

POLITECHNIKA ŚLĄSKA w Gliwicach

WYDZIAŁ INŻYNIERII BIOMEDYCZNEJ

Katedra Informatyki i Aparatury Medycznej

LABORATORIUM PRZETWARZANIA OBRAZÓW MEDYCZNYCH

Transformata Fouriera

Opracowanie: dr inż. Sylwia Pośpiech-Kurkowska, dr inż. Paweł Badura

1 Cel ćwiczenia

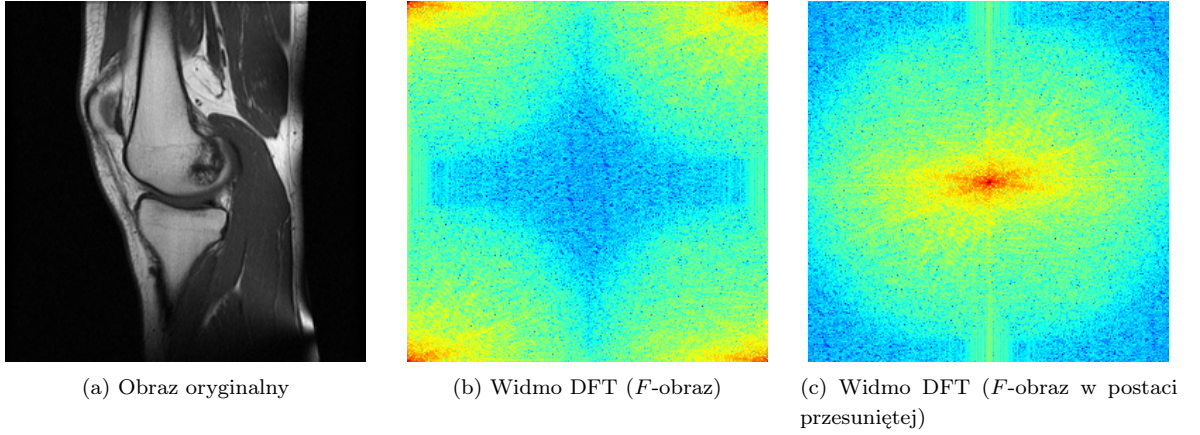
Celem ćwiczenia jest zapoznanie studentów z dwuwymiarową dyskretną transformacją Fouriera i dyskretną transformacją kosinusową oraz ich własnościami i zastosowaniami w przetwarzaniu obrazów przy wykorzystaniu środowiska Matlab.

2 Wprowadzenie

2.1 Transformacja Fouriera

Transformacja Fouriera jest narzędziem pozwalającym na przedstawienie dowolnej funkcji okresowej w postaci szeregu funkcji sinus i cosinus. Stanowi ona podstawę analizy spektralnej (częstotliwościowej) w analizie sygnałów. Transformacja Fouriera jest szczególnie przydatnym narzędziem do analizy procesów zależnych od czasu, takich jak sygnały radiowe, mowa, drgania. W komputerowym przetwarzaniu i analizie obrazów najczęściej korzysta się z transformacji Fouriera w następujących celach:

- filtracja w dziedzinie częstotliwości, czyli wzmocnienie lub usunięcie pewnych składowych odpowiadających szumom, następnie odtworzenie obrazu skorygowanego,
- wydzielenie cech lepiej widocznych w dziedzinie częstotliwości (okresowość, orientacja, rozmycie),
- szybkie obliczanie splotu.



Rys. 1: Ilustracja dyskretnej transformacji Fouriera obrazu. Widma wyświetlono w postaci logarytmicznej

Transformacja Fouriera przenosi obraz do dziedziny częstotliwości. Jednak ta częstotliwość ma interpretację inną, nie tak naturalną jak przy przetwarzaniu sygnałów zmiennych w czasie. Różne składowe częstotliwościowe widma (spektrum) obrazu odpowiadają za różnice w kontraście: wysoki kontrast zwiększa wartości dla wysokich częstotliwości, niski kontrast znajduje odzwierciedlenie w składowych o niskich częstotliwościach. Krawędzie rozdzielające powierzchnie o jednolitych barwach to obszary o wysokiej częstotliwości zmian intensywności; można zatem założyć, że usunięcie z widma obrazu składowych odpowiadających wysokim częstotliwościom spowoduje wygładzenie obrazu.

Zarówno transformacja Fouriera, jak i jej odmiana — transformacja kosinusowa — są przekształceniami globalnymi, czyli wartość piksela w obrazie wynikowym zależy od wartości wszystkich pikseli obrazu pierwotnego. Oznacza to, że transformata Fouriera obrazu nie pokazuje informacji o lokalizacji, natomiast przydaje się do analizy okresowych regularności czy orientacji obiektów.

2.2 Dyskretna transformacja Fouriera

Dyskretna transformacja Fouriera (DFT, ang. *discrete Fourier transform*) obrazu jest z matematycznego punktu widzenia rozszerzeniem na dwa wymiary podstawowych zależności jednowymiarowej DFT. Obraz cyfrowy I o wymiarach $M \times N$ pikseli można traktować jako uporządkowany i ponumerowany zbiór liczb rzeczywistych, co zapisuje się symbolicznie:

$$I = \{I(m, n) \in \mathbf{R} : m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1\}$$

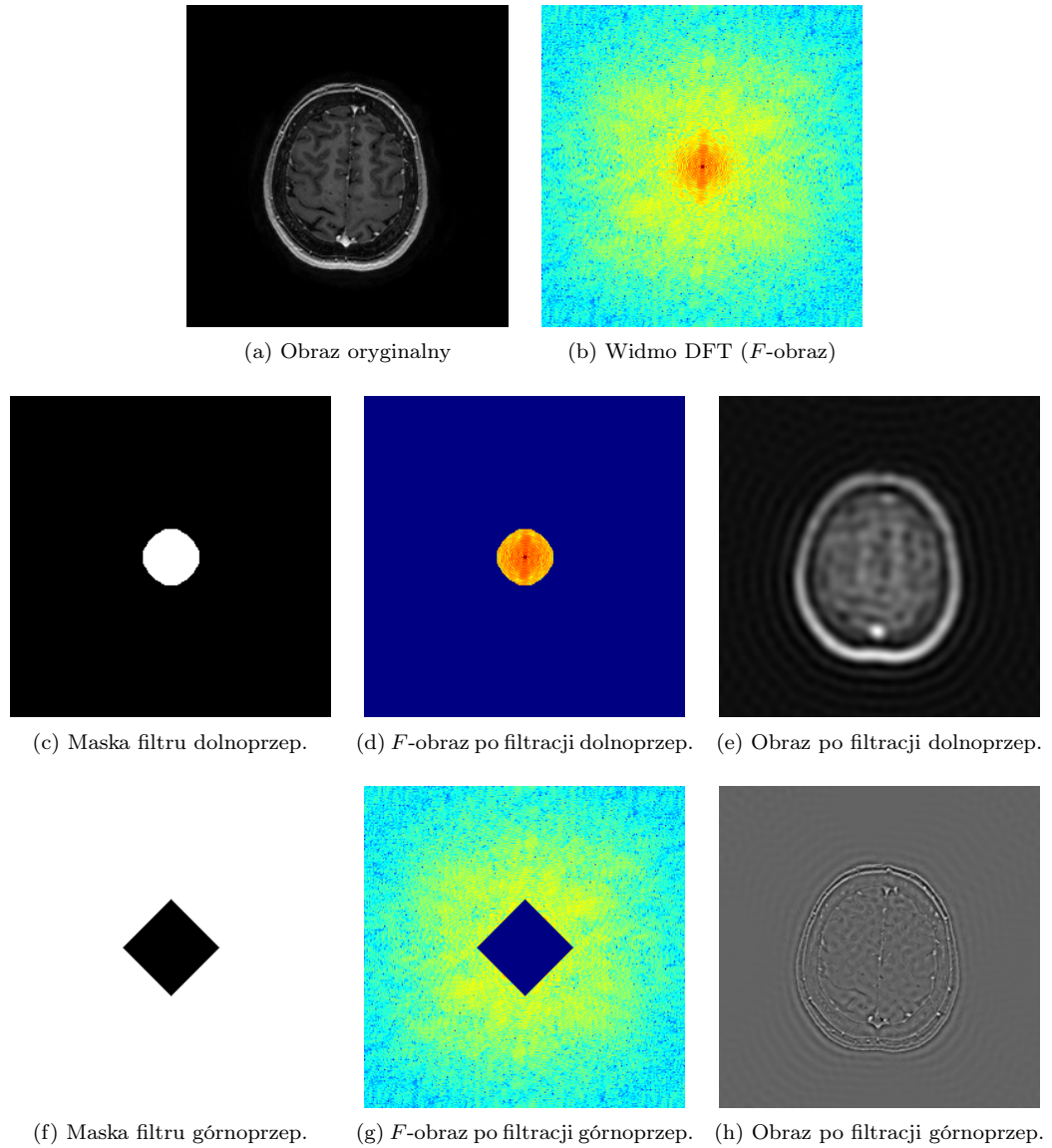
Transformację Fouriera takiego obrazu oraz odwrotną transformację można zapisać w postaci:

$$F(u, v) = \sum_{m=0}^{M-1} \left(\sum_{n=0}^{N-1} I(m, n) e^{-j \frac{2\pi v n}{N}} \right) e^{-j \frac{2\pi u m}{M}} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) e^{-j 2\pi \left(\frac{v n}{N} + \frac{u m}{M} \right)}, \quad (1)$$

$$I(m, n) = \frac{1}{MN} \sum_{u=0}^{M-1} \left(\sum_{v=0}^{N-1} F(u, v) e^{j \frac{2\pi v n}{N}} \right) e^{j \frac{2\pi u m}{M}} = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j 2\pi \left(\frac{v n}{N} + \frac{u m}{M} \right)}. \quad (2)$$

Z powyższych zależności można łatwo wywnioskować, że:

- widmo F jest w ogólności tablicą liczb zespolonych;
- tablica ta (tzw. F -obraz) ma, podobnie jak obraz I , wymiary $M \times N$;
- transformacja jest separowalna, tzn. można osobno policzyć jednowymiarowe transformaty w obrębie wierszy, a następnie wykorzystać je do obliczania jednowymiarowych transformat w obrębie kolumn (lub w odwrotnej kolejności);



Rys. 2: Ilustracja filtracji obrazu w dziedzinie częstotliwości

- ze względu na fakt, że obraz zawiera wartości rzeczywiste, w F -obrazie pojawia się charakterystyczna symetria;
- składowej stałej odpowiada współczynnik $F(0, 0)$, a więc piksel w lewym górnym rogu F -obrazu (Rys. 1b). Ze względu na symetrię widma oraz ogólnie przyjęty standard analizy DFT, bardzo często korzysta się z widma w postaci przesuniętej w taki sposób, że składowa stała znajduje się w środku F -obrazu (Rys. 1c).

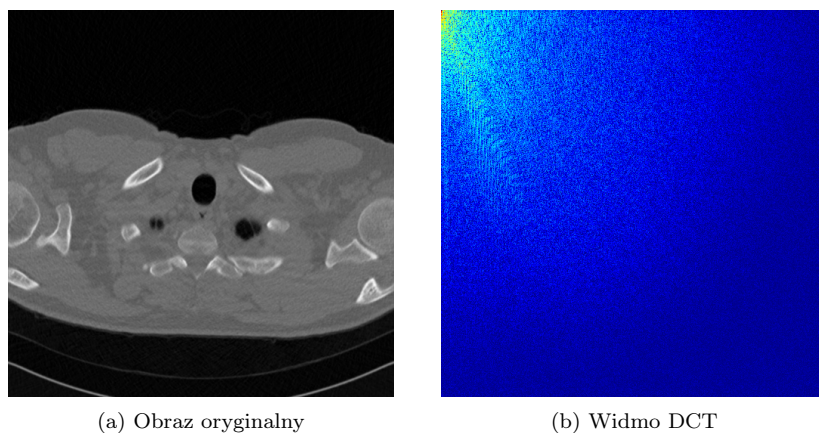
Transformatę Fouriera obrazu można przedstawić jako sumę części rzeczywistej i urojonej:

$$F(u, v) = \text{Re}\{F(u, v)\} + j \cdot \text{Im}\{F(u, v)\} \quad (3)$$

lub – częściej – w postaci amplitudy i fazy:

$$F(u, v) = |F(u, v)| \cdot e^{j\text{Arg}\{F(u, v)\}}. \quad (4)$$

W praktyce do wyznaczania transformaty Fouriera obrazu wykorzystuje się algorytm szybkiej transformacji Fouriera (ang. *Fast Fourier Transform*, FFT), co pozwala na znaczne ograniczenie liczby wykonywanych operacji, w konsekwencji skracając czas obliczeń.



Rys. 3: Ilustracja dyskretnej transformacji kosinusowej obrazu

W środowisku Matlab para transformacji Fouriera jest realizowana przez funkcje¹:

```
F = fft2(I);
I = ifft2(F);
```

Inne przydatne tu funkcje to `fftshift` (przesunięcie F -obrazu do postaci z Rys. 1c), `abs`, czy `angle` (obliczenie amplitudy i fazy widma). Ze względu na fakt, że składowa stała dominuje nad innymi, widmo amplitudowe wyświetla się często w postaci logarytmicznej: `log(abs(fftshift(F))+1)` (jedynekę dodaje się, aby uniknąć obliczania logarytmu zera).

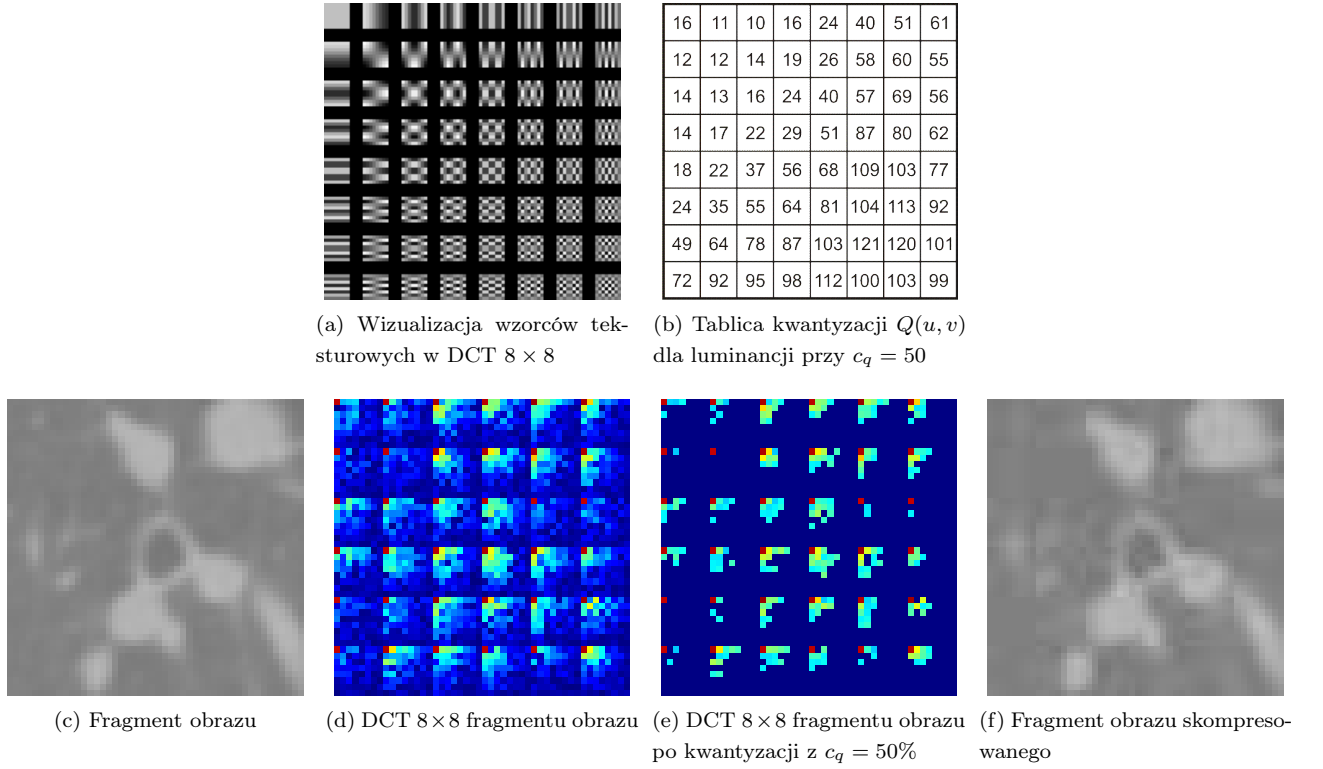
2.3 Filtracja w dziedzinie częstotliwości

Filtracja w dziedzinie częstotliwości wykorzystuje fakt, że transformata spłotu dwóch funkcji jest iloczynem ich transformat, a więc zamiast spłotu macierzy obrazu i maski filtru można dokonać przemnożenia transformaty obrazu i charakterystyki widmowej filtru. W praktyce daje to korzyści obliczeniowe tylko dla filtrów o dużych maskach. Częściej wykorzystuje się transformację Fouriera do opisu filtru poprzez prezentację transformaty jego odpowiedzi impulsowej: filtry dolnoprzepustowe mają małe wartości współczynników dla wysokich częstotliwości, a górnoprzepustowe dla niskich częstotliwości (więcej na ten temat w ćwiczeniu Filtry). Rys. 2 przedstawia przykładowe maski filtrów oraz rezultat ich zastosowania. Zgodnie z ogólnymi zasadami filtracji najczęściej wartości w masce są rzeczywiste z przedziału $[0; 1]$. Tutaj, dla prostoty interpretacji sugerowane są maski binarne, tj. przyjmujące wyłącznie wartości 0 lub 1 (Rys. 2).

2.4 Dyskretna transformacja kosinusowa

Transformata kosinusowa otrzymywana jest przez zastąpienie obrazu pierwotnego o rozmiarach $M \times N$ obrazem o rozmiarach $(2M-1) \times (2N-1)$, powiększonym przez symetryczne odbicie względem zerowego wiersza i zerowej kolumny. Taka symetria sprawia, że w widmie DFT składniki urojone redukują się do zera, a rzeczywiste ulegają podwojeniu. Własności symetrii pozwalają pozostawić tylko jedną ćwiartkę obrazów o indeksach nieujemnych w obydwu dziedzinach. Dla typowych obrazów większość informacji zawarta jest w niewielkiej liczbie współczynników transformaty kosinusowej (Rys. 3), co może stanowić podstawę koncepcji kompresji poprzez zerowanie małych współczynników wysokoczęstotliwościowych, bez większego wpływu na percepcję obrazu. Koncepcja ta będzie poddana analizie w programie ćwiczenia.

¹Funkcje te umożliwiają również np. zmiany rozmiarów macierzy, co jednak nie będzie wykorzystane w tym ćwiczeniu.



Rys. 4: Ilustracja kompresji obrazu opartej o dyskretną transformację kosinusową

Prosta i odwrotna dwuwymiarowa dyskretna transformacja kosinusowa opisana jest parą wzorów:

$$D(u, v) = \alpha_u \alpha_v \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \cos\left(\frac{\pi(2m+1)u}{2M}\right) \cos\left(\frac{\pi(2n+1)v}{2N}\right), \quad (5)$$

$$I(m, n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha_u \alpha_v D(u, v) \cos\left(\frac{\pi(2m+1)u}{2M}\right) \cos\left(\frac{\pi(2n+1)v}{2N}\right), \quad (6)$$

gdzie:

$$\alpha_u = \begin{cases} \frac{1}{\sqrt{M}} & \Leftrightarrow u = 0 \\ \frac{2}{\sqrt{M}} & \Leftrightarrow 1 \leq u \leq M-1, \end{cases} \quad (7)$$

$$\alpha_v = \begin{cases} \frac{1}{\sqrt{N}} & \Leftrightarrow v = 0 \\ \frac{2}{\sqrt{N}} & \Leftrightarrow 1 \leq v \leq N-1, \end{cases} \quad (8)$$

W Matlabie do obliczenia powyższej pary transformacji służą funkcje:

```
D = dct2(I);
I = idct2(D);
```

2.5 Kompresja obrazów oparta o DCT

Transformacja kosinusowa jest podstawą kompresji JPEG (ang. *Joint Photographic Expert Group*). Obraz dzieli się jednak na kwadratowe bloki o rozmiarach 8×8 pikseli i każdy z bloków poddaje operacji DCT. W każdym z małych bloków określa się więc wartość średnią w jego obrębie (składowa stała) oraz 63 współczynniki opisujące składowe o niezerowych częstotliwościach, na różnym poziomie zmienności i orientacji w przestrzeni. Aby uzyskać każdy ze współczynników koreluje się (mnoży) blok z charakterystycznym wzorcem teksturowym

o rozmiarze 8×8 pikseli (Rys. 4a). W efekcie uzyskuje się dla każdego bloku widmo DCT, w którym największa energia jest skoncentrowana w zakresie niskich częstotliwości – w lewym górnym rogu; współczynniki wysokoczęstotliwościowe mają wartości zbliżone do zera (Rys. 4d). Ludzkie oko wykazuje podobny mechanizm wrażliwości na zniekształcenia określonych współczynników: największa na lewą górną część mini-widma DCT. W zasadniczym kroku kompresji współczynniki DCT każdego bloku podlegają kwantyzacji – dzieleniu przez stałe wartości założone w standardzie z tzw. tablicy kwantyzacji Q (Rys. 4b) skalowanej przez współczynnik jakości kompresji c_q i zaokrągleniu do najbliższych liczb całkowitych:

$$D_q(u, v) = \left\lfloor \frac{D(u, v)}{Q_{c_q}(u, v)} \right\rfloor. \quad (9)$$

W efekcie większość współczynników się zeruje, a pozostałe mają relatywnie małe wartości, możliwe do zakodowania na małej liczbie bitów. Następuje tu utrata części informacji, jednak kompresja jest realizowana tak, aby działało się to w sposób najmniej zauważalny dla ludzkiego oka.

Na dalszym etapie kompresji, mniej interesującym z punktu widzenia przetwarzania obrazu, skwantowane współczynniki szereguje się wg tzw. porządku zygzakowego i poddaje kodowaniu algorytmem Huffmana. W procesie dekodowania w uproszczeniu należy wykonać powyższe operacje w odwrotnej kolejności, tj. przemnożyć skwantowane współczynniki $D_q(u, v)$ przez czynnik $c_q \cdot Q(u, v)$ i uzyskane widmo DCT (Rys. 4e) poddać odwrotnej transformacji kosinusowej. Sekwencja operacji: dzielenie-zaokrąglenie-mnożenie skutkuje nieodwracalną utratą informacji, np. współczynnik $D(u, v) = 120$ przy $Q_{c_q}(u, v) = 50$ daje wartość:


$$D_q(u, v) = 50 \cdot \left\lfloor \frac{120}{50} \right\rfloor = 50 \cdot 2 = 100.$$

Zyskiem łańcucha operacji kompresji jest jednak możliwość uzyskania kiludziesięciokrotnego zmniejszenia rozmiaru obrazu.

3 Program ćwiczenia

W ramach ćwiczenia studenci wykorzystują udostępnione przez prowadzącego m-pliki w celu badania wybranych własności transformat DFT i DCT. Wskazane fragmenty kodu wymagają uzupełniania; student powinien również zmieniać wartości wskazanych parametrów algorytmów w celu uzyskania pożądanych rezultatów. Uzyskane wyniki oraz wnioski studenci notują w protokole ćwiczenia, załączonym do instrukcji. Student jest zobowiązany stawiać się na zajęcia z wydrukowanym wzorem protokołu ćwiczenia. UWAGA! Numer aktualnie realizowanego zadania należy wskazać przez nadanie w głównym skrypcie odpowiedniej wartości zmiennej **zadanie**. Na początku ćwiczenia należy również przypisać zmiennej **nrsekcji** własny numer sekcji.

Zadanie 1. Dyskretna transformacja Fouriera

Uruchom skrypt **Lab_06.m**. Z rozwijanego menu w lewym górnym rogu otwartego okna możesz wybrać jeden z przygotowanych obrazów flag państw oraz podejrzeć go w wersji barwnej oraz monochromatycznej. W prawej części okna wyświetlane jest widmo obrazu w jednym z wybranych wariantów (amplituda, logarytm amplitudy, faza). Widmo jest wyświetlane w postaci płaszczyzny, którą można obracać za pomocą narzędzia  w menu okna.

Przyjrzyj się wnikliwie widmom przygotowanych obrazów, najlepiej w kolejności, w jakiej je uszeregowano w menu rozwijanym. Flagi charakteryzują się występowaniem ograniczonej liczby charakterystycznych wzorów/krawędzi, co jest szczególnie pomocne w uczeniu się czytania widma dwuwymiarowego. Odpowiedz na pytania (a)-(f) zadane w części 1. protokołu.

Zadanie 2. Filtracja w dziedzinie częstotliwości

W lewej górnej części okna w zadaniu 2. wyświetlany jest obraz medyczny przypisany Ci na podstawie numeru sekcji². W prawej górnej części pokazano jego widmo amplitudowe i fazowe. W zadaniu 2. realizowana jest filtracja w dziedzinie częstotliwości poprzez maskowanie widma DFT. W środkowej części okna przedstawiono maskę, zainicjowaną zerami wewnątrz funkcji `Fun_FilterFFT`. Efekty filtracji — widmo i obraz po zastosowaniu odwrotnej DFT — jest widoczny w dolnej części okna.

Twoim zadaniem jest nadać wewnątrz funkcji `Fun_FilterFFT` masce `M` postać właściwą temu, aby uzyskać filtrację:

- a) izotropowo dolnoprzepustową,
- b) izotropowo górnoprzepustową,
- c) izotropowo pasmowoprzepustową,
- d) jednocześnie dolnoprzepustową w kierunku poziomym i górnoprzepustową w kierunku pionowym (w przestrzeni obrazu).

Realizując powyższe cele uzupełniaj tabelę w części 2. protokołu.

Zadanie 3. Dyskretna transformacja kosinusowa. JPEG

W zadaniu 3. obserwacji podlega dyskretna transformacja kosinusowa oraz bazujące na niej proste koncepcje kompresji obrazów. Obraz modułów współczynników transformaty DCT obrazu oryginalnego jest wyświetlany w środkowej części okna. Pod nim można obserwować moduł DCT po wyzerowaniu r procent współczynników o najmniejszych wartościach bezwzględnych. Wartość parametru r ustawia się za pomocą suwaka. Obraz skompresowany, odtworzony ze zmodyfikowanej macierzy DCT, umieszczony jest po prawej stronie. W prawym dolnym rogu można obejrzeć obraz różnicowy, w którym intensywność jest wprost proporcjonalna do modułu błędu intensywności danego piksela w skompresowanym obrazie. W zależności od stanu przycisku opcji `DCT 8x8` operacje w dziedzinie DCT są przeprowadzane na całym obrazie bądź na blokach o rozmiarach 8×8 pikseli.

Odpowiedz na pytania zadane w 3. części protokołu, używając kolejno:

- a) obrazu wskazanego Twoim numerem sekcji i pełnej DCT obliczanej na podstawie całego obrazu;
- b) obrazu wskazanego Twoim numerem sekcji i DCT obliczanej na podstawie bloków o rozmiarach 8×8 pikseli;
- c) obrazu flagi wybranej przez siebie z menu rozwijanego i pełnej DCT obliczanej na podstawie całego obrazu;
- d) obrazu flagi jw. i DCT obliczanej na podstawie bloków o rozmiarach 8×8 pikseli;

W trybie `DCT 8x8` masz również możliwość podejrzenia efektów kompresji JPEG obrazu po aktywowaniu przycisku opcji `JPEG`. Pamiętaj, że tym razem im większa wartość na suwaku (c_q), tym wyższa jakość obrazu skompresowanego.

- e) W polu 3e krótko skomentuj rezultaty kompresji JPEG obrazu wskazanego Twoim numerem sekcji; wskaż wartości współczynnika jakości c_q graniczne z punktu widzenia percepcji obrazu.

²Możesz śmiało przyjrzeć się również efektom przetwarzania innych obrazów, w tym kilku flag umieszczonych na końcu listy rozwijanej.

Literatura

- [1] Malina W., Smiatacz M., Cyfrowe przetwarzanie obrazów, AOW EXIT, 2008.
- [2] Tadeusiewicz R., Korohoda P., Komputerowa analiza i przetwarzanie obrazów, Wyd. Fundacji Postępu Telekomunikacji, 1997. <http://winntbg.bg.agh.edu.pl/skrypty2/0098/>
- [3] Koprowski R., Wróbel Z., Praktyka przetwarzania obrazów w programie MatLab, AOW EXIT, 2004.
- [4] Gonzalez R. C., Woods R. E., Digital Image Processing, Prentice Hall, 2002.
- [5] Shih F. Y., Image Processing and Pattern Recognition. Fundamentals and Techniques, Wiley, 2010.
- [6] Joint Photographic Expert Group, <http://www.jpeg.org> [dostęp: 12.10.2015].