



Wydział Elektroniki i Technik Informacyjnych

POLITECHNIKA WARSZAWSKA

Przetwarzanie cyfrowe obrazów

ROZPOZNAWANIE LOGO FEDEX

Sprawozdanie z projektu

Łukasz Kilaszewski 259822

23 stycznia 2018

Wstęp

Podstawowym celem projektu było napisanie aplikacji, która będzie w stanie rozpoznać fragmenty "Fed" oraz "Ex" z logo firmy FedEx. logo to przedstawione zostało na poniższym rysunku.



Rysunek 1: Logo firmy FedEx.

W przedstawionym logo można wyróżnić dwa kolory podstawowe, które jednoznacznie identyfikują jego części. Ten fakt stanowił podstawę do implementacji algorytmu segmentacji.

Etapy przetwarzania obrazu

Przygotowanie obrazu

Obraz po wczytaniu jest konwertowany do przestrzeni barw HSV. Pozwala to wydzielić barwę (ang. *hue*), nasycenie (ang. *saturation*) oraz jasność (wartość - ang. *value*). Przyjęte wartości składowych każdego piksela to liczby całkowite z zakresów:

Składowa	Wartość min.	Wartość max.
Hue	0	179
Saturation	0	255
Value	0	255

W trakcie realizacji projektu, w celu usunięcia szumów z obrazów, zastosowana została filtracja dolnoprzepustowa za pomocą filtra konwolucyjnego i jednej z następujących macierzy:

$$\mathbf{F}_1 = \begin{bmatrix} \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \end{bmatrix}$$
$$\mathbf{F}_2 = \begin{bmatrix} \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} \\ \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} & \frac{1}{25} \end{bmatrix}$$

Operacja ta nie przynosiła jednak zauważalnej poprawy w jakości segmentacji i klasyfikacji, dlatego została usunięta.

Segmentacja

Zgodnie z wcześniejszą obserwacją, elementy rozpoznawanego logo różnią się barwą, dlatego ich segmentacja oparta jest o prosty algorytm, działający na obrazie wejściowym w przestrzeni barw HSV. Z obrazu wydzielane są dwie grupy pikseli. Warunki przynależności piksela p o składowych p.H, p.S, p.V do jednej z grup to:

Grupa "Fed": $(p.H \geq FED_H_VAL - FED_H_EPS) \ \&\& \ (p.H \leq FED_H_VAL + FED_H_EPS)$ $\&\& \ (p.S \geq FED_S_MIN)$

Grupa "Ex": $(p.H \geq EX_H_VAL - EX_H_EPS) \ \&\& \ (p.H \leq EX_H_VAL + EX_H_EPS)$ $\&\& \ (p.S \geq EX_S_MIN)$

Wartości stałych FED_H_VAL oraz EX_H_VAL zostały odczytane z obrazu przedstawiającego logo za pomocą programu do obróbki obrazów rastrowych. Pozostałe stałe zostały dobrane w sposób eksperymentalny. Wartości stałych wynoszą odpowiednio:

Stała	Wartość
FED_H_VAL	122
FED_H_EPS	10
FED_S_MIN	100
EX_H_VAL	5
EX_H_EPS	10
EX_S_MIN	100

Wynikiem segmentacji są dwie mapy bitowe przedstawiające przynależność pikseli do danej grupy.

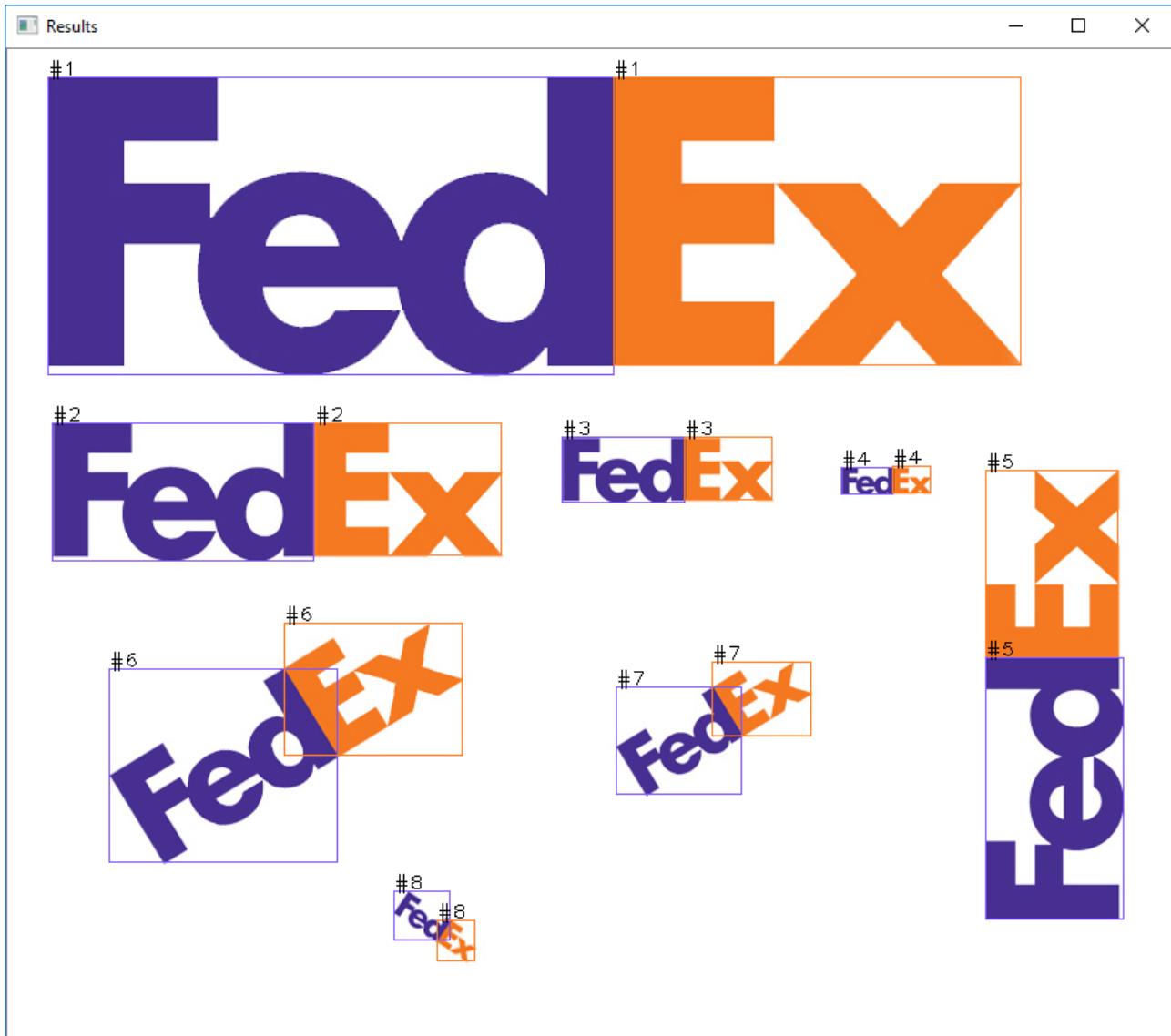
Kolejnym etapem jest filtracja otrzymanych map bitowych. W tym celu najpierw obie mapy są zamknięte za pomocą dylatacji (macierz 5×5) i erozji (macierz 3×3).

Następnie z otrzymanych map ekstrahowane są obiekty, które mogą być rozpoznawanymi elementami logo. W tym celu używany jest algorytm zalewania (ang. *flood fill*), który grupuje sąsiadujące ze sobą piksele. Obiekty złożone z mniej niż $SEGMENT_MIN_PIXELS = 150$ pikseli są odrzucane, ponieważ nie mogą zawierać w sobie elementu logo i stanowią szum. Wartość tego parametru dobrana została w sposób eksperymentalny.

Klasifikacja

Po segmentacji otrzymujemy zbiór grup pikseli (obiektów), które są domniemanymi elementami rozpoznawanego logo. Należy odrzucić te grupy, które nie zostały rozpoznane jako interesujące. W tym celu zastosowane są współczynniki kształtu i momenty geometryczne. W celu wybrania tych parametrów, które pozwolą na rozpoznanie elementów logo, przeprowadzono eksperyment na specjalnie przygotowanym obrazie testowym. Wyniki segmentacji tego obrazu przedstawiono na rysunku 2.

Dla każdego z segmentów wyznaczono wartości pewnych parametrów. Otrzymane wyniki przedstawione zostały w tabeli 1.



Rysunek 2: Wynik działania segmentacji dla obrazu testowego.

Zdecydowano, że do rozpoznawania elementu "Fed" wykorzystane zostaną parametry W6, M1, M2, M7. Ten sam zbiór parametrów początkowo wykorzystywany był do rozpoznawania elementu "Ex", natomiast w trakcie eksperymentów zdecydowano, że konieczne jest dodanie jeszcze jednego warunku. Zbiór parametrów wykorzystywanych do klasyfikacji elementów "Ex" to: W6, M1, M2, M6, M7.

Klasyfikacja polega na sprawdzeniu, czy wszystkie parametry elementu mieszczą się w odpowiednich zakresach. Wartości maksymalne i minimalne dobrano na podstawie tabeli 2 oraz eksperymentów. Są one opisane przez stałe zestawione w tabeli 3.

Grupowanie

Po klasyfikacji elementy stanowiące jedno logo łączone są w pary, a następnie rysowane są prostokąty obejmujące całe logo (ang. *bounding box*). Algorytm łączący w pary wyznacza dla

Segment	W3	W6	M1	M2	M3	M4	M6	M7
#1	2.220393	0.902794	0.385664	0.072226	0.001927	0.000550	-0.000068	0.019128
#2	2.040321	0.903828	0.364132	0.063494	0.001624	0.000460	-0.000055	0.017275
#3	1.633803	0.915984	0.328471	0.050386	0.001184	0.000333	-0.000040	0.014377
#4	0.656992	0.941814	0.272966	0.033802	0.000508	0.000162	-0.000012	0.010177
#5	2.040576	0.904119	0.364710	0.063649	0.001659	0.000165	-0.000067	0.017341
#6	2.378524	0.910247	0.351664	0.058832	0.001419	0.000628	-0.000046	0.016209
#7	1.808266	0.936040	0.317345	0.046307	0.001177	0.000532	-0.000034	0.013600
#8	0.799410	0.941694	0.268004	0.032672	0.000673	0.000206	-0.000011	0.009788
ŚREDNIA	1.697286	0.924277	0.331620	0.052671	0.001271	0.000380	-0.000042	0.014737
MIN	0.656992	0.902794	0.268004	0.032672	0.000508	0.000162	-0.000068	0.009788
MAX	2.378524	0.941814	0.385664	0.072226	0.001927	0.000628	-0.000011	0.019128
#1	1.807662	0.915587	0.305149	0.018762	0.005164	0.001851	0.000032	0.018588
#2	1.639025	0.915780	0.288693	0.016731	0.004153	0.001466	0.000021	0.016653
#3	1.352058	0.915034	0.265987	0.014446	0.003175	0.001103	0.000013	0.014076
#4	0.309014	0.965764	0.202308	0.009096	0.001064	0.000408	0.000006	0.007958
#5	1.661269	0.916043	0.289890	0.016688	0.004158	0.000469	0.000017	0.016837
#6	1.926532	0.917138	0.273685	0.013799	0.003590	0.000404	0.000021	0.015276
#7	1.574777	0.921030	0.245667	0.010118	0.002552	0.000274	0.000013	0.012558
#8	0.723016	0.941940	0.197575	0.005540	0.001118	0.000044	0.000003	0.008374
ŚREDNIA	1.374169	0.926040	0.258619	0.013148	0.003122	0.000752	0.000016	0.013790
MIN	0.309014	0.915034	0.197575	0.005540	0.001064	0.000044	0.000003	0.007958
MAX	1.926532	0.965764	0.305149	0.018762	0.005164	0.001851	0.000032	0.018588

Tabela 1: Wartości wyznaczonych parametrów dla segmentów z obrazu testowego.

każdego z elementów 8 punktów charakterystycznych:

- lewy górny róg prostokąta ograniczającego,
- lewy dolny róg prostokąta ograniczającego,
- prawy dolny róg prostokąta ograniczającego,
- prawy górny róg prostokąta ograniczającego,
- punkt wspólny elementu i lewej krawędzi prostokąta ograniczającego,
- punkt wspólny elementu i dolnej krawędzi prostokąta ograniczającego,
- punkt wspólny elementu i prawej krawędzi prostokąta ograniczającego,
- punkt wspólny elementu i górnej krawędzi prostokąta ograniczającego.

Elementy, których zbiory punktów mają co najmniej 2 pary punktów odległych o nie więcej niż GROUP_MAX_DIST = 8 stanowią jedno logo. Wartość stałej GROUP_MAX_DIST dobrana została na drodze eksperymentów.

Stała	Wartość
FED_W6_MIN	0.902
FED_W6_MAX	0.944
FED_M1_MIN	0.267
FED_M1_MAX	0.397
FED_M2_MIN	0.032
FED_M2_MAX	0.082
FED_M7_MIN	0.009
FED_M7_MAX	0.020
EX_W6_MIN	0.914
EX_W6_MAX	0.966
EX_M1_MIN	0.197
EX_M1_MAX	0.306
EX_M2_MIN	0.005
EX_M2_MAX	0.022
EX_M6_MIN	0.000002
EX_M6_MAX	0.000036
EX_M7_MIN	0.007
EX_M7_MAX	0.019

Rysunek 3: Stałe definiujące zakresy wartości parametrów.

Wyniki

Poniżej przedstawiono wyniki działania programu dla różnych obrazów wejściowych. Program z zadowalającą skutecznością rozpoznaje logo firmy FedEx na obrazach.

Na rysunku 7. - zgodnie z oczekiwaniami - nie został rozpoznany element "Fed" w białym kolorze.

Program nie rozpoznał również jednego z elementów "Ex" na rysunku 8. Prawdopodobnie spowodowane jest to faktem, że element ten jest zniekształcony ze względu na perspektywę.



Rysunek 4: Wynik działania programu dla obrazu testowego.



Rysunek 5: Wynik działania programu dla obrazu *images/1.jpg*.



Rysunek 6: Wynik działania programu dla obrazu *images/2.jpg*.



Rysunek 7: Wynik działania programu dla obrazu *images/3.jpg*.



Rysunek 8: Wynik działania programu dla obrazu *images/4.jpg*.