

Przetwarzanie i rozpoznawanie obrazów: Reprezentacja obrazów

Krzysztof Krawiec

Przetwarzanie i Rozpoznawanie Obrazów

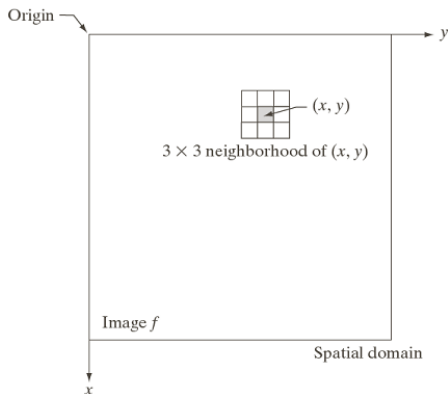
14 czerwca 2018

Cyfrowa reprezentacja obrazów

Obraz rastrowy (ang. *raster image*, *~bitmap*)

Podstawowe właściwości:

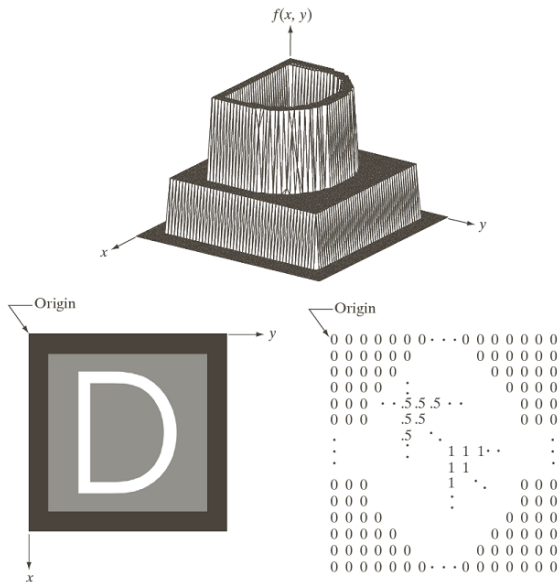
- szerokość (*width*)
- wysokość (*height*)
- [głębokość (*depth*)]



Zależnie od kontekstu, dogodnie będzie interpretować obraz cyfrowy jako:

- macierz, lub
- funkcję f zdefiniowaną na kartezjańskim układzie współrzędnych (tzw. funkcja charakterystyczna)
 - dyskretna w dziedzinie i przeciwdziedzinie,
 - stanowi pewne przybliżenie obrazu rzeczywistego (fizycznego),
- Najczęściej:
 - $f : \mathbb{N} \times \mathbb{N} \rightarrow [0, 1] \ (\mathbb{R})$ lub
 - $f : \mathbb{N} \times \mathbb{N} \rightarrow [0, 255] \ (\mathbb{N})$

Zamienne reprezentacje: macierz i funkcja



- 2D (planarne): $f(x, y)$
 - Pojedynczy element dziedziny: *pixel*, *pel* = picture element, *piksel*, *punkt obrazu*
- 3D
 - $f(x, y, z)$: Pojedynczy element dziedziny: *voxel* = volume element,
 - $f(x, y, t)$: Sekwencje wideo
- 4D, np. $f(x, y, z, t)$, np. sekwencje z fMRI (czynnościowy magnetyczny rezonans jądrowy)

Uwaga: W przeciwieństwie do ścisłego ujęcia matematycznego, pojedynczy punkt dziedziny f (np. (x, y)) ma niezerowe rozmiary.

Liczba kanałów:

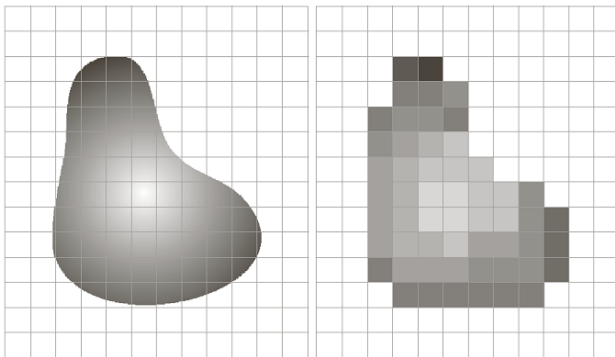
- 1: wartość skalarna:
 - przeciwdziedzina binarna: obraz *czarno-biały* (*bitmapa*)
 - przeciwdziedzina wielowartościowa: obraz *monochromatyczny* (*greyscale*).
- >1 : wektor:
 - obraz wielokanałowy; najczęściej kolorowe (np. RGB),
 - obraz wielomodalny (np. tomograf MRI, obrazy satelitarne),

Alternatywnie przedział wartości charakteryzuje się czasami liczbą bitów przeznaczonych na przechowanie informacji o pojedynczym pikselu (*depth* lub bpp, ang. *bits per pixel*):

- 1bpp: obrazy binarne, bitmapy
- 8bpp: obrazy monochromatyczne (jednokanałowe)
 - czasami także 12 bpp lub 14 bpp
- 24bpp: obrazy barwne (RGB, $3 \times 8\text{bpp} = 24\text{bpp}$)
- 32 bpp i więcej: dodatkowe kanały

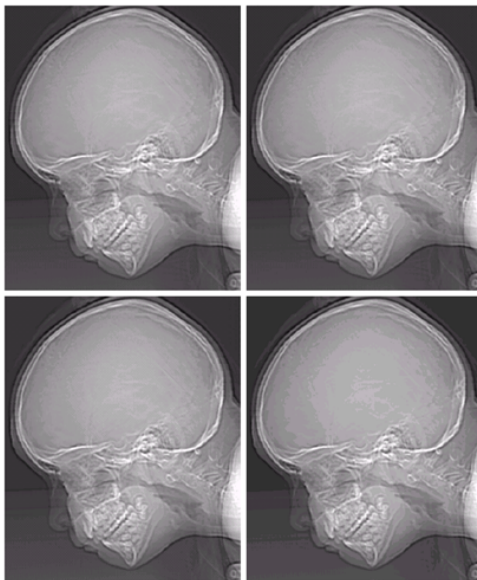
Dwa procesy dyskretyzacji:

- Dyskretyzacja przestrzenna
- Dyskretyzacja wartości sygnału



a b

FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.



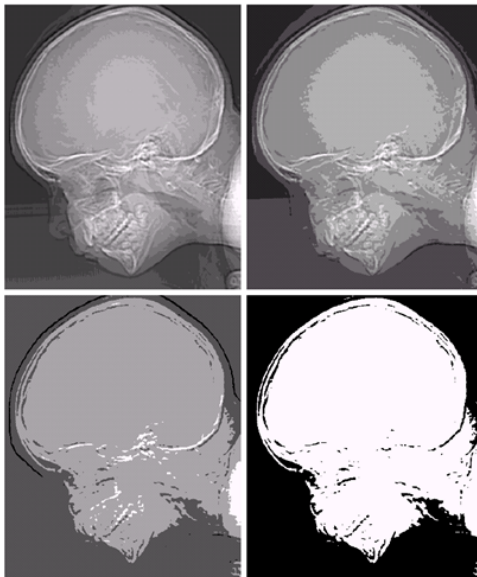
a b
c d

FIGURE 2.21

(a) 452×374 , 256-level image. (b)–(d) Image displayed in 128, 64, and 32 gray levels, while keeping the spatial resolution constant.

c f
g h

FIGURE 2.21
(Continued)
(e)–(h) Image
displayed in 16, 8,
4, and 2 gray
levels. (Original
courtesy of
Dr. David
R. Pickens,
Department of
Radiology &
Radiological
Sciences,
Vanderbilt
University
Medical Center.)



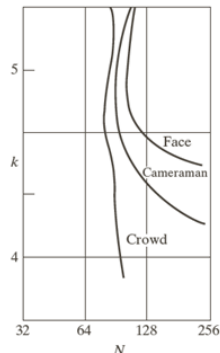
Jak obiektywne parametry obrazu (liczba bitów na piksel k i rozmiar/rozdzielczość N) przekładają się na postrzegana jakość obrazu?

- Krzywe izopreferencji
- Konkluzja: obrazy o większej liczbie detali implikują niższe wymagania odnośnie liczby poziomów szarości.



a b c

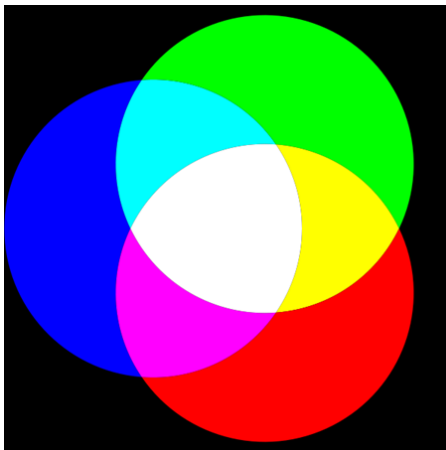
FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)



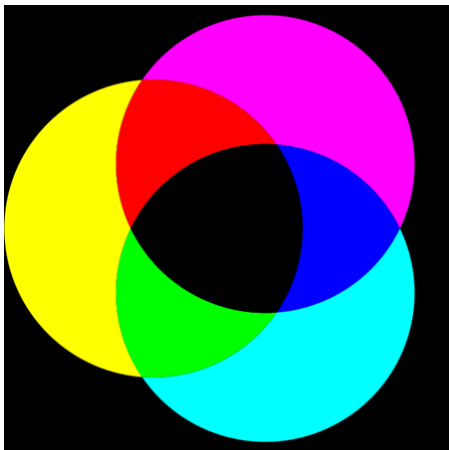
- Co to jest kolor?
- Przestrzenie reprezentacji barw: RGB, HSI, HSV, CMYK,...
- Motywacje dla utrzymywania wielu przestrzeni reprezentacji barw:
 - zgodność z percepcją wzrokową człowieka,
 - zgodność z pewnymi procesami/standardami technologicznymi,
 - zróżnicowanie reprezentacji barw przydatne w przetwarzaniu i wnioskowaniu.

Przestrzeń RGB

Model addytywny. Powszechnie stosowany w urządzeniach projekcyjnych (wrażenie koloru powstaje w efekcie projekcji/emisji światła).



Model subtraktywny. Wykorzystywany głównie w druku (wrażenie koloru powstaje w efekcie odbicia światła).



Konwersja z RGB do CMY

- $C = 1-R$, $M = 1-G$, $Y = 1-B$

Wprowadzenie kanału K (K – *key* (czerni) redukuje ilość niezbędnego barwnika.
Uwaga: konwersja nieodwracalna (stratna).

if $\min\{C', M', Y'\} = 1$

then

$$t_{CMYK} = \{0, 0, 0, 1\}$$

else

$$K = \min\{C', M', Y'\}$$

$$t_{CMYK} = \left\{ \frac{C' - K}{1 - K}, \frac{M' - K}{1 - K}, \frac{Y' - K}{1 - K}, K \right\}$$



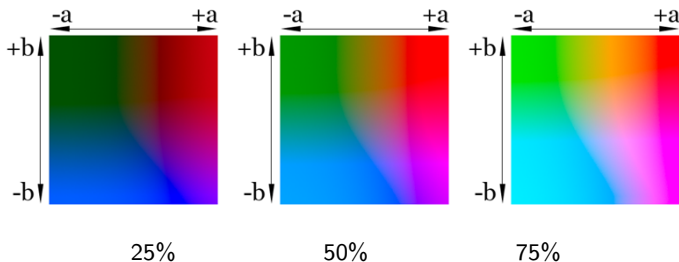
- Przestrzenie RGB i CMY/CMYK nie są przestrzeniami absolutnymi, bo odwołują się do nieustandaryzowanych barw podstawowych – rezultat na monitorze/w wydruku może być różny.
- Inaczej mówiąc: co to jest R? (w sensie fizycznym)

Absolutna przestrzeń reprezentacji barw: L^*a^*b (LAB, CIELAB). Trzy kanały:

- L – jasność (lightness),
- a – pozycja pomiędzy M a G
- b – pozycja między Y a B
- a i b abstrahują od L

Założenie: takie same subiektywne różnice w percepcji barw powinny przekładać się na zbliżone różnice wartości składowych (liniowość percepcyjna, perceptual linearity)

Wizualizacja: L^*a^*b – Jasność:

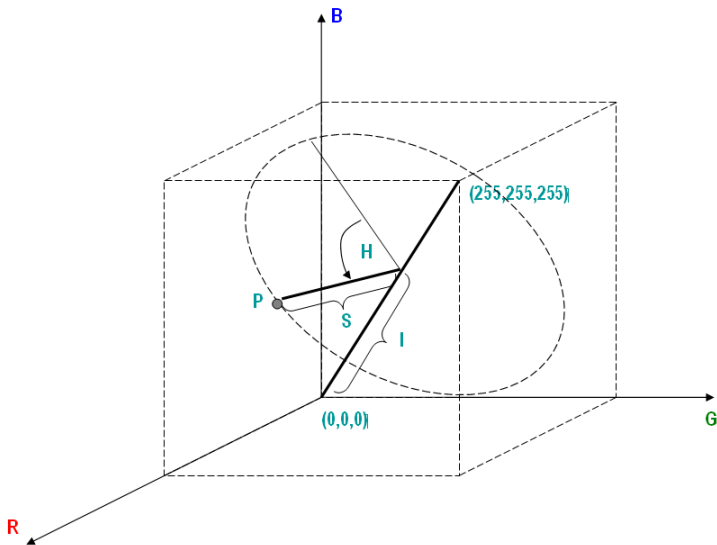


Motywacja: lepsza zgodność z cechami obrazu postrzeganymi przez człowieka (w odróżnieniu od np. RGB)

- *Hue* (barwa)
- *Saturation* (nasycenie, wysycenie)
- *Intensity* [Value] (jasność)

Bardzo popularna w metodach IP i PR, także w grafice komputerowej (tzw. *color pickers*).

Zbliżone przestrzenie: HSV, HSL, HSB



Wymagane przejście ze współrzędnych kartezjańskich do biegunowych:

$$\alpha = \frac{1}{2}(2R - G - B)$$

$$\beta = \frac{\sqrt{3}}{2}(G - B)$$

$$H(\beta, \alpha) = \text{atan2}(\beta, \alpha)$$

W praktyce używa się uproszczonych wzorów które nie angażują trygonometrii:

$$H' = \begin{cases} \text{undefined}, & \text{if } C = 0 \\ \frac{G-B}{C} \bmod 6, & \text{if } M = R \\ \frac{B-R}{C} + 2, & \text{if } M = G \\ \frac{R-G}{C} + 4, & \text{if } M = B \end{cases}$$

$$H = 60^\circ \times H'$$

gdzie C to *chroma*, tj.

$$C = \max(R, G, B) - \min(R, G, B)$$

Wzory przybliżone generują wartości Hue różniące się od wartości dokładnych o nie więcej niż ok. 1° .



Przekształcenia mogą prowadzić do częściowej utraty informacji z powodu:

- dyskretyzacji,
- projekcyjnego (zamiast bijekcyjnego) mapowania (np. RGB- $\rightarrow$ HSI dla obrazów mało intensywnych, np. słabo wybarwionych preparatów mikroskopowych)

Najczęściej:

- wiersz po wierszu, kolumna po kolumnie

Uwagi techniczne

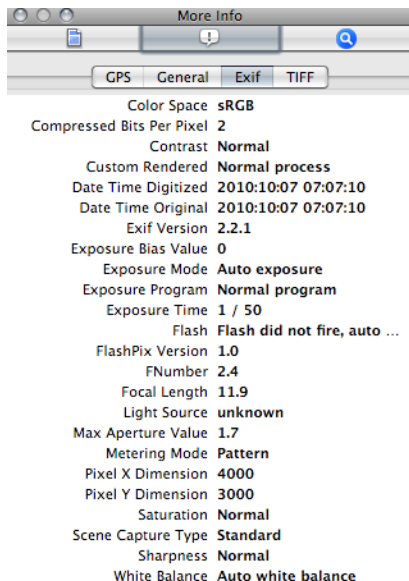
- preferowane przeglądanie zgodne z porządkiem pikseli (ze względu na pamięć podręczną cache),
- w przypadku obrazów wielokanałowych: kanały przeplatane; niekiedy dla przyspieszenia operacji warto odseparować od siebie kanały w pamięci,
- reprezentacja w pamięci często de facto 32 bitowa
- w przypadku nieparzystej szerokości obrazu, wiele bibliotek rezerwuje o jeden bajt na linię obrazu więcej, tak aby każda linia zaczynała się od parzystego adresu.

Inne parametry obrazu opisujące sposób jego pozyskiwania:

- rozdzielczość (DPI),
- rozdzielczość na poszczególnych osiach,
- długość linii w bajtach (np. w formacie BMP)

Zobacz np. Standard EXIF (Exchangeable Image File Format)

- Format przewidujący pewien zestaw metatagów dodawanych do istniejących formatów plików graficznych, np. czas ekspozycji, ustawienia równowagi bieli.
- Stosowany m.in. w JPEG, TIFF, RIFF, nie używany w PNG, JPEG2000, GIF.



Stratne:

- JPEG (Joint Photographic Experts Group) – zasadniczo metoda kompresji a nie format pliku, nazwa formatu to JFIF (JPEG File Interchange Format). Sukcesywne kompresje i dekompresje degradują jakość obrazu.

Bezstratne:

- GIF (Graphics Interchange Format) – 256 kolorów + multiframe, opatentowany (z powodu LZW), obecnie patenty wygasły w większości krajów
- PNG (Portable Network Graphics) – następca GIF, open source
- BMP – Windows bitmap
- RAW – raczej metaformat, nieustandaryzowany

Mieszane:

- TIFF/TIF (Tagged Image File Format) – szeroki standard, może angażować LZW
- JPEG2000 – lepsze właściwości niż JPEG, ale wyższy koszt obliczeniowy kompresji i dekompresji; zastosowaniach profesjonalne (np. niektóre kina)

Inne: PCX, PPM, PGM

Przykład artefaktów wynikających z kompresji obrazu

