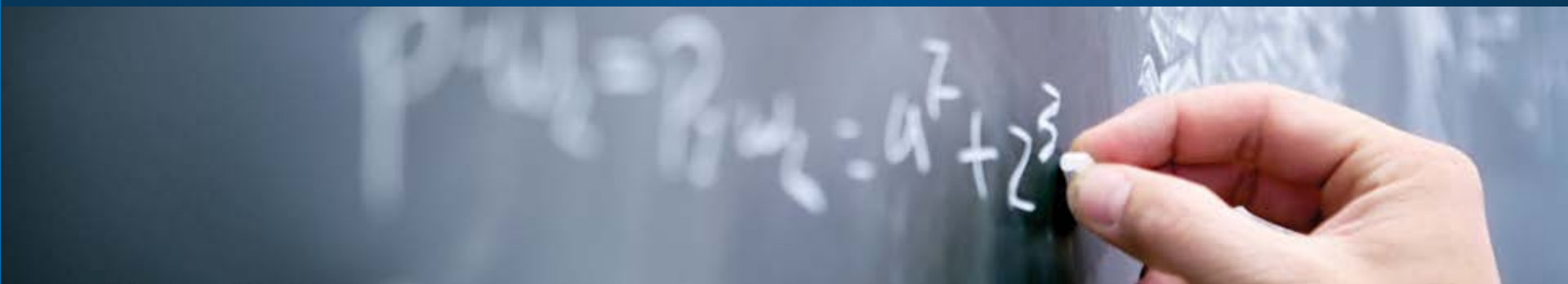
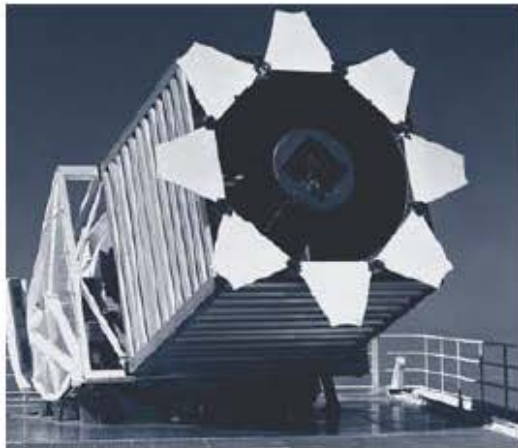


Анализ изображений и видео

Лекция 1: Введение в анализ изображений

Наталья Васильева
nvassilieva@hp.com
HP Labs Russia





Слайд К. Грауман

Что такое анализ изображений?

- *“Computing properties of the 3D world from one or more digital images”* (Trucco and Veri)
- *“Make useful decisions about real physical objects and scenes based on sensed images”* (Shapiro)
- *“The construction of explicit, meaningful description of physical objects from images”* (Ballard and Brown)

Анализ изображений – извлечение значимой информации из изображений



Что такое анализ изображений?

Опишите фотографию словами

Маленькая девочка ест мороженое на улице. Одета в джинсы и футболку с длинным рукавом. На заднем фоне панельные дома. Один из них голубого цвета в белую клетку.

Распознавание лиц

Распознавание позы

Распознавание действия

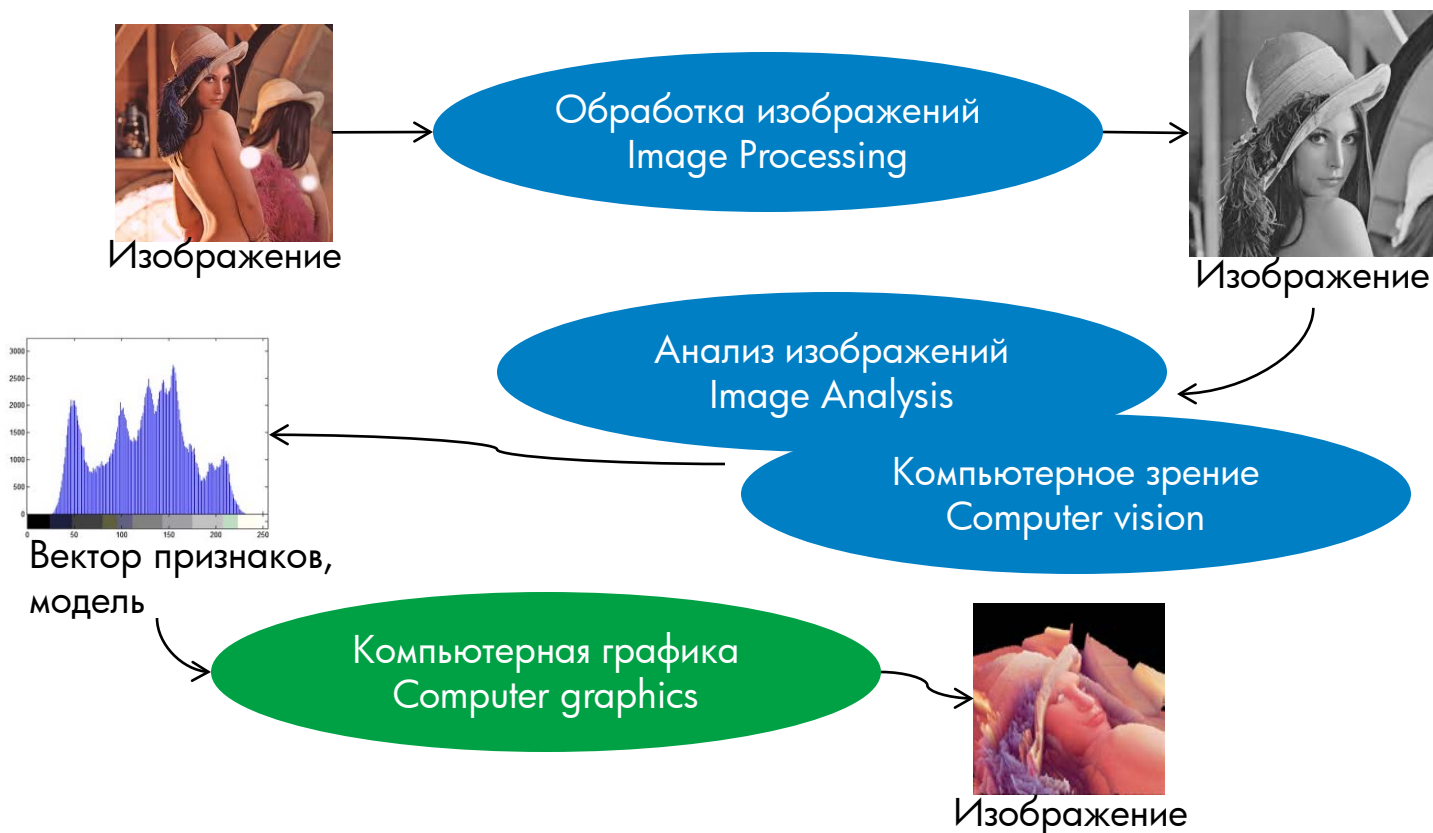
Выделение текстуры

Выделение областей, однородных по цвету

Распознавание объектов



Связь с другими дисциплинами



Распознавание образов (Pattern recognition)

Машинное обучение (Machine learning)

Анализ данных (Data mining)

Обработка сигналов (Signal processing)

Искусственный интеллект (Artificial intelligence)

Статистика

Нейропсихология

План лекции

- Структура курса
- Зачем нужен анализ изображений?
- Сложности
- Зрительное восприятие человека
- Представление цвета
- Цифровое представление изображений



Структура курса

- Лекции и семинары: 1 семинар через каждые 3 лекции
- Во время лекции можно и нужно задавать вопросы
- В начале лекций возможны короткие тесты (коэф. 0.2)

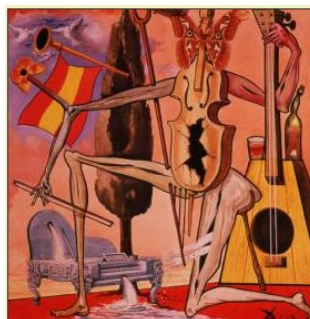
- Практические задания на дом (индивидуальные и групповые) (коэф. 0.35)
 - Разбор домашних заданий на семинарах в группах
- Чтение статей (факультативно) (коэф. 0.35)
- Финальный экзамен (коэф. 0.45)



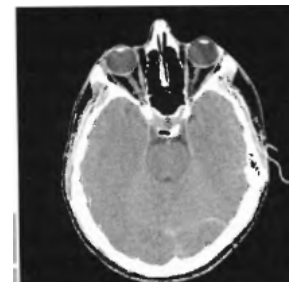
Индексирование и поиск изображений

- Персональные коллекции (тысячи фотографий)
- Коллекции профессиональных фото (миллионы фотографий)
- Коллекции репродукций (миллионы фотографий)

- Навигация
- Поиск дубликатов
- Выбор «лучшей» фотографии
- Автоматическое создание коллажей (<http://www.snapfishlab.com/>)
- Защита авторских прав



Медицина



- Выявление аномалий
- Диагностика заболеваний
- Моделирование организма и предсказание последствий лечения

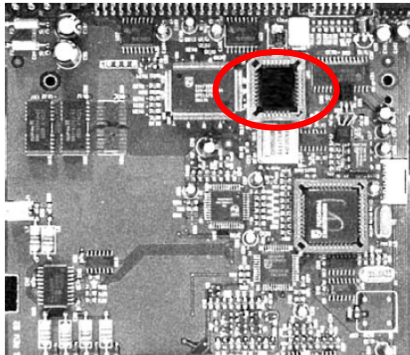
Системы безопасности

- Записи видеонаблюдений
- Отпечатки пальцев, снимки сетчатки глаза, изображения лиц
- Обнаружение «подозрительных» предметов
- Идентификация лиц, находящихся «в розыске»
- Авторизация доступа по отпечаткам пальцев

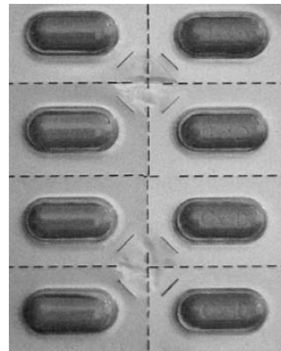


В промышленности

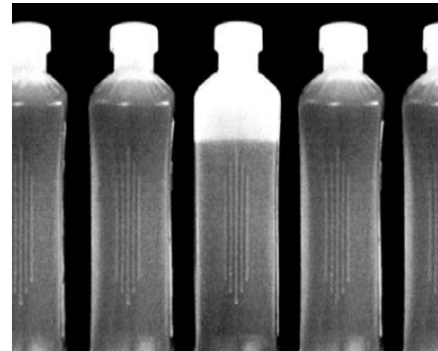
Визуальный контроль качества



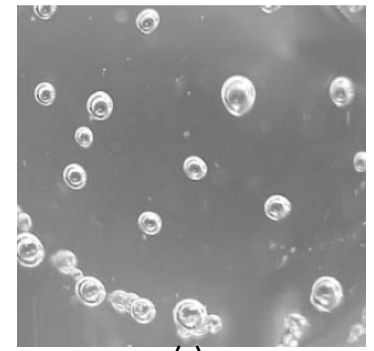
(а)



(б)

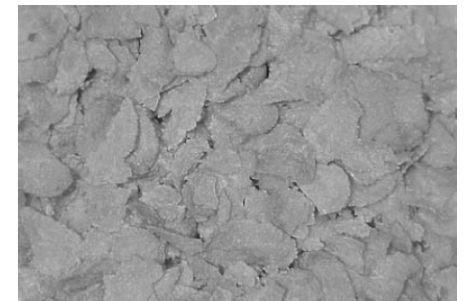


(в)



(г)

- (а) CD-ROM контроллер – все детали на месте?
- (б) Все таблетки на месте?
- (в) Контроль за уровнем жидкости в бутылках
- (г) Контроль за качеством пластика
- (д) Контроль за качеством кукурузных хлопьев!



(д)

Еще приложения

Военная промышленность

- Системы слежения и целенавещения

Автомобилестроение

- Автоматическое управление автомобилем

Кино, компьютерные игры

- Моделирование и 3D-реконструкция

Обработка сенсорных данных

- Зондирование земной поверхности

Оцифровка печатных документов

- Обнаружение и распознавание текста

Робототехника

- Зрение роботов

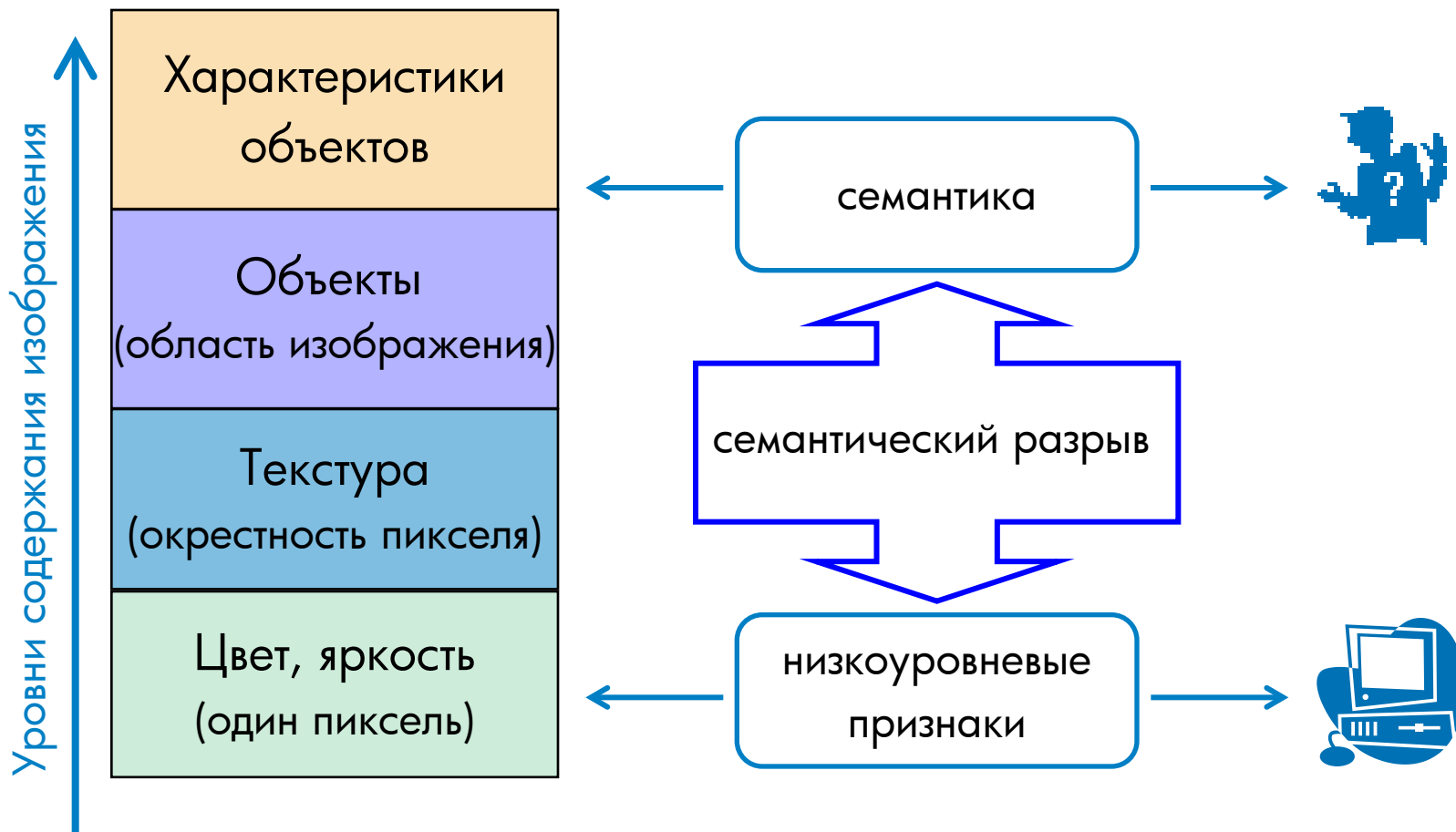


План лекции

- Структура курса
- Зачем нужен анализ изображений?
- **Сложности**
- Зрительное восприятие человека
- Представление цвета
- Цифровое представление изображений



Семантический разрыв



Как компьютеру «понять» изображение?

Что изображено?



Что изображено?



Challenges: robustness



Illumination



Object pose



Clutter



Occlusions



**Intra-class
appearance**



Viewpoint

План лекции

- Структура курса
- Зачем нужен анализ изображений?
- Сложности
- Зрительное восприятие человека
- Представление цвета
- Цифровое представление изображений



Несколько фактов о нашем зрительном восприятии

Наш мозг часто «достраивает» картинку и добавляет семантику
(Мы все можем узнать «что-то» или «кого-то» в очертании облака)

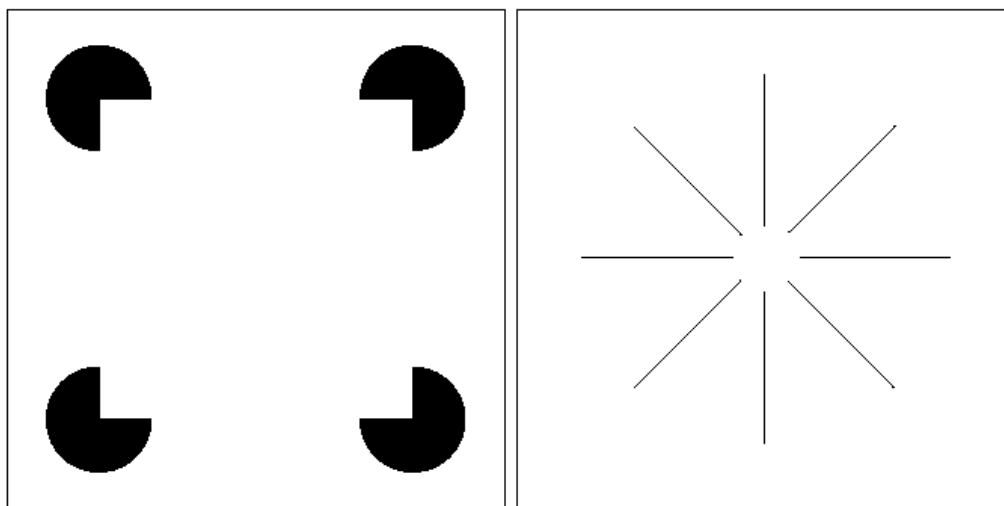
Зрительная система самообучается

- Европейцу сложно различать лица азиатов
- Мы ищем знакомые образы в картинке



Несколько фактов о нашем зрительном восприятии

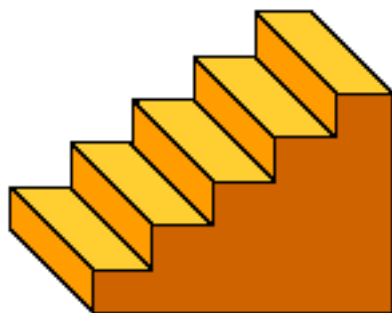
- Мы ищем знакомые образы, мозг «достраивает» картинку



Известные оптические иллюзии

Несколько фактов о нашем зрительном восприятии

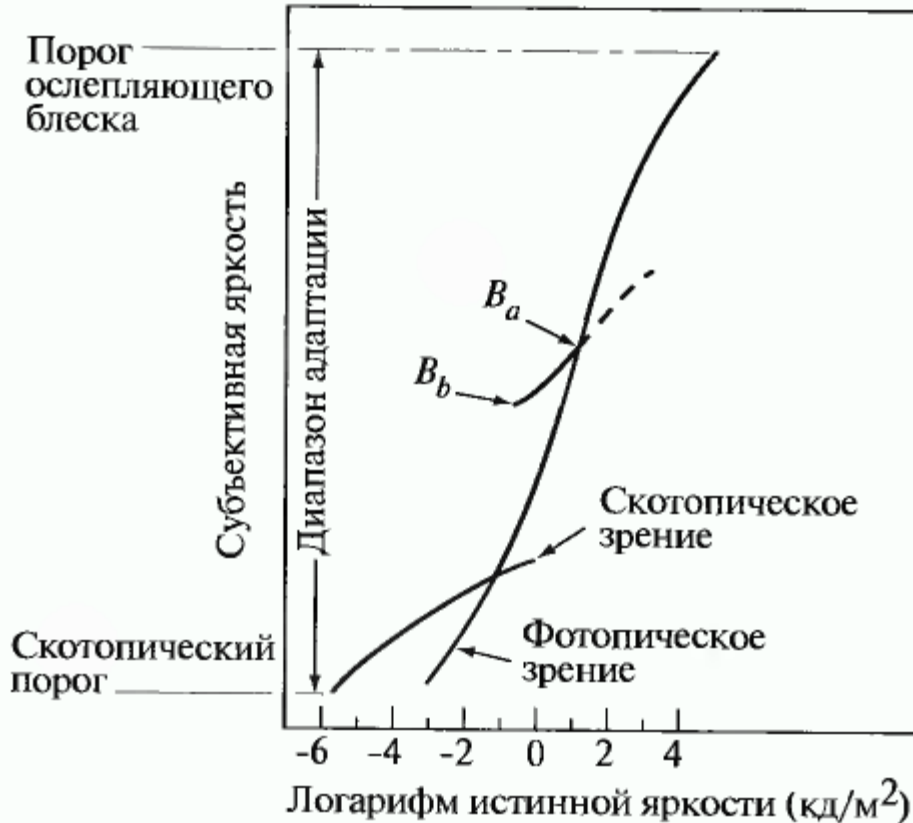
От того, где мы выросли, зависит, как мы «видим» мир вокруг нас



Ступеньки вверх или вниз?

- Арабы ответят «сверху вниз»

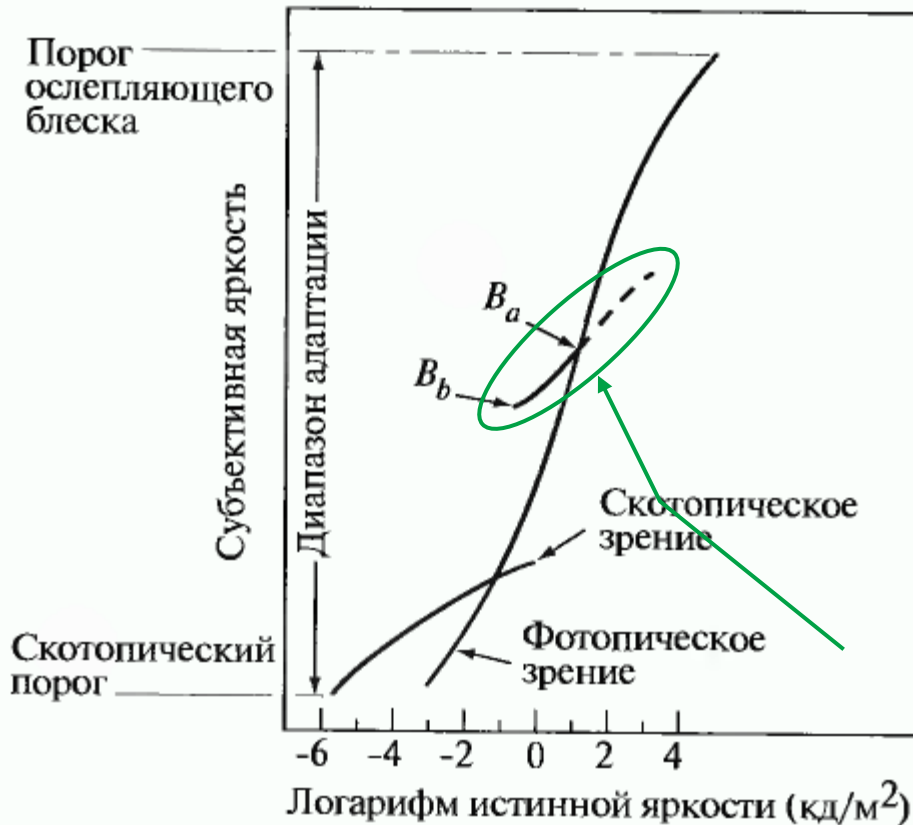
Яркостная адаптация и контрастная чувствительность



Зрительная система способна адаптироваться к диапазону значений яркости порядка 10^{10}

Субъективная яркость — логарифмическая функция от физической яркости

Яркость адаптация и контрастная чувствительность



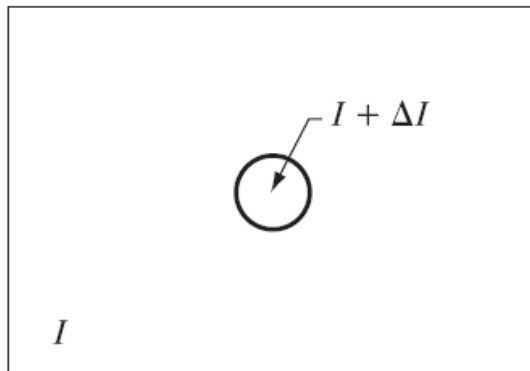
Зрительная система не способна работать во всем диапазоне одновременно

Диапазон субъективной яркости при адаптации к уровню V_a

Уровни яркости ниже V_b субъективно воспринимаются как черные

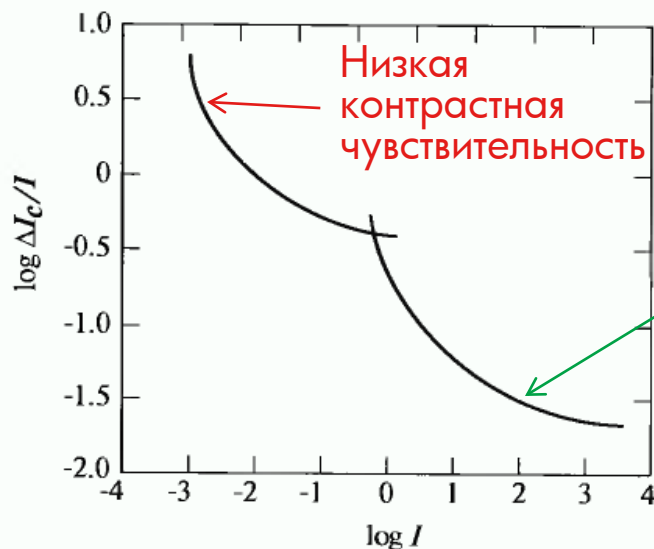
Отношение Вебера

Контрастная чувствительность – отношение Вебера



$\frac{\Delta I_c}{I}$ – Отношение Вебера

ΔI_c – приращение яркости, различимое в 50% случаев на фоне яркости I

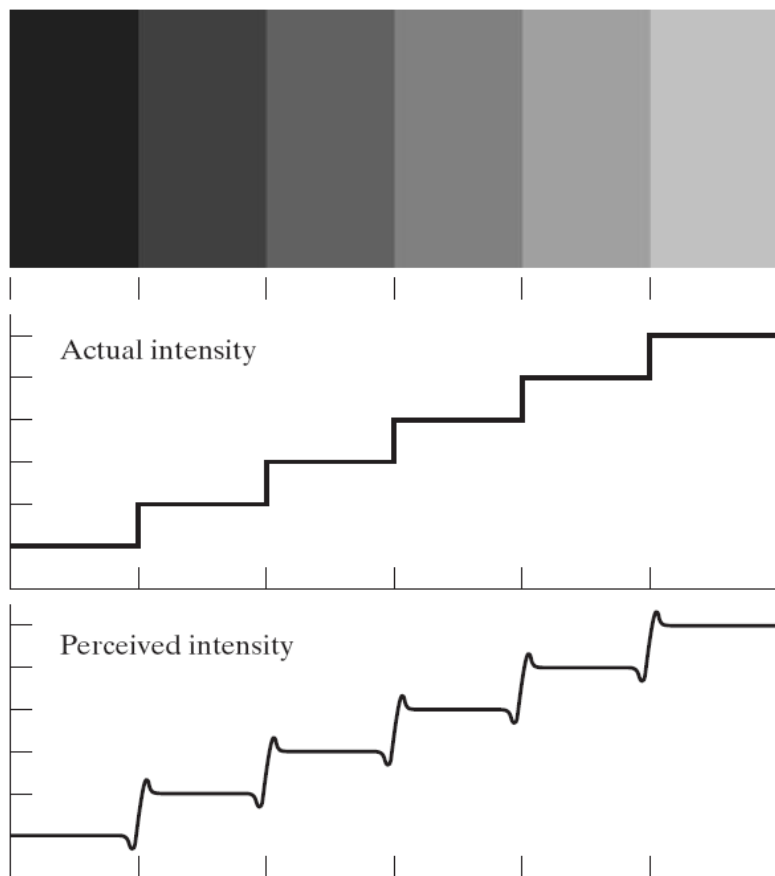


Зависимость отношения Вебера, как функции яркости

Высокая контрастная чувствительность

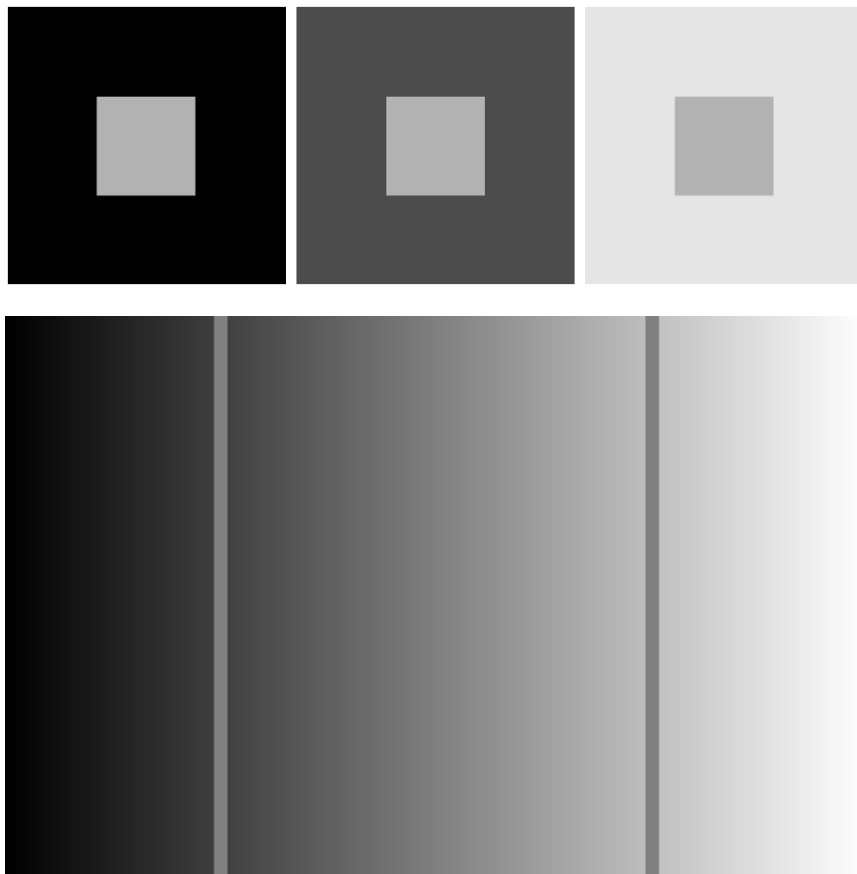
Воспринимаемая яркость как функция истинной яркости

Полосы Маха



Воспринимаемая яркость как функция истинной яркости

Одновременный контраст

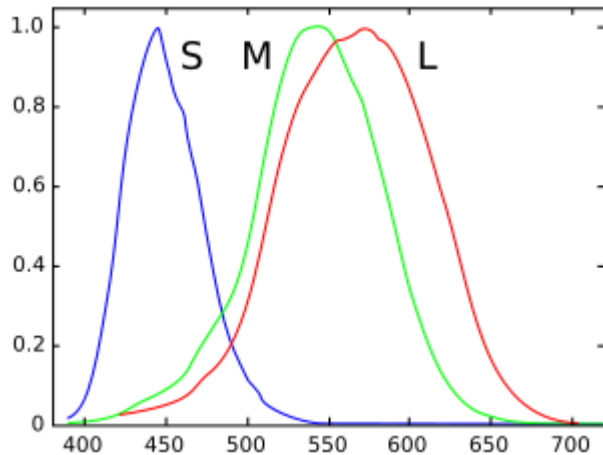


План лекции

- Структура курса
- Зачем нужен анализ изображений?
- Сложности
- Зрительное восприятие человека
- **Представление цвета**
- Цифровое представление изображений



Восприятие цвета



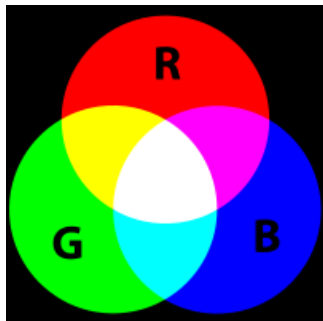
Спектральная чувствительность колбочек трех групп

Три группы колбочек, различающихся чувствительностью к свету с различной длиной волны:

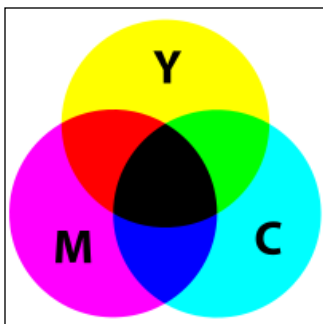
- S-колбочки: короткие волны (blue)
- M-колбочки: средние волны (green)
- L-колбочки: длинные волны (red)

Восприятие цвета

Первичные и вторичные цвета



Аддитивные первичные цвета
(световые источники)



Субтрактивные первичные
цвета (красители)

- Человек воспринимает цвета, как различные сочетания «первичных» цветов: красного, зеленого и синего.
- Сочетание первичных цветов дает вторичные: пурпурный (R+B), голубой (G + B) и желтый (R + G).
- Для красителей – первичный цвет поглощает один первичный основной цвет светового источника.

Светлота, цветовой тон и насыщенность

- **Светлота** связана со зрительным ощущением интенсивности
- **Цветовой тон** характеризует доминирующий цвет
- **Насыщенность** связана с относительной белизной цвета
- Цветовой тон и насыщенность вместе называют **цветностью**



От координат цвета к координатам цветности

- Величины красного, синего и зеленого, необходимые для получения некоторого цвета называются **координатами цвета** и обозначаются как **X, Y, and Z**
- Цветовой тон и насыщенность можно выразить в координатах цветности:

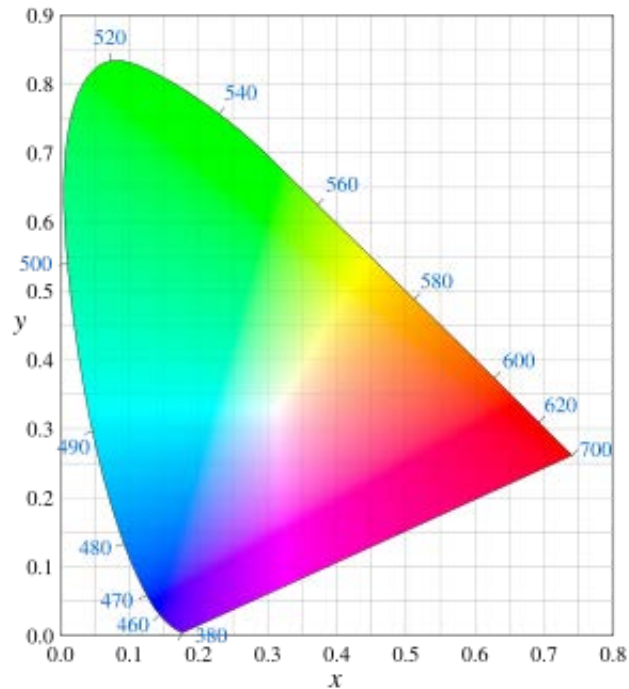
$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}, \quad z = \frac{Z}{X + Y + Z}$$

$$x + y + z = 1$$



Диаграмма цветности

Диаграмма цветности МКО CIE xy Chromaticity Diagram

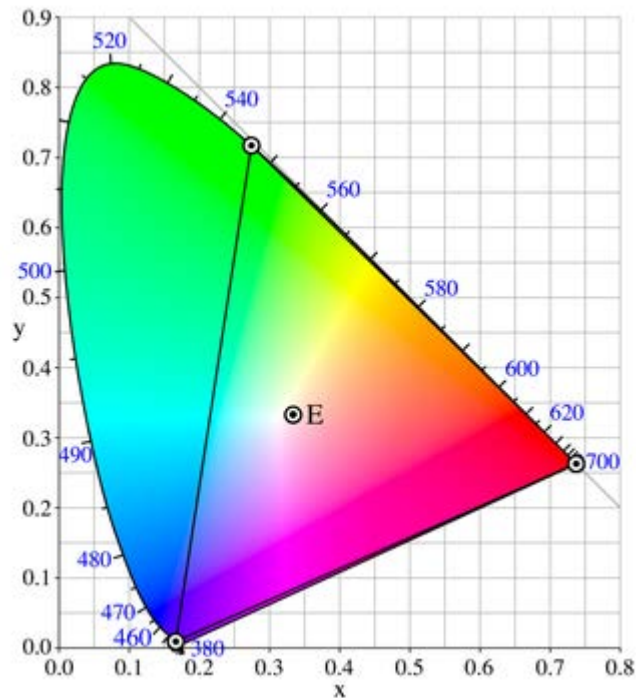


The CIE 1931 chromaticity diagram.

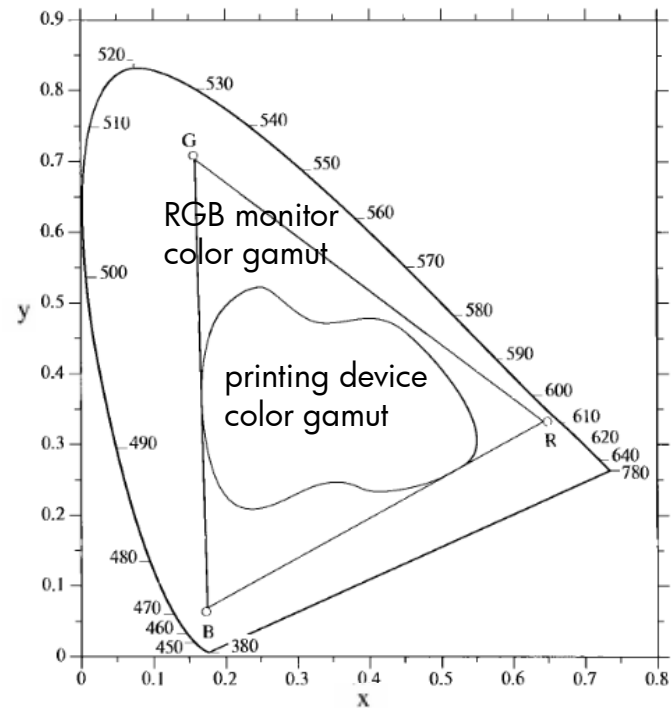
- Создана Международной Коммиссией по Освещению (МКО) в 1931.
- Функция от x (red) и y (green) :
 $z = 1 - (x + y)$.
- Вдоль границы – цвета спектра.
- $(x, y) = (1/3, 1/3)$ - опорный белый цвет стандарта МКО, точка равной энергии.
- Любая точка на границе имеет максимальную насыщенность.
- Граница \rightarrow точка равной энергии:
насыщенность $\rightarrow 0$

Цветовой охват

Color Gamut



Gamut of the CIE RGB primaries and location of primaries on the CIE 1931 xy chromaticity diagram.



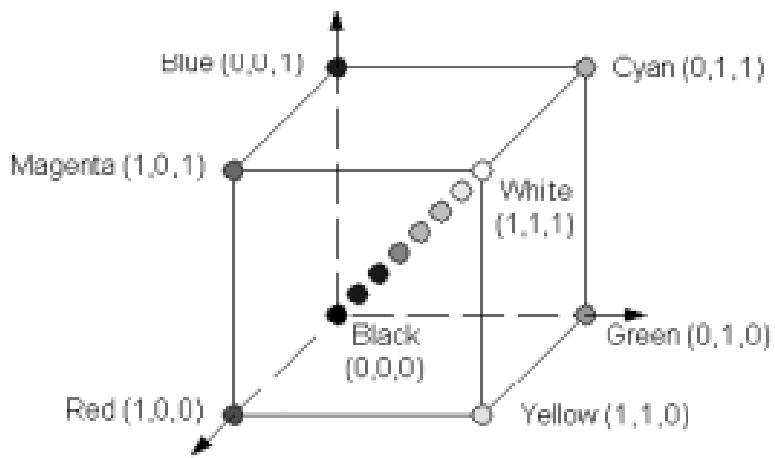
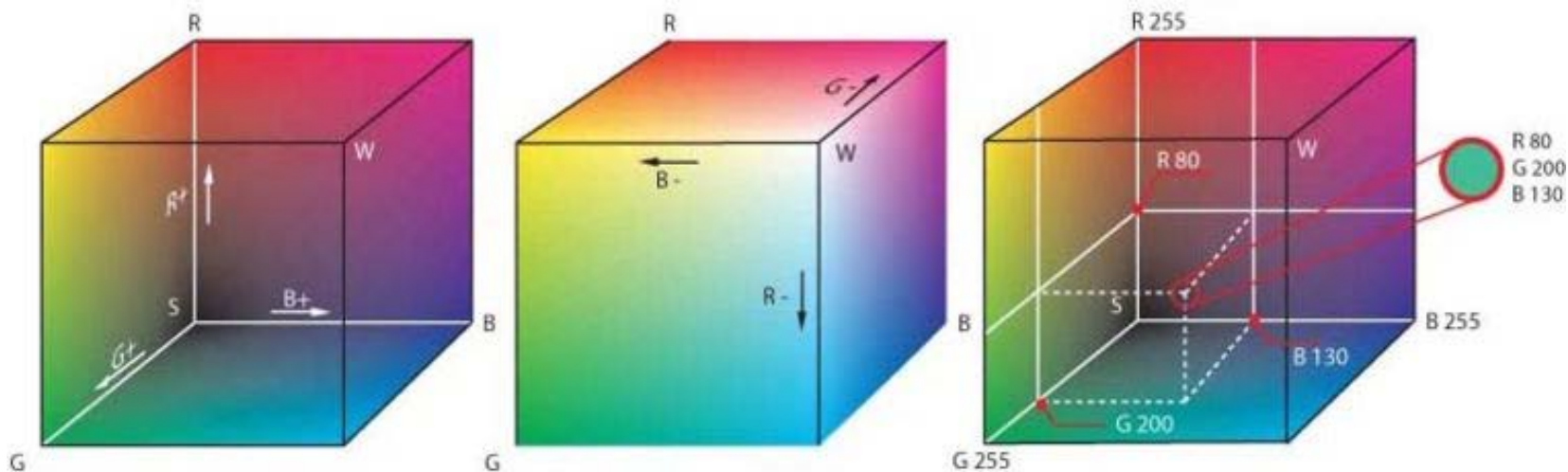
Typical gamuts of a monitor and of a printing device.

Цветовые пространства

- Назначение *цветового пространства (цветовой модели, системы цветов)* – стандартизация описания цвета
- Цветовая модель определяет *систему координат и подпространство* внутри этой системы, в котором каждый цвет представлен единственной точкой
- Распространенные цветовые пространства:
 - RGB (мониторы, видеокамеры),
 - CMY/CMYK (принтеры),
 - HSI/HSV/HSL/HSB (обработка изображений),
 - CIE Lab (обработка изображений).

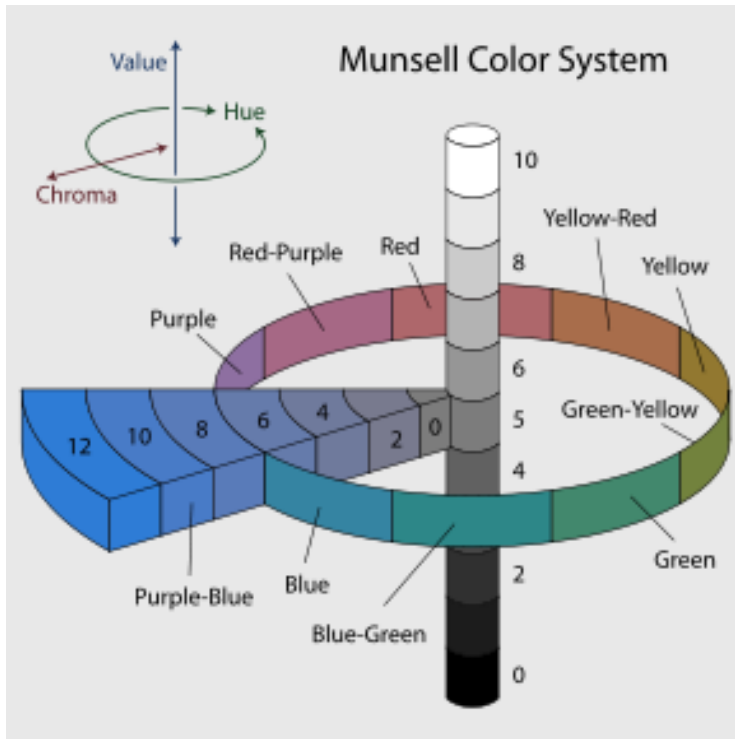


RGB



Если R, G, и B задаются 8 битами (24-битное RGB изображение), то число цветов $(2^8)^3 = 16,777,216$

Цветовая система Манселла



- Разработана профессором Альбертом Манселлом в начале 20 века
- Координаты: цветовой тон, светлота (значение), насыщенность (хрома).

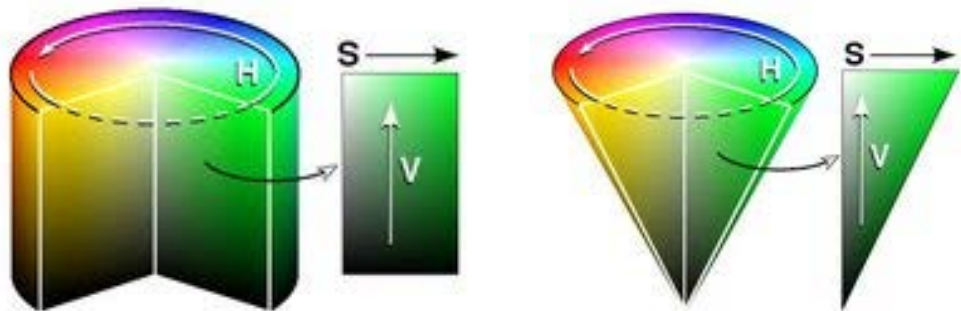


Цветовые пространства HSI/HSL/HSV/HSB

- RGB, CMY/CMYK ориентированы на устройства цветовоспроизведения (мониторы, принтеры)
- HSI/... (Hue, Saturation, Intensity/Lightness/Value/Brightness) соответствуют цветовосприятию человека
- Позволяют отделить цветность (H+S) от яркости (I).



Цветовые пространства HSI/HSL/HSV/HSB

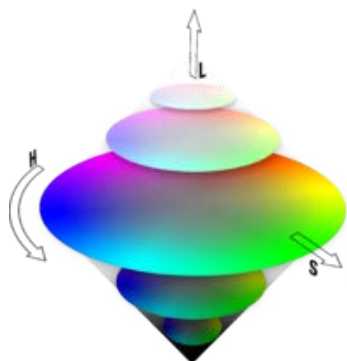


Graphical depiction of HSV (cylinder and cone)

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

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

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



Graphical depiction of HSL

<http://www.easyrgb.com/index.php?X=MATH>

Модель CIE L*a*b

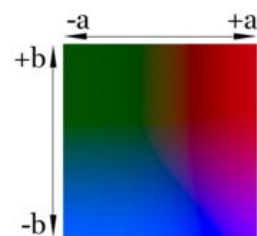
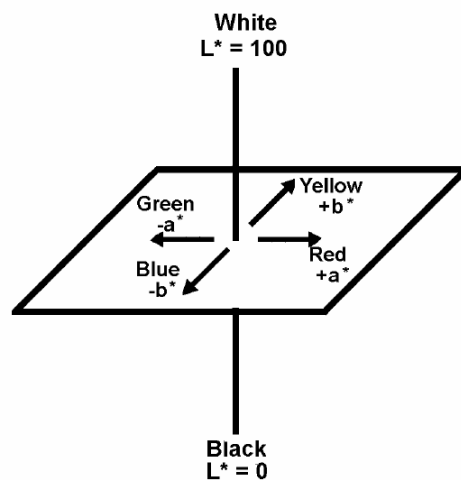
- Аппаратно-независимая и зрительно однородная цветовая модель.
- Позволяет соотносить цветовые охваты мониторов и принтеров
- Компоненты L*a*b* задаются так:

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

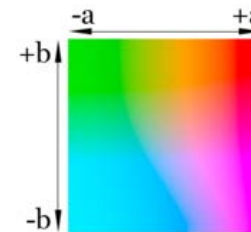
$$a^* = 500 \left[h \left(\frac{X}{X_w} \right) - h \left(\frac{Y}{Y_w} \right) \right]$$

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

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

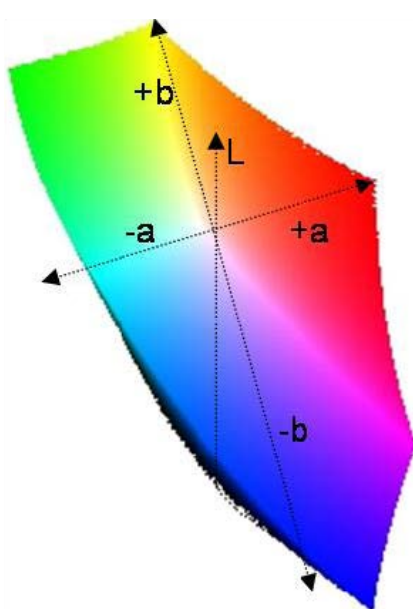


Lightness 75%

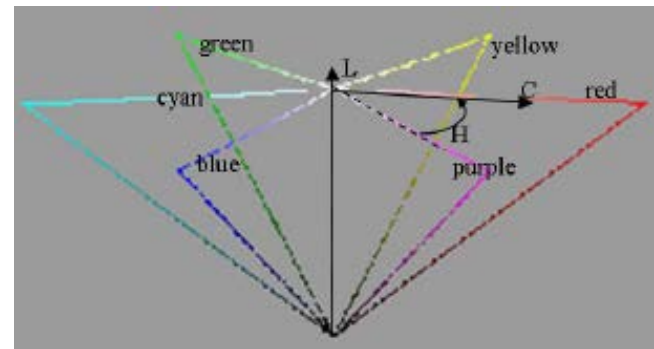
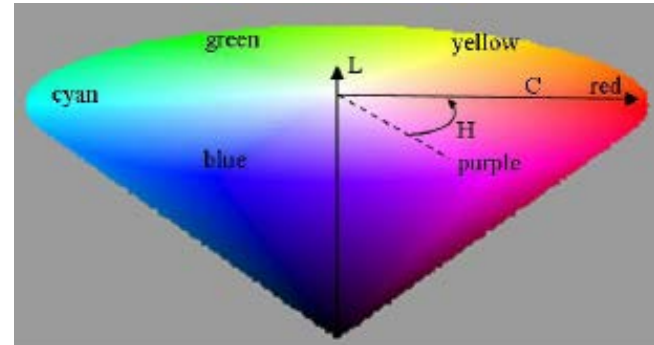
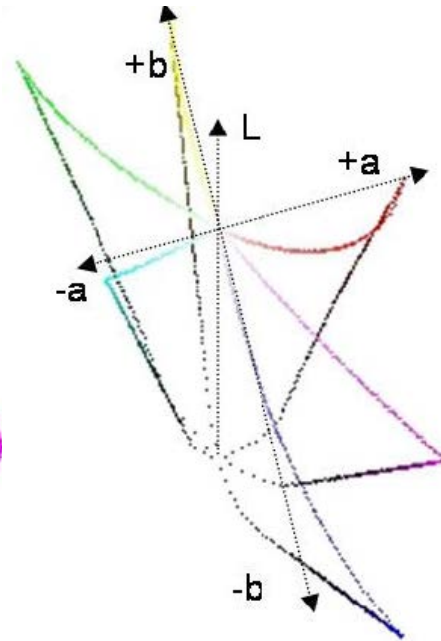


Lightness 25%

Пространство HCL



CIE Lab color space



HCL color space

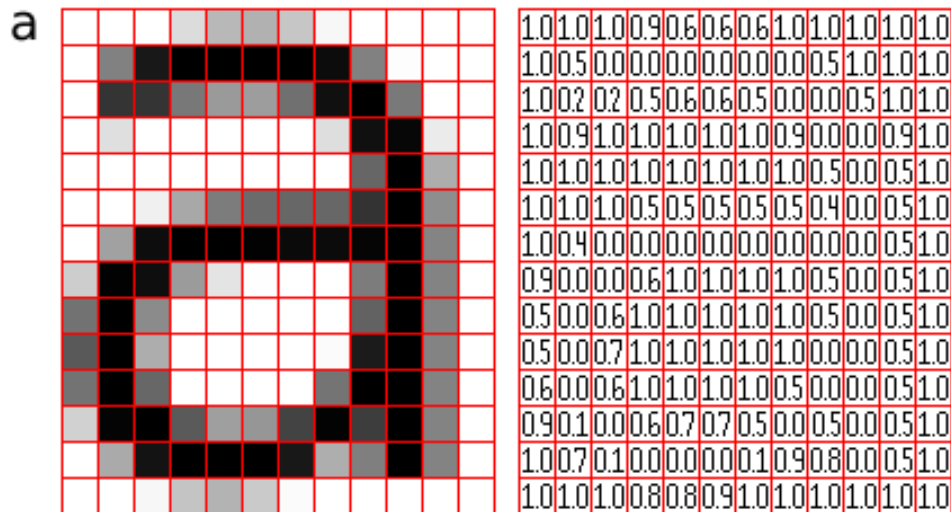
План лекции

- Структура курса
- Зачем нужен анализ изображений?
- Сложности
- Зрительное восприятие человека
- Представление цвета
- Цифровое представление изображений



Цифровое представление изображений

Растровое изображение



$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}$$

$0 \leq f(x,y) \leq L$, and typically $L = 255$

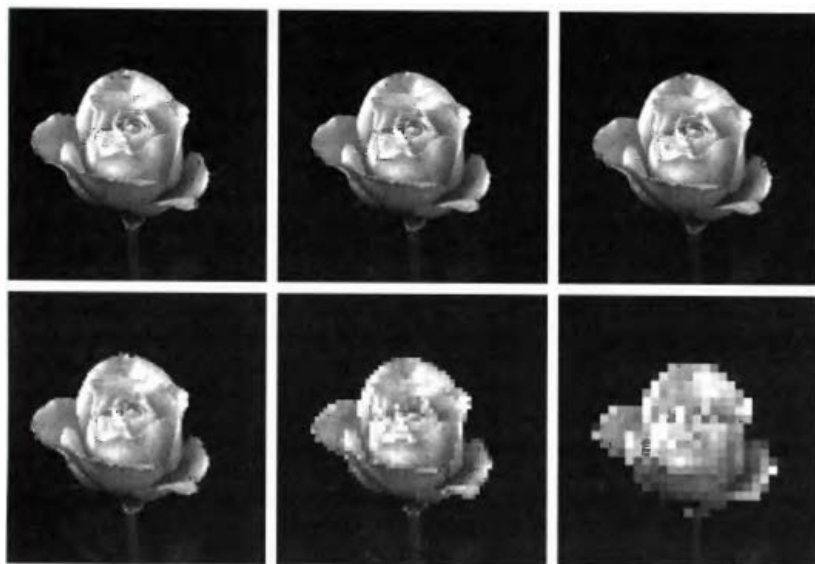


Параметры растрового изображения

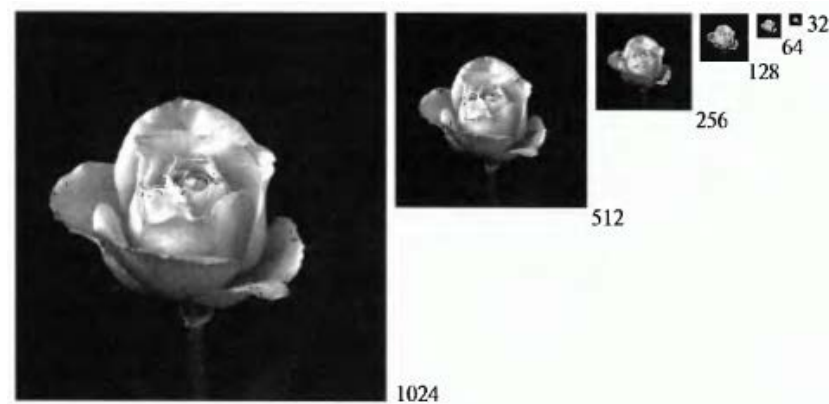
Размерность (Raster dimensions)

Разрешение (Resolution - ppi)

Число уровней (обычно 2^k)



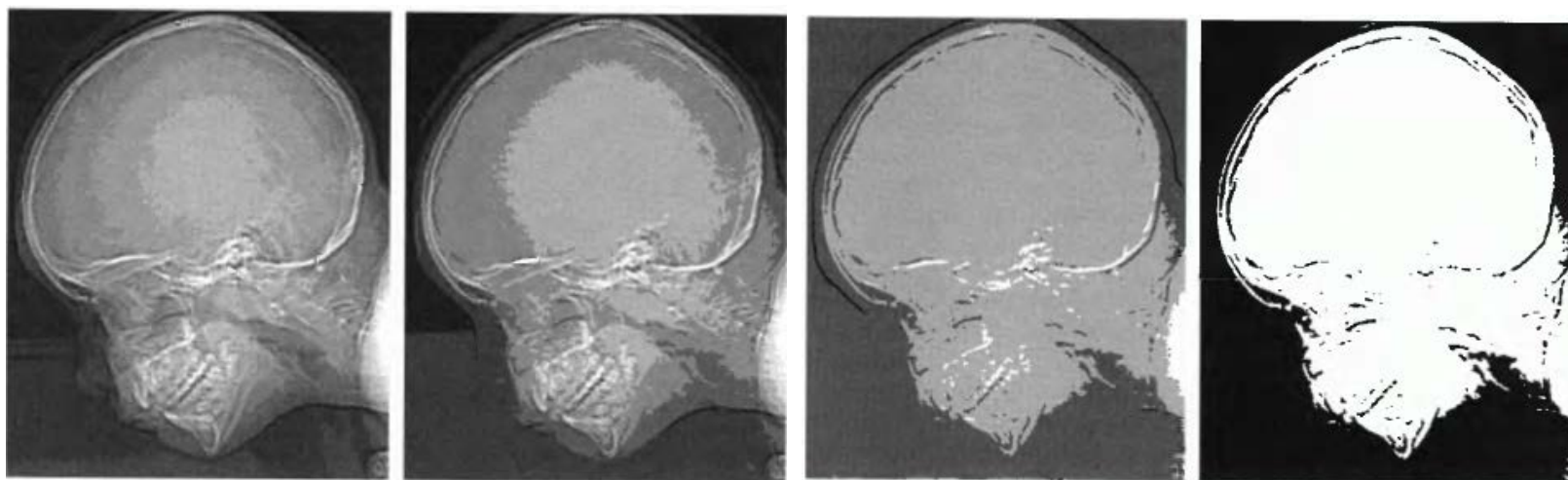
Фиксированная размерность,
изменяем разрешение



Фиксированное разрешение,
изменяем размерность

Параметры растрового изображения

Одно изображение с различным числом уровней:



16 levels

8 levels

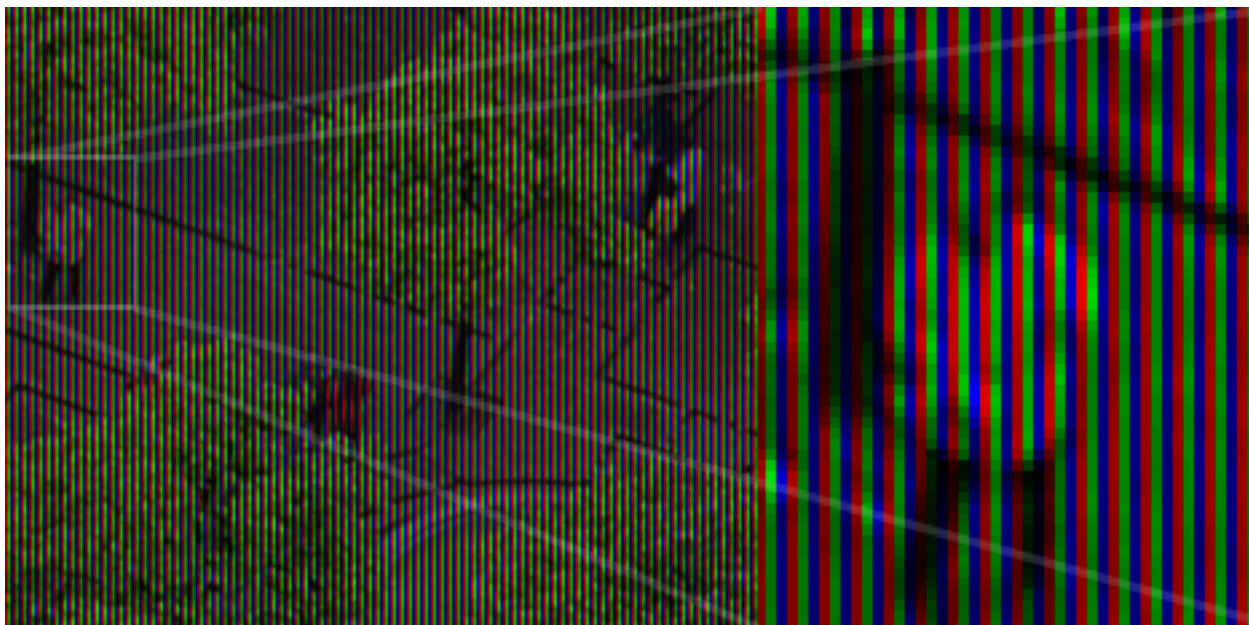
4 levels

2 levels

Распространено: 8 bit (256 уровней), 16 bit – png, tiff

Представление цифровых изображений

Цветное растровое изображение:



$$f(x, y) = \left(\begin{array}{ccc} r_{0,0}, g_{0,0}, b_{0,0} & r_{0,1}, g_{0,1}, b_{0,1} & \dots & r_{0,N-1}, g_{0,N-1}, b_{0,N-1} \\ r_{1,0}, g_{1,0}, b_{1,0} & r_{1,1}, g_{1,1}, b_{1,1} & \dots & r_{1,N-1}, g_{1,N-1}, b_{1,N-1} \\ \dots & \dots & \dots & \dots \\ r_{M-1,0}, g_{M-1,0}, b_{M-1,0} & r_{M-1,1}, g_{M-1,1}, b_{M-1,1} & \dots & r_{M-1,N-1}, g_{M-1,N-1}, b_{M-1,N-1} \end{array} \right)$$

Заключение

- Познакомились со структурой курса
 - Первые домашние задания – после следующей лекции
- Что такое и зачем нужен анализ изображений?
 - Медицина, промышленность, киноиндустрия, поиск и индексирование коллекций
- Основные сложности
 - Семантический разрыв, устойчивость к различиям в освещенности, позе, смене точки съемки, перекрыванию объектов (occlusion), внутривидовому разнообразию
- Зрительное восприятие человека
 - Основные факты
- Представление цвета
 - Рассмотрели различные цветовые пространства
- Цифровое представление изображений
 - Растровые изображений

