

Akademia Górniczo-Hutnicza im. Stanisława Staszica  
Wydział Elektrotechniki Automatyki Informatyki i Elektroniki

Rozprawa doktorska: autoreferat

**Metodyka zrównoleglania algorytmów  
przetwarzania i analizy obrazów  
w systemach przepływowych**

**Mirosław Jabłoński**

Promotor: dr hab. inż. Marek Gorgoń

Kraków, 2009

# 1 Uzasadnienie ważności i celowości podejmowanej tematyki rozprawy doktorskiej

Systemy składowania, transmisji, prezentacji analizy i rozpoznawania sygnałów z otaczającego nas świata stają się coraz bardziej obecne w niemal wszystkich sferach ludzkiego życia.

Od kilku lat obserwuje się zatrzymanie wzrostu maksymalnej częstotliwości taktowania systemów cyfrowych, a w szczególności procesorów ogólnego przeznaczenia. Wzrost mocy obliczeniowej uzyskiwany jest poprzez zwielokrotnienie liczby jednostek obliczeniowych. Przykład stanowią wielordzeniowe procesory ogólnego przeznaczenia [4], niektóre procesory DSP [1][11], procesory graficzne GPU [10] i układy rekonfigurowalne FPGA [6][12]. Autor pracy [7] zaznacza, iż rzeczywisty potencjał równoległych obliczeń nie jest należycie wykorzystany.

Jedną z dziedzin, która ze względu na dynamiczny rozwój jest ciągle domeną prac badawczych i specjalistycznych zastosowań, są szeroko rozumiane systemy wizyjne. Zagadnienie realizacji algorytmów widzenia maszynowego jest obecne w nauce i technice przynajmniej od połowy ubiegłego stulecia. Takie prace były prowadzone również w Akademii Górniczo-Hutniczej [5][9][13][14].

W ostatnich latach, prowadzone przez wiele ośrodków badania w tym obszarze skupiały się na realizacji algorytmów przetwarzania i analizy obrazów w układach rekonfigurowalnych FPGA [2][8]. Podejmowano również próby realizacji kompleksowych algorytmów wizyjnych [3][15].

## 2 Rezultaty i wyniki badań

W trakcie badań i prac projektowych został wykonany przegląd rodzajów równoległości w kontekście strumieniowej realizacji systemów wizyjnych. Analiza została przeprowadzona na reprezentatywnych przykładach torów wizyjnych z uwzględnieniem operacji o różnorodnym kontekście pikseli w obrazie: przestrzenny lokalny, przestrzenny globalny i kontekst temporalny. Następujące algorytmy wizyjne zostały w całości zaimplementowane przez autora rozprawy:

- interpolacja pikseli w barwnym czujniku wizyjnym,
- rozpoznawanie ręcznie pisanych cyfr przy pomocy sieci neuronowej,
- wideodetekcja w ruchu drogowym.

Podczas ich realizacji zastosowano różne rodzaje równoległości (równoległość danych, równoległość operacji) adekwatne do założeń i zidentyfikowanych ograniczeń: działanie strumieniowe, niska częstotliwość taktowania (osiąganie niezbędnych mocy obliczeniowych przy możliwie małym zużyciu energii), różne rodzaje stosowanych kontekstów obrazu, zależność czasu wykonania od treści strumienia wizyjnego, fizyczne ograniczenia platformy (w tym ograniczenia pamięciowe).

Sposób wykorzystania równoległości został przedstawiony etapami. Istotnym elementem badań była identyfikacja i wykorzystanie praw opisujących systemy równoległe podczas projektowania specjalizowanych architektur obliczeniowych dedykowanych do wybranych algorytmów. Autor wskazał

istotne mierniki zrównoleglenia, które są przydatne w efektywnym projektowaniu systemów obliczeniowych dla aplikacji wizyjnych.

Oryginalne osiągnięcie autora rozprawy stanowi opracowanie metodyki oceny stopnia zrównoleglenia algorytmów wizyjnych celem osiągnięcia architektury strumieniowego systemu wizyjnego, wykonalnej dla określonej przepływności źródła sygnału wizyjnego. Metodyka ta obejmuje sposób wyznaczania szeregu powszechnie stosowanych parametrów opisanych w drugiej części rozprawy (rozdział 4 i 5): przyspieszenie, efektywność, opóźnienie transportowe oraz wskaźników wprowadzonych przez autora w drugiej części rozprawy (rozdział 6.1 i 6.4): zapas czasowy<sup>1</sup> strumienia wizyjnego (1), znormalizowany rozstęp<sup>2</sup> czasu wykonania (2) w odniesieniu do nieregularnych algorytmów.

$$V = \frac{T_F - MN}{T_F} \quad (1)$$

gdzie:

- $V$  – zapas czasowy strumienia wizyjnego,
- $M$  – liczba aktywnych wierszy w obrazie,
- $N$  – liczba aktywnych kolumn w obrazie ,
- $T_F$  – całkowity czas akwizycji ramki obrazu liczony w cyklach zegara.

$$v = \frac{T_{max} - T_{min}}{T_{max}} \quad (2)$$

gdzie:

- $v$  – znormalizowany rozstęp czasu obliczeń algorytmu,
- $T_{max}$  – maksymalna liczba cykli algorytmu dla jednego obrazu,
- $T_{min}$  – minimalna liczba cykli algorytmu dla jednego obrazu.

W myśl metodyki opracowanej przez autora rozprawy, podstawowym kryterium wykonalności jest osiągnięcie przepustowości 1 piksel/cykl zegara poprzez zwiększanie przyspieszenia metodą zrównoleglenia obliczeń oraz dodatkowo, w odniesieniu do algorytmów nieregularnych, spełnienie warunku (3).

$$v \leq V \quad (3)$$

---

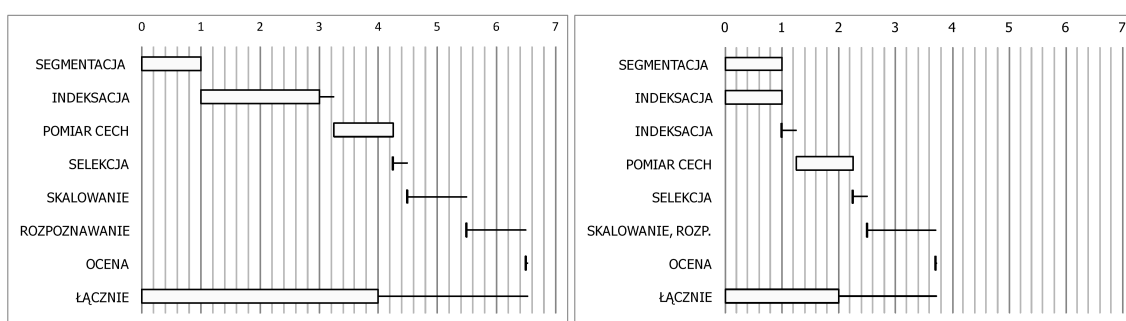
<sup>1</sup>Zapasy czasowy strumienia wizyjnego został wprowadzony przez autora jako miara względnego czasu martwego w strumieniu.

<sup>2</sup>Rozstęp jest pojęciem zaczerpniętym ze statystyki i wykorzystanym przez autora do szacowania granicznych czasów wykonania algorytmów, których działanie zależy od treści obrazu.

## 2.1 Operacje wizyjne o kontekście przestrzennym

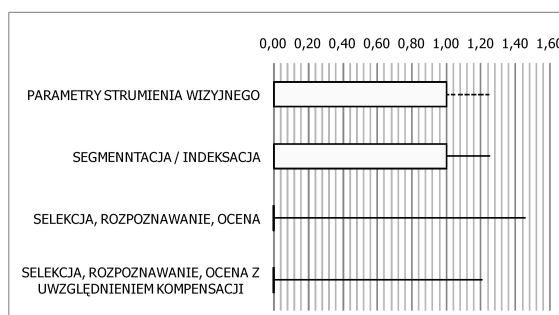
W odniesieniu do operacji wizyjnych o kontekście przestrzennym, istotnym elementem metodyki jest badanie wartości wskaźników dla algorytmów o czasie wykonania zależnym od treści obrazu. Podstawę do oceny wykonalności algorytmów o dużym stopniu nieregularności jest wskazanie takiej treści obrazu dla której czas realizacji algorytmu jest największy. W kolejnych krokach zmniejszany jest rozstęp czasu wykonania do wartości nieprzekraczającej zapasu czasowego strumienia wizyjnego.

Wyniki zastosowania opisanej metodyki toru wizyjnego rozpoznającego ręcznie pisane znaki przedstawiono na rysunku 1. Wykres na rysunku 1a został wyznaczony w oparciu o szacowane rozmiary danych przy założeniu jednego taktu zegara na każdy piksel w danej iteracji.



(a) realizacja sekwencyjna

(b) zrównoleglenie na wysokim i średnim poziomie granulacji



(c) pełne zrównoleglenie

Rysunek 1: Czasy wykonania algorytmów w systemie rozpoznawania znaków.

Poziome osie układu współrzędnych znormalizowane są do czasu transmisji jednego obrazu. Obszary pokryte pustymi prostokątami odpowiadają operacjom o stałym czasie wykonania. Linie poziome wyznaczają granice, w których zawierać się może czas wykonania operacji, zależnie od treści obrazu. Linia przerywana obrazuje zapas czasowy strumienia wizyjnego

Zrównoleglenie na poziomie potoku drobnoziarnistego pozwoliło uzyskać efekty przedstawione na rysunku 1b. Pomimo znaczącego przyspieszenia każdego z etapów algorytmu i uzyskanie przepustowości na poziomie 1 piksel/cykl zegara w wybranych podsystemach toru wizyjnego nie było możliwe płynne przetwarzanie strumienia wizyjnego bez pominięcia danych.

Ostateczny efekt zrównoleglenia aplikacji wizyjnej po zastosowaniu potoku gruboziarnistego został przedstawiony na rysunku 1c. Na szczególną uwagę zasługuje fakt, że (jak to wykazano w rozdziale 6.5 rozprawy) analiza czasu wykonania w zależności od treści obrazu pozwala pozytywnie ocenić cały tor wizyjny, pomimo iż prosta superpozycja rozstępów czasu wykonania na to nie pozwala. Szczegółowa analiza zależności czasu wykonania poszczególnych operacji przetwarzania umożliwiła zaobserwowanie efektu kompensacji rozstępu czasu wykonania w złożonym torze wizyjnym. W analizowanym przypadku rozpoznawania znaków, gdzie kluczowa jest liczba obiektów zidentyfikowanych w obrazie, zjawisko to objawia się wzrostem czasu wykonania indeksacji przy rosnącej liczbie obiektów, podczas gdy spada wówczas czas wykonania selekcji oraz rozpoznawania znaków. W oparciu o przedstawione dane można jednoznacznie stwierdzić, że integracja wielu zrównoleglonych operacji pozwala na przepływową realizację toru wizyjnego złożonego z algorytmów o czasie wykonania zależnym od treści strumienia wizyjnego.

## 2.2 Operacje wizyjne o kontekście temporalnym

W rozdziale 7 opisano realizację operacji wizyjnych o kontekście temporalnym na przykładzie sprzętowej aplikacji wideodetektora w ruchu drogowym. Przedstawione tam algorytmy estymacji tła oraz detekcji pojazdów w obszarach definiowanych przez użytkownika wykorzystują temporalny rodzaj kontekstu. Krytyczne w operacjach wymagających kontekstu temporalnego okazały się organizacja danych w pamięciach oraz sposób komunikacji elementów obliczeniowych z pamięciami. Autor rozprawy zaproponował cztery wzajemnie uzupełniające się rozwiązania, które umożliwiły przepływową realizację systemu wideodetekcji na dostępnej platformie obliczeniowej (rozdział 7.3.2) mimo istniejących ograniczeń zasobów pamięciowych.

Autor zastosował w badaniach parametryczną ocenę równoległości zasobów pamięciowych poprzez równoległość danych pamięci  $Q_{DLP}^{Mem}$  i równoległość operacji pamięci  $Q_{OLP}^{Mem}$ . Celem omówionego w rozprawie procesu projektowego prowadzącego do wykonalności systemu wideodetekcji, było takie zbilansowanie współczynników, aby dostosować aktualne stopnie równoległości do możliwości zasobów pamięciowych. Rozwiązanie to, umożliwiło realizację algorytmu wideodetekcji dla obszarów detekcji o dowolnym kształcie i rozmiarze bez wpływu na opóźnienie transportowe. Wartość opóźnienia transportowego w zaimplementowanym podsystemie wideodetekcji została oszacowana na poziomie  $4\mu s$ .

## 3 Podsumowanie

Zasadniczym osiągnięciem badań opisanych w rozprawie jest opracowanie nowatorskiej metodyki zrównoleglenia algorytmów wizyjnych. Zaproponowany sposób postępowania pozwala na implementa-

cję złożonych systemów wizyjnych w układach rekonfigurowalnych. Zastosowanie ilościowych wskaźników pozwala na ocenę wykonalności bezstratnego, przepływowego systemu wizyjnego z uwzględnieniem wpływu treści strumienia wizyjnego.

## Literatura

- [1] Analog Devices, Norwood USA. ADSP-BF561 Blackfin Embedded Symmetric Multiprocessor, rev. d edition, February 2009.
- [2] Z. Baker, M. Gokhale, and J. Tripp. Matched filter computation on fpga, cell and gpu. In 15th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, pages 207–218, USA, April 2007. IEEE.
- [3] H. Fatemi. Processor Architecture Design for Smart Cameras. PhD thesis, Technische Universiteit Eindhoven, 2007.
- [4] H. Feng, E. Li, Y. Chen, and Zhang Y. Parallelization and characterization of sift on multi-core systems. In IEEE International Symposium on Workload Characterization, pages 14–23. IEEE, Sptember 2008.
- [5] M. Gorgoń. Architektury rekonfigurowalne do przetwarzania i analizy obrazu oraz dekodowania cyfrowego sygnału wideo. Rozprawy Monografie. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, 2007.
- [6] A.K. Gupta, S. Nooshabadi, D. Taubman, and M. Dyer. Realizing low-cost high-throughput general-purpose block encoder for jpeg2000. In IEEE Transactions on Circuits and Systems for Video Technology, volume 16(7), pages 843–858. IEEE, July 2006.
- [7] R. Hartenstein. Reconfigurable hpc: torpedoed by deficits in education? In Seventh International Conference on High Performance Computing and Grid in Asia Pacific Region, pages 428–429. IEEE, July 2004.
- [8] H. Jiang, H. Ardo, and V. Owall. A hardware architecture for real-time video segmentation utilizing memory reduction techniques. In IEEE Transactions on Circuits and Systems for Video Technology, volume 19(2), pages 226–236. IEEE, February 2009.
- [9] Z. Mikrut and Z. Stanek. Low-cost board for image digitalization, remapping and processing. In Automatyka: półrocznik Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie, volume 3(2), pages 487–497. PWN, Kraków, 1999.
- [10] M. Rofouei, M. Moazeni, and M. Sarrafzadeh. Fast gpu-based space-time correlation for activity recognition in video sequences. In IEEE/ACM/IFIP Workshop on Embedded Systems for Real-Time Multimedia, pages 33–38. IEEE, October 2008.

- [11] H. Seo and S. Kim. Openmp directive extension for blackfin 561 dual core processor. In The Sixth IEEE International Conference on Computer and Information Technology, pages 49 – 55. IEEE, September 2006.
- [12] O. Sims and J. Irvine. An fpga implementation of pattern-selective pyramidal image fusion. In International Conference on Field Programmable Logic and Applications, pages 1–4. IEEE, August 2006.
- [13] R. Tadeusiewicz. Systemy wizyjne robotów przemysłowych. Wydawnictwo Naukowo-Techniczne, Warszawa, 1992.
- [14] K. Wiatr. Sprzętowe implementacje algorytmów przetwarzania obrazów w systemach wizyjnych czasu rzeczywistego. AGH Uczelniane Wydawnictwa Naukowo-Dydaktyczne, 2002.
- [15] J. Xu, Y. Dou, J. Li, X. Zhou, and Q. Dou. Fpga accelerating algorithms of active shape model in people tracking applications. In 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools, pages 432–435. IEEE, August 2007.