copyright: M. Dwornik
dwornik@agh.edu.pl

# IPT, Przekształcenia punktowe i geometryczne

## 1 Image Processing Toolbox

### 1.1 Formaty przechowania obrazu

- BIN - macierz binarna (logiczna), przyjmująca wartości true / false
- GRAY - macierz poziomów szarości. Wartości mogą być typu double (0-1), uint8 (0-255) lub uint16.
- RGB - potrójna tablica intensywności (3 tablice typu GRAY)
- IND - tablica indeksowa (składa się z 2 części: mapy i legendy), gdzie legenda jest tablicą o wymiarze Nx3 (gdzie N - ilość dostępnych kolorów), a mapa jest tablicą z numerami kolorów (numerami odpowiadającym wersom w legendzie).

### 1.2 Formaty kodowania koloru

Istnieje wiele sposobów kodowania koloru. Oprócz najpopularniejszego RGB, bazującego na ludzkim sposobie postrzegania barw czy używanym w drukarkach CMYK, istnieją również specyficzne formaty mające na ogół bardzo wąskie zastosowanie. Generalnie podzielić je można ze względu na ilość tablic (kanałów) służących do przechowywania pojedyńczego koloru.

1. Pojedyńcza tablica: Gray
2. Potrójna tablica:
   - RGB (Red, Green, Blue)
   - YCbCr, zwane YUV (Luminance, Chrominance)
   - H S V (Hue, Saturation, Value) - oparte o stożek: H=0:360; S,V=0:100.
   - CMY
   - $L^*a^*b^*$ - (Luminance, a i b)
   - XYZ
   - HLS (Hue, Saturation, Lightness)
3. Poczwórna tablica: CMYK (Cyan, Magenta, Yellow, Key (black))
4. n-kanałowe.

Przykładowe wartości kolorów przedstawia tab.1.

Konwersji pomiędzy tymi formatami i typami używa się przy wykorzystaniu funkcji o nazwie typ2typ, np.: `rgb2gray()` czy `ycbcr2rgb()`.

Dla przykładu, przekształcenie RGB->YIQ polega na macierzowym mnożeniu macierzy przekształcenia przez wektor RGB. Macierz przekształcenia przedstawiano jest poniżej. Format YIQ jest powszechnie wykorzystywany przy analizie barwionych szlifów mikroskopowych

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} \mathbf{0.299} & \mathbf{0.587} & \mathbf{0.114} \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & -0.311 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Z kolei przekształcenie RGB->CMY i odwrotne, polega na odjęciu od największej dopuszczalnej wartości danej palety:
CMY = 255 - RGB (dla uint8) i RGB = 1 - CMY (dla double).

**Tabela 1:** Wartości przykładowych kolorów w wybranych kodowaniach koloru (UWAGA: kolory XYZ, CMY, CMYK zostały określone na podstawie RGB przy użyciu http://www.easyrgb.com/index.php?X=CALC, natomiast HSV, YUV i YIC na podstawie: http://web.forret.com/tools/color.asp)

| format | czarny | biały | czerwony | zielony | niebieski |
|---|---|---|---|---|---|
| RGB | 0, 0, 0 | 1, 1, 1 | 1, 0, 0 | 0, 1, 0 | 0, 0, 1 |
| XYZ | 0, 0, 0 | 95.05, 100, 108.9 | 41.24, 21.26, 1.93 | 35.76, 71.52, 11.92 | 18.05, 7.22, 95.05 |
| CMY | 1, 1, 1 | 0, 0, 0 | 0, 1, 1 | 1, 0, 1 | 1, 1, 0 |
| CMYK | 0, 0, 0, 1 | 0, 0, 0, 0 | 0, 100, 100, 0 | 100, 0, 100, 0 | 100, 100, 0, 0 |
| HSV | 0, 0, 0 | 0, 0, 100 | 0, 100, 100 | 120, 100, 100 | 240, 100, 100 |
| HSL | 0, 0, 0 | 0, 0, 100 | 0, 100, 50 | 120, 100, 50 | 240, 100, 50 |
| YUV | 0, 0, 0 | 100, 0, 0 | 29.9, -14.7, 61.5 | 58.7, -28.9, -51.5 | 11.4, 43.6, -10 |
| YIC | 0, 0, 0 | 100, 0, -62.2 | 29.0, 59.6, 21.1 | 58.7, -27.4, -52.3 | 11.4, -32.1, -31.1 |

### 1.3 Rozdzielczość przestrzenna

Rozdzielczość przestrzenna definiowana jest na wiele sposobów. Jedna z najczęściej spotykanych definicji stanowi, że jest to rozmiar powierzchni jaką zajmuje dany piksel (wielkość terenu). W przypadku wydruków i obrazów często spotykaną miarą rozdzielczości przestrzennej jest DPI (ang. *dot per inch*) czyli ilość punktów (pikseli) na 1 cal wydruku lub skanu.

W MatLABie istnieją 3 metody interpolacji:

- 'nearest' - najbliższego sąsiada,
- 'bilinear' - interpolacja dwuliniowa,
- 'bicubic' - interpolacja kubiczna (dwukwadratowa).

### 1.4 Rozdzielczość barwna

Rozdzielczość poziomów szarości dla obrazów monochromatycznych $R_L$ definiowana jest jako (Wróbel & Koprowski):

$$\frac{1}{R_L} = \frac{l_w - l_n}{P} \tag{1}$$

gdzie:
$l_w$ i $l_n$ - wysoki i niski poziom szarości
P - liczba naturalna $P \in \{0, 1, 2, ..., 2^b - 2, 2^b - 1\}$
b - liczba bitów służących do reprezentacji danego poziomu szarości

## 2 Przekształcenia punktowe

Przekształcenia punktowe (bezkontekstowe) są to przekształcenia dotyczące stopnia szarości lub nasycenia barwy dla każdego punktu oddzielnie, dla których nie mają wpływu wartości w punktach sąsiednich.

### 2.1 Liniowe przekształcenia obrazu

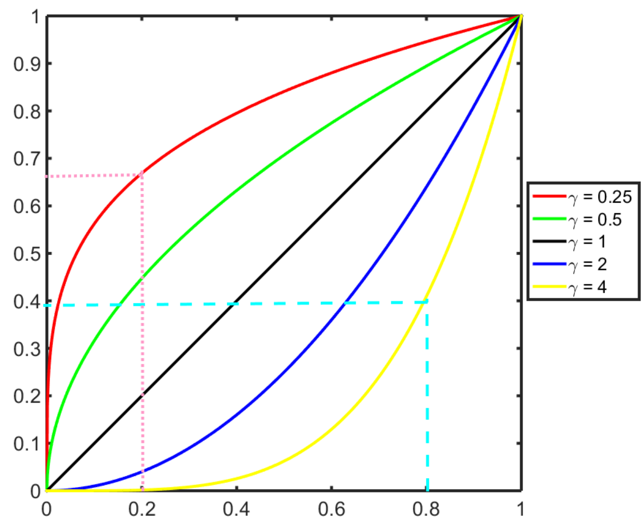
Liniowe przekształcenie polega na zmianie wartości piksela przy użyciu przekształcenia liniowego. Zaliczamy do nich między innymi operacje dodawania, negacji czy mnożenia przez pewną liczbę. Służą one np. do rozjaśniania obrazu. Należy pamiętać, że przy stosowaniu kodowania `uint8` dozwolone wartości są z przedziału 0-255, a dla kodowania `double` dozwolone są wartości z przedziału 0-1.

copyright: M. Dwornik
dwornik@agh.edu.pl

### 2.2 Nieliniowe przekształcenia obrazu

Do nieliniowcych przekształceń zaliczamy np. pierwiastkowanie, potęgowanie logarytmowanie. W niektórych przypadkach jest to znacznie skuteczniejsze niż liniowe korekcje. Szczególne znaczenie ma korekcja gamma $\gamma$ dana wzorem eq. 2. Idee działania współczynnika $\gamma$ na obraz przedstawia Rys. 1.

$$Nowy(m,n) = Stary(m,n)^{\gamma} \tag{2}$$



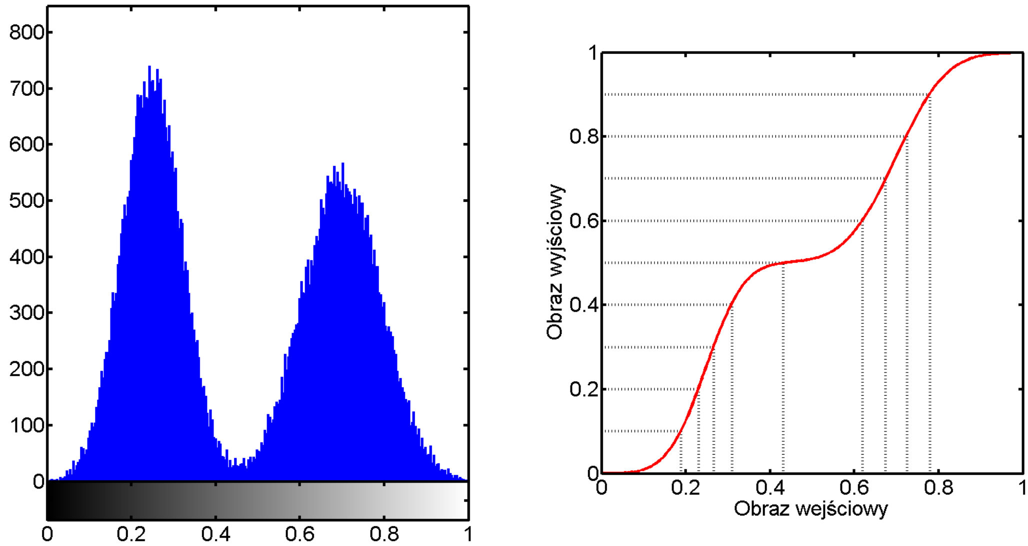

**Rysunek 1:** Idea działania korekcji $\gamma$. Dla $\gamma < 1$ zwiększony jest kontrast dla "ciemnej" części obrazu kosztem "jasnej" części. Dla $\gamma > 1$ efekt jest odwrotny: zyskujemy kontrast w części "jasnej' kosztem "ciemnej".

### 2.3 Wyrównywanie histogramu

Histogram wykreśla się poleceniem `imhist(obraz)`. Histogram jest to wykres słupkowy określający ile punktów przybiera daną wartość (intensywność). Wyrównanie histogramu polega na "upodobnieniu" go do histogramu dla rozkładu równomiernego (wszystkie intensywności mają jednakową częstotliwość występowania) (Rys. 2). Do wyrównywania histogramu służy polecenie `histeq(obraz, ilość_klas)`. Czasami stosuje się wyrównywanie lokalne, tzn. dzieli się obraz na kilka części i każdą wyrównuje oddzielnie. Do tego stosuje się funkcję `adapthisteq(obraz, parametry)`. Do najważniejszych parametrów tego przekształcenia zaliczamy:

- `NumTiles` - Wielkość elementu, dla którego wyrównujemy histogram
- `ClipLimits` - skalarna wielkość decydująca o kontraście (0,1)
- `NBins` - Ilość klas histogramu
- `Range` - zakres wyrównywanych wartości
- `Distribution` - rodzaj histogramu, do którego będzie normowany obraz: 'uniform' (płaski), 'Rayleigh' (dzwonkowy) i 'Exponential' (krzywa eksponenty).
- `Alpha` - parametr rozkładu histogramu, tylko dla Rayleigh i Exponential.

copyright: M. Dwornik
dwornik@agh.edu.pl




**Rysunek 2:** Idea wyrównania histogramu obrazu. 1) Liczymy histogram; 2) Liczymy kumulantę (czerwona linia) i dzielimy ją przez jej maksymalną wartość; 3) Dzielimy Oś OY na wykresie kumulanty na "n" równych odcinków (klas histogramu wyjściowego); 4) Szukamy wartości rzutów wartości poszczególnych progów na Oś OX poprzez kumulantę; 5) Wszystkie wartości z OX z ograniczone dwoma rzutami zamieniamy na odpowiadającą im wartość OY.

## 2.4 Normalizacja

Normalizacja ma na celu rozciągnięcie zakresu tonalnego obrazu. Jeżeli intensywność punktów na obrazie zajmuje tylko część dozwolonej skali, to możliwe jest rozciągnięcie intensywności tak, by zajmowała pełną skalę. Stosowana jest normalizacja zarówno w wymiarze globalnym, jak i lokalnym. Jeżeli przez M oznaczamy minimum w normalizowanym obszarze, przez N maksimum, a przez L i U odpowiednio minimum i maksimum skali, to intensywność punktu wynosi:

$$Nowy(m,n) = \left(\frac{L-U}{M-N} \cdot (Stary(m,n) - M)\right) + L \tag{3}$$

Obraz metodą globalną normalizuje się przy użyciu funkcji:
`imadjust(obraz, [low_in; high_in], [low_out; high_out], gamma)`.

## 2.5 Binaryzacja

Binaryzacja to zamiana dowolnego obrazu na obraz logiczny. Stosuje się kilka rodzajów binaryzacji:

- z dolnym progiem: wszystko poniżej progu: 0, powyżej 1

$$\mathbf{L(m,n)} = \begin{cases} 0 & gdy \quad L(m,n) \leqslant prg \\ 1 & gdy \quad L(m,n) > prg \end{cases}$$

- z górnym progiem: negatyw binaryzacji z dolnym progiem

$$\mathbf{L(m,n)} = \begin{cases} 0 & gdy \quad L(m,n) > prg \\ 1 & gdy \quad L(m,n) \leqslant prg \end{cases}$$

copyright: M. Dwornik
dwornik@agh.edu.pl

- z dwoma progami: przypomina funkcję bramkową lub jej negatyw
- wielokryterialna: kilka progów i tylko niektóre przedziały mają wartość prawda / fałsz.
- z histerezą: wymaga podania 2 wartości progowych (próg1<próg2), wartość piksela L(m,n) dana jest wzorem:

$$\mathbf{L(m,n)} = \begin{cases} 0 & gdy \quad L(m,n) \leqslant prg1 \\ s & gdy \quad prg1 < L(m,n) \leqslant prg2 \\ 1 & gdy \quad L(m,n) > prg2 \end{cases}$$

  s - wartość sąsiadujących punktów.
- automatyczne progowanie metodą maksymalnej entropii
  Wartość entropii **H** dla obrazu **A** definiowana jest poprzez eq. 4.

$$H = -\sum_{k=0}^{N_{kolor}} \left(p(k) \cdot \ln p(k)\right) \tag{4}$$

  gdzie `p(k)` - prawdopodobieństwo wystąpienia intensywności `k` w obrazie **A**, $N_{kolor}$ - maksymalna wartość intensywności występująca w obrazie **A**.
  Binaryzacja metodą maksymalnej entropii polega na znalezieniu takiej wartości progu `k`, dzielącej obraz na tło (piksele o intensywności mniejszej lub równej `k`) i obiekty (o intensywności większej niż `k`), która maksymalizuje funkcję H(k).

# 3 Przekształcenia geometryczne

Do przekształceń geometrycznych zaliczamy między innymi: przesunięcie, obrót, zniekształcenie, odbicia symetryczne, powielanie fragmentów. Mają one wpływ na kształt, wielkość lub wygląd obrazu.

## 3.1 Przesunięcie o wektor

Możliwe są przesunięcia o wektor z dodaniem nowej powierzchni (zwiększenie wymiarów obrazu wyjściowego o wektor) lub przesunięcie w sposób cykliczny z użyciem funkcji `circshift(obraz,wektor)`.

## 3.2 Obrót

Obraz można obracać przy użyciu funkcji `imrotate(obraz, kąt w stopniach, opcje)`. Pierwszą opcją są metody interpolacji nowych danych:

- najbliższego sąsiada (`'nearest'`);
- dwuliniowa (`'bilinear'`);
- bikubiczna (`'bicubic'`).

Drugą opcją jest rozmiar wyjściowy obrazu:

- `'crop'`: Obraz wyjściowy ma taki sam rozmiar jak wejściowy. Rogi są obcięte, wolne miejsca wypełnione zerami.
- `'loose'`: Obraz wyjściowy na ogół jest powiększony tak, by zmieścił się cały obraz. Wolne miejsca wypełnione zerami.

Opcjami domyślnymi jest `'nearest','loose'`. Użycie kąta dodatniego powoduje obrót przeciwnie do kierunku wskazówek zegara.

Rotacja obrazu polega na wyliczeniu nowych współrzędnych danego piksela obrazu. Sprowadza się to do wymnożenia wektora współrzędnych przez macierz rotacji:

$$\begin{bmatrix} x^{obr} \\ y^{obr} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}$$

## 3.3 Odbicia symetryczne

Odbicia symetryczne mogą być względem prostej pionowej (`fliplr (obraz)`) lub poziomej (`flipud (obraz)`). Odbicia te (a zwłaszcza odbicie poziome) mają zastosowanie podczas kalibracji (ang. registering) zdjęcia (przypisania współrzędnych, np. geograficznych, do poszczególnych pikseli obrazu).

## 3.4 Dodawanie wierszy, kolumn do obrazu

Dodawanie kolumn i wierszy do obrazu ma na celu głównie podczas filtracji w celu zmniejszenia efektów brzegowych. Służy do tego polecenie `padarray (obraz, rozmiar, metoda, kierunek);`
Rozmiar w formie wektora [`x y`] mówi, ile wersów `x` i ile kolumn `y` należy dodać.
Do metod zaliczamy:

- `'circular'` - doklejanie cykliczne

copyright: M. Dwornik
dwornik@agh.edu.pl

- `'replicate'` - doklejanie poprzez kopiowanie skrajnych wersów / kolumn
- `'symmetric'` - odbicie symetryczne.

Istnieją 3 kierunki:

- `'pre'`: doklejanie z góry i z lewej;
- `'post'`: zdołu i z prawej;
- `'both'`: z obu jednocześnie.

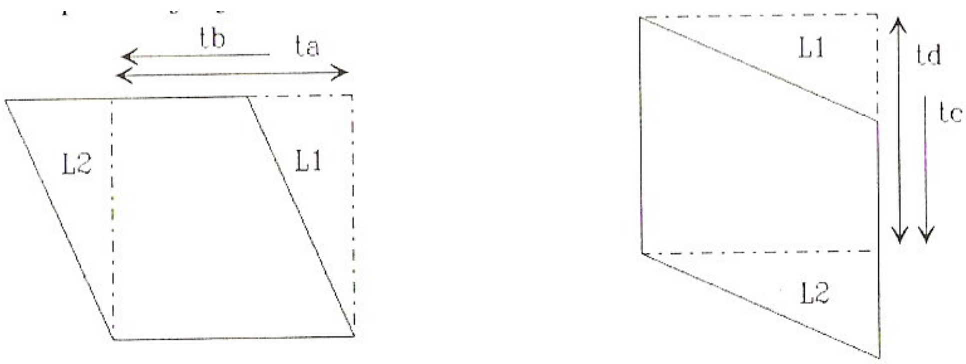
### 3.5 Zmiana kształtu i wymiarów obrazu

Zmiana wymiarów obrazu odbywa się przy wykorzystaniu polecenia `reshape (obraz, [nowy rozmiar], 'opcja');`. Do poprawnego działania funkcji musi być spełniony warunek zachowania powierzchni: $(x^{old} * y^{old} = x^{new} * y^{new})$
Do zmiany kształtu, bez konieczności spełnienia powyższego warunku może służyć polecenie `imresize(obraz, [nowy rozmiar], 'opcja')`.
Zmiana kształtu może odbywać się przy wykorzystaniu polecenia `imwarp(obraz, maska)`. Maska przekształcenia tworzona jest specjalnymi poleceniami.

Wyróżniamy następujące typy transformacji:

- przekształcenie afiniczne `affine2d([ macierz 3x3 ])`: różnowartościowe przekształcenie geometryczne, które wszystkie proste zawarte w dziedzinie tego odwzorowania przekształca na proste (Rys. 3). Ostatnia (trzecia) kolumna musi mieć postać `[0; 0; 1]`. Przekształcenie afiniczne dane jest wzorem:

$$[x_{new}\ y_{new}] = [x\ y\ 1] \cdot \mathbf{T} \tag{6}$$




**Rysunek 3:** Parametry przekształcenia afinicznego (źródło:Wróbel & Koprowski, 2004)

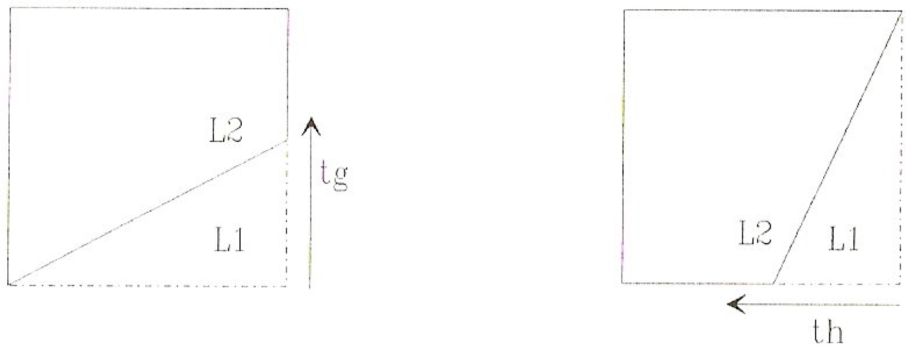
- Projekcję `projective2d([ macierz 3x3 ])`: przekształcenie, gdzie każde punkty leżące na jednej prostej przechodzą w punkty leżące na drugiej prostej (Rys. 4). Przekształcenie afiniczne jest szczególnym przypadkiem projekcji. Projekcja dana jest wzorem:

$$[x_{new}\ y_{new}\ z_{new}] = [x\ y\ w] \cdot \mathbf{T} \tag{7}$$

. W przypadku obrazów 2D, powyższe równania można zapisać w następujący sposób:

$$x_{new} = \frac{T_A \cdot x + T_B \cdot y + T_C}{T_G \cdot x + T_H \cdot y + T_I} \quad y_{new} = \frac{T_D \cdot x + T_E \cdot y + T_F}{T_G \cdot x + T_H \cdot y + T_I} \tag{8}$$

copyright: M. Dwornik
dwornik@agh.edu.pl




**Rysunek 4:** Parametry projekcji (źródło:Wróbel & Koprowski, 2004)

- `rigid2d(rot, przesunięcie)` - odwracalne przekształcenie polegające na rotacji (definiowanej przez macierz 2x2 rotacji) i przesunięciu (poziomy wektor 2 elementowy).

## 3.6 Rozpoznanie transformacji geometrycznej

Polecenie `fitgeotrans` pozwala na estymację wartości parametrów transformacji (np. obrotu czy skalowania). W tym celu konieczne jest określenie par punktów na dwóch obrazach (przed i po operacji) odpowiadającym tej samej lokalizacji. Do tej operacji można użyć funkcji `cpselect`. W zależności od ilości wyznaczonych par punktów, możliwa jest estymacja wartości parametrów przekształceń:

- 'Nonreflective similarity' - dla dwóch i więcej par punktów
- 'Affine' - dla 3 i więcej par punktów
- 'projective' - dla 4 i więcej par punktów

## 3.7 Fraktale

Przekształcenia afiniczne można wykorzystać do generowania fraktali (J. Kudrewicz, *Fraktale i Chaos*, WNT, 1993). W tym celu wykonywane są losowo wybrane przekształcenia afiniczne na pojedynczej współrzędnej. Wartość elementu tablicy o obliczonych współrzędnych zostaje zwiększona o jeden. Przykładowy komplet tablic przekształceń affinicznych zamieszczono poniżej.

$$\begin{bmatrix} -0.67 & -0.02 & 0 \\ -0.18 & 0.81 & 10 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0.4 & 0.4 & 0 \\ -0.1 & 0.4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.4 & -0.4 & 0 \\ -0.1 & 0.4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.1 & 0 & 0 \\ 0.44 & 0.44 & -2 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$