

Politechnika Warszawska  
Wydział Elektroniki i Technik Informatycznych

Warszawa, 29 listopada 2017 r.

D z i e k a n a t

Uprzejmie informuję, że na Wydziale Elektroniki i Technik Informatycznych Politechniki Warszawskiej odbędzie się w dniu 12 grudnia 2017 r. publiczna obrona rozprawy doktorskiej

**mgr inż. Macieja Trochimiuka**

**temat:** „Optymalizacja predykcji międzyobrazowej w kodowaniu danych wizyjnych”.

**promotor – dr hab. inż. Grzegorz Pastuszak, prof. Politechniki Warszawskiej**

**recenzenci:**

**prof. dr hab. Marek Gorgoń z Akademii Górniczo-Hutniczej**

**dr hab. inż. Marek Parfieniuk z Politechniki Białostockiej**

Obrona odbędzie się w dniu 12 grudnia 2017 r. w sali 116 na Wydziale Elektroniki i Technik Informatycznych – Gmach im. Janusza Groszkowskiego, Warszawa, ul. Nowowiejska 15/19; początek godz. 9.00.

Po adresie: [www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje](http://www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje) zapewniony jest na stronie Wydziału dostęp do tekstów streszczenia rozprawy i recenzji, jak również do tekstu rozprawy umieszczonej w Bazie Wiedzy Politechniki Warszawskiej.

Dziekan



prof. dr hab. inż. Krzysztof Zaremba

Rodzaj pracy: rozprawa doktorska

mgr inż. Maciej Trochimiuk

temat: „Optymalizacja predykcji międzyobrazowej w kodowaniu danych wizyjnych”.

promotor – dr hab. inż. Grzegorz Pastuszak, prof. Politechniki Warszawskiej

## Streszczenie

Pomimo olbrzymiego rozwoju algorytmów kompresji danych wizyjnych, podstawową wykorzystywaną przez nie techniką pozostaje predykcja międzyobrazowa oparta na blokowej estymacji ruchu. Jest ona także elementem kodera, który w największym stopniu odpowiada za jego złożoność obliczeniową. Wraz z rozwojem technik kompresji danych wizyjnych zaproponowano wiele szybkich algorytmów predykcji międzyobrazowej, które w pewnych warunkach osiągają suboptymalne wyniki.

Rosnącą rozdzielczości cyfrowych sygnałów wizyjnych wymagają jednak ustawicznego poszukiwania coraz skuteczniejszych algorytmów.

W niniejszej rozprawie przedstawiono szereg algorytmów wyszukiwania wektorów ruchu dla potrzeb predykcji międzyobrazowej. Poza przedstawieniem istniejących metod autor pracy zaproponował własne rozwiązanie mające na celu redukcję złożoności obliczeniowej predykcji międzyobrazowej przy niewielkim zmniejszeniu efektywności kompresji.

Zaproponowane rozwiązanie wykorzystuje w całkowitoliczbowej estymacji ruchu wyniki wyszukiwania wektorów ruchu dla bloków o rozmiarze  $8 \times 8$  w blokach o większych rozmiarach. Ponadto w estymacji z ułamkową dokładnością ograniczono obliczenia do bloków kwadratowych, a zastosowanie oddzielnych list rankingowych oraz wyników cząstkowych pozwoliło na jednoczesne wyszukiwanie wektorów ruchu w pozostałych blokach. Zaproponowane rozwiązania przebadano pod kątem wpływu na efektywność kompresji, implementacji programowej oraz sprzętowej z wykorzystaniem układów typu ASIC i FPGA.

Słowa kluczowe: estymacja ruchu, H.265, HEVC, predykcja międzyobrazowa.

*Prof. dr hab. inż. Marek Gorgoń  
Katedra Automatyki i Inżynierii Biomedycznej  
Wydział Elektrotechniki, Automatyki,  
Informatyki i Inżynierii Biomedycznej  
Akademia Górniczo-Hutnicza  
Al. Mickiewicza 30  
30-059 Kraków*

Kraków, 23 października 2017.

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY  
WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMACYJNYCH  
POLITECHNIKI WARSZAWSKIEJ**

**Tytuł rozprawy: Optymalizacja predykcji międzyobrazowej w kodowaniu danych wizyjnych**

**Autor rozprawy: mgr inż. Maciej Trochimiuk**

Przedmiotem niniejszej recenzji jest rozprawa doktorska zatytułowana „Optymalizacja predykcji międzyobrazowej w kodowaniu danych wizyjnych”, napisana w roku 2017, której autorem jest mgr inż. Maciej Trochimiuk.

Promotorem pracy jest dr hab. inż. Grzegorz Pastuszek prof. Politechniki Warszawskiej. Niniejsza ocena została przygotowana na zlecenie prof. dr hab. inż. Krzysztofa Zaremby, Dziekana Elektroniki i Technik Informatycznych, zawarte w piśmie z dnia 24 marca 2017 roku, w związku z decyzją Rady Wydziału z dnia 21 lutego 2017 roku. Recenzja została przygotowana w formie odpowiadającej kwestionariuszowi przesłanemu przez Zleceniodawcę.

**1. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Tematyka rozprawy skupia się wokół dwóch związanych ze sobą zagadnień: algorytmów predykcji międzyobrazowej stosowanych w kodowaniu wideo oraz sprzętowej implementacji i akceleracji algorytmów. W szczególności Autor podejmuje temat optymalizowania algorytmów predykcji używanych w rozwijanym obecnie standardzie kodowania określonym protokołem H.265 (HEVC), który swoje główne zastosowanie znajduje w kodowaniu obrazu 4K-UHD (*Ultra-High-Definition*). Tematyka związana z implementacją standardu H.265 jest ważna i aktualna. Praca ma charakter interdyscyplinarny: kodowanie obrazu klasyfikowane jest najczęściej w dyscyplinie informatyka jednak architektury i implementacje w układach ASIC i FPGA są zagadnieniami naukowymi klasyfikowanymi w dyscyplinie elektronika.

Autor deklaruje, że celem rozprawy jest analiza istniejących algorytmów predykcji międzyobrazowej i opracowanie algorytmu, który umożliwiłby kodowanie obrazu 4K przy wykorzystaniu układów ASIC i FPGA. Zakłada przy tym, że mimo koniecznych uproszczeń, wyniki działania algorytmu i jego efektywność sprzętowa będą jedynie nieznacznie odbiegać od algorytmu pełnego przeszukiwania.



Sformułowana została następująca teza rozprawy:

1. Wykorzystanie wektorów ruchu znalezionych dla bloków 8x8 z dokładnością pikselową (całkowitoliczbowa) do estymacji ruchu w większych blokach pozwala na znaczną redukcję złożoności obliczeniowej przy niewielkim zmniejszeniu efektywności kompresji.
2. Ograniczenie ułamkowej estymacji ruchu do bloków kwadratowych wraz z wykorzystaniem list rankingowych sprawdzanych kandydatów umożliwia znaczną redukcję złożoności obliczeniowej przy niewielkim zmniejszeniu efektywności kompresji.
3. Architektury sprzętowe algorytmu estymacji i kompresji ruchu o cechach opisanych w poprzednich tezach mogą osiągać wysoki stosunek przepustowości do ilości zasobów sprzętowych.

Recenzent nie ma zastrzeżeń do sformułowania celu i tezy rozprawy.

W rozprawie można wskazać zarówno wątki teoretyczne jak i eksperymentalne. Do pierwszej grupy należy zaliczyć analizę algorytmów predykcji międzyobrazowej. Analiza została uzupełniona wynikami eksperymentów przeprowadzonych przy pomocy koderów programowych. Z przedstawionych w części teoretycznej wniosków, wynikają kierunki dalszych prac doświadczalnych, polegających na opracowaniu architektur elementów obliczeniowych, ich walidacji przy pomocy narzędzi symulacyjnych do projektowania układów ASIC i FPGA.

W pracy przedstawiony został oryginalny dorobek Autora zarówno po stronie teoretycznej – dokonanie przeglądu literatury i szczegółowej analizy oraz sformułowanie wymagań, które powinny być spełnione przez algorytmy przeznaczone do implementacji sprzętowej, jak również po stronie doświadczalnej – przebadanie przy pomocy oprogramowania referencyjnego wpływu rozmaitych parametrów na jakość wyników kodowania wideo i przede wszystkim implementacji wybranych algorytmów w językach opisu sprzętu oraz ich walidacji dla układów ASIC i FPGA. W opinii recenzenta większe znaczenie ma przedstawiony dorobek eksperymentalny. (Prace naukowe – doktorskie czy habilitacyjne – z dziedzin informatyki i elektroniki, w których dominuje dorobek eksperymentalny, są obecnie dość często spotykane i co warto podkreślić, dobrze przyjmowane nie tylko w Polsce, ale również w Europie i na świecie.)

**2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle /świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?**

Rozprawa zawiera omówienie źródeł literatury w kilku grupach zagadnień. Najbardziej obszerny fragment przeglądu literatury zawarty jest w rozdziale 3 i dotyczy różnych algorytmów predykcji międzyobrazowej. Recenzent chciałby podkreślić, że nie spotkał się w literaturze polskojęzycznej z inną pozycją, która by w sposób stosunkowo dokładny i systematyczny omawiała algorytmy predykcji międzyobrazowej. Dokonana w pracy analiza jest więc cennym źródłem informacji na ten temat, choć z drugiej strony, w części eksperymentalnej wykorzystano zaledwie niewielkie fragmenty tych rozważań.

W pracy nie zabrakło przeglądu rozwiązań sprzętowych związanych z implementacją algorytmów kodowania H.264 i H.265 w układach FPGA i ASIC (rozdział 3.3.2). Liczba artykułów dotyczących implementacji sprzętowych (tu: ASIC, FPGA) standardu H.265



cytowanych w pracy jest niewielka co, jak się wydaje, odpowiada stanowi faktycznemu na moment przygotowania pracy. W pracy słusznie zauważono, że ze względu na istotne różnice w sposobie hierarchizacji danych pomiędzy standardami H.264 i H.265 (w tym drugim dane grupowane są w strukturach – zwanych drzewami CTU *Coding Tree Unit* – większych i bardziej różnorodnych niż bloki wykorzystywane w standardzie H.264) metody predykcji międzyobrazowej stosowane we wcześniej opracowanym i obecnie szeroko stosowanym standardzie H.264 nie mogą być wprost przeniesione do predykcji w standardzie H.265.

W końcowej części rozprawy, w rozdziale 5.7, dokonano porównania parametrów zaproponowanego w pracy rozwiązania z architekturami kodera H.265 opisanymi w literaturze. W tekście tego rozdziału opisano pewne wady i ograniczenia ilościowe wcześniejszych rozwiązań. Niektóre fragmenty tego rozdziału, dotyczące cech jakościowych architektur opisanych w literaturze, zdaniem recenzenta, lepiej byłoby umieścić już w rozdziale 3.3.2. Wnioski z rozważań z rozdziału 5.7 powinny być (i niewykluczone, że były) brane pod uwagę w trakcie decyzji projektowych opisanych w rozdziale 4.

Z kolei w rozdziale 4.11 zaprezentowano literaturę opisującą programowe implementacje koderów standardu H.265 i przedstawiono ich efektywności.

Podsumowując, należy ocenić, że analiza źródeł została sporządzona z dużą starannością i dobrze oddaje stan wiedzy z zakresu omawianych zagadnień na moment złożenia pracy.

### **3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?**

Po przedstawieniu w rozdziale 1 założeń i celów, dokonaniu wprowadzenia w rozdziale 2 i dokonaniu przeglądu literatury (rozdział 3), w kolejnych rozdziałach pracy przedstawiono sposób realizacji postawionych celów i dowody tez.

W rozdziale 4, zaprezentowano rozważania, które na drodze metodycznej analizy doprowadzają do wyselekcjonowania algorytmów predykcji międzyobrazowej, które możliwe są do efektywnej implementacji w zasobach sprzętowych. Droga prowadząca do uzyskania ostatecznego rozwiązania była ewolucyjna. Proponowane kolejne rozwiązania – rozmaite warianty algorytmów – implementowano w oprogramowaniu HM. Porównywano wyniki dla sekwencji testowych pochodzących z baz JCT-VC, UVG oraz SVT, stosując opisane w rozdziale 2.6.2 miary jakości i efektywności BD-PSNR BD-RATE (tzw. delty Bjøntegaarda), nie biorąc pod uwagę na tym etapie uzyskiwanych czasów kodowania. Słusznie uzasadniono, że uzyskiwane czasy kodowania nie będą adekwatne do czasów uzyskiwanych na docelowych platformach implementacji (ASIC, FPGA). Przyjętą w pracy metodologię należy uznać za właściwą.

W rozdziałach 4.1.2 i 4.1.3 prawidłowo wskazano na ograniczenia, które należy przyjąć podczas procesu wyboru właściwego algorytmu do implementacji sprzętowej. W dalszej części tego rozdziału badano wpływ struktury grupy obrazów na jakość kodowania, wpływ właściwego doboru wielkości obszaru przeszukiwania, wyniki całkowitoliczbowej estymacji ruchu w zależności od podziału struktur CTU na mniejsze bloki, jakość estymacji ruchu z ułamkową dokładnością w kilku wariantach, jakość kodowania dla bloków kodowanych (CU) i bloków predykcji (PU) o różnym rozmiarze oraz bardziej całościowo, efektywność kodowania w zaproponowanych pracy wariantach, łączących wcześniej przebadane rozwiązania. Dla każdego kroku, w którym dokonywano wyboru wariantu algorytmu lub selekcji parametrów, przedstawiona została argumentacja uzasadniająca podejmowane decyzje projektowe. W rozdziale 4.10 przedstawiono trzy rozwiązania wyłonione spośród licznych rozpatrywanych



przez Autora wariantów i porównano je do niezmodyfikowanego kodera referencyjnego HM 16. W rozdziale 4.11, opisano wyniki badania trzech zaawansowanych koderów standardu H.265, przeznaczonych do wydajnych implementacji w zasobach programowych, z tym, że tym razem dodatkowym parametrem mającym wpływ na wyniki miar BD była szybkość kodowania. Podsumowując wyniki osiągnięte w rozdziale 4, warto zauważyć, że schemat próbkowania opisany w rozdziale 4.4, w którym odwrócona została kolejność przeszukiwania bloków ze schematu od największych do najmniejszych na odwrotny (co pozwoliło zredukować ilość przeszukiwanych wariantów wektorów ruchu, ale przede wszystkim uprościło implementację w zasobach sprzętowych) wypełnia założenia pierwszej tezy rozprawy, mimo iż odnotowano znaczne odchylenia miary BD RATE dla pojedynczych sekwencji.

Dowód drugiej z tez rozprawy okazał się bardzo pracochłonny i skomplikowany. Autor, poszukując docelowego algorytmu, poruszał się niczym w labiryncie złożonej gry komputerowej o bardzo wysokim poziomie trudności, w której gracz napotyka na liczne pułapki: znajdując dobrze rokującą ścieżkę postępowania podąża za nią, po czym okazuje się, że napotyka miejsce, w którym skazany jest na uzyskanie wielu karnych punktów i musi cofnąć się i podążać inną drogą, i tak wielokrotnie. Dla laika labirynt ten byłby nie do pokonania. Autor jednak używa całej swojej wiedzy i doświadczenia w zakresie kodowania w standardzie H.265 oraz doświadczenia zyskanego w projektowaniu zasobów sprzętowych i kojarzy ze sobą różne fragmenty procesu kodowania, tak aby unikać ślepych dróg i ostatecznie, po długim i skomplikowanym dowodzie, udaje mu się zaproponować rozwiązanie spełniające założone kryteria – uzyskuje rezultat spełniający wymagania drugiej tezy rozprawy.

Z punktu widzenia podjętego w pracy doktorskiej zadania ważne jest osiągnięcie rezultatu dowodzącego tezę. Dla nauki i techniki, nie mniej cenny wydaje się zawarty w tej części rozprawy opis przebytej drogi, a zwłaszcza uzasadnienia decyzji projektowych na każdym etapie poszukiwań. Być może, jak to często bywa, inni badacze podążą wytyczonym przez Autora szlakiem.

W rozdziale 5 przedstawiono projekty architektur specjalizowanych elementów obliczeniowych, realizujących fragmenty algorytmów kodera H.265: interpolację oraz estymację ruchu w wersji całkowitoliczbowej i ułamkowej oraz przedstawiono metodę i architekturę do wydajnej komunikacji z zasobami pamięci.

Zaprojektowane architektury zostały zaimplementowane w języku VHDL. Przedstawiono wykresy symulacji post-implementacyjnej. Dokonano syntezy logicznej opracowanych architektur w oprogramowaniu Synopsys Design Compiler dla procesu produkcyjnego TSMC 90 nm oraz przy wykorzystaniu pakietu Altera Quartus II dla rodziny Aria II GX układów FPGA firmy Intel (dawniej Altera). Za najbardziej wartościowe osiągnięcie w tym rozdziale należy uznać propozycje architektur do całkowitoliczbowej i ułamkowej estymacji ruchu oraz rozważania dotyczące rozdzielenia ścieżek całkowitoliczbowej i ułamkowej zawarte w podrozdziałach 5.4 i 5.5. Rozważania w podrozdziale 5.7, wskazują, że rozwiązania przedstawione w podrozdziałach 5.4 i 5.5 (dla analizowanej miary BD-RATE) osiągają wyniki porównywalne lub lepsze od rozwiązań opublikowanych w renomowanych czasopiśmie i w materiałach uznanych konferencji międzynarodowych w latach 2013-2015, w tym lepsze od zaprezentowanych w pracy Jou et al. wydanej w roku 2015 w bardzo wysoko punktowanym i prestiżowym czasopiśmie *IEEE Transactions on Circuits and Systems for Video Technology*.

W podsumowaniu doktoratu odniesiono się bezpośrednio do realizacji 3 tezy rozprawy. Spośród prac, do których Autor porównywał swoje wyniki, w pracy Afonso et al. przedstawiono wyniki o lepszym stosunku przepustowości do zasobów, nie mniej jednak,



w porównaniu z pozostałymi rozwiązaniami z literatury, to wyniki ujęte w doktoracie są wyraźnie lepsze. Warto podkreślić, że rozwiązania zaproponowane przez Autora, oferują dodatkowe funkcjonalności, w tym jako jedyne, mogą realizować tryb *merge*. Jak wiadomo zwiększenie funkcjonalności wpływa zawsze na wzrost zasobów i stąd dość trudno sprowadzić do wspólnego mianownika i bezpośrednio porównać wszystkie aspekty przedstawionych rozwiązań. Przedstawione w rozdziale 5.7 porównanie wyników jest podstawą do uznania, że tezę 3 rozprawy również należy uznać za udowodnioną.

#### **4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?**

Najbardziej wartościowym osiągnięciem przedstawionym w rozprawie są zaproponowane architektury elementów obliczeniowych (modułów) przeznaczonych do wykonania estymacji ruchu dla wektorów całkowitoliczbowych i ułamkowych. Wyniki są porównywalne lub lepsze od opublikowanych na wiodących konferencjach i zamieszczonych w czasopismach indeksowanych JCR. Do innych, mniej znaczących, ale także wartościowych wyników można zaliczyć: rekonfigurowalną architekturę interpolatora oraz zaproponowany sposób organizacji pamięci dla modułu estymacji ruchu. Powstanie wszystkich architektur było możliwe dzięki przeprowadzeniu przez Autora analizy poszczególnych fragmentów procesu kodowania i zestawieniu ich w zaproponowanym algorytmie w sposób umożliwiający efektywne wykonanie operacji w zasobach sprzętowych.

#### **5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?**

Przechodząc do oceny rozprawy od strony redakcyjnej, należy wskazać, że sposób prowadzenia wywodu w rozprawie jest na dobrym poziomie. We wprowadzeniu przedstawiono motywacje do podjęcia prac oraz sformułowano cel, zakres i tezę pracy. Rozdział 2 przedstawia metody kodowania sygnału wideo ze szczególną uwagą poświęconą na redukcję nadmiarowości czasowej danych. W rozdziale 3 dokonano przeglądu algorytmów predykcji międzyobrazowej. Oba rozdziały należy uznać za wprowadzenie do omawianych w pracy zagadnień. Kolejne rozdziały zawierają wyniki prac autorskich oraz opisane we wcześniejszej części recenzji dowody kolejnych tez rozprawy.

Autor prawidłowo buduje narrację pracy i poprawnie konstruuje dowody swoich tez. Mimo dużej złożoności opisywanych metod potrafi zwrócić uwagę czytelnika na fragmenty bardziej istotne. W niektórych fragmentach rozprawy następuje bardzo drobiazgowo przedstawianie rozmaitych aspektów i szczegółów. Ponieważ rozważania te są częścią dowodu, trudno z tego czynić zarzut, ale być może część opisu można by przesunąć do dodatków, bez zasadniczego obniżenia walorów pracy, za to zwiększając płynność jej czytania.

Pod względem językowym rozprawa stoi na dobrym poziomie. Recenzent zwrócił jednak uwagę na kilka błędów w stosowaniu pojęć oraz wskazał drobne usterki, które nie zostały dostrzeżone podczas korekty rozprawy.

Nieprawidłowe posługiwanie się pojęciami:

Str. 5 – rysunek 2.1 – etap akwizycji obrazu w torze wizyjnym błędnie nazwano „analizą”, a wizualizację określono jako „syntezę”. Pierwszy z określeń można odnieść co prawda do sposobu akwizycji obrazu w tzw. „lampach analizujących”, ale technologia ta została wyparta



przez czujniki CCD i CMOS już ponad 25 lat temu. Pojęcie analizy obrazu współcześnie zarezerwowane jest dla etapu pozyskiwania informacji (tworzenia opisu) z postaci rastrowej. Synteza z kolei budzi skojarzenie z syntezą obrazu w grafice komputerowej (w karcie graficznej), ale synteza obrazu nie występuje w cyfrowym torze przetwarzania obrazu, gdzie wartości obrazu nie są zapisywane w formie podobnej do stosowanej w grafice wektorowej.

Str. 5 – użyto pojęcia głębia obrazu, opisując rozdzielczość ze względu na liczbę poziomów jasności (ponownie we wzorze (4.1) str. 68). O ile dopuszczalne wydaje się użycie wyrażenia głębia kolorów, to „głębia obrazu” budzi skojarzenie z pojęciem głębi ostrości obrazu.

Str. 105 – odnosząc się do układów FPGA, Autor stwierdza: „Często są one rekonfigurowalne, tzn. ich strukturę można zmieniać w trakcie pracy.” Zaproponowane zdefiniowanie pojęcia rekonfigurowalności dla układów FPGA świadczy o niezajomości stosowanych w literaturze pojęć w omawianym zakresie. Immanentnymi cechami układów FPGA są reprogramowalność i rekonfigurowalność. Dotyczy to zarówno zasobów logicznych (decydujących o funkcjonalności) i połączeniowych (mających wpływ na architekturę). Rekonfigurowalność układów FPGA nie dotyczy jedynie możliwości przełączania ścieżek danych przy pomocy multiplexerów (tak jak napisano w rozprawie, definiując na str. 109 „filtr rekonfigurowalny”), lecz odnosi się do możliwości zmiany zawartości komórek pamięci konfiguracji SRAM, definiującej sposób pracy zasobów FPGA. Do określenia możliwości rekonfigurowania układów FPGA w trakcie ich działania (po zaprogramowaniu) używa się pojęć rekonfigurowalność dynamiczna (*dynamical reconfiguration*), a niekiedy też rekonfigurowalność częściowa (*partial reconfiguration*). Nota bene pojęcia te nie są całkowicie wymienne, ponieważ istnieją układy FPGA, w których można wymieniać dynamicznie całą zawartość (kontekst) układu. Są to tzw. układy rekonfigurowalne wielokontekstowe – *multi-context FPGA*.

Usterki językowe:

Str. 19 – zdanie niejasno sformułowane ... blok wypełniany jest wartościami pikseli według jednego **określonych standardem trybów ekstrakcji**.

Str. 68 – kolokwializm, żargon: zaimplementowane w **krzemie...**, dla pojedynczej **kości** pamięci.

Str. 68 – niepoprawna forma: ... co **wyeliminowuje...**

Str. 108, 110 – wymienne użycie różnych form dopełniacza słowa filtr: filtra i filtru (obie formy poprawne, ale wymienne użycie niewskazane)

Str. 111 – **ilość** pamięci – wielkość pamięci

Str. 114 – **próbek próbek**

Skład książki jest zrobiony przy pomocy narzędzi profesjonalnych, szata graficzna na dobrym poziomie. Recenzent zauważył następujące usterki edycyjne:

1. Wykaz skrótów jest wybiórczy – brak użytych w pracy pojęć, przykładowo: GOP, DFD, JPEG, CCD, CMOS i wielu innych.
2. Sposób umieszczania w tekście pierwszych odwołań do rysunków lub tabel jest w niektórych miejscach niepoprawny – rysunki występują na stronach wcześniejszych niż pierwsze odwołanie, które ich dotyczy: rys. 2.20, rys. 2.31, tab. 2.4, tab. 2.5, rys. 5.11.



3. Błędne odwołanie – str. 24 (rysunek 2.35 powinno być 2.25).
4. Stosowanie zbyt małej czcionki na rysunkach, która czyni opisy nieczytelnymi: rys. 2.33, rys. 2.34, rys. 5.26.
5. Brak opisu osi na wykresach: rys. 2.41, 6.1.
6. Brak opisu konwencji oznaczeń na rys. 5.3, 5.4, 5.5 (symbol << nie jest oczywisty).

#### **6. Jakie są słabe strony rozprawy i jej główne wady?**

Recenzent nie ma wątpliwości, że przedstawiony doktorat wiązał się z bardzo solidnie wykonaną pracą koncepcyjną i eksperymentalną. Został dobrze napisany, a wyniki opublikowane na konferencjach międzynarodowych i w wydaniu specjalnym czasopisma Journal of Real-Time Image Processing poświęconym tematyce kodowania wideo. To dostateczna nobilitacja i potwierdzenie wysokiej wartości wykonanych prac.

W rozprawie przedstawiono walidację w postaci wyników symulacji post-implementacyjnych. Recenzent nie dostrzegł jednak, dowodów praktycznej weryfikacji zaprojektowanych architektur, które podniosłyby walory doktoratu. Z tego wynikają postawione poniżej pytania. Czy opisana w rozdziale 5 koncepcja przydziału pamięci została zweryfikowana? Czy uruchomiony został (choćby dla opracowanych fragmentów algorytmu koder HEVC), proces kodowania w czasie rzeczywistym dla transmitowanego z kamery strumienia obrazu, czy posługiwano się tylko analizą w koderze programowym i symulacjami? Czy, wykorzystując przykładowo kartę z układem FPGA, sprawdzono, iż koder pracuje z założoną częstotliwością? Jaką metodę zastosowano, aby zebrać dane w tabeli 5.1, przedstawiającej wyniki efektywności koder sprzętowego w implementacji ASIC?

#### **7. Jaka jest przydatność rozprawy dla nauk technicznych?**

Metody kodowania wideo są bardzo złożonymi algorytmami. Zaledwie kilka ośrodków w kraju potrafi prowadzić badania w tym zakresie na światowym poziomie. Realizacja takich badań na poziomie doktoratów, pozwala na poszerzanie grona specjalistów w tej dziedzinie. Wobec ogromnego rozwoju technik przekazu sygnału wideo, powstawanie prac w tym zakresie ma fundamentalne znaczenie dla współczesnych nauk technicznych, a zwłaszcza, że standardy kompresji wideo są esencją współczesnej telewizji cyfrowej i w bardzo ważnym elemencie aplikacji multimedialnych. Implementacja tych operacji w układach ASIC i FPGA wymaga wiedzy interdyscyplinarnej, a w przypadku praktycznego zweryfikowania oraz zastosowania tych metod, może dawać unikalne możliwości aplikacyjne.

Dokonanie przez Autora dogłębnej analizy możliwości implementacji wybranych fragmentów koder HEVC w zasobach sprzętowych (ASIC, FPGA) oraz zaproponowanie ich architektur jest wartościowym i oryginalnym osiągnięciem twórczym.

#### **8. Do której z następujących kategorii Recenzent zalicza rozprawę:**

- a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania
- c/ spełniająca wymagania
- d/ spełniająca wymagania z wyraźnym nadmiarem

**e/ wybitnie dobra, zasługująca na wyróżnienie**

Dokonując podsumowania recenzji, stwierdzam, że mimo uwag dyskusyjnych i krytycznych, które mają drugorzędne znaczenie, przedstawiona rozprawa Pana mgra inż. Macieja Trochimiuka spełnia wymagania ustawowe stawiane rozprawom doktorskim z wyraźnym nadmiarem i wnioskuję o dopuszczenie Autora do dalszych etapów przewodu doktorskiego.

W przypadku przedstawienia przez Doktoranta satysfakcjonujących odpowiedzi na pytania zawarte w części 6, recenzent rozważy postawienie wniosku o wyróżnienie rozprawy.

A handwritten signature in blue ink, consisting of several fluid, connected strokes. The signature is positioned in the right-center of the page.



Białystok, 24 maja 2017 r.

dr hab. inż. Marek Parfieniuk

Katedra Mediów Cyfrowych i Grafiki Komputerowej  
Wydział Informatyki  
Politechnika Białostocka

*KWESTIONARIUSZ- RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY  
WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMACYJNYCH  
POLITECHNIKI WARSZAWSKIEJ*

**Tytuł rozprawy: Optymalizacja predykcji międzyobrazowej w kodowaniu danych wizyjnych**

**Autor rozprawy: mgr inż. Maciej Trochimiuk**

**1. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?**

Przedmiotem rozprawy są algorytmy i układy cyfrowe do międzyobrazowej predykcji klatek wideo na potrzeby kodowania, czyli kompresji danych. Celem predykcji międzyobrazowej jest estymowanie wartości pikseli z aktualnie kodowanej klatki na podstawie sąsiednich klatek, które zostały zakodowane wcześniej. W sekwencjach wideo kolejne obrazy są na ogół bardzo podobne, gdyż radykalne zmiany ujęcia występują względnie rzadko. Zwykle kamera powoli przemieszcza się względem obserwowanych obiektów, a obiekty, lub tylko niektóre ich fragmenty, przemieszczają się stopniowo względem kamery. Wobec tego, zamiast poddawać kompresji niezależnie klatkę po klatce, korzystniej jest kodować różnicę między klatką a jej estymatą. Jeżeli predykcja się udaje, to różnica przyjmuje głównie wartości bliskie zeru, czyli można ją opisać zwartymi kodami entropijnymi, stosując m.in. kodowanie arytmetyczne oraz kodowanie długości serii.

Różnice między klatkami mają zbyt złożoną naturę i zbyt nierównomierny rozkład, by można było opisywać je analitycznie, z wykorzystaniem prostych modeli predykcyjnych. Praktyczne kodery wideo wykorzystują estymację ruchu, która polega na dzieleniu kodowanego obrazu na fragmenty, jednostki predykcji, i określaniu dla każdego fragmentu wektora ruchu: wskazania na obszar klatki wzorcowej, który jest podobny do danego fragmentu. Wektory ruchu opisuje się kodami entropijnymi, które koder umieszcza w wynikowym strumieniu binarnym razem z kodami opisującymi odpowiedni obraz różnicowy.

O skuteczności kodera decyduje w dużej mierze jego zdolność do ustalenia właściwej proporcji między liczbą bitów poświęconych na opisanie obrazu różnicowego a liczbą bitów przeznaczonych do opisanie wektorów ruchu. Tę proporcję koder optymalizuje głównie dobierając sposób podziału obrazu: w statycznych obszarach wykorzystuje duże jednostki predykcji, czyli mniej wektorów ruchu, a napotykać problemy z dokładnością predykcji,



zastępuje dużą jednostkę kilkoma mniejszymi, czyli zwiększa liczbę wektorów. Dodatkowo koder może zwiększać upakowanie danych wykorzystując za podstawę estymacji ruchu piksele ułamkowe, wartości interpolowane z oryginalnych pikseli obrazu odniesienia. Koder może też odstępować od predykcji międzyobrazowej i stosować predykcję wewnątrzobrazową, czyli estymować piksele na podstawie wcześniej zakodowanych fragmentów tej samej klatki.

Dobranie najkorzystniejszego sposobu podziału obrazu i ustalenie dla każdej jednostki podziału najlepszego sposobu predykcji, parametrów estymacji oraz pikseli wzorcowych stanowią trudne zadanie optymalizacji, które nie zawsze ma jednoznacznie określone rozwiązania. Praktyczne kodery wybierają rozwiązanie testując podzbiór wszystkich możliwych podziałów określony pewną heurystyką. Nawet przy takim podejściu predykcja międzyobrazowa jest złożonym problemem obliczeniowym i dlatego od lat prowadzone są badania nad metodami sprawnego określania wektorów ruchu, które są optymalne lub zbliżone do optymalnych, oraz nad wysokowydajnymi sprzętowymi rozwiązaniami do estymacji ruchu.

Ponieważ nie znaleziono dotąd definitywnych rozwiązań problemów związanych z estymacją ruchu, tematyka rozprawy jest bardzo aktualna. Prace nad predykcją międzyobrazową, które doktorant podjął, są zbieżne z obydwojema kierunkami badawczymi, o których wspomniałem wyżej. Postanowił on opracować nowe udoskonalone algorytmy do określania wektorów ruchu i wysokowydajne architektury sprzętowe do implementowania tych algorytmów.

Autor rozprawy wyraźnie sformułował trzy tezy, których lista stanowi zawartość podrozdziału 1.3. Tezy zostały sformułowane czytelnie, podobne jak cele i zakres pracy. Można mieć jedynie zastrzeżenie, że poprzedzający je podrozdział 1.1, który ma wyjaśniać istotę podejmowanego problemu i pobudki autora, jest krótki i ogólny, stanowiąc niezbyt solidne tło szczególnie dla tezy trzeciej, która dotyczy implementacji sprzętowych.

Pierwsza teza dotyczy tego, że wektory ruchu znalezione dla bloków  $8 \times 8$  pikseli można wykorzystać do estymacji ruchu większych bloków i w ten sposób znacznie zredukować złożoność obliczeniową kodowania, pogarszając tylko nieznacznie skuteczność kompresji. Druga teza dotyczy tego, że złożoność obliczeniową kodowania można znacząco zmniejszyć bez istotnego pogarszania skuteczności kompresji, dopuszczając ułamkową estymację ruchu tylko na blokach kwadratowych i wybierając wektory ruchu z list rankingowych. Trzecia teza dotyczy tego, że podejścia ujęte w dwóch pierwszych tezach mogą stanowić podstawę sprzętowych implementacji predykcji międzyobrazowej, które charakteryzują się przepustowością dużą w stosunku do zapotrzebowania na pamięć i powierzchnię układu scalonego.

Rozprawa ma charakter teoretyczno-projektowo-doświadczalny, ponieważ osiągnięcie założonych celów i dowiedzenie słuszności tez było uwarunkowane opracowaniem przez autora nowych koncepcji algorytmicznych, zaprojektowaniem układów cyfrowych i przeprowadzeniem eksperymentów uzasadniających założenia i dowodzących użyteczności zaproponowanych rozwiązań

W rozprawie autor skupił swoją uwagę na standardzie HEVC/H.265, co jest zasadne, chociaż w pracy nie można znaleźć wystarczających argumentów do ograniczenia zakresu badań do jednego koderu. W rozdziale 1.2 doktorant stwierdza tylko pobieżnie, że HEVC jest standardem najnowszym, ukierunkowanym na sekwencje wideo o bardzo wysokich rozdzielczościach i udostępniającym wiele trybów, a mnogość tych trybów utrudnia kodowanie szybkie i



skuteczne. Nie zostało przy tym wyjaśnione, że chodzi o tryby predykcji międzyobrazowej, a nie jest to wcale oczywiste z kontekstu.

Autor powinien użyć konkretniejszych argumentów. Np. powiedzieć, że chociaż problem implementacji sprzętowej dekodera HEVC można uznać za rozwiązany, bo na rynku jest szereg urządzeń, np. niedrogie telewizory klasy UHD, z funkcją odtwarzania strumieni H.265, to otwarte pozostaje zagadnienie opracowania kodera, który maksymalnie upakuje dane, działając w czasie rzeczywistym, wykorzystując jak najmniejszą pamięć i dając się zaimplementować jako układ scalony o niewielkiej powierzchni i poborze mocy. Poza tym rozwiązania opracowane na podstawie HEVC z dużym prawdopodobieństwem znajdują zastosowanie w przyszłych standardach kodowania wideo

## **2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle /świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?**

W przedstawionej mi rozprawie bibliografia liczy 133 pozycje, przy czym w tekście znajdują się powołania na wszystkie pozycje. Dobór źródeł świadczy o tym, że doktorant jest bardzo dobrze zorientowany w literaturze dotyczącej kodowania danych wizyjnych, a szczególnie zagadnień, których dotyczy rozprawa: algorytmów i rozwiązań sprzętowych związanych z predykcją międzyobrazową

Z treści pracy wynika, że autor uważnie przestudiował większość źródeł i na tej podstawie zidentyfikował niedostatki, których powinien ustrzec się w swoich rozwiązaniach, oraz podejścia, które warte były zbadania i zaadoptowania.

Przegląd znanych algorytmów związanych z predykcją międzyobrazową znajduje się w rozdziale 3 rozprawy. Doktorant najpierw wyczerpująco omawia znane algorytmy estymacji ruchu dla pikseli całkowitych, następnie nieco pobieżnie opisuje algorytmy estymacji ruchu z wykorzystaniem pikseli ułamkowych, a na koniec omawia dokładnie algorytmy wyboru sposobu-trybu podziału jednostek kodowania. Ta część rozprawy mogłaby kończyć się jakimś podsumowaniem i zestawieniem wniosków. Autor poprzestał jednak na charakteryzowaniu poszczególnych rozwiązań

Architekturom sprzętowym do predykcji międzyobrazowej poświęcony został paragraf 3.3.2, którym doktorant kończy omówienie algorytmów wyboru sposobu podziału jednostek kodowania. Autor przeanalizował 5 rozwiązań, które dotyczą standardu HEVC, oraz jedno rozwiązanie, które dotyczy standardu H.264/AVC. Wydaje mi się, że przegląd stanu wiedzy o architekturach powinien zostać bardziej wyeksponowany, jako podrozdział wyższego poziomu. Mogłby też być szerszy: obejmować większą liczbę rozwiązań oraz przedstawiać trochę syntetycznych informacji o ogólnych podejściach stosowanych w projektowaniu sprzętowych koderów wideo.

Chociaż nie mam wątpliwości, że bibliografia obejmuje wszystkie publikacje naukowe, w których opisano istotne osiągnięcia powiązane z tematem rozprawy i przydatne dla doktoranta, to dostrzegłem pewne uchybienia w doborze literatury.



Przed wszystkim w bibliografii brakuje mi podstawowych dokumentów ISO/ITU oraz pozycji książkowych, które dotyczą standardu HEVC oraz predykcji międzyobrazowej, czyli:

- [A] ITU, *ITU-Recommendation H.265: High efficiency video coding*, 2013
- [B] ISO, *ISO/IEC 23008-2:2013:Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding*
- [C] Vivienne Sze, Madhukar Budagavi, Gary J. Sullivan: *High Efficiency Video Coding (HEVC): Algorithms and Architectures*, Springer, 2014
- [D] Mathias Wien: *High Efficiency Video Coding: Coding Tools and Specification*, Springer 2015
- [E] Indrajit Chakrabarti, Kota Naga Srinivasarao Batta, Sumit Kumar Chatterjee: *Motion Estimation for Video Coding: Efficient Algorithms and Architectures*, Springer 2015

Nie sposób przyjąć, że doktorant nie korzystał z dokumentów ITU/ISO pracując nad rozwiązaniami do kodera HEVC. Prawdopodobnie pominął w bibliografii dokumentację standardu, zakładając, że czytelnik zdolny do zrozumienia rozprawy z tej dziedziny jest świadomy istnienia tych dokumentów. Założenie takie jest racjonalne, ale z formalnego punktu stanowi uchybienie, szczęśliwie nieistotne dla jakości rozprawy.

Brak powołania na dokumenty ISO/ITU utrudnia stwierdzenie, jakiej wersji standardu HEVC/H.265 dotyczy rozprawa. Standardy wideo podlegają modyfikacjom, które mogą być znaczące, szczególnie w pierwszych latach po wprowadzeniu normy. Aktualnie istnieją 4 wersje dokumentów ITU, z lat 2013-2016, które dotyczą H.265. Doktorant nie pisze nigdzie, którą wersję wybrał do badań. Czytelnik może to określić tylko pośrednio, na podstawie wersji referencyjnego oprogramowania HM 16.0, z którego korzystał autor rozprawy.

Odnośnie książek, to wypadało odnotować ich istnienie, nawet jeżeli nie były przydatne w badaniach.

Ponadto w rozprawie nie znalazłem ani jednej wzmianki o patentach związanych z kodowaniem wideo, a patenty odgrywają bardzo dużą rolę w rozwoju tej dziedziny. Standardy, w tym HEVC, które rozwija grupa ekspercka MPEG pod egidą organizacji standaryzacyjnych ISO i ITU, wykorzystują rozwiązania, które zostały opatentowane. Z kolei Google i SMPTE rozwijają odpowiedniki standardów MPEG, unikając opatentowanych rozwiązań. Niektóre opatentowane rozwiązania zostały przedstawione w publikacjach naukowych, w materiałach konferencji oraz w czasopiśmie, ale sporo rozwiązań pozostaje udokumentowanych tylko opisem wynalazku.

Patentów związanych z predykcją międzyobrazową jest bardzo dużo, a opisy wynalazków z tej dziedziny są często dokumentami złożonymi i skupionymi na niuansach, więc trudno oczekiwać, by doktorant zamieszczał w rozprawie szczegółowy przegląd opatentowanych rozwiązań. Powinien on jednak wyszukać opisy wynalazków, które są zbliżone do jego rozwiązań, i określić relacje między pomysłami swoimi a opatentowanymi.

W rozprawie doktorant nie powołał się na żaden opis patentowy, więc nie sposób stwierdzić, czy wcale nie sprawdzał repozytoriów, takich jak „Google Patent Search”, czy też przeszukiwał je, ale nie znalazł żadnego rozwiązania, które warto byłoby wspomnieć w rozprawie. Wydaje mi się, że ta kwestia powinna zostać wyjaśniona podczas obrony.



### **3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?**

Moim zdaniem doktorant zrealizował wytyczone cele. Przeanalizował dokładnie stan wiedzy na temat predykcji międzyobrazowej i wykorzystał zdobyte w ten sposób wiadomości do opracowania nowych algorytmów, które są ukierunkowane na implementację sprzętową. Następnie opracował zoptymalizowane architektury do realizowania tych algorytmów w formie układów cyfrowych.

Swoje rozwiązania autor szczegółowo scharakteryzował oraz przebadął eksperymentalnie, skrupulatnie odnosząc wyniki do stanu wiedzy. Przedstawione w rozprawie rezultaty i wnioski nie pozostawiają wątpliwości, że postawione tezy zostały udowodnione. Ponadto zaproponowane rozwiązania okazały się konkurencyjne względem znanych wcześniej.

Wyraźnym niedostatkim rozprawy jest to, że tytuł i tezy zostały sformułowane zbyt ogólnie, mówią o rozwiązaniach do kodowania danych wizyjnych, a nie o rozwiązaniach do konkretnego kodeka, zaś opisane badania dotyczyły wyłącznie standardu H.265/HEVC kodowania wideo. Uważam, że właściwe było wybranie tego standardu na podstawie pracy, gdyż obecnie stanowi on najbardziej zaawansowane osiągnięcie z dziedziny kodowania wideo, ale w kontekście tytułu i tez, podczas obrony, doktorant powinien przedyskutować możliwości zastosowania swoich rozwiązań w kodekach wideo innych niż HEVC. Czy rozwiązania można zastosować bezpośrednio, czy wymagają adaptacji? Doktorant powinien też odnieść się do przyszłości algorytmów kodowania wideo.

Autor oparł swoje rozwiązania na spostrzeżeniu, że wśród opcji predykcji międzyobrazowej, których użycie jest dopuszczane przez standard HEVC, są takie opcje, które można zaimplementować sprzętowo wykorzystując pamięć o umiarkowanym rozmiarze i zajmując umiarkowaną powierzchnię układu scalonego, a osiągając dużą przepustowość. Doktorant założył jednocześnie, że znaczne ograniczenie zbioru opcji nie pogorszy istotnie skuteczności predykcji międzyobrazowej, ani skuteczności kodowania.

Spostrzeżenie uważam za bardzo trafne, a założenia za zasadne. To, że podejście użyte przez autora rozprawy okazało się skuteczne, nie wydaje mi się kwestią szczęścia. Treść rozdziału 4 świadczy, że z rozmysłem postawił on zagadnienia badawcze i wybrał metodę rozwiązania, bardzo dobrze poznawszy niuanse predykcji międzyobrazowej oraz problemy związane z jej implementowaniem w układach cyfrowych. Większość decyzji projektowych, ograniczeń i założeń została wyczerpująco uzasadniona, logicznymi wywodami oraz przekonującymi wynikami eksperymentów.

Ekspertyzy polegały głównie na porównywaniu skuteczności kodera wzorcowego do skuteczności kodera z ograniczonym zbiorem opcji predykcji międzyobrazowej. Niektóre ograniczenia autor narzucał odpowiednio modyfikując parametry konfiguracyjne kodera wzorcowego. Inne wymagały zaimplementowania zmodyfikowanego algorytmu predykcji i wbudowania go w we wzorcowy koder HM.

Dodatkowo autor porównał skuteczność kodera wzorcowego do skuteczności istniejących szybszych koderów programowych: NVENC, x265 i Kvazaar. Celem eksperymentów było uzyskanie danych, które pozwoliłyby odnieść te kodery do rozwiązań sprzętowych, a szczególnie do kodera z algorytmem predykcji opracowanym przez doktoranta. Uzyskane wyniki zostały zamieszczone w podrozdziale 4.11.





Autor mierzył skuteczność kodowania z użyciem delty Bjøntegaarda, która jest miarą uznaną i powszechnie stosowaną w badaniach koderów wideo, oraz z użyciem około 30 testowych sekwencji wideo, które są rekomendowane przez zespoły eksperckie zajmujące się rozwijaniem i standaryzacją kodeków wideo. Sekwencje różnią się rozdzielczością, częstotliwością, rodzajem ujęcia i charakterem ruchu obiektów obserwowanych przez kamerę

Nie mam wątpliwości, że uzyskane w ten sposób oceny koderów są obiektywne i stanowią cenne dane, które można wykorzystać w przyszłych badaniach. Należy zauważyć, że uzyskanie tak dużej ilości danych wymagało od autora wielkiego wysiłku, nawet jeżeli wykorzystał jakieś oprogramowanie do automatyzowania eksperymentów i obróbki wyników. W rozprawie mógłby znaleźć się fragment na ten temat, ale autor miał prawo zachować know-how dla siebie.

Implementacje sprzętowe opracowanych algorytmów i architektur autor rozwinął z użyciem języka VHDL, środowisk Synopsys i Altera, oraz biblioteki TSMC, czyli wybrał narzędzia odpowiednie i uznane, z których często korzystają naukowcy i inżynierowie.

Podsumowując, nie mam zastrzeżeń do metod i założeń, z których korzystał autor w badaniach, za wyjątkiem pewnych kwestii dotyczących implementacji algorytmów w technologiach FPGA i ASIC. Odpowiedniejszym miejscem do przedstawienia moich uwag jest jednak punkt 6 recenzji.

#### **4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?**

Ogólnym zamiarem autora było opracowanie algorytmów i architektur do predykcji wewnątrzobrazowej, które byłyby ukierunkowane na implementację sprzętową, wysokowydajną, ale wykorzystującą pamięć o umiarkowanym rozmiarze i zajmującą niewielką powierzchnię układu scalonego. Punktem wyjścia do opracowania takich algorytmów i architektur miało być ograniczenie zbioru dopuszczalnych opcji predykcji, poprzez rezygnację z większości opcji, które dopuszcza standard HEVC.

Te ogólne idee są znane z literatury, więc nie można ich zaliczyć do dorobku autora. Zaproponował on jednak nowe oryginalne sposoby urzeczywistnienia powyższych koncepcji, które pozwoliły mu opracować implementacje sprzętowe predykcji międzyobrazowej konkurencyjne względem znanych rozwiązań tego typu.

Podstawowym oryginalnym osiągnięciem autora jest określenie następującego zestawu głównych ograniczeń-uproszczeń predykcji międzyobrazowej:

- Wykorzystanie wektorów ruchu znalezionych w estymacji ruchu w blokach  $8 \times 8$  do większych bloków.
- Sprawdzanie tylko dwóch kandydatów w trybie "merge".
- Przeszukanie wszystkich pozycji ułamkowych w estymacji ruchu dla całkowitego wektora ruchu bez przekształcenia Hadamarda.
- Analiza wyłącznie trybu "merge" dla jednostki kodowania  $64 \times 64$  pikseli.





- Sprawdzanie wszystkich trybów podziału CU na PU oraz wektorów ruchu o ułamkowej dokładności z użyciem całkowitego wektora ruchu.

które pozwalają sprzętowo implementować estymację ruchu zgodną ze standardem HEVC, przy czym można osiągnąć dużą przepustowość, pomimo wykorzystania pamięci o umiarkowanym rozmiarze i zajęcia umiarkowanej powierzchni układu scalonego.

Każde z ograniczeń autor scharakteryzował z punktu widzenia projektanta układów cyfrowych oraz zbadał pod względem wpływu na skuteczność kodowania. Ograniczenia zostały najpierw zbadane oddzielnie, a następnie w połączeniu. Ponieważ okazało się, że w połączeniu ograniczenia znacząco obniżają skuteczność kodowania, doktorant zaproponował oryginalny sposób jej poprawienia.

Na podstawie znanych metod predykcji międzyobrazowej, ale zwracając większą uwagę na tryb "merge" niż autorzy wcześniejszych rozwiązań, opracował nowy pomysłowy algorytm estymacji ruchu, który kompensuje niekorzystny wpływ wyżej wymienionych ograniczeń na skuteczność standardowego kodera HM 16.0. Algorytm opiera się na następujących głównych postulatach

- Rozpatrywanie wspólnej listy rankingowej kandydatów (sposób podziału CUI i wektory ruchu) dla wszystkich trybów podziału
- Uzależnienie liczby kandydatów, która jest rozpatrywana w celu optymalizacji "rate-distortion", od rozmiaru jednostki kodowania CU.

Ponadto, opracowując architektury do sprzętowego implementowania wyżej wymienionych algorytmów, autor zaproponował szereg oryginalnych rozwiązań układowych:

- W znanych rozwiązaniach ten sam moduł pamięci przechowuje próbki z obrazu kodowanego oraz próbki z klatki odniesienia. Autor zaproponował, by predykcja odbywała się z wykorzystaniem dwóch modułów pamięci: jednego z próbkami aktualnie przetwarzanej jednostki CTU, a drugiego z próbkami z klatki referencyjnej. W ten sposób moduł estymacji i kompensacji ruchu będzie mógł jednocześnie pobierać z pamięci próbki obu rodzajów.
- W znanych architekturach próbki ułamkowe są wyznaczone przed rozpoczęciem estymacji ruchu, przy czym interpolacji poddawany jest cały obszar odniesienia. Autor zaproponował, by najpierw wyszukać wektory ruchu z dokładnością do całego piksela, a następnie zinterpolować tylko podobszary wskazywane przez te wektory i wykorzystać otrzymane próbki ułamkowe do poszukiwania wektorów dokładniejszych. W ten sposób można uniknąć wyznaczania pikseli ułamkowych, które do niczego się nie przydadzą, oszczędzając pamięć, cykle zegara i energię
- W znanych rozwiązaniach próbki referencyjne są przechowywane w jednym module pamięci. Autor zaproponował, by poszczególne próbki z bloku  $8 \times 8$  były umieszczane w 64 odrębnych modułach pamięci zorganizowanych w macierz  $8 \times 8$ . Pozwala to uzyskać dostęp do całego bloku próbek referencyjnych w jednym cyklu zegara.
- Autor zaproponował, by architektura do estymacji ruchu była wyposażona w dwie ścieżki przetwarzania: jedną dla próbek całkowitoliczbowych, a drugą dla ułamkowych, oraz w pamięci dwuportowe. Zwiększa to przepustowość, bo nie trzeba multipleksować wejść i wyjść, tak jak we wcześniej opracowanej architekturze z jedną ścieżką danych.



- W znanych rozwiązaniach do interpolowania próbek na różne sposoby wykorzystywane są zestawy osobnych filtrów. Autor opracował rekonfigurowalny filtr interpolacyjny, który zajmuje mniej zasobów logicznych niż zestaw filtrów. Na tej podstawie opracował innowacyjną wysokowydajną architekturę do interpolowania luminancji i chrominancji, która wykorzystuje 99 filtrów do jednoczesnego wyznaczania wszystkich próbek ułamkowych, które są potrzebne do estymacji ruchu bloku  $8 \times 8$  próbek.
- Autor opracował oryginalny trzyetapowy algorytm poszukiwania wektorów ruchu z dokładnością do ułamka piksela, który jest dopasowany do opracowanej architektury i dzięki temu wykorzystuje cykle zegara efektywniej od znanego algorytmu ENS.

O oryginalności i wartości rozwiązań opracowanych przez autora świadczy to, że publikacje, w których opisał on swoje pomysły i wyniki, zostały przyjęte na dobre konferencje międzynarodowe oraz do prestiżowego specjalistycznego czasopisma, a ponadto były już wielokrotnie cytowane.

**5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?**

Rozprawa została napisana w języku polskim. Jej tekst dobrze się czyta i charakteryzuje się dużą poprawnością językową oraz stylistyczną. Autor przedstawił informacje i przeprowadził wywody w umiejętny sposób, z zachowaniem logicznej kolejności faktów, stwierdzeń i argumentów.

Zasadniczy tekst rozprawy liczy 155 stron, które zostały poprzedzone streszczeniami po polsku i angielsku oraz spisami treści, rysunków, tablic i skrótów. W rozdziale 1, który liczy 3 strony, doktorant przedstawił przesłanki, które skłoniły go do podjęcia badań, określił zakres i cel prac, wymienił postawione tezy, oraz scharakteryzował zawartość rozprawy. Rozdział 2, który liczy 38 stron, wprowadza czytelnika w tematykę kompresji danych wizyjnych i standard HEVC/H.265, a rozdział 3, który liczy 25 stron, przedstawia znane rozwiązania, które dotyczą predykcji międzyobrazowej. W rozdziale 4, który liczy 37 stron, doktorant przedstawił swoje główne osiągnięcie koncepcyjne, czyli nowy, zorientowany sprzętowo algorytm predykcji międzyobrazowej dla standardu H.265/HEVC. W rozdziale 5, który liczy 31 stron, doktorant przedstawił natomiast swoje główne osiągnięcie projektowe, czyli nowe architektury sprzętowe do implementowania opracowanego algorytmu predykcji międzyobrazowej dla standardu H.265/HEVC. Trzystronicowy rozdział 6 zawiera podsumowanie osiągnięć doktoranta oraz listę publikacji, w których przedstawiono wyniki prac. W dodatkach, 7- i 2-stronicowym, autor zamieścił wykaz sekwencji wideo oraz opisał sposób badania efektywności kompresji, które wykorzystał w badaniach. Rozprawę zamyka 9-stronicowa bibliografia.

Układ rozdziałów i pomniejszych jednostek tekstu jest poprawny. Zachowane są rozsądne proporcje pomiędzy ilością tekstu poświęconą na przedstawienie ogólnych informacji wstępnych oraz znanych rozwiązań, a liczbą stron, na których doktorant przedstawił swoje przemyślenia i dokonania.

Moim zdaniem rozprawa była trudna do napisania ze względu na to, że predykcja międzyklatkowa, nie mówiąc o kodowaniu wideo, jest zagadnieniem rozległym, a



implementacje standardu HEVC są bardzo złożonymi rozwiązaniami programowymi, bądź układowymi. Doktorant musiał omówić skomplikowane algorytmy i architektury sprzętowe, przy czym należało nie tylko objaśnić główne idee, ale także zwrócić uwagę na szczegóły, porównując własne rozwiązania do istniejących. Rozprawa świadczy o tym, że doktorant bardzo dobrze poradził sobie z tym zadaniem.

Rozprawa została bardzo starannie sformatowana. Kroje czcionek, wcięcia, odstępy itd. są stosowane konsekwentnie. Rysunki są jednolite i bardzo dopracowane. Podobnie, nie można mieć zastrzeżeń do formatu równań, tablic, oraz bibliografii.

Główne uchybienia, które zauważyłem, mają drugorzędne znaczenie w tym kontekście, ale wyraźnie świadczą o tym, że doktorant powinien zwracać większą uwagę na pozatekstowe elementy publikacji oraz na dobór informacji wprowadzających:

- Wzory zostały osadzone w tekście w sposób odbiegający do znanych mi standardów piśmiennictwa naukowo-technicznego. Doktorant najpierw powołuje się na równanie, wykorzystując numer, a dopiero później zamieszcza równanie między akapitami. Równania powinny być zamieszczane jako części zdań
- Na rysunkach 2.26, 2.37, 2.41 i 6.1 osie wykresów nie zostały opatrzone etykietami.
- Autor rozprawy konsekwentnie zaczyna przypisy małą literą i nie kończy ich kropką, a powinny one zaczynać się i kończyć tak jak zdania.
- Hierarchia nagłówków jest czteropoziomowa, przy czym nagłówki na najniższym poziomie nie zostały ponumerowane. Ich ponumerowanie ułatwiłoby czytelnikom powoływanie się na najmniejsze jednostki tekstu.
- W tablicach 4.1, 4.5 -- 4.11, 4.15 -- 4.21 w nagłówkach kolumn nie podano jednostek, w których zostały wyrażone wielkości PSNR (jakość zrekonstruowanego obrazu) i RATE (intensywność strumienia przenoszącego zakodowane dane wizyjne). Jednostki nie zostały też określone w sąsiadującym tekście, w którym zostało tylko wyjaśnione, że pojedynczy wiersz tablicy przedstawia względną różnicę w skuteczności działania między referencyjnym koderem HM a koderem określonym w nagłówku tablicy: albo różnicę dla wskazanej sekwencji wideo, albo średnią różnicę dla kategorii sekwencji. Nietrudno domyślić się, że tablice przedstawiają wartości delty Bjøntegaarda, czyli w nagłówkach powinny być "dB" i "%", ale charakteryzując tę miarę skuteczności kodera w paragrafie 2.6.2, autor posługiwał się oznaczeniami BD-PSNR i BD-RATE. Te same oznaczenia powinny zostać użyte w tablicach, bo inaczej czytelnik może mieć wątpliwości, czy nie chodzi jednak o standardową miarę PSNR, której definicje doktorant podał w paragrafie 2.6.1.
- Rozdział 2, który ma wprowadzić czytelnika do problematyki badań, przedstawia nazbyt dużo informacji niezwiązanych z nowoczesnymi praktycznymi algorytmami kodowania wideo i predykcją międzyobrazową. Chodzi przede wszystkim o fragmenty na temat transformacji falkowej, kwantyzacji wektorowej, standardu JPEG, kodowania Huffmana, transformacji sinusowej, zjawiska Gibbsa.
- Jednocześnie w rozdziale 2 nie ma zupełnie uwag na temat implementacji sprzętowej i odpowiednich technologii. Odpowiednie informacje wstępne autor przedstawił niefortunnie dopiero w podrozdziałach 4.1 i 5.1, choć wcześniej w rozdziale jest dużo uwag o znanych rozwiązaniach sprzętowych. Rozdzielenie informacji wstępnych między podrozdziały 4.1 i 5.1 też nie było dobrym pomysłem, bo w pierwszym jest mowa o zasadach efektywnego



przetwarzania danych w FPGA, a dopiero w drugim znajduje się wyjaśnienie, co to jest za technologia.

- Trzy główne części rys. 2.30 powinny być oddzielone większymi odstępami i opatrzone etykietami, albo nawet należałoby zastąpić rys. 2.30 trzema prostszymi. Na rys. 2.30 zilustrowano bowiem trzy odmienne aspekty podziału obrazu. Bez etykiet powołania na ten rysunek w tekście są niejednoznaczne. Podobnie jest z rys. 5.12.

## 6. Jakie są słabe strony rozprawy i jej główne wady?

Chociaż pracę charakteryzuje wiele zalet, to występują w niej fragmenty, które są niezbyt udane pod względem merytorycznym. Na szczęście jest ich niewiele i absolutnie nie skłaniają do kwestionowania powodzenia całości badań. Dotyczą jednak praktycznych szczegółów, które zasługują na chwilę dyskusji podczas obrony.

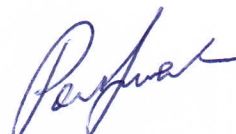
Moim zdaniem, najsłabszym miejscem w rozprawie jest podrozdział 5.6, w którym autor przedstawił wyniki syntezy sprzętowych implementacji opracowanych algorytmów i architektur, rozpatrując technologie ASIC i FPGA. Przede wszystkim brakuje mi tam jasnego sformułowania celów, w jakich synteżowana była implementacja FPGA, oraz analizy uzyskanych wyników. Z kolei wyniki syntezy ASIC mogłyby być nieco obszerniejsze. Autor nie spróbował też odnieść wyników syntezy, które uzyskał odnośnie modułu predykcji międzyobrazowej, do problemu sprzętowej implementacji całego kodera HEVC.

Oczywiście, trudno wymagać od autora, by implementował i synteżował cały koder, ale wydaje mi się, że pewne informacje ogólne mógłby on podać, opierając się na literaturze i swoich wynikach. W szczególności powinna zostać dokonana ocena, czy układy FPGA, które doktorant wykorzystał do badań, zawierają tyle zasobów logicznych, że jeden chip FPGA z zewnętrznymi modułami pamięci wystarcza do zaimplementowania całego kodera HEVC.

Oдноśnie celów syntezy FPGA, to nie zostały one określone także w podrozdziale 5.1, w którym doktorant wprowadza czytelnika do technologii ASIC i FPGA. Mam wrażenie, że synteza została przeprowadzona niejako „z rozpędu”, ponieważ ten sam opis VHDL, który jest potrzebny do syntezy implementacji ASIC, można względnie niewielkim nakładem sił i czasu wprowadzić do środowiska projektowego Altera lub Xilinx. Mogło być jednak tak, że autor pracował nad implementacją FPGA przed eksperymentami z synteżą ASIC, ponieważ środowisko projektowe Synopsys do drugiej technologii było dla niego mniej wygodne do weryfikowania opisów VHDL.

Niezależnie od pobudek, którymi autor się kierował, mógł on rozpatrzyć wyniki syntezy FPGA dokładniej, starając się wyciągnąć pewne ogólne wnioski. Mianowicie:

- Powinien wniknąć w szczegóły wykorzystania zasobów FPGA i sprawdzić nie tylko wykorzystanie tablic LUT, ale też rejestrów, LAB itd.
- Na podstawie szczegółowych raportów z syntezy, oraz tablic 5.2 i 5.3, można było określić, które z opracowanych architektur, a przynajmniej ich opisy VHDL, są niedopasowane do implementowania za pomocą FPGA.
- W jakim stopniu uzyskane implementacje wykorzystwały zasoby, które oferuje rozpatrzony układ Arria II?





Brakuje mi też wywodów na temat dostosowania architektur i opisów VHDL do implementacji w FPGA. Z treści paragrafu 4.4.1 wynika, że autor zdaje sobie sprawę, że jest to konieczne do uzyskania wydajnej implementacji i zaoszczędzenia zasobów. Potem, w podrozdziale 5.6, napisał on jednak tylko zdawkowo, że architektury sprzętowe zostały opisane w języku VHDL. Nie wyjaśnił czy i jak optymalizował kod VHDL, np. czy użył bibliotek Altera, prymitywów i dyrektyw, które pozwalają dokładnie kontrolować odwzorowanie algorytmów na FPGA.

Podobnie doktorant nie wyjaśnił, pod jakim względem optymalizował wyniki syntezy FPGA i ASIC: czy interesowała go szybkość, czy zaoszczędzenie powierzchni układu scalonego? Wiadomo, że na ogół tych kryteriów nie można pogodzić, i nie ulega wątpliwości, że ten problem dotyczy implementowania niektórych architektur, które opracował autor. W szczególności filtry interpolacyjne wykorzystują wiele sumatorów, a sumator można zrealizować na różne sposoby. Sumator z przeniesieniami szeregowymi jest wolniejszy, ale zajmuje znacznie mniejszą powierzchnię układu ASIC, od szybkiego sumatora, który opiera się np. na antycypacji przeniesień. Sposób implementacji sumatora można wymusić stosując niskopoziomowe biblioteki HDL lub odpowiednio konfigurując kompilator logiczny, w szczególności narzucając ograniczenia na pożądaną szybkość działania implementacji.

Cel syntezy implementacji ASIC nie ulega wątpliwości. Autor uzyskał w ten sposób parametry implementacji sprzętowej, które mógł odnieść do wyników opublikowanych przez innych naukowców. Dane, które zostały zestawione w tablicy 5.4 i przedyskutowane w rozdziale 5.7, wystarczają do stwierdzenia, że rozwiązania opracowane przez doktoranta prezentują się bardzo dobrze na tle wcześniejszych i stanowią dla nich poważną konkurencję. Tym niemniej wygląda na to, że doktorant skupił się głównie na funkcjonalności rozwiązań, pomijając milczeniem szereg kwestii związanych z implementacją sprzętową, które nie mają wpływu na wynik porównania architektur, ale mogą być ciekawe dla niektórych czytelników rozprawy. W szczególności:

- Dlaczego doktorant syntezywał implementacje tylko dla procesu produkcyjnego TSMC 90 nm, podczas gdy proces TSMC 65 nm stosowali autorzy 3 z 4 rozwiązań konkurencyjnych?
- Autor porównał implementacje pod względem powierzchni układu scalonego w kategoriach liczby bramek. Ta miara powierzchni jest niezależna do procesu produkcyjnego, ale nie uwzględnia obszarów połączeń. Zatem warto było podać także powierzchnię w mm<sup>2</sup>, nawet jeżeli nie uczynili tego autorzy rozwiązań konkurencyjnych w swoich publikacjach.
- Podobnie, doktorant powinien ocenić zapotrzebowanie swojego rozwiązania na moc, ponieważ jest to wyznacznik jego przydatności do urządzeń zasilanych z baterii.

Inne usterki, które zauważyłem, nie mają znaczenia merytorycznego, ale doktorant powinien potraktować odpowiednie uwagi jako wskazówki przydatne w dalszej działalności naukowej i publikacyjnej:

- Usterki, które wystąpiły we wzorach, wykresach i tablicach, świadczą o tym, że doktorant powinien dążyć do udoskonalenia swojego warsztatu pisarskiego.
- Uwagi, które przedstawiłem odnośnie patentów i wersji standardu HEVC, sugerują, że doktorant w przyszłości powinien zwracać większą uwagę na różne kwestie, które nie mają charakteru naukowego, ale są istotne przy badaniach praktycznych algorytmów i układów cyfrowych, które są związane z komercyjnym oprogramowaniem i sprzętem.
- W tekście rozprawy masowo występuje dosłowne tłumaczenie słowa "frame", czyli "ramka", choć lepiej brzmiałyby słowa "obraz" i "klatka", których autor użył sporadycznie,

ale w ważnych miejscach, np. w tytule, i które są stosowane w polskojęzycznej literaturze fachowej. Zupełnie niefortunny wydaje mi się zwrot "ramka obrazu", który wystąpił na str. 10.

Tę część recenzji kończę listą drobnych usterek edycyjnych, które zauważyłem w pracy:

- Przypis do rys. 2.23 ze str. 22 znajduje się na str. 23.
- W podpisie rys.2.14 jest zbędny apostrof.
- W podpisie rys.2.39 wystąpiła literówka "Kategoie".
- Na str. 43 wystąpiła literówka "Pierwszą ramka".
- Nie widać związku rys. 2.40 z kodowaniem.
- W równaniu 2.25 znak \* został użyty do oznaczenia mnożenia, choć znakiem tym oznacza się splot, tak jak w równaniu 5.1.
- Na str. 43 wystąpiła literówka "unimodlaności".
- W rozdz. 3 sformułowanie "przeszukiwanie wektorów" nie brzmi dobrze. Lepszy byłby zwrot "przeszukiwanie zbioru wektorów", albo po prostu "poszukiwanie wektorów".
- Na str. 69 wystąpiła literówka "nastęcza trudności implementacyjnych".
- Na str. 114, w powołaniach na części rysunku 5.11, występują zbędne kropki. Zamiast np. "5.11.a" powinno być "5.11a".

Należy zauważyć, że takie usterki występują sporadycznie w przedłożonej mi rozprawie.

## 7. Jaka jest przydatność rozprawy dla nauk technicznych?

Kodowanie danych wizyjnych jest dziedziną, która ustawicznie się rozwija. Zapotrzebowanie na nowe, udoskonalone algorytmy i rozwiązania sprzętowe wynika z tego, że

- użytkownicy generują coraz więcej treści multimedialnych;
- pojemność pamięci masowych i przepustowość łączy sieciowych nie zwiększa się odpowiednio szybko;
- wymiana dysku lub łącza na doskonalsze wiąże się z kosztami;
- pojawiły się nowe zastosowania, takie jak telewizja swobodnego punktu widzenia (ang. multiview), oraz nowe rodzaje danych, takie jak wideo UHD lub mapy głębokości 3D (ang. depth maps).

Moim zdaniem wyniki przedstawione w rozprawie mają duże znaczenie, ponieważ mogą znaleźć zastosowanie nie tylko w koderach zgodnych ze standardem HEVC, ale także w przyszłych rozwiązaniach problemów związanych z kodowaniem wideo. Jest bowiem mało prawdopodobne, by w przewidywalnej przyszłości predykcja międzyobrazowa straciła status podstawowego mechanizmu, który służy do upakowywania danych wizyjnych.

O dużym znaczeniu rozprawy dla stanu wiedzy świadczy to, że publikacje doktoranta, w tym najnowsze, z 2016 roku, w których przedstawił on wyniki badań uzyskane w ramach pracy nad



doktoratem, zostały już kilkadziesiąt razy zacytowane przez autorów z różnych stron świata. Można to łatwo sprawdzić za pomocą serwisu scholar.google.com. Co ważne sporo powołań na prace doktoranta znalazło się w artykułach zamieszczonych w prestiżowych czasopismach o wysokich wartościach Impact Factor.

Do specjalistycznego czasopisma tej rangi, Journal of Real-Time Image Processing, wyd. Springer, zostały przyjęte dwie publikacje, których współautorem jest doktorant i które prezentują niektóre wyniki prac opisanych rozprawie. Należy to uznać za kolejny dowód, że eksperci uważają pomysły i rezultaty doktoranta za wartościowe i oryginalne.

**8. Do której z następujących kategorii Recenzent zalicza rozprawę?**

- a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy**
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania**
- c/ spełniająca wymagania**
- d/ spełniająca wymagania z wyraźnym nadmiarem**
- e/ wybitnie dobra, zasługująca na wyróżnienie**

Zaliczam rozprawę mgra. inż. Macieja Trochimiuka do kategorii

*d) spełniająca wymagania z wyraźnym nadmiarem*

ponieważ reprezentuje ona ogólnie bardzo dobry poziom i pozytywnie świadczy o dużej wiedzy, oczytaniu, umiejętnościach i pracowitości doktoranta, a ponadto dotyczy zagadnień praktycznych i aktualnych. Nie mam wątpliwości, że autor poprawnie postawił a następnie rozwiązał zagadnienie badawcze, dowodząc, że zasługuje na awans naukowy. Zauważone przeze mnie usterki merytoryczne i formalne nie umniejszają osiągnięć doktoranta, choć powinny skłonić go do pewnych refleksji i doskonalenia warsztatu badawczo-publikacyjnego. Wnoszę o dopuszczenie autora do publicznej obrony rozprawy.

