

Dr hab. inż. Grzegorz ĆWIKŁA, Prof. PŚ
Politechnika Śląska
Wydział Mechaniczny Technologiczny
Katedra Automatykacji Procesów Technologicznych
i Zintegrowanych Systemów Wytwarzania
44-100 Gliwice, Konarskiego 18a
e-mail: grzegorz.cwikla@polsl.pl

Gliwice, 14.03.2025

RECENZJA

rozprawy doktorskiej mgr inż. Adama SURÓWKI
pt. „Synteza zaawansowanych technik wizyjnych w inteligentnym
systemie nadzoru obiektów użytku publicznego”

Podstawa opracowania recenzji

Opinię opracowano na zlecenie Rady Naukowej dyscypliny *Automatyka, elektronika, elektrotechnika i technologie kosmiczne* Wydziału Inżynierii Elektrycznej i Komputerowej Politechniki Krakowskiej z dnia 15 stycznia 2025 (uchwała 3/2025).

1. Wstęp

Recenzowana praca skupia się na problematyce opracowywania zaawansowanych algorytmów interpretacji obrazów z systemów monitoringu wizyjnego. Podlegająca recenzji rozprawa doktorska została wykonana pod kierownictwem prof. dra hab. inż. Janusza Gołdasza oraz promotora pomocniczego dra Grzegorza Litawy. Zagadnienie analizy informacji wizyjnej w czasie rzeczywistym jest obecnie tematem wielu badań i prac rozwojowych, a skupienie się na automatycznym monitoringu budynków i przestrzeni użyteczności publicznej ma głębokie uzasadnienie społeczne. Wyniki niniejszej pracy budzą nadzieję na poprawę bezpieczeństwa publicznego przy jednoczesnej redukcji obaw o kwestie prywatności oraz ochrony danych i wizerunków osób. Opiniowana rozprawa doktorska mgr inż. Adama Surówki wpisuje się w aktualny trend badawczy i odpowiada na współczesne potrzeby społeczne.

2. Opis rozprawy doktorskiej

Objętość recenzowanej rozprawy doktorskiej to, z uwzględnieniem dodatków i spisów, 223 strony. Praca składa się ze spisu treści, wykazu oznaczeń i skrótów, wykazu zmiennych stosowanych we wzorach, siedmiu głównych rozdziałów, spisów dorobku naukowego Doktoranta (6 pozycji w których jest jedynym autorem, z czego jedna jest jeszcze w fazie recenzji), ilustracji, tabel i literatury, jak również streszczeń w językach polskim i angielskim. Wykaz literatury liczy, łącznie ze źródłami internetowymi, 143 pozycje (Doktorant nie zamieścił tu źródeł własnego autorstwa lub współautorstwa). Do pracy nie dodano załączników zawierających szczegółowo opisane algorytmy ani listingów programów.

W pierwszym rozdziale pracy, zatytułowanym *Wprowadzenie*, Doktorant nakreślił genezę zagadnienia stanowiącego przedmiot rozprawy oraz uzasadnienie zasadności podjęcia tematu

badan, a takze sformulowal cel, tezy i zakres pracy. Akcentowany jest tutaj zamiar nadania jej uzytelnego charakteru poprzez zastosowanie szeroko dostepnego sprzetu i wykorzystanie jej wynikow do podniesienia poziomu bezpieczenstwa publicznego. Opisano tu takze uklad treści rozprawy.

Drugi rozdzial pracy, zatytułowany *Zaawansowane techniki analizy sygnalu wizyjnego*, zawiera systematyczny zagadnień wchodzących w zakres pracy w oparciu o publikacje dotyczące wstepnego przetwarzania obrazów, kamer stosowanych w systemach monitoringu i ich kalibracji, metod detekcji i klasyfikacji obiektów, technik rozpoznawania znaków, metod estymacji sylwetki człowieka, algorytmów śledzenia obiektów, technik analizy aktywności człowieka, a wreszcie wskaźników oceny jakości działania systemu wizyjnego.

Rozdział trzeci pracy to kontynuacja przeglądu literatury, tym razem dotycząca inteligentnych systemów nadzoru wizyjnego. Omówiono tutaj obszary zastosowań, klasyfikację systemów IVSS i rodzaje architektury tych systemów oraz przegląd rozwiązań w obszarze IVSS zestawiony w formie tabelarycznej. W wyniku szerokiej i dogłębnej analizy publikacji naukowych Doktorant wskazał luki badawcze, będące motywacją do prac rozwojowych. Analiza literatury została przeprowadzona poprawnie, można jednak zauważyć, że skupiała się ona na publikacjach naukowych, nieco pomijając przegląd dostępnych rozwiązań komercyjnych takich systemów – można przypuszczać, że część rozwiązań stosowanych w takich systemach i szczegóły ich implementacji nie były opisywane w publikacjach typowo naukowych ze względu na kwestie poufności.

Opis planowanego systemu IVSS zawarto w rozdziale czwartym, poczynając od koncepcji architektury systemu, przez planowany graficzny interfejs użytkownika, charakterystykę stanowiska badawczego, aż po źródła analizowanego przez IVSS sygnału wizyjnego (oprócz własnych nagrań udostępnionych lokalizacji z udziałem ochotników symulujących scenariusze zdarzeń mających być wykrywanymi przez opracowywany system, Doktorant korzystał też z publicznie dostępnych baz materiałów wideo oraz samodzielnie zaprogramowanych symulacji 3D utworzonych w programie Blender).

Rozdziały piąty i szósty dokumentują odpowiednio wstępną i główną część prac nad IVSS.

Piąty rozdział jest poświęcony wstępnemu przetwarzaniu obrazu (zagadnienia kalibracji kamery, korekty jasności obrazu, algorytmów detekcji i śledzenia obiektów, kalibracji sceny). Doktorant zdecydował się na wykorzystywanie istniejących narzędzi, interfejsów i bibliotek, które dogłębnie przeanalizował i przetestował, preferując rozwiązania otwartoźródłowe i dostępne w języku C++, ale też każdorazowo uzupełniał te rozwiązania własnymi algorytmami i indywidualnym podejściem do zagadnień, które nie były wystarczająco dobrze w jego opinii dostosowane do jego potrzeb.

Najobszerniejszy w rozprawie rozdział szósty (około 70 stron) zawiera opisy algorytmów detekcji zdarzeń alarmowych, rozpoznawanych przez opracowywany system po wstępnym przetworzeniu danych z kamery (lub w pewnych sytuacjach z innego źródła) metodami zaprezentowanymi w rozdziale piątym. Każdorazowo prezentowane jest tło problemu i krytyczna analiza rozwiązań znalezionych w źródłach literaturowych, a następnie propozycja

własnego rozwiązania danego problemu (w niektórych przypadkach bardzo obszerna i wymagająca znacznego wkładu pracy) w postaci algorytmu i opisu stosowanych narzędzi lub programów, procedura przygotowań i badań empirycznych, a finalnie opisywane są uzyskane wyniki. Opisy ilościowe uzyskanych wyników są prezentowane w formie tabel podających przyjęte w rozprawie wskaźniki jakości, takie jak liczba poprawnych i niepoprawnych przypadków zadziałania algorytmu (w tym wskazania fałszywie pozytywne i fałszywie negatywne), dokładność, precyzja i czułość konkretnego algorytmu. Każde działanie jest racjonalnie uzasadniane. Ten sposób prezentacji jest ujednolicony i bardzo przejrzysty, każdy podrozdział kończy się podsumowaniem dotyczącym danego zdarzenia alarmowego, co znacznie ułatwia czytanie i zrozumienie pracy. Nie są przedstawiane szczegóły implementacji algorytmów (brak listingów programów), a uzyskane wyniki przedstawiono w formie syntetycznej dla każdego z wykrywanych zdarzeń. Rozdział 6 jest więc kluczową częścią pracy, zawierającą opis rozwiązań problemów, procedury testowania, wyniki badań i krytyczną dyskusję, dokumentującą ich jakość dzięki wskaźnikom jakościowym. Wyniki poszczególnych metod w końcowej części rozdziału w miarę możliwości są zestawiane z wynikami innych rozwiązań, opisanych w literaturze.

Końcówka rozdziału 6 prezentuje i analizuje też wydajność systemu w sytuacji, gdy równocześnie załączone jest wykrywanie wielu rodzajów zdarzeń, co ma udowodnić przydatność systemu w rzeczywistych scenariuszach, w których można się spodziewać losowego pojawiania się różnych sytuacji, wykrywanych przez IVSS. Ma to istotne znaczenie, ponieważ każdy z równocześnie uruchomionych algorytmów wymaga pewnego czasu na przetwarzanie informacji, a niedostateczna wydajność systemu z włączonym wykrywaniem różnorodnych zdarzeń mogłaby wykluczyć go z praktycznych zastosowań – oczekiwane jest działanie takich systemów w czasie rzeczywistym.

Przedstawiona do recenzji rozprawa doktorska zakończona jest rozdziałem zatytułowanym *Podsumowanie*. Stanowi on poprawne zamknięcie tej bardzo obszernej pracy, komentujące uzyskane rezultaty, prezentujące mocne i słabe strony opracowanych rozwiązań oraz zawierające propozycje kierunków dalszych badań. Finalne podsumowanie nie jest szczególnie obszerne, prawdopodobnie dlatego, że wiele wcześniejszych rozdziałów i podrozdziałów kończy się osobnymi podsumowaniami opisanych w nich zagadnień.

Ogólny układ pracy uważam za właściwy i zgodny z typowymi rozprawami doktorskimi z dziedziny nauk technicznych, w dyscyplinie *Automatyka, elektronika, elektrotechnika i technologie kosmiczne*.

3. Uwagi odnośnie do wyboru tematu, celu, oraz zakresu rozprawy

Przedstawiona do recenzji praca doktorska dotyczy problematyki zautomatyzowanego nadzoru wizyjnego, który może odgrywać kluczową rolę w zapewnieniu bezpieczeństwa publicznego i prywatnego, odciążając ludzi od bezpośredniego sprawowania roli obserwatora oczekującego, czasem przez długi czas, aż wydarzy się któraś z sytuacji wymagających reakcji. Proponowany system ma za zadanie wykrywać szerokie spektrum sytuacji, które mogą wymagać interwencji osoby odpowiedzialnej za dany obszar, straży miejskiej, policji, pogotowia ratunkowego lub straży pożarnej. System taki może pozwolić na zmniejszenie liczby

operatorów przy jednoczesnym zwiększeniu liczby kamer, dostarczających obraz do analizy, w razie potrzeby przyciągając uwagę operatora do tego źródła sygnału, którego automatyczne algorytmy klasyfikujące wykryły jedną z poszukiwanych anomalii. Operator może wtedy podjąć decyzję dotyczącą dalszych działań. Istotne jest, by system nadzoru wizyjnego nie generował nadmiernej liczby alarmów, nie mających rzeczywistego uzasadnienia, ale też by był wystarczająco czuły i selektywny, by nie gubić rzeczywistych sytuacji alarmowych. Systemy tego rodzaju zaczynają być wdrażane na coraz szerszą skalę, dla zachowania bezpieczeństwa publicznego obejmując czasem całe miasta, co wskazuje na zasadność zajęcia się tą tematyką.

Cel pracy nie został bezpośrednio sformułowany, ale Doktorant zadeklarował, że przedmiotem badań jest synteza istniejących, zaawansowanych technik przetwarzania i analizy obrazu w celu opracowania inteligentnej aplikacji komputerowej, realizującej automatyczne wykrywanie wybranych zdarzeń alarmowych. Sformułowane zostały za to dwie tezy pracy:

- „Możliwe jest opracowanie metody syntezy zaawansowanych technik widzenia komputerowego opartych na modelach uczenia maszynowego, w celu stworzenia całościowego i skutecznego systemu nadzoru przestrzeni publicznych, zdolnego do automatycznej detekcji i klasyfikacji wybranych zdarzeń alarmowych w środowisku rzeczywistym, z wykorzystaniem sygnału z konwencjonalnych kamer monitoringu wizyjnego.”
- „Opracowane algorytmy przetwarzania danych wizyjnych, ekstrakcji cech, identyfikacji, klasyfikacji oraz śledzenia umożliwiają efektywne wykrywanie i raportowanie kluczowych z punktu widzenia bezpieczeństwa publicznego aktywności związanych z zachowaniem i interakcjami osób oraz wybranych klas zwierząt i obiektów.”

Doktorant prawidłowo przyjął założenia i ograniczenia, dotyczące wykorzystywanego sprzętu (standardowe dostępne na rynku kamery, komputerowe stacje robocze o standardowej w momencie podejmowania tematu mocy obliczeniowej), oprogramowania (w tym bibliotek programowych), rodzajów obserwowanych miejsc lub obszarów i zestawu zdarzeń lub sytuacji, jakie miały być wykrywane.

Niektóre z wykrywanych przez opracowywany system IVSS zdarzeń mają znaczenie głównie porządkowe lub administracyjne (identyfikacja pojazdów, niepożądana obecność zwierząt, nielegalne parkowanie, wejście/wyjście ze zdefiniowanej strefy, naruszenie wirtualnej granicy), ale wiele może być ważnych w kontekście bezpieczeństwa publicznego. Można tu wymienić bieg i analizę aktywności tłumu (masowa ucieczka, zamieszki), upadek, przeludnienie (rozumiane jako potencjalnie nielegalne zgromadzenie ludzi w jednym miejscu), włóczęgostwo oraz naruszenie wirtualnej strefy. Potencjalne najgroźniejsze zjawiska które mogą być wykrywane to pożar, kradzież oraz stosowanie przemocy. Jest to szeroki katalog wykrywanych zdarzeń, odpowiadający współczesnym potrzebom publicznym.

Można uznać, że tytuł rozprawy został sformułowany prawidłowo i odpowiada jej celom i zawartości. Cel, tematyka, zadania i zakres zostały dobrane prawidłowo i spełniają kryteria stawiane pracom doktorskim.

4. Ocena merytoryczna rozprawy

Tematyka rozprawy przyjęta przez Doktoranta jest aktualna w świetle współczesnych trendów rozwoju i stanu wiedzy, dotyczących systemów wizyjnych oraz algorytmów analizy obrazów. Biorąc pod uwagę złożoność omawianych zagadnień oraz potencjalne efekty wdrożenia proponowanych rozwiązań, uważam, że Doktorant prawidłowo przeanalizował czynniki istotne ze względu na zakres pracy, opierając się na współczesnym stanie wiedzy w tej dziedzinie. Analiza literatury, choć pomija niektóre rozwiązania komercyjne, jest przeprowadzona bardzo systematycznie i dogłębnie. Opracowywane własne częściowe rozwiązania są każdorazowo porównywane z tymi, które opisano w publikacjach. Wyniki działania poszczególnych algorytmów są weryfikowane według ujednoliconych kryteriów jakościowych. Dużego wkładu pracy wymagało również przygotowanie materiałów filmowych do analizy przez system, odtwarzających z wykorzystaniem różnych pojazdów i grupy ochotników zarówno sytuacje normalne, jak i zdarzenia które powinny być wykrywane przez IVSS. Doktorant szeroko stosował algorytmy należące do klasy sztucznej inteligencji, dopasowując je i dostrajając do potrzeb opracowywanego systemu. Zaprezentowane przez Doktoranta w rozprawie programy, metody i algorytmy cechują się wysoką jakością działania, mogą jednak być nadal dostosowywane, usprawniane i rozbudowywane, mając już na tym etapie prac potencjał na komercjalizację i szerokie zastosowanie, co zwiększa znaczenie i podwyższa ocenę pracy. Zaproponowane przez Doktoranta kierunki dalszych badań wskazują, że dostrzega on zarówno ograniczenia zastosowanego podejścia, jak również ma pomysły na dalsze usprawnianie proponowanych w rozprawie metod.

4.1. Elementy oryginalne

Do najważniejszych oryginalnych elementów rozprawy zaliczam:

1. opracowanie uniwersalnego algorytmu kalibracji sceny, umożliwiającego poprawne działanie kolejnych zaawansowanych narzędzi analizy obrazu, poprzez korektę geometrii i wzbogacenie płaskiego obrazu z kamery w informacje o głębi lub wysokości obiektów w tylnej części sceny,
2. opracowanie algorytmu automatycznej korekty jasności obrazu,
3. przygotowanie złożonego, modułowego, wielowątkowego algorytmu, umożliwiającego automatyczne wykrywanie zdarzeń alarmowych dzięki syntezie różnych typów informacji wizyjnej; wykorzystywane są tutaj modele szkieletowe wraz z estymacją pozy, metody ekstrakcji i rozpoznawania tekstu, ekstrakcja tła, detekcja krawędzi oraz rzutowanie 3D pozwalające między innymi skompensować fakt, że obiekty dalsze mają mniejszą wysokość zbliżając się do kamery lub odróżnić osobę stojącą od osoby leżącej wzdłuż osi optycznej kamery,
4. opracowanie bazy danych nagrań testowych o łącznym czasie około 4,5 godziny, stanowiących dane uczące lub weryfikujące algorytmy wykrywania różnych rodzajów zdarzeń,
5. opracowanie lub rozwój algorytmów identyfikacji pojazdów, detekcji nielegalnego parkowania, wykrywania obecności zwierząt takich jak pies lub kot, a także detekcji sabotażu kamery,

6. rozwinięcie algorytmu detekcji wejścia/wyjścia z wyznaczonej strefy, pozwalające na precyzyjną ocenę kierunku przekroczenia wirtualnej linii oraz rozwinięcie algorytmu detekcji naruszenia wirtualnej granicy z uwzględnieniem danych historycznych dotyczących ścieżek ruchu osób,
7. rozwinięcie algorytmów detekcji zgromadzeń, biegu i analizy aktywności tłumu oraz wykrywanie włóczęgostwa w oparciu o rzeczywiste parametry przestrzenne sceny 3D,
8. rozwinięcie algorytmu detekcji przemocy, umożliwiającego dzięki temu analizę szybkości zmian kątów orientacji kończyn modelu szkieletowego oraz wykorzystanie rzeczywistych parametrów przestrzennych sceny 3D,
9. opracowanie algorytmu detekcji pożaru z użyciem specjalnie wytrenowanego modelu,
10. rozwinięcie algorytmu detekcji upadku osoby w celu identyfikacji tego zdarzenia, niezależnie od orientacji osoby względem pola widzenia kamery, kierunku oraz szybkości przemieszczania się, jak też interakcji z innymi osobami i obiektami,
11. opracowanie algorytmu detekcji kradzieży obiektów statycznych, z uwzględnieniem możliwości chwilowego maskowania obserwowanych obiektów, które nie powinno wywoływać alarmów.

4.2. Uwagi dyskusyjne

1. Tabela 3.2 (strona 59) wskazuje, że nie ma żadnych publikacji referencyjnych opisujących detekcję sabotażu kamery lub obecności zwierząt oraz identyfikację pojazdów – świadczy to o niesłusznym pominięciu przeglądu rozwiązań komercyjnych w analizie literatury, ponieważ niewątpliwie są w użyciu systemy posiadające przynajmniej niektóre z tych funkcjonalności (zwłaszcza w odniesieniu do identyfikacji pojazdów).
2. Sądząc po zdjęciach zamieszczonych w pracy, wydaje się że testy prowadzono głównie przy bardzo dobrej lub dobrej pogodzie i warunkach oświetleniowych (gorsze oświetlenie testowano głównie w pomieszczeniach). Nie znając całego materiału filmowego trudno to ocenić, ale wydaje się, że uwzględnienie w analizie materiałów wideo zebranych w różnych warunkach oświetleniowych pozwoliłoby dokładniej przetestować opracowane rozwiązania.
3. Co to jest „przeludnienie” – jeśli chodzi o zgromadzenie grupy ludzi w obserwowanej scenie (lub jej wybranym obszarze – nie sprecyzowano tego bezpośrednio), to takie pojęcie wydaje się być odrobinę nieprecyzyjne lub mylące.
4. Jak rozumieć pojęcie „liczby ujęć” występujące w wielu miejscach pracy – czy chodzi o łączną liczbę zarejestrowanych klatek wideo, czy tylko te obrazy, na których występuje obserwowane zjawisko? Być może warto było wprowadzić też jakiś wskaźnik opisujący w jakiej części czasowej danego nagrania jest widoczne poszukiwane zjawisko (zakładając, że nie jest ono obecne w 100% czasu danego nagrania).

5. Czy sprawdzono działanie algorytmu naruszenia granicy przy różnym jej przebiegu, zwłaszcza równoległym do długiego boku obrazu? Wydaje się, że skuteczność algorytmu może zależeć od tego parametru.
6. Czy algorytm przekraczania granic stref prawidłowo odróżnia osobę przechodzącą przez płot od osoby która oddaliła się od kamery, a następnie ponownie przybliżyła? Tutaj ponownie występuje wiele konfiguracji zależnych od kąta ustawienia granicy względem osi obrazu.
7. Z rys. 6.31 można wysnuć wniosek, że testy algorytmu wykrywania ognia/pożaru były prowadzone wyłącznie na materiale filmowym który nie został zebrany przez Doktoranta – wydaje się, że zasymulowanie tych zdarzeń za pomocą np. rozpalonego grilla byłoby możliwe i wystarczająco bezpieczne, a pozwoliłoby lepiej przetestować opracowany model.
8. Czy osoba popełniająca „włóczęgostwo” różni się w sposób możliwy do wykrycia od zwykłego spacerowicza lub osoby odpoczywającej (np. siedzącej na ławce w parku)?
9. W podsumowaniach podrozdziałów opisujących algorytmy detekcji różnych rozpoznawanych zdarzeń Doktorant każdorazowo deklarował osiągnięcie zadowalającej skuteczności działania algorytmu, podczas gdy na rys. 6.43 widać wyraźną różnicę w wartości F1 niektórych algorytmów (obecność zwierząt i przemoc) – czy ta ocena nie jest nadmiernie optymistyczna, zwłaszcza, że testy były prowadzone głównie w dobrych warunkach oświetleniowych?

4.3. Uwagi szczegółowe

Praca jest napisana z dużą starannością i poprawnie językowo. Znalazłem w niej jedynie nieliczne błędy gramatyczne, interpunkcyjne (nadmiarowe przecinki) lub literówki. W przyszlých tekstach Doktorant powinien rozwijać skróty przy pierwszym użyciu (zwłaszcza anglojęzyczne, wraz z rozwinięciem po angielsku), nawet jeśli są one wymienione w wykazie skrótów i oznaczeń – to poprawiłoby czytelność pracy (która i tak dzięki prawidłowej organizacji i systematycznym podsumowaniom jest na wysokim poziomie). Drobną uwagę szczegółową - na stronie 9 są wymienione 4, a nie 3 główne rodzaje zniekształceń obrazu.

5. Wniosek końcowy

Uważam, że podjęty i rozwiązany przez Doktoranta problem badawczy ma istotne znaczenie dla rozwoju właściwej dziedziny wiedzy oraz wpisuje się w aktualne trendy rozwoju nauki i przemysłu. Doktorant wykazał się ogólną wiedzą teoretyczną w dyscyplinie *Automatyka, elektronika, elektrotechnika i technologie kosmiczne* oraz umiejętnością zaplanowania i samodzielnego prowadzenia badań naukowych. Przedstawione w niniejszej recenzji rozprawy uwagi dyskusyjne i szczegółowe nie umniejszają w żaden sposób wartości merytorycznej rozprawy.

Biorąc pod uwagę powyższe stwierdzam, że opiniowana rozprawa doktorska mgra inż. Adama Surówki pt. *Synteza zaawansowanych technik wizyjnych w inteligentnym systemie nadzoru obiektów użytku publicznego* spełnia wymagania stawiane rozprawom doktorskim przez Art. 13 p.1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym

oraz stopniach i tytule w zakresie sztuki (Dziennik Ustaw nr 65, poz. 595) wraz ze zmianami (Ustawa z dnia 21 kwietnia 2017 r. o zmianie ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki oraz niektórych innych ustaw; Dziennik Ustaw, poz. 859) i wnoszę o dopuszczenie jej do publicznej obrony przed Radą Naukową (lub odpowiednią komisją) dyscypliny *Automatyka, elektronika, elektrotechnika i technologie kosmiczne* Wydziału Inżynierii Elektrycznej i Komputerowej Politechniki Krakowskiej.



Dr hab. inż. Grzegorz Ćwikła, Profesor PŚ