

Laboratorium 2

Analiza Semantyczna Obrazu i Dźwięku

autor: dr inż. Jacek Naruniec

Cel laboratorium:

Zapoznanie się z transformatą Hougha w zadaniu normalizacji rotacji zeskanowanej strony tekstu.

Zadania

Stwórz aplikację która pozwoli:

- dla wczytanego obrazu odfiltrować niepożądane obszary mogące zakłócić wynik transformaty Hougha,
- stworzyć transformatę Hougha dla przefiltrowanego obrazu,
- wyznaczyć linię dominującą tekstu,
- obrócić obraz tak, aby tekst był wyrównany przy pomocy wyznaczonej linii,
- wyświetlić linię dominującą, obrazu oryginalnego i obrazu obróconego,
- przy każdej transformacji zapisać tablicę Hougha do pliku graficznego (np. hough.pgm) tak, aby można było obejrzeć jej zawartość.

Uwagi szczegółowe

- Za pomocą dowolnej metody należy odfiltrować wszystkie niepożądane elementy obrazu mogące zakłócić wynik transformaty Hougha (rysunek obok). Etap ten można zrealizować poprzez usunięcie obszarów spójnych o dużej ilości elementów składowych, filtr krawędziowy lub poprzez dowolny, samodzielnie zdefiniowany filtr.
- Transformatę Hougha tworzymy w sposób przedstawiony na wykładzie. Wystarczy wykonać jednostopniową transformatę o odpowiednio dobranych parametrach.
- Linię dominującą można wyznaczyć poprzez wybranie maksymalnego elementu transformaty. Jeśli wyniki takiej analizy byłyby niezadawalające (tzn. nieskuteczne dla większości obrazów) należy zaproponować inną metodę wyznaczania linii dominującej, np. poprzez wybranie kierunku dla którego występuje największa ilość elementów maksymalnych.



- d. Obrótu tekstu dokonujemy poprzez macierz obrotu o określony kąt α względem punktu (x_0, y_0) :

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

Uwagi do aplikacji

Załączona aplikacja umożliwia odczyt pliku w formacie PGM. Obrazy o tym rozszerzeniu obsługuje także m.in. program IrfanView. Po wczytaniu pliku program umożliwia dwie opcje przetwarzania obrazu, które należy zaimplementować.

W aplikacji należy zmodyfikować klasę DominantLineProcessing tak, aby wykonywała założone operacje na obrazie. Główne funkcje do implementacji to:

void HoughDominantLine();

, czyli metoda wyznaczająca parametry linii dominującej. Wynik analizy powinien być wpisany do zmiennej klasy: dominant_line_angle. Po wywołaniu tej funkcji z menu programu („Przetwarzanie”->”Hough–wyznaczanie parametrów linii”), linia opisana kątem dominant_line_angle przechodząca przez środkowy punkt obrazu zostanie narysowana w obrazie.

Dostęp do aktualnych pikseli obrazu odbywa się poprzez zmienną image znajdującą się w deklaracji klasy.

W funkcji powinna znaleźć się wstępna filtracja obrazu, wyznaczenie transformaty Hougha i wyznaczenie linii dominującej.

void RotatImage();

Ta metoda powinna dokonywać obrotu aktualnego obrazu (image) względem wyznaczonego kierunku dominującego (dominant_line_angle). Wynik powinien być wpisany z powrotem do obrazu image.

W aplikacji mogą być użyteczne funkcje writePGM i readPGM służące zapisywaniu i odczytywaniu tablicy bajtów do plików PGM. Można je wykorzystać do podglądu pośrednich wyników przetwarzania (np. obrazu przefiltrowanego) lub zapisywania aktualnej tablicy Hougha.