



华北电力大学
NORTH CHINA ELECTRIC POWER UNIVERSITY

8 图像检索



主要内容

- 图像检索概述
 - ✓ 基于文本的图像检索
 - ✓ 基于内容的图像检索
 - ✓ 常用的图像数据库
- 基于颜色特征的图像检索
 - ✓ 颜色空间
 - ✓ 颜色量化
 - ✓ 颜色特征描述
 - ✓ 特征匹配
 - ✓ 性能评价



主要内容

- 基于纹理的图像检索
 - ✓ 纹理特征
 - ✓ 纹理特征提取
 - ✓ 纹理匹配
- 基于形状的图像检索
 - ✓ 形状特征
 - ✓ 特征匹配
- 基于语义的图像检索
 - ✓ 语义特征
 - ✓ 语义特征表示方法



8.1 图像检索

- 随着计算机网络、社交媒体、数字电视和多媒体获取设备的快速发展，以图像和视频为代表的多媒体数据的生成、处理和获取变得越来越方便，多媒体应用日益广泛，数据量呈现出爆炸性的增长，已经成为**大数据**时代的主要数据对象。
- 如何在海量的图像数据中以较小的时空开销准确地找到一幅**感兴趣的图**像，已经成为近年来多媒体和信息检索领域的重要研究热点。



8.1.1 基于文字的图像检索

- 早期，互联网上常用的搜索引擎，如：百度，Yahoo，必应等主要提供以**文本**关键词作为检索内容的服务，即接受基于文本的检索方式，而对于多媒体信息只能通过人工标注来描述。





8.1.1 基于文字的图像检索

- 由于图像的种类多样和数量庞大，人工标注的成本过于高昂。
- 另一方面，由于人工标注通常带有主观性，关键词难以准确和全面地描述图像信息。



8.1.2 基于内容的图像检索

- 自从20世纪90年代末以来，**基于内容的图像检索**（Content-based Image Retrieval, CBIR）研究开始得到人们的极大关注，并成为信息检索研究领域的热点问题。
- 基于内容的图像信息检索，是对图像进行内容语义的分析和特征的提取，并基于这些特征进行相似性匹配的信息检索技术。



8.1.2 基于内容的图像检索

- 基于内容的图像检索技术特点
 - ✓ (1) 基于内容的检索突破了传统的基于关键字的检索的局限，直接对图像内容进行分析，抽取特征和语义，利用这些内容特征建立索引，并进行检索。
 - ✓ (2) 基于内容的检索是一种相似匹配。在检索过程中，它不像传统检索方法仅支持精确匹配，更主要的是支持相似匹配，如范围查询和最邻近查询，得到最相似对象。

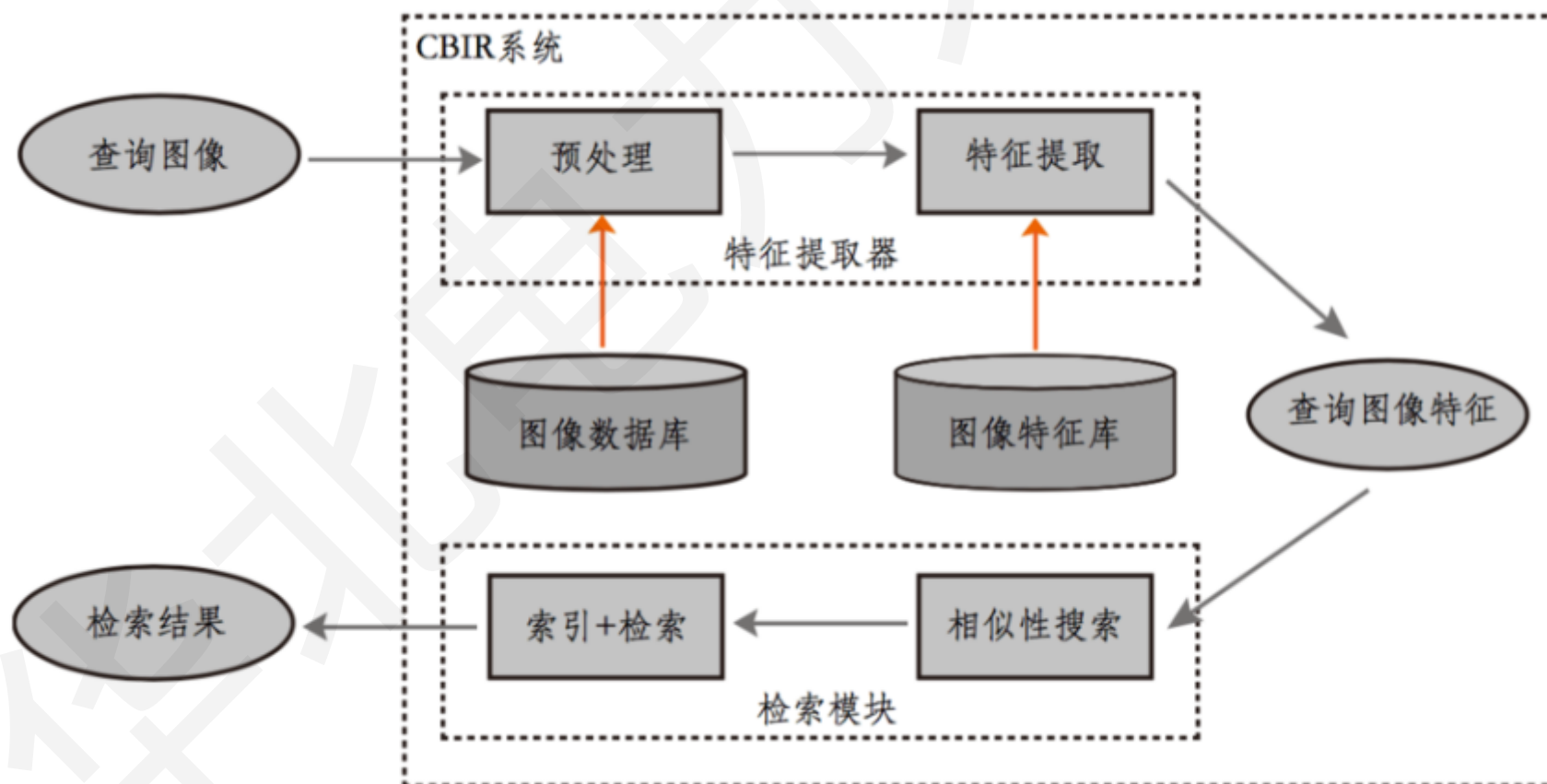


8.1.2 基于内容的图像检索

- 基于内容的图像检索技术特点
 - ✓ (3) 加入相关反馈手段优化检索。计算机语言所能表示对象的颜色和相似性等跟人的感知认识存在一定距离，加入人工反馈机制，检索系统通过学习用户的认识和意图来指导检索过程，提高检索系统的准确性。
 - ✓ (4) 结合机器学习技术，将成熟的学习算法与多媒体检索中的学习过程结合起来以提高检索准确率。

8.1.2 基于内容的图像检索

- 基于内容的图像检索的框架



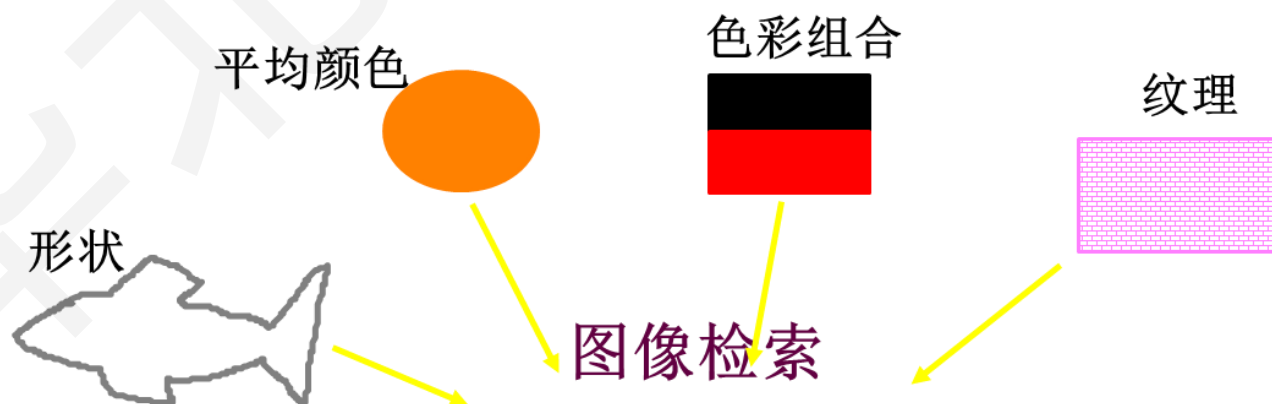


8.1.2 基于内容的图像检索

- 基于内容的图像检索的分类
 - ✓ 低层特征：颜色、纹理和形状等特征
 - ✓ 中层特征：图像中的对象、背景以及空间关系
 - ✓ 语义特征：场景、事件、情感等

8.1.2 基于内容的图像检索

- 基于内容的图像检索的原理
 - ✓ 通过分析图像的内容，提取其颜色、形状、纹理等可视特征
 - ✓ 建立特征索引存储于特征库中
 - ✓ 在检索时，用户只需把自己对图像的模糊印象描述出来，就可以通过多次的近似匹配，在大容量图像库中查询到所需图像



8.1.2 基于内容的图像检索

• 基于内容的图像检索的挑战



(a)光照变化

(b)尺度变化

(c)视角变化

(d)遮挡

(d)背景杂乱

相同物体相似性挑战



(a)类内变化巨大(湖泊)



(b)类间相似性干扰

相似物体相似性挑战



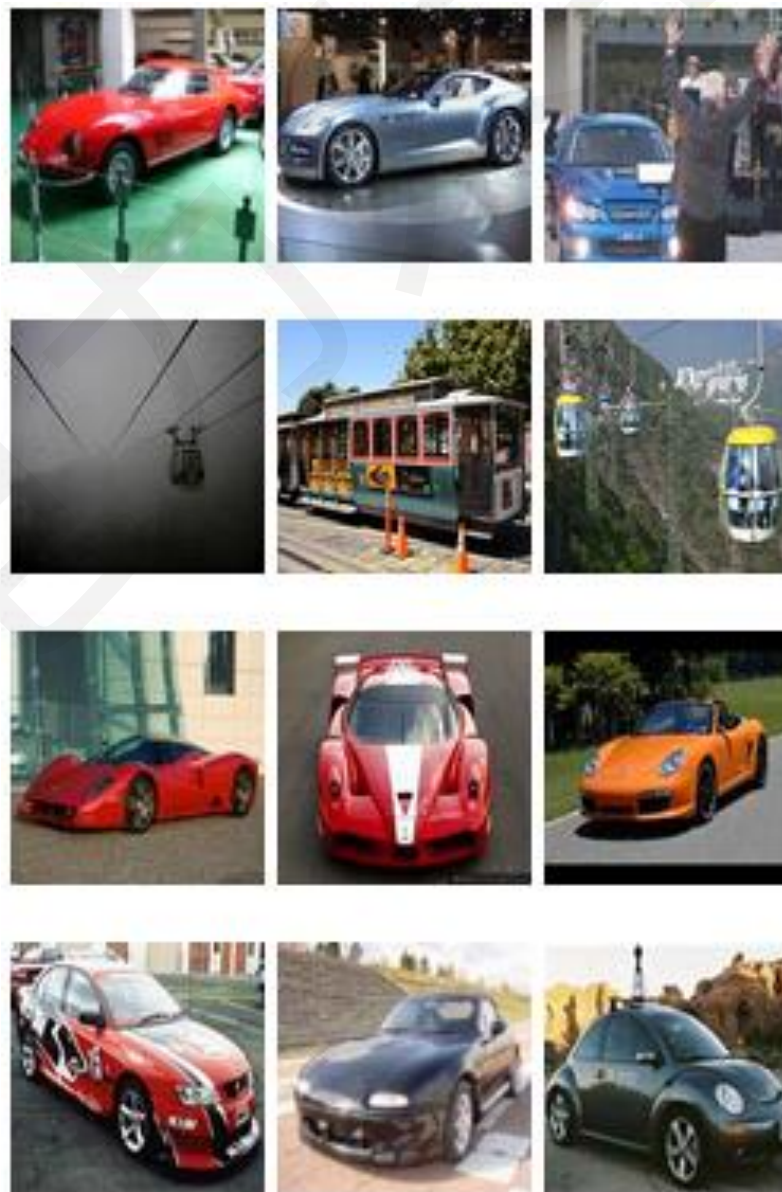
8.1.3 常用的图像数据库

- ImageNet

- ✓ ImageNet是一个计算机视觉系统识别项目，是目前世界上图像识别最大的数据库。是美国斯坦福的计算机科学家李飞飞模拟人类的识别系统建立的。能够从图片识别物体。
- ✓ 目前已经包含14,197,122张图像，是已知的最大的图像数据库。每年的ImageNet大赛更是魂萦梦牵着国内外各个名校和大型IT公司以及网络巨头的心。
- ✓ 需要注册ImageNet帐号才可以下载，下载链接为<http://www.image-net.org/>

8.1.3 常用的图像数据库

- ImageNet





8.1.3 常用的图像数据库

- PASCAL VOC

- ✓ PASCALVOC 数据集是视觉对象的分类识别和检测的一个基准测试，提供了检测算法和学习性能的标准图像注释数据集和标准的评估系统。
- ✓ 包含VOC2007 (430M) ， VOC2012 (1.9G) 两个下载版本。
- ✓ 下载链接为<http://pjreddie.com/projects/pascal-voc-dataset-mirror/>

8.1.3 常用的图像数据库

- PASCAL VOC



2007_000032.jp

g



2007_000033.jp

g



2007_000039.jp

g



2007_000042.jp

g



2007_000061.jp

g



2007_000063.jp

g



2007_000068.jp

g



2007_000121.jp

g



8.1.3 常用的图像数据库

- Labelme

✓ Labelme是斯坦福一个学生的母亲利用休息时间帮儿子做的标注，后来便发展为一个数据集。该数据集的主要特点包括：

- (1) 专门为物体分类识别设计，而非仅仅是实例识别
- (2) 专门为学习嵌入在一个场景中的对象而设计
- (3) 高质量的像素级别标注，包括多边形框 (polygons) 和背景标注 (segmentation masks)
- (4) 物体类别多样性大，每种物体的差异性，多样性也大。
- (5) 所有图像都是自己通过相机拍摄，而非copy
- (6) 公开的，免费的



8.1.3 常用的图像数据库

- Labelme

✓ Labelme是需要通过matlab来下载，一种奇特的下载方式，下载链接为

<http://labelme2.csail.mit.edu/Release3.0/index.php>



coast_arnat59.jpg



coast_art294.jpg



coast_art487.jpg



coast_art1130.jpg



coast_bea1.jpg



coast_bea2.jpg



coast_bea3.jpg



coast_bea4.jpg



8.1.3 常用的图像数据库

- COCO

✓ COCO是一种新的图像识别，分割和加字幕标注的数据集，其主要特征如下：

- (1) 目标分割
- (2) 通过上下文进行识别
- (3) 每个图像包含多个目标对象
- (4) 超过300000个图像
- (5) 超过2000000个实例
- (6) 80种对象
- (7) 每个图像包含5个字幕
- (8) 包含100000个人的关键点

8.1.3 常用的图像数据库

- COCO
 - ✓ COCO支持Matlab和Python两种下载方式，下载链接为<http://mscoco.org/>



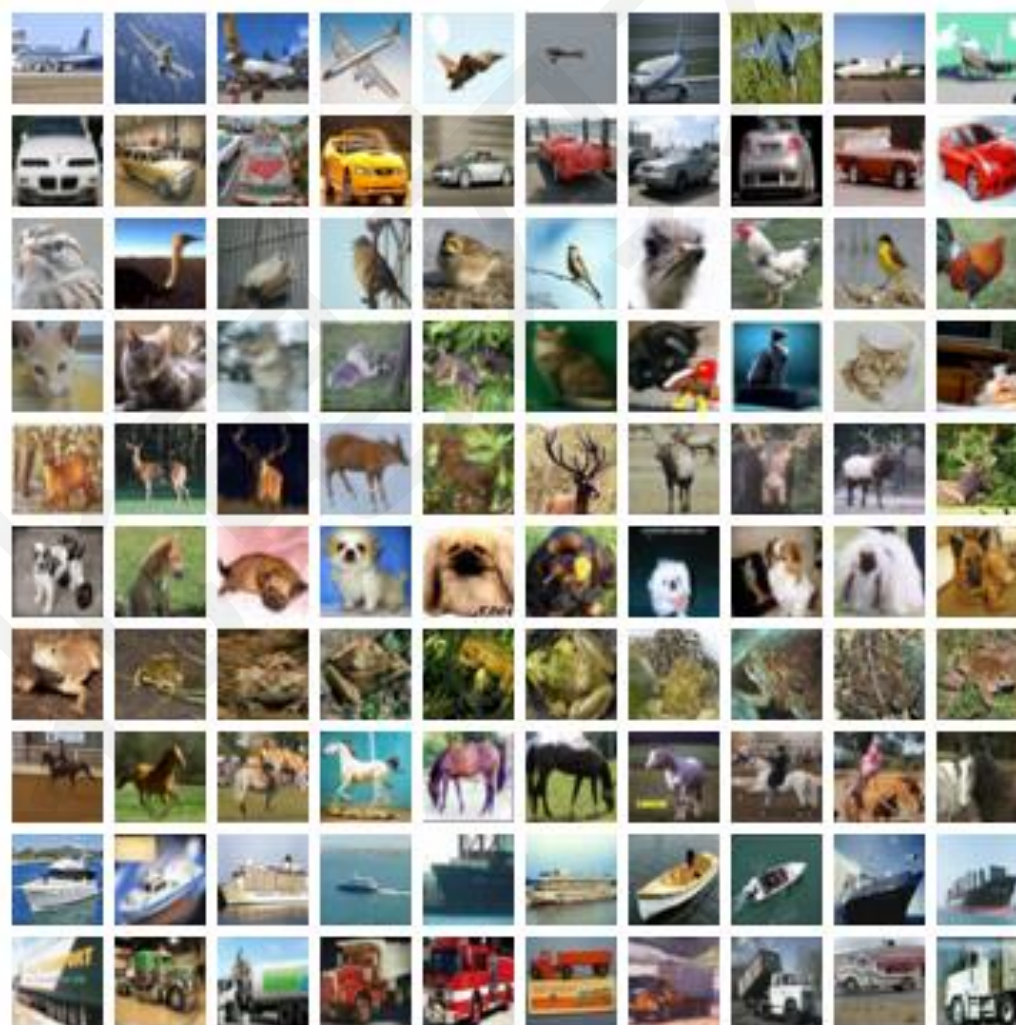


8.1.3 常用的图像数据库

- CIFAR (Canada Institute For Advanced Research)
 - ✓ 是由加拿大先进技术研究院的AlexKrizhevsky, Vinod Nair和Geoffrey Hinton收集而成的小图片数据集。
 - ✓ 包含CIFAR-10和CIFAR-100两个数据集。Cifar-10由60000张32*32的RGB彩色图片构成，共10个分类。50000张训练，10000张测试（交叉验证）。
 - ✓ CIFAR-100由60000张图像构成，包含100个类别，每个类别600张图像，其中500张用于训练，100张用于测试。
 - ✓ 官网提供了Matlab,C, Python三个版本的数据格式。下载链接为<http://www.cs.toronto.edu/~kriz/cifar.html>

8.1.3 常用的图像数据库

- CIFAR (Canada Institute For Advanced Research)





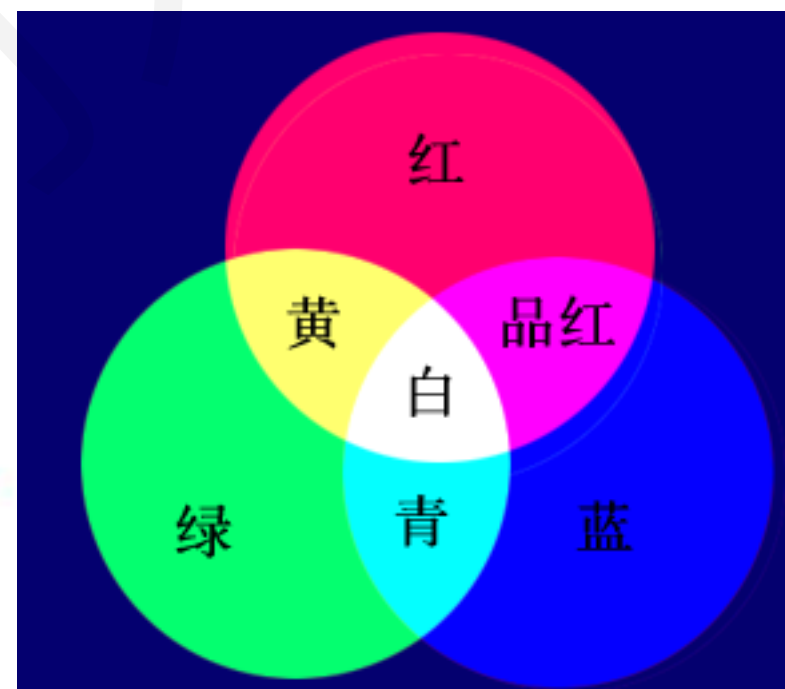
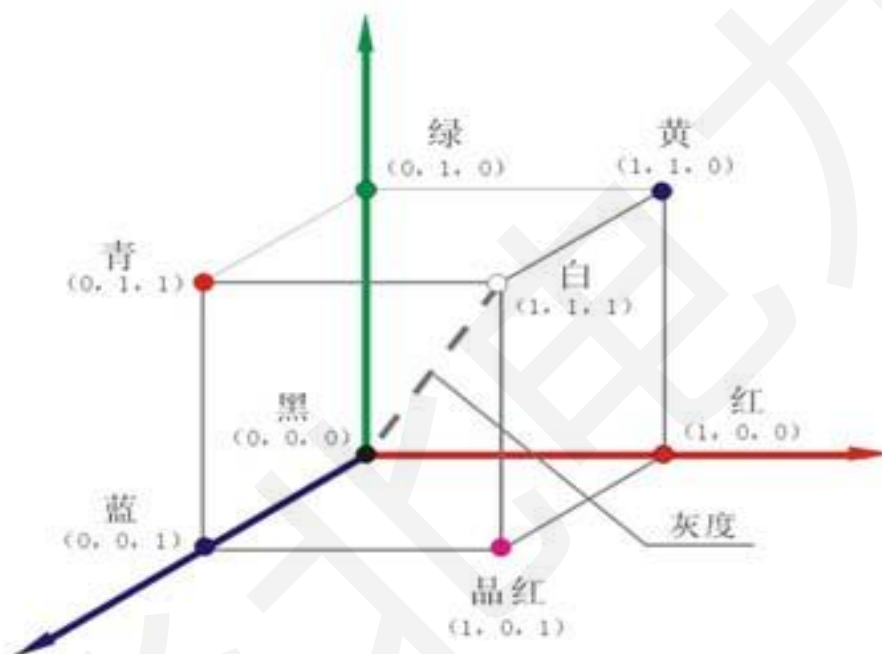
8.2 基于颜色特征的图像检索

- 检索过程

- ✓ 需要选择合适的颜色空间来描述颜色特征
- ✓ 采用一定的量化方法将颜色特征表达为向量的形式
- ✓ 以一定的方式来描述颜色特征
- ✓ 定义一种相似度（距离）标准来衡量特征间的相似性
- ✓ 性能评价

8.2.1 颜色空间

- RGB颜色空间





8.2.1 颜色空间

- RGB颜色空间的缺点

- ✓ 不直观，从RGB值中很难知道该值所表示颜色的认知属性，因此，RGB颜色空间不符合人对颜色的感知心理。

- ✓ 另外，RGB颜色空间是最不均匀的颜色空间之一，两种颜色之间的知觉差异不能采用该颜色空间中两个颜色点之间的距离来表示。

8.2.1 颜色空间

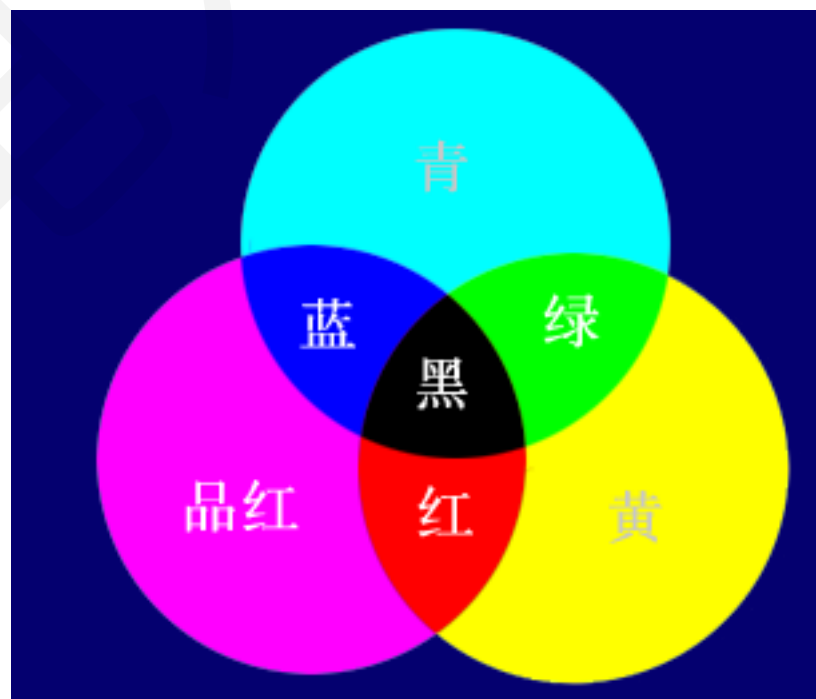
- CMY颜色空间

✓ 青 (Cyan)、品红 (Magenta) 和黄 (Yellow) 三种颜色的简写，加上黑色 (Black)，即为CMYK，相减混色模式。

$$C = 1 - R$$

$$M = 1 - G$$

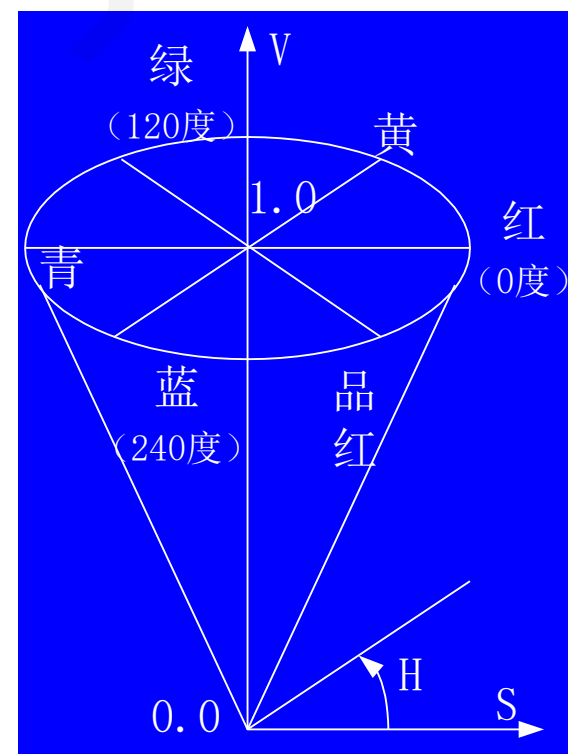
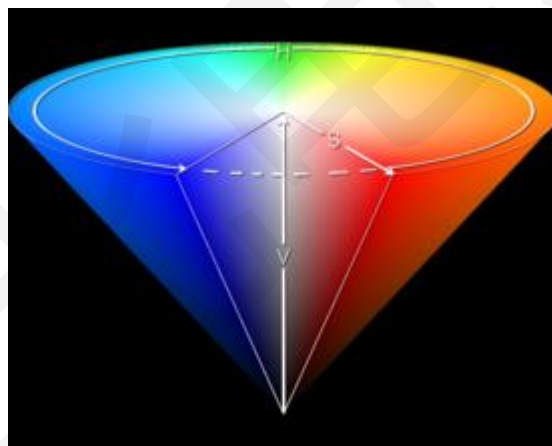
$$Y = 1 - B$$



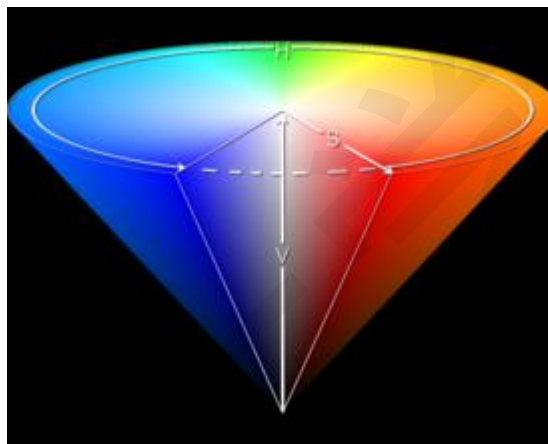
8.2.1 颜色空间

- HSV颜色空间

✓ 用一个三维空间纺锤体将色调 (Hue)，饱和度 (Saturation) 和亮度 (Value) 表示出来。



8.2.1 颜色空间



- HSV颜色空间

- ✓ 色调H参数表示色彩信息，即所处的光谱颜色的位置。该参数用一角度量来表示，取值范围为 $0^{\circ} \sim 360^{\circ}$ 。若从红色开始按逆时针方向计算，红色为 0° ，绿色为 120° ，蓝色为 240° 。它们的补色是：黄色为 60° ，青色为 180° ，紫色为 300° 。
- ✓ 饱和度S：取值范围为 $0.0 \sim 1.0$ 。
- ✓ 亮度V：取值范围为 0.0 (黑色) ~ 1.0 (白色)。



8.2.1 颜色空间

- HSV颜色空间的优点
 - ✓ 亮度分量与图像的彩色信息无关
 - ✓ 色调和饱和度分量与人感受颜色的方式是紧密相连的。
- 这些特点使**HSV颜色空间**非常适合于借助人的视觉系统来感知颜色特性的图像处理算法，因此在基于内容的图像检索中，应用这种彩色模型会更适合用户的视觉判断。



8.2.2 颜色量化

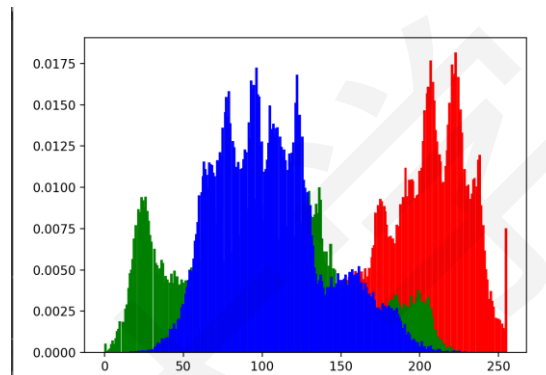
- 颜色量化是图像工程中的一项基本而重要的技术，是进行图像分割和对象提取的基础。
- 对于一幅真彩色图像来说，颜色种类可达到 2^{24} 种颜色，在图像检索过程中，对颜色数目有一定的限制，如何选取有代表性的若干种颜色，并把各种颜色都归并到这些代表色上，就是颜色量化需要解决的问题。



8.2.2 颜色量化

- 分割算法：基本思想是将图像中出现的频率**最高的K种**色彩作为调色板，然后将其余颜色按照最小距离准则映射到调色板中。
- 聚类算法：先选择若干**聚类中心**，然后按某种准则对颜色进行迭代聚合，直到合适的分类为止。

8.2.3 颜色特征描述



- 颜色直方图

✓ 颜色直方图具有**旋转不变性**、**尺度不变性**和**平移不变性**，被广泛的应用到图像检索中。

✓ 核心思想是在颜色空间中采用一定的量化方法对颜色进行量化，然后统计每一个量化通道在整幅图像颜色中所占的**比重**，横坐标表示颜色的色彩归一化的值，纵坐标表示每个小的颜色区间中所对应的图像像素的比重。



8.2.3 颜色特征描述

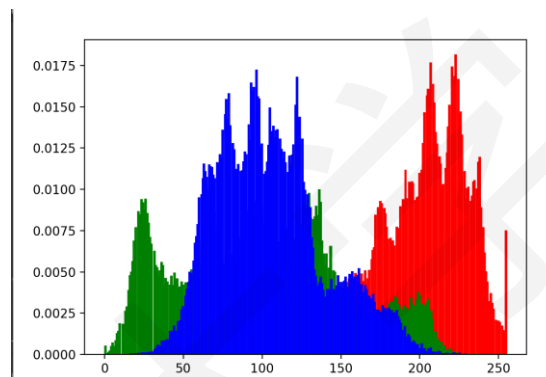
设 $C(x, y)$ 为颜色空间一幅彩色图像，其颜色直方图定义为：

$$H_c(k) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} h(C(x, y)), k = 0, 1, \dots, K$$

$$h(C(x, y)) = \begin{cases} 1 & \text{如果 } C(x, y) \text{ 在变换空间量化后等于 } K \\ 0 & \text{其他} \end{cases}$$

其中， M 、 N 为图像 $C(x, y)$ 垂直和水平方向上的像素数目，并简称为图像的高和宽；而 K 为变换空间的颜色数。

8.2.3 颜色特征描述



- 颜色直方图的缺点

- ✓ 描述的是图像颜色的统计特征，丢失了颜色的空间分布信息，对于两个颜色直方图相似的图像来说，如果颜色的空间分布差别很大，图像的内容会很不相同。
- ✓ 存在维数过高的问题。
- ✓ 由于对图像颜色的量化处理，可能将视觉不同的颜色量化到同一区间，也可能将视觉相同的颜色量化到不同的区间，因而造成误检现象。



8.2.3 颜色特征描述

- 颜色熵

✓ 根据颜色直方图特性和信息论中信息熵的概念，John提出采用图像颜色的信息熵来表示图像的颜色特征，从而将图像的颜色直方图有多维降低到一维，然后采用L1距离进行图像间的距离度量。

✓ 设图像的颜色直方图表示为 (h_1, h_2, \dots, h_n) ，如果我们将图像的颜色直方图看做是图像中不同颜色的像素在空间中出现的概率密度函数，则图像的信息熵可表示为：

$$E = \sum_{i=1}^n h_i \log_2(h_i)$$



8.2.3 颜色特征描述

- 颜色矩

- ✓ 颜色矩 (color moments) 是一种非常简单而有效的颜色特征，这种方法的数学基础在于图像中任何的颜色分布均可以用它的矩来表示。
- ✓ 由于颜色分布信息主要集中在低阶矩中，因此仅采用颜色的一阶矩 (均值, mean)、二阶矩 (方差, variance) 和三阶矩 (斜度, skewness) 就足以表达图像的颜色特征。
- ✓ 与颜色直方图相比，该方法的另一好处在于无需对特征进行量化且降低了颜色特征的维数。



8.2.3 颜色特征描述

- 颜色矩的数学定义

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij}$$

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2}$$

$$s_i = \sqrt[3]{\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3}$$

其中， f_{ij} 表示第*i*个颜色通道分量中灰度为*j*的像素出现的概率，*N*表示灰度级数。每个像素有三个颜色分量，每个分量有三个低阶矩，共有九个分量，与其它的颜色特征相比是非常简洁的。



8.2.4 特征匹配

- 基于内容的图像检索技术通过计算示例图像与数据库中图像之间的相似度进行检索。
- 由于视觉特征大都可以表示成向量的形式，所以常用的相似度匹配算法都是向量空间模型，即将视觉特征看作是向量空间中的点，通过计算两个点之间的接近程度，来衡量图像特征间的相似度。
- 当两幅图像的特征间的距离小于某一个阈值时，图像匹配成功，称为相似匹配。



8.2.4 特征匹配

- Minkowski(明考夫斯基)距离法
是基于 L_p 范数定义的，其数学描述为：

$$L_p(A, B) = \left[\sum_{i=1}^n |a_i - b_i|^p \right]^{\frac{1}{p}}$$

当 $p=1$ 时，曼哈顿距离（城市距离）

$$L_1(A, B) = \sum_{i=1}^n |a_i - b_i|$$



8.2.4 特征匹配

当 $p=2$ 时，欧氏距离

$$L_2(A, B) = \left[\sum_{i=1}^n |a_i - b_i|^2 \right]^{\frac{1}{2}}$$

当 $p=\infty$ 时，切比雪夫距离

$$L_{\infty}(A, B) = \max \left[\sum_{i=1}^{\infty} |a_i - b_i| \right]$$

没有一种方法可以适用于所有特征向量间的相似度量。欧氏距离和城区距离相对于其它相似度量方法具有更好的检索性能。



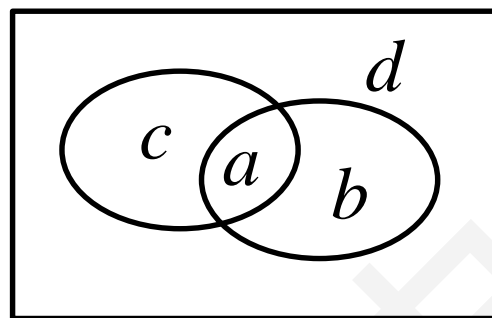
8.2.5 性能评价

- 查准率(Precision)和查全率(Recall)
 - ✓ 查准率是指返回的结果中有效图像的比率，反映了图像检索的**准确性**。
 - ✓ 查全率是指返回的结果中有效图像占数据库中所有相似图像的数量比率，反映了图像检索的**全面性**。



8.2.5 性能评价

- 查准率(Precision)和查全率(Recall)



$$a + c = A$$

$$a + b = B$$

a: 检索正确的样本
b: 检索错误的样本
c: 未被检索的正确样本
A: 应被检索的样本
B: 检索结果的样本

$$Precision = p(A|B) = \frac{p(A \cap B)}{P(B)} = \frac{a}{a + b}$$

$$Recall = p(B|A) = \frac{p(A \cap B)}{P(A)} = \frac{a}{a + c}$$



8.2.6 参考代码

%读取目标图像

```
[file,path] = uigetfile();
```

```
I=imread(strcat(path,file));
```

```
J=rgb2hsv(I); %转化为hsv颜色空间
```

```
h=J(:,:,1);
```

```
s=J(:,:,2);
```

```
v=J(:,:,3);
```

```
histI = zeros(3,256); %三个分量的直方图
```

```
histI(1,:) = imhist(h);
```

```
histI(2,:) = imhist(s);
```

```
histI(3,:) = imhist(v);
```




8.2.6 参考代码

```
%读取图像数据库中所有图像数据  
filePath = '.\images\'; %图像文件夹路径  
imagePathList = dir(strcat(filePath, '*.jpg')); %得到所有文件的名称  
%dir 列出当前文件夹中的文件信息  
imageNum = length(imagePathList) %获取图像总数量  
imageNames = strings(imageNum, 1); %保存每个图像的名称  
similarities = zeros(imageNum, 1); %每个图像与目标图像的相似度  
fprintf('正在读取的图像为: \n');
```



8.2.6 参考代码

```
if imageNum > 0 %有满足条件的图像
    for j = 1:imageNum %逐一读取图像
        imageName = imagePathList(j).name;% 图像名
        fprintf('第%d个: %s\n',j,strcmp(filePath,imageName)); %显示信息
        imageNames(j) = imageName; %保存文件名
        image = imread(strcmp(filePath,imageName)); %图像数据
        imageJ=rgb2hsv(image); %hsv颜色空间
        hj=imageJ(:,:,1);
        sj=imageJ(:,:,2);
        vj=imageJ(:,:,3);
        histJ = zeros(3,256); %图像的HSV直方图
        histJ(1,:) = imhist(hj);
        histJ(2,:) = imhist(sj);
        histJ(3,:) = imhist(vj);
```



8.2.6 参考代码

%计算相似度

```
S1 = histl(1,:)-histJ(1,:);
```

```
S2 = histl(2,:)-histJ(2,:);
```

```
S3 = histl(2,:)-histJ(2,:);
```

% 保存相似度

```
similarities(j) = norm(S1,2) + norm(S2,2) + norm(S3,2);
```

```
end
```

```
end
```




8.2.6 参考代码

%显示检索结果

figure;

subplot(2,3,1);imshow(I);title('目标图像') %绘制目标图像

[v, index] = sort(similarities); %根据相似度排序, index是排序后原序号

for i=1:5 %取距离最小的5个图像

 id = index(i); %获得原始序号

 resultimage = imread(strcat(filePath,imageNames(id)));

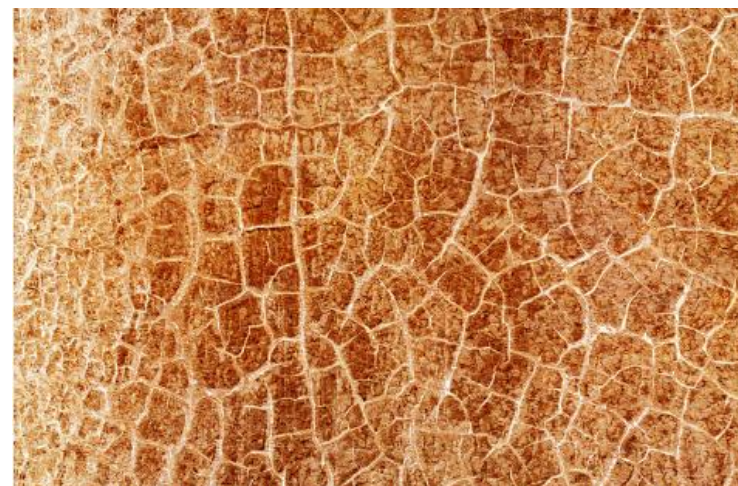
 subplot(2,3,i+1);

 imshow(resultimage); %绘制相似图像

end

8.3 基于纹理特征的图像检索

- 纹理 (Texture) : 把局部不规则而宏观有规律的特性称之为纹理。
 - ✓ 例如云彩、树木、砖等都有各自的纹理特征。





8.3 基于纹理特征的图像检索

- 纹理是一种不依赖与颜色和亮度的反映图像同质现象的视觉特征，体现了物体表面的**内在特性**。
- 纹理还包含了物体表面组织**结构排列**的重要信息以及它们与周围环境的联系。
- 因此，纹理特征在基于内容的图像检索中得到了广泛的应用。



8.3.1 纹理特征

- 是指利用计算机技术从数字图像中计算出来的，可以**定量描述**人对纹理的定性的感知的某些参数。
- 它对区域内部灰度变化或色彩变化的某种规律进行**量化**，能够尽可能地缩小纹理的**类内差距**，同时尽可能增大纹理的**类间差距**。



8.3.1 纹理特征

- 二维灰度直方图

纹理特征：对比度、角二阶矩、熵、均值。

- 灰度共生矩阵

纹理特征：对比度、角二阶矩、相关性、熵、差分矩、逆差分矩、和平均、和方差、和熵。

- 灰度-梯度共生矩阵

纹理特征：小梯度优势、大梯度优势、灰度分布不均匀性、梯度分布不均匀性、能量、相关、灰度熵、梯度熵、惯性。



8.3.2 纹理特征提取

- 统计法纹理提取

- ✓ 根据人们的直观视觉感受，从心理因素出发，形成纹理特征的一种表示方法。
- ✓ 主要思想是：通过图像中**灰度级分布**的随机属性来描述纹理特征，比较适合描述自然纹理，常可提供纹理的平滑、稀疏、规则等性质。
- ✓ 适用于分析木纹、沙地、草坪等纹理细节不规则的图像。



8.3.2 纹理特征提取

- 结构法纹理提取

- ✓ 是建立在纹理基元理论基础上的的一种纹理分析方法。
- ✓ 纹理基元理论认为，复杂的纹理可以由若干简单的纹理基元以一定的有规律的形式重复排列构成。
- ✓ 结构法利用基元排列的规则性来分析纹理，主要包含两个要素：基元特征和基元组合规则。
- ✓ 结构法适用于纹理基元较大而有规则的物体，如布料，砖墙等。



8.3.2 纹理特征提取

- 频谱法纹理提取

- ✓ 是将纹理图像看做二维信号，使用滤波方法对纹理图像进行分析的一种方法。
- ✓ 主要借助于各种变换算法利用图像的频率特性来描述纹理特征。
- ✓ 频谱法主要包括傅里叶变换法、Gabor变换法和小波变换法。



8.3.3 纹理匹配

- 纹理匹配方法

- ✓ 基于统计法纹理提取特征—共生矩阵，心理学特征等
- ✓ 基于模型法纹理提取特征—Markov随机场模型等
- ✓ 基于频谱法纹理提取特征—小波变换，Gabor滤波器等



8.4 基于形状特征的图像检索

- 形状特征是比较颜色和纹理特征更高一层的特征，属于图像的**中间层特征**，人类视觉系统进行物体识别时所识别的关键信息之一。
- 图像的形状信息不随图像颜色等特征的变化而变化，是物体**稳定的特征**。
- 要获取目标的形状参数首先要进行图像**分割**，同时要保证所提取的形状特征不受图像平移、缩放、旋转等变换的影响。



8.4 基于形状特征的图像检索

- 基本思想：
 - ✓ 是通过**图像分割**技术（自动及半自动方法）将图像划分为不同的区域
 - ✓ 对于每一个区域使用**局部特征**来进行描述，综合区域的局部特征从而得到图像的**总体特征**描述
 - ✓ 使用合适的**相似度**度量标准来检索图像

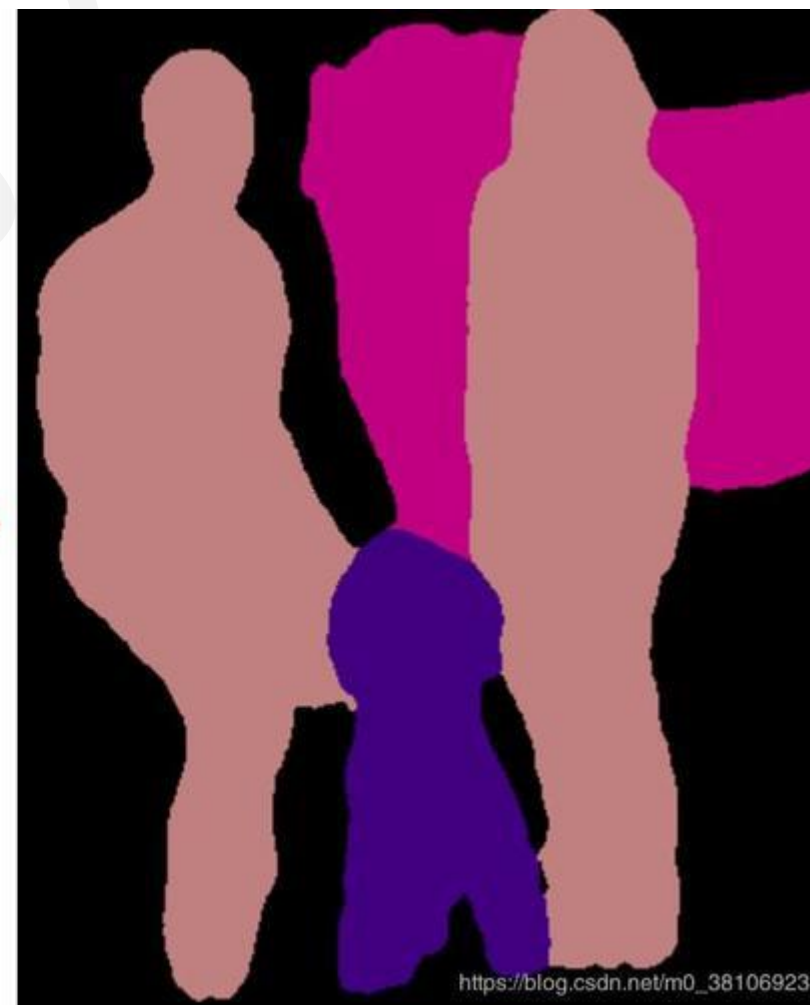


8.4.1 形状特征

- 形状是刻画物体的本质特征之一，通常与目标联系在一起。
- 形状通常分为两种类型：
 - ✓ 一类是基于轮廓特性的，如傅里叶描述
 - ✓ 一类是基于区域特性的，如几何不变矩

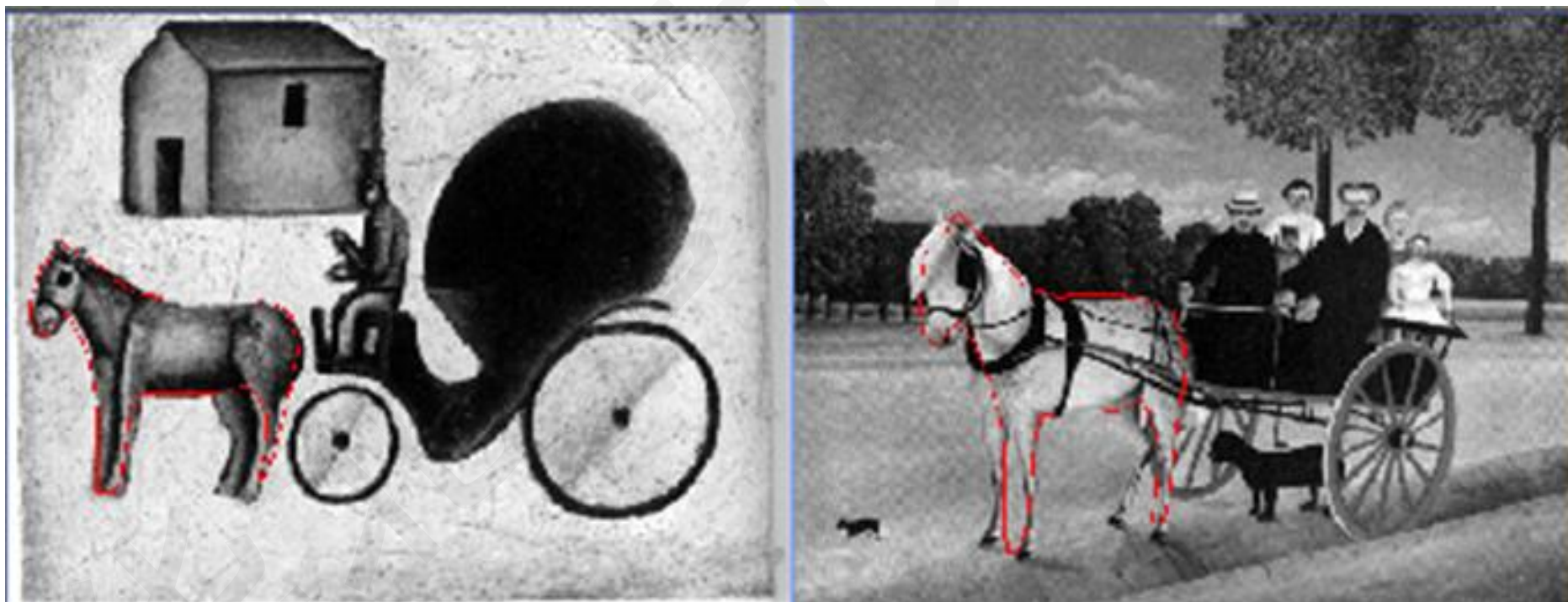
8.4.1 形状特征

- 面积



8.4.1 形状特征

- 轮廓

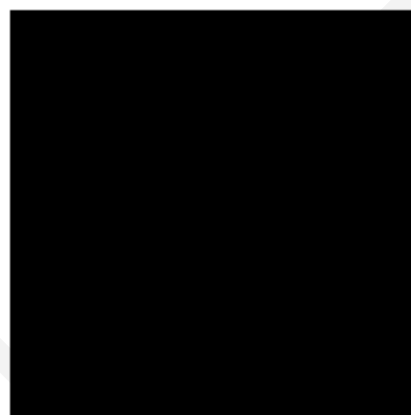




8.4.2 特征匹配

- 面积匹配

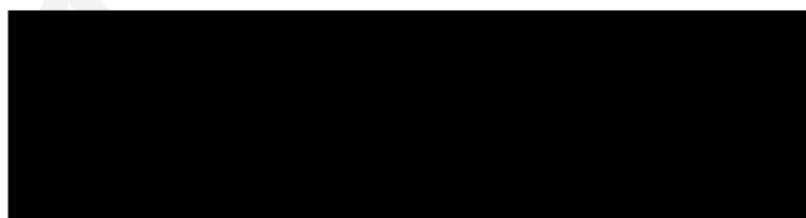
哪2个形状更相似?



200*200



180*180



400*100

8.4.2 特征匹配

- 轮廓匹配



(a) 西藏



(b) 青海



(c) 四川



(d) 云南



8.4.2 特征匹配

- 轮廓匹配-Hausdorff距离

给定欧氏空间中的两点集 $A = \{a_1, a_2, \dots\}$, $B = \{b_1, b_2, \dots\}$, Hausdorff距离就是用来衡量这两个点集间的距离。

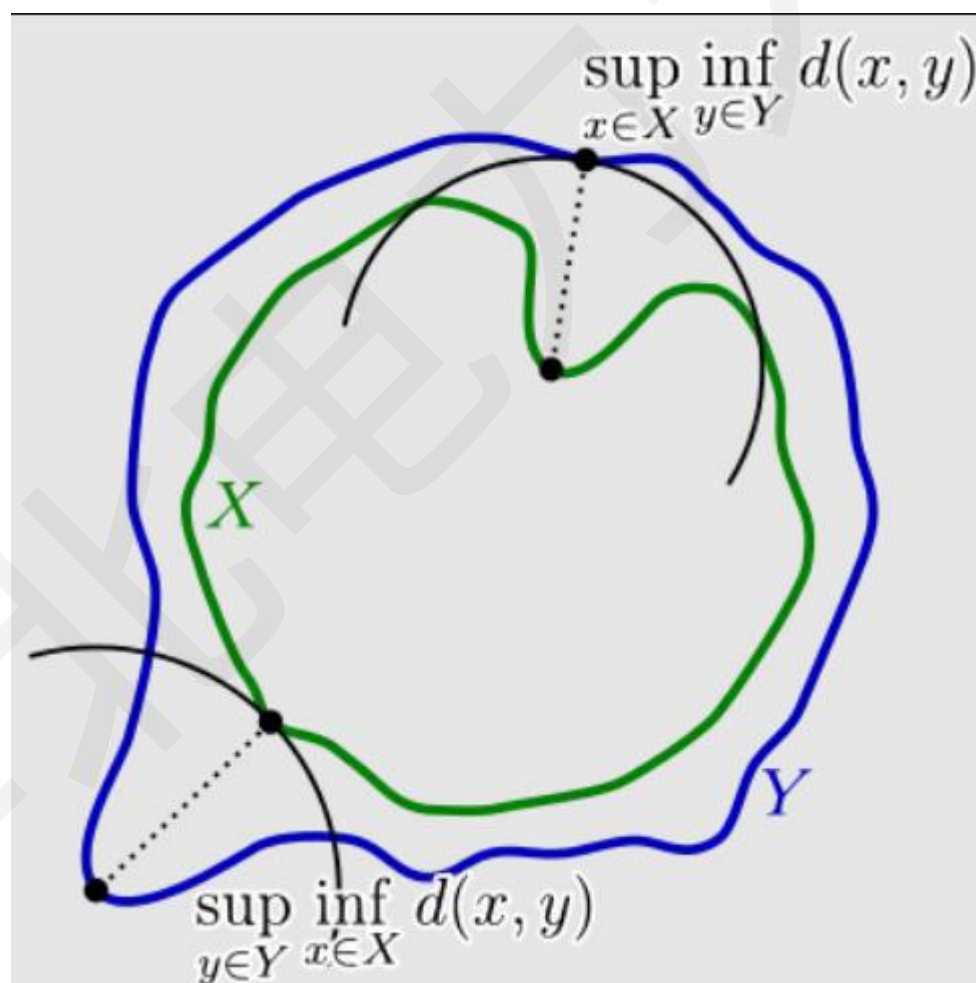
$$H(A, B) = \max[h(A, B), h(B, A)]$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

$$h(B, A) = \max_{b \in B} \min_{a \in A} \|b - a\|$$

8.4.2 特征匹配

- 轮廓匹配-Hausdorff距离





8.5 基于语义特征的图像检索

- 以上基于内容的图像检索技术，多数基于图像底层可视化特征的相似性实现，忽略了图像的语义相似性。
- 而图像的高层语义概念则能更好地表达人眼对图像的理解，它与计算机自动提取的低层可视化特征之间缺乏直接的联系。

8.5 基于语义特征的图像检索

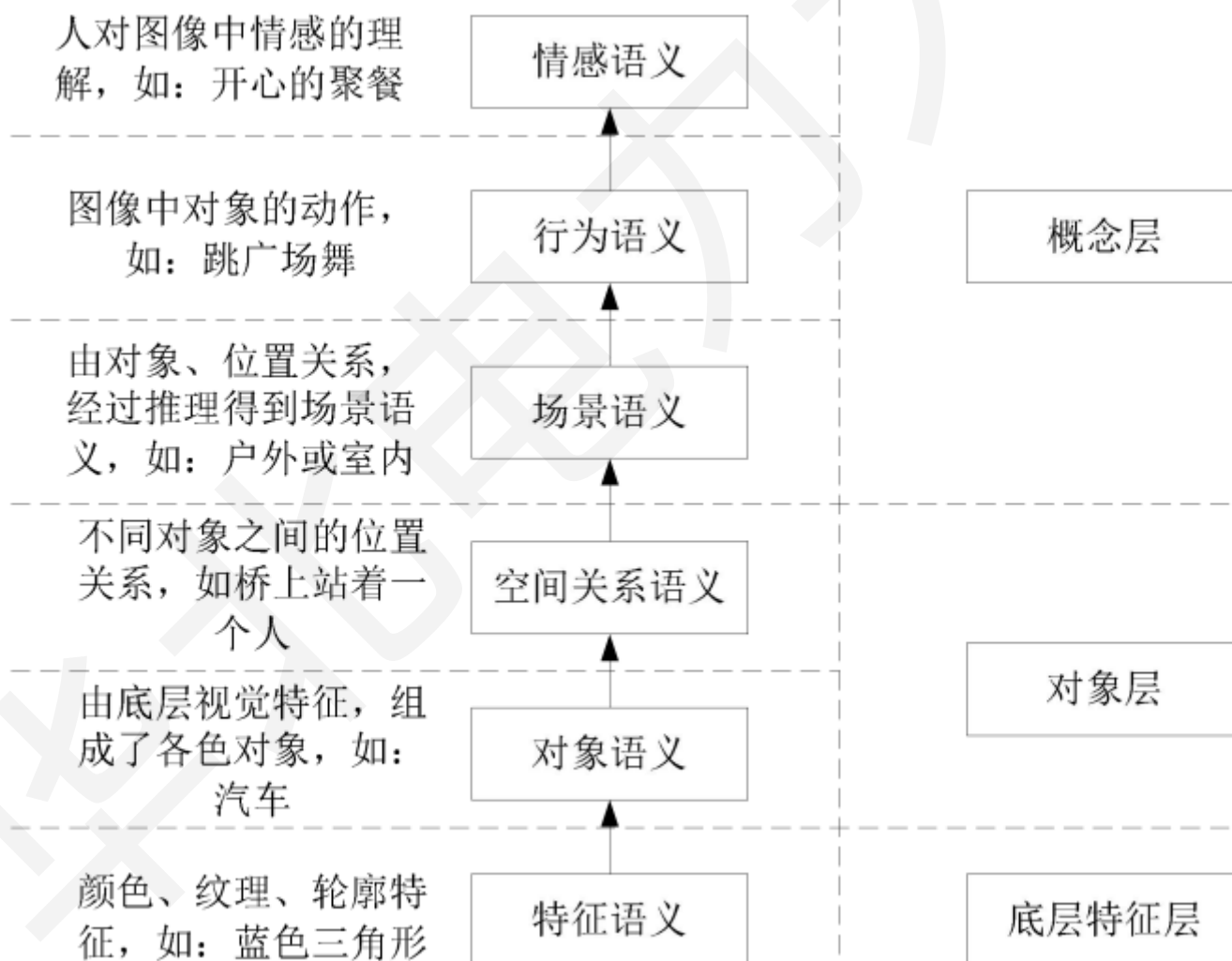
- 具体表现为：可视化特征向量相似度很高的图像可能包含的语义信息并不一致
- 相反，包含语义信息相近的图像可能可视化特征向量相似度很低。
- 由此产生的鸿沟称为“**语义鸿沟**”。





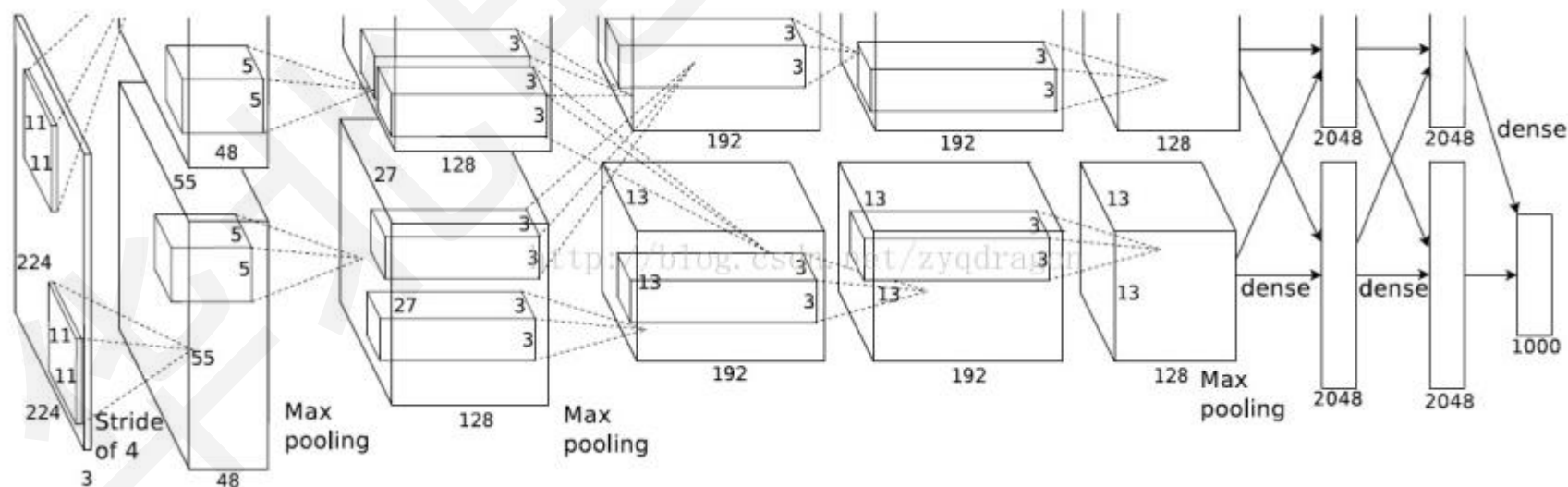
8.5.1 语义特征

- 不同层次的语义



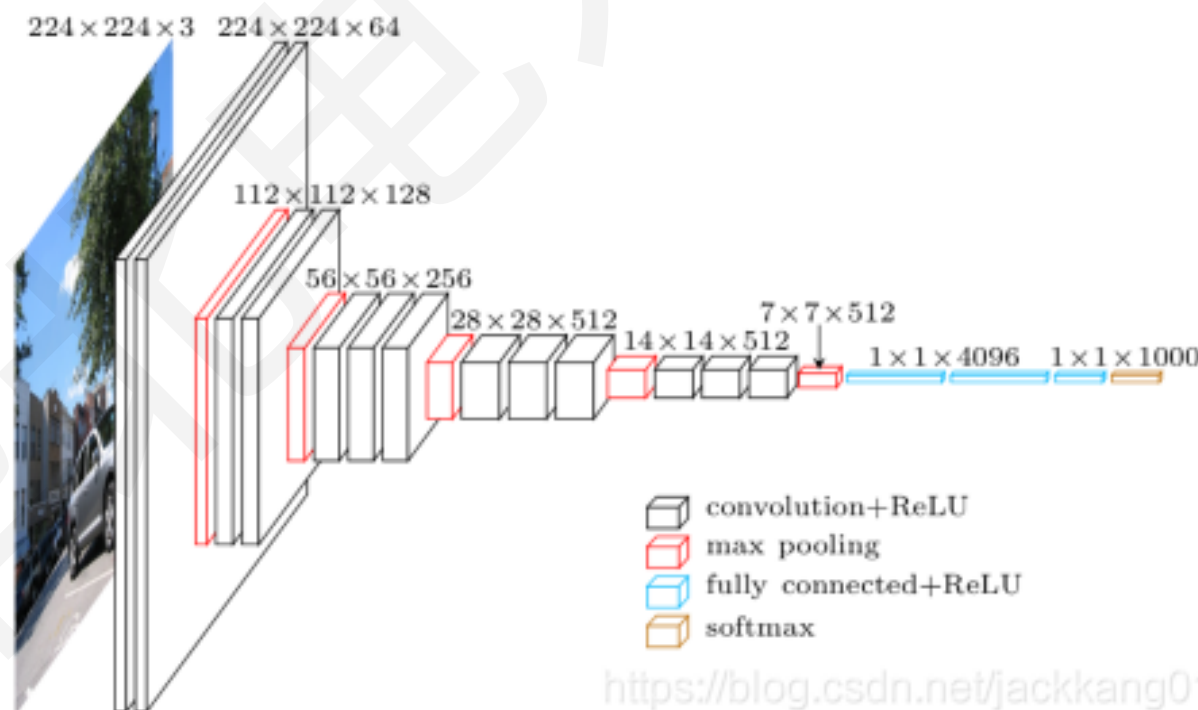
8.5.2 语义特征的表达

- 高层的语义特征多采用深度学习模型表达。
 - ✓ AlexNet, 深度学习的先驱之一——Geoffrey Hinton和他的同事们引入的第一个深层架构。它是一个简单而又强大的网络架构, 它为深度学习的突破性研究铺平了道路。



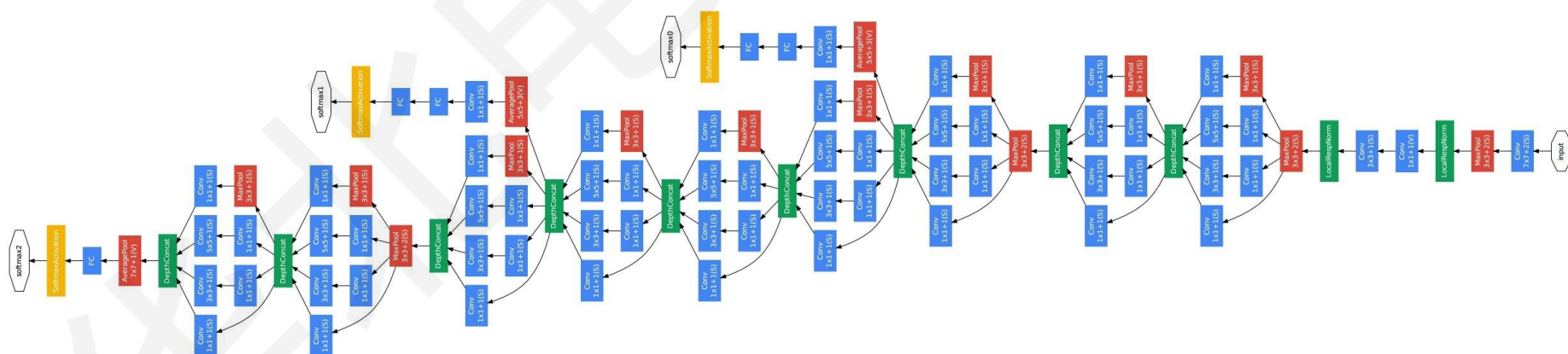
8.5.2 语义特征的表达

- 高层的语义特征多采用深度学习模型表达。
 - ✓ VGG 网络是由牛津大学视觉图形组的研究人员引入的。这个网络的特点是它的金字塔形状，最接近图像的底层是宽的，而最上层是深的。



8.5.2 语义特征的表达

- 高层的语义特征多采用深度学习模型表达。
 ✓ GoogleNet是谷歌研究人员设计的一类架构。谷歌是2014年ImageNet的赢家，在那里它被证明是一个强大的模型。





8.5.2 语义特征的表达

- 高层的语义特征多采用深度学习模型表达。
 - ✓ ResNet是真正定义了深度学习体系结构的深度架构之一。残差网络(简称ResNet)由多个后续残差模块组成，这些模块是ResNet体系结构的基本构件。

