

# Projekt z przedmiotu POBR – sprawozdanie wykrywanie logo Pepsi

Sławomir Zieliński

## Wstęp

W projekcie rozważano problem wyszukiwania logo produktu Pepsi na rzeczywistych, kolorowych obrazach pozyskanych własnoręcznie lub z internetu. Cała baza liczyła 13 obrazów.

## Budowa

Napisany program ma następującą budowę:

- Na samym początku obraz (jeśli tego wymaga) jest zmniejszany do wymiarów takich, aby dalej był czytelny, jednak, aby program przetwarzający nie wykonywał się zbyt długo.
- Następnie wykonywany jest filtr kolorów, wyodrębniający z obrazu kolory występujące w logo produktu Pepsi, a więc czerwony, biały i niebieski. Wynik zwracany jest w postaci trzech masek.
- Następnie wykonywana jest funkcja segmentująca, która dla maski koloru czerwonego grupuje piksele w obszary – te piksele, które leżą bezpośrednio obok siebie będą należeć do jednego obszaru.
- Następnie wykonywana jest warstwa selekcji. Dla wyodrębnionych wcześniej obszarów badane są niezmienniki kształtu M1 oraz M7. Wykluczane są te obszary, które nie spełniają założonych wartości.
- W tym momencie wykonywana jest faktyczna funkcja szukania logo produktu Pepsi. Szczegółowo opisana zostanie w dalszej części tego dokumentu.
- Na obrazie zaznaczane są prostokąty, w których wykryto szukany wzór, obraz jest wyświetlany.

## Opis działania

### Wyszukiwanie kolorów

W tej dziedzinie testowane były trzy metody.

1. Szukanie pikseli o każdej składowej RGB w zadanym przedziale. Np. składowa R między 150 a 255, G – 0 a 120, B – 0 a 120
2. Szukanie pikseli spełniających zadane kryteria, określone w sposób bardziej opisowy np. barwą czerwoną jest piksel, którego składowa R ma wartość co najmniej 80 i jest większa od pozostałych składowych o co najmniej 20. Barwa biała to taka, gdzie wszystkie składowe różnią się maksymalnie o 50 a ich suma jest większa niż 300
3. Szukanie pikseli będących w zadanym zakresie po transformacji do modelu HSV.

Na tym etapie zdecydowano się na dość szeroki wybór barw, gdyż wyniki będą w każdym następnym kroku ograniczane, więc błędem było by ich nadmierne ograniczenie już na początku.

Z przeprowadzonych testów wynika, że najlepiej zadziałał analityczny opis barw (2). Na drugim miejscu szukanie przedziałami po transformacji do modelu HSV.

## Segmentacja

Celem tej warstwy jest zwrócenie informacji które piksele tworzą spójny obszar, a które nie. Wykonywane jest to w następujący sposób. Obraz jest przeszukiwany piksel po pikselu. Gdy natrafi na piksel biały, wykonywana jest funkcja poszukująca obszaru. Piksel ten zostaje dodany do bieżącego obszaru. Zaczynając od tego piksela funkcja iteracyjnie sprawdza sąsiednie piksele dla każdego piksela dodanego już do obszaru. Jeśli sąsiedni piksel również jest biały, dodawany jest do bieżącego obszaru i oznaczany jako szary (sprawdzony).

Podjęto próbę napisania tej funkcji w sposób rekurencyjny, nie iteracyjny, jednak w większości obrazów natrafiono na błąd przepełnienia stosu, z powodu wywołania się zbyt wielu funkcji. Dla każdego piksela w obszarze wykonywana była funkcja.

## Selekcja

Na tym poziomie dla wybranych wcześniej obszarów wybierane są te, które spełniają zadane wartości niezmienników M1 i M7. Wzorcowe przedziały zostały wyznaczone doświadczalnie. Dla czerwonej części logo są to wartości z przedziału 0,15-0,3 dla M1 oraz 0,006 – 0,007 dla M7.

Podjęto próbę zwiększenia tego przedziału dla M7 na 0,006-0,0075, w tym przypadku większy procent log był wykrywany, jednak wzrosła też liczba obszarów niesłusznie uznanych za logo.

Wprowadzenie tej warstwy przetwarzania pozwoliło wykluczyć czerwone obszary, które nie przypominały w żaden sposób kształtu podobnego do półkola. Na poniższym obrazie w środkowej części zostało wykryte logo, z powodu wystąpienia dużego czerwonego obszaru. Jednak obszar ten ma w sobie czarne luki. Warstwa selekcji pozwoliła go wykluczyć.



## Szukanie logo

Do tego poziomu programu dostały się te obszary, które mają kształt przypominający półkole i czerwoną barwę. Dla każdego takiego obszaru szukany jest prostokąt o wymiarach dwukrotnie większych niż prostokąt ograniczający. Dla tego obszaru zliczane są wystąpienia koloru czerwonego, białego i niebieskiego. Jeśli wszystkie te kolory występują i ich liczność stanowi co najmniej 1/6 liczności pikseli czerwonych dla niebieskiego i 1/9 dla białego, obszar ten klasyfikowany jest jako logo Pepsi.

W tym zakresie przeprowadzono próby udoskonalenia algorytmu. Było to na przykład ograniczenie górnej wartości liczności pikseli danego koloru. Ta metoda nie okazała się skuteczna, gdyż butelki z logiem bardzo często są obrócone w ten sposób, że obszar czerwony jest dość mały, co sztucznie ogranicza pozostałe obszary. Jednocześnie mogą występować przebłyski i odbicia w jasnym kolorze, które zwiększać będą licznosc pikseli koloru białego. Butelki napoju Pepsi bardzo często są w niebieskim kolorze, a co za tym idzie, logo często występuje na niebieskim tle, w związku z czym ograniczenie tej barwy od góry mogłoby powodować mylne pominięcie obszaru jako logo.

Kolejną próbą było szukanie obszarów barw niebieskiej i białej oraz liczenie ich niezmienników. Jednak prawdopodobnie z tych samych przyczyn skutkowało to pomijaniem zbyt dużej ilości obszarów. Kolor niebieski z logo zlewał się z kolorem tła i był wykrywany jako jeden obszar, którego kształt nie przypominał półkola, więc obszar ten niesłusznie nie był brany pod uwagę. Dodatkowo, na logo Pepsi bardzo często nałożony jest obrazek owocu, wskazujący na przykład na limonkowy smak napoju. Owoc ten przysłania właśnie niebieski obszar logo, utrudniając tym samym jego właściwe rozpoznanie. W przypadku koloru białego występowanie przebłysków komplikowało tu sprawę, ale również biała obwódka wokół logo, która nie zawsze była widoczna (z powodu obrócenia butelki lub gdy logo było zbyt małe).





Na powyższych obrazach pokazane są odpowiednio wyniki dla liczenia niezmienników jedynie czerwonego koloru, czerwonego i białego oraz wszystkich. Widać na nich, że obliczanie niezmienników jedynie czerwonego koloru przynosi lepsze rezultaty, jednak pojawiają się również obszary wykryte błędnie w lewej części zdjęcia na cenach. Wynika to jednak prawdopodobnie z uznania koloru żółtego jako czerwonego, a więc nie jest to wina tej warstwy, a niedoskonałego dobrania przedziałów wartości składowych koloru czerwonego przy wyszukiwaniu barw.

Wnioskiem, ale również zaskoczeniem płynącym z testowanych sposobów, jest fakt, że metoda z pozoru prymitywna, niedoskonała i narażona na wiele błędów okazała się najlepsza. Może to wynikać z faktu, że w rzeczywistości tak skrajnie różne kolory jak w logo Pepsi, występują blisko siebie dość rzadko, więc przybliżenie do liczności pikseli odpowiedniej barwy okazuje się nie najgorszym pomysłem.

### Pozostałe testy

Testowane było również zastosowanie filtra dolnoprzepustowego przed zastosowaniem wykrywania kolorów. Rozwiązanie to nie przynosiło żadnych efektów a podnosiło czas wykonania, w związku z czym zrezygnowano z tego pomysłu.