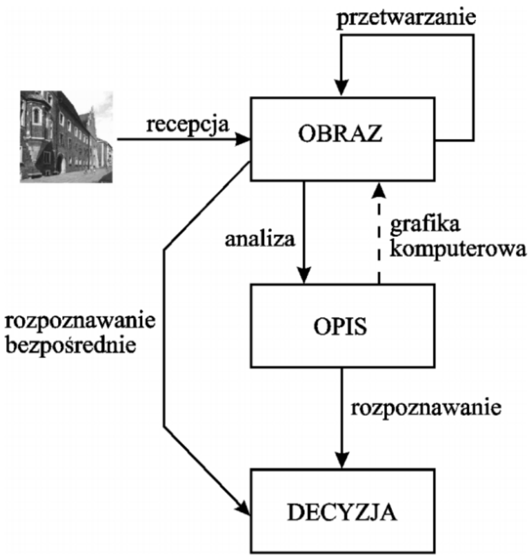
OPRACOWANIE Z PRZEDMIOTU KPO

DYSKRETNA STRUKTURA OBRAZÓW CYFROWYCH

WSTĘP Z WYKŁADÓW

Celem sztucznego przetwarzania lub analizy obrazów jest takie automatyczne ich przetworzenie i przeanalizowanie aby uzyskać z nich użyteczną informację na temat interesujących nas obiektów/ otoczenia systemu. Te informacje mogą wpływać na proces sterowania. Proces widzenia składa się z operacji:

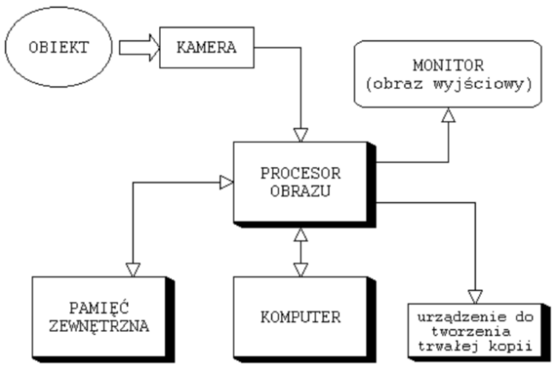
- Recepcja (akwizycja) obrazu,
- Przetwarzanie obrazu (filtracja, kompresja, eksponowanie cech),
- Analiza obrazu (wydobycie informacji i cech)
- Rozpoznanie obrazu i jego interpretacja -> decyzja.



Przetwarzanie obrazu jest etapem najbardziej obciążającym obliczeniowo, zwłaszcza etap filtracji. Trudność natomiast stanowi operacja analizy obrazu.

W systemie wizyjnym wyróżniamy elementy:

- Moduł wprowadzania obrazu,
- Urządzenia do wyświetlania obrazów,
- Urządzenia do tworzenia trwałej kopii,
- Pamięć zewnętrzna,
- Procesor obrazu



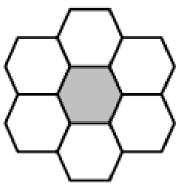
Moduł wprowadzania zamienia obraz do postaci akceptowalnej przez procesor, w etapie przetwarzania na wyjściu i wejściu występują obrazu, wynikiem analizy są dane jakościowe i ilościowe opisujące określone cechy obrazu lub grupy obrazów.

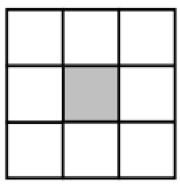
Podczas analizy obrazów gubiona jest część informacji, jednak wynik może informacje które są użyteczne a trudne do wyodrębnienia bez analizy obrazów. Po procesie analizy obraz jest podatny na zastosowanie metod i algorytmów rozpoznawania i oceny merytorycznej zawartość.

Zastosowania: automatyka, medycyna, stalkerzy, kryminalistyka, geodezja i kartografia, komunikacja, laboratoria, wojskowość, zabezpieczenie obiektów, astronomia i astrofizyka, metrologii, mistrzowie PhotoShopa, inwigilacja.

CZĘŚĆ ZASADNICZA PYTANIA

Dyskretna reprezentacja obrazów opiera się na siatkach podziału elementów. Wyróżnia się siatkę (strukturę) heksagonalną i kwadratową (stanrdard).

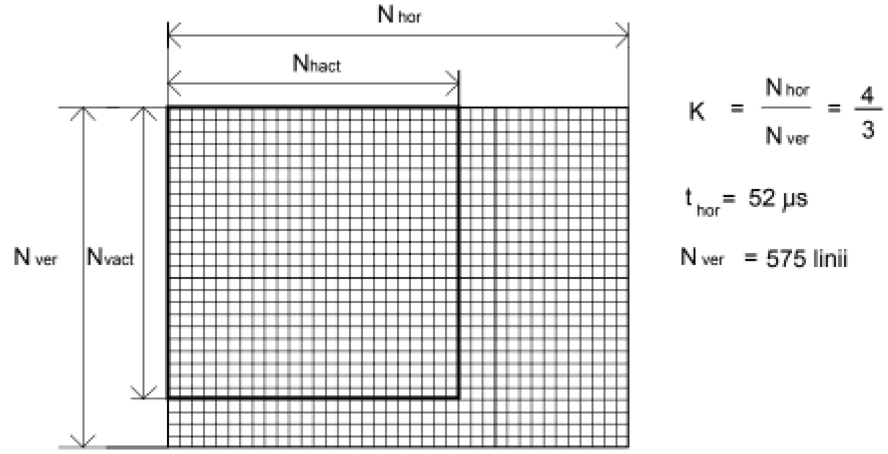




Ograniczenia reprezentacji obrazu mogą być realizowane jednocześnie na wielu płaszczyznach:

- ograniczenie zdolności rozpoznawania szczegółów;
- ograniczenie ilości możliwych do rozróżniania stanów elementu obrazu (kolorów);
- analizowanie obrazu płaskiego zamiast przestrzennego;
- analizowanie obrazu statycznego zamiast dynamicznego.

Rozmiar siatki (rastra) obrazu cyfrowego określa częstotliwość próbkowania sygnału wizyjnego. Dla siatki kwadratowej częstotliwość przetwornika A/C wynosi:



$$f = \frac{K * N_{ver}}{t_{hor}} = 14{,}75\ \text{MHz}$$

$K = N_{hor} / N_{ver} = 4/3$ - proporcje obrazu,
$N_{hor} = 575$ - liczba widocznych na ekranie linii poziomych,
$t_{hor} = 52\ \mu s$ - aktywny czas wybierania linii (PAL).

Wybrawszy określony typ rastra musimy zdecydować, jak wiele elementów ma ten raster zawierać, czyli ustalić, jak duża ma być rozdzielczość obrazu. Wybór właściwej rozdzielczości obrazu jest sprawą bardzo ważną, gdyż rozdzielczość jest miarą zdolności rozpoznawania szczegółów obrazu. Dobór właściwej rozdzielczości jest zawsze kompromisem. Z jednej strony - im większej rozdzielczości jest obraz, tym więcej szczegółów zawiera i więcej informacji można z niego odczytać, a z drugiej strony - liniowy wzrost rozdzielczości obrazu pociąga za sobą kwadratowy wzrost czasu przetwarzania i objętości samej reprezentacji. Sposób postępowania jest zazwyczaj następujący: pozyskuje się obraz o stosunkowo dużej rozdzielczości, by w trakcie kolejnych, wstępnych etapów obróbki obrazu dojść do rozdzielczości mniejszej, praktycznej. Najczęściej spotykane rozdzielczości w analizie obrazu to całkowite potęgi dwójki od 256x256 do 4096x4096.

Rozdzielczość obrazu - Wyraża się ona ilością elementów podstawowych składających się na obraz. Najczęściej przy płaskich obrazach o kwadratowej siatce zapisywana jest ona jako iloczyn ilości elementów w poziomie i pionie obrazu.

Każda komórka naszej siatki (piksel) może przyjmować jeden z pośród ograniczonej liczby stanów. Liczba ta wynika z ilości bitów użytych do zapisania informacji o brawie, tak więc mowa tu o rozdzielczości barwnej obrazu. Przykładowe formaty braw:

- binarny (1bit) – służy głownie do analizy obrazu,
- monochromatyczny (8 bitów) – skala szarości, reprezentuje względną jasność piksela,

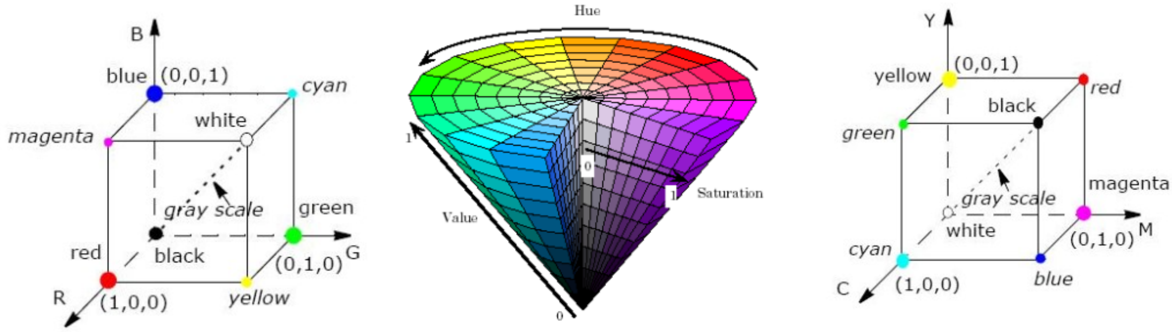
- kolorowy (24 bity lub 32 bitów) – Dla RGB po 8 kolejnych bitów na kolor, format 32 bitowy ma za zadanie poprawić wydajność dla systemów x86, x64 (szyna danych o szerokości Nx32 bitów).

Barwa definiowana jest przez 3 atrybuty:

- odcień barwy (kolor, ton, Hue) – różnica jakościowej barwy, dominująca długość fali,
- Nasycenie (Saturation) – odstępstwo barwy od bieli (tła),
- Jasność (wartość, Value) – wskazuje czy barwa jest bliższa bieli czy czerni.

Modele kolorów:

- RGB – stosowany w grafice komputerowej, bryłą jest sześcian,
- HSV – model intuicyjny dla człowieka, bryłą jest stożek (ostrosłup),
- CMYK – w poligrafii, bryłą jest sześcian, kolor K – czarny ze względu na jego zużycie.



Aby obraz przedstawić w postaci cyfrowej należy go poddać dyskretyzacji przestrzennej (próbkowanie) i dyskretyzacji wartości (kwantowanie). W ten sposób funkcja ciągła L(x, y) przechodzi w funkcję cyfrową L(m, n) o M wierszach, N kolumnach i określonej liczbie poziomów jasności. Rozdzielczość przestrzenna to liczba wierszy i kolumn (wpływa na liczbę szczegółów widocznych na obrazie) a rozdzielczość poziomów szarości to zbiór liczba całkowych opisujących jasność elementów obrazu. Rozdzielczość poziomów jasności można swobodnie zmniejszyć wg wzoru:

$$L'(m,n) = \left[\mathrm{round}\left(\frac{L(m,n) - \left(\frac{delta}{2} - 1\right)}{delta} \right) \right] \cdot delta + \left(\frac{delta}{2} - 1 \right), \quad \text{gdzie} \quad delta = \frac{2^B}{2^{B'}}$$

PRZEKSZTAŁCENIA GEOMETRYCZNE OBRAZU

Przekształcenia geometryczne to przesunięcia, obroty, odbicia czy transformacje (chwała temu kto rozumie imtransform). Używa się ich do korekcji błędów wnoszonych przez system wprowadzania i akwizycji obrazu oraz jako operacje pomocnicze dla bardziej skomplikowanych operacji przetwarzania.

Podstawowe operacje:

Skalowanie – zmiana rozmiaru zdjęcia, może służyć do powiększania/pomniejszania zdjęcia a także jego rozciągania, można podać wymagany wyjściowy rozmiar MxN. Z pozoru prosta operacja jednak problem stanowi metoda interpolacji.	Imresize(image, scale)
Obracanie – operacja obracania obrazu, obraz wynikowy jest obrócony o podany kąt, a kolory pikseli określane są na podstawie wybranej metody	Imrotate(image, angle, method)

interpolacji, dodatkowo obrócone zdjęcie może zwiększyć swój rozmiar (przekątna staje się bokiem),	
Kadrowanie – opracja przycięcia obrazu do zadanaego obszaru,	Imcrop(image, [x, y, dx, dy])
Transformacja przestrzenna – operacja wykonania salta przez obraz, nieintuicyjna metoda, ale używana w obrazowaniu z satelit, zakładając ze mamy google maps ze zdjęciami sprzed 10 lat i chcemy dany obszar zaktualizować, jednak nowe zdjęcia były wykonane pod innym kątem, dlatego łącząc kolejne obrazy musimy obrócić zdjęcia tak aby do siebie pasowały. Pewnie teraz to robi AI ale pierwotnie tak to robili.	Imtransform(image, transformation) Transformation = maketform(metoda, jakies cyferki)

Przekształcenia punktowe obrazu

Przekształcenia punktowe charakteryzują się tym, że poszczególne elementy obrazu (punkty) są modyfikowane niezależnie od otoczenia. Możemy do nich zaliczyć:

- przekształcenia oparte na przetwarzaniu pojedynczych punktów;
- przekształcenia oparte na arytmetycznym przeliczaniu pojedynczych punktów;
- realizacja przekształceń punktowych z użyciem LUT (Look Up Tables);
- wyrównywanie histogramu;
- punktowe operacje wykonywane na dwu obrazach;
- binaryzacja;

Przekształcenia punktowe realizowane są zwykle w taki sposób, że wymagane operacje wykonuje się na poszczególnych pojedynczych punktach źródłowego obrazu, otrzymując w efekcie pojedyncze punkty obrazu wynikowego. Operacje te charakteryzują się następującymi cechami:

- modyfikowana jest jedynie wartość (np. stopień jasności) poszczególnych punktów obrazu. Relacje geometryczne pozostają bez zmian.
- jeżeli wykorzystywana jest funkcja ściśle monotoniczna (rosnąca lub malejąca), to zawsze istnieje operacja odwrotna, sprowadzająca z powrotem obraz wynikowy na wejściowy.
- nie wprowadzają one żadnych nowych informacji do obrazu, jedynie uwypuklają one pewne zawarte w obrazie już informacje.

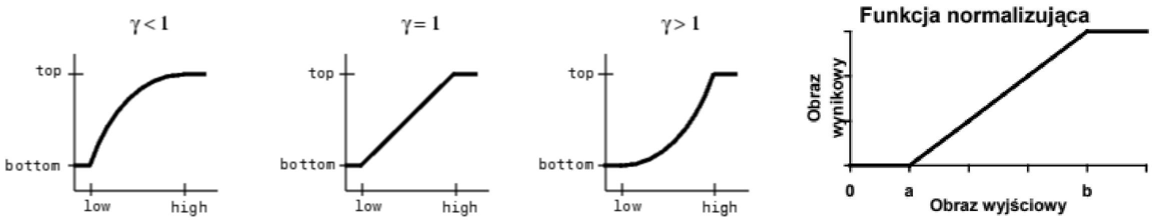
Bezpośrednio widocznym efektem przekształceń punktowych jest więc zawsze zmiana skali jasności obrazu bez zmiany geometrii widocznych na obrazie obiektów. Mimo bardzo prostego matematycznie charakteru przekształcenia punktowe bardzo radykalnie modyfikują subiektywne wrażenie, jakie uzyskujemy oglądając obraz. Czasem prowadzi to do krańcowego zniekształcenia obrazu, czasem jednak pozwala wykryć lub uwypuklić pewne cechy obrazu praktycznie niewidoczne, gdy się ogląda obraz oryginalny.

Dodawanie stałej do obrazu – powoduje rozjaśnienie obrazu,	Imadd(image, stała)
Odejmowanie stałej – powoduje przyciemnienie obrazu,	Imadd(image, -stała)
Mnożenie obrazu przez stałą – powoduje zwiększenie kontrastu? (nie jestem pewien)	Immultiply(image, stała)
Operacje dwuagumentowe, punktowe czyli odejmowanie, dodawanie obrazów do siebie, powodują nałożenie się obrazów, jednak w pewnych punktach wartość poziomu jasności może przekroczyć	Imabsdiff(image1, image2) Imadd(image1, image2)

górną granice i jest wtedy zaokrąglana do górnego progu. Obrazy powinny być tej samej rozdzielczości (możliwe ze matlab sam rozszerza obrazy do odpowiedniej wielkości)

Kolejną operacją punktową jest modulacja gamma (korekcja), usuwa ona zniekształcenia wprowadzone przez urządzenia jak np. skarner poprzez redukcję nadmiernego kontrastu obrazu wejściowego. Podobną operacją jest normalizacja, sprowadza ona jeden zakres zmian do wyjściowego zakresu zmian w sposób linowy.

W matlabie funkcja pod poleceniem imadjust(image, [Lmin Lmax], [L'min L'max], gamma)



Zależność matematyczna korekcji gamma: $L'(m,n) = L(m,n)^{\gamma}$ gdzie gamma jest liczbą rzeczywistą.

Wyrównywanie histogramu ma za zadanie takie przekształcenie jasności punktów aby liczba punktów leżących w danych równych przedziałach jasności była w przybliżeniu taka sama. Histogram jest funkcją pozwalająca scharakteryzować obraz w sposób globalny i dany jest wzorem:

$$h(i) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} p(i/(m,n)) \quad i = 0, 1, \ldots, 2^B - 1 \qquad p(i/(m,n)) = \begin{cases} 1 & \text{gdy } L(m,n) = i \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Histogram obrazu można uzyskać za pomocą funkcji imhist(image), a jego wyrównanie za pomocą polecenia histeq(image).

Kolejną operacją jest binaryzacji obrazu. Polega ona na sprowadzeniu obrazu kolorowego, monochromatycznego do obrazu binarnego. Powoduje znaczną redukcję informacji ale pozwala np. na stosowanie algorytmów rozpoznawania obrazu. Wyróżniamy binaryzacje:

Z progiem dolnym	Z progiem górnym	Z podwójnym ograniczeniem	warunkową
$L'(m,n) = \begin{cases} 0; & L(m,n) \le a \\ 1; & L(m,n) > a \end{cases}$	$L'(m,n) = \begin{cases} 0; & L(m,n) \ge a \\ 1; & L(m,n) < a \end{cases}$	$L'(m,n) = \begin{cases} 0; & L(m.n) \le a_1 \\ 1; & a_1 < L(m,n) \le a_2 \\ 0; & L(m,n) > a_2 \end{cases}$	$L'(m,n) = \begin{cases} 0; & L(m,n) \le a_1 \\ s; & a_1 < L(m,n) \le a_2 \\ 1; & L(m,n) > a_2 \end{cases}$

Gdzie:

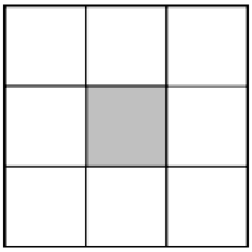
L(m,n) - jasność punktu w obrazie źródłowym $L(m,n) \in [0, 2^B-1]$;
L' (m,n) - wartość odpowiedniego punktu w obrazie wynikowym $L'(m,n) \in \{0,1\}$,
a - próg binaryzacji.

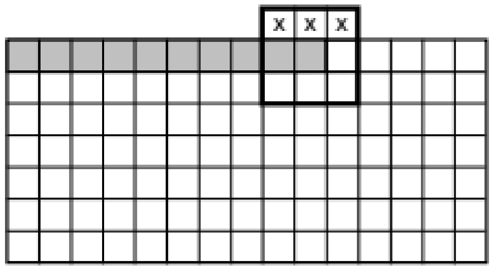
Oraz wielokryterialna – niezależnie przeprowadzana na wielu obszarach obrazu o różnych poziomach jasności. Binaryzacje przeprowadzani się najczęściej znajdując dołek (próg binaryzacji) na histogramie, oddzielający obiekt od tła. Wykorzystuje się polecania im2bw(image, prog), gdzie próg jest liczbą rzeczywistą z zakresu 0 do 1. Lepszą funkcją jest imbinarize(image, metoda_lub_prog).

W operacjach punktowych można wyróżnić przekształcenia liniowe i nieliniowe, zależy to od wykonywanej operacji.

KONTEKSTOWA FILTRACJA OBRAZÓW

Filtry cyfrowe stanowią bardziej złożone narzędzie przetwarzania obrazów niż przekształcenia jednopunktowe, z reguły filtry zakładają kontekstowość operacji, oznacza to że dla wyznaczenia jednego punktu obrazu wynikowe brane pod uwagę są punkty jego otoczenia. Otoczenie badanego piksela może mieć różną formę, ale najczęściej wykorzystuje się kwadratowe okno, zawierające przetwarzany punkt w środku.





Jednak z powodu takiego wyboru otoczenia, punkty na granicach obrazu nie mogą być poddane filtracji. Operacje filtracji są często stosowanie i istotnie zmieniają zawartość obrazu oraz geometrie znajdujących się na nich obiektów. Pozwalają na redukcji szumów, i wydobycie z nich cech nie widocznych na pierwszy rzut oka. Wykorzystywane jest do:

- tłumienia niepożądanych szumów,
- wzmacnianie w obrazi elementów zgodnych ze wzorcem,
- usuwanie wad obrazów,
- poprawa jakości,
- rekonstrukcja obrazów.

Stosowanie filtrów jest proste i intuicyjne, filtry są pewną funkcją przekształcająca jeden obraz w drugi metodą piksel po pikselu. Wyróżnia się filtry liniowe (oparte o liniową kombinację wybranych pikseli) i nieliniowe (oparte na nieliniowych operacjach na wybranych pikselach). Filtry linowe są prostsze w wykonaniu ale nieliniowe dają więcej możliwości. Filtry zaliczane są do liniowych jeśli spełniają warunki:

Addytywność	Jednorodność	Niewrażliwość na przesunięcie
$\varphi(f+g)=\varphi(f)+\varphi(g)$	$\varphi(\lambda f)=\lambda\,\varphi(f),\ \lambda \in \Re$	$\varphi(f_{\vec{h}})=[\varphi(f)]_{\vec{h}}$

Gdzie φ jest funkcja realizującą filtr, f i g są obrazami podlegającymi filtracji a h jest wektorem przesunięcia.

Operacją realizująca filtry liniwe, warto się posłużyć pojęciem konwolucji (splotem funkcji):

$$g(x)=(f \times h)(x)=\int_{-\infty}^{+\infty} f(x-t)h(t)dt$$

Splot (g) jest zdefiniowany w całym zbiorze liczb rzeczywistych, a iloczyn f(x-t)h(t) jest całkowalny w całym R. Funkcja h może mieć skończoną dziedzinę, w takim przypadku konwolucja wykorzystująca funkcję h staję się filtrem. Funkcję realizującą tłumienie szumów na zasadzie lokalnych średnich można zrealizować za pomocą:

$$g(x)=\frac{1}{2a}\int_{-a}^{a} f(t)dt=(f \times h)(x), \qquad \text{gdzie:} \qquad h(u)=\begin{cases}\frac{1}{2a}; & u \in <-a,a> \\ 0; & u \notin <-a,a>\end{cases}$$

Przy rozważaniu właściwości filtrów liniowych znaczeni ma odpowiedź impulsowa filtru czyli reakcja na impuls w postaci delty Diraca. Delta Diraca ma następujące własności a splatając ją z funkcją h otrzymujemy odpowiedź impulsową filtru, stanowiącą jego matryce.

1. $\delta_0 = 0; \quad x \neq 0$

2. $\int_{-\infty}^{\infty} \delta_0(x)dx = 1$

$$(h \times \delta_0)(x) = \int_{-\infty}^{\infty} \delta_0(x-t)h(t)dt = h(x),$$

Własności konwolucji:

- Łącznosć – (f x g) x h = f x (g x h) = f x g x h – pozwala na rozdzielenie filtrowania dowolnie dużą matrycą na kolejne filtrowanie przy pomocy małych matryc,
- Rozdzielnosć – pozwala rozdzielić filtracji dwuwymiarowej funkcji jako złożenie filtracji jednowymiarowych,

Do usuwanie zakłóceń z obrazu wykorzystuje się np. filtry dolnoprzepustowe (uśredniające), usuwają one drobne zakłócenia z obrazu (zanikają, a raczej zostają zamazane) czarne punkty na jasnym tle lub na odwrót), wygładzają zawirowania krawędzi wokół obiektów, usuwają drobne efekty zafalowania. Wadą jest rozmycie konturów i kształtów obiektów, co pogarsza ich rozpoznawalność.

Przykładowa matryca filtru dolnoprzepustowego

1	1	1
1	1	1
1	1	1

W matalbie filtrację obrazu można przeprowadzić za pomocą:
filtr = ones(rozmiar)/sum(sum(ones(rozmiar))
Imfilter(image, filtr)

Aby przypisać większą wagę dla pikseli środkowych można zdefiniować matrycę:

1	2	1
2	4	2
1	2	1

W przypadku filtrów górnoprzepustowych (gradientów) możemy wydobyć z obrazu szybkie zmiany jasności a więc konturów, krawędzi, faktury. Filtry te wyostrzają sygnał (uwypuklają krawędzie). Krawędź można zdefiniować jako linie oddzielająca obszary o różnej jasności, a więc opisać je można skokiem jednostkowym w zerze.

$$i(z) = \begin{cases} 1; z > 0 \\ \frac{1}{2}; z = 0 \\ 0; z < 0 \end{cases} \qquad i(z) = \int_{-\infty}^{z} \delta(t)dt; \quad \text{gdzie } \delta(t) \text{ - delta Diraca}$$

Przyjmując ze linia jest dowolnie zorientowaną prostą można zapisać ja w postaci kanonicznej, a jasność w poszczególnych punktach można zapisać wg wzoru f(x,y):

$$x\sin\alpha - y\cos\alpha + \beta = 0 \qquad f(x,y) = L_1 + (L_2 - L_1)i(x\sin\alpha - y\cos\alpha + \beta)$$

Licząc pochodne cząstkowe otrzymujemy gradient intensywności, a do wykrywania krawędzi wykorzystywany jest jego kwadrat długości:

$$\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 = \left((L_2 - L_1)\delta(x\sin\alpha - y\cos\alpha + \beta)\right)^2$$

Obliczając drugą pochodną funkcji f otrzymujemy Lapasjan funkcji intensywności:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} = (L_2 - L_1)\delta'(x\sin\alpha - y\cos\alpha + \beta)$$

Jest on symetryczny względem obrotu, zachowuje znak krawędzi (mówi czy intensywność za krawędzią jest większa czy mniejsza), jest operatorem liniowym (konwolucja). Filtry górnoprzepustowe realizują zawsze formę różniczkowania, a więc są kierunkowe. Przykładowe matryce gradientów poniżej. Gradient Robertsa daje się opisać za pomocą macierzy 2x2. Realizuje on różniczkowanie pod kątem 45 stopni.

Gradient Robertsa

0	0	0
-1	0	0
0	1	0

Gradient Robertsa 2x2

-1	0
0	1

Gradient Robertsa 2x2

0	1
-1	0

Rozszerzając gradient Robertsa na matryce 3x3 o różnych kierunkach otrzymujemy maski Prewitta.

Maska Prewitta

-1	0	1
-1	0	1
-1	0	1

Maska Prewitta

-1	-1	-1
0	0	0
1	1	1

Maski te mają orientacje horyzontalne i wertykalne. Modyfikując te maski poprzez różne wagi otrzymujemy maski Sobela, należy jednak pamiętać o unormowaniu wag jeśli ich suma nie daje 0. Maski te można obracać co 45 stopni.

Maska Sobela

-1	0	1
-2	0	2
-1	0	1

Maska Sobela

-1	-2	-1
0	0	0
1	2	1

Maski wykrywające narożniki:

Prawy górny narożnik

1	1	1
-1	-2	1
-1	-1	1

Lewy dolny narożnik

1	-1	-1
1	-2	-1
1	1	1

Gdy zachodzi konieczność filtracji górnoprzepustowej bez konkretnego kierunku należy zastosować Laplasjany. Wykrywają one krawędzie, kontury niezależnie od kierunku, do tego celu można użyć wielu metod, szczególnie nieliniowych, jednak w prostych zadaniach laplasjany wystarczają:

$$L'(m,n) = \frac{\partial^2 L(m,n)}{\partial m^2} + \frac{\partial^2 L(m,n)}{\partial n^2}$$

W praktyce w przetwarzaniu obrazów laplasjanem nazywa się wynik konwolucji obrazu z maską.

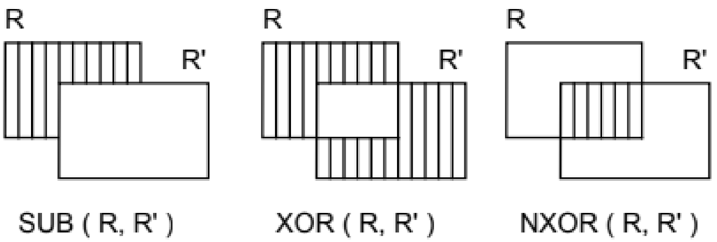
0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Filtry kombinowane mogą wykorzystywać wynik z filtracji pod różnymi kątami a następnie połączenia obrazów za pomocą formuły Euklidesowej lub modułowej.

Filtry nieliniowe mogą wykorzystywać także operacje logiczne. Do podstawowych operacji logicznych na obrazach binarnych zalicza się:

- Zaprzeczenie – NOT,
- Iloczyn logiczny – AND,
- Suma logiczna – OR,
- Rożnica logiczna – SUB,
- Suma rozłączna – XOR,
- Równoważność logiczna – NXOR,



Dla obrazów nie binarnych można stosować logikę rozmytą.

Złożenie dwóch filtrów można wykonać za pomocą:

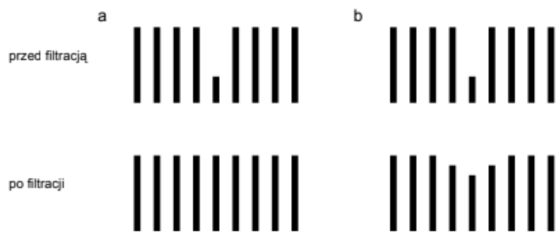
I1 = imfilter(image, filtr1); I2=imfilter(image, filtr2); I3 = imadd(i1, i2);

I1 = imfilter(image, filtr1); I2=imfilter(image, rot90(filtr1)); I3 = imadd(i1, i2);

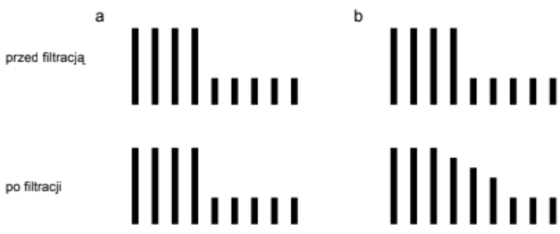
W matlabie istnieje funkcja fspecial pozwalająca zastosować filtry predefiniowane jak:

- Uśredniający,
- Dyskowy,
- Gaussa,
- Laplasjan,
- Laplasjan filtra gaussa,
- Ruchu kamery,
- Prewitta,
- Sobela

Filtry medianowe są znacznie lepsze w odszumianiu niż filtry dolnoprzepustowe, są jednak nieliniowe, ale nie powdują rozmycia i zachowują kontury. Nie wprowadza ona nowych wartości do obrazu, nie wymaga skalowania.



Rys. 3.82. Usuwanie zakłóceń filtrem medianowym (a) i filtrem uśredniającym (b).



Rys. 3.83. Wpływ filtru medianowego (a) i filtru uśredniającego (b) na brzegi obiektu.

Filtry adaptacyjne, to takie które dostosowują swoją charakterystykę do otoczenia. Większość filtrów działa jednostronnie, natomiast adaptacyjne filtry są bardziej wyrafinowane. Na przykład uśredniający filtr adaptacyjny działa dwuetapowo:

- Wyznacza parametr klasyfikujący punkt do krawędzi (np. na podstawie wariancji jasności),
- W drugim etapie filtruje się dolnoprzepustowo tylko te punkty które nie są krawędzia.

TRANSFORMACJA FOURIERA DLA OBRAZÓW CYFROWYCH

Transformacja Fouriera przedstawia zawartość częstotliwościową obrazu. Jest on reprezentowany w postaci liczb zespolonych. Innymi transformacjami w IPToolbox są DCT i Radon transform. W pierwszej wyznaczana jest dyskretna transformata kosinusowa obrazu a w drugiej obraz jest reprezentowany jak zbiór projekcji obrazu względem zmieniającego się kierunku.

Dwuwymiarowa transformata Fouriera i transformata odwrotna:

$$F(\omega_1, \omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m,n) e^{-j\omega_1 m} e^{-j\omega_2 n}$$

$$f(m,n) = \frac{1}{4\pi^2} \int_{\omega_1=-\pi}^{\pi} \int_{\omega_2=-\pi}^{\pi} F(\omega_1, \omega_2) e^{j\omega_1 m} e^{j\omega_2 n} d\omega_1 d\omega_2$$

Częstości są w radianach, F(0, 0) jest sumą wszystkich wartość f(m, n) – odpowiednik składowej stałej dla transformaty Fouriera. W obrazach cyfrowych wykorzystujemy postać dyskretną daną wzorami:

$$F(p,q) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n) e^{-j2\pi/Mpm} e^{-j2\pi/Nqn}$$

$$F(m,n) = 1/MN \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} F(p,q)\, e^{j2\pi/Mpm} e^{j2\pi/Nqn}$$

Gdzie p i m zmieniają się: 0, 1, .. M -1 a q i n: 0, 1, .. N-1

Transformatę możemy traktować jako F-obraz, o takich samych wymiarach, ale możemy zdecydować czy chcemy obserwować amplitudę czy fazę. Ze względu na okresowość funkcji f(m, n) reprezentacje ogranicza się do zakresu -π,π. Transformata posiada 2 osie symetrii i można dokonać przełożenia ćwiartek tak by uzyskać symetryczny obraz. Polecenia w matalbie:

F = fft2(image), F_abs = abs(ff2(image)), F_shifted = abs(fftshift(fft2(image)));

Wykorzystując transformatę obrazu możemy dokonywać filtracji w dziedzinie częstotliwości (a więc splatać obraz z filtrem). Filtracja staje się wtedy mnożeniem F-obrazów. A więc:

F = fft2(image); F_filtr = complex(zeros(size(F))); F_filtr(…) = 1+1i; F = F.*F_filtr; image=uint8(ifft2(f))

Za pomocą filtracji w dziedzinie częstotliwości można dokonać kompresji obrazów. Pomimo usunięcia (wyzerowania) np. 90% amplitud obraz jest rozpoznawalny. Może także służyć do innych operacji jak sumowanie obrazów.

Obliczając odwrtoną transformację obrazu filtru otrzymujemy jego odpowiedz impulsową a na jej podstawie macierz konwolucji.

Transformacja kosinusowa z kolei używana jest do kompresji obrazów (JPEG). Dla obrazu o wielkości MxN zdefiniowana jest następująco:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{matrix}$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2/M}, & 1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2/N}, & 1 \le q \le N-1 \end{cases}$$

Macierz transformaty B_{pq} jes takżę wymiaru MxN. Dla dowolnej macierzy ustalonego rozmiaru można wyznaczyć ogolną macierz funkcji bazowych:

$$\alpha_p \alpha_q \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le p \le M-1 \\ 0 \le q \le N-1 \end{matrix}$$

Przekształcenie odwrotne zdefiniowane jest następująco:

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \le m \le M-1 \\ 0 \le n \le N-1 \end{matrix}$$

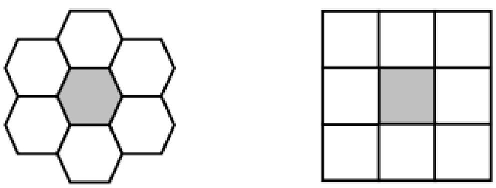
$$\alpha_p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2/M}, & 1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2/N}, & 1 \le q \le N-1 \end{cases}$$

DCT charakteryzuje się tym, że dla obrazu większość istotnej informacji znajduje się w ograniczonym obszarze dziedziny DCT, stwarza to możliwość wykorzystania jej w komprezji stratnej obrazu. W przypadku JPEG obraz dzieli się na tablice 8x8 lub 16x16, dla każdej z nich wyznacza się DCT i dla każdej macierzy DCT zeruje się odpowiednią ilość pól dla danego stopnia komprezji.

PRZEKSZTAŁCENIA MORFOLOGICZNE

Przekształcenia morfologiczne są bardzo ważne w analizę obrazów, pozwalają one na najbardziej złożone operacje związane z analizą kształtu elementów, ich wzajemnego położenia a także umożliwiają złożone procesy symulacji. Są punktem wyjściowym do tworzenia złożonych operacji.

Podstawowym pojęciem jest element strukturalny, jest to pewien wycinek obrazu (zbiór elementów dla reprezentacji dyskretnej) z wyróżnionym punktem centralnym. Najczęściej stosuje się koło o promieniu 1 elementu. Dla różnych siatek przyjmuje one różne postacie:



Ogólny algorytm przekształcenia morfologicznego:

- Element strukturalny jest przemieszczany po całym obrazie i dla każdego punktu obrazu wykonywana jest analiza koincydencji punktów obrazu i elementu, przy założeniu że każdy punkt obrazu jest punktem centralnym elementu,
- W każdym punkcie obrazu następuje sprawdzenie czy rzeczywista konfiguracja pikseli obrazu w otoczeniu tego punktu zgodna jest z wzorcowym elementem strukturalnym,
- W przypadku wykrycia zgodności wzorca pikseli obrazu i szablonu następuje wykonanie pewnej operacji na badanym punkcie (np. zmiana nasycenia)

To samo tylko prościej:

Intuicyjnie można powiedzieć, że każde przekształcenie morfologiczne polega na:

- przyłożeniu centralnego punktu kolejno do wszystkich punktów obrazu;
- sprawdzeniu, czy lokalna konfiguracja punktów odpowiada układowi, zapisanemu w elemencie strukturalnym;
- wykonaniu, w przypadku zgodności konfiguracji punktów, operacji określonej dla danego przekształcenia.

Erozja

Erozja – zakłada się ze istnieje nieregularny obszar X i koło B o promieniu R, które jest elementem strukturalnym. Jako punkt środkowy przyjmuje się środek koła B. Wówczas erozję figury X elementem B można zdefiniować:

- figura zerodowana to zbiór środków wszystkich kol B które w całości są zawarte we figurze X,
- koło B przetacza się po wewnętrznej stronie brzegu figury, kolejne położenia środka kola wyznaczają brzeg figury zerodowanej.

Czyli chodzi o usuniecie wszystkich punktów o wartości 1 które posiadają choć 1 sąsiada o wartości 0. Jest to filtr minimalny z otoczenia zdefiniowanego prze element strukturalny.

$$L'(m,n) = \min_{m_i, n_i \in B(m,n)} (L(m_i, n_i))$$

gdzie: *L(m,n)* - jasność punktu o współrzędnych *(m,n)*;
B(m,n) - element strukturalny z punktem centralnym o współrzędnych *(m,n)*.

Erozja posiada następujące cechy, mające duże znaczenie praktyczne:

- Jest addytywna, co oznacza, że erozję o założonej wielkości można interpretować jako złożenie odpowiedniej ilości erozji o wielkości jednostkowej. Cecha ta ma wielkie znaczenie przy praktycznej realizacji przekształcenia.
- Erozja złożonym elementem strukturalnym jest równoważna złożeniu erozji poszczególnymi elementami tego elementu strukturalnego.
- Położenie punktu centralnego elementu strukturalnego nie ma większego znaczenia. Zmiana położenia punktu centralnego o dany wektor powoduje przesunięcie (translację) obrazu wynikowego o taki sam wektor.
- Erozja ma zdolność do eliminacji drobnych szczegółów i wygładzania brzegu figury.
- Erozja elementami strukturalnymi o podłużnym kształcie pozwala uwypuklić fragmenty obrazu zorientowane liniowo w tym samym kierunku, co element strukturalny.
- Erozja dokonuje generalizacji obrazu. Odizolowane, drobne wyróżnione obszary zostają usunięte. Brzegi wyróżnionych obszarów zostają wygładzone, ich długość zostaje zdecydowanie zmniejszona.

- Zmniejszone zostają także ich powierzchnie. Często większe wyróżnione obszary podzielone zostają na mniejsze30. Spada nasycenie obrazu.

W matlabie:

Se = strel(kształt, rozmiar); bw = imerode(bw, se);

Można ją stosować także do obrazów monochromatycznych oraz kolorowych.

Dylatacja

Dylatacja jest równie ważna jak erozja, jest do niej odwrotna. Zakładając ze znowu mamy obszar X i koło B o promieniu R, to dylatację figury X elementem B można zdefiniować:

- Figura po dylatacji jest zbiorem środków wszystkich kół B dla których choć jeden punkt pokrywa się z jakimkolwiek punktem figury,
- Koło B przetacza się po zewnętrznej stronie figury X, kolejne położenia środka koła B wyznaczają brzeg figury po dylatacji.

$$L'(m,n) = \max_{m_i, n_i \in B(m,n)} (L(m_i, n_i))$$

gdzie: *L(m,n)* - jasność punktu o współrzędnych *(m,n)*;
B(m,n) - element strukturalny z punktem centralnym o współrzędnych *(m,n)*.

Jest to filtr maksymalny, można zdefiniować jako negatyw erozji negatywu obrazu xD. Podstawowe zalety i właściwości dylatacji:

- zamykanie małych otworów i wąskich zatok w konturach obiektów na obrazie;
- zdolność do łączenia obiektów, które położone są blisko siebie;
- brzegi są wygładzane, ich długość się zmniejsza,
- zwiększa powierzchnie obrazów,

W matlabie:

Se = strel(kształt, rozmiar); bw = imdilate(bw, se)

Oba te przekształcenia dokonują generalizacji obrazu.

Otwarcie i zamkniecie obrazu

Są to kombinacje erozji i dylatacji. **Otwarcie** polega na przetaczaniu kola B po wewnętrznej stronie brzegu figury i odrzuceniu wszystkich tych punktów które nie mogą być osiągnięte przez koło, **zamkniecie** polega na przetaczaniu kola B po zewnętrznej stronie brzegu figury i dodanie do niej wszystkich punktów które nie mogą być osiągnięte prze koło.

$$\text{otwarcie} = \text{erozja} + \text{dylatacja}$$
$$\text{zamknięcie} = \text{dylatacja} + \text{erozja}$$

Własności „użytkowe", operacji otwarcia i zamknięcia dla przypadku obrazów binarnych są następujące:

- otwarcie usuwa drobne obiekty i drobne szczegóły, jak półwyspy i wypustki, nie zmieniając wielkości zasadniczej części figury, może też rozłączyć niektóre obiekty z przewężeniami;

- zamknięcie wypełnia wąskie wcięcia i zatoki oraz drobne otwory wewnątrz obiektu, nie zmieniając wielkości jego zasadniczej części, może też połączyć leżące blisko siebie obiekty;
- obydwie operacje nie zmieniają kształtu ani wymiarów dużych obiektów o wyrównanym, gładkim brzegu.

W Matlabie:

Se = strel(kształt, rozmiar);

OTWARCIE: bw1=imerode(bw1, se); bw1 = imdilate(bw1, se); lub imopen(bw, se)

ZAMKNIECIE: bw1 = imdilate(bw1, se); bw1=imerode(bw1, se); lub imclose(bw, se)

Inną opcją jest bmorph(bw, operacja, liczba_powtrzen). Operacja może to być otarcie czy zamkniecie ale także detekcja ekstremów itd. Liczba powtórzeń wykonania operacji może być podana, lub będzie powtarzana aż obraz wynikowy przestanie się zmieniać.

Detekcja ekstremow – top-hat, bot-hat

Aby wyodrębnić z obrazu lokalne maksima (TOP HAT) należy od wyniku otwarcia oddanego obrazu odjąć obraz wyjściowy i dokonać binaryzacji z progiem otrzymanej różnicy. Aby wyodrębnić minima (BOT HAT) jest podobnie ale zamiast otwarcia stosuje się zamkniecie.

$$M(f) = B(O(f) - f),$$

$$m(f) = B(C(f) - f)$$

Efekty zależą od wielkości otwarcia/zamknięcia i progu binaryzacji. W matlabie:

Imtophat(bw, element_str); imbothat(bw, element_str); potem: imadjust(bw, strechlim(bw))

LUB: Bwmorp(bw, ‘tophat, n); bwmorph(bw, ‘bothat’, n)

Scienianie

Jest wspólna nazwa podzbioru przekształceń morfologicznych przeprowadzanych w następujący sposób. Ścienianie obiektu X przy użyciu elementu B polega na przyłożeniu tego elementu do każdego punktu obrazu punktem centralnym i podjęciu decyzji:

- nie zmieniać punktu, gdy element nie pokrywa się z jego sąsiedztwem,
- zmienić wartość na 0 jeśli element pasuje do otoczenia.

Można je przeprowadzać wielokrotnie aż do momentu gdy dalsze ścienianie nie wprowadza zmian.

Bwmorph(bw, ‘remove’)

Cechą charakterystyczna jest to ze figura wyjściowa zawiera się w figurze wejściowej.

Szkieletyzacja

Może być realizowana jako ścienianie z rotowanym elementem strukturalnym. Proces może wprowadzać do obrazu pewne artefakty w postaci bocznego gałązkowania linii szkieletu, efekt ten jest kłopotliwy jeśli oryginalna figura posiada zakłócenia, które mogą mieć siermiężny wpływ na kształt szkieletu.

x	0	0
x	1	x
1	1	1

Bwmorph(bw, ‘skel’, Inf);

Może być realizowana jako ścienianie z rotowanym elementem strukturalnym.

Obcinanie powstałych gałęzi – polega na stopniowym redukowaniu odcinków posiadających wolne zakończenie, w ostateczności pozostają jedynie zamknięte pętle.

0	x	x
0	1	0
1	1	1

Bwmorph(bw, 'spur', inf)

WYZNACZANIE CENTROIDÓW

Centroid jest szkieletem figury ze szczególnie mocno obciętymi gałęziami, operacja ich wyznaczania sprowadza figurę do punktu w przypadku gdy figura nie posiada otworów, lub doi pętli gdy takowe posiada.

Wyznaczanie centroidów realizowane jest jako ścienianie przy użyciu dwóch elementów.

0	x	x
0	1	1
0	0	x

x	0	x
0	1	1
x	0	x

Bwmporh(bw, 'shrink', Inf); bwmorph(bw, 'remove') – usuwa wnętrzne

ZALEWANIE OTWORÓW

Polega na wypełnieniu otworów. Może to być potrzebne z 3 powodów:

- gdy otwory nie odpowiadaj rzeczywistym otworom lecz powstały jako np. odblask światła,
- potrzebne jest wyznaczeni parametrów obiektu bez uwzględniania ich otworów,
- trzeba skupić uwagę wyłącznie na otworach.

Przykładem zastosowani jest komputerowe przetwarzanie obrazów w celu automatycznego rozpoznawania odcisków palca. Funkcja: imfill(image, 'holes'); bwmorph(bw, 'fill');

ORDER CIESIELKI ZA DOTARCIE NA SAM KONIEC.