

審査の結果の要旨

氏名 西村 邦裕

ゲノムサイエンス分野において、膨大な情報をいかに取り扱うかが問題となっている。急速に蓄積されている生命機能情報から新たな知見を見出すには、情報の把握と理解、そして試行錯誤が必要とされている。

本論文では、その膨大なゲノム情報の解析支援として、可視化によるアプローチをとり、インタラクティブ可視化環境を実現することを提案している。具体的には、個々の関係性を見るために、必要な情報を絵とする「試行の可視化」を提案し、その「試行の可視化」された絵同士の関係や作業の履歴などを絵にしてみせるメタレベルの概念として「履歴の可視化」を提案している。さらに、それらを総合的に統合したものを「プロセスの可視化」、実装された環境を「インタラクティブ可視化環境」として提案している。

第1章では、本研究の背景について述べたあとで、本研究の目的と意義について述べている。研究の背景として、ゲノムサイエンス分野の扱う情報の急増について取り上げた後、バーチャルリアリティ技術やヒューマンインタフェース技術、情報の可視化技術の発展により、情報をわかりやすくインタラクティブに提示することが可能になったことを述べている。そして本論文の目的をゲノム情報解析における可視化について論じることと位置づけている。また、最後に本論文の構成について述べている。

第2章では、解析支援のための可視化論として、まず、従来の情報の可視化手法について紹介している。次に、解析支援のための可視化について述べている。解析において、解析時点だけではなく、その時点に至るまでの過程も重要であることに着目し、一つ一つの解析の結果を絵にする可視化の基本動作としての「試行の可視化」と、その絵同士の関係や作業過程、作業履歴を絵にしてみせる「履歴の可視化」について提案している。また、履歴についても分類を行い、「履歴の可視化」手法についての検討を行っている。最後に、「試行の可視化」とそのメタレベルに位置する「履歴の可視化」を総合的に位置づけた「プロセスの可視化」を提案し、それを実装した「インタラクティブ可視化環境」について論じている。

第3章では、ゲノム情報、および、本論文で主に扱うゲノムコピー数情報について紹介している。まず、ゲノム情報解析について整理し、ゲノム情報の特徴、解析の意義、解析への要求事項を述べた後、ゲノムサイエンスの研究者の思考方法を考慮に入れたゲノ

ム情報解析支援のための可視化手法について論じている。その後、ゲノムコピー数取得方法や算出アルゴリズム、ゲノム情報の可視化の関連研究について述べている。

第4章では、第2章で提案した「試行の可視化」および「履歴の可視化」を実現するための可視化手法について論じている。対象を、ヒトゲノムのコピー数情報とし、コピー数異常領域を検出するための可視化手法について述べている。まず、ゲノムコピー数情報の処理について述べたあと、「試行の可視化」としてデータを閲覧するための可視化手法、データの異常・特徴を抽出・比較するための可視化手法、データにアノテーション付けをするための可視化手法について提案している。次に、「履歴の可視化」として、時間軸の利用、作業軸の利用、空間軸の利用、軸を利用しない、の4つにわけて可視化手法をそれぞれ提案している。

第5章では、第4章で実装したシステムについての評価を行っている。まず、データ異常の検出について、正答率と時間から評価し、次に、履歴を利用した場合の作業効率について評価を行っている。その結果、本論文が提案する可視化により、十分な異常値検出力があることが確認できたとともに、作業効率が上がることが示唆され、作業に対する理解や把握が深まることが示されている。

第6章では、本論文で提案する可視化手法を用いて、実際のゲノム解析に応用した結果が述べられている。肺がんにおけるコピー数解析と、正常の人における集団でのコピー数変化 CNP(Copy Number Polymorphism)について解析例を示し、本論文の提案する手法が実用的であることを示している。

第7章では、本論文を統括して研究成果をまとめるとともに、今後の課題と展望に触れ、論文全体の結論としている。

筆者によって提案された「プロセスの可視化」は、従来の可視化と比較して、時間的方向と意味論を含んだ内容となっており、新規的な概念である。本論文で提案したさまざまな可視化手法から有効な知見も得ている。そして、複雑な作業を行う場合の知的作業支援として、本論文で提案するインタラクティブ可視化環境は有効であると期待できる。また、本論文で提案する支援環境により、ゲノム情報解析がより進むと考えられる。

よって本論文は博士(工学)の学位請求論文として合格と認められる。