

Systemy Wizyjne

Opracowanie do zaliczenia cz. 1

Zakres : Lab 1 – Lab 5

Laboratorium 1

Zagadnienia :

- Typy obrazów
- Modele barw

Pytania :

- Czym jest piksel ? Czym jest rozdzielczość ekranu ? Czym jest rozdzielczość obrazu ?
- Czym jest typ obrazu ? Wymień przykładowe.
- Czym są obrazy indeksowane ? Co wykorzystują (jaki sposób odnoszenia się do kolorów) ?
- Czym są obrazy w skali szarości ? Dlaczego się ich używa ?
- Jakie znasz i czym są modele barw ? Omów RGB, HSV, CMYK.

Pojęcia podstawowe – piksel, rozdzielczość ekranu / obrazu

Obraz składa się z **pikseli**, czyli małych kwadracików. Każdy piksel przedstawia konkretny kolor i w ten sposób tworzony jest obraz.

Rozdzielczość ekranu jest to parametr określający liczbę pikseli wyświetlanych na ekranie. Przykładowo matryca full HD o rozdzielczości 1920 x 1080 ma 1920 pikseli w poziomie i 1080 pikseli w pionie. Rozdzielczość jest większa w telefonach komórkowych i smartfonach niż np. w telewizorach ze względu na odległość z jakiej korzystają z nich użytkownicy.

Rozdzielczość obrazu jest to (maksymalna) ilość pikseli obrazu możliwa do wyświetlenia na ekranie jakiegoś urządzenia.

Typy obrazów

Typ obrazu jest to sposób w jaki wartości macierzy obrazu są połączone z kolorami, które mają być wyświetlane dla danego piksela. Przykładowe typy obrazu to **True color images, Scaled Indexed images, Binary images, Grayscale images** lub **Indexed images – obrazy indeksowane**. Na pierwszych laborkach były używane Grayscale i Indexed i zostaną tu omówione.

Obrazy indeksowane (indexed)

W obrazach indeksowanych wartości z macierzy obrazu nie przedstawiają kolorów pikseli wprost. Zamiast tego te wartości używane są jako indeksy do wyszukania kolorów w mapie kolorów związanej z danym obrazem.

Obraz składa się więc z tablicy Look-Up Table, a także mapy kolorów. Aby sprawdzić kolor danego piksela musimy najpierw odczytać jego indeks z tablicy Look-Up Table (macierzy obrazu) , a następnie wiersz o tym indeksie wyciągnąć z mapy kolorów obrazu.

Wiersz ten zawiera trzy składowe R (red) G (green) B (blue), które są reprezentowane liczbowo i które są mieszane w celu osiągnięcia danego koloru.

Natomiast Scaled Indexed Images to obrazy z przeskalowaną reprezentacją kolorów, np. zamiast od 0 – 255 przeskalowane jest na 0 – 36.

// Jakby się zastanowić to daje to oszczędność, ponieważ kilku pikselom można przypisać jeden wiersz i te piksele nie muszą każdy osobno przechowywać kilku liczb, tylko przechowują jeden i ten sam numer wiersza.

Obrazy w odcieniach szarości (Grayscale)

Często aby wydobyć potrzebne informacje z obrazu (np. kształty) nie są nam potrzebne kolory. W takim przypadku lepszą postacią do analizy obrazu jest obraz w skali szarości. Macierz takiego obrazu zawiera informację o poziomie szarości danego piksela w skali od 0 – 255 (256 wartości, 255 - biały, 0 – czarny, 8 bit = 2 do potęgi 8). x, y to współrzędne obrazu , a $L(x,y)$ reprezentuje poziom szarości.

Funkcja `rgb2gray` pozwala w matlabie na przekonwertowanie obrazu z kolorowego obrazu indeksowanego na obraz w skali szarości.

Systemy (modele) barw

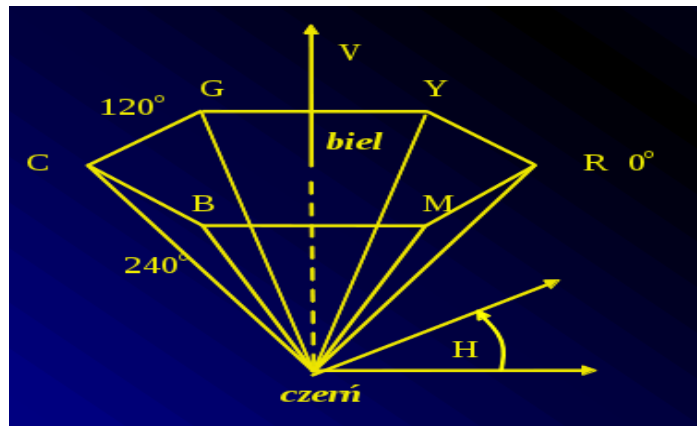
Istnieje wiele modeli zapisywania informacji o barwie w obrazie. Najbardziej znane to RGB, CMYK, HSV, YCrCb. Każdy model jest dobry do czegoś innego.

Model RGB (red-green-blue) – jest najbardziej znany i najpowszechniej używany w grafice m.in. do wyświetlania i akwizycji. Jest to metoda addytywna, oznacza to, że wykorzystuje sumowanie barw podstawowych : czerwonej, zielonej, niebieskiej (zazwyczaj wartości od 0 do 255) .

Model CMYK (Cyan, Magenta, Yellow) – ma szerokie zastosowanie w druku. Jest to model subtraktywny – im większa wartość danej składowej tym kolor jest ciemniejszy, bo każda składowa opisuje jak bardzo dana składowa jest ze światła pochłaniana. Model posiada dodatkową składową czarną ze względu na to, że w drukarkach łatwiej jest wydrukować coś na czarno niż uzyskiwać kolor czarny z pozostałych trzech składowych. Jest to przeciwieństwo RGB, aby uzyskać wartości w CMY wystarczy odjąć od 255 wartości dla RGB.

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Model HSV (Hue - barwa, Saturation – nasycenie, Value – wartość) – Do wartości hue przypisujemy kąt od 0 do 360 (mamy tak jakby koło w którym barwa zmienia się w zależności od stopni – patrz obrazek następna strona). 0 stopni – czerwony, 120 – zielony, 240 – niebieski. Składowa S to tak jakby promień koła, im bliżej koła jesteśmy, to kolor jest bliżej białego, im bliżej obwodu koła tym bliżej czystego koloru. Ostatnia składowa Value tworzy jakby stożek, na czubku jest kolor czarny, a w podstawie biały. Składowa ta oznacza jasność koloru. Model przydatny do prezentacji np. zjawisk w których kolor oznacza natężenie czegoś.



(Można uzupełnić opracowanie o model YCrCb.)

Laboratorium 2

Zagadnienia :

- Rozdzielczość przestrzenna,
- Metody zamiany rozdzielczości przestrzennej,
- Poziomy jasności obrazu

Pytania :

- Czym jest rozdzielczość przestrzenna ? Czym jest skalowanie obrazu ? Jak rozdzielczość jest oznaczana w programach graficznych ?
- Jakie znasz metody zmiany rozdzielczości obrazu ? Opisz.
- Jakie znasz rodzaje zmiany rozdzielczości przez interpolację. Opisz je.
- Czym są poziomy jasności ?

Rozdzielczość przestrzenna

Rozdzielczość przestrzenna to ilość pikseli z których składa się obraz (inaczej gęstość punktów próbkowania). Ilość tych pikseli, czyli rozdzielczość przestrzenną można zmieniać (zmniejszać / zwiększać) co jest nazywane skalowaniem obrazu. Przeskalowanie do innego rozmiaru nie zwiększa ilości informacji, a jedynie ilość pikseli oraz powoduje zniekształcenia. Im obraz ma mniejszą rozdzielczość tym jest mniej wyraźny, ale zwiększając jego rozdzielczość nie sprawimy, że będzie bardziej wyraźny (jakoś muszą zostać wypełnione dodatkowe piksele). W programach graficznych rozdzielczość jest często oznaczana 'Image Size' lub 'Pixel Dimension' . Natomiast pod pojęciem

rozdzielczość 'Resolution' rozumie się parametr DPI (dots per inch), czyli ilość pikseli na cal.

W matlabie do zmiany rozdzielczości przestrzennej służy *imresize*.

Metody

Do wyznaczenie wartości brakujących pikseli po przeskalowaniu stosuje się metody :

- **najbliższego sąsiada** – realizowane poprzez kopiowanie najbliższego piksela, metoda wymagająca najmniejszej mocy obliczeniowej, ale też dla dużych powiększeń wyraźnie widać grupy jednakowych pikseli

- **interpolacji**

a) dwuliniowa – piksele są powielane lub redukowane z uwzględnieniem czterech pikseli stykających się bokami z danym pikselem, daje dość dobry efekt, z rozmytymi krawędziami, wymaga więcej mocy obliczeniowej niż algorytm najbliższego sąsiada

b) dwusześcienna (dwukubiczna) – piksele są powielane lub redukowane z uwzględnieniem wszystkich ośmiu pikseli otaczających dany piksel, daje najlepszy efekt końcowy, ale jest najbardziej złożona obliczeniowo ze wszystkich trzech metod

Ilość poziomów jasności

Dla obrazów w skali szarości pojedynczy piksel zapisuje się zazwyczaj na 8 bitach, co daje 256 rozróżnialnych poziomów szarości. Dla większości zastosowań wartość ta jest wystarczająca. Oko ludzkie nie potrafi rozróżnić wszystkich 256 poziomów jasności (jest za mało czułe). Zazwyczaj człowiek rozróżnia 20-30 poziomów szarości. W matlabie można przeskalować obraz tak aby zawierał mniej poziomów szarości np. 0 – 15. Służy do tego funkcja *imadjust*.

Na laboratorium zostały też omówione transformację przestrzenne (np. rotacja o 45 st.), ale z uwagi na prostotę materiału nie ma o nich tutaj notatki.

Laboratorium 3

Zagadnienia :

- Podstawowe operacje przeprowadzane na obrazach.
- LUT
- Operacje Arytmetyczne
- Kombinacja liniowa
- Operacje Logiczne

Pytania :

- Jak można zakwalifikować operacje na obrazach ze względu na piksele lub ze względu na ilość argumentów ?
- Jakie są podstawowe operacje na obrazach ? (5)
- Jak działa LUT ? Jakie tablice przekodowań wykorzystuje ? Podaj przykłady tablic i zastosowań.
- Jak są realizowane operacje arytmetyczne ? Co się w nich dodaje/odejmuje ?
- Na co trzeba uważać przy dodawaniu, odejmowaniu obrazów ? Co trzeba zrobić aby temu zapobiec ?
- Czy mnożenie obrazów ma jakiś sens ?
- Na jakich obrazach się zazwyczaj wykonuje operacje logiczne ?

Typy operacji na obrazach

Operacje na obrazach mogą być :

- jednopunktowe – piksel wynikowego obrazu obliczany na podstawie jednego piksela punktu pierwotnego,
- kontekstowe – na podstawie wielu pikseli obrazu pierwotnego

Inne kryterium :

- jednoargumentowe – jeden obraz jest argumentem,
- dwuargumentowe – dwa obrazy tworzą obraz wynikowy

Podstawowe operacje na obrazach :

- LUT – Look-Up Table,
- arytmetyczne,
- logiczne,
- binaryzacja,
- kombinacja liniowa

Operacja LUT – Look-Up Table

Jest to przekształcenie punktowe, jednoargumentowe (jeden piksel jest przekształcany na jeden piksel, z jednego obrazu tworzymy inny obraz), piksel wynikowy powstaje poprzez operację na pojedynczym pikselu obrazu pierwotnego. LUT polega na przekształcaniu wartości poszczególnych pikseli obrazu przy użyciu z góry przygotowanych tabel przekodowań (tabel korekcji). Przekształcenie takie jest możliwe, ponieważ obraz wejściowy ma dyskretną i ograniczoną skalę obrazu.



W przetwarzaniu obrazów najczęściej wykorzystuje się następujące funkcje:

- typu kwadratowa, pierwiastek kwadratowy
- typu logarytm, odwrócony logarytm
- typu wykładnicza,
- inne (np. piłokształtna)



Przykładem LUT jest reprezentacja obrazu za pomocą obrazu indeksowanego gdzie przechowywane są kolory (składowe RGB) w innej tablicy, równoważenie histogramu lub np. operacja negacji zrealizowana jak niżej :



Operację LUT realizuje się w matlabie za pomocą funkcji *intlut*. Przy czym najważniejsze jest stworzenie odpowiedniej tablicy przekodowania.

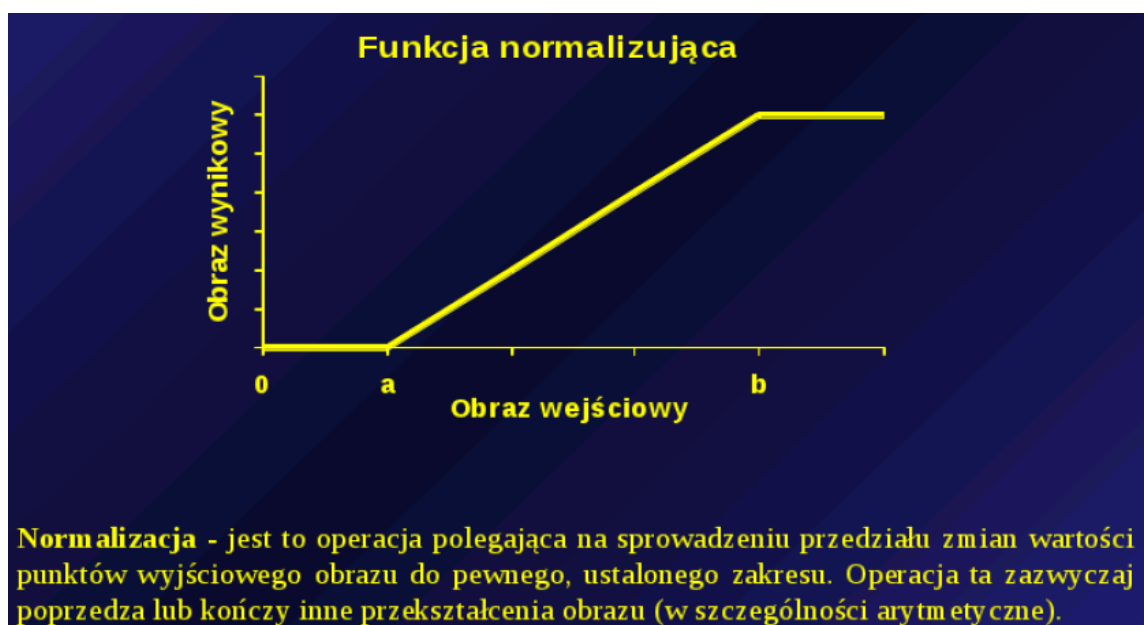
Operacje arytmetyczne na obrazach

Do podstawowych operacji arytmetycznych należą:

- **dodanie dwóch obrazów** $f_3(x,y) = f_1(x,y) + f_2(x,y)$
- **odjęcie dwóch obrazów** $f_3(x,y) = f_1(x,y) - f_2(x,y)$
- **przemnożenie dwóch obrazów** $f_3(x,y) = f_1(x,y) \cdot f_2(x,y)$
- **kombinacja liniowa dwóch obrazów** $f_3(x,y) = p \cdot f_1(x,y) + (1-p) \cdot f_2(x,y) \quad p \in [0,1]$
- **przemnożenie obrazu przez stałą** $f_2(x,y) = k \cdot f_1(x,y)$

O czym należy pamiętać przy dodawaniu / odejmowaniu ?

Przy sumowaniu/odejmowaniu należy pamiętać o normalizacji czyli o przeskalowaniu, bo np. dodając $236 + 210$ otrzymamy liczbę większą od 256 – niemożliwą do zapisania na 8 bitach, dlatego obraz wynikowy pasowałoby zapisać w innym formacie niż int, np. uint, później obraz przeskalować do 0 – 256 i wyświetlić, analogicznie dla innych operacji. Przy odejmowaniu można np. zamiast tego zrobić odejmowanie wartości bezwzględnej.




Obraz wejściowy na tym wykresie to wynik np. odejmowania.

Kombinacja liniowa obrazów

Kombinacja liniowa dwóch obrazów obliczana jest według wzoru:

$$f_3(x,y) = p \cdot f_1(x,y) + (1-p) \cdot f_2(x,y) \quad p \in [0,1]$$

Obraz wynikowy



Mnożenie obrazów – czy to ma sens ?

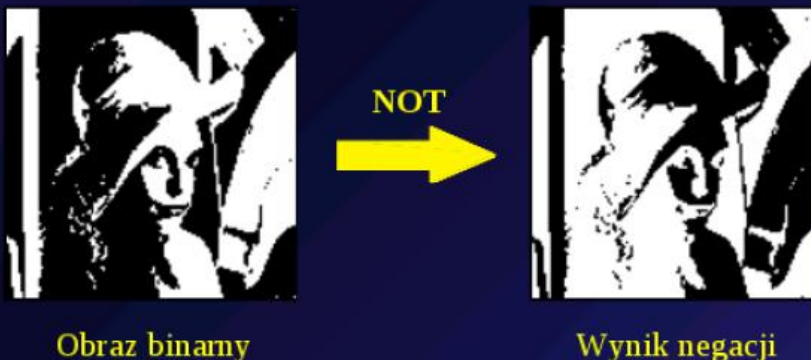
Wymnożone obrazy same w sobie nie zawierają żadnej istotnej informacji. Natomiast wykonywanie operacji mnożenia przez liczbę/maskę może dać już jakiś sensowny efekt.

Operacje logiczne na obrazach

Operacje logiczne wykonuje się zazwyczaj na obrazach binarnych.

Na poszczególnych punktach obrazu (najczęściej binarnego), możemy wykonywać następujące operacje:

- NOT - negacja obrazu
- AND - iloczyn logiczny obrazów
- OR - suma logiczna obrazów
- SUB - różnica logiczna
- XOR - suma rozłączna
- NXOR - równoważność logiczna



Laboratorium 4

Zagadnienia :

- Histogram, histogram kumulacyjny
- Rozciąganie histogramu
- Wyrównywanie histogramu
- Histogram kumulacyjny
- Dopasowywanie histogramu

Pytania :

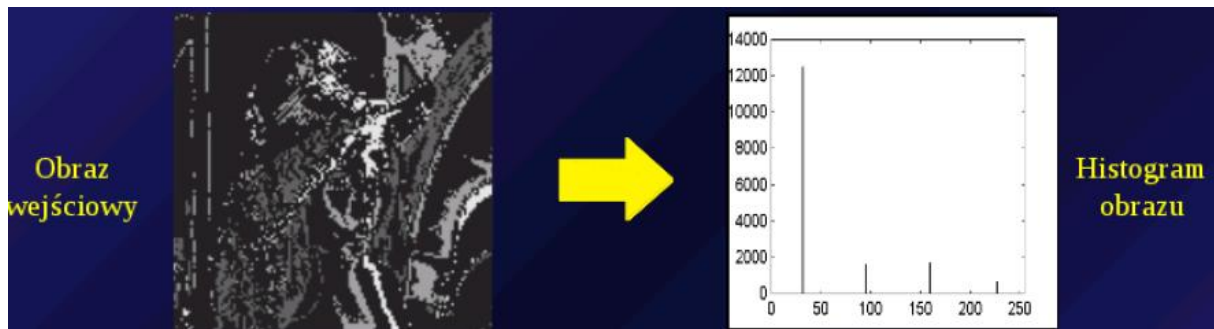
- Czym jest histogram ? Co zrobić aby uzyskać z niego gęstość prawdopodobieństwa (postać znormalizowana) ? Czy może być zdefiniowany dla obrazów kolorowych ? Do czego jest wykorzystywany ?
- Rozciąganie histogramu. Po co się stosuje ? Na czym polega ? Czy jest lepsze od wyrównywania ?
- Wyrównywanie histogramu – jak działa i jakie ma wady / zalety ? Jaki jest algorytm ?
- Czym jest histogram kumulacyjny ? Jaki jest algorytm jego tworzenia ?
- Jaka jest rola histogramu w binaryzacji ?
- Dopasowywanie histogramu – Na czym polega ? Kiedy się stosuje ?

Histogram

Histogram zawiera informację na temat tego ile pikseli o danym poziomie jasności występuje na obrazie (w przypadku obrazu w odcieniach szarości).

Często wykorzystuje się tzw. znormalizowaną postać histogramu - wszystkie wartości $h(i)$ są dzielone przez ilość pikseli na obrazku. Otrzymana w ten sposób wielkość to gęstość prawdopodobieństwa wystąpienia na obrazie piksela o odcieniu "i".

Przykład :



Histogram można zdefiniować dla obrazów kolorowych. Otrzymujemy wtedy 3 histogramy - po jednym dla danej składowej: R,G,B (lub HSV, YCrCb) lub histogram trójwymiarowy

Histogram jest bardzo użyteczny w przetwarzaniu obrazów. Wykorzystywany jest przy binaryzacji (ustalanie progu od którego uznajemy za jedynekę lub zero) oraz do oceny jakości (dynamiki, kontrastu) obrazu.

W idealnym przypadku wszystkie poziomy jasności w obrazie powinny być wykorzystane (i to najlepiej w miarę jednolicie) - obrazowo mówiąc histogram powinien rozciągać się od 0 - 255 (obraz w skali szarości). W przypadku gdy wykorzystujemy jedynie fragment dostępnego zakresu (wąski histogram) lub histogram nie jest jednolity (występują dominujące grupy pikseli) obraz ma dość słaby kontrast. Z histogramu można wnioskować o ostrości i jasności obrazu.

Cechy te można poprawić stosując tzw. rozciąganie albo wyrównywanie histogramu .

Najprostszą metodą poprawienia jakości obrazu jest tzw. rozciągnięcie histogramu - takie przeskalowanie wartości pikseli w obrazie aby wykorzystać cały dostępny zakres (0-255).

Czasem operacja ta mylona jest z wyrównywaniem histogramu. Rozciągnięcie prowadzi do takiej konwersji zakresu wartości składowych, aby histogram obejmował wszystkie wartości składowych. Czyli jeżeli zakres składowej jest równy 0-255, a najmniejsza wartość w obrazie wynosi 4, największa natomiast wynosi na przykład 198, to po operacji rozciągnięcia wartości tej składowej będą w pełnym zakresie 0-255.

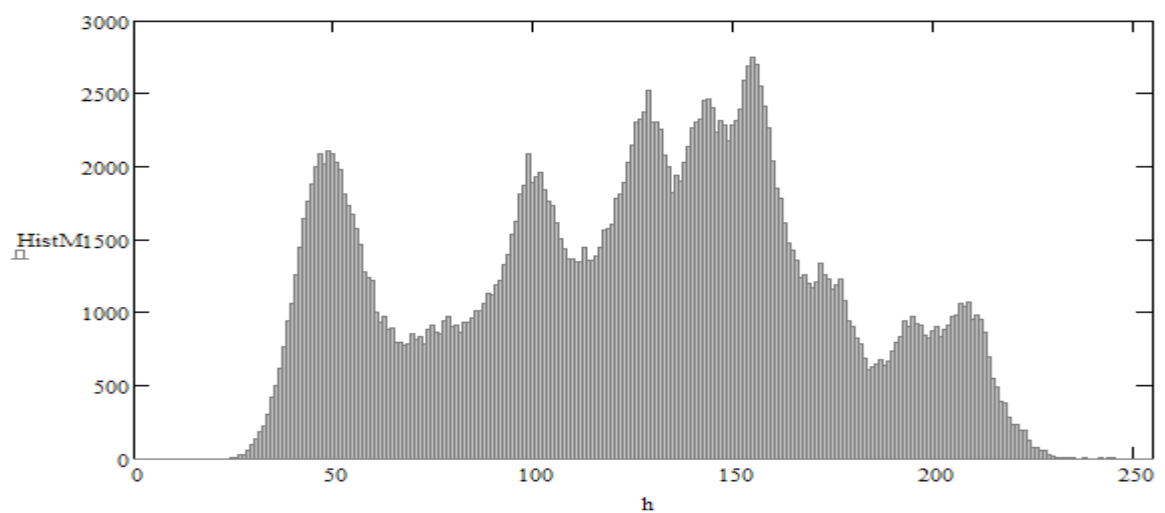
Czyli teraz najmniejsza wartość w obrazie wynosi 0 a największa 255. Operacje rozciągnięcia histogramu można przeprowadzić odpowiednio dobierając jasność i kontrast obrazu.

W praktyce rozciągnięcie histogramu sprowadza się do wykonania przekształcenia obrazu przy pomocy odpowiednio przygotowanej tablicy LUT. Współczynniki w takiej tablicy wyznaczone są wg następującego wzoru:

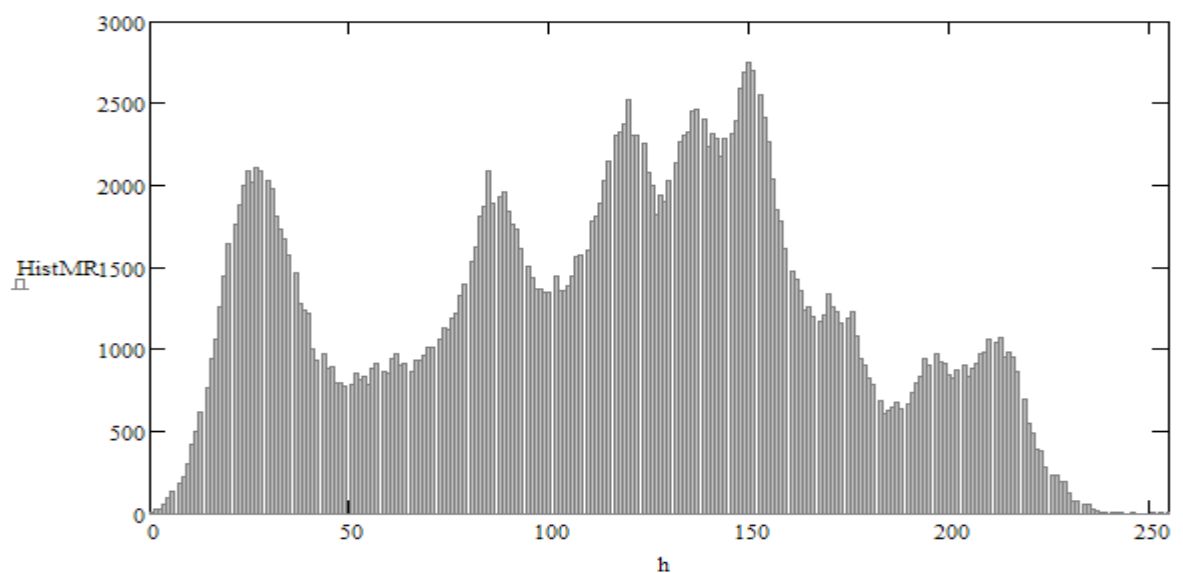
$$L'(w, k) = \frac{255}{\max - \min} (L(w, k) - \min)$$

max - to maksymalna wartość występująca w obrazie źródłowym,
min - to minimalna wartość występująca w obrazie źródłowym

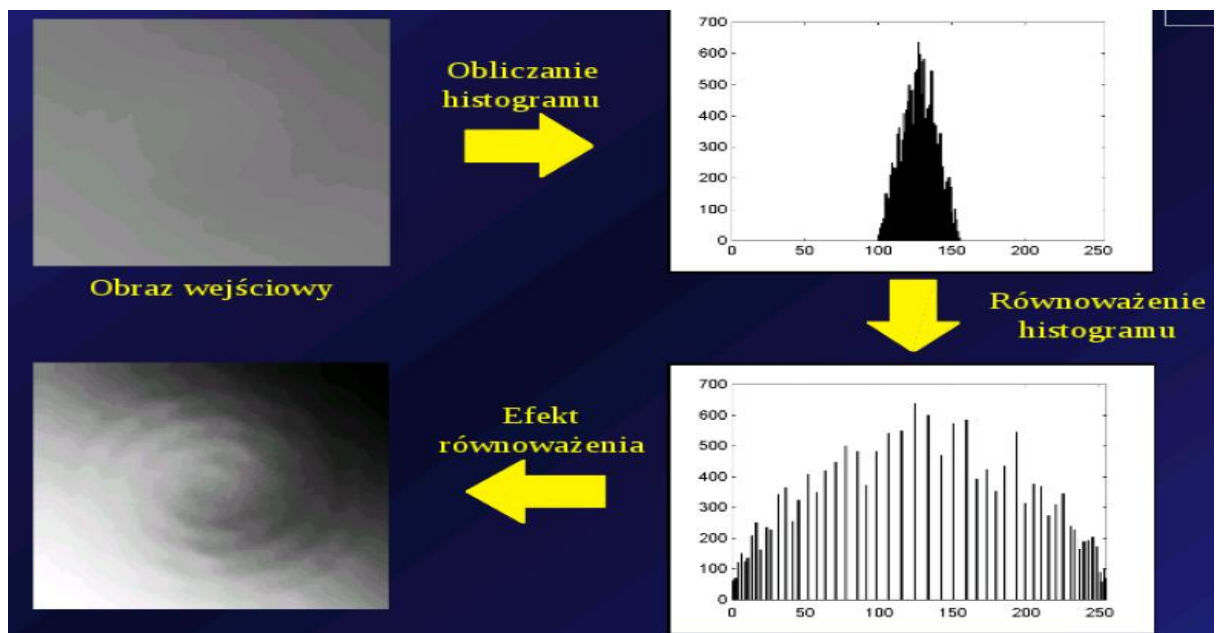
Przed :



Po :



Bardziej zaawansowaną metodą jest tzw. Wyrównywanie (inna nazwa równoważenie) histogramu (HE). Idea jest następująca: z punktu widzenia lepszego wykorzystania dostępnych poziomów jasności lepiej by było gdyby rozciągnąć "szczyty" histogramu, a skompresować "doliny" tak aby taka sama ilość pikseli reprezentowana była przez każdy z dostępnych poziomów szarości – to odróżnia właśnie wyrównywanie od rozciągania. Warto zwrócić uwagę, że takie przekształcenie powoduje częściową utratę informacji o szczegółach w obszarach "dolin" - inaczej mówiąc dążymy do sytuacji aby histogram był jednostajny – płaski, linia prosta pozioma.



Jak wykonać równoważenie ? Przy użyciu histogramu kumulacyjnego.

Histogram kumulacyjny uzyskujemy przez całkowanie zwykłego histogramu. W algorytmie funkcją przejścia (LUT) jest histogram kumulacyjny.



Algorytm i idea (trochę niejasno wyjaśnione, ale pod obrazkiem przykład) :

Ideą stosowania równoważenia histogramu jest „rozzucenie” wartości histogramu, tak, aby zajęły cały dostępny zakres poziomów szarości. Dzięki temu uwydatniane są detale, które wcześniej nie były widoczne.

Czterokrokowy algorytm równoważenia histogramu:

1. Obliczenie histogramu skumulowanego
2. Normalizacja elementów histogramu skumulowanego przez podzielenie ich przez całkowitą sumę pikseli
3. Pomnożenie wartości otrzymanych w kroku drugim przez maksymalną wartość poziomu szarości i zaokrąglenie otrzymanych wartości do liczb naturalnych
4. Utworzenie obrazu wynikowego poprzez zastosowanie tablicy LUT będącej wyliczonym w poprzednich krokach histogramem skumulowanym.

Przykład (poniżej znajduje się obrazek do którego jest poniższy opis – odniesienia do poszczególnych punktów) :

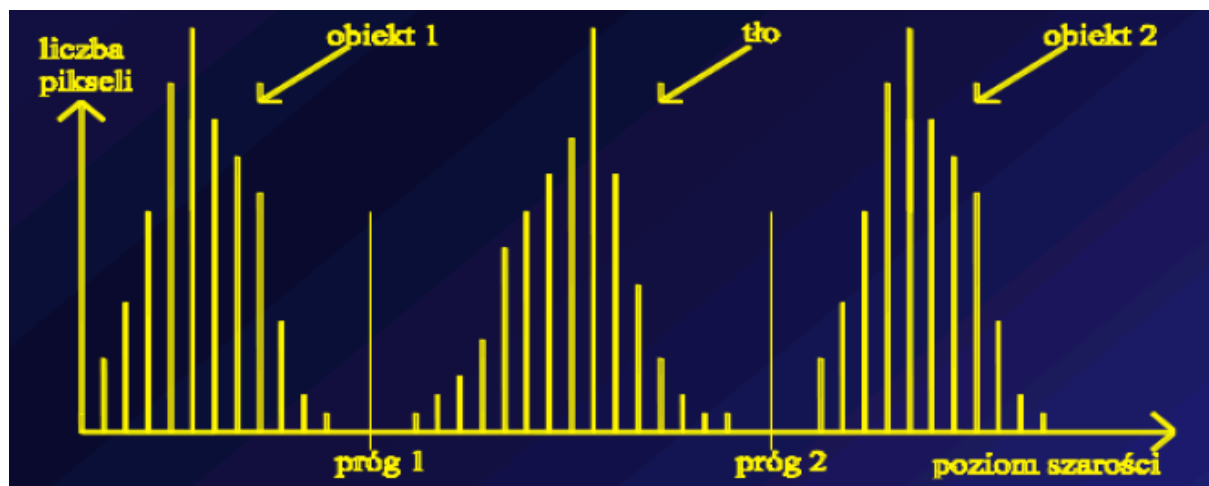
Po lewej stronie wartości obrazka (histogramu) wejściowego.

1. Histogram skumulowany
2. Podzielenie przez całkowitą sumę pikseli
3. Pomnożenie przez maksymalny poziom szarości obrazu pierwotnego i zaokrąglenie – w tym przypadku przez 7. Otrzymuje histogram akumulacyjny wykorzystany do LUT.
4. Przekształcenie LUT.

Po prawej obrazek (histogram) wyjściowy.

		1.	10, 18, 27, 29, 43, 44, 49, 51			
Histogram oryginalny		2.	10/51, 18 /51, 27 /51, 29 /51, 43 /51, 44 /51, 49 /51, 51 /51		Histogram wynikowy	
Poziom szarości	Ilość pikseli	3.	1, 2, 4, 4, 6, 6, 7, 7		Poziom szarości	Ilość pikseli
0	10	4.	Oryginalne wartości pikseli	Nowe wartości pikseli	0	0
1	8				1	10
2	9				2	8
3	2				3	0
4	14				4	11
5	1				5	0
6	5				6	15
7	2				7	7

Histogram w binaryzacji – idea prezentowana przez poniższe obrazki



Klasyczne wyrównywanie histogramu nie zawsze jest najlepszym rozwiązaniem - w przypadkach, kiedy nie uzyskuje się zadowalających rezultatów można spróbować techniki dopasowywania histogramu, w której podaje się docelowy kształt histogramu.

(Na tych laborkach były jeszcze zagadnienia dodatkowe – nadobowiązkowe dla chętnych - , które tu nie zostały opracowane, ale można je dodać)

Laboratorium 5

Zagadnienia :

- Przekształcenia kontekstowe.
- Filtry
- Filtry konwolucyjne
- Filtracja pikseli skrajnych
- Filtr dolnoprzepustowy = uśredniający
- Różnica między filtrami liniowymi, a filtrami nieliniowymi
- Filtr medianowy
- Inne filtry nieliniowe
- Filtry liniowe górnoprzepustowe
- Filtry górnoprzepustowe – Laplace’a
- Unsharp Masking
- Filtry gradientowe / kierunkowe
- Filtry kombinowane

Pytania :

- Przekształcenia kontekstowe – czym są ? Jakiego mają zastosowanie ?
- Filtry – jak można podzielić filtry ?
- Filtry konwolucyjne – Jak działają ? (można wzorek na splot)
- Co robić z brzegami obrazu przy filtracji ?
- Czym jest filtr dolnoprzepustowy inaczej uśredniający ? Jak działa ? Jakiego ma zastosowanie ? Jak wyglądają wykresy przed i po zastosowaniu go ? Jakiego ma wady i zalety ? Czym jest maska Gaussa ? Jak się ma filtr dolnoprzepustowy do częstotliwości ?
- Jaka jest różnica między filtrami liniowymi, a filtrami nieliniowymi ?
- Jak działa filtr medianowy ? Jakiego jest jego działanie w porównaniu do uśredniającego ? (Do czego się nadaje, a do czego gorzej)
- Jakiego znasz inne filtry nieliniowe ?
- Do czego mają służyć filtry górnoprzepustowe ?
- W jaki sposób w odniesieniu do pochodnych obrazu jest realizowana detekcja krawędzi ?
- Czym jest filtr Laplace’a ? Jak działa ? Co wykorzystuje i w związku z tym jakiego wykrywa krawędzie (jeśli chodzi o grubość) ? Jaką operację arytmetyczną należy wykonać aby uzyskać ładnie wyostrojony obraz ?
- Do czego służy Unsharp Masking ? Jak jest realizowane ?

- Do czego służą filtry gradientowe / kierunkowe ? Jakież masz filtry tego typu ? Którą pochodną wykorzystują ?
- Czym są filtry kombinowane ?

Przekształcenia kontekstowe

Przekształcenia kontekstowe są to przekształcenia przy których dla wyznaczenia wartości jednego punktu obrazu wynikowego trzeba dokonać określonych obliczeń na wielu punktach obrazu źródłowego.

Zastosowanie przekształceń kontekstowych :

- Tłumienie szumów
- Wzmocnienie elementów zgodnych ze wzorcem
- Usunięcie wad z obrazu
- Poprawa obrazu o złej jakości technicznej
- Rekonstrukcja obrazu, który uległ częściowemu zniszczeniu
- Wykrywanie krawędzi

Podział filtrów

PRZESZTAŁCENIA KONTEKSTOWE		Podział filtrów	
Filtry liniowe		Filtry nieliniowe	
→ Filtry konwolucyjne		→ Filtry logiczne	
→ Filtry dolnoprzepustowe		→ Filtry medianowe	
→ Filtry górnoprzepustowe		→ Filtry maksymalne	
		→ Filtry minimalne	
		→ Filtry adaptacyjne	

Różnica między liniowymi, a nieliniowymi jest wyjaśniona przy początku omawiania filtrów nieliniowych w dalszej części opracowania.

Filtry konwolucyjne

Filtry konwolucyjne to filtry wykorzystujące specjalną macierz zwaną maską na podstawie której obliczane są wartości pikseli. Zgodnie z kolejnością na laboratorium zostaną omówione najpierw.

Czym jest konwolucja ? Konwolucja to inaczej spłot funkcji. A jak to się ma do obrazu ? Po prostu mnoży się wagi z maski konwolucji przez wartości pikseli i sumuje się wyniki. Na koniec dzieli się przez liczbę wszystkich pikseli.

Konwolucja (spłot funkcji):

$$L'(m,n) = (w \times L)(m,n) = \sum_{i,j \in K} L(m-i, n-j) \cdot w(i,j)$$

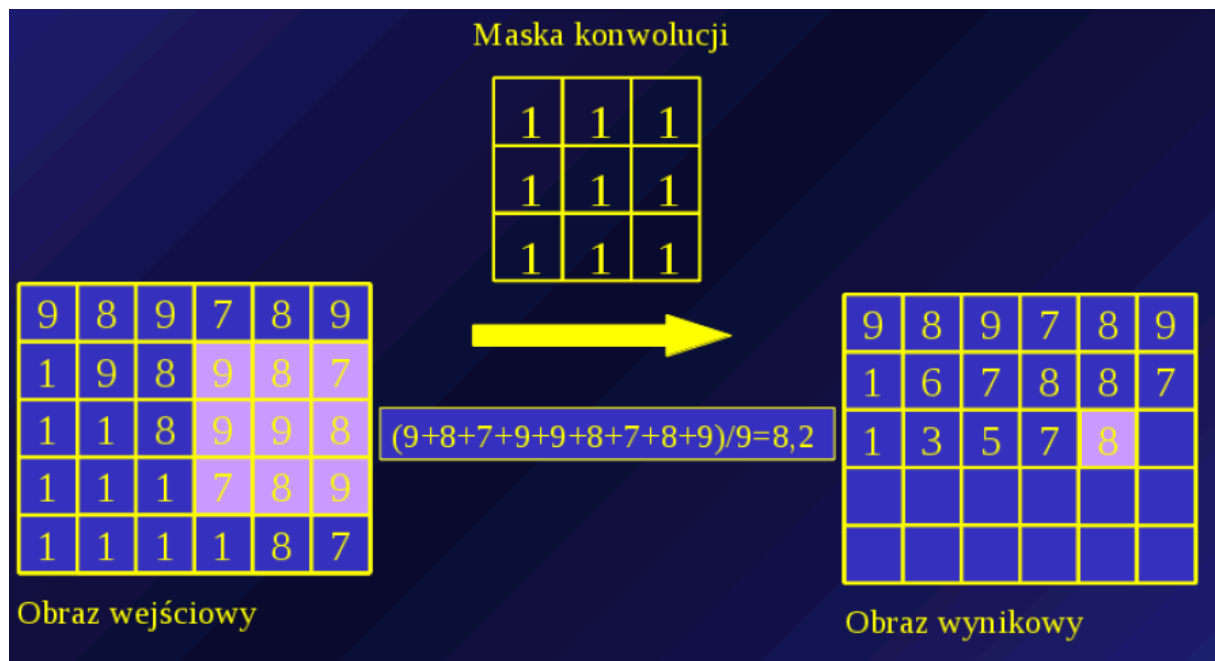
$L(m,n)$ - piksel oryginalny (obraz wejściowy)

$L'(m,n)$ - piksel przekształcony (obraz wyjściowy)

$w(i,j)$ - tablica współczynników (maska konwolucji)

K - zbiór indeksów określających obszar sąsiedztwa $L(m,n)$

Przykład poniżej :



Problemy mogą się pojawić przy przetwarzaniu skrajnych pikseli, czyli brzegów obrazu – brak wtedy wartości skrajnych i trzeba zastosować jedno z poniższych rozwiązań.

Co robić z brzegami obrazu?

- zmiana rozmiarów obrazu
- zmiana sposobu przetwarzania
 - pomijanie brakujących punktów
 - wypełnianie pustych miejsc
 - istniejącymi wartościami
 - nowymi wartościami

Filtry liniowe uśredniające = dolnoprzepustowe

Jest to najprostsza rodzina filtrów stosowana w przetwarzaniu obrazów. Wykorzystuje się je w celu redukcji szumów (zakłóceń) na obrazie, efektami (ubocznymi) stosowania tych filtrów jest rozmazanie obrazu i pogorszenie ostrości – zanik krawędzi. Filtr określony jest przez dwa parametry: rozmiar maski (ang. kernel) oraz wartości współczynników maski.

Filtry te są skuteczne dla obrazów zakłóconych szumem Gaussa, natomiast ich użycie dla obrazów o zakłóceniu 'Salt & Pepper' jest nieuzasadnione, w takim przypadku używa się filtru medianowego.

Tak wygląda obrazek z zakłóceniem szumem Gaussa z którym dobrze poradzi sobie filtr dolnoprzepustowy :

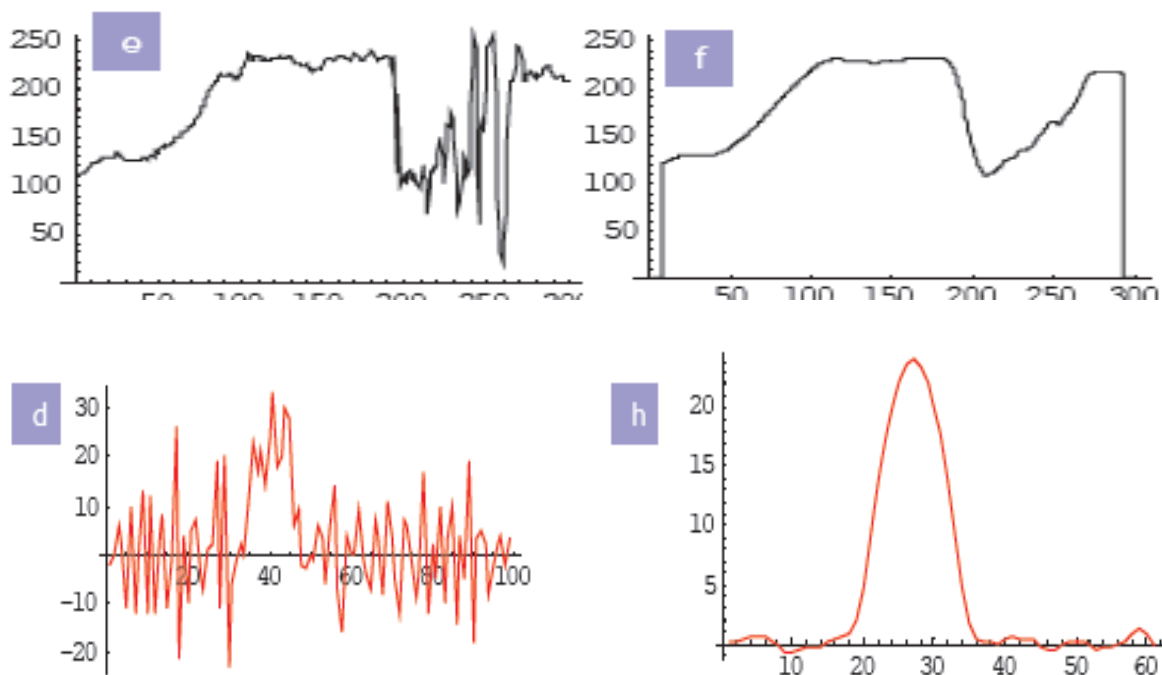


A tak wygląda zakłócenia 'Salt & Pepper' (białe i czarne kropki) z którym filtr dolnoprzepustowy nie poradzi sobie zadowalająco :



Działanie tego filtra opiera się o obliczanie średniej wartości pikseli z otoczenia danego piksela włącznie z tym pikselem. Poprzez ustawienie wartości większej dla danego piksela (np. 2) mocniej uwypukla się znaczenie środkowego piksela, ustawiając dwa dla sąsiadów bocznych mocniej uwypukla się ich wartość i tak można tworzyć różne warianty tego filtra.

Obliczając wartość bezwzględną różnicy obrazu oryginalnego (obrazek e poniżej) i obrazu po filtracji dolnoprzepustowej (obrazek f) otrzymamy obraz prezentujący moduł różnicy między zakłóceniem a średnią, można sobie wyobrazić co uzyskamy gdy zrobimy moduł różnicy, dostaniemy wtedy jakby same krawędzie (wykorzystanie w Unsharp-Masking) , tak jakbyśmy odjęli :e-f.



Przekrój przez wybrany wiersz (rys. d powyżej) i działanie filtra uśredniającego (rys. h)

Przykłady filtrów dolnoprzepustowych i ich wady i zalety :

Filtry dolnoprzepustowe = filtry uśredniające

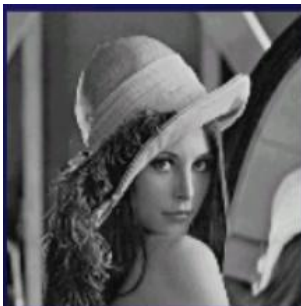
1	1	1
1	1	1
1	1	1

1	1	1
1	2	1
1	1	1

1	2	1
2	4	2
1	2	1

1	1	1
1	0	1
1	1	1

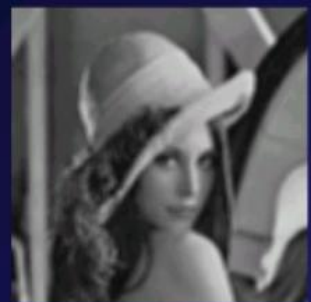
- usuwanie zakłóceń
- wygładzanie obrazu
- eliminacja „szumu”
- rozmycie krawędzi
- pogorszenie ostrości



Obraz wejściowy



1	1	1
1	1	1
1	1	1



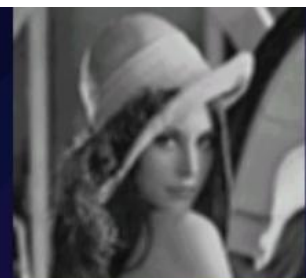
Obraz wynikowy



Obraz wejściowy



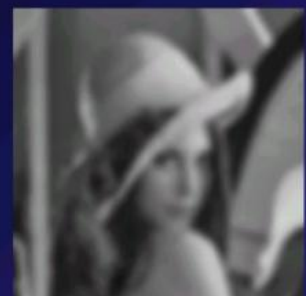
1	1	1
1	1	1
1	1	1



Obraz wynikowy



1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1
1	1	1	1	1

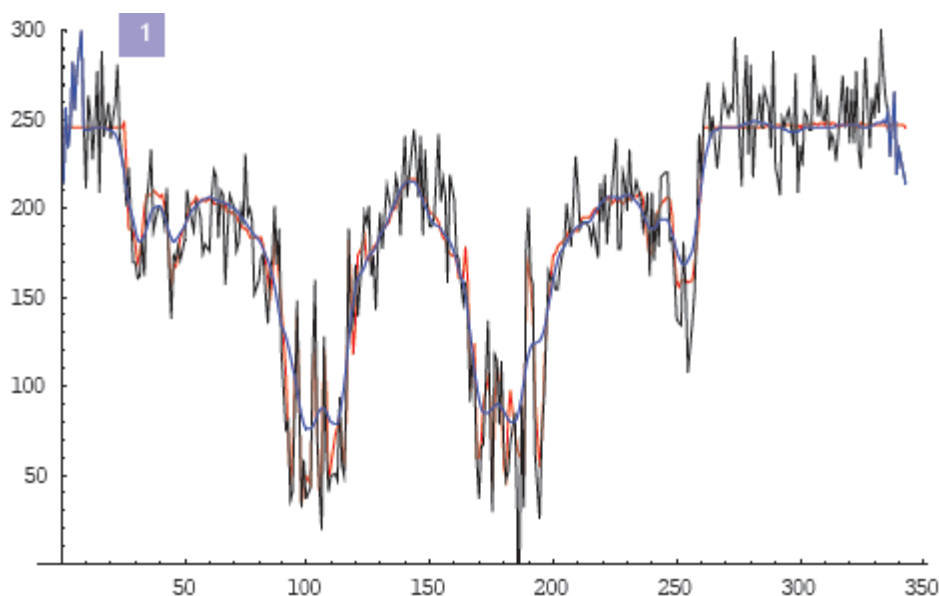


Obraz wynikowy

Jak widać zwiększenie maski powoduje zazwyczaj utratę ostrości. Aby zniwelować w pewnym stopniu efekt rozmycia można zwiększyć wartość piksela środkowego.

Ciekawym rozwiązaniem w tym filtrze jest tzw. Maska Gaussa. Wartość środkowego piksela jest największa na bokach mniejsze, na rogach najmniejsze. Można ustawić te wartości w kolejności takiej aby utworzyły krzywą Gaussa. Podając parametr którym jest odchylenie standardowe określamy wygląd krzywej Gaussa i tym samym maskę.

Przykład działania filtru Gaussa przedstawia poniższy obrazek :



Rys. 1. Rysunek przedstawia przekrój przez wiersz 250 odpowiednio: obrazu oryginalnego (krzywa czerwona), zakłóconego szumem o rozkładzie normalnym ze średnią 0 i wariancją 20.0 (krzywa czarna) oraz obrazu odszumionego filtrem Gaussa z maską 5x5 (krzywa niebieska).

Analiza względem częstotliwości :

Działanie tego typu filtrów opiera się na usuwaniu elementów obrazu o wysokiej częstotliwości (szczegółów, np. dużych różnic w kolorach pomiędzy sąsiadującymi pikselami) i przepuszczaniu elementów o niskiej częstotliwości (ogólnych kształtów, bez szczegółów). Ponieważ większość szumów występujących w obrazach zawiera się w wysokich częstotliwościach, filtry te przeważnie wykorzystuje się właśnie do eliminacji zakłóceń.

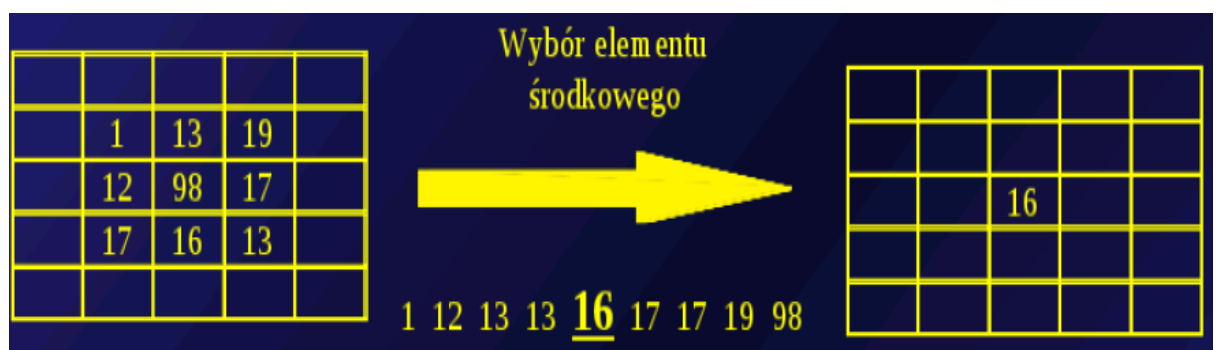
Filtry liniowe górnoprzepustowe zostały w tym opracowaniu zgodnie z kolejnością na laboratorium omówione po filtrach nieliniowych.

Filtr nieliniowy medianowy

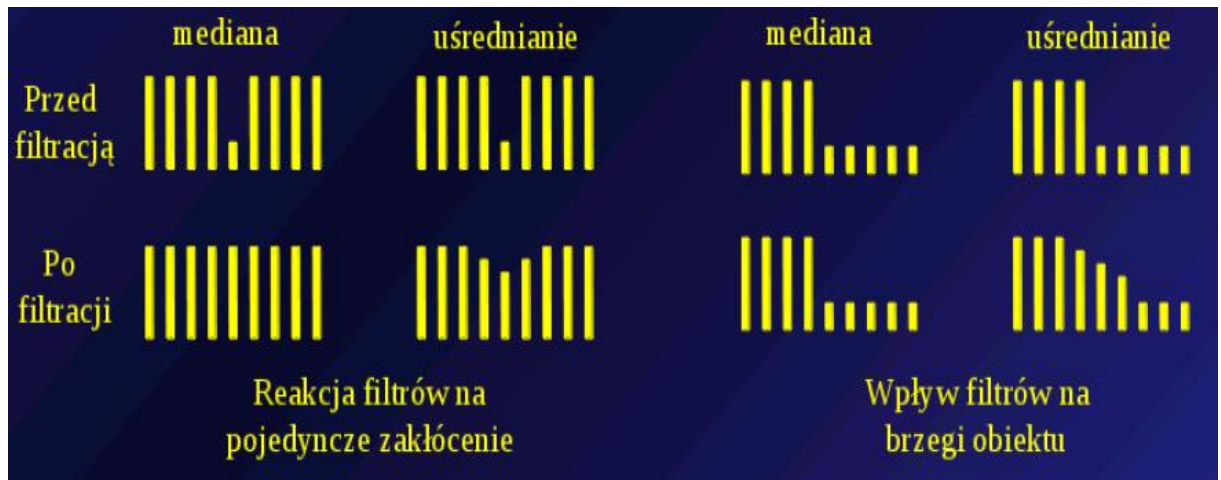
Filtry rozmywające redukują szum ale niekorzystnie wpływają na ostrość obrazu - krawędzie. Dlatego często wykorzystuje się filtry nieliniowe - np. filtr medianowy (mediana - środkowa wartość w posortowanym ciągu liczb).

Podstawowa różnica pomiędzy filtrami liniowymi, a nieliniowymi polega na tym, że przy filtracji liniowej na nową wartość piksela ma wpływ wartość wszystkich pikseli z otoczenia (uśrednianie), natomiast w przypadku filtracji nieliniowej jako nowy piksel wybierana jest któraś z wartości otoczenia - według jakiegoś wskaźnika. Ważną cechą filtrów medianowych jest także to, że zachowują krawędzie.

Przykład poniżej :



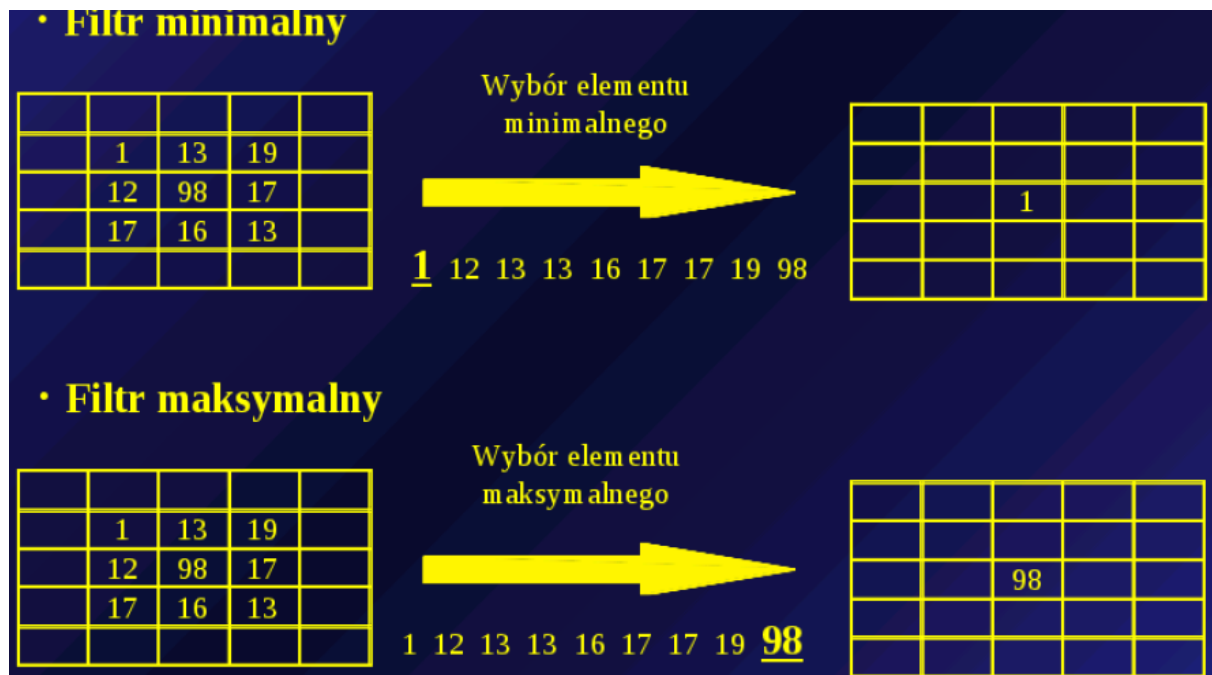
Filtry medianowe zachowują krawędzie. Dobrze radzą sobie z zakłóceniami punktowymi (strona lewa poniżej) , ale gorzej z zakłóceniami ciągłymi – brzegami obiektów (praktycznie brak działania po prawej) :



Jakość filtracji medianowej silnie zależy od wielkości używanego okna (mniejsze okna dają lepszy efekt).

Inne filtry nieliniowe

Filtr minimalny / maksymalny, przykład poniżej :



- filtr modowy - moda (dominanta) zamiast mediany
- filtr olimpijski - średnia z podzbioru otoczenia (bez wartości ekstremalnych)

Filtry liniowe górnoprzepustowe

Zadaniem filtrów górnoprzepustowych jest wydobywanie z obrazu składników odpowiedzialnych za szybkie zmiany jasności - konturów, krawędzi, drobnych elementów faktury. Filtry górnoprzepustowe wyostrzają lub różniczkują sygnał, czyli uwypuklają krawędzie na obrazie.

Filtry te działają w sposób odwrotny do filtrów dolnoprzepustowych, tłumią one niskoczęstotliwościowe elementy obrazu, wzmacniają natomiast elementy o wysokich częstotliwościach (szczegóły). Wynikiem działania tego typu filtrów jest podkreślenie, uwypuklenie elementów obrazu o dużej częstotliwości poprzez zwiększenie ich jasności, koloru, itp.

Filtry górnoprzepustowe można podzielić na:

- **filtry wykrywające krawędzie - laplasjany**
- **filtry wykrywające narożniki**
- **filtry kierunkowe**

Uzupełnienie o pochodnych obrazu przydatne do zrozumienia filtrów górnoprzepustowych

✚ Większość procedur wyostrażania obrazu oraz detekcji krawędzi wykorzystuje operatory pierwszej oraz drugiej pochodnej obrazu, przy czym w takim przypadku obraz jest dyskretną funkcją dwuwymiarową.

Własności pierwszej i drugiej pochodnej:

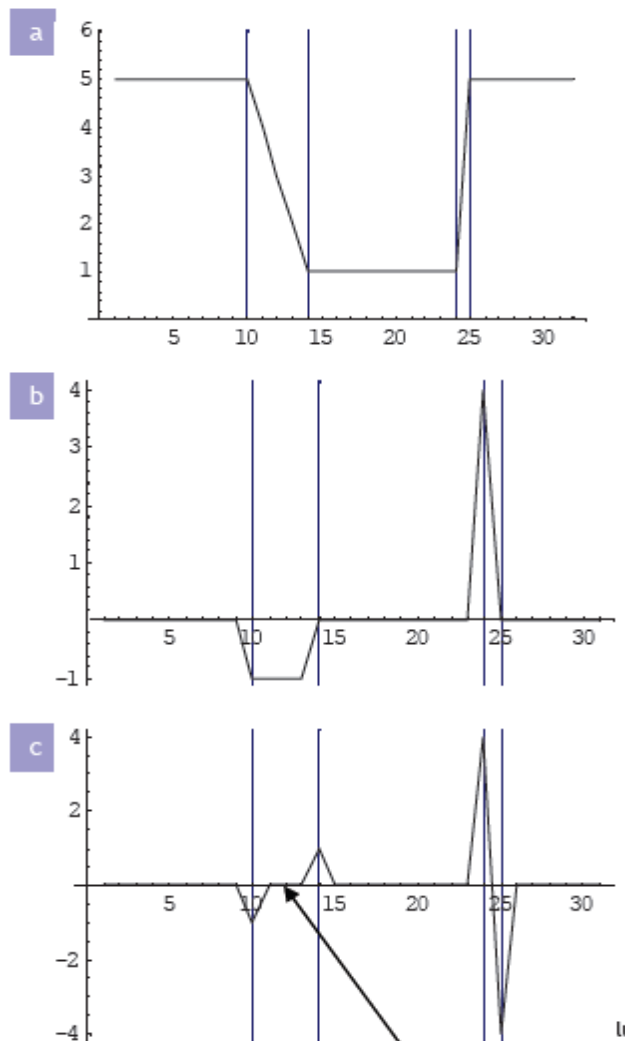
✚ obie generują zero w obszarach o ustalonej wartości,

Poniżej znajdują się wykresy dla przykładowego obrazka :

a – obraz oryginalny

b – pierwsza pochodna

c – druga pochodna



✦ punkty krawędzi w obrazie mogą być rozpoznane na podstawie śledzenia wartości pierwszej pochodnej: piksele dla których wartość pierwszej pochodnej przekracza ustalony próg są interpretowane jako piksele krawędzi,

✦ druga pochodna generuje dwie wartości: ujemną oraz dodatnią dla krawędzi o dowolnej grubości. Jest to cecha niepożądana, odpowiedzialna za zjawisko podwójnego konturu w obrazach po filtracji operatorem drugiej pochodnej,

Wynika to z tego, że gdy pierwsza pochodna wykryje krawędź to najpierw maleje, później ma wartość stałą, a później znów rośnie (rysunek b powyżej) i dlatego druga pochodna ma wartości dodatnie i ujemne (rysunek c)

✦ korzystając z drugiej pochodnej krawędź może być rozpoznana na podstawie śledzenia przejść przez zero (patrz rysunek (c))

✦ dla punktów krawędzi pierwsza pochodna przyjmuje wartości niezerowe (dodatnie lub ujemne w zależności od rodzaju przejścia) w całym obszarze krawędzi z pominięciem ostatniego punktu. Efektem są stosunkowo rozległe obszary o niezerowych wartościach (grube krawędzie),

Problemem przy stosowaniu pochodnych może być wrażliwość na zakłócenia.

Natomiast dla drugich pochodnych mamy wąskie krawędzie (patrz rysunek b i c)

Filtry Laplace'a (Laplasjany)

Stosowane są do wykrywania krawędzi, cechuje je wielokierunkowość, dzięki czemu mogą wykrywać krawędzie o różnych kierunkach. Filtry te wykorzystują drugą pochodną obrazu z czego wynika wąskość krawędzi.

Są ważnym narzędziem wyostrażania krawędzi poprzez dodanie do obrazu oryginalnego obrazu przetworzonego. Efekt wyostrażania można zwiększyć stosując czynnik skalujący :

$$f'(x, y) = f(x, y) \pm k \nabla^2 f(x, y)$$

Istotnym etapem przed zastosowaniem filtru Laplacea jest wygładzenie zakłóceń przy pomocy filtru dolnoprzepustowego.

✦ Laplasjan (druga pochodna) dla dyskretnej funkcji dwuwymiarowej ma postać:

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

i jest najczęściej realizowany jako przekształcenie z maskami 3x3 postaci:

$$\nabla^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \nabla^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Przykładowe maski laplasjanów :

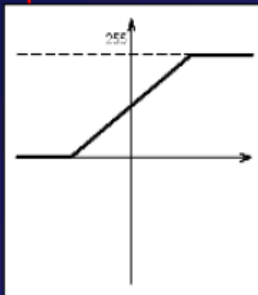
0	-1	0	-1	-1	-1	-1	-2	-1
-1	4	-1	-1	8	-1	-2	12	-2
0	-1	0	-1	-1	-1	-1	-2	-1

0	-1	0	-1	-1	-1	-1	-2	-1
-1	5	-1	-1	9	-1	-2	13	-2
0	-1	0	-1	-1	-1	-1	-2	-1

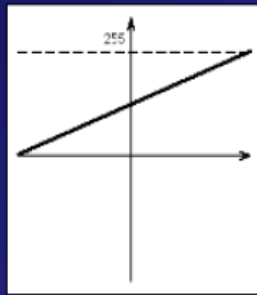
Funkcja normalizująca (dwa obrazki niżej) jest często konieczna w zastosowaniu tych masek, ponieważ mają wartości ujemne. Najczęściej stosowane operacje to skalowanie (np. dodanie +128 do każdego piksela) oraz obliczenie modułu.

Ponieważ w maskach filtrów górnoprzepustowych występują ujemne współczynniki wag, pojawia się problem z normalizacją obrazów wynikowych, gdyż wartości poszczególnych pikseli mogą wykroczyć poza zakres 0 - 255. Poniżej przedstawiono kilka sposobów normalizacji:

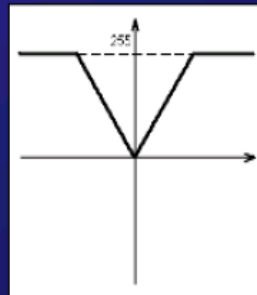
u
r



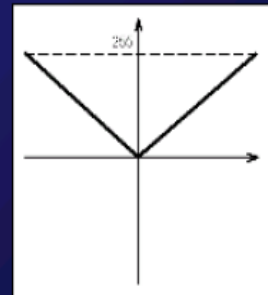
Skalowanie
z obcięciem



Skalowanie
bez obcięcia

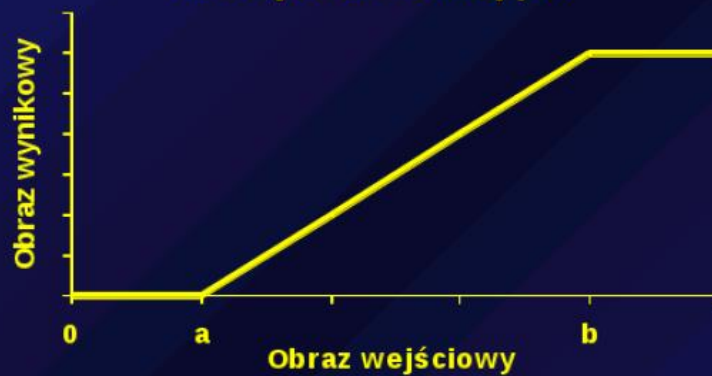


Moduł
z obcięciem

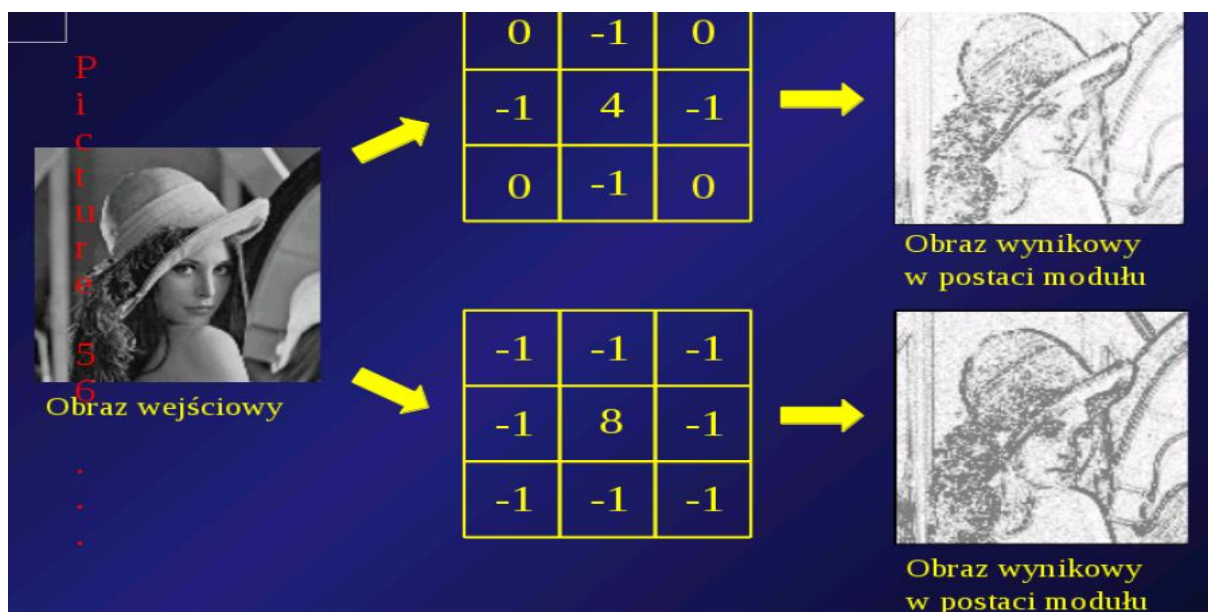
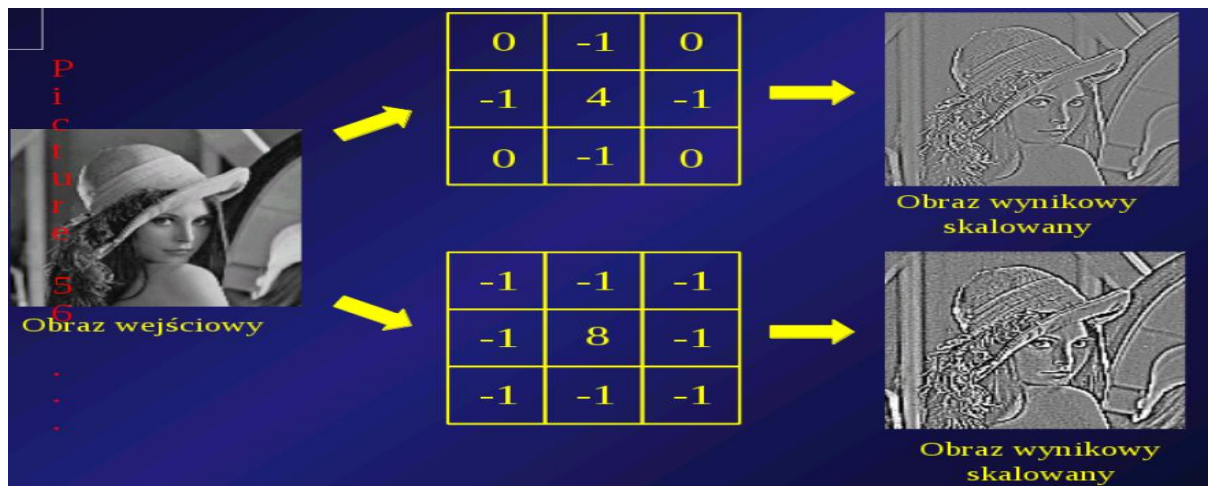


Moduł
bez obcięcia

Funkcja normalizująca



Normalizacja - jest to operacja polegająca na sprowadzeniu przedziału zmian wartości punktów wyjściowego obrazu do pewnego, ustalonego zakresu. Operacja ta zazwyczaj poprzedza lub kończy inne przekształcenia obrazu (w szczególności arytmetyczne).



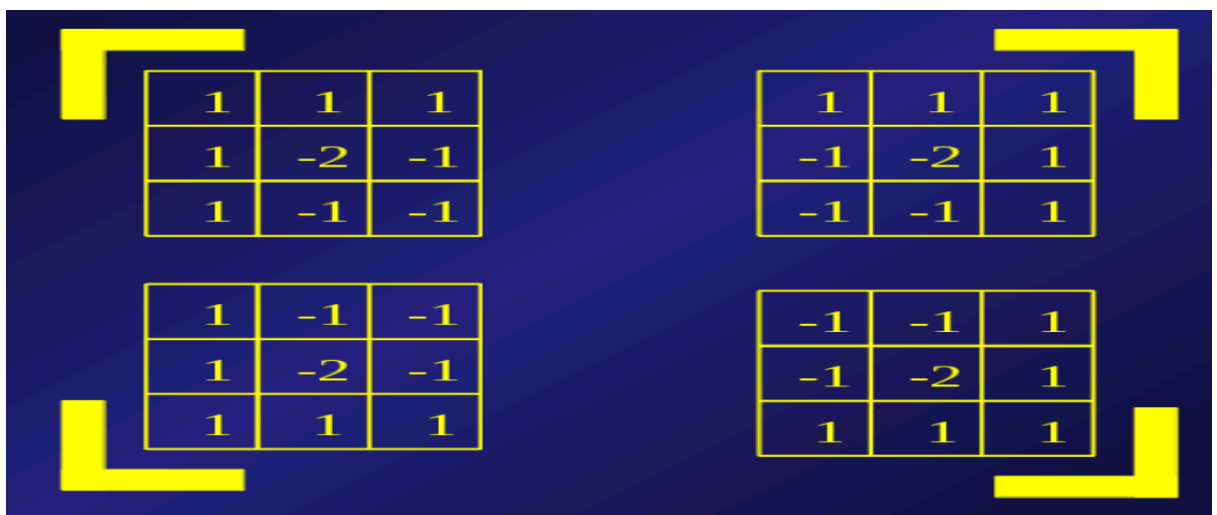
Jak widać efekt końcowy zależy nie tylko od maski, ale i od przyjętej normalizacji.

Unsharp Masking

Jedną z technik wyostrażania obrazu jest tzw. 'unsharp masking'. Polega tak naprawdę na maskowaniu nieostrości. Najpierw rozmywa się obrazek oryginalny (f. dolnoprzepustowy) i odejmuje się rozmycie od oryginału, uzyskujemy w ten sposób maskę dla obrazu oryginalnego. Następnie maskę tą dodajemy do obrazu oryginalnego.

Filtry wykrywające narożniki

Przykładowe maski:



Filtry Kierunkowe

Są wykorzystywane do detekcji krawędzi.

☛ Metody detekcji krawędzi do pewnego stopnia korzystają z narzędzi znanych z wyostrażania obrazu. Cele tych dwóch procesów są jednak inne. Celem wyostrażenia obrazu jest poprawa jakości wizualnej: zwiększenie ostrości, wyrazistości, kontrastu, a z teoretycznego punktu widzenia również zwiększenie entropii czyli ilości informacji przenoszonej przez obraz. Odbiorcą obrazu wyostrażonego jest najczęściej człowiek. Celem detekcji krawędzi jest natomiast ekstrakcja krawędzi oraz usunięcie pozostałych fragmentów obrazu. Obraz po detekcji jest często etapem procesu automatycznej analizy obrazu.

Metody gradientowe

Do ich zalet należy to, że są stosunkowo proste koncepcyjnie, dobrze wykrywają struktury geometryczne. Wadami może być konieczność filtracji w dwóch ortogonalnych kierunkach, generowanie grubych krawędzi (pierwsza pochodna), duża wrażliwość na zakłócenia.

❖ Metody gradientowe stanowią grupę najprostszych metod wykrywania krawędzi w obrazie. Podstawą metod gradientowych jest wyznaczanie pierwszej pochodnej w dwóch ortogonalnych kierunkach (niekoniecznie w pionie i poziomie).

Głównym zadaniem gradientów nie jest jak w przypadku Laplasjanów wyostanie obrazu, ale detekcja krawędzi. Dokonują tego poprzez obliczanie pierwszej pochodnej w przyjętym kierunku.

Gradient Robertsa

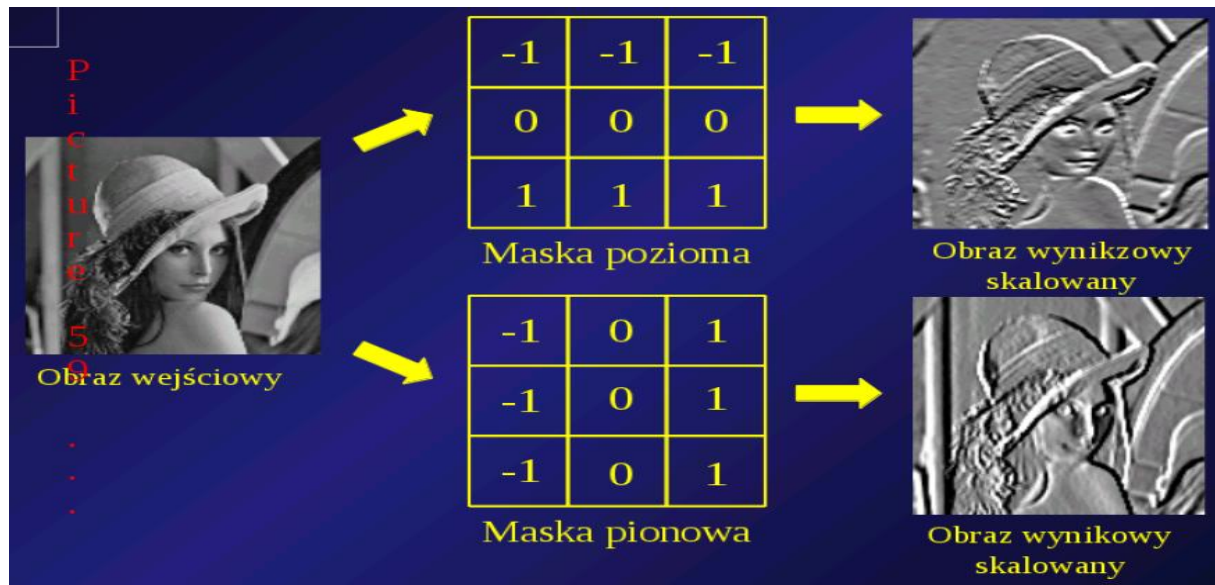
Do wykrywania krawędzi pod kątem 45 stopni (liczy pochodną przekroju obrazu w kierunku 45 stopni)



Umieszczając w inny sposób jedynkę i -1 można osiągnąć inny kierunek detekcji krawędzi.

Gradient Prewitta

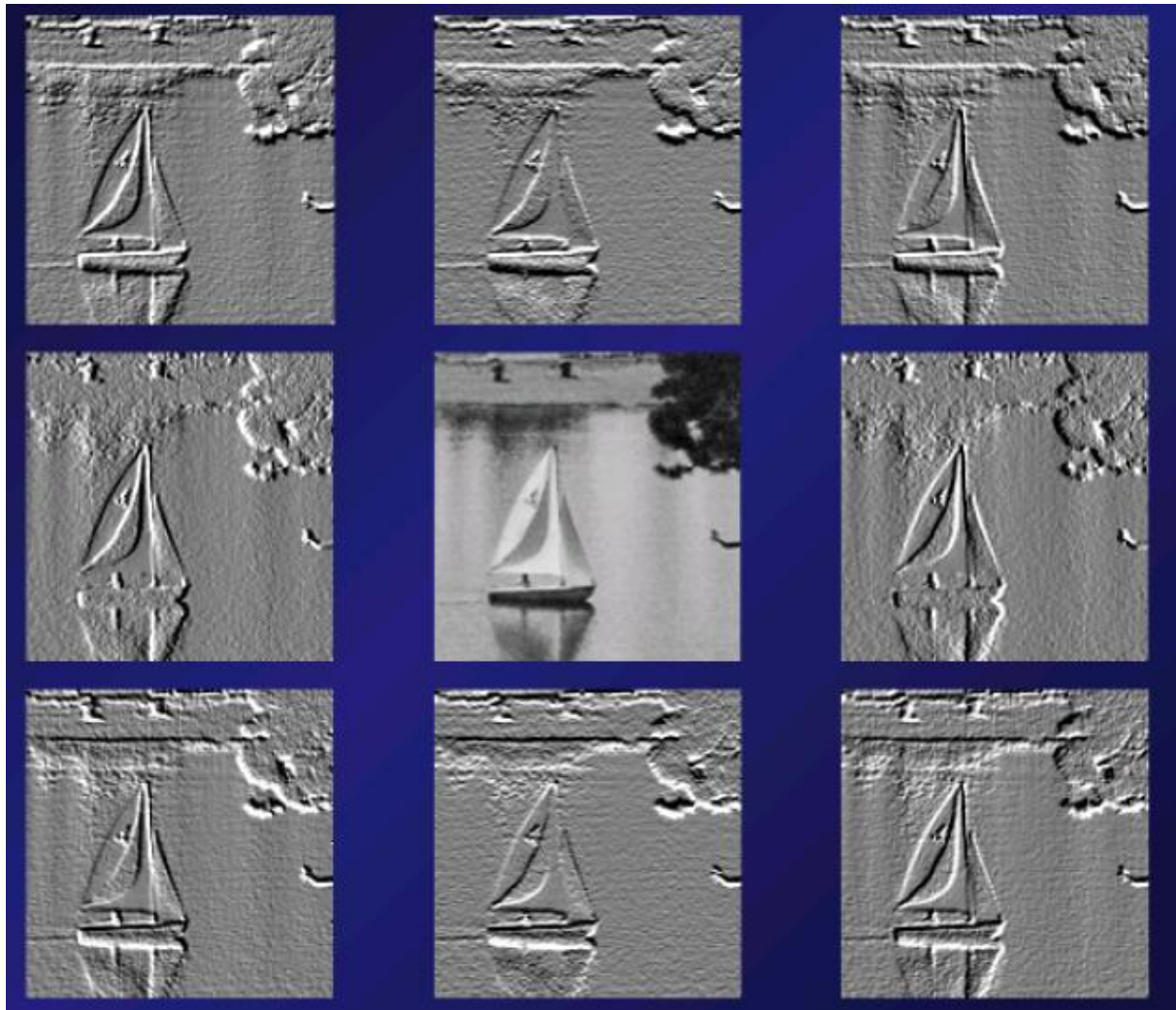
Gradient Prewitta pozwala na wykrywanie krawędzi w pionie lub poziomie w zależności od rodzaju zastosowanej maski.



Gradient Sobela

Jest skonstruowany podobnie do Prewitta, ale np. wartość dla środkowego piksela w boku jest równa -2 zamiast 1 i 2 zamiast 1 na drugim boku.

Istnieje 8 różnych masek dla tego gradientu, każda wykrywa krawędzie w innym kierunku.



Filtry kombinowane

Do tworzenia filtrów kombinowanych wykorzystuje się dwie ortogonalne maski Sobela (wykrywające krawędzie w dwóch ortogonalnych kierunkach).

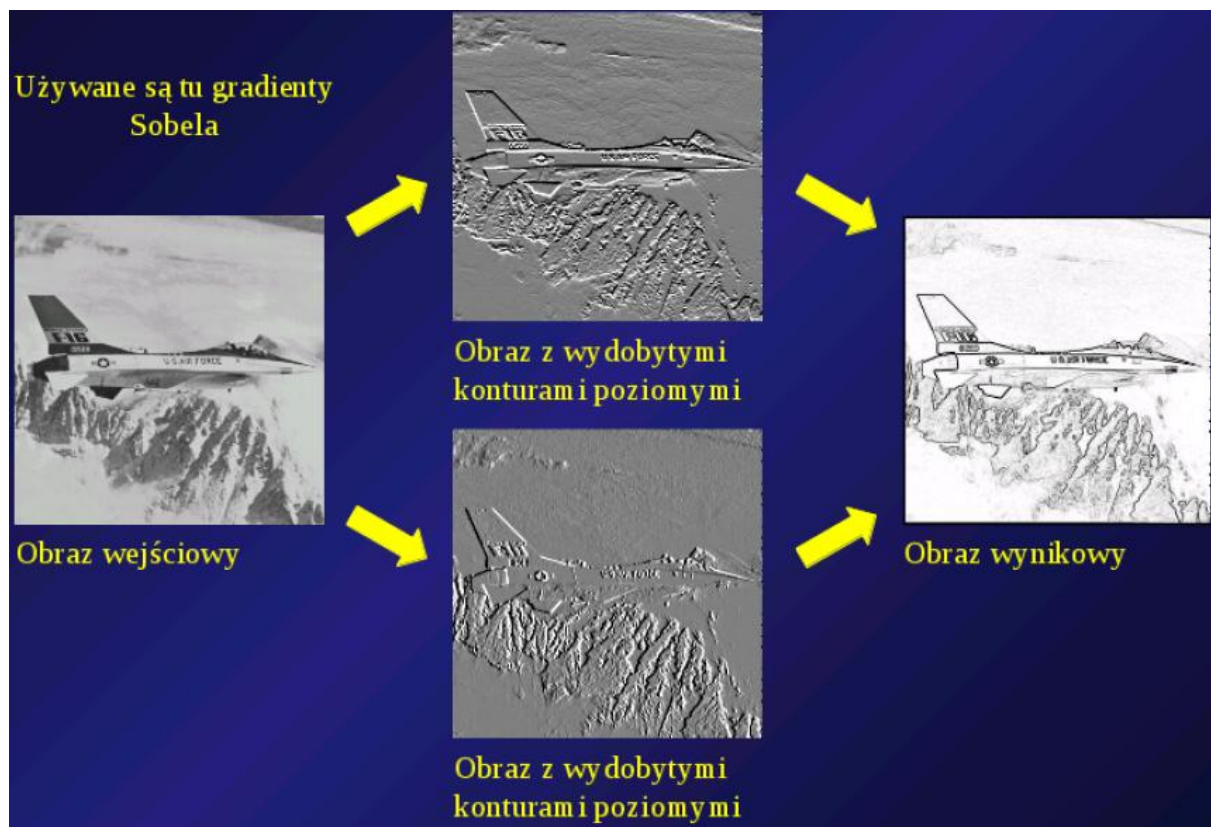
Idea filtrów kombinowanych polega na kolejnym zastosowaniu dwóch gradientów w prostopadłych do siebie kierunkach, a następnie na dokonaniu nieliniowej kombinacji wyników tych gradientów. Dzięki nieliniowej kombinacji rezultatów liniowych transformacji obrazu tworzy się w ten sposób obraz wynikowy o wyjątkowo dobrze podkreślonych konturach niezależnie od kierunku ich przebiegu.

Do połączenia (kombinowania) obrazów można użyć formuły euklidesowej:

$$L'(m,n) = \sqrt{\left(L_1(m,n)\right)^2 + \left(L_2(m,n)\right)^2}$$

$L'(m,n)$ - punkt obrazu wynikowego

$L_1(m,n)$, $L_2(m,n)$ - punkty na obrazach powstałych po zastosowaniu gradientów



Na slajdzie wyżej błąd w podpisie powinno być w drugim przypadku 'pionowymi'.