

東京工業大学 TSUBAME 共同利用  
平成 29 年度利用終了課題 利用成果報告書

東京工業大学 学術国際情報センター  
共同利用推進室

<https://www.gsic.titech.ac.jp/tsubame>



## ■平成 29 年度産業利用

申請課題名 所属機関／利用課題責任者	頁
自動車搭載ミリ波レーダの超大規模電磁界シミュレーション 株式会社豊田中央研究所／松沢晋一郎	1
建築環境の評価に向けた気候変動データの高度化 清水建設株式会社／PHAM VAN PHUC	5
ワイヤレス電力伝送による漏えい電波の環境解析技術の研究開発 株式会社パナソニックシステムネットワークス開発研究所／池田和彦	9
大容量データ伝送用ミリ波アンテナのレドームに関する基礎検討 スタッフ株式会社／千葉修二	15
産業利用のための大規模ディープラーニング 株式会社富士通研究所／山崎崇史	19

## ■平成 29 年度学術利用 利用成果報告書別紙の提出免除課題 (利用成果を論文/学会等にて発表した要旨等の提出により提出免除)

申請課題名 所属機関／利用課題責任者	提出物
高性能計算向け分散メモリ・ストレージ統合システムの研究 成蹊大学／緑川博子	
「高柔軟性と高性能を提供するマルチノードマルチスレッドプログラム向け分散共有メモリシステム」, 研究報告ハイパフォーマンスコンピューティング(HPC), 2018-HPC-163 (10) 1-8 Permalink : <a href="http://id.nii.ac.jp/1001/00185951/">http://id.nii.ac.jp/1001/00185951/</a>	
大型ホログラム記録とホログラフィックディスプレイのためのホログラム計算 国立研究開発法人情報通信研究機構／山本健詞	
「HOPTECH について -波面印刷技術から投影型ホログラフィック 3D 映像技術まで-」 2017 年 第 42 回 光学シンポジウム, 招待講演 <a href="http://myosj.or.jp/event2/opt_symp/2017_042nd/42nd_invited_abst.shtml#no3">http://myosj.or.jp/event2/opt_symp/2017_042nd/42nd_invited_abst.shtml#no3</a>	

## ■平成 29 年度学術利用

申請課題名 所属機関／利用課題責任者	頁
廃棄物最終処分場における間隙内流体挙動の数値解析 埼玉県環境科学国際センター／鈴木和将	21
高解像度画像を使った広域の家屋及び道路の深層学習による自動判別システムの開発 東京大学空間情報科学研究センター／柴崎亮介	25
蛋白質-リガンド複合体の会合・解離の速度定数を平衡論から導く 大阪大学蛋白質研究所／肥後順一	29

申請課題名 所属機関／利用課題責任者	頁
先進的ステンシル・コード技術 東京大学情報理工学系研究科／千葉滋	31
LRnLA アルゴリズムを用いた物理シミュレーション 法政大学 情報科学部／善甫康成	33
凝集誘起発光についての理論的研究 千葉工業大学／山本典史	37
知識に基づく構造的言語処理の確立と知識インフラの構築 京都大学／黒橋禎夫	41
深層学習を利用したリアルタイム宇宙天気予報システムの開発 理化学研究所計算科学研究機構／村主崇行	課題 取り下げ
解析的壁関数を用いた格子ボルツマン法による高レイノルズ数乱流解析コードの開発 大阪府立大学大学院／須賀一彦	45
抗インフルエンザウイルス広域中和抗体のエスケープ変異株由来表面糖タンパク質の構造解析 国立感染症研究所感染病理部／長谷川秀樹	49
スマートデータセンター実現に向けたデータ解析基盤の構築 産業技術総合研究所 RWBC-OIL／松岡聡	51
MEMS 構造の光学応答解析 電気通信大学／菅哲朗	55
GPGPU による長周期地震動シミュレーション 国立研究開発法人防災科学技術研究所／藤原広行	57
都市解像モデルへのドップラーライダー観測データ同化による局地域豪雨発生予測の精度向上 京都大学生存圏研究所／古本淳一	61
第一原理計算と反応速度論による触媒活性予測技術の確立 物質・材料研究機構／石川敦之	65
HPC を利用した自然言語処理技術の研究 情報通信研究機構 ユニバーサルコミュニケーション研究所／鳥澤健太郎	67
有機分子集合体の半導体特性・光応答特性の解明 理化学研究所計算科学研究機構／河東田道夫	69
深層学習によるゲノム上の遺伝子発現制御情報抽出 東京大学・新領域創成科学研究科・メディカル情報生命専攻／木立尚孝	71
ナノ構造界面における熱輸送特性の分子シミュレーション 東京大学工学系研究科機械工学専攻／塩見淳一郎	73
深層学習処理基盤を用いた大規模環境計測データの解析と応用 産業技術総合研究所／李時旭	75
大規模分散深層学習の性能調査 産業技術総合研究所 RWBC-OIL／松岡聡	77



TSUBAME 共同利用 平成 29 年度 産業利用 成果報告書

利用課題名 自動車搭載ミリ波レーダの超大規模電磁界シミュレーション  
 英文: Large scale electromagnetic simulation for automotive millimeter wave radar

松沢晋一郎  
 Shin-ichiro Matsuzawa

株式会社 豊田中央研究所  
 Toyota Central R&D Labs., Inc.  
 URL <http://www.tytlabs.co.jp/>

#### 邦文抄録

車載ミリ波レーダは、前方車などの反射体の相対速度や方位を正確に測定可能という特徴がある。車載ミリ波レーダのフロントガラス背面への搭載を想定し、その主要課題であるフロントガラスの透過損失をシミュレーション評価し、さらにその改善方法を検討した。そのシミュレーションでは、スーパーコンピュータ TSUBAME3.0 を用いて行った。

#### 英文抄録

A millimeter wave radar can measure the relative speed and angle of the target like a front vehicle precisely. This report shows the transmission loss of the glass in the millimeter frequency when the millimeter wave radar is mounted behind the front glass. We have simulated the radiation patterns of the antenna by using the TSUBAME, and studied the improvement method.

*Keywords:* 電磁界シミュレーション、ミリ波レーダ

#### 背景と目的

近年、自動ブレーキなどの先進運転支援システム (ADAS: Advanced Driver Assistance Systems) の開発が進められている。ミリ波レーダは、ADAS のキーセンサの一つである。

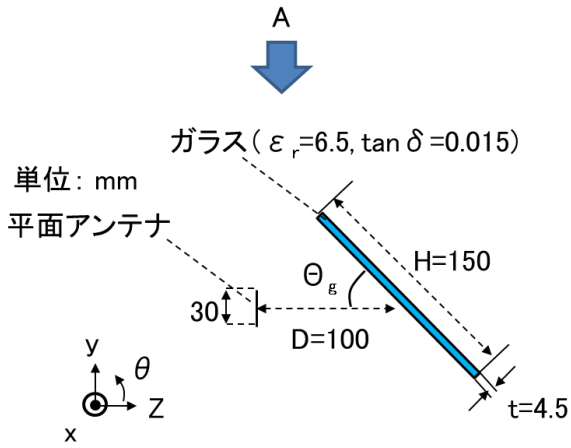
現状の車両前方監視用ミリ波レーダは、一般的に樹脂エンブレムの背面に搭載されている。一方、同じ前方監視用のカメラは、車室内のフロントガラスの背面に搭載されている。将来的にミリ波レーダとカメラを一体化したセンサをフロントガラスの背面に搭載することを検討している。ミリ波レーダをフロントガラスの背面に搭載した場合、フロントガラスの透過損失が大きく、その検出距離が悪化することが事前検討で分かった。そこで、本報告では、ミリ波帯におけるガラスの透過損失をシミュレーション評価し、その改善方法を検討した。

#### 概要

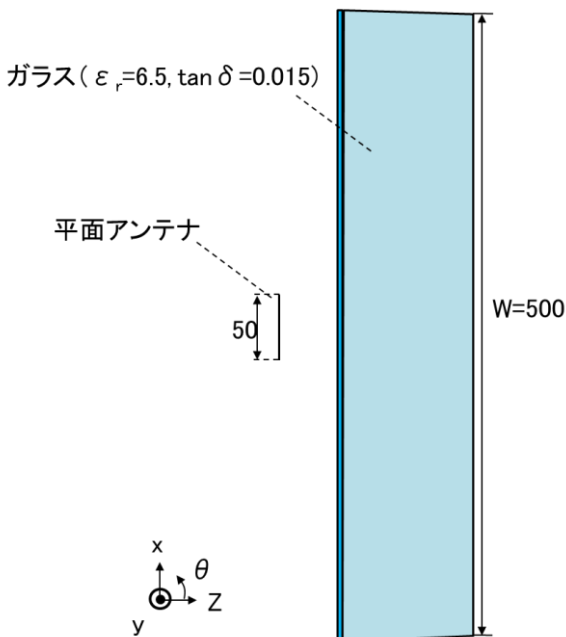
ミリ波レーダは、77GHz の高周波電磁波により、自動車と対象物の相対距離・相対速度・角度を高精度に測定可能である。一方、その波長  $\lambda_0$  は約 4 ミリと短い為、周辺の構造物、例えば車体フレーム、バンパー、フロン

トグリル等による電波散乱も生じやすい。それらは全て測定誤差要因となる為、設計の電磁界シミュレーションに際し、レーダ周辺の車体構造物を忠実にモデル化することが望ましい。しかしながら、FI 法によるシミュレーションをするに際し  $\lambda_0/10$  以下でのメッシュ分割が必要となり、総メッシュ数が約 40 億以上と膨大になるため、PC ワークステーションでは、計算が困難であった。そこで、スーパーコンピュータ TSUBAME3.0 を用いたシミュレーションを行った。

図 1 にシミュレーションモデルを示す。ミリ波レーダ用の平面アンテナの放射方向にフロントガラスを模擬したガラスが置かれている。ガラスの厚みは、一例として、 $t=4.5\text{mm}$  とした。その比誘電率  $\epsilon_r$  および誘電正接  $\tan \delta$  は、それぞれ 6.5、0.015 とした。一般的に車両のフロントガラスは、2 層ガラスの間に中間膜が挿入されているが、本報告書では、初期検討として単層で検討した。ミリ波の平面アンテナとガラスの距離は、一例として、 $D=100\text{mm}$  とした。



(a) 断面図



(b) 上面図(A)

図 1 シミュレーションモデル

現状のミリ波レーダでは、76～77GHz の周波数が用いられているが、次世代レーダとして、高分解能化のための広帯域レーダが検討されている。その周波数は、77～81GHz であるため、それらを含む 76～81GHz をシミュレーションの対象とした。ガラスの傾斜角度  $\theta_g$  は、 $45^\circ$ 、方位の検出角度範囲は、 $-45^\circ \sim +45^\circ$  とした。

ガラスの内部のメッシュサイズは 0.16mm 以下とする必要があり、実際の車両のガラスのサイズでのシミュレーションは計算負荷が高すぎるため、困難である。そこで、水平検出角度  $\pm 65^\circ$ 、垂直検出角度  $\pm 30^\circ$  の範囲にガラスが存在するように、その幅 W、高さ H を決定し、 $W=500\text{mm}$ 、 $H=150\text{mm}$  とした。

図 2 に平面アンテナのモデルを示す。1×9 列のマイクロストリップアレーアンテナとし、各素子を仮想的にワイヤで給電するモデルとした。各素子の形状およびワイヤの給電位置は、 $45^\circ$  偏波を発生するように設定し、その給電電圧は、垂直方向のサイドロープレベルが  $-20\text{dB}$  以下となるように、テーパをつけた。アンテナ基板は、縦 30mm、横 50mm のフッ素樹脂基板とした。実際のレーダでは、本アレーを一つのサブアレーとして、水平方向に複数並べられているが、本解析では、基板の中央にあるアレーアンテナのみの特性を評価した。

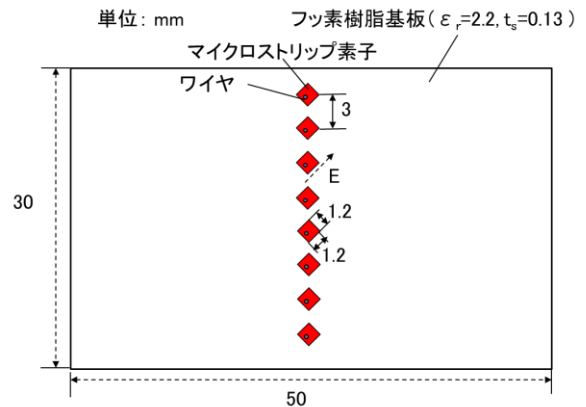


図 2 平面アンテナのモデル

ガラスの損失の主要因は空気とガラスの間の比誘電率の差による反射であるため、その反射を抑える方法として、ガラスの裏面に空気とガラスの中間の比誘電率 3.5 の樹脂フィルムを貼る方法を検討した(図 3)。樹脂フィルムの厚み  $t_s$  については、事前にフレネル係数による解析で最適値を求め、 $t_s=0.48\text{mm}$  としてアンテナの指向性を考慮して評価した。電磁界シミュレーションは、CST社の Microwave Studio Ver.2017 により行い、手法は FI (有限積分法) 法である。

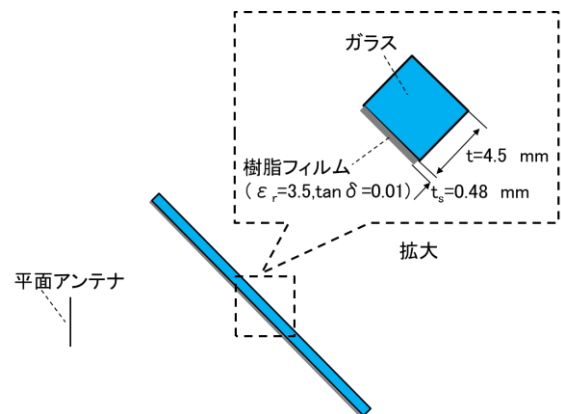


図 3 ガラスの低損失化の方法

### 結果および考察

平面アンテナ単体と前面にガラスを設置した場合の指向性を評価した。図 4(a)および(b)に 76.5GHz における平面アンテナ単体とガラスを設置した場合の水平面および垂直面の指向性の比較を示す。

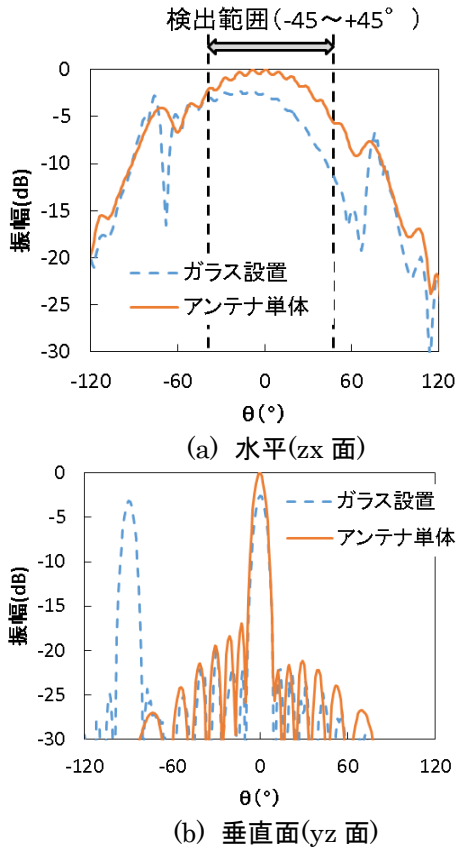


図 4 平面アンテナ単体とガラスを設置した場合の指向性の比較

縦軸は、平面アンテナ単体の利得のピーク値で規格化した相対値 (dB 値) である。水平面指向性では、ガラスを設置した場合、左右非対称の指向性となり振幅が低下していることが分かる。これは、平面アンテナの偏波が  $45^\circ$  傾いているために、水平面内において  $\theta$  が正の時と負の時で、ガラス面への入射偏波が異なるため、非対称となったと考えられる。また、垂直面指向性の結果から、ガラスを設置した場合に  $-90^\circ$  方向に強い反射波が発生していることが分かる。

図 5 に周波数をパラメータとした水平面での  $\theta$  に対する透過損失の関係を示す。透過損失は、図 4(a)に示したガラスを設置した振幅の平面アンテナ単体の振幅からの差分として定義される。検出範囲内では、 $\theta = 45^\circ$  方向の損失が大きく、最大 6.5dB となった。また、

周波数が低いほど、損失が大きくなる傾向となった。

図 6(a)および(b)に、低損失化のため樹脂フィルムの有無と透過損失の関係を示す。

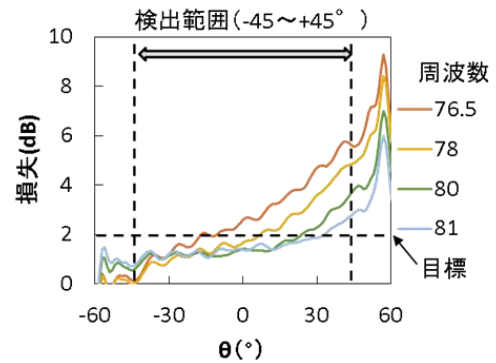


図 5  $\theta$  (水平面) および周波数と透過損失の関係

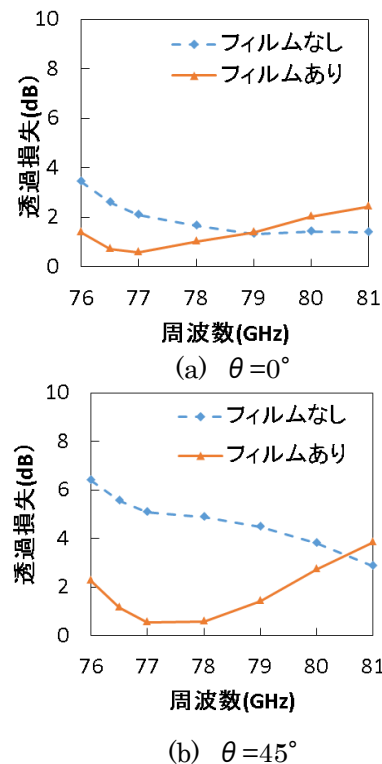


図 6 フィルムの有無と透過損失の関係

(a)は正面である  $\theta = 0^\circ$  および(b)は水平面の検出範囲内で損失の高い  $\theta = 45^\circ$  の場合の結果である。 $\theta = 45^\circ$  の場合に、フィルムを貼ることでの損失低減効果が顕著であった。

最後に計算時間について述べる。本モデルのメッシュ数は 38 億メッシュであり、その CPU タイムは 78 分 (8 ノード) であった。TSUBAME2.5 では、15 億メッシュのモデルで CPU タイムが 7 時間 11 分 (12 ノード) であり、大

幅な高速化を確認した。

#### まとめ、今後の課題

ミリ波レーダをフロントガラス背面への搭載を想定し、その主要課題であるフロントガラスの透過損失を、電磁界シミュレーションで評価し、その改善方法について検討した。

シミュレーションの結果、対策をしていないガラスを設置した場合は、アンテナの水平面指向性が非対称となり、損失が最大 6.5dB と大きいことが明らかになった。その改善方法として、反射防止用の樹脂フィルムを貼る方法を提案し、その効果を確認した。

利用課題名 建築環境の評価に向けた気候変動データの高度化  
英文: Improvement of Climate Change Data for Building Environment Analysis

利用課題責任者  
PHAM VAN PHUC

所属: 清水建設(株)技術研究所  
Affiliation: Institute of Technology, Shimizu Corporation  
URL: <http://www.shimz.co.jp>

本利用課題では、代表的な極端気象現象の一つとして、2017年7月に九州地方で発生した局地豪雨を対象にして、気象モデルを用いた現象の再現計算と、観測結果との比較検証およびその発生メカニズムの解明を行った。また、気候変動データベース d4PDF を用いて、その地域における気候変動による影響を明らかにした。

In this study, meteorological simulation using the WRF model has been carried out to investigate the sensitivity of a heavy localized rainfall over the Kyushu in July, 2017 with a downscaled grid resolution in the nesting method. Simulated results showed good agreement with observation results. Climate change database d4PDF was used to investigate the daily precipitation to find the increase under global warming.

*Keywords:* Heavy Localized Rainfall, WRF, Climate Change

## 背景と目的

異常高温、集中豪雨、大型台風などの極端な気象現象による人的、経済的損失は極めて大きい。近年の気候変動として地球温暖化の影響で一部の地域ではその気象現象の強度はすでに増加傾向にあり、今後も世界規模でさらに増えていく。また、2017年7月5日から6日にかけて、九州北部地方では、総降水量が500mmを超え、7月の月降水量平年値を大幅に超える猛烈的な雨となったところがあった。特に、福岡県朝倉市では、24時間降水量の値が観測史上1位の値を超えて、これまでの観測記録を更新する局地豪雨となった。今回の豪雨による被害については、死者が計36人となり、300カ所以上で山が崩れ、大量の流木や土砂が発生した等、大きな被害が引き起こされた。道路などのインフラ被害では局地激甚災害にも指定された<sup>1)</sup>。建設分野としても、その現象や気候変動の影響などを明らかにする必要がある。適用策などの提案は喫緊の課題である。

そこで、本利用課題は、まず、気象モデルを用いた局地豪雨の再現を試みた。計算から得られた降水量とその観測結果との比較により局地豪雨の特徴とその発生メカニズムを明らかにする。また、局地豪雨の発生地域を対象として気候変動に基づく降水量の変化を評価

することを目的とする。

## 概要

前年度の利用課題では、気象解析モデルの導入や気候変動データベースの整備と、データ可視化システムおよび特徴抽出システムの構築を行った。また、TSUBAME スパコンでは複数並列処理により、過去・現在の気候状態と未来の気候状態における気象パラメータの時間・空間的な変化を確認しており、台風等の極端な気象現象を抽出できるとともに、データベースの特徴を明らかにした。

本利用課題では、これらの解析システムを用い、2017年7月5日から6日にかけて、九州地方福岡県朝倉市で発生した局地豪雨を対象として、再現解析を行った。また、その九州地方の全域を対象として、気候変動による影響を調べた。

なお、図-1には朝倉市の地域気象観測システムAmeDASで、20年間での6月と7月の最大日降水量および今年の7月5日の日降水量の比較を示す。これまでに観測された最大日降水量が250mm以下に対して今回の大雨ではその一日で2倍以上の雨量が降ってきた。また、今回の局地豪雨の発生メカニズムについて

では、気象分野で、図-2 のように考えられている。これは、朝倉市付近の上空では、大気下層にある暖かく湿った空気と、大気上層にある気温の低い寒気が流入したため、積乱雲が次々に発達して、東へ移動することで線状降水帯として形成され、同じ場所に強い雨を継続して降らせたものである。さらに、図-3 は国土交通省の X バンド MP レーダー観測データから得られた当時の雨雲の立体構造である。その雨雲は東西に延び、高度 15 km を超える非常に大きい積乱雲群であった<sup>2)</sup>。



図-1 朝倉市の最大日降水量(気象庁, AmeDAS)

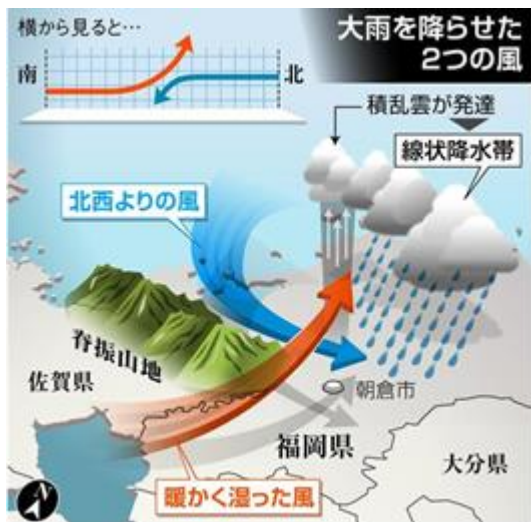


図-2 豪雨の発生メカニズム(出典:朝日新聞)

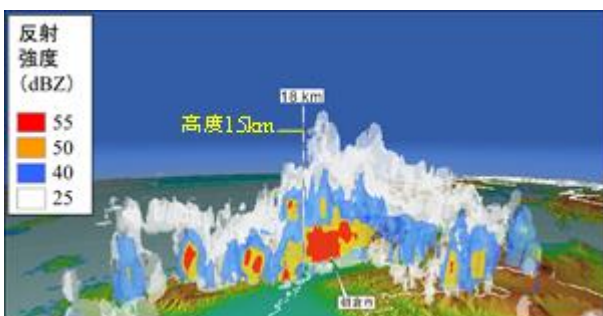


図-3 積乱雲の発達状況(時刻:15時)<sup>2)</sup>

## 計算概要

気象モデルの解析では、米国の大気研究センターが開発したメソ気象モデル WRF (Weather Research and Forecasting)<sup>3)</sup>を採用して、上記の局地豪雨の再現計算を行った。WRFは気圧、風、気温、水蒸気などで表せる大気の状態変化を流体力学や熱力学等の物理法則に基づいて解析できるものである。

計算領域としては、まず、豪雨に起因する積乱雲の再現を目的とし、親領域の水平解像度を 16.2km(d01)として 5 段階ネスティングより最終評価領域(d05)での水平解像度を 200m とした(図-4)。各領域の中心緯度・経度は、朝倉市役所の位置とした。鉛直方向の気圧準拠座標については地表から 60 層を設定した。初期値・境界値には米国の国立環境予測センターの全球客観解析データ NCEP-NFL を用いて、2way ネスティング計算を行った。

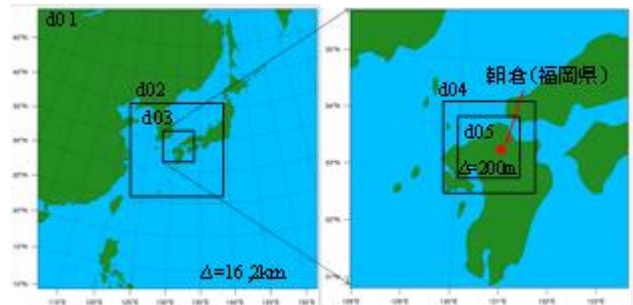


図-4 計算領域

## 結果および考察

再現期間は 局地豪雨の発生した 7月5日 09:00 から 7月6日 09:00 までの 24 時間を対象とした。また、物理過程スキームについては、雨を再現するため、水蒸気・雲・降水過程等を考慮できる雲微物理過程と渦輸送運動量・熱フラックス等を考慮できる境界層乱流混合過程を採用した。

計算精度を検討するため、朝倉観測所の地上観測 AmeDAS データ(ピンポイントの観測結果)とレーダーの解析雨量を用いた。図-5 には観測及び計算から得られた時降水量とその積算降水量を示す。なお、解析雨量は AmeDAS の降水量と比べると 2 倍の差がある。また、図 6 は時刻 15 時の 1 時間降水量の分布の比較である。これより、WRF の計算から得られた降水量は観測結果と比較的良好に一致していることが分かる。



図 7 は、WRF の計算から得られた 15 時の高度 500m の大気 1kg あたり得水蒸気量分布・気圧の等値線・風ベクトル(左図)と、高度約 5500m の気温分布・風ベクトル(右図)である。九州北部には、南西風によって大量の水蒸気が流入していた。また、上空 5500m 付近には、平年よりも約 3℃低い-7℃以下の寒気(福岡の高層観測の7月の平年値:-4.6℃)が流入していた。これらのことから、九州北部付近では積乱雲が非常に発達しやすい不安定な大気状態となっていたことが分かる。

また、図 8 は図 6 の断面図 A-A の雨雲の反射強度を示す。その付近は高度約 15 km の積乱雲が発生した。それらが猛烈に発達しながら、東へ移動することで、局地豪雨の発生をもたらした典型的な「バックビルディング型」の線状降水帯が形成されたことが分かった。

### 気候変動による影響

本章では、「文科省・気候変動リスク情報創生プログラム」に基づく気候変動データベース d4PDF を用いて、九州地方全域における気候変動の影響を評価した。

図-9は2000年(Past)および、全球平均気温が産業革命以降4℃上昇した2100年(Future)における日降水量の頻度分布を示す。2000年頃と比べて、2100年頃の降水量は増加する傾向があり、その増加率が約1.2倍になることが分かる。

### まとめ、今後の課題

本研究では、平成 29 年 7 月九州北部で発生した局地豪雨を対象として気象モデルを用いた再現計算を行った。計算から得られた降水量は観測結果と比較的良好に一致している。また、計算結果の可視化により、局地豪雨発生付近では、大気下層の暖かく湿った空気と大気上層の冷たい空気が流れ込んで、発生した積乱雲が猛烈に発達しながら移動する等、局地豪雨をもたらした典型的な「バックビルディング型」の線状降水帯が形成されたことを確認できた。さらに、九州地方を対象として気候変動による影響を調べており、その降水量が約 1.2 倍に増加することが分かった。今後の建築環境において水害対策に対して注意する必要があると考えられる。

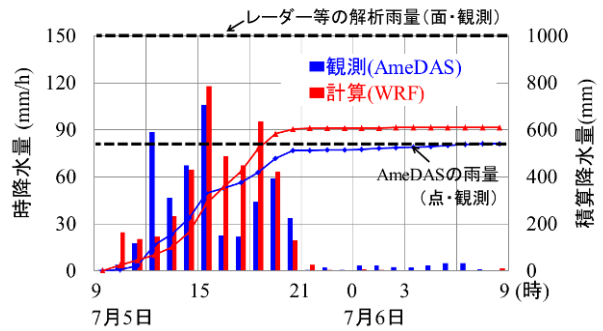


図-5 降水量の変化

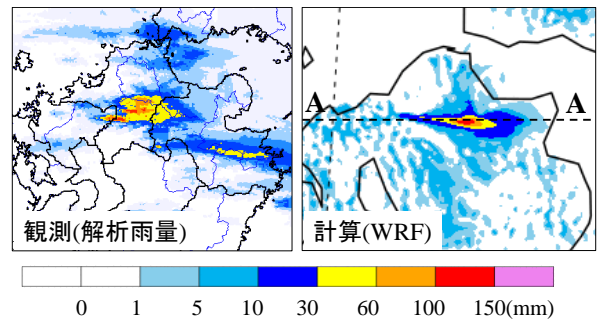


図-6 1時間降水量の分布(時刻:15時)

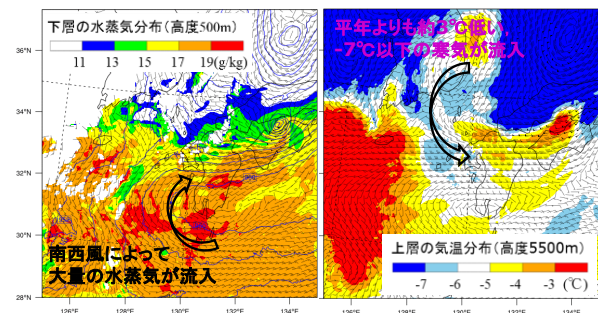


図-7 上空の大気状態(WRF 計算, 時刻:15時)

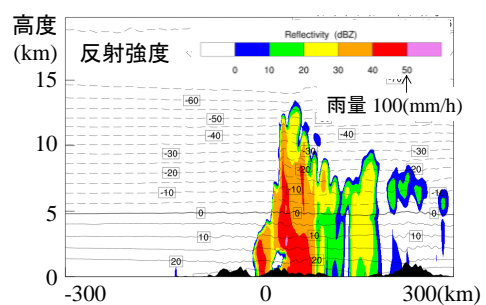


図-8 線状降水帯の発達状況(WRF, 時刻:15時)

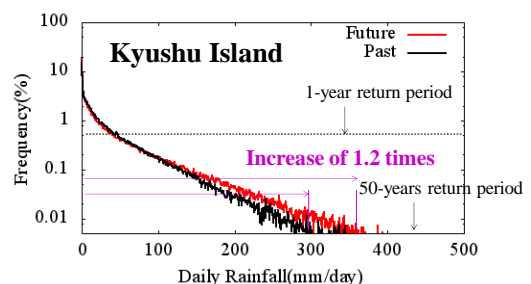


図-9 九州地方の降水量の頻度分布

## 参考文献

- 1) 気象庁: 平成 29 年 7 月 5 日から 6 日に九州北部地方で発生した豪雨の命名について, 2017.
- 2) 防災科学技術研究所: 2017 年 7 月 5 日から 6 日における福岡県・大分県の大雨について, 2017.
- 3) WRF: WRF Users' Guide, Version 3.9, 2017.



TSUBAME 共同利用 平成 29 年度 産業利用 成果報告書

## ワイヤレス電力伝送による漏えい電波の環境解析技術の研究開発

Research and development of environment analysis technique of leaked electromagnetic field from wireless power transfer

池田 和彦

Kazuhiko Ikeda

株式会社パナソニックシステムネットワークス開発研究所

Panasonic System Networks R&amp;D Lab. Co., Ltd

<https://panasonic.co.jp/cns/psnrd/>

家電機器の充電などに用いられるワイヤレス電力伝送 (WPT : Wireless Power Transfer) システムが近年検討されている。WPT システムから漏えいする電磁界は他の様々な電子機器に影響を与える可能性があり、その干渉影響を定量的に評価する必要がある。WPT システムの設置が想定される戸建て住宅や商業施設などの環境では近傍に様々な無線機器が存在するため、それらに影響を及ぼす恐れのある周波数帯での漏えい電磁界特性の評価が必要である。しかしながら、戸建て住宅や商業施設全体の漏えい電磁界を測定するには多大な時間と労力がかかり、シミュレーションでは解析空間サイズが大規模になる課題がある。筆者らはこれまでスーパーコンピュータ TSUBAME 2.5 を用いて戸建て住宅及び商業施設における漏えい電磁界を解析し、鉄筋や外壁などの建材が漏えい電磁界特性に大きな影響を与えることを示した。本利用課題では、より実際に近い WPT 使用環境を想定し、商業施設内に複数台の WPT システムが併設される場合や、商業施設周辺に多数の建造物が存在する場合の漏えい電磁界を解析する。

The wireless power transfer (WPT) system for electronics devices and so on has been actively developed in recent years. Since the electromagnetic field leaked from the WPT system interferes with the other devices, it is required to quantitatively evaluate the electromagnetic interference. In the housing or commercial facilities environment, the WPT system is located generally close to other various devices. If the frequencies of the fundamental and high-harmonic component of the WPT system are overlapped with operating frequency of these devices, the electromagnetic field leaked from the WPT system threatens to interfere with these devices. Thus, it is required to quantitatively evaluate the leaked electromagnetic field at various frequencies. The measurement in the whole house or commercial facilities involves immense amount of time and effort, and the simulation requires the enormous analysis size. Therefore, we have analyzed leaked electromagnetic fields in the housing or commercial facilities environment by using supercomputer TSUBAME 2.5. It was confirmed that architectural materials, the steel mesh structure and exterior wall, especially affect the electromagnetic field in the house simulation. In this paper, assuming an environment close to actual use case of the WPT system, we analyze the leaked electromagnetic field property under simultaneous operating condition of multiple WPT systems by using supercomputer TSUBAME 2.5.

*Keywords: Electromagnetic simulation, Ray launching, Wireless power transfer, Leaked electromagnetic field*

### 1. まえがき

家庭用電子機器や電気自動車の充電に用いられるワイヤレス電力伝送 (WPT : Wireless Power Transfer) システムが近年検討されている[1],[2]。WPT システムから漏えいする電磁界は他の様々な電子機器の誤作動の要因になり得るため、その強度規制値の国際標準化に向けた議論が進められている。多種多様な WPT システムからの漏えい電磁界によって引き起こされる干渉問題が重要な課題であり、様々な設置環境や高次高調波を考慮した定量的な評価が必要不可欠である。例えば住宅環境においては、WPT システムの近傍に様々な無線機器が存在し、それら被干

渉機器に影響を及ぼす恐れのある周波数帯 (kHz 帯から GHz 帯) の漏えい電磁界評価が重要となる。

しかしながら、住宅環境等の WPT 設置環境全体における漏えい電磁界の測定には多大な時間と労力がかかるため、シミュレーションを用いて解析することが望ましい。住宅環境全体を含むシミュレーションの解析空間は大規模なものになるため、並列演算による大規模解析が可能なスーパーコンピュータの利用が有効となる[3]。

平成 27 年度及び 28 年度の利用課題において、筆者らは、WPT システムや各種無線機器などが密集して設置される環境の 1 つとして戸建て住宅または商

業施設を含む空間に WPT 波源を配置し、漏えい電磁界の分布状況を解析することで、住宅を構成する鉄筋や外壁などの建材が漏えい電磁界特性に大きな影響を与えることを示した。本年度は、より実際に近い使用環境を想定し、商業施設内に複数台の WPT システムが併設され同時動作する場合や、商業施設周辺に多数の建造物が立地している場合の漏えい電磁界を解析する。

本利用課題では、ラジオ受信機、携帯端末、無線 LAN ルータへの干渉を想定し、WPT システムの高調波である 87MHz, 3.5GHz, 5.3GHz の漏えい電磁界をスーパーコンピュータ TSUBAME 2.5 [4]を用いて解析する。漏えい電磁界の波源と商業施設をモデリングして商業施設モデル内外の漏えい電磁界を解析し、商業施設内に複数台の WPT 波源がある場合や商業施設周辺に多数の建造物が存在する場合の漏えい電磁界特性を求める。また、商業施設を含む解析空間サイズの変更により解析で考慮する周辺建造物の範囲を変化させ、妥当な解析に必要となる周辺建造物のモデル化範囲を検証する。なお、87MHz の解析には有限積分法を用いた電磁界シミュレータである CST 社の MICROWAVE STUDIO[5]を、3.5GHz 及び 5.3GHz の解析には筆者らが開発した Ray Launching 法を用いた電波伝搬シミュレータ[3]を使用して解析する。

## 2. 電磁界解析による漏えい電磁界の解析

本章では、漏えい電磁界によるラジオ受信機への干渉を想定し、WPT システムの高調波である 87MHz の漏えい電界特性を電磁界解析手法（有限積分法）により解析する。

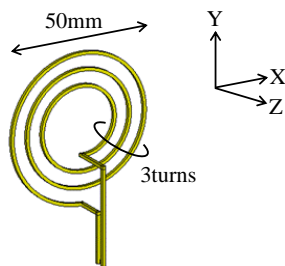


図 2.1 解析モデル（波源）

## 2.1 波源モデル

図 2.1 に本章で用いる波源モデルを示す。直径 50mm, 巻き数 3T のスパイラルコイルを 87MHz の波源として用いる。

## 2.2 商業施設における漏えい電磁界の解析

本節では、商業施設をモデリングして WPT 波源を複数個配置し、複数の波源が同時動作する場合の商業施設モデル内外の漏えい電界を解析する。

### 2.2.1 シミュレーション条件

表 2.1 にシミュレーション条件を示す。計算機システムは TSUBAME 2.5, 有限積分法の電磁界解析を用い、解析周波数は FM ラジオ帯の 87MHz とし、6.78MHz (家電モバイル機器の充電向け) の 13 次高調波を想定した。波源は 2.1 節で説明したモデルを

表 2.1 シミュレーション条件

計算機システム	TSUBAME 2.5	
解析手法	有限積分法電磁界解析	
解析周波数	87MHz	
出力	30dBm	
波源	87MHz	直径 50mm, 巻き数 3T コイル
解析モデル	商業施設モデル	
波源の配置位置	商業施設内イートインスペースのテーブル裏	
波源の配置数	2, 5, 10	
解析対象エリア	商業施設含む 40×40×10m エリア	

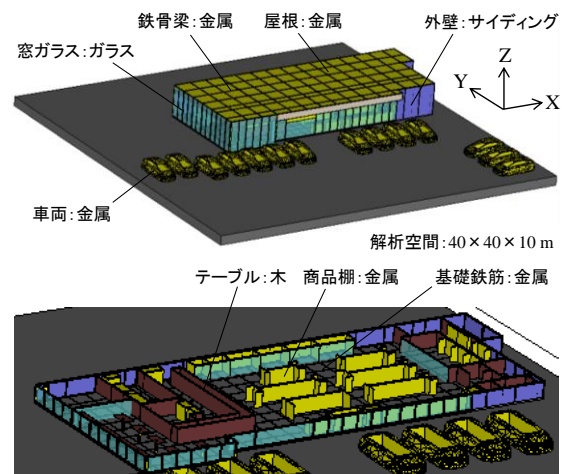


図 2.2 解析モデル（商業施設）

用い、送信電力は 30dBm とした。

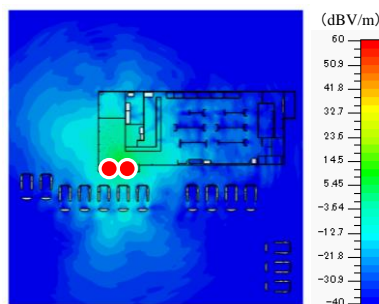
図 2.2 に商業施設モデルを示す。商業施設及びその周辺の駐車場を含む解析空間のサイズは  $40 \times 40 \times 10\text{m}$  である。波源は前述の通り家電モバイル機器の充電向けを想定しているため、商業施設内イートインスペースのテーブル裏に配置する（床からの高さは  $0.75\text{m}$  となる）。波源の向きは図 2.1 及び図 2.2 の座標軸の通りであり、地面に対してコイル開口面が水平となる。波源の数は 2, 5, 10 の 3 通りとした。

### 2.2.2 評価結果

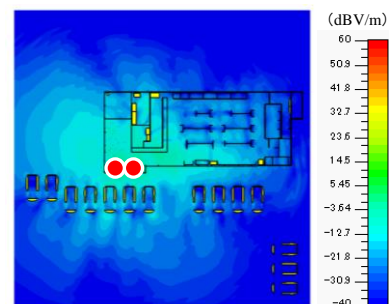
図 2.3 及び図 2.4 に 87MHz における商業施設内とその周辺での漏えい電界分布を示す。地面から  $1\text{m}$  の高さにおける水平面 (XY 面) の分布を記載しており、波源の配置位置を赤丸で示している。2 波源

の場合は駐車場近傍のイートインスペース配置、5 波源の場合は駐車場から最遠の配置、10 波源の場合は最悪条件としてイートインスペース全域に配置した。また、複数の波源が同時動作する場合、波源間の位相差によって漏えい電界強度は変動するため、各波源間の位相差無しの場合と位相差有り(等間隔, 2 波源:  $180^\circ$  刻み/5 波源:  $72^\circ$  刻み/10 波源:  $36^\circ$  刻み)の 2 通りを解析した。

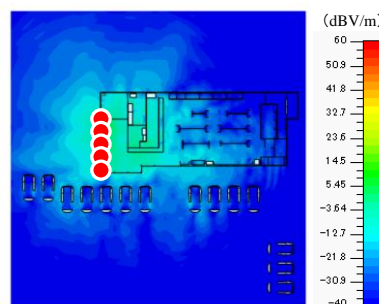
図 2.3, 図 2.4 より、波源の数が増えると漏えい電界強度が増加する傾向や、波源間の位相差によって漏えい電界分布が変化することが確認できる。また、波源として用いたコイルは水平面無指向性のため、漏えい電界は基本的には波源を中心に放射状に分布するが、各構造物の影響による電界強度の変動が見られる。



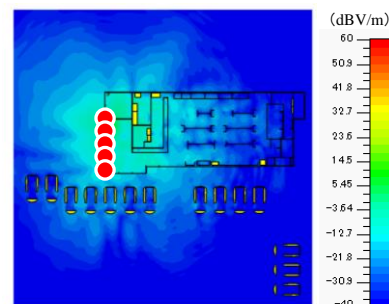
(a) 2 波源



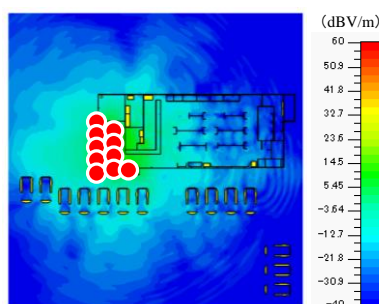
(a) 2 波源



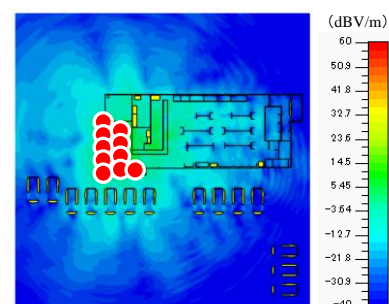
(b) 5 波源



(b) 5 波源



(c) 10 波源



(c) 10 波源

図 2.3 漏えい電界分布 (波源間の位相差無し)

図 2.4 漏えい電界分布 (波源間の位相差有り)

### 3. 幾何光学的電波伝搬解析による漏えい電磁界の解析

本章では、携帯端末及び無線 LAN ルータへの干渉を想定し、WPT システムの高調波である 3.5GHz 及び 5.3GHz の漏えい電界特性を幾何光学伝搬解析手法 (Ray Launching 法) により解析する。

#### 3.1 波源及び受信機器モデル

図 3.1 に本章で用いる波源モデルを示す。各周波数で高い電界強度を放射するために周波数ごとに波源モデルを設計しており、3.5GHz では 20.3×20.3mm 素子のパッチアンテナを波源として用いる。また、無線機器そのものの特性を含まない基本的な漏えい電磁界を解析するため、受信機器をダイポールアンテナで模擬する。

#### 3.2 商業施設における漏えい電磁界の解析

本節では、解析空間サイズをパラメータとして、商業施設モデルに WPT 波源と受信機器を配置し、受信機器を移動させたときのパスロス解析する。

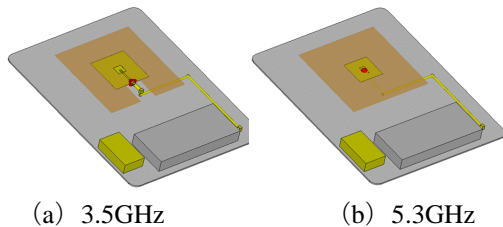


図 3.1 解析モデル (波源)

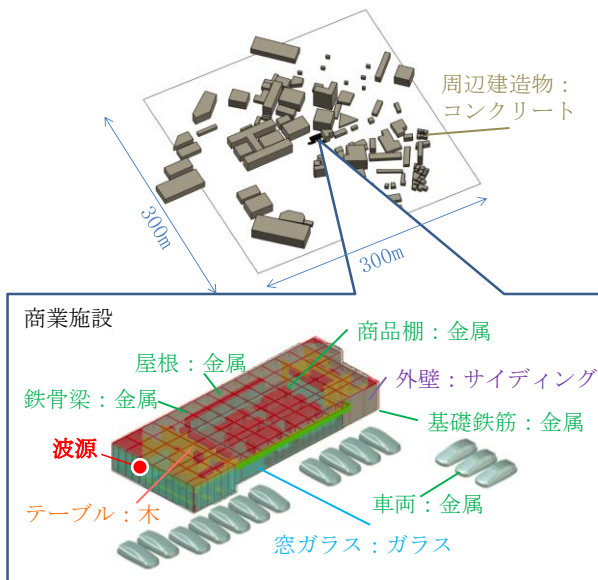


図 3.2 商業施設モデルと波源の配置位置

#### 3.2.1 シミュレーション条件

表 3.1 にシミュレーション条件を示す。波源は 3.1 節で示したパッチアンテナを用い、受信機器はダイポールアンテナとした。なお、送信電力は 10dBm とした。

図 3.2 に商業施設モデルと波源の配置位置を示す。商業施設を中心として、周辺の建造物をモデリングしている。波源は前章同様に、商業施設内のイートインスペースのテーブル裏に配置する。受信機器は図 3.3 に示すように商業施設を含む 40×40m の空間に 0.1m 間隔で格子状に配置しており、受信機器設置位置数は 160,000 点となる。配置高さは 1m とした。図 3.4 に評価する解析空間サイズの条件を示す。解析空間サイズとしては、商業施設を中心として 40×40m, 120×120m, 200×200m, 280×280m の範囲を抽出した 4 種類とする。図 3.5 に示すように、解析空間の境界上にある建造物については解析空間内の壁面のみ残すように設定した。

表 3.1 シミュレーション条件

計算機システム	TSUBAME 2.5	
解析手法	Ray Launching 法電波伝搬解析	
解析周波数	3.5GHz, 5.3GHz	
出力	10dBm	
波源	3.5GHz	20.3×20.3mm パッチアンテナ
	5.3GHz	12.6×12.6mm パッチアンテナ
受信機器	ダイポールアンテナ	
解析モデル	商業施設及び周辺建物モデル	
波源の配置位置	商業施設内イートインスペースのテーブル裏	
解析対象エリア	商業施設含む 40×40×10m エリア	

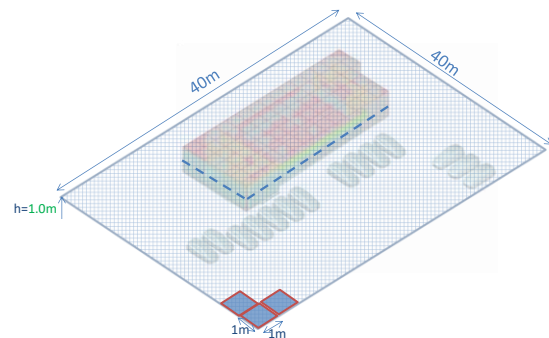


図 3.3 受信機器設置位置



ここで、各受信点ではマルチパスフェージングによるパスロスの瞬時変動が含まれるため、図 3.3 の青色の四角で示すように、1m 四方の領域ごとに 100 受信点のパスロスの中央値を算出する。各解析空間サイズにおいて、図 3.4 に示す全エリア (600×600m) の結果を基準として 1m 四方の領域ごとにパスロス中央値の差分を算出し、その中央値差分の累積確率分布を求める。本解析では、累積確率 10%、50%、90%それぞれにおけるパスロス中央値差分により、解析空間サイズの影響を評価する。

### 3.2.2 評価結果

図 3.6 に解析空間サイズを変更した場合のパスロス中央値差分の累積確率分布を示す。また、表 3.2 に累積確率 10%、50%、90%におけるパスロス中央

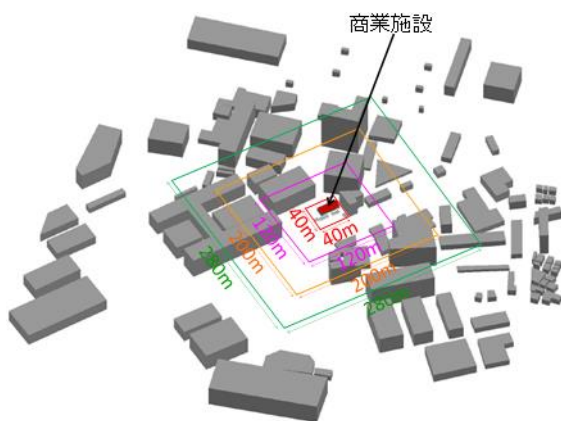


図 3.4 解析空間サイズ (全エリア 600×600m)

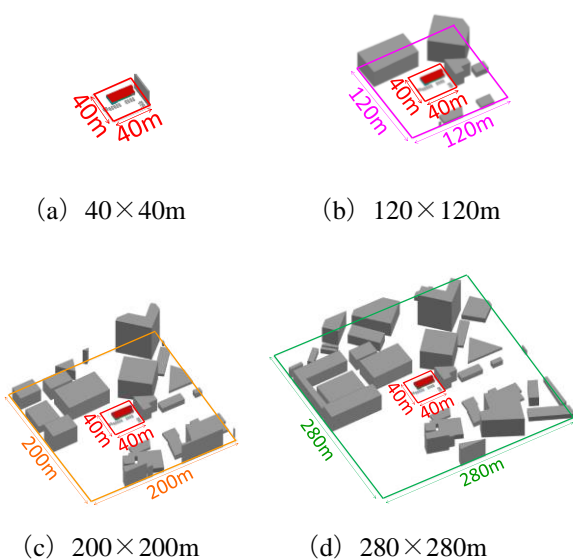
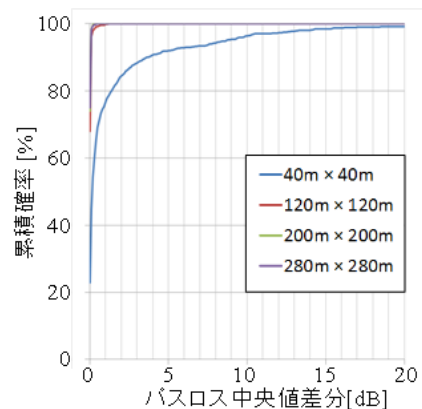


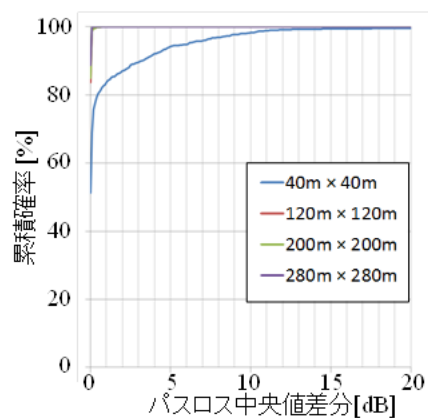
図 3.5 各解析空間サイズの建物モデル

値の差分を示す。これらの結果から、解析空間サイズが 120×120m 以上でパスロス中央値の解析結果が収束することがわかる。

図 3.7 に 3.5GHz の各解析空間サイズにおけるパスロス分布を示す。図 3.7 (a) 40×40m の波源の左側の領域において、他条件と比較してパスロス分布の傾向が異なっているが、これ以外では全条件でほぼ同等の分布となっている。この要因を推測するため、差分が発生している箇所のレイ軌跡を解析する。図 3.8 に 40×40m、120×120m の解析空間サイズにおけるレイ軌跡をそれぞれ示す。40×40m の条件では商業施設の外壁からの透過波のみが到達しているのに対し、120×120m の条件では周辺建築物からの反射波が多数到来することが確認できる。この結果から、波源近傍の建築物からの反射波がパスロス分布の傾向が異なる主要因であるとわかった。また、5.3GHz においても 3.5GHz と同様の分布となる。



(a) 3.5GHz



(b) 5.3GHz

図 3.6 各解析空間サイズでのパスロス中央値差分の累積確率分布

#### 4 まとめ

本利用課題ではスーパーコンピュータ TSUBAME 2.5 を用いて商業施設における漏えい電磁界特性を解析した。より実際に近い WPT 使用環境を想定し、複数の波源が同時動作する場合を解析、波源間の位相差により漏えい電界強度が変動することを確認した。また、解析空間サイズをパラメータとして解析で考慮する商業施設周辺の建造物の影響を検証した。TSUBAME 2.5 を用いることで、商業施設及びその周辺建造物を含む大規模なモデルを解析でき、有限積分法電磁界解析及び Ray launching 法電波伝搬解析における TSUBAME 2.5 の有用性を確認した。

なお、本研究は総務省平成 27 年度電波資源拡大のための研究開発「ワイヤレス電力伝送による漏えい電波の環境解析技術の研究開発」の一部である。

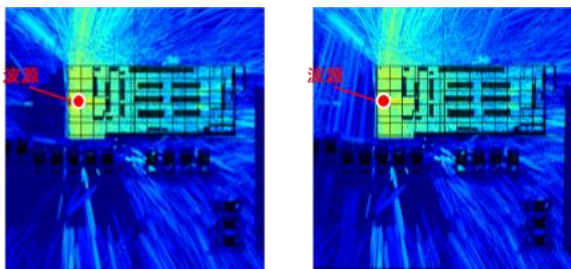
表 3.2 各累積確率におけるパスロス中央値差分

(a) 3.5GHz

解析空間サイズ	累積確率		
	90%	50%	10%
40×40m	4.9dB	0.1dB	0.0dB
120×120m	0.1dB	0.0dB	0.0dB
200×200m	0.1dB	0.0dB	0.0dB
280×280m	0.1dB	0.0dB	0.0dB

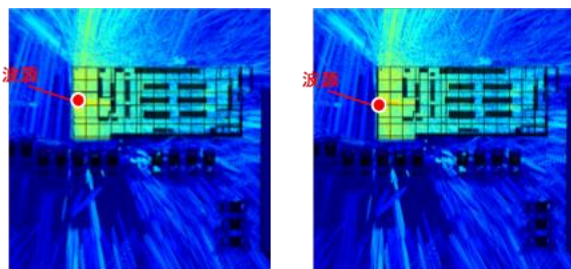
(b) 5.3GHz

解析空間サイズ	累積確率		
	90%	50%	10%
40×40m	4.9dB	0.1dB	0.0dB
120×120m	0.1dB	0.0dB	0.0dB
200×200m	0.1dB	0.0dB	0.0dB
280×280m	0.1dB	0.0dB	0.0dB



(a) 40×40m

(b) 120×120m



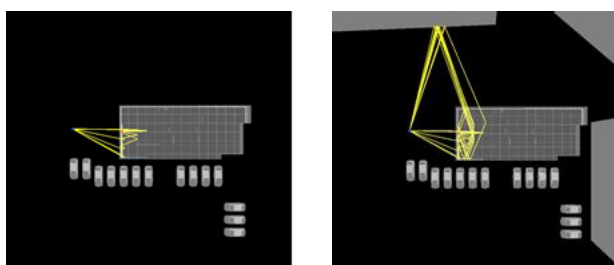
(c) 200×200m

(d) 280×280m

図 3.7 3.5GHz におけるパスロス分布

#### 参考文献

- [1] Q. Chen et al., “Antenna Characterization for Wireless Power Transmission System Using Near field Coupling,” IEEE Antennas and Propagation Magazine, Vol.54, No.4, Aug.2012.
- [2] J. Kim et al., “Electromagnetic Interference and Radiation from Wireless Power Transfer Systems,” 2014 IEEE International Symposium on Electromagnetic Compatibility, Aug. 2014.
- [3] K. Takagi, T. Yui, M. Anada, T. Izumi, H. Uno, H. Watanabe and Y. Saito, “Analysis of Electromagnetic Field Leaked from Wireless Power Transfer System in Case-study House,” WPTC 2015, P2.5, May 2015.
- [4] 東京工業大学 学術国際情報センター, TSUBAME (<http://www.gsic.titech.ac.jp/tsubame>).
- [5] CST MICROWAVE STUDIO (<http://www.cst.com/>)



(a) 40×40m

(b) 120×120m

図 3.8 3.5GHz におけるレイ軌跡

TSUBAME 共同利用 平成 29 年度 産業利用 成果報告書

利用課題名 大容量データ伝送用ミリ波アンテナのレドームに関する基礎検討  
英文 A Study on radome for millimeter-wave antenna

利用課題責任者千葉 修二

Shuji Chiba

所属スタッフ株式会社

STAF corporation.

URL <http://www.staf.co.jp/>

邦文抄録(300 字程度)

屋外で使用される大容量データ伝送用ミリ波帯アンテナに装着されるレドームの形状に着目し検討を行った。

レドーム形状を曲面とし、曲面の曲率を  $R=100\sim 1000$  までの間で変化させた。

TSUBAME を用いた電磁界解析で、レドーム形状の曲面を細かく変化させた時のアンテナ指向性ならびにアンテナ利得を示している。

その結果を元にレドーム形状の曲面について最適値を示した。

英文抄録(100 words 程度)

We studied the radome shape of outdoor millimeter wave antenna.

The shape of the radome was a curved surface, and the curvature of the curved surface was varied between  $R = 100$  and  $1000$ .

The antenna directivity and the antenna gain are shown when the radome-shaped curved surface is finely changed by electromagnetic field analysis using TSUBAME

Based on the result, we showed the optimum value for the radome shape curved surface.

*Keywords:* radome, electromagnetic simulation, horn antenna,

#### 背景と目的

ミリ波帯を使用した大容量伝送用のアンテナは日本国内において、積極的な開発が行われておらず、海外製品がほとんどである。については、アンテナ特性(利得、指向性、VSWR)において、海外製品同等以上のミリ波帯アンテナを無線装置開発メーカー様へ逸早く供給する事は、弊社にとって喫緊の課題であり、この課題解決のために同アンテナ開発を高精度かつ遅滞なく進めることが必要である。屋外で使用されるミリ波帯アンテナにおいて、高いアンテナ性能を維持しつつ、優れた耐候特性を確保する為のレドームは屋外用ミリ波帯アンテナを構成する重要な部品であり、最適化は重要であり大きな課題となる。

本プロジェクトでは、上記課題に対し、電磁界解析を用いて、レドームの影響を明らかにする事を目的としている。平成 26 年度から平成 28 年度にかけて、レドームの厚み、レドーム比誘電率、検討アンテナとレドームの間隔について電磁界解析を実施した。そこで得られた

アンテナ利得、指向性に影響の少なかったレドームの条件を用いて、本年度はレドーム形状(曲面形状)に的を絞り、電磁界解析を用いて基礎的なデータの取得を行い、レドーム形状(曲面形状)の影響を明らかにしている。

#### 概要

本年度は、レドーム形状の曲面の曲率半径の違いによるアンテナ利得、指向性の変化に的を絞り基礎的なデータの取得を行うため、検討アンテナは平成 26 年度から用いている物と同じ条件(図 1)、レドームの厚み、比誘電率、検討アンテナとレドーム間隔は平成 26 年度から平成 28 年度にかけて最適化した結果から、厚み 2mm、比誘電率 4.05 間隔 16.7mm(図 2)とした。レドームの影響を確認する為に電磁界解析実施モデルは、レドームなしモデル図 1 を基本状態とし、次にレドームありモデル図 2 としている。

レドーム形状の曲面の曲率半径を  $R=100\sim 2000$  の範囲で変化させた物とレドームなしモデルの比較を行う

事で E-Band 帯の 71GHz、76GHz、におけるレドームの影響を確認した。

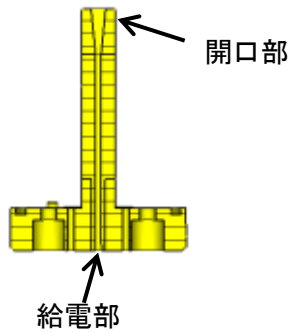


図 1 ホーンアンテナ:角すいホーン  
レドームなしモデル基本状態

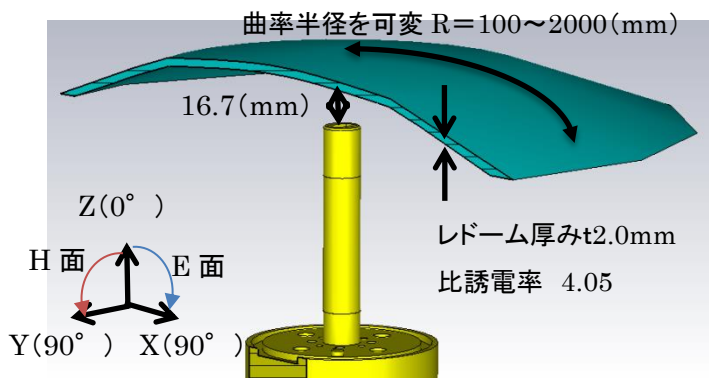


図 2 レドームありモデル

### 結果および考察

図 3(a) (b)に図 1 と図 2 の電磁界解析の結果としてレドーム形状の曲率を変化させた時の利得変化を示す。

71GHz 帯について、曲率を変化させることで利得が変化している事が解る。利得の変化には傾向があり、曲率が大きくなるにつれて利得は低下傾向にある。

76GHzについても、71GHz 同傾向で曲率が大きくなるにつれて利得が低下している事が確認できる。

レドームの影響がない状態=レドームなしの状態とすると、すべての状態において利得の向上が見られる、これは昨年度報告した検討アンテナとレドーム距離において、利得劣化が見られない距離関係を保っている事が起因していると考えられる。

利得の観点から見ると 71GHz 帯では曲率 R=100~200mm では変動が大きい、変動が緩やかになる

R=300 以上が良く、76GHz帯では R=100~500mm の間で利得の変動は少ない。以上のことから R=300 程度が良いと考えられる。

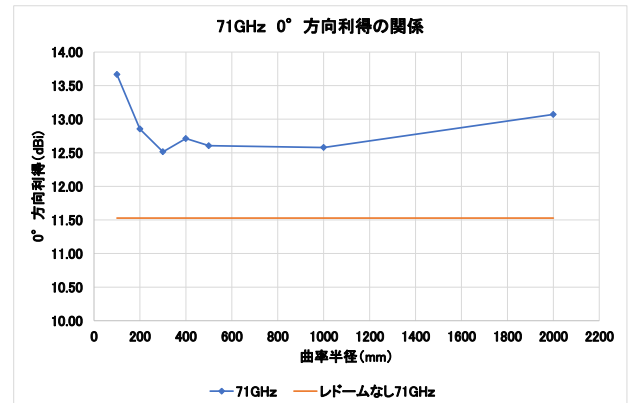


図 3(a) 71GHz 帯利得

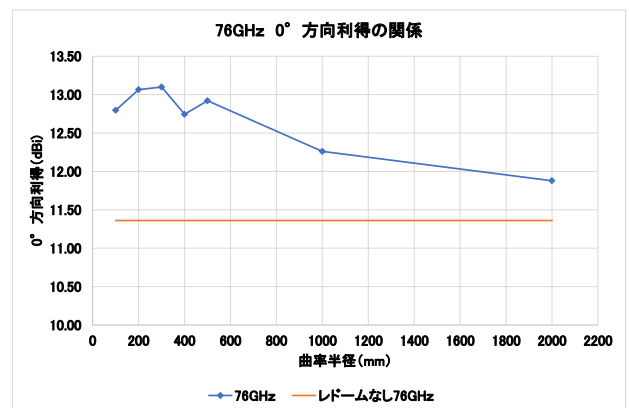


図 3(b) 76GHz 帯利得

次に図 4(a)、(b)、(c)、(d)に 71GHz、76GHzの指向性を示す。

71GHz、76GHz の E 面(ZX 面)、H 面(ZY 面)共に 10~40 度付近はレドーム形状の曲面の違いで指向性が大きく変化する事は無く、レドーム形状の曲面の影響は 0° 付近で現れる。

0° 付近の影響が大きいのは、電波の入射角によるものと考えられ、入射角がレドーム対し垂直になるにつれて、検討アンテナに戻ってくる反射波が多くなり、ちょっとした形状の変化で大きく特性が変わるものと推測する。



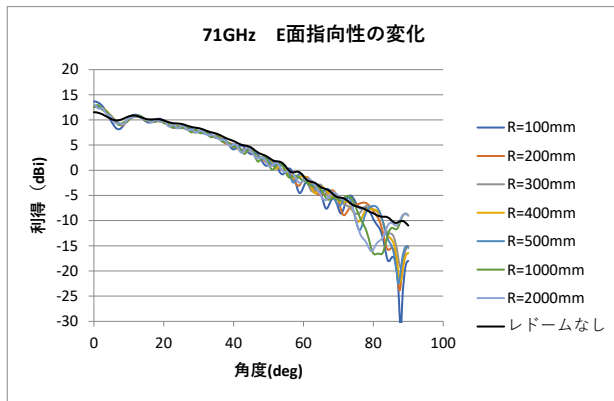


図 4(a) 71GHz 帯 E 面指向性

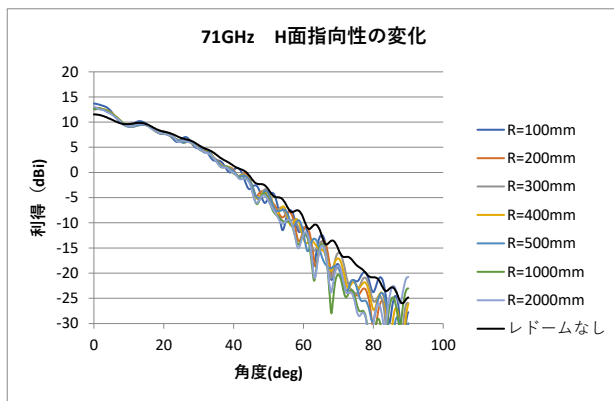


図 4(b) 71GHz 帯 H 面指向性

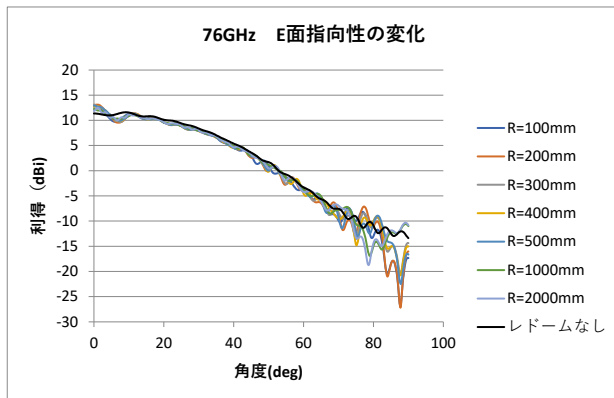


図 4(c) 76GHz 帯 E 面指向性

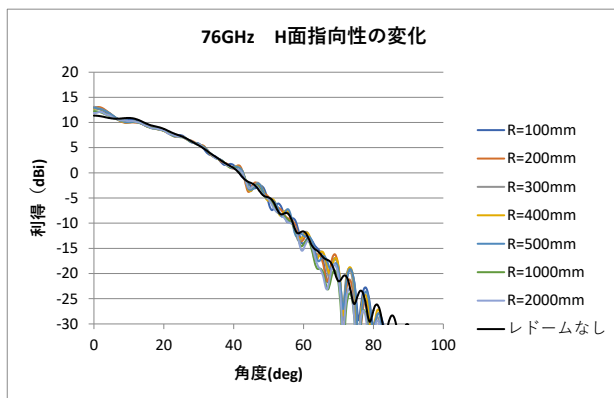


図 4(d) 76GHz 帯 H 面指向性

### まとめ、今後の課題

ミリ波アンテナにおけるレドーム形状曲面の曲率半径の違いによる影響を電磁界解析にて確認した。曲率半径を変化させることで、利得が変化する事が確認できた。また指向性においては、0° 付近での影響が大きい事が確認できた。

今回の条件においては、71GHz、76GHz双方の利得の傾向から利得変動の少ない曲率半径 R=300mm と判断している。

平成 26 年度～今年度に掛けてレドームに関する課題を検討してきた結果、レドームの基礎的なデータは概ね取得できたと考えている。

今後の課題としては検討アンテナの違いによる、レドームの影響があげられる。検討アンテナを変える事でレドームの影響度合いは変化する物と考えている。検討アンテナを変化させ、平成 26 年度から今年度までで最適化を行ったレドームを用いて、レドームあり、なしの特性比較を行う事で、検討アンテナとレドームの関係を明らかにしたい。



TSUBAME 共同利用 平成 29 年度 産業利用 成果報告書

利用課題名 産業利用のための大規模ディープラーニング  
英文: Large-scale Deep Learning for Industrial Use

利用課題責任者  
Takashi Yamazaki

所属  
株式会社 富士通研究所  
<http://www.fujitsu.com/jp/labs/>

邦文抄録(300 字程度)

人工知能研究の中心の一つであるディープラーニング(DL) 技術は、多くのフレームワークが公開されていることもあり、ものづくりから医療まで、様々な産業のフィールドで広く用いられ、特に画像・音声認識に威力を発揮し始めている。しかし、応用の際にはまだ課題も多い。例えば DL の性能は学習データの量や質に大きく依存するが、既存のフレームワークを用いるだけでは、大規模データを効率良く使うことはできない。

我々のプロジェクトでは、DL を大規模並列化し、ものづくりの現場で利用できる効率の良い DL を実装したいと考えている。トライアル的な利用において、既存 DL フレームワークの効率を含む基礎的な調査を行った。

英文抄録(100 words 程度)

Deep learning (DL) technology, which is one of the one of the biggest trend of cutting-edge artificial intelligence technology, has many open frameworks, and is widely used in various industrial fields from manufacturing to medical treatment, and in particular, image and sound recognition. However, there are still many problems in application. For example, DL performance largely depends on the amount and quality of training data, but using existing frameworks cannot efficiently use large-scale data.

In our project, we would like to massively parallelize DL and implement efficient DL code that can be used in the field of manufacturing. A basic survey including efficiency of existing DL code in MPI + GPU environment is conducted in a one-month trial use, and detailed issues are extracted.

*Keywords:* deep learning, artificial intelligence, industry, parallelization

#### 背景と目的

ものづくりの分野においては、ディープラーニング(DL) に代表される AI 技術を用いた、開発過程の効率化や自動化、或いは最適なパラメータやデザインを求めるといった応用がなされ、技術者・施術者の経験値や勘・直感を一般化する取り組みが始まっている。現実の問題に応用する際の最大の課題の一つは、複雑な構造を持つ既存の製品やモックアップを学習データに用いる際に必要となる高解像度化である。この問題を解決するために、大きなデータを扱うことができ、現場で利用できる精度と効率を持つ DL の実装が強く求められている。

本プロジェクトでは、既存の DL フレームワークを用いて、ものづくりの現場で利用できる高効率な DL 実装の評価・検討を念頭に、基礎的な調査を行うことを目的とした。

#### 結果および考察

TensorFlow を TSUBAME 上で実行し、基礎的なフレームワークの環境整備を行うことができた。基礎的な調査として、TSUBAME 上の TensorFlow を用いて、画像データの分類を行うプログラムを実装し、高精度な分類性能を確認した。

#### まとめ、今後の課題

今回の利用では、利用期間が一ヶ月と短かったため、十分な考察を行うことができなかったが、TSUBAME を利用することで、今後の土台となる基礎的なノウハウを得ることができた。今後は、高解像度のデータを用いた学習の精度および効率を評価するとともに、MPI を用いた高並列度の実行に関して調査・検討を進める。



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 廃棄物最終処分場における間隙内流体挙動の数値解析  
 英文: Numerical analyses of fluid dynamics in pores of landfill waste layer

利用課題責任者 鈴木和将  
 Kazuyuki Suzuki

所属 埼玉県環境科学国際センター  
 Affiliation Center for Environmental Science in Saitama  
 URL <http://www.pref.saitama.lg.jp/cess/index.html>

#### 邦文抄録

本研究では、廃棄物最終処分場内部の水やガスの流れ問題の高品質な計算スキームの開発を目的として、数値流体解析の検討を行った。具体的には支配方程式に Navier-Stokes 方程式を用いて SUPG/PSPG 法に基づく安定化有限要素法により離散化した。さらに、GPGPU コンピューティングによる計算の高速化を試みた。連立一次方程式の求解には GPBi-CG 法を導入し、マルチ GPU を利用して高速並列計算によるシミュレーションを行った。

#### 英文抄録

The purpose of this study is to establish a numerical simulation model of fluid flow in a landfill layer with high quality and high precision. A 3D finite element method is a powerful tool for flows having complex geometry such as porous media in landfill. However, it leads a huge amount of computation cost. In this study, we examined to accelerate the 3D FEM by using the Graphics Processing Unit as a general-purpose use (GPGPU).

*Keywords:* landfill, Numerical simulation, FEM, Navier-Stokes equations, GPGPU

#### 背景と目的

我々はこれまで、廃棄物最終処分場内部の水やガスの流れ問題の高品質な計算スキームの開発を目的として、埋立廃棄物のコアサンプル等の CT 画像から間隙形状を抽出する三次元モデリングを行い、実際の形状を反映した内部の流れ解析を検討してきた<sup>1,2)</sup>。モデル化した間隙は複雑な形状をしており、このような中の流れ解析においては、任意形状の適合性に優れた有限要素法による数値解法が有効である。しかし、非圧縮粘性流体の支配方程式である Navier-Stokes 方程式と連続式に対して有限要素法を適用すると、移流項の卓越や非圧縮性に起因して数値的に不安定な挙動を生じることがある。本研究では、この数値的な不安定性を取り除く手法として Tezduyar ら<sup>3,4)</sup>によって提案された SUPG (Streamline Upwind/Petrov-Galerkin) / PSPG (Pressure Stabilizing/Petrov Galerkin)法に基づく安定化有限要素法を適用して離散化を行った。安定化有限要素法は、汎用性が高く、間隙内の流れ解

析においても有用な方法ではあるが、要素数が多く大規模な流れ解析を行う場合、計算量が増加し、計算に莫大な時間を要することとなる。実問題に適用していくためには、大規模化に伴う、計算の高速化が求められている。

近年、画像処理を目的に開発されてきた GPU (Graphics Processing Unit) を画像表示の目的ではなく、科学技術計算に利用する GPGPU (General Purpose Graphics Processing Unit)の研究が、天体や気象等様々な分野で盛んに行われている。GPU 上には、数多くの演算コアが搭載されており、これらのコアを用いた並列計算を行うことにより、大規模な問題に対して計算時間を大幅に短縮することが可能となってきた。数値流体解析においても、このような GPGPU コンピューティングを利用した高速化への試みが図られてきている。非圧縮性流体 Navier-Stokes 方程式を解析する際、解は係数行列が大規模疎行列となるような連立一次方程式を反復解法で解くことで求まるが、このプロセスが計算処

理の最も重い部分となりやすい。このため、連立一次方程式の求解の高速化が重要な課題であり、GPUによる高速化が望まれる。

そこで、本研究では、安定化有限要素法を適用した廃棄物埋立層内の流れ解析に対して、GPU スパコン TSUBAME を用いた計算の高速化を試みた。

## 概要

マイクロフォーカス X 線 CT 装置を用いて撮影した埋立廃棄物のコアサンプル等の CT 画像を画像処理ソフトウェアに読み込み、モデル間隙形状の作成を行った。さらに、作成した間隙形状のデータを STL ファイルへ変換し、この STL ファイルを基に、自動メッシュ生成ソフトウェア Gmsh を用いて有限要素メッシュの生成を行った。三次元モデリングを行った結果を図 1 に示す。

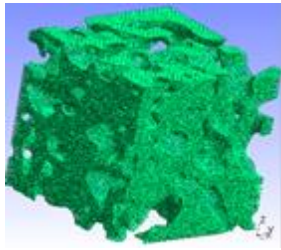


図 1 三次元間隙構造モデル

支配方程式には、非圧縮性 Navier-Stokes 方程式、連続の式を用いた。

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{1}{Re} \frac{\partial}{\partial x_j} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \text{ in } \Omega, \quad (1)$$

$$\frac{\partial u_i}{\partial x_i} = 0 \text{ in } \Omega. \quad (2)$$

ここで、 $u$  は流速、 $p$  は圧力、 $Re$  は Reynolds 数、 $\Omega$  は計算領域である。

上記(1),(2)の支配方程式に対して、安定化有限要素法 (SUPG/PSPG 法) を適用すると以下のように弱形式が導かれる。

$$\begin{aligned} & \int_{\Omega} w_i \left( \frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} \right) d\Omega - \int_{\Omega} \frac{\partial w_i}{\partial x_i} \\ & + \int_{\Omega} \frac{1}{Re} \frac{\partial w_i}{\partial x_j} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \\ & + \sum_{e=1}^{n_{el}} \int_{\Omega} \left( \tau \bar{u}_k \frac{\partial w_i}{\partial x_k} \right) \left( \frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} \right) d\Omega = 0, \quad (3) \end{aligned}$$

$$\int_{\Omega} q \frac{\partial u_i}{\partial x_i} d\Omega + \sum_{e=1}^{n_{el}} \int_{\Omega} \left( \tau \frac{\partial q}{\partial x_i} \right) \left( \frac{\partial u_i}{\partial t} + \bar{u}_j \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} \right) d\Omega = 0, \quad (4)$$

ここで、 $w_i, q$  は、それぞれ式(1),(2)に対する Galerkin 項の重み関数である。また、 $\bar{u}_i$  は移流速度を表し、 $\tau$  は安定化パラメータを表している。これら(3),(4)式に対して Crank-Nicolson 法により時間方向の離散化を行うことで、有限要素方程式が導かれる。導かれた連立一次方程式の解法には GPBi-CG 法を適用した。

本研究では、計算の高速化を達成するために、GPGPU を用いた並列計算を取り入れ、NVIDIA 社の統合開発環境である CUDA とそれらのライブラリ (cuBLAS, cuSPARSE) を用いて GPBi-CG 法を適用した。

## 結果および考察

埋立処分場間隙内の流れの数値シミュレーションを行った。時間刻み幅を  $1.0 \times 10^{-2} \text{sec}$  として 100 ステップを計算し、解析対象とした有限要素メッシュは、節点数 99,560、要素数 437,386 である。計算効率を検討するために、CPU1 コア(1CPU)、CPU1 コアと GPU1 コア(1CPU+1GPU)、CPU2 コア(2CPU)、CPU2 コアと GPU2 コア(2CPU+2GPU)のそれぞれを用いた時の解析に要した時間を計測した。計測結果を表 1 に示す。

表 1 計測結果

数値計算環境	計算時間(sec)
1CPU	$1.48 \times 10^5$
1CPU+1GPU	$1.65 \times 10^4$
2CPU	$7.69 \times 10^4$
2CPU+2GPU	$9.06 \times 10^3$

1GPU では、1CPU と比較して 9.0 倍高速化を実現した。2GPU+2CPU では、2CPU と比較して 8.5 倍、1CPU と比較して 16.3 倍と大幅に高速化した。また、図 2 に数値シミュレーション結果を示す。速度ベクトルの様子は、GPU、CPU どちらも同じ結果となった。以上の結果より、埋立層間隙内流れ問題を対象とした流体解析において GPU 並列計算の有用性を確認

できた。

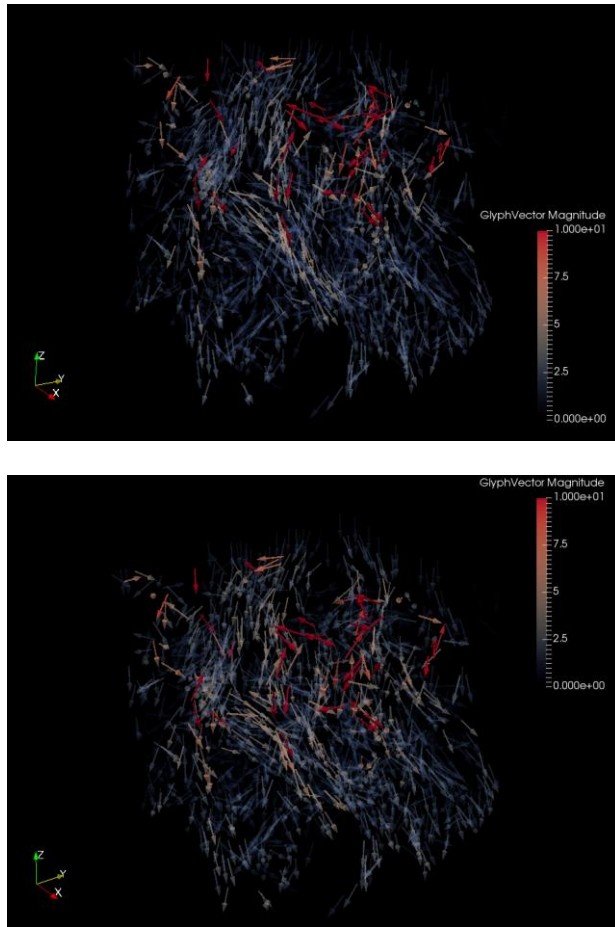


図 2 数値シミュレーション結果(上図:1CPU+1GPU  
計算結果、下図:1CPU 計算結果)

#### まとめ、今後の課題

今後、より長期で広い空間の将来予測を高品質で可能とする数値解析手法の枠組みを構築していく。

#### 参考文献

- 1) 鈴木和将, 水藤寛: 廃棄物最終処分場内部における間隙構造の可視化と間隙中流れの数値的研究、第 26 回廃棄物資源循環学会研究発表会講演原稿, 441-442 (2015)
- 2) 鈴木和将, 水藤寛: 廃棄物最終処分場における埋立層間隙内の気液二相流数値解析、第 27 回廃棄物資源循環学会研究発表会講演原稿, 451-452 (2016)
- 3) Tezduyar, T.E.: Stabilized finite element formulations for incompressible flow computations, *Advanced in Applied Mechanics*, 28, 1-44(1991)
- 4) Tezduyar, T.E., Mittal, S., Ray, S.E. & Shih, R.: Incompressible flow computations with stabilized bilinear and linear equal-order-interpolation velocity-pressure elements, *Computer Methods in Applied Mechanics and Engineering*, 95, 221-242(1992)





利用課題名 高解像度画像を使った広域の家屋及び道路の深層学習による自動判別システムの開発  
英文: System development of automatic building and road network detection using high resolution imagery by deep-learning

柴崎亮介  
Ryosuke Shibasaki

東京大学, 空間情報科学研究センター  
Center for Spatial Information Science, The University of Tokyo  
shiba@csis.u-tokyo.ac.jp

深層学習手法による衛星画像からの特徴量抽出(家屋, 道路, 車両等の検出)についての研究を行った。家屋, 道路, 車両などの検出を検討したが, 今年度は, サンプルデータの準備と検証の容易さから主に車両を対象に研究を進めた。成果の一つとして, Hard Example Mining(HEM)と呼ばれる手法の研究を行った。深層学習手法のトレーニングにおいて, 対象地域内にある車両(ミニバン)中で, 誤検出されやすい画像(ロス値の大きい画像)を対象に学習することという手法である。この手法により, 精度を改善することができた。

Deep learning allows us to detect feature objects (houses, roads, vehicles, etc.) in satellite image, but there is space to improve accuracy and the sample images preparation. We advanced research mainly for vehicles from the ease of preparation and verification of sample data. As one of the accomplishments is a method called Hard Example Mining (HEM). In the training of the deep learning method, it is a method to learn an image (image with a large loss value) that is easily erroneously detected in a vehicle (minivan) in the target area. By this method, accuracy could be improved.

*Object detection, Deep Learning, House-footprint, Vehicle detection*

#### 背景と目的

社会基盤情報として人口分布, 交通ネットワークの整備は必須であるが, 途上国, 僻地の情報基盤の整備は乏しい状況にある。オープンな利用が広まりつつある高解像度衛星画像から機械学習, ディープラーニングなどの手法で家屋や道路ネットワークを自動検出し, 地図データと利用することによる広域・社会基盤情報の整備に寄与する。

宇宙技術の競争が世界的に進む中, その応用利用の社会貢献が注目されている。衛星リモートセンシングの分野では, 衛星画像の利用が容易になり, それらを活用した付加価値のある情報の生成, 社会基盤情報の整備のオープン化が進んでいる。社会基盤情報の一つである地図情報も著作権が自由な情報が整備されつつあるが, 地上情報の整備手法のオープン化はあまり進んでいない。本研究では, 高解像度衛星画像を元に家屋を自動検出し地図データ化する手法を開発することで, 社会基盤情報の整備を加速させる一方で, 画像判別・深層学習・を使った地図情報の整備・更新, 検出

手法の改良(アルゴリズム・サンプリング・対象範囲の条件など)の研究開発を行い, 地上情報の整備手法のオープン化を推進する。

本プロジェクトは少ないサンプルデータで検出精度を高めるために, どのようなサンプルデータを準備するのが最適なのか, 手法と検出精度の向上との関係を調べた。家屋検出, 道路検出を目的とするが, 手法開発においてサンプルデータの準備と検出結果の評価の容易性から今年は, 主に車両を対象として手法の開発を行った。

#### 結果および考察

成果の一つとして, Hard Example Mining(HEM)と呼ばれる手法の研究を行った。

具体的には, 図1に示すように, 車両をターゲットとする深層学習手法のトレーニングにおいて, 間違っ検出しやすい車両(ミニバン)の画像(ロス値の大きい画像)のみを選んで学習することという手法である。これにより, 図2のように, 精度を改善することができた。



図1 HEM のアルゴリズム

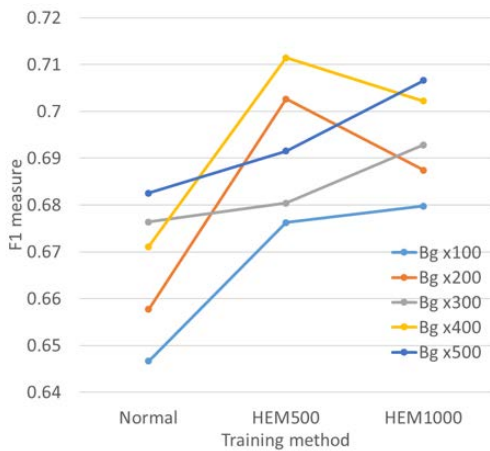


図2 HEM による精度改善

実験にあたっては、実験結果のふらつきを考慮して定量的な検討を行うため、同じ条件での実験を複数回行ったが、TSUBAME の並列計算能力によって、これを効率的に行うことができた。

また、もう一つの成果として、ドメインアダプテーション手法の研究を行った。深層学習を含む機械学習では、異なる分布(ドメイン)のデータがある場合に、一方のドメインでトレーニングしたモデルをもう一方のドメインに適用すると一般的に精度が大きく低下する。ドメインアダプテーションとは、両ドメイン間の共通の特徴空間を見出す等により、教師無しで精度低下を改善する手法である。本研究では、CORAL Domain Adaptation(DA)及び Adversarial DA という手法を設計した(図3, 4)。これにより、精度低下幅の半分以上、大幅に精度を改善させることができた(図5)。

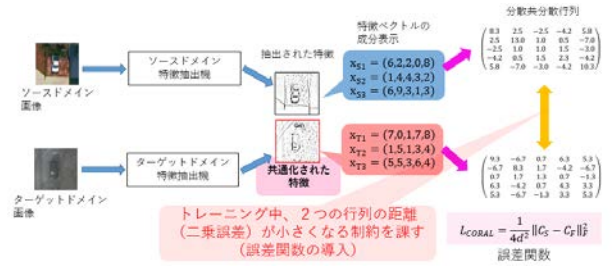


図3 CORAL DA のアルゴリズム

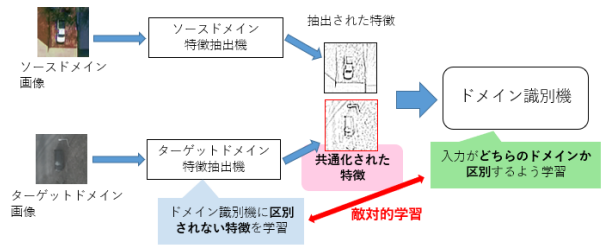
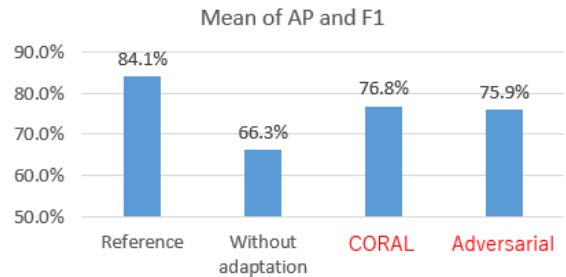


図4 Adversarial DA のアルゴリズム



(a) DA による精度改善幅



(b)車両検出結果の比較

図5 DA による精度改善

まとめ、今後の課題

本実験では、ベースとなる検出手法に比較的新しい手法を用いたほか、DA 手法も計算量が大きく、必要リソースが大きかったが、tsubame3 の最新 GPU スペック

クにより比較的容易に実装することができた。また、DA  
手法の設計にあたってはアルゴリズムのほか、パラメ  
ータ等も様々な検討を行う必要があったが、これも  
TSUBAME の並列計算能力によって効率的に検討を  
行うことができた。



利用課題名 蛋白質-リガンド複合体の会合・解離の速度定数を平衡論から導く  
英文: estimation of protein-ligand association/dissociation rate constant from equilibrium

利用課題責任者: 肥後順一  
First name Surname: Junichi Higo

所属: 大阪大学蛋白質研究所  
Affiliation: Institute for Protein Research, Osaka University

## URL

一般に効率的構造探索法からは、速度定数を見積もることはできない。構造探索効率をあげることを優先することで、実時間で規定される物理量の計算を犠牲にしている。我々は、効率的構造探索法から速度定数を算出することを提案している。具体的には、二つの生体高分子の会合と解離の速度定数を算出するために、アラニン2分子を溶液中に配置し、効率的構造探索法である「アダプティブアンブレラサンプリング法」と「仮想系とカップルしたカノニカ分子動力学法」を、TSUBAME 3.0 上で行った。その結果、多数回の会合と解離を観測し、多数のスナップショットを得た。以上で本プロジェクトの目標を達成した。これにより速度定数を導き出す準備を整えた。

To calculate rate constants for association and dissociation of biomolecules in an explicit solvent, we performed two types of effective conformational sampling methods: *Adaptive umbrella sampling* and *virtual-system coupled molecular dynamics*. The system consisted of two alanine molecules embedded in an explicit solvent. We obtained the association and dissociation events many times for the system during the simulations with using the TSUBAME 3.0 system, and stored a large number of snapshots from the simulation trajectories. These snapshots are used for calculating the rate constants.

*Keywords:* reaction coordinate, enhanced conformational sampling, rate constant, free-energy landscape, molecular dynamics

## 背景と目的

最近、蛋白質をはじめとする生体高分子の立体構造探索が計算シミュレーションによって盛んに行われている。ここ10年、様々な効率的構造探索法が開発されたおかげで、生体高分子の大規模な立体構造変化や生体高分子複合体の会合・解離の自由エネルギー地形の計算が、徐々に可能になってきた。これらの手法では、平衡論としての自由エネルギー地形が得られる。しかし、その方法では速度論としての速度定数は得られない。一方、実験的な研究では、平衡論的な量だけでなく速度定数も得られることがある。一般に速度定数の算出には効率的構造探索法は使えず、従来のカノニカル MD 法で行うしかない。したがって膨大な計算時間がかかる。速度定数を簡単に得られる計算手法が望まれる。

本プロジェクトでは、効率的構造探索法である「アダ

プティブアンブレラサンプリング(AUS)法」や「仮想系とカップルしたカノニカ分子動力学(VcMD)法」を行い、系の自由エネルギー地形を効率的に求めると同時に、それ以上余分な計算をせずに速度定数を求めることを目的とする。我々は、VAUS 法や VcMD 法を TSUBAME3.0 上で実行することで、溶液中でのアラニン2分子の会合・解離の過程を多数回実現することを目指した。

## 概要

アミノ酸の一種であるアラニン2分子を水中に配置し(図1)、AUS 法および VcMD 法を行った。その結果、アラニン2分子が水中で何度も会合と解離を繰り返した。このデータを使い、会合と解離をつかさどる自由エネルギー地形を算出した(図2A)。

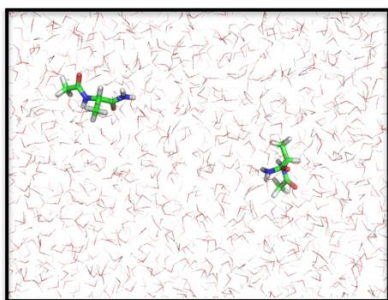


図1: 溶液中のアラニンの二分子

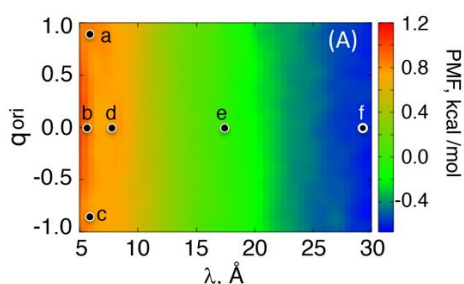


図2A: 300K での自由エネルギー地形。横軸が2分子の重心間距離、縦軸が2分子の相互配向(分子の向きを表す単位ベクトルの内積)。

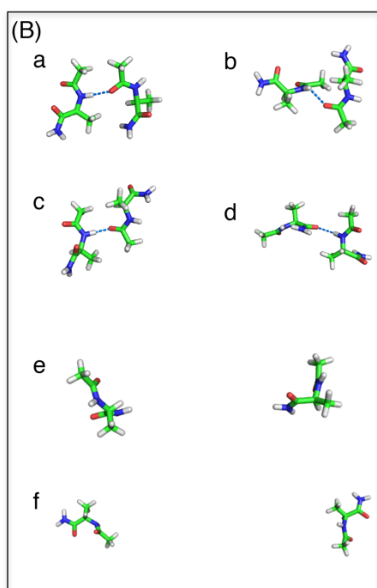


図3B: 図3A で示された地点 (a,...,f) での2分子相互配置。

図2B は、自由エネルギー地形の数力所から抜き出した、二分子の相対配置である。会合・解離の様々な状態が得られたことがわかる。

さて、シミュレーションからは多数のスナップショットが得られるが、スナップショットの時系列は実時間上で構造変化を表していない。一般に効率的構造探索法ではバイアスポテンシャルが導入されており、そのお

かげで構造探索が効率的になる。一方、そのバイアスのおかげで、スナップショットの時系列からは実時間上の構造変化の概念が失われる。

しかし、隣り合うスナップショットの間の構造変化からは、理論上バイアスポテンシャルの影響を消し去ることができる。我々は、シミュレーションから得られた多数のスナップショットを保存した。ここまでの本プロジェクトの目標である。これらのスナップショットから、バイアスポテンシャルの効果を消し去り、速度定数の計算の準備が整った。

### 結果および考察

これらのシミュレーションから、アラニン2分子の会合と解離の様子を図3に示す。これは全計算の一部であり、シミュレーションの全長では多数回の会合・解離が得られた。したがって、このデータから速度定数を導き出すための統計性は充分であると考えられる。ここまでの結果は、J. Comput. Chem. (T. Hayami, K. Kasahara, H. Nakamura, J. Higo. 2018. doi:10.1002/jcc.25196)に発表した。本プロジェクトの目的は、TSUBAME3.0 を使い、多数のスナップショットを作成することであったので、目的は達成された。このデータをもとに、今後、会合・解離の速度定数を算出する。

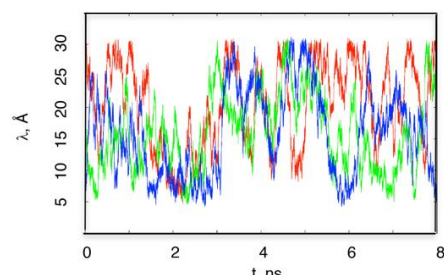


図3: アラニン分子の重心間距離の時間変化。ある時間間隔での3本の trajectories を示した。

### まとめ、今後の課題

前節で記述したように、TSUBAME3.0 を使い、会合・解離という大きな構造変化を表すスナップショットを作成できた。さらに複雑な系の大規模構造変化の速度定数を計算するためには、計算が大規模になる。今後、TSUBAME を使ったシミュレーションに関する経験を蓄積する必要があると考えられる。

TSUBAME 共同利用 平成29年度 学術利用 成果報告書

利用課題名 先進的ステンシルコード技術  
 英文: Advance Stencil-Code Engineering

利用課題責任者  
 千葉 滋  
 Shigeru Chiba

所属  
 東京大学 情報理工学系研究科創造情報学専攻  
 Grad. School of Information Science and Technology  
 The University of Tokyo  
<http://www.csg.ci.i.u-tokyo.ac.jp>

**邦文抄録(300 字程度)**

高度なステンシル計算のためのドメイン専用言語 (DSL) の開発を行い、その実行時性能を検証するため、NAS Parallel Benchmark の CG, LU 相当のプログラムを DSL で書き、TSUBAME 上で元の Fortran によるプログラムと実行時性能の比較をおこなった。本研究課題は JST CREST/独 DFG による日独共同研究 SPPEXA の ExaStencil プロジェクトの中で実施している。ExaStencil プロジェクトで開発している ExaSlang の性能測定の一部も TSUBAME 上でおこなった。さらに DSL の実行基盤として分散システムの資源管理システムの研究もおこなった。具体的には Hadoop のオリジナル実装を MPI 環境で効率よく動かすための研究を TSUBAME をテストベッドとしておこなった。

**英文抄録(100 words 程度)**

We have developed a domain specific language (DSL) for high-performance stencil computing and examined its execution performance on TSUBAME by writing equivalent programs in that DSL to NAS Parallel Benchmark CG and LU. This work was performed under the ExaStencil project of the joint research program SPPEXA of JST CREST and German DFG. Part of the performance evaluation of the ExaSlang language developed in the ExaStencil project was also conducted on TSUBAME. Furthermore, we also studied a distributed resource management system for the execution environment of DSLs. We used TSUBAME to study an efficient MPI-based execution platform for the original implementation of Hadoop.

*Keywords:* domain specific languages, reflection, deep reification, stencil computing, Java.

**背景と目的**

ステンシル計算は多くの科学技術分野で頻繁に使われる計算でこれまでも様々なドメイン専用言語やライブラリが開発されてきた。一方で、ドメイン専用言語はせまい応用範囲 (ドメイン) のために新規にコンパイラやライブラリ、開発環境を整備しなければならず、開発コストの大きさが課題である。ドメイン専用言語の開発コストという課題を解消する手法として注目されているのが埋め込み型ドメイン専用言語 (Embedded Domain Specific Languages) と呼ばれる手法である。この手法では、ドメイン専用言語は汎用プログラミング言語 (ホスト言語とよ

ぶ) のライブラリとして実装される。ライブラリではあるが、そのライブラリを用いた書かれたホスト言語のプログラムが、あたかも専用の言語で書かれたかのように見えるよう、ライブラリのプログラミング・インタフェースを注意深く設計する。しかし、埋め込み型ドメイン専用言語は、ホスト言語のライブラリであるが故に、開発コストが比較的小さくできるが、実行時性能は必ずしも満足いくものではなかった。そのドメイン固有の実行時性能最適化技法などの適用が難しくなりがちだからである。

本研究課題は、最新のプログラミング言語の技術を適用することで、高い実行時性能をもつ埋め込み

型ドメイン専用言語の開発手法を明らかにすることを狙っている。昨年度に引き続き、ホスト言語に Java 言語を用いたドメイン専用言語を開発するためのプラットフォーム Bytespresso の開発を中心に、TSUBAME を用いて研究をおこなった。

## 概要

Bytespresso を用いた簡単なステンシル計算向けドメイン専用言語の性能測定については昨年度に実施し、良好な結果を得た。本年度はステンシル計算に限らずより一般的なドメイン専用言語を作成した場合の性能を実験するため、NAS Parallel Benchmark の中から CG と LU を選び、これを記述するためのドメイン専用言語を Bytespresso を用いて開発し、実際にベンチマークと等価なプログラムを作成し、この性能測定を実施した。特に CG のためには行列およびベクトル演算を MPI 環境で実行するためのドメイン専用言語を新たに開発した。このため、このドメイン専用言語を用いて書き直した CG プログラムは、ベクトルや行列オブジェクトを用いた抽象的な記述になっており、元の Fortran によるプログラムに比べて注釈行を含んだ状態で 30%ほど行数が減少している。

本研究課題は JST CREST/独 DFG による日独共同研究 SPPEXA の ExaStencil プロジェクトの中で実施している。このプロジェクトでは、内部 DSL だけでなく外部 DSL の研究開発もおこなっており、ステンシル計算のための外部 DSL を Scala 言語で開発している。ExaSlang と呼ばれるこの外部 DSL の GPU 搭載スーパーコンピュータ上での実行時性能を TSUBAME を用いて評価した。

加えて、Java 言語で実装された分散処理フレームワークである Hadoop をスーパーコンピュータ上で効率よく動かすための研究を TSUBAME 上でおこなった。通常 Hadoop のようなシステムは各計算ノードにローカルディスクを備えた PC クラスタシステムで動かすが、現在運用中の典型的な大規模スーパーコンピュータでは計算ノードにローカルディスクを備えていないことが多い。このため、Hadoop をそのまま典型的なスーパーコンピュータ上で動かすと、ネットワーク上のファイルサーバのディスクを

ローカルディスクのように用いてしまい、著しく効率が悪くなる。これを避けるためには各ノードのメモリを用いて仮想ディスクを実現すればよいが、各ノードの資源をどのように仮想ディスクに振り向ければ良いかの研究をおこなった。研究のテストベッドに、各ノードに SSD をローカルディスクとして備える TSUBAME を利用することで、各ノードのローカルな物理ディスクを用いる Hadoop の元の実行方法を含め、様々な資源管理方式を比較実験した。

## 結果および考察

Bytespresso で実装したドメイン専用言語で書いた NAS CG と LU の性能を TSUBAME 3.0 で測定した。コンパイラとしては GCC 4.8.5 を使い、OpenMPI をライブラリとして用いた。各ノードで 8 MPI プロセスを実行する設定の下、ノード数を 4, 8, 16, 32, 64 と増やして実験したところ、CG では元の Fortran プログラムとほぼ同等の性能を得た。クラスは C である。一方、LU の場合は 10%から 20%の性能低下が見られた。LU についてはクラス C と D で実験をおこなった。これらの結果の詳細は今後出版予定である。

Hadoop の実装方式に関する実験は TSUBAME 2.5 上で実施し、結果は Thanh-Chung Dao and Shigeru Chiba, SEMem: Deployment of MPI-Based In-Memory Storage for Hadoop on Supercomputers, Euro-Par 2017, Springer, pp. 442-454, 2017 にて公表した。

## まとめ、今後の課題

本研究課題では、我々が研究開発した Bytespresso、ExaSlang、Hadoop のスーパーコンピュータ上での実行方式の実効性を確かめるため、TSUBAME を用いた性能測定をおこなった。おおむね望ましい性能を観察することができた。なお Bytespresso を用いて開発したドメイン専用言語で書き直した NAS LU ベンチマークの実験では、元の Fortran 版の LU ベンチマークのプログラムが TSUBAME 上で動作せず、segmentation fault で終了するという現象を観察した。これは実験当時 TSUBAME 3.0 が稼働開始直後で十分に安定していなかったためと思われる。



利用課題名 LRnLA アルゴリズムを用いた物理シミュレーション  
 英文 Simulation of Physical Processes with LRnLA Algorithms

善甫 康成  
 Yasunari Zempo

法政大学 情報科学部  
 Computer and Information Sciences  
<http://cis.k.hosei.ac.jp/>

従来使用してきた高性能 FDTD コード DTmaxwell を TSUBAME3.0 に移植した。このコードは LRnLA アルゴリズム DiamondTorre を使用しているが、大きな変更の必要はなく本来の性能を引出すことができた。性能をテストでは LRnLA 理論から予測される性能であることが検証できた。これは NVidia GPGPU を使ったことと LRnLA アルゴリズムの成果である。TSUBAME3.0 では完全 4D 時空間での並列化ができれば更に性能を向上できることが判明した。そこで試験的にこの並列化を行う LRnLA アルゴリズム DiamondCandy を開発した。FDTD のコードを使った検証の結果、新しい GPCPU の性能を効率的に向上させることを確認することができた。

The high-performance FDTD code DTmaxwell, which uses LRnLA algorithm DiamondTorre, has been ported to the new TSUBAME3.0 supercomputer, and its performance has been tested. The efficiency of the algorithmic method allowed to predict the performance and the prediction was verified by the test runs. This experience has shown both the advantages of NVidia GPGPU hardware and the advantages of the programming with LRnLA algorithms. The necessity of a full 4D space-time localization has emerged more distinctly with the new hardware. A new FDTD code benchmark was developed with a new LRnLA algorithm DiamondCandy, and it was verified to be more efficient on the new GPGPU.

*Keywords:* LRnLA algorithm, GPU, FDTD, high performance

## 背景と目的

近年、数値シミュレーションにおける時空間局所化が一般的な話題となっている。科学技術計算において必要とされている資源効率の高い、高度な大規模計算を実現できるからである。最先端のハードウェア技術により高性能計算が可能になるため、実施アルゴリズムの進歩が不可欠である。

時空間局所化は、幾つか報告があり、“temporal blocking”、“split tiling”、“diamond tiling”、“time skewing”、“wavefront blocking”という名前で見られている [1]。“polyhedral compilation”と“cache-oblivious algorithms”とも関連している。

これらの技術の考え方は、計算ドメインが時間層ごとに更新される伝統的な「ステップワイズ」アプローチを止めてしまうことである。ステップワイズでは、時間層  $t$  から  $(t + dt)$  への更新は、総てのシミュレーションドメインセルが  $(t - dt)$  から  $t$  に更新された後のみ開始されるためメモリアクセスの局所性と並列スケールが非効率的である。

新しい時空間局所化技術は、 $nD1T[(n + 1)$ 次元時空間]依存グラフの最適な分解をアルゴリズムの局所性および並列性を最適化するブロックを探索することが必要である。LRnLA 方法は、この問題を解決する画期的な方法の 1 つであり、理論的な基盤

に基づき、複数の計算効率向上を図りつつ解析できることを確認されている。

従来、我々は光学解析を行うため FDTD 法による解析のため DiamondTorre アルゴリズムを用いてきた。TSUBAME2.5 でテストを行い、性能が極めて優れていることがわかった [2]。

最近の GPGPU コンピューティング技術は急速に進歩している。新しい GPU アーキテクチャをはじめとし GPU 間の相互接続や CUDA の進歩している。これにより性能向上だけでなく重要な機能も提供されている。優れた互換性に加えて、新しいプラットフォームで古いコードを再コンパイルのみで、最高のパフォーマンスが得られるようになっている。

TSUBAME3.0 を利用することで LRnLA FDTD コード DTmaxwell のパフォーマンスがどのように改善するか、GPGPU スーパーコンピューティング動向に鑑み、どのようなアルゴリズム開発が必要かを検討することが非常に重要となっている。そこで我々の開発した手法を TSUBAME3.0 へ適用して、その性能を分析することとした。高度な内部並列性があればピーク性能が向上すれば、それに比例して性能は向上するはずである。

これまで DiamondTorre を用いた既存の FDTD

コード DTmaxwell を TSUBAME3.0 に移植すると予想される性能向上は 1GPU あたり 3 倍、1 ノードあたり 5 倍となるはずである。

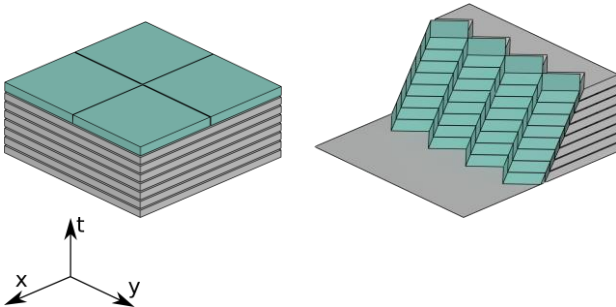


図 1. (2D1T) の stepwise と DiamondTorre のイラスト。緑色の個々のブロックは、同期（データ交換）イベントの前に 4 つの並列プロセッサによる可能な計算を示す。

### TSUBAME3.0 の重要な特徴

TSUBAME3.0 (P100 の 4.76TFlops) であり、TSUBAME2.5 (K20x の 1.3TFlops) より GPU パフォーマンスが高い。また同時に、より大きなメモリスループット (P100 では 732 GB/s、K20x では 250 GB/s) がある。DTmaxwell の FDTD 法はメモリバウンドの領域内に収まる。従ってパフォーマンスとメモリの両方が 3 倍になれば 1 つの GPU でコードのパフォーマンスが 3 倍向上することが期待できる。また、PCI-e に代わって導入された NVlink 相互接続によりノード内のコードの並列効率の向上が期待できる。またメモリサイズも増加しているので (P100 GPU あたり 16GB、K20x GPU あたり 6GB)、最大限のパフォーマンスを得るために 1 つの GPU あたりのアルゴリズムパラメータを以前より自由に選択できるという特徴がある。

一方で GPU 性能の発達に伴い DiamondTorre アルゴリズムの課題が顕在化した。DiamondTorre アルゴリズムでは、データが時間と 2 つの空間次元のみに局所化されているため、第 3 軸は CUDA スレッドを使用してステップワイズ法で処理されるところが大きな課題であった。

最近の GPU には、デバイスにもよるが数多くの SM (Streaming Multiprocessor) が搭載されている。我々のこれまでの研究によると、使用している CUDA ブロックの数が 1 つのデバイスの SM 数またはこの数の整数倍に等しい場合、DTmaxwell の性能が飽和することがわかっている。従ってドメインの規模は、並列分解が可能な程度にとる。

DiamondTorre GPU 実装では、y 軸の分解に CUDA ブロックの並列処理が使用されている。その結果、細長いドメインで最高のパフォーマンスが達成されるような設計である。

最新の GPU でのデバイスあたりの SM の数の増加に伴い、最適なパフォーマンスを得るためには、方向によって異なるアスペクト比が必要になる。高度に局所化された LRnLA アルゴリズムを使うと、オーバーヘッドが増加する。アルゴリズムの中に多彩な機能を追加した新しい実装方法を使うと、局所

性を上げてオーバーヘッドの影響を少なくすることができるはずである。

そこで最近提案された新しいタイプの LRnLA アルゴリズム DiamondCandy を用いることとした [4]。これは 3D1T の局所性を持っているが、CUDA ブロック内のスレッド間の計算は一様ではない。

従って細長いドメインのパフォーマンスは DiamondTorre よりも低い、大規模なキューブ型ドメインでの計算処理では高い性能が出ると期待される。

### パフォーマンスの検討結果

DTmaxwell コードは TSUBAME3.0 で再コンパイルした上テストを行った。これまでに達成された最高の性能は、TSUBAME2.5 の記録を、GPU デバイス 1 個あたり 3.5 倍、1 つのノードあたり 5 倍上回っていた。もちろん非同期ブロックの数が SM の数に達するとパフォーマンスは飽和する。

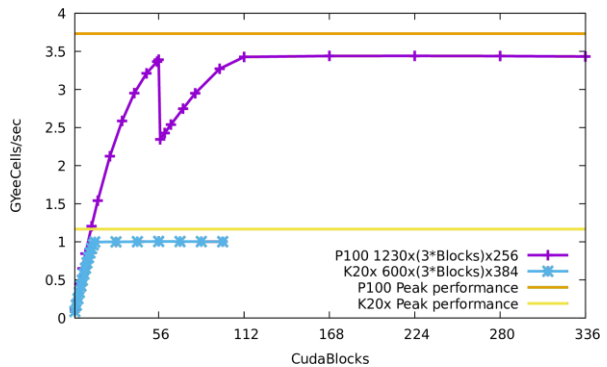


図 2. 使用 CUDA ブロックの関数として TSUBAME2.5 と TSUBAME3.0 上の一の GPU デバイス上で達成される性能

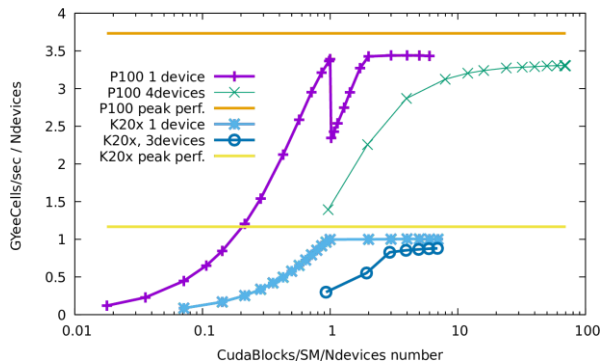


図 3. 使用 CUDA ブロックの関数として TSUBAME2.5 と TSUBAME3.0 上の一のノード上で達成される性能

SM の数は P100 の方が多いので、以前の K20x に比べパフォーマンスの向上を連続して図ることができる高い。ブロック数が SM の数の倍数でない場合、パフォーマンスが低下するので、図 2 と図 3 で横軸が 1 のところで最も高くなり一旦パフォーマンスが低下し再度向上していることがわかる。

また weak scaling についても検証の結果、非常に効果的であることがわかった。ドメインが x 軸または y 軸方向に増加すると、並列効率が異なっ

ている。y方向のノード並列処理は、各ステージの後、すなわち非同期 DiamondTorre 行が実行された後、ノード間の同期が必要である。

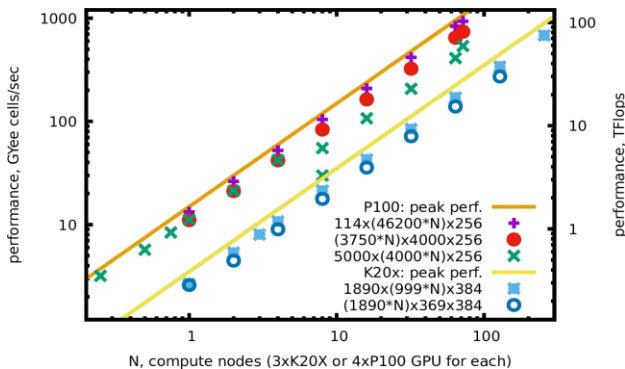


図4. 異なるドメインサイズの設定のための weak scaling 性能

4 つ以上のノードが関与すると、パフォーマンス結果が不安定になるため、ここでは最も成功した実行の結果のみ表示している。この原因は、キュー上の他のジョブの干渉である可能性がある。これは使用可能なすべてのノードが関与すると、パフォーマンスが向上するということから推測できる。

さらに x方向のノード間並列処理のために”moving window”手法が実装されている。使用されているすべてのノードを同時に同期する必要がないためパフォーマンスの低下は見られない。

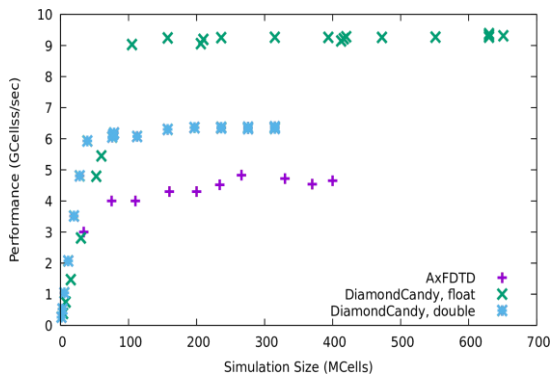


図5. 問題データサイズの間数としての DiamondCandy アルゴリズムの性能テスト結果

また新しく開発した DiamondCandy アルゴリズムでの初期ベンチマーク結果も以下に報告する。図5に問題のデータサイズに対するパフォーマンスの依存性を、図6に P100 の roofline 図と3つのコード(DTmaxwell、DiamondCandy、効率的な最先端の市販ソフト AxFDTD [3])の性能テストの結果を示す。新しいアプローチの利点を示す結果となっている。

## まとめ

TSUBAME2.5 と TSUBAME3.0 の2のマルチGPUスーパーコンピュータでDTmaxwellコードを実行した課題は、GPUハードウェアとCUDAによ

ってできたことである。この移植では、コードの再コンパイル以外の作業は必要なく、パフォーマンスはハードウェア機能の増加に比例して増加した。LRnLA 理論からの性能予測が可能であることが確かめられた。実際1つのGPUで3倍のパフォーマンス向上を得て、1つのノードのパフォーマンスに対する最大5倍の性能向上が得られた。

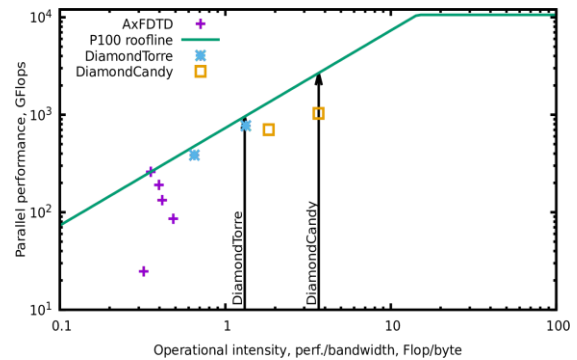


図6. DiamondTorre アルゴリズムと DiamondCandy アルゴリズムのルーライン制限と最大パフォーマンス結果。市販ソフト AxFDTD コードによる結果と比較している。DiamondCandy と DiamondTorre の2つのマーカーは、単精度と倍精度の結果を示す。

また TSUBAME3.0 の72ノードで1秒あたり  $10^{12}$  Yeeセルの更新が達成されたが、同じ結果を得るために TSUBAME2.5 の512ノードが必要であった。ただドメインアスペクトによる効果が最大となるように調整している。また新しいマシンでは並列性を上げることができるので、そのための新しいアルゴリズム DiamondCandy を導入した。3次元の局所性があり、より多くの非同期性が可能になるという特徴がある。今回、この最初のベンチマークを行いその効果を確認した。

今後の課題として新しいアルゴリズムの実装、テスト、実際問題の応用について検討する予定である。

- [1] T. Fukaya and T. Iwashita. Time-space tiling with tile-level parallelism for the 3D FDTD method. In Proceedings of the International Conference on High Performance Computing in Asia-Pacific Region, p. 116–126. ACM, 2018.
- [2] A. Zakirov, V. Levchenko, A. Perepelkina, Y. Zempo, “High performance FDTD algorithm for GPGPU supercomputers”, *J. Phys.: Conf. Ser.* **759** 012100, 2016
- [3] AxFDTD GPU Performance (v11.1.12). <http://www.acceleware.com/fdtd-solvers>, retrieved Mar 2018.
- [4] V. D. Levchenko and A. Y. Perepelkina. Locally recursive non-locally asynchronous algorithms for stencil computation. *Lobachevskii Journal of Mathematics*, **39**(4), 2018.



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名: 凝集誘起発光についての理論的研究  
 英文: Theoretical Study on the Aggregation Induced Emission

山本典史  
 Norifumi Yamamoto

千葉工業大学工学部応用化学科  
 Department of Applied Chemistry, Faculty of Engineering, Chiba Institute of Technology  
 www.appchem.it-chiba.ac.jp/yamamoto/

## 邦文抄録(300 字程度)

近年, 単体では発光しないが, 分子が多数凝集すると発光する特性を持つ蛍光色素が注目されている。この現象は凝集誘起発光と呼ばれている。シアノスチルベン誘導体の CN-MBE は凝集誘起発光を示す蛍光色素であり, 溶液中に分散した状態での蛍光量子収率( $\Phi$ )値は 0.001 であるが, 凝集して微粒子や結晶になると  $\Phi$  値は 0.69 まで増大する。この CN-MBE には E 体と Z 体の 2 種類の異性体があり, この 2 つの異性体のうち凝集誘起発光が観測されるのは E 体のみであり, Z 体は結晶構造でも発光が観測されない。このように CN-MBE は凝集状態における光物性がそれぞれの異性体で大きく異なるが, その分子機構は明らかではなかった。本研究では CN-MBE が示す凝集誘起発光のメカニズムについて, 分子シミュレーションを用いた解析をおこなった。

## 英文抄録(100 words 程度)

The aggregation-induced emission (AIE) of a cyanostilbene derivative, called CN-MBE, was studied theoretically by using electronic structure calculations, together with molecular dynamics (MD) simulations. The *E*-form of CN-MBE has been known to exhibit the AIE, which is non-emissive in dilute solutions but becomes highly emissive in solid or aggregated states; however, its *Z*-isoform is non-emissive even in a crystal. In this study, we have investigated the conformation-dependent AIE mechanism of CN-MBE. The method combining the quantum chemistry calculation and the MD simulation used in this study can not only be utilized for analysis of various other various AIE phenomena but also in the development of new photochemical devices making use of the specificity of molecular crystals.

## Keywords: 5つ程度

凝集誘起発光, 円錐交差, 分子結晶, 電子状態計算, 励起状態ダイナミクス

## 背景と目的

近年, 単量体では発光しないが, 分子が多数凝集すると発光する特性を持つ蛍光色素が注目されている。この現象は凝集誘起発光 (Aggregation Induced Emission: AIE) と呼ばれている。香港科技大学の Tang らは凝集誘起発光特性を持つ分子として tetraphenylethene を見だし, この分子を基本骨格とする AIE 色素を数多く開発している。ソウル国立大学の Park らも CN-MBE と呼ばれるシアノスチルベン誘導体が凝集誘起発光を示すことを発見し, この分子を基本骨格とした機能性蛍光色素を開発している。

シアノスチルベン誘導体の CN-MBE は凝集誘起発光を示す蛍光色素であり, 溶液中に分散した状態での蛍光量子収率( $\Phi$ )値は 0.001 であるが, 凝集して微粒

子や結晶になると  $\Phi$  値は 0.69 まで増大する。この CN-MBE には E 体と Z 体の 2 種類の異性体があり, この 2 つの異性体のうち凝集誘起発光が観測されるのは E 体のみであり, Z 体は結晶構造でも発光が観測されない。このように CN-MBE は凝集状態における光物性がそれぞれの異性体で大きく異なるが, その分子機構は明らかではなかった。そこで本研究では, CN-MBE の凝集誘起発光について分子シミュレーションを用いた理論的解析に取り組み, そのメカニズムを明らかにした。

## 概要

本研究では CN-MBE が示す凝集誘起発光のメカニズムについて, 分子シミュレーションを用いた解析をお

こなつた。その結果, CN-MBE は基底電子状態と励起状態のポテンシャルエネルギーが分子内 C=C 結合軸周りの回転に伴って近接し, わじれ角度  $\theta$  が 90 度付近となる地点で交差 (円錐交差) することが明らかになった。希薄溶液ではこの C=C 結合軸の回転がほぼバリアレスで起こり, 円錐交差を経由した光励起エネルギーの無輻射失活が起こる。しがしながら, 分子同士が密に集合した凝集状態では, E 体の場合, 凝集体中では光励起後, 分子構造の変化 (C=C 結合部分の回転) を経て円錐交差に至る経路がエネルギー的に著しく不利になっていることが明らかになった。このことから, E 体の場合, 凝集構造では無輻射に緩和する経路が著しく抑制されているために強い発光を示すと考えられる。一方で Z 体の凝集体が光励起されたときには, 垂直励起したポテンシャルエネルギー曲面上の地点からバリアレスで円錐交差へとすみやかに緩和することが明らかになった。したがって, Z 体が凝集誘起発光を示さない原因は, 凝集体中であっても比較的容易に C=C 結合部位の回転運動が起こることですみやかにポテンシャル曲面上の円錐交差に到達することが可能であり, 無輻射で緩和するためであることが明らかになった。本研究の成果は, 現在, アメリカ化学会が発行する物理化学専門誌 *Journal of Physical Chemistry C* に投稿中である。

#### 結果および考察

CN-MBE の電子状態について, 時間依存密度汎関数 (TD-DFT) 法を用いて解析した。汎関数には wB97-XD, 基底関数には 6-31+G(d,p) を用いた。S<sub>0</sub>/S<sub>1</sub> 交差点付近の電子励起状態については 8 軌道 8 電子を完全活性空間とする多配置自己無撞着場法 CASPT2/6-31G(d)//CASSCF/6-31G(d) を用いて解析した。電子状態の解析にはプログラムとして Gaussian16 と Molpro を使用した。

CN-MBE 分子の ethylenic C=C 二重結合部位の回転角 ( $\phi$ ) を変化させながら, 量子化学計算による構造最適化を実行した。得られた結果を図 2 に示す。図中では, グラフの横軸に回転角 ( $\phi$ ), 縦軸にポテンシャルエネルギーを示し, 下部は基底 (S<sub>0</sub>) 状態, 上部は励起 (S<sub>1</sub>) 状態のエネルギー変化を示す。図 2 に示すように,

回転角が  $\phi = 90$  度となる点で基底状態および励起状態のポテンシャル曲面が交差することが明らかになった。さらに, Z 型 ( $\phi = 0$  度) および E 型 ( $\phi = 180$  度) を垂直励起した点から, この S<sub>0</sub>/S<sub>1</sub> 交差点まで, ほとんどバリアレスであることも明らかになった。これらの二つの事実から, この分子を光励起した後のふるまいとして, 垂直励起した点から  $\phi = 90$  度となる立体構造までの自発的な変化が起こり, 辿り着いた S<sub>0</sub>/S<sub>1</sub> 交差点を経由してすみやかに基底状態に戻ると予想される。つまり図 2 のエネルギープロファイルからは, 孤立した CN-MBE の場合, S<sub>0</sub>/S<sub>1</sub> 交差点を経由して無輻射緩和するために蛍光を示さないことが明らかになった。

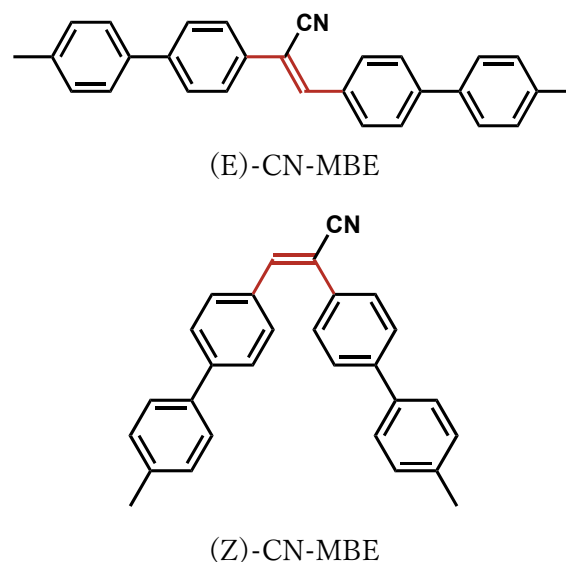


図 1 凝集誘起発光色素 CN-MBE

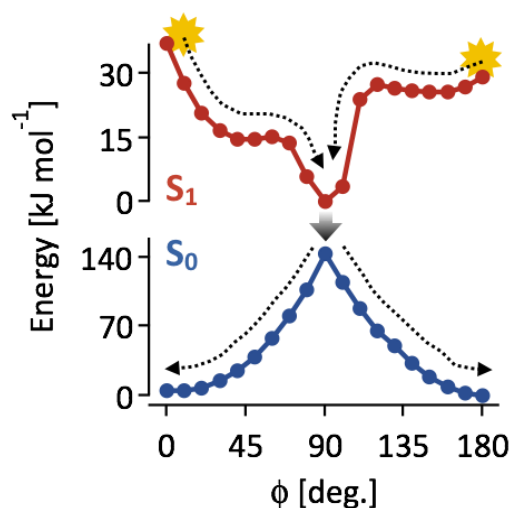


図 2 CN-MBE のポテンシャルエネルギー変化



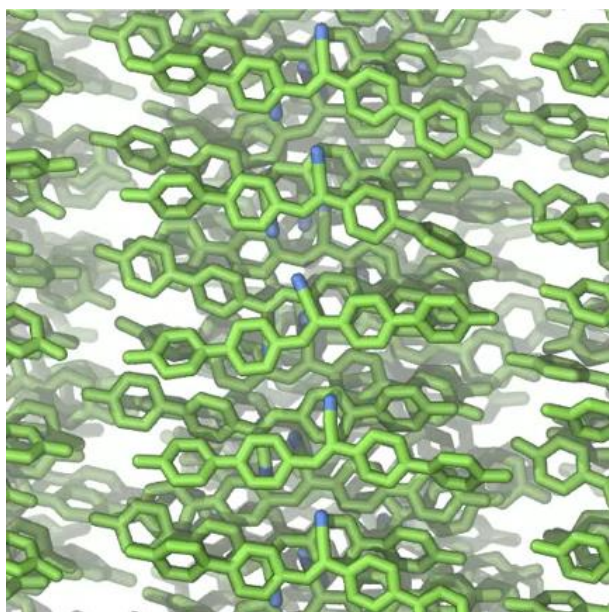


図3 E型 CN-MBE の凝集構造

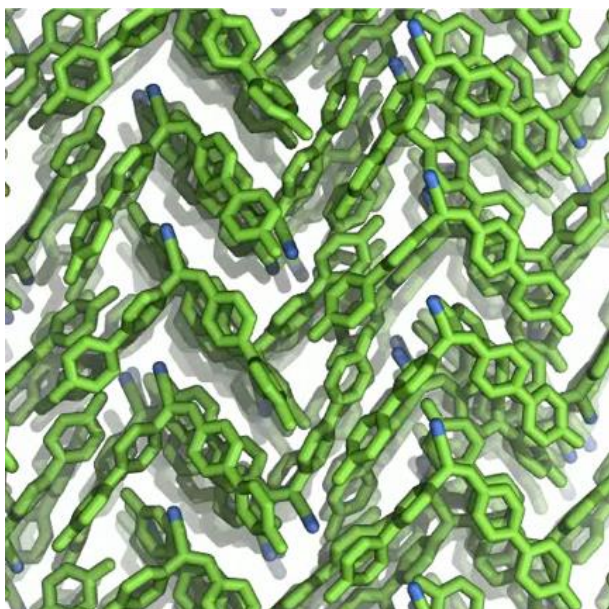


図4 Z型 CN-MBE の凝集構造

図3に示すのはE型 CN-MBE の凝集構造であり、図4に示すのはZ型 CN-MBE の凝集構造である。それぞれの凝集構造で分子動力学シミュレーションを実行し、拡散定数の値( $D$ )を解析した。その結果、E型の凝集体の場合には  $D = 6.07 \times 10^{-6} \text{ nm}^2 \text{ ps}^{-1}$  であり、密に凝集している、分子の運動が抑制された、堅い構造となっていることが明らかになった。一方でZ型の凝集体の場合、 $D = 8.06 \times 10^{-6} \text{ nm}^2 \text{ ps}^{-1}$  であり、比較的疎に凝集している、柔らかい構造となっていることが明らかになった。

図3および4で得られた凝集体でメタダイナミクス分子動力学シミュレーションを実行し、その結果から、CN-MBEのethylenic C=C二重結合部位が回転するときの自由エネルギー変化を解析した。E型 CN-MBEの凝集状態における自由エネルギープロファイルを図5に示す。この場合に得られたエネルギープロファイルは、図2で示す孤立状態の結果と大きく異なる特徴を持つことが明らかになった。図5からCN-MBEの凝集構造中におけるふるまいを読み取ると、E型のCN-MBEを垂直励起した場合には $\phi = 180$ 度となる直線状の構造を有しているが、凝集体中でこのような直線構造の分子がねじれようとすると周囲の分子が障壁となるので、これらを押しつけてethylenic C=C二重結合部分が回転するには大きなエネルギーが必要となる。このように、E型 CN-MBEの場合、凝集体中では垂直励起後、分子構造の変化(ethylenic C=C二重結合部分の回転)を経て $S_0/S_1$ 交差点に至る経路がエネルギー的に著しく不利になっていることが明らかになった。以上のことから、E型 CN-MBEの場合、凝集構造では無輻射に緩和する経路が著しく抑制されているために強い発光を示すと考えられる。

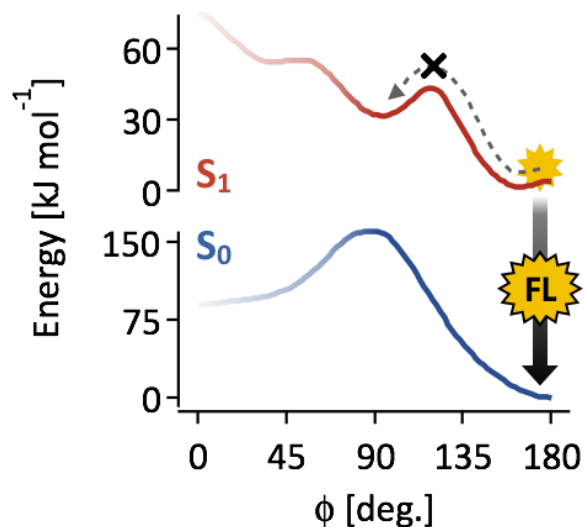


図5 (E)-CN-MBE の凝集状態における C=C 結合軸回転に対する自由エネルギープロファイル

Z型 CN-MBE の凝集構造中におけるエネルギープロファイルを図6に示す。得られた自由エネルギー変化を解析したところ、図5に示すE型 CN-MBEの結果と

は大きく異なることが明らかになった。図6をみると、Z型 CN-MBE の凝集体が光励起されたときには、垂直励起した  $S_1$  ポテンシャルエネルギー曲面上の地点からバリアレスで  $S_0/S_1$  交差点へとすみやかに緩和することを示している。したがって、Z型 CN-MBE が凝集誘起発光を示さない原因は、凝集体中であってもZ型は比較的容易に ethylenic C=C 二重結合部位の回転運動が起こることですみやかにポテンシャルの  $S_0/S_1$  交差点に到達することが可能であり、この交差点を経て無輻射で基底状態に構造緩和するためであることが明らかになった。また、基底状態に至った分子は2つの経路に分岐し、Z型に戻るものとE型に変化するものが現れるために室温中でも容易に異性化反応が起こることも明らかになった。

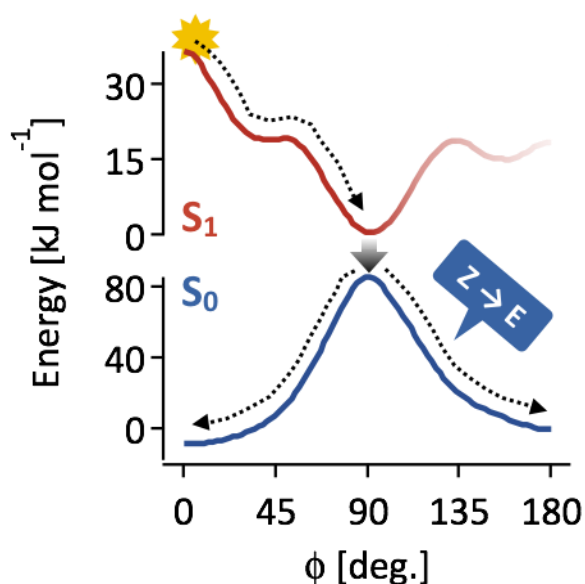


図6 (Z)-CN-MBE の凝集状態における自由エネルギープロフィール

#### まとめ、今後の課題

本研究では、CN-MBE が示す凝集誘起発光現象のメカニズムの詳細を明らかにするために、量子化学と分子動力学シミュレーションを組み合わせた理論的な解析をおこなった。この結果、これまで不明であったCN-MBE の凝集誘起発光のメカニズムを明らかにし、この現象を利用した分子設計をおこなうための基礎を確立した。

シアノスチルベン誘導体の他にも、凝集誘起発光現

象を示す蛍光物質(AIEgen: AIE luminogen)が報告されている。これらの AIEgen について、基本的なメカニズムは今回解析したシアノスチルベン誘導体と同様だと考えられるが、その詳細は不明である。そこで今後の課題として、本研究課題で確立した手法を駆使して、これまでに発見されている AIEgen の発光過程を調べることで、AIE の基本的なメカニズムの解明に取り組む予定である。



## TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 知識に基づく構造的言語処理の確立と知識インフラの構築  
 英文: Establishment of Knowledge-Intensive Structural Natural Language Processing and  
 Construction of Knowledge Infrastructure

利用課題責任者 黒橋 禎夫  
 First name Surname Sadao Kurohashi

所属 京都大学 大学院情報学研究科  
 Affiliation Graduate School of Informatics, Kyoto University  
 URL <http://nlp.ist.i.kyoto-u.ac.jp/>

## 邦文抄録(300 字程度)

本利用課題では、知識に基づく頑健で高精度な構造的言語処理を実現し、これによって様々なテキストの横断的な関連付け、検索、比較を可能とする知識インフラを構築することを目的とする。そのために、大規模 Web コーパスに対して言語解析を適用し、その解析結果から膨大な知識を獲得する。この知識獲得の処理には膨大な計算量が必要になるが、TSUBAME の大規模並列計算資源を利用することによって、極めて短期間で達成することができた。

## 英文抄録(100 words 程度)

In this project, we aim to establish robust and accurate knowledge-intensive structural NLP, and to construct a knowledge infrastructure which can relate, retrieve and compare knowledge from various texts. We perform linguistic analysis to a large-scale Web corpus and use the resulting analyses to acquire wide-coverage knowledge. We accomplished several knowledge acquisition processes quite rapidly using TSUBAME.

## Keywords: 5つ程度

*natural Language processing, Web, predicate-argument structure analysis, parsing, similarity learning*

## 背景と目的

テキストは、専門家によるデータの分析結果や解釈、ステークホルダーの批判・意見、種々の手続きやノウハウなどが表出されたものであり、人間の知識表現の根幹である。テキストとして表現された知識を計算機によって抽出・関連付けすることができれば、社会における知識循環を円滑化し、異なる分野間での知識の相互関連性の発見や、新しい知識・法則の発見を支援することが可能となる。言語情報処理はウェブをはじめとする大規模テキストの活用によって長足の進歩を遂げつつあるが、本利用課題ではこれをさらに発展させ、知識に基づく頑健で高精度な構造的言語処理を実現し、これによって様々なテキストの横断的な関連付け、検索、比較を可能とする知識インフラを構築する。

本利用課題では、TSUBAME の大規模計算機環境を用いて、上記の目的を実現するために必要となる知識を大規模テキストコーパスから自動獲得する。具体的には述語項構造解析・共参照解析の同時学習、論理式と

定理証明を用いたフレーズ公理生成、関係知識獲得ならびに英語コーパスの構文解析を行った。

Entity-Centric な述語項構造解析・共参照解析の同時学習

## 背景

述語項構造解析とはテキスト中の「誰が何をどうした」を明らかにする解析である。述語項構造解析は情報抽出や質問応答など様々なアプリケーションにとって重要な基礎解析である。

述語項構造解析は格解析とゼロ照応解析からなる。格解析は述語と項が係り受け関係にある場合に格関係を明らかにする解析であり、ゼロ照応解析は述語と項が係り受け関係にない場合に項を同定する解析である。ゼロ照応解析はさらに文内ゼロ照応解析(述語と項が同一文にある)と文間ゼロ照応解析(述語の文より前の文に項がある)に大別することができる。述語と項の依存構造上のパスなどの手がかりが使えない、項の候補がたくさん

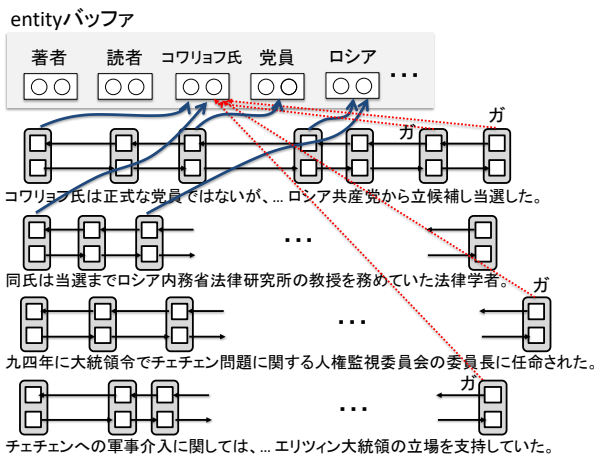


図 1 述語項構造解析と共参照解析の同時学習

あるなどの理由から文間ゼロ照応解析は非常に難しいタスクである。文間ゼロ照応解析の精度を向上させるためには文章中で何が話題の中心であるかなどを捉える必要がある。

本研究では話題の中心を捉えるために entity という概念を導入する[柴田+ 18]。entity を考えるためには共参照解析を行う必要がある。Wiseman らは RNN (Recurrent Neural Network) で entity の embedding を計算しており [Wiseman+ 2016]、本研究でもこの手法にならない entity の embedding を計算し、話題の中心を捉える。

#### 方法

提案手法の概要を図 1 に示す。各 entity には embedding を割り当て、図の上部に示すバツファで管理する。入力テキストが与えられると、まず Bi-directional LSTM を用いて文脈を考慮した句の表現を得る。次に、文章の先頭から対象の句が名詞句の場合は共参照解析を、動詞句の場合は述語項構造解析を行う。共参照解析・述語項構造解析ともにニューラルネットワークを用いてスコアを計算する。その際に mention の embedding (Bi-directional LSTM で得られるもの) だけでなく entity の embedding を考慮する。そして、両解析の解析結果を用いて RNN で embedding を更新し、情報を蓄積する。

#### 結果

評価は京都大学ウェブ文書リードコーパス(ウェブ、約 5,000 文書)と京都大学テキストコーパス(新聞、約 550 文書)の 2 種類のコーパスで行った。entity を考慮しないベースライン手法と提案手法を比較した。

格解析と共参照解析の精度はほぼ変化ないが、ゼロ照応解析の精度はベースラインに対してウェブで 3.8 ポイント(0.523 → 0.561)、新聞で 7.5 ポイント(0.280 → 0.355)の精度向上がみられた。特に文間ゼロ照応解析の精度がウェブで 27.2 ポイント(0.046 → 0.318)、新聞で 15.5 ポイント(0.028 → 0.183)と大幅な精度向上がみられた。

訓練とテストは 1GPU で半日程度の時間を要する。様々な条件での実行が必要となるが、TSUBAME を用いることで単時間での実行が可能となった。

#### まとめ、今後の課題

提案手法は文間ゼロ照応解析の精度を大きく向上することができた。今後の課題として橋渡し参照などを行うことや事態間関係知識などの外部知識を取り入れることなどがあげられる。

### 論理式と定理証明を用いたフレーズ公理生成

#### 背景

含意関係認識 (Recognizing Textual Entailment, RTE) は、テキスト T が仮説文 H を含意するか否かを自動判定する自然言語処理のタスクであり、意味処理のもっとも重要な基盤技術の一つである。文の意味を論理式で表現し、論理推論によって高度な意味解析を行う手法は、論理式による意味表現と整合性の高い組合せ範疇文 (CCG) による頑健な統語解析の発展に伴い、RTE や文間類似度学習のタスク [Yanaka+ 2017] において高精度を達成している。

論理推論によるアプローチでは、否定表現や数量表現など機能語の意味解析を正確かつ効率的に扱うことができるが、内容語の語彙知識が必要な推論については課題が残されている。特に、論理ベースの手法において、WordNet など既存の大規模データベースの語彙知識を利用することで、単語間知識の扱いは一定の有効性が認められているが、フレーズ間知識の扱いは十分に検討されてこなかった。

#### 方法

そこで本研究では、論理ベースの深い意味解析の手法に基づいてフレーズ間の語彙知識を獲得し、論理推論と組み合わせて用いる手法について検討した。フレーズ間知識を論理推論で扱うためには、文中のフレーズ

ズの適切な対応付け(フレーズアライメント)を特定する必要がある。先行研究では文の表層情報を利用してフレーズアライメントを行う手法が研究されているが, *cut ... into pieces* など, 非連続なフレーズを扱うことが難しいという問題がある。文の意味を論理式で表現することで, こうした非連続なフレーズを含む広範囲の推論・言い換えが扱えるようになると期待される。

より具体的には, 本研究では, 自然演繹に基づく文間の含意関係の証明の実行過程からフレーズアライメントを特定し, 論理推論に適用する手法を提案した。提案手法を用いてフレーズ間知識の公理を生成し, RTEの問題に適用することで, 単語間知識だけでは推論できなかったパラフレーズを含む文間の含意関係が推論可能となるか, RTE 評価用データセットによる評価を行った。

## 結果

提案手法の評価には RTE 評価用データセットである SemEval2014 SICK を用いた。SICK は 2 つの文の含意関係(yes, no, unknown)が人手で付与されている。訓練データは 5000 件, テストデータは 4927 件である。CCG 解析によって文を高階論理式に変換し, 自動推論を行うシステム *ccg2lambda* に提案手法であるフレーズアライメント機能, フレーズ間の公理生成機能を拡張し, 実験を行った。まず, 訓練データ中の RTE 正解ラベルが yes か no の全文例について, 提案手法を用いて含意関係の証明を行った上で, フレーズアライメントを特定し, フレーズ間知識の公理を収集した。次に, テストデータ中の全文例で含意関係の証明を行い, 評価を行った。

実験の結果, フレーズ間・単語間知識の公理生成を組み合わせた場合の正答率は 84.3%であり, 公理生成なし(76.6%), 単語間公理生成のみ(83.1%)の場合と比べて向上が見られた。また, SemEval2014 のトップスコア(83.6%)を上回る結果となった。フレーズ間知識の公理収集時・評価時の証明時間はいずれも 1 文ペアあたり平均 3.0 秒だった。

TSUBAME の分散並列処理の利用により, CCG による統語解析, *ccg2lambda* による意味解析と自動証明の各処理に関して, きわめて短期間に成果を得ることができたと言える。今後, 訓練データから収集したフレーズを用いて公理を生成するか否かを判定する分類器を作成

することで, さらなる精度向上が見込まれる。

## オートエンコーダとの同時学習による関係知識共有概要

WordNet, YAGO, Freebase などのような大規模知識ベースは, 事実を(ヘッドエンティティ, 関係, テールエンティティ)の三つ組で表現し, 意味解析, 情報抽出, 質問応答などに広く応用されてきた。しかし事実を完全網羅するのが難しいため, 欠損されたエンティティを予測する知識ベース補完が盛んに研究され, これらの研究では, エンティティと関係を含む知識ベースの要素を連続ベクトル空間に埋め込むことで, 格納された事実(三つ組)の背後に存在する統計的な規則性をモデル化し, 記述漏れした事実を導き出す。

この連続空間埋め込みを行う際の研究課題の一つに, 関係間の知識共有がある。直感的に, 例えば「(人物が)(作品を)監督した」と「(人物が)(作品を)製作した」という関係はどちらも「人物」と「作品」を結びつく概念であるように, 多くの関係はいくつかの概念を共有するので, このような直感を知識ベースのモデル化に取り入れることが望ましい。本研究は, 知識ベースの埋め込みを訓練すると同時に, 関係に対するオートエンコーダを合わせて学習することで関係間の知識共有を促す柔軟な枠組みを提案し, 評価実験を行なった。この手法は, 関係のモデル化自体に何の制約も与えずに, 知識共有の強さ度合いをオートエンコーダのコード長で柔軟に調整できる利点を持つ。

## 方法

具体的には, これまでに開発した VecDCS モデル [Tian+ 2016]をベースにした学習手法を知識ベースの埋め込みに適用し, その学習と同時に関係を表す行列を低次元コードから復号化するオートエンコーダを加えた。

## 結果および考察

知識ベース補間の標準的な評価データである WN18, FB15k, WN18RR, FB15k-237, YAGO3-10 などを用いてモデルを訓練し, 評価を行った。訓練は TSUBAME の *f\_node* を利用し, CPU 28 並列で行なった。その結果, VecDCS モデルの学習手法を利用したベースモデルは, 既に多くのデータセットで state-of-the-art を達

成しており、今回提案のオートエンコーダを加えると更なる精度向上が見られた。これら結果の報告は[高橋+ 2018]で行なった。

#### まとめ、今後の課題

今後は、パラメータ・チューニングによる更なる精度向上と、知識ベース補間と含意関係認識の統合を目指すよう研究を進めていきたい。

### 英語コーパスの構文解析

#### 概要

本研究は、単語の分散表現に感情極性や文書トピックなど、単語生成に関わるいろんな因子を取り込むことを目指す。そのために用いたデータセットは Multi-domain Sentiment Dataset で、25 カテゴリの商品に対する 140 万レビューを集めたコーパスである。このコーパスに対して、SyntaxNet を用いた品詞タグ推定、及び依存構造解析を行なった。

#### 結果および考察

構文解析は TSUBAME の q\_node を使い、GPU を用いて行なった。24時間以内に全コーパスの解析が終わった。この解析結果を用いて訓練した単語分散表現のモデルから、各因子だけに関わる部分の特徴が現れた。これらの結果は[田+ 2018]で報告した。

#### まとめ、今後の課題

今後は、訓練した単語表現を他分野への転移学習や含意関係認識に応用する予定である。

#### 参考文献

[柴田+ 2018] 柴田知秀, 黒橋禎夫. Entity-Centric な述語項構造解析・共参照解析の同時学習. 言語処理学会第 24 回年次大会, 2018.

[Wiseman+ 2016] Wiseman, Sam, Rush, Alexander M. and Shieber, Stuart M. Learning Global Features for Coreference Resolution. Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016.

[Yanaka+ 2017] Hitomi Yanaka, Koji Mineshima, Pascual Martínez-Gómez and Daisuke Bekki. Determining Semantic Textual Similarity using

Natural Deduction Proofs. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP2017), Copenhagen, Denmark, September 7-11, 2017.

[谷中+ 2018] 谷中瞳, 峯島宏次, Pascual Martínez-Gomez, 戸次大介. 自然演繹に基づく文間の含意関係の証明を用いたフレーズアライメントの試み. 言語処理学会第 24 回年次大会, 2018.

[馬目+ 2018] 馬目華奈, 谷中瞳, 吉川将司, 峯島宏次, 戸次大介. RNN 系列変換モデルを用いた高階論理式からの文生成, 言語処理学会第 24 回年次大会, 2018.

[Tian+ 2016] Ran Tian, Naoaki Okazaki, Kentaro Inui: Learning Semantically and Additively Compositional Distributional Representations; In Proceedings of ACL, 2016.

[高橋+ 2018] 高橋諒, 田然, 乾健太郎, オートエンコーダとの同時学習による知識共有, 言語処理学会第 24 回年次大会, 2018.

[田+ 2018] 田然, 渡邊研斗, 乾健太郎, 多因子単語埋め込みを用いる複数分野感情極性の転移学習, 言語処理学会第 24 回年次大会, 2018.

TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 解析的壁関数を用いた格子ボルツマン法による高レイノルズ数乱流解析コードの開発  
 英文: Development of Lattice Boltzmann Method with Analytical Wall Function for Higher Reynolds  
 number Turbulent Flow Simulations

利用課題責任者: 須賀 一彦  
 First name Surname: Suga Kazuhiko

所属: 大阪府立大学  
 Affiliation: Osaka Prefecture University  
 URL: <http://www2.me.osakafu-u.ac.jp/htlab/>

#### 邦文抄録

格子ボルツマン法(LBM)を用いて航空機周りに見られる高レイノルズ数乱流の解析を行うために、LBM に壁関数法を組み合わせた解析を検討した。壁関数法で得られた壁面せん断応力と実際の壁面せん断応力が合致するように、壁面に滑り速度を与えて解析を行った。手法評価のために、D3Q27 多緩和時間格子ボルツマン法を用いて摩擦レイノルズ数2,000のラージ・エディ・シミュレーションを行い、壁関数法には対数測を用いる標準壁関数法と、剥離・再付着流れにも適用可能な解析的壁関数法を用いた。標準壁関数法で得られた結果は対数速度分布と良く合致することは確認されたが、壁面近傍の過大な圧力・速度変動により解析的壁関数法では正しく壁面せん断応力が見積もることができなかった。

#### 英文抄録

To simulate high Reynolds number turbulent airfoil flows by the lattice Boltzmann method, a wall-function method is introduced to the LBM. To impose the wall-shear stress determined by the wall-function, a slip velocity is imposed to the wall. For the validation, large eddy simulation of turbulent channel flow at friction Reynolds number of 2,000 is carried out by the D3Q27 multiple-relaxation-time lattice Boltzmann method. For the wall function, we test the standard wall-function and the analytical wall-function. It is found that the standard wall-function successfully works while the analytical wall function does not provide precise wall-shear stress due to the overpredicted pressure and velocity fluctuations at the wall adjacent cells.

*Keywords: Lattice Boltzmann method, turbulence, standard wall-function, analytical wall-function large-eddy-simulation*

#### 背景と目的

近年の格子ボルツマン法(LBM)の目覚ましい発展により、高いレイノルズ数の乱流解析においても高精度で安定に解析することが可能になった。しかし、等間隔格子の使用を前提とする格子ボルツマン法では、壁面近傍の乱流を詳細に解析するためには膨大な解析格子が要求されるために、壁関数法などの導入が望まれる。壁関数法はこれまで、レイノルズ平均モデルを用いた差分法に用いられていたが、近年では計算機能力の向上とともに、より高精度なラージ・エディ・シミュレーションにも組み合わせも検討されている。また、乱流境界層での対数壁法則を前提する標準壁関数は、衝突・剥離流れなどの工学的な流れ場で予測精度が大きく下がるために、解析的壁関数法[1]をはじめとして多くの汎用的な壁関数法が提案されてきた。

しかし、格子ボルツマン法に壁関数モデルを組み合わせた解析例は少なく、解析的壁関数法などの高度な壁モデルの導入には未だ多くの検討が必要である。そこで本課題では、標準壁関数法に加えて解析的壁関数法などの高度な壁モデルを格子ボルツマン法に導入する方法に関して検討を行った。

#### 概要

計算密度が高く並列計算に適したLBMは、その解析精度や妥当性向上の為に多くの研究が行われ、目覚ましい発展を遂げてきた。特に、乱流解析においては、高い精度で安定に解析が可能である衝突演算モデルや、乱流モデルを用いた解析手法、高い精度を持つ局所細密格子法などが提案されてきた。しかし、LBM は等間隔格子を用いるために、船舶や航空機周りなどの高い

レイノルズ数を対象とする乱流解析においては、壁面近傍の格子解像度の不足は原理上避けることができない。そこで、本課題では、これまで有限差分・体積法などの乱流解析で広く用いられてきた壁関数法を格子ボルツマン法に組み合わせることで、LBM による高いレイノルズ数乱流解析を可能とすることを目的とする。さらに本研究では、壁面近傍の流れに対数則を仮定する標準的な壁関数法に加えて、解析的に積分した運動量式を用いることで剥離・再循環・衝突流れを含む複雑な流れ場を汎用的に扱える解析的壁関数法を用いる。

有限差分法・体積法においては、壁関数法で求められる壁面せん断応力は、そのまま差分式に代入することができる。しかし、密度分布関数の時間発展を解く格子ボルツマン法は、壁関数法で得られた壁面せん断応力を直接導入することができない。そこで、図1に示すように、壁面隣接点の速度  $u_{p1}, u_{p2}$  とせん断応力  $\tau_w$ 、粘性係数  $\mu$  を用いて壁面の疑似的な滑り速度  $u_{wp}$  を算出し、LBM へと導入する。滑り速度  $u_{wp}$  は

$$u_{wp} = \frac{4u_{p1} - u_{p2} - 2\Delta \frac{\tau_w}{\mu}}{3}$$

と与えられ、バウンスバック法を用いて LBM に導入する。開発した手法の評価を行うために、摩擦レイノルズ数 2,000 のチャンネル乱流ラージ・エディ・シミュレーションを行った。壁面垂直方向には 100 点の計算格子を設定し、乱流モデルにはスマゴリンスキーモデル[2]を用いた。解析は D3Q27 多緩和時間格子ボルツマン法[3]を用いた。

結果および考察

図2に摩擦速度で無次元化した平均速度分布を直接数値解析のデータ[4]と比較した結果を示す。なお、標準壁関数・解析的壁関数の2つの壁関数を導入して解析を行ったが、解析的壁関数法を用いた解析では数値

振動が生じ、正しい解を得ることができなかったので、標準壁関数法の結果のみを示している。標準壁関数で得られた結果は、DNS と非常に良い一致を示していることが確認できる。図3と図4にそれぞれ速度・圧力変動強度分布を示す。図3より、壁面近傍では壁面垂直方向やスパン方向に大きな速度変動が生じていることが確認できる。さらに、図4より圧力変動に関して非常に大きな値を予測してしまうことが確認できる。解析的壁関数法は瞬時の速度・圧力を薄層近似した運動量式に代入し壁面せん断応力を求めるが、過大な速度・圧力変動のために、解析的壁関数法では正しく壁面せん断応力が見積もれなかったと考えられる。

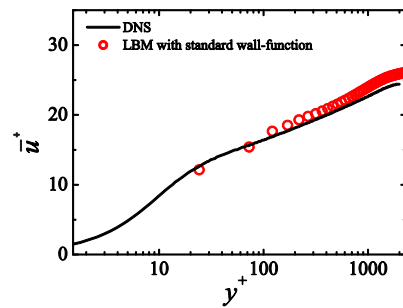


図2 平均速度分布

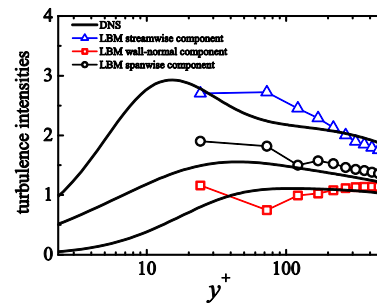


図3 壁面近傍の乱流強度分布

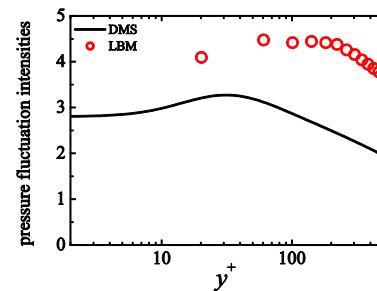


図4 圧力変動強度分布

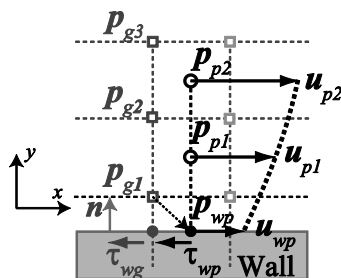


図1 壁面隣接セルの速度・圧力

#### まとめ、今後の課題

格子ボルツマン法(LBM)を用いて航空機周りに見られる高レイノルズ数乱流の解析を行うために、LBM に壁関数法を組み合わせた解析を検討した。壁関数法で得られた壁面せん断応力を LBM に導入するため、壁関数で得られた壁面せん断応力と実際の壁面せん断応力とが合致するように、壁面に滑り速度を与えて解析を行った。手法評価のために、D3Q27 多緩和時間格子ボルツマン法を用いて摩擦レイノルズ数2,000のラージ・エディ・シミュレーションを行い、壁関数法には対数測を用いる標準壁関数法と、剥離・再付着流れにも適用可能な解析的壁関数法を用いた。標準壁関数法で得られた結果は対数速度分布と良く合致することは確認されたが、壁面近傍では圧力・速度変動は過大に見積もってしまうことが確認された。また、解析的壁関数法を用いた解析では数値振動が生じてしまい LBM への導入には課題が残った。今後は、数値振動を抑える方法の検討や、壁面せん断応力の導入方法に問題がないか詳細に調べる必要がある。

#### 参考文献

- [1] K.Suga, T.J.Craft and H.Iacovides, An analytical wall-function for turbulent flows and heat transfer over rough walls. *Int. J. Heat Fluid Flow*, Vol. 27, pp852-866, 2006.
- [2] J.Smogorinsky, General circulation experiments with the primitive equations. *Mon. Wea. Rev.*, Vol.91, pp.99-164, 1963.
- [3] Suga, K. et al., A D3Q27 multiple-relaxation-time lattice Boltzmann method for turbulent flows, *Compt. Math. Appl.*, Vol.69, No.6, pp.518-529, 2015.
- [4] Hoyas, S. and Jimenez, J., Scaling of the velocity fluctuations in turbulent channels up to  $Re = 2003$ . *Phys. Fluids* 18 (1), 011702, 2006.





TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 抗インフルエンザウイルス広域中和抗体のエスケープ変異株由来表面糖タンパク質の構造解析  
英文: Structural analysis of surface glycoproteins derived from anti-influenza virus broadly neutralizing antibody-escape mutants influenza A viruses

利用課題責任者 長谷川 秀樹  
Hideki Hasegawa

所属 国立感染症研究所  
Affiliation National Institute of Infectious Diseases  
URL <https://www.niid.go.jp/niid/ja/>

邦文抄録(300 字程度)

我々は、有効性の高いインフルエンザワクチンや抗体医薬の開発を目指している。これまでに、様々なウイルス株の感染を阻止できる抗 HA 広域中和抗体 F11 を単離し、さらにその耐性変異を同定した。しかし、抗体による中和と変異による中和回避の分子機序は全く不明である。そこで本研究では、分子動力学解析により、変異が HA タンパク質構造に与える影響を調べた。その結果、耐性変異は、HA 表面の疎水性に富んだ窪み (hydrophobic groove) の構造特性変化を誘起することがわかった。この領域が、F11 のエピートプである可能性を示唆する。

英文抄録(100 words 程度)

We aim to develop effective vaccine and antibody preparation against influenza virus. Thus far, we have isolated an anti-HA monoclonal antibody F11, which neutralizes various influenza virus strains and identified two resistant mutations against F11. However, underlying mechanisms for the F11-mediated neutralization and antibody resistance remains unclear. To address this issue, we here examined effects of resistant mutations on the structural properties of HA protein by using molecular dynamics simulations. The study revealed that the mutations could induce structural changes around a particular hydrophobic groove of HA trimer surface, suggesting that this region may be an epitope of F11

*Keywords:* 5つ程度

influenza virus, broadly neutralizing antibody, neutralization escape, HA protein, molecular dynamics simulation

背景と目的

季節性インフルエンザ、および鳥インフルエンザの流行は社会に大きな損失をもたらす。我々は、有効性の高いインフルエンザワクチンや抗体医薬の開発に向けて、様々なウイルスの感染を阻止できる広域中和抗体の開発に取り組んでいる。これまでに、様々なウイルス株の感染を阻止できる抗 HA 広域中和抗体 F11 を単離し、その耐性変異を同定した。しかし、抗体による中和と変異による中和回避の分子機序は全く不明である。本研究では、耐性変異が標的分子 (HA タンパク質三量体) の構造特性にどのような影響をもたらすかについて分子動力学解析により調べた。

概要

H1N1 HA 分子 三次元情報 (PDB ID:3LZG

A/California/04/2009/H1N1 California 株) を鋳型に、ホモロジーモデリング法を用いて、中和実験に用いた株 (A/Narita/01/2009/H1N1 株)、および F11 抗体低感受性株 (C1 株、G6 株) の HA 三量体モデルを構築した。それぞれのモデルについて、Amber 16 プログラムパッケージを用いて 1 atm, 310 K, 150 mM NaCl 溶液環境下で MD 計算を 100 ns まで実施した。各分子の力場は、ff14SB (タンパク質)、TIP3P (水分子) を用いた。時系列三次元情報を用いて、RMSD (根平均二乗偏差) と RMSF (根平均二乗揺らぎ) を算出した。

結果および考察

RMSD 値の推移から、各 HA 三量体モデルは 70 ns 以降に熱的平衡状態に達すると推察された。そこで、100 ns 時の各 HA 三量体モデルを比較したところ、

C1 株では変異により HA 表面の疎水性に富んだ窪み (hydrophobic groove) の表面積が減少することがわかった。また HA タンパク質の個々のアミノ酸残基の 70 ns~100 ns の間の RMSF を解析したところ、C1、G6 株は、共に hydrophobic groove および fusion peptide 周囲のアミノ酸残基の RMSF の絶対値とプロファイルが変化していることがわかった。これらの結果より、F11 抗体は HA タンパク質の hydrophobic groove を標的とする可能性が考えられた。

#### まとめ、今後の課題

F11 抗体は、HA タンパク質の hydrophobic groove 周囲をエピトープとする可能性が提示された。既報のインフルエンザウイルス広域中和抗体は hydrophobic groove 周囲に結合することが知られており、F11 抗体も類似の領域を認識する可能性が高まった。次年度は、より生理的関連性の高いモデル(糖鎖付加型 HA 三量体分子モデル)を MD を用いて構築・解析し、今回の発見の再現性を調べる。

TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 スマートデータセンター実現に向けたデータ解析基盤の構築  
 英文: Towards a data analysis framework for smart data centers

利用課題責任者 松岡聡  
 Satoshi Matsuoka

所属 産総研・東工大 実社会ビッグデータ活用オープンイノベーションラボラトリ  
 AIST – TokyoTech Real World Big-Data Computation Open Innovation Laboratory  
<https://unit.aist.go.jp/rwbc-oil/>

データセンターの大規模化・複雑化に伴い、人力でシステムを監視し、運用設定を変更する従来型手法は困難になりつつある。本研究では従来型運用管理手法の問題を解決すべく、データセンターを構成する各種機器が生成するセンサーデータを蓄積し、解析することにより運用のトレンドを把握し、運用改善を行うシステムパラメータを提案し、運用にフィードバックするシステムの実現を目指す。TSUBAME3.0 の利用を通じて、センサーデータ、機械学習手段の調査を行い、センサーデータ収集基盤の基本設計を行った。

Operating datacenters by traditional manual operations becomes unfeasible because they become larger and more complicated systems than before. Our focus in this research is to develop a system that collects sensor data generated from components that compose datacenters, analyze the data and propose system parameters that can improve system efficiency. In this work, we surveyed sensor data of a datacenter and evaluated machine learning libraries on TSUBAME3.0 and then propose a basic design of data collector for datacenters.

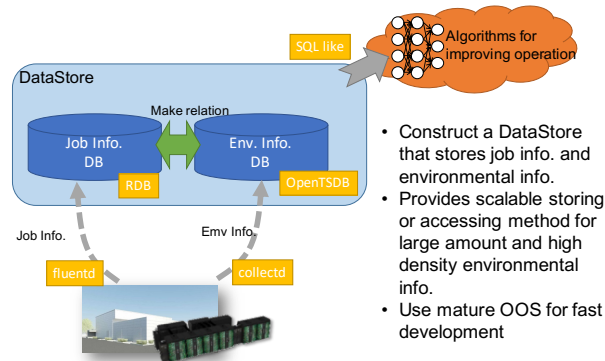
*Keywords:* smart data center, machine learning, system operation

## 背景と目的

データセンターの大規模化・複雑化に伴い、人力でシステムを監視し、運用設定を変更する従来型手法は困難になりつつある。その問題を解決すべく、近年ではデータセンターが生成する各種センサー情報をもとにデータ解析を行い、運用改善するパラメータを提案し、システムに自動適用する研究が行われていたり、Google など先進的な大規模クラウドデータセンターを運用する企業ではその技術の一部実用化を進めている。しかしながら、それら技術の詳細は公開されていないこと、また、データセンターのハードウェア構成や、提供するサービスごとに取得すべきセンサー情報が異なるため、他データセンター事業者が模倣することは困難である。本プロジェクトでは、近年注目されている HPC・AI 向けサービスを提供するデータセンターを事例に、オープンな技術を用いて、データセンターのデータ収集・解析・運用へのフィードバックを行うシステムの実現を目指す。TSUBAME3.0 を用いてセンサーデータの解析、機械学習手段の調査を行い、センサーデータ収集基盤の基本設計を行った。

## センサーデータ収集基盤基本設計

下図にデータストアの基本設計概念図を示す。



データ取得対象とするデータセンター構成要素としては、以下を想定する。

- 計算ノード
- ジョブスケジューラ(兼、資源管理サービス)
- 付帯設備(電力、冷却システム)

計算ノードについては、各種センサーデータとログデータを時系列データとして取得する。センサーデータについては collectd にて、分単位の粒度で取得することを考える。これは計算ノードで実行されるジョブ 1 つに

つき、ワークロードとして評価されるべきジョブは最低 1 分以上の実行時間は有すべきであること、その時に最低 1 データポイント以上のデータを取得することが必要という判断からである。より細粒度で取得することも不可能ではないが、ネットワーク転送量、蓄積するデータ量が増える。なお、TSUBAME2.0/2.5 計算ノード相当のデータを産総研 ABCI 規模(1088 ノード)のシステムで、1 分間隔で取得した際のデータ量は以下の通りとなる。

- 950MB/Node/年
- 1010GB/システム/年

ログデータについては fluentd にて、収集対象のログ出力のたびに取得することを想定する。取得するログ内容については、GPU, SSD などの各種デバイスのヘルスチェック結果、重要なプロセスの死活チェック結果などが考えられるが、運用状況に応じて判断すべき項目である。

ジョブレコードについては、利用者のジョブ完了のたびに取得する。1 つのジョブレコードは複数のデータ型の異なるデータから構成されるため、fluentd にてテキストベースで収集する。付帯設備については、各種センサーデータを時系列データとして取得する。付帯設備は使用する機材によってサポートするデータ取得用のプロトコルが異なるため、サポートされているプロトコルに応じたデータコレクタを採用する。

データストアを構成するデータベースには次の 2 種類を用いる。

- ジョブレコード保存用の Relational データベース
- センサーデータ、ログデータ保存用の時系列データベース

ジョブレコードを格納するために Relational データベースを用いる。利用者数・利用状況に応じて、ジョブレコード数は年間数万～100 万規模まで変動すると予想されるが、PostgreSQL や MySQL, MariaDB などの枯れたサーバクライアント型 DB であれば性能・機能的に十分と考えている。Relational データベースでは fluentd にて送られてきたデータをレコード形式に変換して DB に登録する。DB 上の 1 レコードサイズを仮に 4KB とすると、100 万レコードの場合には 4GB のサイ

ズとなる。5 年間運用したとしても 20GB であり、1 ノード構成で問題ないと考えられる。

ジョブレコード以外のデータは時系列データであり、その格納には時系列データベースを用いる。現状では、実装が十分に枯れている OpenTSDB の採用を考えているが、近年時系列データベースの開発競争が活発であることから、システム構築時の状況に応じて変更することも考えている。OpenTSDB は collectd、Fluentd の両方に対応したプラグインを有するため、それぞれで送られてきたデータはそのまま格納できる。ただ、格納データ形式は数値に限定されるため、ログデータを Fluentd で転送する場合には、意味のある数値データに変換して送る必要がある。「過去の運用データ解析を通じてのシステム運用改善」を目的としていることより、時系列データベースでは格納したデータの時間圧縮は行わない。これにより、ABCI 規模のシステムを 5 年間運用した場合、計算ノードのセンサーデータだけおよそ 5TB、ログデータや付帯設備のデータも含めると 7TB ほどになると思われる。1 ノードでも蓄積不可能なデータ量ではないが、ジョブデータとの関連づけのためには検索性能が重要となるため、分散システムとして構築した方が良いと考えている。

データ解析用プログラム、運用改善アルゴリズムからのデータストアへのアクセスについては、本データストア用に専用の API を用意するのではなく、様々な側面からデータを解析する必要があると考え、ジョブレコードを保存する Relational データベース、時系列情報を保存する時系列データベースのそれぞれの API をそのまま提供することを考えている。ただし、頻出する検索項目、例えば、特定ジョブが使用した計算ノード群における資源利用データの取得、に関しては、その機能を実現する API を用意する。

まとめ、今後の課題

本研究では TSUBAME3.0 上での TSUBAME2.0 のセンサーデータの解析、機械学習ライブラリの評価を通じて、大規模データセンター用センサーデータ収集基盤の基本設計を行った。今後、性能評価を含め詳細設計を行い、産総研 ABCI システムへの配備を行う。ABCIでの運用を通じて運用データを蓄積し、その解析

技術を確立するとともに、得られた成果は広く公開する。



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 MEMS 構造の光学応答解析  
 英文: Analysis on Optical Responses of MEMS Structures

菅 哲朗  
 Tetsuo Kan

電気通信大学  
 The University of Electro-Communications  
<http://www.uec.ac.jp>

本研究課題は、通常はシリコン単体では検出できない赤外光を、ナノピラー構造による光の高効率な吸収と、金属・シリコン界面に形成されるショットキー障壁を用いることで、電気的に検出可能とする赤外検出デバイスの光学応答解析を目的としている。受光素子の性能向上には計算による構造探索が必要である。しかし、通常のワークステーションでは計算速度に限界があり、有効な探索が難しかった。そこで、有限要素アプリケーションを高速な計算機上で用いることで、本課題の本格的な展開に取り組むために、TSUBAME での計算の方法の検討を進めた。

In this research, numerical analysis of optical responses of Si nano-pillar structures coated by Au thin film was performed. Our research group is investigating on a Si based infrared photodetector based on the nano-pillar structure, and finite element method calculation is used to analyze its photo response. This year, we mainly focused on an investigating regarding how to construct a calculation environment using TSUBAME.

*Keywords:* Si infrared photodetector, surface plasmon resonance, FEM, nano-pillar, MEMS

#### 背景と目的

本研究課題は、通常はシリコン単体では検出できない赤外光を、ナノピラー構造による光の高効率な吸収と、金属・シリコン界面に形成されるショットキー障壁を用いることで、電気的に検出可能とする赤外検出デバイスの光学応答解析を目的としている。学術的には、シリコンデバイスを利用した、新たな光センシング方法の確立と位置付けることができる。具体的構造としては、赤外光を波長選択的に吸収する金属ナノピラーを用いた、シリコン製の MEMS 構造である。入射した赤外光が金属表面の自由電子を励振することで表面プラズモン共鳴 (SPR) が発生することで、光が吸収される。ナノピラー構造の組成・形状により SPR が発生する波長や偏光、および入射角度の条件が異なるため、受光素子の性能向上には計算による構造探索が必要である。しかし、通常のワークステーションでは計算速度に限界があり、有効な探索が難しかった。

#### 概要

有限要素アプリケーションを高速な計算機上で用いる

ことで、本課題の本格的な展開に取り組むために、TSUBAME での計算の方法の検討を進めた。利用者は、光学的応答解析のソフトウェアとして、商用のアプリケーションソフトである COMSOL Multiphysics を利用しており、COMSOL の TSUBAME 上での実行可能性についての調査を進めた。

#### 結果および考察

調査の結果、ライセンスサーバの設置方法の具体案が明確化した。来年度以降、ライセンスサーバの TSUBAME 内のインストールを進め、本格的な計算実行へと着手する予定である。本年度は計算実行として本格的運用は未着手である。

#### まとめ、今後の課題

本年度は、計算機としての本格利用までは到達せず、商用アプリケーションを TSUBAME 上でどのように運用するかについての方針決定を進めた。来年度以降、本格的な計算に着手する予定である。





TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 GPGPU による長周期地震動シミュレーション  
 英文: Long-period ground motion simulation using GPGPU

藤原広行  
 Hiroyuki Fujiwara

国立研究開発法人 防災科学技術研究所  
 National Research Institute for Earth Science and Disaster Resilience  
 URL: <http://www.bosai.go.jp/>

本課題では、南海トラフで発生する海溝型巨大地震を対象として、3次元差分法を用いた大規模地震波伝播シミュレーションにより長周期地震動の検討を行った。具体的には、震源パラメータの不確実性や破壊不均質性を考慮して112ケースの震源モデルを設定し、関東から九州までを計算対象領域とし、申請者らがTSUBAMEのGPU環境に対応させた地震動シミュレータGMSによる地震波伝播シミュレーションを実施し、それらの結果をもとに震源モデルの不確実性や不均質性を考慮した長周期地震動評価を行った。

We simulated long-period ground motion of Nankai-Trough megathrust earthquakes by the 3-D FDM of the Ground Motion Simulator (GMS) adapted to GPU on TSUBAME. We used 112 cases of seismic source model which have different source parameters such like hypocenter and asperity and simulated long-period ground motion for vast area from Kanto to Kyushu. These simulation results enable us to estimate long-period ground motions by considering uncertainties of source model on long period ground motions.

*Keywords:* 長周期地震動、3次元差分法、GMS、南海トラフ地震

## 背景と目的

大地震の際に超高層ビルや石油タンクなどの長大構造物が長周期地震動による被害を受けることが知られている。2011年東北地方太平洋沖地震の際に、大阪平野内の超高層ビルが長周期地震動による被害を受けた事例は、震源から遠く離れた平野や盆地においても、複雑な地下構造の影響によって増幅された長周期地震動が、構造物に被害を及ぼし得ることを示している。3次元差分法などの数値シミュレーションにより複雑な地下構造が長周期地震動に及ぼす影響を評価することが可能だが、長周期地震動の予測のためには、地下構造の情報に加え、震源からの地震波の励起特性（震源モデル）に関する情報も必要となる。

巨大地震の震源モデルを事前に予測することが極めて困難なことは、2011年東北地方太平洋沖地震や2016年熊本地震からも明らかであり、長周期地震動の予測では、さまざまな可能性を考慮して多数の震源モデルを設定し、各震源モデルに対する予測に加え、予測結果がどの程度のばらつきを持つかを定量

的に評価することが重要となる。そのためには計算コストの高い大規模シミュレーションを多数行う必要がある。

そこで、本課題では、TSUBAMEのGPU環境を利用して、多数回の大規模地震波伝播シミュレーションにもとづいた長周期地震動のハザード評価を行う。

## 概要

南海トラフで発生する海溝型巨大地震を対象として、震源パラメータの不確実性や破壊不均質性を考慮して112ケースの震源モデルを設定し、関東から九州までを計算対象領域とし、3次元差分法を用いた大規模地震波伝播シミュレーションにより長周期地震動の検討を行った。シミュレーションには申請者らがTSUBAMEのGPU環境に対応させた地震動シミュレータGMSを使用した。

本検討で対象とした震源モデルは、地震調査研究推進本部の長期評価 ([https://www.jishin.go.jp/main/chousa/kaikou\\_pdf/nankai\\_2.pdf](https://www.jishin.go.jp/main/chousa/kaikou_pdf/nankai_2.pdf)) で例示された15の震源域のうち、マグニチュード8程度に相当する5つの震源域 (図

1) に対応するもので、それぞれ破壊開始点やアスペリティの配置を変えた計 112 モデルを設定した。

地下構造モデルは防災科学技術研究所で作成している全国の 3 次元地下構造モデルにプレート構造を付与したものをを用いた。南北 660km、東西 950km、深さ 100km までの領域を水平方向 140m、深さ報告 100m (深さ 8km 以深ではその 3 倍の大きさ) の格子でモデル化した。格子数は約 36.8 億である。

地震波伝播シミュレーションでは、地下構造モデルを水平方向に 6×8 に分割し、それぞれの部分領域に 1 つの GPU を割り付けて計算処理を行っている。計算結果は、GMS の 1D Dump 機能を利用し、地動速度の 1 成分を 1 ファイルとして東西、南北、上下の 3 成分について出力している。1 ファイルには、陸域に約 2km 間隔で設定した約 5 万観測点の毎タイムステップ (全部で 103200 ステップ=430 秒) の地動速度値が記録されており、1 ファイルで約 90GB である。震源モデルの大きさにもよるが、TSUBAME3.0 の 48GPU (12 ノード) を使用した実行時間は、2.1~2.5 時間であり、90%以上は 2.2 時間以下であった。

### 結果および考察

出力された 1D Dump ファイルに対して、フィルタ、ダウンサンプリングのポスト処理を行い、各ケースについて速度応答スペクトルの空間分布を求め、さらに、結果を統計処理し、5 つの震源域に対する平均の速度応答スペクトル (周期 3 秒、7 秒) の空間分布を求めた (図 1)。

震源域が主として西側に位置している AS010、TS010 では、震源域に近い太平洋側の紀伊半島南部や四国南部で揺れが大きくなるとともに、震源から

離れた大阪平野や濃尾平野でも局所的に揺れが大きい。震源域が主として東側に位置する SF010、S0010 では震源域に近い御前崎周辺での揺れが大きくなっていると同時に、震源域から離れた地域としては、大阪平野、濃尾平野に加え関東平野でも揺れが大きく、特に震源域が駿河湾まで広がる SF010 では 100cm/s を超える領域も現れている。震源域が浅い A0100 では SF010 に近い分布を示しており、更に日本海側でも揺れの大きな地域が見られ、震源が浅いために長周期地震動に寄与する表面波が効果的に励起されたためと考えられる。震源モデルの不確かさを考慮した計算結果から、長周期地震動が大きい領域が関東平野などの平野・盆地部に概ね限定されており、長周期地震動評価において地下構造の影響を適切に評価することの重要性を示している。

### まとめ、今後の課題

TSUBAME の GPU 環境を用いて、南海トラフの海溝型巨大地震を対象とした 112 ケースの震源モデルに対する長周期地震動シミュレーションを行い、震源モデルの違いによるばらつきを含んだ長周期地震動評価を行った。

今後は、本研究で行った検討を他の地域で想定されている巨大地震に対して適用していくと同時に、本課題で設定した多数の震源モデルが長周期地震動のばらつきを評価する上で十分な母集団となっているかについて検討することが重要な課題である。また、シミュレーションによって得られる膨大な計算結果から、防災、減災に有用な情報を抽出し、効果的に可視化する手法を開発することも今後の重要な課題である。

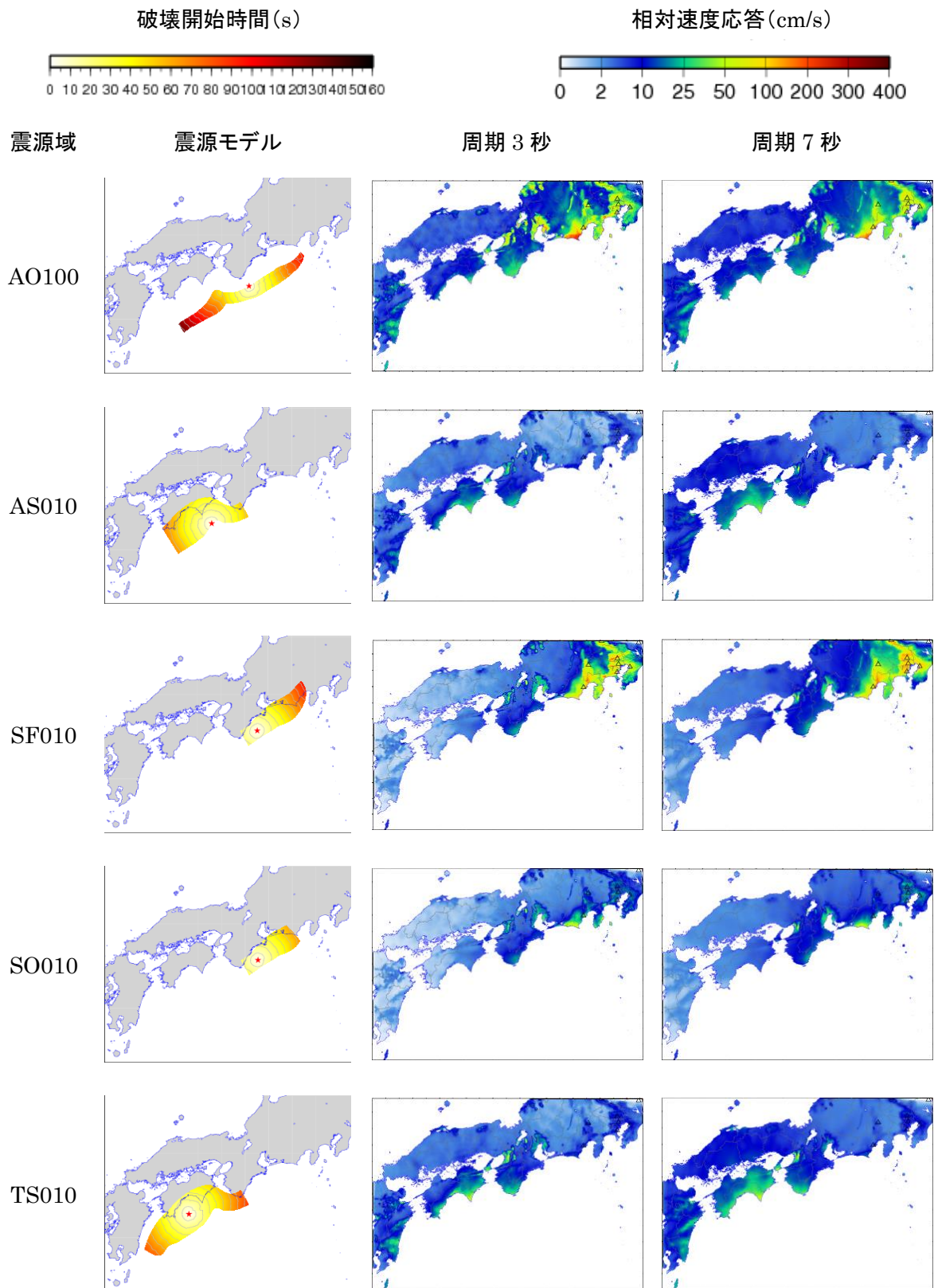


図1 5つの震源域の断層面（左列）と、速度応答値（減衰5%）の平均値空間分布（中列：周期3秒、右列：周期7秒）。



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

都市解像モデルへのドップラーライダー観測データ同化による局地域豪雨発生予測の精度向上  
Data assimilation of Doppler lidar data into the Large Eddy Simulation aiming at the improve  
convective generation forecasts in the Metropolitan area

古本 淳一

Jun-ichi Furumoto

京都大学生存圏研究所

Research Institute for Sustainable Humanosphere, Kyoto University

<http://www.rish.kyoto-u.ac.jp>

邦文抄録

人間活動に影響を及ぼす集中豪雨等の極端気象現象の中でも、都市部で発生する局地的豪雨は予測の社会的需要の大きさにも関わらずその発生予測は極めて難しい。この予報精度向上には、その発生メカニズムを明らかにすることが必要不可欠である。

本研究では積雲対流活動発生メカニズムを表現できる LES (Large Eddy Simulation) を用いて東京都市域の実地形を考慮したシミュレーションを行うとともに、詳細な街区レベルの地表面データを入力するだけでなく、CDL による稠密な風況観測データを組み合わせ、局地的豪雨の発生時における大気境界層内の大気構造を再現した。

英文抄録

This study focused on the characteristics of convective initiation by combining Coherent Doppler Lidar (CDL) and Large Eddy Simulation (LES), aiming to explicate the developing mechanism of the active convective cell, which often causes localized heavy rainstorms. The convergence near the ground is one of the most important triggers of the convective activity. The observed data using with CDL were assimilated into the LES model by using the three-dimensional variational assimilation method.

*Keywords:* large eddy simulation, Doppler lidar, boundary layer, localized heavy rain, metropolitan area

背景と目的

本研究では積雲対流活動発生メカニズムを表現できる LES (Large Eddy Simulation) を用いて東京都市域の実気象シミュレーションを実現する。詳細な街区レベルの地表面データを入力するだけでなく、CDL (Coherent Doppler Lidar) による稠密な風況を同化を行うことで、極端気象の予測精度向上を目指し、気象予報モデルの空間分解能を 2m まで向上させる。

メソ気象モデルとして実用されている WRF (Weather Research and Forecasting) と独・ハノヴァー大学が開発した LES モデルである PALM を用いて、空間分解能を多段階で向上させ、対流発生に重要な役割を果たすとされる下層風収束をモデル内部で表現するため、CDL 観測で得られた風速データを LES に同化することを試みる。

概要

人間活動に影響を及ぼす集中豪雨等の極端気象現象の中でも、都市部で発生する局地的豪雨は予測の社会的需要の大きさにも関わらずその発生予測は極めて難しい。この予報精度向上には、その発生メカニズムを明らかにすることが必要不可欠である。

本研究では積雲対流活動発生メカニズムを表現できる LES (Large Eddy Simulation) を用いて東京都市域の実地形を考慮したシミュレーションを行うとともに、詳細な街区レベルの地表面データを入力するだけでなく、CDL による稠密な風況観測データを組み合わせ、局地的豪雨の発生時における大気境界層内の大気構造を再現した。

結果および考察

都市域における大気境界層内の風の流れについて

詳細に表現するため、大きな渦を陽に表現できる LES モデルである PALM Ver.4.0 [1] を使用し、東京都環境科学研究所の地表面データを用いて図 1 に示す水平分解能 2m の高解像度都市モデルを作成した。計算領域は 2km x 2km で CDL が設置されていた東京都中央区日本橋が含まれるように設定した。

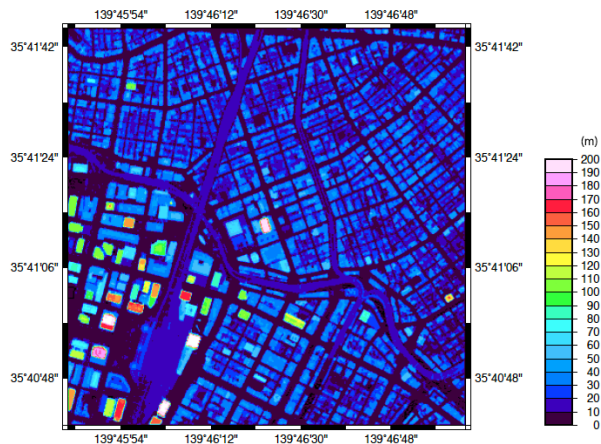


図 1: 東京都中央区日本橋付近の街区

CDL は距離分解能が 100m、角度分解能は 1 度 で観測範囲は半径 4km、地上 200m の高さの視線風速を観測した。観測結果を VAD 法を用いて解析し、解析された平均風を PALM に初期値として用いた。

本実験では CDL で観測を行っていた 2015 年の夏季において発生した局地的豪雨の事例 (2015 年 7 月 24 日) について解析を行った。

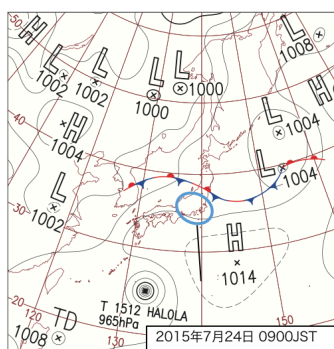


図2:2015 年 7 月 24 日午前 9 時の地上天気図

図2にこの日の午前 9 時の地上天気図、図3に世田谷で 22.5mm/10min の激しい降雨があった時のレーダー画像を示す。この日は停滞前線が東北地方に停滞し、関東付近は太平洋に存在する高気圧の縁辺をまわりこむ南西風場になっており、湿った空気塊が流入しやすい状況であった。図3に示すレーダーエコーの時系列

図によると、西から積乱雲が発達しながら都市部に接近し、CDL が設置されている日本橋周辺に近づくにつれて発達していく様子がわかる。

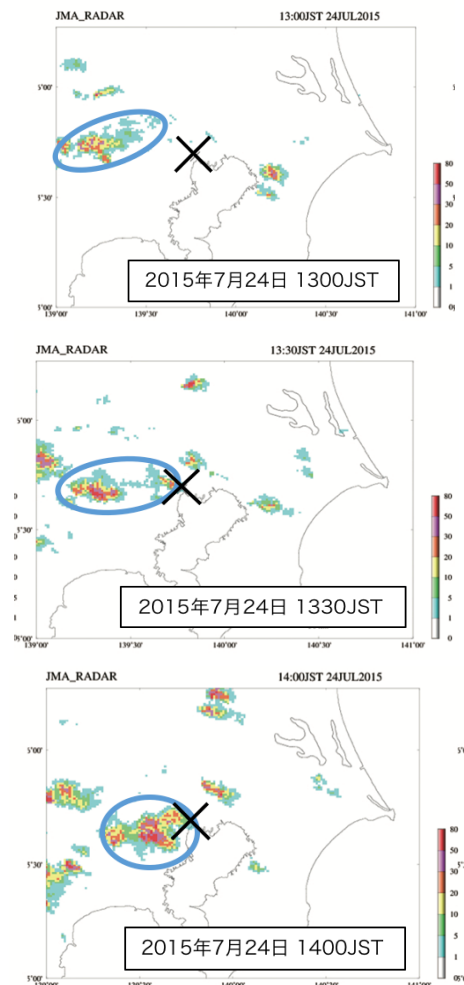


図 3: 関東周辺のレーダーエコー分布図 (2015 年 7 月 24 日 1300JST~1400JST)

図 4 に計算開始 1 時間後における地上 200m での水平風 (矢印) と鉛直流 (上向きが正) を示す。計算領域内の境界層内では空間スケールが数 100m スケールのオーダー局所的な収束域と上昇流域が再現されており、積雲対流が接近すると局所的に対流活動が活発化しやすい環境が形成されていたことが示唆された。



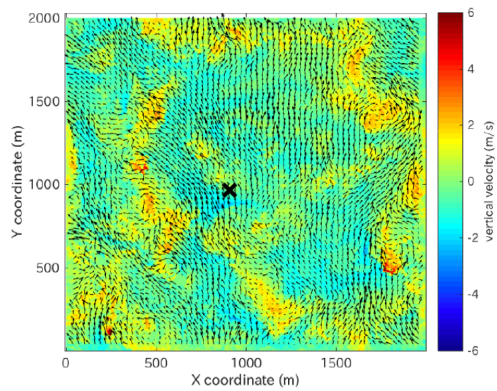


図 4: 地上 200m の水平風(矢印) と鉛直流(上向き正)  
×印は CDL 観測点を示す。

### まとめ、今後の課題

モデルの分解能を 2m まで向上させ、ビルを解像することによって都市部の細かな収束発散を表すことが可能になった。境界層内の局所的な収束発散が、自由対流高度より上空の大气に及ぼす影響を明らかにするには、より多くの事例について実験を行い、CDL 観測との比較検証が必要であると考えられる。

### 参考文献

[1] Siegfried Raasch and Michael Schroter, "PALM- A large-eddy simulation model performing on massively parallel computers", Meteorologische Zeitschrift, Vol. 10, No. 5, pp,363-372, 2001.



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 第一原理計算と反応速度論による触媒活性予測技術の確立  
 英文: Development of combined method of first-principle method and  
 chemical kinetics for catalyst activity prediction

利用課題責任者 石川 敦之  
 First name Surname Atsushi Ishikawa

所属 物質・材料研究機構  
 Affiliation National Institute of Materials Science

邦文抄録 本研究では、Ru, Os, Rh 触媒による  $\text{NH}_3$  生成反応を研究対象として第一原理計算と反応速度論を用いることにより  $\text{NH}_3$  合成の活性予測を行なった。理論計算の結果、Ru が触媒として最も優れた役割を示すことがわかり、実験と一致する結果となった。また、活性を決める要因は(1)活性化エネルギー、(2)活性サイトの占有率、(3)活性サイトの数 の3つが主であることが示された。

英文抄録  $\text{NH}_3$  synthesis on Ru, Os, and Rh nanoparticle catalysts was investigated using density functional theory calculations. For all metal species, step sites appeared at nanoparticle diameters ( $d$ )  $>2-4$  nm. The calculated activation barriers ( $E_a$ ) were small at step sites, and Ru and Os step sites exhibited similar  $E_a$  values despite the former having a higher turnover frequency. This is likely due to the surface coverage of vacant sites being higher on Ru. Our results show that  $E_a$  values, surface vacant sites, and the number of step sites are important factors for  $\text{NH}_3$  synthesis. The Ru nanoparticles exhibited high activity due to satisfying all three factors.

Keywords: 触媒、第一原理計算、表面、化学反応、電子状態

## 背景と目的

近年、水素社会の到来により優れた  $\text{NH}_3$  合成触媒が求められている。本研究では、第一原理計算をベースとした理論計算によって  $\text{NH}_3$  合成に適した金属触媒を検討した。

## 概要

本研究では、Ru, Os, Rh 触媒による  $\text{NH}_3$  生成反応を研究対象として第一原理計算と反応速度論を用いることにより  $\text{NH}_3$  合成の活性予測を行なった。第一原理計算には平面波基底関数による密度汎関数法を用い、プログラムソフトウェアとして Vienna ab initio simulation package (VASP)を利用した。

## 結果および考察

図1に、第一原理計算により求められた Ru, Os, Rh のステップ面による  $\text{N}_2$  解離反応の  $\text{N}_2$  吸着状態と遷移状態の構造を示す。Ru や Os においては吸着状態と遷移状態の  $\text{N}_2$  結合長変化が Rh に比べて小さく、 $E_a$  が

小さいことが予測される。実際に算出された  $E_a$  は、Ru, Os, Rh に対して 0.39, 0.41, 1.20 eV と、予測した通りの結果となった。これらの活性化エネルギーを用いて反応

速度論解析を実行したところ、Ru がもっとも  $\text{NH}_3$  生成反応速度が高く、次いで Os, Rh となった。これは、Os では  $\text{NH}_2$  が活性サイトの占有率が高く被毒作用を引き起こしていることが原因であった。

## まとめ、今後の課題

これらの研究成果により、活性化エネルギー同様に被毒作用も触媒活性を左右する上で重要な因子であることが、第一原理計算と速度論解析により明らかとなった。今後はさらに多くの金属系や他の触媒反応系における活性の予測に取り組む。

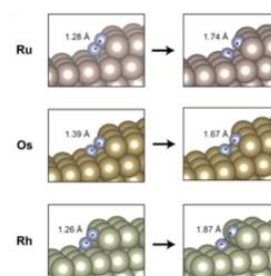


図1. Ru, Os, Rh における  $\text{N}_2$  解離反応の遷移状態



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 HPC を利用した自然言語処理技術の研究

英文: High Performance Computing for Natural Language Processing Technology Research

利用課題責任者

鳥澤 健太郎

所属

国立研究開発法人情報通信研究機構 ユニバーサルコミュニケーション研究所

データ駆動知能システム研究センター

<http://www2.nict.go.jp/direct/>

邦文抄録(300 字程度) 質問応答や対話等の高度な自然言語処理技術には、大量のテキストの意味的に深い分析が必要であるが、数十億規模の Web 文書の分析には、多数の計算機を用いた並列実行が必須である。本課題では、各種の分析プログラム群を TSUBAME 上で多数並列実行し、大量の Web テキストを高速に処理することを目指した。その結果、大規模 Web 情報分析システム WISDOM X で必要となる分析を、1 日 1 億 Web ページ以上の速度で実行できた。また、200 億レコード(1.5TB)超のデータに対して、対話システム WEKDA で用いられる深層学習プログラムを適用し、最大 100 以上の GPGPU を使用し、16 時間で処理できた。さらに、ニューラルネットワークの自動分割による、複数 GPU を用いた並列処理を実現し、同じく WEKDA で用いられる深層学習プログラムの学習を約 2 倍高速化した。

英文抄録(100 words 程度) Advanced natural language processing applications including question-answering and dialogue systems require semantic deep analyses on a large amount of texts. Such analyses should be distributed on many compute nodes because the semantic analyses consist of computationally heavy programs. In this project, we introduce our middleware RaSC and tuned the configurations for TSUBAME. As the result, we processed 100 million web pages per day with analysis programs for WISDOM X. We also processed 20 billion records in 16 hours with our deep learning programs for WEKDA using more than 100 GPUs. We also automated automatic model parallelism on multiple GPUs and doubled the training speed of the deep learning programs.

*Keywords:* 自然言語処理, 大規模情報分析, テキスト分析, モデルパラレル

## 背景と目的

質問応答や対話等の高度な自然言語処理技術には、大量のテキストの意味的に深い分析が必須である。例えば、情報通信研究機構が開発している大規模 Web 情報分析システム WISDOM X<sup>1</sup> や対話システム WEKDA<sup>2</sup> には、40 億以上の Web ページの解析結果が用いられている。こうしたアプリケーションのためのテキストの分析は、計算機負荷が高く、数十億規模の Web 文書の分析には、多数の計算機を用いて並列に分析を実行することが必須である。

そこで本課題では、当機構の持つ分析プログラム群を TSUBAME 上で多数並列実行し、大量の Web テキストを高速に処理することを目指した。

## 概要

当機構に蓄積された Web テキストについて、TSUBAME 環境で各種の分析プログラムを適用した。本課題では、以下に示す 3 つのケースを実施した。  
ケース 1) ストリーミング処理によるプログラム連携  
当機構で用いている分析プログラムは、速度や使用メモリ量が様々であり、複数の分析プログラムを連携させるにあたって、単純に全てのプログラムを同一プロセス数だけ起動してパイプラインを構成すると、メモリ消費量等の点で非効率である。従って、当機構が開発したミドルウェア RaSC を用い、各分析プログラムに最適な数のプロセスを起動してプールし、それらをストリーミング通信によって連携する構成とした。

バッチジョブの中でこの構成を実現するため、バッチジョブの冒頭で、分析プログラムの実行環境構築と、

<sup>1</sup> <http://wisdom.nict.jp>

<sup>2</sup> <https://www.nict.go.jp/press/2017/10/24-1.html>

それらのプログラムのプロセスをプールする RaSC サービスの起動等を行った。RaSC を用いる場合、全体の実行を管理する中央サーバによって、分析プログラムの複数の実行ノードへの配備や、プログラムの死活監視等を一括して行える。しかし、TSUBAME 環境においては動的に計算ノードが割り当てられるため、複数のジョブにまたがって持続的に中央サーバを起動しておくことが難しい。そのため、各ジョブが一つの計算ノードを用いるものとし、その計算ノードに中央サーバと全ての分析プログラムを設置・起動してから、分析対象データを送信することとした。

#### ケース 2) 深層学習プログラムのバッチ実行

本課題では、学習済みの深層学習モデルを用いて、Web テキストに対する推論を行うプログラムを実行した。深層学習のためのライブラリには、実行時に GPGPU のための実行コードを生成し、コンパイルを行うものがある。コンパイルされたコードは、ディスク上の一時領域に保存される。本課題で用いた深層学習プログラムでは、深層学習のためのライブラリに Theano 及び PyTorch を用いたが、前者はユーザのホーム領域を一時領域として、GPGPU 用コードの出力とコンパイルを行う。多数の計算ノードが同一のパスを参照することになるため、ロック開放待ちやサイズの小さいファイルが多数読み書きされるため、パフォーマンスが低下する。そのため、コンパイル結果の配置は、ノードごとに独立したキャッシュ領域に書き込むように設定して実行した。

#### ケース 3) モデルパラレルによる並列化

上記で用いたニューラルネットワークが、並列実行可能な分岐構造を持つことから、ニューラルネットワークを分割し、各部を異なる GPGPU 上で計算するモデルパラレルの自動化を試みた。このため、ニューラルネットワークを構成する各計算モジュールの計算時間をプロファイリングするとともに、最適な分割方法をスケジューリングによって求めるようにした。

#### 結果および考察

上述のケース 1) においては、対象となる全データは圧縮状態で数百 TB 規模となるため、本課題ではその一部を対象とした上、200 万 Web ページ相当のデータ (約 30GB) を 1 単位として TSUBAME にコピーし、1

ジョブとして実行した。各ジョブ内においては、RaSC を用いた並列実行が行われるため、資源リソースには多数のコアを持つ `f_node` を用いた。最大でおよそ 50 ジョブが同時実行され、一日あたり 1 億 Web ページを超える分析が実現された。

またケース 2) においては、対象となる Web テキスト約 1.5TB (200 億レコード) を全て TSUBAME のストレージにコピーした上で、学習済み深層学習モデルによる推論を行った。また、全データを 200 に分割し、それぞれを 1 つのジョブとして実行した。各ジョブは GPU を 1 つのみ使用するため、資源タイプには `q_node` を用いた。その結果、最大で 100 ジョブ以上を同時稼働することができ、およそ 20 時間で全ジョブを完了した。

ケース 3) については、WEKDA で用いる深層学習プログラムの学習を、2 倍以上高速化することに成功した。同一の計算機上の GPGPU を用いるため、資源タイプには 4 枚の GPGPU が利用可能な `f_node` を用いた。高速化効果には、GPGPU 間の通信速度が極めて大きく影響するため、TSUBAME において GPGPU が NVLink で接続されていることが大きく寄与したと思われる。

#### まとめ、今後の課題

TSUBAME の計算機リソースが、ストリーミング処理による分析プログラム連携を行うケースや、深層学習プログラムを用いる行うケースなど、幅広い状況で大量データの処理に有用であることが確認された。

一方で、TSUBAME 上に配置不可能な大規模データを処理するため、データを分割してそれぞれをバッチジョブで処理していることから、入出力データのディスク読み書きのコストを生じている。今後、TSUBAME 環境に起動した分析プログラムに対し、ストレージを介することなく、直接ネットワークでの入出力データの送受信を行うよう拡張することを検討している。

また本課題で実施したモデルパラレルの自動化は、1 台の計算機上の GPGPU を用いることを前提としたものであり、スケーラビリティに限界があった。今後、TSUBAME が多数の GPGPU 計算機を高速に結合していることを生かし、複数の GPGPU を用いるように拡張していく予定である。

利用課題名 有機分子集合体の半導体特性・光応答特性の解明

英文: Elucidation of semiconducting and photo response properties of organic molecular assemblies

河東田 道夫  
Michio Katouda

国立研究開発法人理化学研究所 計算科学研究機構  
RIKEN Advanced Institute for Computational Science  
[https://researchmap.jp/michio\\_katouda/](https://researchmap.jp/michio_katouda/)

有機分子集合体を対象に、半導体特性や光応答特性の発現機構を第一原理電子状態計算や古典分子動力学シミュレーションなどの大規模計算化学シミュレーションにより解明する。研究対象として、従来型の有機半導体や有機発光材料だけでなく、外力などの外部刺激に反応して発現および変化する動的な半導体特性やメカノクロミック発光特性を示す分子を対象に、圧力に誘起されて発現するこれらの電子機能の発現機構を解明する。さらに、仮想的な計算機実験を実施し、新奇有機分子材料の理論的分子設計、外力などの外部刺激に寄る動的な半導体特性や光応答特性の制御方法の提案も試みる。

The object of this project is to elucidate mechanisms of hole and electron transport and photoluminescence processes in organic molecular assemblies responded to the external stimulus such as pressure, rubbing, light, and heat. In this project, we use TSUBAME for large-scale massively parallel electronic structure calculations of organic molecular assemblies.

*Keywords:* 有機半導体、有機発光材料、有機太陽電池、第一原理電子状態計算、分子動力学シミュレーション

#### 背景と目的

近年、材料作成コストの低さ、材料の軽量性と加工性より、有機エレクトロニクス材料研究に関心が持たれている。特に、有機太陽電池、有機発光材料、有機半導体は次世代エネルギー・電子デバイス材料として精力的に研究開発が進められている。

これらの有機エレクトロニクス材料は、光や電場を加えた際の励起子の生成、励起子の電荷分離、ホール・電子移動が材料の性能に大きく影響するが、マクロな材料中の高速かつ微視的な物理的機構を伴うため、物理化学的測定による材料の動作機構の理解は難しく、詳細な物理的動作機構は未解明な材料も多い。そのため、計算物理、計算化学を活用し、材料の物理的動作メカニズム解明や材料設計・探索を微視的・直接的に検討する研究アプローチが強く求められている。

本研究課題では、有機分子集合体を対象に、半導体特性や光応答特性の発現機構を第一原理電子状態計算や古典分子動力学シミュレーションなどの大規模計算化学シミュレーションにより解明することが目的である。

#### 概要

本研究では、従来型の有機半導体や有機発光材料だけでなく、外力などの外部刺激に反応して発現および変化する動的な半導体特性やメカノクロミック発光特性を示す分子を対象に、圧力に誘起されて発現するこれらの電子機能の発現機構を第一原理電子状態計算により解明するために研究計画を策定し共同利用を申請した。

申請が採択された後、実験担当の共同研究者と相談し、TSUBAME を用いた計算を行うための結晶構造モデルを作成し、第一原理電子状態計算プログラム Quantum Espresso を用いた計算を行うための入力ファイルを作成し、計算の実施準備を行った。

#### 結果および考察

本務との兼ねいで課題を実施する時間が確保できずシステムを全く利用しなかったため、利用成果はない。

#### まとめ、今後の課題

今後、本課題を実施する研究時間が確保できた際に

は、TSUBAME 共同利用を申請し、本課題を実施したい。



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 深層学習によるゲノム上の遺伝子発現制御情報抽出  
英文: Extracting regulatory code on genome by deep learning

木立 尚孝  
Hisanori Kiryu

東京大学 新領域創成科学研究科 メディカル情報生命専攻  
Dept. of Computational Biology and Medical Sciences, GSFS, Univ. of Tokyo  
<https://sites.google.com/edu.k.u-tokyo.ac.jp/kiryulab>

邦文抄録(300 字程度)

細胞中の遺伝子の発現量は主にシスエレメントと呼ばれるゲノム上の DNA 配列によって制御されている。本研究では、ゲノム中のシスエレメントを精度よく検出するために、遺伝子周辺のシスエレメントを含むと予想される DNA 配列を入力としその遺伝子の発現パターンを出力とする畳み込みニューラルネットワーク(CNN)モデルを作成した。本手法を線虫(*C. elegans*)の遺伝子発現量データ(single cell RNA-seq)に対して適用したところ、CNN のパラメタの一部である畳み込みフィルタに転写因子の結合モチーフを含む、シスエレメントと予想される配列特徴が学習された。

英文抄録(100 words 程度)

Gene expression in a cell is controlled mainly by DNA sequences on genome called cis-regulatory elements. In this study, we developed a convolutional neural network (CNN) model to predict an expression pattern of a gene from DNA sequence around the gene, which may harbor cis-regulatory elements. We have applied our model to a gene expression data (single cell RNA-seq) of worm (*C. elegans*). The CNN model learned sequence features including transcriptional factor binding motifs, which may be cis-regulatory elements.

*Keywords:* deep learning, convolutional neural network, single cell RNA-seq, cis-regulatory elements, worm

## 背景と目的

細胞中の遺伝子の発現量は、主にシスエレメントと呼ばれるゲノム上の DNA 配列によって制御されている。このシスエレメントを遺伝子の発現量データを用いて検出する試みがなされてきた。また、近年、画像認識等の分野で成果を上げた畳み込みニューラルネットワーク(CNN)が生物の配列解析にも用いられるようになってきた。CNN は予測を行う上で有用な特徴を入力の中から自動的に抽出するという特徴を持っている。

本研究では、ゲノム中のシスエレメントを精度よく検出するために、遺伝子周辺の DNA 配列を入力とし、その遺伝子がどの細胞で発現するかという発現パターンを出力とする CNN モデルを作成した。発現パターンを予測するように学習を行った CNN モデルは、入力 DNA 配列中からシスエレメントを含む発現パターンの予測に有用な特徴を抽出していると期待出来る。

## 概要

線虫(*C. elegans*)の遺伝子の転写開始点周辺の DNA 配列を入力とし、その遺伝子の発現パターン(single cell RNA-seq データより作成)を出力とする CNN モデルを作成した。線虫の遺伝子のうち、11865 遺伝子を訓練データ、2966 遺伝子をテストデータとし、訓練データに対し予測誤差を最小化するように学習を行った。学習後のモデルの畳み込み層において認識されている配列特徴を計算した。

## 結果および考察

上記のデータを用いて学習を行い、テストデータに対する予測精度を評価した結果、CNN モデルが入力配列から発現パターン予測に有用な特徴を抽出していることが示唆された。学習後のモデルの畳み込み層において認識されている配列特徴を計算したところ、一部が既知の転写因子結合モチーフなどのシ

スエレメントと一致した(図 1)。

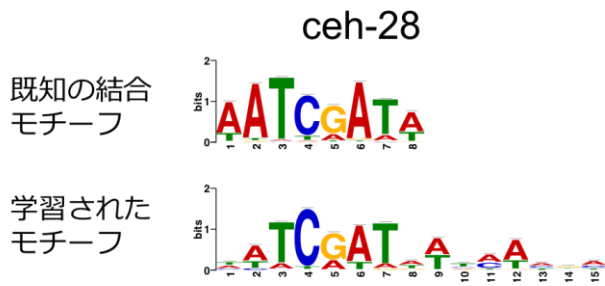


図 1. 転写因子 ceh-28 の既知結合モチーフ(上段)と CNN モデル中の量み込みフィルタに学習されたモチーフ(下段)。

#### まとめ、今後の課題

遺伝子周辺 DNA 配列から遺伝子発現パターンを予測する CNN モデルを作成した。このモデルは発現パターン予測に有用な配列特徴の一部を抽出したことが確認された。今後の課題としては、この CNN モデルを用いた新規シスエレメントの発見等が挙げられる。

TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 ナノ構造界面における熱輸送特性の分子シミュレーション  
 英文: Thermal transport properties at nanostructured interfaces

利用課題責任者  
 塩見淳一郎  
 Junichiro Shiomi

所属 東京大学 工学系研究科 機械工学専攻  
 Affiliation: Department of Mechanical Engineering, The University of Tokyo  
 URL: <http://www.phonon.t.u-tokyo.ac.jp>

近年ナノ加工技術の発達により、ナノスケールにおける熱・電子輸送特性制御が可能になりつつある。一方、ナノ構造は複数の特性に影響するため、それぞれの特性を独立に制御することが困難である。例えば、熱電変換を目的とした加工ではナノ構造が熱・電子両方の特性を抑制してしまい、熱伝導率抑制の効果が限定的になってしまう場合が多い。そこで本研究では、グリーン関数法を用いた熱・電子輸送特性解析とベイズ最適化を組み合わせることで、熱電変換特性向上を目的としてグラフェンナノリボンの最適ナノ構造を探索した。その結果、構造最適化により熱伝導特性抑制と電子輸送特性向上を両立させ、熱電変換特性を大きく向上させることが可能であることが明らかになった。

Recent development of nanotechnology enables to tune thermal and electronic properties at nanoscale. Because nanostructures affect different properties, independent tuning of these transport properties is difficult. For instance, nanostructures for thermoelectric applications suppress both of heat and electron transport. Here, we perform optimization of nanostructures of thermoelectric graphene nanoribbons by combining phonon and electron transport simulations with a Green's function method and a Bayesian optimization. We found that the multifunctional structural optimization can independently tune phonon and electron transport properties and considerably accelerate nanostructure design for thermoelectric properties.

*Keywords: Nanostructure, thermoelectric property, Green's function method, Bayesian optimization*

## 背景と目的

材料開発ではしばしば複数の材料特性間のジレンマに直面する。例えば、熱電変換では高い電気伝導率、高いゼーベック係数、低い熱伝導率が必要になるが、一般にはそれらは材料構造に対して相反する傾向を示す。ナノ構造技術はこれらの特性をある程度独立に制御することを可能にするが、最適構造の決定は自由度が大きく困難である。本プロジェクトでは、熱電材料としてグラフェンナノリボン(GNR)を例に、複数(フォノンと電子)の輸送特性解析とベイズ最適化による構造最適化を交互に行うことにより、熱・電子輸送特性の独立制御を可能にした。その結果、ランダム探索と比較して約 5 倍の速さで構造最適化が可能になり、熱電変換効率を 10 倍以上向上することに成功した。また、この手法は熱・電子輸送特性のみでなく、機械、化学特性等、様々な物性の最適化に利用することが可能である。

## 概要

近年のナノ技術の発達により、ナノスケールにおける熱・電子輸送特性制御が可能になりつつある。そこで、本研究では、従来の熱・電子輸送特性解析技術(材料科学)と情報科学を組み合わせたマテリアルズ・インフォマティクス(MI)を駆使することで、種々の目的にかなった最適なナノ構造を提案する。特に、熱電変換材料として注目されているナノ構造体(ナノ粒子の焼結材やナノ粒子を母材に埋め込んだ構造体)を対象とし、第一原理解析やグリーン関数法を用いた材料特性解析とベイズ最適化を組み合わせることで、最適なナノ構造を効率よく探索し、設計指針を確立する。

## 結果および考察

まず、ベイズ最適化の効率を評価するために、ベイズ探索とランダム探索それぞれにおいて全ての構造の内上位  $k\%$  に含まれる構造を見つけるまでに必要な試

行回数を比較することでベイズ最適化の効果を評価した。この評価では全ての候補構造が解析可能なよう、図 1 挿図のように、幅方向に炭素原子鎖が  $m$  列 ( $m = 6-10$ ) 並んだ zigzag 型 GNR ( $m$ -ZGNR) の比較的小さなユニットセルの周期構造を用い構造最適化を行い、完全な六員環の有無により各サイトを“0”, “1”でラベル付けしベイズ探索の descriptor とした。

その結果図 1 に示すように、全ての場合においてベイズ最適化により構造最適化が加速された。また、探索構造が限定される程 ( $k$  が小さい程) ベイズ探索の効率が向上していることから、探索構造が多い場合などのより厳しい条件において、ベイズ最適化によりより効率良く構造最適化が可能であることが期待できる。なお、 $m = 6$  の場合に特異点が生じているのは、構造の変化に対して熱電特性が不連続に変化しているためであると考えられる。

さらに、より現実的な系として  $m$ -ZGNR ( $m = 6-10$ ) の長さ 30 nm の領域に直径 2.9 Å のポーラスを作製することで熱・電子輸送特性を制御した。図 2A に 7-ZGNR のナノ構造領域に周期的にポーラスを導入した構造と非周期にポーラスが導入された最適構造を示した。図 2B に示すように、ポーラスを導入していない構造と比較して、無次元熱電特性は約 5 倍、11 倍となった。さらに、パワーファクター、熱抵抗ともに構造最適化により向上していることが分かる。

詳細なフォノン・電子輸送特性の解析を行った所、フォノンに関してはポーラスの導入による界面散乱、ポーラス間における干渉効果、さらに非周期性によるフォノンモードのミスマッチの効果が構造最適化により最大限に利用されていると考えられる。また電子に関しては、非周期性によりエッジ構造以外の状態の透過が抑制されることで、エッジ状態を利用することで熱電変換効率が向上していることが明らかになった。

### まとめ、今後の課題

以上のようにベイズ最適化により、複数の材料特性を独立に制御することが可能であることが明らかになった。一方で、本研究ではモデルのサイズや手法において現実を完全には再現できていない部分もあるため、今後、計算速度と現実的な計算のバランスをとりつつ、より現実的な系の計算を進めていく予定である。なお、

本プロジェクトの成果は参考文献[1]にまとめられている。

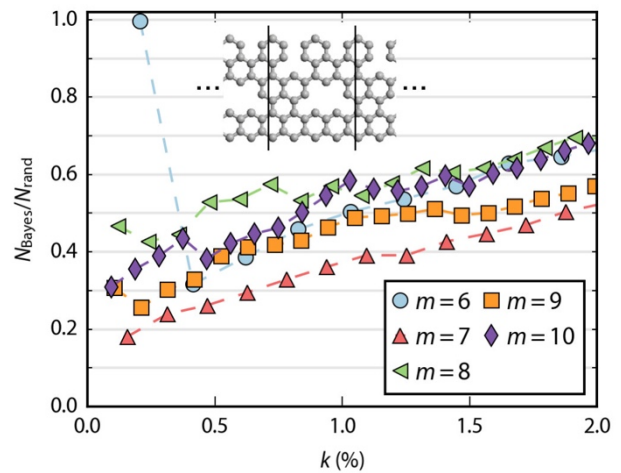


図 1. ベイズ最適化の効果.  $N_{\text{Bayes}/\text{rand}}$  はベイズ/ランダム探索において、全ての候補構造の内上位  $k\%$  に含まれる構造を見つけるのに必要な試行回数. [1]

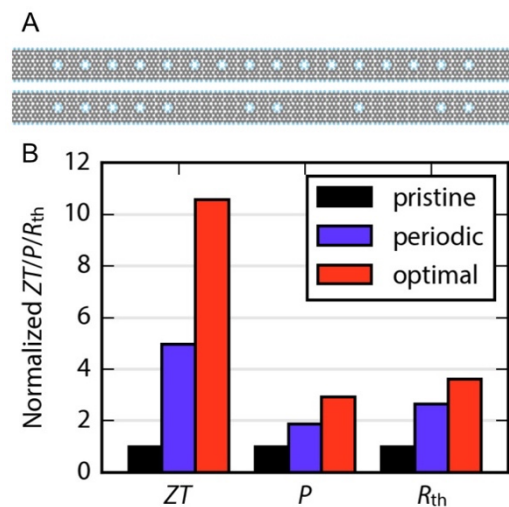


図 2. GNR 熱電材料のナノ構造最適化. A. 規則的なポーラス構造 (上) とランダムにポーラスを作製した最適構造 (下). B. ポーラスのない構造 (黒), 周期的なポーラス構造 (青), 最適化された非周期的なポーラス構造 (赤) の熱電変換特性. 無次元性能指数 ( $ZT$ ), パワーファクター ( $P$ ), 熱抵抗 ( $R_{th}$ ). [1]

### 参考文献

[1] M. Yamawaki, M. Ohnishi, S. Ju, and J. Shiomi, Sci. Adv.4, eaar4192 (2018).

TSUBAME 共同利用 平成 30 年度 学術利用 成果報告書

利用課題名 深層学習処理基盤を用いた大規模環境計測データの解析と応用  
英文: Developing Large-Scale DNN for Real-World Data

利用課題責任者 李 時旭  
First name Surname Shi-wook Lee

所属 産業技術総合研究所  
Affiliation National Institute of Advanced Industrial Science and Technology  
URL <http://www.aist.go.jp/>

邦文抄録(300 字程度)

マルチメディア情報の時系列処理に適した統合的な回帰ニューラルネットワークの高性能化を研究目的とする。既存の技術では、画像・音声・テキストなどの情報を個別に処理するシステムや深層学習の研究開発が進められてきた。本研究課題では、異種性の高い画像・音声・テキストの情報を総合的に処理可能な深層学習の開発やシステム統合手法を研究する点に学術的な意義がある。研究では、大規模な実データに対する実証実験を通してその有効性を確認する。また、深層学習に対する実験的・数理的な解釈を伴い、マルチメディア情報処理の基盤技術の研究も並行する。

英文抄録(100 words 程度)

We aim to achieve high performance by recurrent neural network, which is suitable for time series of multimedia information, speech and video. Research and development of deep neural network-based systems have been significantly advanced in the last five years. In this research project, we develop a novel system combination method which manages heterogeneous image, speech and text information.

*Keywords:* multimedia data, time series, system combination, recurrent neural network

背景と目的

本プロジェクトでは、社会に埋め込まれる様々な高精度センサから得られる、映像、音響、音声を含むマルチモーダルデータの処理を深層回路網によって行った。

った。また、深層学習に対する実験的・数理的な解釈を伴い、マルチメディア情報処理の基盤技術の研究も並行した。

概要

社会に埋め込まれる様々な高精度センサから得られる、映像、音響、音声を含むマルチモーダルデータを統合的に処理できる深層学習処理基盤の構築を行う。ドライブレコーダや監視カメラ等から所得される映像、音響、音声のデータの収集及び整備、大規模スーパーコンピュータを用いた個別データにおける深層学習処理の高性能化を進める。映像、音響、音声等の異種マルチモーダルデータを統合的に処理できる深層学習処理基盤の構築を行い、評価実験に向けて実フィールド、実データの整備を進める。

まとめ、今後の課題

大規模マルチメディアデータベースの特徴抽出や深層学習を実施した。二か月の短期間であるため、まだ顕著な成果は得られなかったが、設定などの工夫を含めた進捗からこれからの成果を期待する。

結果および考察

研究では、大規模な実データに対する実証実験を行



TSUBAME 共同利用 平成 29 年度 学術利用 成果報告書

利用課題名 大規模分散深層学習の性能調査  
 英文: Performance Studies for Large-scale Distributed Deep Learning

利用課題責任者 松岡 聡  
 First name Surname Satoshi Matsuoka

所属 産業技術総合研究所 RWBC-OIL  
 Affiliation National Institute of Advanced Industrial Science and Technology, RWBC-OIL  
 URL <https://unit.aist.go.jp/rwbc-oil/>

邦文抄録(300 字程度) 将来のスーパーコンピュータでの最重要カーネルである分散深層学習を対象に、大規模環境下での性能評価を ImageNet1k のデータセットに対して行った。実行の結果、ネットワーク通信など性能上のボトルネックなどが明らかになった。

英文抄録(100 words 程度) We conducted performance studies for distributed deep learning to ImageNet1K datasets. The results basically show good scalability on TSUBAME3, however, we observe performance degradation derived from network interconnects.

*Keywords:* HPC, AI, Big Data, Deep Learning

#### 背景と目的

将来のスーパーコンピュータでの最重要カーネルである分散深層学習を対象に、大規模環境下での学習の際のアルゴリズムやハイパーパラメータの設定の精度への影響の調査、ネットワーク通信やストレージ I/O などシステムに対するバンド幅や遅延などの性能要求の調査を行うことで、将来の AI/ビッグデータ処理に特化したクラウド上での実社会ビッグデータを活用するアプリケーションの実行やシステムソフトウェアの要素技術の要求要件・設計情報となることを目指す。

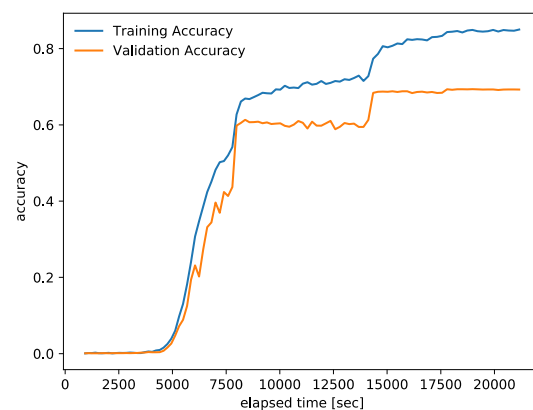
#### 概要

現在広く行われている ImageNet 1K(1000 カテゴリ・100 万枚の画像)の学習において、これまで、NVIDIA Tesla K20 GPU で 21 日程度かかる処理が並列化を含むアルゴリズムの改善や実装技術の向上により、NVIDIA Tesla P100GPU を 1024 台用いることで 15 分程度に大幅に削減できることが知られている。しかし、分散深層学習を幅広く汎用的に応用し高速化するためには、様々な学習シナリオでのアルゴリズムやハイパーパラメータの評価、システムに対する性能要求の明確化が必須となる。そこで、我々は、ImageNet 1K のデータセットに対して ChainerMN で

分散深層学習を用いて性能を計測した。学習の際のアルゴリズムは、Goyal et al. 「Accurate, Large Minibatch SGD: Training ImageNet in 1 Hour」(<https://arxiv.org/abs/1706.02677>)と同等のものをを用いた。

#### 結果および考察

実行の結果、TSUBAME3.0 の 64 ノード、256GPU を用いることで、100 epoch の学習に 21152 秒 (5.87 時間)を要し、OmniPath を用いた際の通信の最適化が必要であることを確認した。



#### まとめ、今後の課題

今後の課題としては、他のシステムでの性能評価などが挙げられる。

# 東京工業大学 TSUBAME 共同利用 平成 29 年度利用終了課題 利用成果報告書

発行 : 令和元年 6 月

国立大学法人 東京工業大学 学術国際情報センター 共同利用推進室

住所 : 〒152-8550 東京都目黒区大岡山 2-12-1 E2-6

E-mail : [kyoyo@gsic.titech.ac.jp](mailto:kyoyo@gsic.titech.ac.jp)

URL : <https://www.gsic.titech.ac.jp/tsubame>

本書に記載の記事・写真等の二次利用を禁じます。これらの情報は著作権法上認められた「私的利用」または「引用」の条件をみたした場合を除いて、著作権者に無断で転載、複製、放送、公衆送信、翻訳、販売、貸与等の利用を禁じます。