

Oponentský posudek diplomové práce

Diplomová práce: Implementace algoritmů pro zpracování obrazu na IBM Cell

Diplomant: Bc. Václav Klecanda

Vedoucí práce: Mgr. Václav Krajíček

Oponent: Ing. Miroslav Skrbek, Ph.D.

Diplomová práce Bc. Václava Klecandy se zabývá implementací algoritmů pro zpracování obrazu na platformě IBM Cell. Konkrétně se jedná o algoritmy pro segmentaci obrazu cílené do oblasti analýzy lékařských dat.

Práce obsahuje rešeršní část, kde je popsána architektura Cell Broadband Engine (dále jen Cell), vývojové SDK, vývoj programů a možnosti paralelizace na této architektuře. Součástí rešerše je popis metod pro segmentaci obrazu a jejich implementace na SIMD architekturách. Další část práce popisuje návrh a implementaci zvoleného algoritmu na procesoru Cell a jeho integraci do prostředí MedV4D. K diplomové práci je přiloženo CD se zdrojovými soubory a dokumentací.

Práce si klade dva cíle:

- zmapovat způsoby vývoje aplikací, nástroje pro vývoj programového vybavení a možnosti paralelizace algoritmů,
- implementovat vybraný algoritmus pro segmentaci obrazu na platformě Cell.

Musím konstatovat, že žádný z cílů není jednoduchým problémem. Splnění prvního cíle představuje studium značného množství literatury, návodů, poznámek a doporučení na Internetu, pochopit funkci a vzájemné vazby jednotlivých vývojových nástrojů a zejména překonat mnoho drobných (ale časově náročných) problémů týkajících se zejména různých nekompatibilit, neúplnosti dokumentace apod.

Na tomto místě musím konstatovat, že tuto část diplomatu velice dobře zvládl a ve své práci dokumentoval, což bude dobrým vodítkem pro jeho následovníky.

Druhý cíl práce představoval výběr segmentačního algoritmu a jeho implementaci na platformě Cell. V rešeršní části bohužel nejsou popsána kritéria výběru, ani důvody proč student vybral právě Level Set algoritmus. Není jasné, zda byla při výběru algoritmu zvažována jeho vhodnost pro danou implementační platformu nebo byl vybrán bez ohledu na architekturu (např. po dohodě s vedoucím práce nebo prostě proto, že je nejpoužívanější).

Diplomant zvolil metodu portace existujícího algoritmu, nikoli jeho novou implementaci s ohledem na architekturu. Tuto metodu nepovažuji za nejlepší, protože ne vždy vede k optimálním výsledkům. Na druhou stranu je to dnes legitimní a často užívaný způsob, jak rychle získat funkční implementaci z již existující implementace pro jinou platformu. Výběr implementační metody tedy nepovažuji za chybný, diplomant si vybral s ohledem na charakter a objem celé diplomové práce časově méně náročnou metodu, jejichž prověření má praktický význam.

Segmentační algoritmus byl integrován do existujícího prostředí MedV4D, doplněn o možnost vzdáleného spouštění na výpočetním serveru a vybaven uživatelským prostředím.

Celek tvoří aplikaci s grafickým uživatelským rozhraním, které je prakticky použitelné a dovoluje snadnější experimentování. Z hlediska dosaženého výsledku je to jednoznačně přednost a správná cesta, kterou musím ocenit.

Podle přiložených pokynů se mi aplikaci podařilo bez problémů přeložit, a to díky pečlivě provedenému Makefile. Podařilo se mi spustit jak aplikaci, tak server, nahrát snímek a spustit výpočet, ale s implicitními hodnotami parametrů se mi nepodařilo získat rozumný výsledek, případně aplikace skončila chybou bus error v závislosti na vybraných datech. Diplomant by měl ověřit, zda je k diplomové práci přiložena opravdu aktuální funkční verze. V dokumentaci jsem postrádal popis, případně jednoduchý příklad, použití aplikace pro konkrétní data.

V diplomové práci jsou prezentovány výsledky měření algoritmu na architektuře Cell a i686, které značně vycházejí v neprospěch Cellu a patrně nenaplnily očekávání diplomanta. Negativní výsledek nic nemění na celkově pozitivním hodnocení celé práce. Důležité je, že v kapitole 6.2 diplomant rozebírá příčiny pomalého běhu algoritmu na platformě Cell. Oceňuji snahu diplomanta v kap. 6.3 charakterizovat algoritmy vhodné pro platformu Cell.

Se závěry v kapitole 6.2. lze souhlasit, mám k nim však následující poznámky:

- přenos po jednotlivých 32b slovech je neefektivní, přenos okolí 3^3 lze provést DMA přenosem $4 \times 3 \times 3$ (místo $3 \times 3 \times 3$), tj. DMA list s 9 položkami délky 16B, případně $8 \times 3 \times 3$, tj. DMA list s 9 položkami délky 32B, když je okolí na rozhraní alignmentu [předpokládám obraz uložený po řádcích, 32b/pixel]. Výběr jednotlivých pixelů lze efektivně provést např. makrem `spu_sel`,
- postrádám přesná měření na SPU s použitím čítačů a vysledovat chování DMA přenosů. Pro jednoduché přenosy platí, že přenos prvních 16 bytů může trvat několik mikrosekund, což je několikanásobně více než přenos následujících dat. Pokud čteme ze stejné nebo blízkých adres je přenos podstatně kratší (souvislost s překladem adres). Záleží tedy silně na přístupovém vzoru,
- je třeba zkontrolovat, zda nebyla v algoritmu použita nějaká funkce, která je typu „PPE supported“. Tím se výpočet radikálně zpomalí. Obvykle se jedná o systémovou funkci (zjištění času, `printf` apod.),
- při použití mailboxů lze komunikaci zrychlit použitím atributu `SPE_MAP_PS` ve funkci `spe_context_create`.

Vlastní text práce je psán přehledně a srozumitelně. Oceňuji, že je práce napsána v anglickém jazyce.

Předložená diplomová práce představuje velký kus práce, kterou student musel vykonat a na které prokázal, že je schopen vyřešit složitý problém. Výsledky práce jasně ukázaly na problémy portace algoritmů na platformu Cell a budou cenným východiskem pro další implementace.

I přes výše uvedené připomínky považuji práci za kvalitní a doporučuji ji k obhajobě.

V Praze 2.9.2009

Ing. Miroslav Skrbek, Ph.D.