

POLITECHNIKA WROCLAWSKA

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

Opiekun pracy:

Dr inż. Bartosz Jabłoński W04/K8

OCENA PRACY:

WROCLAW 2019

Streszczenie

...

Spis treści

Streszczenie.....	2
Analiza problemu, cel i zakres pracy	2
Metody detekcji obiektów na obrazach panoramicznych	3
Algorytmy łączenia obrazów w panoramę.....	4
Metody klasyczne.....	4
Metody bazujące na przekształceniu Fouriera	5
Metody bazujące na punktach charakterystycznych (kluczowych)	5
Łączenie obrazów	7
Segmentacja obrazów	7
Wykrywanie na podstawie koloru	8
Wykrywanie na podstawie kształtu	9
Inne podejścia	10
Identyfikacja obiektów.....	10
Program.....	10
Publiczne bazy danych znaków	10
Propozycja rozwinięcia/konstrukcji metody.....	10
Przygotowanie środowiska badawczego i plan badań.....	10
Wyniki badań	10
Podsumowanie	10
Bibliografia	10

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 okół 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 kilkakrotnie 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.

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 o 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 rozwiązywania 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, translację pionową lub skalę obrazów wejściowych względem siebie. Istotną cechą poszukiwanego algorytmu była odporność na różnicę w jasności w obrazach wejściowych oraz szumy spowodowane lekkimi drganiem, 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 ekstrakcję 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 tworzenia 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 rotację. 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 przesuwają 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 niezmiennosc 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 niezmiennie 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:

$$M = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}} & \left(\frac{\partial I}{\partial x}\right)^2 \left(\frac{\partial I}{\partial y}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}} \\ \left(\frac{\partial I}{\partial x}\right)^2 \left(\frac{\partial I}{\partial y}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}} & \left(\frac{\partial I}{\partial y}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}} \end{bmatrix}$$

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 ($P(x, y) = \det(M) - k(\text{trace}(M))^2$) będzie mniejsza od zera $P(x, y) < 0$ to punkt $P(x, y)$ jest uważany za krawędź. Jeżeli zaś $P(x, y) > 0$ to punkt $P(x, y)$ uważany jest za narożnik. Jeżeli punkt jest bliski zera $P(x, y) \sim 0$ to uważa się, że obszar nie posiada znaczących zmian. Dodatkowo dla większej dokładności algorytmu wykrywania narożników wprowadza się próg $P(x, y) > T$. 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 funkcję oceniającą $P(x, y) = \det(M) - k(\text{trace}(M))^2$. W przypadku GFTT funkcja ta ma postać

$$P(x, y) = \min\left(\left(\frac{\partial I}{\partial x}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}}, \left(\frac{\partial I}{\partial y}\right)^2 e^{\frac{-(x^2+y^2)}{2\sigma^2}}\right)$$

Dzięki swojej modyfikacji algorytm znajduje najbardziej użyteczne wierzchołki obrazu. Dla każdego wykrytego wierzchołka 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 interesujące 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 kontury obiektów D_n , 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

$$D_n = \frac{1}{2\pi(\sigma k^{n+1})^2} e^{-\frac{(x^2+y^2)}{2k^2\sigma^2}} \cdot P(x, y) - \frac{1}{2\pi(\sigma k^n)^2} e^{-\frac{(x^2+y^2)}{2k^2\sigma^2}} \cdot P(x, y)$$

Zazwyczaj w za parametr odchylenia standardowego przyjmuje się $\sigma = 1,6$. Stała $k = \sqrt{2}$ 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: $x = [x_1, x_2, \dots, x_n]$. Dla ułatwienia opisu obiektów zostają również wprowadzone klasy wzorców, zawierające wzorce z podobnymi wektorami cech. Klasę wzorców oznacza się za pomocą $\omega_1, \omega_2, \dots, \omega_M$ (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: $\sigma_{x,y} = \frac{\frac{1}{P} \sum_{i=1}^P (x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \sigma_y}$ gdzie P oznacza liczbę

klasyfikowanych obiektów, μ wartości średnie, a σ oznacza odchylenie standardowe zbioru cech. Jeżeli współczynnik $\sigma_{x,y} = 1$ (-1) 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 najważniejszych zostanie omówione poniżej. Wpierw zostanie omówione podejście na podstawie koloru, a następnie kształtu. W dużej ilości przypadków te dwa podejścia są ze sobą łączone co daje dobre wyniki. Wpierw na obrazie zostaje wykonane progowanie w określonym kolorze, a następnie wykrywa się kształty w wykrytych regionach.

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 czynności 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. 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.

Warto wspomnieć o pracy [8] 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 [9]. Algorytm wielu słabych klasyfikatorów tworzy jeden mocny klasyfikator. W przypadku wcześniej wspomnianej pracy [8] 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 kąta 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

Jest to o tyle ciekawe podejście, że możliwe jest użycie raz już wykrytych krawędzi. 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 próg lokalny 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 między sąsiednimi pikselami w obrazie. Niestety jest on bardzo czuły na zakłócenia w postaci szumów.

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 wzorców na obrazach binarnych. W 1993 roku Anagnou przedstawił 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 $\rho = x \cos \theta + y \sin \theta$. Początkowo zostaje zaimplementowana tablica (komórki nazywa się akumulatorami) z wartościami zerowymi. Następnie każdy punkt obrazu przekształcany jest w dyskretną krzywą sinusoidalną w przestrzeni $\text{Op}\Theta$. Oblicza się wartości dla $\Theta=(0,360^\circ)$. 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 określa parametry prostej zaś wielkość ρ określa długość odcinka. Jeżeli wykryte proste mają podobną długość i przecinają się pod kątem prostym to można uznać, że kształt jest kwadratowy.

Inne podejścia

Wartym wspomnienia podejściem detekcji znaków jest tworzenie mapy cieplnej obszarów występowania znaków. Taka mapa działa na zasadzie algorytmów biologicznych. Autorzy artykułu [10].

Identyfikacja obiektów

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, kwiecień 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] 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.
- [9] Y. F. R. E. Schapire, „Experiments with a New Boosting Algorithm,” w *Machine Learning: Proceedings of the Thirteenth International Conference*, New Jersey, 199.
- [10] T. M. T. B. J. F. C. G. R. Kastner, „Attention-based traffic sign recognition with an array of weak classifiers,” czerwiec 2010.
- [11] P. Guzik, „Metody wyszukiwania punktów charakterystycznych i wyznaczania ich cech,” *Praca dyplomowa inżynierska Politechniki Warszawskiej*, 2014.