. 66, 2017, 36-43

DOI: 10.1016/j.compbiolchem.2016.10.009

H. Tsujikawa, K. Sato, W. Cao, G. Saad, K. Sumikoshi, S. Nakamura, T. Terada, K. Shimizu, Development of a protein-ligand-binding site prediction method based on interaction energy and sequence conservation, Journal of Structural and Functional Genomics, 査読有, Vol. 17, 2016, 39-49
DOI: 10.1007/s10969-016-9204-2

〔学会発表〕(計3件)

K. Sumikoshi, T. Ikeya, Y. Ito, K. Shimizu, Reconstruction of NMR spectra using compressed sensing with multiple regularization terms, 第53回日本生物物理学会年会, 2015年09月15日, 金沢大学角間キャンパス(石川県、金沢市)

K. Sumikoshi, T. Ikeya, Y. Ito, K. Shimizu, Effect of basis selection in reconstructing NMR spectra using compressed sensing, 第52回日本生物物理学会年会, 2014年09月25日, 札幌コンベンションセンター(北海道、札幌市)

K. Sumikoshi, T. Ikeya, Y. Ito, K. Shimizu, Reconstruction of NMR spectra using compressed sensing, 第51回日本生物物理学会年会, 2013年10月29日, 国立京都国際会館(京都府、京都市)

6. 研究組織

(1) 研究代表者

角越 和也 (SUMIKOSHI, Kazuya)
東京大学・大学院情報学環・特任助教
研究者番号: 90431832