

システム情報工学研究科修士論文概要

年 度	平成 26 年度	学位名		修士(工学)
専 攻	知能機能システム	専攻	著者氏名	曾我部 陽光
指導教員氏名 丸山 勉				
論文題目				
FPGA を用いたショートリードマッピングの高速化				
論文概要				
<p>DNA シーケンシングとは、DNA の塩基配列を読み取る技術であり、生物学や医療などに幅広く利用される。次世代シーケンスマシン(NGS)の登場により、劇的に低コストで高速な DNA シーケンス解析が可能となった。NGS は単に、ショートリードと呼ばれる DNA を断片化した数十～数百の短い塩基配列の形式で DNA を読み出す。このままでは、断片化前の DNA は得られず、ショートリードを断片化前の形に並び替える処理が必要である。</p> <p>ショートリードマッピングとは、ショートリードの(予め用意した同じ生物種の既知の配列、例えばヒトゲノムなど)参照配列上での最も類似した位置を決定することである。この位置決定には技術的な課題が主に2つある。</p> <p>第一に、データ量が膨大なため、処理速度が重要である。例えば、ヒトの遺伝子の場合には 1000 億のショートリードを処理することも必要となる。第二に、参照配列と完全に一致しない箇所でも、うまくマッピングできることが重要である。リード対象と参照配列に違いがある部分は、両者の遺伝的な違いであり、これを発見することがショートリードマッピングの目的の一つでもある。現状では、さまざまなショートリードマッピングソフトウェアが利用されている。しかしながら、新世代シーケンサーは、ショートリードの検出を大規模かつ並列的に行うため、ショートリードマッピングソフトウェアが処理のボトルネックになっている。そのためショートリードマッピング問題を高速化することは重要な課題である。マッピング処理はショートリードごとに並列に処理することができる。そこで、Field Programmable Gate Array (FPGA) を用い並列に処理することで高速化を実現した。提案手法では、初めにソーティングして同種のデータを集め、集まったデータを FPGA 上に 1000 個以上実装された簡単な比較器で同時に比較する。提案手法を従来手法(ソフトウェアおよび FPGA システム)と比較し、処理速度、マッピング精度の両方で優れていることを示した。</p>				
審査日	平成 27 年 1 月 29 日			
審査員	(大学名 職名)	(学位)	(氏名)	
主査	筑波大学 教授	工学博士	丸山 勉	
副査	筑波大学 教授	工学博士	白川 友紀	
副査	筑波大学 准教授	博士(工学)	延原 肇	