

課題名 (タイトル) :

領域抽出法の性能評価に関する研究

利用者氏名 :

竹本 智子

所属 :

本所 知的財産戦略センター V C A D システム研究プログラム 生物基盤構築チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

(背景)

画像の領域抽出 (Image segmentation) は、過去何十年にも渡り膨大な成果発表がなされているにも関わらず、いまだに Computer Vision 研究のチャレンジングな課題の一つである。領域抽出は定量解析の第一歩であることから Computer Vision に限らず、生物学・医療分野などにおいて、ますます重要性が高まっている。現象の定量化を行うことは、自然科学研究に欠かすことはできない。

領域抽出が難しい理由には、アルゴリズムの汎用性欠如が挙げられる。高精度での領域抽出を望むならば、アルゴリズムは各アプリケーションに特化される必要があり一般性を欠く。結果として研究者は、アプリケーションごとにアルゴリズムのチューニングや検討を行う必要があり、定量解析の大きな妨げになっている。

(目的)

前述の問題を解決するため、本研究では領域抽出アルゴリズムの性能評価基準を定義し、各アプリケーションに対するアルゴリズムの評価システムについての研究・開発を行っている。システムは、ユーザー (領域抽出を必要とする人) によって示された Ground-truth 領域と、あらかじめシステムに用意されているアルゴリズム群による自動抽出領域との類似性を評価し、アルゴリズムをランク付けするものである。ランクが良いアルゴリズムは、Ground-truth 領域と類似した画像特徴の新たな入力画像に対しても、高精度で領域抽出を実現することができる。

(関係プロジェクト)

利用者は、平成 20～24 年度、文部科学省科学研究費補助金・新学術領域研究「細胞内ロジスティクス」の研究分担者である。この研究は、主に共焦点レーザー顕微鏡で撮影した細胞の時空間画像を対象としている。ロジスティクスとは「物流」を意味し、細胞内に生じる細かな物流現象を解析することによって、細胞内生命現象の定量化を目指している。またこの研究は、細胞生物学・分子生物学・生物化学・情報科学のはじめての本格的融合研究であり、新しい学術領域確立のさきがけとして研究成果が各方面から大いに期待されている。

2. 具体的な利用内容、計算方法

本研究で取り扱う画像は最大で 512 x 512 x 100 voxel 程度の高分解能ボクセル (のタイムラプスデータ) である。また、領域抽出や評価計算は各 voxel のスカラー値のみを対象とするのではなく、輝度由来の高次元ベクトルを対象とするため、計算データは数 GB～数十 GB のデータ量になる。評価計算ではこれらの入力データに対し、複数のアルゴリズム (及び大量のパラメータセッティング) による計算が実行される。従って、本研究では RICC の超並列 PC クラスタによるジョブ並列を行っている。

なお本研究では、開発した評価システムの妥当性を示すため、現在 Computer Vision 研究において state-of-the-art とされる領域抽出アルゴリズムの性能を評価する必要がある。今年度は抽出アルゴリズムとしてまず、Support Vector Machine 及び Approximate Nearest Neighbor をシステムに実装した。また評価計算には交差検定法 (Cross-validation) を用いた。

具体的にはまず、ユーザーによって与えられた Ground-truth 領域の元画像に対し、各アルゴリズムによる領域抽出を実施する。その後、各抽出結果を学習データとテストデータに分割し、Ground-truth を用いた誤差検定を行うことでアルゴリズムの識別性能をランク付けする。このとき、学習データとテストデータの分割数が多くなると計算コストが高くなることは良く知られている。この方法はすなわちモデル汎化性概念に基づいており、ランクが高いアルゴリズムは、Ground-truth と似た画像特徴を持つ対象の領域抽出に高い抽出精度を示すことが保障される。評価計算は最低でも、アルゴリズム数及び各アルゴリズムに関わるパラメータ数の総当たり回数が実行される。さらに、例えば Leave-one-out Cross-validation を行った場合には、上記の総当たり回数に対してさらにボクセル数倍の計算量が必要となる。以上の計算について本システムでは、各計算を各ノードに分散し、一定時間ごとにログインノードから計算終了についてリクエストし、終了している場合は計算結果をログインノードに回収している。

3. 結果

利用者は、平成 21 年 6 月までに上記システムのプロトタイプの開発、システムのテスト・性能評価、簡易実験について、RSCC を利用して実施していた。その

後、RICC への変更に伴い本研究のシステムに一部変更を行う必要が生じた。そのため平成 21 年 7 月～10 月に変更を行い、10 月～11 月に変更後システムのテストを実施した。その結果、データによっては各計算ノードにジョブを投入した際、1 ジョブが 10 分を超えるパラメータセッティングの計算では、ログインノードに正しく計算結果が返ってこないバグが発覚し、現在もその原因を調査中である。そのため、以下に列挙した研究結果についての結果の多くは、RSCC を使用した結果を用いたものである。

- ✓ 細胞内画像について、ユーザーの注目領域を精度良く抽出するアルゴリズムを選択するための評価関数を定義した。
- ✓ 評価関数が実画像実験において再現性、安定性を獲得するものであることを確認した。
- ✓ ユーザーの要求を反映する評価関数に、画像領域の「形状」及び「テクスチャ」に基づく距離関数を導入することが有効であることを示した。

4. まとめ

利用者は、領域抽出アルゴリズムの性能評価基準を定義し、各アプリケーションに対するアルゴリズムの抽出精度の評価システムについて、RICC を利用した研究・開発・テストを行った。

評価システムはそれぞれ独立した計算を大量に実施する必要があり、利用者は RICC の超並列 PC クラスタを利用したジョブ並列を行った。その結果、システムに実装されている抽出アルゴリズムの評価関数に再現性があることを確認し、一部の計算結果を国際学会、国内発表等で発表した。

なお、平成 21 年度は RSCC から RICC への移行に伴い、評価システムの一部変更が必要であった。現時点で一部の計算にバグが発生しており、現在その原因について調査中である。

5. 今後の計画・展望

(今後の計画)

先に述べたように、一部の計算について、1 ジョブの計算時間制限を超えていないにも関わらず、計算結果がログインノードに正しく返ってこないバグについての原因調査・修正を行う予定である。その後、現在実装されている 2 つの抽出アルゴリズムに加え、Graph Cut や Mean Shift などを実装した評価システムを用いて、抽出性能の評価計算を実施する。本システムは将来的に数千ノードを用いたジョブ並列を予定していることから、今後は一般利用申請に向けたデータ収集も行う予定である。

(研究の展望)

細胞内画像 (Intracellular images) を対象としたアルゴリズム性能の評価計算を行い、細胞内の「輸送」や「形状変化」など、現象ごとに適したアルゴリズムの発見を目指した実験分析を行う。細胞内画像に

限らず、実際の観測画像を対象とした領域抽出アルゴリズムの評価及び適用指針を示した例は少なく、研究成果は Computer Vision や定量解析を必要とする自然科学の研究にとって極めて重要なものとなると考えている。

6. RICC の継続利用を希望の場合は、これまで利用した状況 (どの程度研究が進んだか、研究においてどこまで計算出来て、何が出来ていないか) や、継続して利用する際に行う具体的な内容 (研究進捗)

3 の研究成果に記載のように、細胞内画像を対象とした領域抽出について、ユーザーの要求を満たすアルゴリズムの評価関数を定義した。RICC での計算は、主に定義した評価関数の安定性・再現性を確認するためのものであった。なお、計算時間が 10 分を超える一部の計算については正しい結果を得られないことがあり、現在はその原因について調査中である。

(継続利用の場合の具体的な内容)

現在発生しているシステムバグについて、原因の調査・修正を行う。その後、本システムは将来的に数千ノードを用いたジョブ並列を予定していることから、一般利用申請に向けたデータ収集を行う。

また、利用者は平成 22 年 4 月より現評価システムの拡張システムを稼働予定である。上記に並行して、拡張システムのテスト・評価を実施する予定である。

平成 21 年度 RICC 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

Satoko Takemoto, Hideo Yokota, Taketoshi Mishima and Ryutaro Himeno, "Segmentation of Anatomical Structure by Using a Local Classifier Derived from Neighborhood Information," Human-Computer Systems Interaction, Series: Advances in Intelligent and Soft Computing, Vol. 60, pp.171-180, Springer-Verlag, Berlin Heidelberg, 2009.

【国際会議などの予稿集、proceeding】

Satoko Takemoto and Hideo Yokota, "Algorithm Selection for Intracellular Image Segmentation based on Region Similarity", International Conference on Intelligent Systems Design and Applications, pp. 1413-1418, November 30-December 2, 2009, Pisa, Italy.

【国際会議、学会などでの口頭発表】

Satoko Takemoto, Hideo Yokota, "Parallelized Performance of Image Segmentation Algorithm by using the Linux Cluster on RICC," Super Computing 2009, Poster Presentation, November 14-20, 2009, Portland, Oregon, USA

Satoko Takemoto, "Algorithm Selection for Intracellular Image Segmentation," 2nd Joint Workshop on Computational Science, July 9-10, 2009, RIKEN Yokohama Institute, Yokohama, Japan.

竹本智子, 横田秀夫, “ダイヤモンドを考慮した生物・医用画像の領域抽出システム,” Biomedical Interface Workshop, 宮古島, 2010 年 3 月, 沖縄

竹本智子, 横田秀夫, 牧野内昭武, "細胞内画像のための領域抽出アルゴリズムの性能評価基準," 精密工学会 2010 年春季大会, 3 月, 埼玉 (発表予定)

