

Компьютерное
Зрение Лекция №1,
осень 2024

Введение в цифровые изображения



План лекции

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

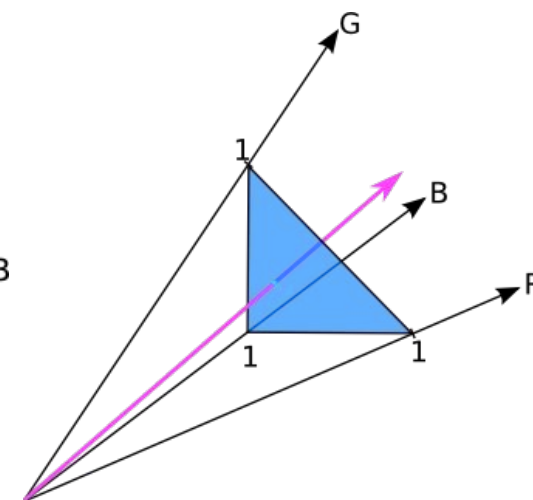
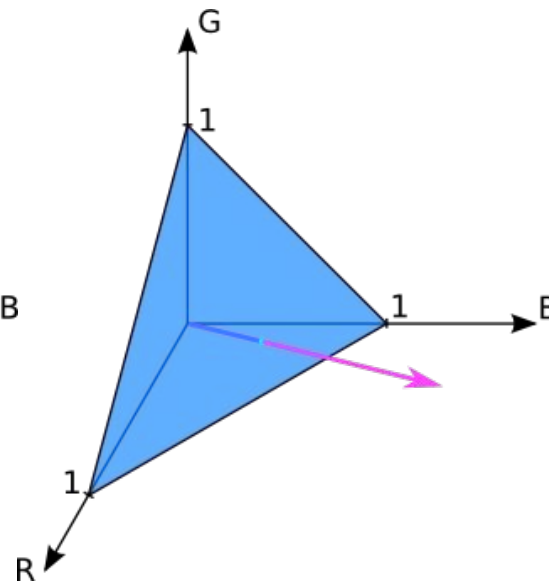
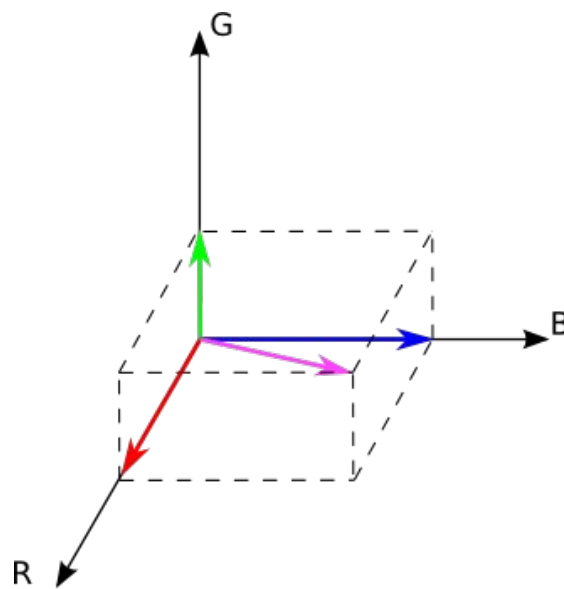
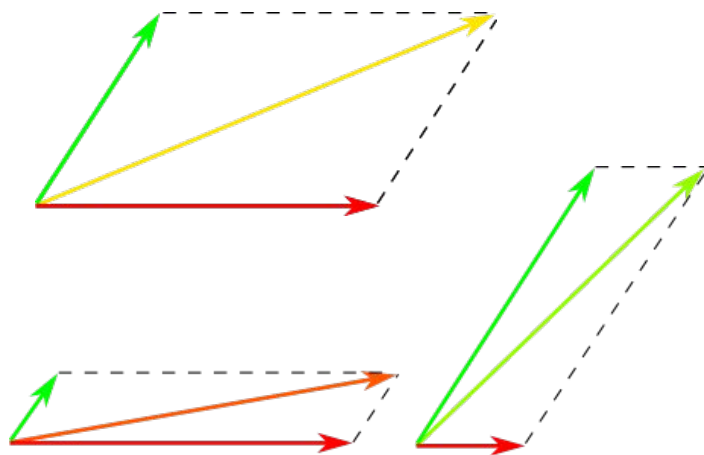
План лекции

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

Координатная система

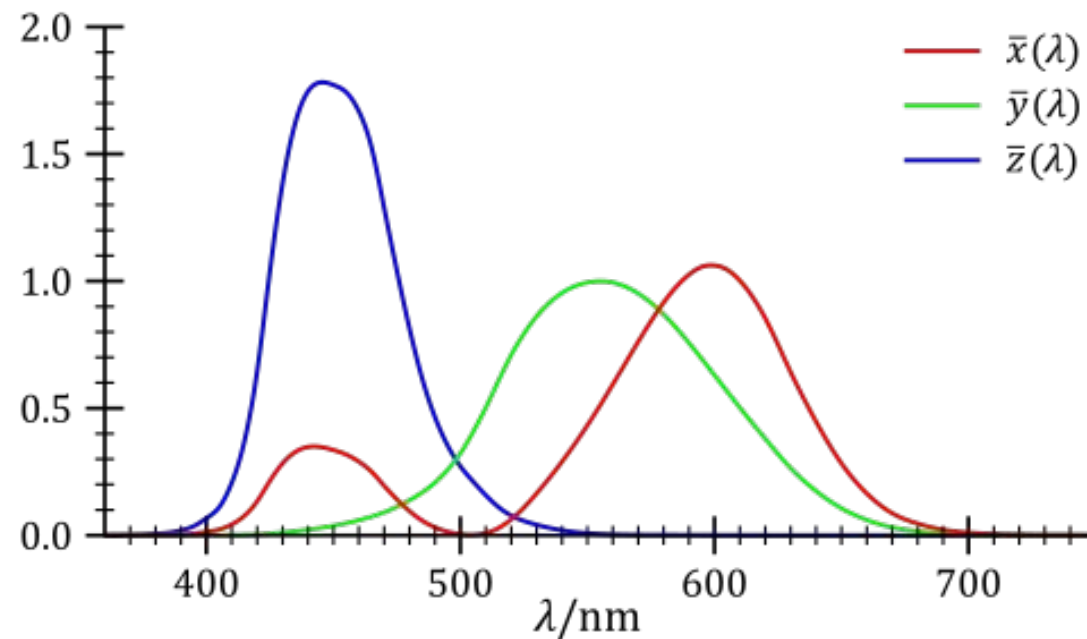
RGB

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



Система CIE

XYZ



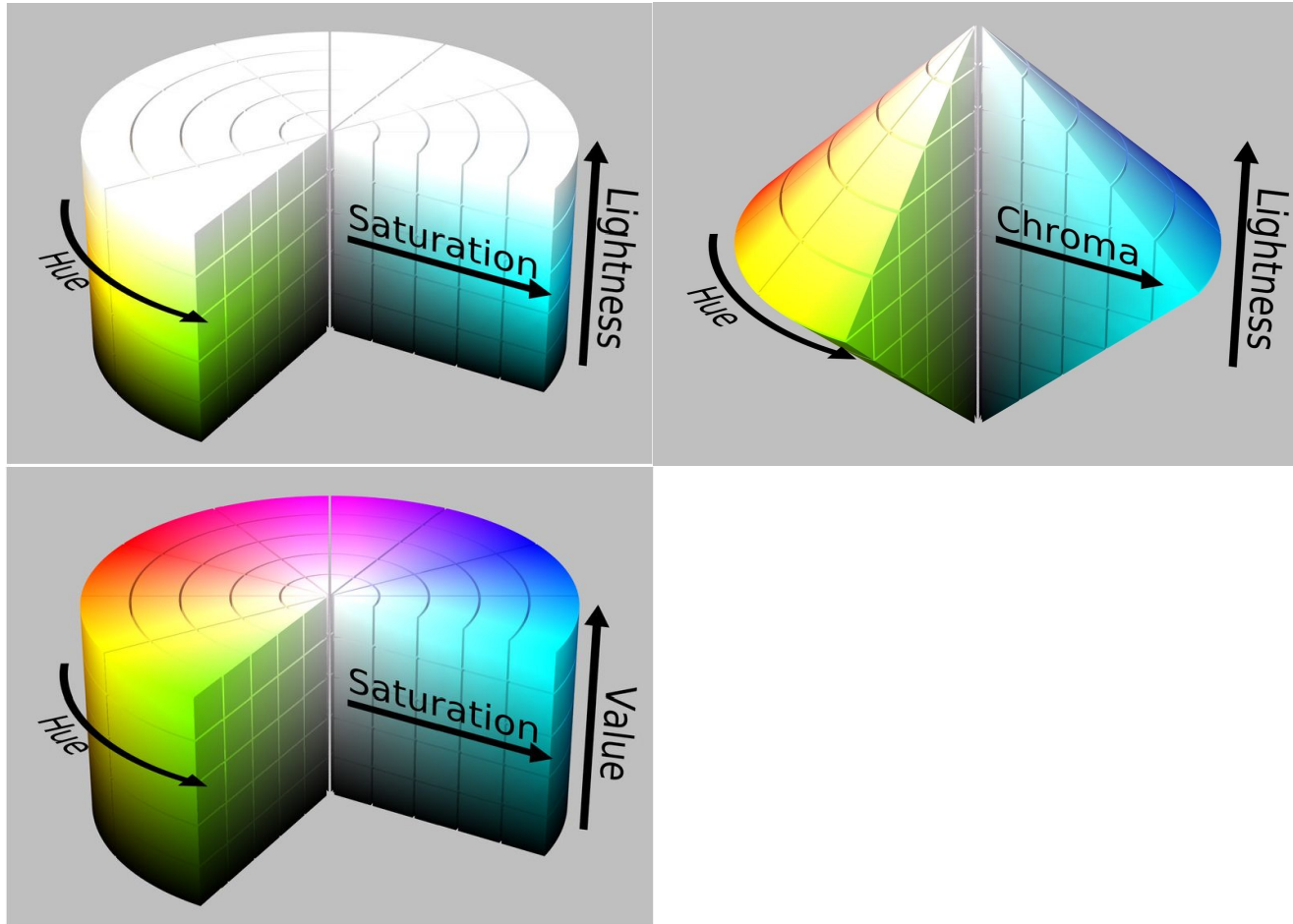
Свойства системы:

- 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/HSV

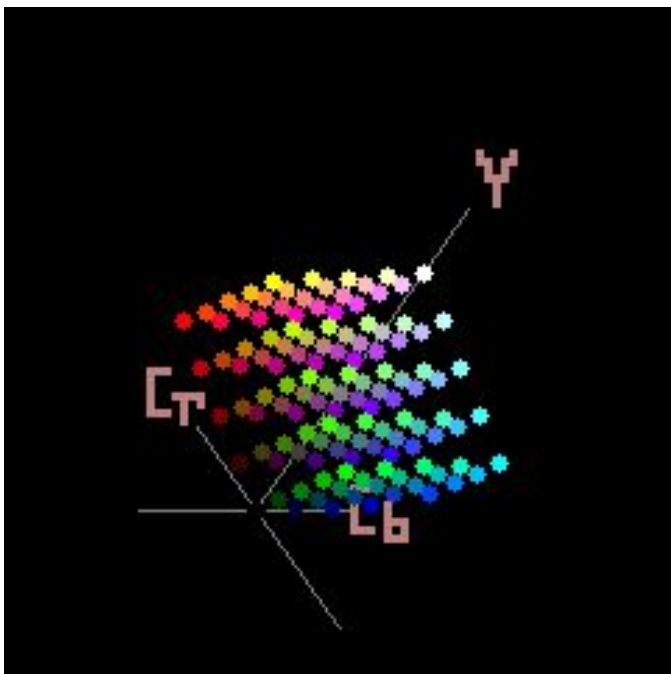


$$I = \frac{R + G + B}{3}$$

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

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

Цветовое пространство YCbCr



Преобразование в пространство
YCbCr:

$$Y = k_r R + k_g G + k_b B$$

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

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

$$k_r + k_g + k_b = 1$$

k_r, k_g, k_b – весовые
коэффициенты

Источники для погружения в теорию цвета

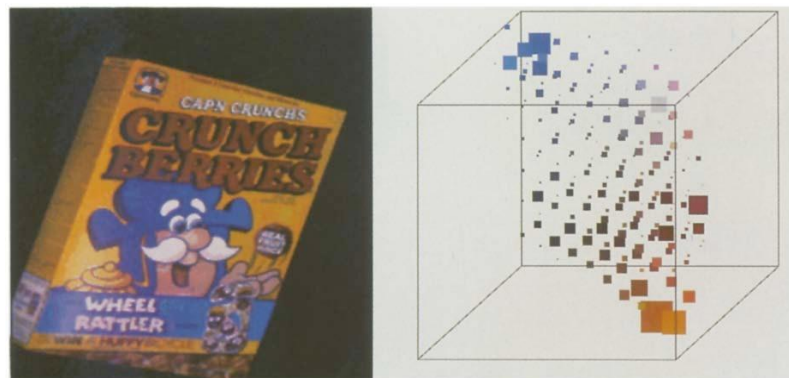
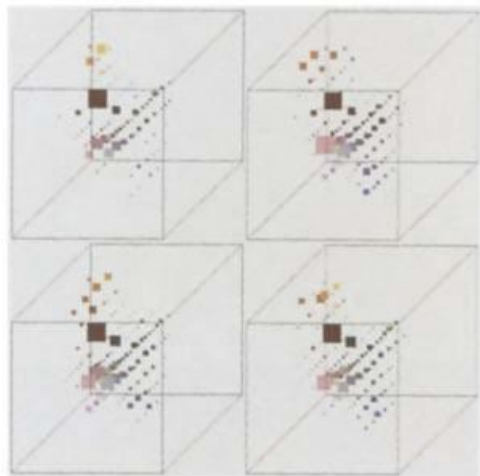
Лекция: [Как устроен цвет - Дмитрий Николаев, заведующий сектором зрительных систем ИППИ РАН](#)

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

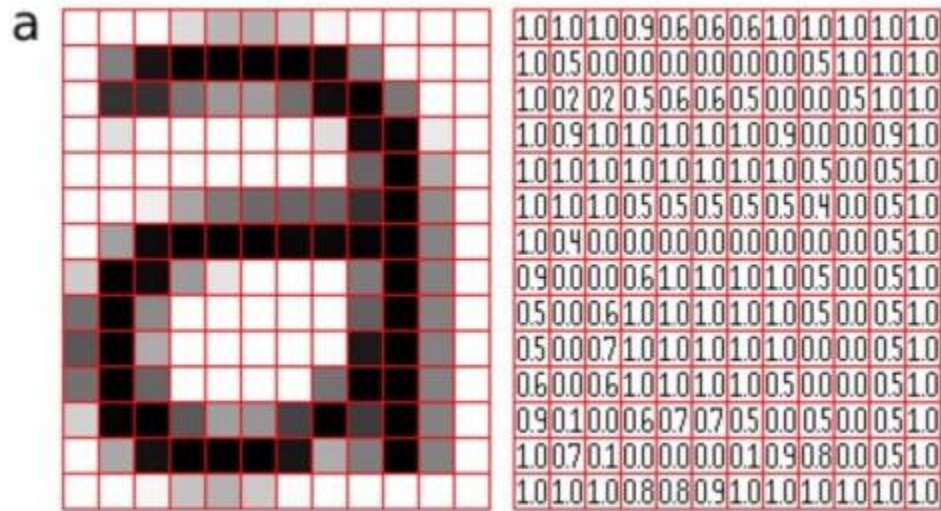
Поиск по заданному цвету – кожа
человека



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

Бинарно

е

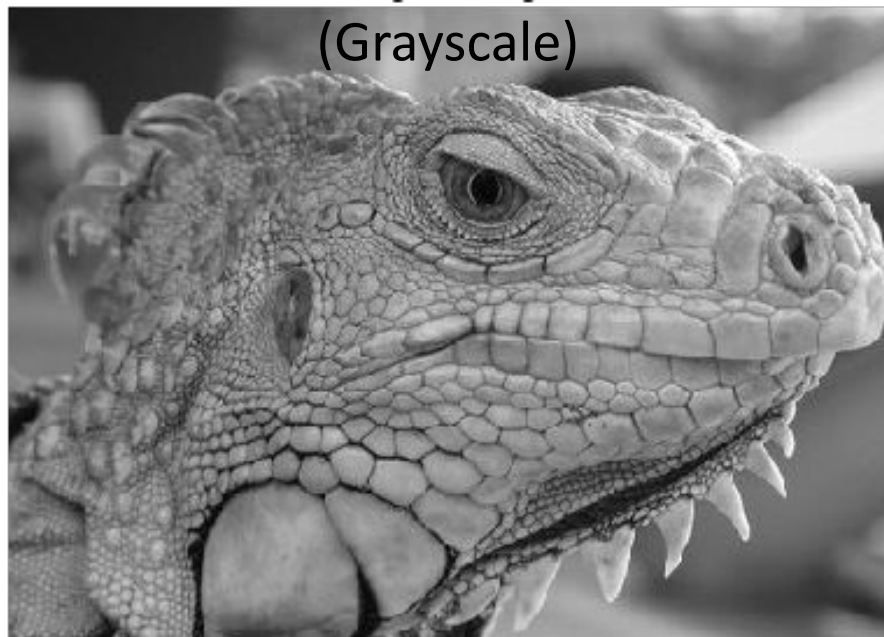
(Binary)



В градации

серого

(Grayscale)



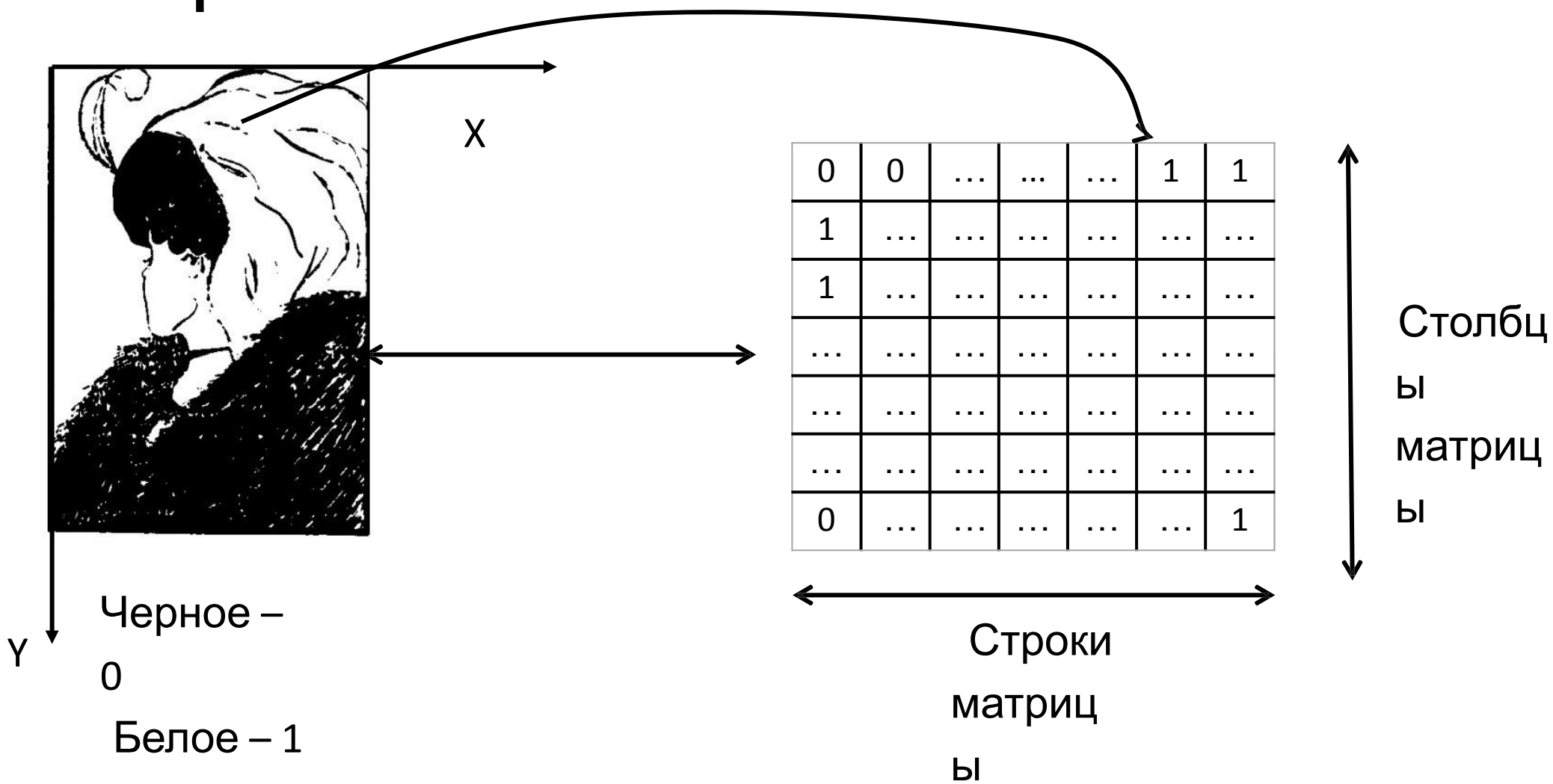
Цветно

е

(Color)



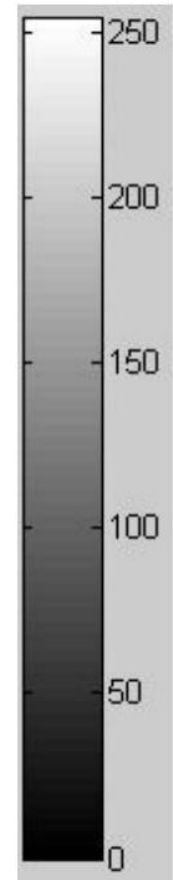
Вінарны прадстаўленне ізображэння



Grayscale представление изображения



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



Color представление изображения – один канал



Color представление

Из



Phil Noble / AP



Phil Noble / AP



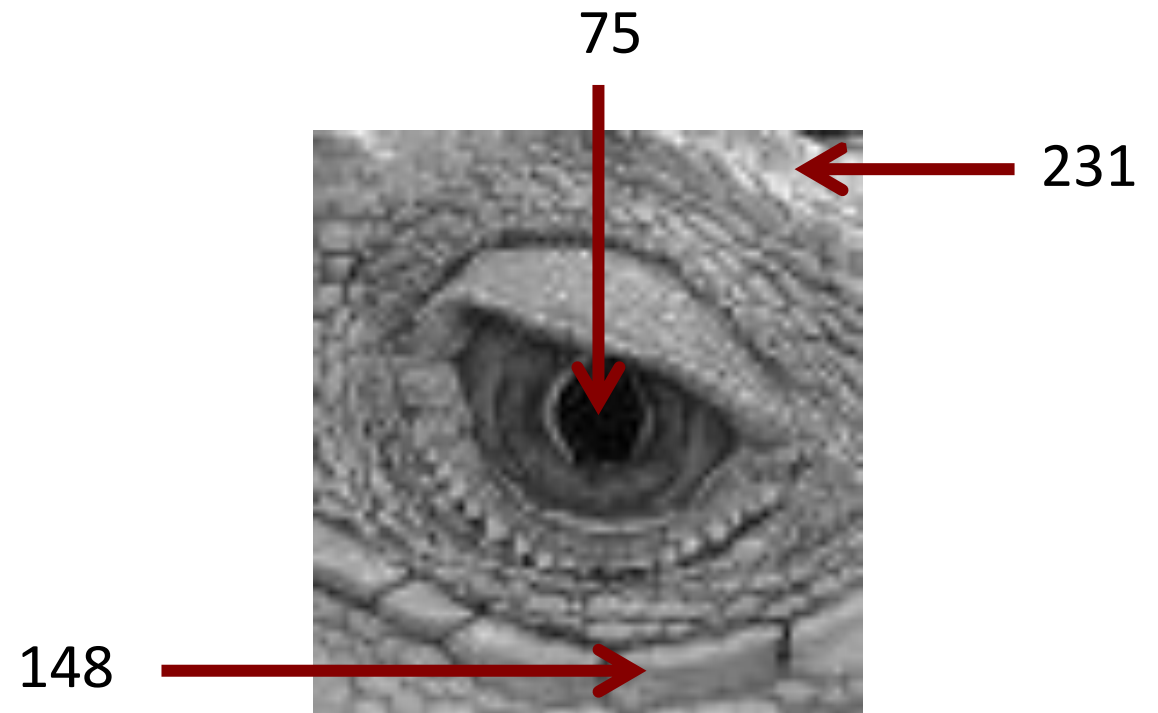
Phil Noble / AP

Представление части изображения

Изображение содержит дискретное количество пикселей

Значение пикселя:

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

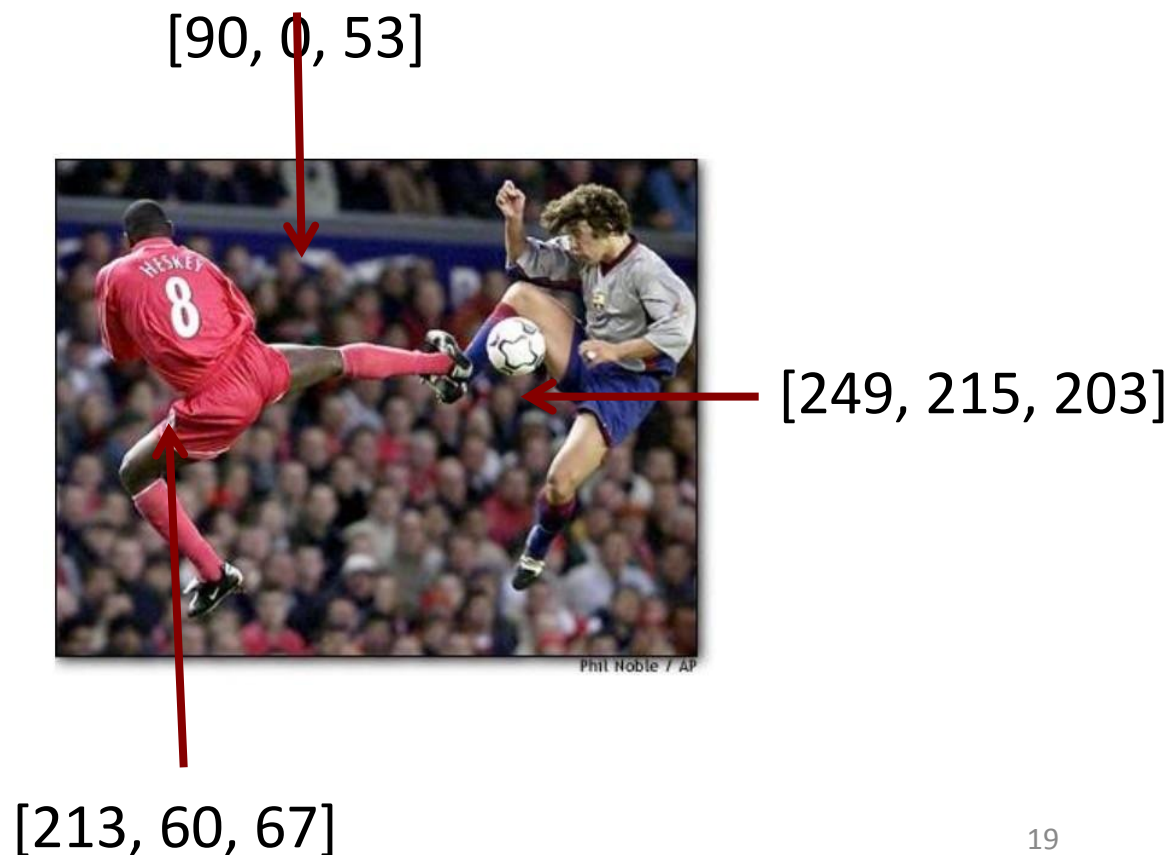


Представление части изображения

Изображение содержит дискретное количество
пикселей

Значение пикселя:

- «grayscale»
(или «интенсивность»): $[0, 255]$
- «color»
 - RGB: $[R, G, B]$
 - Lab: $[L, a, b]$
 - HSV: $[H, S, V]$

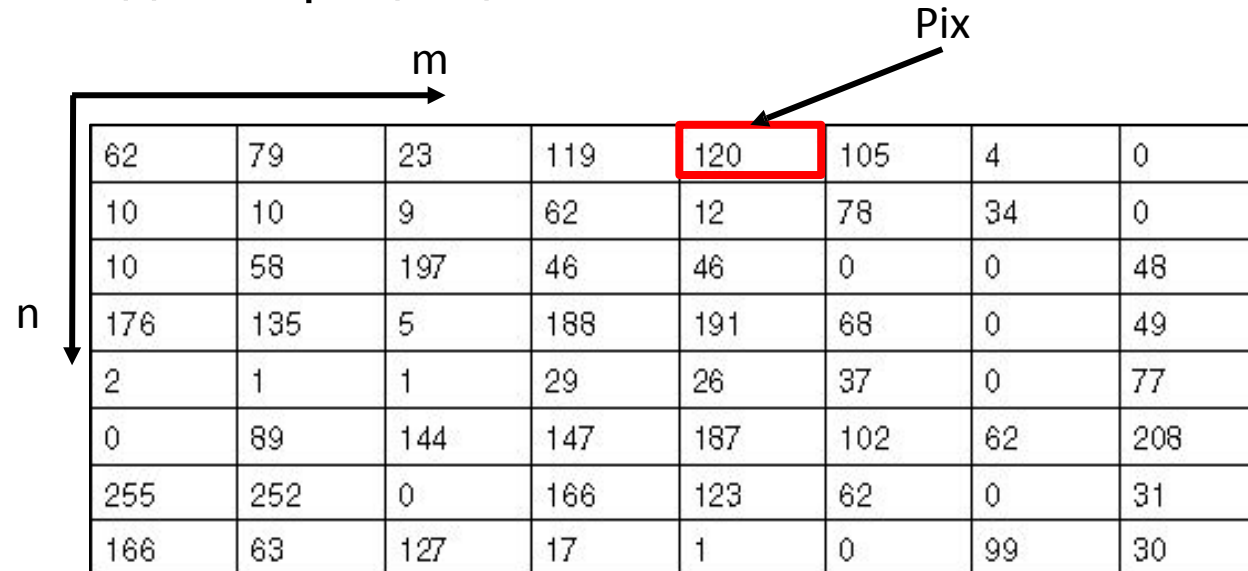


План лекции

- 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



The diagram illustrates a 2D pixel grid. A horizontal arrow labeled 'm' points to the right, indicating the column index. A vertical arrow labeled 'n' points downwards, indicating the row index. The grid consists of 8 rows and 8 columns of numerical values. The value 120, located at row 1 and column 5, is highlighted with a red border. An arrow labeled 'Pix' points to this highlighted cell.

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

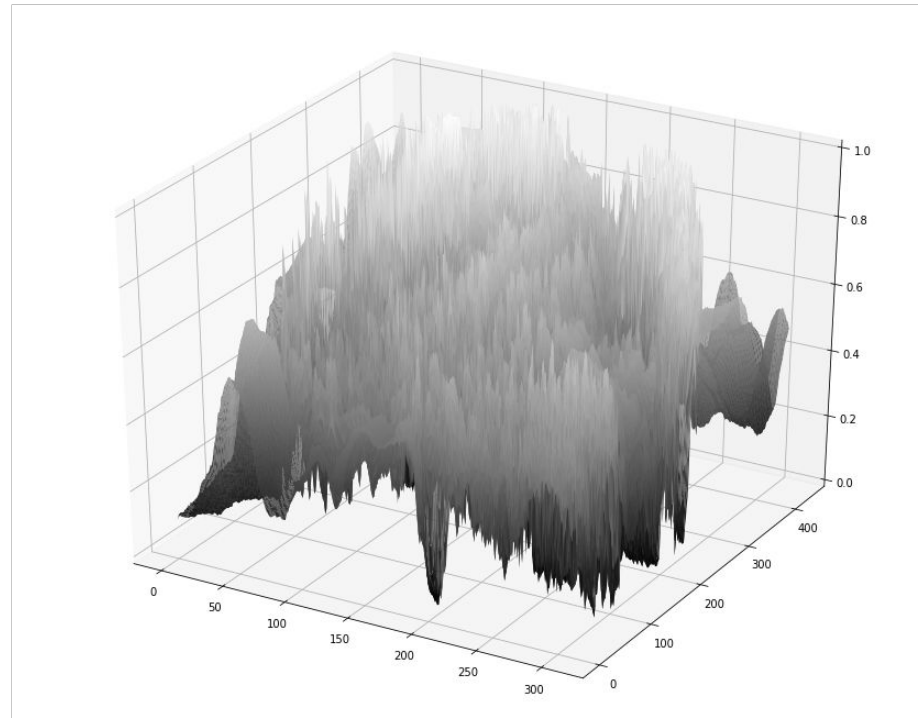
$$f[n, m] = \begin{bmatrix} \ddots & & \vdots & & \\ & f[-1, 1] & f[0, 1] & f[1, 1] & \\ \dots & f[-1, 0] & \underline{f[0, 0]} & f[1, 0] & \dots \\ & f[-1, -1] & f[0, -1] & f[1, -1] & \\ & & \vdots & & \ddots \end{bmatrix}$$

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 \mathbb{R}^2 до \mathbb{R}^M :

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

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

- 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 $s = 1$

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

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

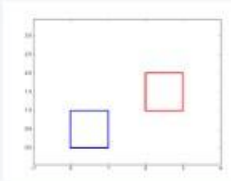
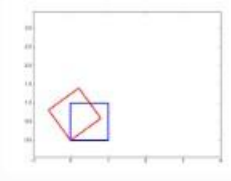
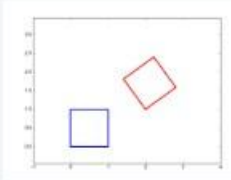
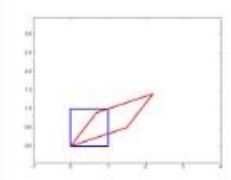
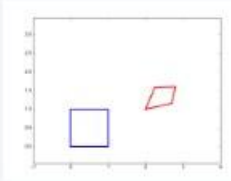
Аффинные
трансформации

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

Трансформация
перспективы

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

Трансформац

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

Совмещение преобразований.

Аффинное преобразование

- Поворот

$$R = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Итоговую матрицу получаем, умножив матрицы преобразований в следующем порядке:

- Масштабирование

$$S = \begin{pmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$M = T \cdot S \cdot R$$

- Сдвиг

$$T = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix}$$

Итоги

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