

Sprawozdanie projektu z przedmiotu POBR

1. Założenia projektu

W ramach projektu zrealizowano wykrywanie loga firmy SubWay (rys.1) dla reprezentatywnego zestawu pięciu obrazów wejściowych (folder *obrazy/*). W tym celu zaimplementowano procedury wstępnego przetworzenia, segmentacji, wyznaczania cech oraz identyfikacji obrazów cyfrowych.



rys. 1 rozpoznawane logo

Program został wykonany w środowisku Microsoft Visual Studio 2012 (32 bit) przy użyciu języka C++ oraz biblioteki OpenCV 2.4.11.

2. Wstępne przetworzenie

Na pierwszy krok przetwarzania składała się konwersja obrazu z modelu RGB na HLS. Następnie dokonano progowania z naciskiem na kanał H (żółta barwa) oraz L (jasność odpowiadająca bieli). W ten sposób otrzymano binarny obraz, gdzie piksele spełniające dany warunek (potencjalnie należące do loga) zostały pomalowane na białą, pozostałe zaś na czarno.

3. Segmentacja

Segmentacja czarno-białego obrazu została przeprowadzona z wykorzystaniem algorytmu *Flood Fill*. Proces polega na przechodzeniu przez kolejne punkty obrazu, po napotkaniu białego piksela tworzony jest nowy obiekt Segment i przyporządkowuje mu się pseudolosowy kolor. Dla wybranego białego piksela wywoływany jest algorytm *Flood Fill*. Analizowane są sąsiednie piksele i jeśli są koloru białego, to ich współrzędne zostają wstawiane do pomocniczej listy punktów pełniącej rolę kolejki. W każdej iteracji algorytmu usuwa się pierwszy punkt z pomocniczej listy punktów, wstawia do listy punktów należących do segmentu i maluje na przyporządkowany wcześniej segmentowi kolor. Następnie przechodzi się do analizy sąsiadów wstawionego punktu. Algorytm kończy się w momencie osiągnięcia pustej pomocniczej listy punktów.

Z otrzymanej listy segmentów usuwane są te, które składają się z mniej niż 300 pikseli. Pozwoliło to na znaczną redukcję liczby segmentów, które w następnym kroku poddawane są analizie.

4. Analiza i wyznaczanie cech

Dla każdego z segmentów zostają policzone i zapisane wartości niezmienników momentowych M1, M2, M3 oraz M7.

5. Identyfikacja loga

Logo jest identyfikowane na dwa sposoby: jako całość lub przy pomocy jego liter składowych. Zostało to wymuszone przez to, że w przypadku jednego obrazu (subway3.jpg) logo trochę odbiega od wzorca i w wyniku segmentacji cały napis stanowi jeden segment. W przypadku reszty obrazów (subway.jpg, subway2.jpg, subway4.jpg oraz subway5.jpg) każda z liter (S, U, B, W, A, Y) staje się osobnym segmentem.

Uznanie danego segmentu za element loga następuje poprzez sprawdzenie czy wyliczone wcześniej dla niego wartości niezmienników momentowych $M1$, $M2$, $M3$ oraz $M7$ mieszczą się w którymś z zakresów wartości przyjmowanych przez wzorcowe litery S, U, B, W, A, Y lub cały napis. Pozwoliło to na jednoznaczną identyfikację loga.

Na wyjściu program wyświetla dwa obrazy. Pierwszy stanowi obraz wyjściowy z czerwonym obramowaniem znalezionej loga. Drugi jest zaś czarny z pozostawionym tylko kolorowym napisem Subway (w przypadku odnalezienia loga).



Przykładowy wynik programu - obrazy z rozpoznanym logiem dla pliku subway2.jpg

6. Wnioski

Zastosowane metody przetwarzania obrazów pozwoliły na identyfikację loga firmy SubWay na każdym z obrazów testowych.

Kluczowe dla wykonania zadania projektowego okazało wstępne przekonwertowanie obrazu z modelu RGB na HLS. Dzięki temu w łatwiejszy sposób udało się w ramach progowania wyszczególnić piksele o kolorze białym oraz żółtym składające się na logo. Warto zauważyć, że przy progowaniu należało uwzględnić dosyć szeroki zakres, gdyż barwa oraz jasność liter loga różniła się dosyć mocno między obrazami. Dodatkowo w niektórych przypadkach litery nie miały jednolitej barwy, z powodu nierównomiernego oświetlenia lub tekstury loga. Dobrym pomysłem okazało się również usunięcie zbędnych, małych segmentów jeszcze przed fazą ich analizy. Spowodowało to przyspieszenie wykonania programu. Zastosowanie algorytmu *Flood Fill* pozwoliło na wygodną i czytelną implementację procedury segmentacji. Pewne wyzwanie stanowił dobór niezmienników momentowych będących rozpatrywanymi

cechami analizowanych obiektów. W celu jednoznacznej identyfikacji wymagane było zastosowanie zestawu czterech niezmienników.