

RECENZJA

rozprawy doktorskiej pt.

Automatic vision quality enhancement of CCTV data

Pana mgr inż. Adama Konieczki

Streszczenie recenzji

- | | |
|------------------------------------|---------------------|
| 1. Krótka charakterystyka rozprawy | 5. Poprawność |
| 2. Zawartość merytoryczna rozprawy | 6. Wiedza kandydata |
| 3. Problem badawczy - znaczenie | 7. Inne uwagi |
| 4. Wkład autora | 8. Podsumowanie |

1. Krótka charakterystyka rozprawy

Recenzowana rozprawa doktorska dotyczy zagadnienia automatycznego poprawiania jakości obrazów w systemach telewizji przemysłowej CCTV (*Closed Circuit TeleVision*). Koncentruje się ona na dwóch problemach: 1) redukcji szumu impulsowego, występującego samodzielnie lub w obecności dodatkowego, addytywnego szumu gaussowskiego, oraz 2) lokalnej poprawie zakresu dynamicznego obrazu. W pracy doktorant udowadnia tezę, sformułowaną na stronie 4 (po polsku na str. vii):

„Opracowane i eksperymentalnie zweryfikowane metody redukcji szumu i powiększenia lokalnego zakresu dynamicznego obrazu w systemach wizyjnych zwiększają skuteczność ekstrakcji zawartości informacyjnej ze strumienia wideo zarówno w systemach nadzorowanych przez człowieka oraz systemach automatycznych”.

W rozprawie zaproponowano nowe, oryginalne, efektywne rozwiązania algorytmiczne, dedykowane do pracy w czasie rzeczywistym. Zwraca uwagę rozległość i szczegółowość przeprowadzonych badań eksperymentalnych, dotyczących testowania („kalibrowania”) nowych narzędzi oraz ich porównania, zakończonego sukcesem, ze znanymi rozwiązaniami światowymi.

Przedstawione w pracy wyniki uzyskano w ramach projektu INDECT (*„Intelligent information system supporting observation, searching and detection for security of citizens in urban environment”*, 01.01.2009-30.06.2014, <http://www.indect-project.eu/>), finansowanego przez 7 Program Ramowy Unii Europejskiej. Doktorant jest autorem 18 publikacji, związanych z tym projektem, z których 9 bezpośrednio dotyczy jego rozprawy doktorskiej (4 publikacje w czasopiśmie: *Przegląd Elektrotechniczny*, 2015, 2016, i *Elektronika*, 2011, 2014, 5 referatów konferencyjnych: IEEE SPA-Poznań, 2013, 2015, 2016, MCSS-Kraków 2011, DCPS-Cottbus 2016).

2. Zawartość merytoryczna rozprawy

Praca jest napisana w języku angielskim. Składa się ona ze: spisu treści (2 strony), streszczenia w języku polskim i angielskim (2 strony), wykazu skrótów i oznaczeń (4 strony), siedmiu rozdziałów (157 stron), w tym wprowadzenia (rozd. 1) i podsumowania (rozd. 7), bibliografii (169 pozycji, 8 stron) oraz 3

dotatków ze szczegółowymi wynikami (35 stron) - łącznie liczy ona 208 stron (!). Do tekstu pracy jest dołączona płyta CD.

Zawartość poszczególnych części pracy jest następująca.

W rozdziale pierwszym, wprowadzeniu (6 stron), przedstawiono rosnącą popularność i zastosowania zdalnego nadzoru wizyjnego, często występujący problem niskiej jakości obrazów, uniemożliwiający ich użycie, oraz tendencje rozwoju współczesnych kamer do monitoringu. Na tym tle postawiono cele pracy, czyli opracowanie szybkich algorytmów obliczeniowych do odsumiania i lokalnej poprawy dynamiki obrazów. Następnie sformułowano tezę pracy, przedstawiono zawartość kolejnych rozdziałów i pokazano powiązanie otrzymanych wyników pracy doktorskiej z dotychczasowymi publikacjami kandydata.

W rozdziale drugim „*State of the art*” (31 stron) na początku, w rozdziale 2.1, krótko opisano wymagania stawiane przed systemami wizyjnymi do zdalnego nadzoru, poszczególne elementy toru akwizycji i przetwarzania wstępnego obrazów, ograniczenia wpływające na złą jakość obrazów i szybki rozwój metod do automatycznego określania ich jakości.

W rozdziale 2.2 szczegółowo scharakteryzowano źródła i rodzaje szumu, występującego w obrazach. Podkreślono występowanie szumu impulsowego oraz mieszanego, impulsowego i gaussowskiego równocześnie [67, 71-79]. Przedstawiono dwa podziały metod redukcji szumu. W pierwszym z nich, metody odsumiania są podzielone na (rys. 2.9 z [69]): przestrzenne i częstotliwościowe, adaptacyjne i nieadaptacyjne, liniowe i nieliniowe (wymieniono liniowe filtry uśredniające i wienerowskie oraz nieliniowe filtry medianowe). W drugiej klasyfikacji wyodrębniono [60]: 1) lokalne modelowanie bayesowskie/gaussowskie, 2) progowanie współczynników transformat, 3) rzadkie kodowanie, 4) lokalne/blokowe uśrednianie. Metody opracowane w rozprawie należą do czwartej grupy.

Podkreślono popularność metod bayesowskiego odsumiania obrazów w dziedzinie transformacji falkowej (z nieliniowym progowaniem, wykorzystującym ukryte modele Markowa i mikstury gaussowskie) oraz metod wykorzystujących ICA (*Independent Component Analysis*), ale zwrócono uwagę na złożoność obliczeniową tych metod i niemożliwość (lub trudność) ich zastosowania w interwencyjnych systemach nadzoru wizyjnego. Następnie skupiono się na metodach prostszych i dokładniej omówiono odmiany „inteligentnych”, przełączanych filtrów medianowych, stosowanych do odsumiania obrazów. Na końcu omówiono najnowsze rozwiązania hybrydowe [95-100].

W rozdziale 2.3 szczegółowo zaprezentowano techniki poprawy zakresu dynamicznego obrazów (HDR). Rozwijają się one ostatnio bardzo szybko i mają na celu zmniejszenie lokalnego niedoświetlenia lub prześwietlenia fragmentów obrazów, wynikających z zastosowanej przesłony i czasu otwarcia migawki. Omówiono pojęcie zakresu dynamicznego, opisano następujące metody stosowane do przeliczania wartości pikseli (*tone mapping*): Millera & Hoffmana [117], Reinharda [118, 113], Drago & Myszkowskiego [119], Ashikhmina [113], Duranda & Dorsey [121], Fattala [122], Mantiuka [123], Pattanaika [124]. Podkreślono rosnące praktyczne zainteresowanie metodami HDR w systemach nadzoru wizyjnego w inteligentnych samochodach [130, 131, 132] oraz rozwój metod ewaluacji technik HDR [141-150].

Należy stwierdzić, że rozdział 2 jednoznacznie potwierdza, że doktorant ma bardzo szeroką i bardzo dobrze usystematyzowaną wiedzę z zakresu tematyki rozprawy. Potwierdza to imponujący, rozległy przegląd literatury, dobrze dobranej i umiejętnie cytowanej w rozdziale 2. Robi wrażenie nie tylko liczba cytowanych pozycji, ale też umiejętność: sprawnego porządkowania zawartej w nich informacji/wiedzy, łączenia faktów i wyciągania poprawnych wniosków.

Rozdział trzeci „*Proposed method for impulse and Gaussian-impulse noise removing*” (24 strony) jest jednym z dwóch najważniejszych rozdziałów pracy. Przedstawiono w nim zaproponowaną metodę poprawy jakości obrazów, w których występuje szum impulsowy dwuwartościowy (0/1, „*salt and pepper*”) lub impulsowy losowy-wielopoziomowy. Nowa metoda łączy w sobie zalety metod uśredniających i medianowych. W sposób iteracyjny znajduje się w niej piksele zaszumione, a następnie zastępuje się je średnią wartością tych pikseli sąsiednich obrazu oryginalnego, które uznano, po wielu iteracjach, za niezaszumione (rozmiar sąsiedztwa „*j*” może być stały lub adaptowany).

Na początku wszystkie piksele obrazu są traktowane jako niezaszumione, a następnie decyzja ta jest weryfikowana i zmieniana w kolejnych iteracjach na podstawie porównania bezwzględnej wartości różnicy pomiędzy konkretną próbką obrazu, a średnią wartością próbek sąsiednich, zaliczonych jako „niezaszumione” w poprzedniej iteracji (3.4)(3.9). Ogólny schemat zasady pracy zaproponowanego algorytmu jest przedstawiony na rysunku 3.2. Na początku, na podstawie dwóch wersji tego samego obrazu – niezaszumionego i zaszumionego, są wyznaczone zbiory wartości progów decyzyjnych „piksel zaszumiony/niezaszumiony” dla kolejnych etapów poszukiwania pikseli zaszumionych – rys. 3.5. Potem obliczone wartości progowe są stosowane do iteracyjnego odsumiania jakiegoś obrazu – tego samego lub innego – według schematu pokazanego na rys. 3.6. „Promień” sąsiedztwa „ j ” może być wybierany arbitralnie lub adaptacyjnie – rysunek 3.7.

Wstępnie zbadano wpływ przyjętych wartości promienia sąsiedztwa „ j ” (1, 3, 5) oraz tolerancji „ U ” (0, 12, 24) (3.8) na otrzymywane wyniki (tabela 3.1). W większości przypadków najskuteczniejszym okazywał się adaptacyjny dobór wartości „ j ”, a wybór wartości „ U ” był nieistotny.

Następnie wstępnie sprawdzono poprawność działania algorytmów (rys. 3.9-3.17): sprawdzono histogramy alokacji pikseli, wyznaczone progi detekcyjne oraz poprawność podejmowanej decyzji dla progu „ T ” (3.8) o różnej wartości.

Pod koniec rozszerzono metodę o możliwość wyznaczania zbiorów progów detekcyjnych, wieloma parami różnych obrazów o różnym poziomie szumu (rys. 3.18) oraz o możliwość odsumiania dowolnego obrazu (rys. 3.19). W tym przypadku algorytm odsumiający sam estymuje poziom szumu (3.11)(3.12) i na tej podstawie dobiera odpowiednie zbiory progów decyzyjnych. Skuteczność tego podejścia zweryfikowano na rysunkach 3.21-3.23.

Ostatecznie rozszerzono algorytm do wersji obrazów kolorowych, 14-bitowych (nowe kamery CCTV), dopasowano jego parametry (m.in. wartość progu „ T ” (3.8)) i powtórzono powyższe testy, ale tylko dla schematu odsumiania pojedynczego obrazu R, G, B (rys. 3.24 – 3.27). Obraz referencyjny uzyskano jako sumę wielu zaszumionych obrazów.

W rozdziale czwartym „De-noising experiments for test images” (53 strony) przedstawiono wyniki bardzo dokładnego testowania opracowanej metody odsumiania. Wybrano standardowe obrazy testowe (rys. 4.1), cztery czarno-białe: *Barbara*, *Boat*, *Goldhill*, *Lena*, oraz cztery kolorowe: *Airplane*, *Baboon*, *Fruits*, *Peppers*, oraz samemu przygotowano dodatkowych 8 obrazów (rys. 4.2). Obliczenia przeprowadzono w programie Matlab R2010b. Do zaszumiania wykorzystano program `impulsenoise.m` ze strony Mathworks [152] oraz funkcję Matlaba `imnoise()` (miksury gaussowskie). Każdy z 16 obrazów był zaszumiany 20 różnymi typami szumów (ile realizacji dla każdego przypadku?). SNR dla każdego sygnału jest podany w tablicach 4.3 – 4.6, średni w tabeli 4.7. W przypadku nagrań CCTV użyto 6 obrazów testowych (rys. 4.8), dla których obrazy referencyjne obliczono metodą uśredniania 30 kolejnych ekspozycji (tabela 4.9 - przetestowano 6 różnych metod uśredniania, wybrano pierwszą z nich jako porównywalną z innymi, a najprostszą). Ponieważ testy przeprowadzono dla czułości ISO {400, 800, 1600, 3200, 6400} łącznie rozpatrywano 30 przypadków.

Dla przypadku odsumiania pojedynczych obrazów dobrano najlepszą wartość tolerancji „ U ” (3.8) oraz przebadano zmienność PSNR obrazu odsumionego w funkcji numeru iteracji (rys. 4.17-4.23). Wykazano, że zbiory wartości progowych, obliczonych dla jednego obrazu, nie muszą być dobre dla innego obrazu – nie ma jednej prostej reguły ich wyboru (podmiany) (tabele 4.11 – 4.14). Najlepsze wyniki uzyskiwano, kiedy stosowano zbiór progów, obliczony dla konkretnego obrazu (ograniczenie tej metody!) (rys. 4.24-4.25).

Następnie przetestowano metodę z wyznaczaniem zbiorów wartości progowych wieloma obrazami.

Rozważono cztery przypadki (S - obrazy standardowe, P - obrazy przygotowane, U - uczenie, O - odsumianie): UO = SS, SP, PP, PS. Pokazano, że obrazy z jednego zbioru można skutecznie odsumiać, stosując zbiory progów, wyznaczonych dla innego zbioru obrazów (rys. 4.27-4.30). Sprawdzono zmianę PSNR w funkcji iteracji (rys. 4.31-4.34). Pokazano, że zastosowanie rozszerzonych zbiorów progów jest korzystniejsze (4.35-4.36).

Potem przetestowano odsumianie obrazów z CCTV. Wyznaczono optymalne wartości: tolerancji $U=25$ (rys. 4.41-4.42) oraz liczby iteracji $i=2$. Potwierdzono, że zbiory progów, wyznaczone dla jednego obrazu, mogą być stosowane dla innego, ale jednak najlepsze wyniki są uzyskiwane, kiedy stosuje się progi obliczone dla konkretnego obrazu (tabela 4.16).

Na końcu rozdziału przedstawiono porównanie opracowanych metod z alternatywnymi podejściami znanymi literatury (!). Wyniki zebrano w tabelach 4.17, 4.19, 4.20 i 4.21 dla czterech różnych kombinacji szumów. **W przypadku szumów wyłącznie impulsowych, zaproponowana metoda okazała się najlepsza w największej liczbie rozpatrywanych przypadków (tab. 4.17 i 4.19).** W sytuacji, kiedy dodatkowo występował szum gaussowski, zawsze najlepsze były metody Lmin lub MK-SVD, obie z [99], jednak o wiele bardziej złożone obliczeniowo, a zaproponowana metoda plasowała się nieznacznie za metodami złożonymi przed metodami medianowymi. Jest to ważne osiągnięcie recenzowanej pracy.

W rozdziale piątym „*Proposed method for local dynamic range increasing*” (21 stron) zaproponowano metodę do lokalnego zwiększania zakresu dynamicznego obrazów. W metodzie stosuje się dwa obrazy: niedoświetlony (DK) oraz prześwietlony (BR). Na początku są one wyostrzane (5.1) w celu zwiększenia szczegółowości, a następnie analizowane pod względem jasności (5.3)(5.4) w celu znalezienia obszarów pikseli, które muszą być koniecznie wykorzystane na dalszych etapach przetwarzania (gdyż nie są niedoświetlone lub prześwietlone). Z dwóch wyostrzonych obrazów jest składany jeden obraz przejściowy IMS (5.2) oraz obliczana jest mapa istotności jego pikseli (5.6). Potem tworzona jest wstępna mapa tonalna T_1 (5.7)(5.8), zawierająca tylko informację dotyczącą kontrastu w otoczeniu każdego piksela obu obrazów. Następnie mapa ta jest przetwarzana ($T_1 \rightarrow T_2$, rys. 5.15): informacja o kontraście i jasności w każdym fragmencie obrazu jest łączona i uśredniana. Z kolei jest przeprowadzana detekcja krawędzi, aby zapewnić ich odpowiednią jasność w obrazie wynikowym ($T_2 \rightarrow T_3$). Otrzymana w ten sposób mapa tonalna T_3 zawiera dokładną informację, które części obrazów DK, BR i IM mają być wzięte do obrazu wynikowego i w jakich proporcjach. Końcowe równanie mapowania opisane jest przez (5.13). Schemat blokowy algorytmu jest przedstawiony na rys. 5.8. Uproszczenie przetwarzania mapy tonalnej $T_2 \rightarrow T_3$ zaproponowano na rys. 5.22.

W Matlabie przetwarzano własne obrazy DK i BR, małe (1200x900) i ośmiobitowe, które wcześniej tworzono w oprogramowaniu *Digital Photo Professional* ze zdjęć wielopikselowych i wielobitowych.

W rozdziale szóstym „*Experiments and results for the method of increasing local dynamic range*” (20 stron) przedstawiono wyniki walidacji algorytmu z rozdziału 5. Testowano algorytm pełny i uproszczony. Eksperymentalnie dobrano wartości progów T_L (5.4), T_D (5.6) oraz T_E (5.9) (Tabela 6.1). Dobrano wartości parametrów dla metod z literatury (tabela 6.2). Zwiększenie zakresu dynamiki było oceniane przez ludzi w zadanym czasie (tabela 6.3) oraz bez limitu czasowego (rys. 6.5 – 6.10). Obserwatorzy musieli udzielić odpowiedzi szczegółowo-ilościowej na podstawie zsyntezowanych obrazów HDR. Metody opracowane były najlepsze lub bardzo dobre dla każdego testowanego obrazu. W przeprowadzonych testach MOS (*Mean Opinion Score*) (tabele 6.4-6.7, rys. 6.11) metody opracowane też wypadły bardzo dobrze, co ciekawe **zdecydowanie wygrały pod względem oceny szczegółowości obrazów (tab. 6.5), co jest najważniejszym wymaganiem w monitoringu (!).** Użycie opracowanych metod do rozpoznawania tekstu (rozd. 6.3.1) dało porównywalne wyniki z najlepszą z metod (tabela 6.8), a zastosowanie ich do wzrokowej rozpoznawania twarzy (rozd. 6.3.2), według recenzenta, dało najlepszy wynik (rys. 6.18), ze względu na największą szczegółowość uzyskanych obrazów.

W ostatnim, siódmym rozdziale, podsumowaniu (2 strony), krótko podkreślono zalety zaproponowanych w pracy rozwiązań.

Pracę zamyka **wykaz cytowanej literatury** (8 stron), liczący 169 pozycji. Jest on bardzo rozbudowany, reprezentatywny, aktualny i dobrze dobrany. Zawiera on bardzo różne materiały źródłowe, od książek poprzez artykuły w czasopismach, referaty konferencyjne, dokumentację techniczną urządzeń i oprogramowania, aż po informacyjne strony internetowe. Powyżej, podczas omawiania rozdziału 2, podkreślono, że bibliografia jest bardzo dobrze wykorzystywana w pracy – umiejętnie omówiona i cytowana. **Całościowo praca daje bardzo dobry przegląd aktualnego stanu rozwiązań światowych w zakresie rozpatrywanych w niej zagadnień.**

W dodatku 1 (7 stron) przedstawiono 12 rozbudowanych ilustracji graficznych - przykładów odsumowania obrazów testowych metodą SNP (*single noise profile*).

W dodatku 2 (24 strony) przedstawiono 24 tabele z wynikami odszumiania obrazów testowych dla różnych wartości parametry U (od wartości 0 do 48 co 4).

W dodatku 3 (4 strony) przedstawiono 4 tabele z wynikami klasyfikacji warstw RGB obrazów testowych do różnej liczby grup obrazów, ze względu na poziom szumu.

3. Problem badawczy i jego znaczenie

Jak już było powiedziane w punkcie 1 recenzji, w rozprawie są rozpatrywane dwa problemy naukowe, mające bardzo duże znaczenie praktyczne, związane z przetwarzaniem obrazów: 1) redukcja szumu impulsowego, występującego samodzielnie lub w obecności dodatkowego, addytywnego szumu gaussowskiego, oraz 2) lokalna poprawa zakresu dynamicznego obrazu. Są to dwa bardzo ważne, trudne i aktualne problemy. Pierwszy z nich jest ciągle aktualny, a drugi – wyjątkowo ciekawy i aktualny obecnie, w czasach wchodzenia nowych standardów wizualizacji obrazów HDR 10 (*High Dynamic Range*) i *Dolby Vision*.

W dzisiejszych czasach, z jednej strony zagrożonych terroryzmem, a z drugiej – wypełnionych marzeniami o bezpiecznych samochodach autonomicznych, umiejętność szybkiej i dokładnej analizy zawartości danych wizyjnych jest bezcenna. Niestety, zbyt często spotykamy się obecnie z sytuacją, że pomimo posiadania zdalnych systemów nadzoru wizyjnego (monitoringu), ich użytkownicy nie są w stanie zidentyfikować sprawców lub zdarzeń z powodu złej jakości nagrań, głównie występującego w nich szumu (zła jakość kamery, złe warunki oświetleniowe) oraz lokalnych prześwieleń lub niedoświeleń obrazów (zła ekspozycja zdjęć – za mała lub za duża przesłona oraz czas akwizycji). Z tego powodu recenzowana praca trafia w 10-tkę, gdyż dotyczy jednego z najważniejszych problemów współczesnego monitoringu – poprawy jakości uzyskiwanych obrazów już po ich akwizycji, na etapie przetwarzania wstępnego – przed właściwym rozpoznawaniem (automatyczną interpretacją) obrazów. Za szczególnie aktualne uważam zajęcie się przez doktoranta zagadnieniem lokalnej poprawy zakresu dynamicznego obrazów, często niedostrzeganym lub ignorowanym, a obecnie niezwykle „gorącym” (obrazy HDR – *High Dynamic Range*, np. w smartfonie Samsung Galaxy S8 aparat szybko wykonuje serię 3 zdjęć o różnej ekspozycji, z których metodami obliczeniowymi uzyskiwany jest jeden obraz lokalnie uzdatniony).

Oczywiście, pierwszy problem badawczy rozprawy, dotyczący odszumiania obrazów, **nie jest nowy** i jest bardzo intensywnie rozwijany od wielu lat. Wynika to z faktu, że dawniej urządzenia pozyskiwania obrazów (aparaty, kamery) były technicznie znacznie gorsze i z tego powodu zdjęcia koniecznie wymagały cyfrowej obróbki już po etapie akwizycji. Jednak **znaczenie tego problemu gwałtownie rośnie** z powodu faktu, że obecnie monitoring wizyjny jest stosowany powszechnie. Bardzo często, z powodów historycznych i finansowych, wykorzystywany obecnie sprzęt nie jest najwyższej jakości. Dlatego jeśli można, to trzeba wspomóc ten sprzęt obliczeniowo i poprawić jego parametry użytkowe. I tym właśnie zajmuje się naukowo doktorant.

Ze względu na szybko rosnące możliwości techniczne cyfrowych kamer i aparatów fotograficznych, drugi problem rozprawy, dotyczący lokalnej poprawy dynamiki, wydaje mi się **szczególnie aktualny i ważny**. Pokazane w pracy wyniki, uzyskane w tym zakresie, są imponujące.

Istnieje bardzo wiele metod odszumiania obrazów, od bardzo prostych do bardzo złożonych. Mają one różne zastosowania. Doktorant koncentruje się wyłącznie na prostych, szybkich obliczeniowo metodach analizy i przetwarzania pikseli obrazów, a nie na złożonych metodach analizy i przetwarzania współczynników transformat obrazów, mając na uwadze zastosowanie tych metod w systemach monitoringu czasu rzeczywistego. W pracy wykazuje, że zaproponowane przez niego rozwiązania mogą z sukcesem konkurować z metodami znanymi z literatury, zarówno w kwestii odszumiania obrazów jak i lokalnej poprawy ich zakresu dynamicznego.

4. Wkład autora

W pracy zaproponowano dwa nowe, autorskie rozwiązania, które dokładnie przetestowano i porównano z metodami znanymi z literatury. Dokładnie wykazano, że nowe metody nie są gorsze niż znane, stosowane rozwiązania, a często nawet lepsze. Nową metodę odsumiania przedstawiono w rozdziale 3 i przetestowano w rozdziale 4. Nową metodą poprawy zakresu dynamicznego obrazów przedstawiono w rozdziale 5 i przetestowano w rozdziale 6. Zawartość tych rozdziałów i uzyskane wyniki dokładnie scharakteryzowano w punkcie drugim recenzji. Uważam, że szczególnie wartościowa jest opracowana druga metoda, gdyż **okazała się ona najlepsza ze wszystkich testowanych metod, jeśli chodzi o szczegółowość przetworzonych obrazów, a według recenzenta jest to właśnie najbardziej pożądana cecha w monitoringu.**

5. Poprawność

Badania zostały poprawnie zaplanowane i przeprowadzone (z małymi wyjątkami, o tym dalej). Rozprawa doktorska została wzorowo napisana jako szczegółowy raport z wykonanych prac badawczych. Bardzo dobrze przedstawiono w niej: problemy naukowe, istniejący stan wiedzy, zaproponowane metody, ich walidację i porównanie z rozwiązaniami konkurencyjnymi oraz wnioski, wynikające z przeprowadzonych prac.

W części dotyczącej odsumiania mam jednak pewne uwagi **dyskusyjne.**

1) Zaproponowany algorytm jest matematycznie (programowo) prosty, co wcale nie znaczy, że jest on szybki (efektywny obliczeniowo). Doktorant obiecuje „czas rzeczywisty”, ale w rozdziale z wynikami eksperymentalnymi nie zajmuje się pomiarem (oszacowaniem) czasu obliczeń.

2) Zaproponowany algorytm jest trochę za mało matematyczny, a za bardzo heurystyczny (iteracyjnie wypełniony, z wartościami parametrów obecnie dobieranymi arbitralnie, wymagający obrazów treningowych „czystych” i „zaszumionych”). Czyli jest za bardzo „*hard*”, a za mało „*soft*” (adaptacyjny, inteligentny). Za bardzo dedykowany, za długo kalibrowany, za mało uniwersalny. W pracy widać dużą zależność otrzymanywnych wyników od wartości „nastaw”. Dlatego słowo „automatic”, występujące w tytule rozprawy, moim zdaniem nie do końca jest uzasadnione.

Przykłady:

str. 42: przyjęto $v=10$ dla obrazów 8-bitowych,

str. 43: przyjęto 150 jako maksymalną wartość średniego bezwzględnego odchylenia pikseli dla obrazów 8-bitowych,

str. 45: wzór (3.10) – dlaczego przyjęto, że promień „ r ” to tylko wartości 1, 3, 5 (?),

str. 48: założenie 1500 wartości optymalnych progów,

str. 43: stwierdzenie: „Wartość U musi być eksperymentalnie dobrana dla każdego rodzaju i poziomu szumu”.

3) Wydaje mi się także, że algorytm odsumiania nie jest do końca poprawnie zbudowany i testowany. Monitoring to sekwencja obrazów, a doktorant testuje pojedyncze obrazy i w ogóle nie wykorzystuje faktu, że obrazy ciągle same do niego „płyną”, co powinno poprawić skuteczność odsumiania (brak jest przetwarzania obrazów w osi czasu). Jeśli CCTV to dynamicznie zmienna scena, to czy zaproponowana metoda jest odpowiednia dla CCTV z powodu konieczności uśredniania wielu obrazów w celu otrzymania jednego obrazu referencyjnego?

4) Według mnie metoda odsumiania, projektowana dla monitoringu, powinna mieć mechanizmy, które pozwolą jej zaadaptować się samej do mniej lub bardziej zmiennej sceny obserwacji i zmiennego, najczęściej powoli, oświetlenia.

5) Nie jestem też przekonany, że eksperyment odsumiania prawie idealnych pojedynczych obrazów, sztucznie zaszumionych idealnymi wzorcami szumów, jest lepszy niż eksperyment odsumiania sekwencji obrazów z monotonicznie pogarszającymi się warunkami obliczeniowymi.

6) Aby wyciągać wnioski statystyczne trzeba przebadać wiele realizacji szumu. W doktoracie dodawano do obrazów testowych pojedyncze realizacje szumu, jak zrozumiałem. Przedstawione wyniki mogą więc być tylko poglądowe. Czy na str. 66 tylko jedna realizacja szumu była użyta dla każdego przypadku z tablicy 4.2? Jeśli tak było, to czym to było umotywowane? Według mnie zbiór obrazów zaszumionych powinien być większy i zawierać różne realizacje szumu.

7) Brakuje mi w rozprawie jednej tabeli, w której zbiorczo byłyby przedstawione dla wielu metod średni zysk PSNR, uzyskany po odszumieniu wielu typów obrazów z wieloma realizacjami określonych typów szumów. Obecnie praca liczy ponad 200 stron, z których zdecydowana większość zawiera tabele z wynikami dla mnogości przypadków. Duża liczba tabel przekonuje czytelnika o wysiłku włożonym w znalezienie dobrego rozwiązania, ale nie koniecznie o jego znalezieniu. W tym aspekcie w pewnych częściach rozprawa jest napisana jak raport z projektu, w którym najważniejsze jest samo „sprawozdanie”.

8) W pracy w rozdziale 2 omawia się obiektywne i subiektywne metody do pomiaru jakości obrazów, ale w rozdziałach 3 i 4 stosuje się tylko najprostsze miary „obiektywne” PSNR i MSE, które nie zawsze są najlepsze.

9) Uważam, że tytuł rozdziału 3 jest niepoprawny: metoda redukuje szum, a nie kompletnie go usuwa.

6. Wiedza kandydata

Przeładowy rozdział 2 pracy oraz bogata i umiejętnie cytowana w nim literatura robią bardzo dobre wrażenie. Doktorant bardzo dobrze czuje się w każdym poruszonym temacie, widzi i rozumie rzeczy małe i duże, trafnie ocenia i porównuje, podejmuje poprawne decyzje. W konsekwencji tego opisane w pracy badania są przeprowadzone logicznie i konsekwentnie oraz zakończone sukcesem, a sam tekst rozprawy doktorskiej jest wzorowo napisany pod względem merytorycznym oraz edycyjnym. Nie powinno to dziwić. Pan mgr inż. Adam Konieczka to już bardzo doświadczony badacz, autor 28 publikacji, który zajmował się wieloma problemami (pierwsza publikacja w 2008 roku). Świadczą o tym dane biblio-metryczne, przedstawione w tabeli poniższej (zebrane w dniu 2 sierpnia 2017 roku). Liczba i „widzialność” publikacji kandydata jest rzeczywiście bardzo duża. Jednak, a szkoda, że są to prawie wyłącznie publikacje krajowe (lub międzynarodowe wydawane w kraju) i wieloautorskie (konsekwencja pracy w dużym, prężnym zespole, szybko zmierzającym do celu i bez zbytej zwłoki sprawozdającym wyniki swoich prac projektowych). Równocześnie, jak na dużą liczbę prac, prace te nie są często cytowane przez innych (za to często przez współpracujący zespół).

Baza danych	Liczba publikacji	Liczba cytowań (bez autocytowań)	h-indeks (bez autocytowań)
WoS	16	18 (3)	2 (?)
Scopus	20	45 (3)	4 (1)
Google scholar	28	89 (?)	5 (?)

7. Inne uwagi

Zalety i mocne strony rozprawy:

- 1) opracowana metoda poprawy zakresu dynamiki daje najbardziej szczegółowe obrazy,
- 2) opracowanie i przebadanie dwóch metod, a nie jednej,
- 3) bardzo dobry opis stanu wiedzy, poparty bogatą literaturą,
- 4) staranne porównanie się z wieloma, nowymi metodami alternatywnymi,
- 5) rozległość i skrupulatność przeprowadzonych badań eksperymentalnych,
- 6) wzorowa edycja pracy.

Wady i słabe strony rozprawy

- 1) za mało całościowych wyników, za duża objętość pracy (ponad 200 stron),
- 2) zaproponowane metody odszumiania są mało elastyczne (wybór wielu parametrów).
- 3) drobne błędy metodologiczne, moim zdaniem.

8. Podsumowanie

Po starannej lekturze rozprawy nie mam żadnej wątpliwości, że Pan mgr inż. Adam Konieczka zasługuje na stopień naukowy doktora nauk technicznych. W recenzji szczegółowo omówiłem zawartość rozprawy jak i uzyskane wyniki. Jedno i drugie robi duże wrażenie: starannie i logicznie napisana praca, licząca ponad 200 stron oraz ogrom przeprowadzonych badań i wartościowość uzyskanych wyników. Uważam, że spokojnie każda z dwóch opracowanych przez kandydata metod zasługuje osobno na doktorat. Osobiście bardziej sobie cenię nową metodę poprawy jakości zakresu dynamicznego obrazu. **Przedstawione w rozprawie wyniki w pełni potwierdzają postawioną w niej tezę.**

Zdecydowanie uważam, że:

- praca zawiera oryginalne rozwiązanie dwóch problemów naukowych,
- kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie Automatyka i Robotyka (Systemy wizyjne),
- kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej.

Dlatego, uwzględniając bardzo bogaty dorobek publikacyjny doktoranta z zakresu rozprawy (9 publikacji, w tym 4 w czasopismach krajowych) oraz opracowanie dwóch metod, a nie jednej, nie tylko **wnioskuję o dopuszczenie go do dalszych etapów przewodu doktorskiego, w tym do oficjalnej obrony**, ale także o **wyróżnienie jego pracy**, jeśli tylko spełnia ona inne, formalne wymagania Wydziału Informatyki Politechniki Poznańskiej.

Kraków, 7.08.2017

Tomasz Zieliński