

Sprawozdanie z laboratorium:
Komunikacja Człowiek Komputer

Wyszukiwanie i rozpoznawanie tablic rejestracyjnych.

8 listopada 2012

Prowadzący: dr inż. Wojciech Jaśkowski

Autorzy:	Michał El Fartas	inf100160	I1b	m.elfartas@gmail.com
	Marcin Biernacki	inf98740	I1b	phisikus@o2.pl

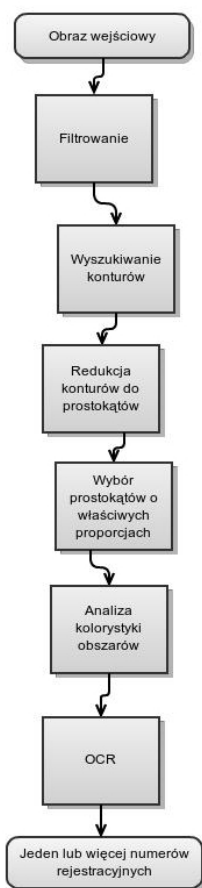
Zajęcia piątkowe, 08:00.

1 Wstęp

Głównym celem projektu było przedstawienie jednego z możliwych rozwiązań problemu rozpoznawania tablic rejestracyjnych pojazdów na zdjęciach. Systemy realizujące ALPR (Automatic License-Plate Recognition) są powszechnie wykorzystywane przez różnego rodzaju służby do określania sprawców wykroczeń drogowych, jak również przez właścicieli parkingów i autostrad jako pomoc w identyfikacji klientów.

Przygotowany program przyjmuje na wejściu zdjęcie pojazdu z widoczną tablicą, analizuje je i przedstawia na wyjściu jeden lub więcej ciągów znaków, które najbardziej pasują do tablicy ze zdjęcia.

Generalny algorytm jest następujący:



Rysunek 1: Schemat algorytmu

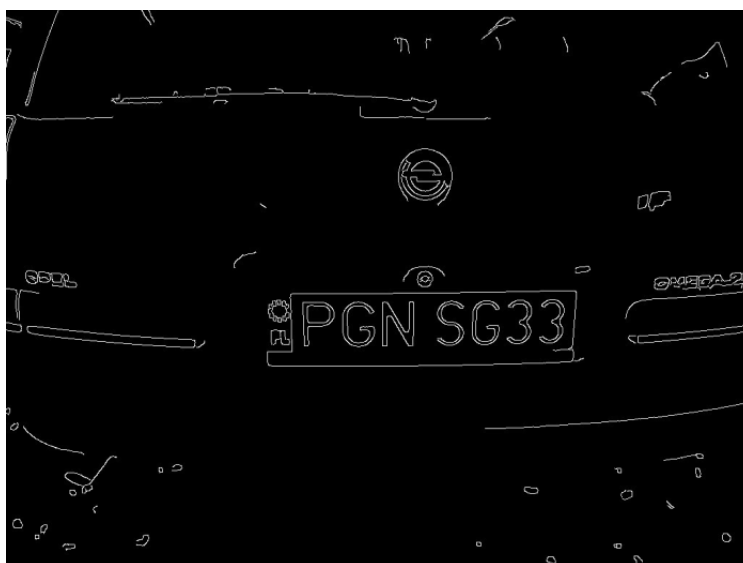
Program realizujący to zadanie powstał z użyciem biblioteki OpenCV oraz biblioteki Tesseract.

2 Filtry

Mając na uwadze fakt, że na wejściowym obrazie należy znaleźć obiekt o dość regularnych proporcjach wskazane było zastosowanie filtrów, które pozbyłyby się szumów oraz uwypukliły cechy szukanego obszaru. Na początku użyto rozmycia Gaussa do "wygładzenia" niepożądanymi elementami tła. W drugim kroku zastosowano funkcję dilata, która uwypukla jasne elementy (takie jak tło tablicy). Na koniec zastosowano algorytm Canny otrzymując czarno-biały obraz zawierający kontury.



Rysunek 2: Obraz po rozmyciu Gaussa



Rysunek 3: Końcowy filtr - algorytm Canny

3 Wyodrębnianie tablic

Na przefiltrowanym obrazie wykonujemy funkcję `findContours` z biblioteki OpenCV, która pozwala na wyznaczenie na obrazie konturów niezależnych obszarów. Ponieważ niektóre z tych obszarów mogą zawierać tablice rejestracyjne, konieczne jest podjęcie taktyki eliminującej zbędne obszary. W pierwszym kroku za pomocą funkcji `approxPolyDP` dokonujemy "optymalizacji" obszarów zapisanych w postaci listy punktów. Połączenie tych punktów daje figurę która dokładnie przylega do konturu, ale nie pozwala na bezpośrednie obliczenie proporcji. Eliminacja sąsiadujących punktów zmniejsza dokładność, ostatecznie sprowadzając obszary do postaci prostokątów. Ponieważ znane są proporcje tablic rejestracyjnych, możliwe jest odrzucenie wszystkich tych obszarów, które przez swoje wymiary na pewno tablicami nie są. Zmniejsza to znacznie ilość koniecznych do przeanalizowania algorytmem OCR obrazów.



Rysunek 4: Zaznaczone ramkami obszary potencjalnie zawierające tablice.



Rysunek 5: Tablica wykryta przy zdjęciu nachylnym w pionie.



Rysunek 6: Powyżej wzorcowo wykryty pojedynczy obszar.



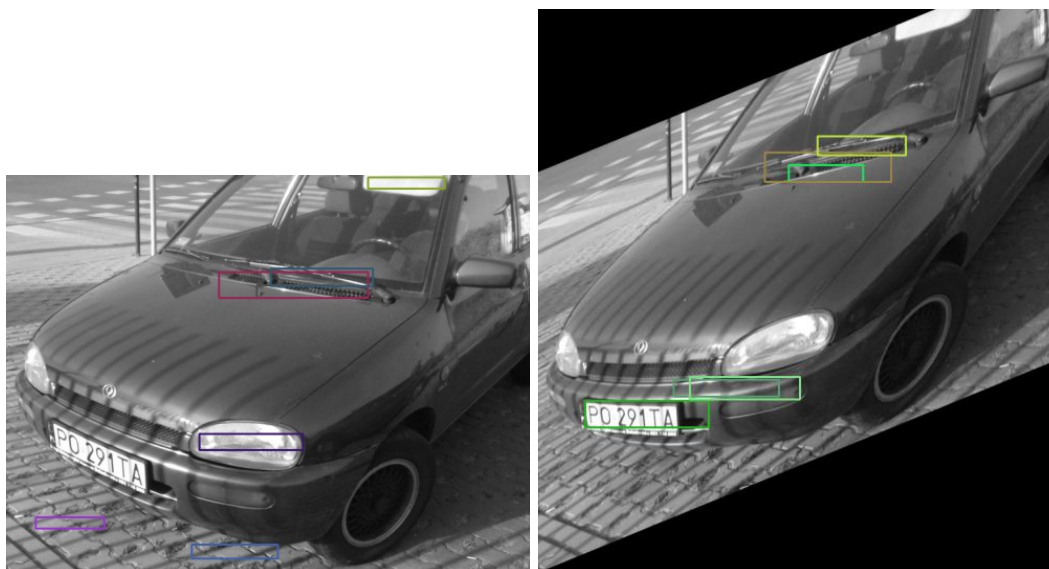
Rysunek 7: W tym przypadku oba obszary pokrywają tablicę rejestracyjną.

3.1 Krzywe tablice

Ponieważ w niektórych przypadkach zdjęcia tablicy są robione od boku, to algorytm powinien przewidywać możliwość zaburzenia proporcji. W tym celu został zaimplementowany system "pochyleń" - oraz zostaje przetransformowany w taki sposób że jego prawa krawędź znajdzie się częściowo powyżej lub poniżej jego lewej krawędzi. W ten sposób niwelujemy zaburzenie proporcji odpowiednio ją "prostując".

4 Selekcja tablic

Niestety wykrywanie tablic przez dopasowanie proporcji często zwracało dodatkowe obiekty zupełnie nie związane z tablicami. Pośród wyników można było znaleźć poza tablicami zdjęcia fragmentów karoserii, podłoża oraz inne elementy tła. Aby pozbyć się niepożądanych



Rysunek 8: Przykład zdjęcia wykonanego od boku w którym "pochylenie" przyniosło pożądany efekt.

obrazów stworzona została funkcja oceniająca jakość zdjęcia. Przez "jakość" rozumiemy tu prawdopodobieństwo że na zdjęciu znajduje się tablica rejestracyjna.

Ponieważ dopasowywanie na podstawie kształtu i proporcji zostało już wykonane w poprzednim etapie, to w tym etapie zastosowane zostało wybieranie na podstawie kolorów. W tym celu należało odnaleźć charakterystyczne cechy tablic rejestracyjnych, które będą stosunkowo łatwe do potwierdzenia na wielu zdjęciach. Najbardziej widocznymi cechami tablic rejestracyjnych są: kolorystyka mieszcząca się w skali szarości oraz stosunkowo wysoki kontrast między czarnymi literami oraz białym tłem.

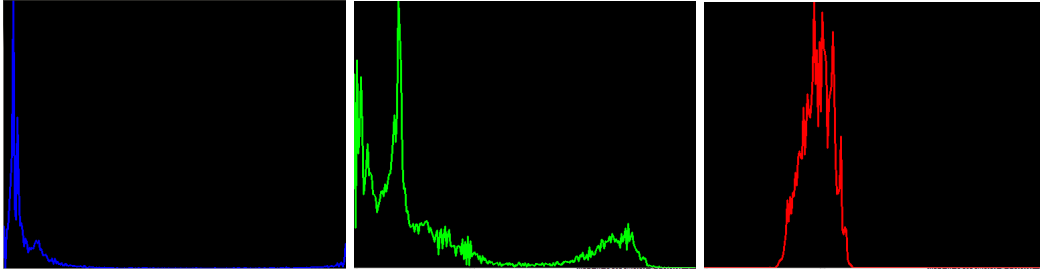
4.1 Wybór modelu

Aby ułatwić rozpoznawanie tych cech model przestrzeni barw został przekonwertowany z RGB na HSV. Kolor w modelu HSV jest reprezentowany przez trzy parametry: Hue(kolor), Saturation(nasycenie) oraz Value(jasność). Sprawdzając czy dane zdjęcie jest zdjęciem tablicy rejestracyjnej przyjmujemy że jej obraz będzie przede wszystkim składał się z kolorów - białego i czarnego (właściwych kolorów tablicy) oraz odcieni szarego wynikających z niedokładności zdjęć oraz różnego oświetlenia.

4.2 Analiza koloru

Funkcja oceniająca jako pierwszy parametr analizuje nasycenie(saturation) koloru $qSat$. Dzięki zastosowaniu modelu HSV wiemy że jeżeli wartość drugiej składowej (S) będzie niska to kolor będzie w skali szarości. Ponieważ większość zdjęć samochodów jest robiona na zewnątrz przy różnobarwnym oświetleniu oraz cieniowaniu przyjęty został margines błędu.

Jako główną składową jakości saturacji $qSat1$ przyjęty został procent pikseli posiadający składową s mniejszą od 20% maksymalnej wartości. Druga składowa $qSat2$ oparta została na średniej $avgSat$ która reprezentuje średnią wartość nasycenia dla całego

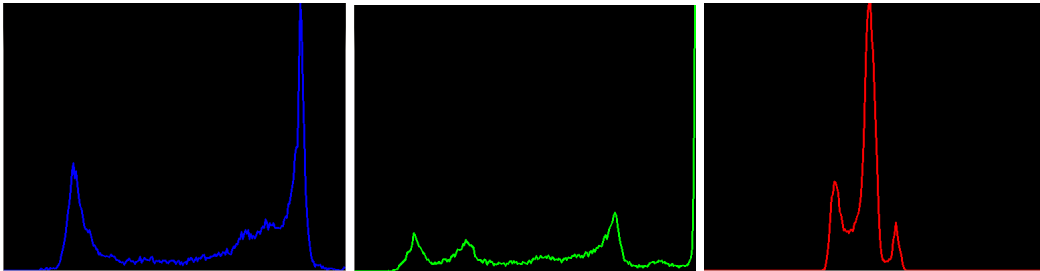


Rysunek 9: Wykresy ilości pikseli od wartości Saturation dla trzech rodzajów zdjęć: poprawnie wyciętej tablicy, tablicy z elementami karoserii, samej karoserii kolorowego auta

zdjęcia. Aby zachować jednolitość jednostek $avgSat$ została przekształcona do wartości procentowych tak, że dla minimalnej wartości średniego nasycenia wartość $qSat2$ będzie dążyć do 100% natomiast dla maksymalnej wartości otrzymamy 0%. Ostateczna wartość $qSat$ jest wyliczana jako średnia wartości $qSat1$ i $qSat2$.

4.3 Analiza kontrastu

Niestety samo badanie nasycenia nie wystarcza, ponieważ w przypadku samochodów koloru białego, szarego lub czarnego program nadal wykrywa fragmenty karoserii. Aby zapobiec takim sytuacjom należy badać jakość kontrastu, którą można wyznaczyć z składowej jasności (v). Charakterystyczną cechą kontrastu jest, że dla jego wysokiej wartości rozkład (histogram) jasności formuje się, na kształt litery U - większość pikseli przyjmuje wartości skrajne.



Rysunek 10: Wykresy ilości pikseli od wartości Value dla trzech rodzajów zdjęć: poprawnie wyciętej tablicy, tablicy z elementami karoserii, samej karoserii kolorowego auta

Do wyznaczenia jakości składowej jasności $qVal$ należy najpierw wyznaczyć wartość średnią jasności $avgVal$. Po wyliczeniu wartości średniej rozkładu jasności można podzielić rozkład na część powyżej oraz poniżej $avgVal$. Następnie dla obu części należy wyliczyć jakość która jest odmianą średniej ważonej, gdzie waga jest większa im bardziej skrajna jest wartość. Z wartości jakości dla obu tych części wykonujemy średnią geometryczną, która zwróci nam wysoką wartość jedynie jeżeli obie części będą miały wysoką jakość jednocześnie. Wartość $qVal$ jest już na tym etapie dość skutecznym miernikiem kontrastu aczkolwiek aby zabezpieczyć skuteczność tej metody, wartość ta jest dodatkowo normalizowana przy wykorzystaniu odchylenia $avgVal$ od połowy maksymalnej wartości jasności.

4.4 Wybór tablic

Tak utworzona funkcja oceniająca zwraca wartości `qSat` i `qVal`, przykładowo:

poprawnie wycięta tablica:	<code>qSat=0.888357</code>	<code>qVal=0.291081</code>
tablica wycięta z elementami karoserii:	<code>qSat=0.714162</code>	<code>qVal=0.276065</code>
fragment maski czerwonego auta:	<code>qSat=0.00956175</code>	<code>qVal=0.007453</code>
fragment maski szarego auta:	<code>qSat=0.986233</code>	<code>qVal=0</code>

Opierając się na otrzymanych wynikach, ustalone zostały następujące progi: 0.8 dla `qSat` oraz 0.1 dla `qVal`.

5 Rozpoznawanie numerów

W ostatnim kroku konieczne było rozpoznanie symboli znajdujących się na obszarach domniemania zawierających tablice. W tym celu pomocna okazała się biblioteka Tesseract implementująca technologię OCR. Biblioteka ta jest uważana za jedną z najlepszych "wolnych" bibliotek tego typu i potrafi rozpoznawać symbole z różnych alfabetów. W większości przypadków otrzymane numery rejestracyjne były zgodne z oczekiwanymi, jednakże w niektórych sytuacjach zwracane były dodatkowe i niewłaściwe znaki, które nie należą do alfabetu. Błędy w rozpoznawaniu znaków wynikały z niedokładności zaznaczonych obszarów. Poza białym tłem, tablica zawiera eurotag (niebieski) oraz obramowanie, które często samo zawiera jakiś tekst. Zdarzało się, że analizowany obraz zawierał na krawędzi fragment obramowania, rozpoznawany jako "—" bądź "I". Występowanie tego zjawiska zredukowało dostrojenie parametrów algorytmów użytych we wcześniejszych etapach analizy oraz ustawienie w parametrach biblioteki ciągu dopuszczalnych znaków. Ten ostatni zabieg wyeliminował wszystkie niealfanumeryczne znaki. W trakcie testów powstała również funkcja ewaluacyjna, określająca jakość otrzymanych rezultatów. Wszelkie odstępstwa od normy takie jak: nadmierna ilość znaków, ciągi zawierające zbyt dużo/zbyt mało liter lub cyfr, były punktowane. Taktyka ta pozwalała na określenie, które wyniki OCR pochodzą z obszarów niewłaściwie wybranych w poprzednich etapach. W późniejszych pracach dostrajających algorytm znaczenie tej funkcjonalności zmalało.

6 Testy

Pierwsza baza zdjęć testowych zawierała ok. 60 pozycji, z których niemal połowa została wykluczona - były to zdjęcia zrobione pod dużym kątem, nieostre, mające na celu oddać charakter wszystkich możliwych zaburzeń. Ostatecznie za podstawę do testów posłużyło 35 zdjęć przedstawiających samochody o różnych rozmiarach, w różnym otoczeniu i oświetleniu. Pojazdy były fotografowane pod różnymi, małymi kątami od przodu i z tyłu. Na ok. 60% wszystkich zdjęć odnaleziono tablicę rejestracyjną. Na połowie z odnalezionych tablic wyznaczone numery były bezbłędne, pozostałe zawierały drobne błędy (m.in. przez obecność fragmentów obramowań w analizowanych wycinkach). Na około 30% zdjęć tablice nie zostały prawidłowo odnalezione. Tak pozornie niska wykrywalność wynika z różnorodności zdjęć. Porównując ten algorytm z komercyjnymi systemami stosowanymi np. w policji należy pamiętać, że są one zwykle zintegrowane z urządzeniami wykonującymi zdjęcia o dużej rozdzielczości i ustawionymi pod stałym kątem do jezdni. Dodatkowe czujniki pozwalają na wykonanie zdjęcia w określonej odległości. Uświadamia to, jak duże znaczenie na analizę obrazu ma jakość danych wejściowych oraz jak dalece zdolność człowieka do rozpoznawania kształtów wyprzedza rozwiązania programowe.