

Sprawozdanie

Przetwarzanie obrazów
Projekt – „Rozpoznawanie loga Pepsi”
Paweł Witkowski
14.06.2010

Wstęp

Celem projektu było zaznajomienie się z technikami przetwarzania obrazów począwszy od pozyskania obrazu do rozpoznania konkretnych obiektów na obrazie. Obiektem, który będzie rozpoznawany jest logo Pepsi (rys. 1). Dodatkowo zrealizowano rozpoznawanie loga w obrazach w odcieniach szarości, a także rozpoznawanie kilku log na jednym obrazie.



Rysunek 1. Logo Pepsi

Etapy

Wstępna poprawa jakości obrazów

We wstępnym etapie badany obraz poddawany jest procesowi poprawy kontrastu, dzięki czemu obraz jest wyraźniejszy.

Segmentacja

W zależności od rodzaju obrazu (kolorowy lub w odcieniach szarości) są różne ścieżki postępowania.

Obraz kolorowy:

1. Progowanie
2. Filtry: minimalny, maksymalny, medianowy.

Obraz w odcieniach szarości:

1. Filtr Prewitta.
2. Filtry: minimalny, medianowy.

Progowanie

Obraz progowany jest dla każdego koloru, stanowiącego element loga, oddzielnie (czerwony, niebieski i biały). Wartość piksela ustawiana jest na „1” (tak naprawdę (0, 0, 0), a tło (255, 255, 255)), gdy jedna ze składowych RGB, odpowiadająca danemu kolorowi jest większa od pozostałych o ustaloną wartość. W przypadku koloru białego, gdy wszystkie składowe nie różnią się zbytnio od siebie.

Filtr medianowy

Sprogowany obraz poddawany jest filtracji medianowej, usuwając tym samym losowe zakłócenia. Zastosowana jest tu maska 3x3.

Filtr minimalny i maksymalny

Filtracja minimalna, a następnie maksymalna na sprogowanym obrazie ma na celu usunięcie małych zakłóceń oraz zamknięcie krawędzi.

Filtr Prewitta

Dla obrazów w odcieniach szarości stosowany jest filtr Prewitta, aby odnaleźć krawędzie. Filtracja następuje w czterech kierunkach: północnym, południowym, wschodnim i zachodnim. Po progowaniu obrazów uzyskanych z filtracji w każdym kierunku, obrazy są składane w jeden (dodawane).

Indeksacja

Do indeksacji użyto najpopularniejszej techniki, tzn. przeglądania obrazu piksel po pikselu i, w zależności od sąsiednich pikseli, nadawania etykiet; użyto również tablicy sklejeń.

Ekstrakcja cech

Dla każdego znalezionej obiektu wyznaczane jest 5 cech: współczynnik M_z , współczynnik Malinowskiej oraz momenty M_1 , M_4 , M_7 . Aby nie brać pod uwagę zbyt małych i zbyt dużych elementów, które błędnie zostały wyznaczone na etapie segmentacji, przyjęto, że jeżeli obiekt jest mniejszy niż 0.05% obrazu lub większy niż 60% to jest nieistotny w etapie klasyfikacji. Uzyskano to poprzez ustawienie cech tych obiektów na wartość maksymalną.

Klasyfikator

Do klasyfikacji wykorzystano klasyfikator 1-NN – obiekt należy do tej klasy, do której odległość jest najmniejsza w przestrzeni cech. Każda klasyfikowana próbka, na podstawie zbioru uczącego, otrzymuje jedną z czterech etykiet: czerwony, niebieski, biały albo żaden z wcześniejszych, gdy odległość próbki od każdej ze zbioru uczącego jest większa niż określona (doświadczalnie przyjęto, że wynosi 1).

Końcowe rozpoznanie

Logo uznaje się za znalezione, gdy przynajmniej dwa elementy składowe (czerwony z białym lub czerwony z niebieskim) znajdują się w bliskim sobie otoczeniu (prostokąty opisane na tych elementach), w szczególności nachodzą na siebie.

Klasyfikacja

Zbiór uczący

Zbiór uczący stanowi 45 próbek, po 15 na każdy element loga (czerwony, niebieski, biały), umieszczone w pliku *uczacy.bmp*.

Obrazy testowe

Do testów użyto 23 obrazów kolorowe i 3 w odcieniach szarości w formacie JPEG. Łącznie na wszystkich obrazach znajdowało się 82 loga Pepsi (75 w obrazach kolorowych i 7 w odcieniach szarości).

Uzyskane wyniki

Program sklasyfikował poprawnie 66 log Pepsi (60 kolorowych i 6 w odcieniach szarości). W związku z tym, jakość klasyfikacji wynosi ok. 83%). Na jednym obrazie jeden obiekt, który nie jest logiem, został sklasyfikowany pozytywnie.

Komentarz

Uważam, że uzyskane wyniki są bardzo dobre, biorąc pod uwagę, że niektóre loga były tak małe lub zdeformowane, że nawet człowiek miałby trudność z ich rozpoznaniem, jak również wielkość zbioru uczącego, który jest ponad 5-krotnie mniejszy od zbioru testowego (15 próbek zbioru uczącego jest zdecydowanie za mało, żeby uzyskać lepsze rezultaty).