

# **Transformaty Fouriera i Hough'a, kompresja obrazów**

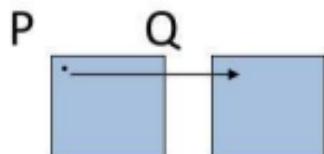
WYKŁAD 5  
Dla studiów niestacjonarnych  
2021/2022

Dr hab. Anna Korzyńska, prof. IBIB PAN

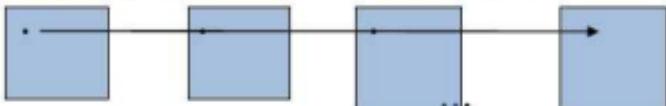
# Operacje na obrazach

- Operacje punktowe (jednopunktowe):

Jednoargumentowe  
 $[q(i, j)] = f[p(i, j)]$



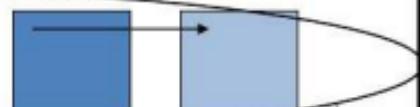
Wieloargumentowe

$$[q(i, j)] = f[p_1(i, j), p_2(i, j), \dots, p_k(i, j)]$$


- Operacje sąsiedztwa (kontekstowe)

$$[q(i, j)] = f[p(i, j), p(i-1, j-1), p(i+1, j+1), \dots]$$

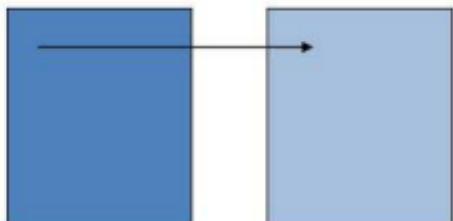

- Operacje globalne

$$[q(i, j)] = f[P]$$


# Operacje globalne

Są to operacje, w których na wartość zadanego piksela obrazu wynikowego  $q$  o współrzędnych  $(i, j)$  mają wpływ wartości wszystkich pikseli obrazu pierwotnego  $p$ , czyli są to operacje kontekstowe dla których otoczeniem piksela jest cały obraz  $p$ .

$$[q(i, j)] = f[p]$$



# Operacje globalne na obrazach

- Transformacje
  - Fouriera (matematyczny pryzmat)
  - kosinusowa
  - Falkowa (matematyczny mikroskop)
  - Hough'a
  - Odległościowa (morfologia matematyczna)
- Inne globalne
  - Skalowanie rozmiarów
  - Rejestracja multymodalna i multisesyjna

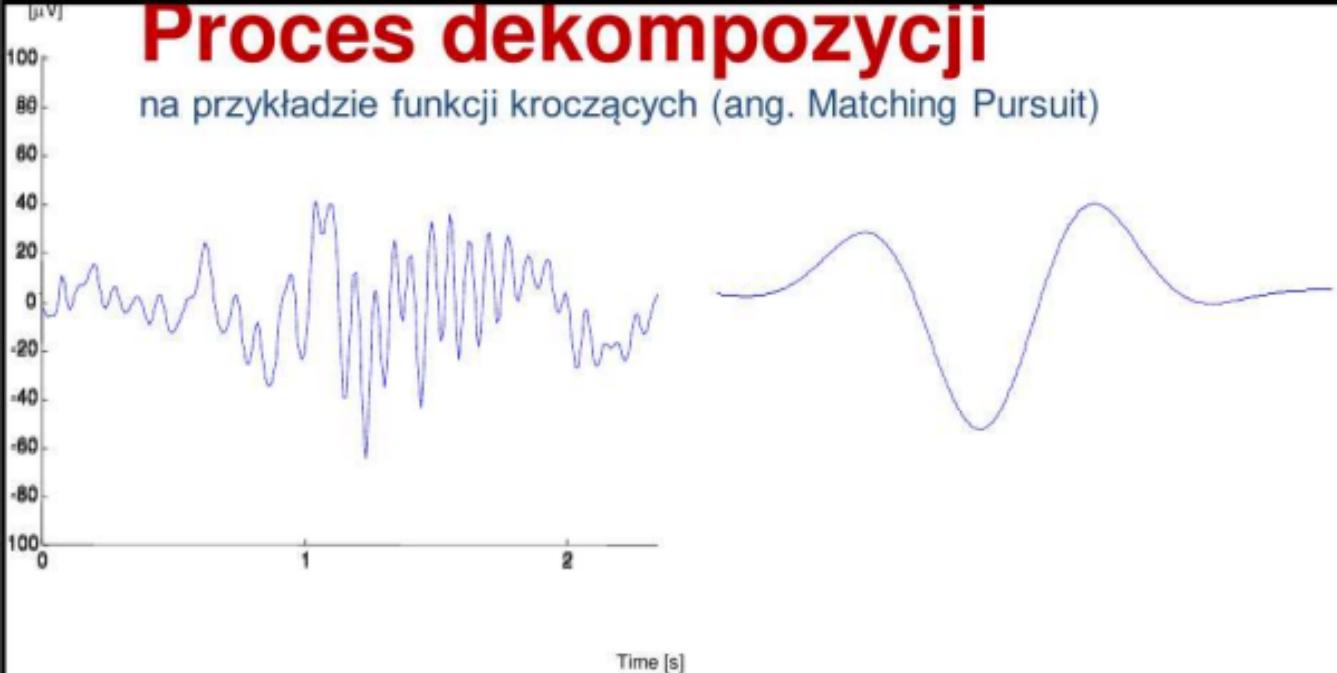
## Transformata

Przekształcenie matematyczne **odwracalne i bezstratne**, przenoszące sygnał z jednej przestrzeni w inną, w której wygodniej dokonywać pewnych analiz lub procesów (np.: kompresji informacji), ponieważ w przestrzeni docelowej uwypuklane są cechy sygnału istotne z punktu widzenia celu analizy lub przetwarzania .

W przestrzeni docelowej sygnał jest **dekomponowany ze względu na zbiór funkcji bazowych tej przestrzeni**.

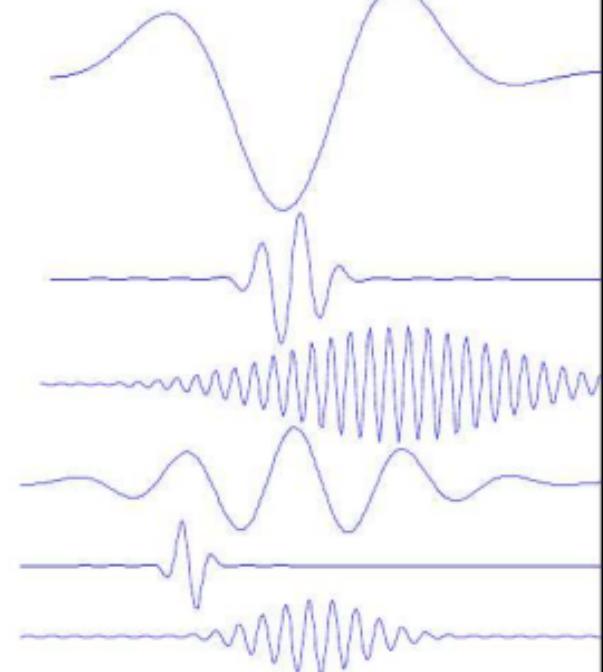
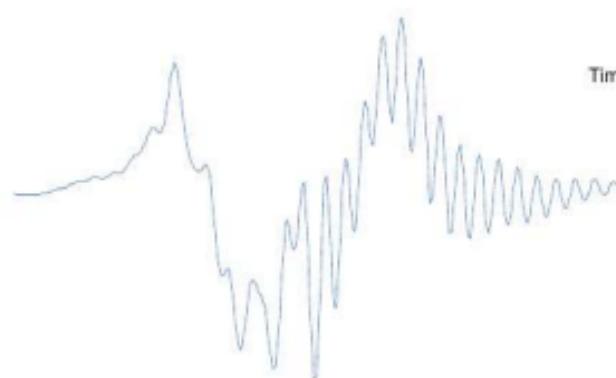
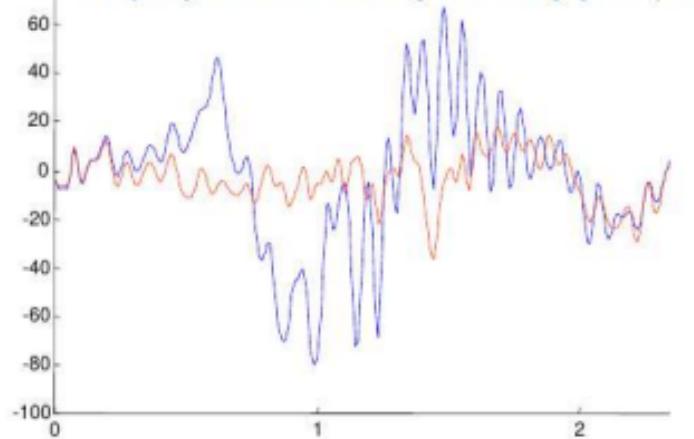
# Proces dekompozycji

na przykładzie funkcji kroczących (ang. Matching Pursuit)



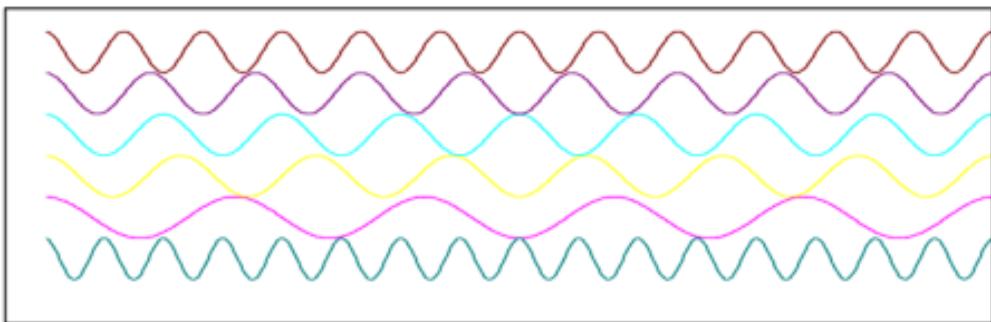
# Proces dekompozycji

na przykładzie funkcji kroczących (ang. Matching Pursuit)



# Transformacja Fouriera

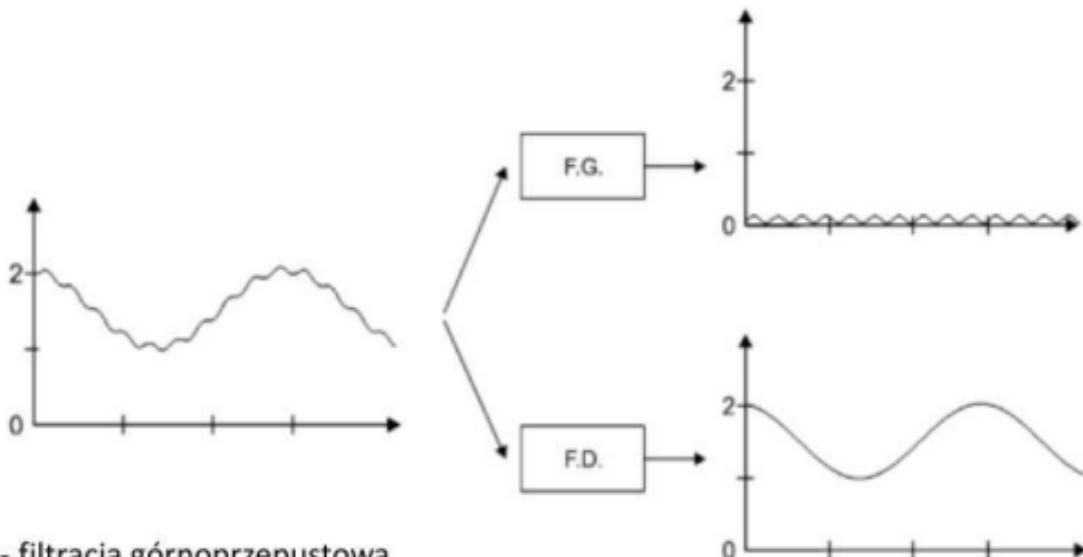
Funkcje bazowe dla przestrzeni jednowymiarowej 1D



# Transformata Fouriera

Główne zastosowanie: – poprawa jakości sygnału według  
przyjętego kryterium

Interpretacja częstotliwościowa filtracji



F.G. - filtracja górnoprzepustowa

F.D. - filtracja dolnoprzepustowa

# Transformata Fouriera (opis matematyczny)

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp\left(-j \frac{2\pi u x}{N}\right)$$

gdzie:  $u = 0, 1, \dots, N-1$ ,  $j = \sqrt{-1}$        $u$  - argument widma

**Funkcje bazowe**  
- składniki sumy  
w wyrażeniu  
opisującym  
transformatę

## Złożoność obliczeniowa:

$O(N^2)$       - tzn.  $N^2$  dodawań i mnożeń

$O(N * \log_2 N)$       - szybka transformata Fouriera  
(Fast Fourier Transform (FFT))

## Transformata odwrotna:

$$f(x) = \frac{1}{N} \sum_{u=0}^{N-1} F(u) \exp\left(j \frac{2\pi u x}{N}\right)$$

gdzie:  $x = 0, 1, \dots, N-1$ ;       $x$  - argument funkcji

## Dwuwymiarowa transformata Fouriera

$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp\left(-j2\pi(ux + vy)\right)$$

gdzie:  $u, v = 0, 1, \dots, N-1$ ;  $u, v$  - argumenty widma  $F$

## Dwuwymiarowa odwrotna transformata Fouriera

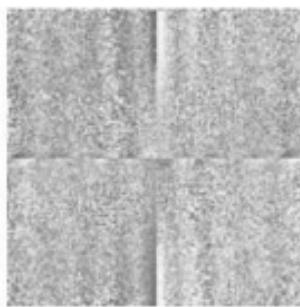
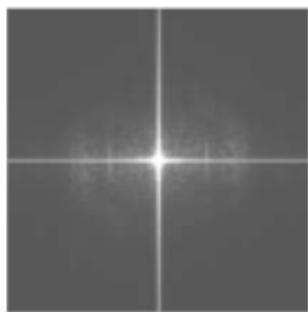
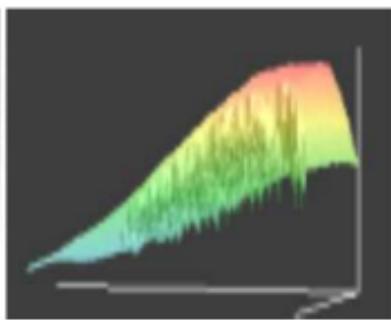
$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp\left(\frac{j2\pi(ux + vy)}{N}\right)$$

gdzie:  $x, y = 0, 1, \dots, N-1$   $x, y$  - argumenty funkcji  $f$

Duża złożoność obliczeniowa.

# Transformata Fouriera

To transformacja która przenosi obraz (dyskretny sygnał dwuwymiarowy) z dziedziny oryginalnej do dziedziny częstotliwości przestrzennych.



Dziedzina przestrzenna jest taką dziedziną obrazu, w której odległości między obiektemi na obrazie odpowiadają odległościom między odpowiadającymi im obiektemi w przestrzeni (z uwzględnieniem skrótów perspektywicznych, - ang. spatial or pixels domain, którą przedstawia obraz).

W dziedzinie częstotliwości przestrzennych wartość w określonej pozycji przedstawia, jaki jest udział zmian poziomów szarości zgodnie z funkcją sinusoidalną, o kierunku i częstotliwości odpowiadającej pozycji tego punktu w układzie współrzędnych widma amplitudowego.

# Wyliczanie transformaty Fouriera dla obrazów

Dzięki separowalności:

- Krok 1 po wierszach

$$P(k, m) = \frac{1}{N} \sum_{i=0}^N f(i, m) \exp\left[-j \frac{2\pi}{N}(ki)\right]$$

- Krok 2 po kolumnach

$$F(k, l) = \frac{1}{M} \sum_{m=0}^M P(k, m) \exp\left[-j \frac{2\pi}{M}(lm)\right]$$

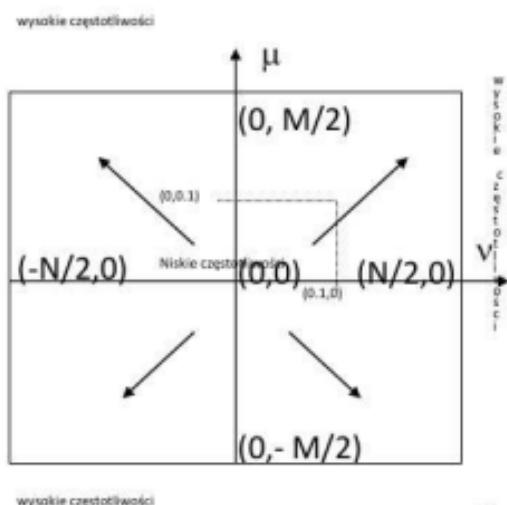
Gdzie:

$$j = \sqrt{-1} \quad \exp\left[-j \frac{2\pi}{N}(ki)\right] = \cos\left(\frac{2\pi}{N}(ki)\right) - j \sin\left(\frac{2\pi}{N}(ki)\right)$$

Algorytmy liczenia szybkiej transformaty Fouriera FFT

Wartość transformaty Fouriera jest liczbą zespoloną.  
Informacja o obrazie jest zawarta w dwóch tablicach lub odpowiadających im obrazach:

- albo części rzeczywistej i części urojonej liczb zespolonych
- albo amplitudy i fazy liczb zespolonych.



## Interpretacja w odniesieniu do operacji na obrazach

$u, v$  - interpretuje się jako częstotliwości przestrzenne obrazu  $f(x,y)$  odpowiednio wzdłuż osi  $x, y$ . Oznacza to, że **drobne przedmioty, detale, szum i gwałtowne** przejścia od czerni do bieli widoczne na obrazie są odpowiedzialne głównie za fragmenty funkcji  $F(u,v)$  występujące dla **dużych** wartości jej argumentów. Przeciwnie, składowe  $F(u,v)$  odpowiadające **małym** wartościom  $u,v$  kodują **duże obiekty lub obszary tła** obrazu.

## Schemat filtracji:

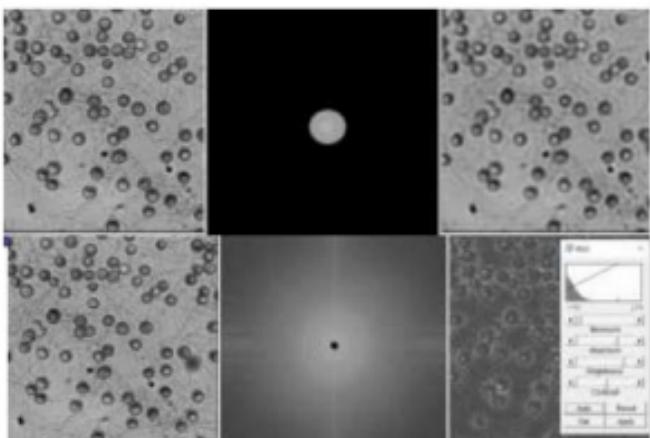
- $f(x, y) \xrightarrow{TF} F(u, v)$  funkcja filtru (usuwanie zakłóceń; podnoszenie kontrastu).
- $F' (u, v) = F(u, v) \cdot H(u, v)$  zmodyfikowana funkcja widma
- $F' (u, v) \xrightarrow{OTF} f' (x, y)$

$H(u, v)$  realizuje filtrację:

dolnoprzepustową (FD)

lub

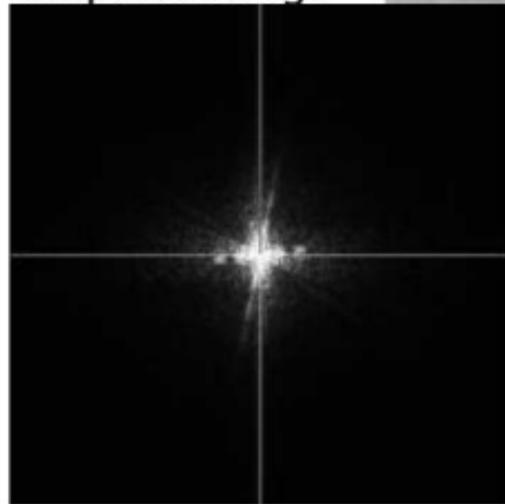
górnoprzepustową (FG).



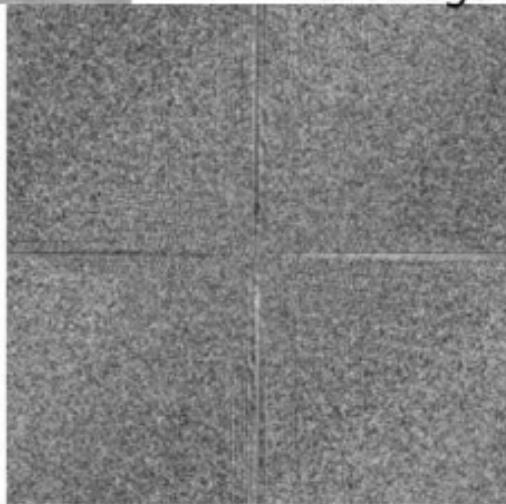
## Przykład

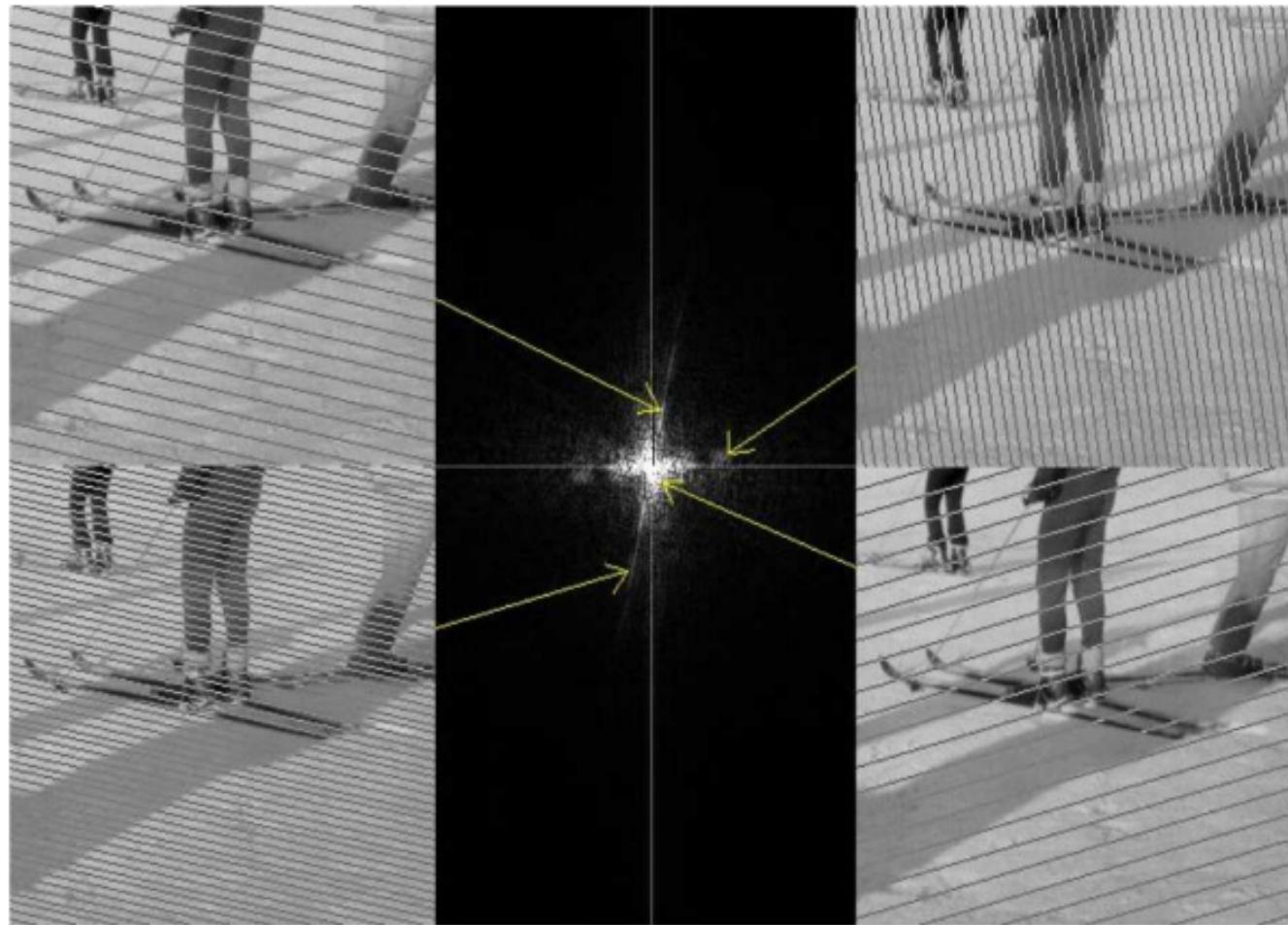


Obraz widma  
amplitudowego



Obraz widma  
fazowego



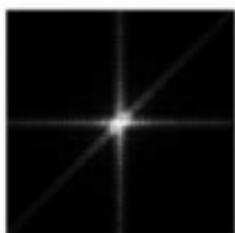
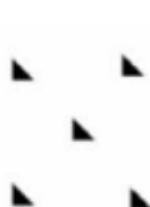


# **Do czego wykorzystujemy FFT w przetwarzaniu obrazów**

- Do obserwacji **periodyczności** w obrazie
- Do wyznaczania **kierunku struktur** w obrazie
- Do **wytłumiania lub wzmacniania** pewnych kategorii informacji
- Do **wygładzania i wyostrzania** obrazu

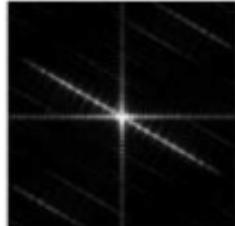
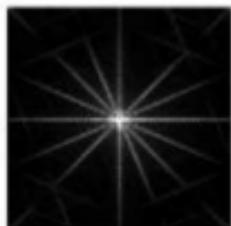
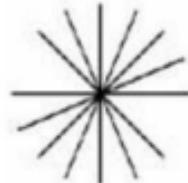
## Przykład 1

Dwa zestawy periodycznie ułożonych trójkątów oraz ich widma amplitudowe



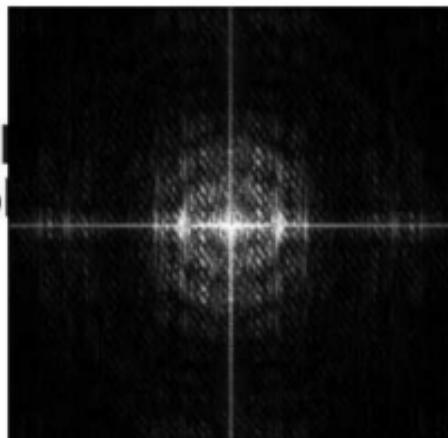
## Przykład 2

Dwa zestawy linii o grubości 2 piksele i długości 150 pikseli każda oraz ich widma amplitudowe. Obraz z liniami równoległyymi ma widmo z jednym dominującym kierunkiem, prostopadłym do kierunku linii na obrazie pierwotnym.



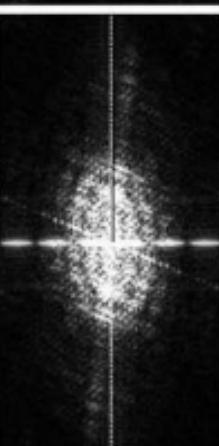
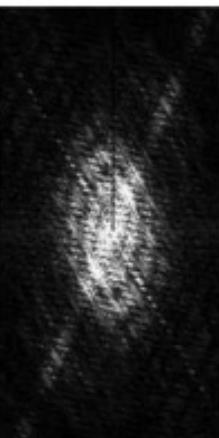
To jest tekst do  
jest tekst do pro  
tekst do prob.

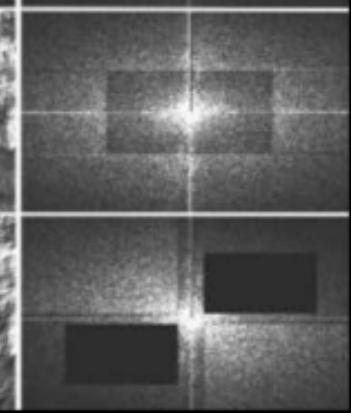
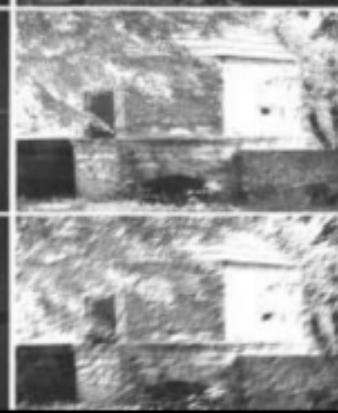
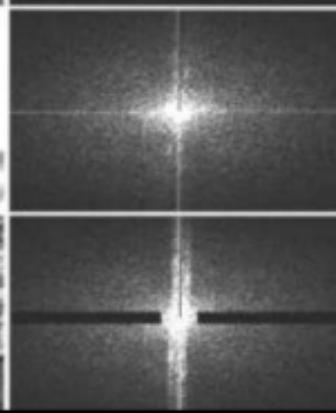
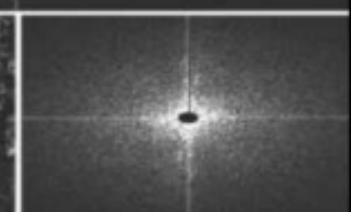
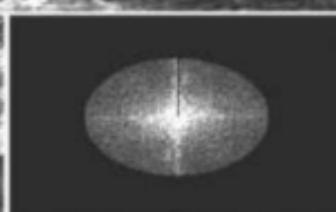
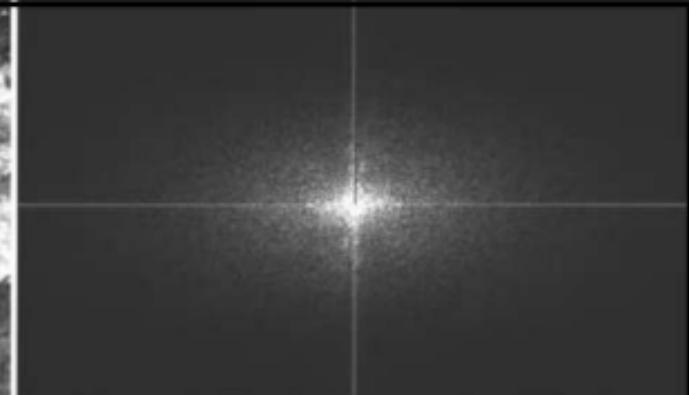
To jest tekst do pr  
est tekst do prob.  
t do prob.



ekst do pr  
ekst do prob

est tekst do  
tekst do prob  
do prob



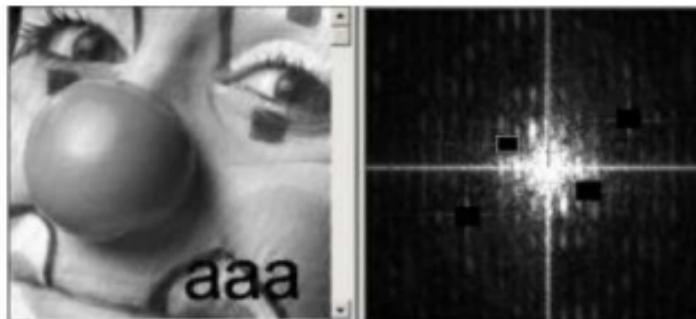


## Przykład 4

Usunięcie zakłóceń w postaci regularnie powtarzających się skośnych prążków poprzez wytłumienie odpowiednich obszarów widma



Obraz i jego widmo amplitudowe po usunięciu zakłóceń



W przestrzeni docelowej sygnał jest dekomponowany ze względu na zbiór funkcji bazowych tej przestrzeni.

## Do czego jeszcze może przydać się analiza częstotliwości w obrazach?

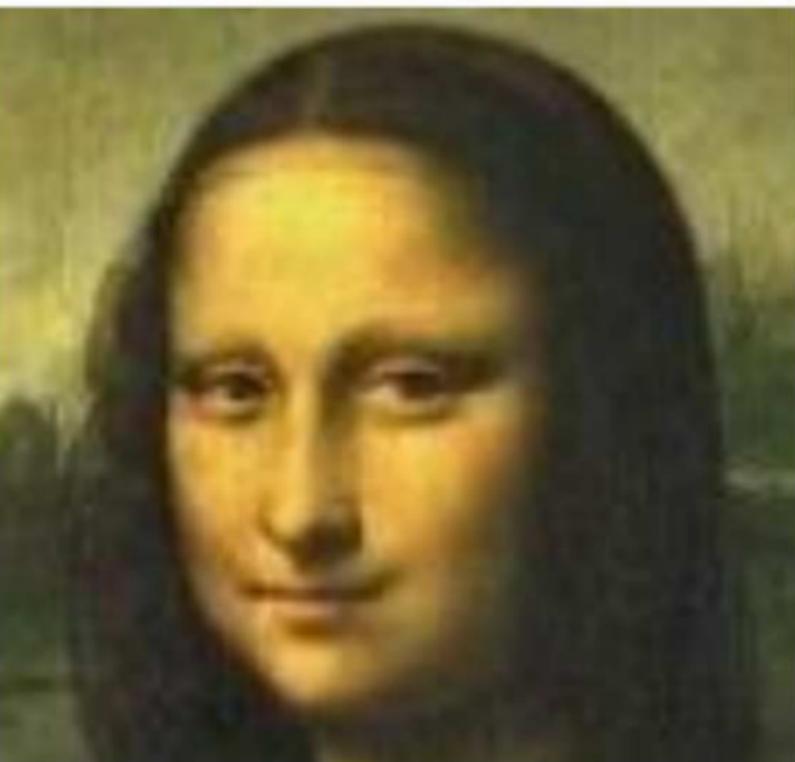
- Kompresji obrazów
- Rozszyfrowania tajemnicy uśmiechu Mony Lizy

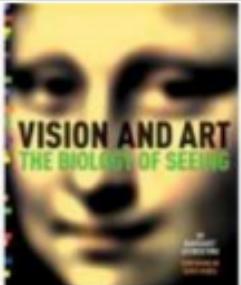
Inne cele niż te typowe dla przetwarzania obrazów

# Tajemnica obrazu Mony Lizy



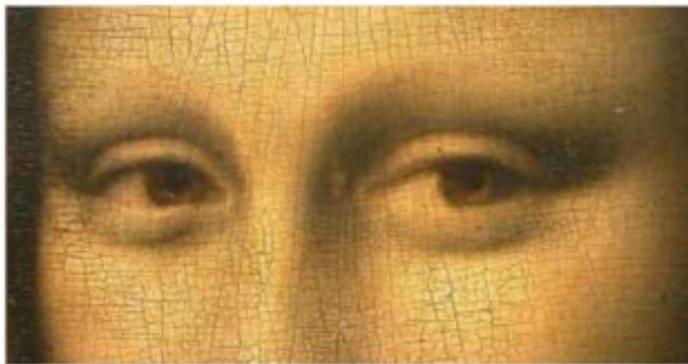
Notatki na marginesie książki malarza, zrobione w październiku 1503 roku, mówią, że to Liza del Giocondo żona florenckiego kupca Francesco del Giocondo



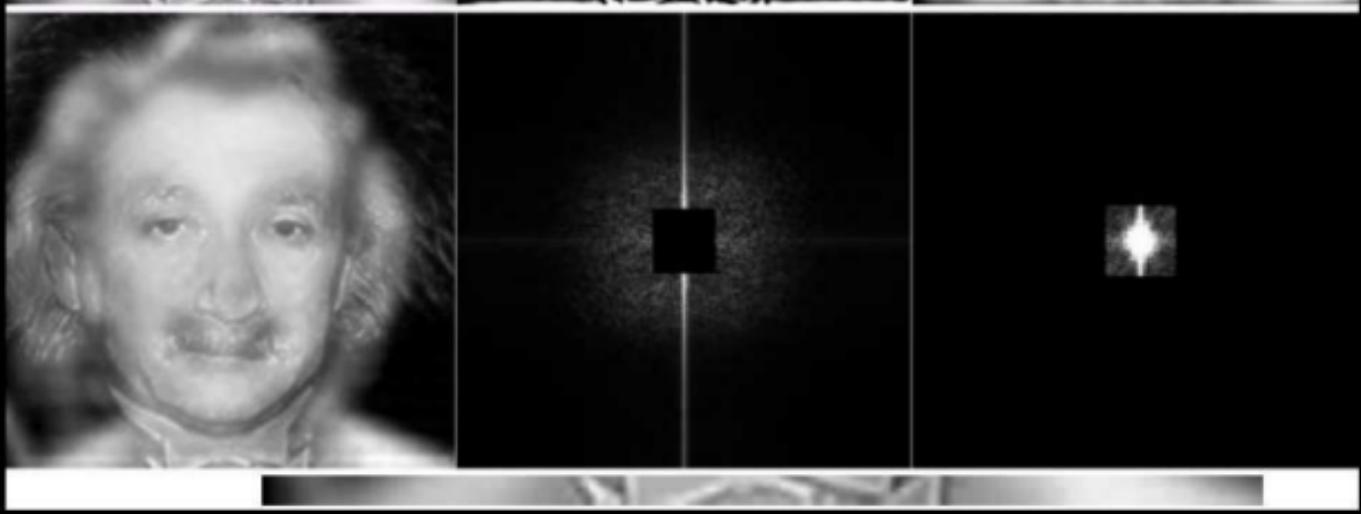


Dr. Margaret Livingstone nurobiolog wyjaśniła fenomen uśmiechu Mony Lisy tym,  
jak działa ludzki wzrok.

Francuz, Pascal Cotte za pomocą urządzenia o rozdzielczości **240 mln pikseli/cal** zbadał różne warstwy obrazu w świetle białym, ultrafioletowym i podczerwonym.



Stwierdził, że kiedyś na obrazie były namalowane brwi i rzęsy kobiety, że jej dekolt ozdobiony był białą koronką, oraz że, miała **bardziej wyrazisty uśmiech**. Detale te zniknęły prawdopodobnie na skutek starzenia się obrazu.



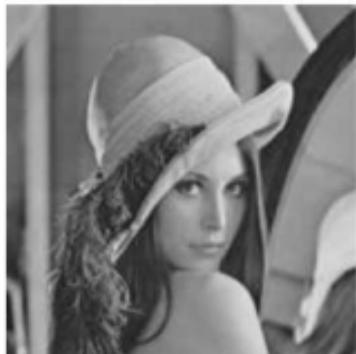
**Transformata Hougha**

# Detekcja krawędzi z wykorzystaniem Transformaty Hougha

Detekcja linii daje rezultaty w postaci  
**fragmentów prostej lub krzywej, grupy**  
**pikeseli (piksele) poszukiwanej krawędzi.**

**Kolejny etap:**

aproksymacja linii (prostych i krzywych) na



## Transformata Hough'a

Metoda detekcji krzywych analitycznych  
(jak w zapisie wektorowym obrazu) nie pikseli (!))  
przez transformatę wynikającą z **dualności** pomiędzy  
**punktami** na krzywej a **parametrami** tej krzywej.  
Krzywa analityczna o postaci  $f(x, a) = 0$   
gdzie  $x$  – punkt obrazu,  $a$  - wektor parametrów.

### Zaleta:

działa dobrze nawet wówczas, gdy ciągłość krawędzi  
nie jest zachowana i oprócz punktów krawędzi  
występują dodatkowe elementy (najczęściej z powodu  
szumów).

## Założenia:

Piksele o niezerowej wartości są elementami krawędzi (obraz po krawędziowaniu).  
Jeśli piksel  $(x, y)$  leży na prostej  $\rightarrow$  znaleźć zbiór wartości  $(\rho, \phi)$  w przestrzeni parametrów tej prostej.

$(x, y)$  - stałe,  $(\rho, \phi)$  - zmienne  $\rightarrow$  równanie normalne prostej przedstawia relację pomiędzy krzywą w przestrzeni parametrów, a punktem w obrazie.

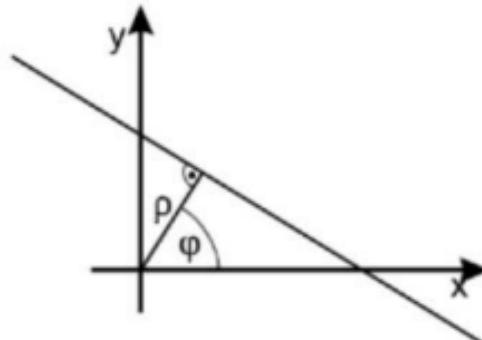
Punkt  $(x, y)$  leży na prostej  $\rightarrow$  krzywa w przestrzeni parametrów to sinusoida

TH - transformacja pomiędzy punktami obrazu a przestrzenią parametrów poszukiwanej krzywej.

Równanie *normalne* prostej:

$$x \cdot \cos\varphi + y \cdot \sin\varphi = \rho$$

Przestrzeń obrazu



# Transformacja pomiędzy punktami obrazu, a przestrzenią parametrów poszukiwanej prostej

Właściwości transformacji punkt

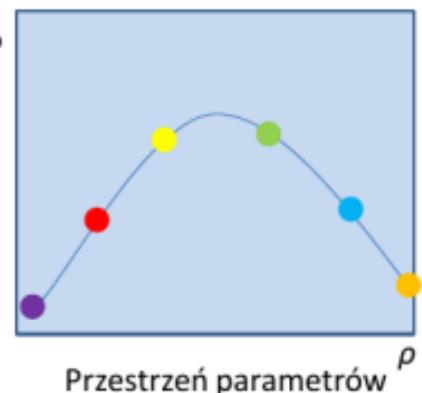
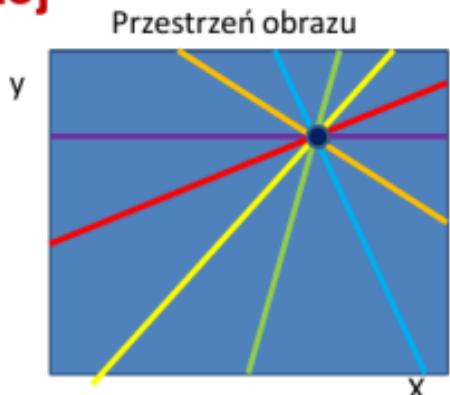
- sinusoida (przy detekcji prostej):

- punkt **obrazu** koresponduje z sinusoidą w przestrzeni parametrów,

- punkt **w przestrzeni parametrów** koresponduje z linią prostą w obrazie,

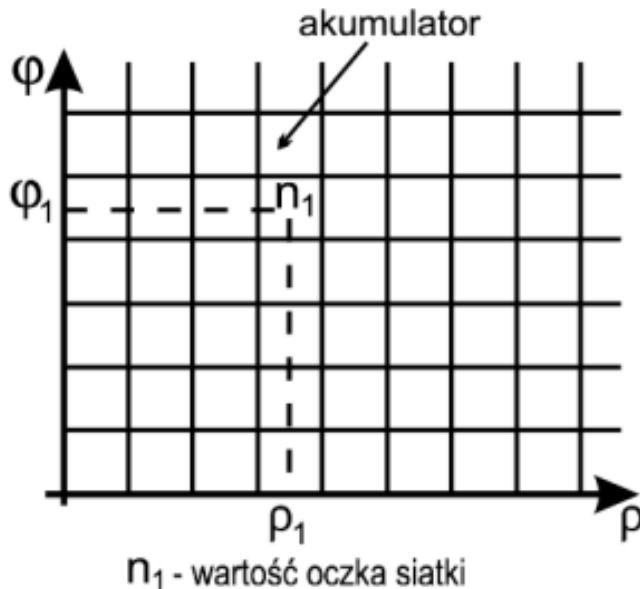
- punkty leżące na tej samej prostej w obrazie korespondują z sinusoidami przechodzącymi przez wspólny punkt w przestrzeni parametrów,

- punkty leżące na tej samej sinusoidzie w przestrzeni parametrów korespondują z liniami prostymi przechodzącymi przez ten sam punkt na obrazie.



**Redukcja złożoności obliczeniowej**  $\Rightarrow$  **dyskretyzacja przestrzeni parametrów  $\rho, \phi$** , utworzenie regularnej siatki ich wartości

### Zdyskretyzowana przestrzeń Hougha dwuwymiarowa tablica akumulatorów:



## 1 krok:

Dla każdego punktu  $(x_1, y_1)$  obrazu oznaczonego jako należący do prostej korespondująca krzywa (sinusoida dana równaniem

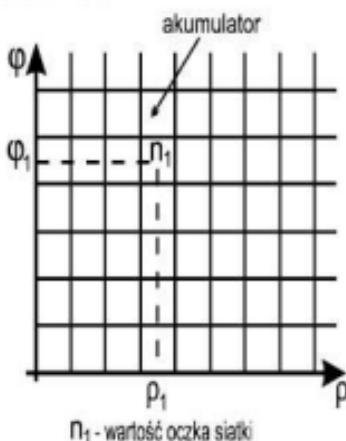
$$x_1 \cdot \cos \phi + y_1 \cdot \sin \phi = \rho$$

jest wprowadzana do tablicy poprzez powiększenie o 1 wartości oczek siatki w punkcie  $(\phi, \rho)$ .

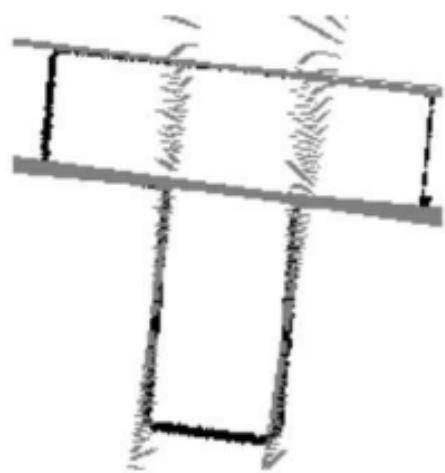
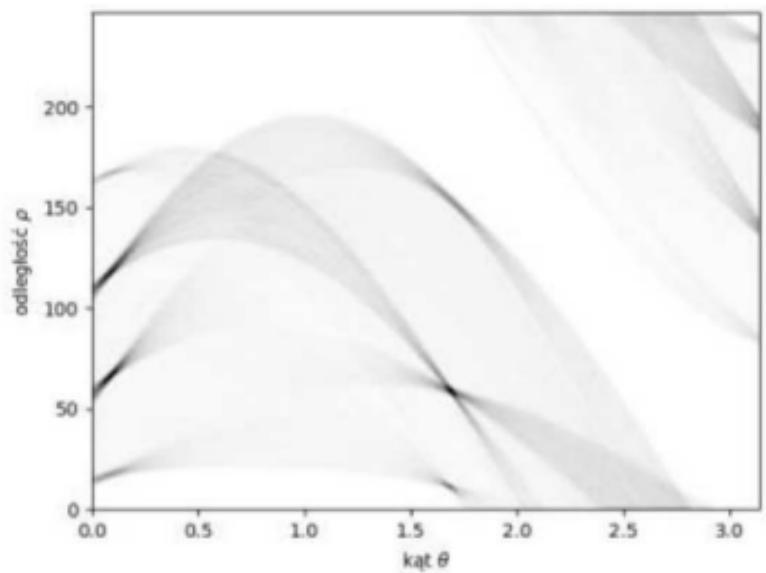
Dzięki temu każde oczko  $(\phi, \rho)$  tablicy przechowuje liczbę równą liczbie sinusoid przechodzących przez to oczko.

## 2 krok:

Przeszukiwanie tablicy w celu znalezienia oczek o największych wartościach (miejsca przecięć największej liczby krzywych). Jeśli któryś element tablicy akumolatorów ma wartość  $k$ , oznacza to, że dokładnie  $k$  punktów obrazu leży wzdłuż linii, której parametrami są  $\phi$  i  $\rho$  (pomijając błąd kwantyzacji).



$\Pi_1$  - wartość oczka siatki



**Dla okręgu:**

$$(x - a)^2 + (y - b)^2 = c^2$$

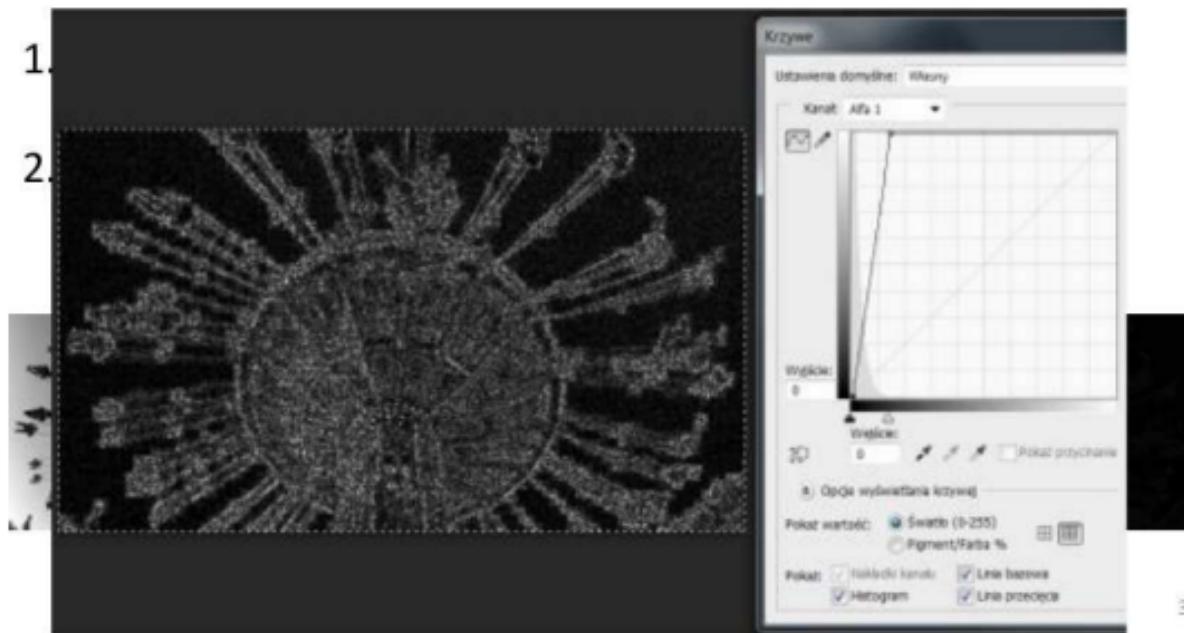
⇒ trójwymiarowa tablica akumulatorów.

współrzędne oczek: (a,b,c)

# Kompresji obrazów

# Jak ocenić podobieństwo obrazów?

Liczenie różnic pomiędzy obrazami – 2 sposoby



# Obraz jako wektor (punkt w przestrzeni n-wymiarowej)

$$\underline{d} = [d_1, \dots, d_n]^T \quad \underline{d} \in D$$

*n* – liczba pikseli obrazu , np.  $N^2$

*D* – przestrzeń obrazów

Wartości odpowiednich składowych wektora zależą od sposobu przeglądu obrazu.

Przykładowe sposoby przeglądu obrazu:

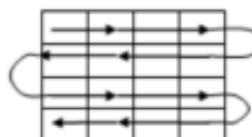
- "linia po linii" - 1-szy sposób

Bez przeplotu

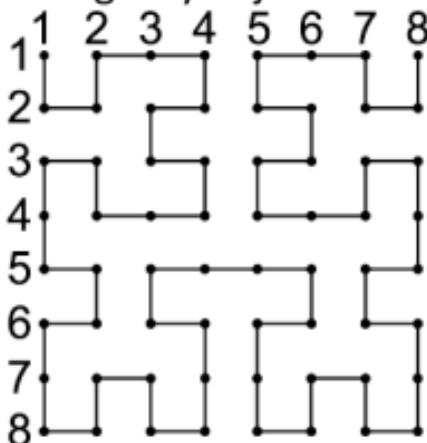


- "linia po linii - 2-gi sposób

Z przeplotem



- według krzywej Hilberta



Krzywa Hilberta rzędu 3. Obraz o rozmiarach 8x8 ( $N=8$ )

**Metryka** odwzorowanie  $\rho : X \times X \rightarrow R_*$  spełniające dla wszystkich wektorów  $\underline{x}^\mu \in X$  ( $\mu$  założenia) (warunki):

$$\rho(\underline{x}^\mu, \underline{x}^\nu) = 0 \Leftrightarrow \underline{x}^\mu \equiv \underline{x}^\nu \quad - \text{tożsamość}$$

$$\rho(\underline{x}^\mu, \underline{x}^\nu) = \rho(\underline{x}^\nu, \underline{x}^\mu) \quad - \text{symetria},$$

$$\rho(\underline{x}^\mu, \underline{x}^\nu) \leq \rho(\underline{x}^\mu, \underline{x}^\eta) + \rho(\underline{x}^\eta, \underline{x}^\nu) \quad - \text{warunek trójkąta.}$$

Gdzie:

$X$  - przestrzeń wektorów

$R_*$  zbiór liczb nieujemnych

### Zastosowanie praktyczne

obliczanie **różnic** pomiędzy poszczególnymi obrazami

## Metryka Euklidesowa

$$\rho_1(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{\nu=1}^n (x_\nu^\mu - x_\nu^\eta)^2}$$

Wady i zalety metryki Euklidesowej

- odpowiada obiegowej definicji odległości
- ignorowanie składowych o bardzo małych wymiarach, długie czasy obliczeń (pierwiastkowanie, podnoszenie do kwadratu)

## Uogólniona metryka Euklidesowa

$$\rho_2(\underline{x}^\mu, \underline{x}^\eta) = \sqrt{\sum_{\nu=1}^n [\lambda_\nu (x_\nu^\mu - x_\nu^\eta)]^2}$$

$\lambda_\nu$  - mnożniki normalizujące

## Metryka uliczna (Manhattan, city block distance):

$$\rho_3(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n |x_v^\mu - x_v^\eta|$$

## Uogólniona metryka uliczna

$$\rho_4(\underline{x}^\mu, \underline{x}^\eta) = \sum_{v=1}^n \lambda_v |x_v^\mu - x_v^\eta|$$

## Metryka Czebyszewa (maksymalna)

$$\rho_5(\underline{x}^\mu, \underline{x}^\eta) = \max_{1 \leq v \leq n} |x_v^\mu - x_v^\eta|$$

Oznaczenia       $\underline{x}^\mu = [x_1^\mu, \dots, x_v^\mu, \dots, x_n^\mu]^T$      $\underline{x}^\eta = [x_1^\eta, \dots, x_v^\eta, \dots, x_n^\eta]^T$

Obliczyć różnicę (stosując metrykę euklidesową) pomiędzy dwoma obrazami trzypikslowymi     $\underline{d}^1 = [d_1^1, d_2^1, d_3^1]^T = [2, 3, 5]$ ,     $\underline{d}^2 = [d_1^2, d_2^2, d_3^2]^T = [1, 4, 5]$

$$\rho(\underline{d}^1, \underline{d}^2) = \sqrt{\sum_{i=1}^3 (d_i^1 - d_i^2)^2} = \sqrt{(2-1)^2 + (3-4)^2 + (5-5)^2} = \sqrt{2}$$

# KOMPRESJA OBRAZÓW

## Cele kompresji

- archiwizacja,
- przesyłanie.

## Stopień kompresji obrazu

$$SK = \frac{KP}{KW}$$

**SK** - stopień kompresji obrazu.

**KP** [abajt] - obszar pamięci zajmowany przez **kod pierwotny** obrazu.

**Kod pierwotny** - reprezentacja rastrowa lub wektorowa.

**KW** [abajt] - obszar pamięci zajmowany przez **kod wynikowy** obrazu.

## Rodzaje kompresji:

- kompresja bezstratna (*lossless compression*)
- kompresja stratna (*lossy compression*)

# Rodzaje kompresji (kodowania)

- **Bezstratna** (ang. lossless coding) odwracalna (ang. reversible) redukcja redundancji statystycznej (ang. statistical redundancy) w czasie i przestrzeni
- **Stratna** (ang. lossy coding) nieodwracalna (ang. irreversible) redukcja redundancji subiektywnej (ang. subjective redundancy), dotycząca nieistotnej informacji (ang. irrelevancy), z lub bez uwzględnienia charakterystyki słuchu i wzroku człowieka
- „Prawie” bezstratna lub **percepcyjnie bezstratna**, tzn. stratna, ale poziom zmian względem oryginału nie jest odczuwany przez człowieka

# Kompresja bezstratna i stratna

**lossless compression**

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) = 0$$

**lossy compression**

$$\rho(\underline{x}^{\mu}, \underline{x}^{\nu}) \neq 0$$

$\underline{x}^{\mu}$  - wektor reprezentujący obraz pierwotny

$\underline{x}^{\nu}$  - wektor reprezentujący obraz odtworzony (zrekonstruowany)

$\rho$  - wartość metryki

## Wybrane metody kompresji bezstratnej obrazów rastrowych:

- Kompresja obrazów z obszarami o jednolitej jasności
- Kompresja obrazów z obszarami o niejednolitej jasności

## Wybrane metody kompresji stratnej obrazów rastrowych

- kodowanie różnic,
- kodowanie blokowe

**Kompresja stratna** - uzyskiwany jest większy stopień kompresji (SK) niż przy stosowaniu kompresji bezstratnej.

# Kompresja obrazów z obszarami o jednolitej jasności

## Kodowanie ciągów identycznych symboli (RLE)

ciąg identycznych symboli - para zawiera 1 symbol i liczbę jego powtórzeń

### Przykład:

Kodowanie dziesiętne:

7,7,8,9,10,10,10,10,9,9,9,8,7,7,7

7(2), 8(1), 9(1), 10(4), 9(3), 8(1), 7(3).

Kodowanie binarne: 111, 111 .....

Długość kodu zależy:

- **od rozmiaru obrazu ( $N_1 \times N_2$ ),**
- **od liczby poziomów jasności obrazu (M)**

**Uwaga:** dla obrazów o dużych obszarach o jednolitej jasności przegląd według krzywej Hilberta daje z reguły dłuższe ciągi identycznych symboli niż w przypadku przeglądu "linia po linii"

### Praktyczna realizacja

metody kodowania ciągów identycznych symboli:

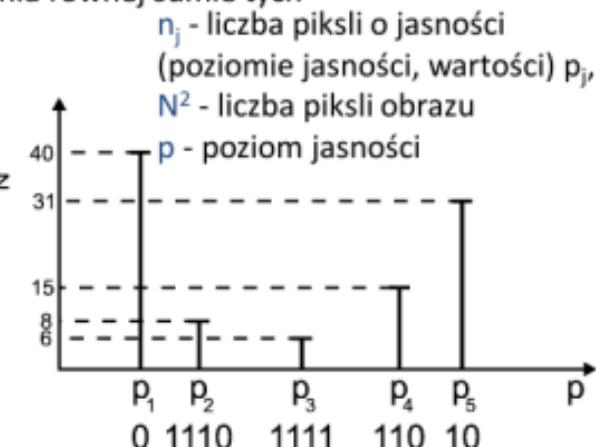
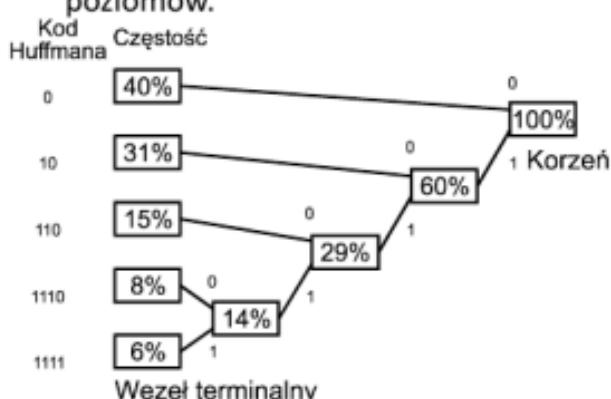
- formaty obrazowe: \*.PCX, \*.PIC

# Kompresja obrazów z obszarami o niejednolitej jasności

## Kod Huffmana - kod o zmiennej długości słowa.

Sposób postępowania:

- Przypisanie każdemu poziomowi jasności częstości występowania pikseli o tym poziomie jasności (na podstawie histogramu).
- Wyszukanie 2 poziomów o najmniejszej częstości występowania i połączenie w jeden o częstości występowania równej sumie tych poziomów.



Schemat łączenia - drzewo Huffmana.

Praktyczna realizacja kodu Huffmana - format obrazowy \*.TGA

# Metody słownikowe

Kodowanie *ciągów symboli* (pikseli) za pomocą odwołań do *słownika* zawierającego takie ciągi.

Stopień kompresji (SK) rośnie, gdy **długość ciągu pikseli** możliwych do zastąpienia *indeksem* do słownika rośnie.

Podział metod słownikowych:

- *statyczne* (słownik nie zmienia się w trakcie kodowania),
- *adaptacyjne* (słownik zmienia się w trakcie kodowania)

# Kodowanie różnic

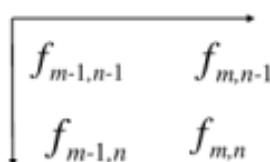
$$\epsilon_{m,n} = f_{m,n} - f_{m,n-1}$$

gdzie:

$f_{m,n}$  - poziom jasności piksela o współrzędnych  $m,n$ ,

$f_{m,n-1}$  - poziom jasności piksela (**kolejnego**) o współrzędnych  $m,n-1$ .

Położenie pikseli



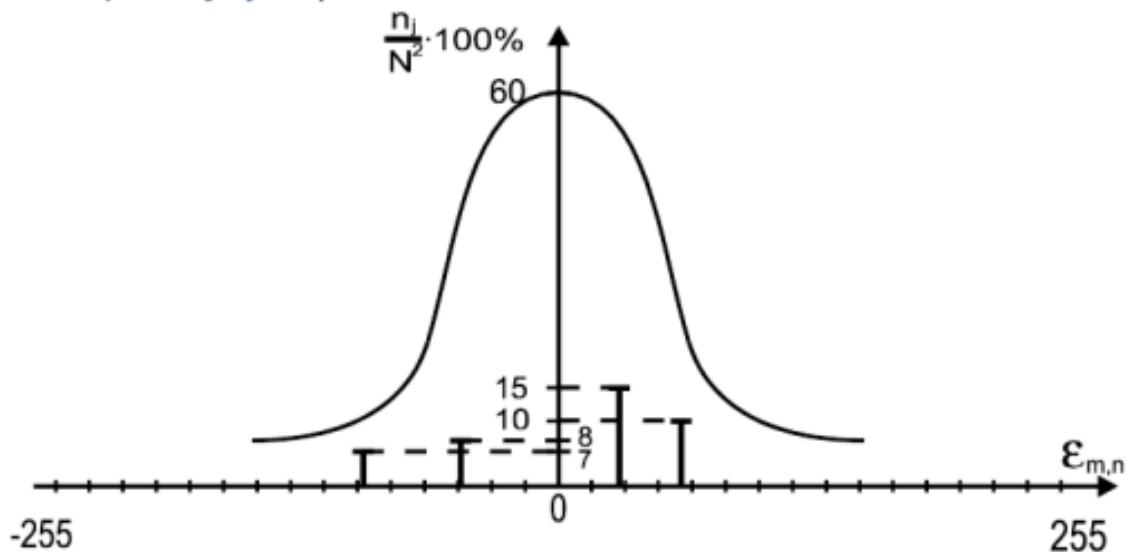
Przykład:

Obraz pierwotny :  $L = 255$  ( $M = 256$ )

- kodowanie wartości pikseli:  $I = 0,1,2,3,\dots,255$ ; zajętość pamięci: 8 bitów/piksel,
- kodowanie różnic pomiędzy wartościami kolejnych pikseli:

$$\epsilon_{m,n} = -255, -254, \dots, 0, \dots, 254, 255 \quad \text{zajętość pamięci: 9 bitów / piksel.}$$

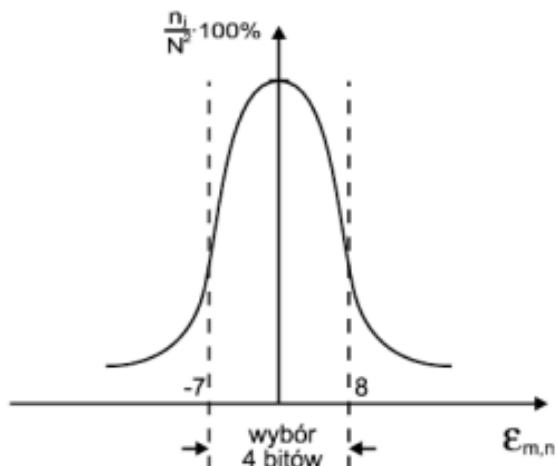
*Histogram różnic  $\varepsilon_{m,n}$  pomiędzy wartościami (poziomami jasności) kolejnych pikseli*



Histogram różnic ma charakter *krzywej Gaussa* z maksimum w okolicy zera.

# Kompresja wynika z zakodowanie najczęściej występujących różnic

Histogram różnic z zaznaczonym obszarem najczęściej występujących różnic



## Kryteria wyboru obszaru:

- wymagania na wielkość stopnia kompresji (SK),
- wymagania na dokładność rekonstrukcji (określoną wielkością „ $p$ ”)

$$KP = 9, KW = 4$$

$$SK = KP/KW = 9/4 = 2,25$$

Obraz zakodowany  $\epsilon_{m,n} = -7, -6, \dots, 7, 8$  - 16 symboli zamiast 512

Niepożądany efekt: zależnie od rodzaju obrazu - większe lub mniejsze rozmywanie (*blurring*) ostrych krawędzi.

## Kodowanie blokowe

Podział obrazu na jednakowe bloki, najczęściej  $4 \times 4$  piksele. Obliczenie dla każdego bloku średniej arytmetycznej jasności. Podział pikseli na dwie grupy:

- a) o jasności większej lub równej jasności średniej,
- b) mniejszej niż jasność średnia.

Obliczenie nowej jasności średniej dla każdej z grup (**wartość góra** dla (a) i **wartość dolna** dla (b)). Przypisanie wszystkim pikselom danej grupy obliczonej jasności średniej (górną lub dolną), stąd blok zostaje zakodowany jako *mapa bitowa* określająca podział na grupy, plus dwie wartości jasności.

The diagram illustrates the process of block-based coding. On the left, a 4x4 grid of pixel values is shown. An arrow points from this grid to a 4x4 binary matrix on the right, which represents the bit map for the upper group. The pixel values are: Row 1: 13, 11, 10, 13; Row 2: 8, 9, 12, 15; Row 3: 6, 8, 11, 14; Row 4: 4, 7, 9, 12. The binary matrix has columns separated by vertical lines and rows by horizontal lines. The first column contains 1, 0, 0, 0. The second column contains 1, 0, 0, 0. The third column contains 0, 1, 1, 0. The fourth column contains 1, 1, 1, 1.

13	11	10	13		1	1	0	1
8	9	12	15		0	0	1	1
6	8	11	14		0	0	1	1
4	7	9	12		0	0	0	1

Średnia=10,125

Wartość góra= 12,625≈13  
Wartość dolna = 7,625≈ 8

# Kompresje obrazów oparte na transformatach

- Wybieram takie transformaty, które dają w dziedzinie transformacji takie upakowanie informacji, aby większość współczynników była zerowa (bliska零)
- Po kwantyzacji niezerowych współczynników dają minimalne zniekształcenia obrazu
- Prowadzą do dekorelacji sygnału
- Są ortogonalne, bo to zapewnia pełną transmisję informacji
- Są separowalne, bo to zapewnia szybki algorytm dekompozycji

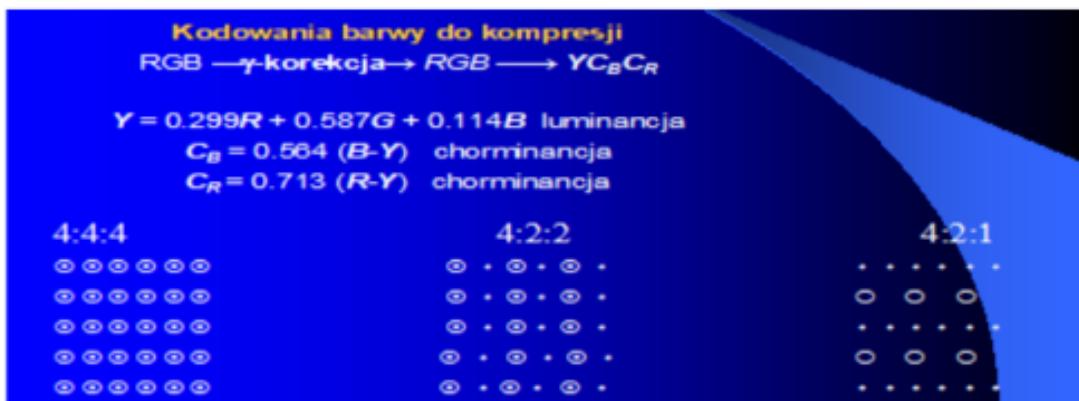
# Kompresja stratna oparta na transformacjach

- Transformacja kosinusowa w kompresji JPEG
- Transformacja falkowa w kompresji JPEG 2000

## Kolejność działań (i):

Konwersja obrazu kanałów (RGB) na jasność (luminancję) i 2 kanały barwy (chrominancje).

Odrzucenie części pikseli kanałów barwy, podział kanałów na bloki  $8 \times 8$  pikseli.



## Kolejność działań (ii)

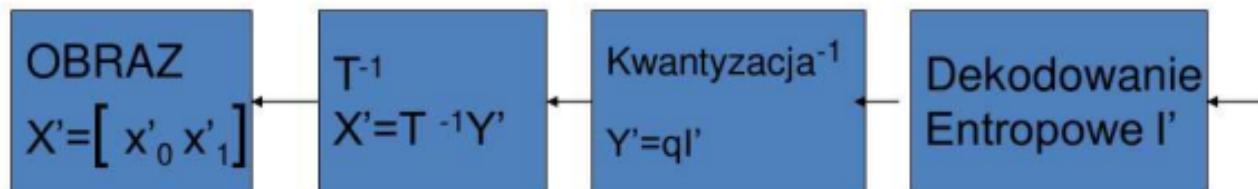
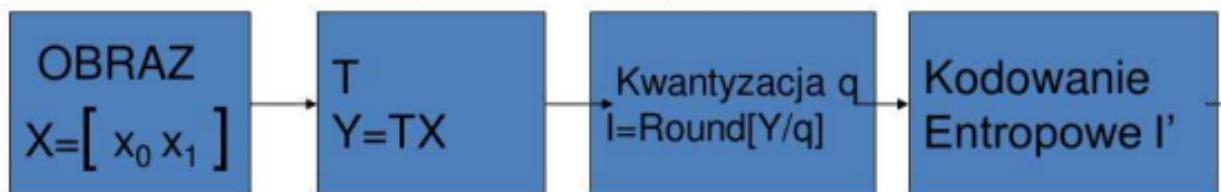
Na blokach wykonywana jest dyskretna transformata kosinusowa (DCT).

Zastąpienie średnich wartości bloków przez różnice wobec wartości poprzedniej (DPCM). Poprawia to w pewnym stopniu współczynnik kompresji.

Kwantyzacja, czyli zastąpienie danych zmiennoprzecinkowych przez liczby całkowite.

# Schemat kompresji opartej na transformacjach

Kodowanie (ang. encoder)



Dekodowanie (ang. decoder)

# Kompresja JPEG

- Odpowiada kompresji stratnej
- Oparty na dyskretnej transformacji kosinusowej (DCT)

$$\exp\left[-j \frac{2\pi}{N} (ki)\right] = \cos\left(\frac{2\pi}{N} (ki)\right) - j \sin\left(\frac{2\pi}{N} (ki)\right)$$

$$F(u, v) = \frac{4c(u, v)}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \frac{(2x+1)\pi u}{2N} \cos \frac{(2y+1)\pi v}{2N}$$

$$c(u, v) = \begin{cases} \frac{1}{2} & u = 0 \text{ lub } v = 0 \\ 1 & \text{pozostało} \end{cases}$$

- Wraz ze wzrostem stopnia kompresji (SK) coraz bardziej widoczny jest podział na bloki w obrazie zrekonstruowanym

## Przykład



Obraz	Jakość	Rozmiar (abajty)	Współczynnik kompresji	Komentarz
	Pełna jakość (Q = 100)	83,261	2,6:1	Niezauważalnie małe artefakty



	Niska jakość (Q = 10)	4,787	46:1	Duże straty rozciągłości tonalnej; artefakty na krawędziach obiektów
--	-----------------------	-------	------	--

Innowacyjność algorytmu JPEG: możliwość kontroli stopnia kompresji w jej trakcie

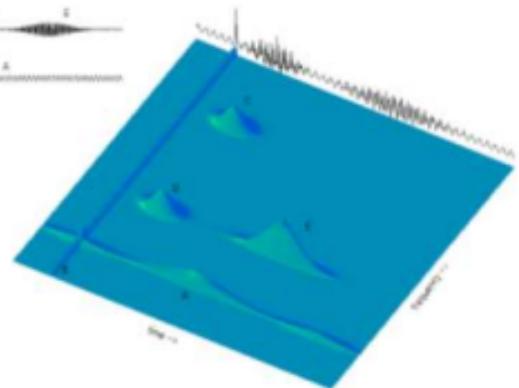
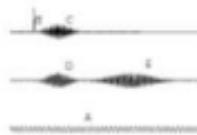
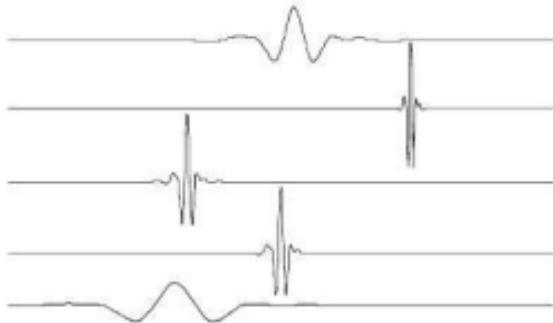
# Kompresja JPEG200

- Koduje różne typy obrazów i filmów: binarne, szare, kolorowe, wielospektralne o różnych właściwościach, transmisja w czasie rzeczywistym, archiwizacja biblioteki obrazów.
- Umożliwia progresywne dekodowanie: pozwala widzowi zobaczyć wersję danego obrazu o niższej jakości, podczas gdy cały plik nie został w pełni odebrany lub pobrany. Jakość obrazu stopniowo się poprawia, gdy więcej bitów danych jest pobieranych ze źródła.
- Zapewnia zachowanie przejrzystości obrazów.
- Daje niezwykle wysoki poziom skalowalności.
- Obrazy mogą być kodowane w dowolnej jakości od bardzo skompresowanych do kompresji bezstratnej.
- Firma udostępnia oprogramowanie open-source tzn. OpenJPEG kod napisany w C do wykorzystania przez programistów.
- Stosuje transformację falkową

# Co to są falki?

**Komplet funkcji umożliwiających hierarchiczną wielorodzicielną reprezentację** aproksymowanej funkcji, składający się z :

- Funkcji skalującej  $\Phi(x)$
- Rodziny falek generowanych na podstawie falki podstawowej



# Falki i transformacja falkowa w przetwarzaniu obrazów

- Kompresja obrazów i sekwencji wideo
- Odszumianie obrazów
- Analiza tekstury
- Segmentacja
- Rozpoznawanie np.: ręcznego pisma, biologicznego wieku dzieci, charakterystycznych cech biologicznego sygnału (EEG)
- Rekonstrukcja obrazów biomedycznych

# Metody oceny jakości kodowania

- Subiektywne – notowanie wrażeń widzów lub słuchaczy
- Obiektywne – porównanie ilościowe, na podstawie matematycznie wyznaczonych parametrów, opisujących różnicę między sygnałem oryginalnym i po dekompresji
- Metody oparte na modelach percepcji informacji przez człowieka:
  - perceptualny model, zapewniający wierność percepcji słuchowej
  - modelowanie układu percepcji wzrokowej człowieka  
(HVS ang. Human Visual System)

# Subiektywna ocena jakości kodowania

- MOS (ang. Mean Opinion Score) – uśredniona ocena słuchaczy lub widzów, nie będących ekspertami

Norma dla sygnałów telewizyjnych ITU-R, zalecenie BT.500:

15 obserwatorów ogląda nie dłużej niż 30 min szereg 10-15 sekundowych prezentacji obrazów statycznych i sekwencji wizyjnych, będących międzynarodowymi standardami.

Zastosowane wymuszenia: pojedyncze (bez obrazu odniesienia) lub podwójne (z obrazem odniesienia), oddzielane szarym obrazem.

Bezpośrednio po prezentacji oceniają w ciągłej lub skokowej skali pięciostopniowej dwie cechy: jakość obrazu i poziom zakłóceń.

*Inne protokoły miar: absolutnych i porównawczych (szeregujących)*

**Wady:** zależność oceny od kontekstu i kolejności

**Zalety:** lepsze niż obiektywne

# Obiektywna ocena jakości kodowania

- Średnia różnica
- Maksymalna różnica
- Błąd średniokwadratowy (ang. Mean Square Error)
- Szczytowy błąd średniokwadratowy (ang. Peak Mean Square Error)
- Znormalizowany błąd średniokwadratowy (ang. Normalized Mean Square Error)
- Laplasjanowy błąd średniokwadratowy (ang. Laplacian Mean Square Error)
- Jakość korelacji (ang. Correlation Quality)
- Znormalizowana korelacja skośna (ang. Normalized Cross-Correlation)
- Zawartość strukturalna (ang. Structural Content)
- Wierność obrazu (ang. Image Fidelity)
- Ważone odległości obrazów według różnych norm, np. normy Minkowskiego
- Inne, oparte na graficznych miarach jakości obrazów, liczone np. na podstawie histogramu lub wykresu Hosaka, obrazującego degeneracje jakości