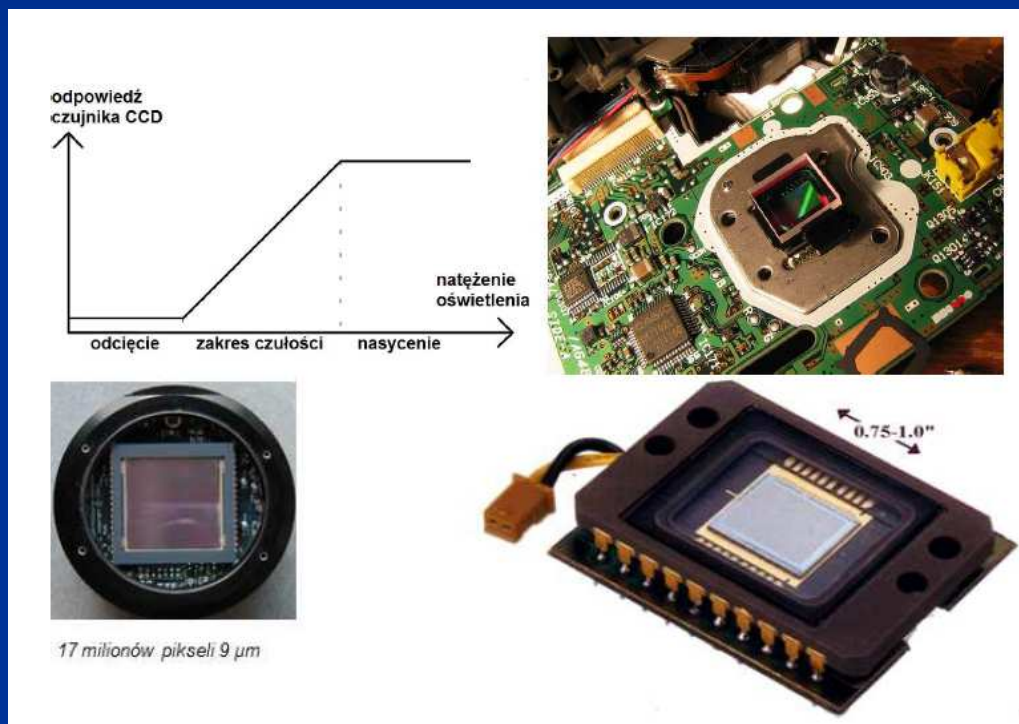


Akwizycja obrazów. Zagadnienia wstępne



Wykorzystane materiały:

R. Tadeusiewicz, P. Korohoda, Komputerowa analiza i przetwarzanie obrazów, Wyd. FPT, Kraków, 1997

A. Przelaskowski, Techniki Multimedialne, skrypt, Warszawa, 2011

Obrazy źródłowe, postrzegane przez ludzki zmysł wzroku, opisywane są generalnie **funkcją jasności**

$$C(x, y, z, t, \lambda),$$

reprezentującą przestrzenny rozkład energii promieniowania widzialnego

- x, y, z – współrzędne przestrzenne
- t – czas trwania obserwacji
- λ – długość fali optycznej

Dziedzina tej funkcji ograniczona jest **polem rejestracji** (tj. umownym polem widzenia obserwatora), skończonym **czasem obserwacji** oraz **zdolnością percepcji** określonego zakresu promieniowania.

Odpowiedź w zakresie intensywności i koloru standardowego oka ludzkiego na funkcję jasności obrazu:

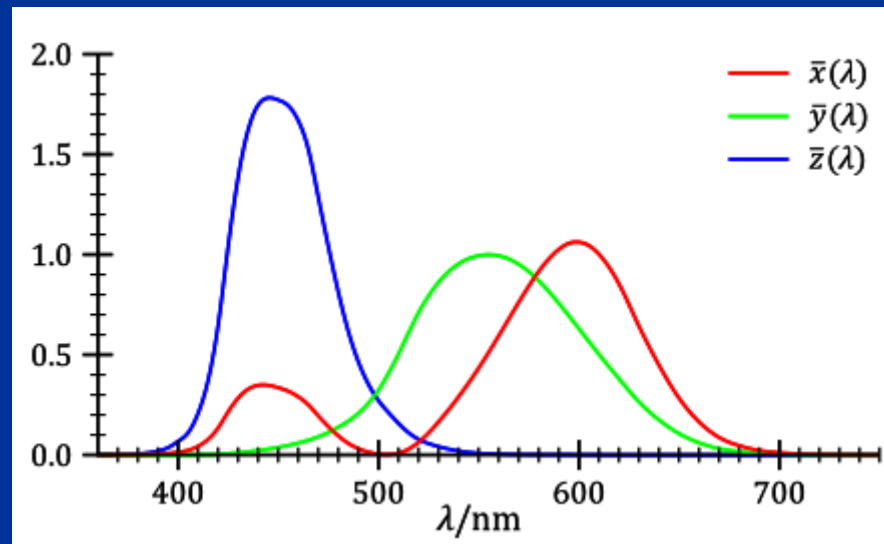
- **luminancja** – zróżnicowanie jasności (intensywności)
- **chrominancja** – zróżnicowanie odcienia oraz nasycenia kolorów (barw)

Funkcja pola obrazu:

$$\mathcal{F}_i(x, y, z, t) = \int_0^{\infty} \mathcal{C}(x, y, z, t, \lambda) S_i(\lambda) d\lambda, \quad i = 1, 2, \dots$$

- $S_i(\lambda)$ – **widmowa odpowiedź czujnika i** , nadająca charakter odbieranemu obrazowi

- **Czopki** – światłoczułe receptory siatkówki
- W ludzkim oku występują **trzy rodzaje czopków**, o **odmiennej charakterystyce widmowej**, z dominującymi barwami:
 - **czterwoną**
 - **zieloną**
 - **niebieską**



■ Czopki

- Szybka rejestracja obrazów
- Duża ostrość rejestrowanych obrazów (wysoka rozdzielczość)
- Niewielka czułość
- Wrażliwość jedynie na światło bezpośrednie

■ Pręciki – receptory pomocnicze

- działają przy słabym oświetleniu
- rejestrują światło rozproszone
- nie rejestrują kolorów, ale kształty i ruch
- ograniczona ostrość widzenia

W świetle systemu recepcji obrazu przez ludzkie oko, naturalne jest wykorzystanie do rejestracji obrazów **czujników wrażliwych na trzy podzakresy** długości fal promieniowania widzialnego:

- **Red - czerwony** ($i = 1$)
- **Green - zielony** ($i = 2$)
- **Blue - niebieski** ($i = 3$)

Otrzymujemy w wyniku funkcję rozkładu kolorów **w przestrzeni RGB**:

$$\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3) = (\mathcal{F}_R, \mathcal{F}_G, \mathcal{F}_B)$$

- Obrazy opisywane są zwykle w **przestrzeni RGB**
 - przede wszystkim ze względu na budowę ludzkiego oka oraz działanie typowych urządzeń służących do rejestracji i prezentacji obrazów (kamery, aparaty fotograficzne, monitory, telewizory, skanery)
- Jednak model percepcji obrazów naturalnych opisujący podobieństwo kolorów, a także wrażliwość na jakość odbieranej informacji wygodniej jest tworzyć w **innych przestrzeniach barw**
- Także **przetwarzanie obrazów** wygodniej jest zwykle dokonywać w innej przestrzeni barw.

- Modele sprzętowe (ważne przy konstrukcji urządzeń):
 - **RGB** – monitory
 - **CMY(K)** – drukarki
 - **HSV** – procesory grafiki komputerowej
- Modele percepcji (pozwalają rozróżnić i precyzyjnie definiować barwy widoczne i wyobrażane):
 - **XYZ**
 - **YIQ**
 - **YUV**
 - **YCrCb**
 - wykorzystywane m.in. w kodowaniu JPEG, MPEG

- Opisany macierzowo **przestrzenny rozkład intensywności** – jasności, lub zestaw kilku komponentów barwowych o określonej dynamice oraz lokalizacji w przestrzeni i czasie
- **Obrazy cyfrowe** rejestrowane są w różnorodnych systemach akwizycji z **bezpośrednią konwersją** funkcji jasności C na postać cyfrową (np. za pomocą matryc CCD) lub też z **konwersją pośrednią** (np. skanowanie klasycznego zdjęcia analogowego)

■ próbkowanie

- dyskretyzacja chwil czasowych lub współrzędnych przestrzennych, w których rejestrowany jest sygnał)

■ kwantyzacja

- dyskretyzacja zbioru wartości rejestrowanego sygnału

■ kodowanie

- ustalenia postaci binarnej reprezentacji kolejnych próbek sygnału

- funkcja jasności obrazu jest rejestrowana w **skończonym przedziale czasowym** (tj. realnym czasie akwizycji całego obrazu, po którym uśrednione zostają wartości pikseli – zwykle dąży się do minimalizacji tego czasu)
- funkcja jasności jest jednocześnie próbkowana **wzdłuż obu współrzędnych przestrzennych**,
- wartości tej funkcji są kwantowane **w każdym punkcie** dyskretnego pola obrazu,
- każdej skwantowanej wartości przypisana jest w sposób jednoznaczny odpowiednia **reprezentacja bitowa**,

to otrzymujemy obraz cyfrowy reprezentujący **rzeczywiste, ciągłe obrazy naturalne**

Sekwencja wizyjna

WIDZENIE
KOMPUTEROWE

Proces kwantyzacji, próbkowania i kodowania, realizowany w określonym systemie akwizycji i powtarzany w kolejnych ustalonych chwilach czasowych, dostarcza sekwencji obrazów naturalnych nazywanej **sekwencją wizyjną**.



- **Reprezentacja pojedynczego obrazu** powstaje na bazie funkcji pola obrazu

$$F(x, y, z) = F_{i=u}(x, y, z, t = t_u),$$

zarejestrowanej w danej chwili czasowej t_u , przy ustalonej charakterystyce czujników, określonej indeksem u .

- Niech $f(x, y, z) = F(x, y, z)$ będzie funkcją jasności obrazu (czyli rozkładem luminancji) zdefiniowaną przestrzennie – trójwymiarowo.
- W przypadku płaskim mamy $f(x, y)$.

- Rejestrowana cyfrowo funkcja jasności obrazu podlega **dyskretyzacji przestrzennej** względem współrzędnych x i y (w systemach akwizycji 3D – również z)
- Ponieważ w zdecydowanej większości systemów rejestracji pole obrazu jest prostokątne (lub prostopadłościennie w systemach 3D), wykorzystywany model obrazu bazuje na **prostokątnym (prostopadłościennym) polu obrazu**
- Jeśli aktywna lub użyteczna powierzchnia detektora ma inny kształt, zwykle wpisywana jest w referencyjną strukturę prostokątną, stanowiącą układ odniesienia względem bardziej typowych urządzeń i systemów

Obraz cyfrowy – formalnie (1)

Formalnie obraz cyfrowy jest zbiorem pikseli o współrzędnych dziedziny (pola) obrazu, z przypisanymi wartościami funkcji jasności. Niech $f_{\mathcal{O}} : \mathcal{P} \rightarrow \mathcal{F}$ oznacza funkcję jasności obrazu \mathcal{O} , która przypisuje dyskretnemu polu $\Omega_f \in \mathbb{Z}^2$ tego obrazu dyskretny zbiór wartości $\mathcal{F} \in \mathbb{R}$. Zakładamy, że dziedzina obrazów jest powłoką (otoczką) wypukłą skończonego zbioru rejestrowanych pikseli $[\Omega_f]$, czyli przyjmuje postać wielokąta, najczęściej prostokąta.

Dziedziną przekształcenia jest wtedy zbiór wszystkich punktów pola (pikseli)

$$\Omega_f = \{(k, l) \in \mathbb{Z}^2 : k_{\min} \leq k \leq k_{\max}, l_{\min} \leq l \leq l_{\max}\}$$

gdzie $k_{\min} = \min_{(k,l) \in \mathcal{O}} \{k\}$, $k_{\max} = \max_{(k,l) \in \mathcal{O}} \{k\}$ oraz analogicznie l_{\min} , l_{\max} ograniczające odpowiednio szerokość i wysokość obrazu. Stąd szerokość $K = k_{\max} - k_{\min} + 1$, a wysokość: $L = l_{\max} - l_{\min} + 1$. Zwykle, ze względu na konwencję zapisu liczb w rejestrach i komórkach pamięci cyfrowej, kładziemy $k_{\min} = l_{\min} = 0$, wtedy $K = k_{\max} + 1$, $L = l_{\max} + 1$.

Obraz cyfrowy – formalnie (2)

Każda z wartości $f(k, l)$ – elementów uporządkowanego zbioru wartości pikseli $\mathbf{f} = \{f(k, l) : k = 0, \dots, K-1, l = 0, \dots, L-1, \}$ należy do uporządkowanego rosnąco zbioru (alfabetu) wartości możliwych, tj. $f(k, l) \in A_f = \{a_0, \dots, a_{M-1}\}$. Zakładając kod dwójkowy jako naturalną regułę reprezentacji rejestrowanych obrazów cyfrowych otrzymujemy $k = \lceil \log_2 M \rceil$. bitowy obraz cyfrowy $\mathcal{O} \in \{0, 1, \dots, 2^k - 1\}^{\Omega_f}$. W przypadku obrazów o q składowych barwowych mamy $f_{\mathcal{O}} : \mathbb{R}^2 \rightarrow \mathbb{R}^q, \mathcal{P} \rightarrow \{\mathcal{F}_i\}_{i=1}^q$. Wtedy alfabet wartości pikseli (f_1, f_2, \dots, f_q) jest iloczynem kartezjańskim zbiorów wartości poszczególnych komponentów obrazu:

$$A_{f_1, \dots, q} = A_{f_1} \times \dots \times A_{f_q}.$$

Ze względu na postać dyskretnego zbioru wartości pikseli oraz liczbę komponentów, obrazy cyfrowe można podzielić na:

- **dwupoziomowe** (*bilevel*): $M = 2$ i $q = 1$
 - Czarno-białe faksy, skanowane dokumenty tekstowe, szkice graficzne
- **wielopoziomowe** (*continuous-tone, multilevel*): $M > 2$
 - Wartości pikseli mogą być wyrażone w skali szarości ($q = 1$) – mamy wtedy obrazy **monochromatyczne** ze skalą szarości
 - Wartości pikseli mogą być wyrażone w przestrzeni wielokomponentowej ($q > 1$), które nazywamy obrazami **kolorowymi** (barwnymi)
 - Podstawowy zakres wartości dla komponentów obrazów wielopoziomowych, naturalnych wynosi 0–255 (obrazy bajtowe, $M = 256$), podczas gdy dla obrazów kolorowych mamy zwykle $q = 3$ (np. w przestrzeniach barw RGB, YUV, YCrCb) lub $q = 4$ (np. CMYK).