**Przetwarzanie Cyfrowe Obrazów**

**Sem 2021Z, Łukasz Popek**

**Temat: Detekcja logo sieci McDonald**

1. **Wprowadzenie**

Głównym celem projektu było stworzenie programu znajdującego logo sieci McDonald na rzeczywistych zdjęciach. Do tego celu posłużyła semantyczna analiza obrazu. Wszystkie algorytmy oprócz pobierania i zmiany rozmiaru zostały zaimplementowane przez autora. Większość funkcji ma wykonane testy jednostkowe potwierdzające skuteczność implementacji.

1. **Opis poszczególnych kroków**
   1. **Czytanie plików graficznych**

Pierwszym krokiem było wczytywanie pliku i w przypadku, gdy jego wielkość przekraczała na którejś osi 1500 pikseli zostawał on transformowany do mniejszych rozmiarów z zachowaniem proporcji.

* 1. **Konwersja do przestrzeni HSV oraz progowanie**

Wczytany obraz zostawał transformowany do przestrzeni HSV w celu oddzielenia chrominancji od luminacji – dzięki czemu można było wydzielać obszary o jednakowej barwie (Hue). W przyjętym rozwiązaniu pikselom o kolorze H=22 ± 8 i saturacji wyższej niż 124 była przypisywana wartość 1 natomiast pozostałym pikselom wartości 0. W ten sposób otrzymywany był obraz binarny. 

**Porównanie obrazu przed i po progowaniu**

* 1. **Operacje binarne**

Kolejnym krokiem było pozbycie się małych artefaktów pozostających na obrazie. Do tego posłużyła operacja erozji. Aby przywrócić kształt obiektom po operacji została zastosowana operacja dylacji, w efekcie cały proces można nazwać zamykaniem. Na etapie preprocessingu zastosowano również testy z wykorzystaniem filtrów w celu poprawy jakości, ale znacznego efektu nie uzyskano. Do operacji zamykania zastosowano różne rodzaje jądra – najlepiej wypadały o rozmiarach 7x7 o typie rectangle.

Obraz zawierający tekst, roślina

Opis wygenerowany automatycznieObraz zawierający tekst

Opis wygenerowany automatycznie

Obraz zawierający zamknąć

Opis wygenerowany automatycznie

Wycięte przybliżenie na efektem operacji zamykania. Jak widać bez tej operacji podczas segmentacji zostałyby zakalikowane jako dwa oddzielne obiekty podczas gdy po zastosowaniu zamknięcia pozostają jednym. Wszystkie funkcje opisane w podpunkcie b i c znajdują się w module preprocessing.

* 1. **Segmentacja i wyznaczanie konturów**

Do segmentacji został zaimplementowany algorytm Floodfill w formie kolejkowej z lekką modyfikacją – gdy segment ma poniżej ustalonej liczby pikseli jest wcielany do tła. Pozwoliło to na ostateczne pozbycie się małych artefaktów pozostałych po progowaniu. Na wyjściu z z funkcji użytkownik otrzymuje listę znalezionych segmentów oraz obraz z ich oznaczeniem – każdy kolejny segment ma etykietę wyższą o 1. Finalny efekt segmentacji można zobaczyć na obrazku poniżej.



Przed wyliczeniem niezmienników obiektowych należy wyznaczyć również kontur każdego obiektu – do tego został zaimplementowany algorytm Moore Neighborhood przeszukujący sąsiedztwo piksela w poszukiwaniu elementu o zadanej wartości. W efekcie można policzyć obwód danego segmentu oraz zaznaczyć piksele należące do konturu.

* 1. **Klasyfikacja w oparciu o wyznaczenie momentów geometrycznych i niezmienników.**

Ostatnim krokiem była weryfikacja czy dany segment jest logiem czy nie. Do tego stworzono wektor wyliczonych cech na podstawie momentów geometrycznych (9) oraz współczynnik Malinowskiej. Na podstawie stworzonego datasetu (obrócone i zniekształcone binarne obrazy logo) wyznaczono uśrednione wartości, natomiast za dopuszczalne przedziały błędu przyjęto wielokrotność odchylenia standardowego.

Sama klasyfikacja oparta została o modyfikacje k-nearest-neighbour – każda wartość wektor cech była porównywana z wartością odniesienia. W przypadku gdy mieściła się w granicach dodawano do punktacji klasyfikacji 1, natomiast gdy wypadała poza granicę: -1. Dla wyniku klasyfikacji większego niż 7 (co znaczy ze co najmniej 9 cech mieściło się w przedziale) stwierdzano że dany obiekt jest logiem sieci McDonald



1. **Rezultaty i dyskusja.**

Podczas testów potwierdzono skuteczne działanie programu, jak i zauważono 2 poważne problemy:

* Oparcie się na progowaniu na podstawie barwy jest zawodne dla obrazów wykonanych w nocy
* Należałoby zaimplementować lepsze rozwiązanie dotyczące klasyfikacji

Na poniższych obrazach można zaobserwować powyższe efekty:

Obraz zawierający tekst, niebo, zewnętrzne, znak

Opis wygenerowany automatycznieObraz zawierający tekst, niebo, zewnętrzne, zachmurzenie

Opis wygenerowany automatycznie

Poprawnie znalezione logo

Obraz zawierający tekst, drzewo, zewnętrzne, jazda na nartach

Opis wygenerowany automatycznieObraz zawierający tekst, budynek, zewnętrzne, znak

Opis wygenerowany automatycznie

Widać problemy, gdy progowanie wychwytuje również światła.

Obraz zawierający tekst, wewnątrz

Opis wygenerowany automatycznieObraz zawierający tekst, wewnątrz, torba

Opis wygenerowany automatycznie

Obraz zawierający tekst

Opis wygenerowany automatycznie

Po prawej stronie jest obraz wynikowy z detekcją – a raczej jej brakiem. Widać poprawność segmentacji a błąd w klasyfikacji.