

Oponentní posudek disertační práce

Název práce: Návrh a optimalizace obrazových klasifikátorů

Autor: Ing. Filip Kadlček

Školitel: doc. Dr. Ing. Otto Fučík

Předložená dizertační práce se zabývá návrhem architektury pro klasifikaci (rozpoznávání) obrazových dat. Navržená architektura je založena zejména na programovatelných obvodech FPGA, nicméně je zde možnost efektivně využít „spolupráci“ s CPU. Zde se přímo nabízí použití v moderních systémech na čipu (SoC), což autor v důsledku plně využil při implementaci v dnes moderní Xilinx Zynq platformě.

Ve výsledku navržená architektura dosahuje značně lepších výsledků, než současná architektury - klasifikační schopnosti jsou srovnatelné, přičemž je dosahováno podstatně větší propustnosti, tedy rychlosti klasifikace.

Toto téma je dnes velice aktuální, klasifikace obrazových dat je zapotřebí ve stále více praktických aplikacích. S nástupem tzv. chytrých kamer se tato potřeba jen více zdůrazňuje. Výsledky práce jsou tedy přímo aplikovatelné v praxi, nicméně znamenají i nezanedbatelný teoretický přínos.

Text práce je dobře strukturovaný, velice dobře čitelný a srozumitelný i pro neodborníka.

V úvodu práce je nastíněna problematika a problémy související s praxí v oblasti rozpoznávání (klasifikace) obrazových dat. Je řečeno, že využití FPGA pro klasifikaci dat by mohlo přinést zlepšení výkonu. Dále jsou stanoveny cíle práce.

V druhé kapitole autor uvádí do problematiky klasifikace (detekce) objektů a prezentuje současné přístupy (state-of-the-art) s jejich teoretickým základem. Zde bych možná uvítal více detailů o klasifikátorech založených na neuronových sítích.

Kapitola 3 představuje současné techniky implementace „boostingových metod“, na kterých je dizertační práce založena. Pojednává tedy o praktické, implementační, stránce. Jednotlivé metody jsou zde porovnány z hlediska rychlosti. Dle mého názoru je výčet současných metod vyčerpávající.

Kap. 4 představuje první část vlastního přínosu práce – nové tvary LBP operátorů a genetický algoritmus pro jejich návrh. Musím ocenit poctivé testování GA (zejména z hlediska konvergence). Zdá se, že opravdu funguje.

V kapitole 5 je uveden způsob vyhodnocování AdaBoost klasifikátoru v FPGA (resp. pouze způsob vyhodnocování jednoho detekčního okna), tedy na paralelním hardwaru. Je zde efektivně využit princip proudového zpracování. V důsledku je architektura schopna zpracovávat jeden pixel obrazu v jednom hodinovém taktu, přičemž i latence je relativně malá.

Zde oceňuji zejména možnost volby nejjvhodnější implementace klasifikátoru v FPGA, podle způsobu jeho natrénování. Tento proces je navíc automatizovaný (iterativní proces návrhu klasifikátoru).

Dále oceňuji způsob rozdělení klasifikace na dvě části – preprocessing a postprocessing, přičemž každá část je implementována v jiných HW prostředcích (FPGA, CPU). Ač to pravděpodobně znamená nezanedbatelné navýšení plochy celkového návrhu, je dosahováno značné úspory energie.

Návrh celého klasifikátoru je pak popsán v kap. 6. Zde jsou popsány dva přístupy – sekvenčně paralelní klasifikátor a multiparalelní detektor objektů. Druhý přístup pak byl vybrán pro implementaci. Výsledná architektura je pak schopna zpracovávat obraz rychlostí

až 300 Mpx/s, což je, jak autor tvrdí, téměř o řád výše, než jsou schopny dosahovat současné klasifikátory.

Kapitola 7, Závěr, shrnuje dosažené cíle a navrhuje možná pokračování práce.

Veškerá experimentální vyhodnocení byla provedena velice precizně a značně vypovídajícím způsobem. V tomto ohledu mohu jen chválit.

Práce obsahuje pouze minimum chyb a překlepů. Mám jen několik více méně formálních připomínek:

- V kap. 3.3.1 je uvedeno, že „Vyrovňovací paměť konstruována tak, že má jeden čtecí port a 21 výstupních portů“. Kudy se tedy do ní zapisuje?
- V legendě k Tab. 2 jsou uvedeny zdroje „Slice Register“ a „DSP“. Ty však v tabulce nejsou uvedeny. Proč?
- Seznam zkratk (i dále v textu): GP-GPU General Purpose GPU – překlep (purpose).
- Zvažte rozdělování/nerozdělování některých složených slov. Např. „jedno jádrový“, „pod-okna“, ...
- Opatrně s pojmem „problém“. V textu je často problémem označována instance problému (např. kap. 4.3).
- V kapitolách 5 a 6 je zkratka „LUT“ používána ve dvou kontextech (Look-up Table v FPGA a tabulka s uloženými výsledky trénovacího procesu).

Výsledky práce byly prezentovány na třech prestižních mezinárodních konferencích a na doktorandském workshopu. V tomto ohledu je práce mírně podprůměrná. Otázkou je, proč. Nicméně tyto publikace potvrzují dobrou vědeckou erudici uchazeče i předložené práce, nevidím v tom tedy velký problém.

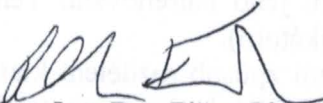
Otázky k diskuzi

1. V jakých aplikacích je opravdu potřeba vyhodnocovat snímky takovou rychlostí?
2. Neuvažoval jste použití FMGA (Fast Messy Genetic Algorithm) nebo nějaké jeho modifikace, namísto standardního GA? Zde se přímo nabízí – části chromozomu (oblasti) jsou na sobě pozičně nezávislé.
3. Jak dlouho trvá výpočet fitness funkce, vzhledem k celkovému času běhu GA?
4. Kdy je tedy nutné pouštět GA? Za jakých okolností?
5. Bylo by možné provést srovnání s klasifikátory založenými na neuronových sítích? Máte představu, jak by dopadlo, případně jaké jsou výhody/nevýhody těchto dvou přístupů?

Závěr

Lze konstatovat, že předložená práce odpovídá obecně uznávaným požadavkům k udělení akademického titulu Ph.D. a doporučuji ji k obhajobě.

V Praze, 24. 1. 2017



doc. Ing. Petr Fišer, Ph.D.

České vysoké učení technické v Praze
Fakulta informačních technologií