

The background features two wireframe hands, one in the upper right and one in the lower left, reaching towards the center. The hands are composed of a grid of white lines. The background is a solid blue color with several bright, circular lens flare effects in the center.

计算机视觉基础

周文晖

计算机学院

自我介绍

- 办公地址：三教南楼505-507室
- 联系电话：13336096310 电子邮件：zhouwenhui@hdu.edu.cn
- 个人主页：<https://faculty.hdu.edu.cn/jsjxy/zwh/main.htm>
- 研究方向：
 - 计算成像
 - 立体视觉和三维建模
 - 视觉感知和视觉导航
 - ...

前导知识结构

- 基础理论
 - 高等数学、概率论与数理统计、几何学、图论、...
- 专业理论
 - 数字图像处理、数字信号系统、...
- 编程能力
 - Python、C/C++、Matlab、Java、...
- 硬件与系统知识
 - 嵌入式系统、硬件语言、系统结构、...

如何学？

- 计算机视觉是一门多学科交叉性课程，涉及计算机、信息、电子、仿生、数学等。
- 如何去学？
- 学习的阶段和层次
 - 本科生是导论，概念、发展和思想建模。
 - 研究生阶段涉及算法、编程和实现。
- 学习的方法
 - 学习必要性：智能化是社会发展趋势。
 - 学习的方法：日常生活中培养兴趣。

教材

- 章毓晋编著 《计算机视觉教程》 人民邮电出版社 2017年3月第2版



主要参考书目

- Rafael C. Gonzalez. Digital Image Processing (3rd Edition). Prentice Hall.
- 马颂德,张正友. 计算机视觉—计算理论与算法基础. 科学出版社. 2003.
- Richard I. Hartley, Andrew Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press. 2004.
- ...

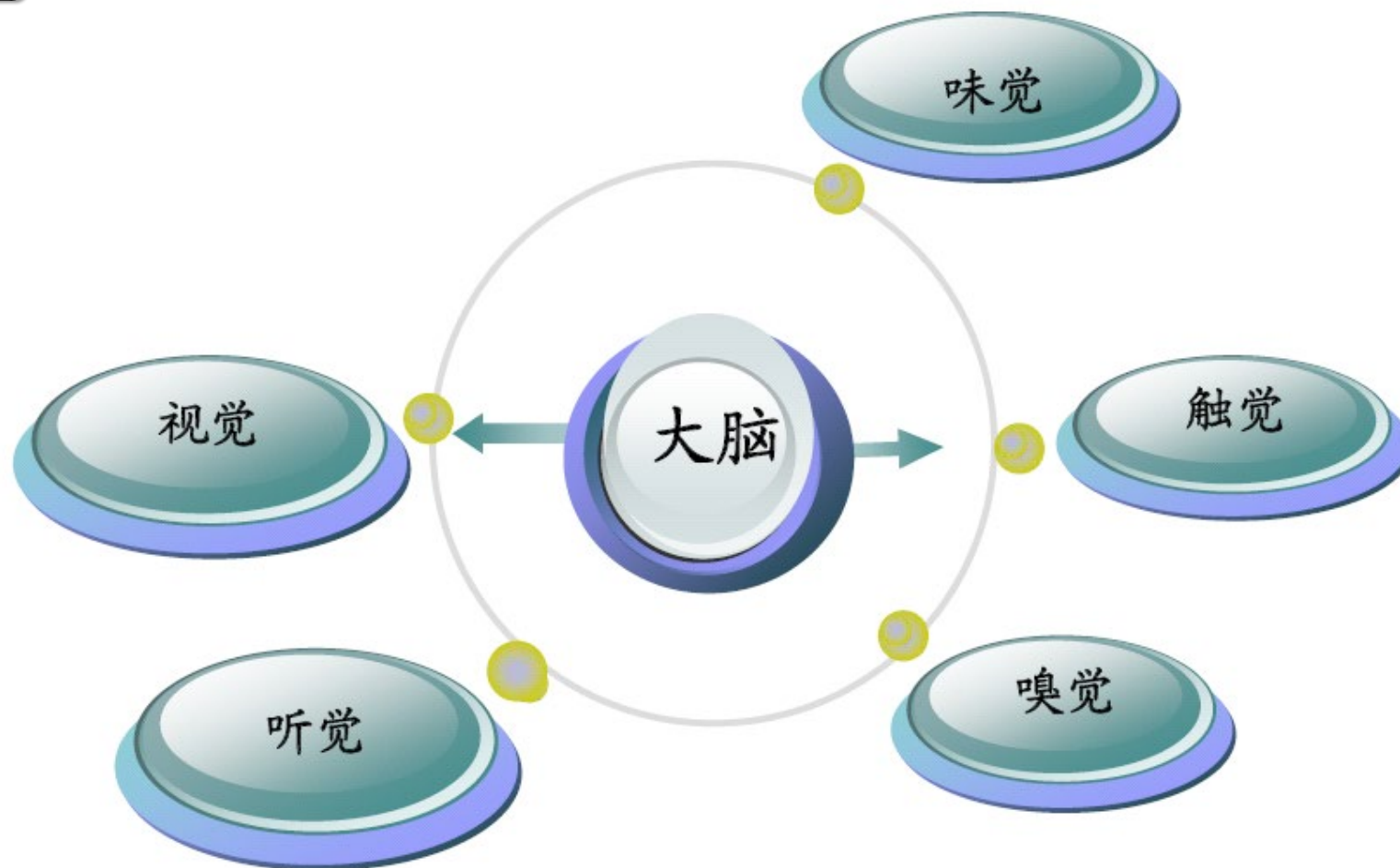
第一章 绪论

1.1 计算机视觉

人们是怎样认识世界



人的五感



百闻不如一见

- 人类从外界获取的信息约60~90%来自视觉系统

视感觉

- ① 光的物理特性;
- ② 光刺激视觉感受器官的程度;
- ③ 光作用于视网膜后经视觉系统加工而产生的感觉。

视知觉

- ① 接收到视觉刺激后的反应;
- ② 将分散刺激加以组织, 构成具有一定形状的整体以认识世界。

视觉的最终目的是**对客观场景的解释和描述**

什么是计算机视觉？

- 教材定义：采用计算机实现人类视觉功能。



让计算机理解图像和视频



What kind of scene?

Where are the cars?

How far is the building?

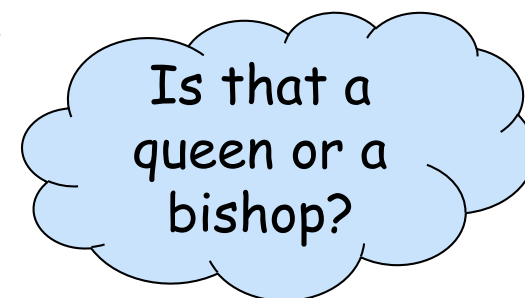
...

与数字图像处理的区别

- 数字图像处理
 - 图像/视频 -> 图像/视频
 - 图像变换、图像滤波、图像复原、图像压缩、 ...
- 计算机视觉
 - 图像/视频 -> 模型
 - David Marr 视觉计算理论的创始人
 - 二维基素图 -> 2.5维要素图 -> 三维模型表征

Vision is really hard

- 视觉是大自然智慧的一个奇迹
 - 视觉皮层面积约占整个大脑皮层面积的50%
 - 视力正常的人类大脑中视觉活动占整个大脑活动的主导地位



CVPR2021

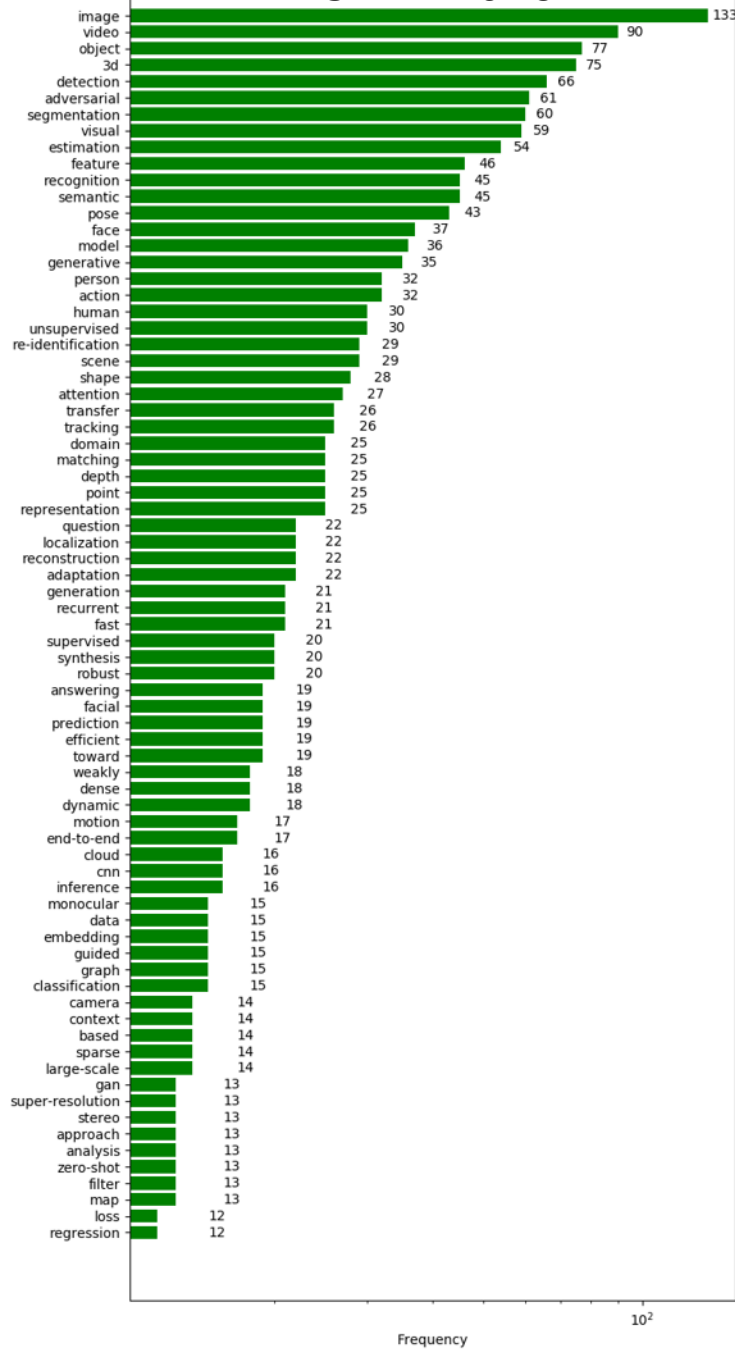
- 2D object recognition
- 3D computer vision
- 3D object recognition
- Action and behavior recognition
- Adversarial learning, adversarial attack and defense methods
- Biometrics, face, gesture, body pose
- Computational photography
- Datasets and evaluation
- Efficient training and inference methods for networks
- Explainable AI, fairness, accountability, privacy, transparency and ethics in vision
- Image and video retrieval
- Image and video synthesis
- Image classification
- Low-level and physics-based vision
- Machine learning architectures and formulations
- Medical, biological and cell microscopy
- Motion and tracking
- Optimization and learning methods
- Pose estimation
- Representation learning, deep learning
- Scene analysis and understanding
- Transfer, low-shot, semi- and un-supervised learning
- Video analysis and understanding
- Vision + language, vision + other modalities
- Vision applications and systems, vision for robotics and autonomous vehicles

ICCV2021

- 3D from a single image and shape-from-x
- Action and behavior recognition
- Adversarial learning
- Biometrics
- Computational photography
- Datasets and evaluation
- Detection and localization in 2D and 3D
- Efficient training and inference methods
- Explainable AI
- Faces
- Fairness, accountability, transparency, and ethics in vision
- Gestures and body pose
- Image and video synthesis
- Image and video retrieval
- Image and video manipulation detection and integrity methods.
- Low-level and physics-based vision
- Machine learning architectures and formulations
- Medical, biological, and cell microscopy
- Motion and tracking
- Neural generative models
- Optimization and learning methods
- Recognition and classification
- Representation learning
- Scene analysis and understanding
- Scene text and document understanding
- Segmentation, grouping and shape
- Stereo, 3D from multiview and other sensors
- Transfer/Low-shot/Semi/Unsupervised Learning
- Video analysis and understanding
- Vision + language
- Vision + other modalities
- Vision applications and systems
- Vision for robotics and autonomous vehicles
- Visual reasoning and logical representation

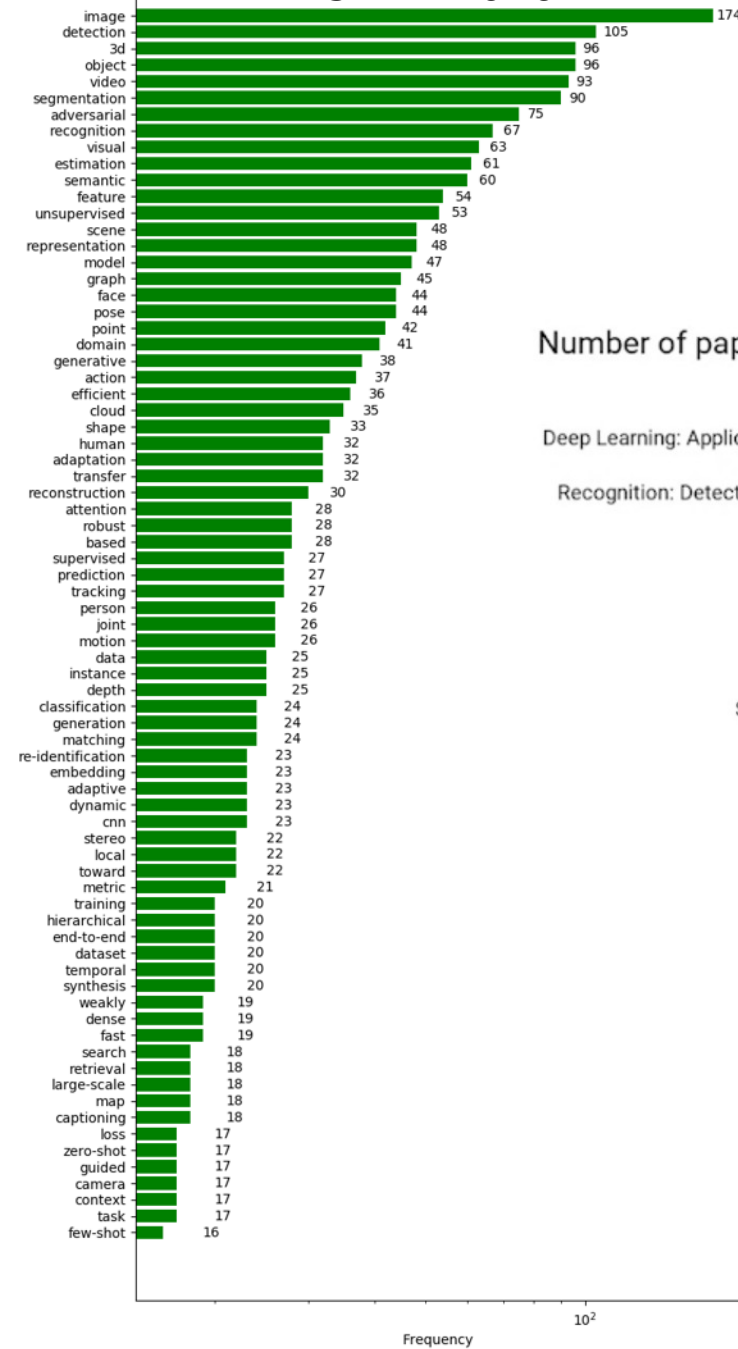
CVPR 2018 Submission Top 75 Keywords

CVPR2018



CVPR 2019 Submission Top 75 Keywords

CVPR2019



ECCV2020

Number of papers per area



Computer Science and Technology 计算机学院 周文晖

深度学习已经成为计算机视觉中的标配？

- 首要原因是深度学习可以做到传统方法无法达到的精度；
- 深度学习算法的通用性强；比如faster RCNN在人脸、行人、一般物体检测任务上都可以取得非常好的效果；
- 工程开发、优化、维护成本低。深度学习计算主要是卷积和矩阵乘，针对这种计算优化，所有深度学习算法都可以提升性能。

深度学习面临的挑战

Deep Learning: A Critical Appraisal
<https://arxiv.org/abs/1801.00631>

- 1、截至目前，深度学习依然严重依赖于数据集的质量和数量；
- 2、截至目前，深度学习迁移能力有限；它所学会的，只是特定场景下的特例。
- 3、截至目前，深度学习还不够透明；神经网络的“黑箱”性质至今没有解决。
- 4、深度学习还未与先验知识结合
-

- MIT 联合 IBM 研究团队提出了一个数据集，图像识别 SOTA 模型的性能下降了 40 多个点

ImageNet

Chairs



ObjectNet

Chairs by rotation



Chairs by background



Chairs by viewpoint



Teapots



T-shirts



Nature发文：深度学习系统为什么这么好骗？

2019年10月15日 14:46:20

来源：机器之心 大风号

0人参与

0评论

FOOLING THE AI

Deep neural networks (DNNs) are brilliant at image recognition — but they can be easily hacked.

These stickers made an artificial-intelligence system read this stop sign as 'speed limit 45'.

Stop

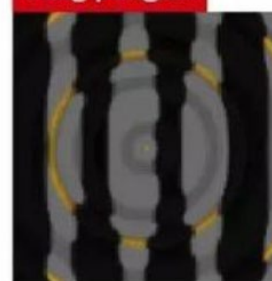


Speed limit 45



Scientists have evolved images that look like abstract patterns — but which DNNs see as familiar objects.

King penguin



Starfish



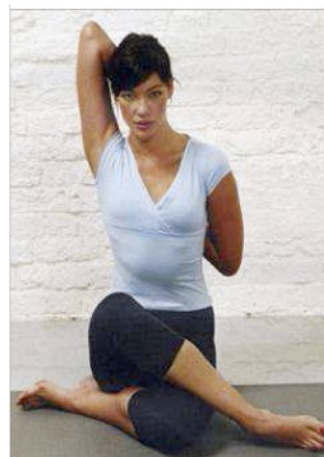
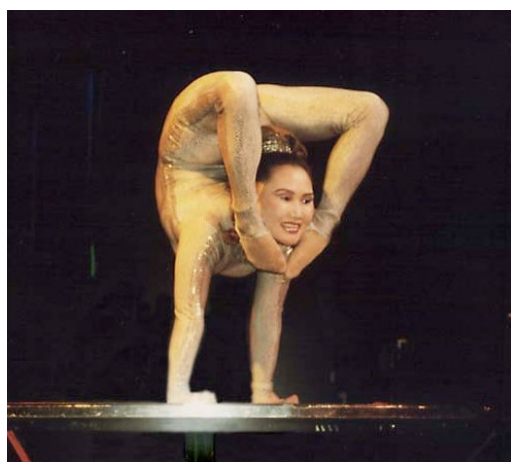
©nature

传统计算机视觉面临的挑战：外观、大小和形状



图像中的人存在各种形状、大小和外观

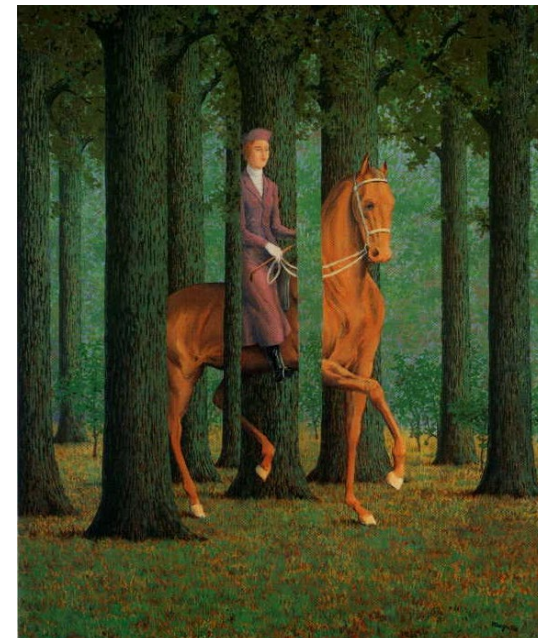
传统计算机视觉面临的挑战：复杂姿态/运动



传统计算机视觉面临的挑战：复杂和不可预测的行为



传统计算机视觉面临的挑战：噪声和遮挡



传统计算机视觉面临的挑战：外观变化



Hang

院 周文晖

传统计算机视觉面临的挑战：上下文间依赖性



传统计算机视觉面临的挑战：视点变化

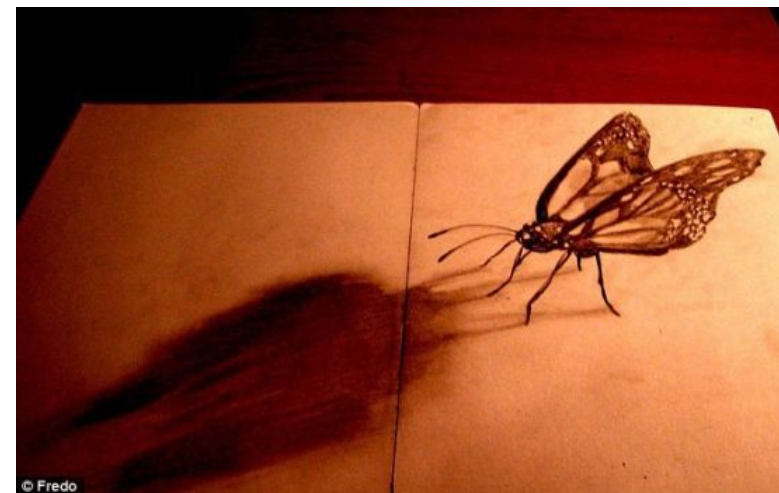
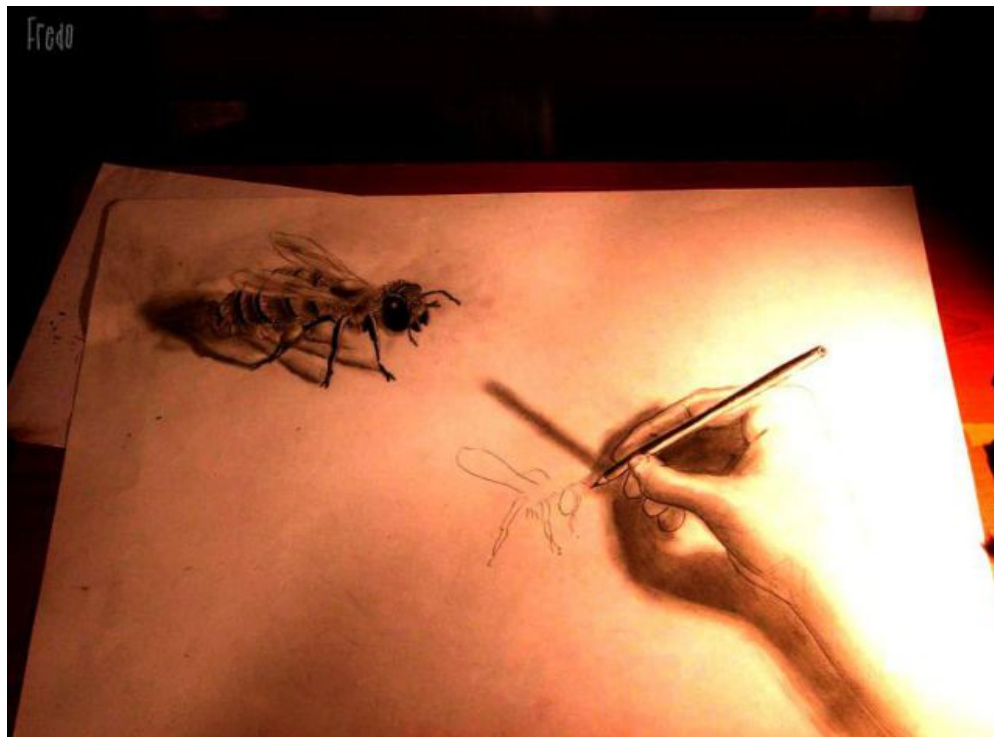


挑战或机遇？

- 虽然计算机视觉面临诸多挑战，但图像隐含着许多有用的线索，可以解释和描述场景结构。



智利Fredo的3D铅笔素描



深度线索：直线透视

- 平行直线在图像平面上相交



深度线索：空间透视

- 浓淡远近法



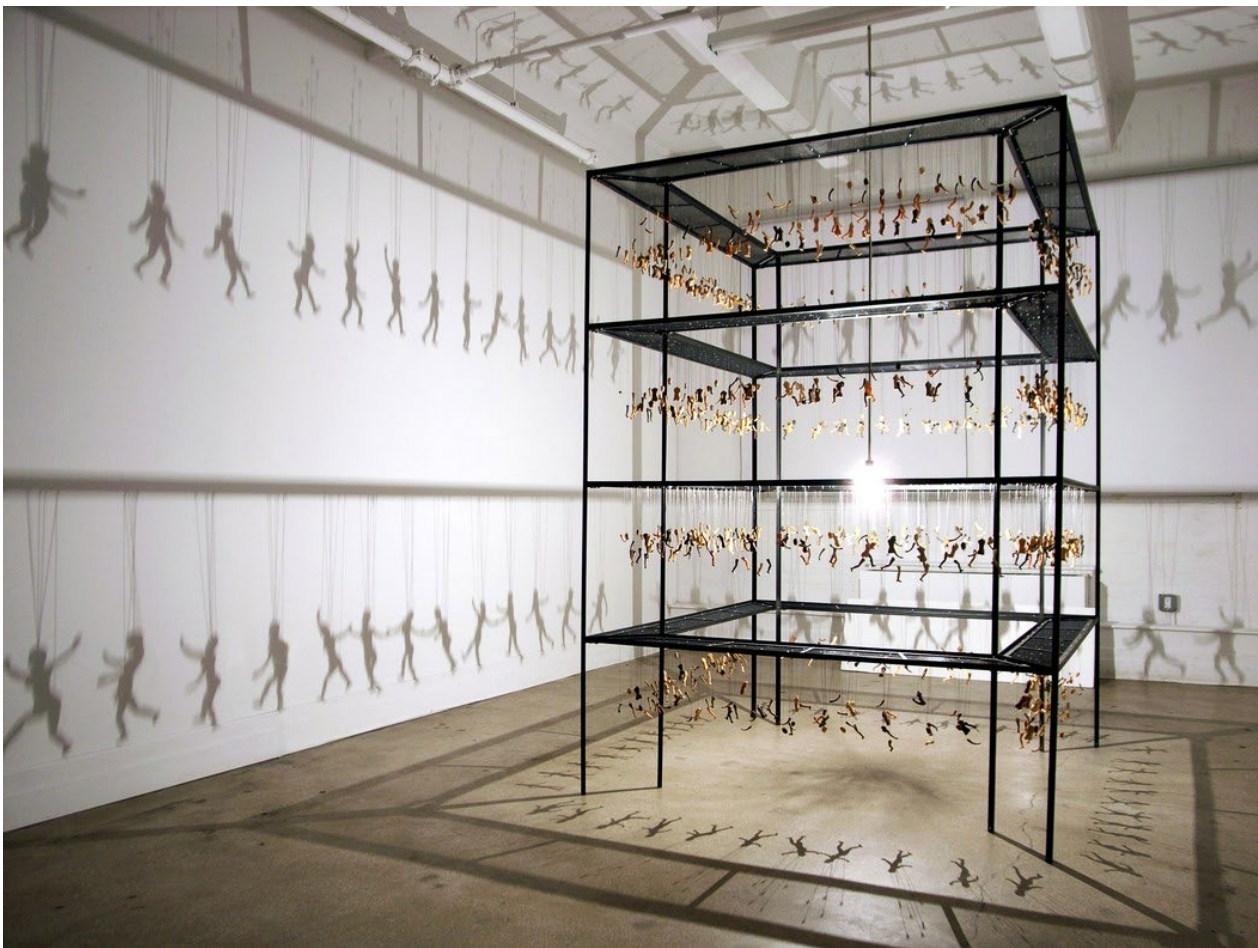
远近顺序线索：遮挡



形状线索：纹理梯度



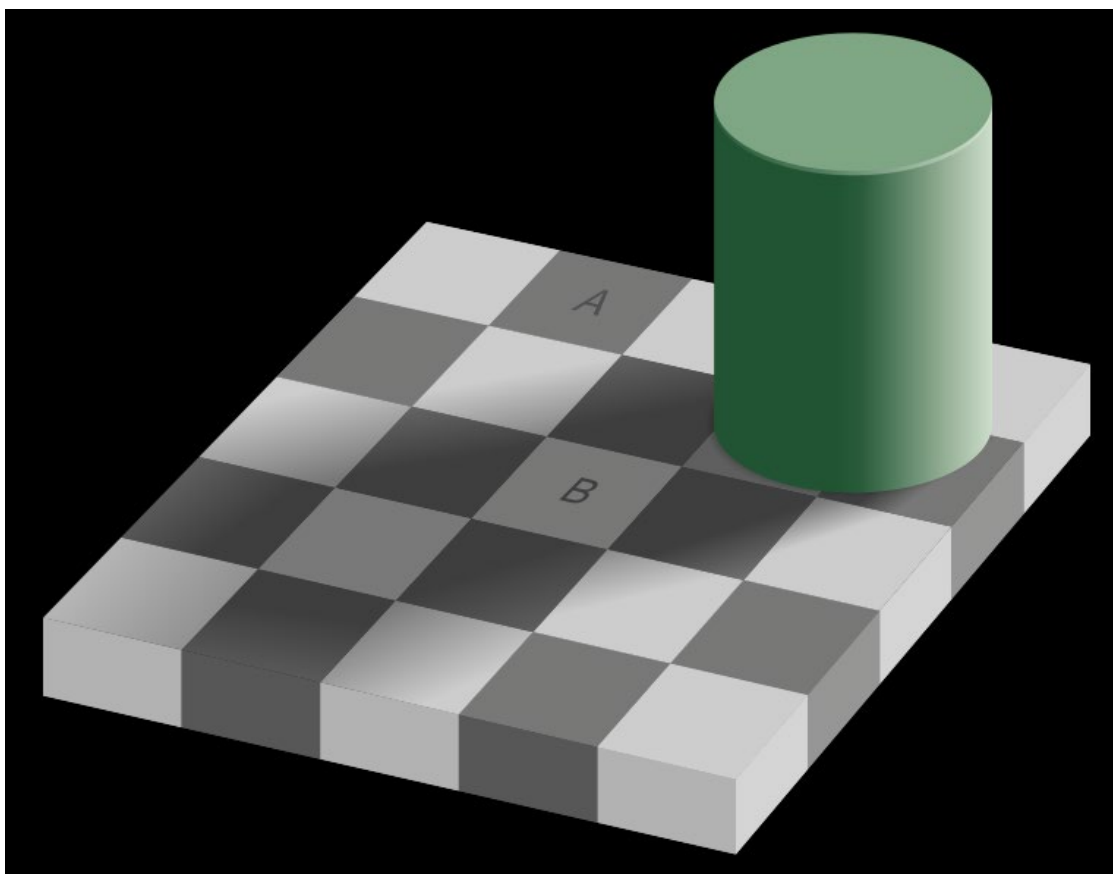
形状和光照线索：阴影



位置和光照线索：投影



视觉错觉



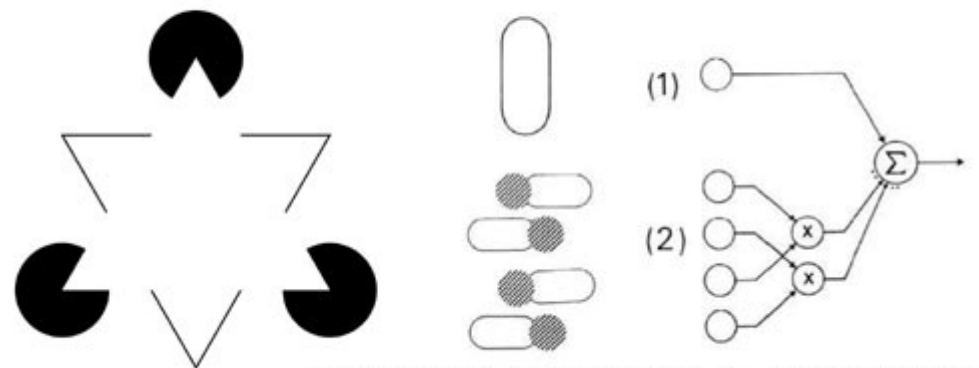
棋盘格亮度错觉。图中A，B两块区域的亮度是一致的

https://commons.wikimedia.org/wiki/File:Grey_squares_optical_illusion.svg

视觉错觉

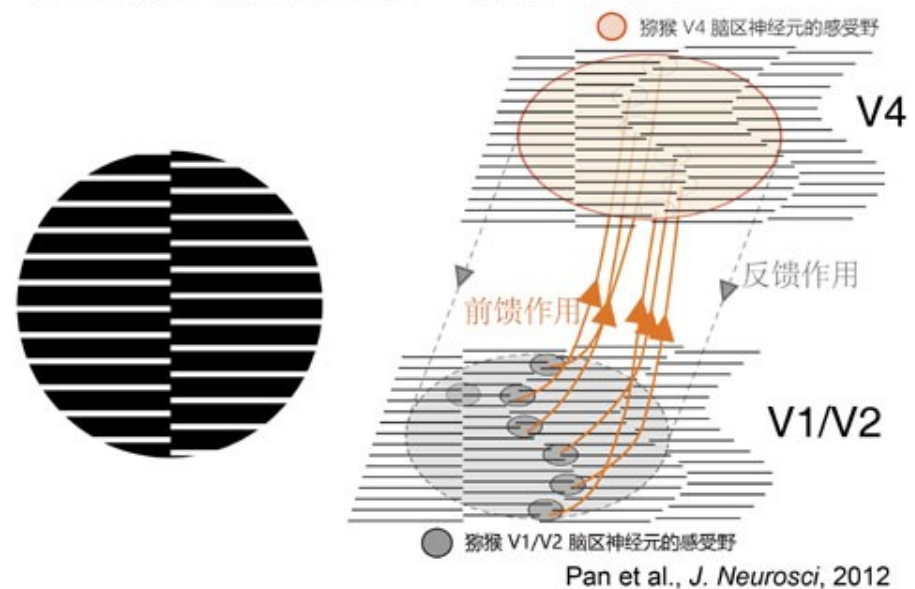
两个著名的错觉轮廓。a图中能够看到一个比背景亮度更亮的白色三角形。b图中能够看到一根竖直的细线轮廓。人们已经发现了灵长类的视觉皮层存在编码对两种错觉轮廓有反应的神经元。

a. Kanisza 三角形以及可能的神经信息处理机制



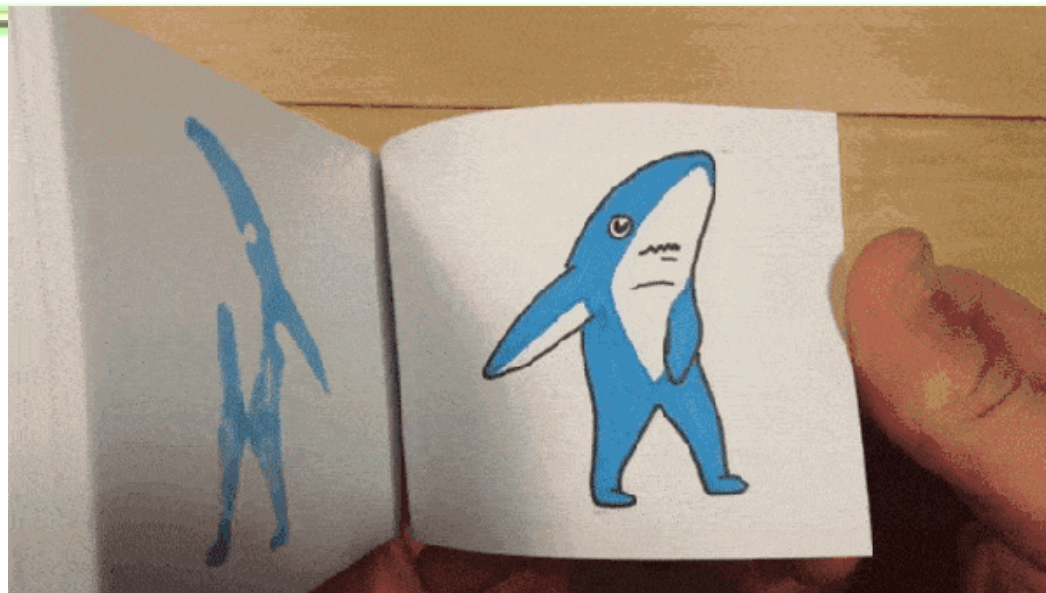
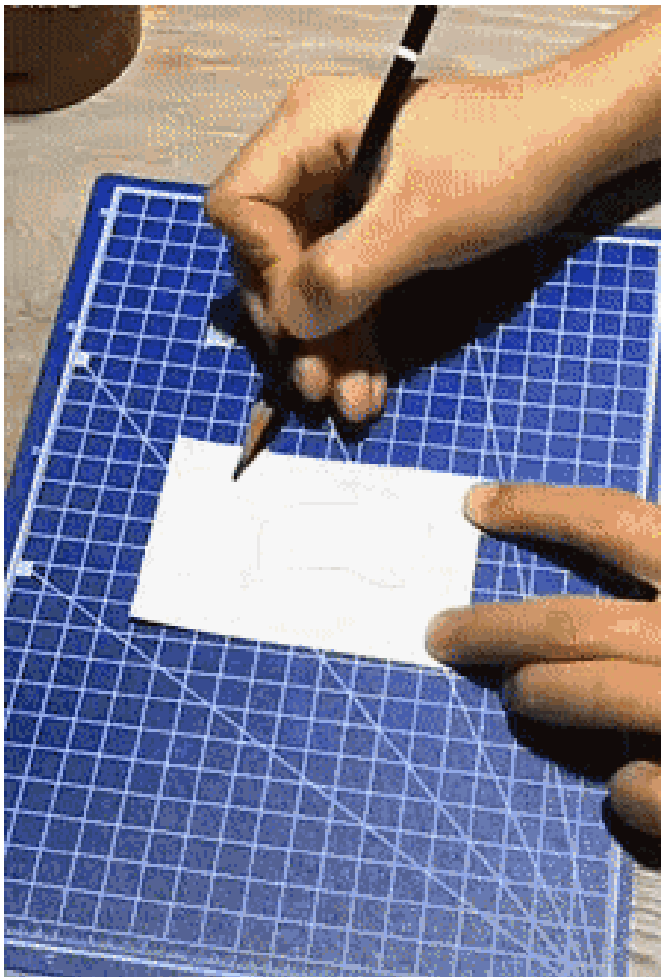
Peterhans, E. and von der Heydt, R., *J. Neurosci.* 1989

b. 短线错觉以及可能的神经信息处理机制



Pan et al., *J. Neurosci.* 2012

手绘翻页动画



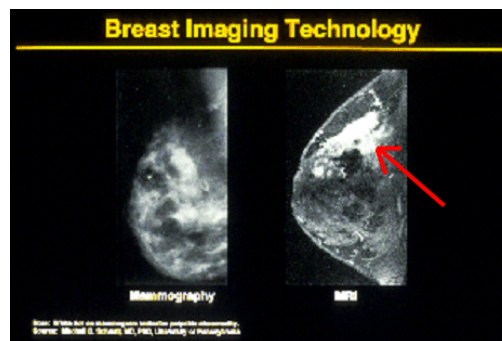
研究方法

- 仿生学方法
 - 仿生模型 -> 计算模型
- 工程方法
 - 数学、信号处理等方法

计算机视觉应用领域



交通安全



生物医学



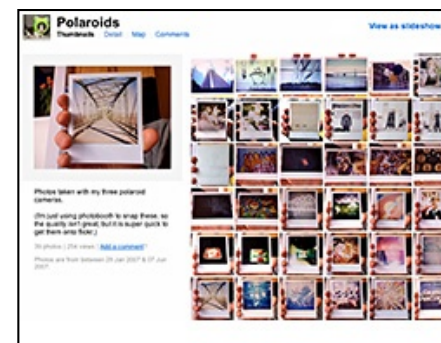
安防监控



辅助服务

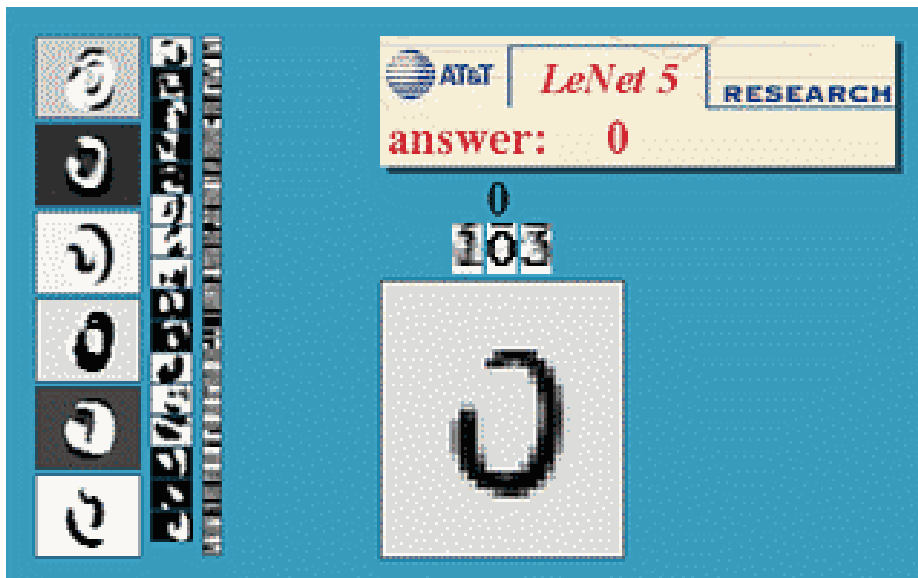


影视娱乐



人机交互

OCR（光学字符识别）



Digit recognition, AT&T labs

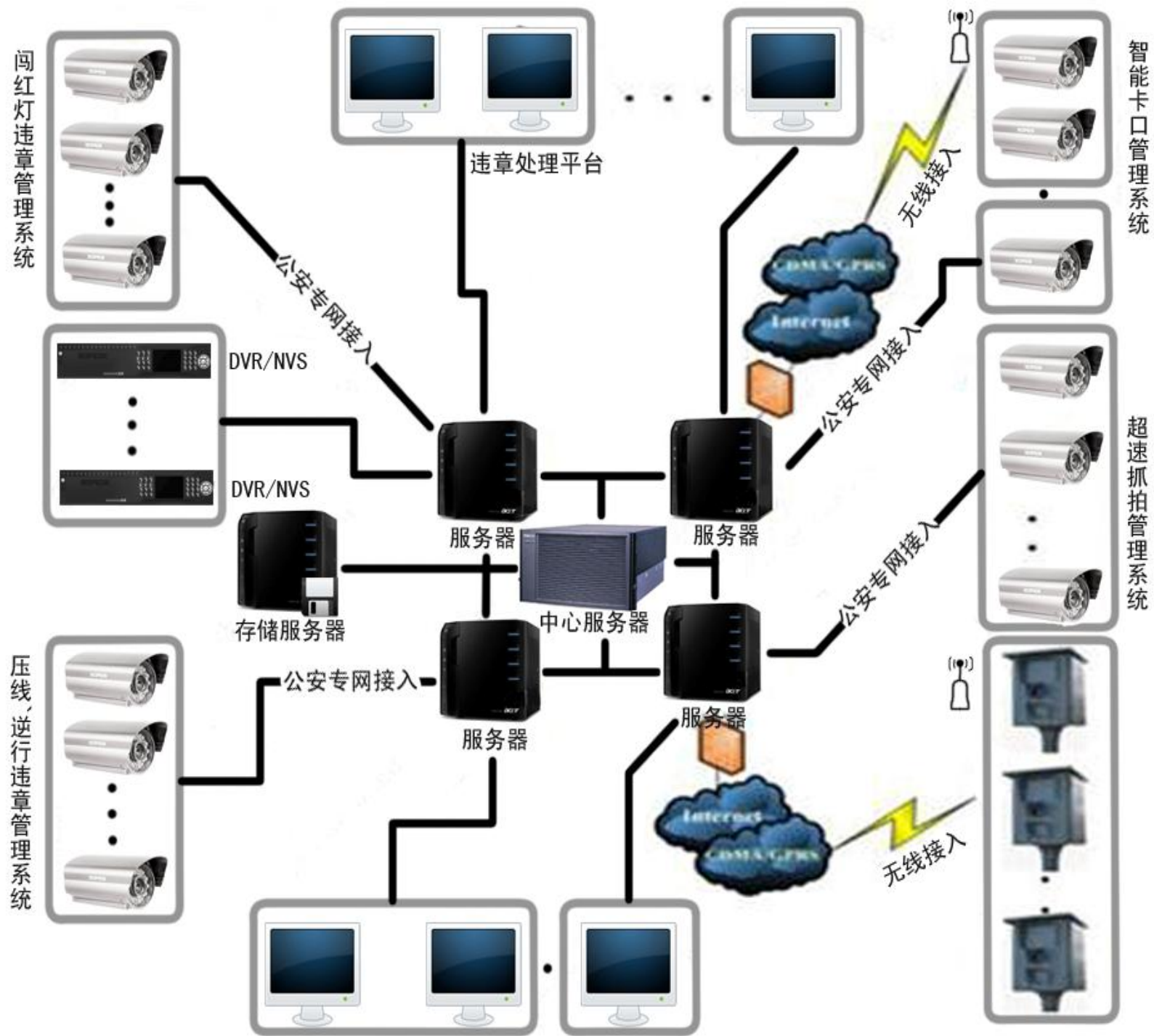
<http://www.research.att.com/~yann/>



License plate readers

http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

智能交通



人脸检测

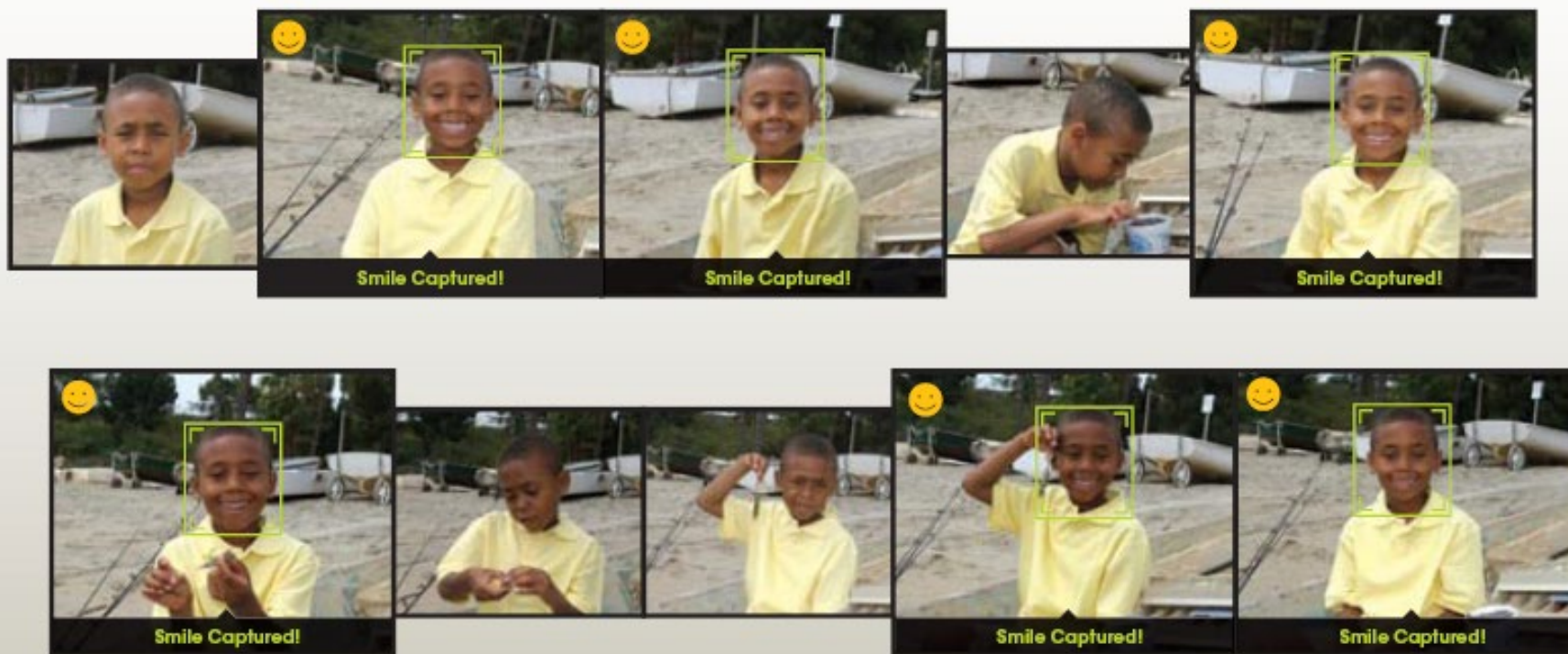


- ▶ 大部分数码相机都具有人脸检测和定位功能
 - Canon, Sony, Fuji, ...

表情识别

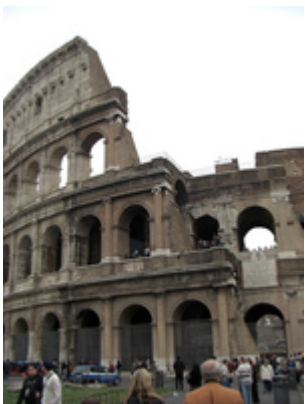
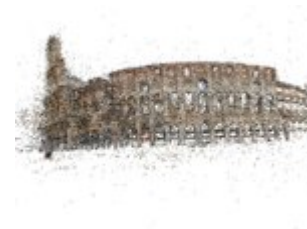
The Smile Shutter flow

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.

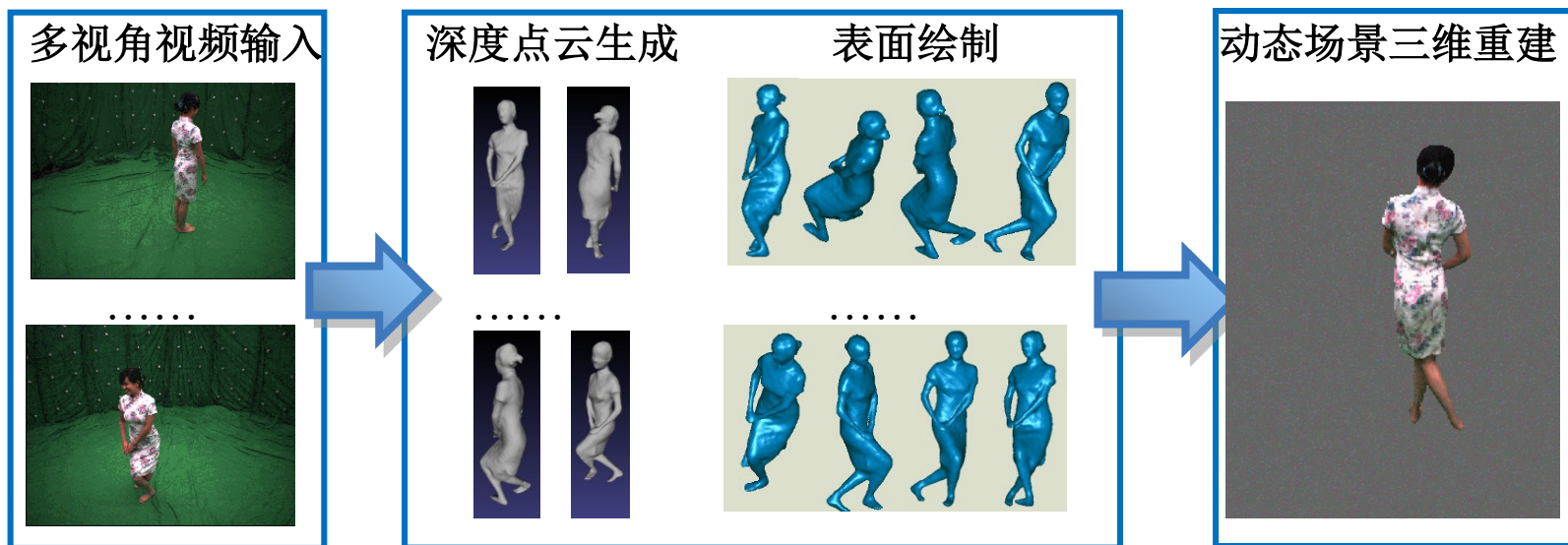


多视点三维重建

- Building Rome in a Day



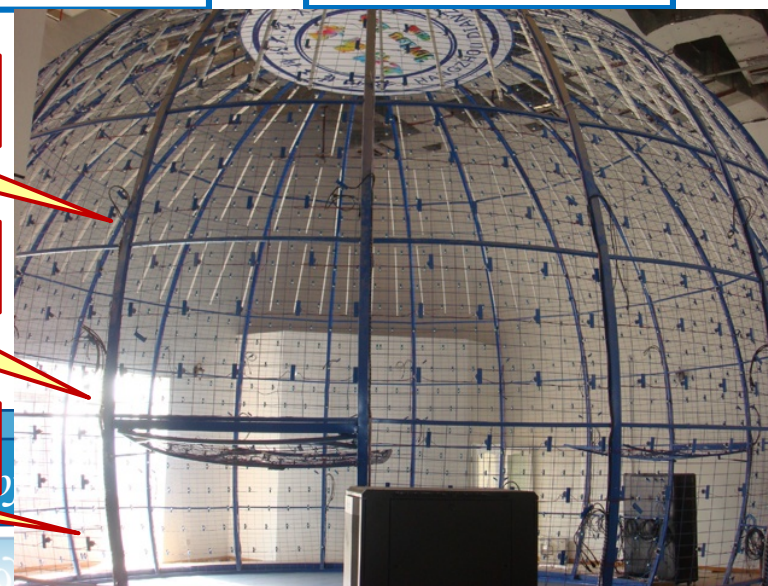
多视点三维重建



直径10米、高7米

50个CCD相机

2000个LED



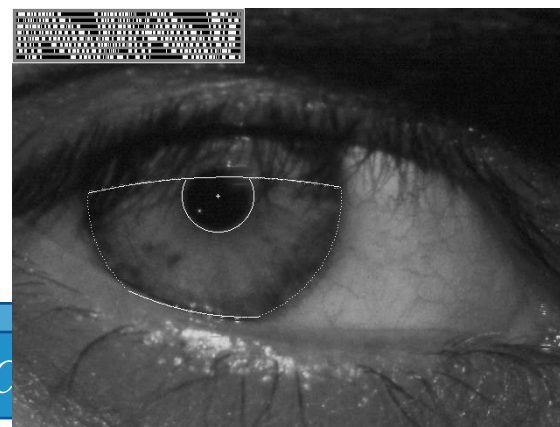
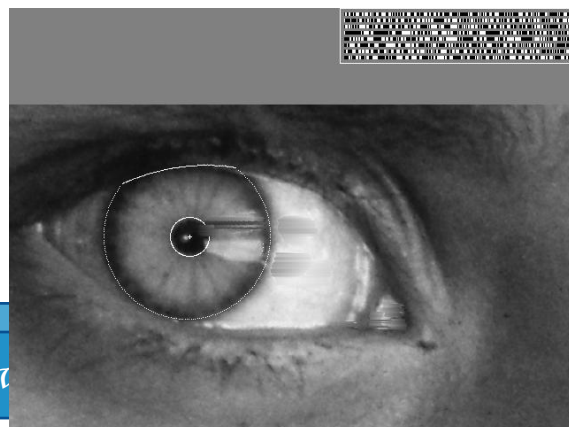
基于视觉的生物识别

- 虹膜识别



“How the Afghan Girl was Identified by Her Iris Patterns”

Read the [story wikipedia](#)

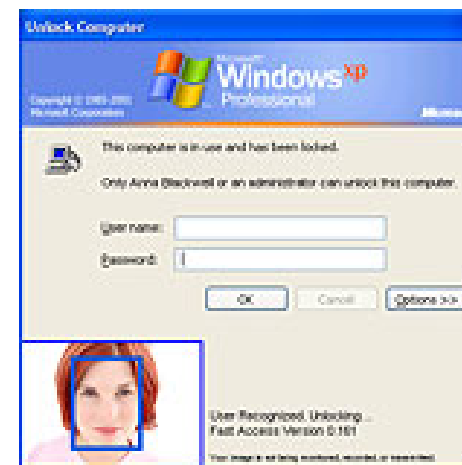
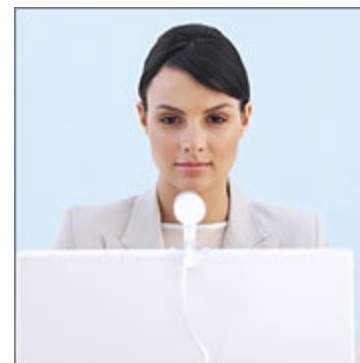


基于视觉的生物识别

- 指纹识别



Fingerprint scanners on
many new laptops,
other devices



Face recognition systems now
beginning to appear more widely

<http://www.sensiblevision.com/>

辅助驾驶

The screenshot displays the Mobileye website's main navigation and product highlights. At the top, there are tabs for 'manufacturer products' and 'consumer products'. The central banner features the slogan 'Our Vision. Your Safety.' above a top-down view of a car with four camera fields of vision: 'rear looking camera', 'forward looking camera', and 'side looking camera'. Below the banner, three main product sections are visible: 'EyeQ Vision on a Chip' with an image of the chip, 'Vision Applications' showing a pedestrian on a crosswalk, and 'AWS Advance Warning System' with a circular sensor graphic. On the right side, there are sections for 'News' and 'Events', each containing headlines and a 'read more' link.

manufacturer products consumer products

Our Vision. Your Safety.

rear looking camera forward looking camera side looking camera

EyeQ Vision on a Chip

Vision Applications Road, Vehicle, Pedestrian Protection and more

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

all news

Events

- Mobileye at Equip Auto, Paris, France
- Mobileye at SEMA, Las Vegas, NV

read more

► Mobileye

Hangzhou Dianzi University 杭州电子科技大学 School of Computer Science and Technology 计算机学院 周文晖

已应用在宝马、英菲尼迪、沃尔沃等的高端车型。

无人驾驶汽车

- DARPA “Grand Challenge” 2004/2005

10小时内跑完132英里



无人驾驶汽车

- DARPA “Urban Challenge” 2007

要求在都市环境自动驾驶，遵守交通规则。并根据交通状况做出适当反应。



Google无人驾驶汽车



<http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence>

Autonomous Driving

Google's modified Toyota Prius uses an array of sensors to navigate public roads without a human driver. Other components, not shown, include a GPS receiver and an inertial motion sensor.

LIDAR

A rotating sensor on the roof scans more than 200 feet in all directions to generate a precise three-dimensional map of the car's surroundings.

POSITION ESTIMATOR

A sensor mounted on the left rear wheel measures small movements made by the car and helps to accurately locate its position on the map.

VIDEO CAMERA

A camera mounted near the rear-view mirror detects traffic lights and helps the car's onboard computers recognize moving obstacles like pedestrians and bicyclists.

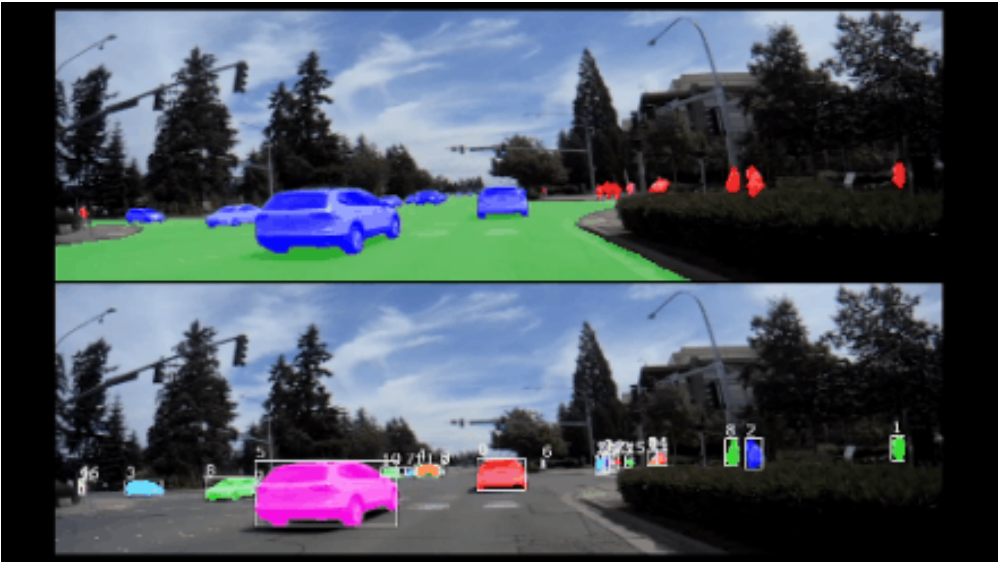


RADAR

Four standard automotive radar sensors, three in front and one in the rear, help determine the positions of distant objects.



自动驾驶/辅助驾驶Adas



上图中预测的目标和目标类别（蓝色=汽车；绿色=可驾驶空间；红色=行人）

基于视觉的人机交互

- Kinect



<http://www.microsoft.com/en-us/kinectforwindows/develop/>



智能机器人

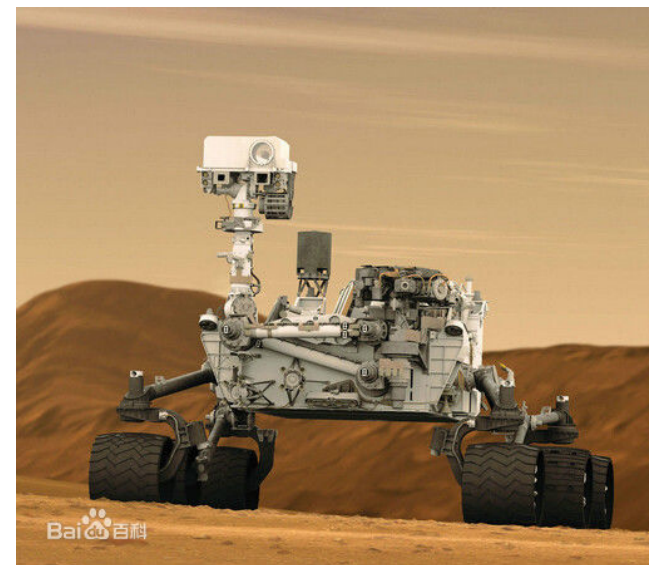
▶ 美国“雅典娜”系列火星探测器



NASA'S Mars Exploration Rover Spirit

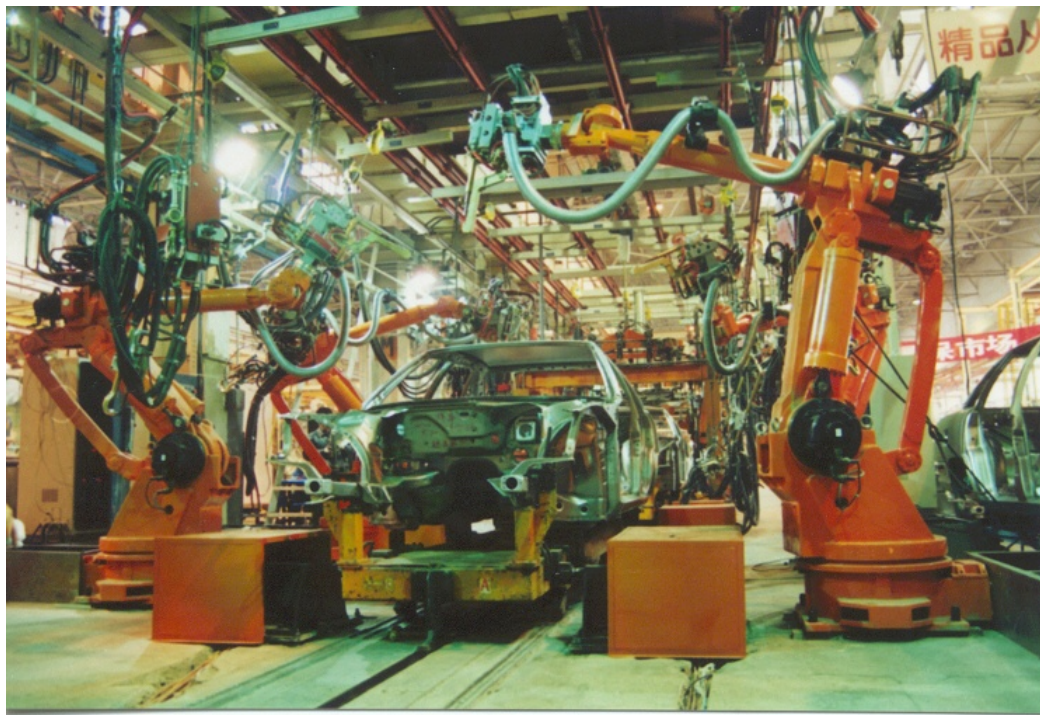
Matthies et al. “Computer Vision on Mars”

- 好奇号火星探测器，于2011年11月发射，2012年8月成功登陆火星表面。它是美国第七个火星着陆探测器，第四台火星车，也是世界上第一辆采用核动力驱动的火星车。

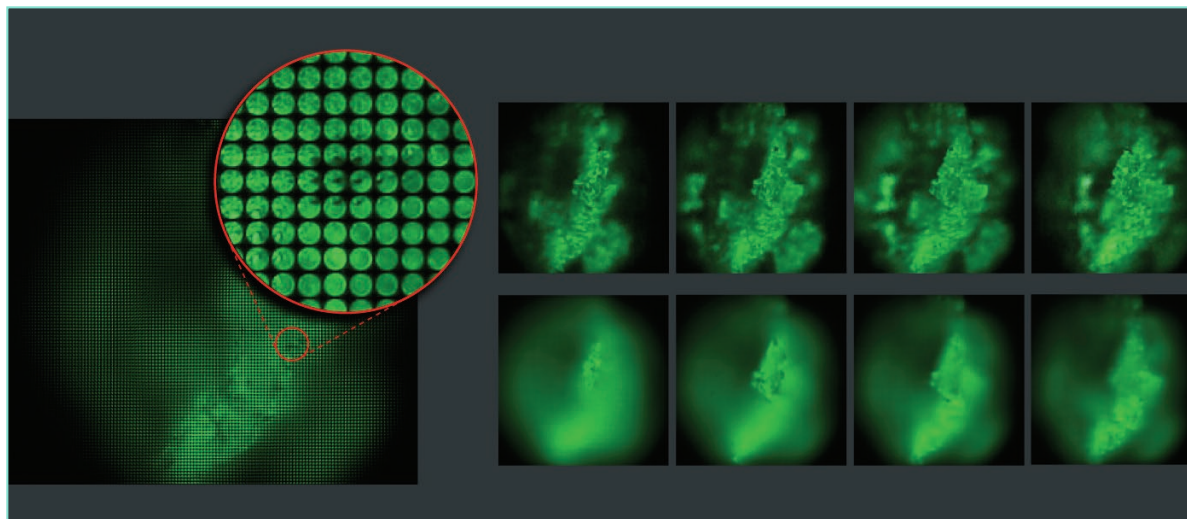
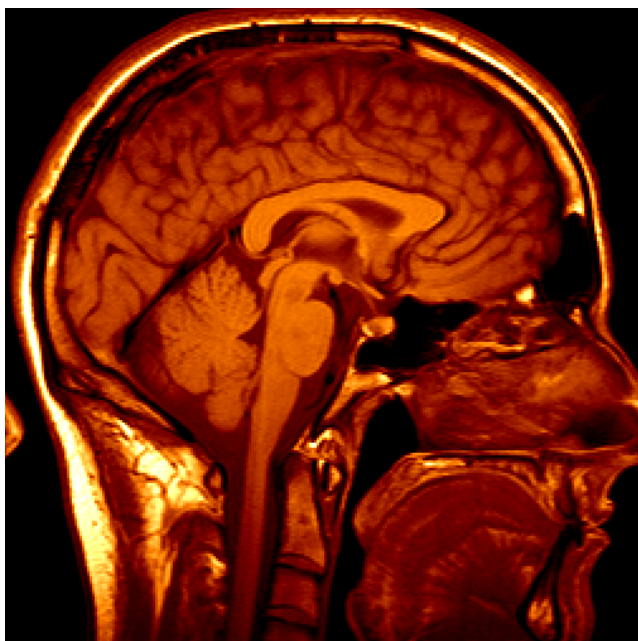


工业机器人

- 一汽和二汽点焊线



医学成像

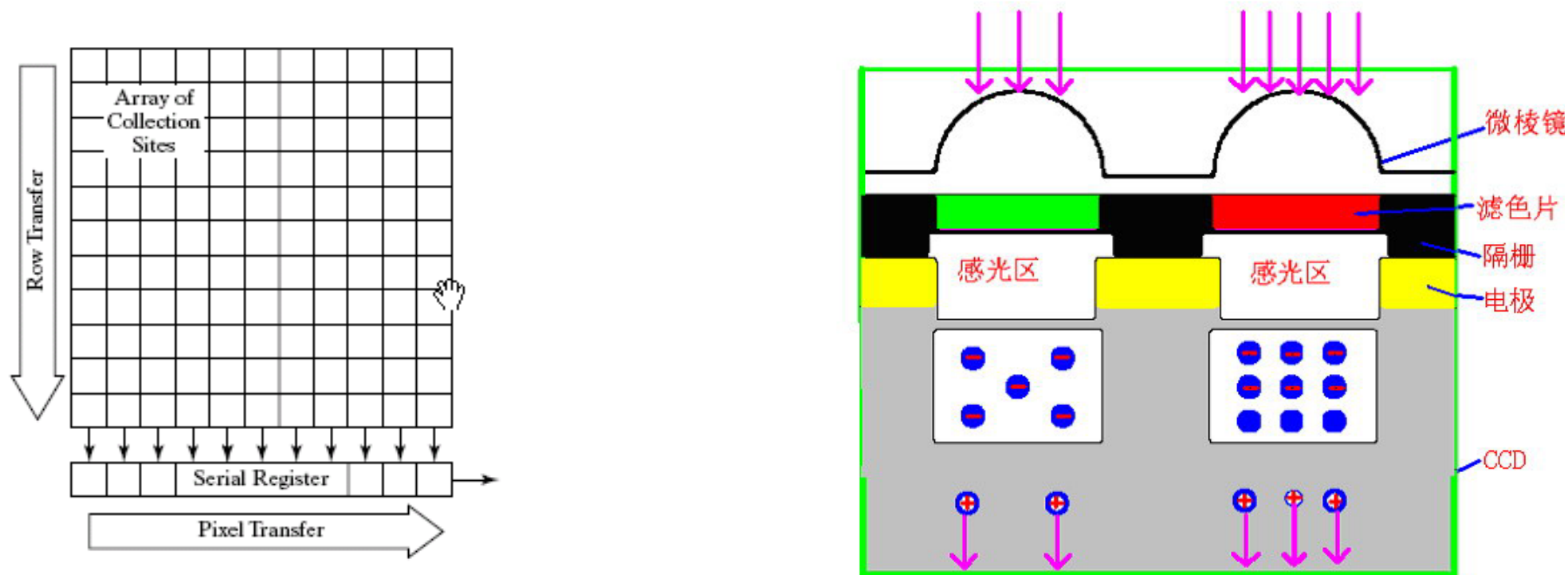
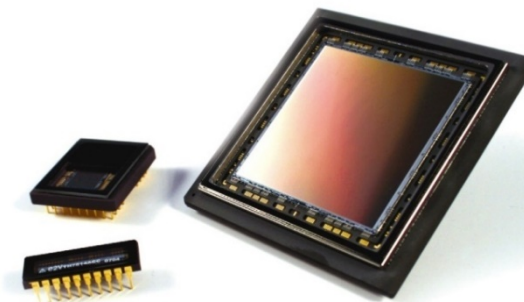


第一章 绪论

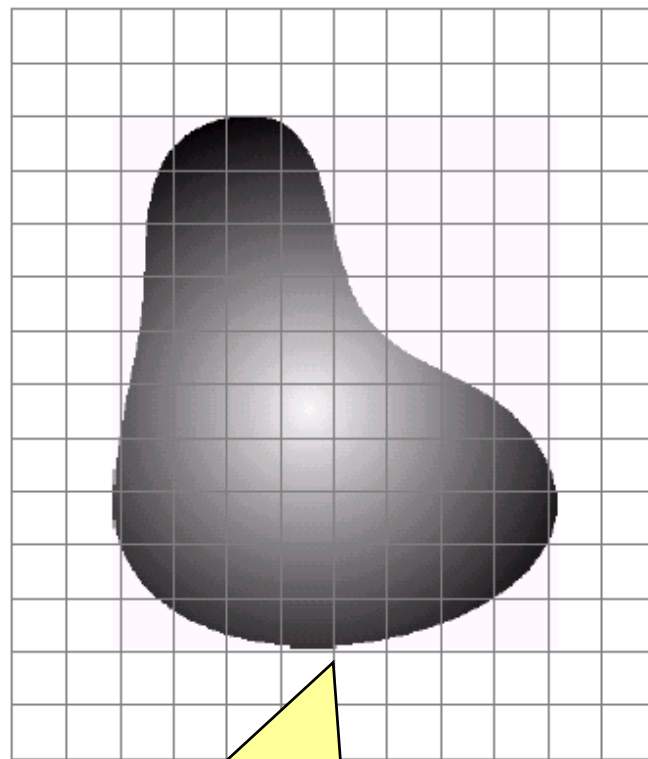
1.2 图像基础

CCD/CMOS成像原理

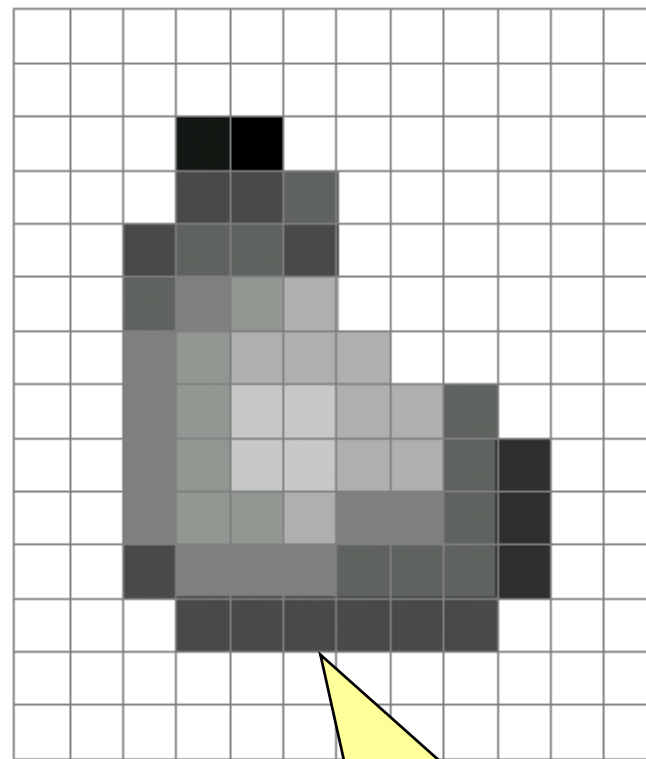
- 光电转换/光积分过程



采样和量化

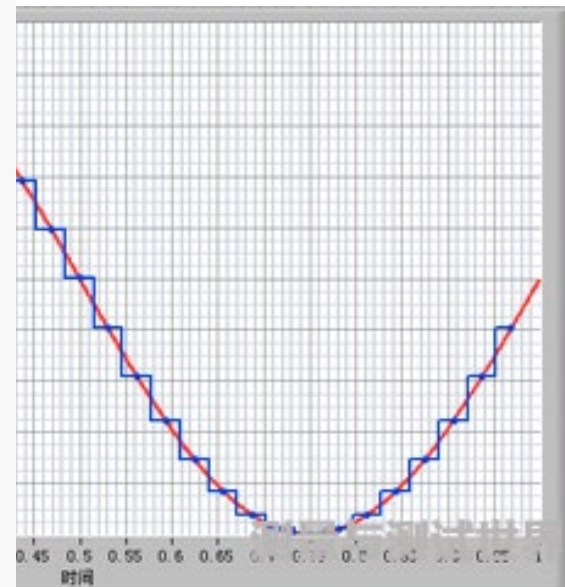
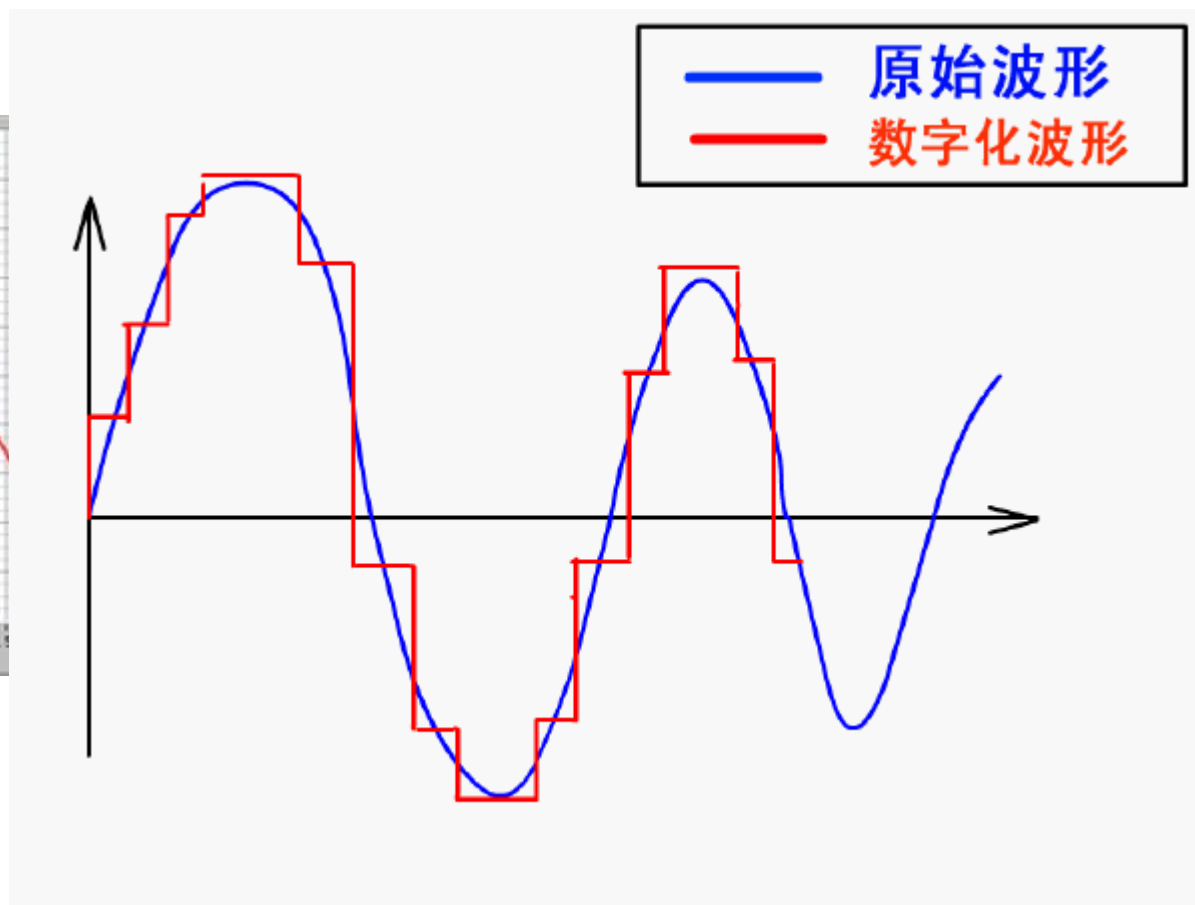
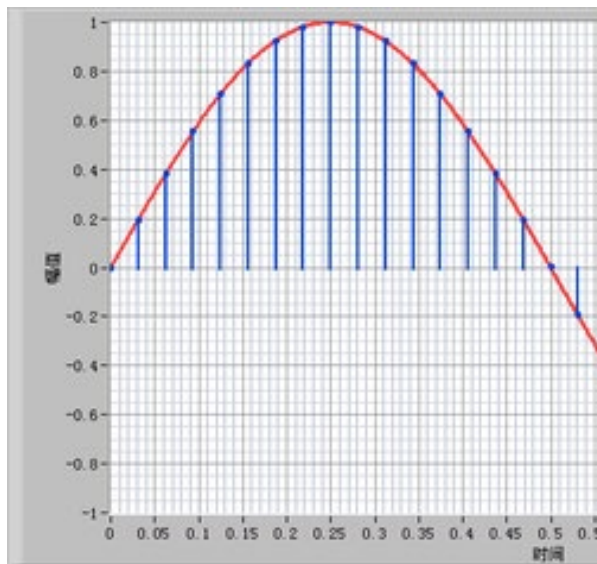


投影到传感器阵列
上的连续图像

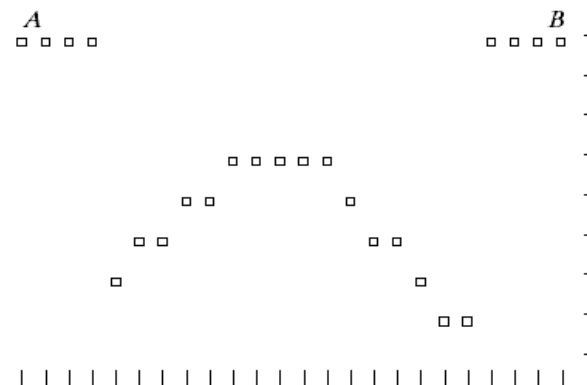
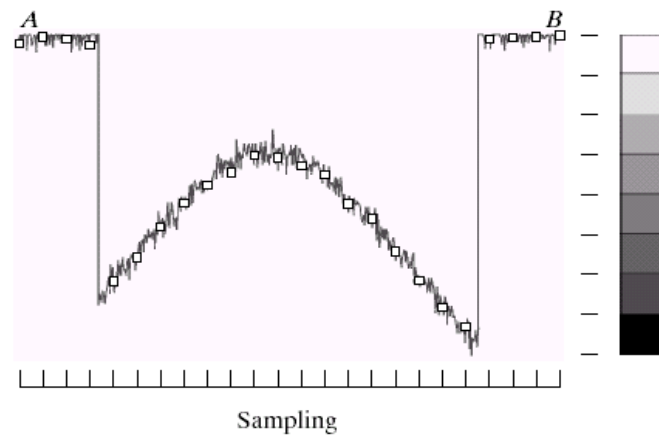
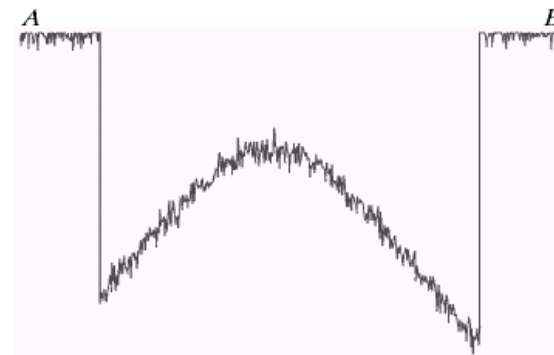
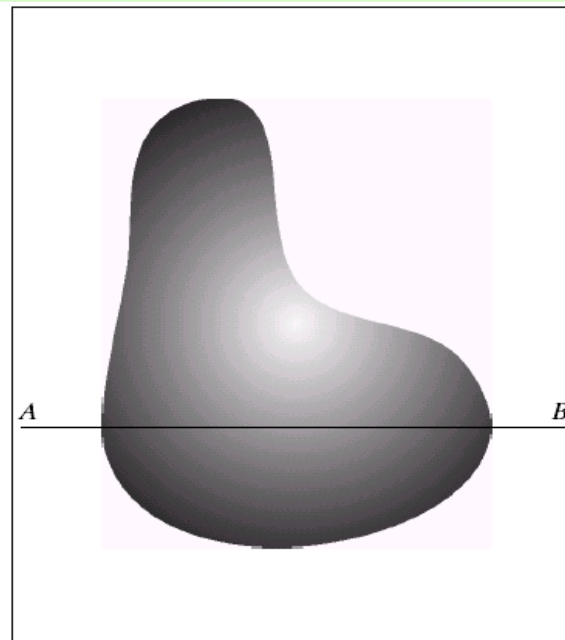


取样、量化后
的数字图像

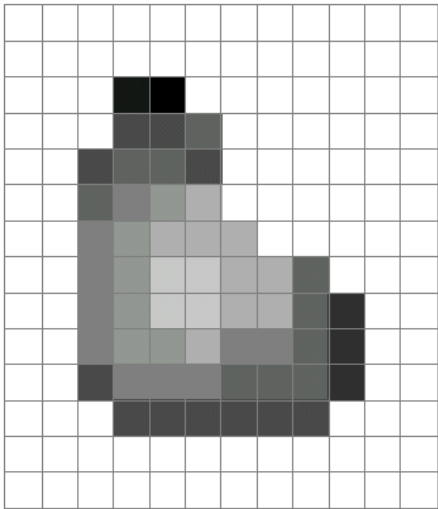
采样和量化



采样和量化



采样和量化 → 图像表示



=

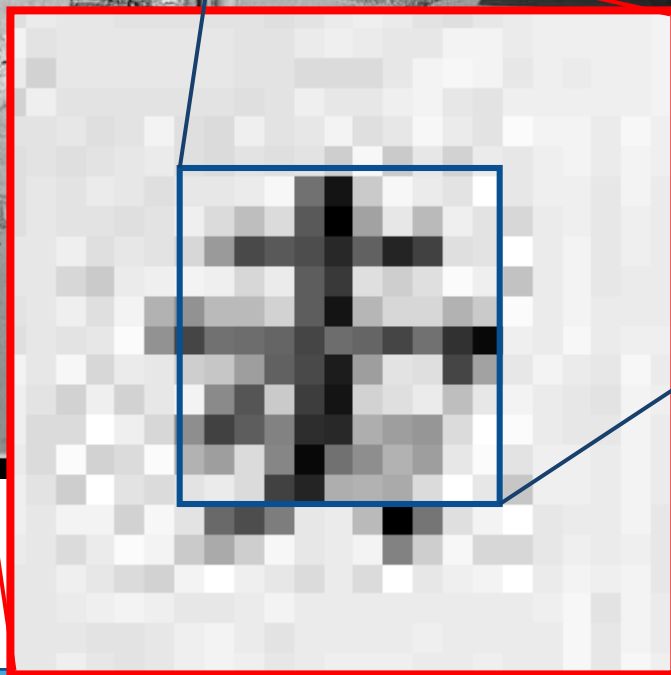
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	20	0	255	255	255	255	255	255	255
255	255	255	75	75	75	255	255	255	255	255	255
255	255	75	95	95	75	255	255	255	255	255	255
255	255	96	127	145	175	255	255	255	255	255	255
255	255	127	145	175	175	175	255	255	255	255	255
255	255	127	145	200	200	175	175	95	255	255	255
255	255	127	145	200	200	175	175	95	47	255	255
255	255	127	145	145	175	127	127	95	47	255	255
255	255	74	127	127	127	95	95	95	47	255	255
255	255	255	74	74	74	74	74	74	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255	255	255	255	255

每个值为一个字节，0 = black, 255 = white

图像表示



卖
部



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

不同采样率测试

图像分辨率



256x256像素



128x128像素

空间分辨率：数字化的空间采样点数



64x64像素



32x32像素

不同量化灰阶测试

像素取



256级灰度



32级灰度

幅度分辨率：采样点值的量化级数

图

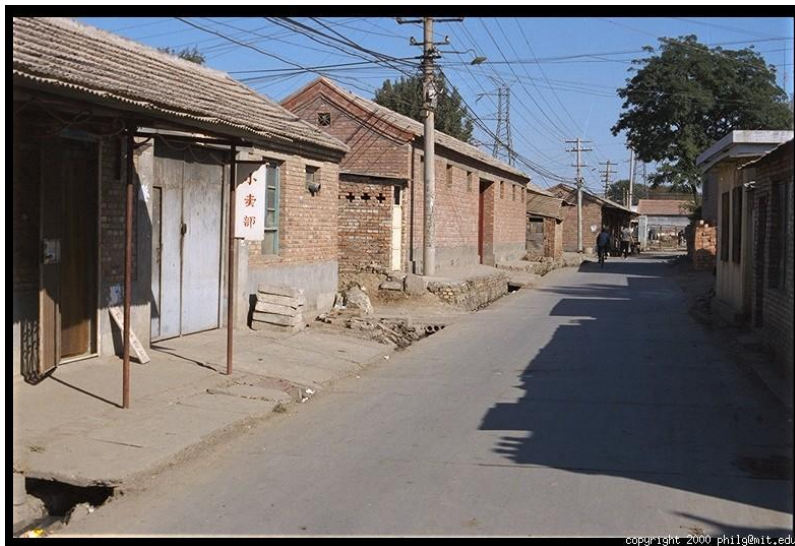


8级灰度

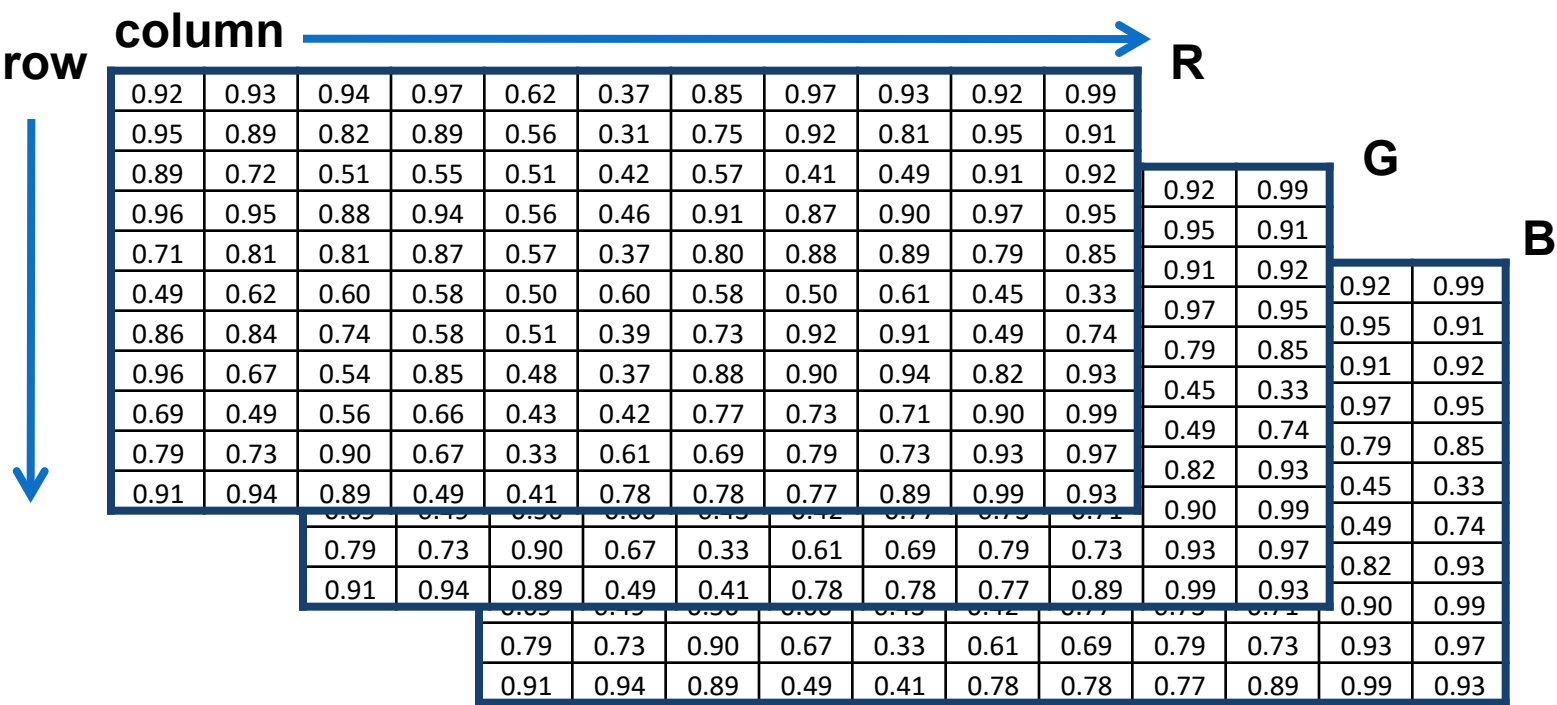


2级灰度

彩色图像



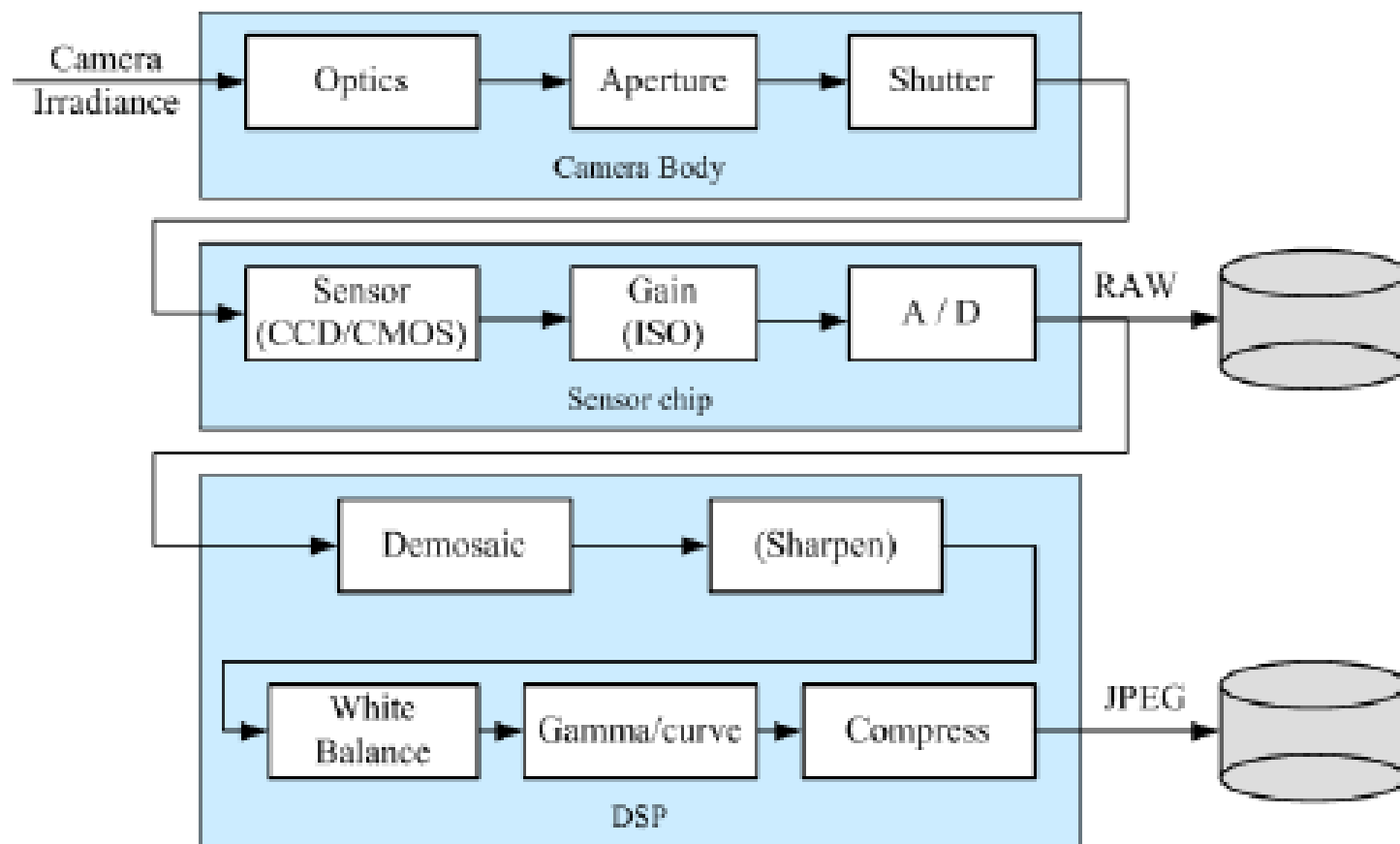
彩色图像表示



常见颜色的RGB组合

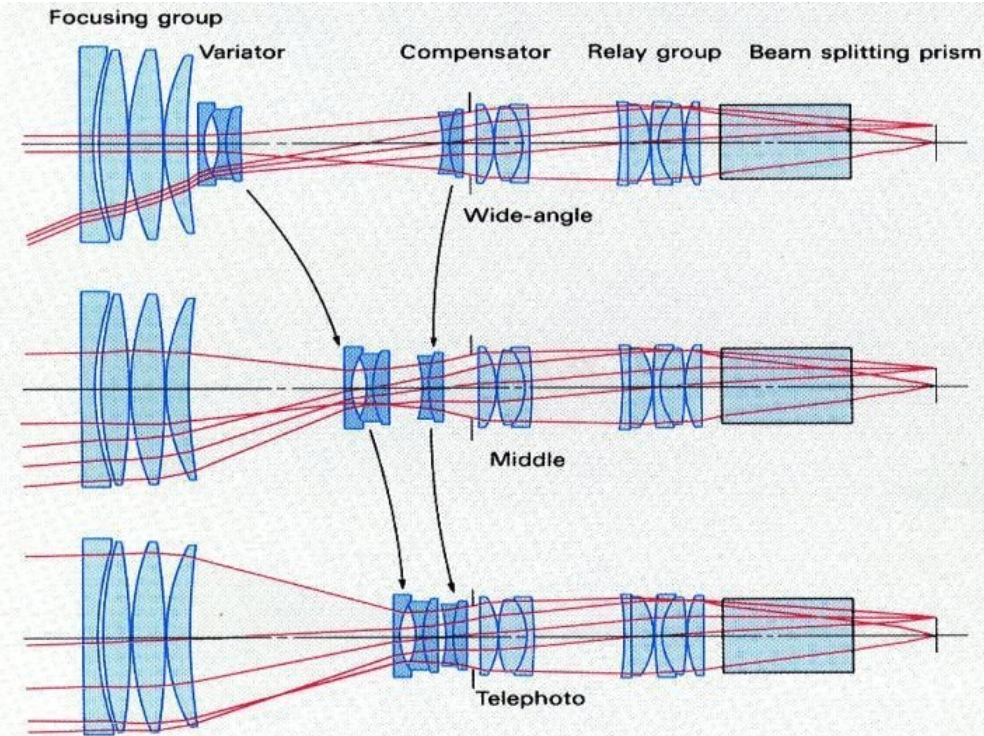
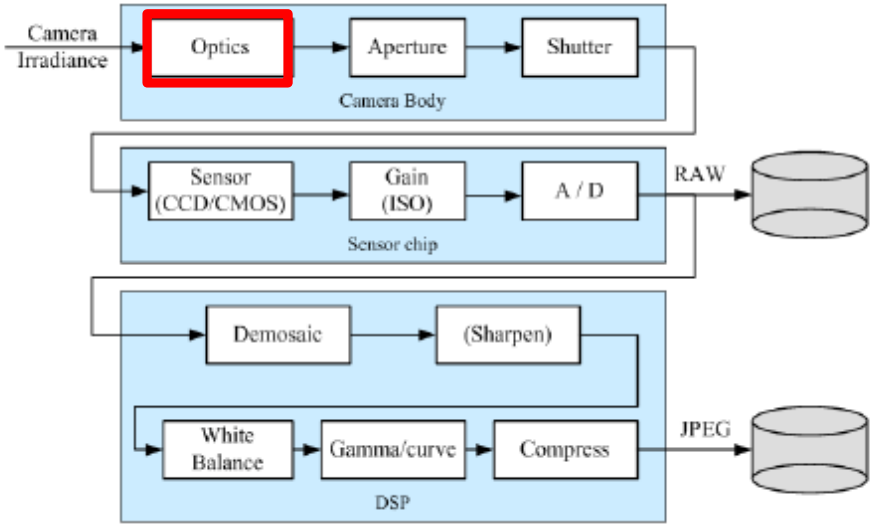
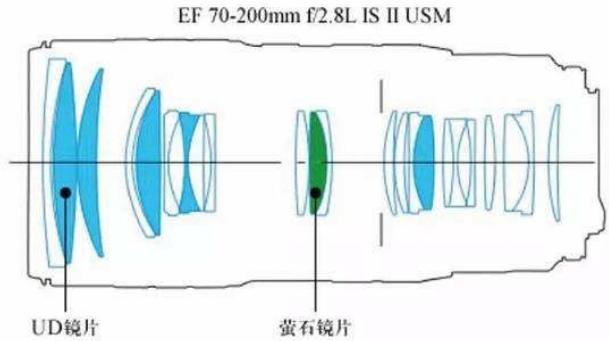
颜色	红色分量R	绿色分量G	蓝色分量B
黑色	0	0	0
白色	255	255	255
灰色	128	128	128
红色	255	0	0
绿色	0	255	0
蓝色	0	0	255
青色	0	255	255
紫色	255	0	255
黄色	255	255	0

像机成像流程



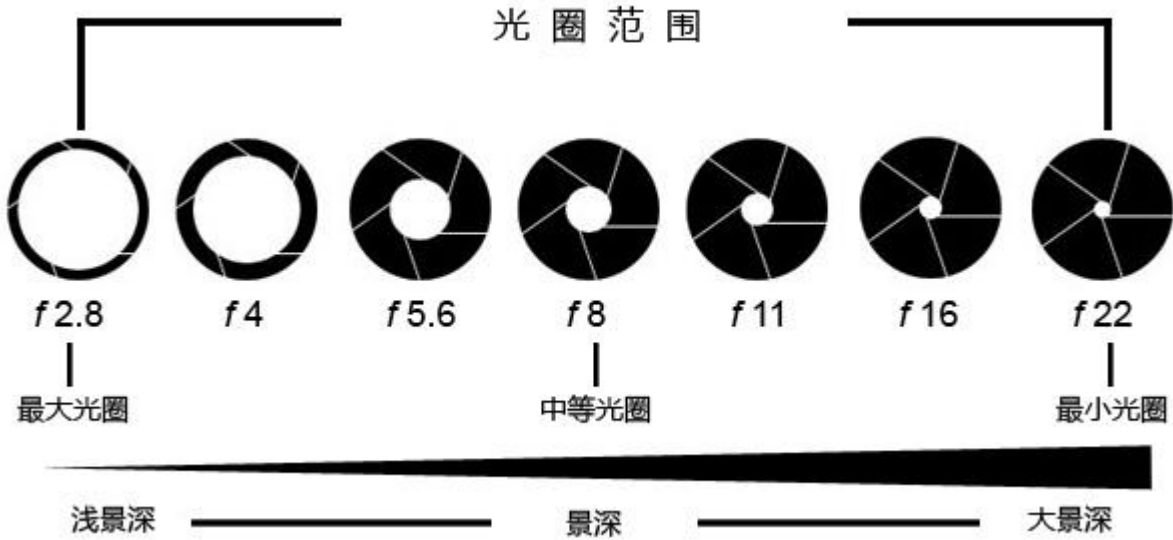
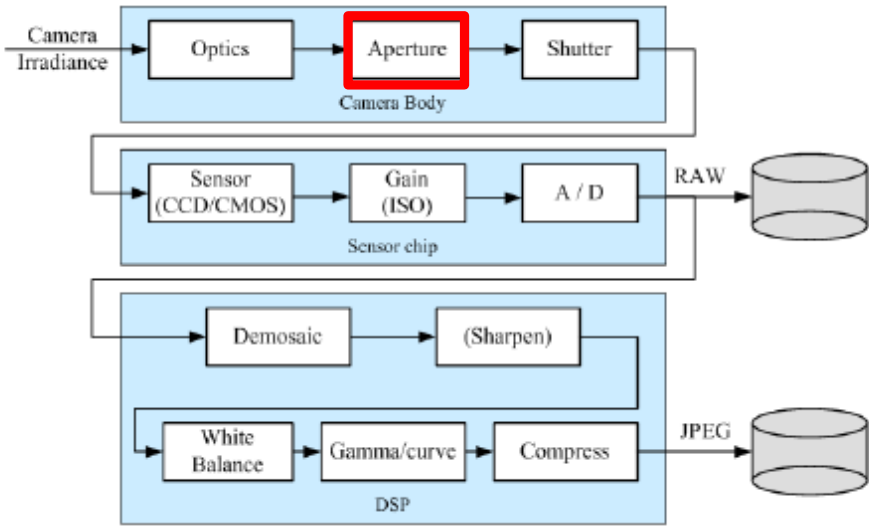
像机成像流程——镜头组

- 镜头组



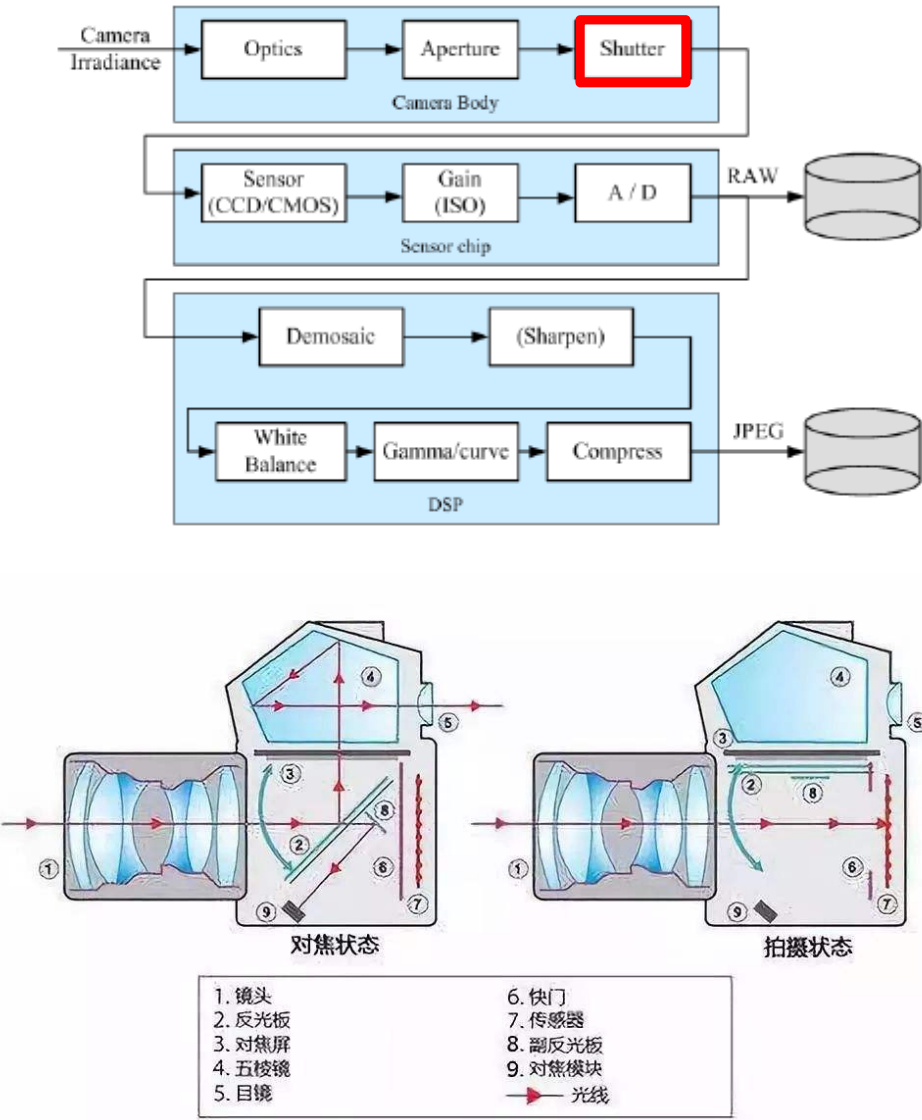
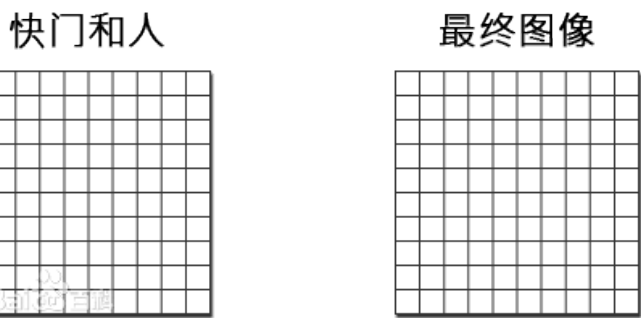
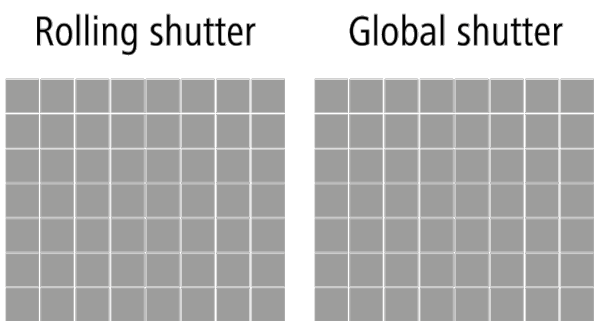
像机成像流程——光圈

- 光圈



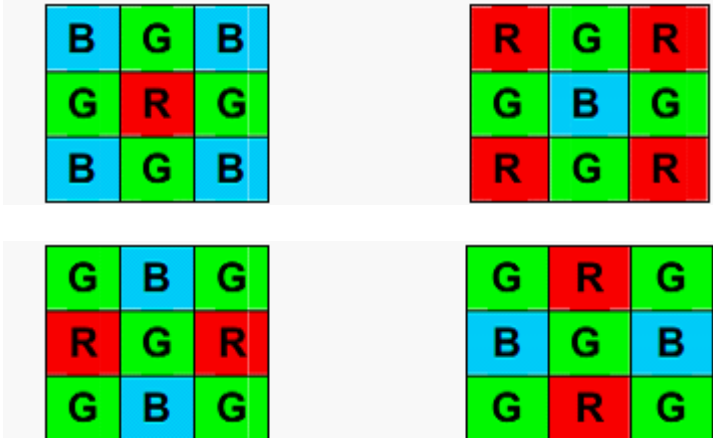
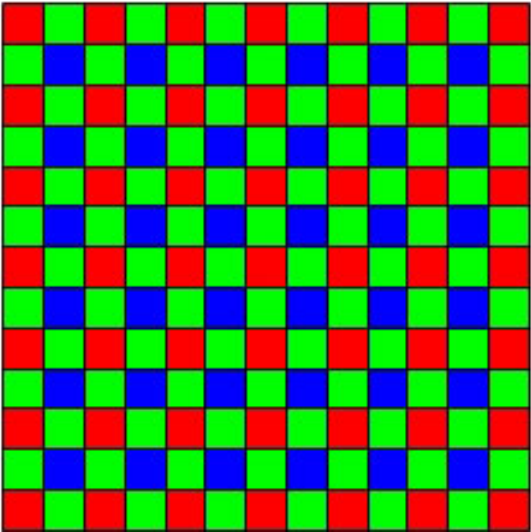
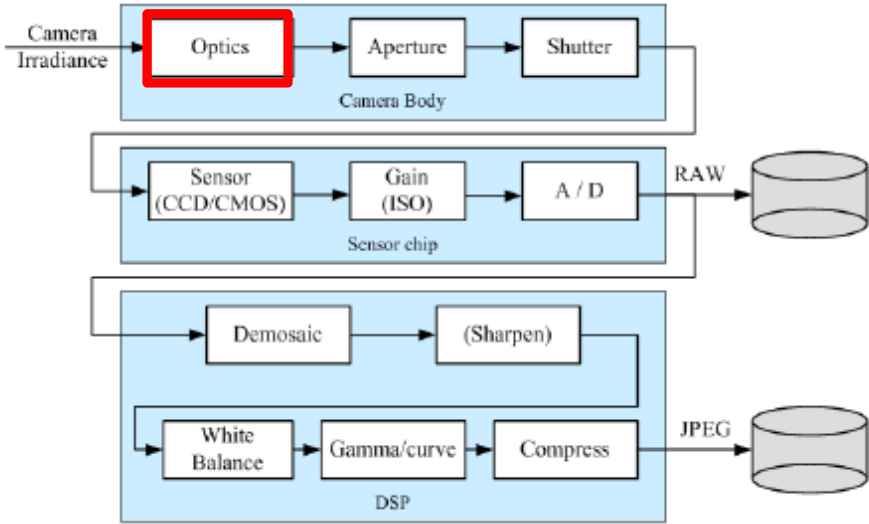
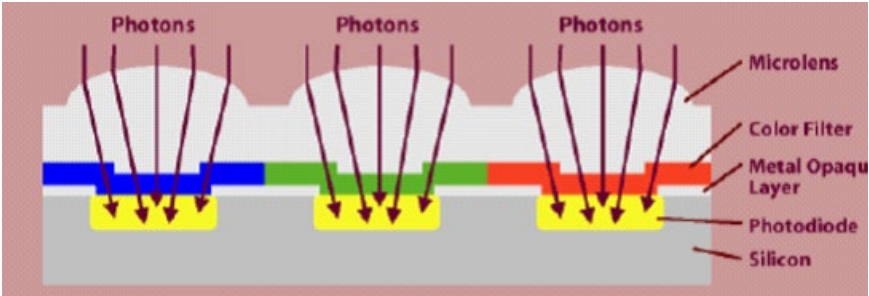
像机成像流程——快门

• 快门



像机成像流程——Bayer格式

- RAW和Demosaic



数字图像表示

- 一幅图像可以定义为一个二维函数 $f(x,y)$ ，空间坐标 (x,y) 上的值 f 称为该点图像的灰度。

- 矩阵表示:

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

矩阵中的每个元素对应着图像的一个点或区域，称为像素（**pixel**）。

数字图像的存储

- 数组存储方式
 - 二维数组 $f(x,y) = I[y][x]$
 - 一维数组 $f(x,y) = I[y*width+x]$
- 数组类型（8bits）
 - BYTE 或 unsigned char
- 指针表示：内存分配同数组存储方式
 - 指针 `BYTE* pimg;`
 - 指向指针的指针 `BYTE** ppimg;`

数字图像坐标系

- 图像左上角为图像坐标系原点，即(0,0)点



数字图像格式

- 图像文件格式
 - BMP格式
 - GIF格式
 - TIFF格式
 - JPEG格式
 - ...

BMP图像格式

- BMP文件组成
 - 位图文件头、位图信息头、调色板和数据
- 位图文件头
 - 包含BMP文件的类型、文件大小和位图起始位置等信息

```
typedef struct tagBITMAPFILEHEADER {  
    WORD bfType;      // 位图文件的类型，必须为BM  
    DWORD bfSize;     // 位图文件的大小，以字节为单位  
    WORD bfReserved1; // 位图文件保留字，必须为0  
    WORD bfReserved2; // 位图文件保留字，必须为0  
    DWORD bfOffBits;  // 位图数据起始地址相对文件头的偏移  
} BITMAPFILEHEADER;
```

BMP图 像格式 (2)

- 位图信息头:用于说明位图的尺寸等信息

```
typedef struct tagBITMAPINFOHEADER {  
    DWORD   biSize;      // 本结构所占用字节数  
    LONG    biWidth;     // 位图的宽度，以像素为单位  
    LONG    biHeight;    // 位图的高度，以像素为单位  
    WORD     biPlanes;   // 目标设备的级别，必须为1  
    WORD     biBitCount; // 每个像素所需的位数，必须是1(双色),  
                        // 4(16色)，8(256色)或24(真彩色)之一  
    DWORD   biCompression; // 位图压缩类型，必须是 0(不压缩),  
                        // 1(BI_RLE8压缩类型)或2(BI_RLE4压缩类型)之一  
    DWORD   biSizeImage;  // 位图的大小，以字节为单位  
    LONG    biXPelsPerMeter; // 位图水平分辨率，每米像素数  
    LONG    biYPelsPerMeter; // 位图垂直分辨率，每米像素数  
    DWORD   biClrUsed;    // 位图实际使用的颜色表中的颜色数  
    DWORD   biClrImportant; // 位图显示过程中重要的颜色数  
} BITMAPINFOHEADER;
```

BMP图像格式(3)

- 调色板:用于说明位图中的颜色

```
typedef struct tagRGBQUAD {  
    BYTE rgbBlue;      // 蓝色的亮度(值范围为0-255)  
    BYTE rgbGreen;     // 绿色的亮度(值范围为0-255)  
    BYTE rgbRed;       // 红色的亮度(值范围为0-255)  
    BYTE rgbReserved;  // 保留, 必须为0  
} RGBQUAD;
```

- 位图数据
 - 记录了位图的每一个像素值,
 - 记录顺序是在扫描行内是从左到右,扫描行之间是从下到上
 - 彩色分量顺序是GBR

读取BMP图像(1)

```
// 打开BMP文件
```

```
FILE *fp_BMP; fopen_s(&fp_BMP, filename, "rb");
```

```
// 读位图文件头
```

```
fread(&bmfh.bfType,sizeof(unsigned short),1,fp_BMP);
```

```
if(bmfh.bfType != (unsigned short)0x4D42) // 若文件头两个不是"BM"
```

```
    return FALSE;
```

```
fread(&bmfh.bfSize,sizeof(unsigned long),1,fp_BMP);
```

```
fread(&bmfh.bfReserved1,sizeof(unsigned short),1,fp_BMP);
```

```
fread(&bmfh.bfReserved2,sizeof(unsigned short),1,fp_BMP);
```

```
fread(&bmfh.bfOffBits,sizeof(unsigned long),1,fp_BMP);
```

读取BMP图像(2)

// 读位图信息头

```
fread((char*)&bmih,sizeof(BITMAPINFOHEADER),1,fp_BMP);
```

```
bmpwidth = (int)bmih.biWidth;           // 读取图像宽度
```

```
bmpheight= (int)bmih.biHeight;          // 读取图像高度
```

```
if (bmih.biBitCount == 8)      nbit = 1;    // 读取图像位数
```

```
else if (bmih.biBitCount == 24) nbit = 3;
```

```
else      return FALSE;
```

// 距离文件起始点偏移bmfh.bfOffBits，开始读图像

```
fseek(fp_BMP, bmfh.bfOffBits, SEEK_SET);
```

读取BMP图像(3)

```
// 读取数据

for(i=bmpheight-1; i>=0; i--) {

    addr = i*bmih.biWidth*nbit ;

    fread(linebuf,sizeof(BYTE),linebuf_size,fp_BMP); // 读入一行数据

    if (bmih.biBitCount == 8)

        memcpy((pbuf->Img+addr), linebuf, bmpwidth*sizeof(BYTE));

    else if (bmih.biBitCount == 24) { // 存放顺序为B,G,R,B,G,R,...

        for(j=0; j<bmpwidth; j++) {

            pbuf->Img[addr+j*3+2] = linebuf[j*3];

            pbuf->Img[addr+j*3+1] = linebuf[j*3+1];

            pbuf->Img[addr+j*3] = linebuf[j*3+2];

        }

    }

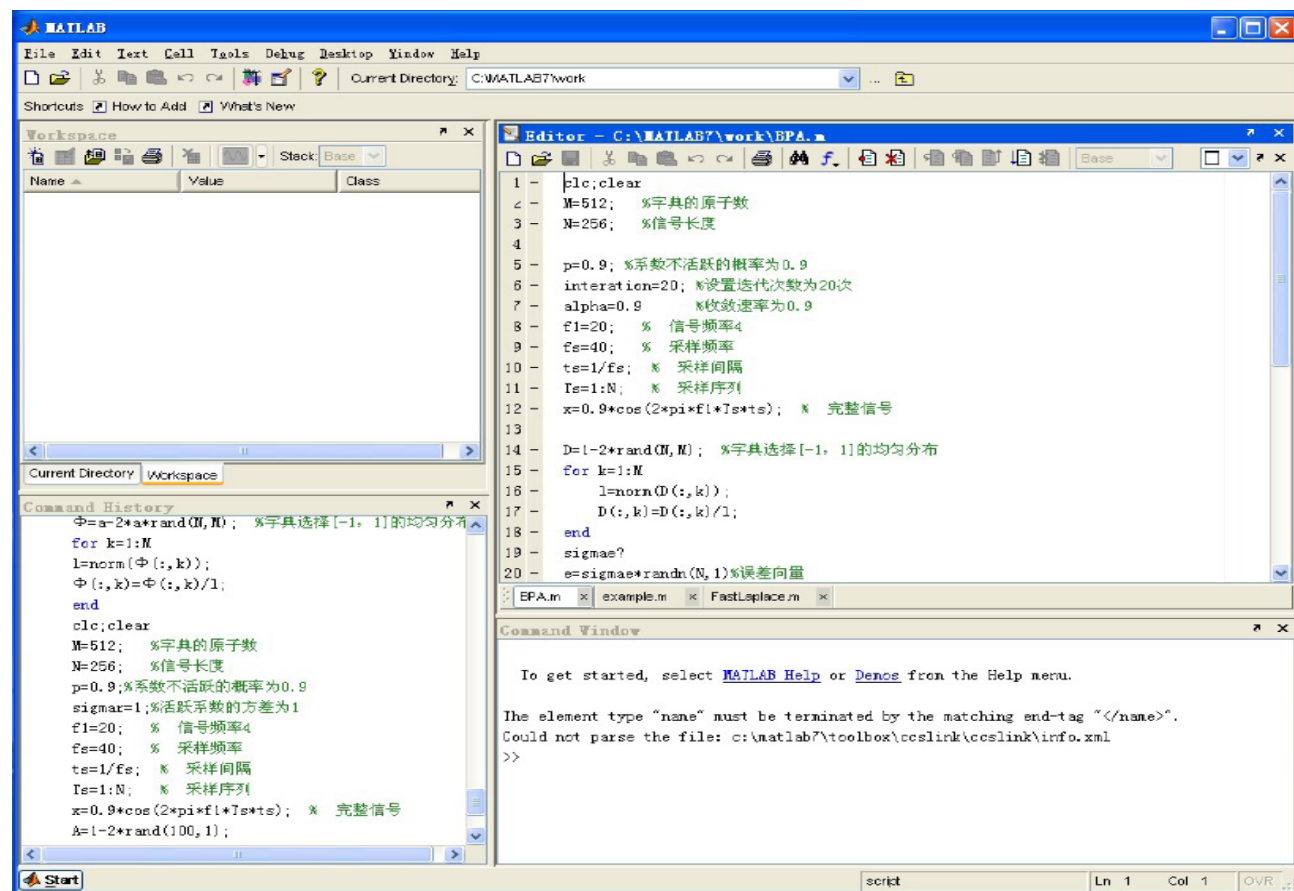
}
```

Matlab及其图像处理库

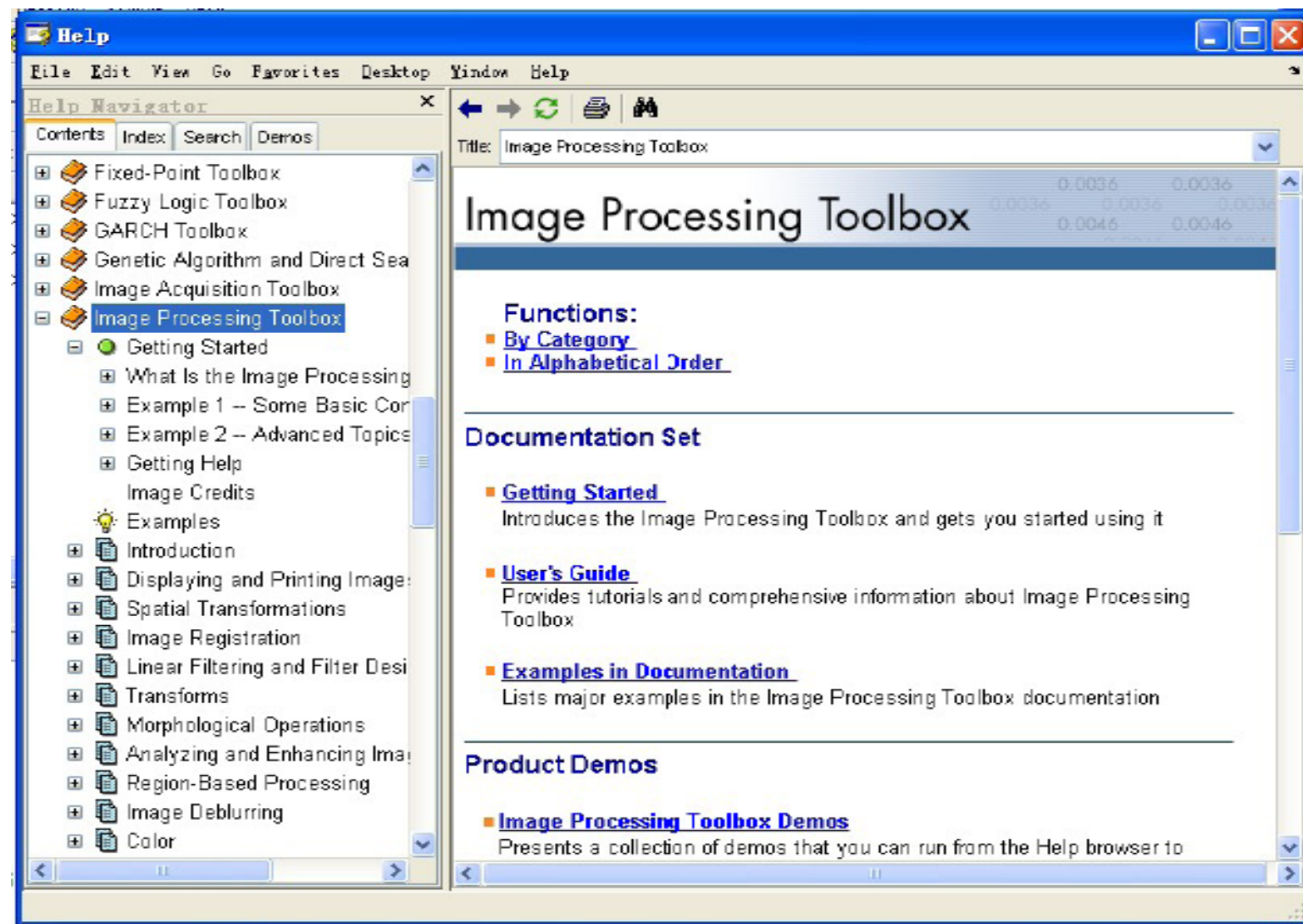
- MATLAB是Math Works公司开发的科学计算工具软件包。
- 主要特色：
 - “草稿纸式”的编程语言
 - 提供了大量的进行数组和矩阵运算的函数
 - 提供了各种科学计算工具箱
 - 具有良好的用户界面
 - 具有良好的软件复用性

Matlab桌面

- Matlab桌面是Matlab应用的主窗口
- 包括5个子窗口：
 - 命令窗口
 - 工作空间窗口
 - 当前目录窗口
 - 历史命令窗口
 - 图形窗口



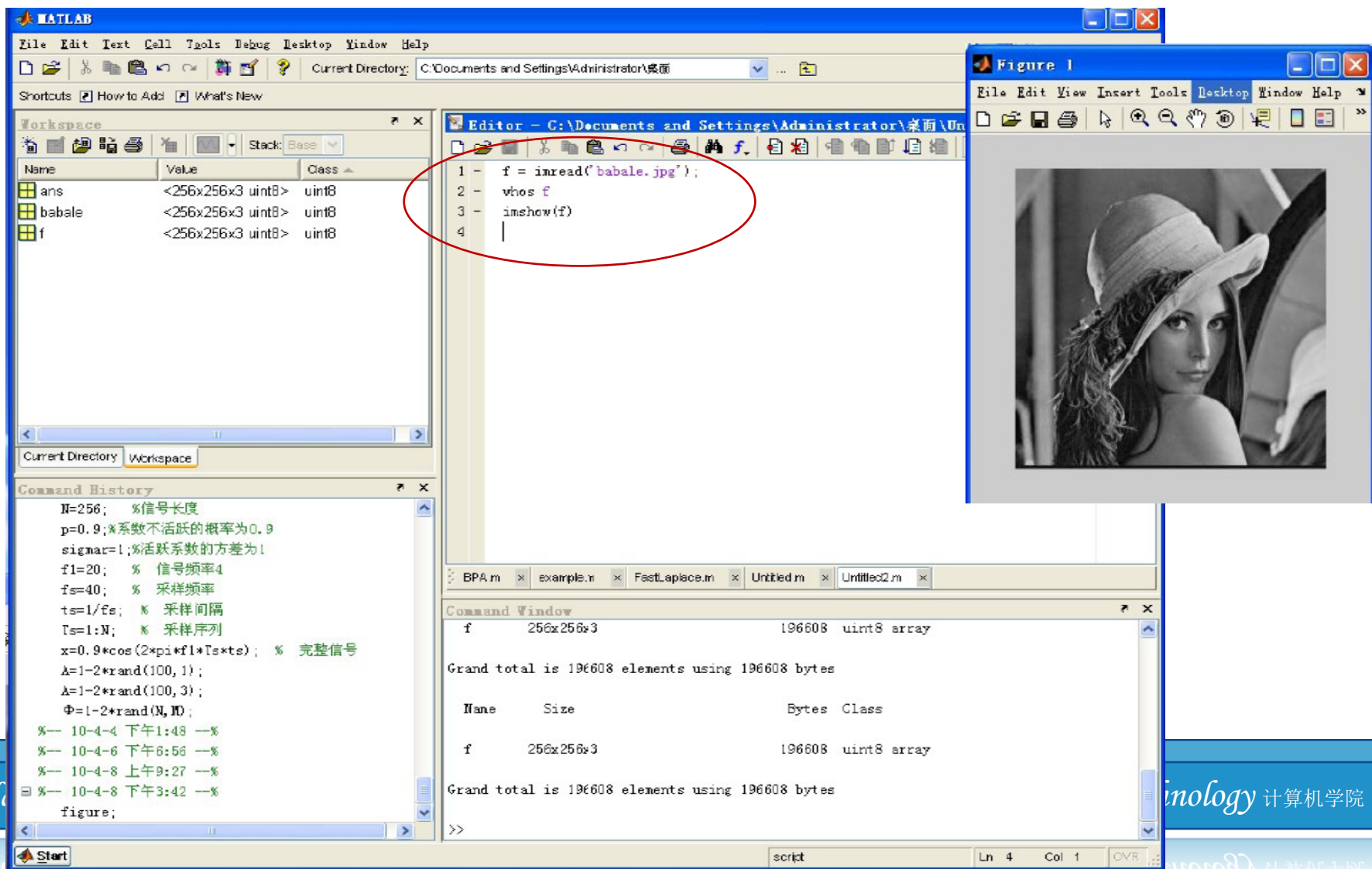
Matlab的Image Processing TOOLBOX



图像处理工具箱中常用函数

imread	读图像文件
imwrite	写图像文件
imshow	显示图像
image	创建和显示图像对象
imnoise	给一幅图像添加噪声
dct2	二维离散余弦变换
fft2	二维快速傅里叶变换
wave2gray	显示小波分解系数
imfilter	滤波二维和N维图像
wiener2	执行二维去噪滤波
.....	

例：读入和显示图像



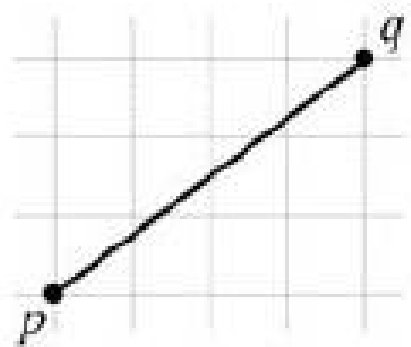
第一章 绪论

1.3 像素间关系

像素距离：欧氏距离

• 欧氏距离

$$D_E(p,q) = [(x-s)^2 + (y-t)^2]^{1/2}$$

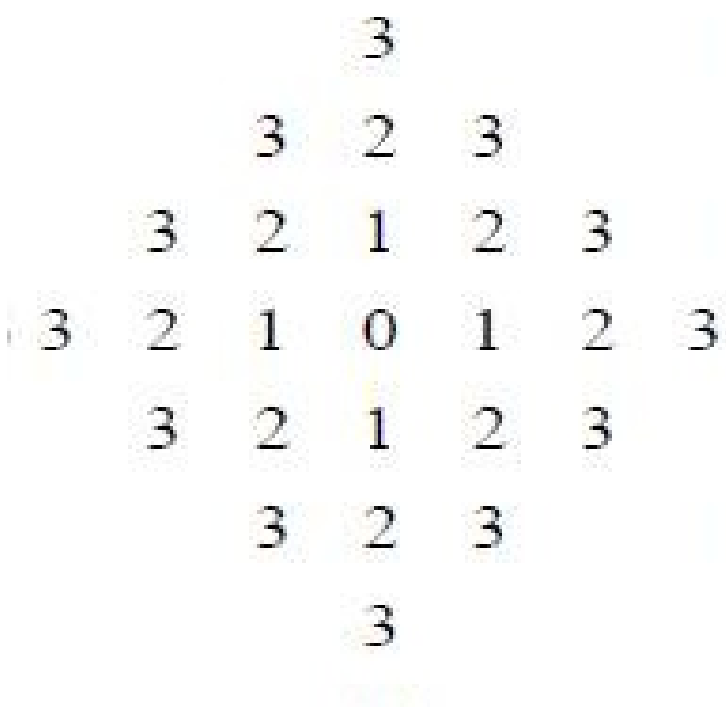
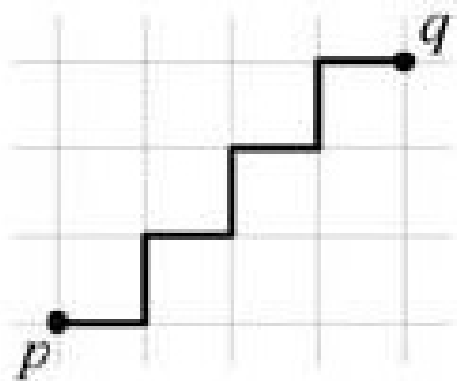


			3			
	2.8	2.2	2	2.2	2.8	
	2.2	1.4	1	1.4	2.2	
3	2	1	0	1	2	3
	2.2	1.4	1	1.4	2.2	
	2.8	2.2	2	2.2	2.8	
			3			

像素距离：城区距离

- 城区距离， 又称街道距离或曼哈顿距离

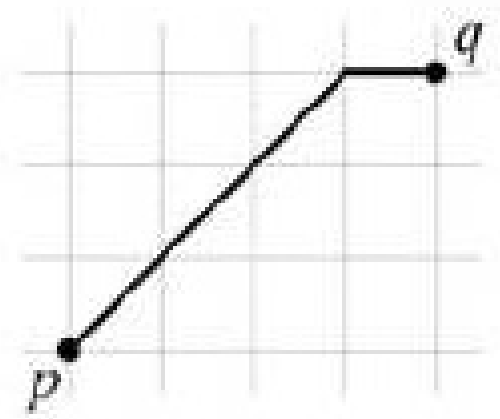
$$D_4(p,q) = |x - s| + |y - t|$$



像素距离：棋盘距离

• 棋盘距离

$$D_8(p, q) = \max (|x - s|, |y - t|)$$



3	3	3	3	3	3	3
3	2	2	2	2	2	3
3	2	1	1	1	2	3
3	2	1	0	1	2	3
3	2	1	1	1	2	3
3	2	2	2	2	2	3
3	3	3	3	3	3	3

像素邻域

4邻域 $N_4(p)$

	r	
r	p	r
	r	

对角邻域 $N_D(p)$

s		s
	p	
s		s

8邻域 $N_8(p)$

s	r	s
r	p	r
s	r	s

$$N_4(p) = \{ r \mid D_4(p,r)=1 \} = \{ r \mid D_E(p,r)=1 \}$$

$$N_8(p) = \{ r \mid D_8(p,r)=1 \}$$

习题

第一章 绪论

习题1.2 P19

- 计算机视觉要达到的目的有哪些？
- 计算机通过图像和视频对客观世界的感知、识别和理解。
 - 对场景进行解释和描述
 - 根据对场景的解释和描述制定行为规划

习题1.9 P19

- 因为8-邻域中近邻像素个数是4-邻域中近邻像素个数的两倍，那么某个像素的 D_8 距离小于或等于1的像素个数**是否等于**该像素 D_4 距离小于或等于2的像素个数？

3	3	3	3	3	3	3
3	2	2	2	2	2	3
3	2	1	1	1	2	3
3	2	1	0	1	2	3
3	2	1	1	1	2	3
3	2	2	2	2	2	3
3	3	3	3	3	3	3

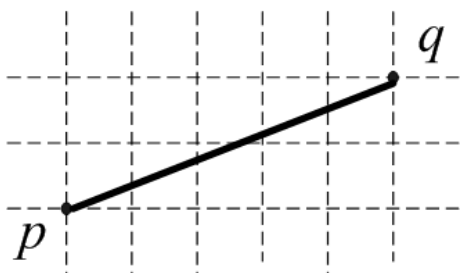
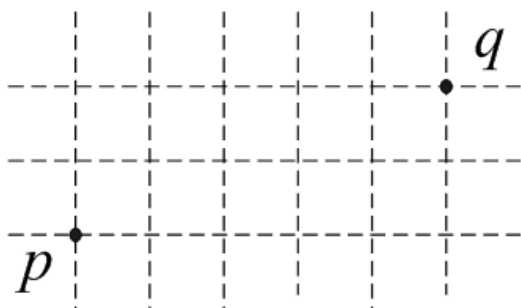
D_8 距离

			3			
		3	2	3		
	3	2	1	2	3	
3	2	1	0	1	2	3
	3	2	1	2	3	
		3	2	3		
			3			

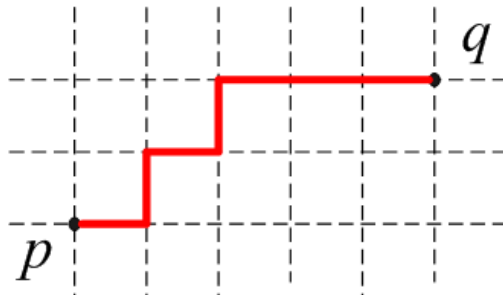
D_4 距离

习题1.10 P19

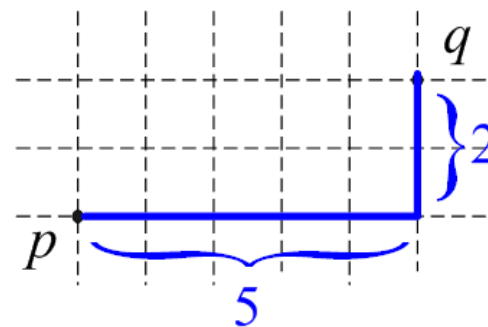
- 计算下图中两个像素 p 和 q 之间的 D_E 距离、 D_4 距离和 D_8 距离。



$$D_E = \sqrt{5 \times 5 + 2 \times 2} \approx 5.39$$



$$D_4 = 7$$



$$D_8 = \max\{5, 2\} = 5$$

The End