# Inżynieria obrazów

Programowe przetwarzanie obrazów

Marcin Lasak 272886

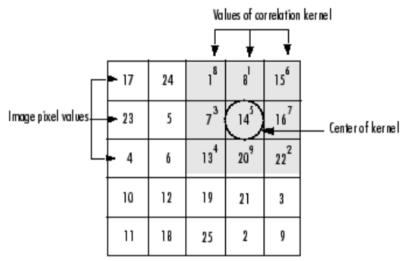
## Spis treści

1. Wstęp teoretyczny	3
1.1 Filtracja obrazów	3
1.2 RGB	4
1.3 YCbCr – model luminancja-chrominancja	4
1.4 DVB	4
1.5 Błąd średniokwadratowy	5
2. Zadania	6
2.1 Filtr górnoprzepustowy	6
2.2 Przekształcenia kolorów	
2.3 Konwersja z RGB do YCbCr	9
2.3 Symulacja transmisji DVB	12
2.5 Błąd średniokwadratowy	16
3. Zmiana struktury programów	17
Źródła	18

#### 1. Wstęp teoretyczny

#### 1.1 Filtracja obrazów

Jest to przekształcanie obrazu, w którym wartości pikseli wyjściowych są wynikami operacji matematycznych na pikselach sąsiednich. Piksele są traktowane jak sygnały, analogicznie do metod cyfrowego przetwarzania sygnałów zadanych funkcjami (filtrowanie np. sygnałów dźwiękowych).



Rysunek 1

Przy zadanej masce (rysunek 1<sup>1</sup>), obliczamy nową wartość piksela, poprzez obliczenie sumy ważonej składowej punktu i wszystkich sąsiadów zgodnie z wagami wskazanymi przez maskę filtra. Następnie tak otrzymana suma jest dzielona przez sumę wag.

W kontekście obrazów, częstotliwością nazywamy częstotliwość zmian wartości pikseli w obrazie, które są interpretowane w kontekście przestrzennym. Tj. o niskiej częstotliwości mówimy w przypadku wolnych zmian w obrazie; wartości pikseli zmieniają się w sposób łagodny i stopniowy. Możemy uprościć to zagadnienie do ogólnych kształtów, bez szczegółów. Natomiast o wysokiej częstotliwości mówimy, gdy zmiany wartości pikseli są nagłe, gwałtowne. Są to ostre krawędzie, detale, tekstura. To zagadnienie możemy uprościć do wszelkich szczegółów obrazu.

Tak więc filtry dolnoprzepustowe będą szczegóły z obrazu usuwać (rysunek 2<sup>2</sup>), a górnoprzepustowe – je uwypuklać.



Rysunek 2

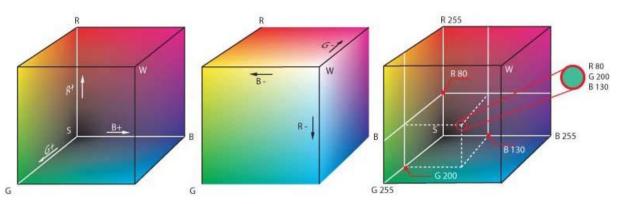
3

<sup>&</sup>lt;sup>1</sup> https://multimed.org/student/pdio/pdio09 filtracja obrazu poprawa jakosci tekstu.pdf

<sup>&</sup>lt;sup>2</sup> http://www.algorytm.org/przetwarzanie-obrazow/filtrowanie-obrazow.html

#### **1.2 RGB**

RGB (Red Green Blue) to model przestrzeni barw, opisywanej trzema współrzędnymi: Red – czerwień, Green – zieleń, Blue – niebieski. Każda współrzędna zapisywana jest na ośmiu bitach; w zakresie 0-255. Wartość 0 oznacza brak koloru, natomiast 255 jego pełną intensywność. Rysunek 3<sup>3</sup> przedstawia wizualizacje przestrzeni RGB.



Rysunek 3

System RGB posiada również wariant zmiennoprzecinkowy, gdzie wartości współrzędnych przyjmują wartości z zakresu 0-1.

#### 1.3 YCbCr – model luminancja-chrominancja

YCbCr to model przestrzeni kolorów bazującą na modelu RGB, w której współrzędne przechowują informacje o luminancji – mierze natężenia światła, oraz chrominancji składowej odpowiedzialnej za odcień i nasycenie koloru. Przestrzeń luminancjachrominancja, tak jak przestrzeń RGB, zadana jest trzema współrzędnymi:

- Y składowa luminancji; obraz w czerni i bieli odpowiadający za natężenie światła;
- Cb składowa różnicowa chrominancji odpowiadająca różnicy między luminancją, a kolorem niebieskim.
- Cr składowa różnicowa chrominancji odpowiadająca różnicy między luminancją, a kolorem czerwonym.

Nałożenie takich trzech macierzy pikseli pozwala na uzyskanie pełnobarwnego obrazu.

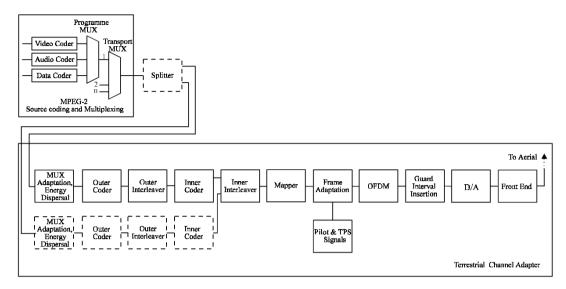
Model ten jest wykorzystywany przy kompresjach obrazu, ponieważ zmniejszanie rozdzielczości składowych chrominancji, przy zachowaniu wejściowej składowej luminancji, zmniejsza rozmiar obrazu w pamięci, zachowując pozornie swoją pierwotną jakość. Wynika to z faktu, iż ludzkie oko jest dużo bardziej wrażliwe na natężenie światłą, niż na kolor.

#### **1.4 DVB**

DVB (Digital Video Broadcasting) to zbiór standardów transmisji telewizji cyfrowej. Przy takiej transmisji wykorzystywany jest model luminancja-chrominancja, ponieważ pozwala na kompresję obrazów, co jest konieczne aby efektywnie wykorzystać dostępne pasmo. Po redukcji chrominancji obrazy kompresowane sa do formatu MPEG. Rysunek 4<sup>4</sup> przedstawia schemat modelu transmisji DVB.

<sup>&</sup>lt;sup>3</sup> https://pl.wikipedia.org/wiki/RGB

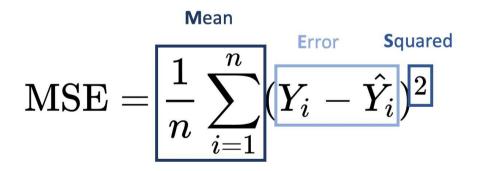
<sup>4</sup> https://www.etsi.org/deliver/etsi\_en/300700\_300799/300744/01.01.02\_60/en\_300744v010102p.pdf



Rysunek 4

#### 1.5 Błąd średniokwadratowy

Błąd średniokwadratowy (Mean Squared Error) jest wartością oczekiwaną kwadratu błędu między estymatorem, a wartością estymowaną (rysunek 5<sup>5</sup>). W przypadku przetwarzania obrazów jest to miara różnicy między dwoma obrazami (np. oryginałem, a obrazem po kompresji lub rekonstrukcji). Oblicza się go jako średnią kwadratów różnic między odpowiadającymi sobie pikselami. Im niższa wartość MSE, tym lepsza jest jakość przetwarzania; obraz wynikowy jest bliższy oryginałowi. Jeśli wartość MSE wynosi 0, oznacza to, że obrazy są identyczne.



Rysunek 5

<sup>&</sup>lt;sup>5</sup> <u>https://suboptimal.wiki/explanation/mse/</u>

#### 2. Zadania

#### 2.1 Filtr górnoprzepustowy

W tym zadaniu należało wykonać filtr górnoprzepustowy, tzw. detektor krawędzi, o zadanej masce. Zadanie wykonano w języku Python. Skorzystano z bibliotek numpy, do właściwego przetwarzania tablic reprezentujących obraz w przestrzeni RGB oraz Pillow, do operacji na plikach graficznych (odczyt, zapis).

Filtracja odbywać się będzie przy pomocy zaimplementowanej funkcji hpfilter(), przyjmującą jako argument nazwę pliku graficznego poddawanego filtracji.

Funkcja ta wpierw otwiera zadany plik (rysunek 6), a następnie konwertuje go do modelu RGB, by następnie zapisać go w tablicy numpy. Dalej zapisujemy podaną maskę jako tablicę numpy, by móc ją wykorzystywać w późniejszych obliczeniach. Zapisujemy także wymiary maski oraz rozmiary wyrównania (padding). Tworzymy także wyjściową tablicę, wstępnie wypełniona zerami.

Rysunek 6

Przed obliczeniem nowych wartości pikseli (rysunek 7), każdy kanał jest wyrównywany: dodawane są piksele, aby poprawnie przetworzyć skrajne punkty obrazu. Zastosowano wypełnienie z odbiciem, zamiast dopełniania zerami.

Następnie obliczana jest nowa wartość danego kanału dla każdego piksela obrazu. Taki krok powtarzany jest dla każdego kanału, w rezultacie czego otrzymujemy tablicę z przetworzonymi współrzędnymi RGB.

Rysunek 7

Finalnie normalizujemy wartości do zakresu 0-255 (ponieważ obliczone wartości mogły wyjść za ten zakres) i zapisujemy jako typ uint8 (liczba całkowita zapisana na 8 bitach).

Do sprawdzenia wyników filtrowania korzystamy z funkcji pomocniczych save() oraz show\_image() (rysunek 8).

Rysunek 8

Program Python lab1zad1.py i korespondujący z nim plik wykonywalny, wyświetlają najpierw obraz input1.jpg, przetwarzają go i zapisują, a następnie wyświetlają przetworzony obraz. Obraz wejściowy dany jest rysunkiem nr 9, natomiast wyjściowy rysunkiem nr 10.



Rysunek 9



Rysunek 10

#### 2.2 Przekształcenia kolorów

Kolejnym zadaniem, było przekształcenie kolorów obrazu z wykorzystaniem konwersji na zmiennoprzecinkowy format RGB, z wykorzystaniem wzoru z rysunku nr 11.

$$\begin{bmatrix} R_{new} \\ G_{new} \\ B_{new} \end{bmatrix} = \begin{bmatrix} 0.393 & 0.769 & 0.189 \\ 0.349 & 0.689 & 0.168 \\ 0.272 & 0.534 & 0.131 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Rysunek 11

W tym zadaniu również korzystamy z bibliotek numpy oraz Pillow. W tym programie zdecydowano się na rozdzielenie etapu odczytu pliku graficznego z konwersją na tablicę numpy oraz właściwego przetwarzania obrazu. Dlatego też zaimplementowano funkcje file\_to\_array(), odpowiedzialną za odczyt i konwersję obrazu, oraz inttofloat(), przetwarzającą obraz w formacie zmiennoprzecinkowym (rysunek 12).

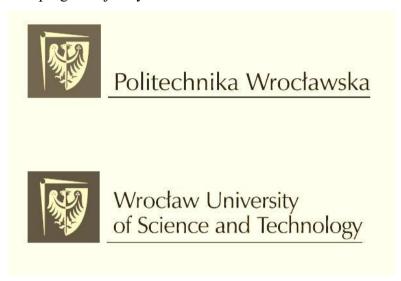
Rysunek 12

W funkcji inttofloat(), wpierw tworzymy tablicę numpy wypełnioną zerami. Co ważne, używanym typem danych w tablicy jest float32 (trzydziestodwubitowa liczba zmiennoprzecinkowa). Następnie w pętli dokonujemy konwersji każdego piksela na format zmiennoprzecinkowy oraz jego przekształcenia wg wzoru z rysunku 11. Wpierw obliczamy współrzędne RGB w nowym formacie, a następnie dokonujemy przekształceń kolorów. Tak

otrzymane dane zapisujemy w tablicy numpy, z uwzględnieniem możliwego przekroczenia wartości 1.

W celu sprawdzenia działania tej funkcji, przetwarzamy obraz z zadania pierwszego, zapisujemy go i wyświetlamy. W tym celu również korzystamy z ww. funkcji pomocniczych, z tym że funkcja zapisu została zmodyfikowana, aby przed zapisem format zmiennoprzecinkowy został zamieniony na całkowity (uint8). Takie operacje wykonuje również korespondujący z programem Python lab1zad2.py plik wykonywalny.

Wynikiem działania programu jest rysunek nr 13.



Rysunek 13

#### 2.3 Konwersja z RGB do YCbCr

W tym punkcie, należało zaimplementować konwersję obrazu z przestrzeni RGB, do przestrzeni YCbCr wg wzoru danego rysunkiem nr 14.

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.229 & 0.587 & 0.114 \\ 0.500 & -0.418 & -0.082 \\ -0.168 & -0.331 & 0.500 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Rysunek 14

Ponownie skorzystano z bibliotek numpy i Pillow oraz ww. funkcji pomocniczych. Dodatkowo użyto biblioteki openCV do konwersji z YCbCr do RGB.

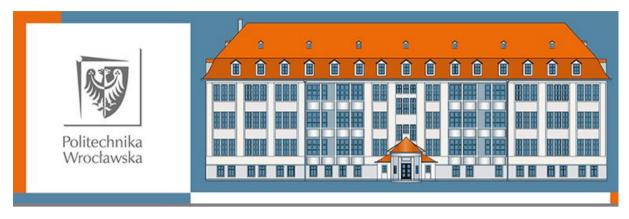
W celu konwersji tablicy numpy zawierającą współrzędne RGB, do tablicy zawierającej współrzędne YCbCr zaimplementowano funkcję RGBtoYCBCR() daną rysunkiem nr 15. Tworzymy w niej nową tablicę numpy wypełnioną zerami, a następnie w zagnieżdżonej pętli przechodzimy po każdym pikselu obrazu w modelu RGB, obliczamy współrzędne w modelu YCbCr i zapisujemy w odpowiednim miejscu tablicy. Współrzędne YCbCr są zaokrąglane tak, aby nie przekraczały wartości 1.

Rysunek 15

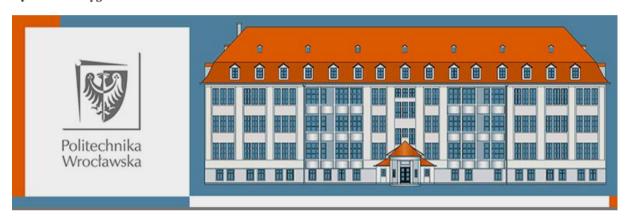
W celu sprawdzenia poprawności działania tej konwersji, zaimplementowano funkcję przeprowadzającą konwersję odwrotną i zapisującą tak przetworzony obraz oraz funkcję zapisującą każdą składową obrazu w modelu YCbCr w odcieniach szarości (rysunek 16).

Rysunek 16

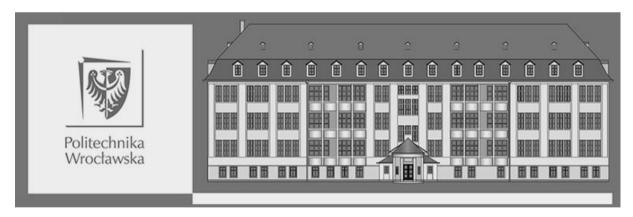
Program wykonuje więc konwersję z RGB do YCbCr oraz konwersję odwrotną, zapisuje wynik tej operacji i składowe powstałe przy pierwszej konwersji. Obraz wejściowy dany jest rysunkiem nr 17. Wyniki działania programu dane są rysunkami nr 18, 19, 20, 21.



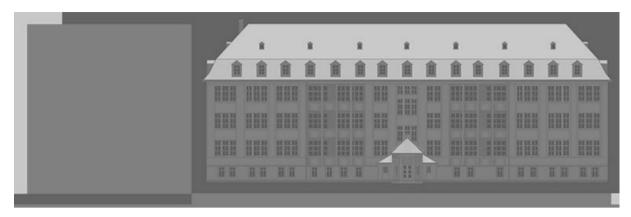
Rysunek 17 Oryginał



Rysunek 18 Przetworzony obraz



Rysunek 19 Składowa luminancji – Y



Rysunek 20 Składowa chrominancji – Cb



Rysunek 21 Składowa chrominancji – Cr

#### 2.3 Symulacja transmisji DVB

W tym zadaniu należało zasymulować transmisję obrazu w systemie DVB. Symulowana będzie część kodowania i dekodowania sygnału. Symulację przeprowadzono w języku Python, z wykorzystaniem bibliotek numpy, Pillow oraz openCV.

W tym celu na pliku graficznym dokonujemy operację konwersji z modelu RGB na model YCbCr. Następnie przeprowadzamy operacje podpróbkowania kanałów chrominancji (po stronie nadawcy) oraz nadpróbkowania (po stronie odbiorcy). Po tym składowe są łączone w jeden obraz i konwertowane z powrotem do modelu RGB.

Do konwersji z modelu RGB do YCbCr wykorzystano funkcję z zadania 2 (rysunek 15).

W celu przeprowadzenia konwersji odwrotnej, zaimplementowano funkcję YCBCRtoRGB() daną rysunkiem nr 22.

W poprzednim zadaniu konwersja ta była przeprowadzana przy użyciu narzędzi biblioteki openCV. W tym przypadku zdecydowano się na ręczną implementację tej konwersji, co znacznie wydłuża działanie programu. Jest to spowodowane dużo lepszą optymalizacją narzędzi biblioteki openCV.

Rysunek 22

Następnie zaimplementowano funkcje wykonujące operacje podpróbkowania i nadpróbkowania dane rysunkiem nr 23.

Rysunek 23

Funkcja downsample() wykonuje podpróbkowanie kanałów chrominancji. Operacja ta polega na zmniejszeniu rozdzielczości tablicy składowej chrominancji, przez pobieranie tylko współrzędnych znajdujących się w co drugim wierszu w co drugiej kolumnie (elementów macierzy  $a_{ij}$ , gdzie  $i=2k, j=2l, i, j \in \{0,1,...,\frac{liczba wierszy/kolumn}{2}\}$ ).

Funkcja upsample() wykonuje operację nadpróbkowania kanałów chrominancji. Jest to czynność dużo bardziej złożona, wykorzystującą interpolację. Z tego powodu korzystamy z narzędzia biblioteki openCV. Funkcja resize pozwala na wypełnienie brakujących punktów z wykorzystaniem interpolacji liniowej. Istnieje możliwość wykorzystania innych rodzajów interpolacji, które dały by lepszy wynik, lecz postanowiono wykorzystać najprostsze narzędzie, aby zobaczyć jakie negatywne efekty może za sobą nieść nadpróbkowanie.

Finalnie zaimplementowano funkcję process(), wykonującą kolejne operacje symulujące transmisję DVB (rysunek 24).

#### Rysunek 24

Wynikiem działania programu jest "przesłany" obraz oraz jego składowe w modelu YCbCr w odcieniach szarości.

Obraz wejściowy dany jest rysunkiem nr 9. Obrazy wyjściowe dane są rysunkami 25, 26, 27 i 28.



Rysunek 25 Obraz po transmisji



Rysunek 26 Składowa luminancji – Y



Rysunek 27 Składowa chrominancji – Cb



Rysunek 28 Składowa chrominancji – Cr

#### 2.5 Błąd średniokwadratowy

Ostatnim zadaniem było zaimplementowanie obliczanie błędu średniokwadratowego pomiędzy dwoma obrazami. Biblioteka numpy pozwala nam na sprawne działania na macierzach, więc z jej wykorzystaniem stworzono funkcję calculate\_mse() daną rysunkiem nr 29.

Rysunek 29

Obiekty image1 i image2 to tablice numpy, które możemy swobodnie odejmować. Następnie podnosimy do kwadratu każdy element różnicy i obliczamy z nich średnią. W ten sposób otrzymujemy błąd średniokwadratowy między dwoma obrazami.

Następnie zaimplementowano funkcję process() (rysunek 30), konwertującą dwa obrazy do tablic numpy i obliczającą błąd średniokwadratowy.

Rysunek 30

Wynik działania programu dany jest rysunkiem nr 31.

```
C:\Users\marci\PycharmProjects\IOLab1\.venv\Scripts\python.exe C:\Users\marci\PycharmProjects\IOLab1\lab1zad5.py
Mean Squared Error (MSE): 80.93746994338427

Process finished with exit code 0
```

Rysunek 31

#### 3. Zmiana struktury programów

Zgodnie z zaleceniem prowadzącego, przekształcono programy dedykowane każdemu zadaniu, w jeden zbiorczy program, umożliwiający uruchomienie wybranego zadania. Stąd w folderze dist, brak wspominanych plików wykonywalnych lab1zad1, lab1zad2 itd. Zamiast nich, znajduje się tam plik lab1.exe, umożliwiający wykonanie każdego z zadań. Co ważne, kody w języku Python zostały odpowiednio zakomentowane, aby umożliwić ich proste importy, więc same w sobie nie wykonują żadnej czynności. W przypadku chęci ręcznego uruchomienia kodów w Pythonie, należy odkomentować odpowiednie ostatnie linie.

Program lab1 działa w ciągłej pętli. Wybór zadania odbywa się przez podanie cyfry 1-5, natomiast zamknięcie programu możliwe jest przez podanie na wejściu znaku "q". W folderze dist znajdują się wszystkie pliki potrzebne do poprawnego działania programu.

Plik lab1.exe został wygenerowany przy użyciu pyInstaller na podstawie kodu Python lab1.py.

### Źródła

https://fajne.studio/przestrzen-barwna-rgb-i-cmyk/

https://pl.wikipedia.org/wiki/Digital\_Video\_Broadcasting

https://multimed.org/student/pdio/pdio09\_filtracja\_obrazu\_poprawa\_jakosci\_tekstu.pdf

http://www.algorytm.org/przetwarzanie-obrazow/filtrowanie-obrazow.html

https://pl.wikipedia.org/wiki/RGB

 $\frac{\text{https://www.etsi.org/deliver/etsi\_en/300700\_300799/300744/01.01.02\_60/en\_300744v01010}{2p.pdf}$ 

https://suboptimal.wiki/explanation/mse/

https://pwr.edu.pl/uczelnia/informacje-ogolne/materialy-promocyjne/logotyp

 $\underline{https://wmat.pwr.edu.pl/wydarzenia/uroczyste-otwarcie-budynku-wydzialu-matematyki-173.html}$