Częstochowa, dn. 26.07.2021

Dr hab. inż. Łukasz Szustak, prof. PCz
Katedra Informatyki
Politechnika Częstochowska
ul. Dąbrowskiego 69
42-201 Częstochowa

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr Anny Gajos-Balińskiej

„Metodyka optymalizacji czasowej algorytmu fastICA w elektroencefalografii gęstej matrycy z wykorzystaniem możliwości wielordzeniowych architektur"

Promotor: dr. hab. Grzegorz Wójcik

Niniejsza recenzja przygotowana została na zlecenie Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Rzeszowskiej na podstawie uchwały z dnia 19 maja 2021 r.

1. Ocena wyboru tematu i celu rozprawy

Jesteśmy obecnie świadkami bezprecedensowych zmian w obszarze systemów komputerowych wielkiej mocy. Globalnym trendem, który obecnie napędza rozwój systemów komputerowych jest stosowanie architektur masywnie wielordzeniowych, co bezpośrednio przekłada się na nieustający przyrost oferowanej mocy obliczeniowej. Jednakże obecny rozwój systemów komputerowych wielkiej mocy rodzi szereg problemów badawczych optymalnego ich wykorzystania. W rezultacie zauważalna jest wyraźna potrzeba opracowania nowoczesnych metod i algorytmów adaptacji wybranych aplikacji komputerowych do nowoczesnych systemów komputerowych.

W tym kontekście należy pozytywnie ocenić wybór tematyki rozprawy zmierzający do opracowania metodyki poprawy wydajności algorytmu fastICA w elektroencefalografii gęstej matrycy z wykorzystaniem możliwości wielordzeniowych architektur komputerowych, w tym akceleratorów GPU. Rozpatrywane problemy stanowią bez wątpienia istotne wyzwanie metodologiczny i algorytmiczne, gdyż ze względu na złożoność problemu wydobycie potencjału obliczeniowego z obecnych systemów komputerowych jest niebagatelnym wyzwaniem o teoretycznym i praktycznym charakterze.

Tematyka przedłożonej do recenzji rozprawy doktorskiej mgr Anny Gajos-Balińskiej dotyczy ważnego obszaru badawczego w dziedzinie informatyki, a podjęte wyzwania badawcze dobrze wpisują się we współczesne trendy rozwoju nowoczesnych metod wykorzystania systemów wielordzeniowych. Podsumowując ten punktu recenzji, można stwierdzić, iż podjęta problematyka jest zasadna a rozważane w niej problemy są istotne i aktualne, zarówno dla teorii, jak i praktyki.

WPŁYNĘŁO
06. SIE. 2021

2. Koncepcja i redakcja rozprawy

Recenzowana praca doktorska obejmuje formalnie 4 rozdziały poprzedzone wstępem oraz zakończone podsumowaniem. Zasadnicza część rozprawy liczy łącznie 89 stron oraz zawiera bibliografię składającą się z 78 pozycji.

Praca rozpoczyna się wstępem, który w połączeniu z pierwszym rozdziałem wprowadza czytelnika w dziedzinę jaką jest neuroobrazowanie, a w szczególności elektroencefalografia. We wstępie pracy Autor sformułował tezę, przedstawił szczegółowe cele oraz wykazał nowatorskie elementy pracy. Następnie zostały przedstawione pokrótce techniki neuroobrazowania oraz został zawarty opis niezbędnych urządzeń umożliwiających rejestrowania (akwizycji) sygnału EEG z wykorzystaniem mikrowoltomierzy różnicowych.

Kolejny rozdział charakteryzuje metodę ślepej separacji źródeł oraz algorytm analizy niezależnych składowych w kontekście ich zastosowania do przetwarzania danych EEG. W tym rozdziale Autor dokonuje również przeglądu literatury prezentując aktualny stan wiedzy na temat istniejących sposobów implementacji przedstawianych wcześnie algorytmów i metod.

Rozdział trzeci rozpoczyna w pełni oryginalną część dysertacji i poświęcony jest opracowaniu autorskiej metodzie dostosowania wybranego algorytmu typu ICA do nowoczesnych systemów komputerowych. Rozdział ten zawiera całościowy proces skorelowania urządzania EEG z równoległą wersją algorytmu typu ICA przy wsparciu bibliotek matematycznych, technologii Intel i CUDA

W rozdziale czwartym Autor w sposób obszerny zweryfikował skuteczność proponowanych rozwiązań wykorzystując do tego celu systemy komputerowe dużej mocy. W tej części pracy zostały omówione wyniki wydajnościowe obrazujące możliwości obliczeniowe autorskiej implementacji algorytmu typu ICA.

W kolejnej części rozprawy dokonano zwięzłego podsumowania realizowanych badań, w tym jej oryginalnych elementów, a także wskazano kierunki przyszłych prac w obszarze tematyki rozprawy.

3. Najważniejsze osiągnięcia rozprawy

Uwzględniając powyższe omówienie zawartości pracy oraz pozytywną ocenę jej zawartości merytorycznej, za bezsporne osiągnięcia Autora należy uznać zastosowanie nowoczesnych systemów komputerowych do poprawy wydajności algorytmu typu ICA. Najważniejszymi elementami rozprawy decydującymi o jej wartości naukowej są badania nad dostosowaniem algorytmu fastICA do architektur wielordzeniowych. W badaniach została opracowana metodologia pozwalająca nie tylko na wykorzystanie wielordzeniowości ale również wektoryzacji, których połączenie stanowi klucz do wydobycia pełnego potencjału obliczeniowego obecnych procesorów ogólnego przeznaczenia CPU. Proponowana metodologia została również rozszerzone o możliwości wykorzystania akceleratorów GPU, pozwalając tym samym na pełne użycie obecnych systemów wysokowydajnych w obliczeniach równoległych. Należy przyznać, że Autor podjął się realizacji bardzo ciekawego oraz istotnego z punktu widzenia

praktycznych zastosowań problemu badawczego.

Uzyskane wyniki zostały opublikowane w kilkunastu pracach w języku angielskim, w ramach których znajdują się wysoko punktowane pozycje czasopism z tzw. listy filadelfijskiej. Prowadzona aktywność publikacyjna oraz zakres tematycznych zagadnień poruszanych w powyższych pracach świadczy bardzo pozytywnie o stopniu weryfikacji uzyskanych rezultatów przez międzynarodową społeczność badawczą. Przedstawiona praca świadczy o dużej wiedzy Autorki w zakresie poruszanej tematyki badawczej, popartej szerokim doświadczeniem praktycznym.

4. Poprawność pracy i uwagi krytyczne

Poprawność treści pracy nie wzbudza zastrzeżeń, a stwierdzenia w niej zawarte wydają się być godne zaufania, co wynika w szczególności z dosyć szczegółowych uzasadnień, popartych wynikami przeprowadzonych badań eksperymentalnych. Generalnie sposób i jakość przeprowadzenia badań eksperymentalnych stanowi bardzo wartościowy element pracy i zasługuje na podkreślenie. W szczególności, dotyczy to różnorodnych eksperymentów przeprowadzonych z wykorzystaniem zróżnicowanych systemów komputerowych.

Jednocześnie Autor nie ustrzegł się pewnych braków i słabości. Wśród uwag o charakterze krytycznym, a po trosze dyskusyjnym, wymienić należy:

1. W pracy zabrakło szczegółowego omówienia adaptacji proponowanych rozwiązań do architektury hybrydowej CPU-GPU. W szczególności zabrakło opisu dla zastosowanej metodologi dystrybucji i realizacji obliczeń równoległych z wykorzystaniem CPU i GPU. W konsekwencji, nasuwa się tu szereg pytań, np. które fragmentu kodu zostały oddelegowane do wykonania na GPU, czy CPU wykonuje obliczenia w tym samym momencie co GPU, jaki został zastosowany model zrównoleglania obliczeń pomiędzy i wewnątrz CPU i GPU: „task parallelism" czy raczej „data parallelism".
2. Praca zyskałaby na wartości poprzez dodanie np. opisu analizy i oceny proponowanego rozwiązania redukcji ilości przesyłanych danych między CPU a GPU. Czy jest możliwe aby określić w jakim stopniu został zredukowany transfer danych pomiędzy CPU i GPU poprzez zastosowanie proponowanych w pracy modyfikacji (str. 43)?
3. Czy jest możliwe określenie czy też oszacowanie stopnia wykorzystania teoretycznej wydajność testowanych systemów komputerowych?
4. Obecnie można zaobserwować, że zastosowany standard Intel Cilk Plus został wyparty przez np. dostępną w ramach kompilatorów czy też standardu OpenMP automatyczną wektoryzacje obliczeń. Nasuwa się tutaj ogólne pytanie czy automatyczna wektoryzacja dostępna powszechnie w ramach obecnych kompilatorów mogłabym skutecznie zastąpić zastosowany w pracy standard Intel Cilk Plus?

5. Prezentacja wyników zrealizowanych badań zyskałaby na wartości poprzez wprowadzenie np. wpływu wektoryzacji obliczeń na wydajność, poprzez pomiar czasu wykonania algorytmu z włączoną i wyłączoną wektoryzacją. Nasuwa się tutaj naturalne pytanie jaki zysk wydajnościowy płynie z zastosowania wektoryzacji obliczeń?
6. Chciałbym także zapytać o zastosowaną w testach metodologię przypisywania wątków OpenMP do rdzeni (thread affinity) i jej wpływ na wydajność i skalowalność obliczeń.
7. Czy jest możliwe aby zaadoptować zaproponowane rozwiązania do systemów równoległych z pamięcią rozproszoną używając np. standardu MPI?
8. Drobne uwagi szczegółowe:
   - W pracy zabrakło również testów porównawczych opracowanej metodologii z wykorzystaniem nowoczesnych infrastruktur dużej skali, np. procesorów AMD EPYC Rome czy też procesorów Intel Cascade Lake
   - Opisy rysunków, a w szczególności tabel w rozdziale 4 są niejasne dla czytelnika. Każdy rysunek/tabela powinna zawierać klarowny opis czego dotyczy, np. Pomiar czasu [s] wykonania aplikacji xyz wykonany na platformie xyz z użyciem różnej liczby wątków
   - Określenia zawarte w rozdziale 4 typu „dla *architektury* Lomond" warto byłoby zstąpić określeniem np. „dla platformy Lomond" lub „dla klastra Lomond"
   - Rys. 2.1 i 3.2 wydaje się być mało czytelny dla czytelnika
   - Określenie „*solidna* maszyna" (str. 36) warto byłoby zstąpić określeniem np. „maszyna o wysokiej wydajności"
   - Autor powinien wskazać wersje oraz rodzaj użytych kompilatorów i bibliotek
   - Brak odwołania i opisu do listingu np. 3.6 i 3.7
   - W sekcja 4.1 wkradł się błąd ponieważ klaster Eagle, tak jak Altair udostępniany jest w ramach infrastruktury PCSS, a nie AGH Cyfronet
   - Rysunki 4.1 - 4.5 powinny zajmować zdecydowanie mniej miejsca
   - Określenie typu „+12 *wirtualnych*" warto byłoby zstąpić „+12 logicznych" (np. str. 53)
   - Określenie „Po pierwszych *obiecujących próbach*" (str. 43) warto byłoby zstąpić np. „Wstępne badania wykazał możliwość skrócenia czasu obliczeń poprzez ...."
   - Brak konsekwencji w nazewnictwie tabel, w opisie Autor używa słowa *Tabela*, natomiast słowo*Tablica* używane jest do opisu „nad" tabelą.
   - Str. 39, wydaje się, że zdanie *„To pozwoliło wykorzystać funkcję kompilatora Intel mm malloc, która usprawnia przechowywanie i strumieniowanie danych oraz zapobiega tak zwanym „cache misses"."* może wprowadzić w błąd

czytelnika. Warto zauważyć, że funkcja „__mm_malloc" pozwala na alokacje danych dopasowanych do wymaganej granicy w pamięci komputera umożliwiając tym samym bardziej efektywny ich transfer CPU-GPU lub CPU-cache. Niepożądany efekt „cache misses" wiąże się bardziej ze sposobem odwoływania do danych w tablicy co bezpośrednio zależy od danej aplikacji.

5. Podsumowanie

Przytoczone wyżej uwagi dyskusyjne nie umniejszają zasług Autora ani nie kwestionują przedstawionych osiągnięć. Recenzowana praca zasługuje na wysoką ocenę merytoryczną i wnosi istotny oraz oryginalny wkład w dziedzinie informatyki. Podstawowe cele i zadania pracy zostały zrealizowane, a jej tematyka dobrze wpisuje się we współczesny nurt badań w tym zakresie.

Stwierdzam zatem z pełnym przekonaniem, że opiniowana rozprawa Pani mgr Anny Gajos-Balińskiej zawiera samodzielne rozwiązanie ważnego i trudnego problemu naukowego, spełniając wszystkie wymagania przewidziane dla rozpraw doktorskich w aktualnie obowiązującej Ustawie o Tytule Naukowym i Stopniach Naukowych. W związku z tym stawiam wniosek o dopuszczenie Autora do dalszego toku przewodu doktorskiego.

Szustak
.....

Dr hab. inż. Łukasz Szustak