

Компьютерное Зрение
Лекция №1, весна 2021

Введение в цифровые изображения



Кафедра
технологий
проектирования
сложных
технических
систем

План лекции

- Теория цвета
- Представление изображения в компьютере
- Представление изображения в виде функции

План лекции

- **Теория цвета**
- Представление изображения в компьютере
- Представление изображения в виде функции

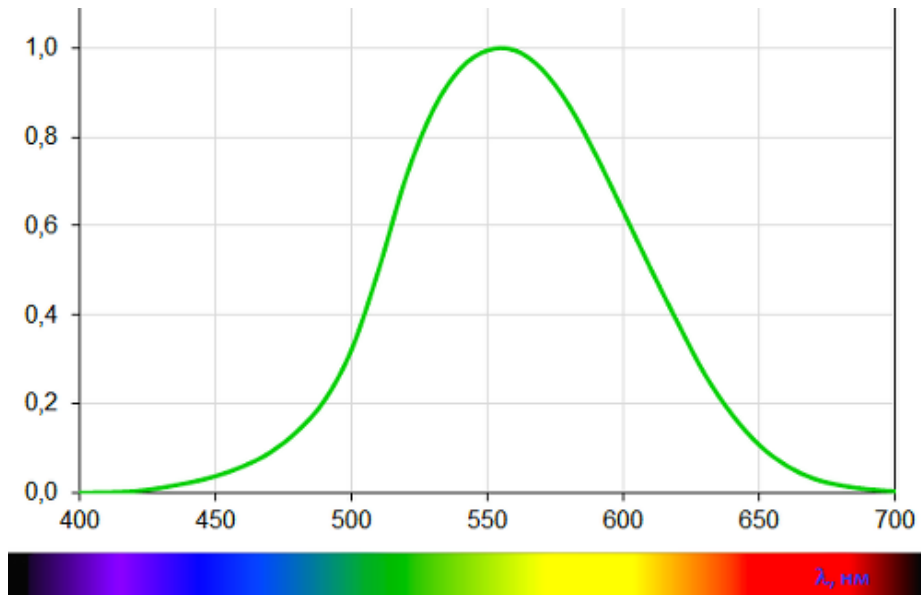
Формализация цвета

Попробуйте дать ответ на следующие вопросы:

- почему физика не может дать определения понятию цвета?
- какая из семи основных единиц измерений СИ основывается на свойствах зрительной системы человека?
- какого цветового тона нет в спектре?
- как удалось измерить ощущение цвета человеком ещё 90 лет назад?
- где используются цвета, которые не имеют яркости?

Цвет — это *ощущение*, которое получает человек при попадании ему в глаз световых лучей, а не физическое свойства объектов и света

Палочки и колбочки



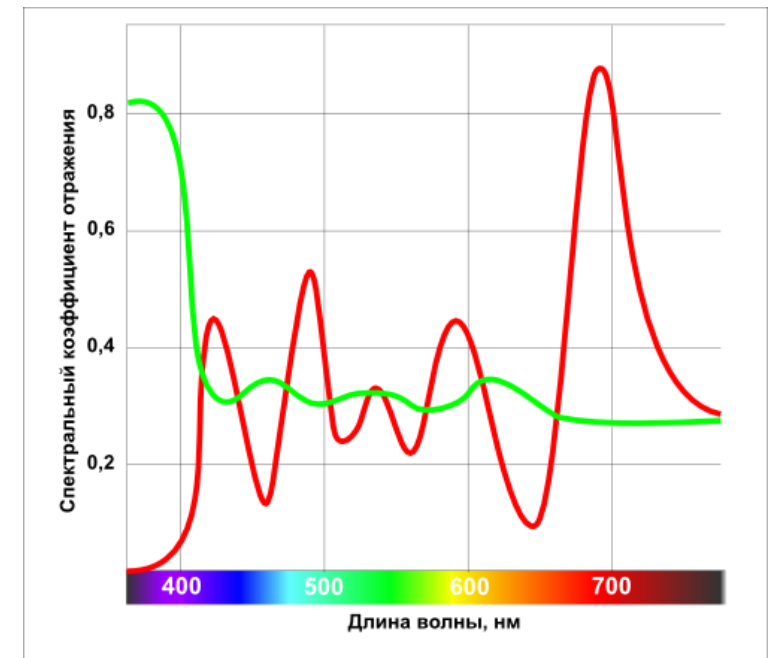
$$I_v = 683 \int_0^{\infty} \bar{y}(\lambda) \frac{dI_e(\lambda)}{d\lambda} d\lambda$$

I_v – сила света в канделах (кд),

I_e – интенсивность излучения в ваттах настерадиан (Вт/ср),

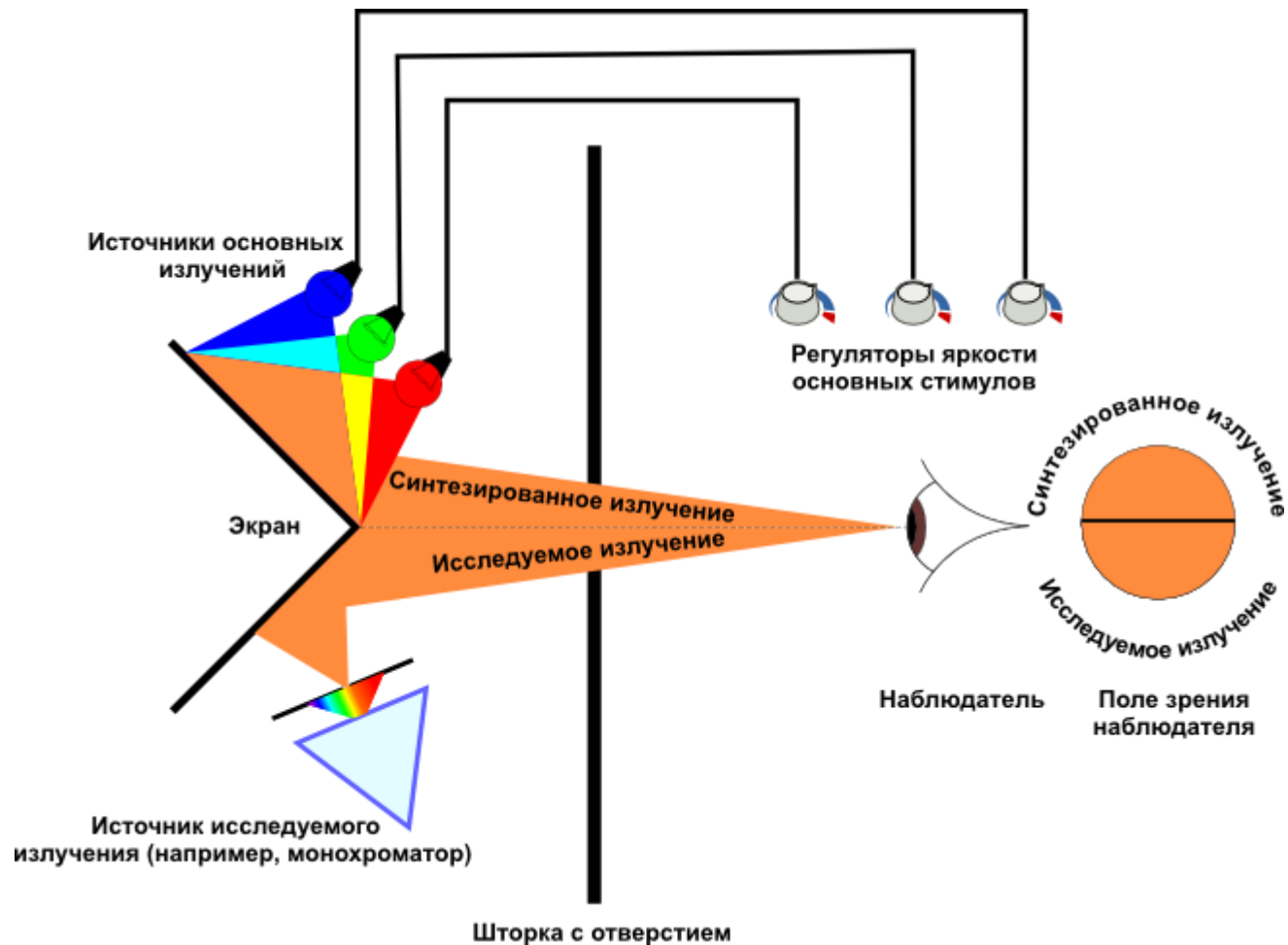
$\bar{y}(\lambda)$ – стандартная функция светимости.

Метамеры – разные спектры
одного цвета

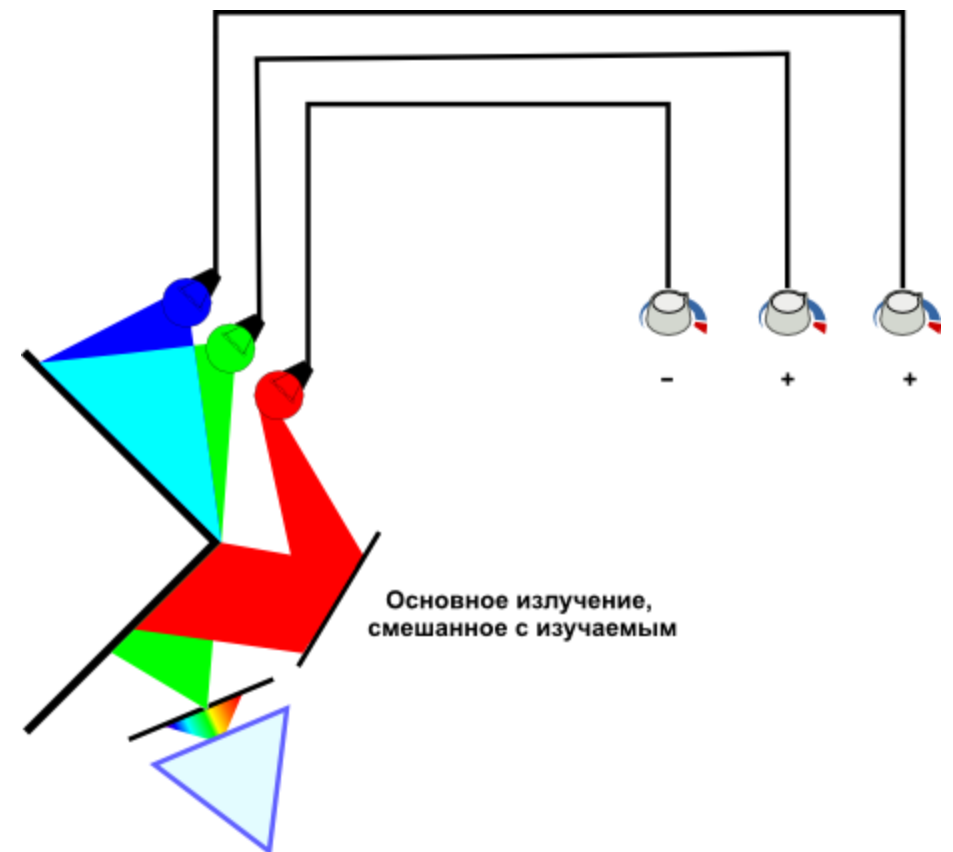


Система CIE RGB

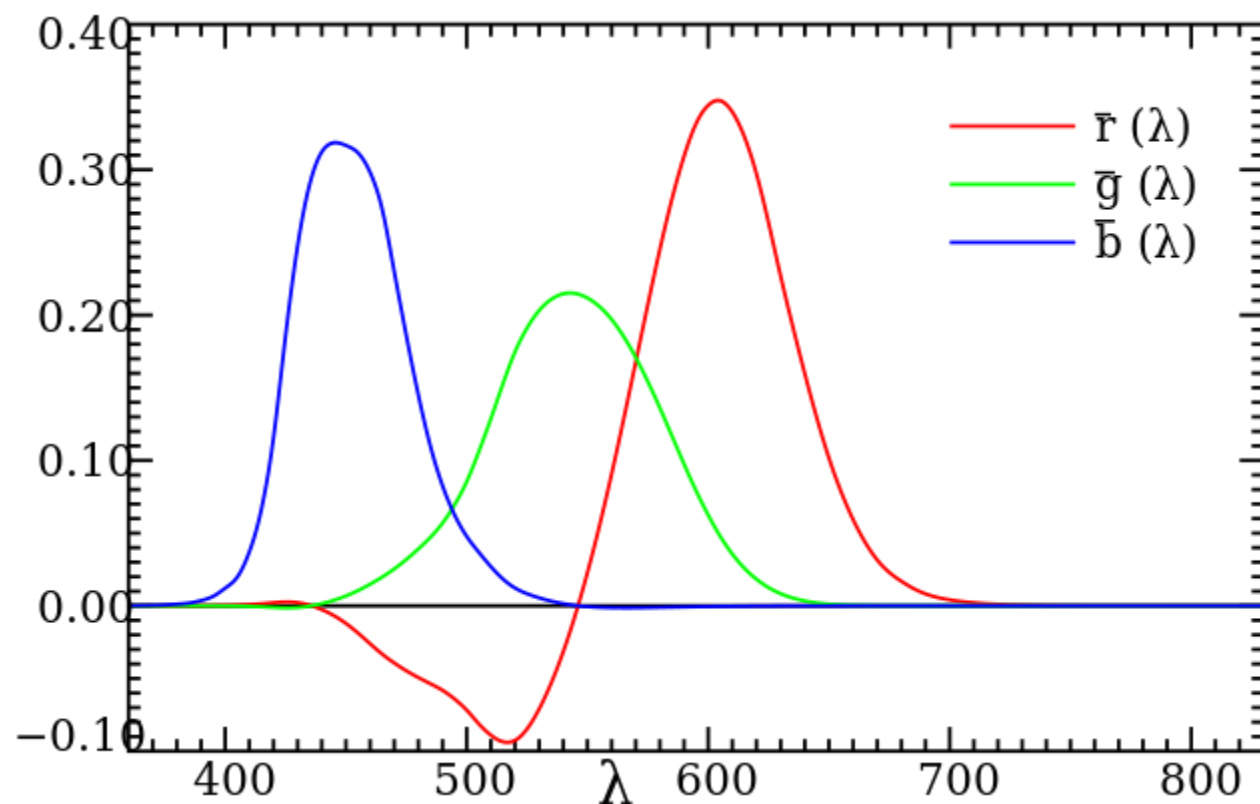
Первый вариант измерения



Второй вариант измерения



Система CIE RGB



Кривые нормализованы:

$$\int_0^{\infty} \bar{r}(\lambda) d\lambda = \int_0^{\infty} \bar{g}(\lambda) d\lambda = \int_0^{\infty} \bar{b}(\lambda) d\lambda$$

Масштаб $R:G:B = 1:4,5907:0,0601$

Значения трехцветного RGB для цвета со спектральным распределением мощности $S(\lambda)$:

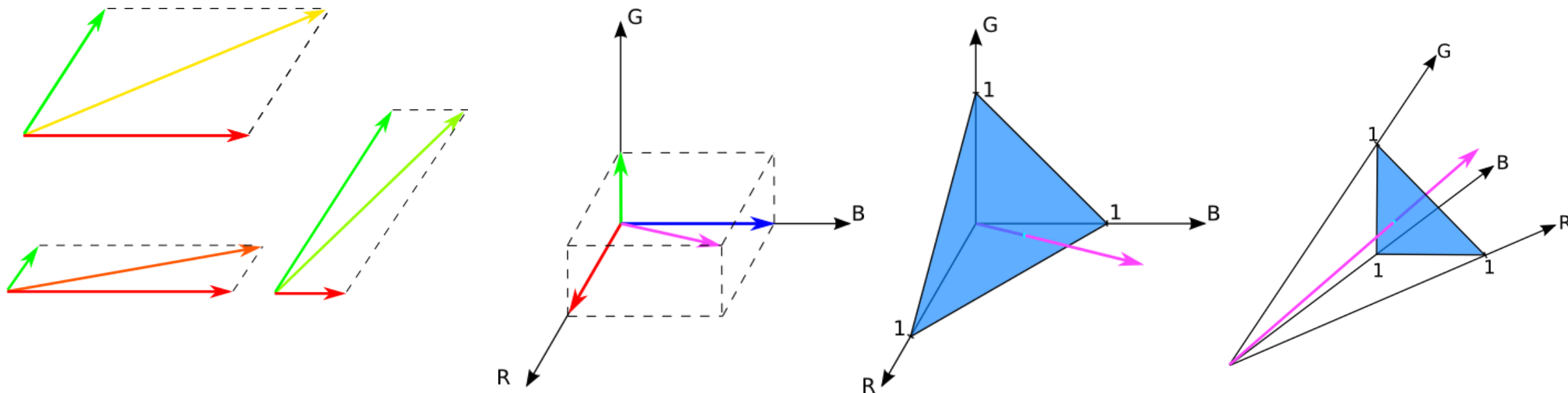
$$R = \int_0^{\infty} S(\lambda) \bar{r}(\lambda) d\lambda$$

$$G = \int_0^{\infty} S(\lambda) \bar{g}(\lambda) d\lambda$$

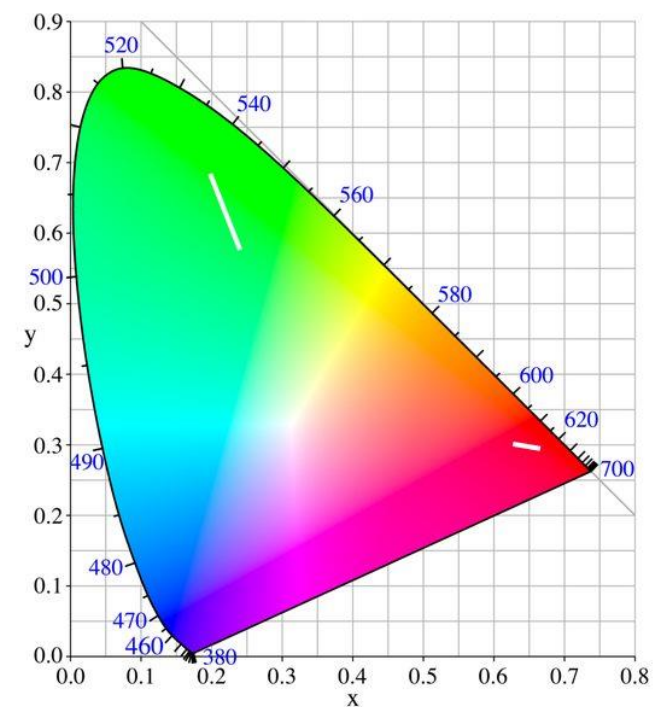
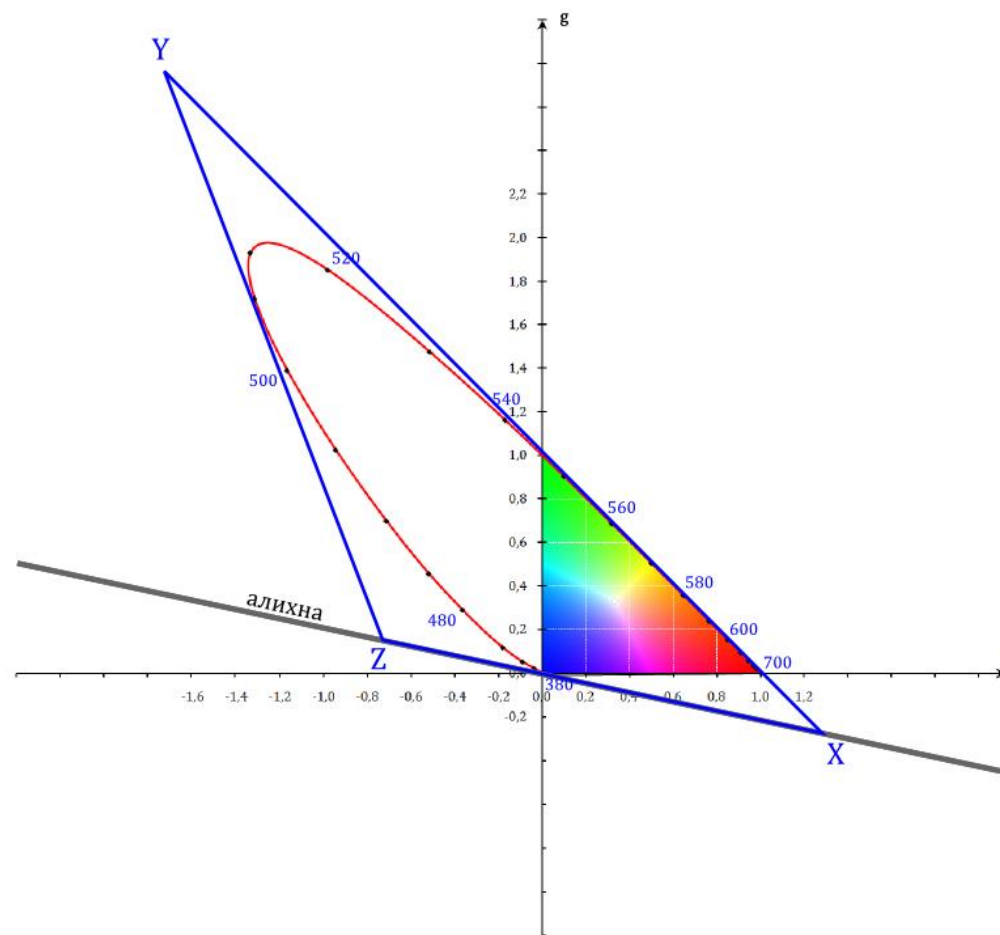
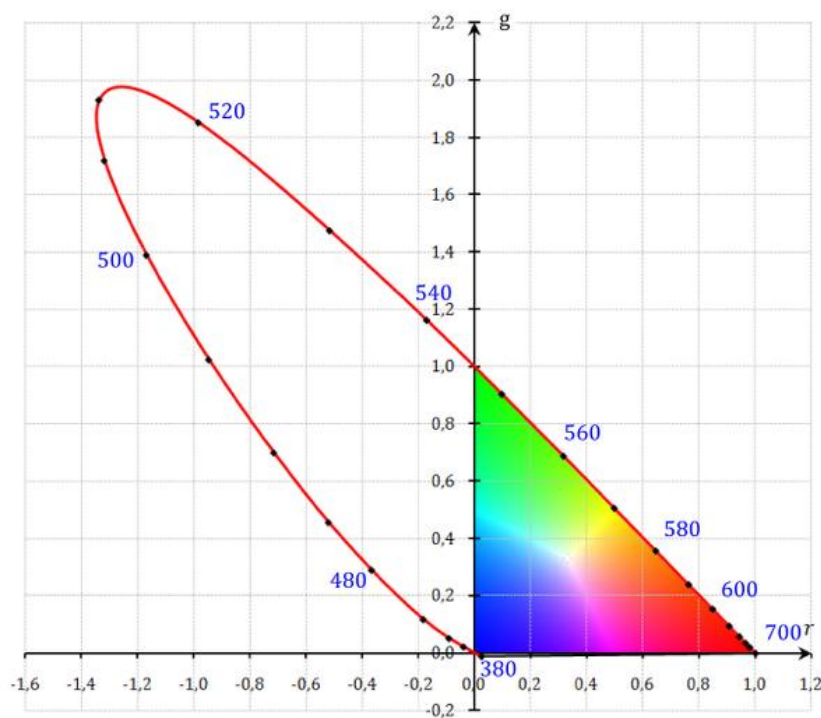
$$B = \int_0^{\infty} S(\lambda) \bar{b}(\lambda) d\lambda$$

Координатная система RGB

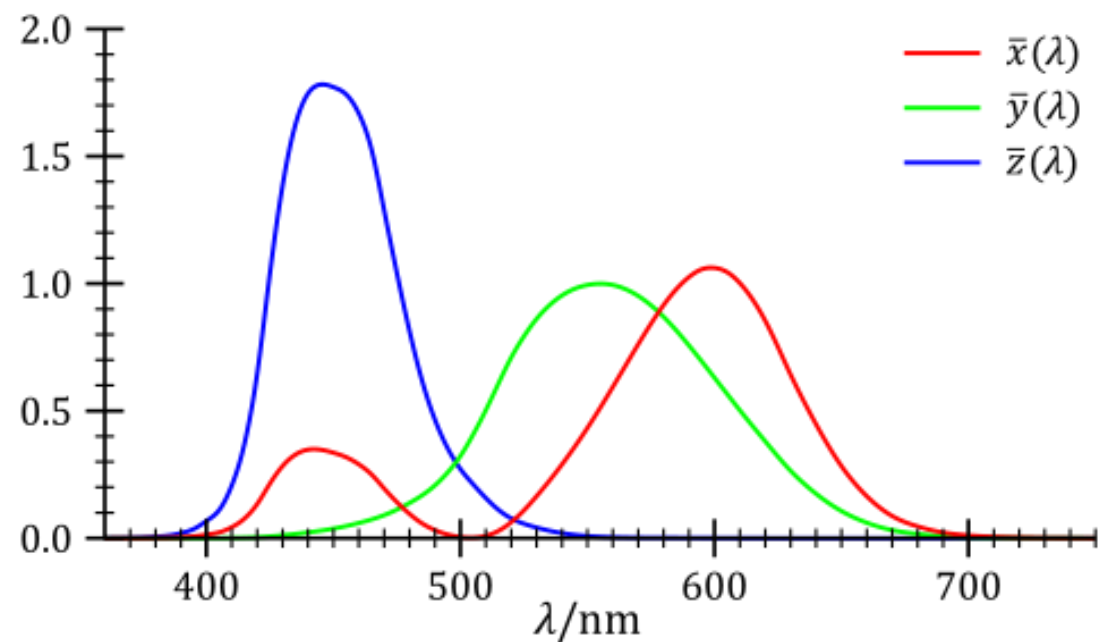
Закон Грассмана – эмпирическое наблюдение, что восприятие хроматической составляющей цвета описывается примерно линейным законом



Переход к системе CIE XYZ



Система CIE XYZ



Свойства системы:

- Y соответствует видимой части спектра
- X и Z описывают хроматическую компоненту
- Точки (1,0,0), (0,1,0), (0,0,1) мнимые базовые цвета
- X, Y, Z изменяются от 0 до ∞

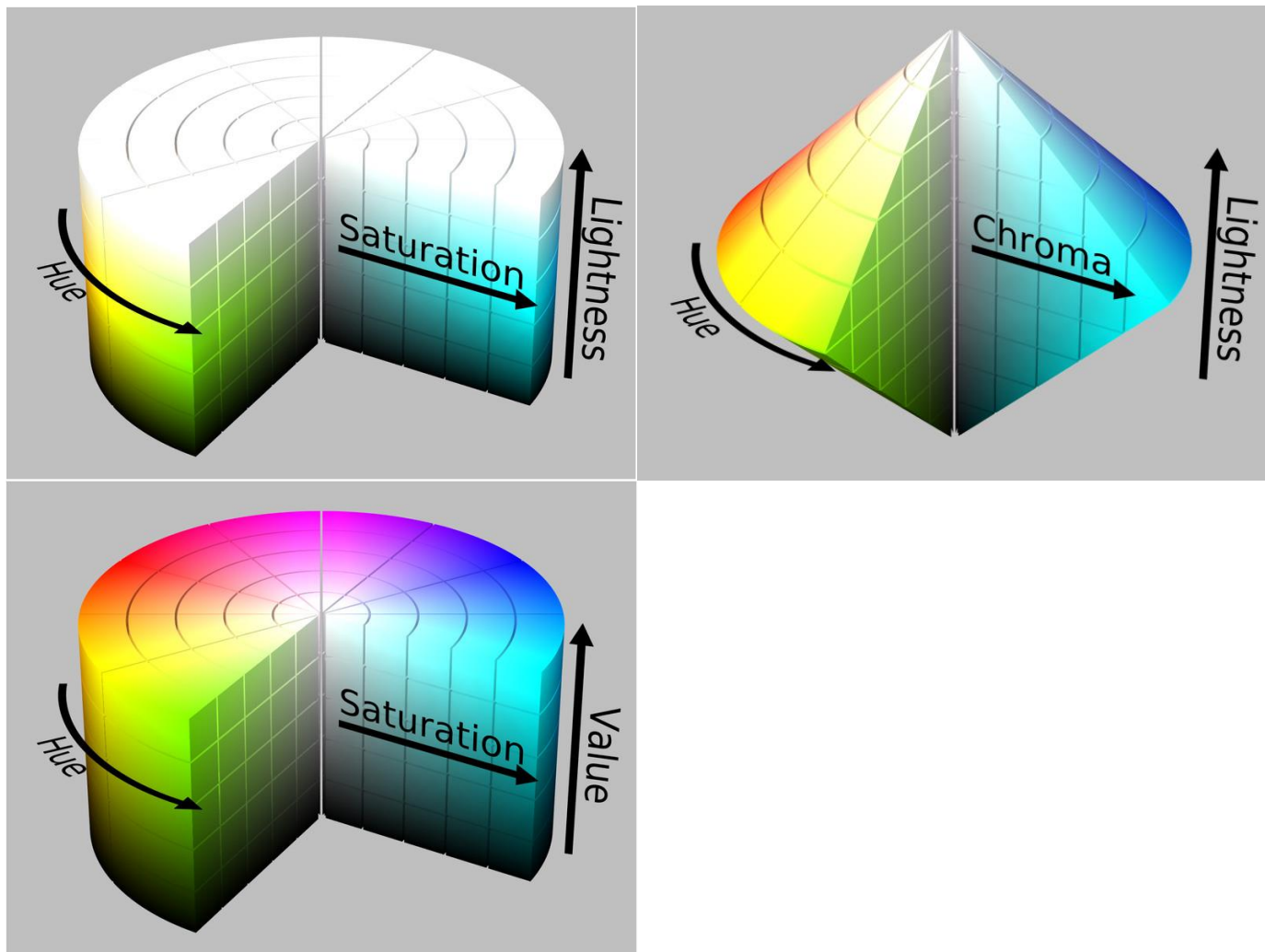
Значения трехцветного XYZ для цвета, где $I(\lambda)$ – спектральная плотность какой-либо энергетической фотометрической величины :

$$X = \int_{380}^{780} I(\lambda) \bar{x}(\lambda) d\lambda$$

$$Y = \int_{380}^{780} I(\lambda) \bar{y}(\lambda) d\lambda$$

$$Z = \int_{380}^{780} I(\lambda) \bar{z}(\lambda) d\lambda$$

Цветовое пространство HSV (HSB)/HSI/HSL



$$I = \frac{R + G + B}{3}$$

$$L = \frac{\max(R, G, B) + \min(R, G, B)}{2}$$

$$V = \max(R, G, B)$$

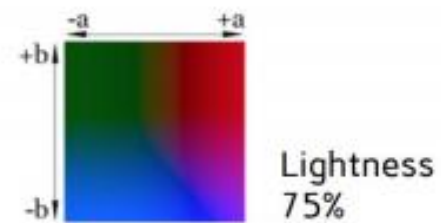
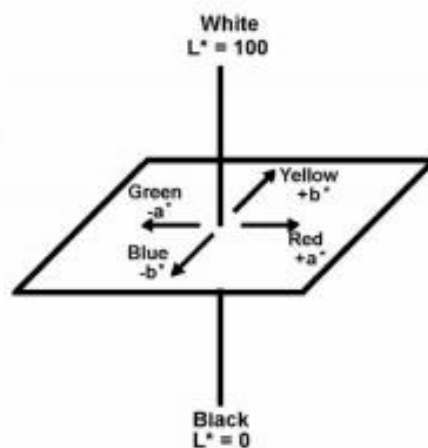
Цветовое пространство CIE L*a*b

$$L^* = 116 * h\left(\frac{Y}{Y_w}\right) - 16$$

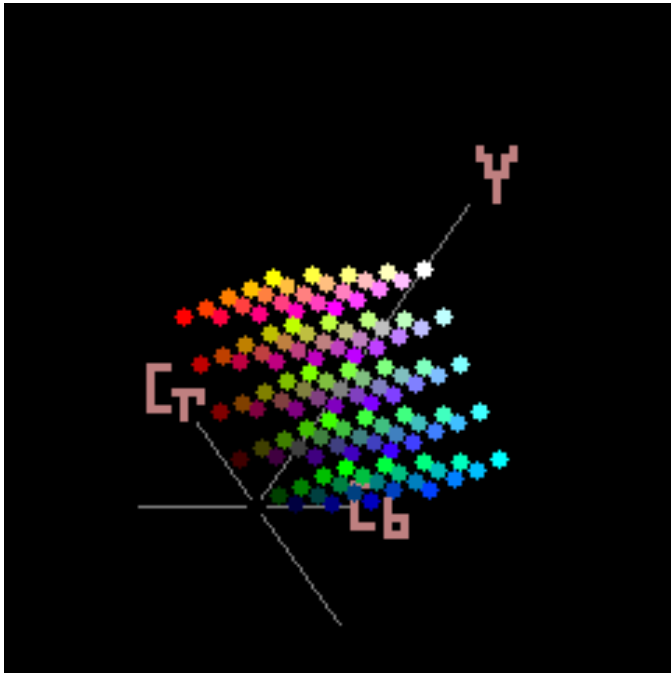
$$a^* = 500 \left[h\left(\frac{X}{X_w}\right) - h\left(\frac{Y}{Y_w}\right) \right]$$

$$b^* = 200 \left[h\left(\frac{Y}{Y_w}\right) - h\left(\frac{Z}{Z_w}\right) \right]$$

$$h(q) = \begin{cases} \sqrt[3]{q} & \text{if } q > 0.008856 \\ 7.787q + 16/116 & \text{otherwise} \end{cases}$$



Цветовое пространство YCbCr



Преобразование в пространство YCbCr:

$$Y = k_r R + (1 - k_g - k_b)G + k_b B$$

$$C_b = \frac{0,5}{1 - k_b} (B - Y)$$

$$C_r = \frac{0,5}{1 - k_r} (R - Y)$$

$$k_r + k_g + k_b = 1$$

k_r, k_g, k_b – весовые коэффициенты

Источники для погружения в теорию цвета

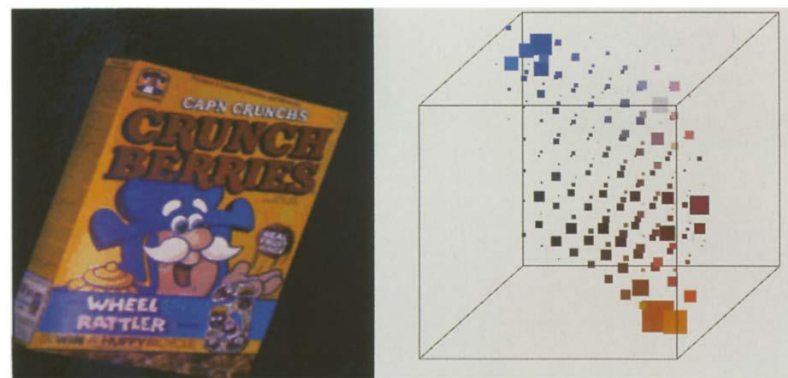
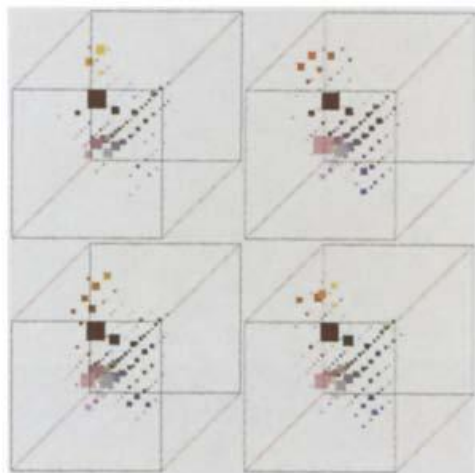
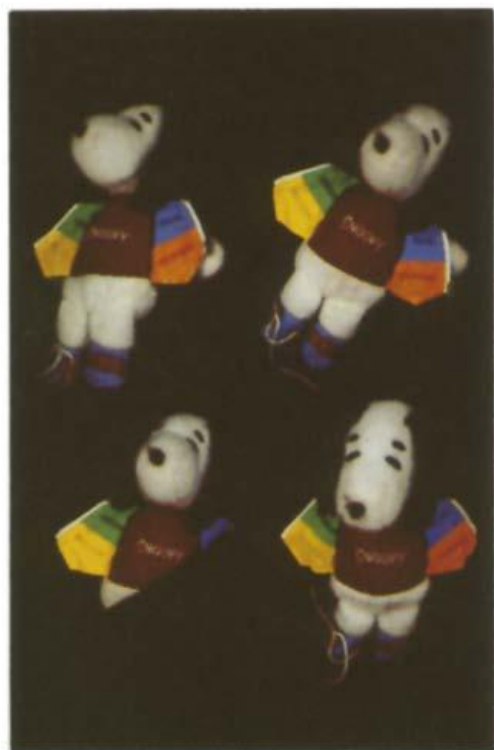
Лекция: [Как устроен цвет - Дмитрий Николаев, заведующий сектором зрительных систем ИППИ РАН](#)

Статья: [У цветового треугольника не два, а один угол](#)

Статья: [Как устроен формат JPEG](#)

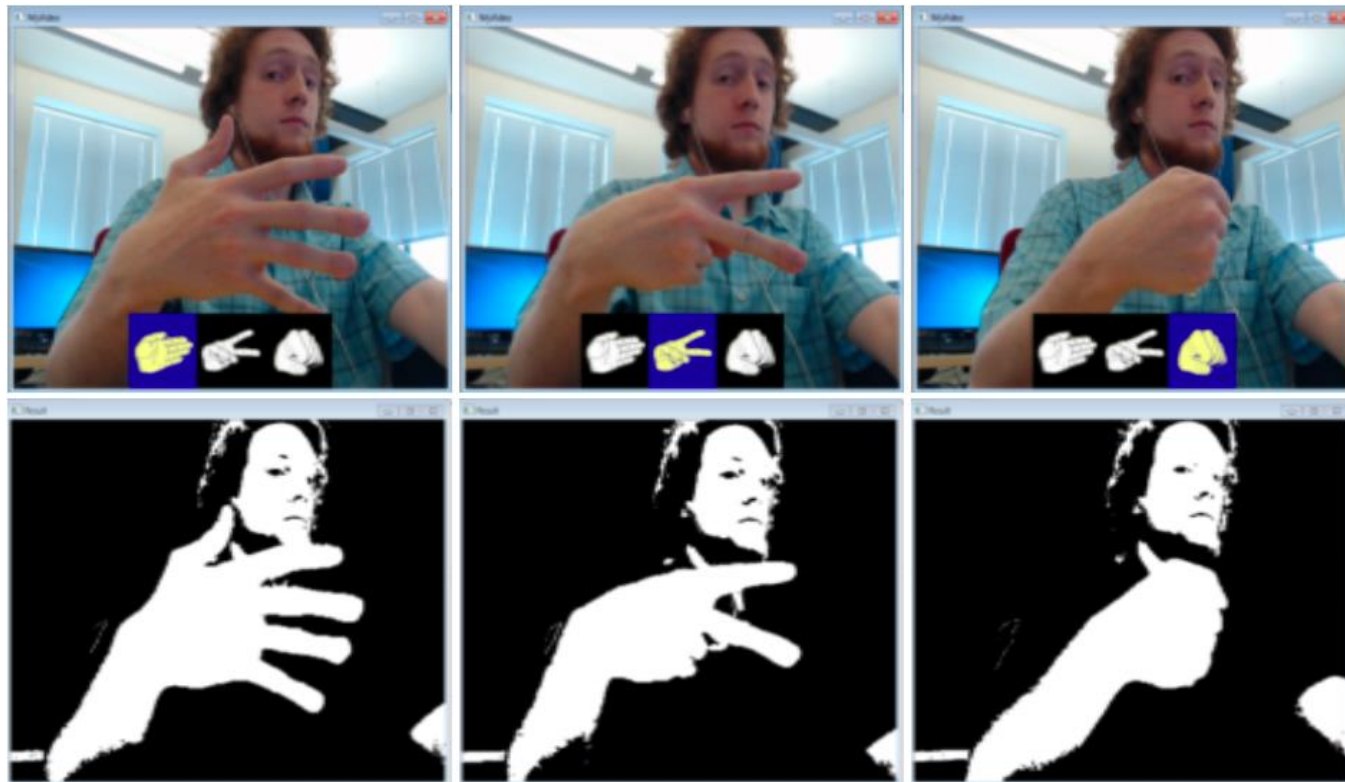
Применение цвета в задачах

Построение гистограмм по цветам для
индексированного поиска



Применение цвета в задачах

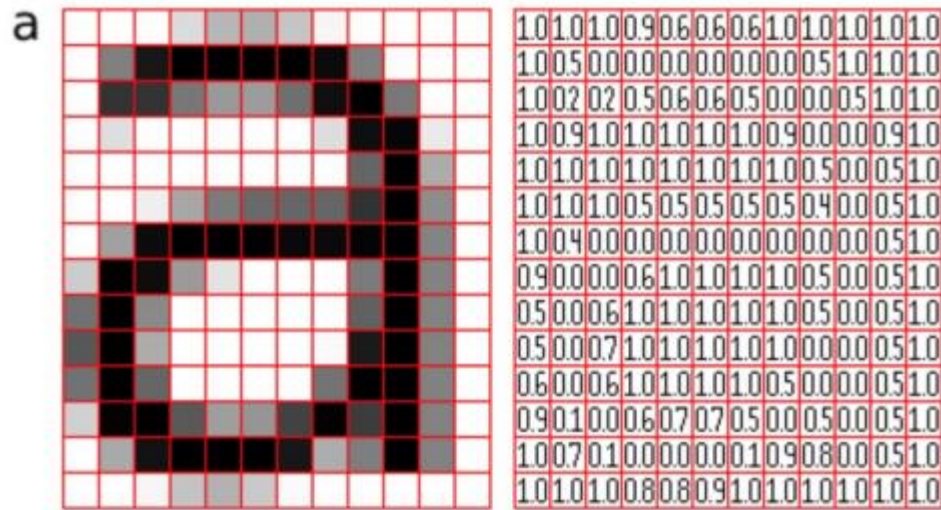
Поиск по заданному цвету – кожа человека



План лекции

- Теория цвета
- **Представление изображения в компьютере**
- Представление изображения в виде функции

Цифровое изображение



$$f(x, y) = \begin{bmatrix} f(0,0) & \cdots & f(0, n-1) \\ \vdots & \ddots & \vdots \\ f(m-1, 0) & \cdots & f(m-1, n-1) \end{bmatrix}$$

$$0 \leq f(x, y) \leq L$$

Обычно $L = 255 - \text{uint8}$

Типы изображений

Бинарное
(Binary)



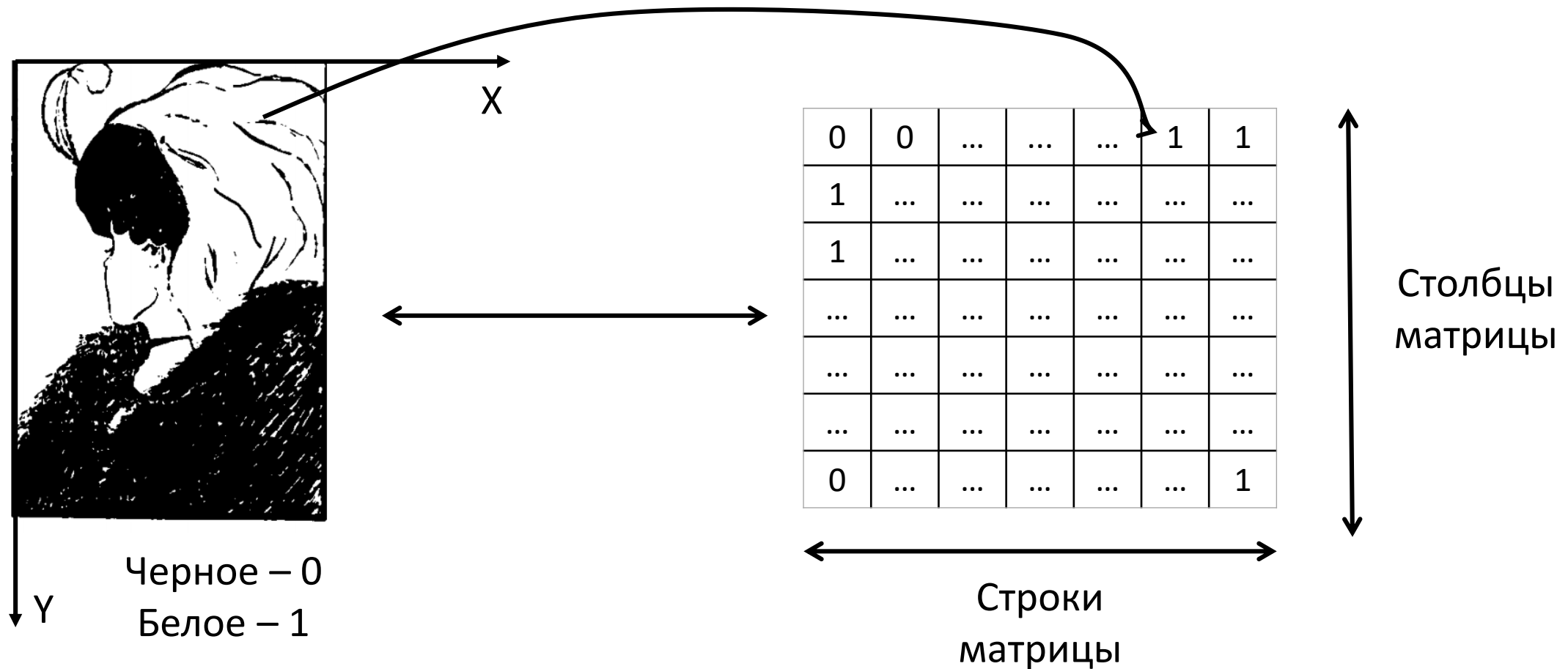
В градации серого
(Grayscale)



Цветное
(Color)



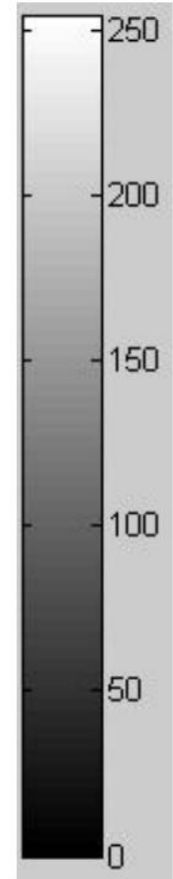
Binary представление изображения



Grayscale представление изображения



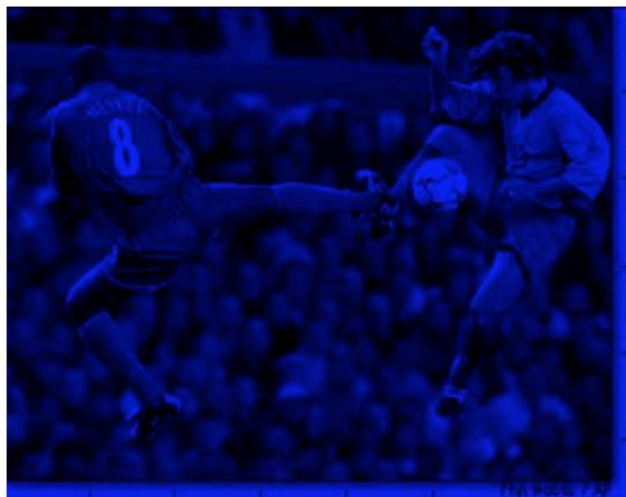
11	10	136	130
78
15
...
...
...
46	200



Color представление изображения – один канал



Color представление изображения

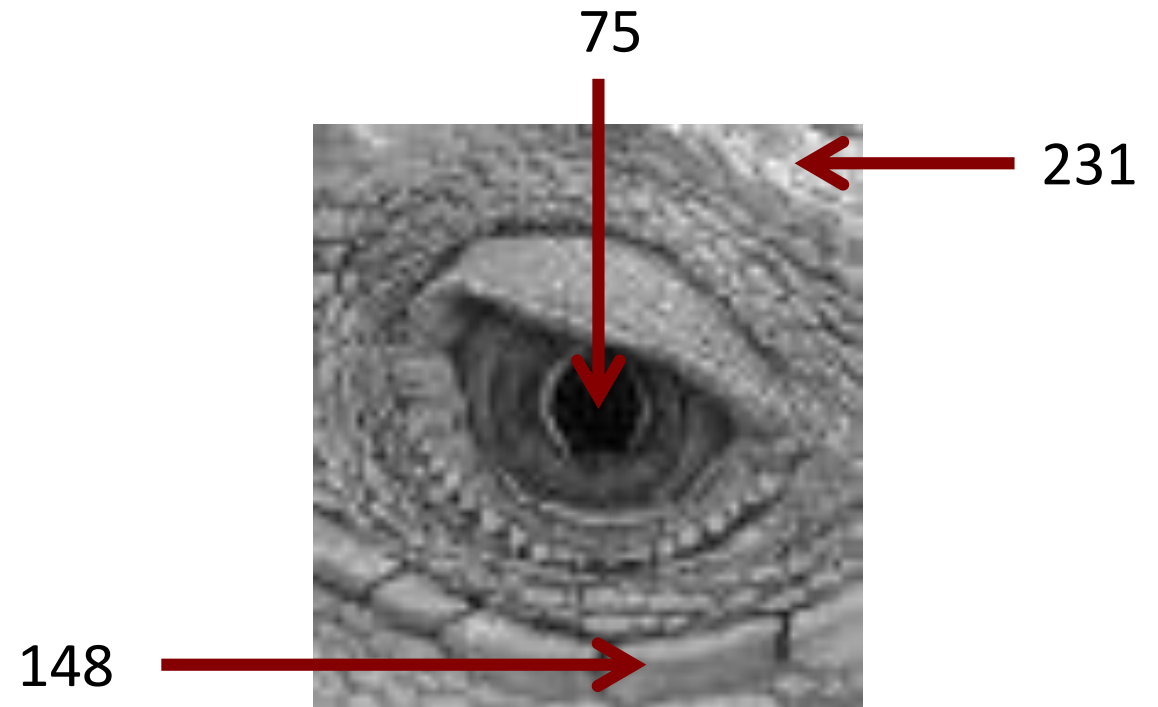


Представление части изображения

Изображение содержит дискретное количество пикселей

Значение пикселя:

- «шкала серого»
- (или «интенсивность»): $[0, 255]$



Представление части изображения

Изображение содержит дискретное количество пикселей

[90, 0, 53]

Значение пикселя:

- «grayscale»

- (или «интенсивность»): [0,255]

- «color»

- RGB: [R, G, B]

- Lab: [L, a, b]

- HSV: [H, S, V]



[249, 215, 203]

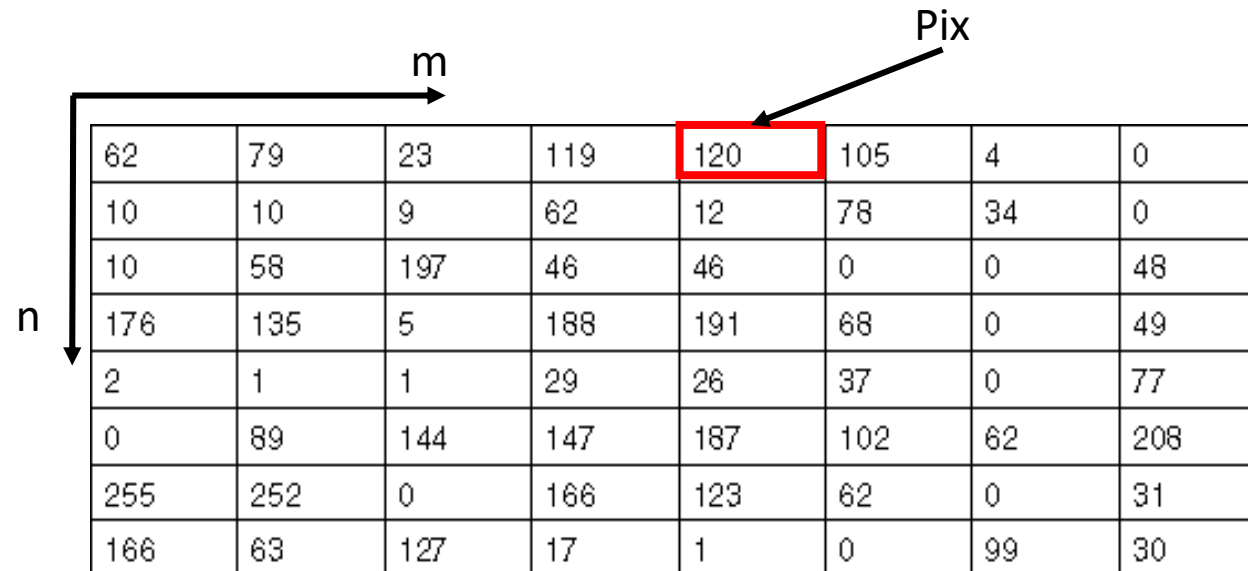
[213, 60, 67]

План лекции

- Теория цвета
- Представление изображения в компьютере
- **Представление изображения в виде функции**

Изображение как дискретная функция

- Изображения обычно цифровые (дискретные):
 - Пример 2D пространства на регулярной сетке
- Представлено в виде матрицы целочисленных значений



62	79	23	119	120	105	4	0
10	10	9	62	12	78	34	0
10	58	197	46	46	0	0	48
176	135	5	188	191	68	0	49
2	1	1	29	26	37	0	77
0	89	144	147	187	102	62	208
255	252	0	166	123	62	0	31
166	63	127	17	1	0	99	30

Изображение как дискретная функция

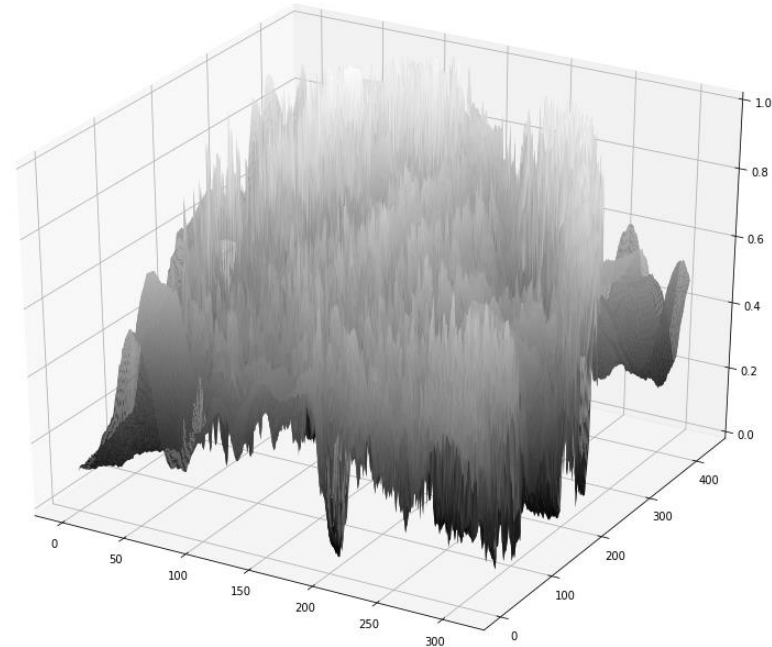
Декартовы координаты

$$f[n, m] = \begin{bmatrix} \ddots & & \vdots & & \\ & f[-1, 1] & f[0, 1] & f[1, 1] & \\ \dots & f[-1, 0] & \underline{f[0, 0]} & f[1, 0] & \dots \\ & f[-1, -1] & f[0, -1] & f[1, -1] & \\ & & \vdots & & \ddots \end{bmatrix}$$

Изображение как дискретная функция

Изображение как функция f от \mathbb{R}^2 до \mathbb{R}^M :

- $f(x, y)$ дает интенсивность в позиции (x, y)
- Определяется через прямоугольник, с конечным диапазоном:
 $f: [a, b] \times [c, d] \rightarrow [0, 255]$



Изображение как дискретная функция

Изображение как функция f от R^2 до R^M :

- $f(x, y)$ дает интенсивность в позиции (x, y)
- Определяется через прямоугольник, с конечным диапазоном:

$$f: [a, b] \times [c, d] \rightarrow [0, 255]$$

- Цветное изображение: $f(x, y) = \begin{bmatrix} r(x, y) \\ g(x, y) \\ b(x, y) \end{bmatrix}$

Гомогенные координаты

Обычные координаты

$$(x \ y)^T$$

$$R_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$$

Гомогенные координаты

$(sx \ sy \ s)^T$, где $s \neq 0$, но обычно $s = 1$

$$\tilde{x} = \begin{pmatrix} x \\ 1 \end{pmatrix}$$

$$R_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Аффинные трансформации

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos(\phi) & -\sin(\phi) & t_x \\ \sin(\phi) & \cos(\phi) & t_y \\ 0 & 0 & 1 \end{pmatrix}$$

Трансформация перспективы

$$P = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \quad P_{33} = 1, \text{ т.к. } P \sim aP \ \forall a \neq 0$$

Трансформации

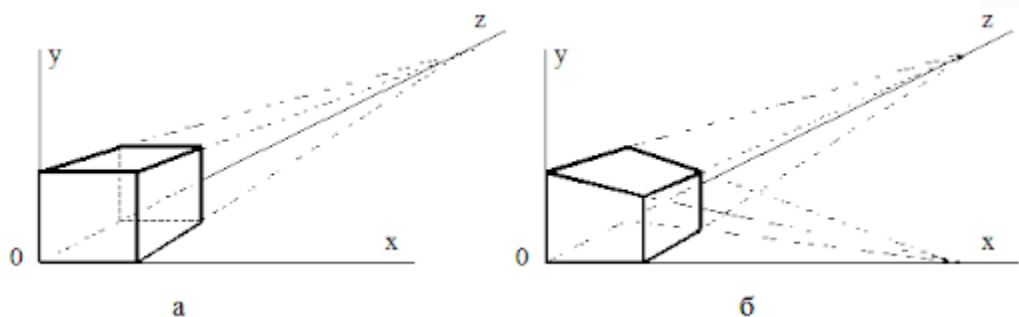
Transform of unit square	Name	Transformation matrix	DoF
	Translation	$\begin{pmatrix} 1 & 0 & t_1 \\ 0 & 1 & t_2 \\ 0 & 0 & 1 \end{pmatrix}$	2
	Rotation	$\begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$	1
	Rigid Body	$\begin{pmatrix} \cos(\phi) & -\sin(\phi) & t_x \\ \sin(\phi) & \cos(\phi) & t_y \\ 0 & 0 & 1 \end{pmatrix}$	3
	Affine	$\begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix}$	6
	Projective Transform	$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix}$	8

DoF – Degrees of Freedom

Оценка параметров. Аффинное преобразование

$$\begin{aligned}
 & \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = \begin{pmatrix} x_i & y_i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_i & y_i & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \quad \longrightarrow \\
 & \longrightarrow \begin{pmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_2 & y_2 & 1 \\ x_3 & y_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_3 & y_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} \quad \longrightarrow \quad \mathbf{q} = M\mathbf{p} \quad \longrightarrow \quad \mathbf{p} = (M^\top M)^{-1} M^\top \mathbf{q}
 \end{aligned}$$

Оценка параметров. Перспективное преобразование



$$s \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

$$\tilde{\mathbf{x}}' \sim P \tilde{\mathbf{x}}$$

$$x' = \frac{ax + by + c}{gx + hy + i}$$

$$y' = \frac{dx + ey + f}{gx + hy + i}$$

$$\begin{pmatrix} x & y & 1 & 0 & 0 & 0 & -x'x & -x'y & -x' \\ 0 & 0 & 0 & x & y & 1 & -y'x & -y'y & -y' \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1x_1 & -x'_1y_1 & -x'_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1x_1 & -y'_1y_1 & -y'_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2x_2 & -x'_2y_2 & -x'_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2x_2 & -y'_2y_2 & -y'_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3x_3 & -x'_3y_3 & -x'_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3x_3 & -y'_3y_3 & -y'_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4x_4 & -x'_4y_4 & -x'_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4x_4 & -y'_4y_4 & -y'_4 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \\ i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{M}\mathbf{p} = 0$$

$$\min \|Mp\|, \|p\| = 1$$

$$\min \|UDV^T p\|, \|p\| = 1$$

$$\min \|DV^T p\|, \|p\| = 1, q = V^T p$$

$$\min \|Dq\|, \|Vq\| = 1$$

ИТОГИ

- Рассмотрены цветовые пространства: RGB, XYZ, HSV, Lab, YCbCr
- Показаны виды представления изображений: Binary, Grayscale, Color
- Изучена интерпретация изображения в виде двумерной дискретной функции