# POLITECHNIKA WROCŁAWSKA

WYDZIAŁ ELEKTRONIKI

KIERUNEK: Automatyka i Robotyka

SPECJALNOŚĆ: Technologie informacyjne w systemach automatyki

PRACA DYPLOMOWA

MAGISTERSKA

Badanie metod rozpoznawania znaków drogowych z wykorzystaniem wielu źródeł obrazu  
Research on traffic sign recognition methods using multiple image source

autor : BARTOSZ LENARTOWICZ

autor : BARTOSZ LENARTOWICZ

Opiekun pracy:

Dr inż. Bartosz Jabłoński W04/K8

OCENA PRACY:

WROCŁAW 2019

# Streszczenie

…

**Spis treści**

[Streszczenie 2](#_Toc1583629)

[Analiza problemu, cel i zakres pracy 3](#_Toc1583630)

[Metody detekcji obiektów na obrazach panoramicznych 4](#_Toc1583631)

[Algorytmy łączenia obrazów w panoramę 4](#_Toc1583632)

[Metody klasyczne 4](#_Toc1583633)

[Metody bazujące na przekształceniu Fouriera 5](#_Toc1583634)

[Metody bazujące na punktach charakterystycznych (kluczowych) 5](#_Toc1583635)

[Łączenie obrazów 8](#_Toc1583636)

[Segmentacja obrazów 8](#_Toc1583637)

[Wykrywanie na podstawie koloru 8](#_Toc1583638)

[Wykrywanie na podstawie kształtu 9](#_Toc1583639)

[Wykrywanie na podstawie tekstury 10](#_Toc1583640)

[LBP i HOG 10](#_Toc1583641)

[ACF 11](#_Toc1583642)

[Podejście hybrydowe 11](#_Toc1583643)

[CNN 11](#_Toc1583644)

[Algorytm Violi i Jonesa 11](#_Toc1583645)

[Inne podejścia 12](#_Toc1583646)

[Wnioski 12](#_Toc1583647)

[Identyfikacja obiektów 12](#_Toc1583648)

[Sieci neuronowe 12](#_Toc1583649)

[Uczenie maszynowe 12](#_Toc1583650)

[AdaBoost 13](#_Toc1583651)

[Program 13](#_Toc1583652)

[Publiczne bazy danych znaków 13](#_Toc1583653)

[Propozycja rozwinięcia/konstrukcji metody 13](#_Toc1583654)

[Przygotowanie środowiska badawczego i plan badań 13](#_Toc1583655)

[Podsumowanie 13](#_Toc1583656)

[Bibliografia 13](#_Toc1583657)

# Analiza problemu, cel i zakres pracy

Rozpoznanie znaków drogowych przez algorytmy przetwarzania obrazów jest aspektem niosącym ze sobą wiele korzyści, szczególnie jeżeli analiza następuje w czasie rzeczywistym. Pierwszą korzyścią jaka się nasuwa jest wspomaganie kierowców podczas jazdy. Po nieco dłuższym zastanowieniu można wywnioskować, że rozpoznawanie znaków jest również istotną kwestią dla pojazdów. Dodatkowo rozpoznanie znaków drogowych może realnie przyczynić się do polepszania bezpieczeństwa na drodze. Przykładowo po rozpoznaniu znaku drogowego można zbadać w jakim stopniu jest on czytelny przez porównanie z wzorcem (np. czy nie został zamazany, zakrzywiony) i jeżeli tak się stało poinformować o tym służbę drogową. Istnieje szereg dziedzin gdzie analiza znaków drogowych może wprowadzić wiele innowacji. Prawdopodobnie dlatego inżynierowie od ponad dwóch dekad starają się sprostać temu zadaniu. [kiedy pierwsza próba]

Starając się wyjaśnić zakres pracy magisterskiej należy również rozszerzyć kilka terminów, które zostały w niej poruszone. W literaturze problem rozpoznawania znaków drogowych przedstawia się jako problem TSR (ang. Traffic Sign Recognition). Jest on składową większego zagadnienia jakim jest analiza sytuacji w okuł samochodu podczas jazdy, a czasem nawet wewnątrz niego. W takim przypadku obszar wokół samochodu traktowany jest jako scena w, której należy rozpoznawać i analizować obiekty różnego typu. Zostaje to wspomniane, ponieważ niektóre systemy prócz rozpoznania znaków drogowych są w stanie rozpoznawać obiekty takie jak samochody, rowerzystów i pieszych. Taki system został przedstawiony w pracy „Fast Detection of Multiple Objects in Traffic Scenes With a Common Detection Framework” [1]. Dodatkową zaletą pewnych systemów jest możliwość obiektów np. w celu poinformowania kierowców o kursie kolizyjnym oraz zapobiegnięcie wykrycia kilkukrotnie tego samego obiektu. Według autorów pracy „"Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey” [2] ostatni aspekt jest szczególnie ważny dla systemów w których pojazd jest kierowany przez kierowcę (a nie przez algorytm). Wielokrotne informowanie kierowcy o znaku może rozproszyć jego uwagę powodując tym samym zagrożenie. Autor w swojej pracy wspomina, że traktowanie kierowcy jako integralną część środowiska może polepszyć parametry systemu.

Chcąc zrealizować w skończonym czasie temat pracy należało wprowadzić pewne ograniczenia. Pierwszym z nich jest problem wykrywania obiektów zawężony do problemu TSR. Dodatkowo w pracy główny nacisk został położony na wykrywanie znaków pionowych. Obrazy zostały zarejestrowane przez kamery cyfrowe obejmujące obszar przed samochodem [oraz części prawego pobocza?]. Klatki utrwalone w tym samym czasie zostały połączone w obraz panoramiczny, na którym wystąpiła detekcja obiektów.

Skupiając się na problemie badawczym czyli wpływie zastosowania kilku kamer, zamiast kamery pojedynczej, należało przyjąć kryteria oceny. Problem oceny został ujęty jako zagadnienie statystyczne. Hipotezą zerową było stwierdzenie, że w przeszukiwanym obszarze nie ma znaków drogowych. Z pozoru nielogiczne założenie wprowadziło kilka udogodnień w rozumowaniu. Błędami pierwszego rodzaju w takim przypadku było wykrycie znaków drogowych gdzie w rzeczywistości ich nie ma. Błędem drugiego rodzaju było ominięcie znaku drogowego na obszarze gdzie w rzeczywistości on występuje. W pracy została przebadana skuteczność rozpoznawania znaków po wprowadzeniu dodatkowych kamer, dlatego mniejsze znaczenie przypisano czasom trwania poszczególnych algorytmów.

Warto zwrócić też uwagę na fakt, że system, który jest w stanie wykryć kilka różnych typów znaków lecz jest obarczony dużym błędem wykrywalności, w niektórych zastosowaniach może być uważany za lepszy niż system, który może wykryć tylko jeden określony typ znaku, ale robi to bardzo dobrze. Oczywiście możliwa jest również sytuacja odwrotna.

# Metody detekcji obiektów na obrazach panoramicznych

Problem identyfikacji obiektów na obrazach panoramicznych pociąga za sobą szereg zagadnień do zrealizowania. Pierwszym z nich jest stworzenie obrazu panoramicznego z kilku obrazów. Następnie należy rozpoznać co znajduje się na obrazie. To zagadnienie realizowane jest w trzech krokach. Pierwszym z nich jest przetwarzanie niskiego poziomu. Podczas takiego przetwarzania następuje akwizycja obrazu oraz przetwarzanie wstępne. Takie przetwarzanie obejmuje poprawę jakości obrazu poprzez eliminację zakłóceń, poprawę kontrastu, filtrację itd. Kolejny krok nazwany przetwarzaniem średniego poziomu dotyczy segmentacji obrazu. Segmentacja jest przeprowadzona na podstawie wydzielonych cech obrazu. Przykładowo znajduje się punkty kluczowe na obrazie, po czym następuje znalezienie obszarów, w których mogą wystąpić obiekty. Po segmentacji następuje próba rozpoznania i klasyfikacji obiektu. Gdy mowa obrazach dostarczanych przez kamery należało by również dodać etap śledzenia. Ponieważ każdy z tych etapów przetwarzania obrazów można traktować jako oddzielny problem, poniżej w osobnych podrozdziałach zostały przedstawione metody rozwiązania każdego z powyższych zagadnień.

## Algorytmy łączenia obrazów w panoramę

Istnieje szereg algorytmów starających się sprostać problemowi dopasowania kilku obrazów do siebie. Wiele algorytmów jest odporne na duże rotację, translacje pionowe lub skalę obrazów wejściowych względem siebie. Istotną cechą poszukiwanego algorytmu była odporność na różnicę w jasności w obrazach wejściowych oraz szumy spowodowane lekkimi drganiami, rotacją czy zabrudzeniami typu, błoto, kurz, krople wody.   
W wielu pracach fuzja obrazów zostaje podzielona na dwa etapy. Pierwszym z obrazów jest faza dopasowania obrazów do siebie. Następuje to poprzez ekstrakcje cechy szczególnych klatki. Takie cechy mogą posłużyć w kolejnych etapach np. do wykrywania obiektów. Po dopasowaniu obrazów następuje faza fuzji właściwej czyli łączenie obrazu w panoramę.

### Metody klasyczne

Omawianie algorytmów dopasowania obrazów należy zacząć od metody klasycznych. Takie metody wymagają wyznaczenia maksimum miary dopasowania. Takie dopasowanie wykonuje się głównie metodami optymalizacji. Do metod klasycznych można zaliczyć wielowymiarową metodę Powella z użyciem jednowymiarowej metody Brenta [3]. Innym podejściem jest metoda Gaussa-Newtona, Levenberga-Marquardta [4]. Klasyczną metodą dopasowania dwóch klatek do siebie jest również metoda template matching. Dla tej metody jako wzorzec przyjmuje się jeden z obrazów wejściowych do którego stara się dopasować drugi obraz. Najpopularniejszą miarą dopasowania dwóch obrazów suma kwadratów błędów. Metoda została przedstawiona np. w pracy [5]. Opisywanie dokładnie całej metody mija się z celem, ponieważ posiada ona dużą złożoność obliczeniową oraz jest słabo odporna na rotację i szumy jednak przez jej trywialność należało o niej wspomnieć. Do metod klasycznych można zaliczyć również metody bazujące na informacji wzajemnej.

#### Szybki algorytm dopasowania obrazów dla potrzeb fuzji w czasie rzeczywistym

W pracy [6] opisano „Szybki algorytm dopasowania obrazów dla potrzeb fuzji w czasie rzeczywistym”. Jest to algorytm szybszy od wymienionych poprzednio. Został zaproponowany do łączenia klatek filmowych tego samego obszaru nagrywanych różnymi kamerami. Z powodzeniem można go zastosować do tworzenie panoramy. Autorzy testowali algorytm na dwóch kamerach – TV i IR. Początkowa faza algorytmu została pominięta ponieważ na potrzeby pracy klatki z filmowe mają taką samą wielkość oraz rotacje. Algorytm został przedstawiony następująco. Oba obrazy zostają poddane działaniu filtru krawędziowego. Następnie pionowe kolumny obrazów skanuje się względem siebie. Najlepsze dopasowanie wyznacza się za pomocą sumy najmniejszych kwadratów lub sumy modułów różnic, identycznie jak w metodach klasycznych. Następnie zostaje wyznaczona wynikowa pionowa translacja „y” z wykorzystaniem metody statycznej tj. średnia, mediana lub dominanta. Ta ostatnia daje najlepsze wyniki dopasowania. W następnym kroku obraz drugi zostaje przesunięty o wartość „y” względem pierwszego. Następnie czynność powtarza się dla wierszy i skanuje się obraz poziomo, oraz przesuwa względem osi „x”. Algorytm jest słabo odporny na rotację lecz jego największą zaletą jest szybkość.

### Metody bazujące na przekształceniu Fouriera

Istnieją metody opierające się na przekształceniu Fouriera. Takie metody działają w dziedzinie częstotliwości i są znacznie szybsze od metod klasycznych. Podstawową metodą wykorzystującą przekształcenie Fouriera jest metoda Fouriera-Mellina. Jak zostało wspomniane wykorzystuje ona twierdzenie Fouriera o przesunięciu. Pozwala to metodą korelacji fazowej wyznaczyć przesunięcie dwóch obrazów względem siebie. Metoda ma liniową złożoność obliczeniową i dobrze radzi sobie z rotacją obrazów.

Warta uwagi jest również metoda zaproponowana przez grupę Waterfall Solutions, która polega na wstępnym dopasowaniu obrazów. Następnie używając metod bardziej złożonych dopasowuje oba obrazy dokładnie. Wstępne przetwarzanie polega na podzieleniu obrazu na małe okna. Każde z tych okien dopasowuje się do obrazu drugiego z wykorzystaniem miary statystycznej korelacji. Następnie łączy się segmenty w całość i znajduje najlepsze globalne dopasowanie. Powyższy opis jest jedynie jedną z iteracji algorytmu, a kolejne iteracje bazują na poprzednich. Zazwyczaj algorytm kończy się po kilku iteracjach. Autorzy algorytmu w opisie przedstawiają, że algorytm wykorzystuje metodę quasi-Newtona, z piramidą obszarów. Jest to popularnym rozwiązaniem. Takie łączenie obrazów pozwala na dużą eliminację błędnych dopasowań oraz jest bardzo szybka.

### Metody bazujące na punktach charakterystycznych (kluczowych)

W metodach bardziej odpornych dopasowanie obrazów do siebie wymaga ekstrakcji cech punktowych. Cechy takie rozumiane są jako specyficzne konfiguracje pikseli układające się w struktury. Przykładami takich konfiguracji mogą być: linie, zakończenia linii, krawędzie czy kąty. Takie struktury wykrywa się zazwyczaj za pomocą filtrowania obrazów. Ważnym elementem punktów kluczowych jest ich niezmienność względem siebie. Oznacza to, że po wykryciu kilku punktów kluczowych/ charakterystycznych na dwóch obrazach, można takie obrazy do siebie dopasować. Dodatkowo punkty kluczowe można śledzić na kolejnych klatkach obrazu ponieważ są one niezmienne względem bliskiego otoczenia. Opisuje się je jako skalo-niezmiennicze.

#### Detektor Harrisa

Na potrzeby łączenia obrazu w panoramę za punkty kluczowe, charakterystyczne można przyjąć narożniki. Do wykrywania krawędzi i narożników można zastosować detektor Harrisa, który jest ulepszoną wersją detektora Moraveca. Koncepcja detektora Moraveca polega na przeszukiwaniu obrazu z wykorzystaniem okna przeszukiwania. Gdy podczas przeszukiwania w jednym kierunku zostanie zauważona duża zmiana w jasności pikseli obszar zostaje zakwalifikowany jako krawędź. Jeżeli zmiana zostanie również zauważona w kierunku prostopadłym to obszar kwalifikuje się jako narożnik. Zmiany intensywności sprawdza się co 45˚. W detektorze Moraveca wadami są również zaszumiona odpowiedź z uwagi na binarną funkcję okna oraz minimum jako kryterium.   
Chcąc poprawić algorytm Chris Harris w 1988 roku zaproponował ulepszoną wersję algorytmu z powodzeniem stosowaną do dnia dzisiejszego. W detektorze Harrisa wykorzystywana jest macierz autokorelacji w postaci:

Gdzie I jest funkcją intensywności, a σ jest odchyleniem standardowym funkcji Gaussa. Jeżeli wyznaczona różnica między wyznacznikiem macierzy M, a kwadratem jej śladu przemnożonym przez stałą k ( ) będzie mniejsza od zera to punkt jest uważany za krawędź. Jeżeli zaś to punkt uważany jest za narożnik. Jeżeli punkt jest bliski zera to uważa się, że obszar nie posiada znaczących zmian. Dodatkowo dla większej dokładności algorytmu wykrywania narożników wprowadza się próg . Zaleca się również wybrania punktów z lokalnym maksimum. Detektor Harrisa wykrywa narożniki z wiele większą dokładnością niż detektor Moraveca oraz sprawdza narożniki pod każdym kątem. Algorytm może z powodzeniem zostać wykorzystywany do tworzenia panoram.

#### GFTT

Algorytm GFTT (ang. - Good Features To Track) jest ulepszeniem jakie wprowadzili Jianbo Shi oraz Carlo Tomasi w 1994 roku dla detektora Harrisa. Obraz wejściowy powinien być przetworzony do skali szarości. Zmianie podległ wzór wyliczający funkcje oceniającą . W przypadku GFTT funkcja ta ma postać

Dzięki swojej modyfikacji algorytm znajduje najbardziej użyteczne wierzchołki obrazu. Dla każdego wykrytego wierzchołka no należy zadać minimum w procentach. Określa to minimalny próg jasności jakości wierzchołka i jeżeli minimum nie zostanie osiągnięte wierzchołek jest odrzucany. Dodatkowo określa się minimalny dystans pomiędzy wykrytymi wierzchołkami. Powyższe kryteria pozwalają wyłonić najlepiej procentujące wierzchołki, a więc intersujące punkty kluczowe.

#### FAST

Detektor FAST (ang. Features from Accelerated Segment Test) w swoim kryterium porównuje jasność piksela z jasnością pikseli oddalonych o ustalony promień. W przypadku gdy jasność większości pikseli będzie się różnić od piksela centralnego o określoną wartość to punkt kwalifikowany jest jako narożnik. Przykładowo dla promienia równego 3 piksele jeżeli 12 z 16 pikseli będą jaśniejsze niż piksel centralny pomniejszony o pewną wartość to algorytm wykrył narożnik. Na potrzeby algorytmu promień powinien być dostatecznie mały (od 2 do 6 pikseli) oraz wartość progu kwalifikacji i ilości pikseli odpowiednio przeskalowana do promienia. Algorytm FAST osiąga przeciętne wyniki w wykrywaniu punktów kluczowych. Jego zaletą jest szybkość.

#### SIFT

Algorytmem który wykorzystuje punkty charakterystyczne na którego należało by zwrócić szczególną uwagę jest zaproponowany przez David Lowe algorytm SIFT (ang. Scale Invariant Feature Transform) [7]. Ponieważ jedno z podejść pracy wykorzystywało właśnie algorytm SIFT jego działanie zostanie przybliżone w tym rozdziale. Jak większość opisywanych algorytmów w tym rozdziale tak i ten prócz tworzenia panoramy z powodzeniem można zastosować np. do rozpoznawania bądź śledzenia obiektów. Algorytm jest zrealizowany w czterech krokach. Kroki zostaną przedstawione a następnie szczegółowo opisane. Pierwszy nazwany „scale space extrema detection” wykrywa punkty ekstremalne na dwóch obrazach. W kroku drugim „accurate keypoint location” następuje dokładna lokalizacja punktów charakterystycznych. Następnie przypisuje się orientację wykrytym punktom w przestrzeni. Krok nazwany jest „keypoint orientation assignment”. Czwarty etap zawiera tworzenie deskryptorów dla punktów charakterystycznych.   
Szczegółowe omawianie algorytmu należy zacząć od procesu skalowania obrazu wejściowego do różnych wielkości. Dla każdego obrazu w skali zostaje użyty filtr Laplace’a który pozwala uzyskać kontóry obiektów , obliczany za pomocą odjęcia dwóch obrazów rozmytych filtrem Gaussa z różnymi parametrem σ. Zazwyczaj większe rozmycie obrazu tworzy się poprzez zwiększenie potęgi do której zostanie podniesiona stała k. Dla ułatwienia zrozumienia poniżej opisano wzór

Zazwyczaj w za parametr odchylenia standardowego przyjmuje się . Stała podnoszona do kolejnych potęg w następujących po sobie operacjach. Następnie przez binaryzację wykrywa się lokalne maksima i minima. Powoduje to wykrycie ogromnej ilości punktów ekstremalnych dlatego należy zastosować dwa kryterium które pozwoli wyłonić najlepsze możliwe punkty maksymalne. Pierwsze kryterium odrzuca płytkie minima bądź niewielkie maksima za pomocą rozwinięcia funkcji w szereg Taylora. Pozwala to odrzucić większość punktów które powstały na obszarze np. nieba. Drugie kryterium sprawdza czy punkt nie leży na odcinku i używa metody wykrywania kątów, rogów metodą Harris’a. Po otrzymaniu punktów niezmienniczych względem skalowania czyli umiejscowionych w środkach okręgu, końcach odcinków itd. algorytm przypisuje punktom orientację. W wyniku tej operacji punkty stają się niezmiennicze również względem orientacji. Można więc ustawić obraz tak by punkty wskazywały jednakowy kierunek. Realizowane jest to poprzez wyznaczenie gradientu w punkcie charakterystycznym. Następnie wyznacza się gradienty w małym otoczeniu punktu charakterystycznego i z odpowiednio wyważonych gradientów tworzy się histogram. Z histogramu powstaje deskryptor który pozwala wyznaczyć orientację obszaru przetwarzanego w końcowym etapie. Etap końcowy polega na tworzeniu ostatecznych deskryptorów. Najczęściej przetwarzaniu podlega rozmyty wejściowy obraz. Etap przypomina ten poprzedni lecz otoczenie punktu charakterystycznego jest dużo większe i podzielone na cztery obszary. W każdym obszarze wyznacza się osobne deskryptory co kończy algorytm. Uzyskuje się w ten sposób punkty charakterystyczne wraz z dokładną orientacją.

### Łączenie obrazów

Chcąc zakończyć tematykę tworzenia panoramy należało by wspomnieć w jaki sposób za pomocą metod klasycznych oraz tych bazujących na przekształceniu Fouriera lub na punktach kluczowych zostaje utworzony jednolity obraz panoramiczny. Jak łatwo się domyślić w każdej z metod wykryte cechy charakterystyczne na obu obrazach dopasowuje się do siebie. Nakłada się jeden obraz na drugi w taki sposób by występowała największa liczba pasujących elementów np. wykrytych linii, krawędzi, narożników czy deskryptorów. Następnie obrazy zostają połączone w jeden tz. Dokonuje się ich fuzji.

## Segmentacja obrazów

Jak zostało wspomniane wcześniej segmentacja obrazów polega na znalezieniu obszaru, w którym występuje poszukiwany obiekt. W przypadku systemu TSR segmentacja obrazów polega na znalezieniu na obrazie znaku drogowego. By takie wykrywanie mogło mieć miejsce należy wyznaczyć zbiór cech tworzący jakościowy i ilościowy opis obiektu (znaku drogowego). Taki opis nazywany jest wzorcem. Zapisuje się go za pomocą wektora cech: ]. Dla ułatwienia opisu obiektów zostają również wprowadzone klasy wzorców, zawierające wzorce z podobnymi wektorami cech. Klasę wzorców oznacza się za pomocą (M - numer klasy). Rozpoznanie wzorców nazwane klasyfikacją polega na przyporządkowaniu wzorców do ich klas. Wyznaczenie wzorców powinno się opierać o starannie dobrane cechy. Cechy takie powinny mieć specyficzne własności. Ważne jest by wybrane cechy przypisane do obiektu przyjmowały różne wartości dla różnych klas obiektów. Ważnym atrybutem każdej cechy powinna być jej niezawodność. Niezawodność powinna polegać na przyjmowaniu podobnych wartości dla wszystkich obiektów danej klasy. Dodatkowo cechy wybrane wzorca powinny być nieskorelowane ze sobą. Skorelowanie można opisać za pomocą współczynnika korelacji: gdzie P oznacza liczbę klasyfikowanych obiektów, μ wartości średnie, a σ ozacza odchylenie standardowe zbioru cech. Jeżeli współczynnik To cechy x i y uważa się za silnie skorelowane i jedną z nich można odrzucić. Ważne jest by dla obiektu nie przyjmować zbyt dużej ilości cech, ponieważ złożoność obliczeniowa detektorów obiektów rośnie wraz z ilością cech. Wykrywanie obiektów na obrazach może się odbywać na wiele sposobów. Kilka metod wykrywania i rozpoznawania obiektu zostały przedstawione poniżej. Wpierw zostało omówione podejście na podstawie koloru, a następnie kształtu. Zazwyczaj metoda wykrywania obiektu jest ściśle połączona z metodą rozpoznawania obiektu. Niekiedy używa się połączenia metod opartych na różnych podejściach przykładowo wpierw zostaje wykonane progowanie w określonym kolorze, a następnie rozpoznaje się kształty wykryte przez progowanie.

### Wykrywanie na podstawie koloru

Wykrywanie obiektów na podstawie koloru stosuje się zazwyczaj w przypadku obiektu którego kształt jest skomplikowany. Czasem wykrywanie na podstawie koloru stosuje się dla obiektów których kształt jest dobrze określony ze względu na szybkość działania takiego algorytmu. Oczywiście ważnym aspektem jest to by kolor obiektu znacząco odróżniał się od pozostałych obiektów. Określenie koloru może być zrealizowane za pomocą pobrania próbek z kilku wzorcowych obiektów. Takie czynność wykonywane jest zazwyczaj za pomocą ręcznego oznaczenia obiektu z którego kolor ma być pobrany. Następnie wyznacza się średni kolor i dopuszczalny zakres odchyłki. Następnie na obrazach testowych wykonuje się progowanie w ustalonym kolorze. Mimo prostoty metody w praktyce takie wykrywanie obiektów jest zazwyczaj mało skuteczne, ponieważ kolor obiektu w dużej mierze zależy od oświetlenia. Możliwe jest polepszenie wykrywania obiektów wykrywając ich teksturę. Uwzględnia się wtedy regularność powierzchni zazwyczaj przez analizę częstotliwościową występowania zmian, które są charakterystyczne.

#### Zmiana przestrzeni barw

Wiele osób uważa, że przestrzeń barw RGB jest bardzo zmienna w zależności od warunków dlatego wiele grup analitycznych proponuje przejście na przestrzeń HSI (ang. Hue, Saturation, Intensity). Pozwala to w pewnym stopniu zniwelować różne wahania koloru w zależności od pory dnia, rodzaju atmosfery czy innych zaburzeń świetlnych. Niestety przestrzeń HSI również nie jest wolna od wszelkich wad. Wiele autorów uważa, że przestrzeń nie uwzględnia temperatury barwy. Dodatkowo koszt obliczeniowy transformacji obrazu z RGB do HSI jest bardzo duży ponieważ transformacja jest nieliniowa. Nieco poprawionym podejściem jest przejście na przestrzeń LCH (ang. Lightness, Chroma, Hue) uzyskanej za pomocą modelu CIECAM97, które to podejście uwzględnia temperaturę barw. Część autorów w swoich pracach jednak stwierdza, że zmiana przestrzeni barw jest zbędna ponieważ unormowana przestrzeń RGB jest wystarczająco dobra. Niektóre sieci neuronowe [8] o których wspomniano w późniejszych rozdziałach wykorzystują przestrzeń barw YUV.

Warto wspomnieć o pracy [9] gdzie model został oparty na kolorach ale wykorzystuje nie tylko progowanie. Wykorzystuje on kaskadowy klasyfikator wyszkolony w AdaBoost. AdaBoost to oparty na boostingu algorytm zaprezentowany 1996 roku przez  Yoav Freund i Robert Schapire [10]. Algorytm wielu słabych klasyfikatorów tworzy jeden mocny klasyfikator. W przypadku wcześniej wspomnianej pracy [9] autorzy algorytmu zaproponowali pomniejszenie rozdzielczości dostępnych kolorów. Wszystkie odcienie kolorów występujących na znakach są rzutowane na ich podstawowe odpowiedniki. Tam gdzie skupiska kolorów odpowiadają tym na znakach (np. czerwony i biały) tam jest zaznaczany obszar z potencjalnym znakiem.

### Wykrywanie na podstawie kształtu

Równie powszechnym podejściem jak wykrywanie obiektów na podstawie koloru jest wykrywanie obiektów na podstawie kształtu. W przypadku znaków drogowych kształt znaku jest konkretnie określony. Znaki informacyjne są kwadratowe, nakazu lub zakazu okrągłe, a ostrzegawcze trójkątne. Wykrywanie kształtów jest problematyczne ze względu na zmienność w kształcie znaku w zależności z jakiego konta znak jest obserwowany oraz tego, że znak może być częściowo przykryty. Wykrywanie figur geometrycznych na obrazie może być zrealizowane na kilka sposobów. Te ważniejsze zostały przedstawione poniżej.

#### Wykrywanie krawędzi

Wykrywanie krawędzie jest najbardziej popularną metodą stosowaną w wykrywaniu obiektów na obrazie. W poprzednim rozdziale opisywano wykrycie krawędzi przez algorytmy klasyczne na potrzeby łączenia obrazów w panoramę. Nic nie stoi na przeszkodzie by raz wykryte krawędzie stosować do wykrywania obiektów. Najczęściej krawędź identyfikowana jest jako przejście z obszaru ciemniejszego do jaśniejszego bądź na odwrót. Można stwierdzić, że jest to granica pomiędzy dwoma obszarami o różnych jasnościach. Takie podejście wymaga ustalenia progu zmiany jasności jaki powoduje wykrycie krawędzi. Większość metod bazuje na operatorach gradientowych ustalając lokalny próg wykrywania krawędzi zamiast globalnego. Metody te wykorzystują zmiany pierwszej lub drugiej pochodnej obrazu przekonwertowanego do skali szarości. Przykładem takiego algorytmu wykrywania krawędzi jest Operator Robertsa sprawdzający różnicę miedzy sąsiednimi pikselami w obrazie. Niestety jest on bardzo czuły na zakłócenia w postaci szumów. W większości przypadków do wykrywania krawędzi stosuje się operator Canny lub innej bardzo podobnej metody [here] (opisać Canny i Haara). Warto wspomnieć, że pierwsze metody używały Operatora Sobela czyli splotu macierzy z obrazem [11].

#### Detektor Canny

#### Wykrywanie rogów

Tak jak metody klasyczne łączą się z wykrywaniem obiektów na podstawie krawędzi tak metody wykrywania narożników jako punktów kluczowych łączą się z metodami detekcji obiektów na podstawie rogów. Narożniki wykryte w poprzednim etapie mogą posłużyć do wykrywania obiektów w obecnym.

#### Transformata Hougha

Metodą na którą należało by zwrócić uwagę zastosowanie filtru Hougha, która pozwala identyfikować regularne kształty. Została ona zaprezentowana w 1962 roku jako metoda wykrywania linii prostych ale udało się ją zastosować również dla wykrywania regularnych kształtów takich jak okręgi. W 1993 roku Anagnou przestawił udoskonaloną metodę pozwalającą uzyskanie większej rozdzielczości kątowej oraz radialnej znajdowanej linii. Omówienie metody zostało zrealizowane na przykładzie wykrywania kształtu kwadratowego. W tej metodzie każda prosta opisana jest równaniem . Początkowo zostaje zaimplementowana tablica (komórki nazywa się akumulatorami) z wartościami zerowymi. Następnie każdy punkt obrazu przekształcany jest w dyskretną krzywą sinusoidalną w przestrzeni Oρϴ. Oblicza się wartości dla ϴ=(0,360˚). Obliczenie wartości ρ uzyskuje się przez dodanie jedynek w odpowiednie pola tablicy. Po wykonaniu inkrementacji dla wszystkich punktów komórki z największą wartością ρ. Położenie komórki określają parametry prostej zaś wielkość ρ określa długość odcinka. Jeżeli wykryte proste mają podobną długość i przecinają się pod kontem prostym to można uznać, że kształt jest kwadratowy. Transformacja Hougha odnosi sukcesy w wykrywaniu linii czy okręgów lecz do wykrycia innych kształtów jak np. trójkąty wymaga skomplikowanych obliczeń i jest obciążająca pamięciowo.

## Wykrywanie na podstawie tekstury

Wykrywanie obiektów na podstawie koloru i kształtu nie zawsze daje dobre wyniki, zaczęto szukać innych rozwiązań mogących poprawić skuteczność wykrywania.

### LBP i HOG

LPB (ang. local binary patterns) jest podejściem, które zostało zaprezentowane w 1990, a opisano 1994 roku [12]. Algorytm podczas działania dzieli obraz na mniejsze części np. 16x16 px). W każdej części każdy piksel porównuje się z pikselami sąsiadującymi. W najprostszej wersji jest to 8 pikseli. Następnie w ustalonej kolejności porównuje się każdy piksel sąsiadujący z pikselem centralnym i w zależności czy jego wartość jest większa lub równa czy mniejsza wpisuje się odpowiednio 1 lub 0. Tak powstałe słowo (np. 00101001) zamienia się na liczbę w systemie dziesiętnym. Dzięki takiemu rozwiązaniu gdy zostaje podniesiona jasność całego obrazu, skuteczność algorytmu jest taka sama. W każdej łacie tworzy się histogram. Metoda LPB jest zazwyczaj ściśle połączona z metodą HOG (ang. Histogram Oriented Gradient) dlatego z tak otrzymanych liczb dziesiętnych tworzy się histogram i normalizuje go, a następnie ustala gradient okna. Po stworzeniu gradientów w każdym oknie następuje połączenie podobnych gradientów ze sobą i próba rozpoznania obiektu. Jeżeli wykryte gradienty są podobne do tych, które posiada obiekt jaki chciano wykryć algorytm uznaje, że obiekt został wykryty. Algorytm odniósł duży sukces w wykrywaniu twarzy dlatego postarano się go użyć do wykrywania innych obiektów. Ponieważ algorytm do swojego działania wykorzystuje dużą ilość obliczeń wprowadza się obszar ROI, który zawęża strefę poszukiwań co poprawia szybkość algorytmu. Takie podejście zaproponowano w [13]. Wpierw algorytm ustala potencjalne obszary poszukiwań za pomocą okna przesuwanego o niewielkich rozmiarach. Etap nazywa się filtrowaniem zgrubnym i używa algorytmu HOG z klasyfikacją LDA. Następnie podczas dokładnego filtrowania weryfikuje się okna odnalezione w etapie poprzednim oraz za pomocą metody NMS (ang.  non-maximal suppression) wykonywane jest filtrowanie okien mogących odnosić się do tego samego obszaru. Klasyfikacja obiektu wykonywana jest za pomocą maszyny wektorów nośnych, która została opisana w następnym rozdziale.

## ACF

ACF łączy informacje o kolorach i gradientach, co jest lepsze od HOG [14]

## Podejście hybrydowe

Podejście hybrydowe polega na połączeniu metod opartych na kolorze, kształcie i fakturze. Zazwyczaj w początkowym etapie następuje progowanie w określonych kolorach co zawęrza obszar poszukiwań. Następnie w odnalezionych regionach wykorzystuje się podejścia oparte na krztałcie i fakturze w celu poprawy wydajności wykrywania.

## CNN

Splotowa siec neuronowa

[15] P. Sermanet, Y. LeCun, "Rozpoznawanie znaków drogowych za pomocą wieloskalowych sieci splotowych", *Proc. Int. Joint Conf. Neural Netw.*, str. 2809-2813, 2011.

## Algorytm Violi i Jonesa

Niekiedy do wykrywania znaków w czasie rzeczywistym stosuje się algorytm Viola–Jones zaproponowany w 2001 przez Paul Viola i Michael Jones do wykrywania twarzy [16]. Wielką zaletą metody jest wykrywanie obiektów w czasie rzeczywistym. Algorytm używa niewielkiej ramki przesuwanej po obrazie. W każdym położeniu ramki za pomocą cech Haara stara się odnaleźć miejsce występowania znaku. Cechy Haara sprawdzane są kaskadowo co pozwala od razu odrzucić rejony gdzie znak nie występuje. By przyspieszyć obliczenia wprowadza się macierz o rozmiarach rozdzielczości obrazu oryginalnego i oblicza się dla każdej wartości macierzy sumę złożoną z odpowiadającego wartości macierzy piksela i lewej oraz górnej wartości macierzy co pozwala zmniejszyć ilość obliczeń dla każdej cechy Haara. Metoda wykrywania znaków drogowych oraz samochodów a nawet rowerzystów została zaimplementowana w pracy [1]. Metoda zaproponowana w tej pracy oprócz wykrywania znaków wykrywa dwa inne typy obiektów dlatego wprowadzono dodatkowo funkcje subkategoryzacji.

## Inne podejścia

Wartym wspomnienia podejściem detekcji znaków jest tworzenie mapy cieplnej obszarów występowania znaków. Taka mapa działa na zasadzie algorytmów biologicznych. Autorzy artykułu [17]. [here] (opisać artykuł)

Sieci neuronowe radialne

## Wnioski

**Fast Detection of Multiple Objects in Traffic Scenes With a Common Detection Framework** - Podejścia oparte na fakturze najpierw wyodrębniają ręcznie spreparowane funkcje obliczone na podstawie tekstury obrazów, a następnie wykorzystują te wyodrębnione funkcje do szkolenia klasyfikatora. Popularne ręczne funkcje obejmują HOG, LBP, ACF, itp. [2] , [7] , [11] . Niektóre podejścia [35] , [51] , [63] wykorzystują funkcje HOG z SVM, inne [42] używają funkcji ACF z klasyfikatorem AdaBoost. Oprócz powyższych podejść, splotowa sieć neuronowa (CNN) jest stosowana do wykrywania znaków drogowych i osiąga doskonałe wyniki w [56] .

Zdecydowanie najbardziej popularne są krawędzie i gradienty, ale zbadano inne opcje, takie jak fale fal HOG i Haar. Etap wykrywania jest zdominowany przez transformatę Hough i jej pochodne, ale dla funkcji falowych HOG i Haar zastosowano również SVM, sieci neuronowe i kaskadowe klasyfikatory.

Operator Sobela jako prosta metoda wykrywania krawędzi algorytm genetyczny

## Identyfikacja obiektów

### Sieci neuronowe

Sieci neuronowe były stosowane do wykrywania znaków drogowych od początku. Już artykuł z 1996 roku [11] używa sieci do wykrycia typu znaku drogowego. Oczywiście każda sieć na wejściu posiada określoną liczbę neuronów dlatego należy obraz wejściowy przeskalować, znormalizować do tej liczby. Do normalizacji obrazu można stosować wiele metod. Najprostszą z nich jest odrzucenie części pikseli. Możliwa jest również interpolacja dwuliniowa.

### Uczenie maszynowe

#### SVM

Część systemów TSR do klasyfikacji znaków drogowych korzysta z maszyny wektorów nośnych w literaturze nazywanej po prostu SVM (z ang. support vector machine). Przykładem takiego systemu jest wcześniej omawiany system z 2013 roku przedstawiony w pracy [13]. Teoria działania systemu SVM polega na znalezieniu hiperpłaszczyzny oddzielającej dwie klasy. Do wytrenowania SVM należy w pierwszym etapie wprowadzić oznaczone dane uczące. Następnie należy znaleźć taką funkcję, która oddzieli inaczej oznaczone klasy. Jeżeli taka funkcja nie istnieje należy wprowadzić funkcję jądra, która przutuje dane na funkcję o wyższym wymiarze. W przypadku wspomnianej pracy autorzy wyuczyli algorytm bazujący na maszynie wektorów do klasyfikacji znaków ostrzegawczych i informacyjnych. W przypadku tej pracy algorytm był szkolony za pomocą 600 różnych danych testowych pochodzących z zbioru GTSBD oraz testowany na 300 innych obrazach. Osiągnął dobre wyniki wykrywania niestety nie pozwalające pracować w czasie rzeczywistym. Implementacja została wykonana za pomocą programy Matlab i na jednostce Core 13 3.3 GHz. Możliwe, że po implementacji w języku C/C++ i na bardziej wydajnym procesorze szybkość wzrosła by na tyle, że można by było uznać że algorytm pracuje w czasie rzeczywistym.

### AdaBoost

Klasyfikator AdaBoost używany mi w pracy [1] jako słabe klasyfikatory przyjmuje drzewa decyzyjne

Y. Aoyagi and T. Asakura, "A study on traffic sign recognition in scene image using genetic algorithms and neural networks, " in Proceedings of the 1996 IEEE IECON 22nd International Conference on Industrial Electronics, Control, and Instrumentation, vol. 3, 1996, pp. 1838- 1843.

 Show Context [View Article](https://ieeexplore.ieee.org/document/570749)[Full Text: PDF](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=570749) (560KB) [Google Scholar](https://scholar.google.com/scholar?as_q=A+study+on+traffic+sign+recognition+in+scene+image+using+genetic+algorithms+and+neural+networks&as_occt=title&hl=en&as_sdt=0%2C31)

**10.** A. d. l. Escalera, J. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles, " Image and Vision Computing, vol. 21, no. 3, pp. 247-258, 2003.

 Show Context [CrossRef](https://doi.org/10.1016/S0262-8856(02)00156-7" \t "_blank)[Google Scholar](https://scholar.google.com/scholar?as_q=Traffic+sign+recognition+and+analysis+for+intelligent+vehicles&as_occt=title&hl=en&as_sdt=0%2C31)

# Program

## Publiczne bazy danych znaków

Istnieje kilka publicznie dostępnych zestawów danych znaków drogowych:

1. Niemiecki test TSR (GTSRB),
2. KUL Belgium Zestaw danych znaków drogowych (zestaw danych KUL,
3. Szwedzki zestaw danych znaków drogowych (zestaw danych STS),
4. Baza danych obrazów RUG Traffic Sign (zestaw danych RUG),
5. Baza danych Stereopolis

# Propozycja rozwinięcia/konstrukcji metody

# Przygotowanie środowiska badawczego i plan badań

Wyniki badań

# Podsumowanie

# Bibliografia

|  |  |
| --- | --- |
| [1] | S. P. C. S. A. v. d. H. a. F. P. Q. Hu, „Fast Detection of Multiple Objects in Traffic Scenes With a Common Detection Framework,” *IEEE Transactions on Intelligent Transportation Systems,* pp. 1002-1024, kwieceń 2016. |
| [2] | M. M. T. a. T. B. M. A. Mogelmose, „Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey,” *Transactions on Intelligent Transportation Systems,* pp. 1484-1497, grudzień 2012. |
| [3] | C. A. V. D. M. G. S. P. Maes F., „Multimodality image registration by maximization of mutual information,” *Transactions on Medical Imaging,* p. 187–198, 1997. |
| [4] | F. J. Zitová B., „Image registration methods, a survey. Image and Vision Computing,” p. 977–1000, 2003. |
| [5] | T. Szymczyk, „Metoda dopasowania wzorców w rozpoznawaniu obrazów – ograniczenia, problemy i modyfikacje,” w *Automatyka*, 2008, pp. Tom12, Zeszyt 2. |
| [6] | B. P. M. B. Marcin Kondej, „Szybki algorytm dopasowania obrazów dla potrzeb fuzji w czasie rzeczywistym,” *Projekt rozwojowy MNiSW nr. O R00 0019 07,* 25 listopad 2010. |
| [7] | D. Lowe, „Scale-invariant feature transform,” *University of British Columbia,* 1999. |
| [8] | Y. L. Pierre Sermanet, „Traffic sign recognition with multi-scale Convolutional Networks,” w *The 2011 International Joint Conference on Neural Networks*, San Jose, 31 lipca - 5 sierpnia 2011. |
| [9] | M. S. K. D. I. I. H. M. D. Deguchi, „Intelligent traffic sign detector: Adaptive learning based on online gathering of training samples,” *Proc. IEEE IV Symp.,* pp. 72-77, 2011. |
| [10] | Y. F. R. E. Schapire, „Experiments with a New Boosting Algorithm,” w *Machine Learning: Proceedings of the Thirteenth International Conference*, New Jersey, 199. |
| [11] | L. E. M. M. A. S. J. M. A. Arturo de la Escalera, „Road Traffic Sign Detection and Classification,” *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS,* grudzień 1997. |
| [12] | M. P. a. D. H. T. Ojala, „Performance evaluation of texture measures with classification based on Kullback discrimination of distributions,” w *ICPR*, 1994. |
| [13] | G. R. Z. W. Y. Z. L. J. G. Wang, „A robust coarse-to-fine traffic sign detection method,” w *Proc. Int. Joint Conf. Neural Netw*, 2013. |
| [14] | Z. T. P. P. S. B. P. Dollár, „Integral channel features,” w *Proc. Bri. Conf. Mach. Vis.*, 2009. |
| [15] | Y. L. P. Sermanet, „Traffic sign recognition with multi-scale convolutional networks,” w *Proc. Int. Joint Conf. Neural Netw.*, 2011. |
| [16] | M. J. J. P. Viola, „Robust real-time face detection,” Maj 2004. |
| [17] | T. M. T. B. J. F. C. G. R. Kastner, „Attention-based traffic sign recognition with an array of weak classifiers,” czerwiec 2010. |
| [18] | P. Guzik, „Metody wyszukiwania punktów charakterystycznych i wyznaczania ich cech,” *Praca dyplomowa inżynierska Politechniki Warszawskiej,* 2014. |