

Введение в технологии компьютерного зрения.

Преподаватель: Колкер А.Б.
a.kolker@corp.nstu.ru

Курс для магистров второго года обучения
факультет АВТ НГТУ, группы ААМ

Теоретический курс-содержание

- Введение в курс. Техническое зрение в современных информационных системах. Обзор технологий, решений, задач. Задача анализа изображения. Изображение как цифровой сигнал, особенности его анализа, получения, хранения. Алгоритмы сжатия данных, обратимые, необратимые преобразования.

Теоретический курс-содержание

Хранение данных изображений. Многомерные массивы. Потoki данных. Распространенные форматы хранения изображений. Цветовые пространства и особенности работы с ними. Форматы хранения видеопотоков.

- Высокотехнологичное оборудование и матобеспечение для решения задач компьютерного зрения (прочитает Д.С.Худяков)

Теоретический курс-содержание

- Методы получения потоковых данных изображений и видеосигналов с цифровых систем. Библиотеки низкого уровня. Библиотеки высокого уровня. Шина FireWire IEEE1934. Шина USB. Импорт данных с Ethernet камер.
- Библиотека OPENCV. Загрузка данных. Внутренний формат хранения файлов. Отображение данных встроенными средствами.

Теоретический курс-содержание

- OPENCV:
- Обработка изображения:
сглаживание, фильтрация,
масштабирование, морфология, альфа-
смешивание, пороги, преобразования:
свертка, Операторы Собеля, Лапласа.
Детектор Канни, фильтры Хаара. Поиск
объекта на изображении. Нахождение
контуров. Моменты.

Теоретический курс-содержание

Системы трехмерного зрения —
стереовидение.

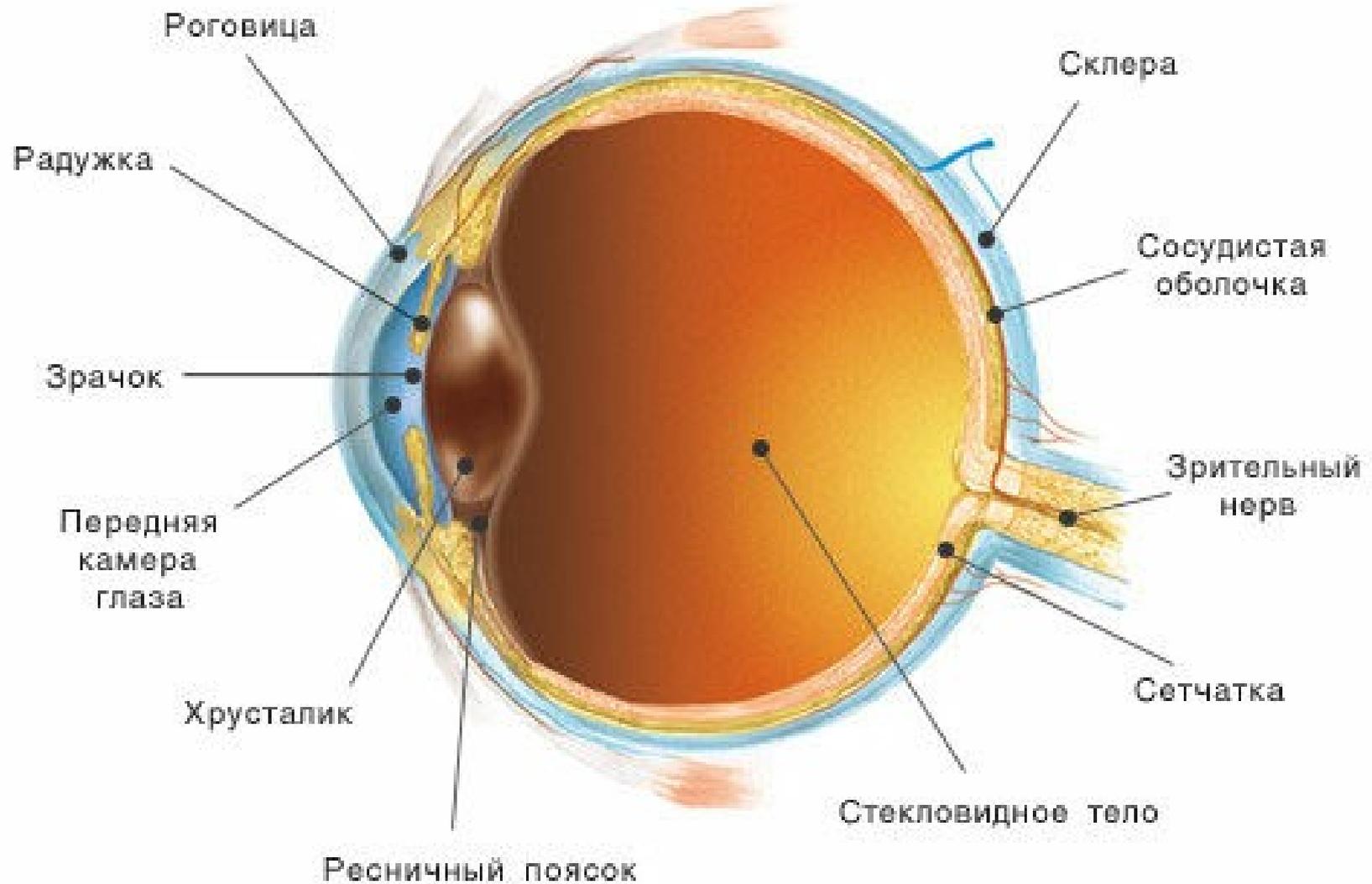
- Инструментарий OpenCV. Kinect. Freenect. OpenNI. Штриховые камеры. Работа с данными.
- Аппаратные средства трехмерного зрения.

Практический курс

- Изучение форматов изображений. Преобразование изображений. Цветовые пространства. Сжатие данных
- Работа с ImageMagick
- OpenCV - HelloWorld
- OpenCV - Работа с данными

- Индивидуальное задание.
Потребуется- а) OpenCV, web-камера.

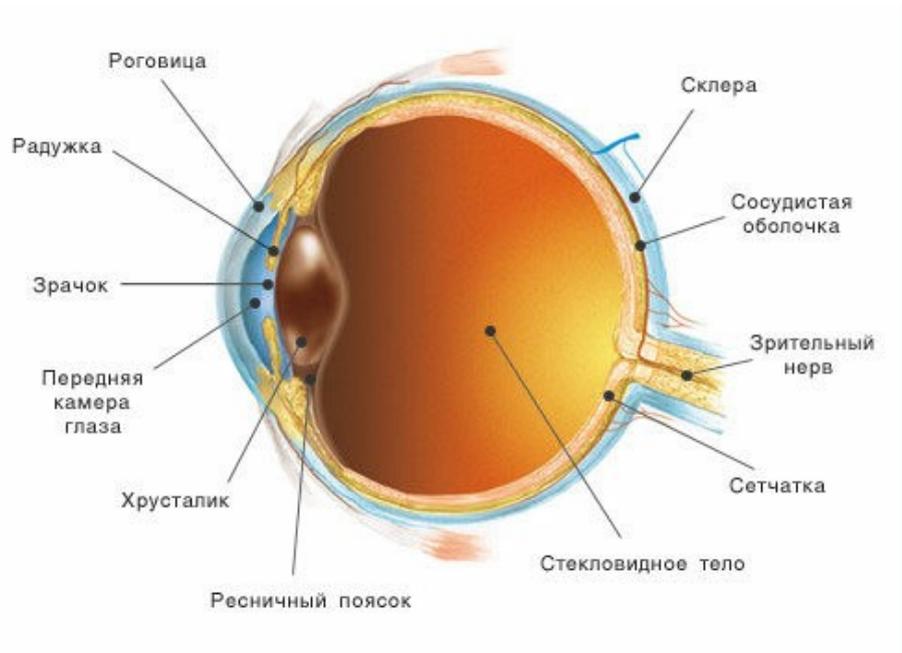
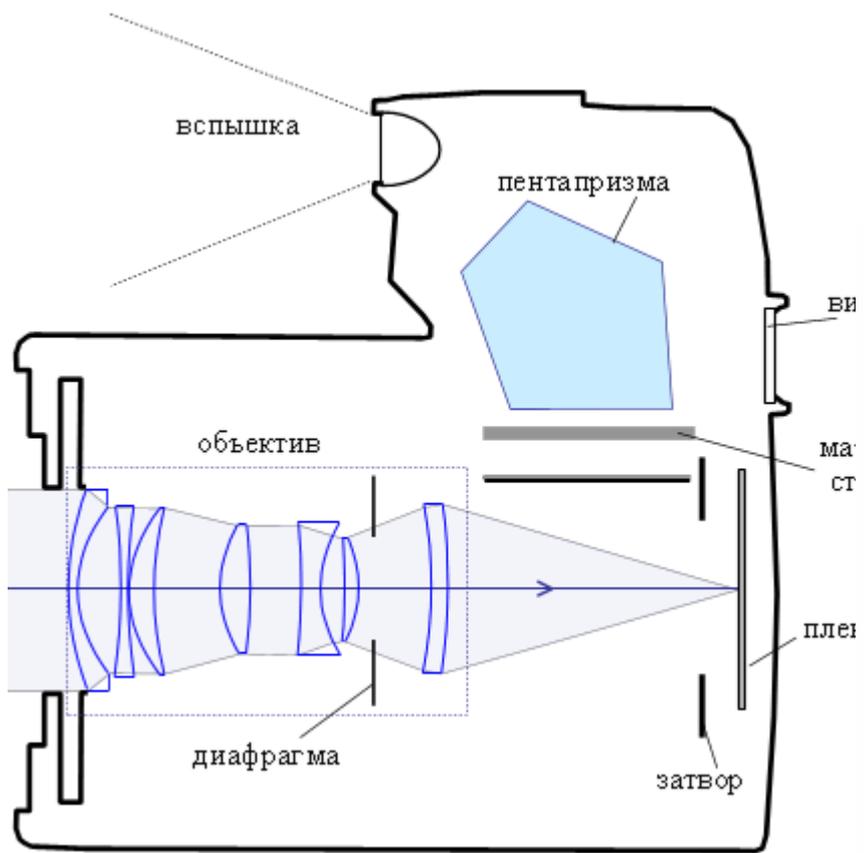
Строение человеческого глаза- Чуток анатомии



- Роговица — прозрачная оболочка, покрывающая переднюю часть глаза. В ней отсутствуют кровеносные сосуды, она имеет большую преломляющую силу. Входит в оптическую систему глаза. Роговица граничит с непрозрачной внешней оболочкой глаза — склерой. См. строение роговицы.
- Передняя камера глаза — это пространство между роговицей и радужкой. Она заполнена внутриглазной жидкостью.
- Радужка — по форме похожа на круг с отверстием внутри (зрачком). Радужка состоит из мышц, при сокращении и расслаблении которых размеры зрачка меняются.
- Зрачок — отверстие в радужке. Его размеры обычно зависят от уровня освещенности. Чем больше света, тем меньше зрачок.

- Хрусталик — "естественная линза" глаза. Он прозрачен, эластичен — может менять свою форму, почти мгновенно "наводя фокус", за счет чего человек видит хорошо и вблизи, и вдали. Располагается в капсуле, удерживается ресничным пояском. Хрусталик, как и роговица, входит в оптическую систему глаза. Прозрачность хрусталика глаза человека превосходна - пропускается большая часть света с длинами волн между 450 и 1400 нм.

- Стекловидное тело — гелеобразная прозрачная субстанция, расположенная в заднем отделе глаза. Поддерживает форму глазного яблока, участвует во внутриглазном обмене веществ. Входит в оптическую систему глаза.
- Склера — непрозрачная внешняя оболочка глазного яблока, переходящая в передней части глазного яблока в прозрачную роговицу. К склере крепятся 6 глазодвигательных мышц.
- Зрительный нерв — при помощи зрительного нерва сигналы от нервных окончаний передаются в головной мозг.



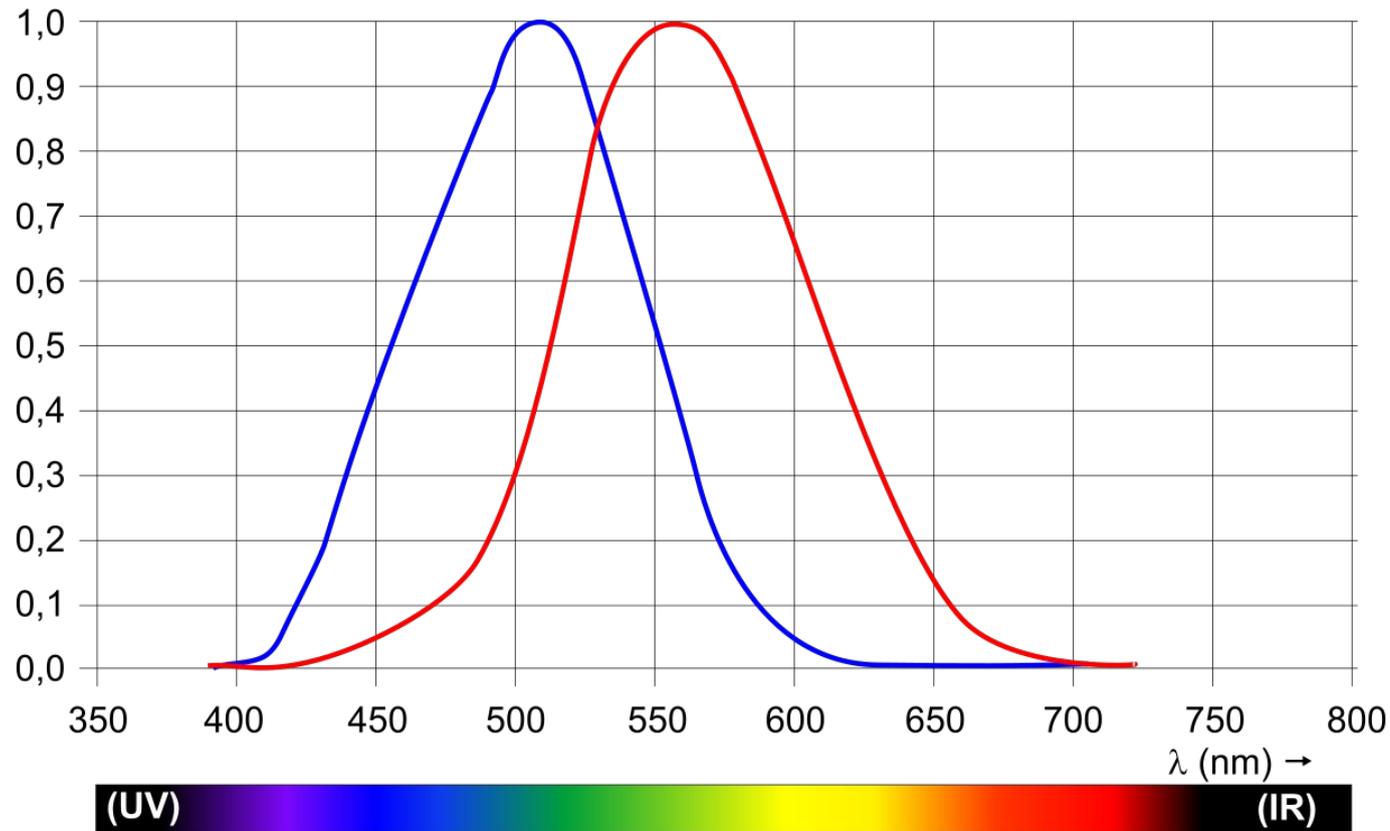
Световая чувствительность глаза-колбочки

- Максимальная световая чувствительность достигается после достаточно длительной темновой адаптации
- В человеческом глазе расположено три вида «колбочек» - светочувствительных элементов (6-7 млн). В фигурных скобках максимум светочувствительности
- S 400-500 нм (420-440 нм) ~ В
- M 450-639 нм (534-555 нм) ~ G
- L 500-700 нм (564-580 нм) ~ R

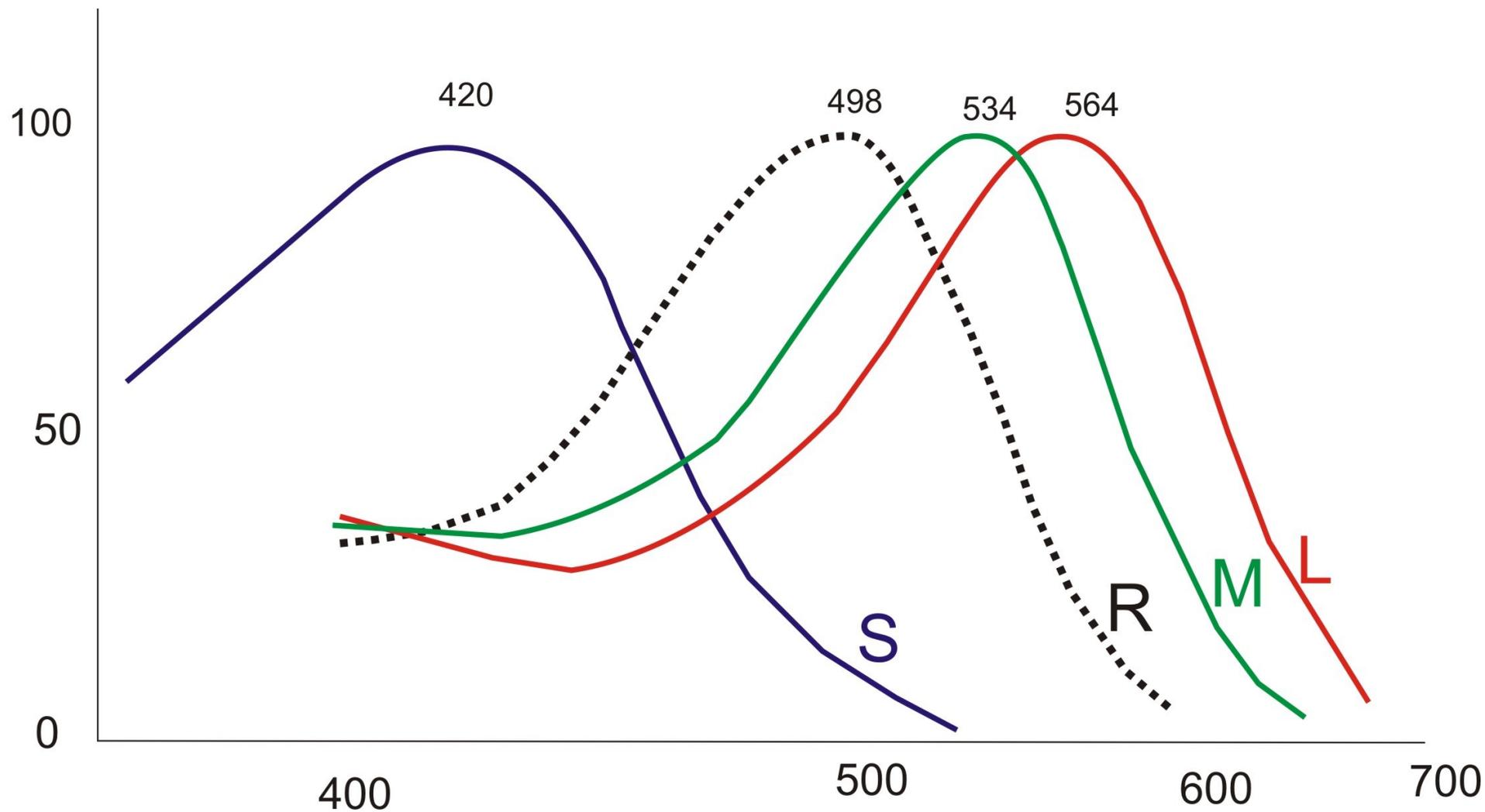
Ночное зрение-палочки

- Приблизительно 120 млн палочек (0.06x0.002мм)
- Высокочувствительные клетки. Чувствительность достаточна, чтобы зарегистрировать 2-3 фотона.

Сравнение цветного(красная) и чб(синяя) зрения человека

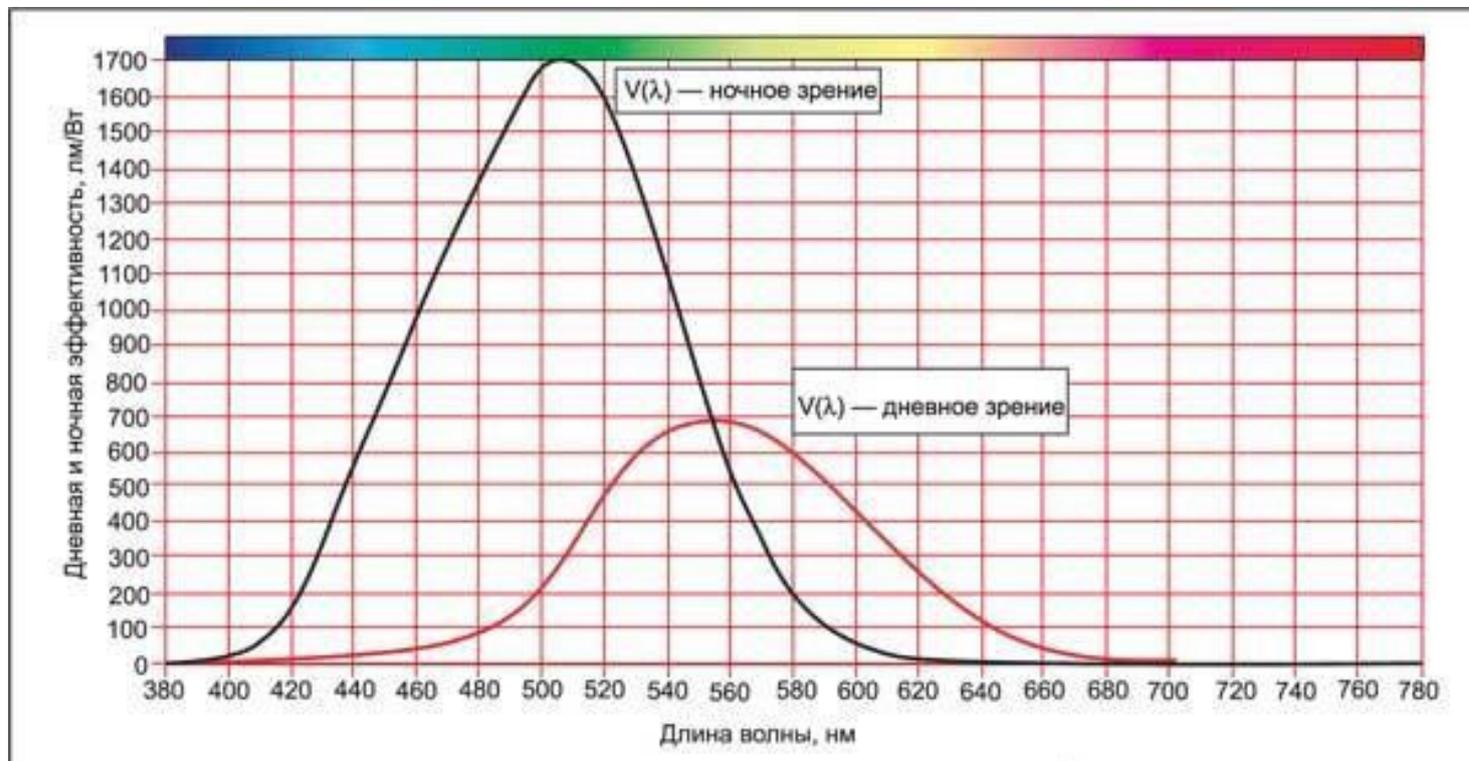


Графики цветовой чувствительности



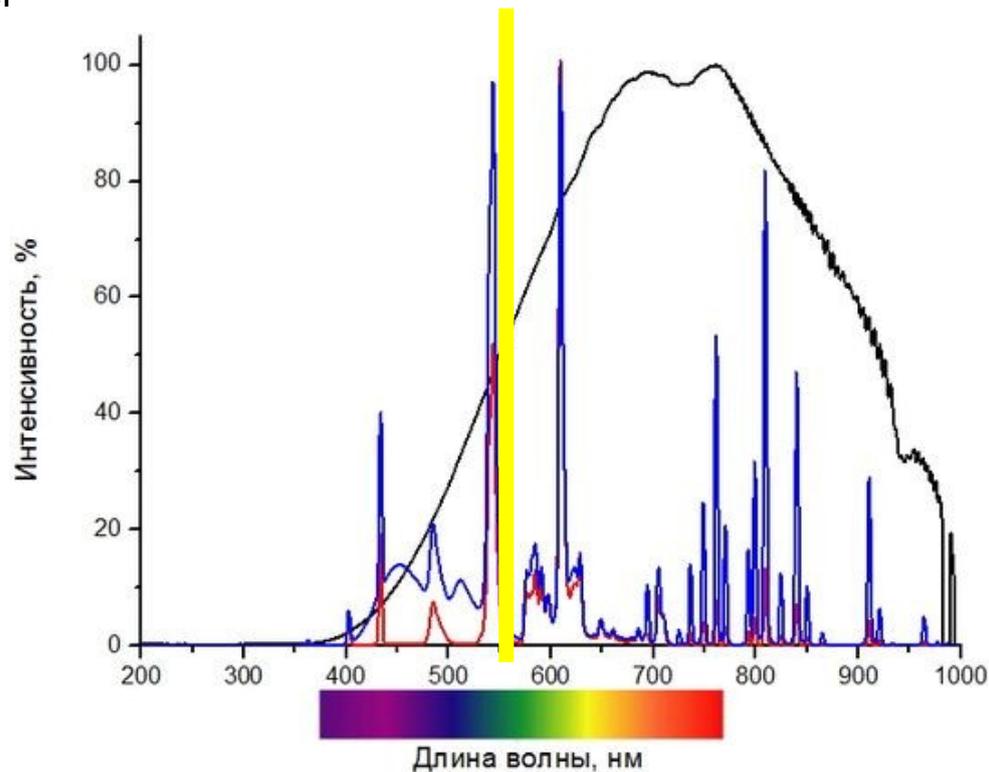
Почему так получилось?

Естественный источник излучения — солнце.
Источник свечения- газ, разогретый до температуры 5800К. Спектральная чувствительность глаза за эволюцию подстраивалась под эту ситуацию.

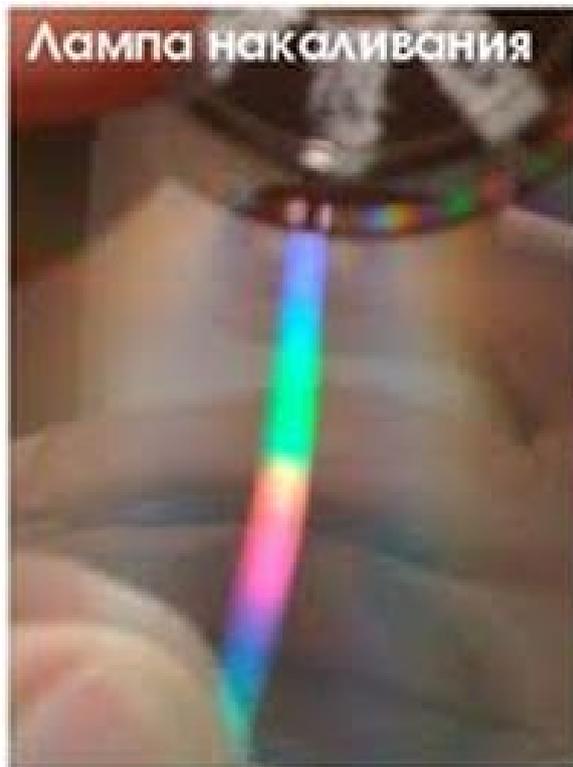


Почему глаза устают от «энергосберегающих» ламп?

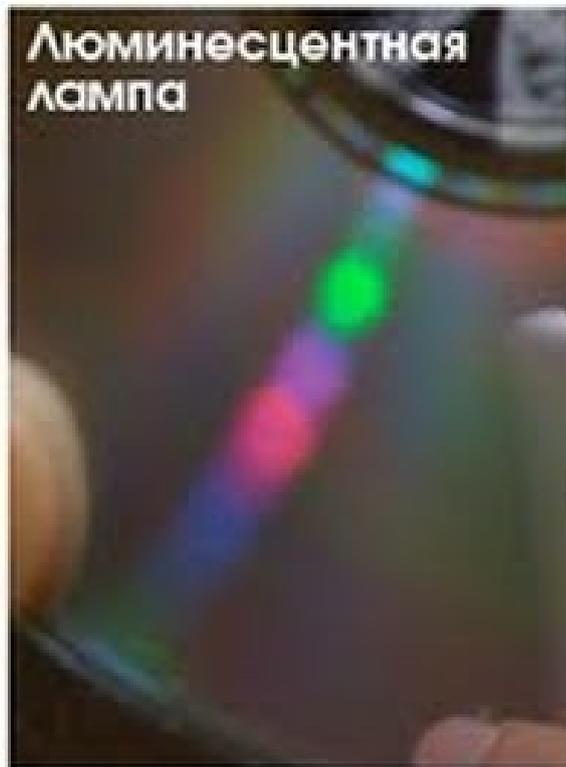
- Спектр двух разных энергосберегающих ламп и одной накапивания (черная линия). Местоположе



Лампа накаливания



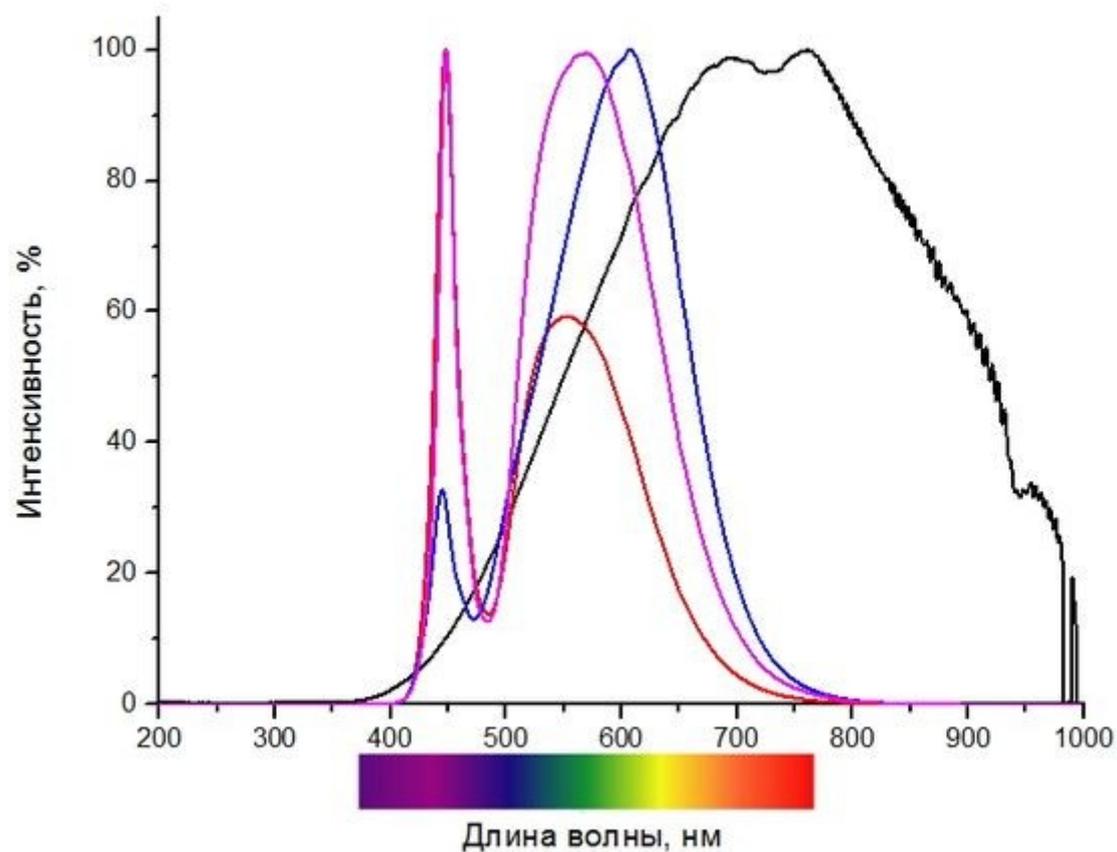
Люминесцентная
лампа



Люминесцентная
лампа



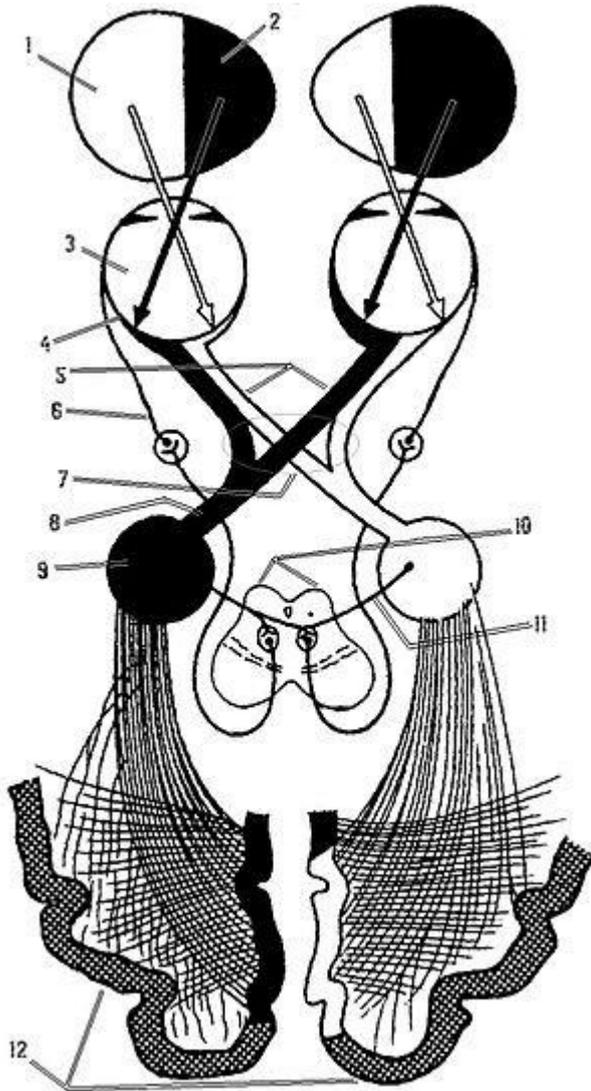
а Спектр трех разных СДЛ и одной ЛН (черная линия). Пики излучения смещаются, что свидетельствует о разном составе люминофора



Цветовая температура

Интегральная характеристика спектра любого источника света — его цветовая температура (ЦТ). Она определяется как температура абсолютно черного тела, при которой оно испускает излучение того же цветового тона, что и рассматриваемый ИС.

Бинокулярное (стереоскопическое зрение)

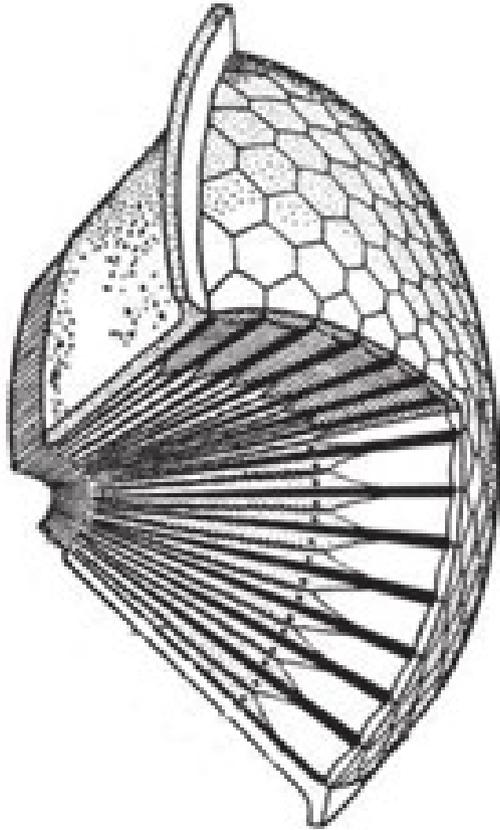


- Проводящие пути зрительного анализатора
- 1 — Левая половина зрительного поля, 2 — Правая половина зрительного поля, 3 — Глаз, 4 — Сетчатка, 5 — Зрительные нервы, 6 — Глазодвигательный нерв, 7 — Хиазма, 8 — Зрительный тракт, 9 — Латеральное коленчатое тело, 10 — Верхние бугры четверохолмия, 11 — Неспецифический зрительный путь, 12 — Зрительная кора головного мозга.[2]

Ощущение глубины

- Оценка расстояния до объекта опирается **не только** на данные бинокулярного зрения. В наивысшей степени способность оценивать расстояние до объекта и воспринимать картину мира принадлежит именно мозгу-центральному процессору. (если закрыть один глаз, то мозг включит «резервные механизмы» и вторичные признаки- данные аккомодации, размеры,перспектива итд.

Фасеточное зрение



- Система без «движущихся частей»
- Обеспечивает круговой обзор
- Стрекоза — 30000 омматидий
(Человек 6-7 млн колбочек+)

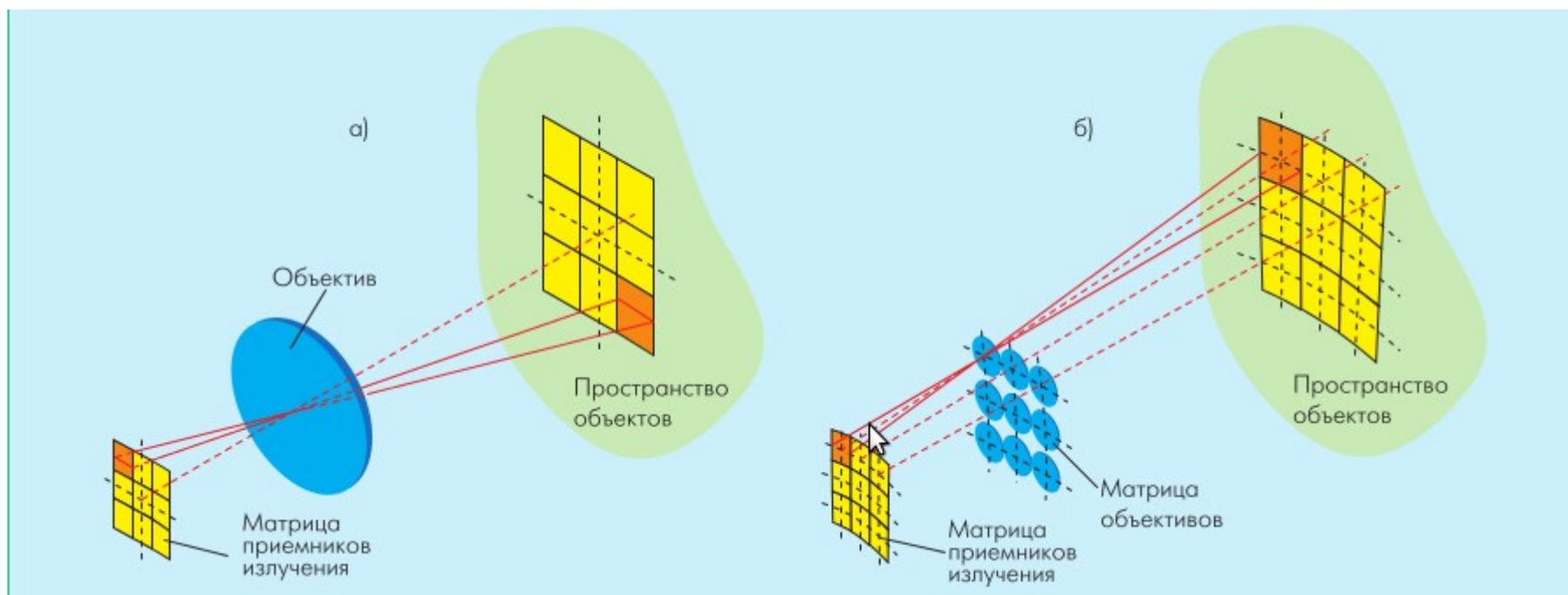
Свойства фасеточного глаза

- Большое угловое поле, малые габариты
- Быстрая реакция на движение (в 15 раз быстрее глаза человека)
- Большая «глубина резкости»
- Чувствительность к поляризации света
- Простота механизмов стереовидения

Но

- Низкое пространственное разрешение, обусловленное числом «линз»
-

Фокальная и фасеточная системы



Два метода создания систем технического зрения

- 1. Копирование «природы»: механизмов зрения человеком, животными, насекомыми.
- 2. Создание систем по своей сути отличных, от того что создала эволюция.



Пример копирования- фотокамера

- Объектив = хрусталик
- Радужка+зрачок = диафрагма
- Механизм изменения фокусного расстояния = хрусталик+ресничный поясок.

3D без биноккулярной системы = Kinect

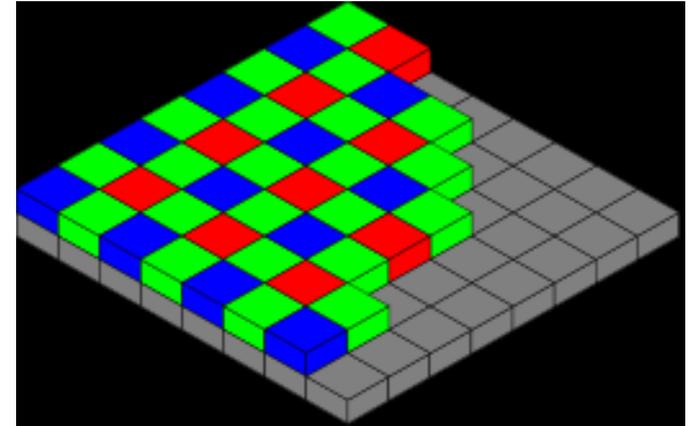


- IR подсветка с патентованной мозаичной структурой
- IR камера
- RGB камера
- процессор

Конструкция матрицы, маска Байера

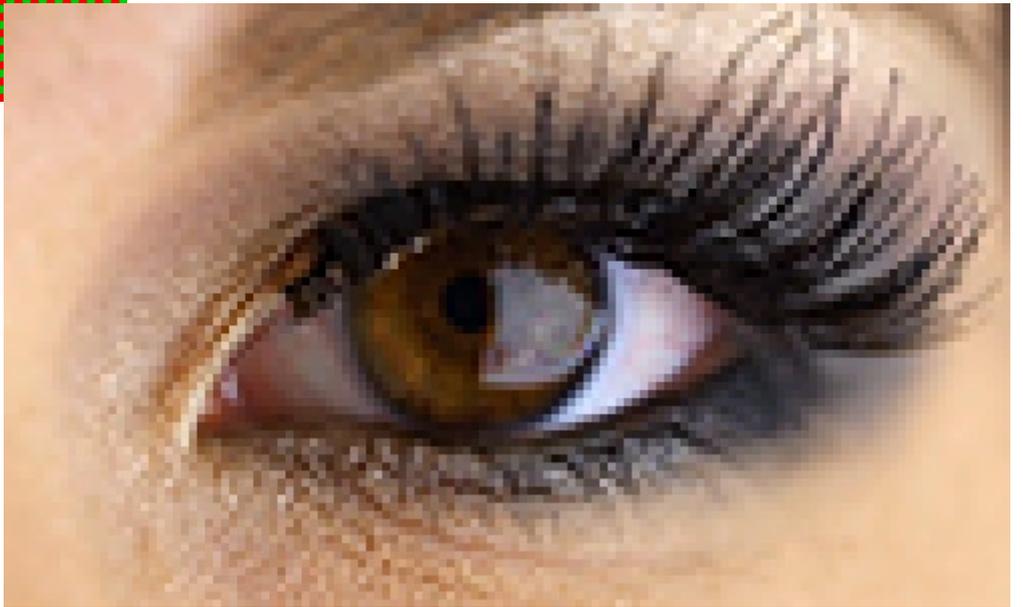
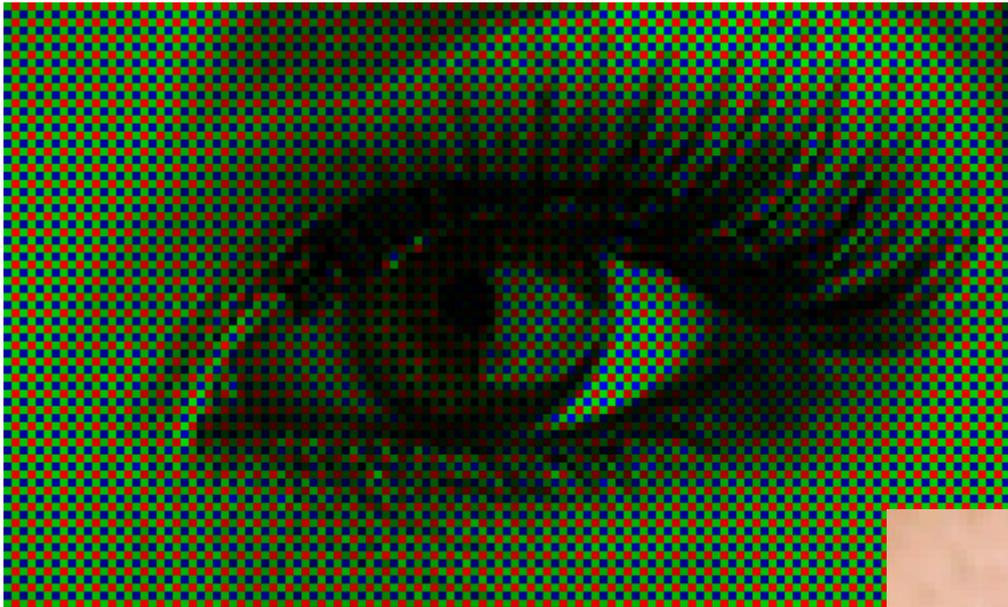
- Электронная матрица-светочувствительный элемент обычно содержит «ячейки» чувствительные к широкому диапазону волн. Для восприятия цвета камера покрывается маской. Чаще всего используется шаблон Байера (патент Kodak 1976).

- 25%R
- 25%B
- 50%G

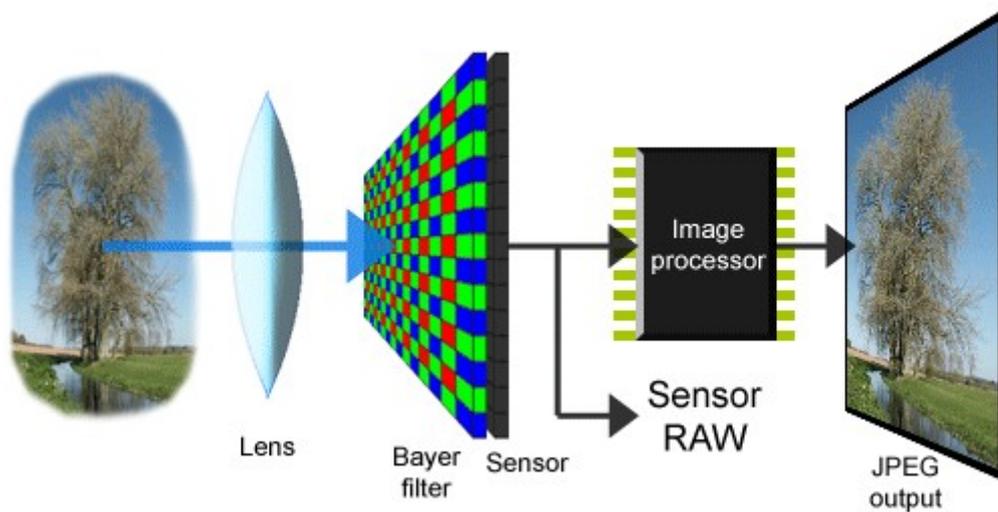


-
-
- Для восполнения отсутствующих цветовых значений используются интерполированные данные соседних ячеек. Для формирования данных 1 пиксела используется 9 или более соседних ячеек.

-
- GR
- BG - классическая маска Баера



Побочный эффект от преобразования Байера-размытость



Алгоритмика

Восстановление цветовой информации сложная задача, обычно объявляемая «нохау» производителей. Однако, большинство простых устройств используют алгоритмы, реализованные в библиотеке dscraw

Искажения

- ЧБ камера будет иметь лучшую чувствительность и четкость «при прочих равных»
- Алгоритм восстановления данных размывает изображение
- Требуются дополнительные ресурсы GPU камеры для обработки данных

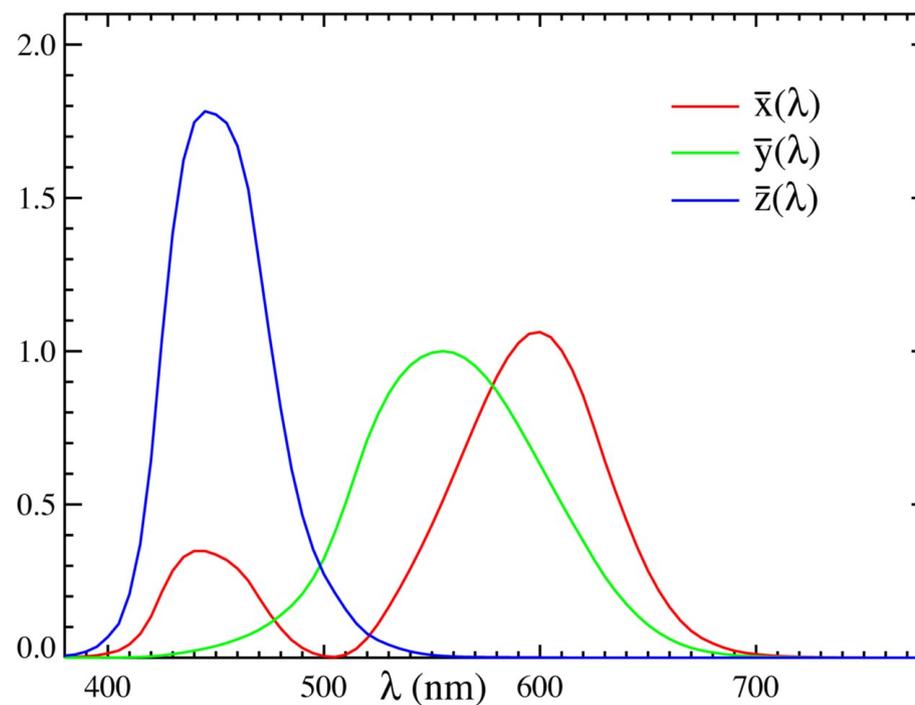
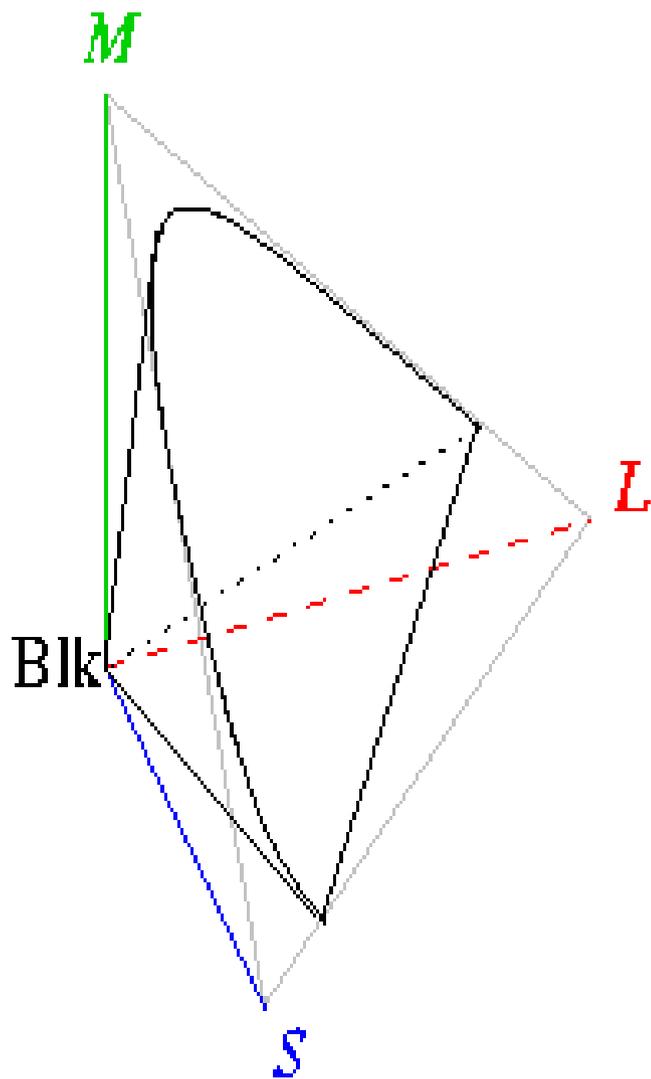
RAW режим

Некоторые промышленные камеры (RGB) позволяют выключить алгоритмы восстановления Байера и передавать «сырые» данные.

Цветовое пространство

- Цветовое пространство- модель представления цвета в многомерной системе координат.
- Цветовое зрение человека неортогонально (диапазоны перекрываются), сл-но цвет с разными спектрами может вызвать один и тот же цветовой стимул (метаметрия)

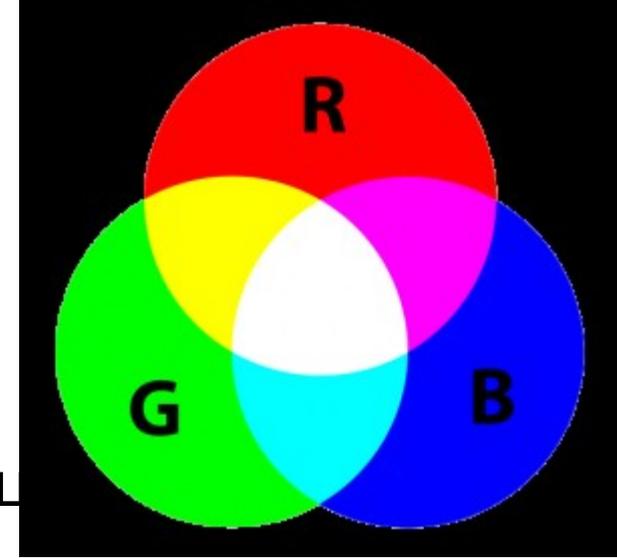
Цветовое пространство XYZ



Особенности восприятия цветов ЧЕЛОВЕКОМ

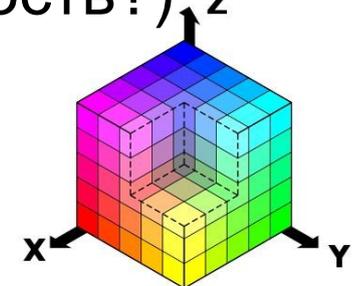


RGB



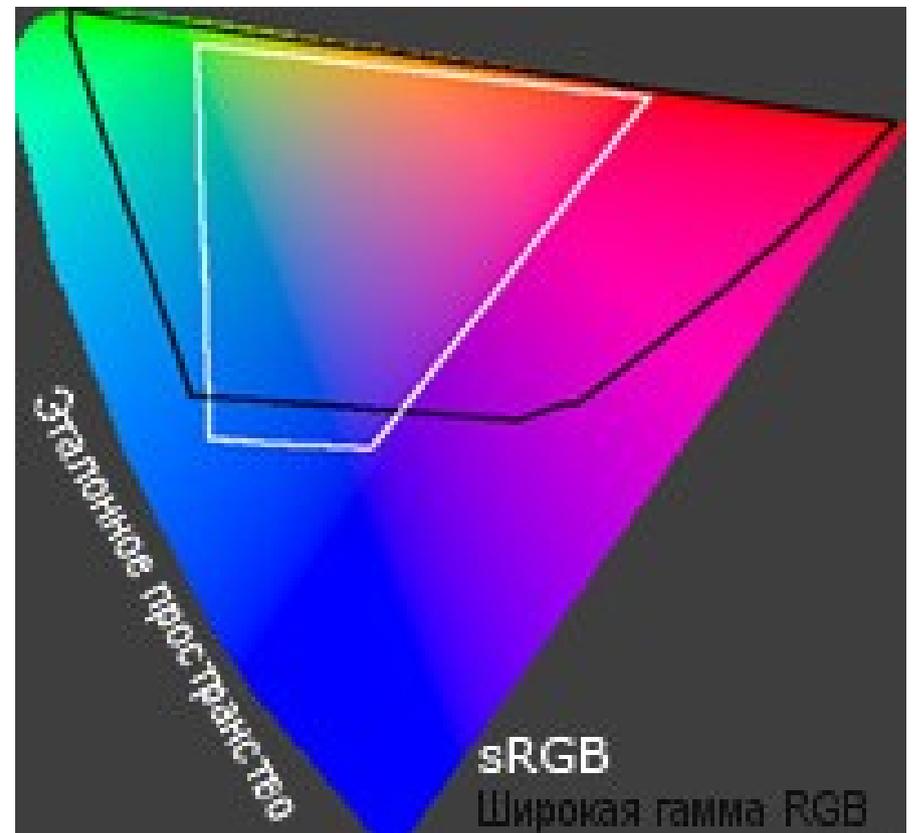
- Red Green Blue — аддитивная цветовая модель, обеспечивающая синтез света для цветного воспроизведения и цветоописания.
- Используется в мониторах, телевизорах. Наиболее простая технически реализуемая цветовая схема.
- Как правило, применяется «глубина цвета» 8 бит на канал или 24 бит, однако глубина может достигать и 16 бит на канал (48 бит на точку)

Проблематика. #000000 — черный. #FFFFFF — белый. Но что есть белый и насколько он белый (какая яркость?)



sRGB

- Вводится понятие «цветовая температура». Предложено HP и Microsoft — оптимизирована под типовые возможности мониторов

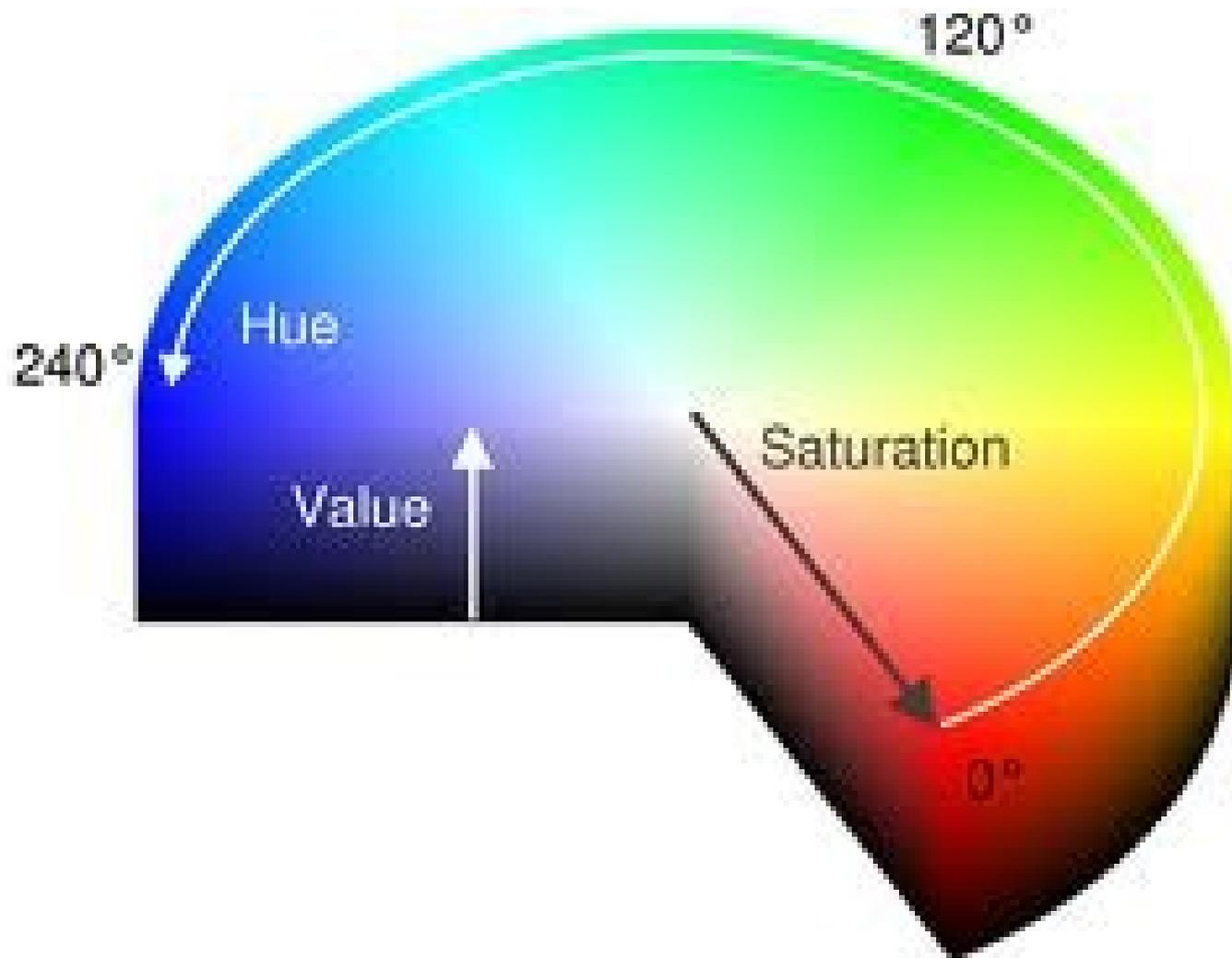


HSV

- Hue (тон) Saturation(насыщенность) Value — значение. Нелинейная цветовая модель, предложенная Элви Реем Смитом, одним из основателей Pixar в 1978 году. Представляет собой нелинейное преобразование RGB.
- Hue- цветовой тон (оттенок) диапазон 0-360 или иногда 0-1
- Saturation Насыщенность 0-100 (0-1) насколько близок тот или иной цвет к нейтральному серому.
- Value — яркость. 0-100 или 0-1.

HSV модель крайне удобна для использования в системах технического зрения, т.к. Существенно менее чувствительна к освещенности.

HSV



YUV

3 компоненты: яркость и две цветоразностных

$$R = Y + 1.13983 * (V - 128);$$

$$G = Y - 0.39465 * (U - 128) - 0.58060 * (V - 128);$$

$$B = Y + 2.03211 * (U - 128)$$

$$Y = 0.299 * R + 0.587 * G + 0.114 * B;$$

$$U = -0.14713 * R - 0.28886 * G + 0.436 * B + 128;$$

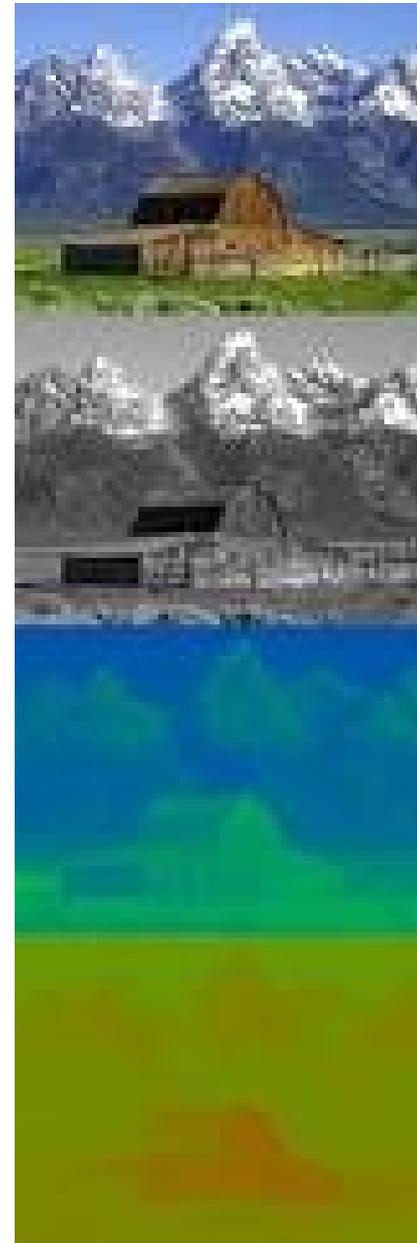
$$V = 0.615 * R - 0.51499 * G - 0.10001 * B + 128;$$

Другие цветовые модели

- CMYK (Cyan Magenta Yellow Black) — полиграфия, офсет. Нам не интересна.
- Lab Cie L*a*b линейная, привязанная к цветоощущению человека. Популярна в фотографии.

UYV

Стержневая идея: цветовые компоненты областей изображения с минимальной яркостью могут быть опущены.



Другие цветовые модели

- CMYK (Cyan Magenta Yellow Black) — полиграфия, офсет. Нам не интересна.
- Lab Cie L*a*b линейная, привязанная к цветоощущению человека. Популярна в фотографии.
- YUV линейная цветоразностная модель
$$Y = 0.299 * R + 0.587 * G + 0.114 * B; \rightarrow \text{яркость}$$
$$U = -0.14713 * R - 0.28886 * G + 0.436 * B + 128; \rightarrow \text{цветоразность}$$
$$V = 0.615 * R - 0.51499 * G - 0.10001 * B + 128;$$
 - используется в аналоговом телевидении, где важно иметь возможность транслировать чб сигнал.
 -
 -
 -

Lenna story

<http://www.cs.cmu.edu/~chuck/lennapg/>



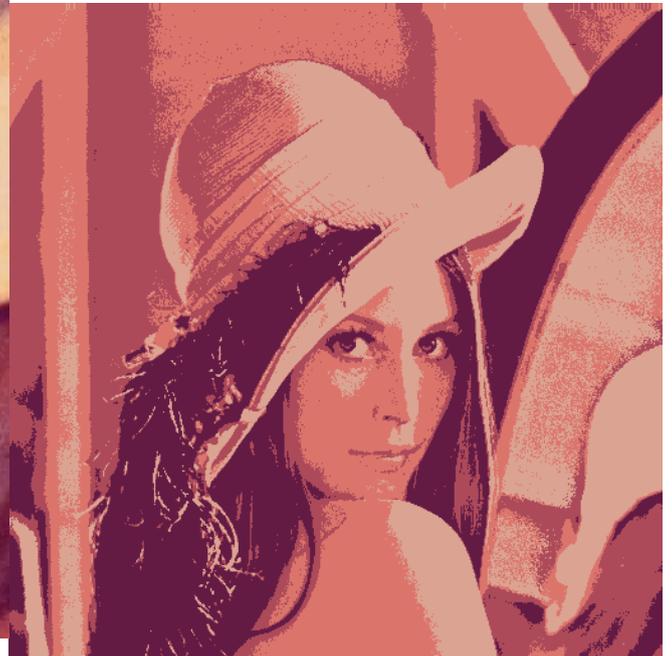
Lenna (Lina) — Оцифрованный портрет шведской модели Лины Седерберг — фрагмент из журнала Playboy за ноябрь 1972 года. В 1973 году Александру Савчуку (Университет Северной Калифорнии) понадобился фотопортрет с хорошим динамическим диапазоном. Не долго думая, он отсканировал постер из журнала Playboy, который попался под руку. Вскоре эта картинка превратилась в стандарт для тестирования алгоритмов обработки изображений и известно всякому, кто работает в области Image Processing. Есть и стандарты размеров: 512x512 или 720x576

http://www.tecnick.com/public/code/cp_dpage.php
- база тестовых изображений

Форматы хранения графических данных. BitMap (BMP)

- Растровый формат, разработка Microsoft. Глубина цвета 1,4,8,18,24,32,48 бит.
- Если глубина цвета меньше 16 бит, хранится палитра RGB цветов (индексная таблица), а в поле данных указывается номер индекса.
- Хранит только один слой
Обычно bmp — несжатый массив данных.

Лепна , глубина цвета 8, 4 бит, 2



GIF

- Graphics Interchange Format

Может отображать не более 256 оттенков цветов (размер индексной палитры жестко фиксирован)

Не уменьшает детальность изображения (LZW сжатие) — без потерь.

- Стандарт GIF89a модифицирован с учетом прозрачности и анимации.
- Хранит только один слой.

Патенты истекли 7 July 2004.

Gif True color

- Формально Gif хранит только 256 цветов, однако можно разбить изображение на кадры, используя заложенные технологии анимации и в каждом из них будет своя таблица.

TIFF

- Tagged Image Format — формат хранения растровых данных с большой глубиной цвета. Первоначально разработан Aldus Corporation(в дальнейшем слилась с Adobe) в кооперации с Microsoft.
- возможность применения сжатия без потерь
- возможность применения сжатия с потерями
- поддерживает разнообразные палитры
- Хранит несколько слоев изображения
- Поддерживает привязку к другим информационным слоям, например гео-привязку.

Форматы хранения графических данных. BitMap (BMP)

- Растровый формат, разработка Microsoft. Глубина цвета 1,4,8,18,24,32,48 бит.
- Если глубина цвета меньше 16 бит, хранится палитра RGB цветов (индексная таблица), а в поле данных указывается номер индекса.

Обычно bmp — несжатый массив данных.

Сжатие с потерями

- Сжатие без потерь: преобразования обратимы.
- Сжатие с потерями а) цветовая информация- постеризация
- Сжатие с потерями б) искажающее сжатие
главная задача сжатия с потерями-
искажение данных несущественны с точки
зрения их дальнейшего применения.

Jpeg

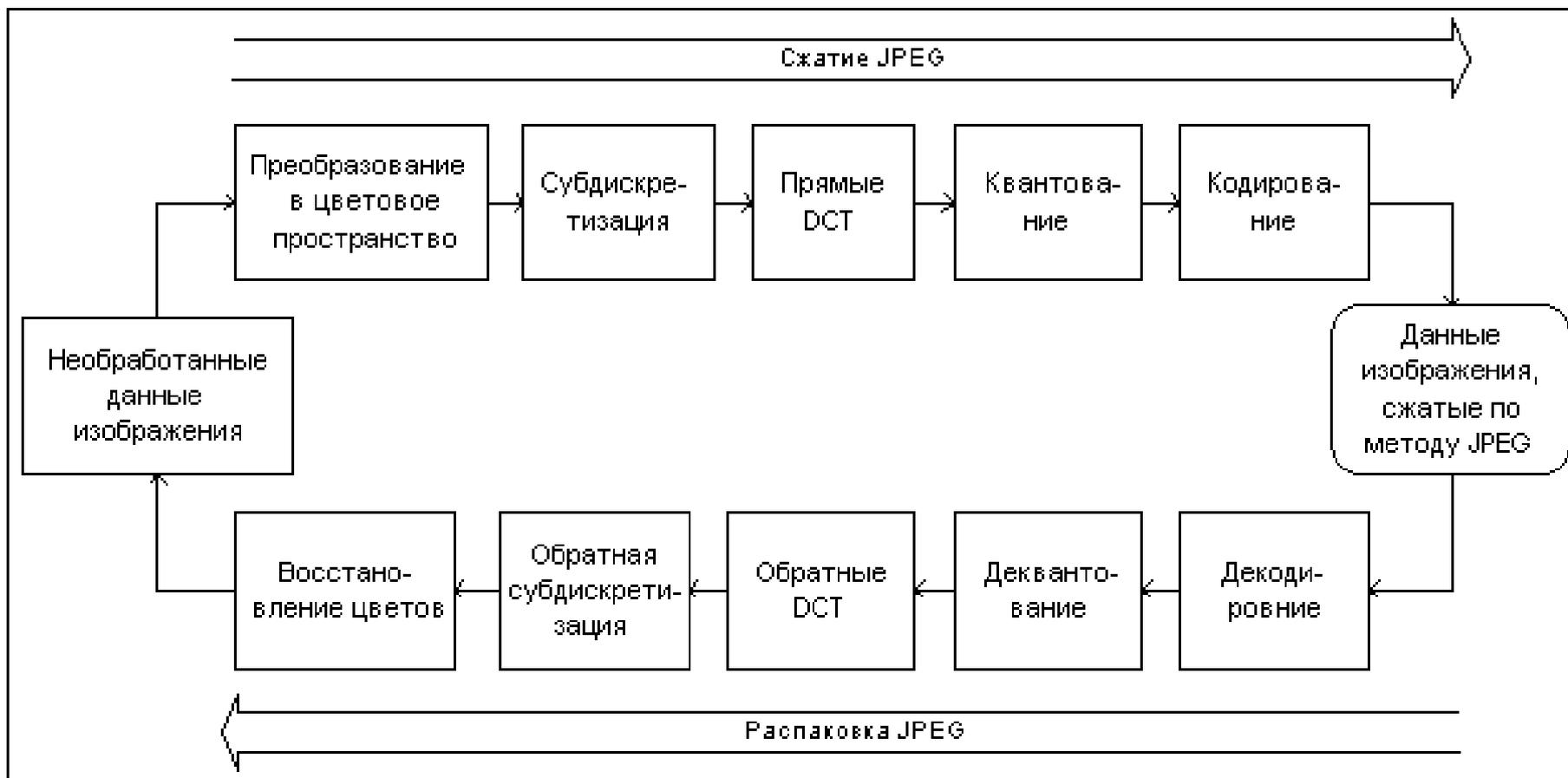
- Joint Photographics Experts Group — популярный графический формат сжатия изображений (в основном фотографических данных). Использует необратимое сжатие с потерями(возможен режим без потерь). Макс размер 65535x65535 точек

Сжатие с потерями

- Сжатие без потерь: преобразования обратимы.
- Сжатие с потерями а) цветовая информация- постеризация
- Сжатие с потерями б) искажающее сжатие
главная задача сжатия с потерями-
искажение данных несущественны с точки
зрения их дальнейшего применения.

Јрег- алгоритмика

1. Преобразование в YUV модель
2. Прореживание каналов Cb и Cr 2x2
3. Яркостный компонент — дискретное косинусное преобразование блоков 8x8
4. Полученные коэффициенты квантуются и пакуются кодом Хаффмана. Высокочастотные коэффициенты квантуются более сильно, чем низкочастотные, что приводит к «смазыванию мелких деталей» и появлению «артефактов».

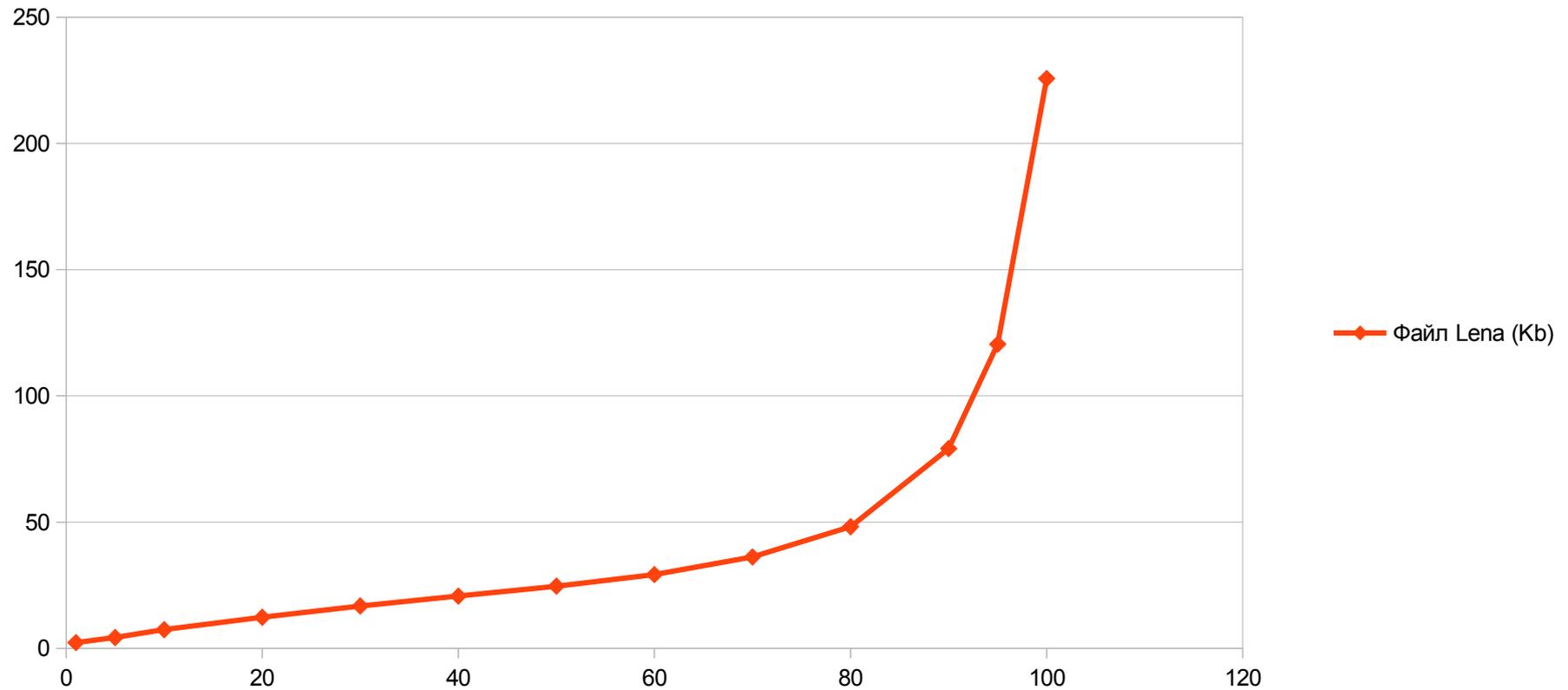


Фактор сжатия



Фактор 5- 50

Зависимость размера файла от коэф. качества



Дрег и постобработка

- Артефакты
- Потеря мелких деталей
- Уменьшение цветового охвата
- Размывание фронтов
 - > для решения задачи цифровой обработки сигналов предпочтительнее применять камеры, способные передавать данные в RAW(сыром) формате.

Типы устройств(некоторые)

- USB камера



- Ethernet камера



- IEEE1394 FireWire



USB камера

- Для устройств USB 2.0 регламентировано три режима работы:
- Low-speed, 10—1500 Кбит/с (используется для интерактивных устройств: клавиатуры, мыши, джойстики)
- Full-speed, 0,5—12 Мбит/с (аудио-, видеоустройства)
- Hi-speed, 25—480 Мбит/с (видеоустройства, устройства хранения информации)

USB web камера



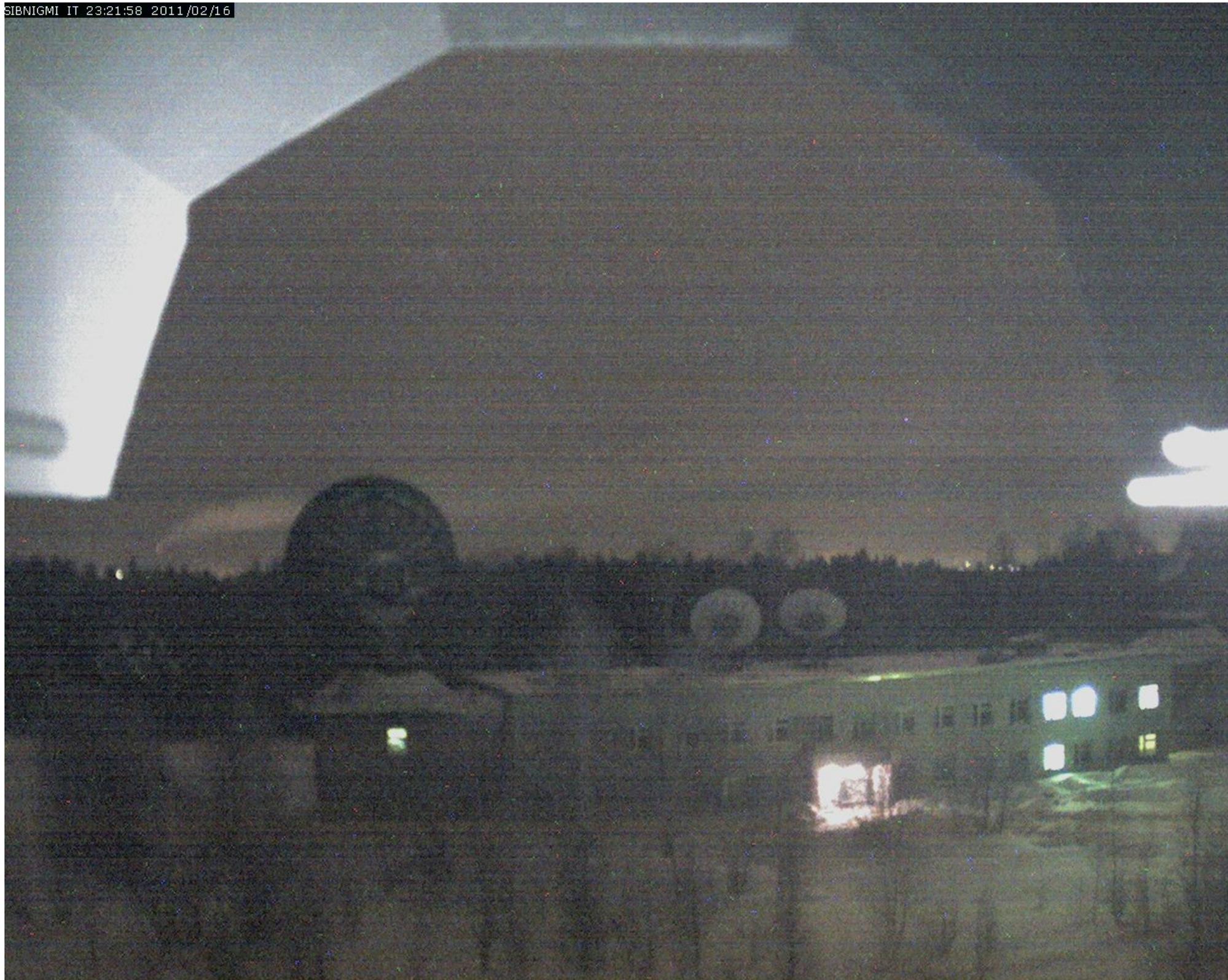
- Реальное разрешение матрицы «скрыто» и зачастую это про...
- Управлять параметрами камеры (диафрагма, ISO) затруднительно (невозможно)
- Лучшие образцы обеспечивают не более 25 Fps
- Получить данные в Raw формате нельзя (только сжатые с потерей)
- Сигнальный процессор камеры в режиме «единичного» кадра может не успеть настроить «диафрагму» и фокус (где это возможно)
- Управлять ROI невозможно
-
- + Дешевое решение из коробки, предоставляющее данные удовлетворительного качества

Ethernet камера

- Актуально все что сказано для USB камер, с поправкой на «публичный канал связи»
-
- Интересны «охранные камеры», способные
в тч Видеть в ИК диапазоне (например,









400 m OK



10000 m OK



500 m OK

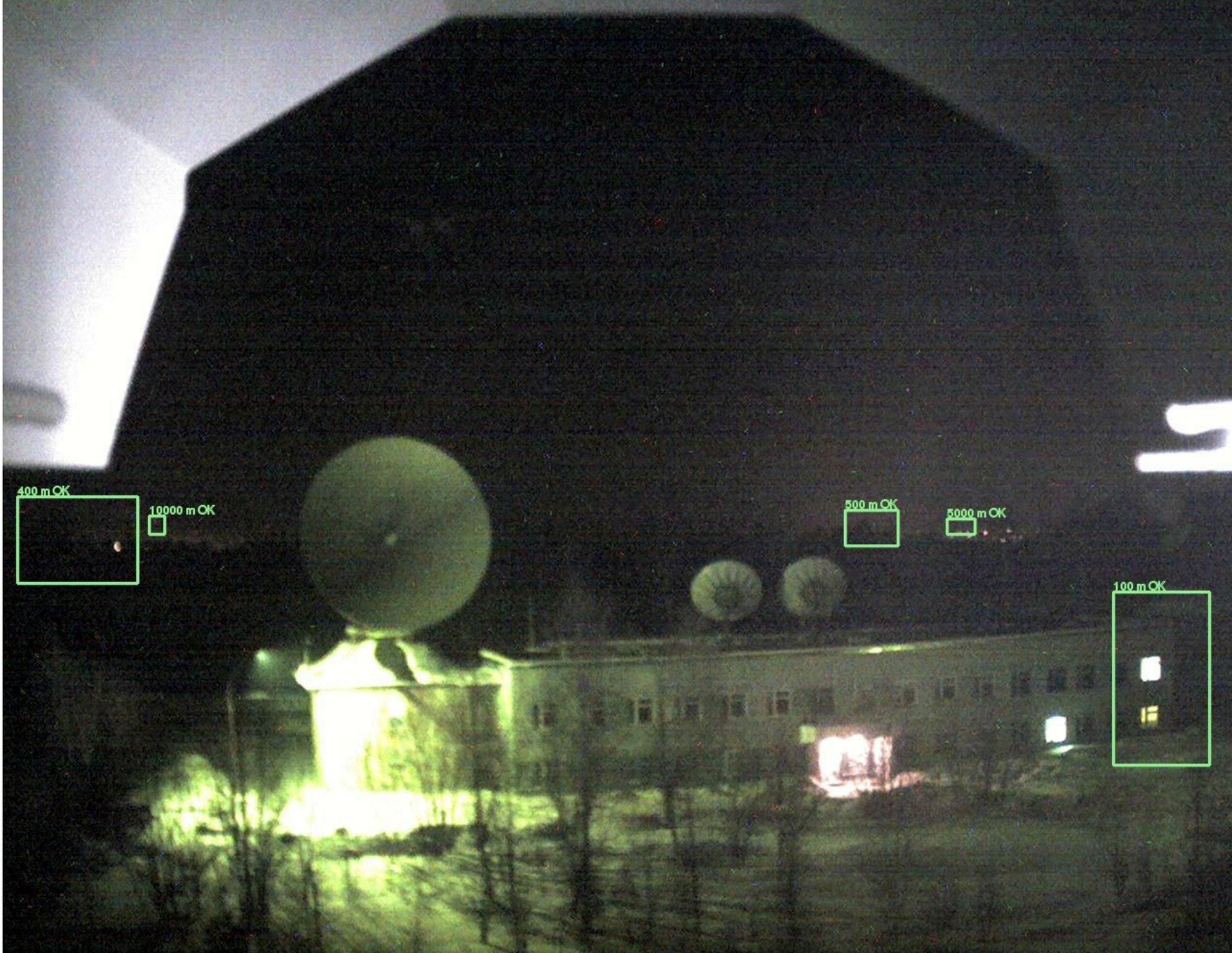


5000 m BAD



100 m OK





400 m OK



10000 m OK



500 m OK



5000 m OK



100 m OK



FireWire IEEE1394



4 pin

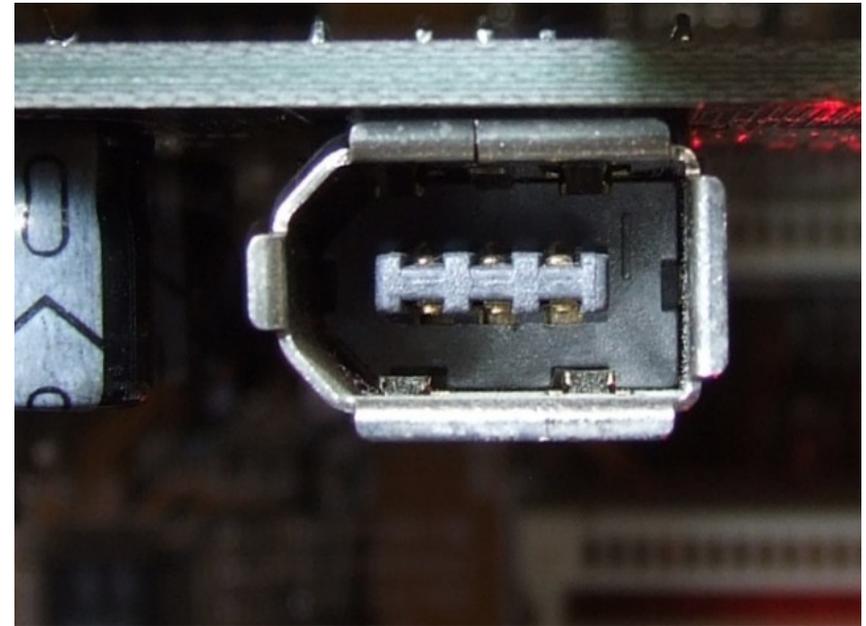


9 pin

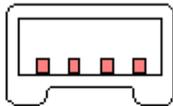


6 pin

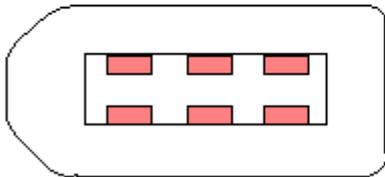
From Computer Desktop Encyclopedia
© 2008 The Computer Language Co., Inc.



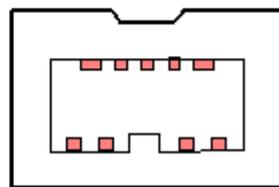
4-Pin iLink



FireWire 400
6-Pin Alpha Connector



FireWire 800
9-Pin Beta Connector



FireWire

- Различная скорость передачи данных — 100, 200 и 400 Мбит/с в стандарте IEEE 1394/1394a, дополнительно 800 и 1600 Мбит/с в стандарте IEEE 1394b и 3200 Мбит/с в спецификации S3200.
- Наличие питания прямо на шине (маломощные устройства могут обходиться без собственных блоков питания). До полутора ампер и напряжение от 8 до 40 вольт.
- Высокая скорость — возможность обработки мультимедиа-сигнала в реальном времени
- Изохронный трафик - Поток данных, передаваемый с постоянной скоростью, в котором все последовательно передаваемые блоки данных строго взаимно синхронизированы с большой точностью.
----> профессиональные скоростные камеры

Marlin Cam

- Пример линейки профессиональных камер.
- До 100 fps
- Возможность установки Roi
- Shutter
- Gain
- Pixel depth