

Prof. dr hab. inż. Zbigniew Banaszak
Politechnika Koszalińska
Śniadeckich 2
75-453 Koszalin

Koszalin, 28.01.2015 r.



RAU	Biuro Dziekana	
	Wpłynęło dnia	02.02.2015
	Nr	200 / zał.

Recenzja

rozprawy doktorskiej mgr inż. Tomasza Grzejszczaka pt.: „**Wykrywanie i lokalizacja punktów charakterystycznych dłoni w obrazach cyfrowych**”.

1. Obszar problemowy rozprawy

Tematyka rozprawy dotyczy metod automatycznego przetwarzania obrazów cyfrowych gestów dłoni, metod wykorzystywanych m.in. do rozpoznawania znaków języka migowego sygnalizowanych układami dłoni. Tematyka ta aczkolwiek znana od wielu lat jest ciągle bardzo aktualna, gdyż wśród wielu istniejących metod rozpoznawania kształtów żadna nie okazuje się w pełni skuteczna we wszystkich możliwych zastosowaniach.

W przedstawionym kontekście, opiniowana rozprawa koncentruje się na zagadnieniach lokalizacji, ekstrakcji i pomiaru tych cech obrazu (wektora cech opisujący dłoń) które decydują o efektywności implementujących je metod rozpoznawania. Przedstawione w niej podejście można zrekonstruować w następujący sposób: Dany jest obraz dłoni oraz zbiór występujących na nim arbitralnie wybranych cech morfologicznych. Poszukiwana jest metoda wykorzystująca elementy rozwiązań już istniejących algorytmów rozpoznawania umożliwiającą dostrajanie swoich parametrów do specyfiki obrazów cyfrowych rozważanych obiektów. Poszukiwane rozwiązanie winno stanowić atrakcyjną, w sensie jakości wyników i czasochłonności związanych z nimi obliczeń, alternatywę dla istniejących, implementowanych w metodach rozpoznawania wykorzystujących parametryczne modele obrazu dłoni. Reasumując, opiniowana rozprawa koncentrując się na zagadnieniach ekstrakcji punktów charakterystycznych dłoni stawia sobie za cel opracowanie „algorytmu hybrydowego łączącego klasyczne metody przetwarzania obrazów z procesami decyzyjnymi, opartymi na drzewach decyzyjnych, zawierającymi oryginalne reguły wnioskowania”. Zaproponowane podejście posiada cechy nowości w stosunku do metod znanych z literatury, wiążą się one m.in. z: wykrywaniem i lokalizacją nadgarstka,

rozpoznawaniem zagiętych palców dłoni, dostrajaniem parametrów algorytmu wykrywania i lokalizacji punktów charakterystycznych, itp.

Tezę rozprawy sformułowano następująco: „Zaproponowane w pracy metody sztucznej inteligencji polepszają dokładność lokalizacji i detekcji wybranych punktów charakterystycznych dłoni.” Sformułowanie to jest niestety mało precyzyjne, a w świetle wcześniej przedstawionej rekonstrukcji problemu badawczego, nawet mylące. Lepiej oddającym istotę rozprawy byłoby następujące jej sformułowanie: *Lepsza dokładność, w zakresie detekcji i lokalizacji wybranych punktów charakterystycznych dłoni, jest możliwa do uzyskania w sposób automatyczny z użyciem algorytmu hybrydowego łączącego wybrane, heurystyczne metody przetwarzania obrazów z procesami decyzyjnymi.* W mojej ocenie, autorską wersję tezy rozprawy można uznać za wykazaną, ale nie w tak szerokim zakresie, jaki wynikałby z jej sformułowania.

Tematyka rozprawy, noszącej metodologiczno-eksperymentalny charakter, leży na styku co najmniej trzech dyscyplin automatyki i robotyki, informatyki i elektroniki. Uważam, że jej podjęcie jest uzasadnione zarówno ze względów poznawczych, jak i możliwości wielu praktycznych zastosowań związanych z systemami komunikacji człowiek-maszyna, diagnostyki obrazowej, itp.

2. Kompozycja i treść rozprawy

Opiniowana praca liczy 112 stron i składa się z 6 rozdziałów, spisu treści, streszczenia, spisów tabel i ilustracji, a także spisów skrótów oraz symboli i oznaczeń. W załączonym, liczącym 94 pozycje, wykazie cytowanej literatury występuje 5 współautorskich i dwie samodzielne (jedna z nich opublikowana w 2011 w Przeglądzie Elektrotechnicznym, wówczas jeszcze występującym na tzw. liście filadelfijskiej) publikacje Doktoranta. W tym miejscu warto zauważyć, że cytowany w rozprawie dorobek Doktoranta stanowi wybór spośród jego wszystkich 11 anglojęzycznych publikacji. Załączona lista, obejmuje ważniejsze publikacje z zakresu przedmiotu pracy, warto jednak odnotować brak istotnych pozycji z obszaru przedstawianych w rozprawie zagadnień, a pochodzących z innych krajowych ośrodków akademickich, jak np.:

Mamík J., Rozpoznawanie znaków polskiego alfabetu palcowego z wykorzystaniem morfologii matematycznej i sieci neuronowych, Rozprawa doktorska, AGH, Kraków, Wydział EAliE, Kraków, 2002.

Kapuściński T., *Rozpoznawanie polskiego języka migowego w systemie wizyjnym*, Rozprawa doktorska, Uniwersytet Zielonogórski Wydział Elektrotechniki, Informatyki i Telekomunikacji, Zielona Góra, 2006.

Myśliński Sz. *Rozpoznawanie obrazów dłoni za pomocą gramatyk klasy ETPL(k) w systemach wizyjnych analizy języka migowego*. Rozprawa doktorska, AGH, Kraków, Wydział EAIiE, Kraków, 2009.

czy też prac przeglądowych:

Rafiqul Zaman Khan, Noor Adnan Ibraheem, *Hand Gesture Recognition: A Literature Review*, *International Journal of Artificial Intelligence & Applications (IJAIA)*, Vol.3, No.4, July 2012, 161-174, DOI : 10.5121/ijaia.2012.3412

Noor Adnan Ibraheem, Aligarh Uttar Pradesh, Rafiqul Zaman, *Survey on Various Gesture Recognition Technologies and Techniques*, *International Journal of Computer Applications*, Vol. 50, No.7, 2012, 38-44.

Z merytorycznego punktu widzenia, w rozprawie wyodrębnić można trzy zasadnicze części. W części pierwszej, posiadającej wprowadzająco-systematyzujący charakter (obejmującej rozdziały 1-3) przedstawiono wybrane zagadnienia z zakresu rozpoznawania gestów, wykrywania punktów charakterystycznych dłoni oraz wybranych metod sztucznej inteligencji. I tak w rozdziale 1, bardzo powierzchownie uzasadniono potrzebę podjęcia tematu, w szczególności, w zakresie możliwości bezdotykowego sterowania urządzeniami technicznymi, a także oczekiwań związanych z rozwojem technik uczenia maszynowego wykorzystywanych w rozwiązaniach samouczących się algorytmów rozpoznawania gestów w językach palcowych i/lub migowych. W podrozdziale tego rozdziału: **1.2 Układ pracy**, zupełnie pominięto przedstawienie zakresu rozdziału **3. Wybrane metody sztucznej inteligencji i przetwarzania obrazów**.

Rozdział **2. Analiza stanu wiedzy** wyróżnia trzy podrozdziały: **2.1 Przegląd literatury** - ograniczający się do systemów sensorycznych wizyjnych i bez wizyjnych, 12-to wierszowy podrozdział **2.2 Założenia dotyczące zakresu badań** – pomijający omówienie domyślnie przyjętych założeń ograniczających prowadzone badania do gestów prawej ręki, nie noszącej zmian morfologicznych, bez ozdób, postrzeganej w jednej prostopadłej dla obserwatora płaszczyźnie, a także podrozdziału **2.3 Przegląd istniejących metod wykrywania punktów charakterystycznych**. Uważam, że z perspektywy treści całej rozprawy, rozdział **Przegląd literatury** ograniczający się do systemów sensorycznych mógłby zostać poświęcony, zgodnie ze swoim tytułem, przeglądowi metod modelowania gestów dłoni i związanymi z nimi metod rozpoznawania w czasie rzeczywistym.

Na Rozdział 3 składają się dwie części – część pierwsza obejmuje podrozdział **3.1 Metody sztucznej inteligencji** odwołujący się, na jednej stronie tekstu, do kwestii wnioskowania w przód i wstecz oraz do 14-wierszowego akapitu o statusie podrozdziału: **3.1.2 Drzewo decyzyjne**. Na część drugą składa się podrozdział **3.2 Metody analizy obrazu** koncentrujący się na wybranych zagadnieniach segmentacji obrazu dłoni, transformacji odległościowej, poszukiwania wzorca oraz analizy tekstury maski związanej z wyznaczaniem tzw. obrazu kierunkowego.

Lekturę tej części rozprawy utrudniają liczne błędy redakcyjne począwszy od niezręczności stylistycznych, np. 26₁₇ „*Po pierwsze, nowo wykryty punkt nie może znajdować się zbyt blisko innych, po drugie, punkt czubka palca nie może znajdować się wewnątrz dłoni lub wewnątrz palca.*”, na stronie s. 16₇ stosowany jest termin „*mapa głębi*” z kolei na stronie 27¹⁰ „*mapa odległościowa*”, poprzez błędy typograficzne, np. s.27¹⁰ „*...barw YCbCr.*”, zależność (3) $N_{W1} = N_{W1} + 1$, po nieprecyzyjny opis omawianych kwestii, np. s. 24 oprócz punktu typu A, typu B występują jeszcze punkty typu AB, brak jest również definicji zmiennej kroczącej (liczonej w stopniach promienia wodzącego lub długości odcinka konturu), a także występujące kolizje oznaczeń, np. symbol D raz oznacza *ciąg* s.21₁, innym razem zaś *odległość* s.30₄.

Podjmując w tej części próbę swoistej systematyzacji tak różnych obszarów, jak np. klasyfikacji systemów sensorycznych, metod sztucznej inteligencji, jak i różnych zagadnień obejmujących metody segmentacji obrazu, samoorganizującej się mapy neuronowej itp., Doktorant nie skorzystał z okazji przedstawienia swoich autorskich uwag komentujących wady i zalety różnych, dostępnych w literaturze przedmiotu, modeli dłoni np. w kontekście efektywności (dokładności) implementujących je metod rozpoznawania obrazów, a w szczególności klas rozpoznawanych obiektów. Szkoda, że przedstawiane w tej części rozprawy rozdziały nie podsumowywał stosownym komentarzami.

W części drugiej (rozdział 4), zawierającej główne, poznawcze wyniki rozprawy, przedstawiono badania składające się na autorski algorytm wykrywania punktów charakterystycznych dłoni. W odróżnieniu od już istniejących, opracowany algorytm analizuje rozpoznawany obraz tak pod kątem cech charakterystycznych konturu jego maski, jak i tekstury jego wnętrza, obrazu kierunkowego. Działanie algorytmu determinuje sekwencja trzech etapów związanych odpowiednio z: wykryciem i lokalizacją nadgarstka, lokalizacją punktów charakterystycznych oraz identyfikacją punktów charakterystycznych dłoni umożliwiającą rozpoznanie związanego z nią gestu.

Lekturę tej części rozprawy utrudniają liczne błędy redakcyjne na które składają się niezręczności stylistyczne, np. 58⁴ „Podsumowując, celem przedstawionej metody jest wykrycie pary punktów T i B za pomocą odnalezionego punktu ekstremum P lub porzucenie tego punktu.”, na stronie s.57⁴ „Dlatego wyznaczono statystycznie średnią długość palca zgiętego, która w przybliżeniu równa jest promieniowi regionu śródreżca.”, poprzez błędy typograficzne, np. s.42 formuła (10) jest błędnie zapisana, po nieprecyzyjny opis omawianych kwestii, np. s.42⁸ sformułowanie: „... C jest kombinacją bez powtórzeń n elementów z k elementowego zbioru...” winno być zapisane jako: C_n^k liczba kombinacji k -elementowych ze zbioru n -elementowego, a także występujące kolizje oznaczeń, np. symbol A w formule (10) oznacza liczebność zbioru punktów konturu, na s.24⁷ zaś typ punktu. Warto też zwrócić uwagę na brak opisu stosowanych symboli i oznaczeń, m.in. w zależnościach: 6 -9, 12, 13 i 19, a także na nieuzasadnione powoływanie się na „reguły wnioskowania” s.53² bez ich wcześniejszego, standardowego sformułowania w postaci schematu: **IF...THEN...**

W części trzeciej, obejmującej rozdziały 5 i 6, stanowiącej próbę jakościowej i ilościowej weryfikacji zaproponowanej metodyki ekstrakcji punktów charakterystycznych dłoni przedstawiono przykłady analizy porównawczej etapów autorskiego algorytmu z wybranymi, zrekonstruowanymi na podstawie doniesień literaturowych metodami wykrywania i lokalizacji, odpowiednio nadgarstka i innych punktów charakterystycznych dłoni. Wyniki badań przedstawiono w postaci wykresów i zestawień tabelarycznych wykorzystując arbitralnie wybrane miary oceny. Ważną rolę w tej części odgrywa podrozdział **5.3.5 Wyznaczanie parametrów sterujących**, poświęcony identyfikacji parametrów opracowanego algorytmu ekstrakcji cech, tzn. dedykowany wyznaczaniu wartości parametrów minimalizujących błędy wykrywania i lokalizacji punktów charakterystycznych. Niedostatkami tej części jest brak odniesień zaproponowanego rozwiązania do jednej wybranej klasy obiektów (np. grupy gestów), w kontekście której można by ocenić praktyczną konkurencyjność algorytmu autorskiego. Mimo tych zastrzeżeń należy uznać, że przedstawione wyniki badań eksperymentalnych są oryginalnym dorobkiem Doktoranta potwierdzającym sformułowaną we **Wstępie** tezę rozprawy. W rozdziale 6 podsumowano uzyskane wyniki oraz przedstawiono propozycje przyszłych badań.

Podobnie jak poprzednio, lekturę tej części rozprawy utrudniają liczne błędy redakcyjne na które składają się niezręczności stylistyczne, np. 74⁵ „Jako, że niektóre algorytmy poszukiwania punktów charakterystycznych bazują na metodach poszukiwania nadgarstka, do ich implementacji użyto punktów odnajdywanych za pomocą przedstawionego algorytmu

bazującego na zliczaniu profilu obrazu odwróconej o kąt najdłuższej cięciwy spośród losowo wybranych par punktów, a następnie wyznaczającego punkt nadgarstka za pomocą metody poszukiwania minimum wzdłuż kierunku z kontrakcją przedziału, oraz z poprawieniem lokalnym wyniku.”, poprzez błędy typograficzne, np. s.74₆ rozdziale 2.3.4 i w =[74]=[74], s.63₉ ...nazwy pików..., po nieprecyzyjny opis omawianych kwestii, np. s.63₁₆ „W skład tej bazy wchodzi kilka gestów prostych do wykrycia, jednak głównie znajdują się w niej gesty skomplikowane oraz bardzo podobne, wybrane z Polskiego i Amerykańskiego Języka Palcowego.”, a także występujące kolizje oznaczeń, np. symbol N na s.70¹⁷ oznacza liczbę przedziałów błędu, a na s.28⁷ wartość błędów nagromadzonych.

Reasumując, należy zwrócić uwagę na rzetelne dokumentacje bardzo licznych analiz i ocen wielu autorskich rozwiązań, a także zwrócić uwagę na brak przykładów ilustrujących możliwości praktycznego wykorzystania proponowanego algorytmu oraz bardzo liczne błędy redakcyjne.

3. Oryginalne osiągnięcia

Zmierzając do osiągnięcia zamierzonych przez siebie celów, Doktorant uzyskał szereg nowych rezultatów. Do ważniejszych z nich, wyróżniających je spośród dostępnych w literaturze przedmiotu, można zaliczyć:

1. Rekonstrukcję i komputerową implementację wybranych, dostępnych w publikacjach z zakresu rozpoznawania gestów, algorytmów wykrywania i lokalizacji nadgarstka i punktów charakterystycznych dłoni.
2. Opracowanie i komputerową implementację autorskiego algorytmu hybrydowego wykrywania i lokalizacji punktów charakterystycznych dłoni, algorytmu wykorzystującego oryginalną procedurę wykrywania i lokalizacji nadgarstka oraz detekcji ułożenia zagiętych palców wewnątrz konturu obrazu maski dłoni.
3. Opracowanie prototypu środowiska wspomagającego eksperymenty komputerowe i wykorzystanie go do przeprowadzenia eksperymentów weryfikacyjnych.
4. Zaplanowanie i przeprowadzenie szeregu eksperymentów i analiz porównawczych, weryfikujących efektywność wybranych, dostępnych w literaturze przedmiotu i zrekonstruowanych algorytmów przetwarzania obrazów, w kontekście opracowanego algorytmu autorskiego.

Opracowany algorytm stanowiący rozwiązanie rozważanego w rozprawie problemu badawczego umożliwia budowę zadaniowo zorientowanych, dedykowanych metod rozpoznawania gestów dłoni. Oznacza to, że zaproponowany mechanizm identyfikacji parame-

trów algorytmów decydujących o dokładności lokalizacji wyznaczanych punktów charakterystycznych pozwala na adaptacyjne dostrajanie się algorytmu do wybranej klasy obiektów.

Uzyskane wyniki potwierdzają wysokie kwalifikacje Doktoranta umożliwiające Mu swobodne poruszanie się zarówno w obszarach zagadnień z zakresu modelowania i rozpoznawania gestów dłoni, metod rozpoznawania obrazów cyfrowych, a także technik programowania i planowania eksperymentów komputerowych. Wymienione fakty potwierdzają, że Doktorant potrafi podejmować i samodzielnie realizować zaplanowane cele badawcze.

4. Uwagi

Do głównych wad rozprawy należy zaliczyć jej liczne błędy redakcyjne, brak krytycznej analizy dostępnych metod rozpoznawania gestów dłoni, a w szczególności brak eksperymentów potwierdzających przydatność opracowanego algorytmu do rozpoznawania określonych, występujących w praktyce klas gestów.

Opracowane rozwiązanie pomija kontekst klasy obiektów dla potrzeb rozpoznawania których ma być ono ostatecznie wykorzystane. Z jednej strony stanowi to o jego uniwersalności, z drugiej zaś wymusza potrzebę każdorazowej kalibracji algorytmu uwzględniającej specyfikę rozpoznawanej klasy obiektów. Warto tutaj podkreślić, to co sam Doktorant uważa, że jego algorytm nie pozwala rozróżnić wszystkich spotkanych w praktyce układów dłoni. Konstatacja ta, zgodna z tezą „*No free lunch theorem*”, potwierdza fakt, braku jednej, uniwersalnej metody umożliwiającej rozpoznawanie wszystkich możliwych gestów z dowolną, wystarczającą dla zastosowań praktycznych, dokładnością.

Lekturę pracy bardzo utrudnia brak pełnej listy stosowanych w jej tekście skrótów, symboli i oznaczeń. Konsekwencje tego braku widoczne są m.in. w licznych kolizjach oznaczeń. Podobna uwaga dotyczy braku stosownego glosariusza terminów technicznych wykorzystywanych w tekście pracy.

Rozprawa sprawia wrażenie pisanej w pośpiechu, bardzo widoczny jest tutaj brak korekty tak autorskiej, jak i redakcyjnej. Brak tej staranności widać m.in. w wielu usterkach typograficznych, niezręcznościach i/lub błędach stylistycznych, przykłady niektórych z nich wymienione zostały już wcześniej.

5. Konkluzja

Mimo dostrzeżonych i wyżej omówionych niedostatków opiniowanej pracy uważam, że przedstawiony w niej dorobek naukowy mgr inż. Tomasza Grzejszczaka uzasadnia **dopuszczenie** do publicznej obrony rozprawy zmierzającej do nadania ww. stopnia doktora

nauk technicznych w dyscyplinie „automatyka i robotyka”, tzn. spełnia warunki stawiane przez obowiązującą w zakresie ustawę o stopniach i tytule naukowym w określeniu do rozpraw doktorskich (Dz. U. Nr. 65 z dnia 14 marca 2003, ze zm. w Dz. U. z 2005 r. Nr 164, oraz w Dz. U. z 2011 r. Nr 84).

