

ポスト「京」重点課題 2

個別化・予防医療を支援する統合計算生命科学

NEWS LETTER

Vol.

7

Contents

• Research Report

何してるの?なぜ?どうやって?ちょっと解りやすく
教えて!に研究者が応える

"生きている"状態での細胞の動きと
形状の変化の検出法の開発

• Information

お知らせやイベント情報

INTEGRATED
COMPUTATIONAL
LIFE
SCIENCE
TO SUPPORT
PERSONALIZED AND
PREVENTIVE MEDICINE



ポスト「京」
重点課題2
個別化・予防医療を支援
する統合計算生命科学



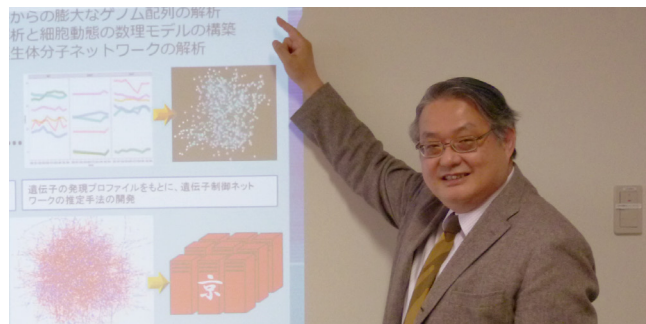
■ Research Report

Subtheme **B**

"生きている"状態での細胞の動きと形状の変化の検出法の開発

サブ課題B

大阪大学大学院情報科学研究科 松田秀雄



顕微鏡による生態観察技術の進展により、生物の体内での微細な活動の様子が、“生きている”状態で観察できるようになりました。これにより、今まで誰も見たことがなかった細胞等の活動の様子をリアルタイムで見ることができるようになり、大量の動画像データが産生されています。しかし、そのような大規模画像データの解析には、従来の物体認識のための画像処理技術がそのままでは使えず、新たな解析技術の開発が求められています。本稿では、“生きている”状態での細胞の活動、特に動きと形状の変化を「京」を使って解析した結果について紹介します。

生体内の観察技術の進展

「未だかつて誰も見たことのないものを見てみたい」のは生体内の観察技術の主たる動機付けとなってきました。一つの方向性は「より小さいものを見たい」というもので、光学顕微鏡、電子顕微鏡に続いて、2014年のノーベル化学賞の受賞テーマとなった超解像顕微鏡に向かう進展があります。別の方向性としては「生きている状態の体の中を見てみたい」というもので、多光子励起顕微鏡 [1] の出現により可能となったものです。

光学顕微鏡では、注目する対象からの光を観察します。さらに、蛍光顕微鏡では観察対象を蛍光分子で標識し、光（光子）をあてることで蛍光分子がエネルギーを吸収して励起状態になった後、励起前の元の状態に戻るときに放出される光を観察するものです。これらに対して、多光子励起顕微鏡では、光子1個ではなく複数（通常は2個）の光子を蛍光分子に同時に当てることにより励起させます。この現象は光子密度が極めて高い点のみでしか起こらず、それ以外の箇所では励起が起こらないため、観察対象のみに焦点を設定して観察することが可能となり

ます。さらに、エネルギーの低い（つまり、波長が長くて浸透性が高い）光子を複数個（通常は2個）集めて高いエネルギーにして励起できるため、深部組織に焦点をあてることで通常の顕微鏡では見えないような組織内部を観察することができます。また、照射光のエネルギーを低くできるため体に与える影響が抑えられる点も観察に有利となります。

多光子励起顕微鏡により、“生きている”状態での体の奥の微細な細胞などを観察することが可能となりました。“生きている”状態とは、血流など生体の環境が維持されている状態を指し、このような条件での顕微鏡による画像取得は intravital imaging（決まった訳語はまだありませんが直訳すれば、生きている体内の画像化）[2] と呼ばれ、ここ数年の間に急速に注目を集めています。

バイオイメージ・インフォマティクスの形成

顕微鏡により生体から取得した画像データの処理には特有の難しさがあります。画像処理で広く行われている物体認識を例にとると、従来の方法では、対象

となる物体の特徴情報（例えば、人の顔であれば瞳、鼻下、口角などの特徴点）を利用して、まず、画像中で特徴情報をもとに対象となる物体の位置の候補点を求めて、その物体のモデル（人の顔であれば多数の顔画像が登録されたデータベースを使ったりします）との照合を行います。しかし、生体内の観察対象は千差万別で、かつ特徴も十分にわかっておらず、蛍光以外の特徴情報をあらかじめ決めることは困難です。蛍光についても、自家蛍光と呼ばれる意図しない発光が観察されることがあり、1枚の静止画像を見ただけでは何が映っているのかわからないこともしばしばです。このような観点から、生体内の微細な対象の画像データを解析して、そこから新たな知見を得ようとする学問領域である「バイオイメージ・インフォマティクス」の重要性が指摘されています [3]。

例えば、経時的に観察した動画像から細胞などの移動を検出する方法を考えます。前述のように細胞には、人の顔の持つ瞳、鼻下、口角のような特徴点がなく、単に蛍光で光っている領域としか見えなことがほとんどです。このように、物体の特徴情報が使えない状況でそれら

■ Research Report

Subtheme **B**

の動きを求める方法の一つに、オプティカルフロー検出法があります。オプティカルフローとは物体やカメラの移動によって生じる、時間的に隣接するフレーム（動画像を静止画像のコマ落としとみなしたときの1枚分の画像）間での物体の位置のずれを、あるフレームから次のフレームへの変位ベクトルで表します。つまり与えられた個々の画像で物体を認識するのではなく、隣接するフレームで位置が変化している（つまり移動している）領域を「物体」とみなすわけです。ただ、位置が変化している領域というだけでは一意に特定できないため、オプティカルフロー検出法では、以下のような仮定を設定しています。

- 1) 画素が移動しても輝度（画素値）は不変
- 2) 画像は滑らか（空間的・時間的に微分可能）
- 3) 画素の移動量は小さい

このような仮定の下に、原理的には1画素単位で、移動している領域の変位ベクトルをオプティカルフローとして検出するのですが、実際の画像ではどの仮定も完全に満たされるわけではなく、例えば、個々の画素の輝度はフレームごとにある程度変化することがあり、また近傍の画素が同じ輝度を持つこともあります。

そこで、オプティカルフロー検出法では、1画素単位ではなく、縦横数画素で構成される領域単位で相互に輝度を比較し、輝度が類似して位置も近い領域どうしを対応付けて、領域間の変位ベクトルをオプティカルフローとして求めるのが一般的です。しかし、画素を領域に拡大したことにより、領域内のすべての画素の移動を一つの変位ベクトルで表すため、

- 4) 領域内のすべての画素は同じ移動をする（平行移動のみで回転や変形はしない）

という仮定も加わります。

深層マッチングによる細胞の動きの検出

前項のオプティカルフロー検出法では、動画像中の物体を追跡するのに、時間的に隣接するフレーム間において物体の位置や形状が微小にしか変化しないと仮定し、局所的な探索により物体の領域どうしのマッチングを行っていました。しかし、この仮定はある種の細胞では成り立ちません。

例えば、マウスの血管内の免疫細胞は、血流に乗る細胞と血管の内皮に接着する細胞で動きが大きく異なるため、オプティカルフロー検出法では変位ベクトルの検出が容易ではありません。図1の(a)と(b)に実際にマウスの血管内の免疫細胞（白血球の一種である好中球）を二光子励起顕微鏡で観察した画像を示します。(a),(b)は時間的に隣接するフレームの画像で、赤い領域が血管、緑の領域が免疫細胞を表しています。(a)と(b)を比べると、細胞によって動きにかなりばらつきがあり、特に黄色の丸で囲まれた細胞は大きく動いていることがわかります。

本研究では、細胞の動きを検出するのに深層マッチング(deep matching)[4]を用いています。深層マッチングは、コンピュータビジョンの分野で開発された動画像中での物体の追跡手法で、深層学習(deep learning)の代表的な手法である畳み込みニューラルネットワークに着想を得ています。具体的には、次のような手順で隣接フレームでの領域の対応付けを行います。

① 隣接するフレーム間で、フレームをそれぞれ小領域（例えば、8画素×8画素の領域）に分割し、フレーム*i*の小領域と、フレーム*i+1*の小領域どうしのすべての可能な組み合わせごとに畳み込み演算を実行

② ①の演算結果から類似度が極大となる小領域のペアを局所的に探索し、さらに領域を縦横それぞれ2倍（16×16画素）に拡大して、マッチング結果のマップを作成

③ ①と②を領域の縦横のサイズを倍々に上げながら繰り返して、多階層の局所最適なマッチング結果のマップを小領域から順に領域のサイズが画像全体になるまで拡大

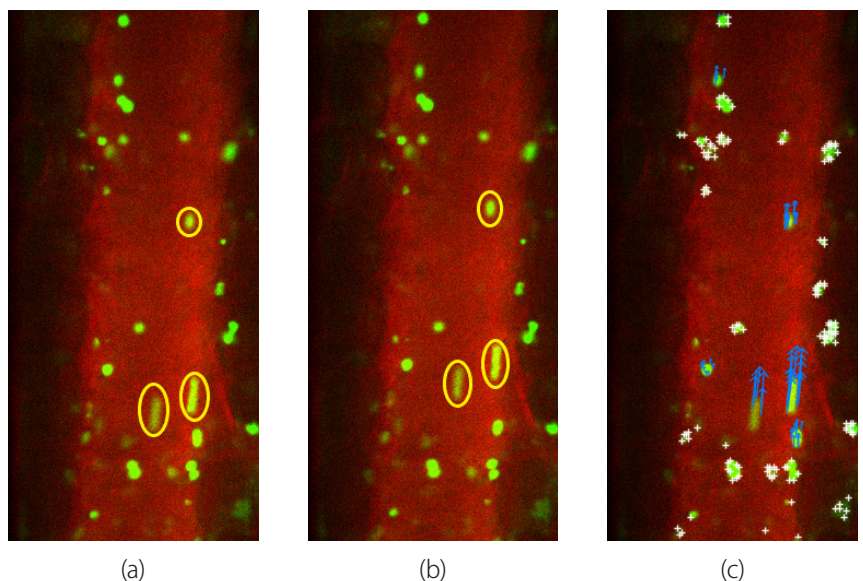


図1: マウスの血管内の免疫細胞を二光子励起顕微鏡で観察した画像（画像提供：大阪大学医学系研究科 石井優教授）

■ Research Report

Subtheme

B

④ 逆に全体画像から小さな領域に向けて、局所最適なマッチング結果を探索しながらつなげて、画像全域に渡るマッチングにより対応付けられた領域ペアを検出

領域間の各画素の輝度を直接比較するのではなく、領域のサイズを変えながら畳み込み演算で類似性を判定するため、物体の変形や急激な動きに強い方法となっています。

本研究では、この深層マッチングを細胞の動きの検出に利用しています。図 1(a),(b) の 2 枚の画像に対して深層マッチングを行った結果を図 1(c) に示します。図 1(c) で、青い矢印は深層マッチングにより細胞の移動を検出した変位ベクトル、白い十字が移動しない細胞を検出したものとなっています。図 1(a),(b) で目視で特に動きが大きいと見られた細胞（黄色の丸で囲まれた細胞）から大きな変位ベクトルが検出されており、それ以外の細胞では一部に微小な青矢印、他は白い十字で動きが検出されなかった細胞を表しています。血流内の免疫細胞（白血球）は、血流に乗って移動する、血管内皮に接触してローリングと呼ばれる転がるような動きをする、血管内皮に接着する、血管内皮に浸潤して遊出する、などの動きを示します [6]。深層マッチングによりそれ

ぞれの細胞の動きを定量的に識別することができると考えられます [7]。

脂肪組織内での免疫細胞の動きと形状変化の検出

体重の増加により肥満状態が進むと、やがてメタボリックシンドロームに陥りますが、これは体型の問題ではなく、体内で起こる現象が問題と考えられています。肥満状態が進むと、体を守るはずの免疫細胞が“暴走”し始めると言われています。「免疫細胞の暴走」とは、メタボの人の脂肪細胞がある種の信号伝達物質を放出し、それを受け取った免疫細胞が活性化して、自らも同じ物質を放出することで免疫細胞の活性化が広がっていくことだと考えられています [8]。

実際に、マウスに高脂肪食を与えたときに脂肪組織内での免疫細胞（マクロファージ）の活動を二光子励起顕微鏡で観察した例を図 2 に示します [8]。図 2(a) は観察した画像のフレームの一つで、青い領域が脂肪細胞で、緑色に光るのが免疫細胞、赤い筋が微小な血管を表しています。(b) および (c) は、(a) の黄色い四角で囲まれた領域を、隣接したフレームからそれぞれ抽出した画像です。(b) と (c) を見ると、図 1 とは違って細胞の形状が丸くなく、非常に伸びた形をしていま

すが、これは脂肪細胞と脂肪細胞の間隙間にはまりこんでいると思われます（特に大きな変形をしている細胞を黄色で囲っています）。高脂肪食ではなく通常の食事を与えたマウスでは、免疫細胞は動かずにじっとしていますが、高脂肪食を与えてしばらくすると活発に動き出し複雑に形状を変えていきます。

図 2(b) と (c) に対して深層マッチングを適用した結果を (d) に示します。(d) では (b),(c) の黄色で囲った領域周辺部分を拡大して表示しています。深層マッチングにより、細胞の動きだけでなく、形状の変化も変位ベクトル（白い矢印）の大きさとして検出されていることがわかります。このように脂肪組織中の免疫細胞の変化を二光子励起顕微鏡で観察し、深層マッチングを適用することで、脂肪組織での免疫細胞の活性化のレベル、すなわちメタボの進行度合いが定量的に求められることが期待されます。

「京」での深層マッチングの計算

深層マッチングのプログラムは C++ 言語でコーディングされており、それを「京」に移植しました。畳み込み演算の部分はスレッド並列版の BLAS（富士通社の数

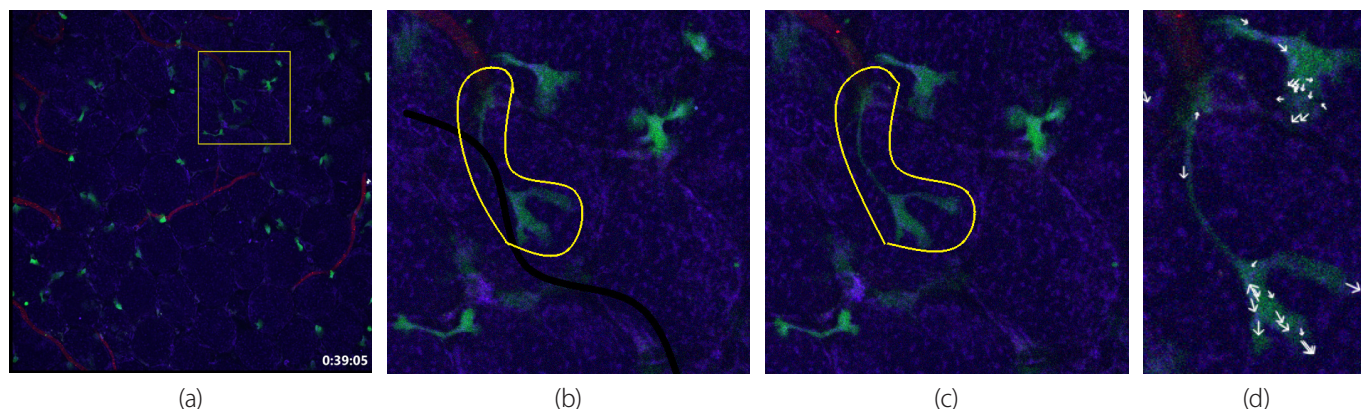


図2: 高脂肪食を与え続けたマウスの体内で活動している免疫細胞（画像提供：大阪大学医学系研究科 石井優教授、水野紘樹助教）

■ Research Report

Subtheme **B**

値計算ライブラリ中のSSL2BLAMP)を使用し、さらにOpenMPによりスレッド並列を実装しています。また、東京工業大学の秋山研究室で開発されたMPIDPを使用して、経時観察した動画像を100フレーム程度のブロックに分割して、各ブロックの深層マッチングをMPIによるノード並列で処理するようにしており、スレッド並列とノード並列を組み合わせたハイブリッド並列処理となっています。

「京」での実行時間は、1ノード8スレッド並列で実行したとき、 512×512 画素の画像で隣接2フレーム間の1回のマッチングを計算するのに約13.6秒かかりました。動画像のフレームを「京」の各ノードに分配してMPIによるノード並列処理を実装して、フレーム数を増やしていくことで「京」の全ノードを使った並列処理を達成できました。1ノード当たりマッチングを100回実行するように固定したWeak Scalingで、「京」全ノードを使った並列処理では約829万フレームという膨大な量のマッチングを約41分で実行

できました。マッチング1回あたりの処理時間は約24.6秒となっています。

利用している二光子励起顕微鏡では30フレーム/秒で画像を取得できますので、「京」の約738ノード($=30 \times 24.6$)を使えば、画像の取得速度と同程度の処理速度で深層マッチングの計算ができることになります。実際には、画像ファイルの「京」への転送に時間がかかるため、計算前に「京」に移しておく必要があります。免疫細胞の「暴走」は、病態としては炎症という形で現れますが、炎症の進行は急性で3～4時間以上持続するため、フレーム数にすると急性で32.4万～42.3万フレームの計算が必要です。さらに慢性の炎症は1週間以上かけて徐々に進行するとされています。常に30フレーム/秒で観察する必要はないと思われますが、一定の間隔で週単位の時間をかけて画像を取得する必要があり、膨大な量のフレームを解析する必要があります。また、本稿では2次元の画像しか示していませんが、二光子励起顕微鏡はXYの

2次元だけでなく深さ方向のZ軸で焦点を設定できるため、得られる画像は3次元+時間軸の4次元画像となります。3次元画像では2次元画像と比べて膨大な数のフレームが生成されるため、より大きな計算資源が必要となります。さらに、1例だけの解析ではなく、将来的に数千から数万例のコホートで処理することが考えられます。以上を合わせて考えると、ポスト「京」規模の計算資源の利用が必要になると考えられます。今後、intravital imagingの画像データのスパコンによる解析が進んでいくのではないかと期待されます。

<謝辞>

二光子励起顕微鏡の画像データをご提供頂いた、大阪大学医学系研究科免疫細胞生物学教室の石井優教授、水野紘樹助教に深く感謝いたします。

<参考文献>

- [1] Winfried Denk, James H. Strickler, Watt W. Webb, Two-photon laser scanning fluorescence microscopy, *Science*, 248, pp.73-76, 1990.
- [2] Mikael J. Pittet, Ralph Weissleder, Intravital imaging, *Cell*, 147(5), pp.983-991, 2011.
- [3] Gene Myers, Why bioimage informatics matters, *Nature Methods*, 9(7), pp.659-660, 2012.
- [4] Philippe Weinzaepfel, Jerome Revaud, Zaid Harchaoui, Cordelia Schmid, DeepFlow: Large displacement optical flow with deep matching, *IEEE International Conference on Computer Vision (ICCV)*, 2013.
- [5] Jerome Revaud, Philippe Weinzaepfel, Zaid Harchaoui, Cordelia Schmid, DeepMatching: Hierarchical deformable dense matching, *International Journal of Computer Vision*, 120(3), pp.300-323, 2016.
- [6] 理科年表オフィシャルサイト「白血球の血管外遊出にかかわる白血球と血管内皮細胞上の接着分子」
https://www.rikanenpyo.jp/kaisetsu/seibutsu/sei_003.html
- [7] Hideo Matsuda, Hironori Shigeta, Shigeto Seno, Junichi Kikuta, Masaru Ishii, Analyzing cell migration dynamics from intravital imaging by deformable image matching, *International Conference on Intelligent Systems for Molecular Biology and European Conference on Computational Biology (ISMB/ECCB 2017)*, B-383, 2017.
- [8] NHKスペシャル「人体」第2集「脂肪と筋肉」(2017年11月15日放映)
<https://www.nhk.or.jp/kenko/jintai/programs/2>

■ Information

News & Events

開催報告：ポスト「京」重点課題2ワークショップ

日程：12月4日(月) 10:00～16:50

場所：東京大学医科学研究所

ポスト「京」重点課題2は、プロジェクト参加機関の研究開発について理解を深めるとともに情報交換の場として、全体ワークショップを毎年1回開催しています。第2回を迎えた今回は、ポスト「京」重点課題2に参加している研究者ら約40名が一堂に集い、17名の研究者が取り組んでいる課題の進捗について発表を行いました。発表後には総括として諮問委員により改善提言や評価がなされました。その後、サブ課題責任者により、当総括と文部科学省からの中間評価結果の内容を踏まえ、成果の創出、研究開発の一層の発展、また課題間連携の推進へ向けた今後の取り組み体制と方向性等について、さらなる意見交換が行われました。この内容は実施計画に反映していく予定です。



今年も笑顔でバチリ

参加報告：SC17

2017年11月12日から17日にかけて、SC17が米国コロラド州デンバーで開催されました。この国際会議はHPC（ハイパフォーマンス・コンピューティング）分野におけるトップカンファレンスの一つであり、毎年6月頃にドイツで開催されるISC（International Supercomputing Conference）とともにTOP500が発表される場としても有名です。

今年も世界各国から多くの研究者や企業が参加し、活発な議論が交わされました。2017年の総括となるTOP500等の発表においては、理化学研究所計算科学研究機構のスーパーコンピュータ「京」がベンチマーク「HPCG」および「Graph500」において前回に引き続き世界1位を獲得しました。また、昨今注目されている対消費電力性能ランキングのGreen500では、上位TOP3を含む7機の日本製スパコンがTOP10ランク入りし素晴らしい結果を残しました。



会場のコロラド・コンベンション・センター



TOP500発表直前のメインホールの様子



文部科学省 ポスト「京」開発事業

重点的に取り組むべき社会的・科学的課題に関するアプリケーション開発・研究開発
重点課題2 個別化・予防医療を支援する統合計算生命科学

Integrated Computational Life Science to Support Personalized and Preventive Medicine

■ 問い合わせ先

国立大学法人東京大学医科学研究所 ヒトゲノム解析センター DNA情報解析分野
ポスト「京」重点課題2 個別化・予防医療を支援する統合計算生命科学 事務局

〒108-8639 東京都港区白金台 4-6-1 TEL：03-5449-5615 FAX：03-5449-5442

E-mail：icls-office@hgc.jp URL：http://postk.hgc.jp/



ポスト「京」重点課題は、国家基盤技術としてスーパーコンピュータ「京」の後継機となるポスト「京」を活用し、国家的に解決を目指す社会的・科学的課題に戦略的に取り組み、世界を先導する成果の創出を目指す文部科学省の事業です。重点課題2「個別化・予防医療を支援する統合計算生命科学」は、東京大学医科学研究所を代表機関として、ポスト「京」によって初めて実現できる「情報の技術」、「物理の原理の応用」、そして「ビッグデータの活用」により、病態の理解と効果的な治療の探索法の研究を行い、その成果を個別化・予防医療へ返す支援基盤となる統合計算生命科学を確立することを目的としています。