

法律声明

- 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。



关注 小象学院

第六讲



图像数据挖掘

--Robin

目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- K-Means聚类
- 实战案例5：Fashion-MNIST图片分类

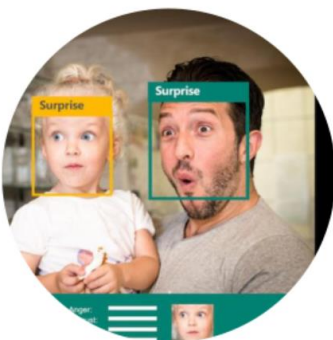
目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- K-Means聚类
- 实战案例5：Fashion-MNIST图片分类

计算机视觉

什么是计算机视觉

- 从**图像**和视频中提取数值或符号信息的计算系统
- 让计算机能够和人类一样“看到并理解”图像
- 主要应用



图像理解—高级视觉

识别
鉴别
监测



图像分析—中级视觉

运动
分割
跟踪
多视图几何



图像处理—低级视觉

线性滤波
边缘检测
纹理

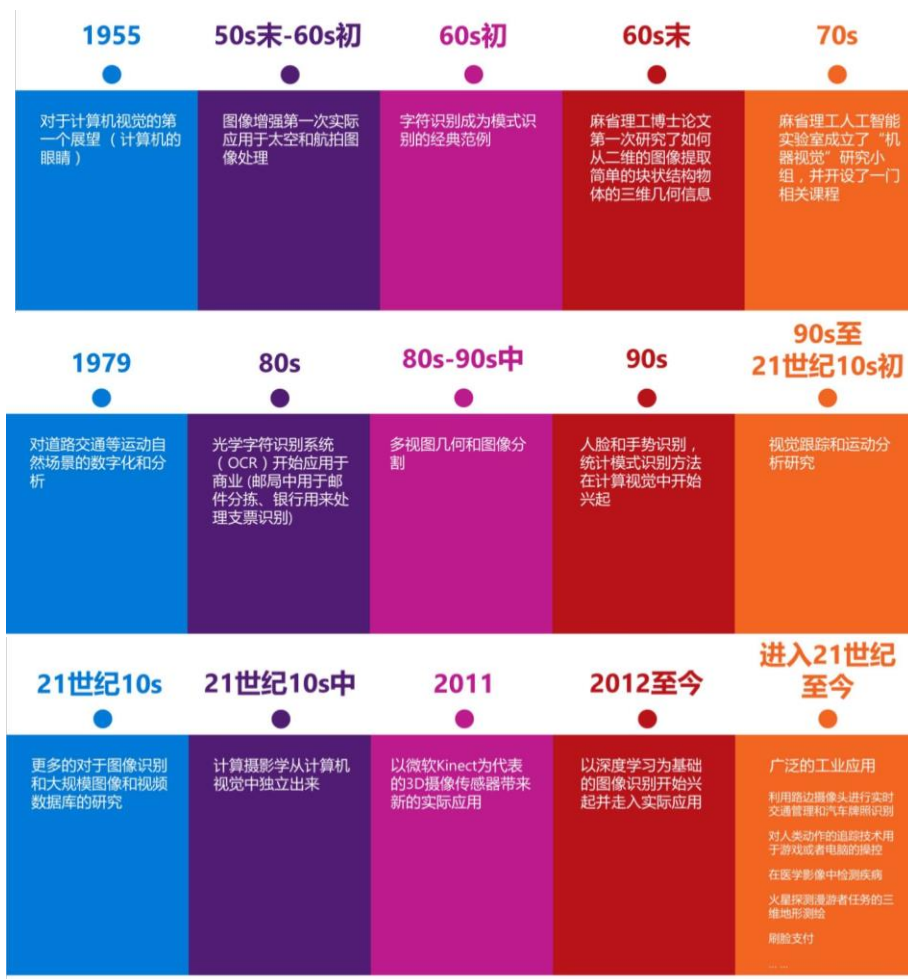


图像获取—成像

相机模型
相机标定
辐射测定
颜色

计算机视觉

CV发展历史



计算机视觉库OpenCV

OpenCV

- 开源的跨平台计算机视觉库
- 可运行在Linux、Windows、Android和MacOS操作系统上
- 轻量且高效，提供了Python、Ruby、Matlab等语言的接口
- 实现了图像处理和计算机视觉方面的很多通用算法



OpenCV-Python

- 面向Python的OpenCV接口/库
- 支持NumPy数据结构

安装

- `pip install opencv-contrib-python`

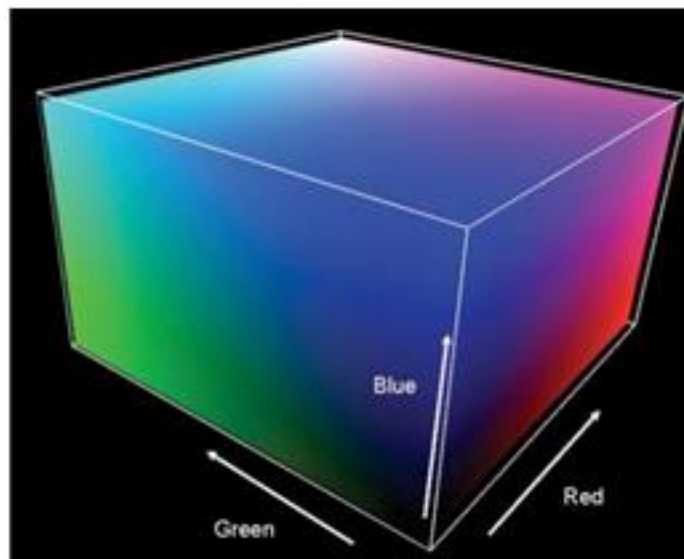
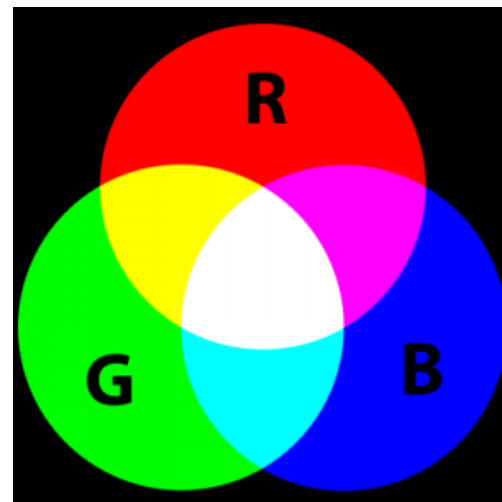
目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- K-Means聚类
- 实战案例5：Fashion-MNIST图片分类

图像数据基本概念及操作

RGB颜色空间

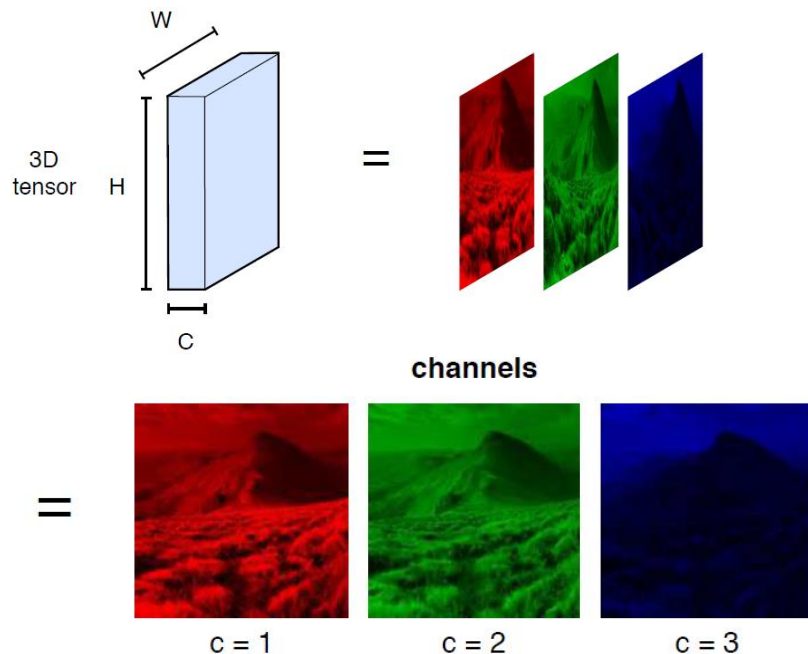
- 加法混色，彩色显示器
- 3通道
 - Red通道
 - Green通道
 - Blue通道
- 一个像素颜色值
 - (b, g, r)
- 取值范围
 - $[0, 255]$
 - $[0.0, 1.0]$



图像数据基本概念及操作

主流颜色空间

- RGB三通道彩色图
 - 图片→3维矩阵 $([0,255])$



- 单通道灰度图
 - 亮度信息 $([0,255])$
 - $\text{Gray} = R \cdot 0.3 + G \cdot 0.59 + B \cdot 0.11$



灰度化



图像数据基本概念及操作

opencv的图像数据

- 图像数据是由NumPy的**多维数组 (ndarray)** 表示的
- 由opencv加载的图像数据可以调用其他常用的包进行处理和计算，如 matplotlib, scipy等

数据类型和像素值

- CV中图像的像素值通常有以下两种处理范围
 1. 0 - 255, 0: 黑色, 255: 白色
 2. 0 - 1, 0: 黑色, 1 : 白色
- OpenCV读取的图像像素值范围是0-255

lect06_eg01.ipynb

图像数据基本概念及操作

OpenCV图像IO

- 读取图像 `cv2.imread()`
`cv2.IMREAD_COLOR` 读取彩色图像数据
`cv2.IMREAD_GRAYSCALE` 读取灰度图像数据
- 显示图像 `cv2.imshow()`, `cv2.waitKey(0)`, `cv2.destroyAllWindows()`
- 保存图像 `cv2.imwrite()`

`lect06_eg01.ipynb`

图像数据基本概念及操作

图像数据

- 图像数据是多维数组，前两维表示了图像的高、宽第三维表示图像的通道个数，比如RGB，第三个维度为3，因为有三个通道；灰度图像没有第三个维度
- 分割和索引
 - 像操作ndarray一样操作即可

色彩空间, RGB, HSV, Gray...

- RGB转Gray, `cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)`

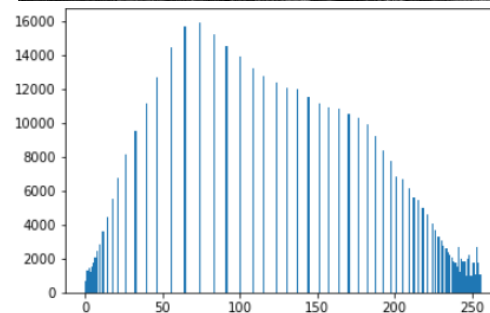
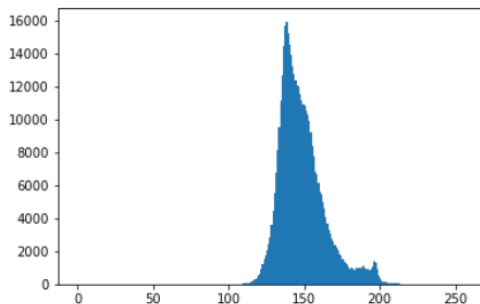
颜色直方图

- 直方图是一种能快速描述图像整体像素值分布的统计信息，
`skimage.exposure.histogram`
 - 如：可以根据直方图选定阈值用于调节图像对比度

图像数据基本概念及操作

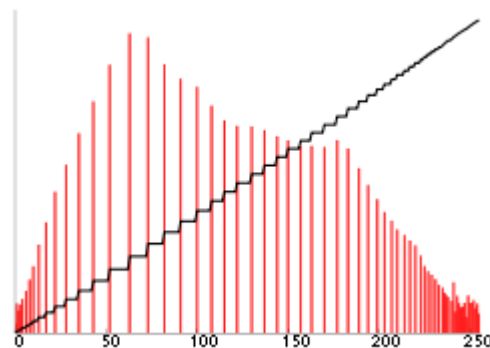
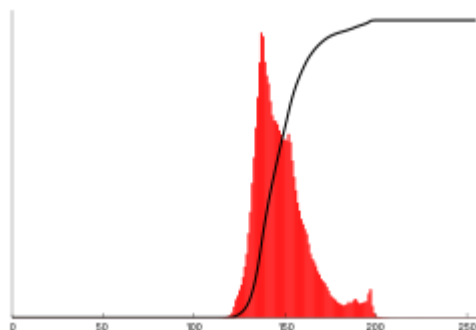
直方图均衡化

- 增强图像数据的对比度有利于特征的提取，不论是从肉眼还是算法来看都有帮助
- 自动调整图像的对比度 `cv2.equalizeHist(image)`



图像数据基本概念及操作

直方图均衡化



图像数据基本概念及操作

滤波/卷积

- 在每个图片位置 (x, y) 上进行基于邻域的函数计算

- 滤波函数 → 权重相加

- 卷积核、卷积模板
 - 滤波器、滤波模板
 - 扫描窗

$$\boxed{\text{滤波结果}} \quad h[x, y] = \sum_{k, l} \boxed{\text{滤波函数}} f[k, l] \boxed{\text{像素邻域值}} I[x + k, y + l]$$

- 不同功能需要定义不同的函数

- 图像增强

- 平滑/去燥
 - 梯度/锐化

- 信息提取、检测

- 边缘、显著点、纹理
 - 模式

图像数据基本概念及操作

滤波/卷积

参数解释

- x, y 是像素在图片中的位置/坐标
- k, l 是卷积核中的位置/坐标
 - 中心点的坐标是 $(0,0)$
- $f[k, l]$ 是卷积核中在 (k, l) 上的权重参数
- $I[x+k, y+l]$ 是与 $f[k, l]$ 相对应的图片像素值
- $h[x, y]$ 是图片中 (x, y) 像素的滤波/卷积结果

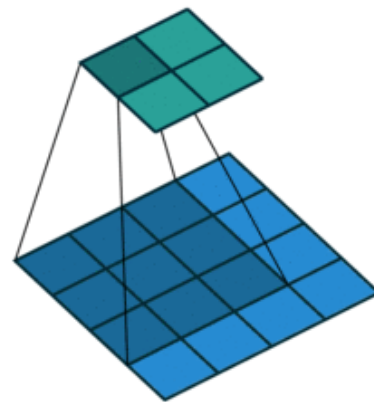
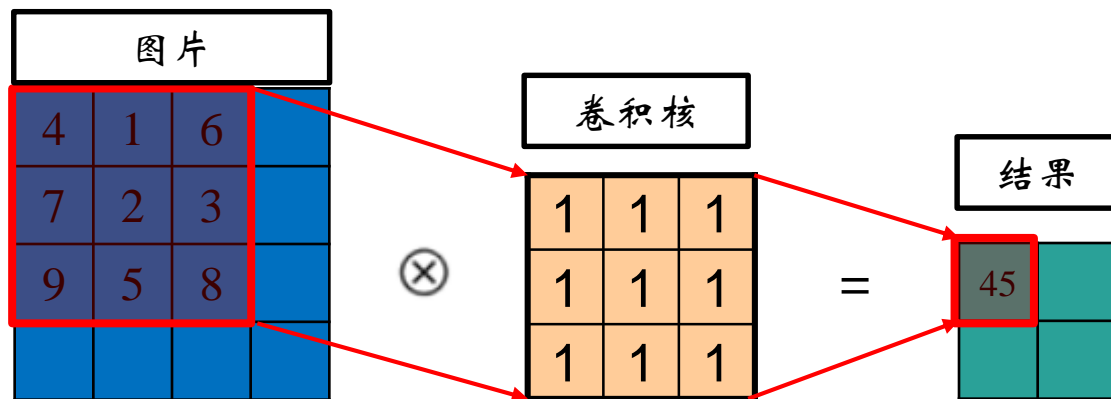
滤波结果

滤波函数

像素邻域值

$$h[x, y] = \sum_{k, l} f[k, l] I[x + k, y + l]$$

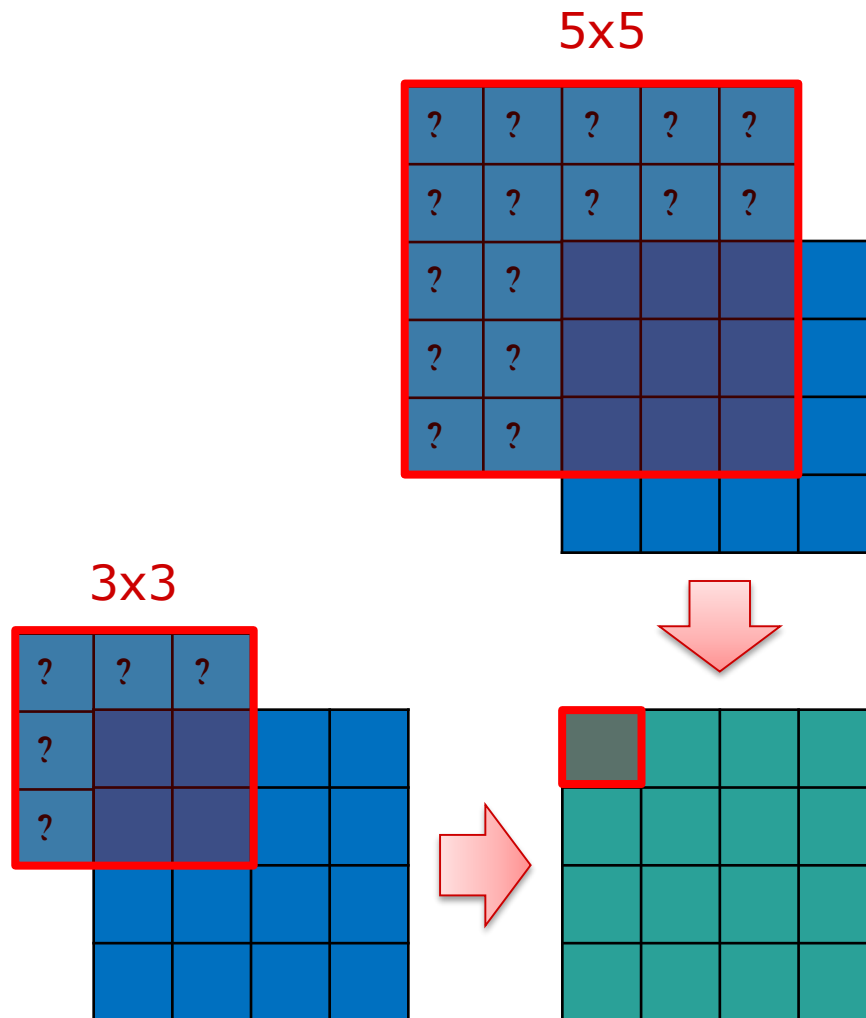
$(-1,-1)$	$(-1,0)$	$(-1,1)$
$(0,-1)$	$(0,0)$	$(0,1)$
$(1,-1)$	$(1,0)$	$(1,1)$



图像数据基本概念及操作

滤波/卷积

- 边界补充 (Padding)
 - 获得同尺寸输出的情况下
 - 卷积核越大，补充越多
- 补充类型
 - 补零 (zero-padding)
 - 边界复制 (replication)
 - 镜像 (reflection)
 - 块复制 (wraparound)



图像数据基本概念及操作

滤波/卷积

- 补零 (zero-padding)
- 边界复制 (replication)
- 镜像 (reflection)
- 块复制 (wraparound)

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

图像数据基本概念及操作

滤波/卷积

- 7x7卷积: $3 \times 3 \rightarrow 9 \times 9$

补零 (zero-padding)

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

边界复制 (replication)

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

图像数据基本概念及操作

滤波/卷积

- 7x7卷积: $3 \times 3 \rightarrow 9 \times 9$

镜像 (reflection)

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

块复制 (wraparound)

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

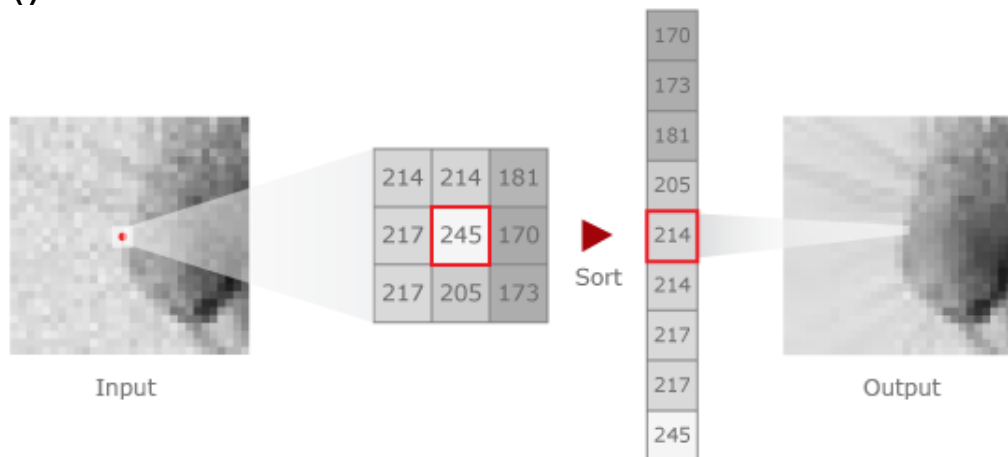
图像数据基本概念及操作

图像滤波

- 滤波是处理图像数据的常用基础操作
- 滤波操作可以去除图像中的噪声点，由此增强图像的特征

中值滤波

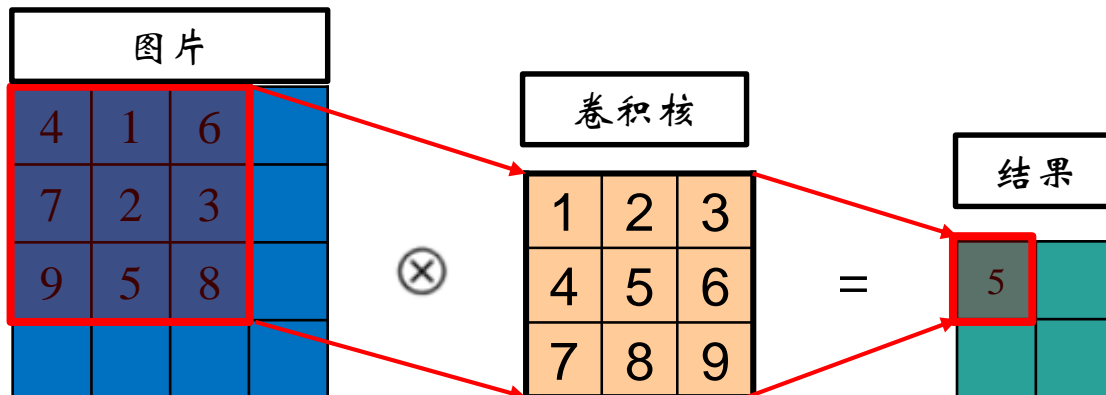
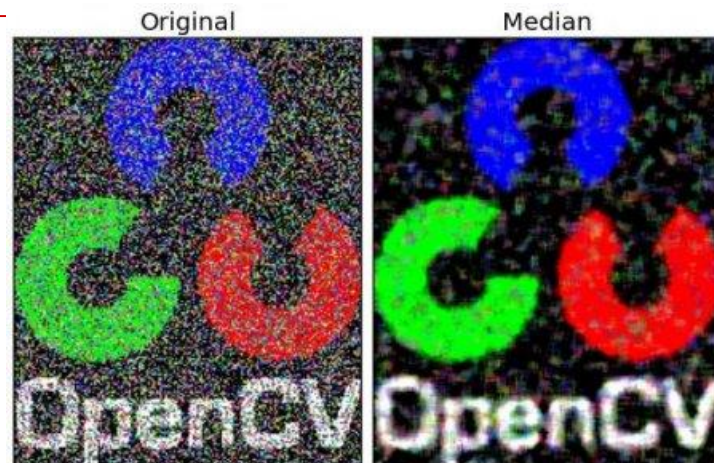
`cv2.medianBlur()`



图像数据基本概念及操作

中值滤波

- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 操作原理
 - 卷积域内的像素值从小到大排序
 - 取中间值作为卷积输出
- 有效去除椒盐噪声



图像数据基本概念及操作

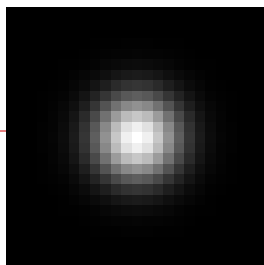
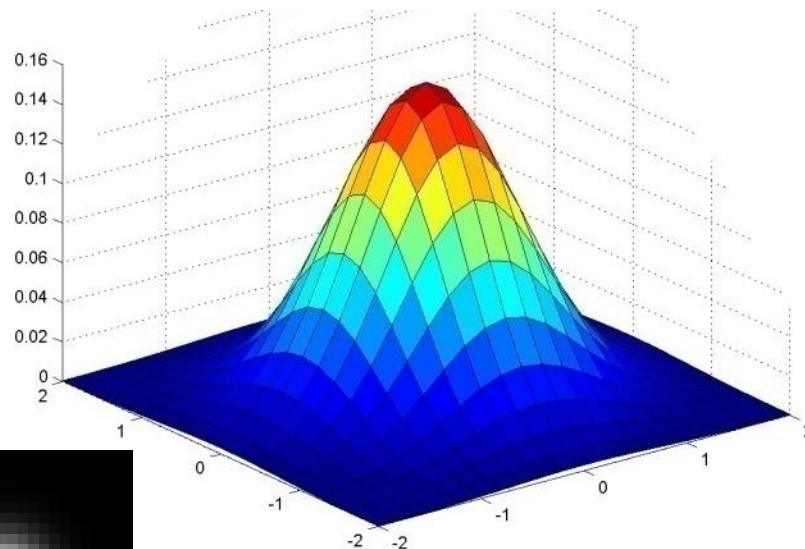
高斯滤波 `skimage.filters.gaussian()`

- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 模拟人眼, 关注中心区域
- 有效去除高斯噪声
- 参数
 - x, y 是卷积参数坐标
 - 标准差 σ

0.003	0.013	0.022	0.013	0.003
0.013	0.059	0.097	0.059	0.013
0.022	0.097	0.159	0.097	0.022
0.013	0.059	0.097	0.059	0.013
0.003	0.013	0.022	0.013	0.003

5 x 5, $\sigma = 1$

$$G_{\sigma} = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

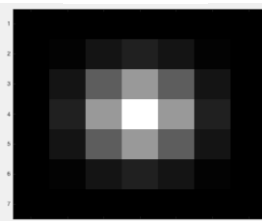


图像数据操作

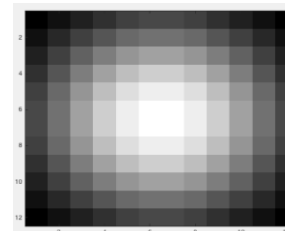
高斯滤波

- σ 越小
- 关注区域越集中

$\sigma=1$



$\sigma=4$



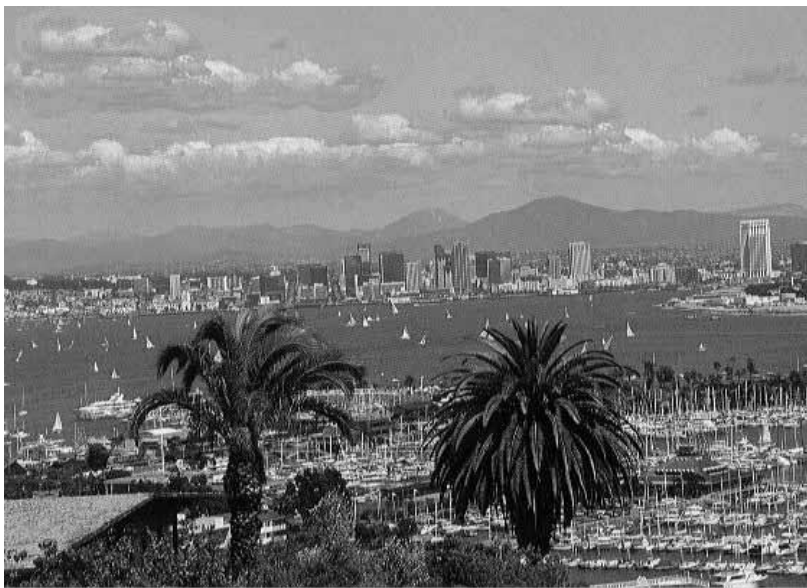
原图



图像数据操作

高斯滤波

- 人眼特性：离关注中心越远，感受精度越模糊



摄像头

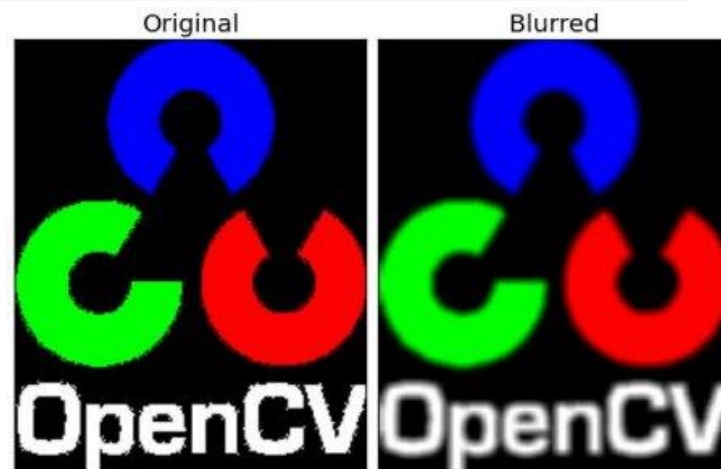


人类视觉

图像数据操作

均值滤波 `cv2.blur()`

- 奇数尺寸
 - 3x3, 5x5, 7x7, $2n-1 \times 2n-1$
- 参数和为：1



$$\otimes \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} =$$



图像数据操作

边缘检测

- 图像导数：灰度图像 I 中亮度值的变化可以描述为x方向、y方向上的导数 I_x, I_y
- 图像梯度为向量 $\nabla I = [I_x \ I_y]^T$ 。图像梯度包含两个重要的属性：
 - 大小（magnitude） $|\nabla I| = \sqrt{I_x^2 + I_y^2}$ ，描述亮度值变化的强度
 - 梯度角（gradient angle） $\alpha = \arctan2(I_y, I_x)$ ，描述每个像素点最大亮度变化的方向，可通过numpy.arctan2()计算。[arctan2定义](#)
- 图像导数的计算是通过离散估计得出的，即在x方向，y方向上进行以下的卷积操作： $I_x = I * D_x$ and $I_y = I * D_y$

图像数据操作

边缘检测

- 常用的 D_x , D_y 有:
 - Sobel算子

$$D_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

- Laplacian 算子

$$\text{kernel} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

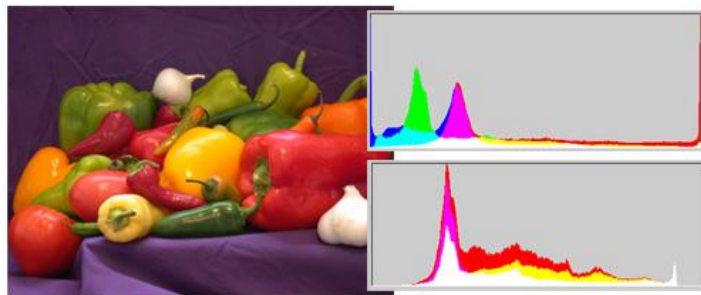
`lect06_eg01.ipynb`

目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- K-Means聚类
- 实战案例5：Fashion-MNIST图片分类

常用的图像特征描述

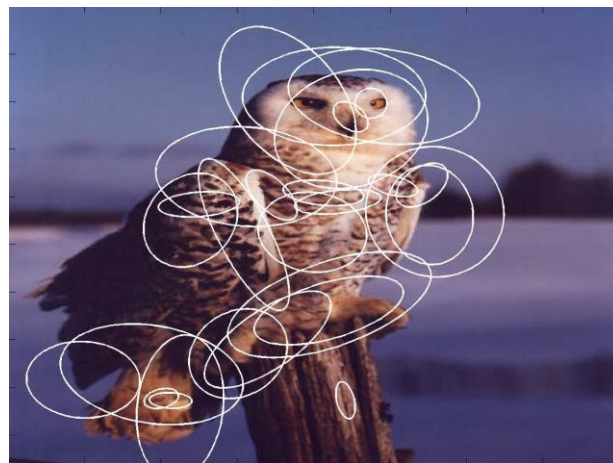
1. 颜色特征



2. 纹理特征



3. 形状特征



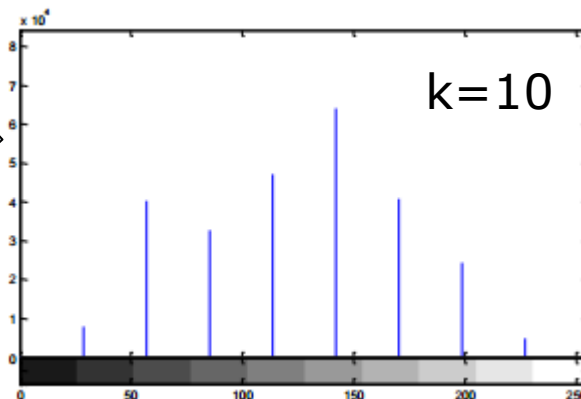
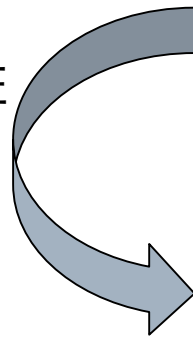
4. opencv中的特征方法

- https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_table_of_contents_feature2d/py_table_of_contents_feature2d.html

常用的图像特征描述

颜色特征

- 图像检索中应用最为广泛的视觉特征
 - 颜色直方图，如：

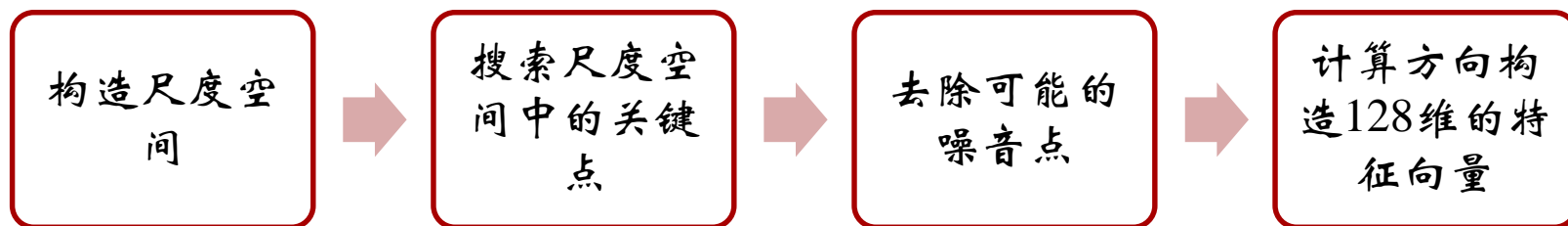


- 从512x512的灰度图像中提取维度为k的颜色直方图，就是讲256种灰度颜色分为k个区间，然后计算每个区间中像素点总数为多少。
- 如果k=20，则得到的20维直方图向量维 (0, 0, 750, 14613, 24233, 11126, 12943, 19345, 22012, 23122, 27978, 33309, 25312, 15992, 9563, 12967, 8045, 828, 6, 0)

常用的图像特征描述

图像形状特征

- 形状特征的表达必须以对图像中物体或区域的分割为基础
- **SIFT** (Scale-invariant feature transform), 在尺度空间中所提取的图像局部特征点。SIFT特征点提取较为方便, 提取速度较快, 对于图像的缩放等变换比较鲁棒, 因此得到了广泛的应用。
- http://docs.opencv.org/trunk/da/df5/tutorial_py_sift_intro.html



常用的图像特征描述

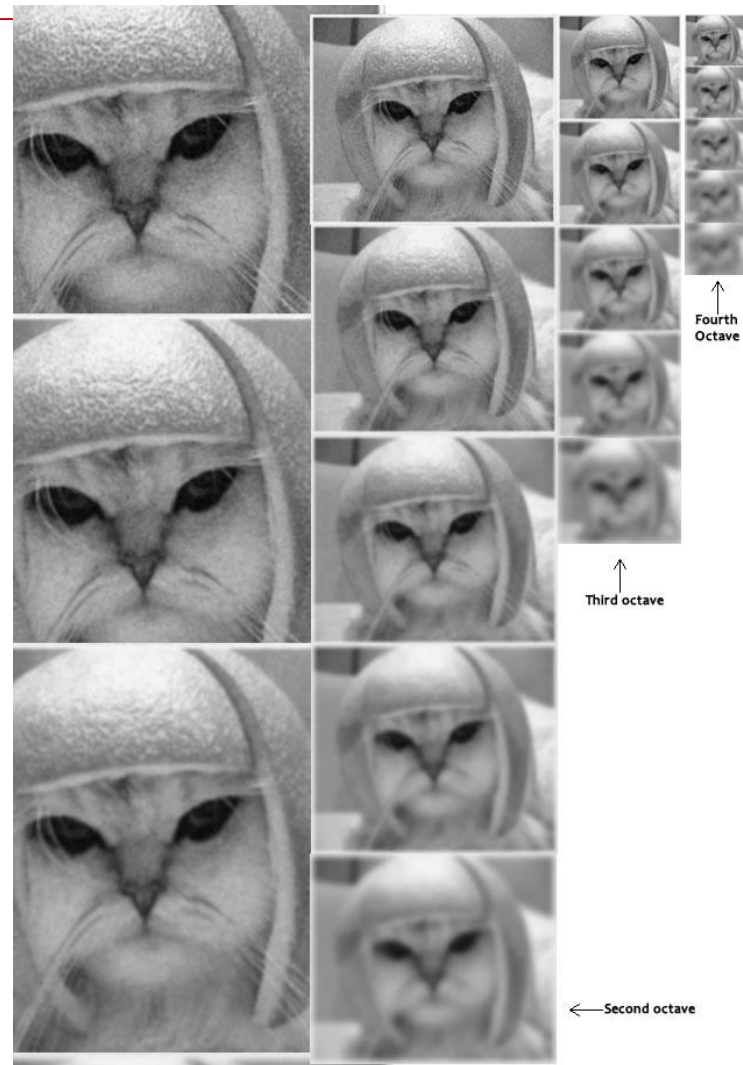
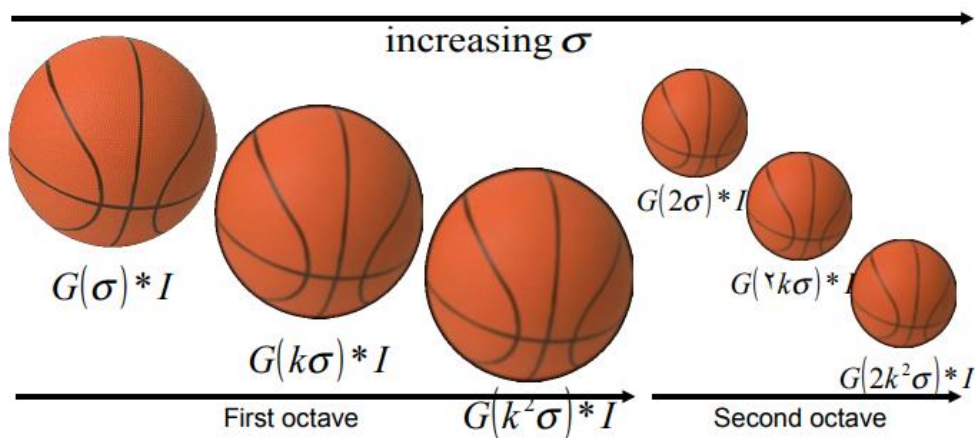
图像形状特征

- **SIFT** (Scale-invariant feature transform)
- 构造尺度空间

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$



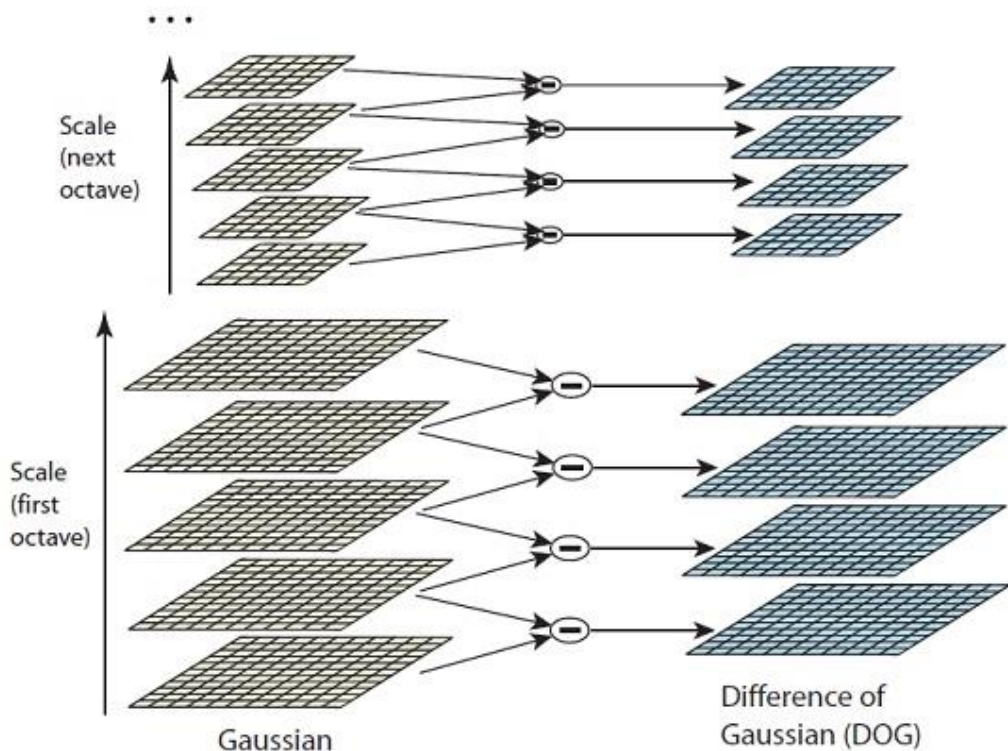
$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$



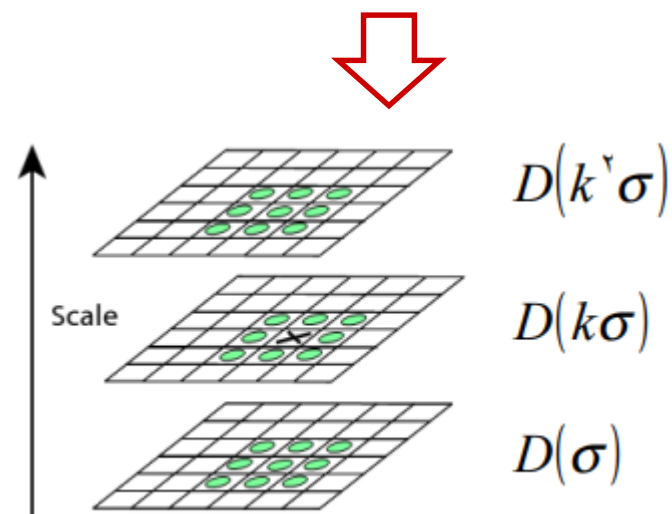
常用的图像特征描述

图像形状特征

- SIFT (Scale-invariant feature transform)
- 搜索关键点



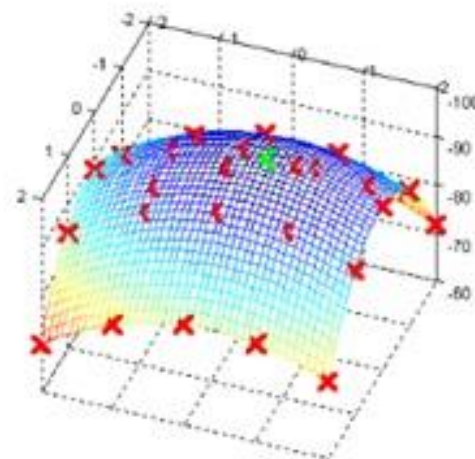
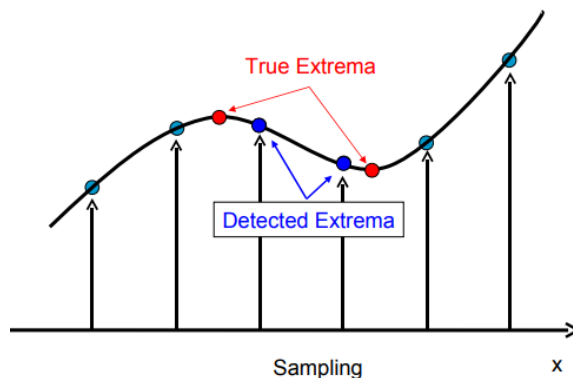
搜索3x3x3邻居的极值点



常用的图像特征描述

图像形状特征

- SIFT (Scale-invariant feature transform)
- 搜索关键点



二阶泰勒展开

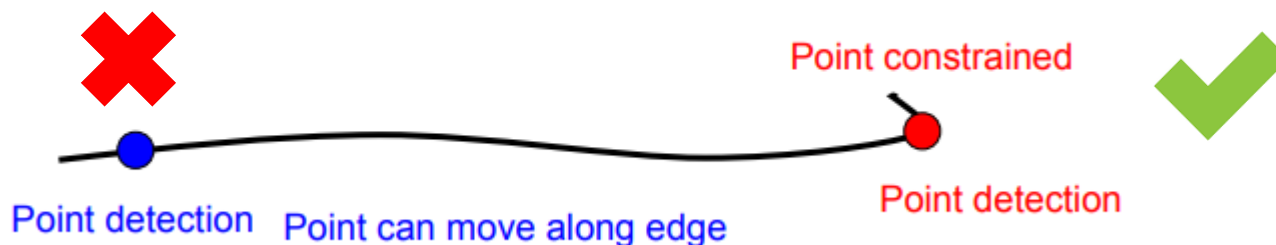


$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x$$

常用的图像特征描述

图像形状特征

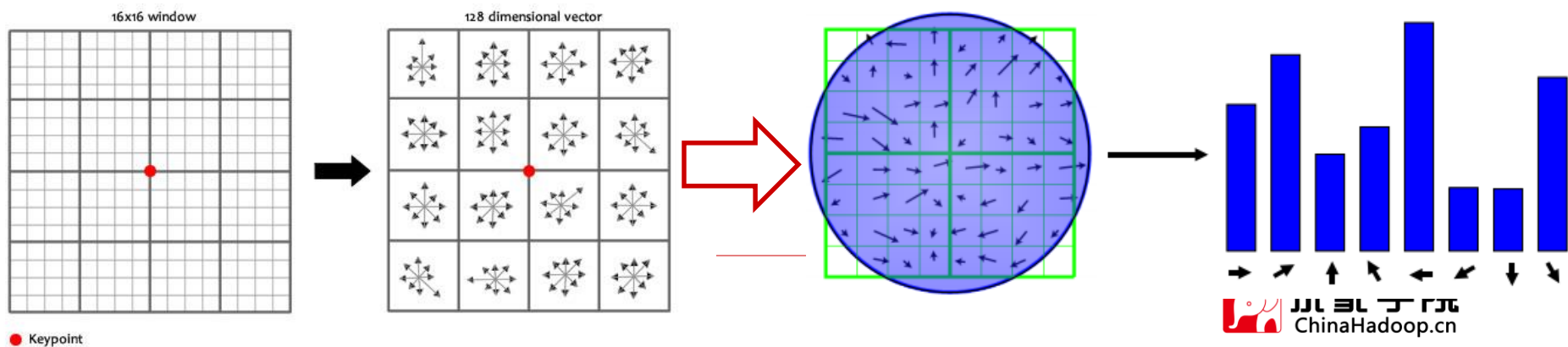
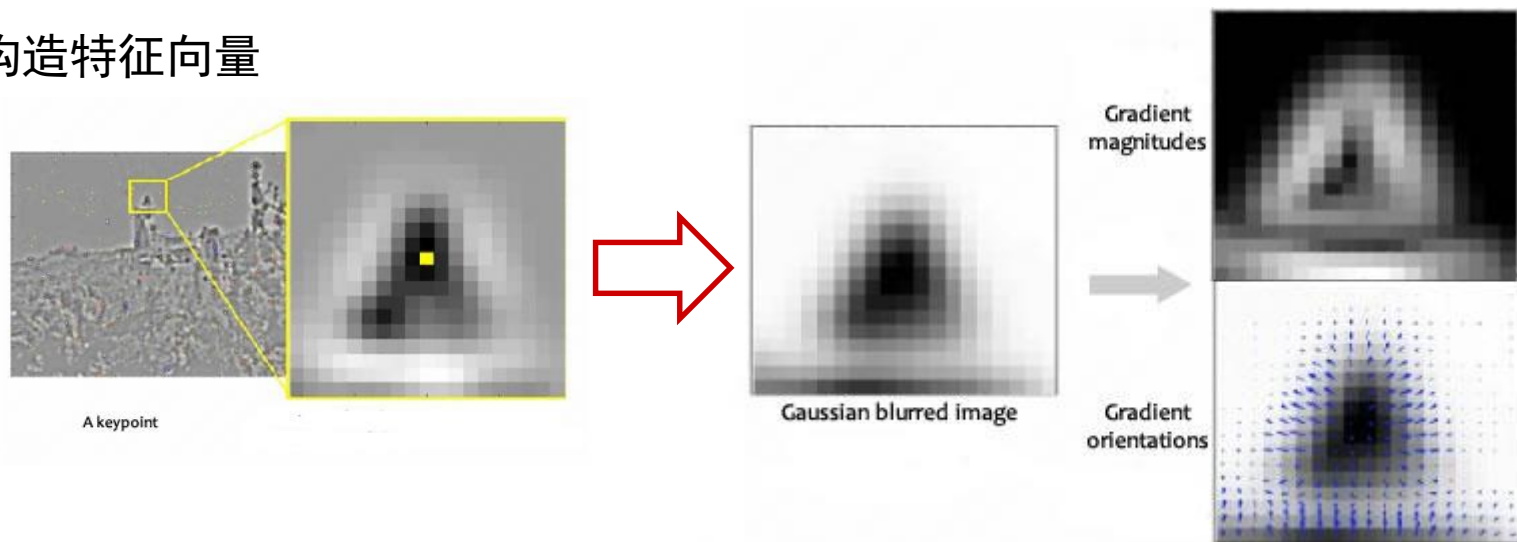
- SIFT (Scale-invariant feature transform)
- 去除噪声点
 1. 梯度值过小 (<0.03)
 2. 位于边缘的点



常用的图像特征描述

图像形状特征

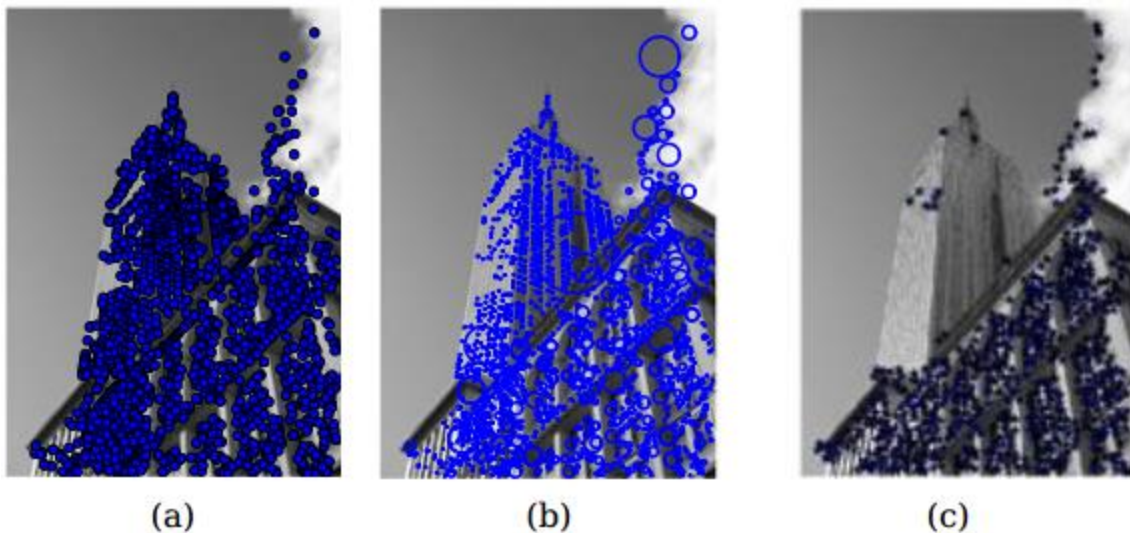
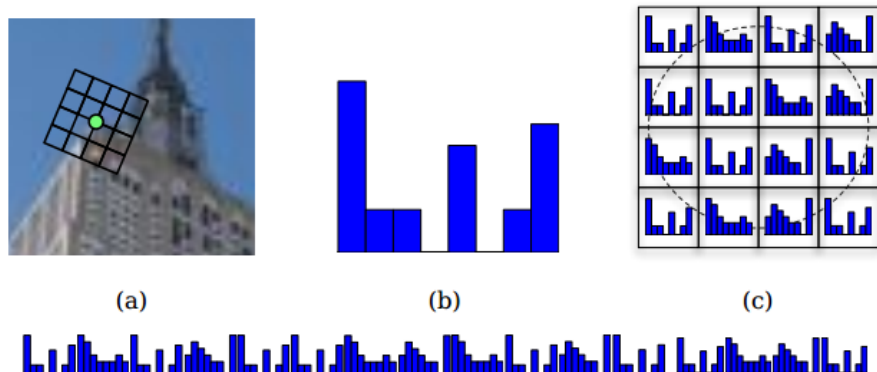
- SIFT (Scale-invariant feature transform)
- 构造特征向量



常用的图像特征描述

图像形状特征

- SIFT

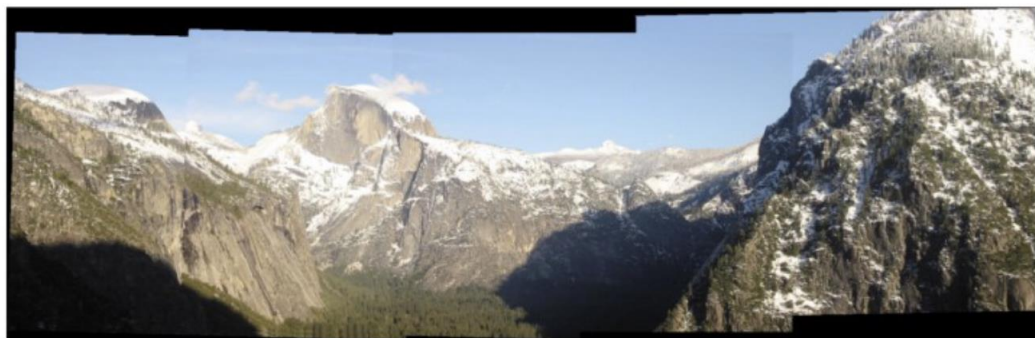
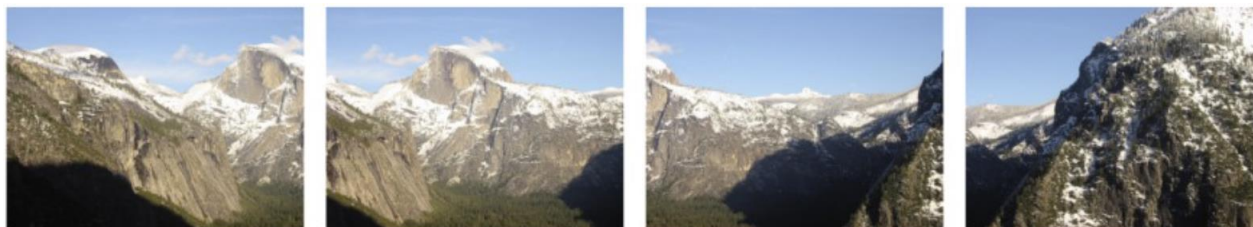
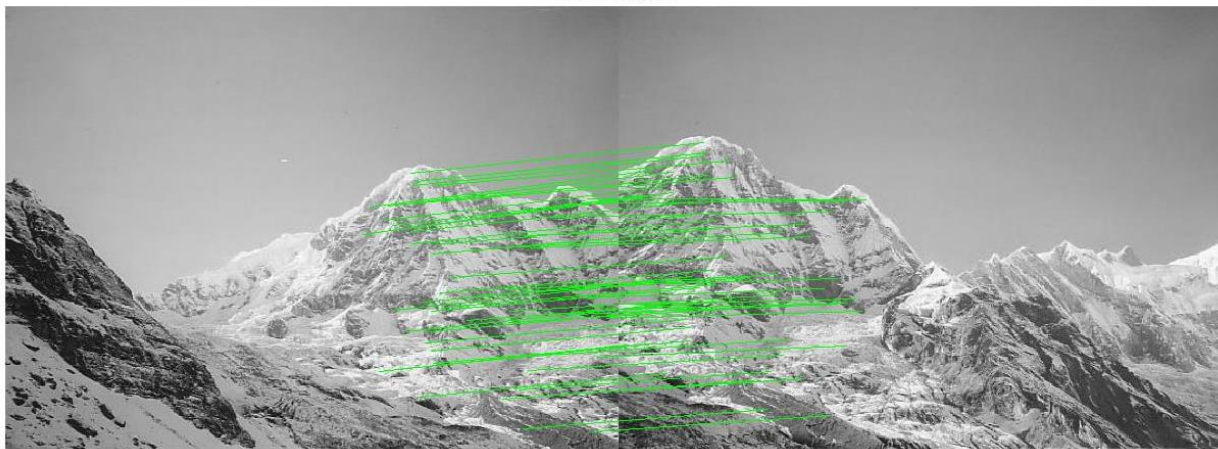


- (a) SIFT特征
- (b) SIFT特征对应的尺度
- (c) SIFT特征的关键点

常用的图像特征描述

图像形状特征

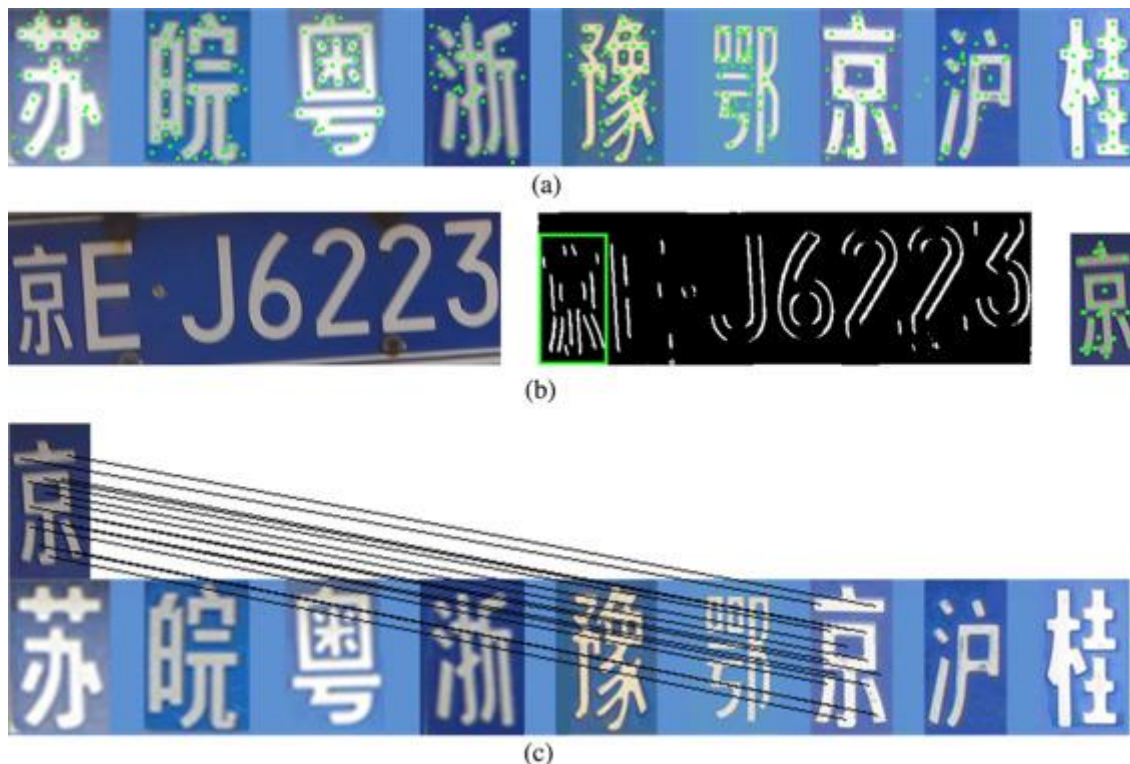
- SIFT应用
 - 图像拼接



常用的图像特征描述

图像形状特征

- SIFT应用
 - 图像识别



常用的图像特征描述

图像形状特征

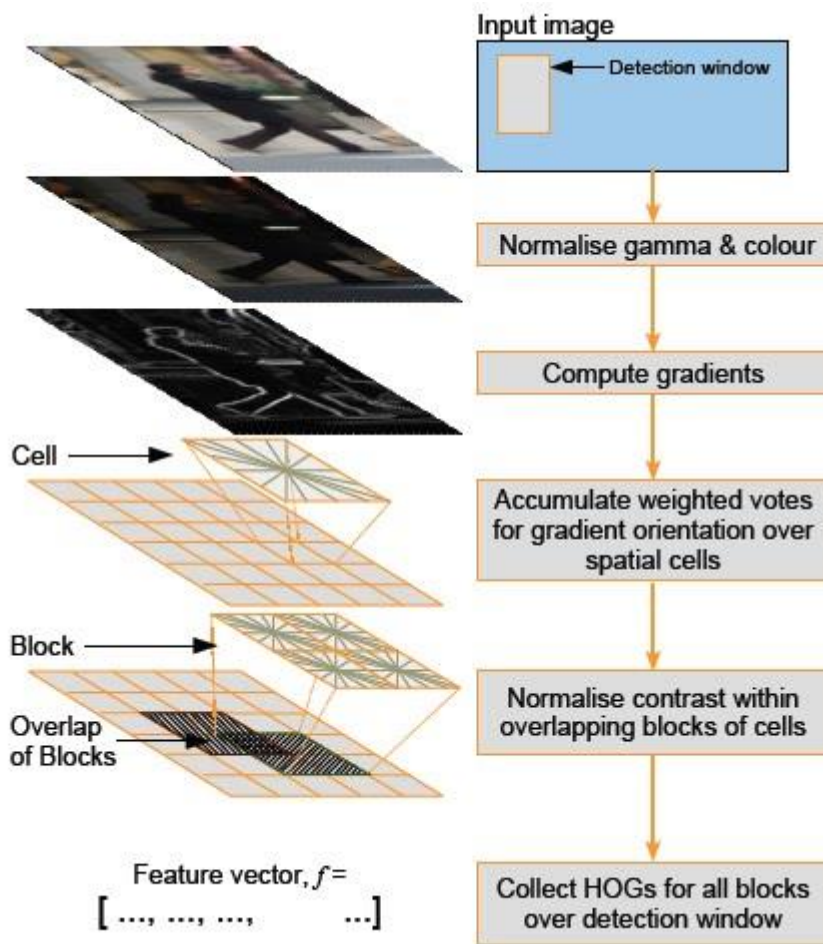
- **HOG** (Histogram of Oriented Gradient), 用于检测物体的特征描述, 通过计算和统计图像局部区域的梯度方向直方图来构建特征
- 由于HOG是在图像的局部方格单元上操作, 所以它对图像几何的和光学的形变都能保持很好的不变性
- 在粗的空域抽样、精细的方向抽样以及较强的局部光学归一化等条件下, 只要行人大体上能够保持直立的姿势, 可以容许行人有一些细微的肢体动作, 这些细微的动作可以被忽略而不影响检测效果
- HOG特征特别适合于做图像中的人体检测

<http://mccormickml.com/2013/05/09/hog-person-detector-tutorial/>

常用的图像特征描述

图像形状特征

- HOG



目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- **K-Means聚类**
- 实战案例5：Fashion-MNIST图片分类

K-Means聚类

- 聚类（clustering）属于无监督学习（unsupervised learning）
- 无类别标记
- 在线demo <http://syskall.com/kmeans.js/>
- 数据挖掘经典算法之一
- 算法接收参数k；然后将样本点划分为k个聚类；同一聚类中的样本相似度较高；不同聚类中的样本相似度较小
- 算法思想：
以空间中k个样本点为中心进行聚类，对最靠近它们的样本点归类。通过迭代的方法，逐步更新各聚类中心，直至达到最好的聚类效果

K-Means聚类

- 算法描述：
 1. 选择k个聚类的初始中心
 2. 在第n次迭代中，对任意一个样本点，求其到k个聚类中心的距离，将该样本点归类到距离最小的中心所在的聚类
 3. 利用均值等方法更新各类的中心值
 4. 对所有的k个聚类中心，如果利用2,3步的迭代更新后，达到稳定，则迭代结束。

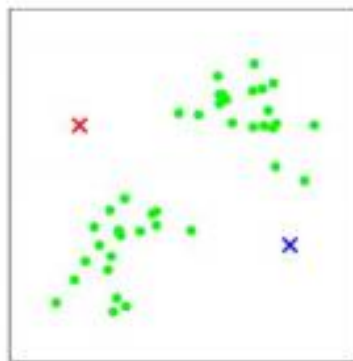
优点	缺点
<ul style="list-style-type: none">• 速度快，简单	<ul style="list-style-type: none">• 最终结果和初始点的选择相关，容易陷入局部最优• 需要给定k值

K-Means聚类

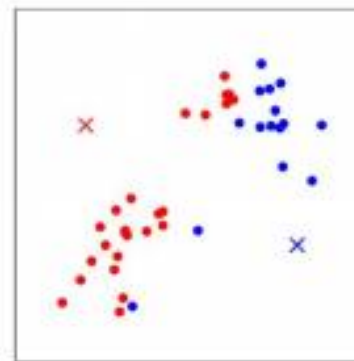
- 算法演示：



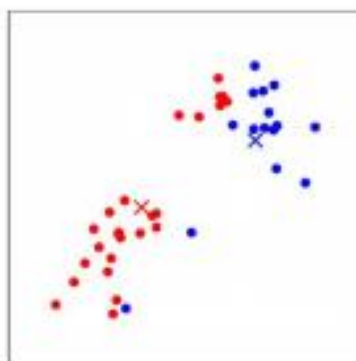
(a)



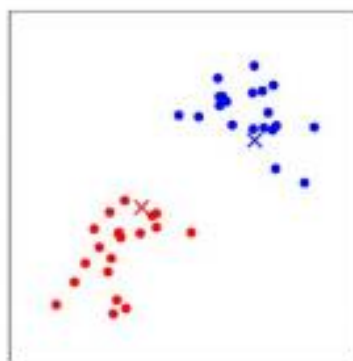
(b)



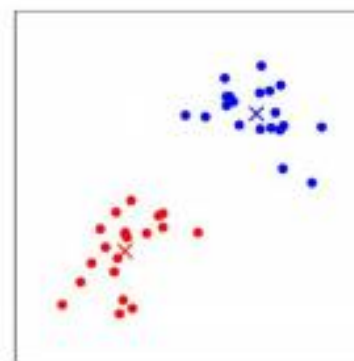
(c)



(d)



(e)



(f)

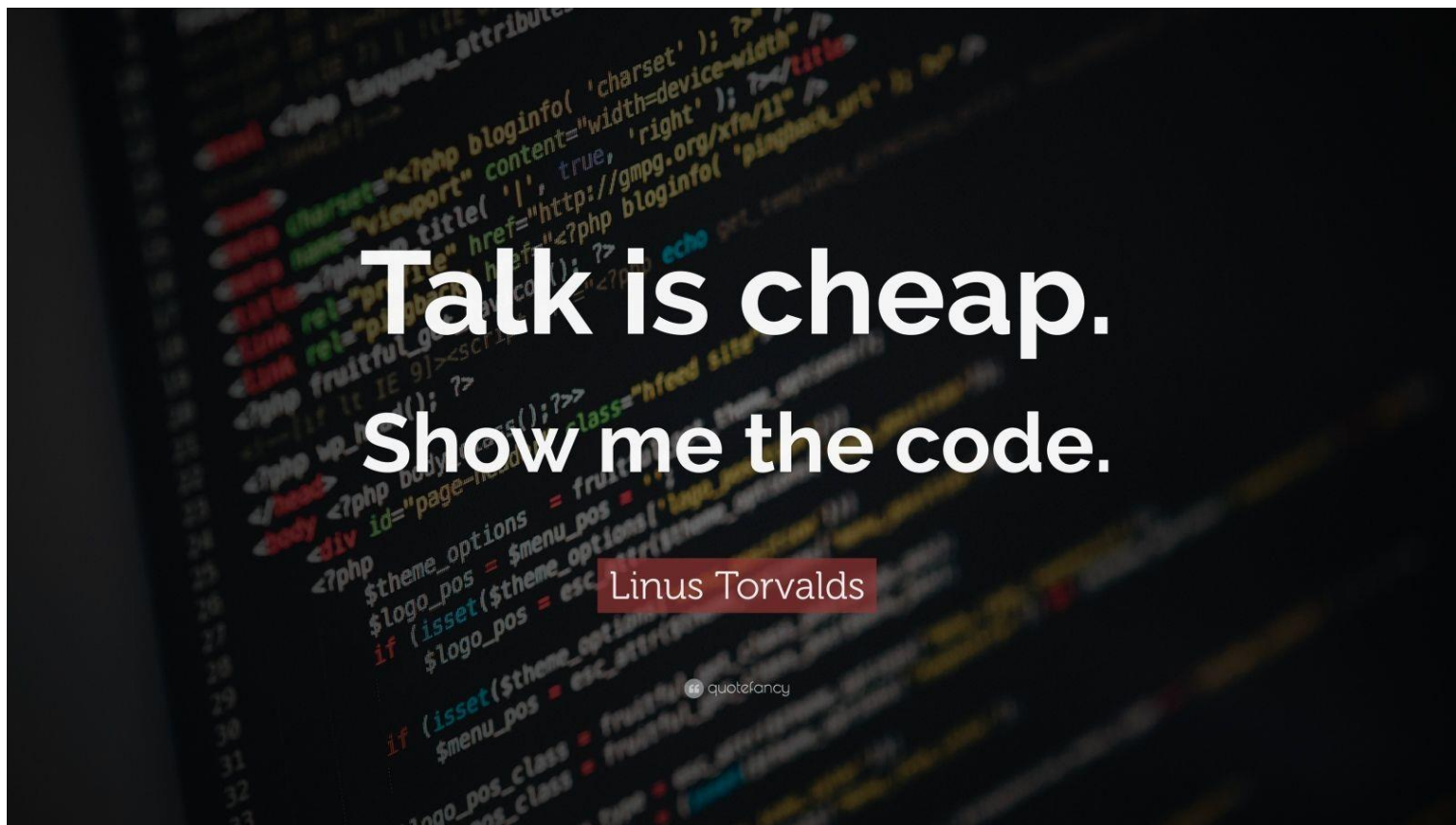
目录

- 计算机视觉库OpenCV
- 图像数据基本概念及操作
- 常用的图像特征描述
- K-Means聚类
- 实战案例5：Fashion-MNIST图片分类

实战案例 5

项目名称：Fashion-MNIST图片分类

- 请参考相应的配套代码及案例讲解文档



问答互动

在所报课的课程页面，

- 1、点击“全部问题”显示本课程所有学员提问的问题。
- 2、点击“提问”即可向该课程的老师 and 助教提问问题。



联系我们

小象学院：互联网新技术在线教育领航者

— 微信公众号：**小象学院**

