

Dr hab. inż. Paweł Badura  
Politechnika Śląska  
Wydział Inżynierii Biomedycznej  
Katedra Informatyki i Aparatury Medycznej  
ul. Roosevelta 40, 41-800 Zabrze

Zabrze, 3 września 2018 roku

---

## RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy:

**Keypoint matching for object detection in 2.5D images**

Autor rozprawy: **mgr inż. Karol MATUSIAK**

Promotor rozprawy: **prof. dr hab. inż. Paweł STRUMIŁŁO**

Promotor pomocniczy: **dr inż. Piotr SKULIMOWSKI**

Dziedzina: **nauki techniczne**

Dyscyplina: **informatyka**

---

# 1 Temat, zakres i cel rozprawy

Przedłożona rozprawa doktorska porusza temat detekcji i dopasowania wzorców obiektów w obrazach planarnych scen trójwymiarowych, dla których określono również mapę głębi. Zakres tematyczny jest aktualny w naukach informatycznych, a konkretnie w skrzydle obejmującym analizę obrazów. Autor opracował własną metodę detekcji punktów charakterystycznych na podstawie posiadanych danych, które traktuje się jako przestrzenne dane 2,5-wymiarowe (2.5D). Dla punktów o określonym charakterze Doktorant zaproponował autorską technikę deskrypcji, a następnie algorytm dopasowania wyznaczonych punktów. Metodę obejmującą powyższe składowe nazwano Depth-based feature descriptor (DBFD). Dodatkowym wkładem Autora jest metoda oceny jakości detekcji punktów charakterystycznych wykorzystująca operacje obrotu obrazu z transformacją punktów, wraz z wprowadzeniem wskaźnika powtarzalności detekcji. Opracowana metodyka została zwalidowana w kilku eksperymentach, uwzględniających porównanie z istniejącymi narzędziami, a uzyskane wyniki poddano dyskusji.

Autor sformułował 2 tezy rozprawy:

- T1. Deskrypcja punktów charakterystycznych, uwzględniająca dane o lokalizacji w odniesieniu do krawędzi w mapie głębi sceny, zwiększa efektywność dopasowania.
- T2. Eliminacja punktów charakterystycznych zlokalizowanych w rejonie krawędzi obiektów zwiększa powtarzalność detekcji.

Rozprawa napisana w języku angielskim zawiera 73 strony i jest podzielona na 6 rozdziałów, poprzedzonych spisem treści i wykazem skrótów. Bibliografia pracy obejmuje 64 pozycje, w tym 5 artykułów Autora; bibliografię dobrano poprawnie, sięgając zarówno do źródeł kanonicznych, jak i aktualnych prac w rozpatrywanej dziedzinie.

W rozdziale 1. przedstawione są motywacje Autora oraz wprowadzenie w dziedzinę z wyjaśnieniem podstawowych pojęć, mechanizmów i uwarunkowań sprzętowych i fizycznych. Doktorant formułuje tezy rozprawy i streszcza jej zawartość.

Rozdziały 2. i 3. są w założeniu przeglądowe i obejmują zestawienie i opis algorytmów detekcji, deskrypcji i dopasowania punktów startowych (rozdział 2.) oraz ewaluacji powyższych (rozdział 3.) W rozdziale 2. poprawnie podzielono i pokrótce scharakteryzowano detektory punktów charakterystycznych w obrazach 2D, wymieniając najważniejsze spośród nich, w tym gradientowy SIFT (scale-invariant feature transform), szczególnie istotny z punktu widzenia opisu metody Doktoranta. Opisano również detektory punktów w obrazach 2.5D, w tym bezpośrednio wykorzystujące mapę głębi. W podsumowaniu rozdziału Doktorant motywuje podjęcie próby zaprojektowania detektora w obrazach 2.5D i czyni to w sposób przekonujący. W rozdziale 3. Autor przedstawia miary jakości detekcji i dopasowania bazujące na macierzy błędów oraz przykładowe znane z literatury rozwiązania problemu wiarygodnej ewaluacji algorytmów. W rozdziale 3.3 Doktorant specyfikuje własną metodę oceny detekcji punktów oraz przedstawia i poddaje dyskusji wyniki jej zastosowania w przypadku 8 metod detekcji.

Rozdział 4. stanowi najistotniejszą część rozprawy, wprowadzając technikę detekcji, deskrypcji i dopasowania punktów oraz protokół jej ewaluacji w każdym z etapów. Algorytm detekcji bazuje na detektorze SIFT. Punkty charakterystyczne wyznaczone w efekcie zastosowania SIFT podlegają weryfikacji pod kątem odległości od krawędzi obiektów szacowanych na podstawie mapy głębi. Usuwane są punkty, które w ujęciu odległości euklidesowej są zlokalizowane zbyt blisko krawędzi określonych w mapie głębi przez detektor Canny. Deskryptor punktów charakterystycznych wyznaczony przez SIFT jest uzupełniony przez Doktoranta o miary odległości do najbliższych krawędzi w mapie głębi w czterech podstawowych kierunkach znormalizowanych względem orientacji SIFT punktu

oraz wartość głębi. W procedurze dopasowania po kroku wykorzystującym klasyczny mechanizm SIFT następuje weryfikacja autorskiego warunku, porównującego w każdym z czterech kierunków odległości do krawędzi mapy głębi z uwzględnieniem stosunku wartości głębi punktów w obu rozpatrywanych obrazach. Następnie Autor opisuje stanowisko akwizycyjne, procedurę kalibracji oraz protokół przeprowadzonych eksperymentów testowych.

W rozdziale 5.: (1) za pomocą metody zaproponowanej przez Autora w rozdziale 3.3 oceniono detekcję punktów uzyskanych za pomocą DBFD i SIFT; (2) porównano dwie ww. metody pod kątem dopasowania punktów w ujęciu precyzji i stopy detekcji fałszywie pozytywnych; (3) przeprowadzono porównanie z innymi metodami (w tym SIFT i PIFT) zgodnie z protokołem i na zbiorze testowym wprowadzonym przez autorów techniki PIFT. W ostatnim kroku tego rozdziału Autor przedstawia wyniki czułości i precyzji wariantu swojej metody, które są wyraźnie lepsze od wskaźników pozostałych metod.

Rozdział 6. podsumowuje pracę, m.in. wskazując wyniki uzasadniające przyjęcie prawdziwości obu postawionych tez rozprawy.

## 2 Ogólna ocena rozprawy

Przedłożona rozprawa i algorytm DBFD zawierają szereg oryginalnych osiągnięć Doktoranta, mianowicie:

1. Opracowanie efektywnej metody detekcji punktów charakterystycznych w obrazach 2.5D. Autor zaproponował wykorzystanie krawędzi mapy głębi sceny do usunięcia leżących zbyt blisko punktów charakterystycznych zlokalizowanych przez detektor SIFT.
2. Sformułowanie protokołu deskrypcji oraz dopasowania punktów charakterystycznych. Oryginalny wkład w deskrypcję stanowi uwzględnienie wartości głębi punktu oraz odległości do najbliższej krawędzi mapy głębi sceny w czterech podstawowych kierunkach znormalizowanych względem orientacji SIFT punktu. Nowe elementy deskryptora są wykorzystywane w procesie dopasowania jako decydujące o dopasowaniu poprzez porównanie odległości do krawędzi mapy głębi skorygowanej o stosunek głębi punktów w dwu analizowanych scenach.
3. Opracowanie metodyki zobiektywizowanej oceny jakości detekcji punktów charakterystycznych wraz ze statystycznym wskaźnikiem powtarzalności detekcji.

Porównanie opracowanego algorytmu DBFD z istniejącymi narzędziami wypada zdecydowanie korzystnie: powtarzalność detekcji punktów i precyzja ich dopasowania są lepsze od wykazywanych przez algorytm SIFT o dwucyfrowe wskaźniki procentowe, podobnie jest z precyzją dopasowania w odniesieniu do algorytmu PIFT. Na uwagę zasługuje praktyczna eliminacja detekcji fałszywie pozytywnych. Drobna wątpliwość może budzić jedynie kwestia liczebności zbiorów testowych. W pierwszej części (7 wzorców, 61+49 obrazów testowych) ilość danych można uznać za wystarczającą, lecz nie znalazłem informacji o liczbie obrazów wykorzystanych w analizie porównawczej DBFD z PIFT (rozdział 5.3).

W podsumowaniu tej części recenzji oraz w obliczu wymienionych dalej uwag pragnę napisać, że wkład Doktoranta w dziedzinę oraz oryginalność pomysłów i implementacji oceniam pozytywnie, uzyskane wyniki porównania z uznanymi technikami potwierdzają to odczucie, zaś moje uwagi koncentrują się głównie na charakterze i jakości przekazu, choć i ten w ogólnym ujęciu jest poprawny.

## 2.1 Uwagi

Praca nie jest zbyt obszerna, jednak zawiera niezbędne informacje o charakterze wprowadzającym, przeglądowym, oryginalnym, ewaluacyjnym, dyskusję oraz podsumowanie. Pozwolę sobie na uwagę, że czytając rozprawę odnosiłem wrażenie, jakbym czytał mocno rozbudowany artykuł naukowy, co uznaję za pozytywny efekt. Uważam wszakże, że charakter wprowadzanych w rozprawie (rozdziały 2-4) zagadnień prosi się o bogate ilustrowanie przykładami obrazowymi, co w rozdziale 4. wykonano poprawnie, w rozdziale 3. — w zakresie minimalnym, jednak w rozdziale 2. dość skąpo. Mam również zastrzeżenie odnośnie umiejscowienia w pracy rozdziału 3.3, traktującego o autorskiej metodzie oceny detekcji punktów charakterystycznych. Rozdział 3. ma jeszcze charakter przeglądu zagadnienia ewaluacji rozpatrywanych narzędzi i pojawienie się tam opisu autorskiej metodyki, a na dodatek wyników eksperymentalnych, jest cokolwiek nieoczekiwane; spodziewałbym się raczej rozłożenia tego materiału między rozdziały 4.2 i 5.

W nawiązaniu do ogólnej oceny rozprawy proszę Doktoranta o ustosunkowanie się do wymienionych poniżej uwag merytorycznych.

1. Operator Canny daje dość zaskakujący rezultat przedstawiony na Rys. 4.3c. Co jest tam krawędzią mapy głębi? Co reprezentują obszary zaznaczone na szaro? To, że punkty charakterystyczne wyznaczone przez algorytm SIFT w obszarach szarych (Rys. 4.3a) zostają usunięte (Rys. 4.3d) sugeruje, że szare obszary algorytm traktuje jak krawędzie obiektów w mapie głębi. Nie jest to dla mnie oczekiwane zachowanie algorytmu. Jeśli w opisie ominięto jakiś krok eliminacji punktów na podstawie mapy głębi, to proszę o jego skomentowanie podczas obrony rozprawy.
2. Wątpliwości wzbudzają również wizualizacje wybranych umieszczonych w rozprawie map głębi. Np. mapy przedstawione na Rys. 4.8b,d wyglądają na trójwartościowe, podczas gdy kształt czajnika elektrycznego czy blatu z Rys. 4.8d sugeruje bardziej zniuansowaną i nieco inną postać mapy. Na ile jest to wina (a) trybu wizualizacji, (b) specyfikacji urządzenia rejestrującego, czy (c) innych czynników? Czy wyraźna krawędź prowadząca wzdłuż zupełnie nieoczekiwanej trajektorii przez korpus czajnika nie wpływa niekorzystnie na detekcję punktów charakterystycznych i ich dopasowanie?
3. W dyskusji wyników ostatniego eksperymentu przedstawionych na Rys. 5.13 stanowczo brakuje specyfikacji parametrów wariantu DBFD v2 (a może raczej różnicy ich wartości w odniesieniu do wariantu podstawowego DBFD). Czytelnik rozprawy z pewnością jest zainteresowany tym, jakie środki prowadzą do uzyskania metody dopasowania deklasującej pozostałe w ujęciu obu obserwowanych miar jakości. Przy okazji — ile obrazów wykorzystano do porównania metod w tym eksperymencie?
4. W dyskusji wyników w rozdziale 5.2 Autor formułuje hipotezę, mówiącą że poziom niekorzystnego efektu nadmiaru dopasowań w DBFD może być zmniejszony za pomocą wprowadzenia mechanizmu adaptacyjnego doboru wartości progu  $\epsilon$  odległości do krawędzi mapy głębi. Czy Doktorant prowadził badania w tym kierunku? Czy rozważał użycie dodatkowej zależności  $\epsilon = f(Z_s, Z_t)$  od głębi dopasowywanych punktów w metrach? Bardzo proszę o bardziej szczegółowy komentarz na ten temat podczas obrony rozprawy.

W pracy stwierdzam również szereg usterek redakcyjnych oraz niestaranności w zapisie matematycznym, które nie wymagają komentarza Doktoranta, jednak jako utrudniające percepcję powinny zostać wskazane w celu unikania podobnych mechanizmów w przyszłych opracowaniach Autora.

1. W kilku miejscach rozprawy rysunki są uporządkowane i numerowane w sposób niezgodny z odwołaniami w treści i porządkiem narracji. Np. omówienie i odwołanie do Rys. 3.4 znajduje się na str. 30, a sam rysunek na str. 33, zaś pomiędzy tymi stronami (str. 32) umieszczono Rys. 3.3 wraz z odwołaniem. Podobne sytuacje można odnotować w przypadkach (podano w kolejności odwołań):

- Rys. 4.3, 4.1, 4.2,
- Rys. 4.6, 4.4, 4.5,
- Rys. 4.12, 4.11,
- Rys. 5.8, 5.6, 5.7.

2. Mam zastrzeżenia odnośnie poprawności, spójności i kompletności zapisu matematycznego w rozprawie. Dotyczy to niestaranności (we wzorze (2.6) brakuje oznaczenia  $I$  przed pierwszym nawiasem zarówno w liczniku, jak i w mianowniku), doboru oznaczeń i spójności w ich stosowaniu ( $x = (x, y)$  w pierwszym wierszu na str. 17, oznaczenie punktów obrazu małymi literami w rozdziale 2., i wielkimi w rozdziale 4.), zaniedbań w formułowaniu iteratorów/zakresów operatora sumy (wzory (2.8), (2.10)). Na Rys. 4.6 pojawiają się oznaczenia niewystępujące w treści (indeksy dolne). Pewnym problemem jest wykorzystanie symbolu  $\epsilon$  na oznaczenie dwóch różnych parametrów algorytmu, co w szczególności w części poświęconej wynikom i dyskusji wprowadza niepotrzebny zamęt. Uwaga,  $\epsilon$  wprowadzone w zależności (4.12) jako progowa wartość odległości euklidesowej wyrażonej liczbą pikseli nie powinna być podawana w procentach w dyskusji wyników na przełomie strony 50. i 51., pomimo że z wcześniejszych ustępów można wywnioskować, że chodzi o odsetek odległości  $d_t$ .

3. Autor przeważnie odwołuje się do równań nie biorąc ich numerów w nawiasy ( $\cdot$ ), co jest uznaną praktyką. Z kolei na str. 36 Autor z nieznanym przyczyn ujął w nawias łączne odwołanie do Tabel 3.1, 3.2, 3.3.

4. W znacznej części wypunktowań Autor nie stosuje niezbędnych znaków interpunkcyjnych na końcach pozycji (np. na str. 15, 17/18, 31, 32, 38, 39, 52).

Poziom znajomości i umiejętności posługiwania się językiem angielskim Doktoranta zasługuje na uznanie i generalnie nie nasuwa uwag krytycznych. Nieliczne literówki (co ciekawe najczęstsze w polskojęzycznym streszczeniu) nie są istotne, wyjąwszy zmieniające przekaz pomyłki wymienione poniżej:

1. Str. 32: czytelnik nie wie po przeczytaniu całego przedostatniego akapitu, ile zbiorów obrazów testowych wykorzystano do ewaluacji detekcji punktów startowych: 22 czy 20.

2. Str. 50: w przedostatnim wierszu odwołano się do warunku (4.10), podczas gdy chodzi o (4.12).

3. Str. 69: pozycja 3 wypunktowania stanowi raczej dowód na prawdziwość tezy nr 1, a nie 2.

### 3 Podsumowanie i konkluzja

Przedstawiona do recenzji rozprawa zawiera sformułowany i rozwiązany problem badawczy oraz stanowi realny wkład w dziedzinę rozpoznawania wzorców poprzez detekcję i dopasowanie punktów charakterystycznych w obrazach 2.5D z określoną mapą głębi. Opisana w pracy autorska metodologia badań umożliwiła udowodnienie tezy pracy oraz realizację postawionego celu — opracowanie

efektywnej i co najmniej porównywalnej z kanonicznymi metody detekcji i dopasowania punktów charakterystycznych. W procesie dochodzenia do przedstawionego stanu badań Autor przedstawiał jej kroki pośrednie w środowisku naukowym, publikując szereg artykułów w ramach konferencji naukowych, rozdziałów w monografiach oraz w czasopiśmie indeksowanych na liście JCR. Sformułowany problem badawczy, jego realizacja oraz kompetencje Autora pozwalają na stwierdzenie, że rozprawa spełnia warunki określone w art. 13 Ustawy z dn. 14 marca 2003 o stopniach naukowych i tytule naukowym oraz o stopniach w zakresie sztuki (Dz.U. 2003 nr 65, poz 595 z późn. zm., tekst jednolity: Dz.U. 2017, poz. 1789) skłaniają do postawienia wniosku o dopuszczenie mgr. inż. Karola Matusiaka do dalszych etapów przewodu doktorskiego.

A handwritten signature in black ink, appearing to be 'J. Paweł', written in a cursive style.