

Яндекс



Анализ изображений и видео

Антон Конушин





Лектор



Антон
Конушин

- Научный сотрудник лаборатории компьютерной графики и мультимедиа ВМК МГУ имени М.В. Ломоносова, к.ф.-м.н.
- Руководитель группы компьютерного зрения
- Научный консультант стартапов «Технологии мобильного картографирования» и «Технологии видеоанализа»

E-mail: ktosh@graphics.cs.msu.ru





Ассистенты



Антон
Слесарёв

- Аспирант МФТИ, выпускник ШАД
- Сотрудник Яндекс
- ton4a@list.ru



Артём
Бабенко

- Администратор Школы анализа данных Яндекса,
- Аспирант МФТИ
- руководитель практики ШАД
- arbabenko@yandex-team.ru



План лекции

- Введение в компьютерное зрение
 - Что это такое
 - Почему это сложно
 - История компьютерного зрения
 - Современные достижения
- Компьютерное зрение и зрение человека
 - Изображение
 - Камера и глаз
 - Цвет и модели цвета



Задача компьютерного зрения

Яндекс

- Понять, что запечатлено на изображении



Мы видим

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

Компьютер видит

Source: S. Narasimhan



Задача компьютерного зрения



- «To see means to know what is where by looking»
 - *David Marr, Vision, 1982*
- «Тест Тьюринга» - компьютер должен ответить на любой вопрос об изображении, на который может ответить человек
- Что это в действительности обозначает?
 - Зрение - источник семантической информации о мире
 - Зрение - источник метрической информации о трехмерном мире



Семантическая информация

Яндекс



Slide credit: Fei-Fei, Fergus & Torralba



Классификация сцены

Яндекс

- вне помещения
- город
- уличное движение
- Пекин, Китай
- Пл. Тяньаньмэнь



slide credit: Fei-Fei, Fergus & Torralba



Выделение объектов

Яндекс





Выделение границ объектов

Яндекс





Признаки (атрибуты) объектов

Яндекс





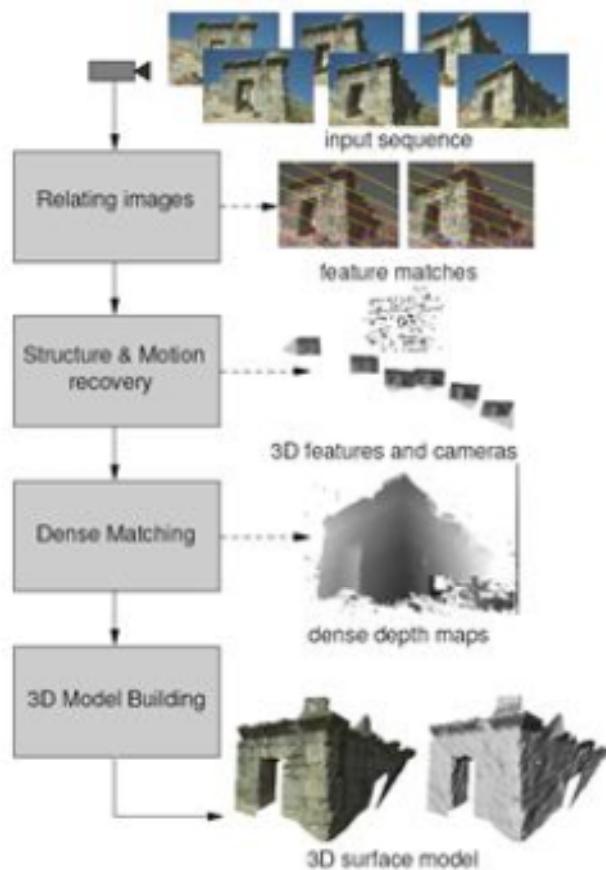
Измерения

Яндекс

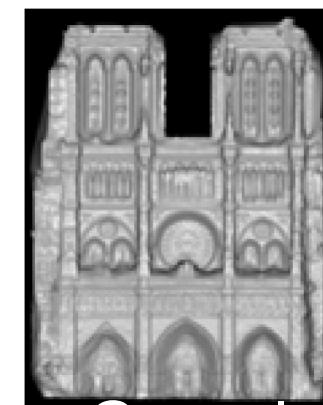
Стерео реконструкция



Структура из движения



Моделирование по пользовательским снимкам





Зрение... принятые названия

- **Фотограмметрия (Photogrammetry)**
 - Исторически измерение расстояний между объектами по 2D изображениям
- **Компьютерное зрение (Computer vision)**
 - Изначально восстановление 3д структуры по 2д изображениям, сейчас шире, как принятие решений о физических объектах, основываясь на их изображениях
- **Машинное зрение (Machine vision)**
 - Обычно понимается как решение промышленных, производственных задач (сложилось исторически)
- **Обработка изображений (Image processing)**
 - На входе и выходе изображение
- **Анализ изображений (Image analysis)**
 - Фокусируется на работе с 2D изображениями
- **Распознавание образов (Pattern recognition)**
 - Распознавание, обучение на абстрактных числовых величинах, полученных в том числе и из изображений



Зачем?

- Наглядное применение многих математических методов анализа данных
- Много практических применений
- Много нерешенных задач
 - Стимул для развития методов анализа данных и высокопроизводительных вычислений
- Сложно
 - 25+% мозга человека отвечает за зрение
 - «ИИ-полная» задача – решение задачи зрения на уровне человека равносильно решению задачи искусственного интеллекта



Почему зрение – это сложно?

Яндекс



Точка наблюдения (ракурс)

Яндекс



Michelangelo 1475-1564



slide credit: Fei-Fei, Fergus & Torralba



Освещение

Яндекс

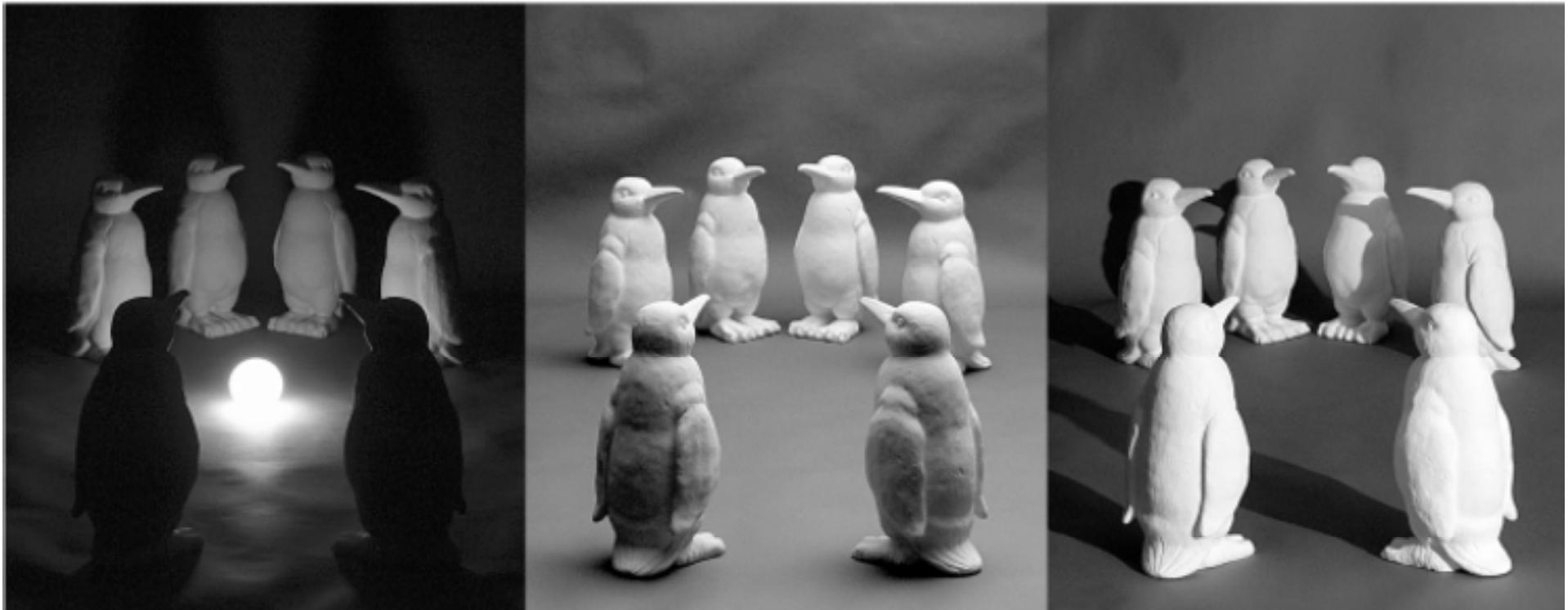


image credit: J. Koenderink



Масштаб



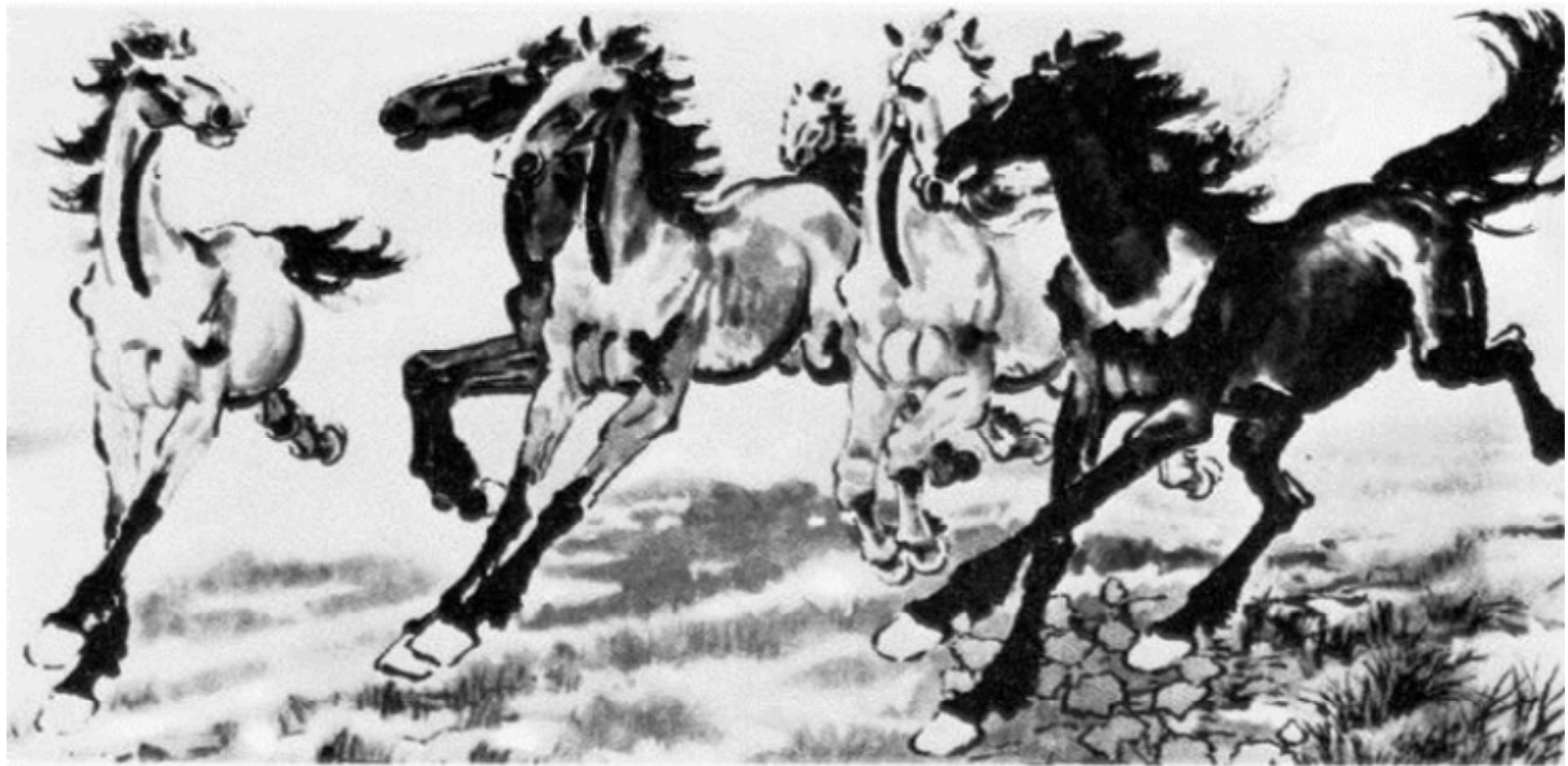
Яндекс

Slide credit: Fei-Fei, Fergus & Torralba



Деформация

Яндекс



Xu, Beihong 1943

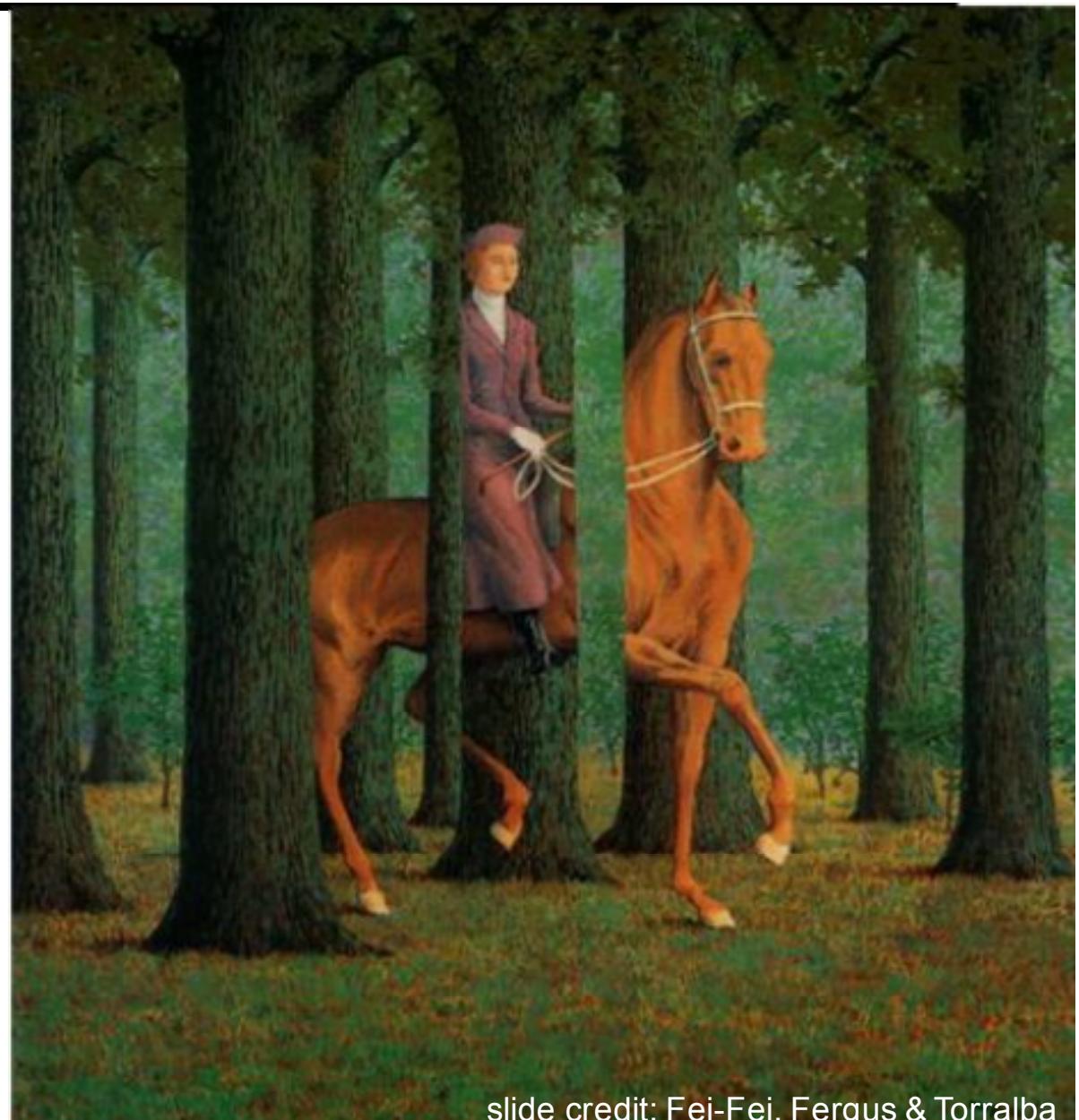
Slide credit: Fei-Fei, Fergus & Torralba



Перекрытие

Яндекс

Magritte, 1957



slide credit: Fei-Fei, Fergus & Torralba



Маскировка

Яндекс





Движение

Яндекс





Внутриклассовая изменчивость

Яндекс

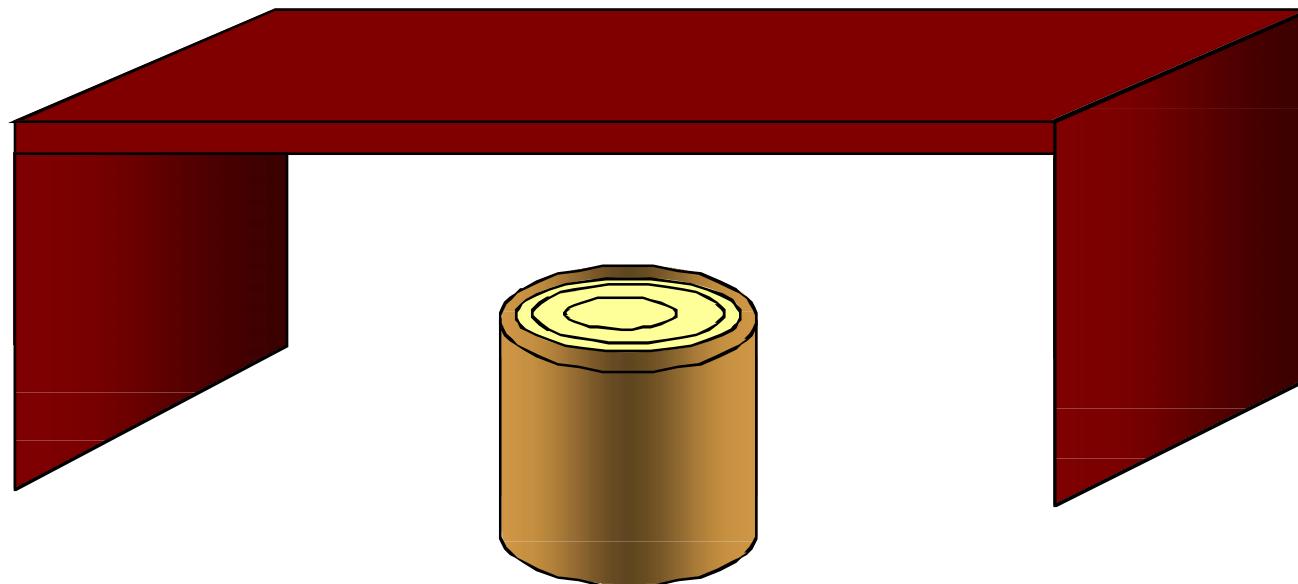


Slide credit: Fei-Fei, Fergus & Torralba



Контекст

Яндекс



~~Полено~~

Стул



Сложности или возможности?

Яндекс

- Изображение запутывает, но дает много подсказок
- Наша задача – интерпретировать подсказки



Image source: J. Koenderink



Цвет

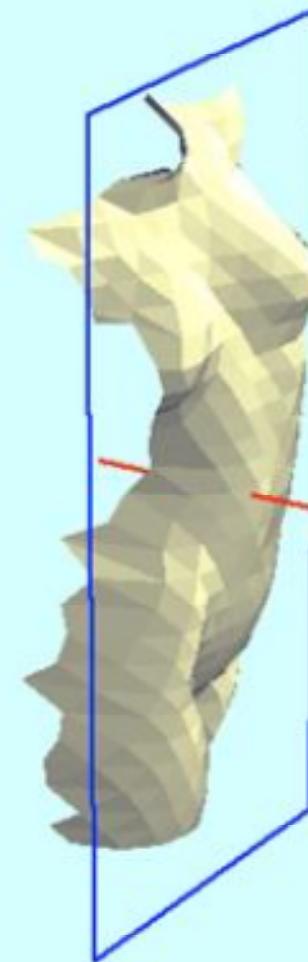
Яндекс





Тени и освещение

Яндекс



Source: J. Koenderink



Отбрасываемые тени

Яндекс



Source: J. Koenderink



Текстура

Яндекс





Перспектива

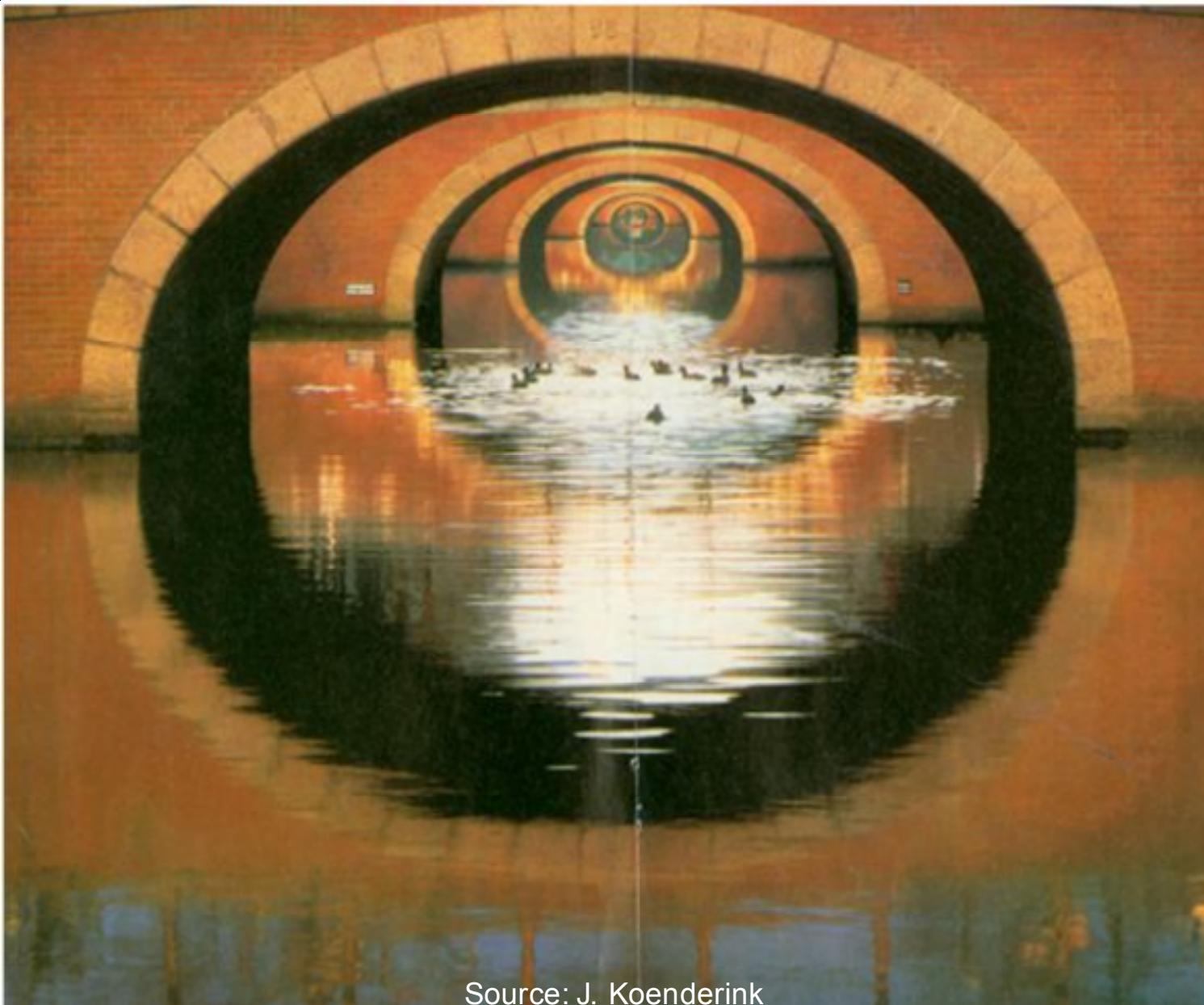
Яндекс





Упорядочивание по глубине

Яндекс



Source: J. Koenderink



Туман и фокусировка

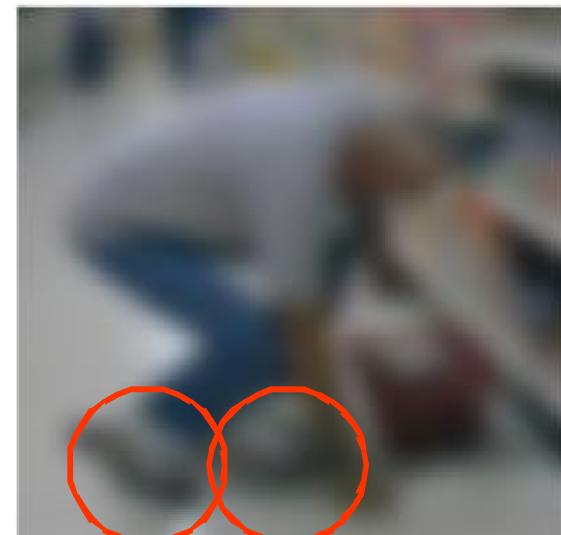
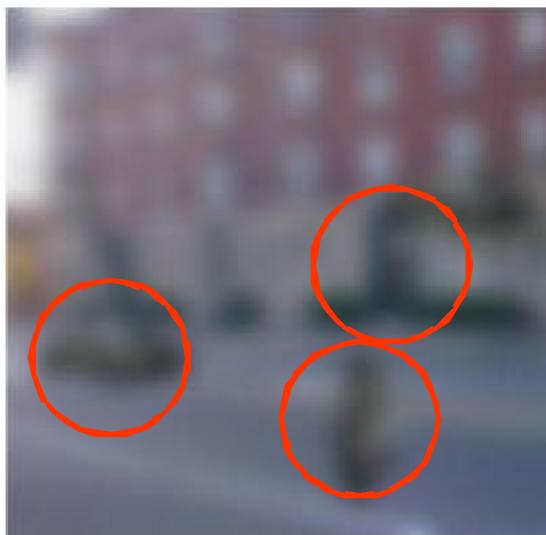
Яндекс





Контекст

Яндекс



Slide credit: Fei-Fei, Fergus & Torralba



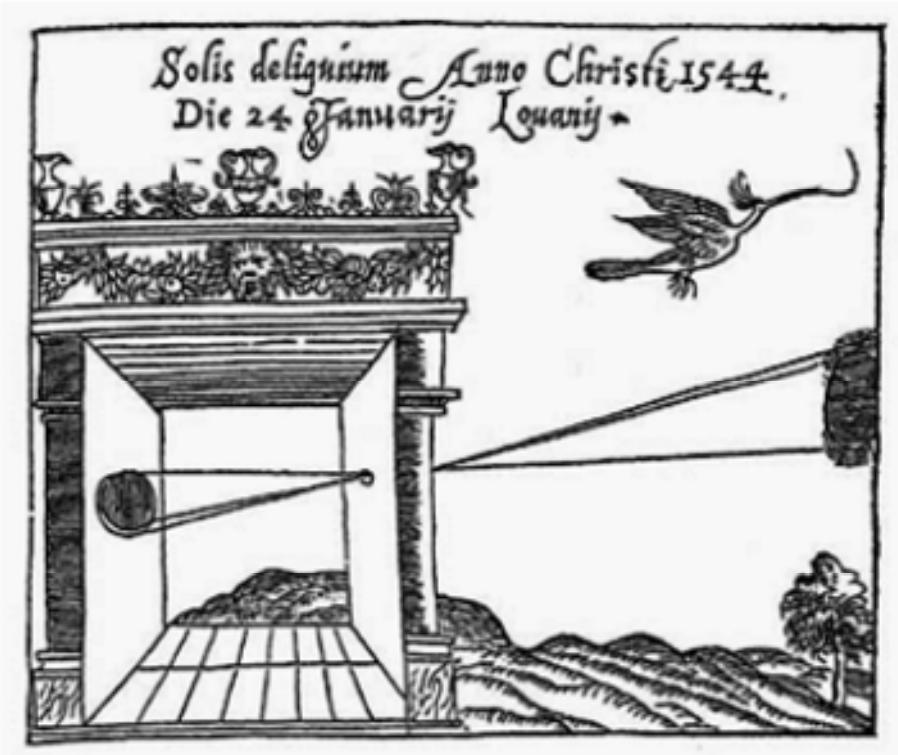
- Зрение изначально нечеткая задача
 - Разные 3D сцены дают одно и то же 2D изображение
 - Необходимы априорные знания о структуре и свойствах мира
 - Нам нужно сопоставлять наблюдения и априорные знания



Image source: F. Durand



История: Камера-обскура



- Принцип был известен еще Аристотелю (384-322 до Н.Э.)



Первая фотография

Яндекс



Самая первая фотография
1825 год



Figure 5. J. N. Niepce.

Требовала 8 часов проявки



Фотограмметрия

Яндекс



Figure 6. Jacques Daguerre.

1837 – первые практически применимые фотографии

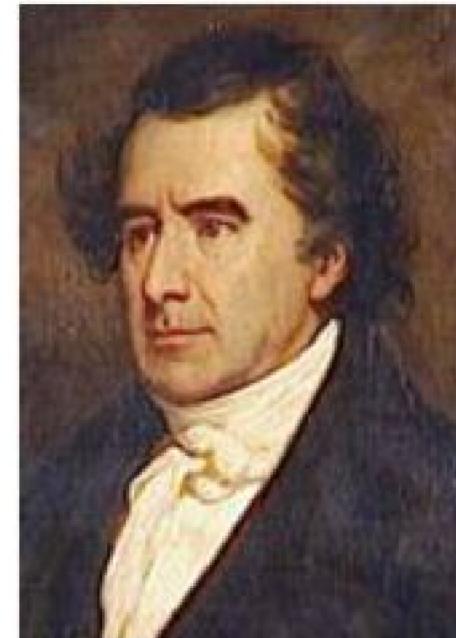


Figure 8. Dominique François Jean Arago.

1840 – «Фотограмметрия – будущее геодезии»



Видео

Яндекс



1878 – первая скоростная съемка, Eadweard Muybridge

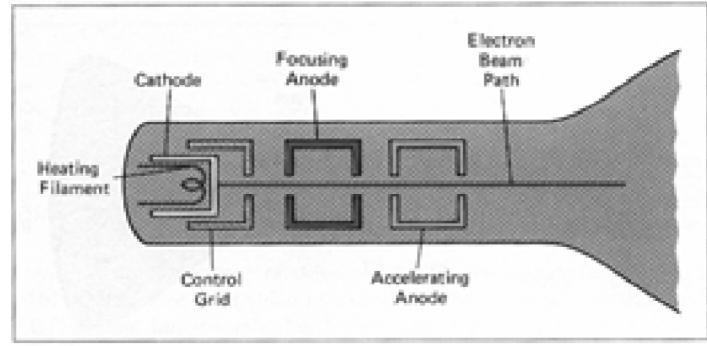


1888 – первое кино на плёнке, Louis Le Prince

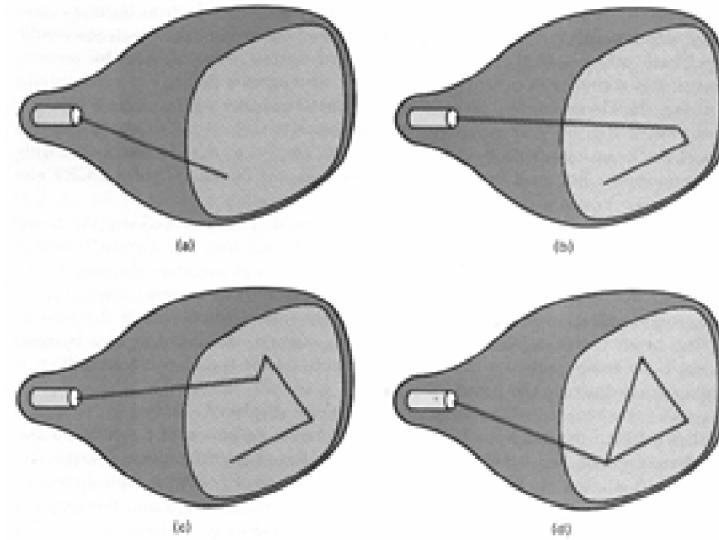


Электронно-лучевая трубка(CRT)

Яндекс



1885 – изобретение CRT

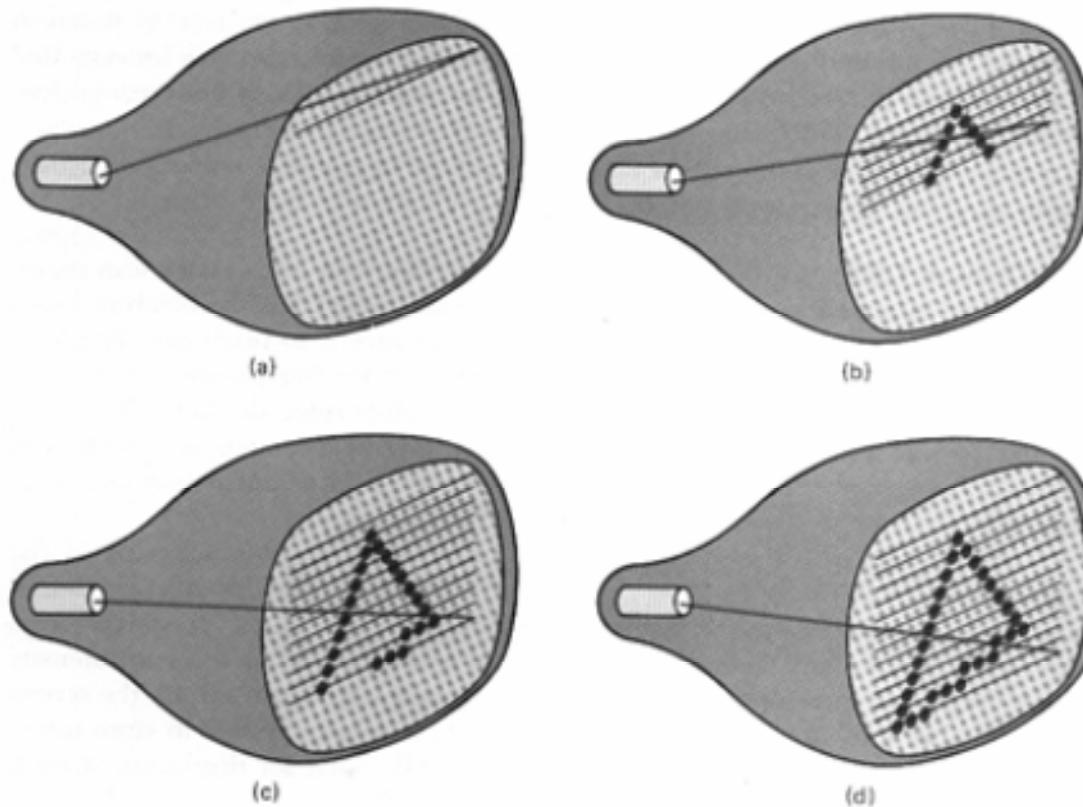


1897 – CRT с
флуоресцентным
экраном



Растровый дисплей – 1927 год

Яндекс



Philo Farnsworth – 60-строчный растровый дисплей



Whirlwind, MIT, 1951

Яндекс

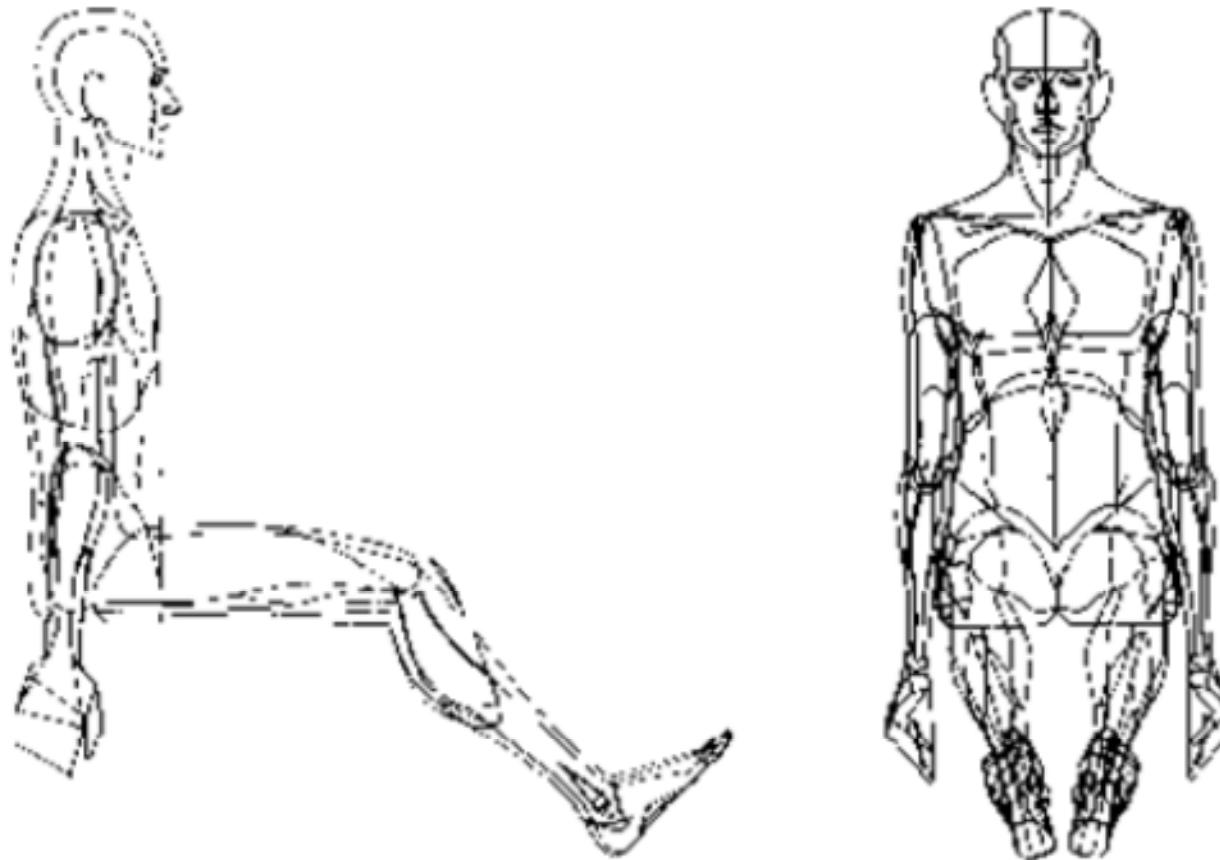


- Первый компьютер, отображающий текст и графику в реальном времени на мониторе
- Точками карту, значком самолёт.
- «Световое перо» для взаимодействия с экраном (запрос информации об объекте)



“The Boing man”, 1960

Яндекс

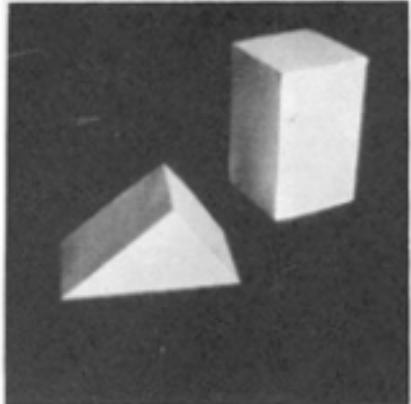


Первое компьютерное изображение человека

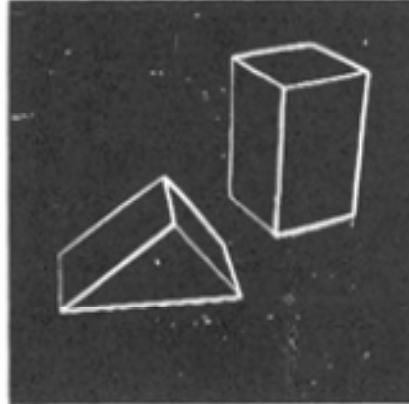


Зарождение компьютерного зрения

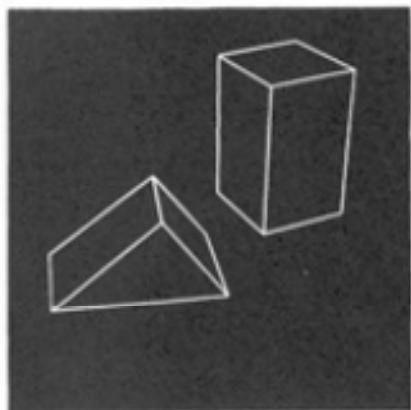
Яндекс



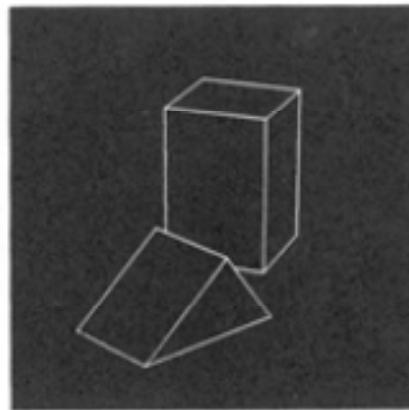
(a) Original picture.



(b) Differentiated picture.



(c) Line drawing.



(d) Rotated view.

L. G. Roberts, *Machine Perception of Three Dimensional Solids*,
Ph.D. thesis, MIT Department of
Electrical Engineering, 1960



Spacewar, MIT, 1961

Яндекс



- Steve Russell, 200 человеко-месяцев



SketchPad, MIT, 1963

Яндекс

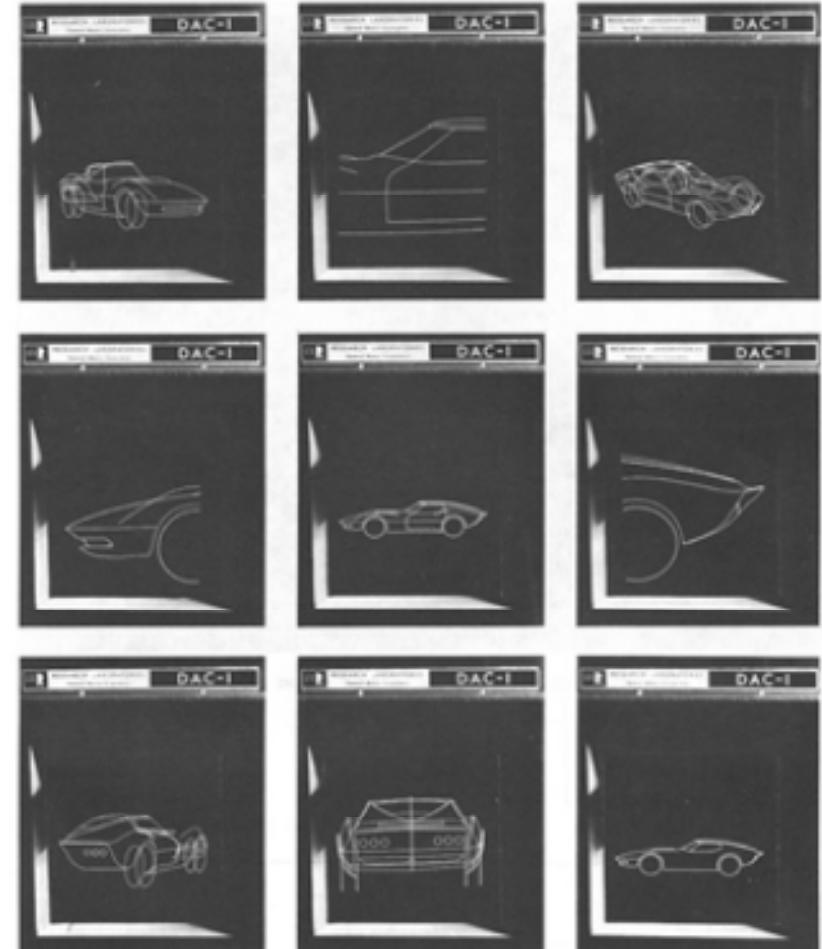


- Ivan Sutherland демонстрирует интерактивный графический редактор SketchPad



CAD, IBM + GM, 1964

Яндекс



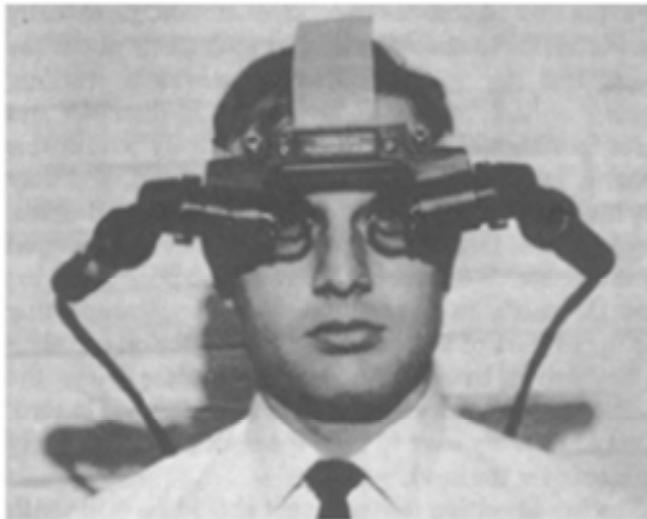
Первая CAD-система,
геометрические
преобразования (поворот,
вращение)

FIG. 1-8 Scale expansion, rotation, and partial views in a DAC-1 design exercise.



Virtual Reality, Harvard, 1968

Яндекс

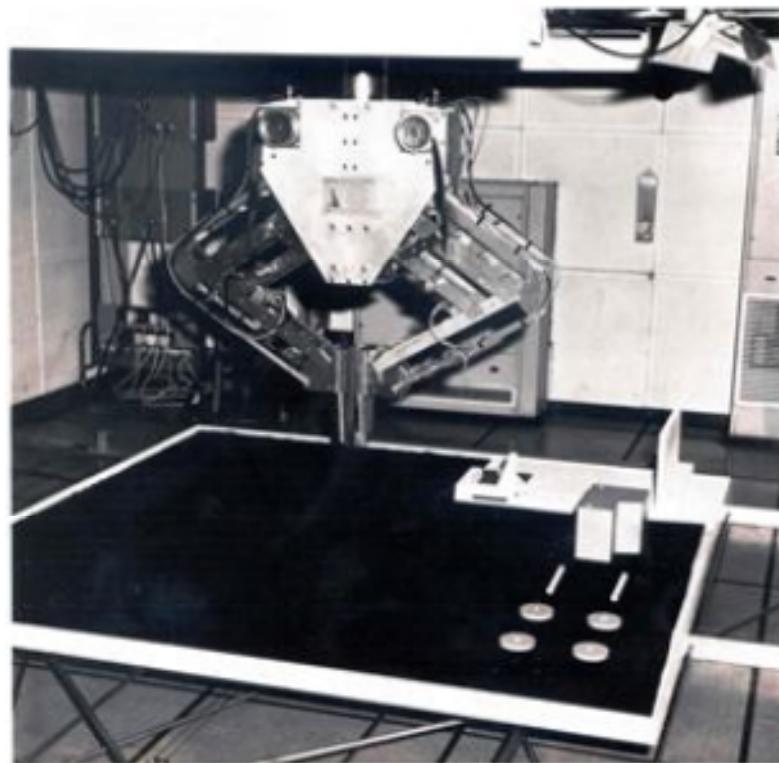


- Ivan Sutherland перешел в Гарвард, где разработал первый Head Mounted Display (HMD)
- Виртуальная комната (wireframe), в которую можно войти



Freddy II, 1973

- Университет Эдинбурга
- Один из первых роботов с системой машинного зрения
- 5 степеней свободы
- Умеет собирать машинки из кубиков, разбросанных по столу
- 384КБ RAM в управляемом компьютере

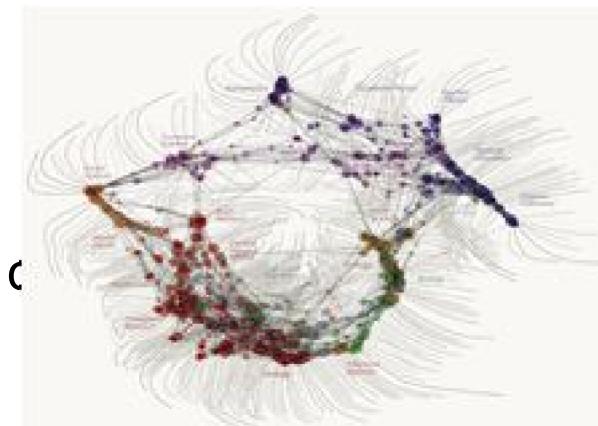
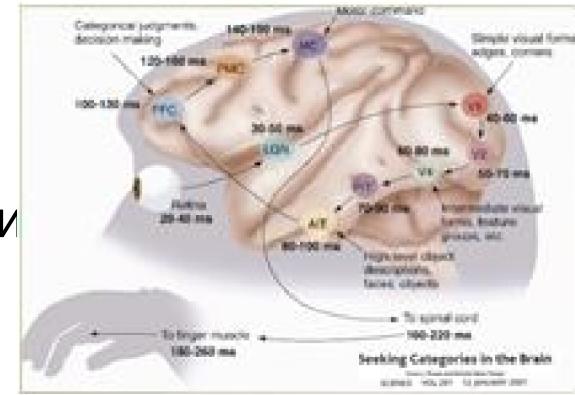




Давид Марр (1970е)

Яндекс

- «Primal sketch»
 - Низкоуровневые («low-level») свойства изображения: направленные края, отрезки и т.д.
- «2.5D sketch»
 - Упорядочивание по глубине (бинокулярное стерое), учёт текстуры и т.д.
- «3D model»
 - Распознавание объектов и представление в 3х мерном мире





Детектор лиц Viola-Jones (2001)

Яндекс



Алгоритм Viola-Jones – первый быстрый и надежный алгоритм поиска лиц. Демонстрация силы машинного обучения.

Source: S. Seitz



Практическое применение



- Автоматизация обработки текстовых данных привела к революционным изменениям в организации бизнеса и жизни
- Изображения дают 90+% информации, но пока обрабатываются вручную
- В перспективе, компьютерное зрение – один из главных компонентов робототехники





Области применения



- **Распознавание и обработка документов**
- **Обработка фотографий** – улучшение качества, ретушь, изменение размера и формы, композиция
- **Интернет** – поиск, аннотация, поиск дубликатов, распознавание объектов
- **Системы безопасности** – видеонаблюдение, отслеживание, распознавание объектов, распознавание жестов и событий
- **Дистанционное зондирование и ГИС** – карты, анализ спутниковых данных
- **Неразрушающий контроль** - диагностика, контроль качества
- **Медицинские системы** – анализ томограмм
- **Спецэффекты в кино** – композиция, монтаж фонов, захват движения

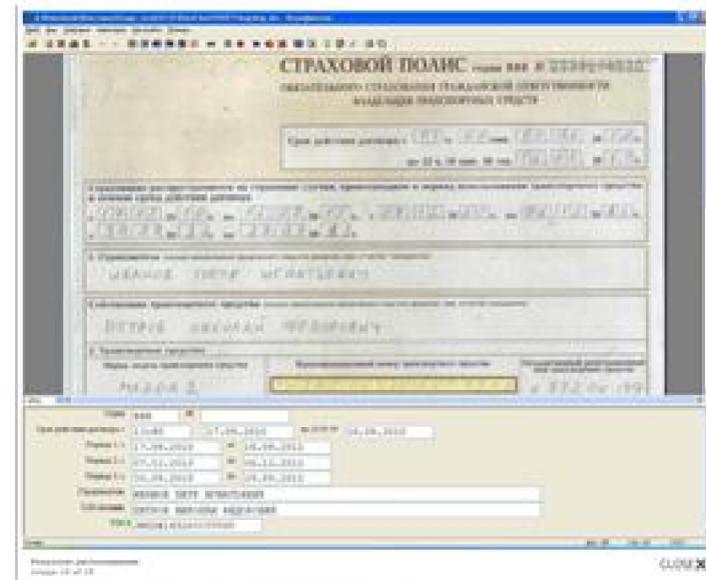


Распознавание текста

Яндекс



FineReader, ABBYY, Россия
<http://www.abbyy.ru/finereader/>

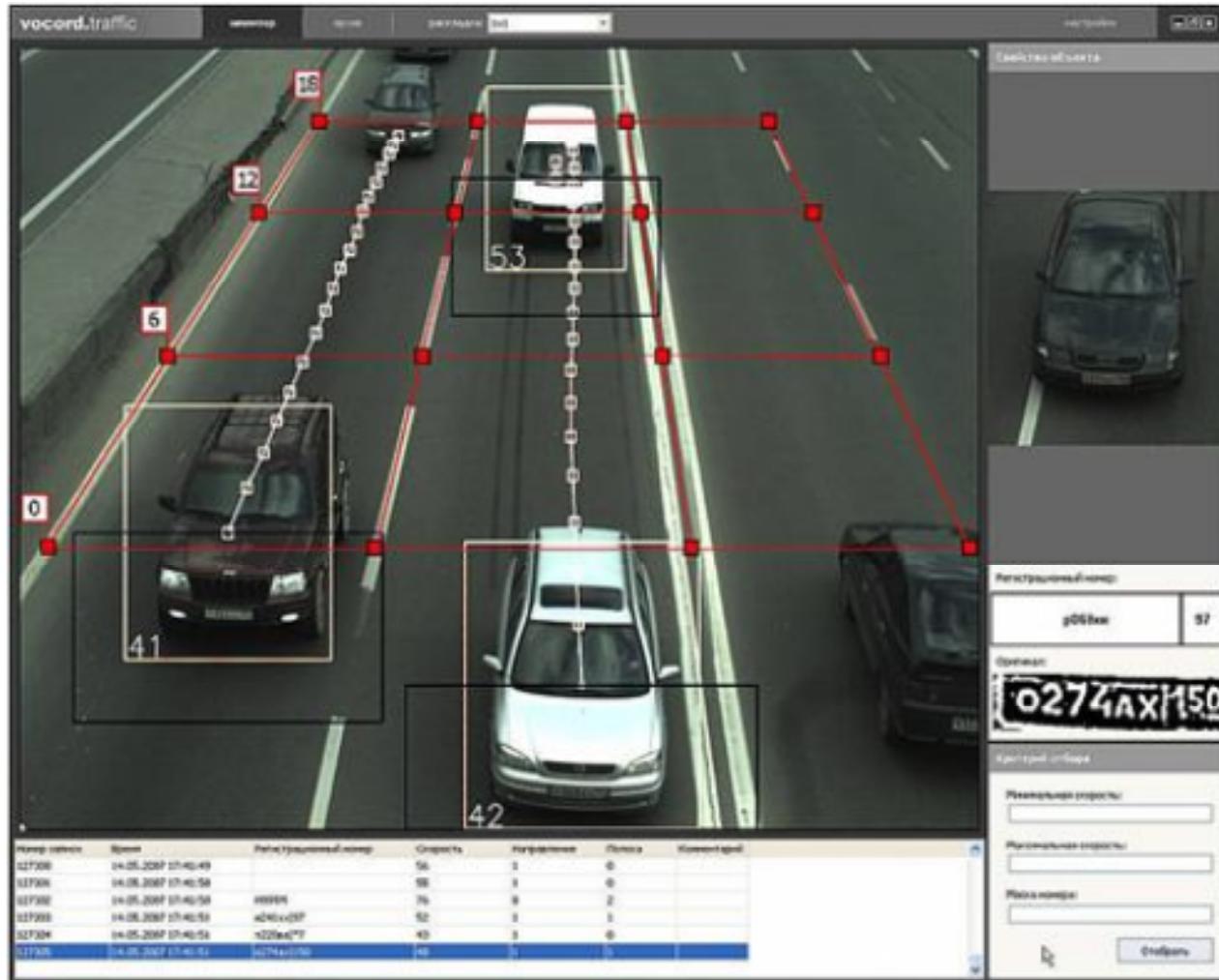


Cognitive Technologies, Россия
<http://cognitiveforms.ru/products/cognitive-forms/>



Видеонаблюдение

Яндекс



VOCORD Traffic, Vocord, Россия

<http://www.vocord.ru/218/>



Обнаружение лица и улыбки

Яндекс

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



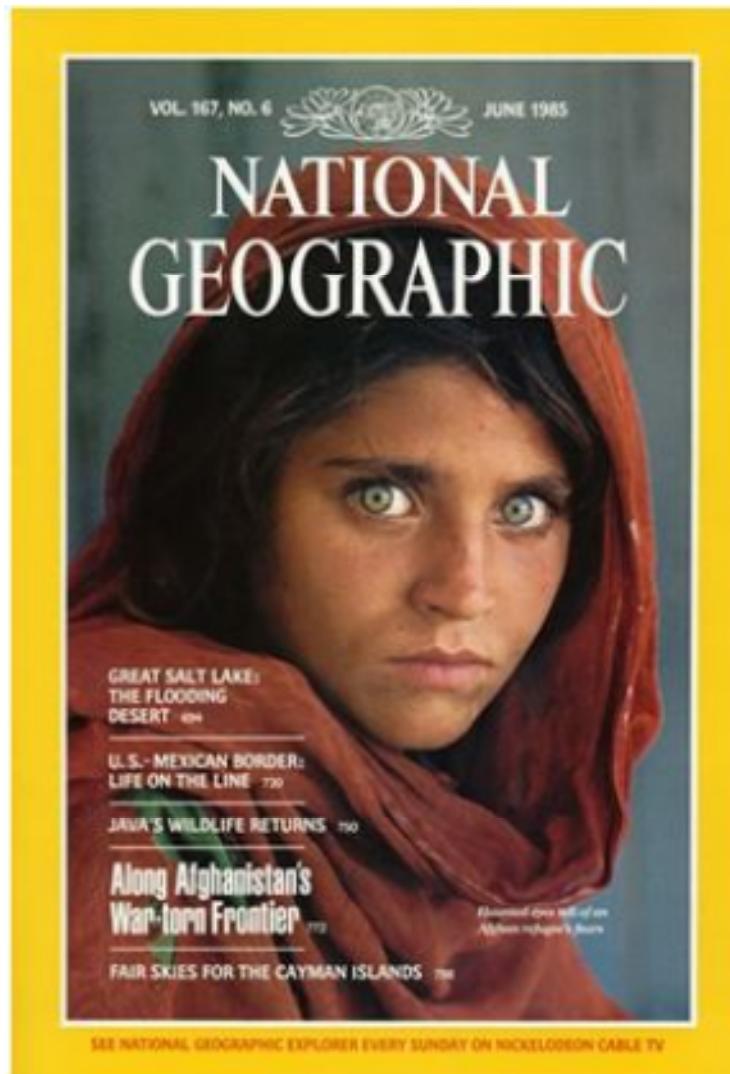
[Sony Cyber-shot® T70 Digital Still Camera](#)

Source: S. Seitz



Биометрия

Яндекс



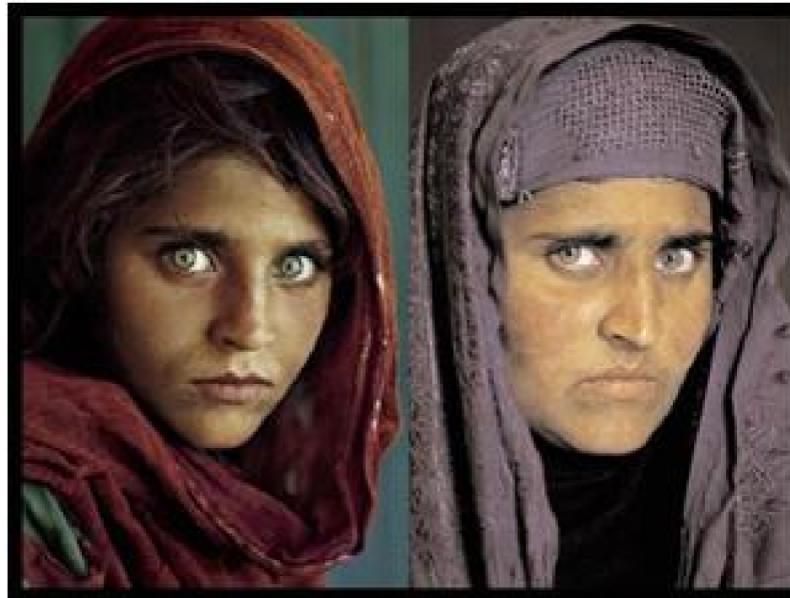
Кто она?

Source: S. Seitz

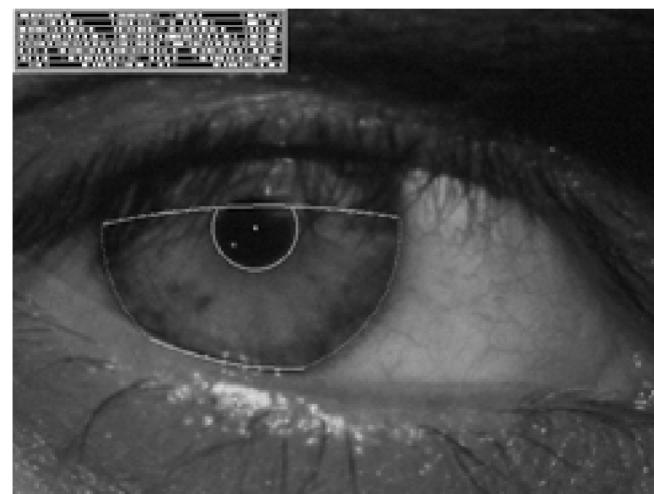
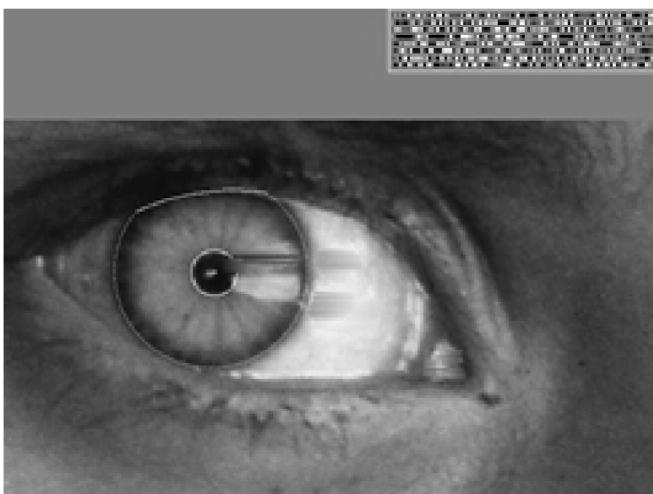


Биометрия

Яндекс



“Как девушку из Афганистана идентифицировали по радужке глаза” Читайте [историю](#)



Source: S. Seitz



Биометрия

Яндекс



Biolink, Россия
<http://www.biolink.ru>

Сканеры отпечатков пальцев для контроля доступа



Распознавание лиц

Яндекс

The world's largest and most accurate face recognition platform

CHECK OUT OUR FREE API >

KLIK

Jack

Kyle

Chloe

Face Recognition Camera
Tagging your Facebook friends has never been easier!

Available on the App Store

Была куплена Facebook за \$100M

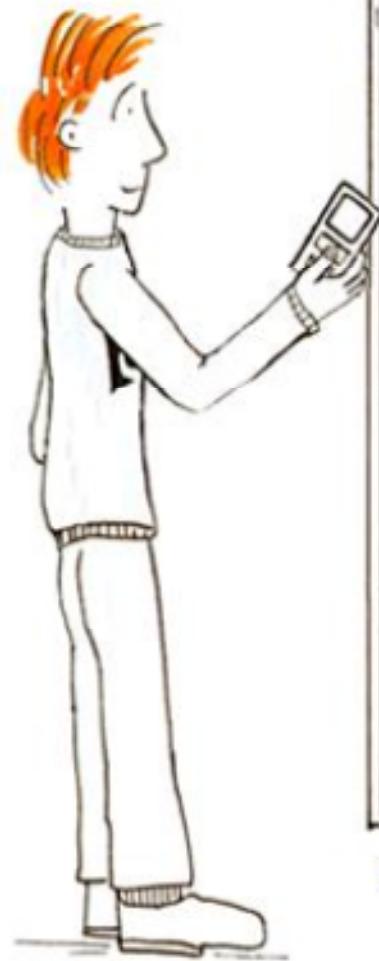


Мобильные приложения

Яндекс

kooaba

MOBILE IMAGE RECOGNITION
TRY IT OUT NOW!!!



www.kooaba.com

1. POINT
YOUR MOBILE
PHONE CAMERA TO
THE MOVIE
POSTER.

2. SNAP A
PICTURE AND SEND
IT:

IN SWITZERLAND:
MMS TO 5555 (OR
079 394 57 00
FOR ORANGE
CUSTOMERS)

IN GERMANY:
MMS TO 84000

EVERYWHERE:
EMAIL TO
M@KOOABA.COM

3. FIND ALL
RELEVANT INFOR-
MATION ABOUT THE
MOVIE ON YOUR
MOBILE PHONE



Умные машины

Яндекс

The screenshot shows the Mobileye website homepage. At the top, there are two tabs: "manufacturer products" (highlighted in blue) and "consumer products". Below the tabs, the slogan "Our Vision. Your Safety." is displayed. A central image shows a car from above with three cameras highlighted: "rear looking camera" on the left, "forward looking camera" on the right, and "side looking camera" at the bottom. Below this, there are three main sections: "EyeQ Vision on a Chip" featuring a chip image, "Vision Applications" featuring a pedestrian crossing image, and "AWS Advance Warning System" featuring a circular display image. To the right, there is a "News" section with a thumbnail of a hand on a steering wheel and a "Events" section with thumbnails for "Mobileye at Equip Auto, Paris, France" and "Mobileye at SEMA, Las Vegas, NV".

- > [EyeQ Vision on a Chip](#)
- > [Vision Applications](#)
- > [AWS Advance Warning System](#)

News

- > [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- > [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)

Events

- > [Mobileye at Equip Auto, Paris, France](#)
- > [Mobileye at SEMA, Las Vegas, NV](#)

- [Mobileye](#)
 - Топ-модели от BMW, GM, Volvo
 - К 2010: 70% производителей машин

Source: S. Seitz



Умные машины

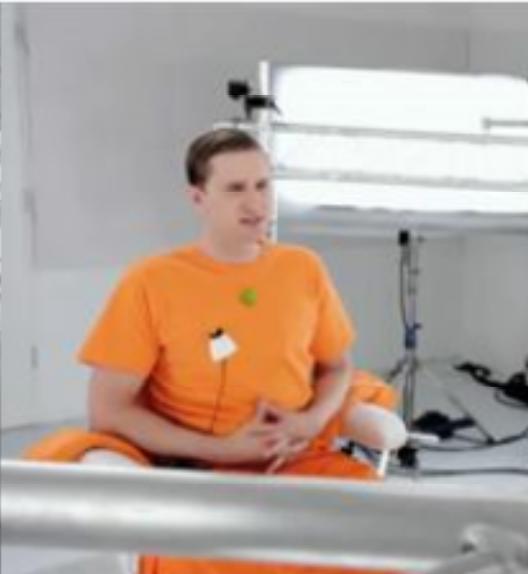
Яндекс





3D модели и захват движения

Яндекс



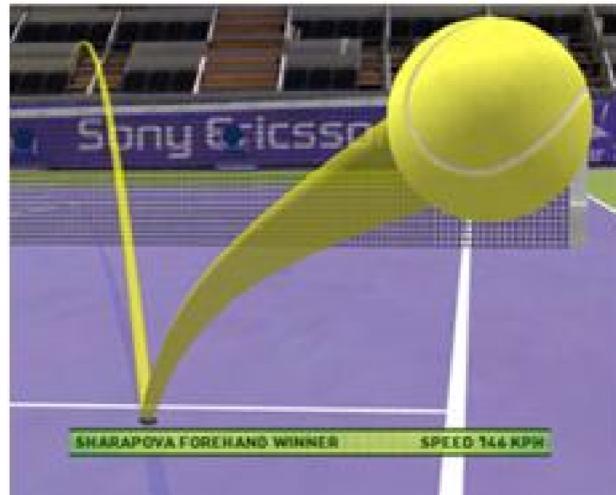
L.A. Noire, Team Bondi/Rockstar, 2011





Спортивные соревнования

Яндекс



<http://www.hawkeyeinnovations.co.uk>



Зрение в космосе

Яндекс



[NASA'S Mars Exploration Rover Spirit.](#)

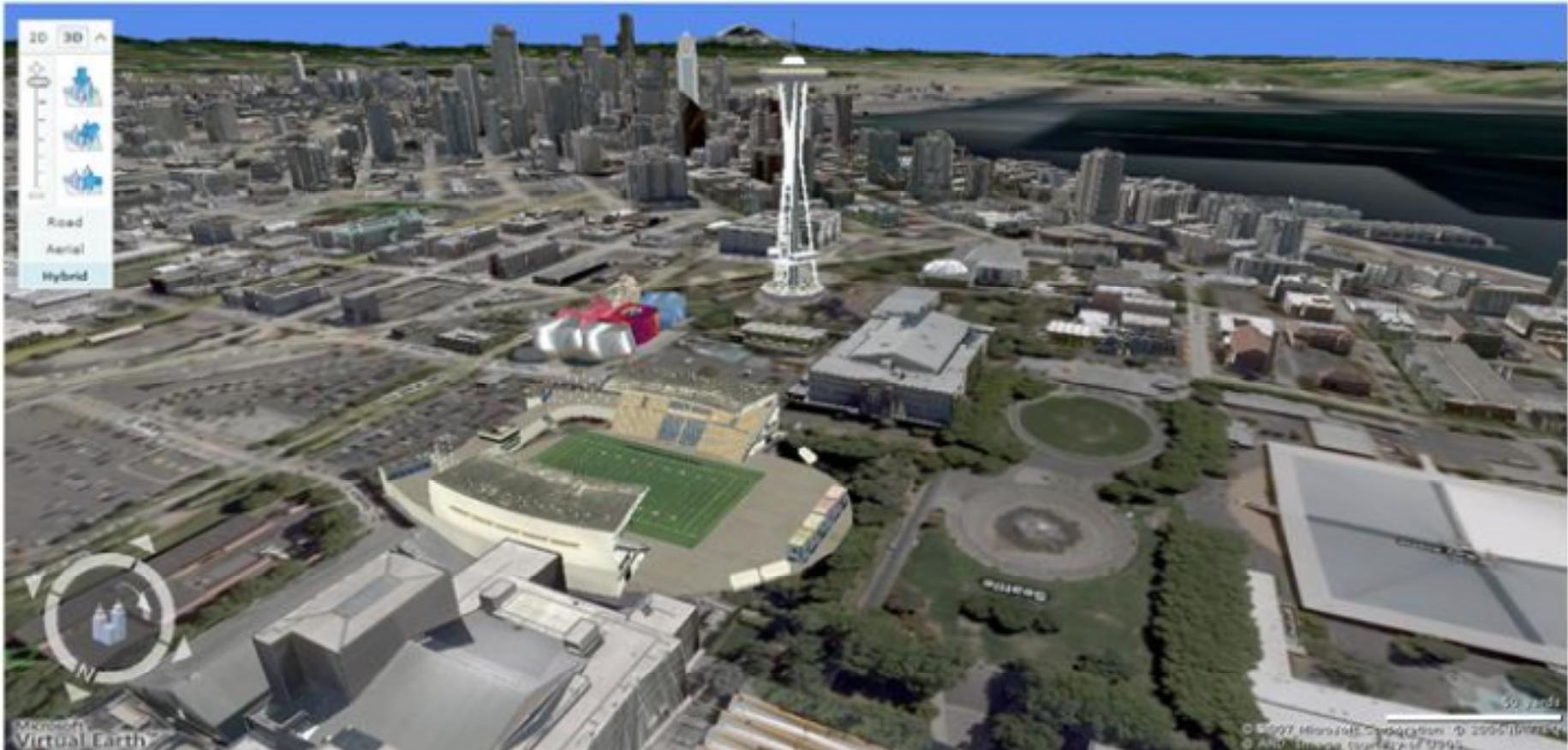
Системы зрения использовались для:

- Склейка панорам
- 3D моделирование местности
- Поиск препятствий, определение местоположения
- Подробнее см. “[Computer Vision on Mars](#)” by Matthies et al.



Трехмерные карты

Яндекс



Изображение из Microsoft's [Virtual Earth](#)
(аналогичные [Google Earth](#))

Source: S. Seitz



PhotoSynth

Яндекс

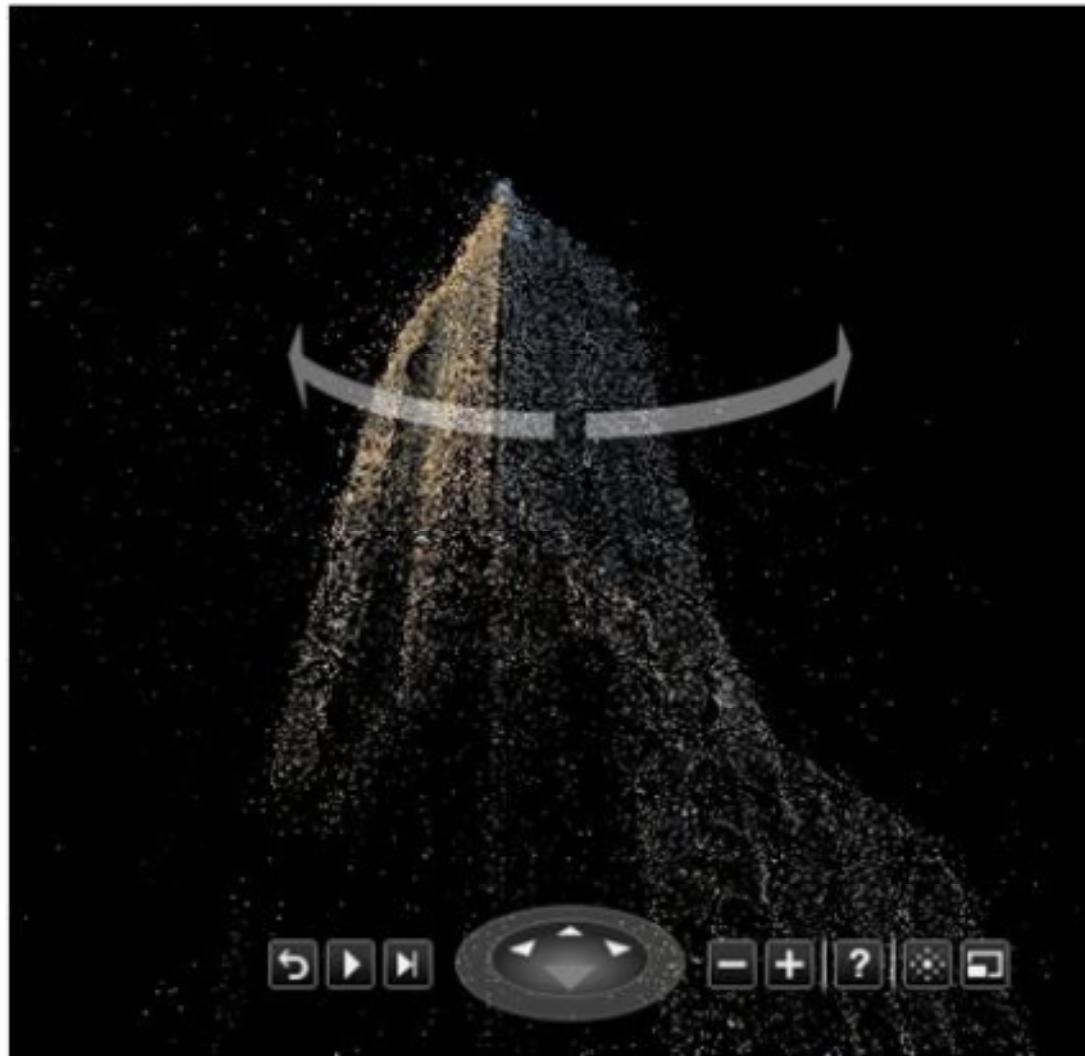


Прогулки по коллекциям изображений из одной сцены



PhotoSynth

Яндекс





Kinect (2010)

Яндекс



Первая потребительская система взаимодействия с
компьютером с помощью жестов