

Politechnika Warszawska
Wydział Elektroniki i Technik Informatycznych

Warszawa, 20 września 2016 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 11 października 2016 r. publiczna obrona rozprawy doktorskiej

mgr inż. Piotra Graszki

temat: „Modelowanie tła oraz detekcja obiektów pierwszoplanowych w sekwencjach obrazów pochodzących ze statycznych kamer wideo”

promotor – prof. dr hab. inż. Jan Zabrodzki z Politechniki Warszawskiej

recenzenci:

**prof. dr hab. inż. Marek Gorgoń z Akademii Górniczo Hutniczej
dr hab. inż. Waldemar Rakowski, prof. Politechniki Białostockiej**

Obrona odbędzie się w dniu 11 października 2016 r. w sali 116 na Wydziale Elektroniki i Technik Informatycznych – Gmach im. Janusza Groszkowskiego, Warszawa, ul. Nowowiejska 15/19; początek godz.9.00.

Po adresem: 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

ROZPRAWA DOKTORSKA

mgr inż. Piotr Graszka

Modelowanie tła oraz detekcja obiektów pierwszoplanowych w sekwencjach obrazów pochodzących ze statycznych kamer wideo

Promotor

prof. dr hab. inż. Jan Zabrodzki

Streszczenie

Głównym celem rozprawy jest zaproponowanie nowych metod modelowania tła oraz detekcji obiektów pierwszoplanowych. Proponowane metody łączą zalety najczęściej wykorzystywanych rozwiązań problemu modelowania tła starając się nie powielać ich wad. W tym celu nowe metody biorą pod uwagę i odpowiednio wykorzystują informacje o kolorze oraz korelacje przestrzenne obrazu, a także odpowiednio aktualizują model tła tak, aby skompensować refleksy świetlne oraz cienie występujące w analizowanych sekwencjach wideo, równocześnie nie dopuszczając do powstawania nadmiernych szumów oraz przekłamań.

W pracy przeprowadzono analizę oraz testy najczęściej spotykanych metod modelowania tła, a następnie na podstawie wniosków płynących z przeprowadzonych eksperymentów zaproponowano dziewięć nowych metod, które zarówno jakością zwracanych wyników jak i szybkością pracy przewyższają wiele spośród znanych dotąd i często wykorzystywanych rozwiązań omawianego problemu.

Wszystkie opisane w pracy metody zostały zaimplementowane w aplikacji testowej „Background Modeling”, która posłużyła do wykonania testów jakościowych i ilościowych. Testy zostały przeprowadzone z wykorzystaniem 30 sekwencji wideo pochodzących z publicznych baz danych oraz 16 sekwencji własnych. W sumie przetestowanych zostało 51 różnych metod modelowania tła. Na podstawie otrzymanych wyników wyciągnięto i opisano wnioski oraz wytypowano sześć najlepszych pod względem jakości i szybkości metod modelowania tła spośród wszystkich przetestowanych – trzy z nich to autorskie rozwiązania zaproponowane w ramach tej rozprawy. Zaproponowano i omówiono również sposób doboru jednej z sześciu wytypowanych wcześniej najlepszych metod modelowania tła w zależności od wymogów potencjalnego zastosowania.

dr hab. inż. Waldemar Rakowski

Białystok, 1.08.2016

prof. nadzwyczajny

Politechnika Białostocka

Wydział Zarządzania

Katedra Informatyki Gospodarczej i Logistyki

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

Tytuł rozprawy: **Modelowanie tła oraz detekcja obiektów pierwszoplanowych w sekwencjach obrazów pochodzących ze statycznych kamer wideo**

Autor rozprawy: **mgr inż. Piotr Graszka**

1. Zagadnienie naukowe będące przedmiotem rozprawy należy do dziedziny widzenia komputerowego i komputerowego wspomaganie nadzoru wizyjnego (monitoringu). Autor postawił tezę, iż możliwa jest poprawa efektywności i skuteczności istniejących algorytmów modelowania tła w sekwencjach obrazów pochodzących ze statycznych kamer wideo i w konsekwencji detekcja obiektów pierwszoplanowych w monitorowanych scenach. Celem Autora było również zaproponowanie nowych metod, w szczególności wykorzystujących informacje o kolorze w obrazach i korelacje przestrzenne pomiędzy pikselami, umożliwiającą kompensowanie występujących cieni bez wprowadzania nadmiernych szumów. Tak sformułowane zagadnienie badawcze nadaje rozprawie charakter algorytmiczno-eksperymentalny i poprzez aplikacyjny charakter algorytmów oraz eksperymenty obliczeniowe ściśle wiąże się z informatyką w dziedzinie nauk technicznych.
2. Autor przeprowadził rzetelną analizę literatury światowej. W Bibliografii zamieścił 87 pozycji i odwołał się do najważniejszych publikacji światowych dotyczących przedmiotu badań. Przegląd literaturowy przedstawił w formie opisu najważniejszych metod modelowania tła i detekcji obiektów pierwszoplanowych w rozdziale 2 rozprawy, liczącym 22 strony.
3. Autor założył, że możliwa jest poprawa skuteczności istniejących metod modelowania tła i detekcji obiektów pierwszoplanowych poprzez modyfikacje tych metod pod kątem zwiększenia trafności określenia przynależności pikseli do tła i do obiektów pierwszoplanowych lub poprzez skrócenie czasu obliczeń czyli zwiększenie szybkości metody. W tym celu Autor modyfikował istniejące metody w bardzo różny sposób, np. zastępując histogramy rozkładami statystycznymi typu Gaussa lub rozkłady statystyczne zastępując medianami i wykorzystując do klasyfikacji piksela jeden lub

wiele klasyfikatorów. Z przeprowadzonych eksperymentów obliczeniowych na danych ogólnodostępnych w internecie jak i specjalnie przygotowanych przez siebie testowych sekwencji wideo, Autor wyciągnął wnioski odnośnie skuteczności modyfikacji istniejących metod i zestawił je w formie tabelarycznej w Podsumowaniu. Ze względu na oryginalność niektórych modyfikacji, można uznać je za autorskie propozycje nowych sposobów modelowania tła oraz detekcji obiektów pierwszoplanowych. W ocenie Autora najwyższą trafność klasyfikacji tła i obszarów pierwszoplanowych ma opisana w punkcie 3.4.2 metoda modelowania tła przez głosowanie, która jest oryginalną metodą opracowaną przez Autora. W porównaniu z innymi metodami, czas obliczeń jest stosunkowo długi, ale w związku ze stale zwiększającą się mocą obliczeniową komputerów, w zastosowaniach gdzie wymagana jest wysoka trafność klasyfikacji, metoda ta może być konkurencyjna w stosunku do innych.

4. Oryginalność rozprawy polega na zbadaniu i przedstawieniu wniosków z wyników badań metod modelowania tła i detekcji obiektów pierwszoplanowych poprzez bardzo różnorodne i oryginalne, autorskie modyfikacje istniejących metod, prowadzące w szczególności do opracowania nowych metod. Zaproponowane przez Autora modyfikacje istniejących metod jak i nowo opracowane metody odpowiadają aktualnemu stanowi wiedzy w zakresie modelowania tła i detekcji obiektów pierwszoplanowych w sekwencjach wideo.
5. Układ i strona językowa rozprawy nie budzi istotnych zastrzeżeń. Można jednak wskazać sformułowania, których należałoby unikać. Np. Autor oprócz sformułowania „wynik działania metody”, które jest powszechnie używane, w wielu miejscach używa sformułowania „wynik pracy metody”. Moim zdaniem lepiej jest konsekwentnie używać tego pierwszego sformułowania. Zdecydowana większość zdań to są zdania klarowne, przyjazne dla czytelnika. Jednakże na początku strony 139 rozprawy, w Podsumowaniu, napotkałem zdanie, które chyba jest najdłuższym zdaniem jakie kiedykolwiek przeczytałem, bo liczy ono 60 słów. Jeśli idzie o precyzję i jasność omawianych zagadnień to nie zawsze opis jest przyjazny czytelnikowi. I tak, np. w przedstawieniu istniejących algorytmów modelowania tła, na stronie 23 i następnych, Autor używa terminu „histogram czasowy”. Nie jest to termin powszechnie znany i uważam, że należało precyzyjnie go zdefiniować. Jest to istotne dla wyobrażenia sobie przez czytelnika wielkości pamięci potrzebnej do wykonania algorytmu. Innym ważnym i często występującym w rozprawie terminem jest termin „metoda”. Można mieć wątpliwości czy Autor właściwie używa tego terminu w każdym kontekście. Np. na stronie 76 rozprawy jest zdanie: „W związku z tym w sumie przetestowanych zostało 120 różnych metod uruchomieniowych w różnych konfiguracjach”. Moją wątpliwość zilustruję następującym przykładem. Istnieją metody kontekstowe kodowania obrazów, w których wybierany jest kontekst. Nie można powiedzieć, że każdy nowy kontekst daje nową metodę kodowania. Z powodu dość „rozmytego”

użycia w rozprawie terminu „metoda”, trudno jest doszukać się precyzyjnej i bezspornej liczby opracowanych przez Autora nowych metod modelowania tła w sekwencjach wideo. Autor w zamierzeniu opisywał istniejące algorytmy dotyczące przedmiotu badań pod kątem możliwości pokazania sposobów poprawy ich efektywności i skuteczności. Dlatego istotnym było pokazanie w opisie algorytmów tych ich kroków, które decydują o możliwości ich realizacji, w szczególności w czasie rzeczywistym. Tymczasem np. w opisie metody modelowania tła opartej na analizie składowych głównych (punkt 2.8) pominięty został bardzo istotny krok przekształcenia liniowego zmniejszający wymiar obliczanych wektorów własnych, bez którego to kroku metoda praktycznie jest nierealizowalna. Nie wątpliwie walory rozprawy, z punktu widzenia czytelnika, podniosłoby wskazywanie explicite w opisie metody roku jej opublikowania. Dostrzegłem przeoczenie Autora dotyczące umieszczenia w Bibliografii dwukrotnie tego samego artykułu (pozycja 21 i 42) wynikające prawdopodobnie z różnic w zapisie nazwisk autorów publikacji (litery spoza alfabetu łacińskiego).

6. Rozprawa nie zawiera istotnych słabych stron. Jednakże czytelnik analizujący rozprawę z punktu widzenia informatyka może odczuwać pewien niedosyt dotyczący szacunku zasobów konsumowanych przez metody, tj. wielkości wymaganej pamięci i złożoności obliczeniowej bądź czasu obliczeń w odniesieniu do aktualnie dostępnego sprzętu. Podobnie można mieć niedosyt rozważań teoretycznych w podejściu do konstruowania nowych metod modelowania tła. I tak np. w odniesieniu do metod wykorzystujących przekształcenie Fouriera lub kosinusowe, opisanych w podrozdziale 3.2, Autor nie wskazuje przesłanek teoretycznych zastosowania tych przekształceń. Transformata Fouriera jest rozwinięciem sygnału w szereg harmonicznych (sinusów i kosinusów) i można rozważać na ile taka reprezentacja sygnału jest użyteczna w modelowaniu tła i detekcji obiektów pierwszoplanowych.
7. Rozprawa bardzo dobrze wpisuje się w nauki techniczne. Autor nie tylko opisał, udoskonalił istniejące i opracował nowe metody modelowania tła i detekcji obiektów pierwszoplanowych, ale również je zaimplementował i przedstawił wyniki testów wraz z wnioskami wynikającymi z przeprowadzonych testów. Napisana przez Autora w technologii obiektowej aplikacja działająca w środowisku systemu operacyjnego MS Windows, wykorzystująca aktualnie dostępne biblioteki MFC, OpenCV oraz OpenMP, jest wartościowym narzędziem informatycznym. Umożliwia ona eksperymentowanie z dostępnymi w internecie metodami modelowania tła z biblioteki BGS Library jak i z opracowanymi przez Autora. Wyniki testów wraz z wpływającymi z nich wnioskami zamieszczone w tabeli 31 (strony 139-141) są dowodem rzetelnie przeprowadzonych badań eksperymentalnych, bardzo istotnych w naukach technicznych i świadczą o osiągnięciu przez Doktoranta sformułowanego na początku rozprawy celu. Materiał zawarty w rozprawie może być pomocny w implementowaniu systemów monitoringu, w szczególności specjalizowanych, w

których można scharakteryzować właściwości tła i ewentualnej interakcji obiektów pierwszoplanowych i na tej podstawie wybrać, w oparciu o oceny przedstawionych metod, właściwą metodę modelowania tła.

8. Rozprawę uznaję za spełniającą z wyraźnym nadmiarem wymagania stawiane rozprawom doktorskim w dziedzinie nauk technicznych w dyscyplinie informatyka.

Na wyróżnienie zasługuje bardzo obfity materiał eksperymentalny i wnioski z przeprowadzonych eksperymentów obliczeniowych, przedstawione finalnie w zwartej postaci tabelarycznej, pokazującej zalety, słabe strony, szybkość obliczeń i ocenę najważniejszych zbadanych metod.

Jednym z głównych wniosków wypływających z przeprowadzonych badań jest powszechna i stosunkowo duża wrażliwość metod modelowania tła i detekcji obiektów pierwszoplanowych na zmiany (w szczególności szybkie) oświetlenia sceny, co stanowi wskazówkę do ewentualnie dalszych badań mających na celu poprawienie ich odporności na zmiany oświetlenia.

W konkluzji wnoszę o dopuszczenie mgr inż. Piotra Graszki do dalszych etapów postępowania w przewodzie doktorskim.

Aleksander Radecki

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, 7 września 2016.

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

Tytuł rozprawy: Modelowanie tła oraz detekcja obiektów pierwszoplanowych w sekwencjach obrazów pochodzących ze statycznych kamer wideo

Autor rozprawy: mgr inż. Piotr Graszka

Przedmiotem niniejszej recenzji jest rozprawa doktorska zatytułowana „Modelowanie tła oraz detekcja obiektów pierwszoplanowych w sekwencjach obrazów pochodzących ze statycznych kamer wideo”, napisana w roku 2016, której autorem jest mgr inż. Piotr Graszka.

Promotorem pracy jest prof. dr hab. inż. Jan Zabrodzki z Politechniki Warszawskiej. Niniejsza ocena została przygotowana na zlecenie prof. dr hab. inż. Krzysztofa Zaremby, Dziekana Elektroniki i Technik Informatycznych, zawarte w piśmie z dnia 8 czerwca 2016 roku, w związku z decyzją Rady Wydziału z dnia 24 maja 2016 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 i zagadnienie naukowe rozpatrywane w rozprawie, zostało dobrze ujęte w jej tytule. Tematyka jest ważna i aktualna. Algorytmy analizy obrazu bardzo często klasyfikuje się jako dział informatyki, ale też niekiedy, w zależności od kontekstu prowadzonych badań, w dyscyplinach pokrewnych tj. elektronice, automatyce i robotyce, inżynierii biomedycznej czy telekomunikacji.

Doktorant, już w pierwszym zdaniu streszczenia swojej rozprawy stawia sobie jasny cel: cyt. "zaproponowanie nowych metod modelowania tła oraz detekcji obiektów pierwszoplanowych". We wstępie rozwija tą myśl w następujący sposób: cyt. "W szczególności metody te będą brały pod uwagę i odpowiednio wykorzystają informacje o kolorze oraz korelacje przestrzenne obrazu, a także będą odpowiednio aktualizować model tła tak, aby skompensować występujące cienie i nie będą prowadziły do powstawania nadmiernych szumów oraz obiektów duchów."

W rozprawie przytoczone jest dwuzdaniowe sformułowanie tezy rozprawy: cyt. "... możliwe jest opracowanie nowych metod modelowania tła oraz detekcji obszarów pierwszoplanowych. Metody takie mogą stanowić rozwinięcie istniejących rozwiązań jednocześnie przewyższając je pod względem jakości i szybkości pracy oraz lepiej rozwiązywać problemy, z którymi nie radzą sobie najczęściej wykorzystywane obecnie metody modelowania tła." W odczuciu recenzenta cel i teza pracy zostały

sformułowane poprawnie, z tym że wskazując precyzyjnie na pewien obszar badawczy nie wskazały zbyt precyzyjnie oczekiwań wobec spodziewanych efektów. Wskazane byłoby również zadbać o lepsze uwidocznienie tezy w tekście rozprawy oraz postarać się zawrzeć tezę w jednym zdaniu.

Rozprawa ma charakter doświadczalny, ale zawiera niezbędne analizy teoretyczne: opis matematyczny i słowny analizowanych metod. W omawianych zagadnieniach jest to wariant stosowany najczęściej, zarówno w publikacjach jak i pracach monograficznych, choć warto odnotować, że nie jest to jedyna możliwa droga w analizie obrazów tj. publikowane są również prace na gruncie czysto teoretycznym zwykle potwierdzone eksperymentalnie.

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?

Rozdział 2 rozprawy zawiera przegląd metod modelowania tła. Rozdział ten wprowadza podstawowe pojęcia z zakresu modelowania tła i detekcji obiektów pierwszoplanowych. Autor zdecydował się pokazać ewolucję metod modelowania tła od elementarnych do bardziej zaawansowanych, cytując źródła literaturowe. Na końcu tego rozdziału sformułował wnioski krytyczne, które, w dalszej części rozprawy stanowiły wytyczne dla prowadzenia badań. Natomiast w rozdziale 4, poświęconym porównaniu jakości i szybkości metod własnych z metodami z literatury, wziął pod uwagę szerszą grupę nowszych metod modelowania tła zawartą w bibliotece *Background Subtraction Library (BGS Library)* w wersji z 2013 roku.

Niestety, brak w pracy odniesienia do metod modelowania tła i detekcji obiektów pierwszoplanowych powstałych po 2013 roku. Wobec bardzo dynamicznego postępu dokonującego się w tej dziedzinie, należy zwrócić uwagę, że w okresie 2013-2016 powstało wiele nowych metod wysokiej jakości. Dobrym źródłem informacji o jest strona internetowa profesora Thierry'ego Bouwmansa z Université de La Rochelle, gdzie można znaleźć odwołania do kilkuset starszych i nowszych publikacji dotyczących metod modelowania tła, uporządkowanych wg sposobu działania (<https://sites.google.com/site/backgroundsubtraction/Home>) jak również strona changedetection.net, na której znajduje się, aktualizowane na bieżąco, porównanie metod modelowania tła z czerwca 2014. Nota bene, na tej stronie autorzy nowopowstałych metod rejestrują ich wyniki i mogą porównać je z konkurencyjnymi. W kontekście przedstawionej w pracy metody głosowania, warto też odnotować brak w spisie literatury stosunkowo nowej publikacji autorstwa S. Bianco, G. Ciocca, R. Schettini "How far can you get by combining change detection algorithms?", zgłoszonej do IEEE Transactions on Image Processing, w listopadzie 2015 i opublikowanej w całości na stronie <http://arxiv.org/pdf/1505.02921v2.pdf>. (Być może jednak, w momencie przygotowania przez Autora materiałów do druku rozprawy doktorskiej, artykuł ten nie był jeszcze upubliczniony.)

W odczuciu recenzenta, mimo powyższych uwag krytycznych, odnoszących się do przedstawionego przeglądu literatury, autor posiadał wystarczającą wiedzę na temat problemów występujących w trakcie modelowania tła. W pracy dokonał dokładnej analizy metod opisywanych w rozdziale 2 i wskazał na ich słabe i mocne punkty. Szkoda, że nie zostały opisane bardziej szczegółowo choćby wybrane metody z biblioteki BGS wykorzystane do testów w rozdziale 4.

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

Przechodząc do merytorycznej analizy wykonanych prac należy wrócić do sformułowanego w rozprawie celu. Autor założył że powstaną nowe metody modelowania tła które cyt. „będą brały pod uwagę i odpowiednio wykorzystają informacje o kolorze oraz korelacje przestrzenne obrazu, a także będą odpowiednio aktualizować model tła tak, aby skompensować występujące cienie i nie będą prowadziły do powstawania nadmiernych szumów oraz obiektów duchów."

Autor w rozdziale 3 przedstawił 9 opracowanych metod modelowania tła i zaproponował osiągnięte wyniki. W rozdziale 4 przedstawił wyniki dokonanych testów zawierające porównanie metod własnych z metodami zawartymi w bibliotece BGS wg stanu z 2013 roku. W rozdziale 5 przeanalizował wyniki testów, a w rozdziale 6 dokonał podsumowania.

Ważną decyzją podjęta przez Autora, która w dużej mierze rzutuje na wnioski wyciągnięte z testów, był wybór metody oceny algorytmów modelowania tła (rozdział 4.3).

Założenia przedstawionej metody są analogiczne do metod opisanych w literaturze: maska obiektów pierwszoplanowych jest obrazem binarnym, można łatwo porównać wyniki osiągnięte przez poszczególne metody detekcji obiektów pierwszoplanowych z maską idealną, uzyskaną na skutek prawidłowej klasyfikacji dokonanej przez eksperta. Prowadzi to do wskazania prawidłowo i nieprawidłowo zaklasyfikowanych pikseli do tła bądź do pierwszego planu, a następnie do oceny ilościowej przy pomocy typowych statystyk klasyfikacyjnych.

Autor zaproponował własną metodę oceny algorytmów, definiując propozycję miary statystycznej, która posłużyła mu do stworzenia rankingów metod. Przyjął również arbitralnie próg tej miary na poziomie 80 i do końcowej grupy najlepszych rozwiązań dopuścił dwanaście algorytmów, które przekroczyły wartość tego progu. Z jednej strony konstrukcja zaproponowanej miary została uzasadniona w pracy, a takie podejście dało bardzo klarowne wyniki rankingów, lecz z drugiej strony powstaje pytanie, czy taka metodologia nie jest zbyt daleko idącym uproszczeniem. Strona changedetection.net, która jest powszechnie uznawanym, miarodajnym źródłem informacji o metodach modelowania tła, zawiera rankingi metod z przedstawieniem wszystkich podstawowych miar statystycznych. Metody oceniane są wg średniej pozycji w rankingach podstawowych miar statystycznych. Na stronie podkreślono, że przedstawienie wszystkich podstawowych miar statystycznych pozwala budować rankingi w zależności od wyzwań jakie stoją przed metodą modelowania tła. Ograniczenie się do opublikowania jedynie wyników miary autorskiej, utrudnia potencjalnym użytkownikom bardziej wszechstronną ocenę metod autorskich przedstawionych w rozprawie. Ponadto w literaturze, do oceny jakości klasyfikacji metod detekcji obiektów pierwszoplanowych często stosowana jest standardowa miara dokładności „*Accuracy*” $ACC = (TP + TN) / (TP + FN + FP + TN)$, która jest miarą symetryczną, a więc podobną w istocie swoich założeń do miary zaproponowanej przez Autora w rozprawie. Dla osób zajmujących się modelowaniem tła przydatne byłoby porównanie rankingów metod wyliczonych przy pomocy miary zaproponowanej przez Autora i miary dokładności „*Accuracy*”.

W podsumowaniu tej części recenzji należy z jednej strony docenić zaproponowanie nowej miary oceny jakości metod modelowania tła, ale wytknąć brak przedstawienia pełnego zestawienia miar podstawowych dla wszystkich badanych metod, które były przecież wyliczane jako wyniki pośrednie. (Takie zestawienie mogłoby znaleźć się w formie dodatku i ewentualnie w arkuszu kalkulacyjnym zapisanym na płycie DVD, którą Autor dołączył do rozprawy).

Dodatkowo, słowo komentarza należy się do wspomnianej, przyjętej wartości progu autorskiej miary oceny. Na miejscu Autora, recenzent rozważyłby próg 78 lub 79. Dzięki temu, na wykresie na rys. 65 pojawiłoby się dane na temat najbardziej popularnej grupy metod modelowania tła tj. GMM, a także kilku metod autorskich, przedstawionych w rozdziale 3, które uplasowały się tuż poniżej progu 80. Dałoby to możliwość łatwej oceny miar autorskich i wartościowe odniesienie do metod GMM, ciekawe w zwłaszcza w kontekście ich porównania z metodami medianowymi. Jak wyglądałby wtedy wykres z rys. 65 i czy dalsza selekcja najlepszych metod dałaby te same wyniki?

Autor podkreśla, że użycie filtracji końcowej poprawia jakość maski obiektów pierwszego planu, a w tabelach w rozdziale 4 rozróżnia wyniki uzyskane przed i po filtracji. Stosowanie metod filtracji końcowej – zwykle mediany – jest często spotykane w publikacjach z literatury. Czy maski obiektów pierwszoplanowych, uzyskane przy pomocy metod z biblioteki BGS, podlegały przefiltrowaniu

podobnie jak maski z metod autorskich? Jeśli nie, to czy Autor analizował, czy dodanie filtracji końcowej poprawiłoby uzyskane przez te metody wyniki?

Na tle rozważań odnoszących się do przyjętej metodologii łatwiej będzie obecnie recenzentowi przedstawić ocenę poszczególnych metod zaproponowanych przez Autora w rozdziale 3.

Metoda najbliższego sąsiedztwa bierze pod uwagę historię analizowanego piksela, ale też odwołuje się do wartości bieżącej i historycznych jego najbliższych sąsiadów. Metoda przebadana została w różnych wariantach. Wyniki przedstawione w rozdziale 4 przy pomocy miary autorskiej, pozwoliły sklasyfikować ją nieznacznie poniżej progu 80 punktów. Na podobnych założeniach oparta jest i podobne wyniki uzyskuje, znana i cytowana w tym rozdziale metoda PBAS. Jednak metoda autorska okazała się kilkukrotnie od niej szybsza. Nie mniej jednak warto wspomnieć, że z wykresu na rys. 65 wynika, że porównywalne do metody najbliższego sąsiada, metody z biblioteki BGS osiągają nieznacznie lepsze wyniki jakościowe. Pewien niedosyt w opisie powoduje brak precyzyjnego określenia pojęcia najbliższego sąsiedztwa – czy Autor za najbliższe sąsiedztwo uznaje obszar zdefiniowany parametrami „x” i „y”? Czy jeśli tak, to jest to „najbliższe sąsiedztwo” czy raczej „zdefiniowane bliskie otoczenie”. Czy wprowadzenie pojęć „wiek piksela” i „szczegółowość” nie lepiej zastąpić standardowymi określeniami „wielkość bufora czasowego” i „skala obrazu”. Czy szczegółowość „75%” oznacza stopień przeskalowania analizowanego obrazu? W metodzie warto podkreślić *in plus* implementację parametru odpowiadającego za czas inkorporacji nieruchomych obiektów do tła, choć w typowych aplikacjach definiuje się podobne reguły.

Dwie przedstawione próby implementacji metod częstotliwościowych nie przyniosły wartościowych rezultatów. Porównanie wyników transformat częstotliwościowych dla historycznych wartości piksela i modelu tła, a następnie wykonanie transformacji odwrotnej, a także dokonanie transformacji częstotliwościowej całej analizowanej ramki i porównanie różnymi metodami z transformatami modelu tła, po czym wykonanie transformacji odwrotnej na różnicowym wyniku, dało wyniki nie lepsze niż bezpośrednie zastosowanie metod modelowania tła bez wyliczania transformat. Czy jednak wykonanie operacji modelowania tła, nie w przestrzeni czasu, ale w przestrzeni częstotliwości, z wykorzystaniem liniowych transformat częstotliwościowych w tył i w przód (dość kosztownych obliczeniowo), mogło przynieść wartość dodaną (dodatkową informację) poprawiającą rezultaty? Czy widmo tła lub pierwszego planu było możliwe do wyseparowania?

Kolejny fragment rozprawy przedstawia 4 warianty metod opartych o wyliczenie mediany jako podstawy ustalenia wartości tła. Punktem wyjścia tych rozważań jest przedstawiona w rozdziale 2.3, znana z literatury metoda przybliżania mediany (*Approximated Median*). Autor zainspirował się drogą jaką przeszły metody wykorzystujące rozkład Gaussa od prostej formy do postaci bardziej złożonych czyli popularnych metod mieszania rozkładów Gaussa. Zaproponował 4 analogicznie skonstruowane metody medianowe i wykazał, że są one lepsze pod względem szybkości i dokładności niż metody wykorzystujące rozkłady Gaussa. I to był bardzo dobry, godny pochwały pomysł Autora! Szczególnie wartościowa, wg miary oceny zaproponowanej w pracy, okazała się metoda mieszania median (*Median Mixture*) w różnych wersjach implementacji. Warto podkreślić, że Autor opublikował osiągnięte wyniki dotyczące metod medianowych w materiałach konferencyjnych w roku 2014. Oryginalność zaproponowanych metod nie budzi wątpliwości.

Ostatnia grupa zaproponowanych metod to dwa algorytmy hybrydowe. Pierwszy, nazwany metodą łączenia mieszanin median (*United Median Mixture - UMM*) zakładał powstanie maski obiektów pierwszoplanowych na podstawie median liczonych dla obrazu kolorowego, różnicowego obrazu kolorowego i gradientu obrazu. Metoda stanowi dalsze rozwinięcie metod medianowych i jest w pełni oryginalnym osiągnięciem Autora. W rozważaniach przedstawiony został również autorski algorytm eliminacji cieni, oparty o spostrzeżenie, że granica obszaru cienia jest w większości przypadków nieostra, a więc jest słabo widoczna w porównaniu do granic rzeczywistych obiektów na obrazie gradientowym. Mechanizm integracji wyników skutecznie usuwał nieostre cienie, zakłócenia

i szumy. Przeprowadzona implementacja wielowątkowa tej metody pracowała wolniej od innych mniej złożonych metod medianowych, ale wyróżniała się lepszą jakością.

Druga z metod hybrydowych to metoda głosowania. Autor wykorzystał znaną i opisaną w literaturze własność poprawy wyników klasyfikacji binarnych na podstawie łącznej analizy wyników z kilku słabych klasyfikatorów. W zaproponowanej metodzie o wynikach klasyfikacji decydowało głosowanie. Autor przebadał wiele wariantów głosowania i wybrał ten, który dał najlepsze wyniki. Konkretna metodologia głosowania opracowana przez Autora jest oryginalna. Wszystko wskazuje na to, że propozycja Autora powstała równolegle do metod opisanych we wzmiankowanej w recenzji publikacji S. Bianco, w której również proponuje się głosowanie jako jeden z kilku wariantów łączenia klasyfikatorów. Pod względem jakości modelowania tła, metoda głosowania, w wersji która wykorzystywała klasyfikatory z biblioteki BGS i zaproponowaną w pracy filtrację końcową, dała najlepsze wyniki spośród wszystkich metod zaproponowanych przez Autora. Wg tabeli 27 i rysunku 75 metoda głosowania wypada nieco lepiej niż najlepsza pod względem jakości metoda z biblioteki BGS (*Multi Layer BGS*). Według rysunku 66 zajmuje 3 miejsce. Jednak, aby uzyskać pełny obraz co do jakości przedstawionej metody, rozważania należałoby poszerzyć, biorąc pod uwagę najnowsze metody modelowania tła przedstawione w zestawieniu na stronie [changedetection.net](http://wordpress-jodoin.dmi.usherb.ca/cdw2014/) (<http://wordpress-jodoin.dmi.usherb.ca/cdw2014/>). Pozwoliłoby to ocenić jakość przedstawionej metody głosowania na tle obecnego stanu wiedzy w zakresie modelowania tła.

Podsumowując realizację tezy rozprawy warto wskazać, że jak sam przyznaje Autor, dodanie informacji o korelacji pomiędzy pikselami ma znaczenie jedynie w metodach prostych, metody zaawansowane radzą sobie z tym problemem. Najbardziej wartościowe uzyskane efekty dotyczyły metod medianowych pracujących w czasie rzeczywistym. Przedstawiona została procedura redukcji nieostrych cieni dla metody łączenia mieszanin median. Metoda mieszanin median radziła sobie z nieostrych cieniami, ale Autor nie wskazał precyzyjnie dzięki któremu fragmentowi algorytmu. Natomiast dużo uwagi Autor poświęcił zagadnieniom filtracji końcowej (usuwanie szumów), uzyskując satysfakcjonujące wyniki przy pomocy pewnych kombinacji mediany i prostych metod morfologicznych. Recenzent nie zauważył nowatorskich pomysłów dotyczących sposobu usuwania tzw. „obiektów duchów”. Osiągnięte wyniki pozwalają uznać, że teza rozprawy została udowodniona.

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?

Chcąc wskazać na najbardziej wartościowe i w pełni oryginalne dokonania Autora należy raz jeszcze wymienić:

1. Opracowanie metod medianowych, w szczególności mieszania przybliżonych median wektorowych i mieszania median, ze względu na ciekawy pomysł, konsekwentną rozbudowę koncepcji, dobre wyniki jakości modelowania tła w klasie metod pracujących w czasie rzeczywistym.
2. Opracowanie metody hybrydowej łączenia mieszanin median, ze względu na oryginalne połączenie elementów składowych i zawartą w niej procedurę eliminacji cieni zintegrowaną z przetwarzaniem końcowym.

Wymienione powyżej metody medianowe stanowią oryginalny wkład Autora w stan wiedzy na temat metod modelowania tła.

Warto też podkreślić stworzenie przez Autora przydatnego narzędzia do oceny jakości testowanych metod modelowania tła oraz wykonanie systematycznych badań wszystkich przedstawionych metod w wielu wariantach. Również słowa uznania należą się za dokonanie implementacji metod

modelowania tła w sposób wielowątkowy. Wartościowe są nagrane przez Autora sekwencje testowe, zwłaszcza ciekawy był pomysł nagrania podobnych ujęć w różnych porach dnia i w różnych porach roku. Niektóre z sekwencji nagranych przez Autora cechuje bardzo wysoki poziom trudności z punktu widzenia prawidłowego określenia masek obiektów pierwszoplanowych.

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 generalnie na dobrym poziomie. Rozdział 2 ma niewątpliwe walory dydaktyczne. Opis metod własnych w rozdziale 3 pozwala prześledzić pomysły Autora, choć nie dostarcza wszystkich szczegółów, które niekiedy można odnaleźć w rozdziale 4. Z kolei w rozdziale 4 bardzo liczne szczegółowe wykresy ilustrujące wyniki dla każdej z sekwencji testowej budzą mniejsze zainteresowanie niż ciekawsze tabele podsumowujące. Bardzo przejrzysto zredagowany jest rozdział 5 zwłaszcza rysunki 65 i 66. Dobrym pomysłem było zamieszczenie obrazów dla sekwencji testowych. Jest ich dość sporo, ale wszystkie wydają się potrzebne. Na rysunku 21d można było zastosować negatyw, dla lepszego uwidocznienia krawędzi. Zabrakło ilustracji krok po kroku metody eliminacji cieni opisanej na stronie 56. Warto za to wyróżnić staranne przygotowanie tabel 4 i 5, zwłaszcza za zredagowanie kolumny „Trudność do pokonania”, pomocnej przy analizowaniu wyników.

Pod względem językowym rozprawa stoi na wysokim poziomie. Biorąc pod uwagę trudność i zawłość zagadnień, czytanie i analiza tekstu przebiega sprawnie. W całym tekście recenzent odnalazł pojedyncze literówki czy błędy językowe – głównie interpunkcyjne, co na tle innych recenzowanych prac doktorskich wyróżnia pracę bardzo pozytywnie. Pewnym mankamentem jest tylko brak korekty pojedynczych liter na końcu wiersza (tzw. sierotek).

W rozprawie znalazło się kilka nieprecyzyjnych sformułowań. Po analizie całej rozprawy recenzent jest przekonany, że są to pewne niezręczności językowe:

1. Str. 11. „Właściwie wszystkie komputerowe systemy wspomagania nadzoru wideo mają jedną podstawową wspólną cechę: swoją pracę opierają na algorytmach detekcji ruchu, czyli wydzieleniu obszarów pierwszoplanowych z tła. Przy czym „tło” jest tu rozumiane jako zbiór pikseli o stałych cechach, na przykład barwie lub częstotliwości zmian intensywności [3].”

Nieścisłości są dwie: obiekty pierwszoplanowe mogą być okresowo lub długotrwale statyczne (Autor sam podkreślał problem wtapiania się obiektów pierwszoplanowych w tło). Z kolei piksele tła krótko- i długoterminowo zmieniają swoje wartości a nawet cechy.

2. Str. 11 „Idea detekcji obiektów pierwszoplanowych poprzez usuwanie tła opiera się na założeniu, że różnice pomiędzy obecnie analizowaną klatką wideo a wyznaczoną klatką tła spowodowane są wyłącznie obecnością obiektów pierwszoplanowych.”

Ściśle biorąc, w metodach wielowariantowych (tj. w takich, w których dopuszcza się więcej niż jedną reprezentację modelu tła - przykład GMM), nie wyznacza się i nie przechowuje klatki tła w całości (jako pojedynczego obrazu), a jedynie klasyfikuje się, czy dany piksel należy do pierwszego planu bądź tła na podstawie jego podobieństwa do kilku wariantów (w GMM do rozkładów Gaussa)

i dalej:

„Oznacza to, że piksele z obecnej klatki, które nie zmieniły się w porównaniu z pikselami z klatki

tła, będą zaklasyfikowane jako tło, pozostałe natomiast jako obiekty pierwszoplanowe (nazywane również obiektami ruchomymi z racji tego, że tylko obiekty pierwszoplanowe się poruszają).”

Przekonanie, że tylko obiekty pierwszoplanowe się poruszają, jak już wspomniano powyżej, jest nieuzasadnione – przykładem może być drzewo kołysane wiatrem, fontanna, rzeka, wodospad i inne ruchome obiekty, które zalicza się do tła.

3. Str. 16. „Binarny obraz zawierający idealny wynik modelowania tła nazywany jest wzorcem modelowania tła”

Można mówić jedynie o wzorcu maski pierwszoplanowej, jak już wspomniano, w metodach wielokontekstowych, czy choćby hybrydowych, nie powstaje „integralna klatka tła” lecz maska obiektów pierwszoplanowych.

4. Str. 45. „Warto tutaj zauważyć, że w metodzie tej moduł różnicy wartości piksela i mediany jest porównywany z sześciokrotnością wariancji w przeciwieństwie do dwuipółkrotności użytej w metodzie GMM. Wartość ta została dobrana empirycznie jako optymalna dla końcowych wyników modelowania tła.”

Rozumiem, że Autor użył słowa „optymalny” w sensie potocznym, tj. miał na myśli najlepszą wartość spośród uzyskanych w eksperymentach (np. metodą prób i błędów), a nie wartość optymalną w sensie ścisłym tj. w sposób jaki używa się tego pojęcia w automatyce.

5. Str. 54. „Strukturalnie metoda łączenia mieszanin median, nazywana również *United Median Mixtures* (UMM), składa się z trzech pracujących niezależnie od siebie instancji metody mieszania median, zwanych dalej rdzeniami.”

Nazwa rdzeń oznacza jeden z równolegle pracujących elementów obliczeniowych we współczesnych procesorach. Przypisanie i utożsamienie określenia „rdzeń” do zadań algorytmu, wykonywanych równolegle, nawet jeśli każde z nich wykonywane jest w osobnym rdzeniu procesora, wydaje się niewłaściwe.

Punktem dyskusyjnym w trzech przypadkach są zaproponowane nazwy nowopowstałych metod. Nazwa „metoda najbliższego sąsiada”, kojarzona jest od wielu lat w rozpoznawaniu obrazów z klasyczną grupą metod klasyfikacji. Jest ona też używana w kontekście sposobu zmiany rozdzielczości obrazu. (Już używanie tej samej nazwy w tych dwóch przypadkach prowadzi do niekiedy do nieporozumień.) Z kolei określenie metoda czasowo-częstotliwościowa używane jest dla grupy transformat czasowo-częstotliwościowych np. bardzo popularnej transformaty falkowej. Zdaniem recenzenta również zaproponowana nazwa „metoda przestrzenie-częstotliwościowa” mogłaby zostać nazwana po prostu „metodą częstotliwościową”.

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

Krytyczne uwagi do rozprawy przedstawione są w dużej mierze w tekście recenzji. Głównym zarzutem, na który zwrócił uwagę recenzent, jest brak odniesienia do metod modelowania tła powstałych po 2013 roku. Uwaga ta dotyczy aspektu porównania zaproponowanych w pracy metod pod względem jakości uzyskiwanej maski obiektów pierwszoplanowych w stosunku do aktualnego stanu wiedzy.

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Oceniając przydatność rozprawy dla nauk technicznych należy wskazać, że opracowane metody

medianowe mogą zostać zastosowane w aplikacjach działających w czasie rzeczywistym w miejsce bardzo popularnych metod wykorzystujących rozkład Gaussa. Metody medianowe mogą się okazać również bardziej przydatne (niż metody oparte na rozkładzie Gaussa) w tzw. implementacjach sprzętowych czyli wykonywanych w układach ASIC i FPGA.

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 pomimo uwag dyskusyjnych i krytycznych rozprawa **spełnia stosowne kryteria określone w Ustawie o stopniach naukowych i tytule naukowym** stawiane pracom na stopień doktora nauk technicznych. i wnioskuję o dopuszczenie mgr inż. Piotra Graszki do dalszych etapów przewodu doktorskiego, w tym do publicznej obrony pracy.

A handwritten signature in blue ink, appearing to read 'Piotr Graszka', is written across the lower right portion of the page.