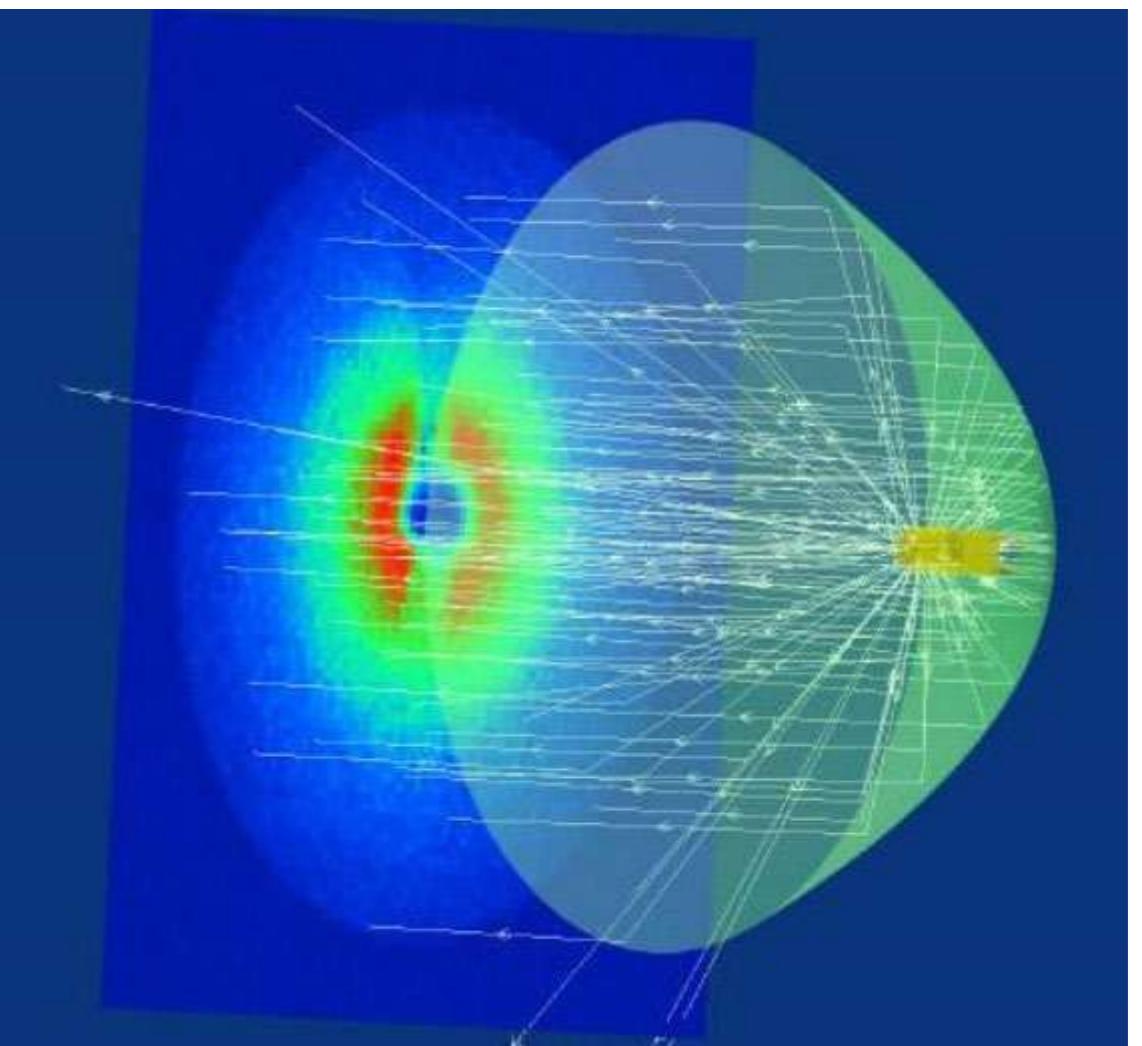
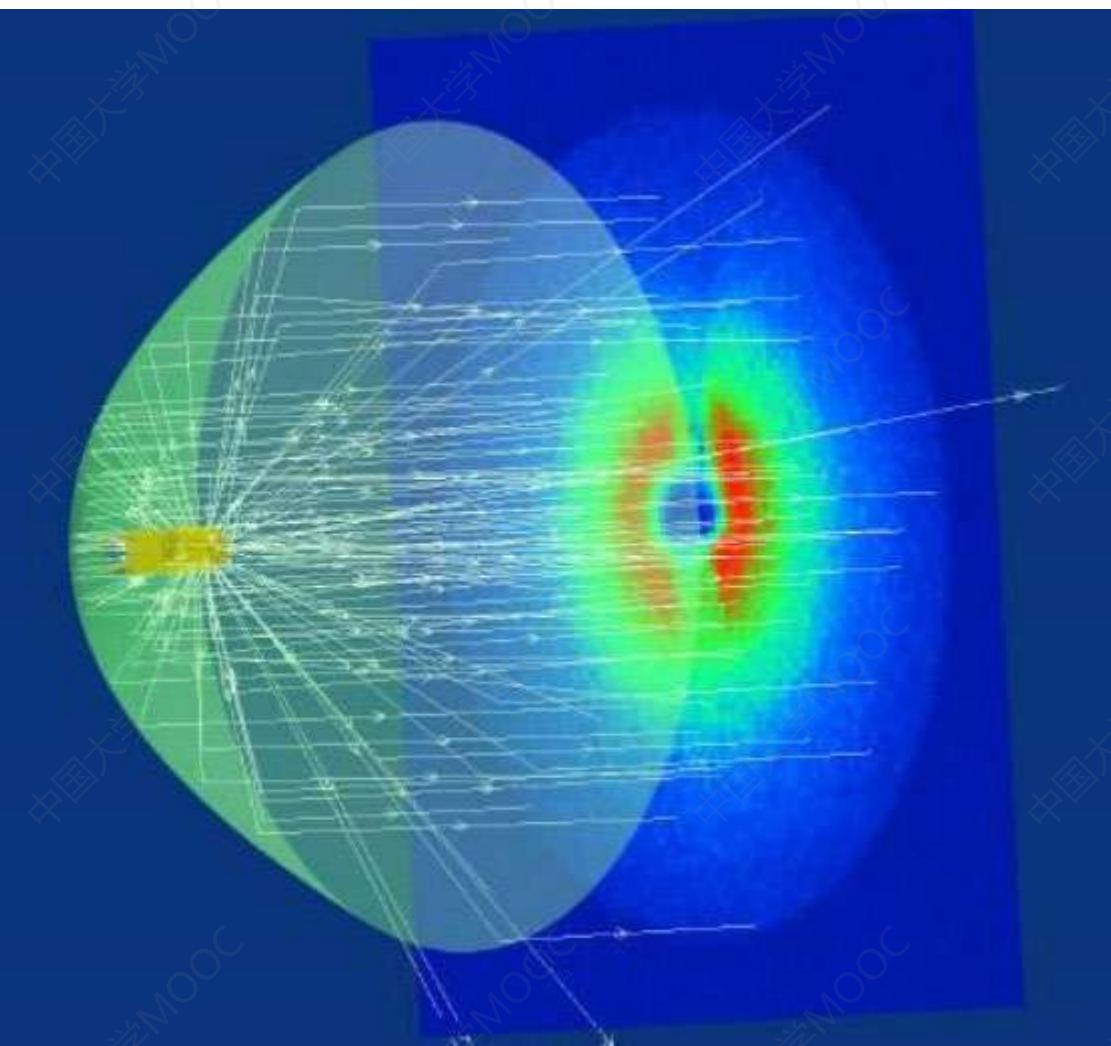


第一章 绪论



1.0 图像处理技术的重要性



人类感觉器官

- 视觉是人类最重要的感官，人类75%感知来自视觉；
- 视觉系统是一个典型的图像处理系统；
- 图像处理技术与我们的生活密切相关——智能手机、监控视频、互联网视频、人脸识别、门禁、车牌识别等等

第一章 绪论

- 1.1 什么是图像，数字图像
- 1.2 图像处理方式分类
- 1.3 数字图像处理起源和发展
- 1.4 数字图像处理核心技术

第一章 绪论

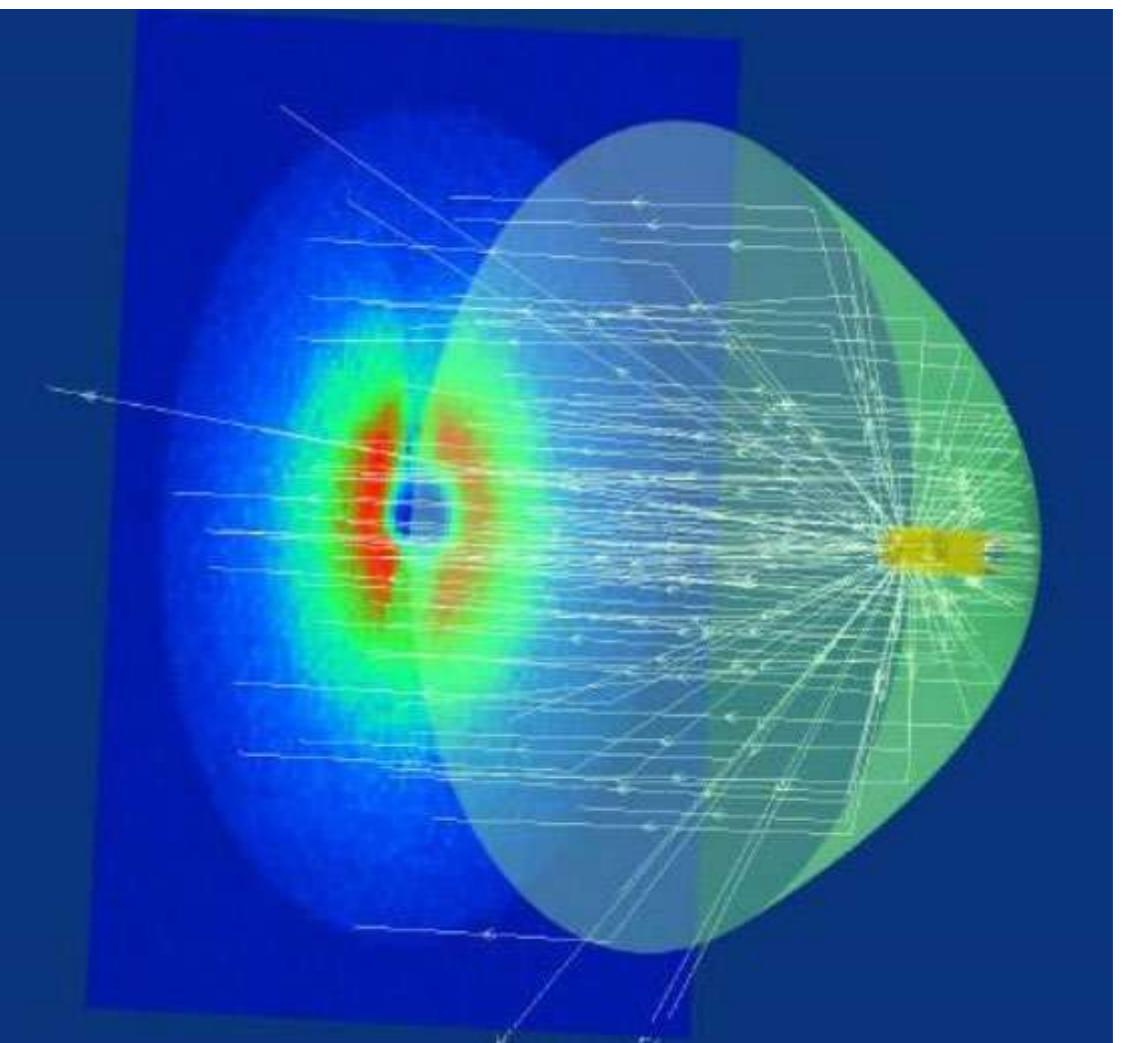
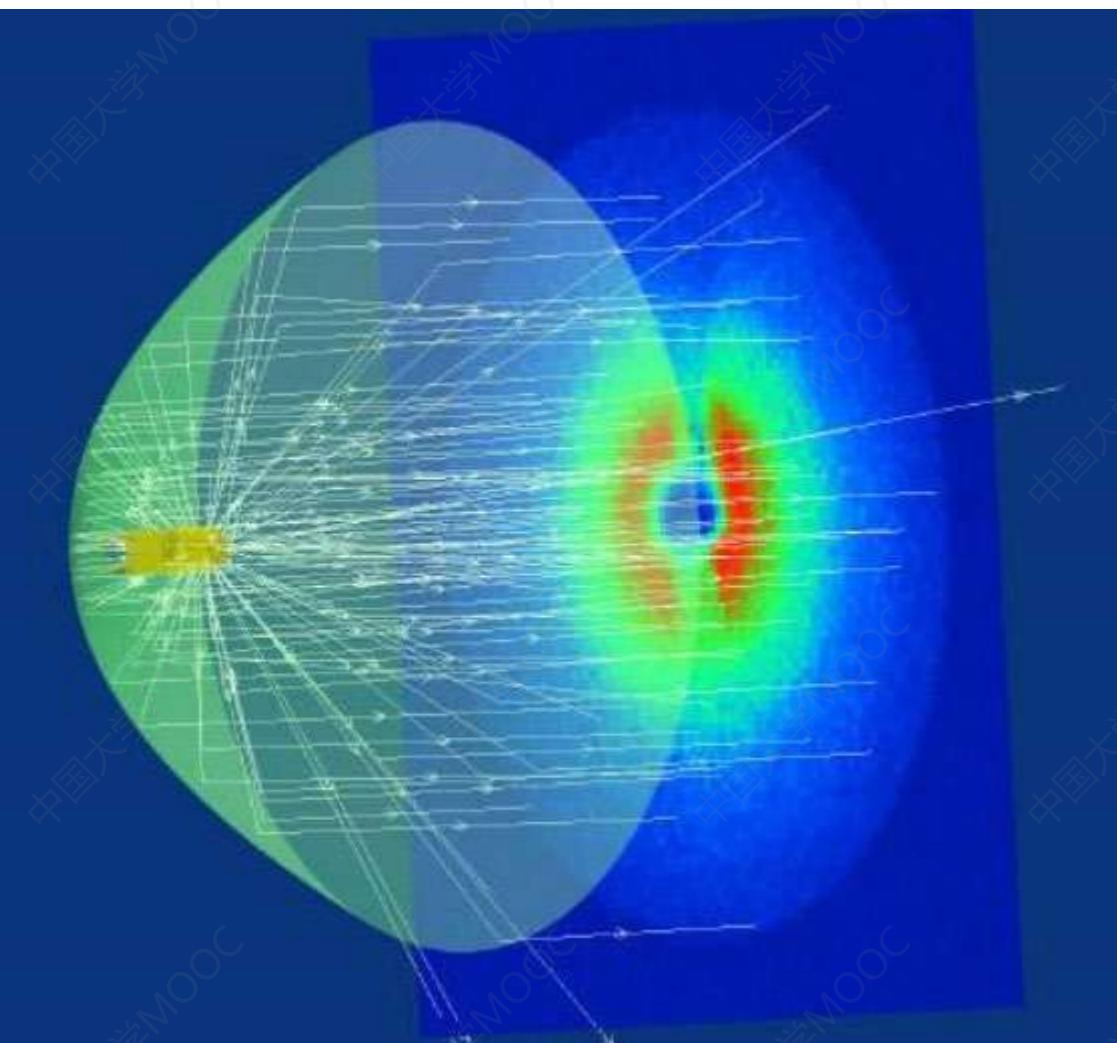
1.1 什么是图像，数字图像

1.2 图像处理方式分类

1.3 数字图像处理起源和发展

1.4 数字图像处理核心技术

1.1 什么是图像，数字图像

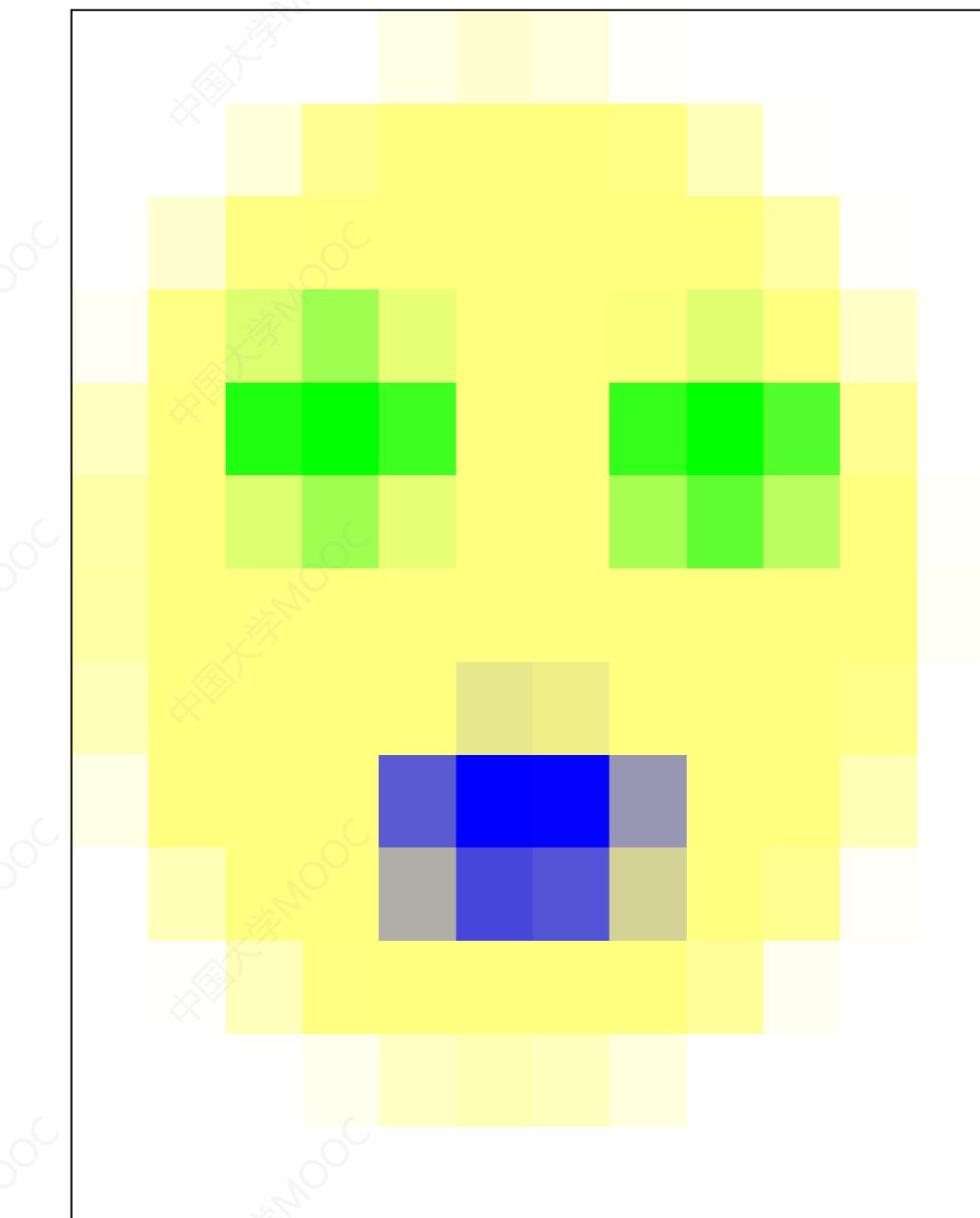
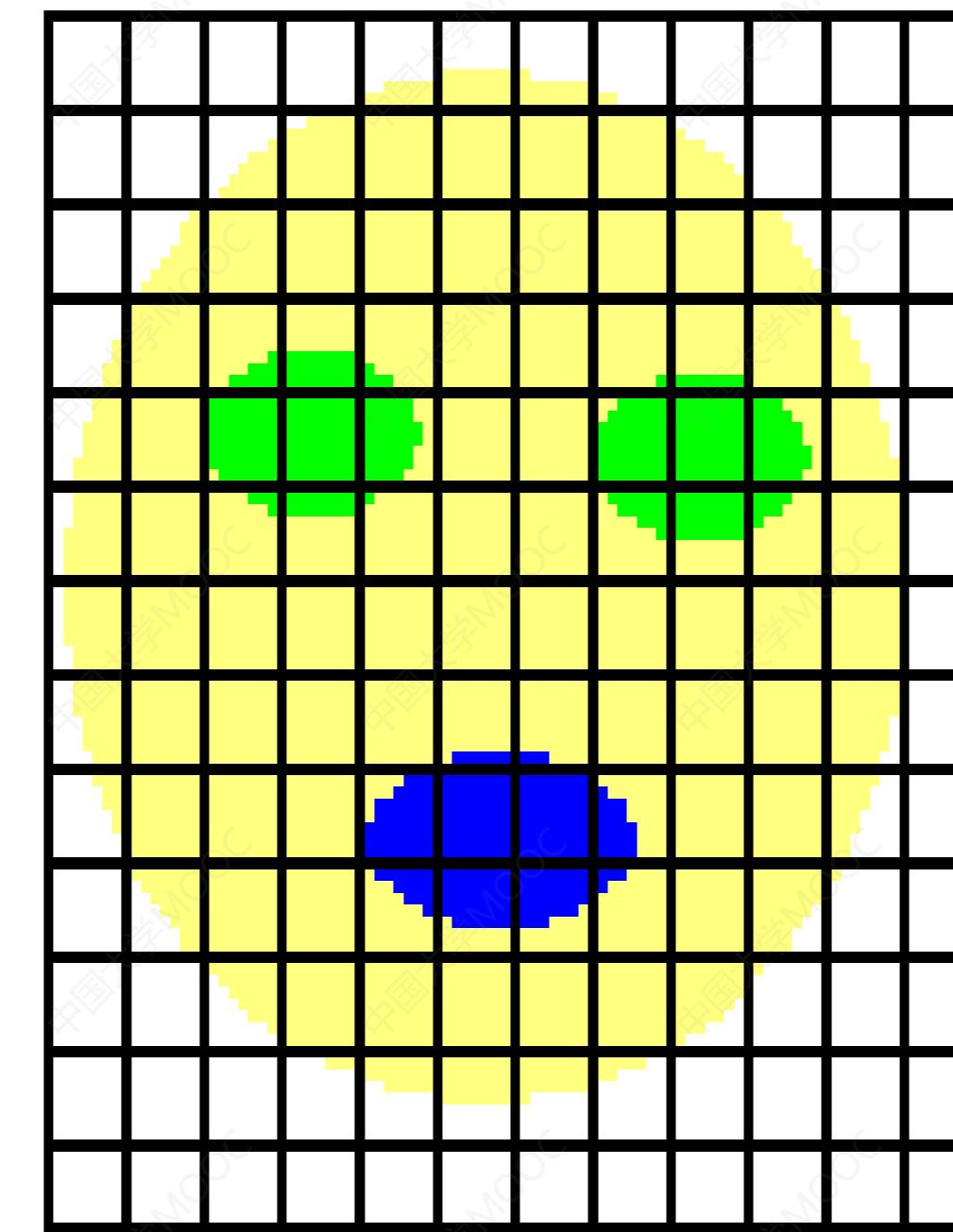
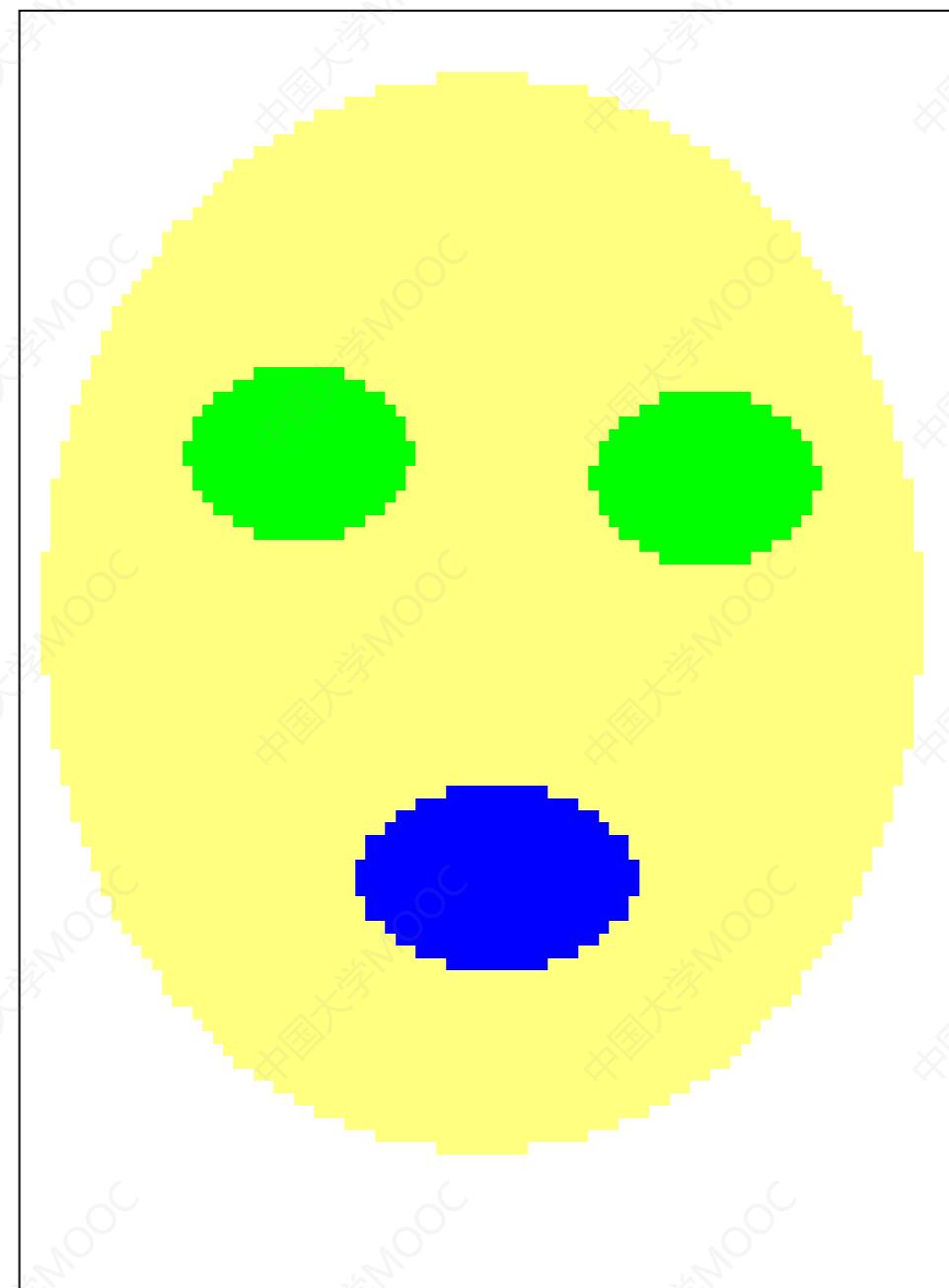


■ 1.1 什么是图像，数字图像

- 《数字图像处理学》，阮秋琦，电子工业出版社
 - 图像 (image) : 以某一技术手段被再现于二维画面上的视觉信息
 - 活动的、立体的、彩色的图像表示为 $f(x, y, z, \lambda, t)$
 - 静止的、平面的、彩色的图像表示为 $f(x, y, \lambda)$
 - 静止的、平面的、单色的图像表示为 $f(x, y)$
- 《Digital Image Processing》(Second Edition) Rafael C. Gonzalez and Richards E. Woods, 电子工业出版社
 - An Image may be defined as a two-dimentional function $f(x,y)$, where, x and y are spatial coordinates and the amplitude of f at any pair of coordinates (x,y) is called the intensity or gray level of the image at that point.

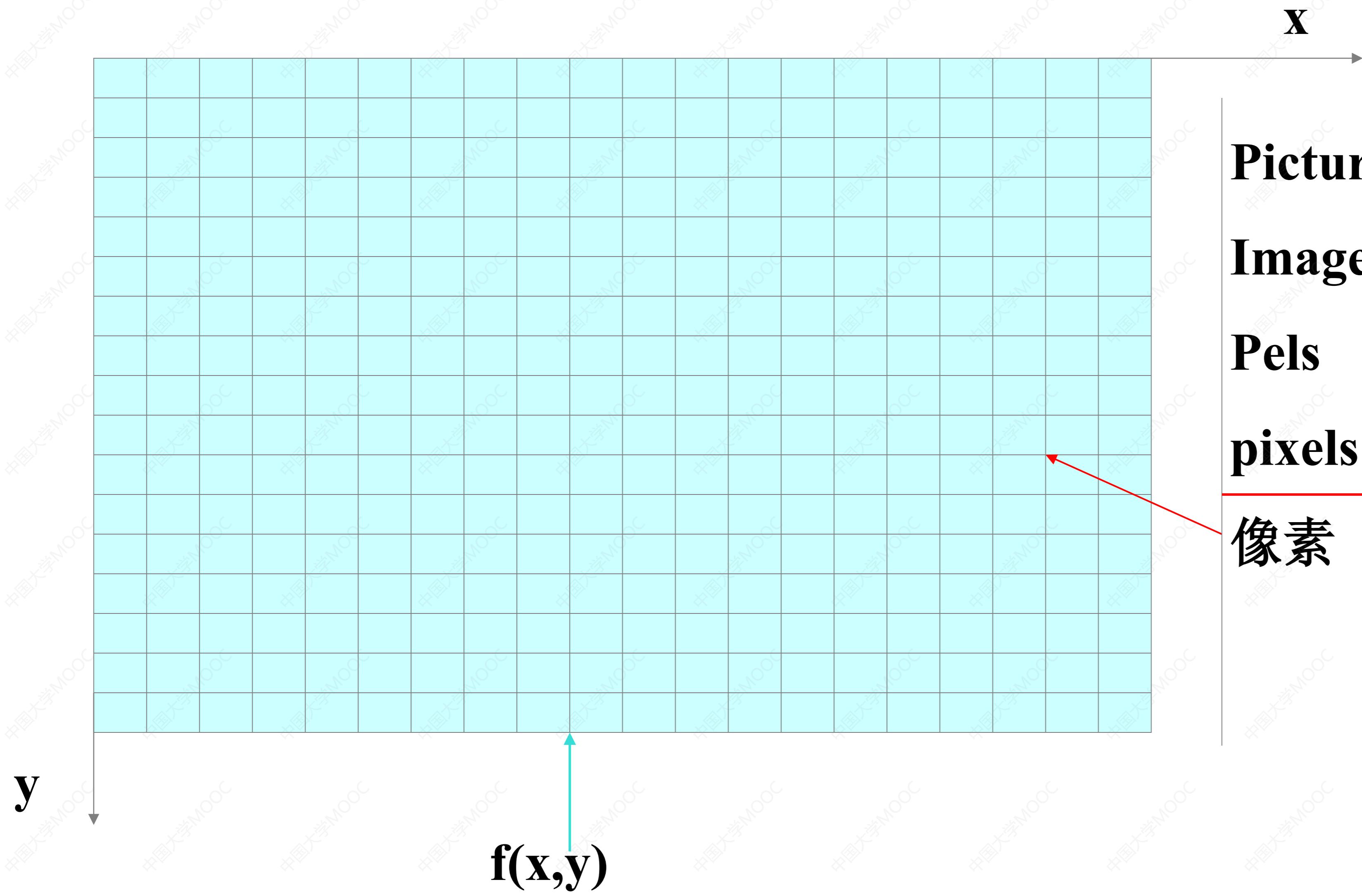
■ 1.1 什么是图像，数字图像

- 数字图像：空间域离散、幅度上量化的图像，将模拟图像经过空间域采样，幅度上量化，即可得到数字图像。





1.1 什么是图像，数字图像



Picture elements

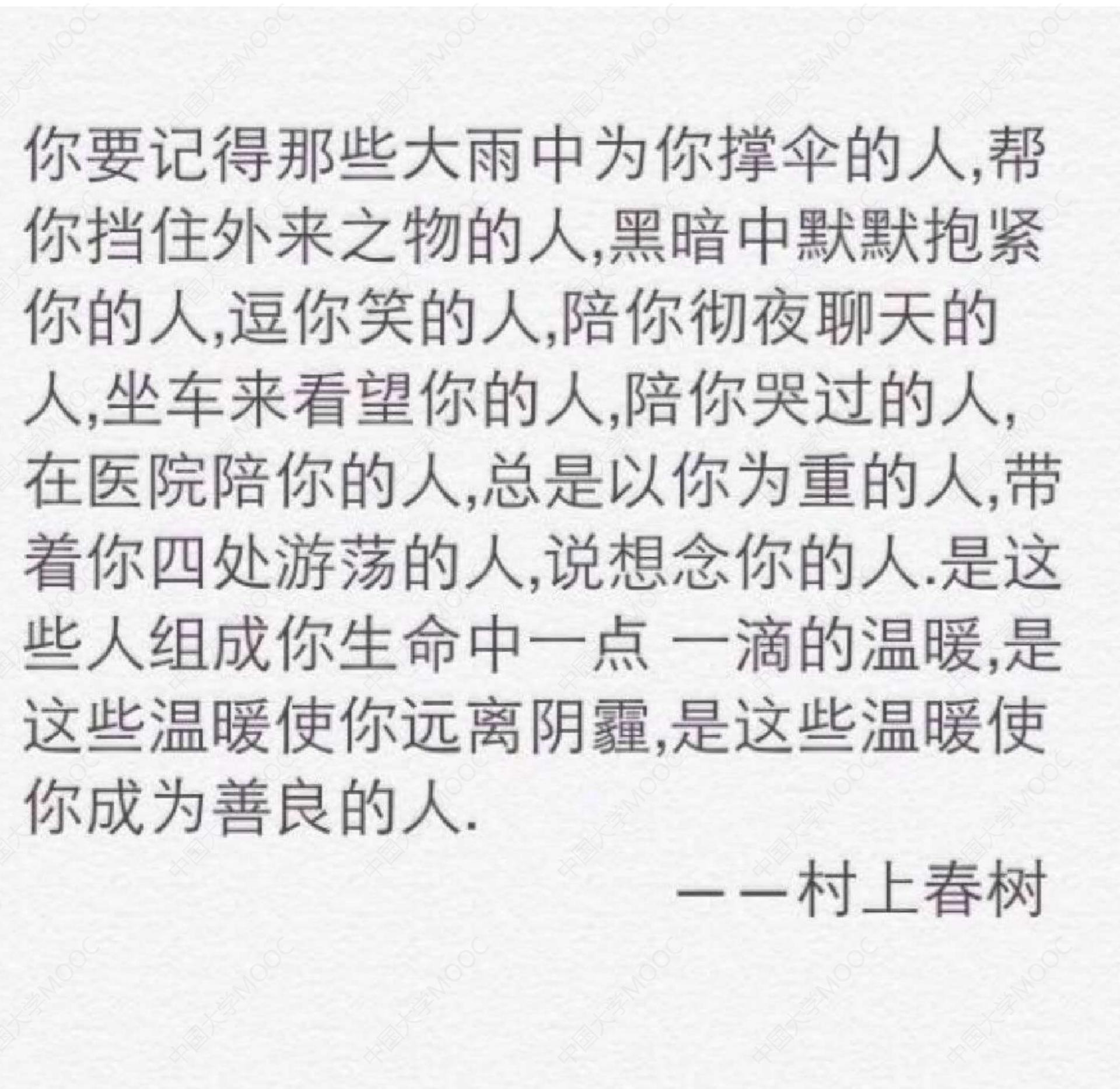
Image elements

Pels

pixels

像素

1.1 什么是图像，数字图像



二值图像



黑白图像

1.1 什么是图像，数字图像



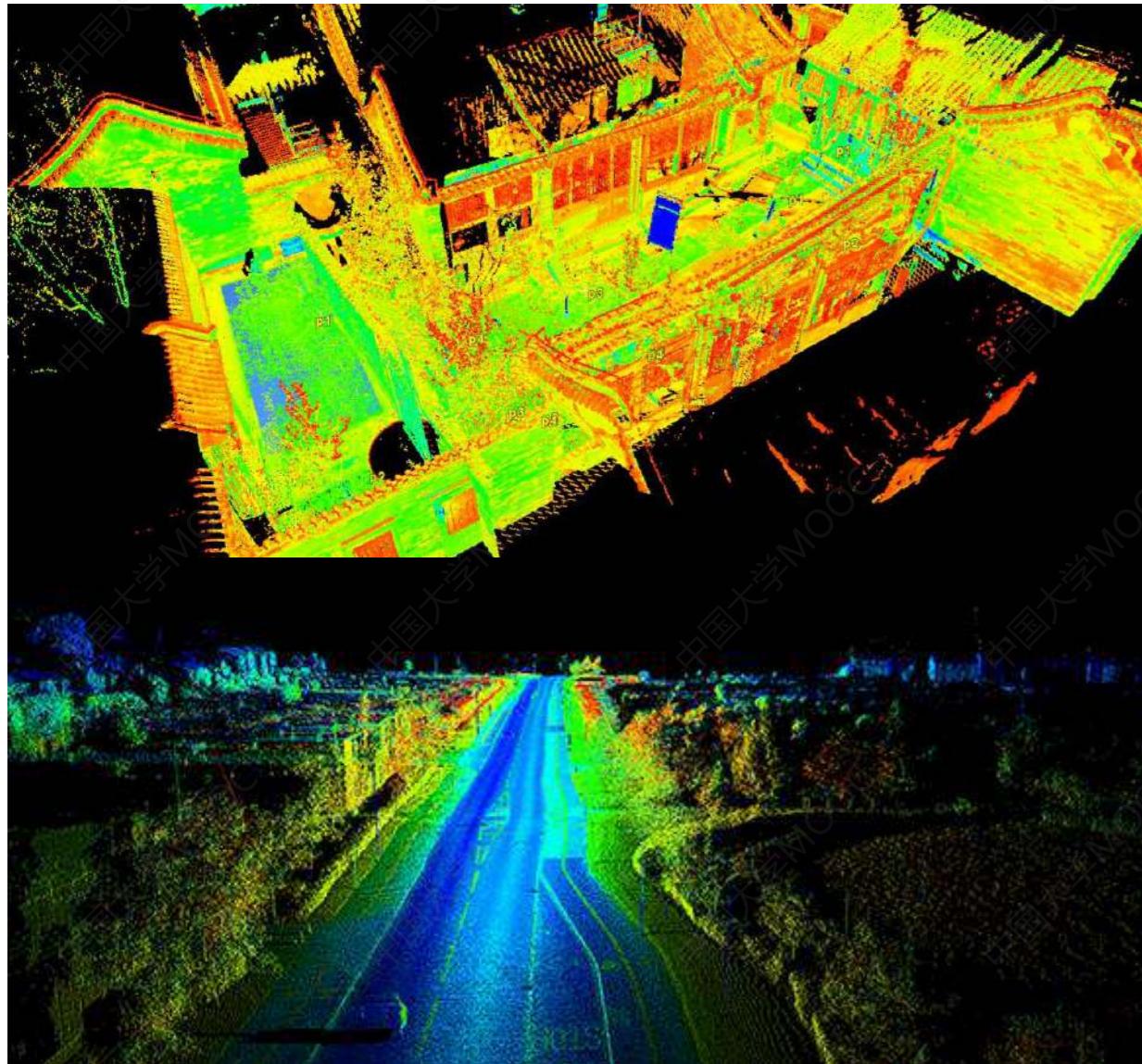
灰度图像



彩色图像



1.1 什么是图像，数字图像

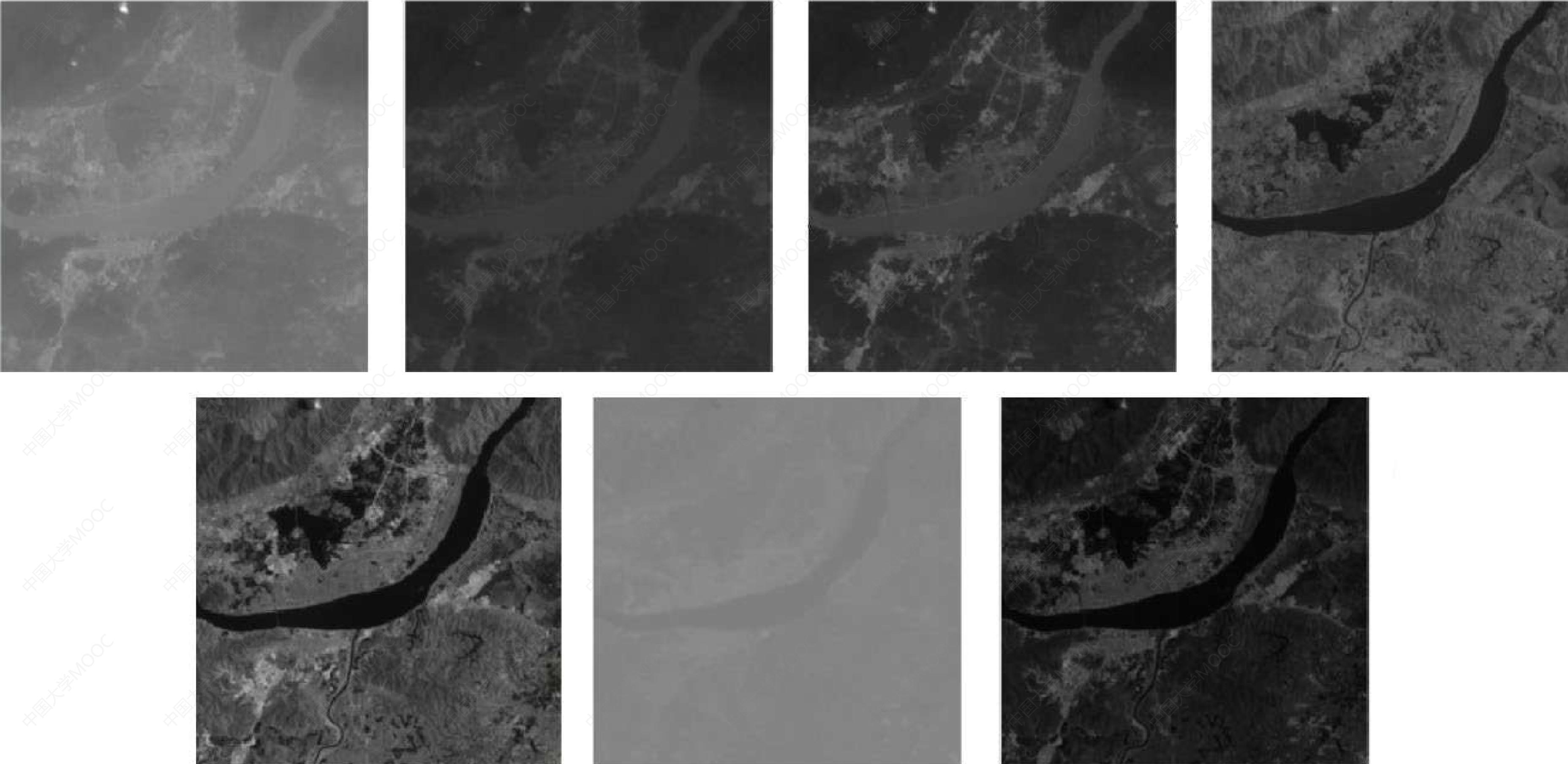


点云图像



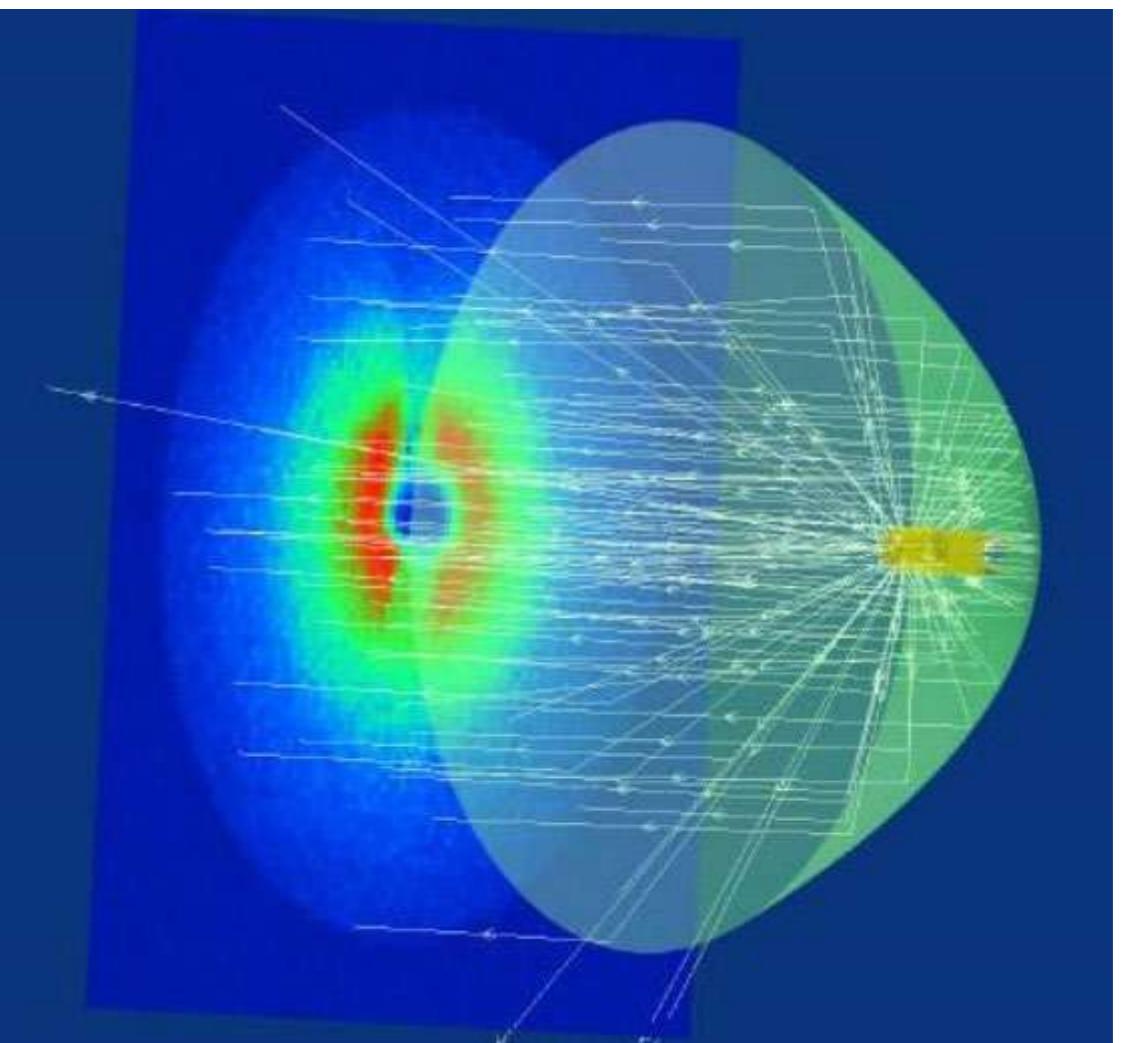
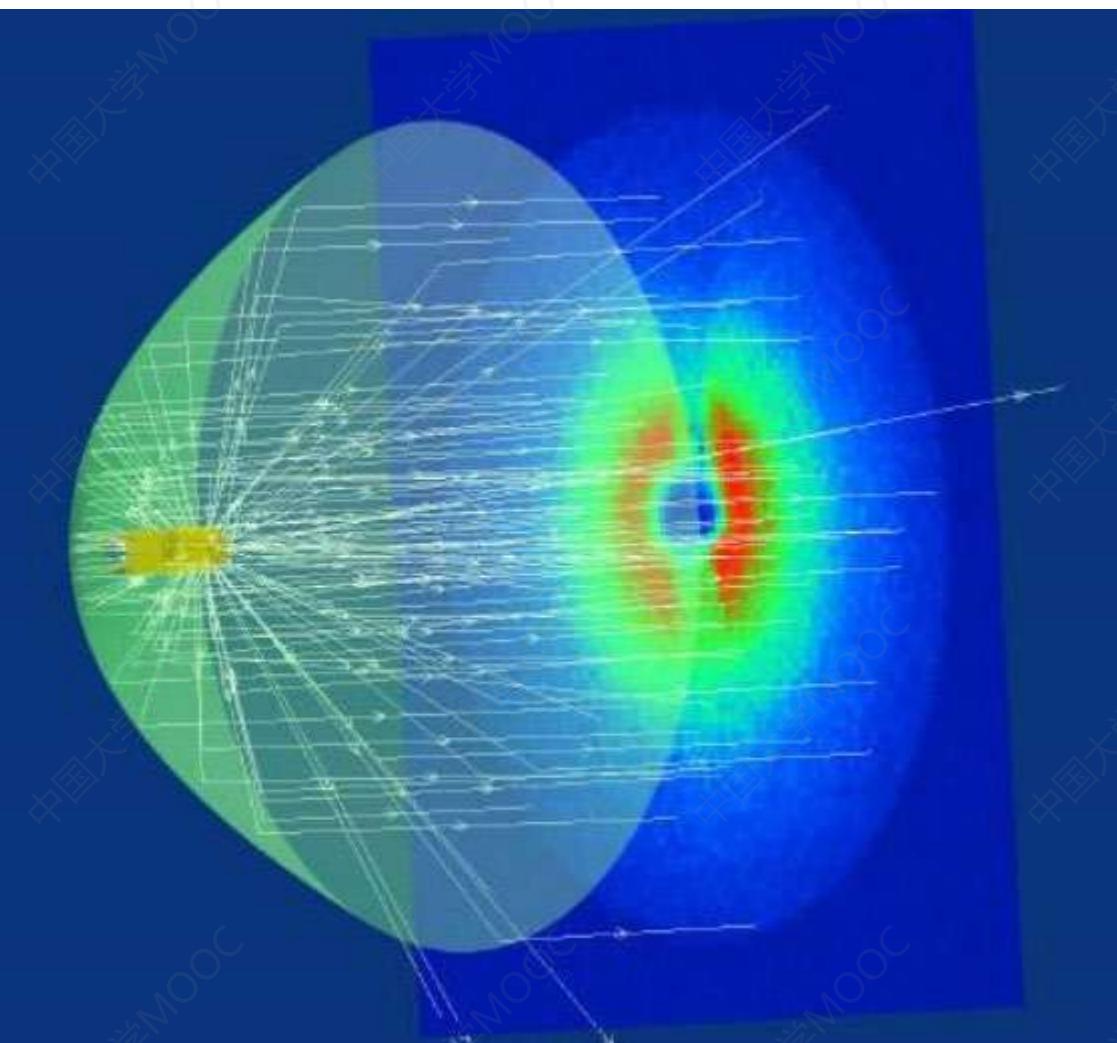
彩色图像+深度图像

■ 1.1 什么是图像，数字图像



多光谱图像

1.2 图像处理方式分类



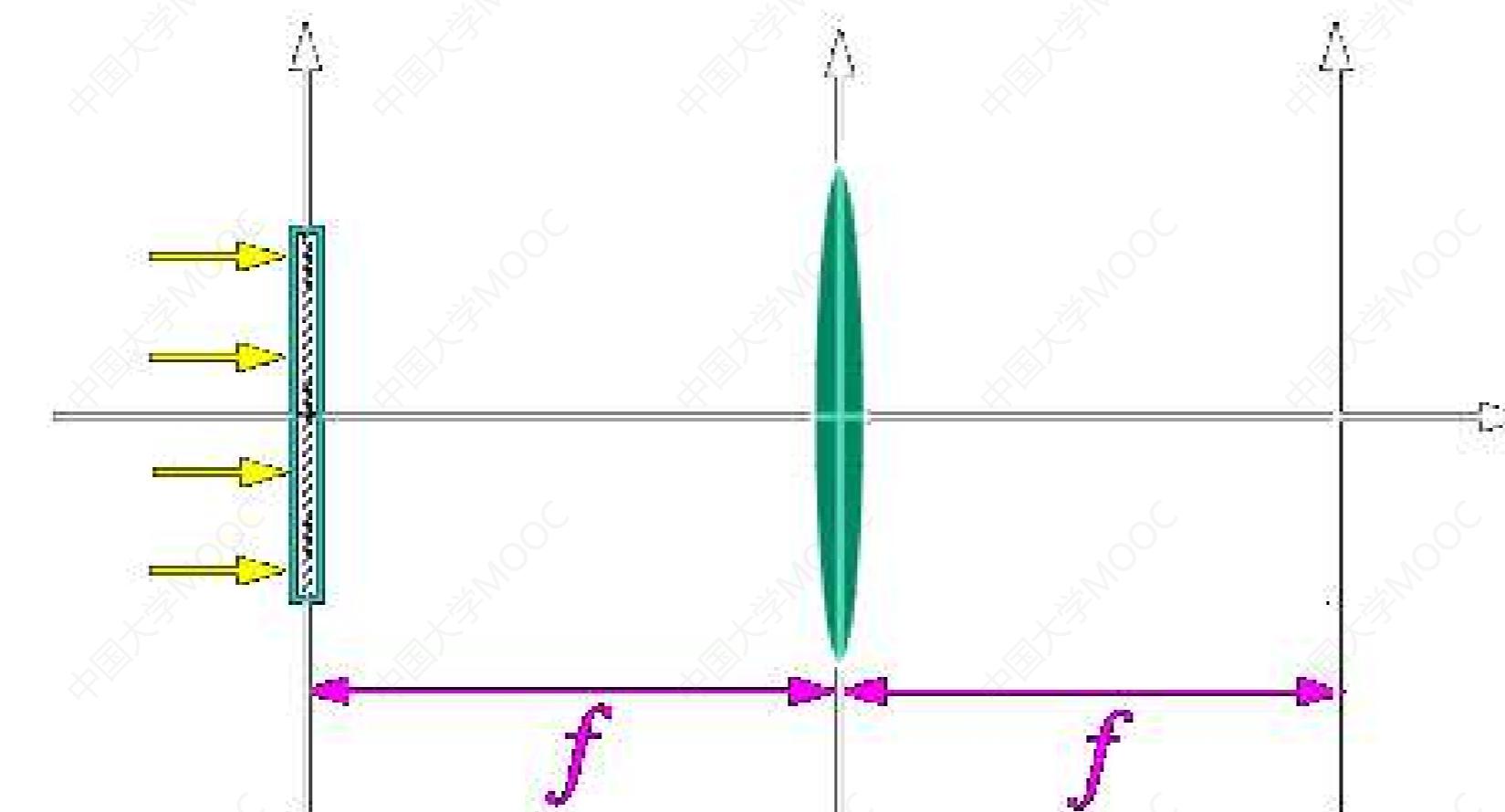
■ 1.2 图像处理方式分类

- 光学图象处理
- 模拟电信号图象处理
- 数字图像处理

■光学图象处理

- 定义

- 利用光学器件直接对光学图像进行处理。



两个常见的光学现象

- 1) 入射光为在焦点上的一个点光源，通过透镜后变为平行光线；
- 2) 入射光为平行均匀光，通过透镜后汇聚到为焦平面上的一点；

■ 1.2.1 光学图象处理

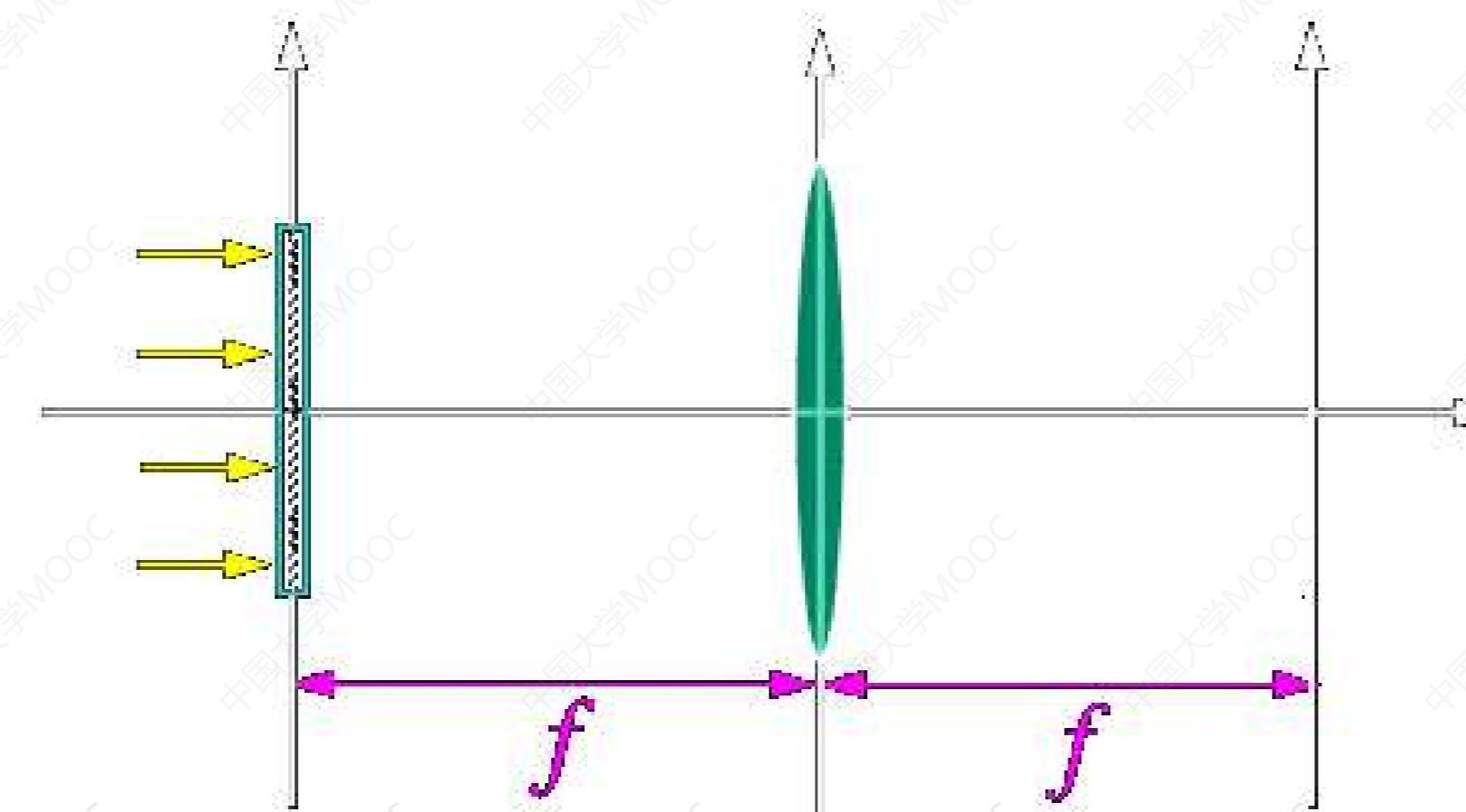
- 两个常见的光学现象

- 入射光为在焦点上的一个点光源，通过透镜后变为平行光线；

$\delta(x,y)$ (输入) ——> 常量 (输出)

- 入射光为平行均匀光，通过透镜后汇聚到为焦平面上的一点；

常量 (输入) ——> $\delta(x,y)$ (输出)

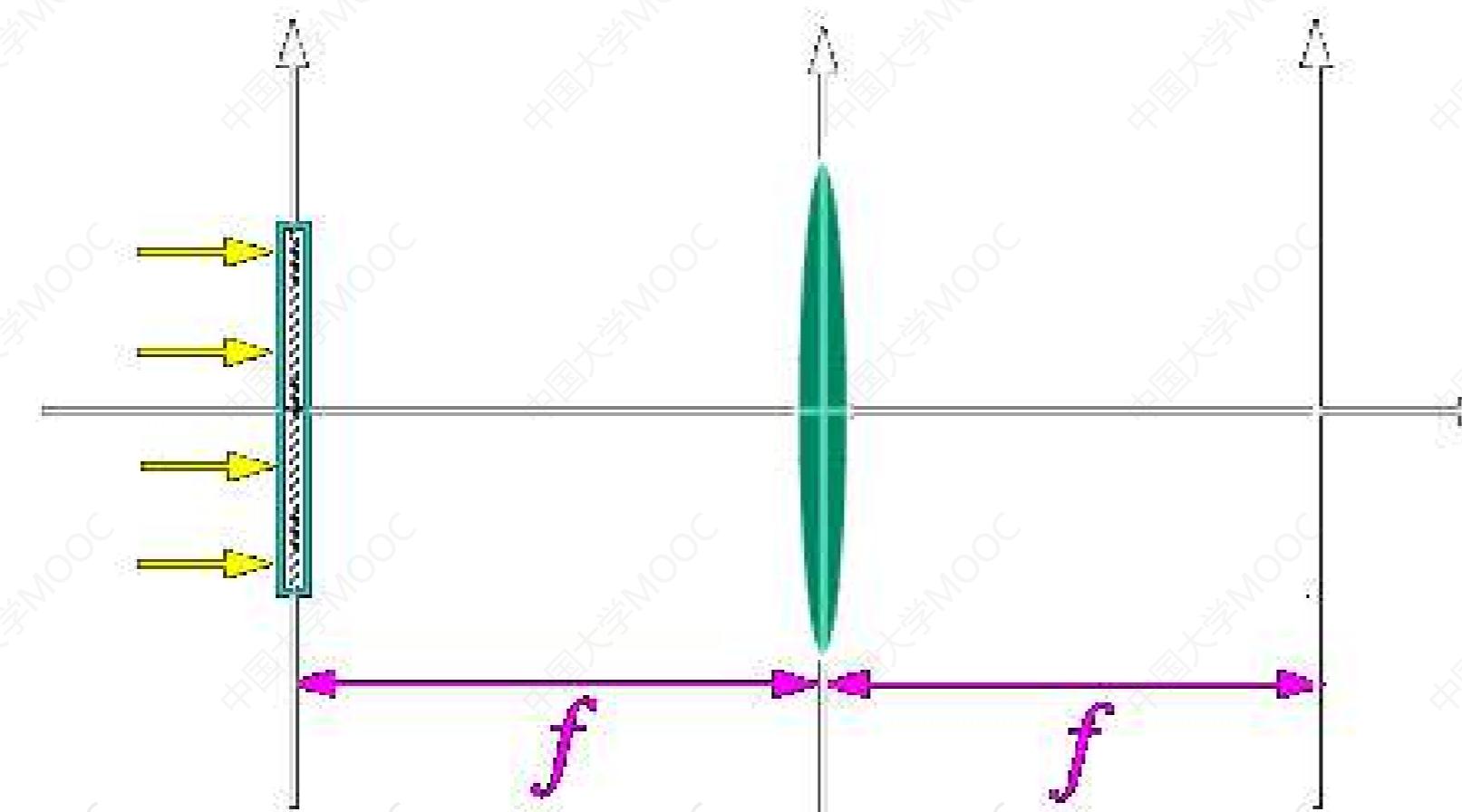


前焦平面：图像平面（物面）
后焦平面：频谱面

■ 1.2.1 光学图象处理

透镜的傅立叶变换性质

物体放在焦距为 f 的透镜的前焦面，用波长为 λ 的单色平面波垂直入射照明，在透镜后焦平面上得到物体的准确的傅立叶变换。

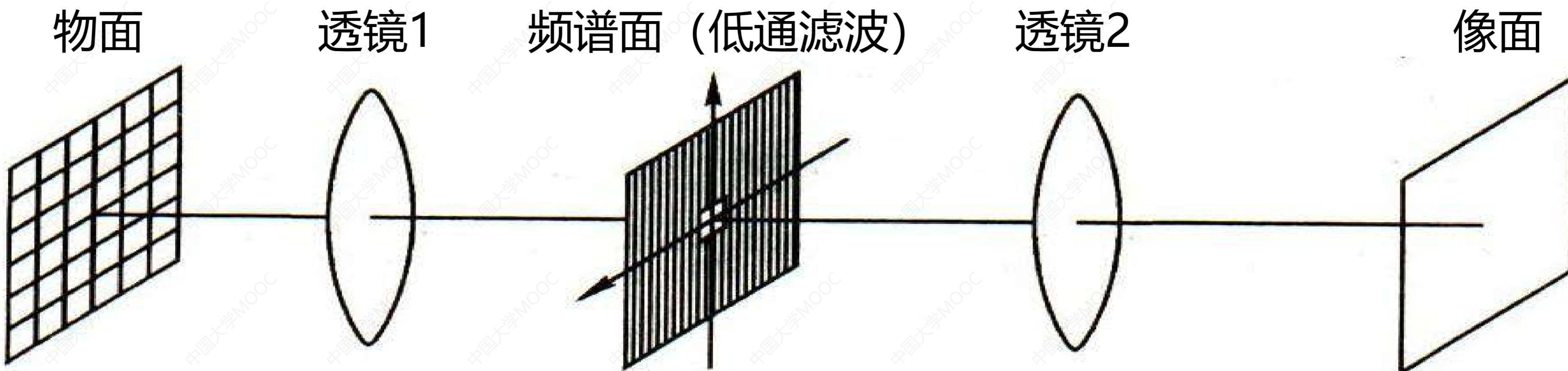


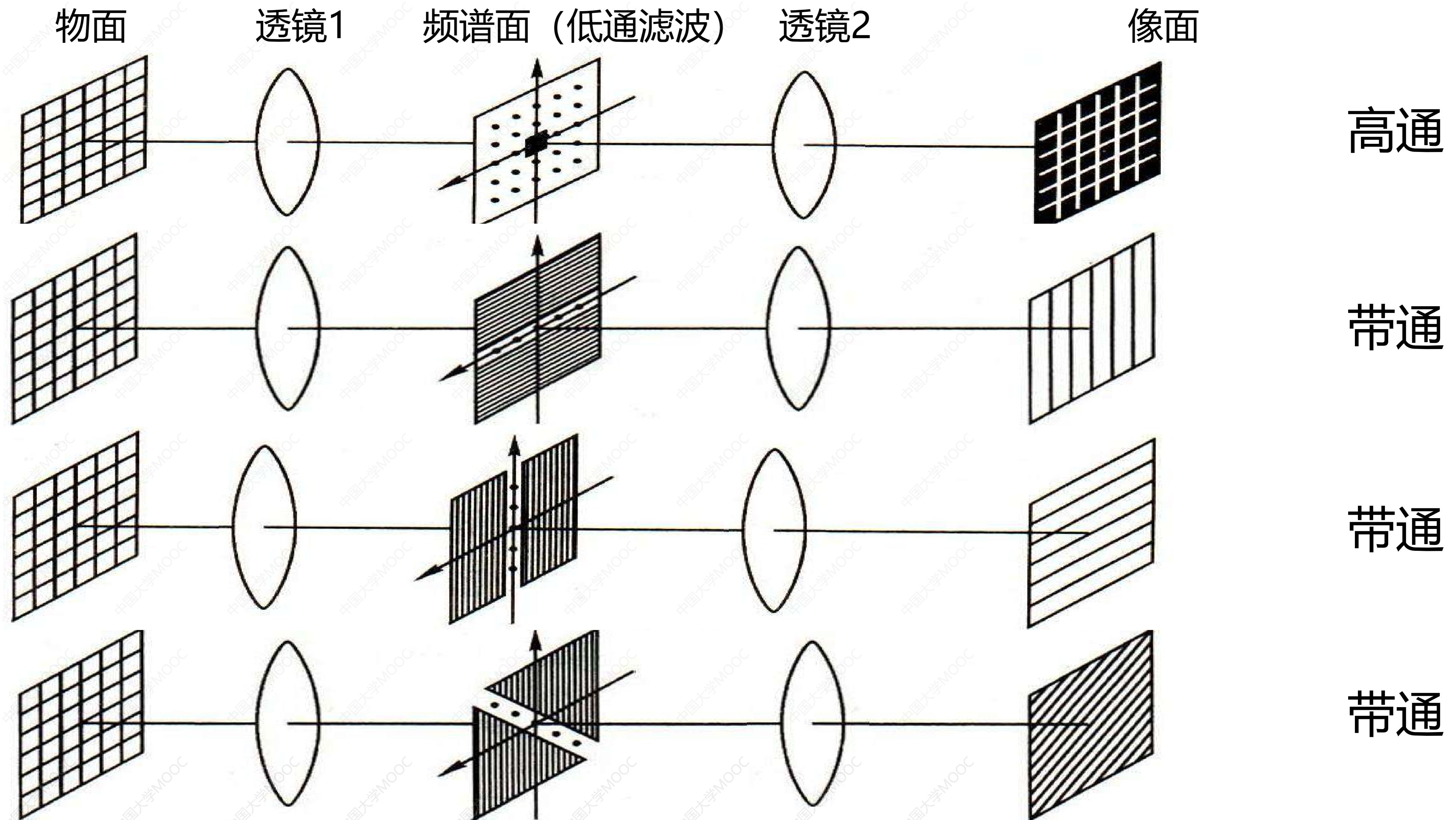
前焦平面：图像平面（物面）
后焦平面：频谱面

1.2.1 光学图象处理

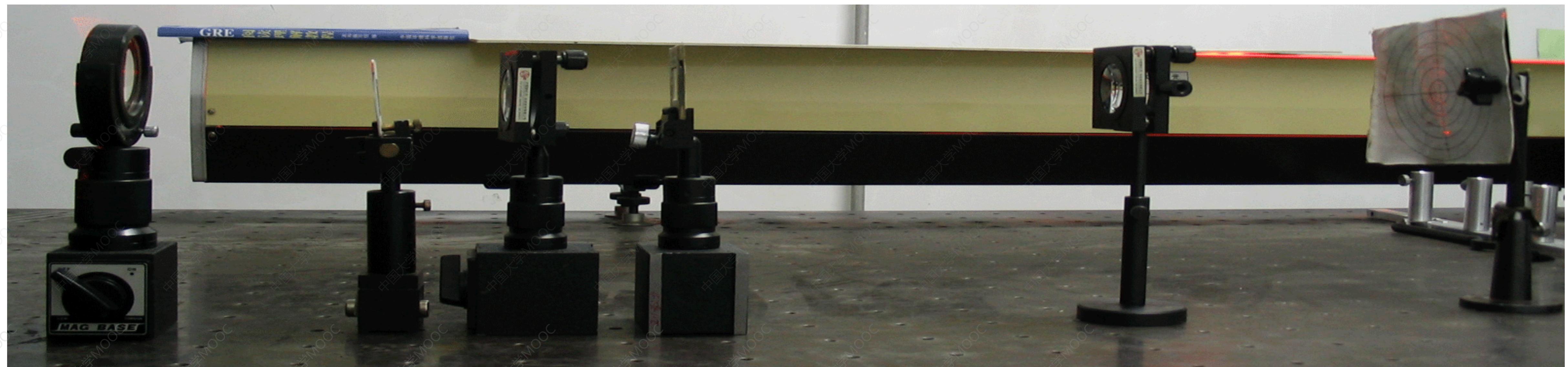
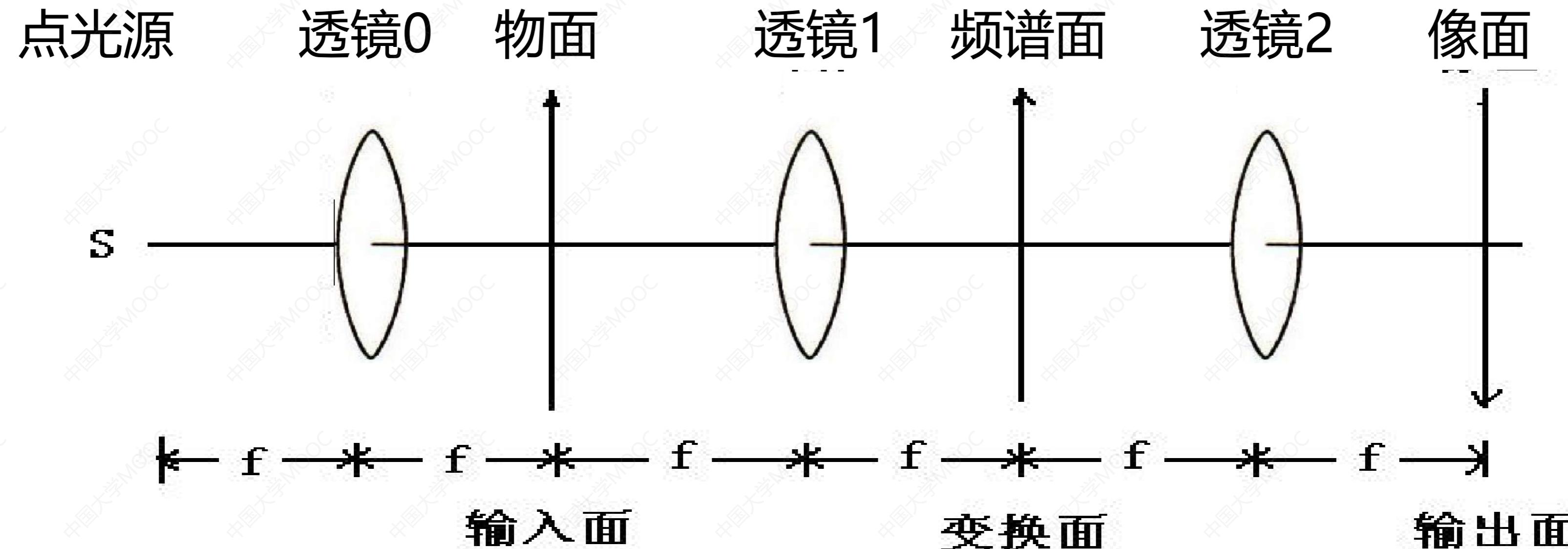


(a) 低通滤波器结构





1.2.1 光学图象处理



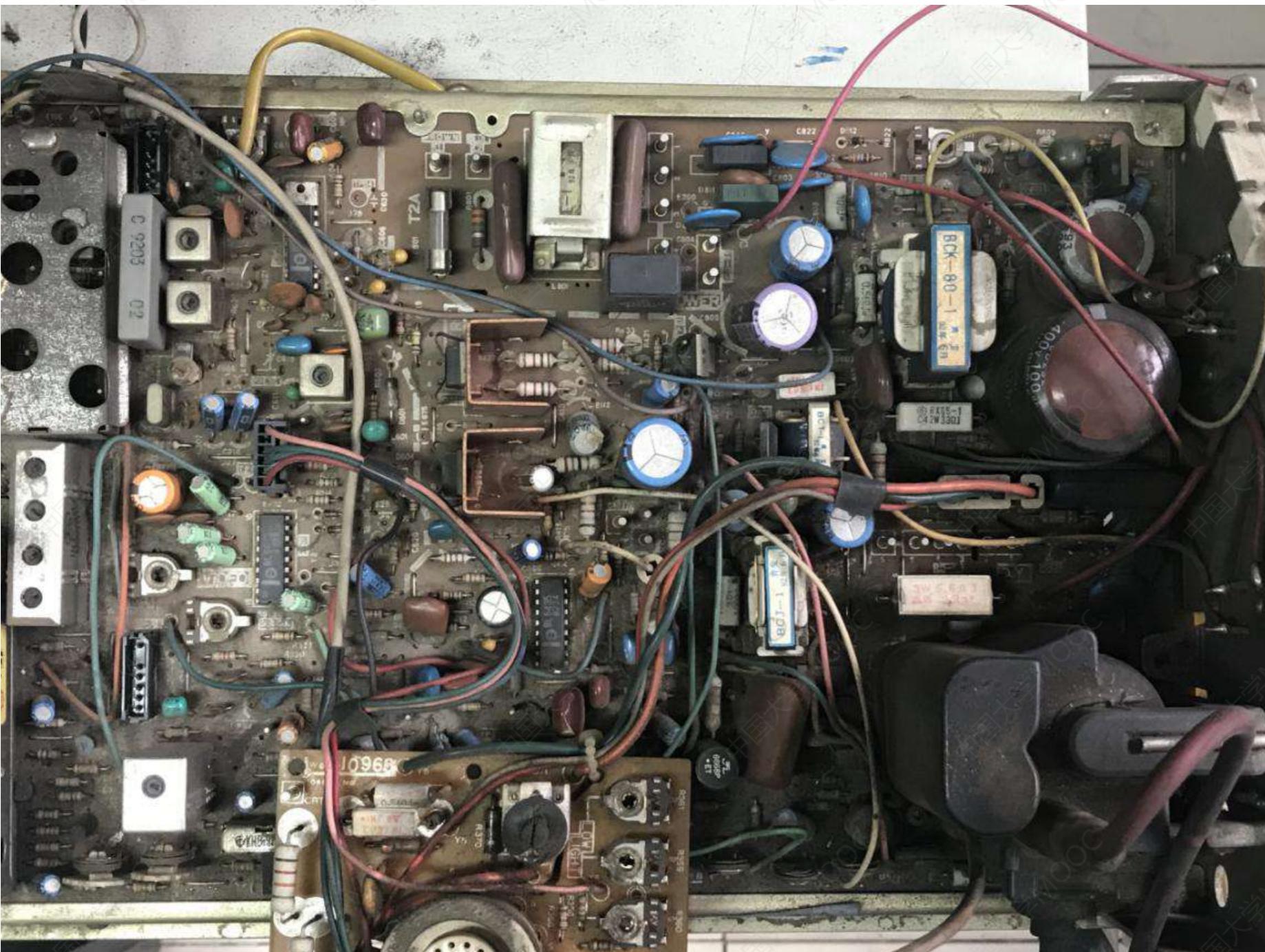
■ 光学图象处理

- 并行处理；
- 只能对图像进行**频谱分析、滤波、图像加减运算、相关识别**等基本处理，处理方式有限；
- 更换处理方法，需要更换相应的光学器件或光路；

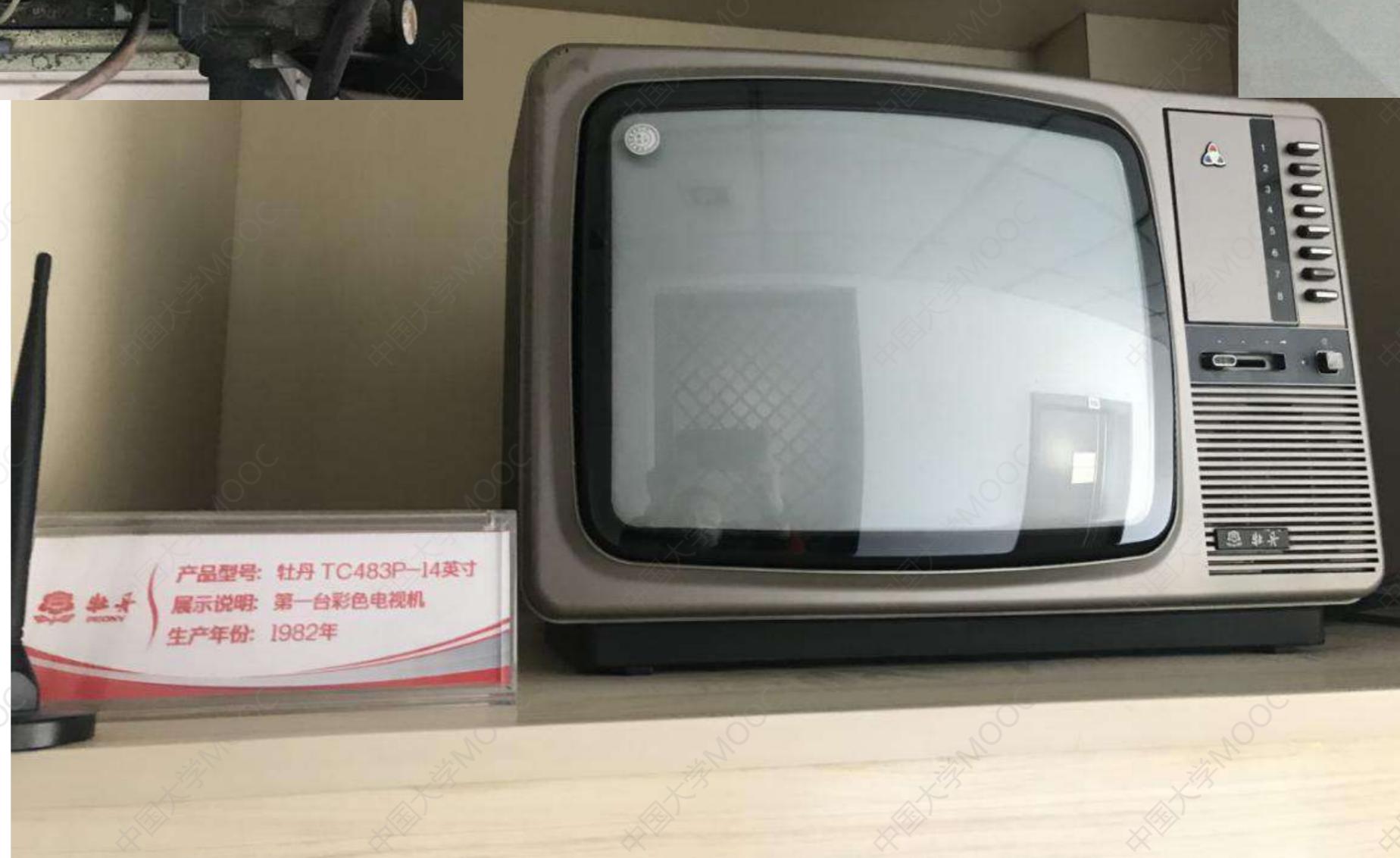
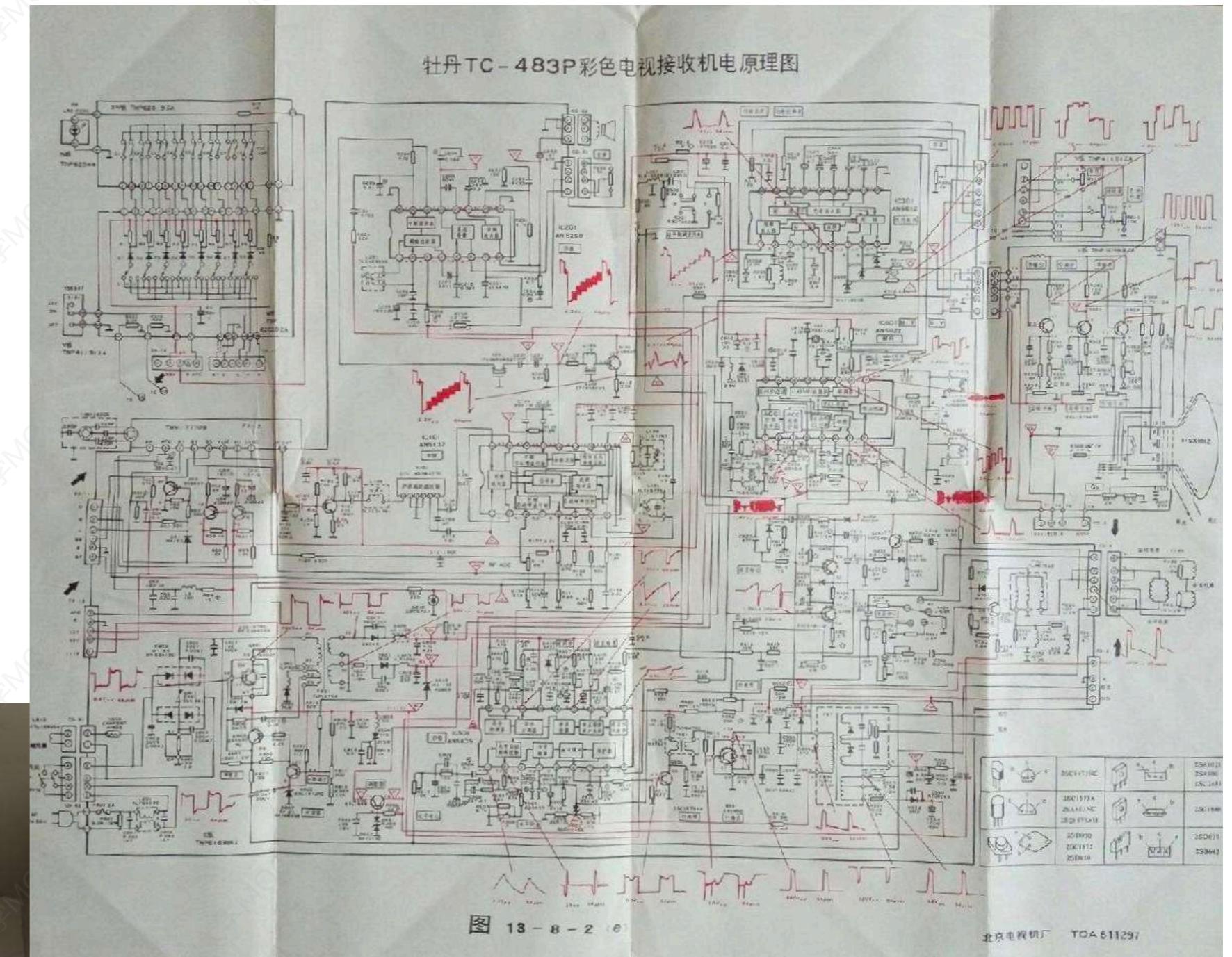
■ 1.2.2 模拟电信号图象处理

- 定义
 - 将光学图像信号转变成模拟电信号，用模拟电路进行处理。
- 代表性应用：**模拟电视。**
- 对图像进行处理，图像信号的**采集、调制、解调、重现等**，处理复杂程度远超过光学图像处理；
- 更换处理方法，需要更换相应的**电路或者元器件**；
- 处理精度依赖于基本元器件电容、电感、电阻和电路结构；

■ 1.2.2 模拟电信号图象处理



牡丹14英寸
模拟电视机



模拟电视机
电路原理图

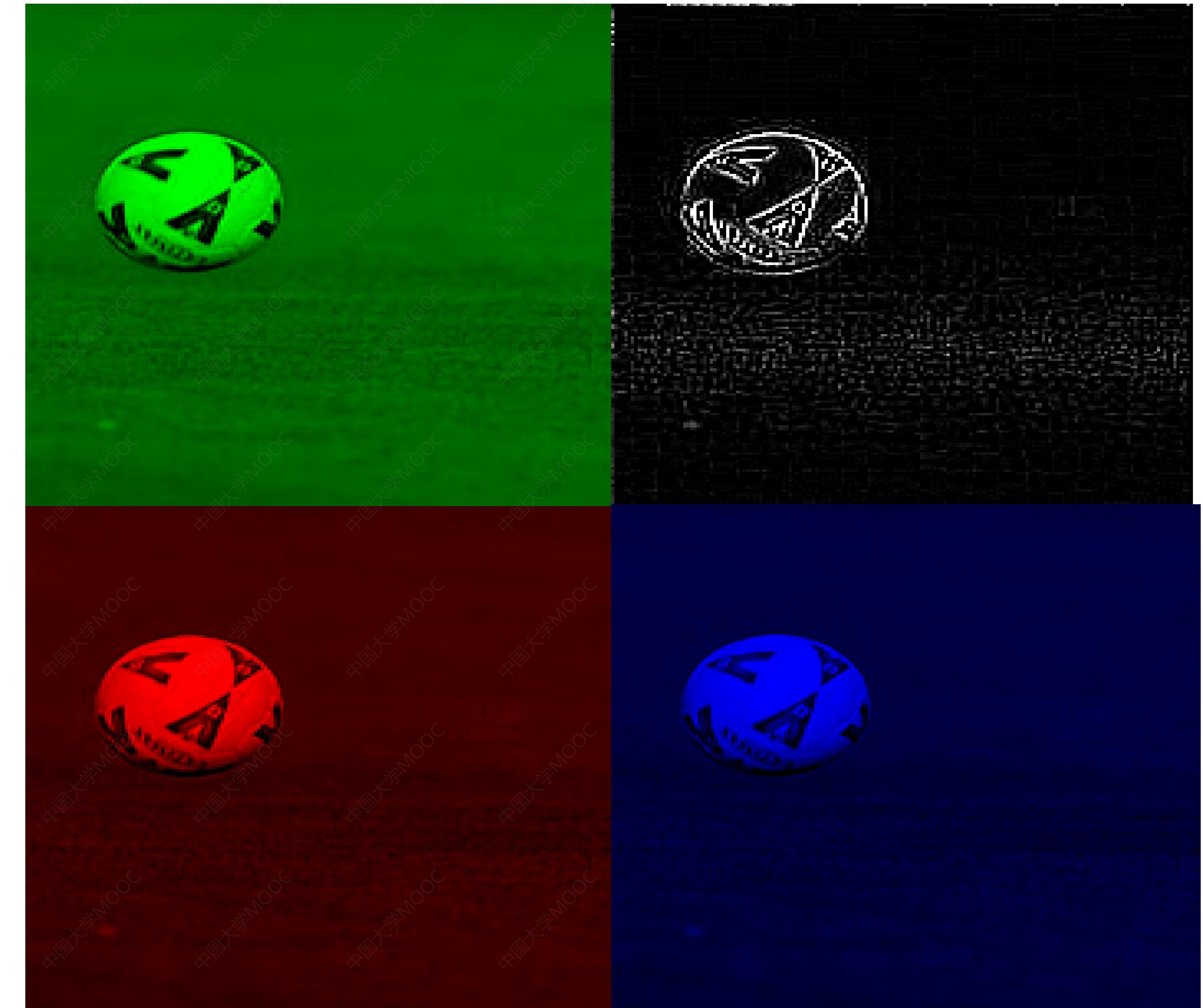
■ 1.2.3 数字图像处理

- 定义：
 - 依托**数值计算设备**如计算机，以**数值计算方式**对**数字化的图像**进行处理。
- 对图像进行处理，只需要编写软件实现该处理方法对应的数学模型（公式）即可；**方便**
- 更换处理方法，只需要更新处理算法对应的代码即可；**灵活**
- 处理精度依赖于计算机的字长；**精度高**
- 数字图像处理是数字计算机发展的产物，是计算机应用重要方向之一；

1.2.3 图像处理的三个层面——狭义图像处理

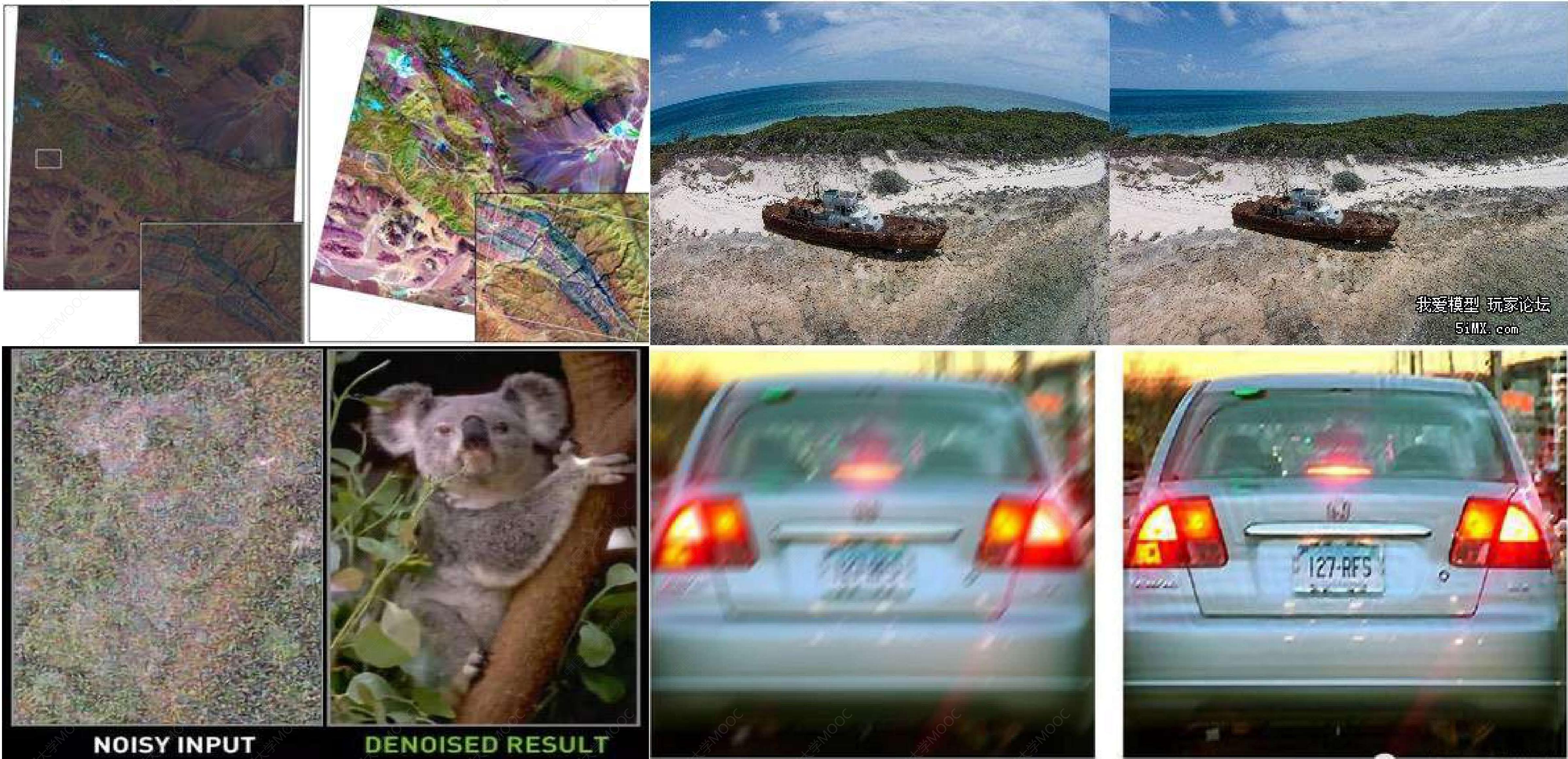


输入图像



输出图像

1.2.3 图像处理的三个层面——狭义图像处理



1.2.3 图像处理的三个层面——图像分析与理解

➤ 输出特征参数——图像分析

➤ 背景颜色：



- 颜色范围： $90 < H < 100$,
- 纹理特征：

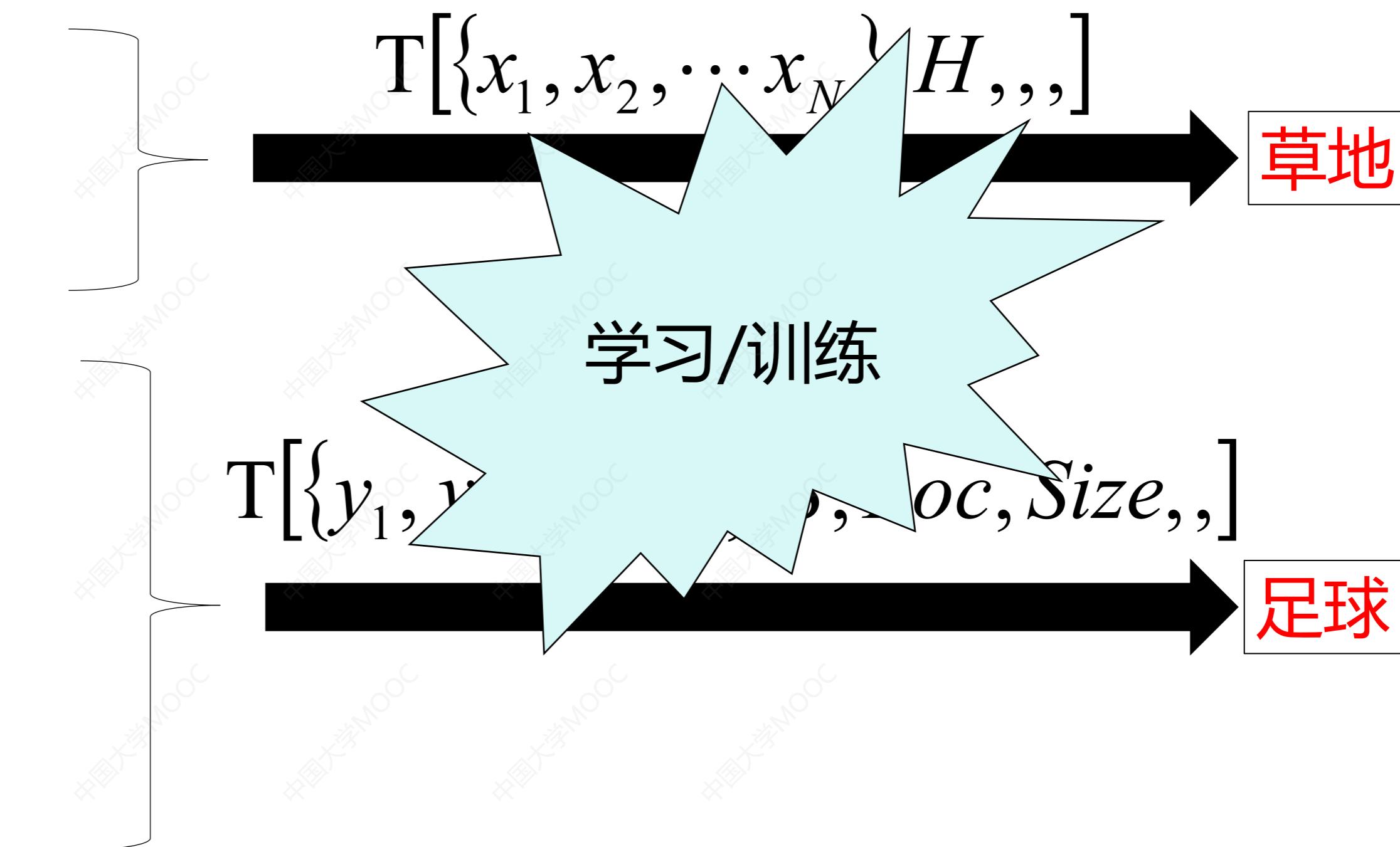
$$\{x_1, x_2, \dots, x_N\}$$

➤ 圆：

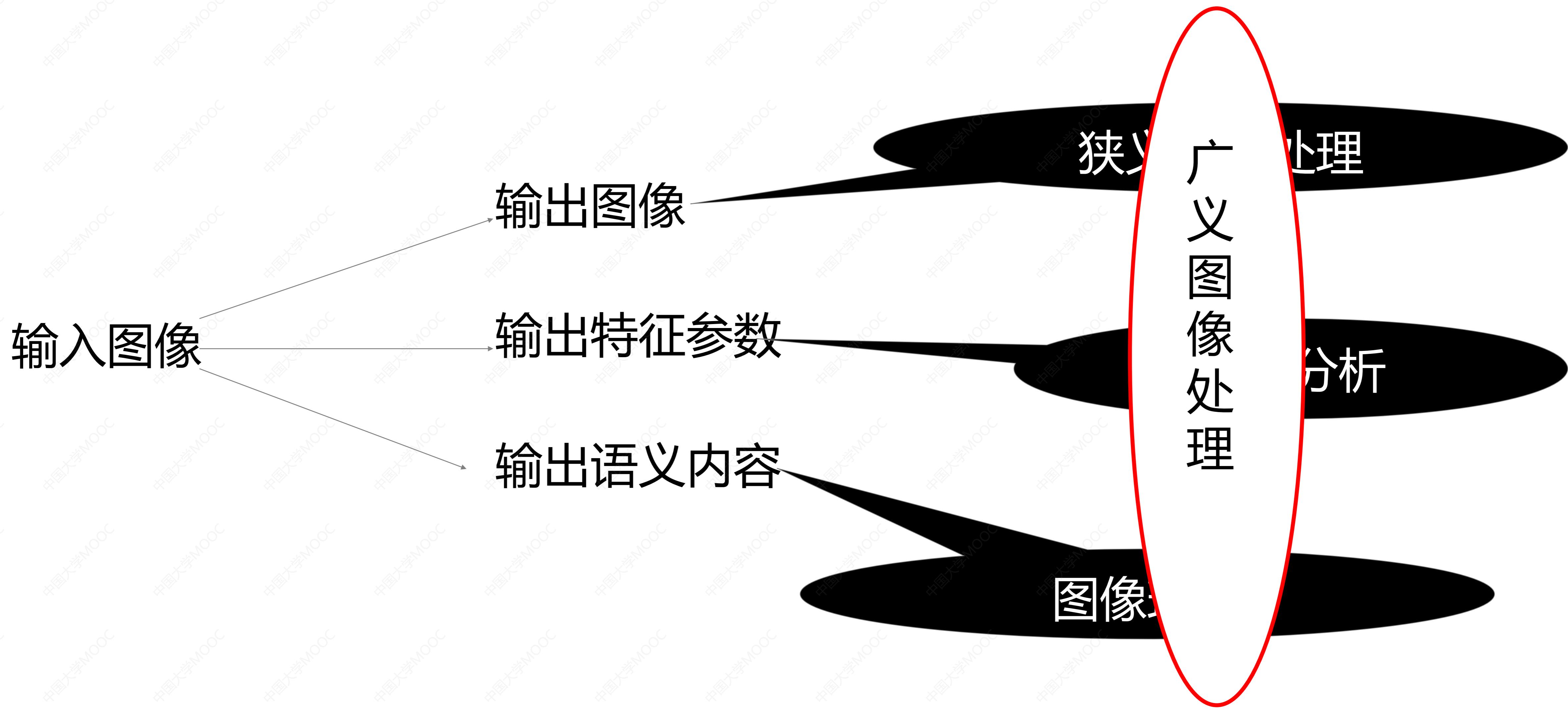
- 形状特征 S
- 位置 Loc: $(50, 79)$,
- 大小 Size: $(60*60)$
- 纹理特征: $\{y_1, y_2, \dots, y_N\}$

➤ 输出语义、语言——图像理解

➤ 绿色的草地上飞一个足球



1.2.3 图像处理的三个层面

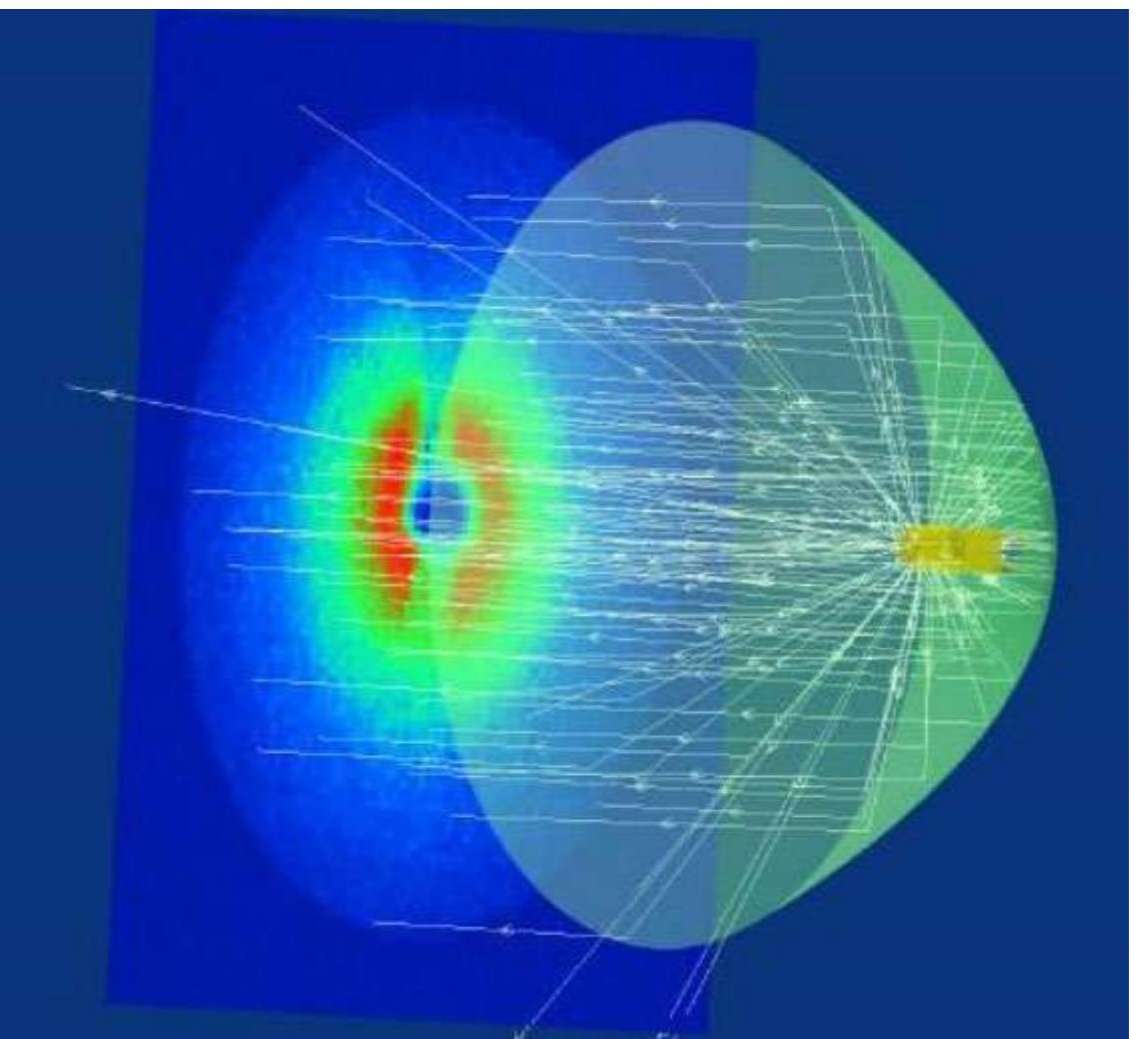
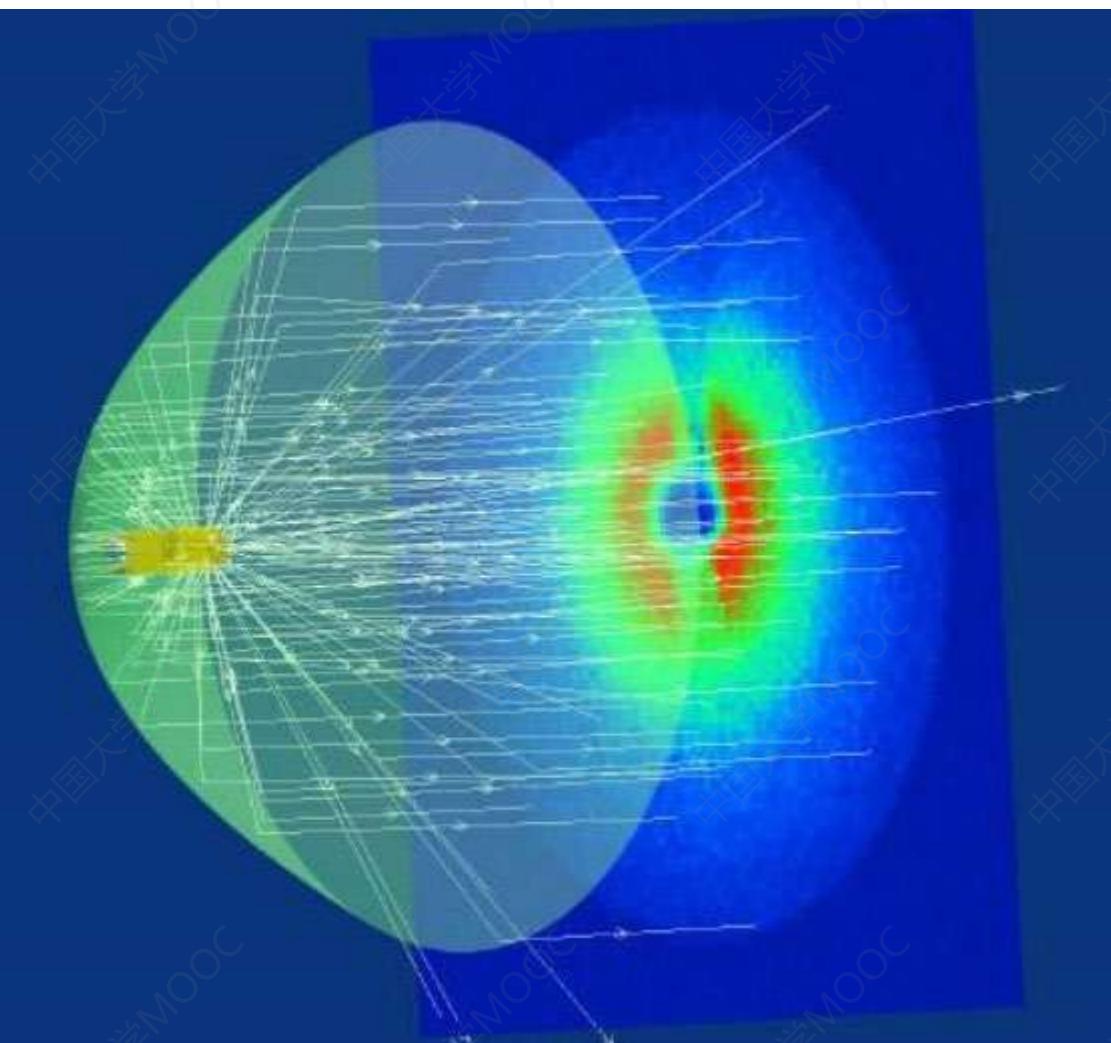


■ 1.2.4 三种图像处理方式对比

图像处理方式	处理速度	灵活性	是否可以智能处理
光学图像处理	超快（并行处理）	差	不可以
模拟电信号图像处理	快	差	不可以
数字图像处理	依赖硬件	好	可以

- 光学图像处理、模拟电信号图像处理有其特殊的应用场景，数字图像处理有广阔的应用前景！

1.3 图像处理技术 的起源和发展



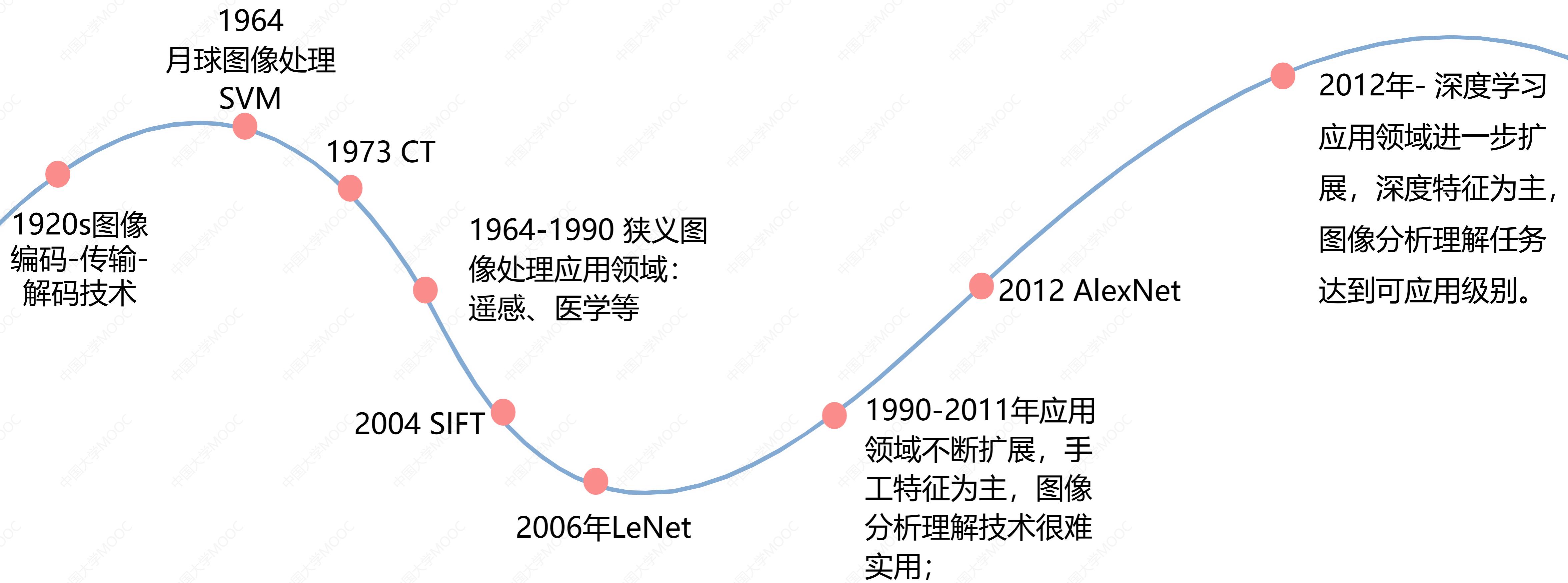
■计算机技术

数字图像处理技术的发展，是由一个个里程碑式的事件组成的。组成了成像/显示技术、狭义图像处理、广义图像处理等阶段，另外数字图像处理技术的发展依赖于计算机的发展

■计算机技术



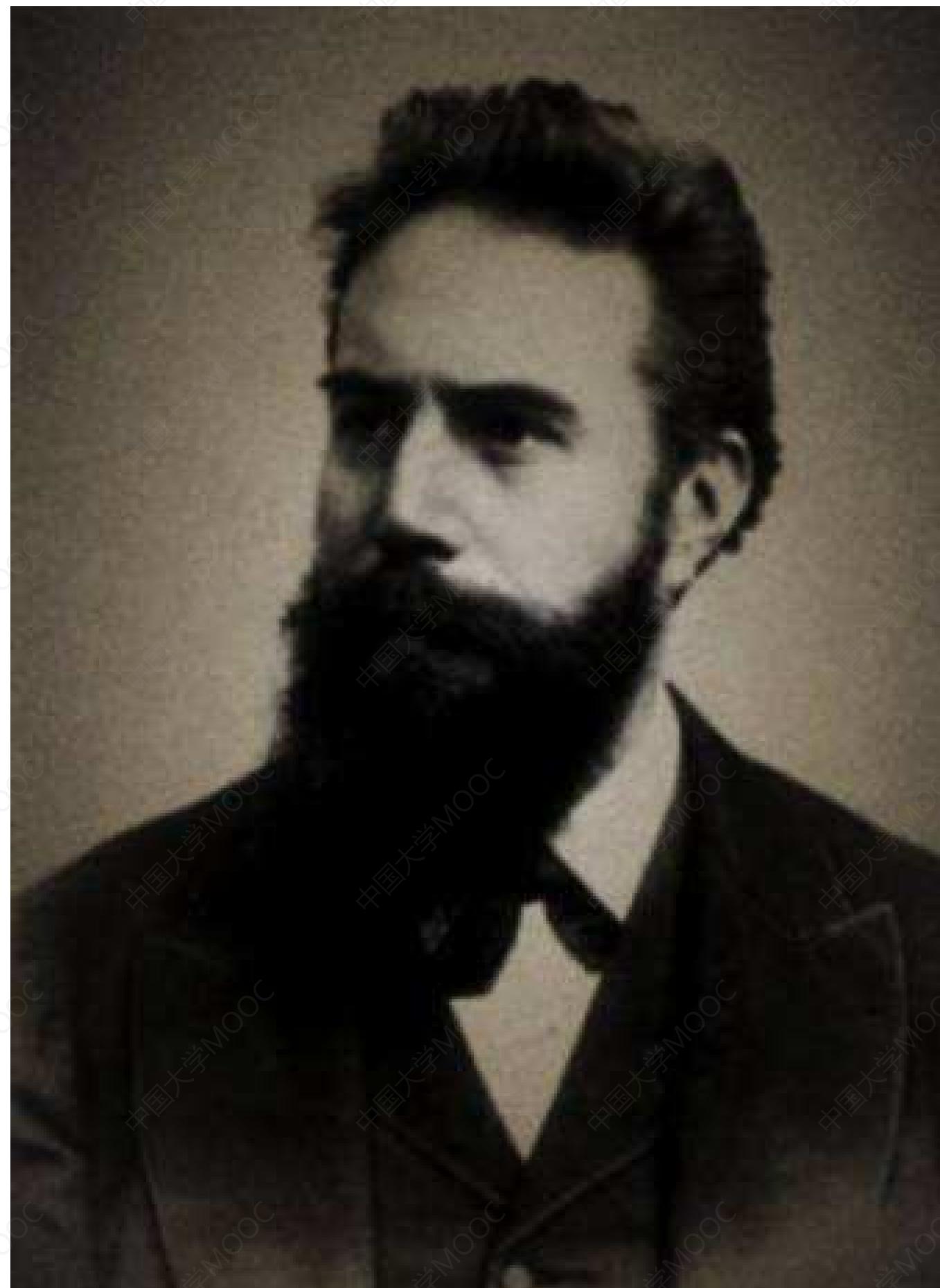
图像处理技术



成像/显示技术



1.3 图像处理技术的起源和发展

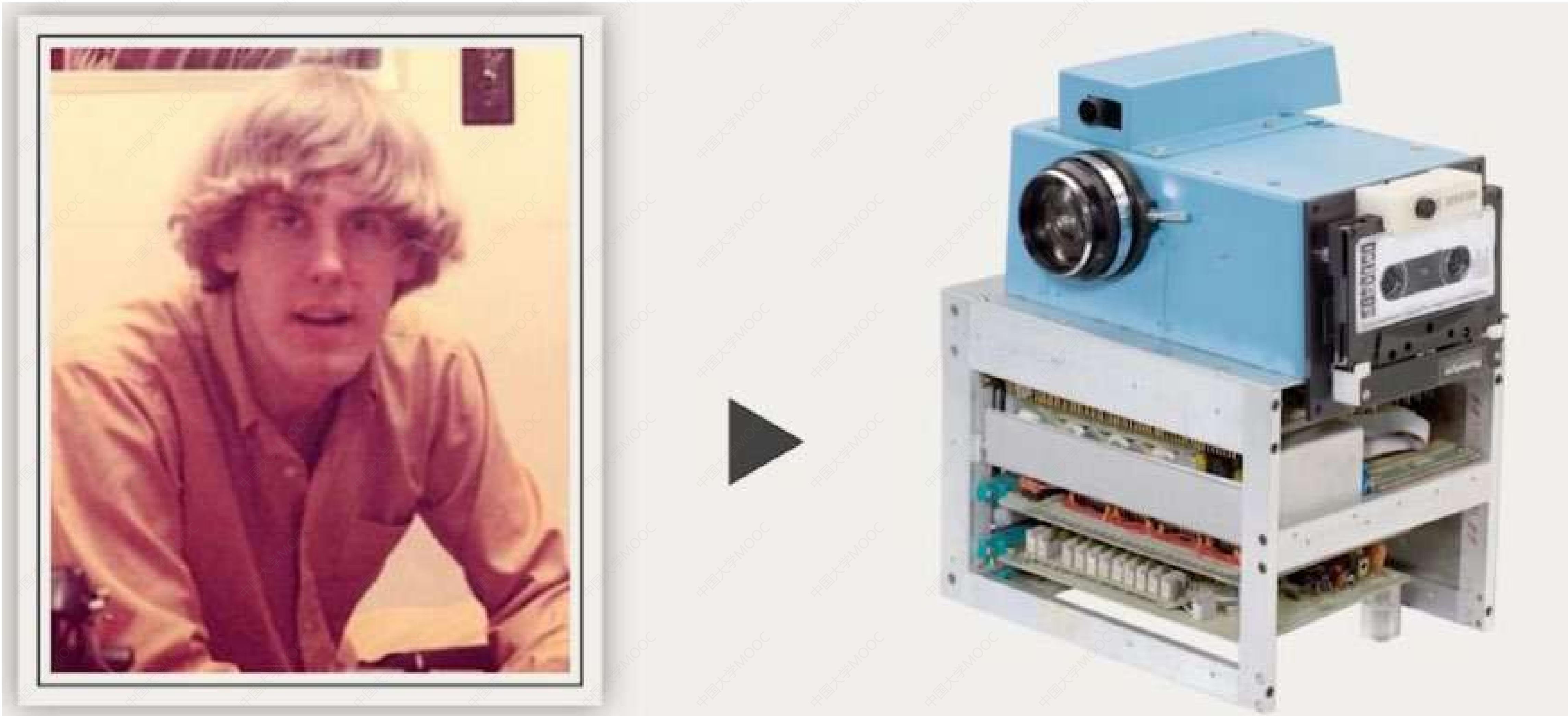


伦琴和他拍摄的第一张X光照片（1895）



安培公司研制的第一台摄像机（1955）

■ 1.3 图像处理技术的起源和发展



柯达工程师Steven Sasson和他研制的第一台数码相机（1975）

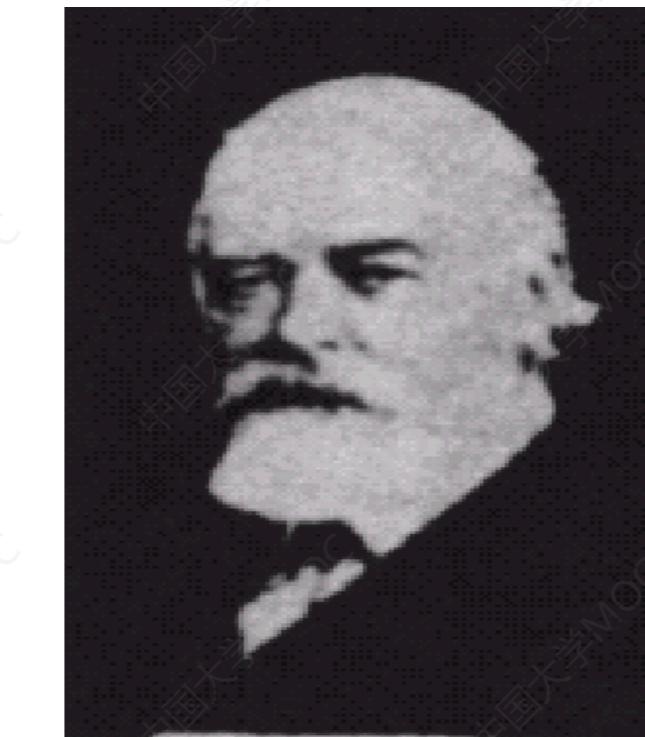
1.3 图像处理技术的起源和发展

1. 图像编码-传输-解码技术-1920s

- Bartlane有线图片传输系统——在纽约和伦敦之间经过大西洋建立海底电缆，传输数字新闻照片。为了用电缆传输图像，首先要进行编码，然后用电报码的方式进行编码，通过海底电缆进行传输，接受到信号后进行解码并通过特殊的打印设备重构该图片。传输一幅数字图像所需的时间由一周多减少到小于3个小时。
- 涉及到的技术：**图象数字化（采样、量化）** ——》 **图象编码** ——》 **图象解码** ——》 **穿孔纸带系统** ——》 **照片复现**



2灰度图像



5灰度图像

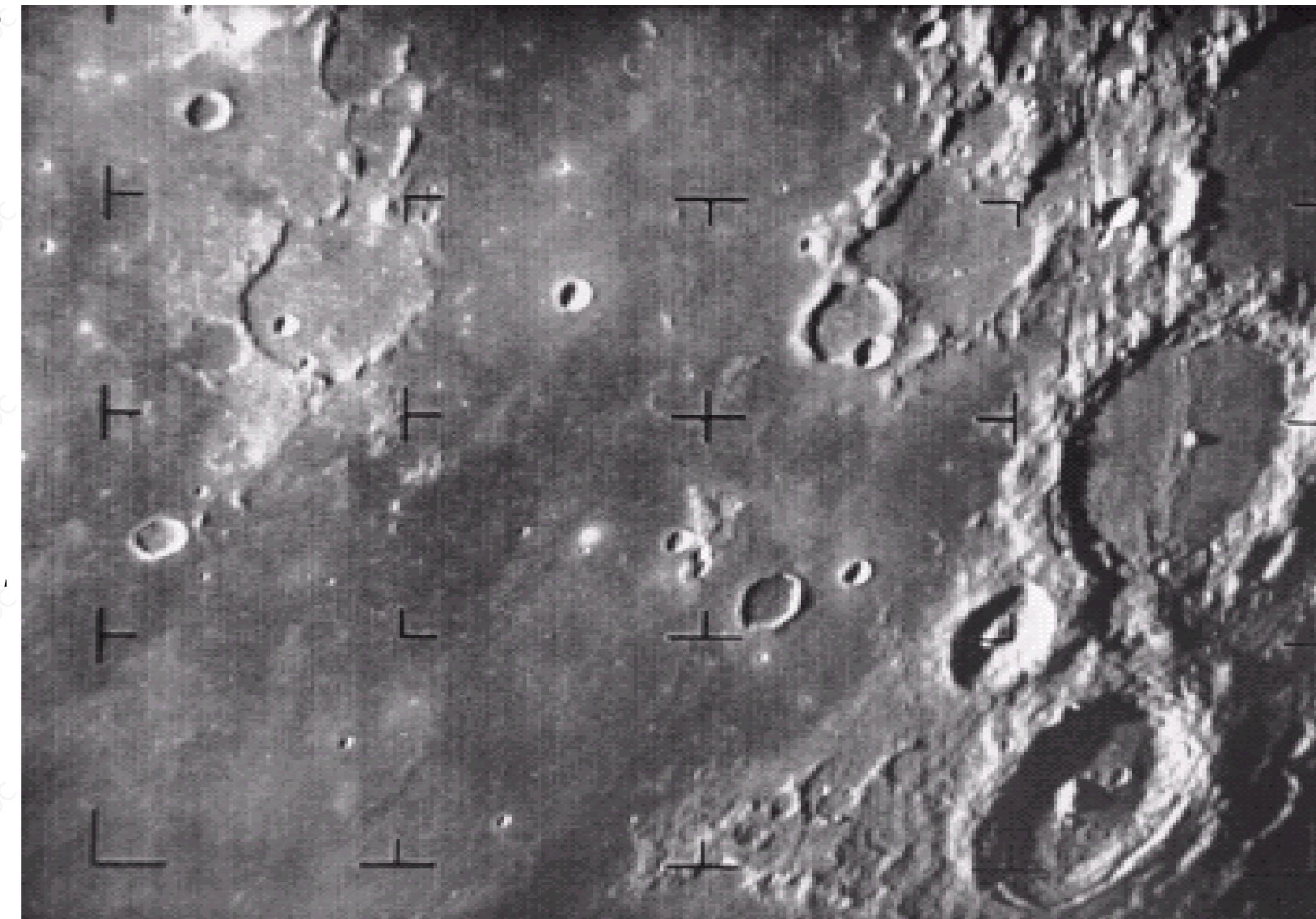


15灰度图像

■ 1.3 图像处理技术的起源和发展

2. 月球图像处理——1964年

- 美国喷气推进实验室 (JPL) 对航天探测器太空漫步者7号 (Ranger 7) 在1964年发回的几千张月球照片使用了图像处理技术，如几何校正、灰度变换、去除噪声等方法进行处理，并考虑了太阳位置和月球环境的影响，由计算机成功地绘制出月球表面地图，获得了巨大的成功。随后又对探测飞船发回的近十万张照片进行更为复杂的图像处理以致获得了月球的地形图、彩色图及全景镶嵌图，获得了非凡的成果，为人类登月创举奠定了坚实的基础，也推动了数字图像处理这门学科的诞生。
- 在以后的宇航空间技术，如对火星、土星等星球的探测研究中，数字图像处理技术都发挥了巨大的作用。



Ranger 7 拍摄的第一张月球照片

■ 1.3 图像处理技术的起源和发展

3. Computerized Tomography (CT) —— 1972

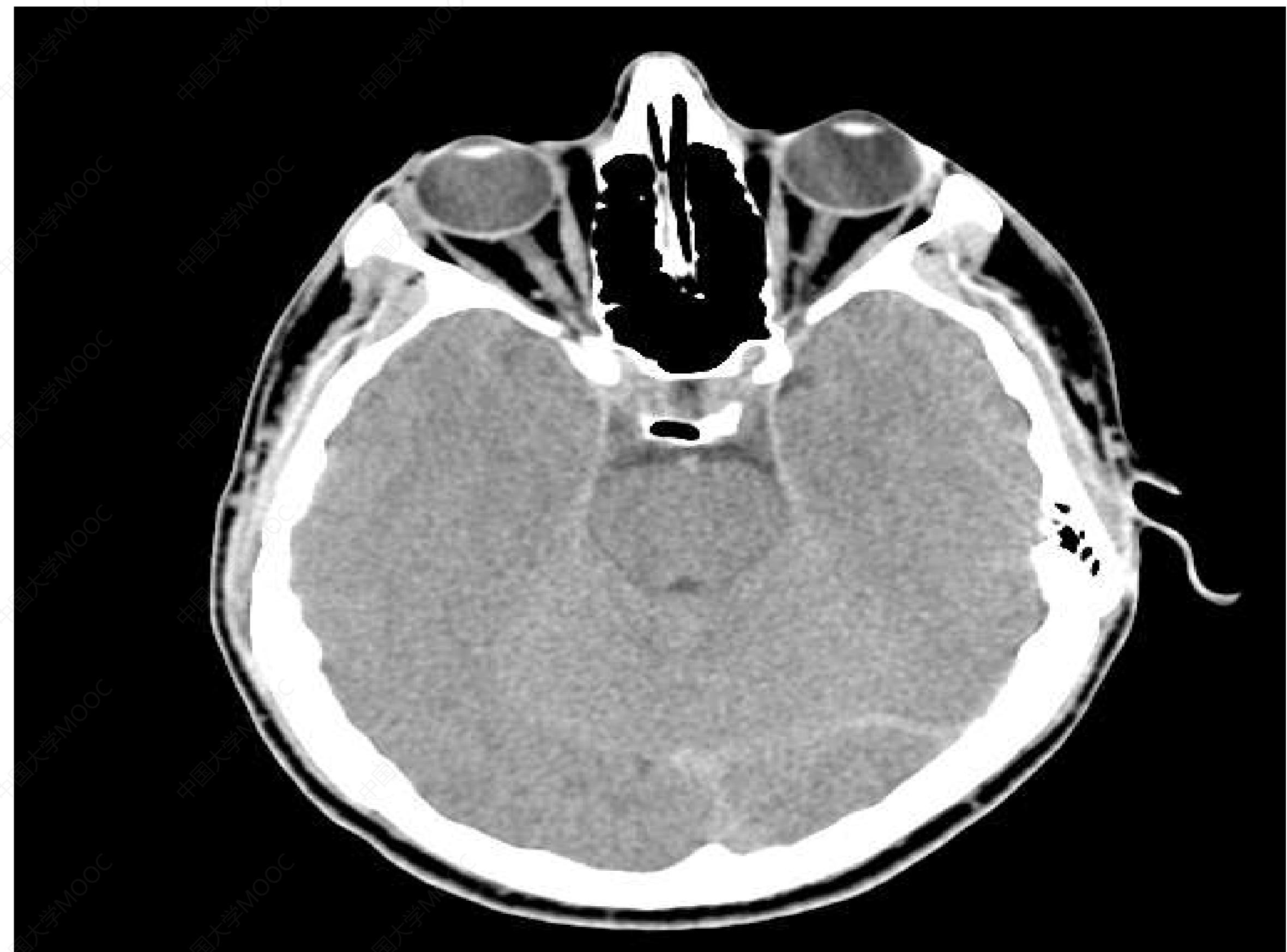
1917年 Radon 变换

1971年9月 英国EMI公司工程师霍斯费尔德与一位神经放射学家合作制造了一种X射线断层摄影装置

1972年4月，霍斯费尔德在英国放射学年会上公布了这一结果

1974 年被正式命名为计算机断层摄影术
(CT: Computerized Tomography)

1979年，获得Nobel医学奖



1.3 图像处理技术的起源和发展



G. N. Hounsfield

Central Research Laboratories,
EMI
London



A. M. Cormack

Tufts University
Medford, MA, USA

■ 1.3 图像处理技术的起源和发展

- CT技术的原理 (图像重建)
- Radon理论：任意二维或三维物体可以用其无穷多个投影数据来表征。

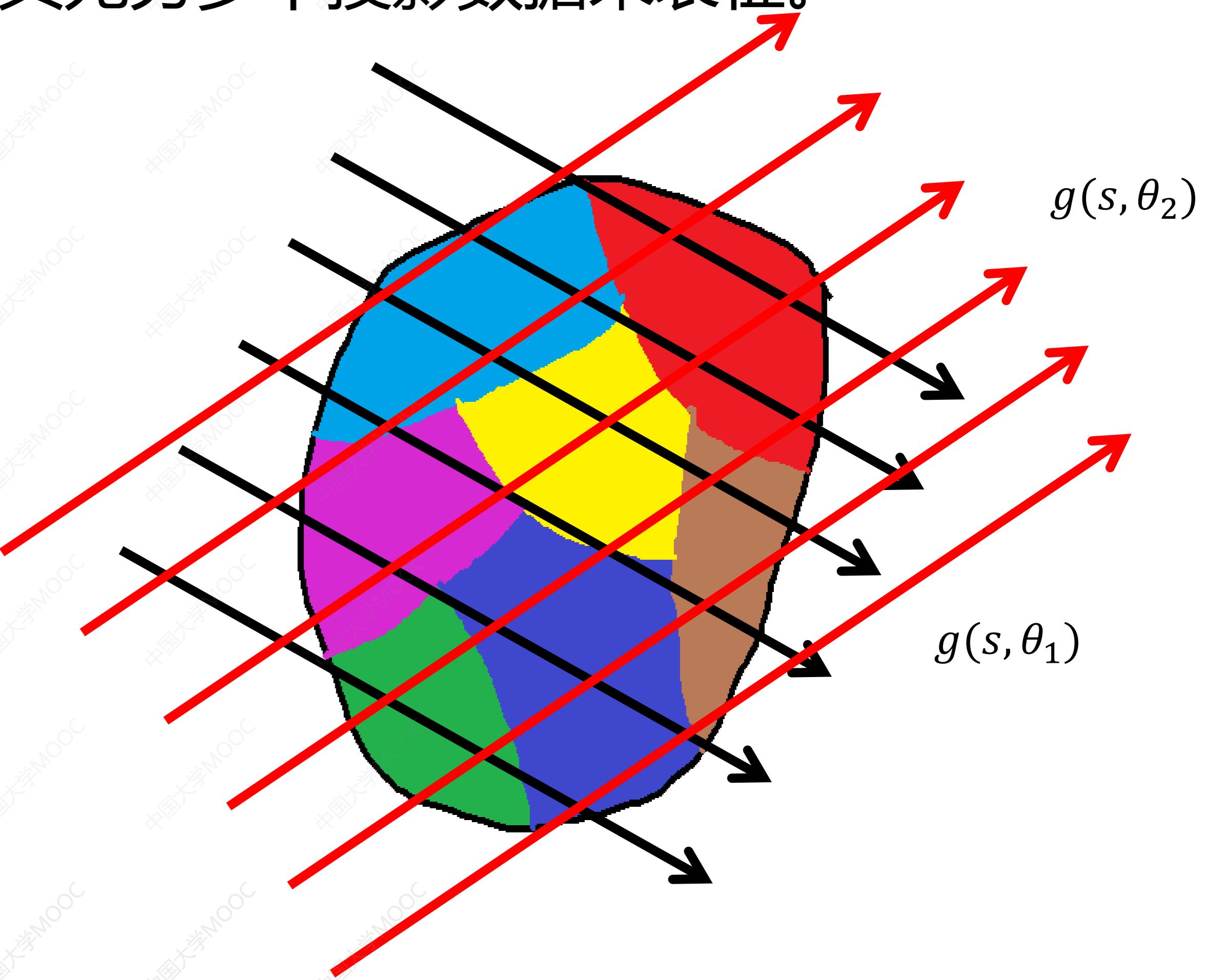
$$f(x, y) \Leftrightarrow \{g(s, \theta_1), g(s, \theta_2), \dots\}$$

- 对于数字化的图像

$$f(x, y) \Leftrightarrow \{g(s, \theta_1), g(s, \theta_2), \dots, g(s, \theta_N)\}$$

- 投影数据的物理意义：以CT为例

- 不同组织对X射线的衰减系数不同
- 一条射线的总衰减量是不同组织衰减量的累加
- 投影数据与组织之间呈线性积分关系



■ 1.3 图像处理技术的起源和发展

- Allan MacLeod Cormack: 1944年南开普敦大学物理学学士，1946年南开普敦大学晶体学硕士，1947年剑桥大学卡文迪实验室Research Student，1957年秋，美国塔夫斯（Tufts）大学物理系教授。
- 1963年发表论文“函数的直线积分表示及其放射学应用”。
 - (1) 确定二维区域内X射线吸收率的变化；
 - (2) 确定物质中正电子湮灭的非均匀分布；
 - (3) 确定恒定成分物质中的密度变化。

■ 1.3 图像处理技术的起源和发展



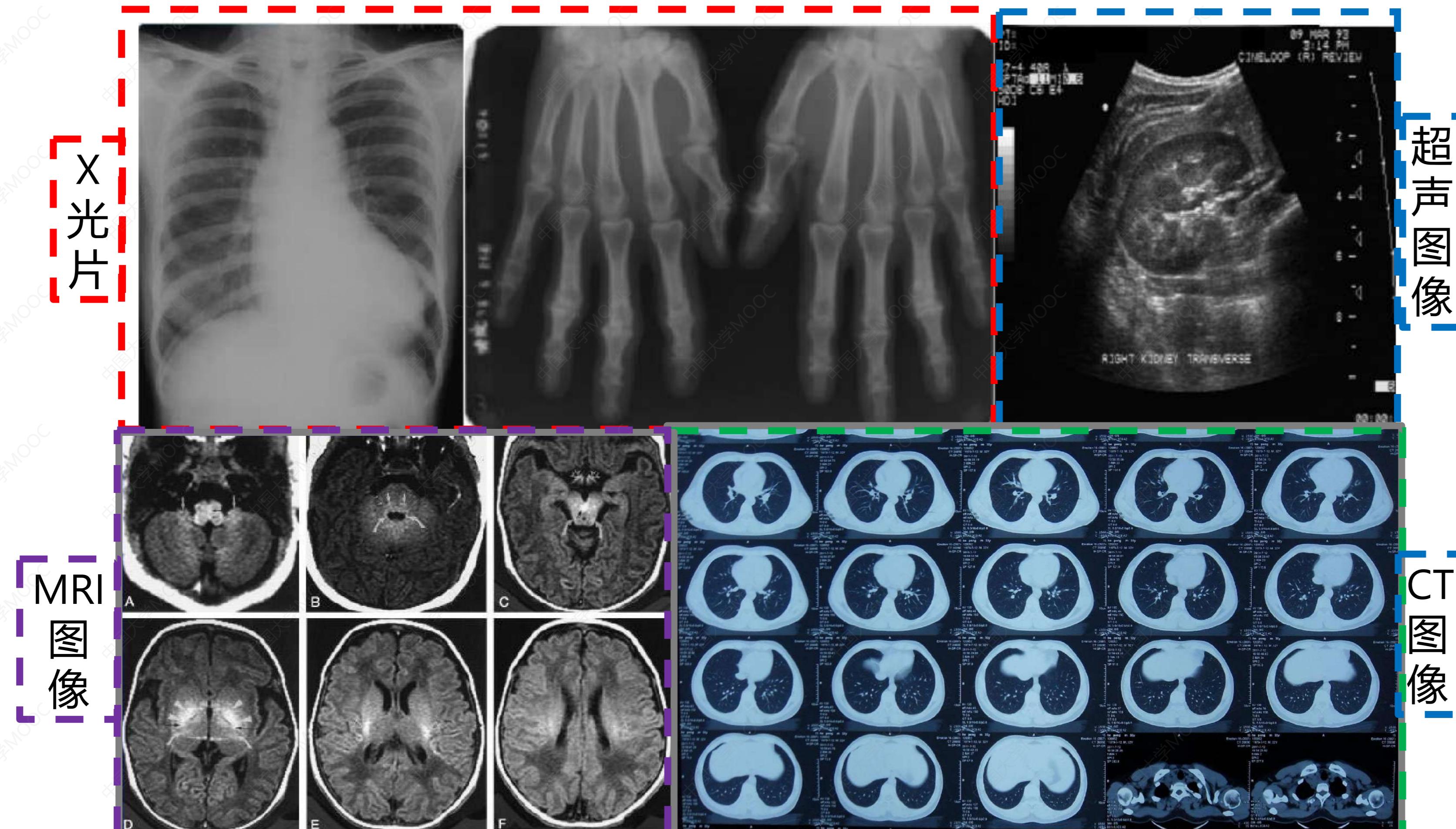
一个CT设备

1.3 图像处理技术的起源和发展



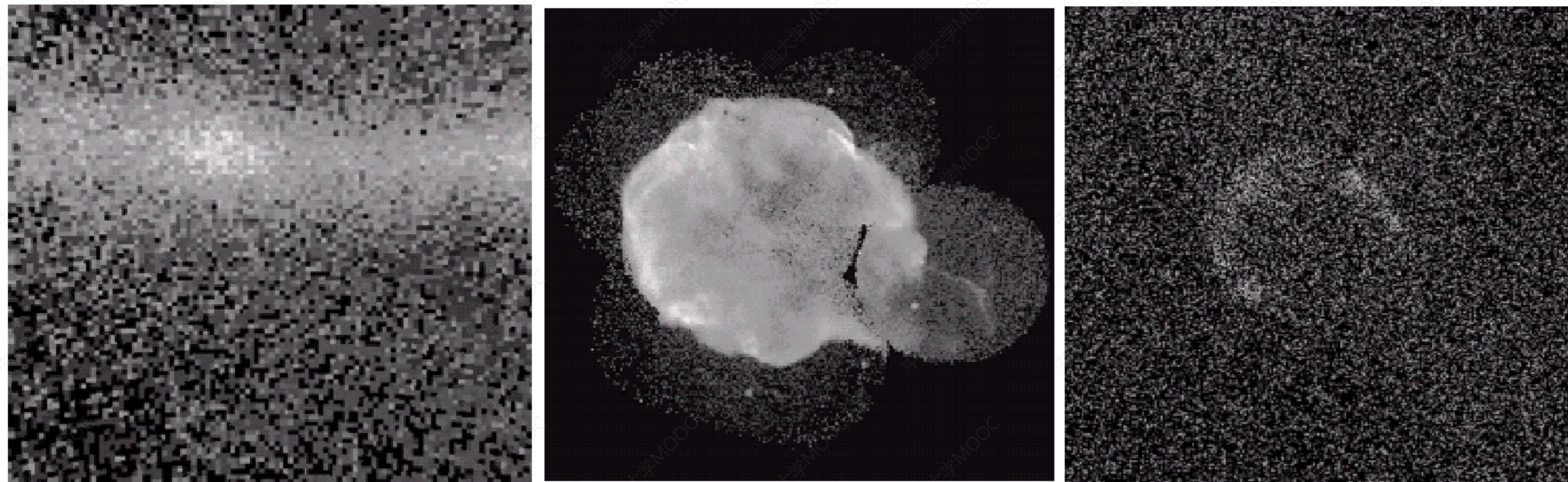
自此以后，图像处理技术在**医学、遥感、航空航天**等重要领域受到重视并取得了重大的开拓性成就，后来随着个人计算机的发展，数字图像处理在我们生产生活的各个领域的有广泛应用，如航空航天、生物医学工程、工业检测、机器人视觉、公安司法、军事制导、文化艺术等，使图像处理成为一门引人注目、前景远大的新型学科。

1.3 图像处理技术的起源和发展

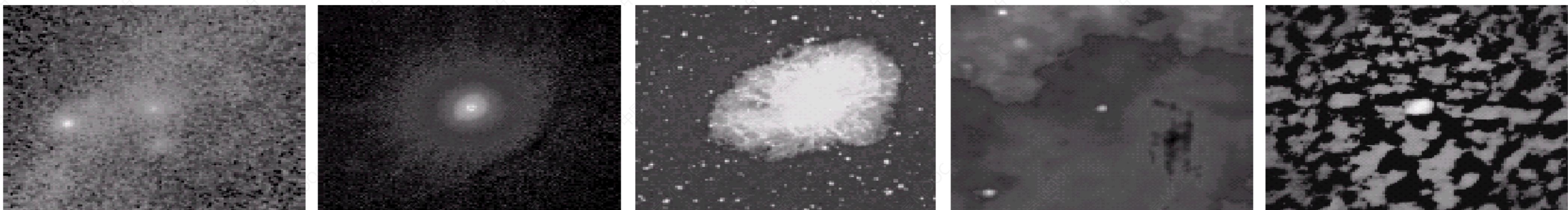


医学领域

■ 1.3 图像处理技术的起源和发展



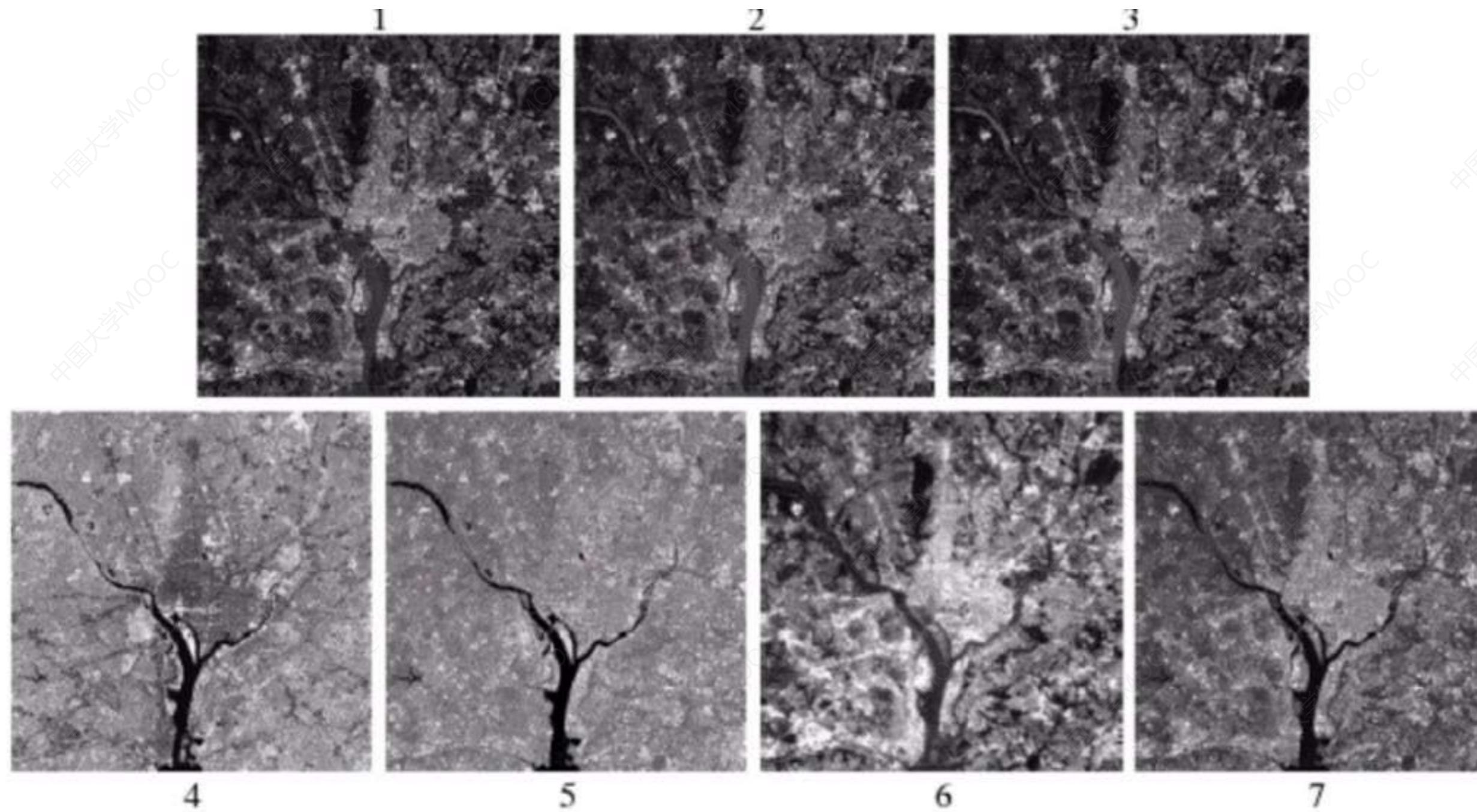
(a) 伽马波段图像 (b) X射线波段图像 (c) 紫外线波段图像
天鹅环不同波段下的图像



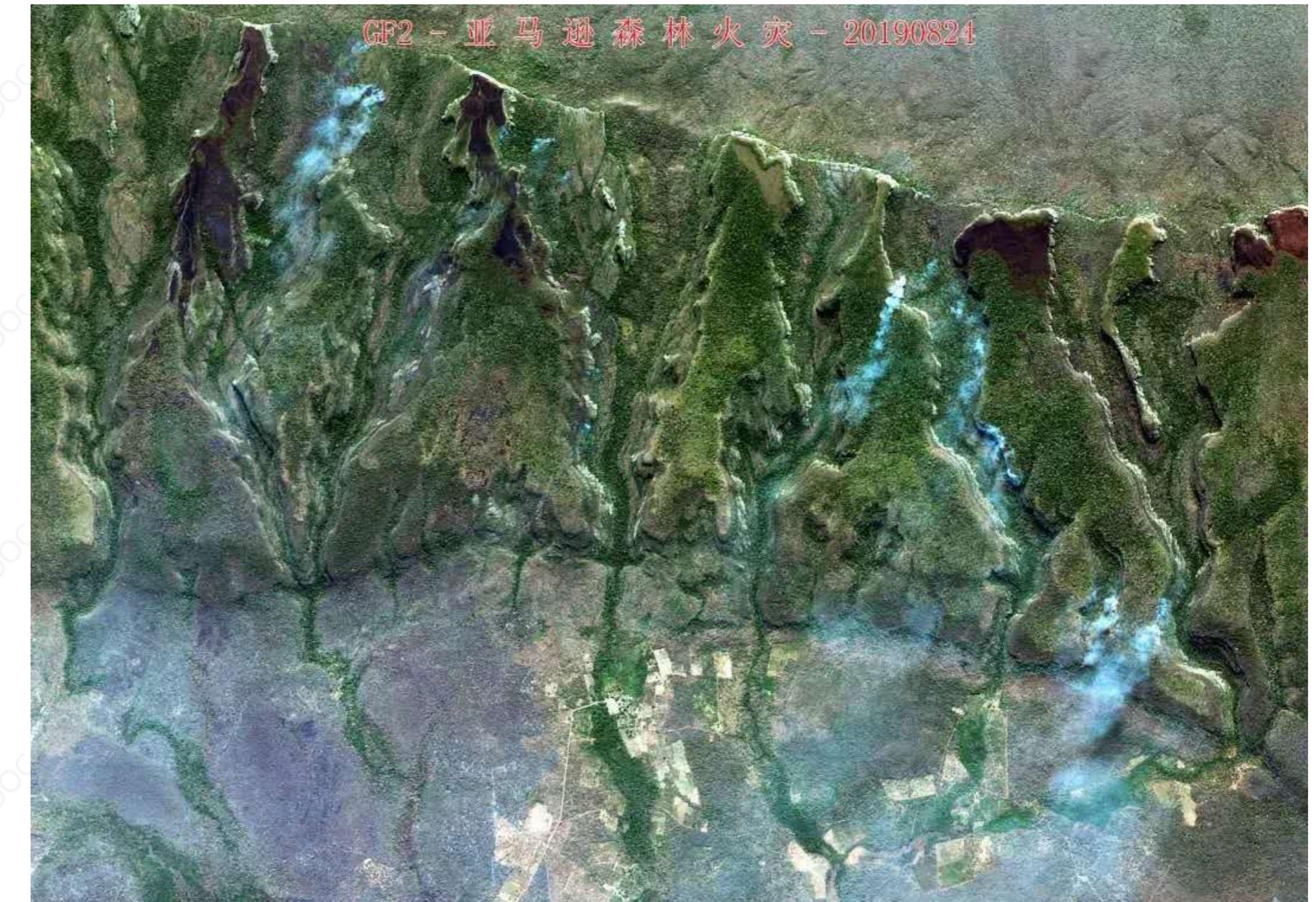
(a) 伽马波段 (b) X射线波段 (c) 可见光波段 (d)红外波段 (e)无线电波段

蟹状星云脉冲星不同波段下的图像

天文领域



(a) 华盛顿地区的一组主题映射 (TM) 图像

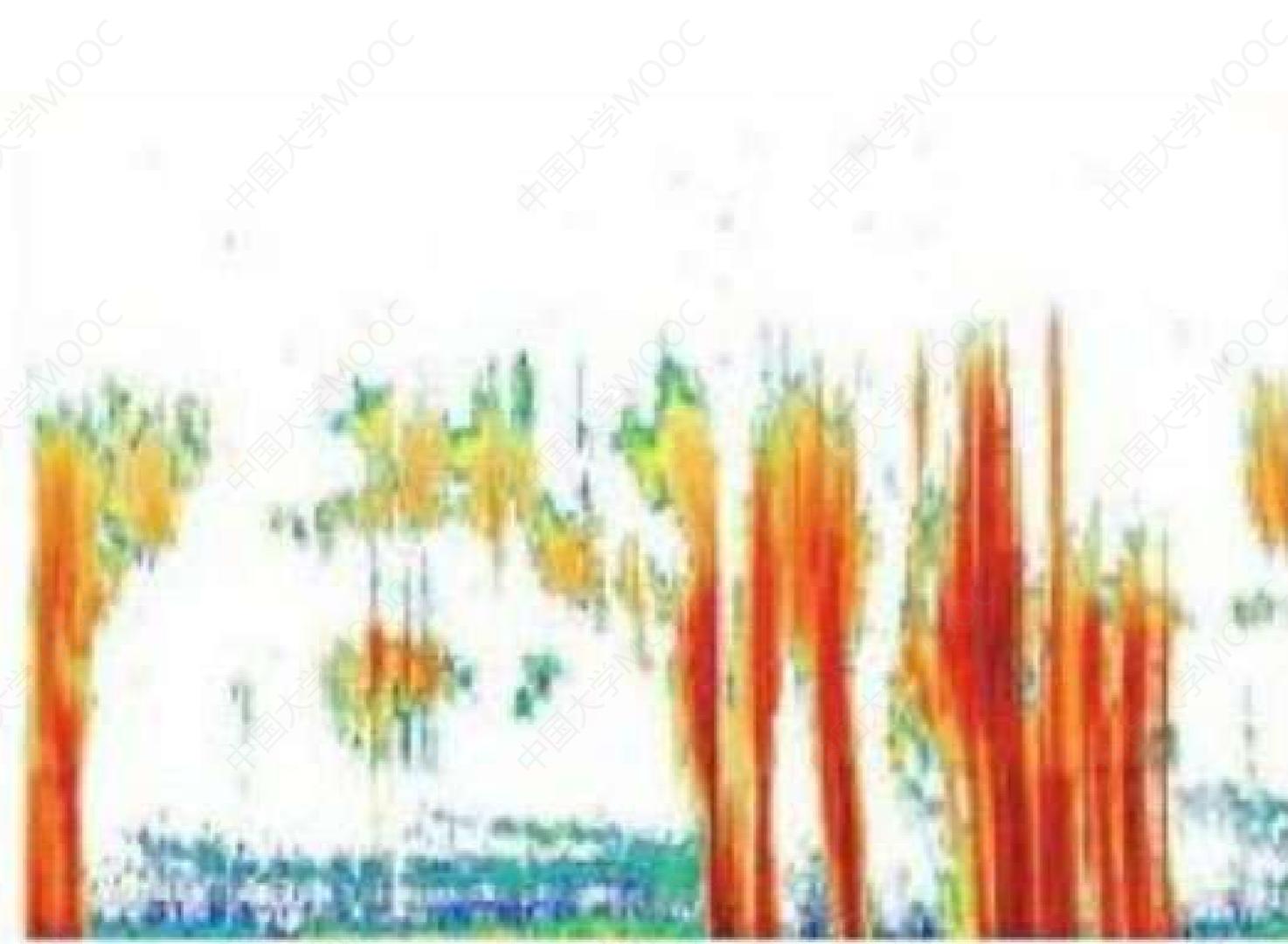


(b) 亚马逊区域大火检测

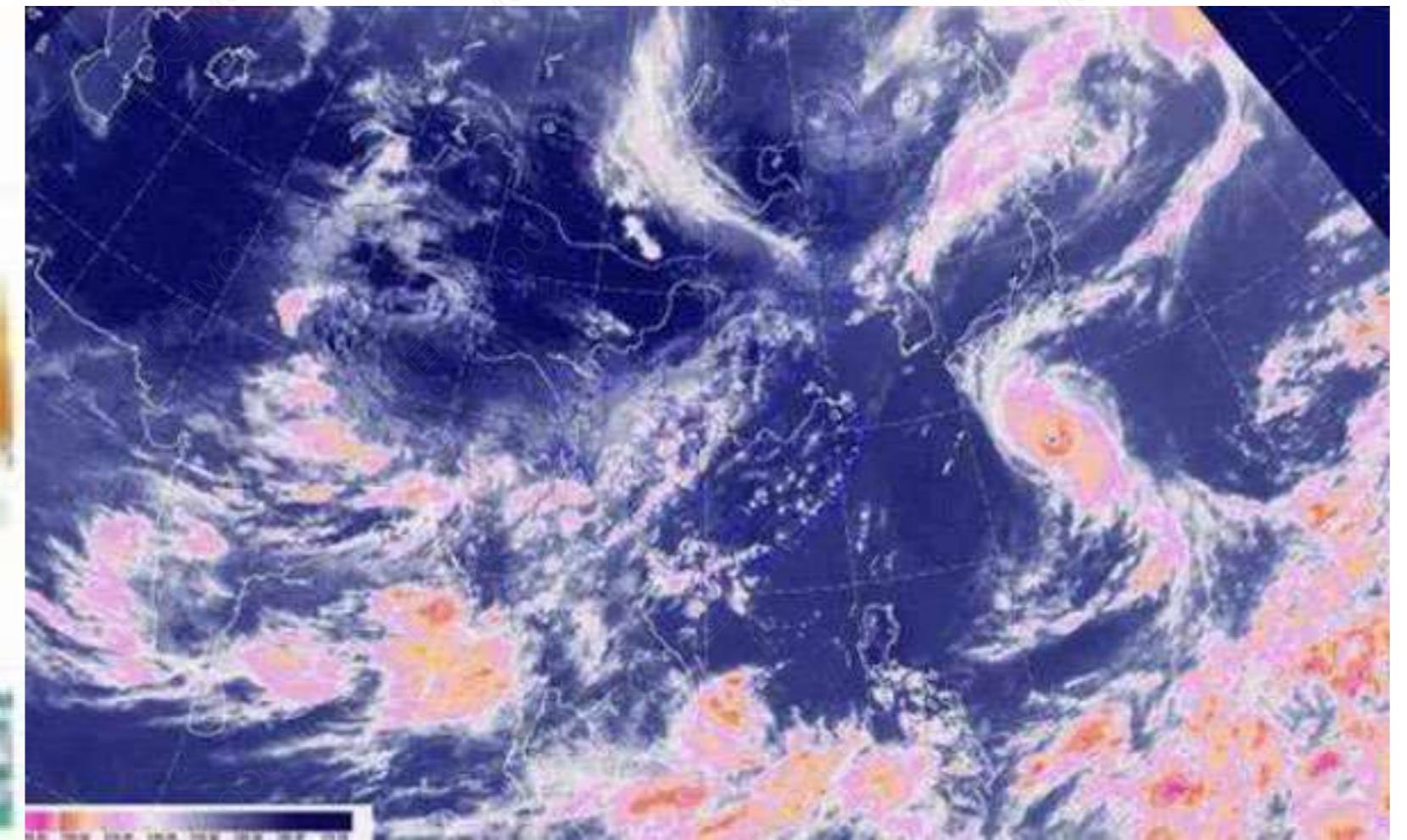
遥感领域

气象领域

(c) 农业气象图像采集和分析



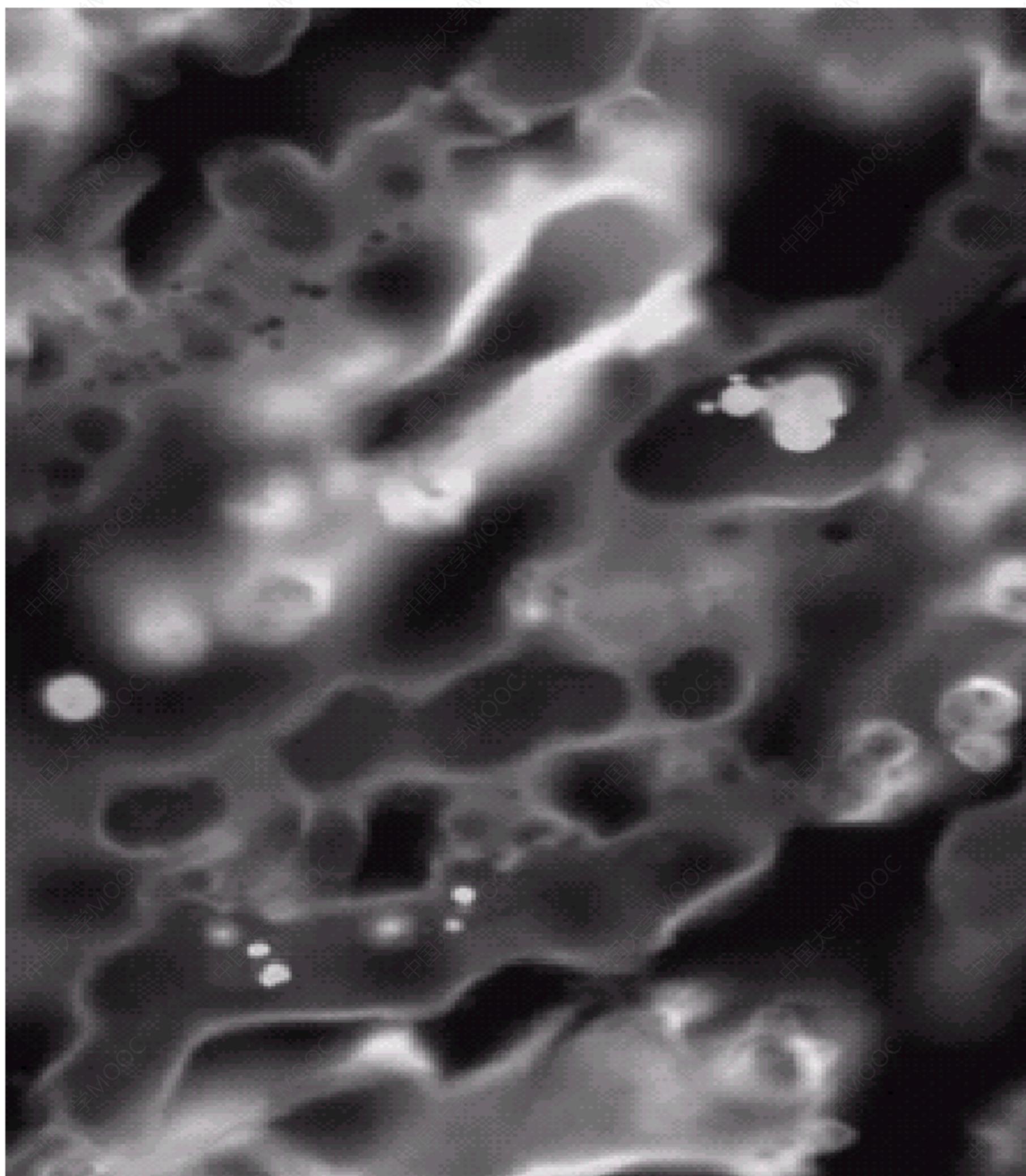
(a) 毫米波雷达气象云图



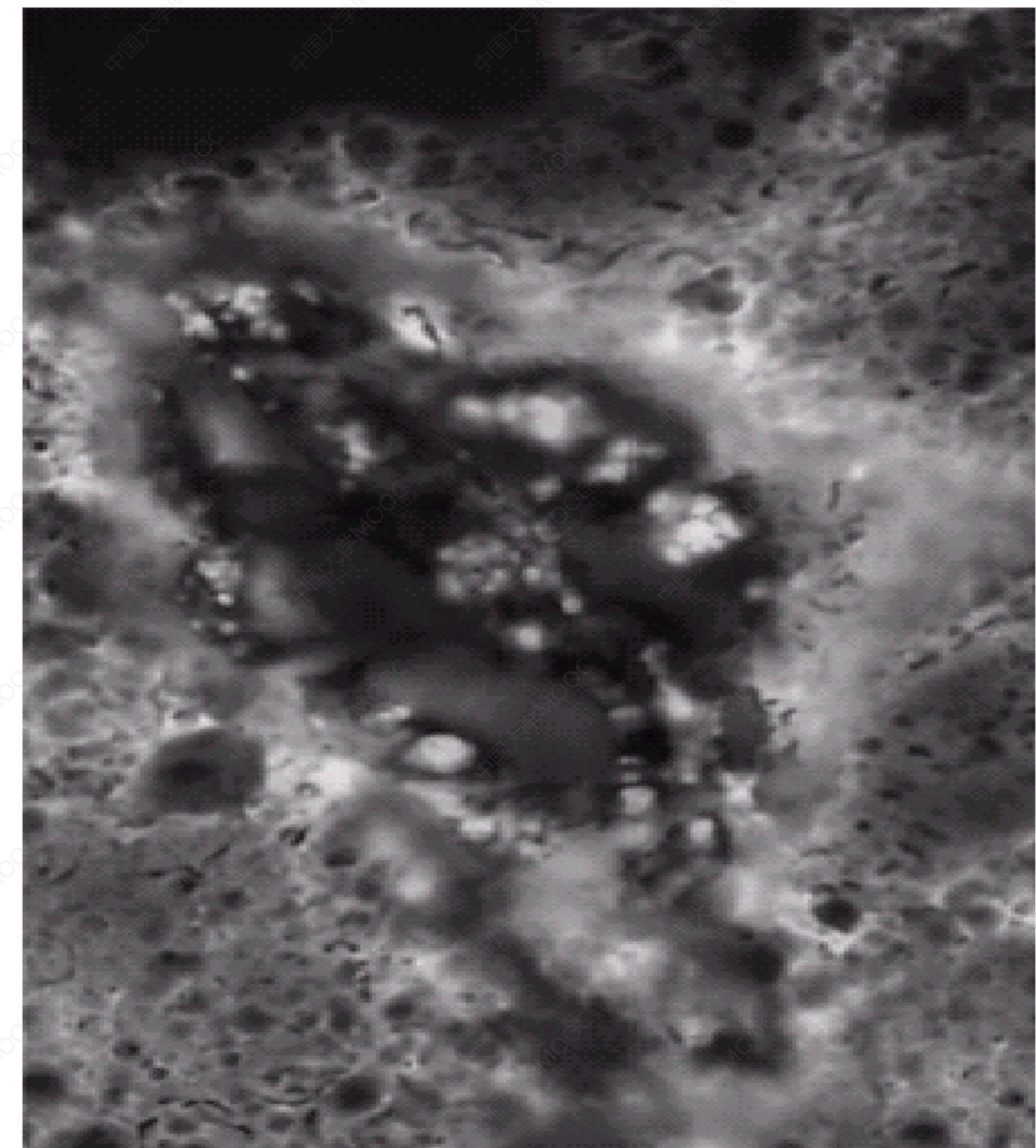
(b) 气象卫星云图



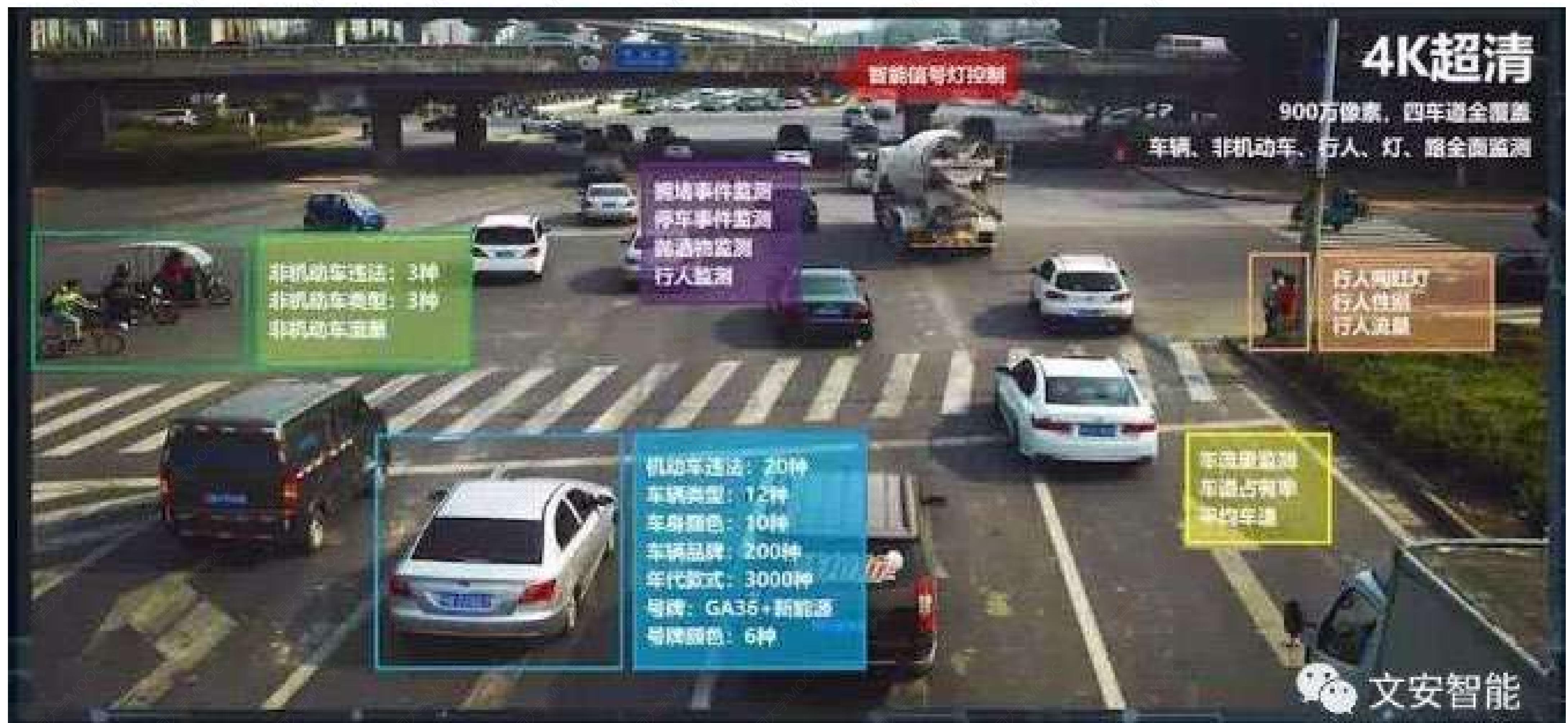
农业领域



(a) 正常玉米胚胎紫外光图像



(b) 有黑穗病的玉米胚胎的紫外图像



交通领域

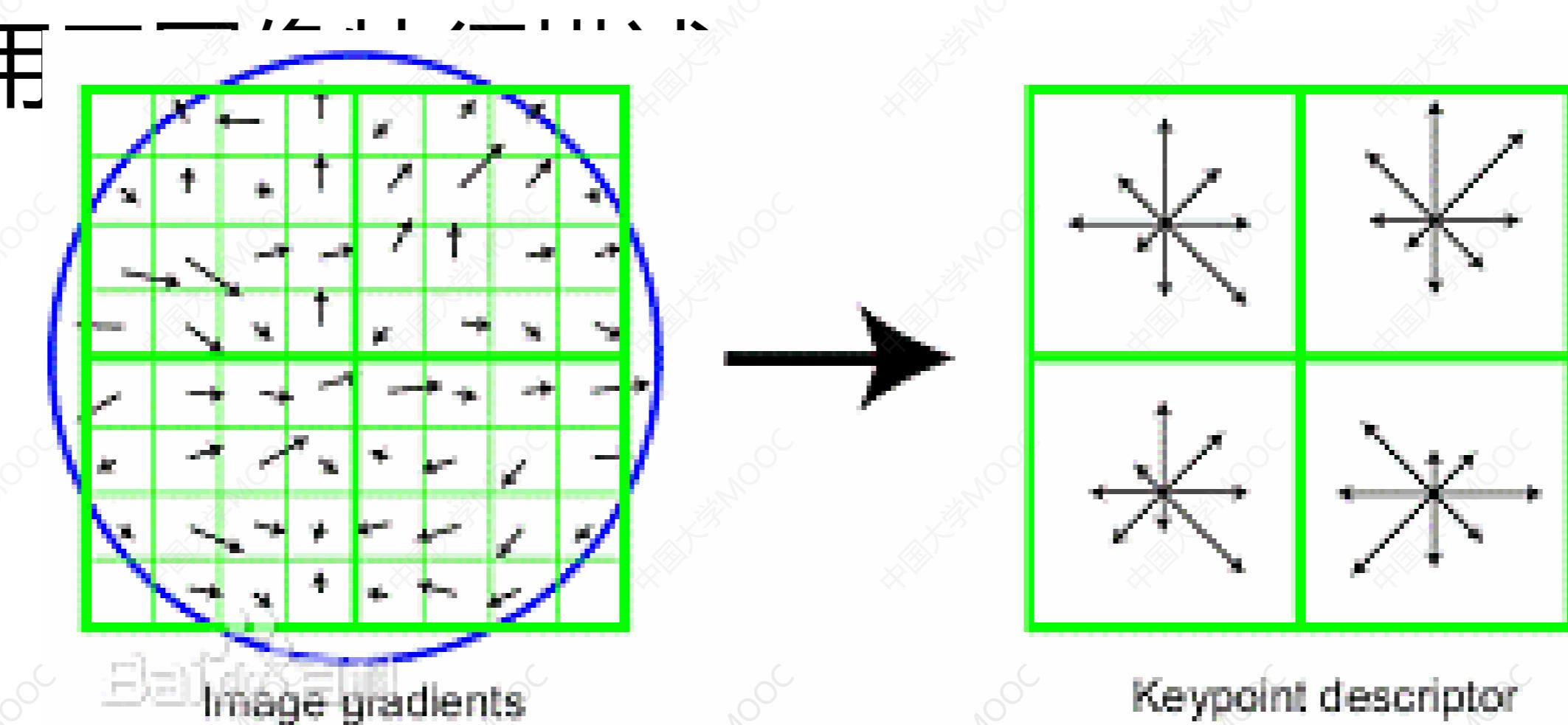
■ 1.3 图像处理技术的起源和发展

- 关于Lena图像
- Lena是一个瑞典模特，1972为某杂志拍摄插页，William K. Pratt博士截取呈现在的Lena图像，此后Lena图像在图像处理领域流行起来。
 - 1) 这是一个令人赏心悦目的图像；
 - 2) 这是一副很好的测试图像！它包含了各种细节、平滑区域、阴影和纹理等要素，这些要素对测试各种图像处理算法很有用。
- 1997年第五十届IS&T，Lena受邀参会，并与很多用Lena图像作为实验图像的论文合影留念。



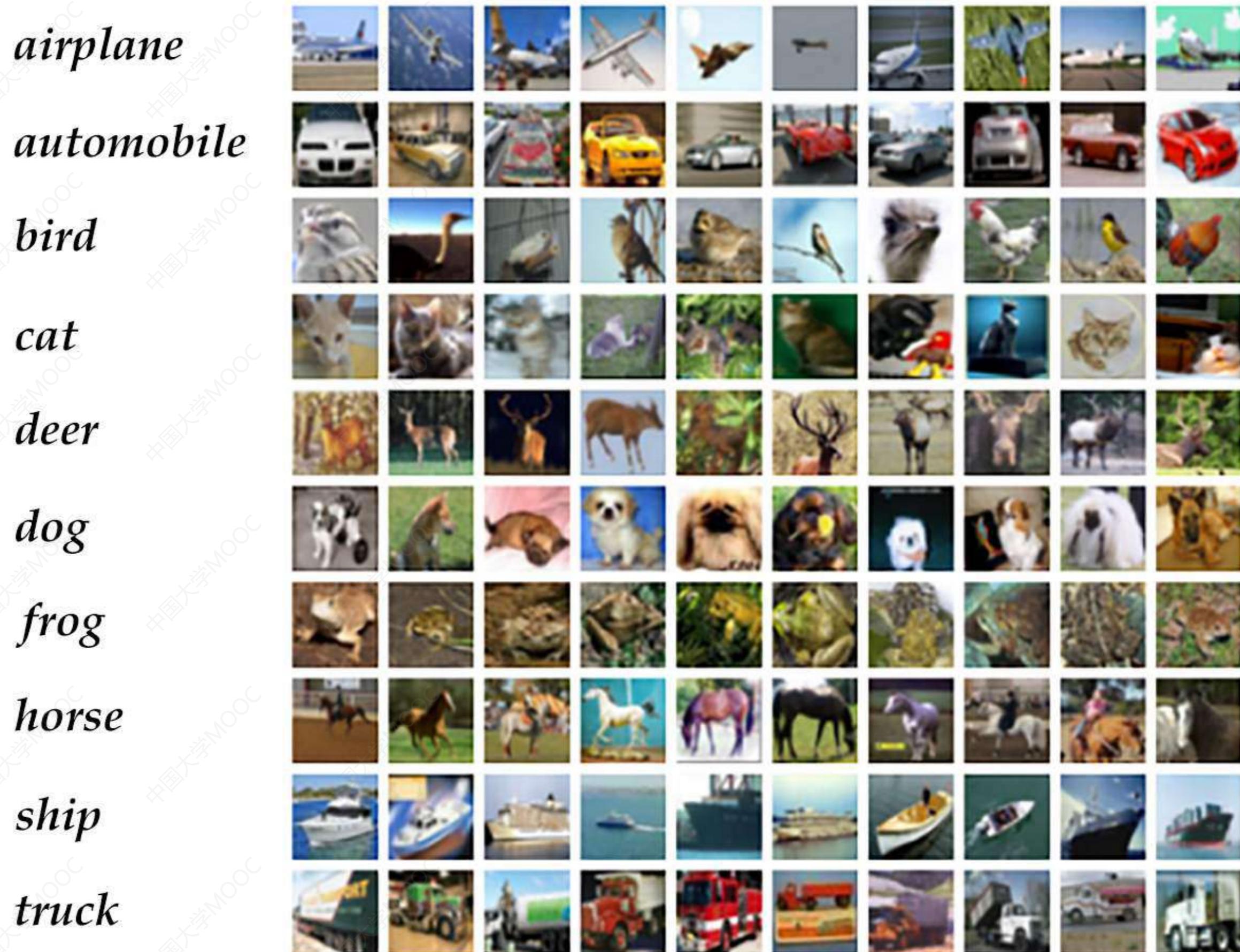
■ 1.3 图像处理技术的起源和发展

- SIFT (Scale-invariant feature transform) 特征
- 又称尺度不变特征变换，是一种图像局部特征描述子。David Lowe [2] 于 ICCV2019首次提出，并于2004年整理完善后发表于IJCV。它可在图像中检测出关键点，并对关键点的局部特征进行描述，由于特征描述子的具有尺度、方向不变性，对于不同视角、不同尺度、有畸变等物体匹配都有很强鲁棒性。与特征点位置一起常用



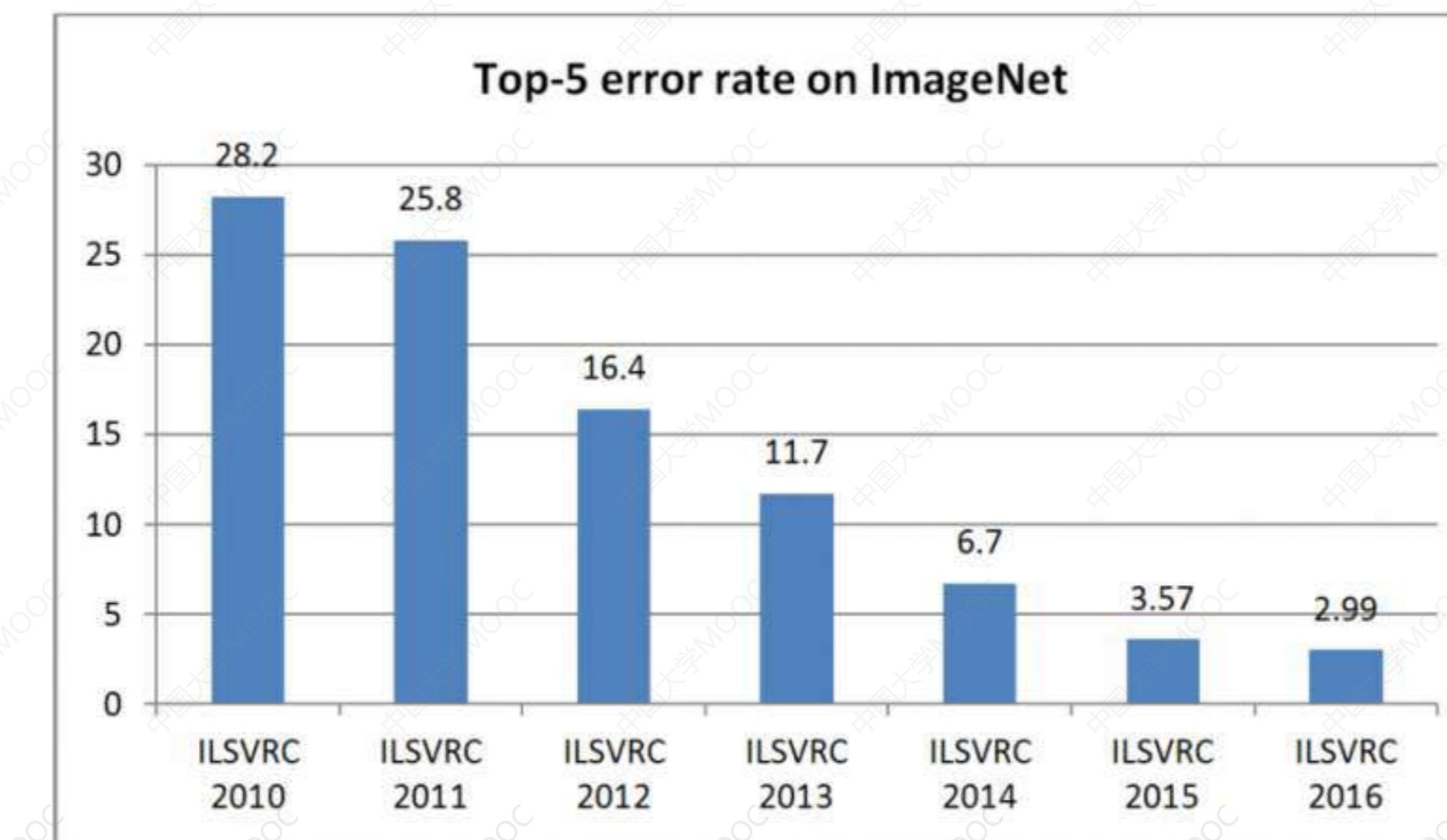
■ 1.3 图像处理技术的起源和发展

- ImageNet竞赛与AlexNet
- ImageNet图像数据集始于2009年，斯坦福大学李飞飞教授等在CVPR2009上发表了一篇名为《ImageNet: A Large-Scale Hierarchical Image Database》的论文，根据WordNet层次结构组织的图像数据集（每一个具有相同意义的字条组称为一个synset），对每个synset提供平均1000个图像，图像均来自于互联网，大小、尺度、视角各异。包含14197122幅图像，总共分为21841个类别。



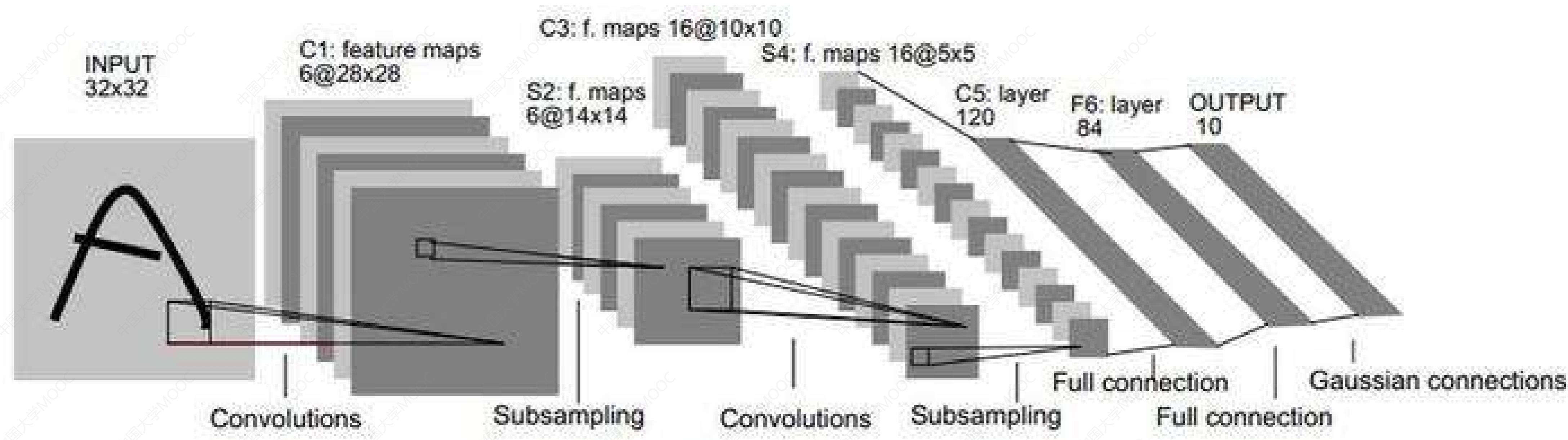
■ 1.3 图像处理技术的起源和发展

- ImageNet挑战赛ImageNet Large Scale Visual Recognition Challenge(ILSVRC)(2010-2017)
- 竞赛内容最早只有分类，后来加入目标检测、场景分类等任务。
- 训练集：1000类120万+张图片训练，验证集：50000张，测试集：10万张测试



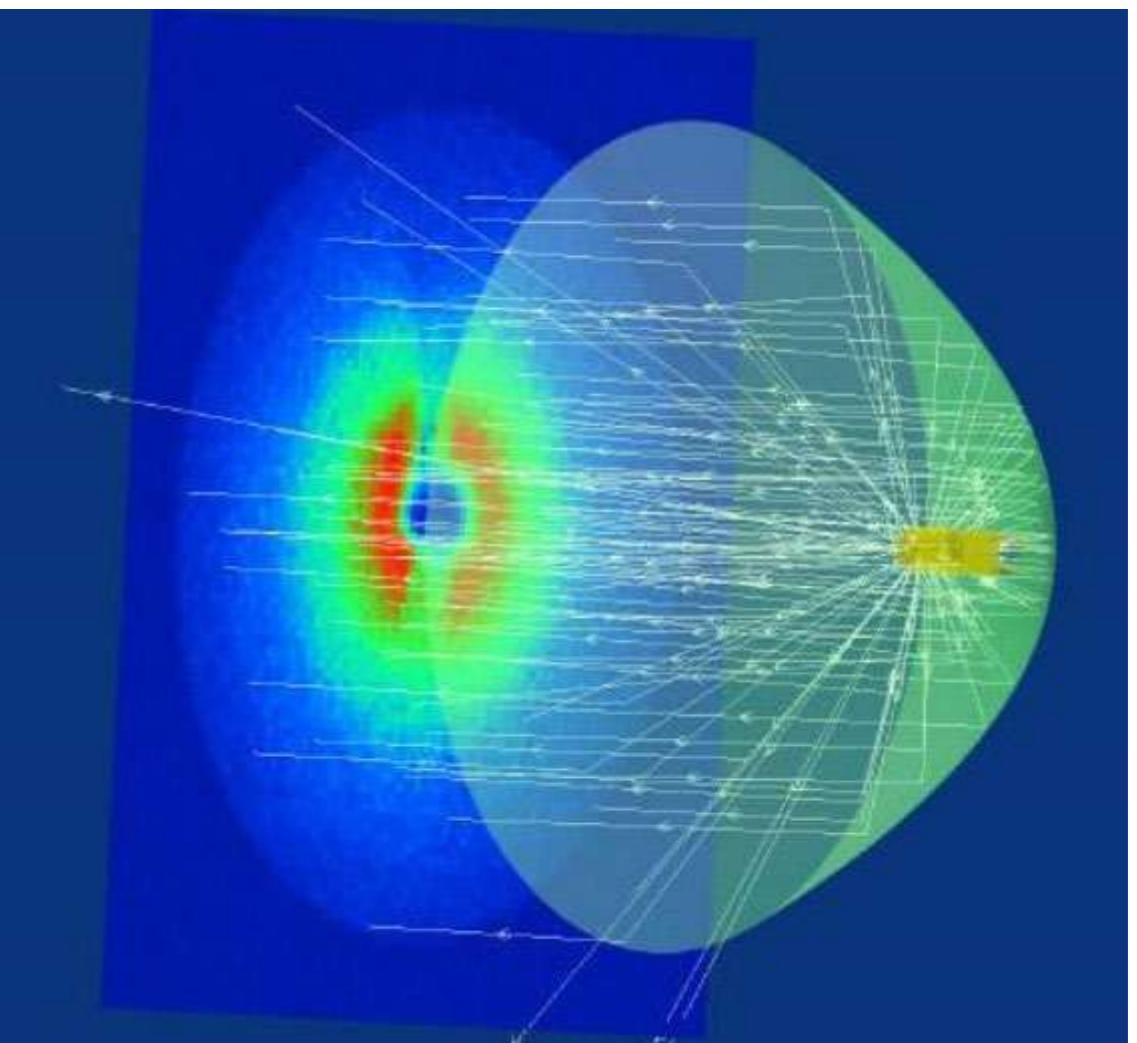
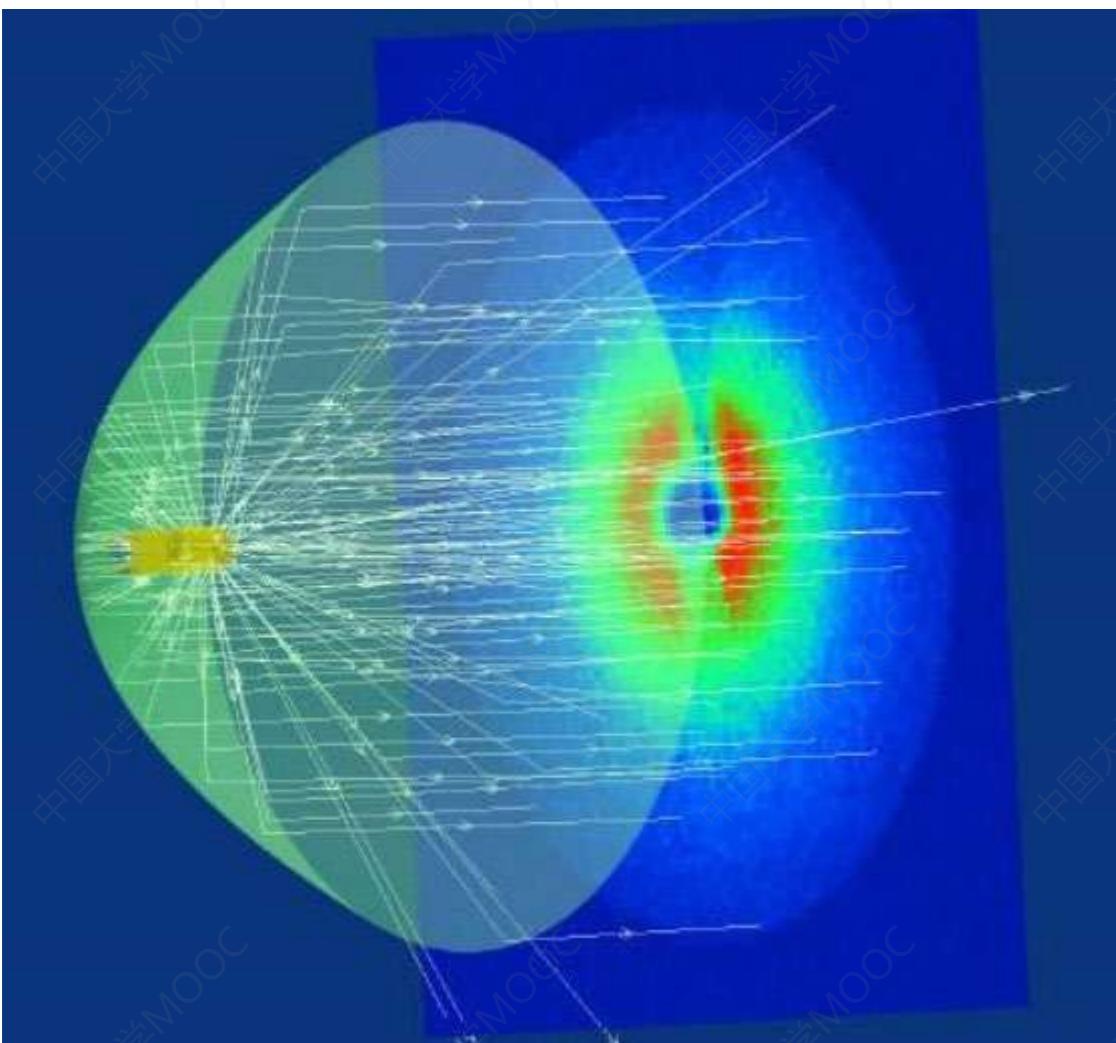
2012: AlexNet (多伦多大学)
2013: OverFeat
2014: GoogleNet (Google)
/VGG
2015: ResNet(微软)
2016: 公安部三所/ResNeXt;
2017: SENet

■ 1.3 图像处理技术的起源和发展



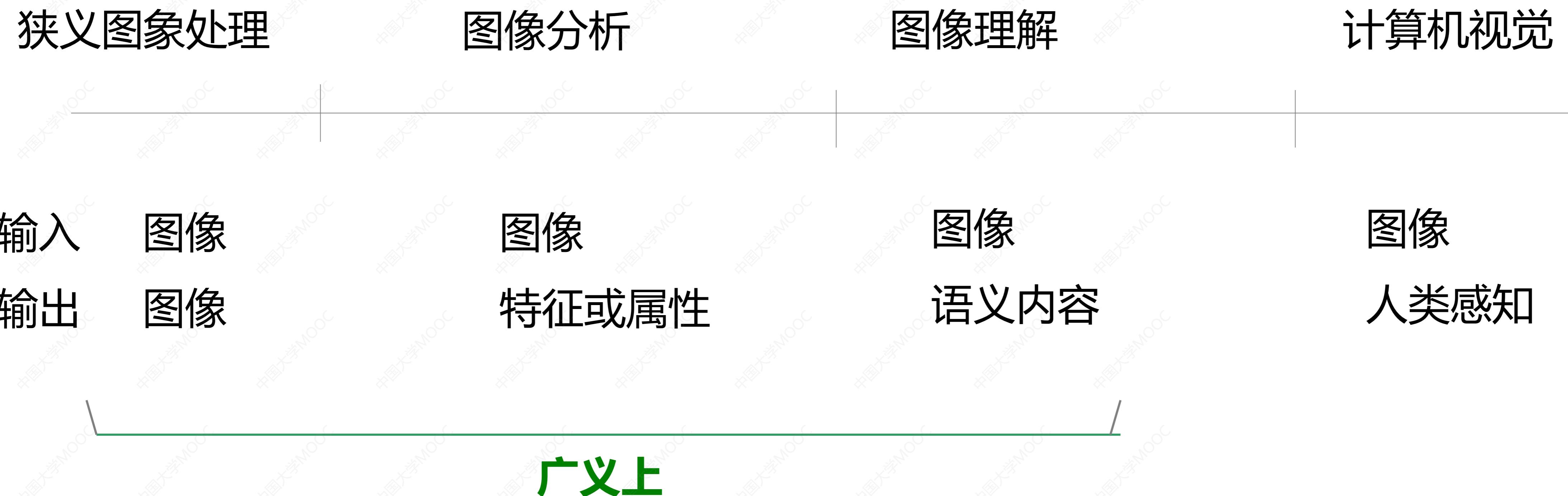
- 1) 深度学习在视觉分析任务中好用，推动了一些任务从实验室研究走到应用；如车牌识别、人脸识别
- 2) 深度学习缺乏可解释性，我们的努力方向。

1.4 数字图像处理 核心技术

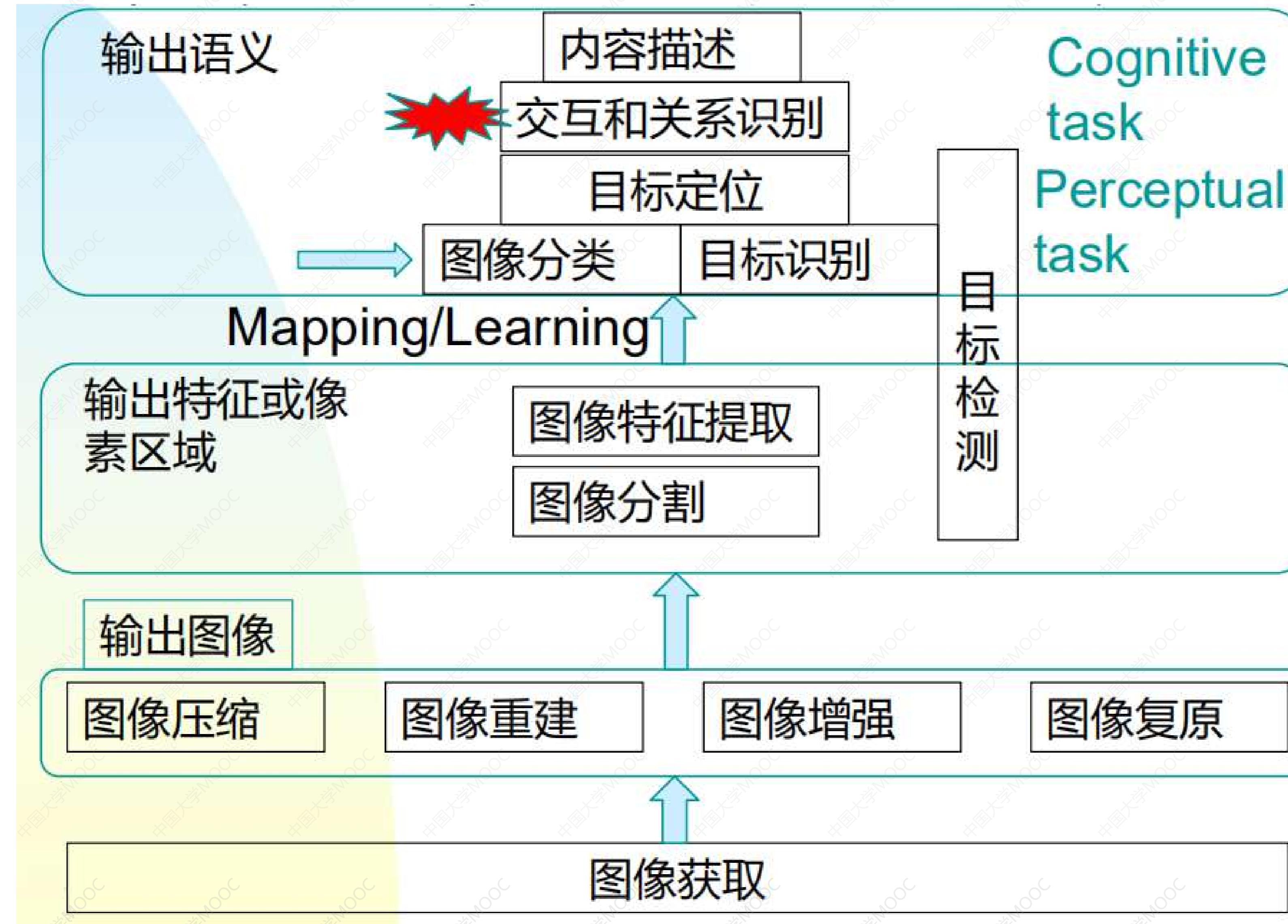


1.4 数字图像处理核心技术

• 范围



■ 1.4 数字图像处理核心技术



图像理解

图像分析

狭义图像处理

1.4 数字图像处理核心技术

图像获取 (Image Acquisition)

利用图像采集设备获取一幅图像。

模拟图像：视觉传感器 ——CCD、CMOS；

数字图像：视觉传感器+数字化器。数字化器包括采样和量化两部分

图像增强 (Image enhancement)

面向某一目标对图像感兴趣的部分进行处理，以改善图像的视觉质量。

面向目标，不考虑降质原因（基于人类主观偏好）设计方法；

方法设计更主观

处理结果图像存在局部优化的特点

图像复原 (Image restoration)

对于降质图象去除其降质因素，尽可能恢复图像的本来面目。

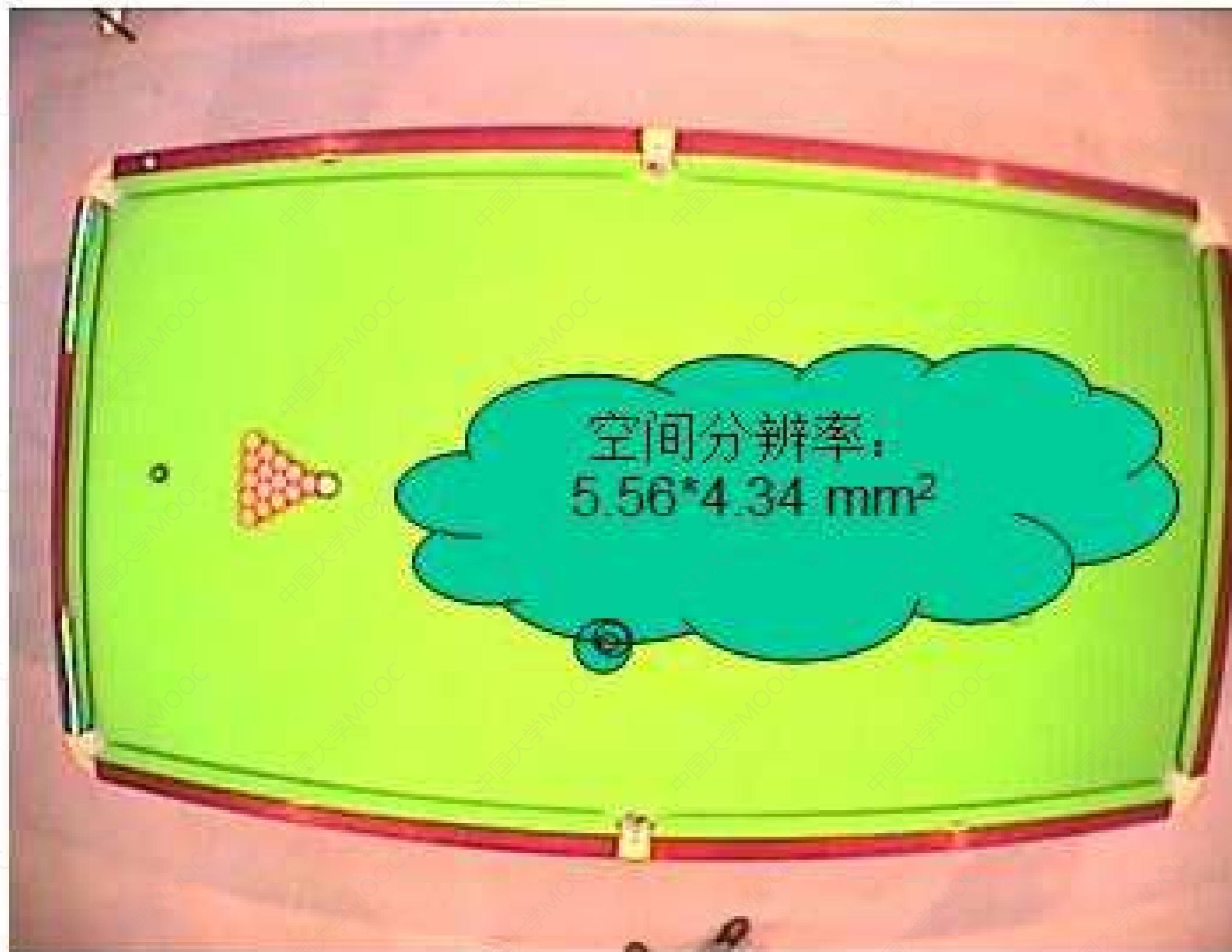
面向原因，对图像降质过程求逆。

方法设计更客观

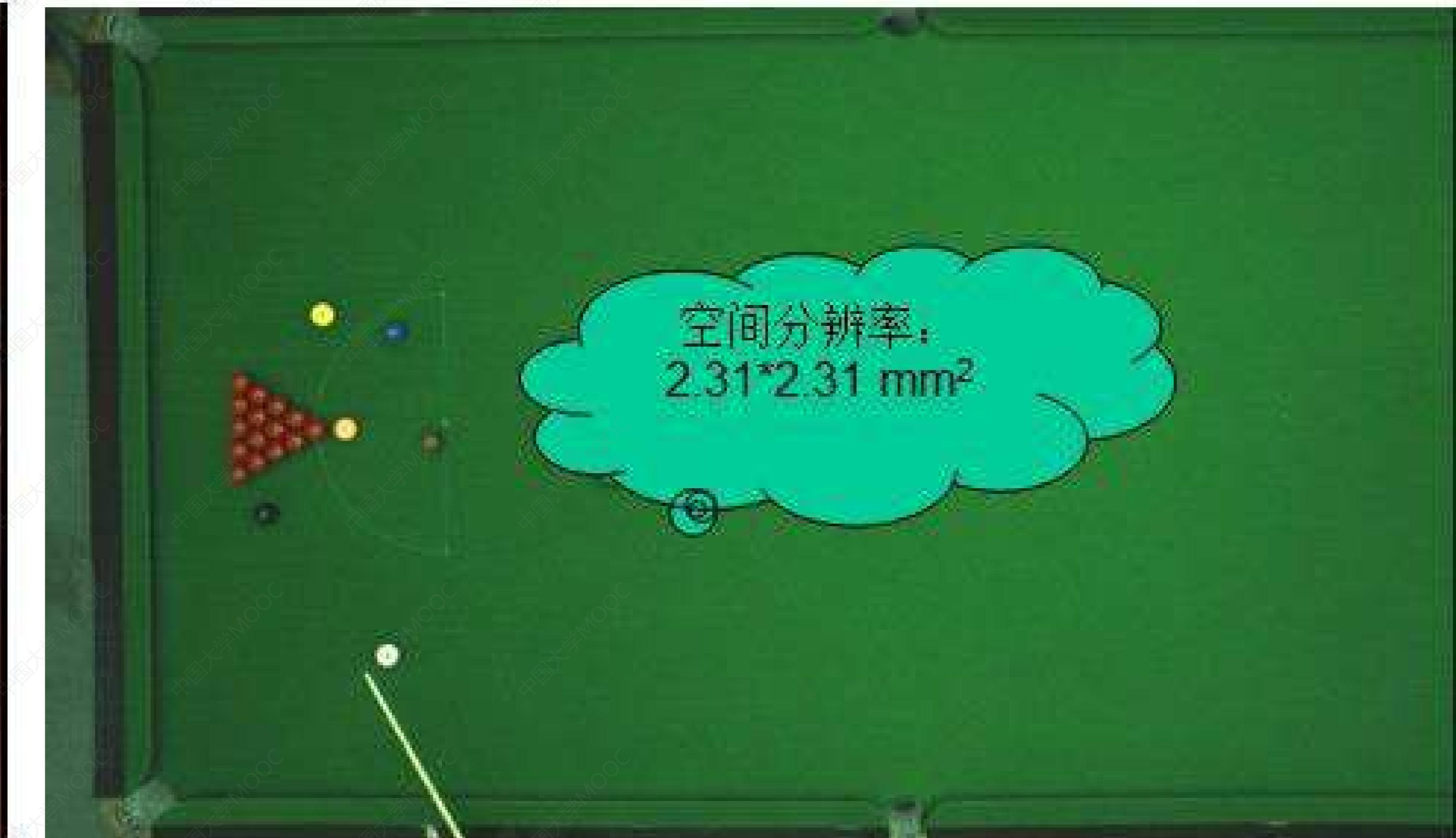
处理结果图像存在全局优化的特点

■ 1.4 数字图像处理核心技术

普通摄像头+镜头+C30



AC1300+Computar 5mm+ RGB21A



■ 1.4 数字图像处理核心技术



标准图像



运动模糊图像

■ 1.4 数字图像处理核心技术



边缘增强图像



差分复原图像

1.4 数字图像处理核心技术

图像获取 (Image Acquisition)

利用图像采集设备获取一幅图像。

模拟图像：视觉传感器 ——CCD、CMOS；

数字图像：视觉传感器+数字化器。数字化器包括采样和量化两部分

图像增强 (Image enhancement)

面向某一目标对图像感兴趣的部分进行处理，以改善图像的视觉质量。

面向目标，不考虑降质原因（基于人类主观偏好）设计方法；

方法设计更主观

处理结果图像存在局部优化的特点

图像复原 (Image restoration)

对于降质图像去除其降质因素，尽可能恢复图像的本来面目。

面向原因，对图像降质过程求逆。

方法设计更客观

处理结果图像存在全局优化的特点

核心问题：降质过程建模是关键设计过程，式逆过程

1.4 数字图像处理核心技术

图像重建 (Image reconstruction)

图象重建即由投影数据重建图象的技术。对于一些三维（二维）物体其内部内容未知，通过其他手段获取其投影数据，由投影数据重建出三维数据（二维图像）。

典型应用：

CT (Computerized Tomography)、工业无损检测、MRI (Magnetic Resonance Imaging 磁共振成像) 等

成像原理：

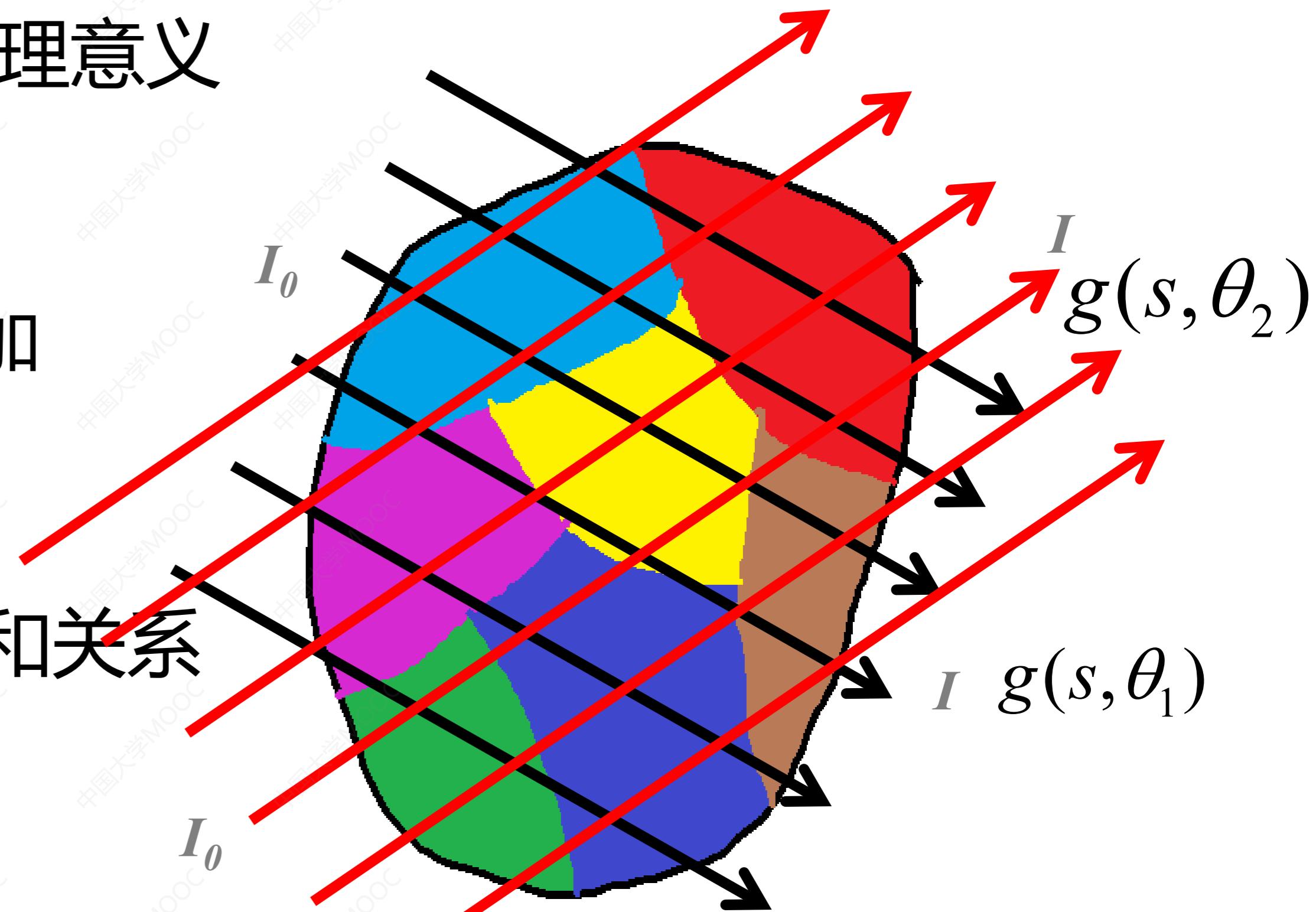
CT、工业无损检测等由X射线获取投影数据

MRI通过对静磁场中的人体施加某种特定频率的射频脉冲，使人体中的氢质子受到激励而发生磁共振现象。停止脉冲后，质子在弛豫过程中产生MR信号。通过对MR信号的接收、空间编码和图像重建等处理过程，即产生人体切片图像。

■ 1.4 数字图像处理核心技术

- CT技术的原理（图像重建），投影数据的物理意义
 - ✓ 不同组织对X射线的衰减系数不同
 - ✓ 一条射线的总衰减量是不同组织衰减量的累加
 - ✓ 投影数据与组织之间呈线性积分关系
- 投影数据与不同组织衰减系数之间是线性求和关系
- 由投影重建图像的本质是解线性联立方程组

$$g(s_i, \theta_j) = \ln(I / I_0) = \sum_{(x,y) \in s_i} f(x, y)$$



1.4 数字图像处理核心技术

图像压缩 (Image compression)

图象压缩又称图象编码，即去除图象冗余，减少图像表示的数据量，以便节省图象传输、处理时间及存储容量。

图象压缩包括有损压缩编码 (Lossy) 和无损压缩编码(Lossless)两大类。

无损压缩：可以无失真的重构原始图象，但压缩比较低，一般为2-3；通常具有重要的保存价值的图象如医学图象、遥感图象等采用无损压缩。

有损压缩：重构图象与原始图像存在误差，但很多时候视觉感知上与原始图像差别不大。有损压缩的图像压缩比较高，日常生活图像多为有损压缩图像。

图像压缩标准：JPEG, JPEG LS, JPEG2000

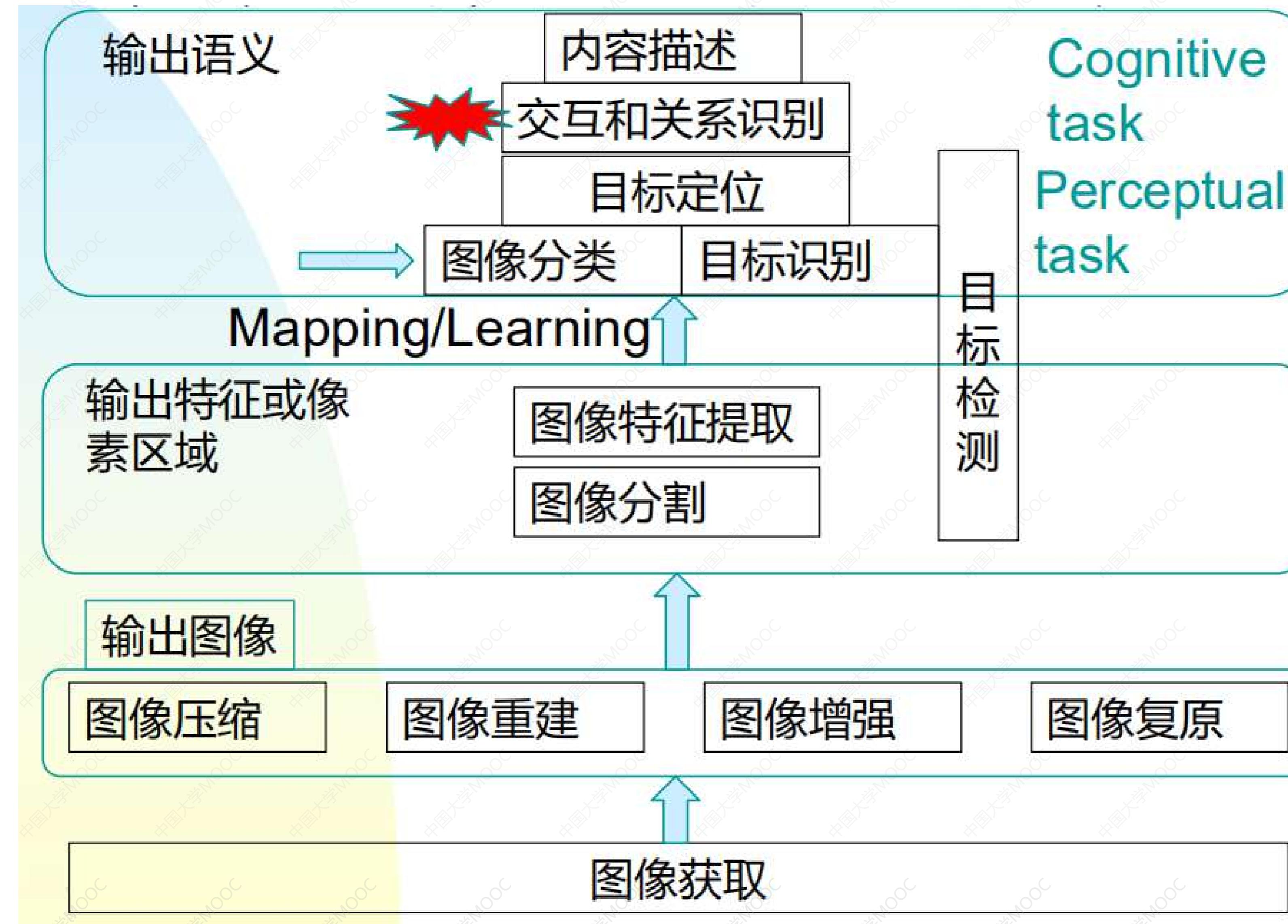
1.4 数字图像处理核心技术

看图说话的基本范式：**对象在**场景下做***事。对象识别、场景识别、关系识别



一个老爷爷和一个小朋友在西瓜地里下象棋；
地里有很多西瓜；
一只鸡妈妈带着一群小鸡；
一只小鸡和大家不在一起；
天空中飘着三只气球；
还有一只蝴蝶在飞。

■ 1.4 数字图像处理核心技术



图像理解

图像分析

狭义图像处理

1.4 数字图像处理核心技术

➤ 图像分割 (Image Segmentation)

把图像分成若干个特定的、具有独特性质的区域的技术。

图像分割是由狭义图像处理到图像分析理解的关键步骤。是像素级处理到高层语义理解的关键环节。



■ 1.4 数字图像处理核心技术

- 图像（对象/目标）特征提取 (Image(Object) Feature Extraction)
 - ✓ 对图像分割得到的每个区域提取其特征并形成特征表示向量。
- 常见特征
 - ✓ **形状特征**: 获取描述物体形状的特征空间并进行映射，常见有形状几何参数、圆度、傅立叶形状描述子、形状不变矩等；
 - ✓ **纹理特征**: 获取描述物体内部纹理的特征空间并进行映射，常见有灰度共生矩阵、LBP等
 - ✓ **颜色特征**: 获取描述物体颜色统计特性或颜色分布的特征空间并进行映射，常见有颜色直方图、主颜色、主颜色分布等；
 - ✓ **深度特征**: 通过深度学习获取描述物体的特征并进行映射；

1.4 数字图像处理核心技术

- 图像（对象）识别/分类 (Image(Object) Recognition)

✓ 根据区域的特征通过学习将其映射为区域所对应的语义概念，是由图像底层特征到语义概念映射的过程，这是一个典型的模式识别问题；

✓ 通常的学习都是基于有限种类的目标进行学习，因此目前的目标识别任务的本质等价于目标分类；

✓ 传统层次化方法：分割/检测——》 特征提取——》 识别；

✓ 深度学习：端到端



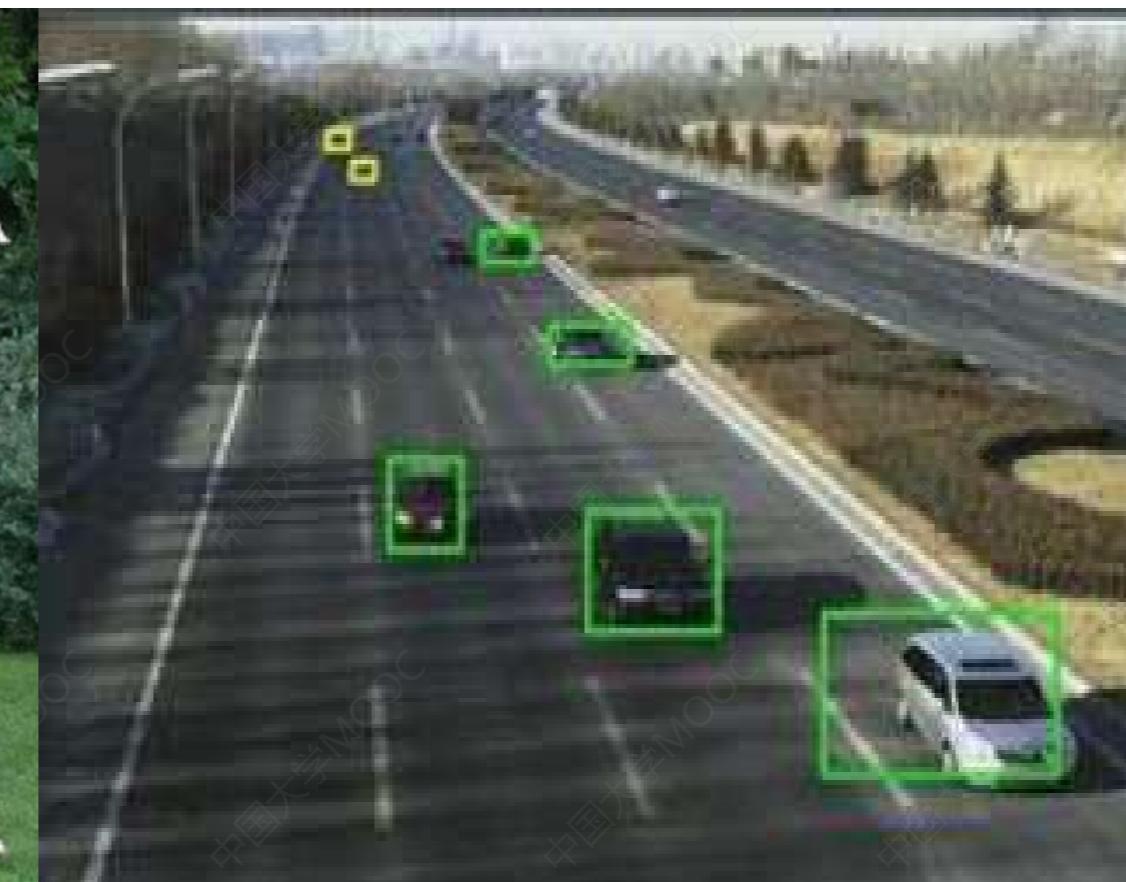
■ 1.4 数字图像处理核心技术

• 目标（对象）检测 (Object Detection)

- ✓ 自动判断图像中是否存在指定一个或多个类别对象，并在图片中标记其位置，目标检测是一个从像素到语义对象的过程。
- ✓ 目标检测需要从像素级特征到语义对象并给出其位置，因此存在一个由底层特征到高层语义的映射过程，某种程度上一次性完成了分割——》特征提取——》对象识别系列任务。



人脸检测



车辆检测

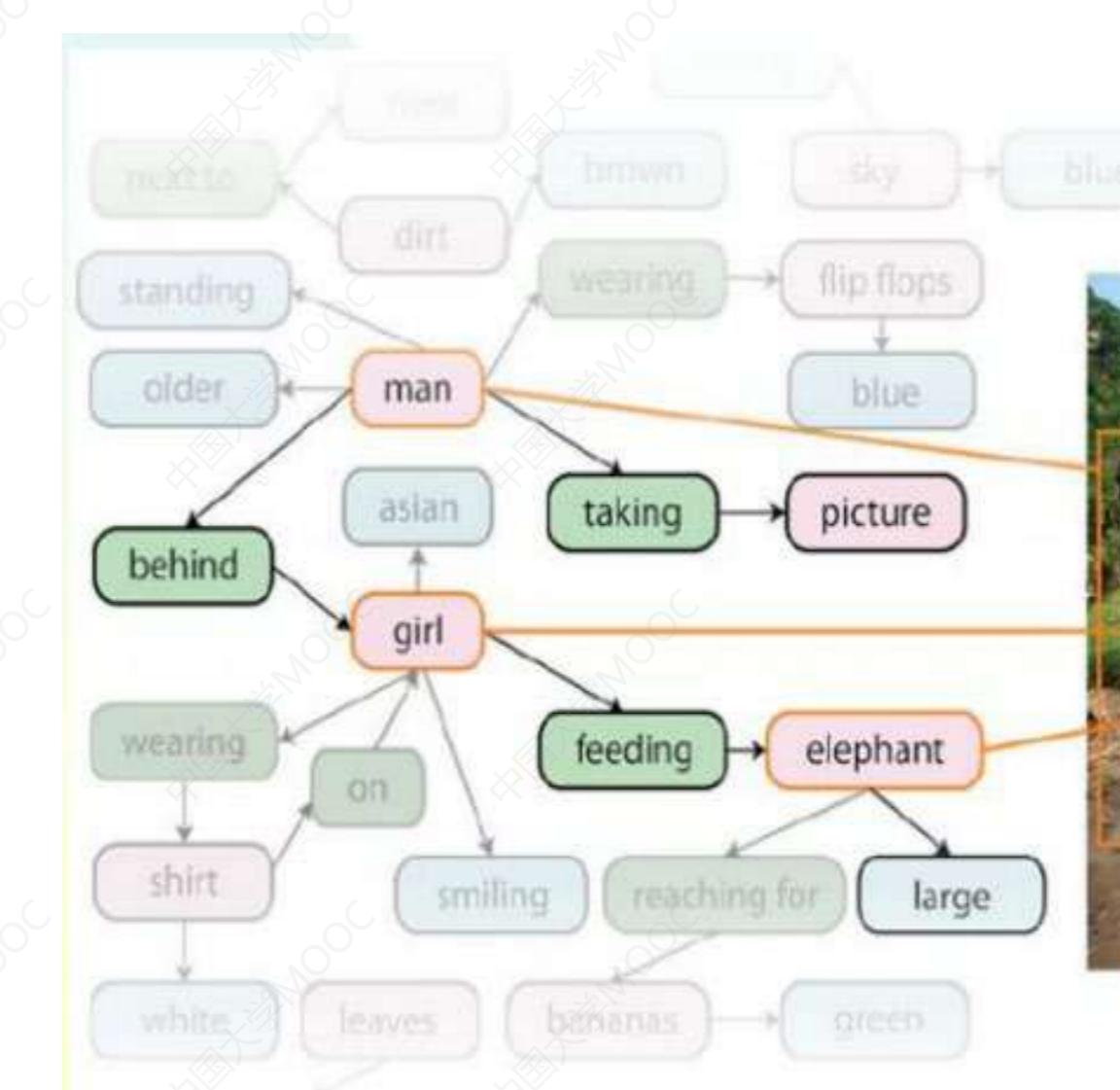


行人检测

■ 1.4 数字图像处理核心技术

• 视觉关系表达 (Visual Relation representation)

- ✓ 表示对象之间的空间位置关系、交互关系等。
- ✓ 位置关系挑战在于由二维图像估计三维空间的位置关系存在一定的歧义；
- ✓ 交互关系本质上是将两个对象用动词连接起来；
- ✓ 同样的两个对象可能有不同的交互关系；



■ 1.4 数字图像处理核心技术

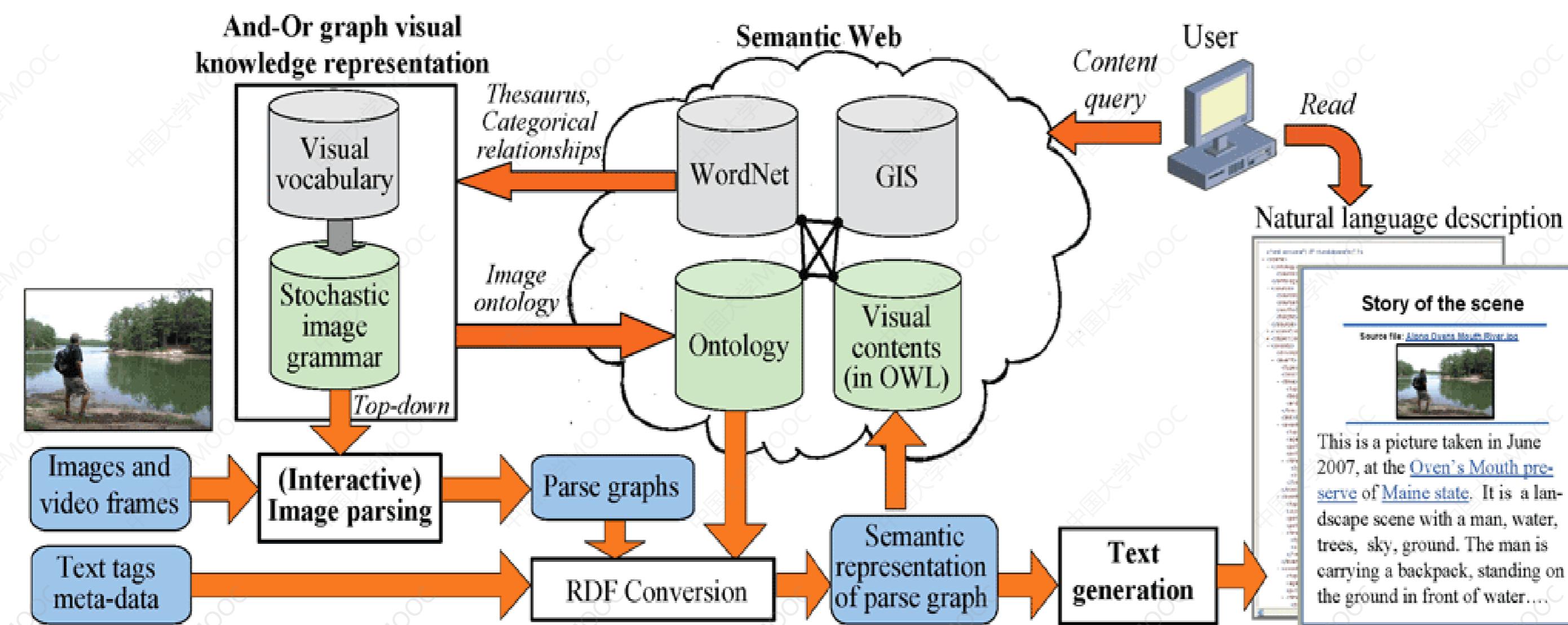
- 行为识别 (Action Recognition)
- 即基于对象之间的交互关系对对象的行为进行识别 (分类)
 - ✓ 有些行为有时域动态特征，因此基于视频才能完成；有些行为可以从一幅关键帧图像直接得到；
 - ✓ 行为包括个体行为和群体行为，个体行为和群体行为有关系也有区别；



■ 1.4 数字图像处理核心技术

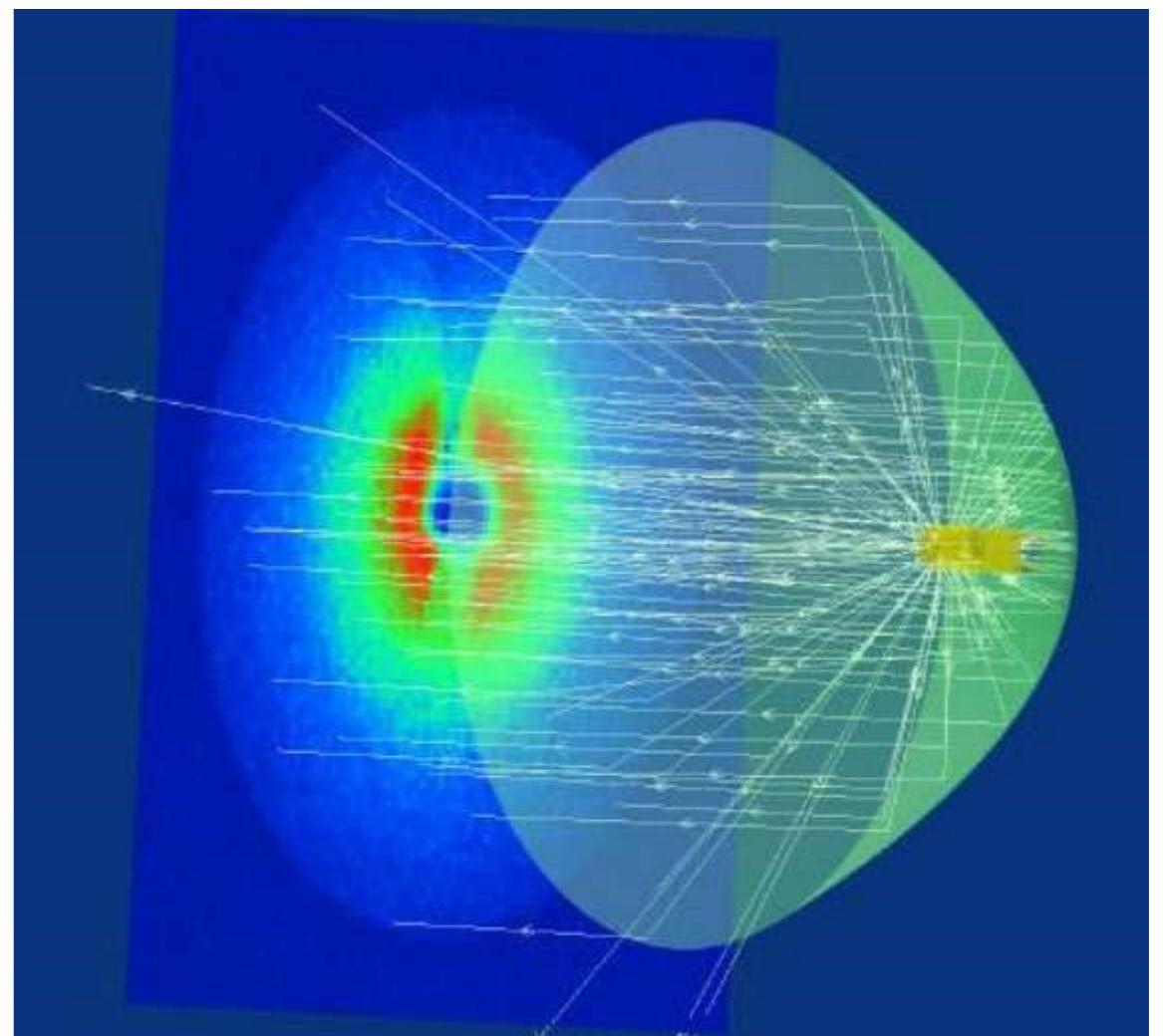
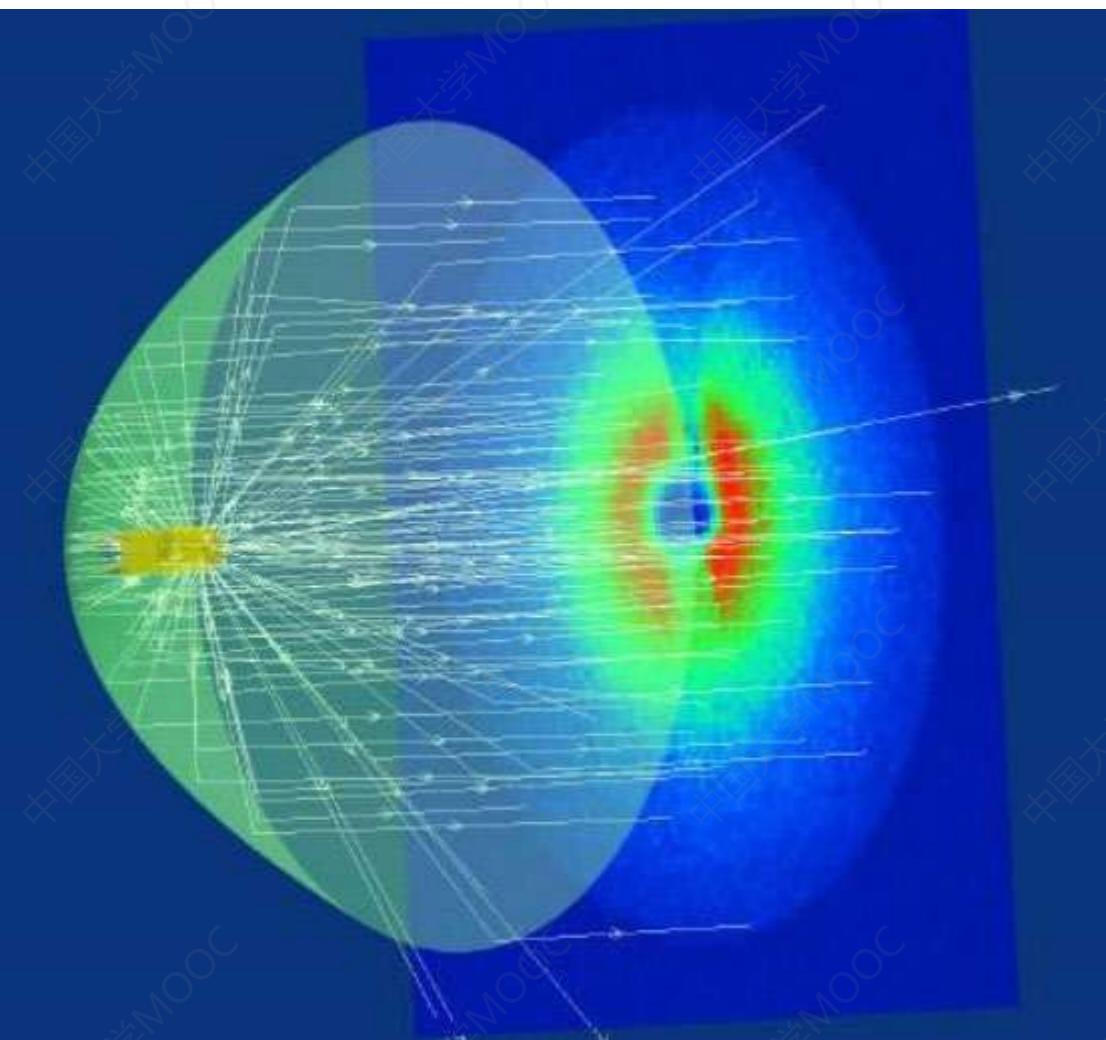
• 图像内容描述 (Image Caption)

- ✓ 对一幅输入图像自动生成一段语言文字来描述图像内容，类似于看图说话。
- ✓ 图像内容描述是图像理解的最高境界，其包含了对象空间位置关系描述，对象交互关系描述、对象行为描述，而且还要注意描述语言的流畅性，语句之间的转折连贯关系等,因此是图像理解和自然语言处理的交叉与综合



第二章

数字图像处理基础



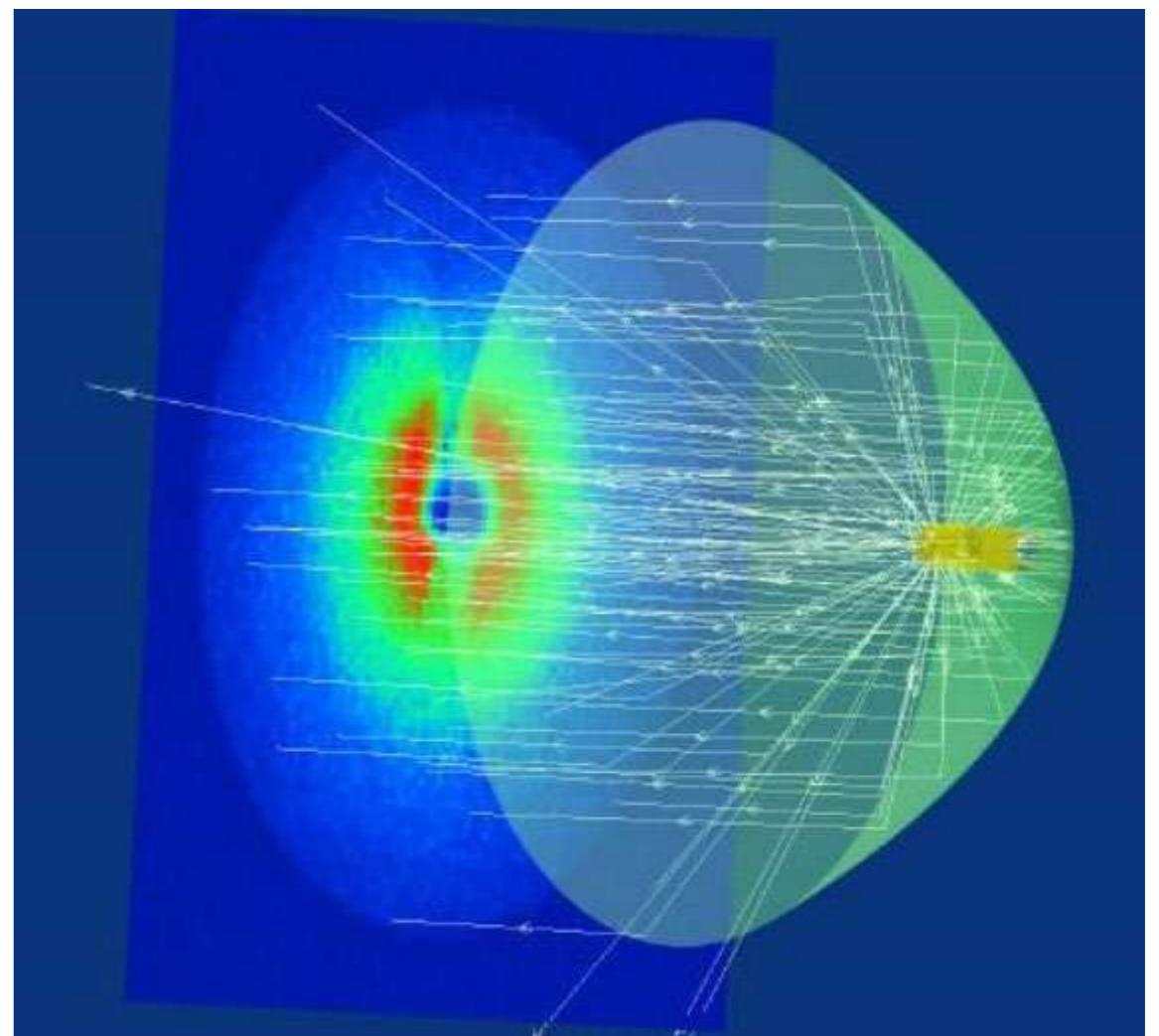
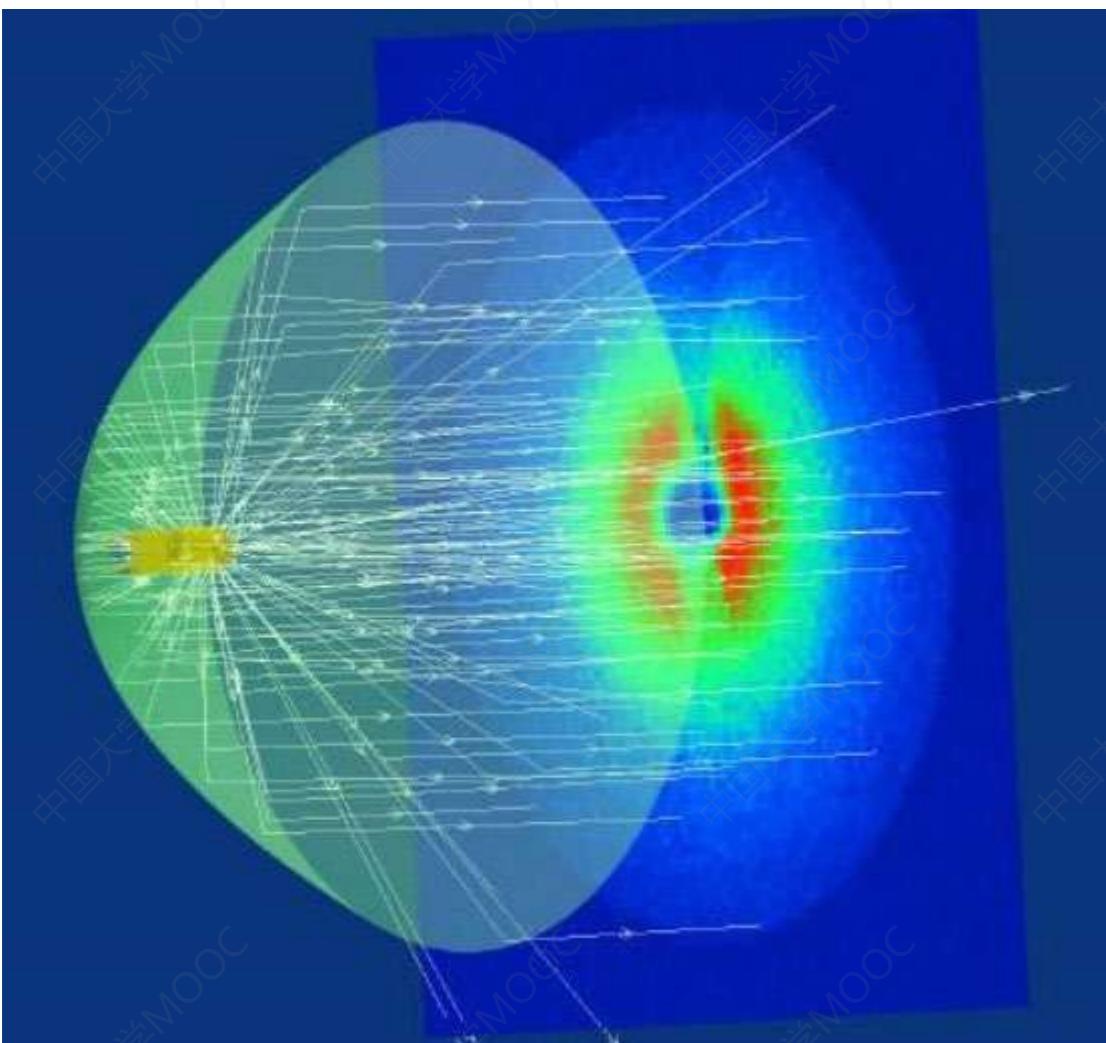
目录

- 2.1 人类视觉系统
- 2.2 视觉特性
- 2.3 成像和数字化
- 2.4 图像中的常用术语
- 2.5 图像质量评价

目录

- 2.1 人类视觉系统
- 2.2 视觉特性
- 2.3 成像和数字化
- 2.4 图像中的常用术语
- 2.5 图像质量评价

2.1 人类视觉系统



2.1 人类视觉系统

Visual pathway

Lateral Geniculate
Nucleus (外侧膝状
核 (LGN))

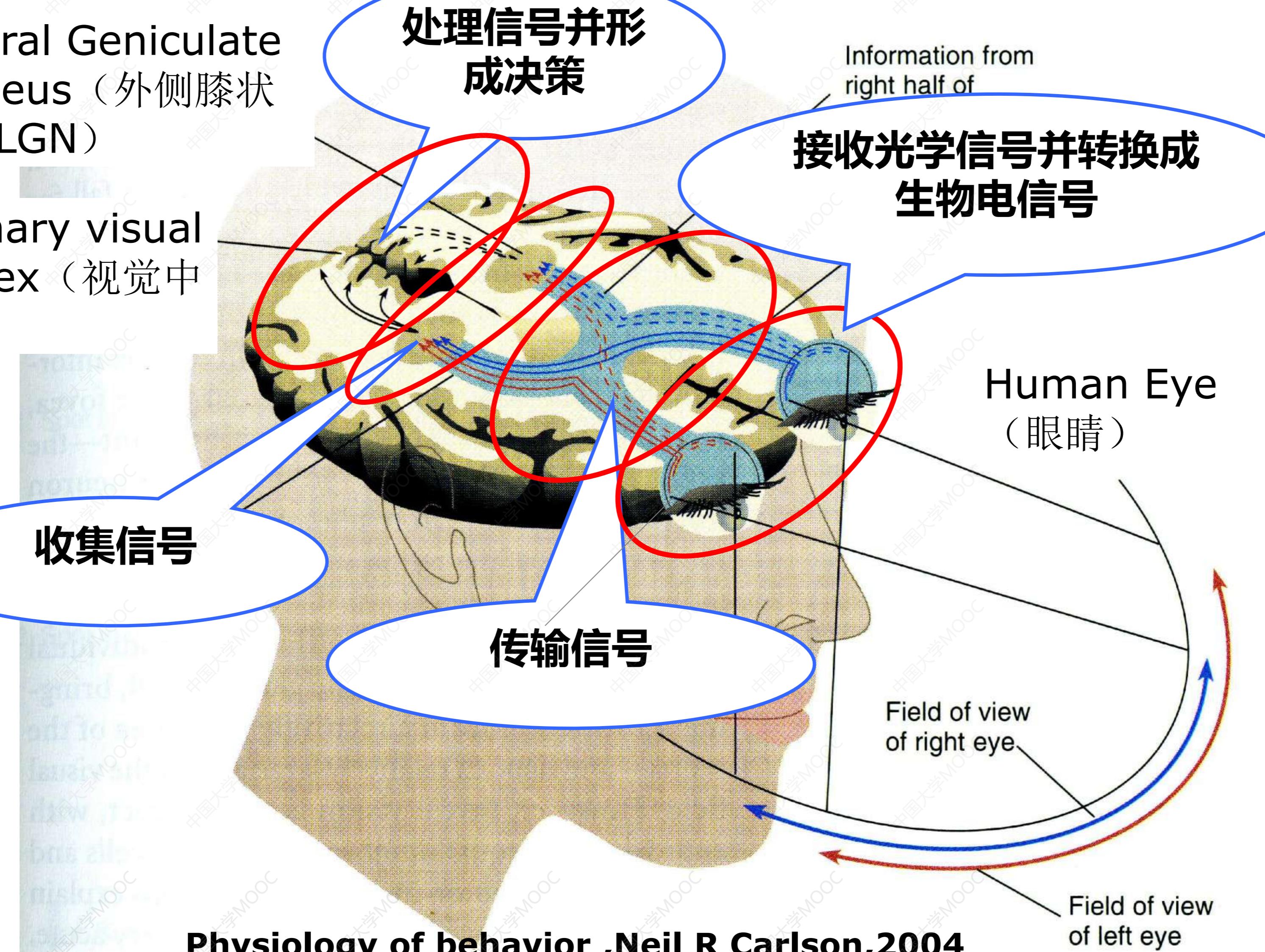
Primary visual
cortex (视觉中
枢)

处理信号并形
成决策

Information from
right half of

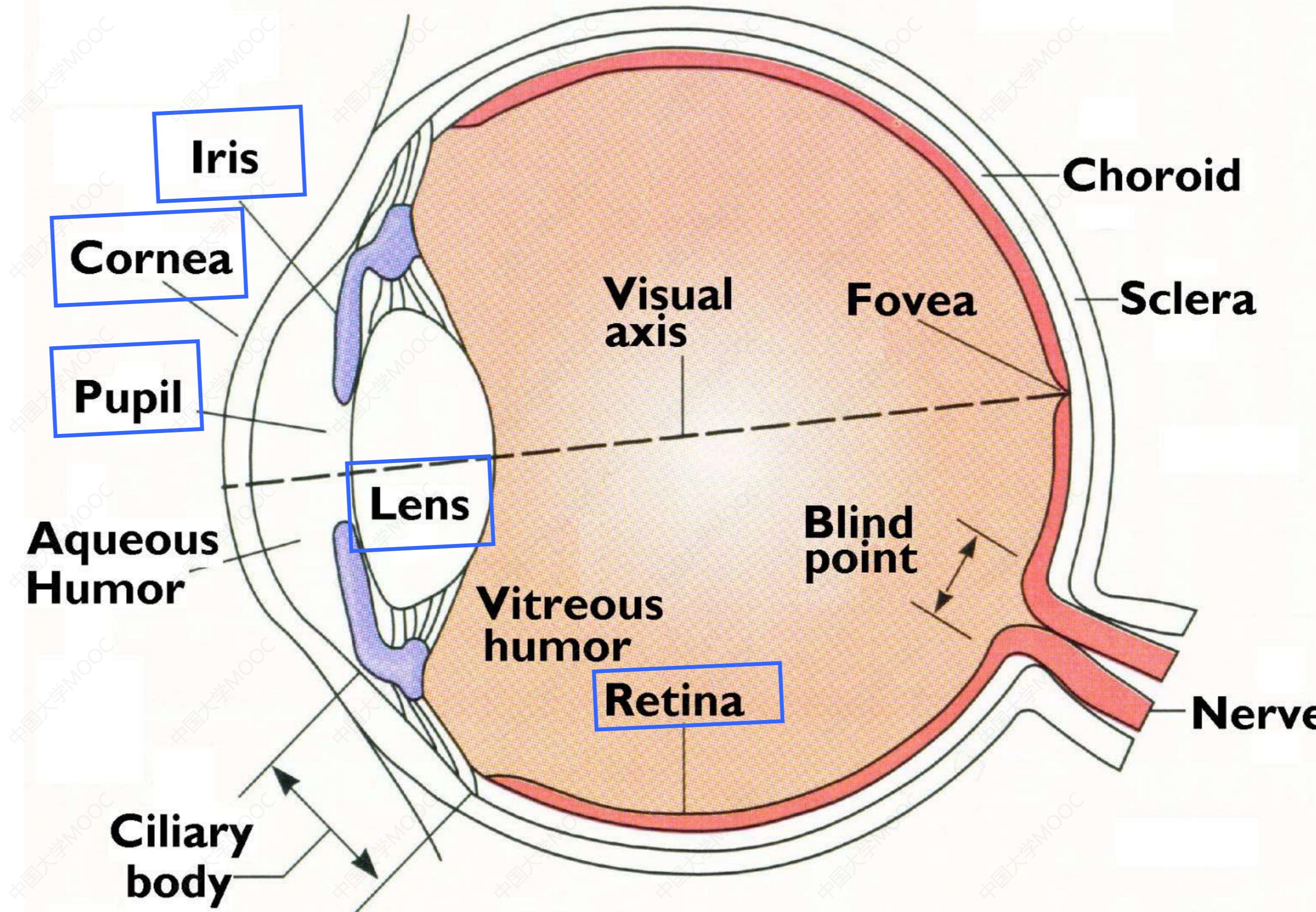
接收光学信号并转换成
生物电信号

Human Eye
(眼睛)



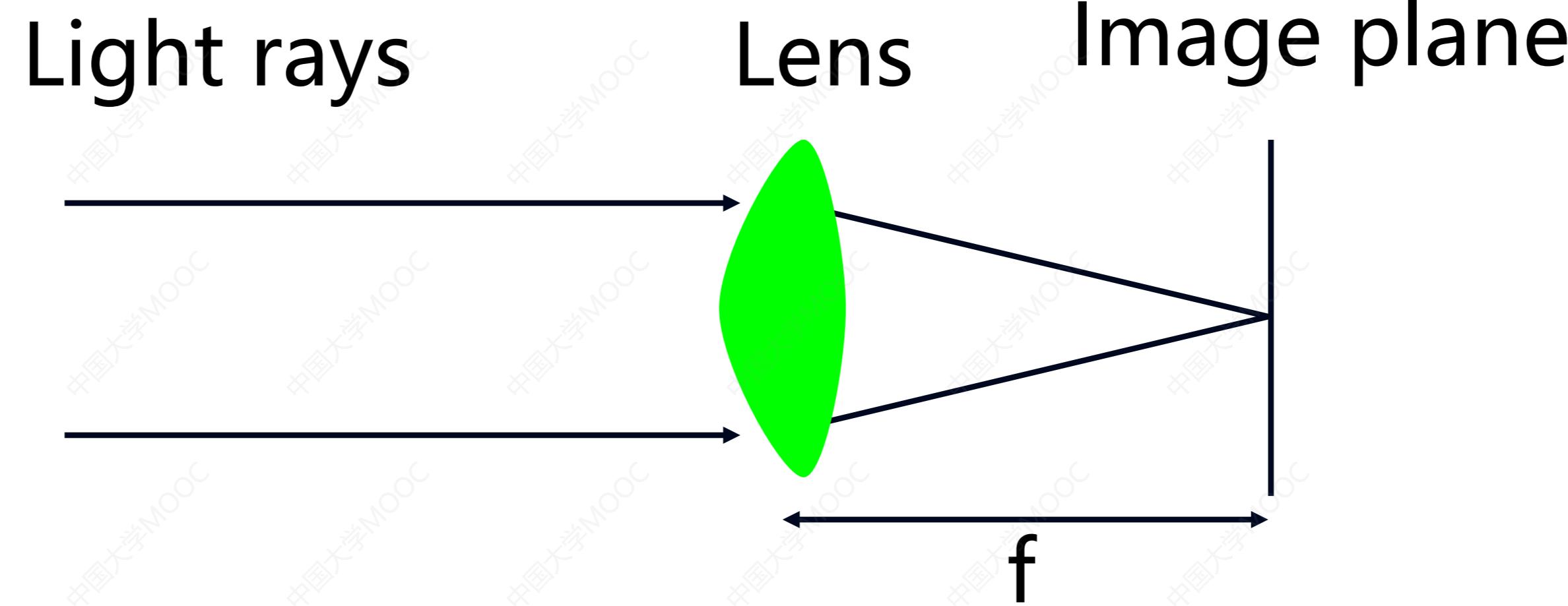
2.1 人类视觉系统

Structure of the eye

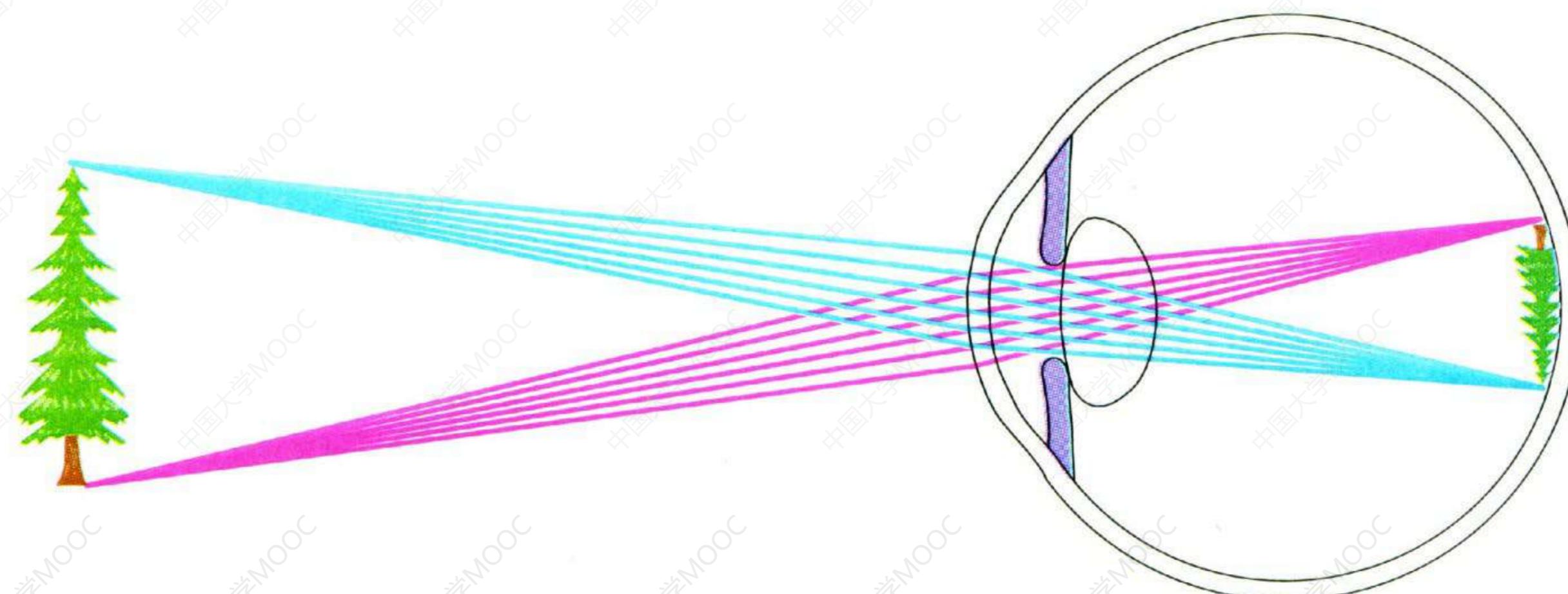


2.1 人类视觉系统

光学成像

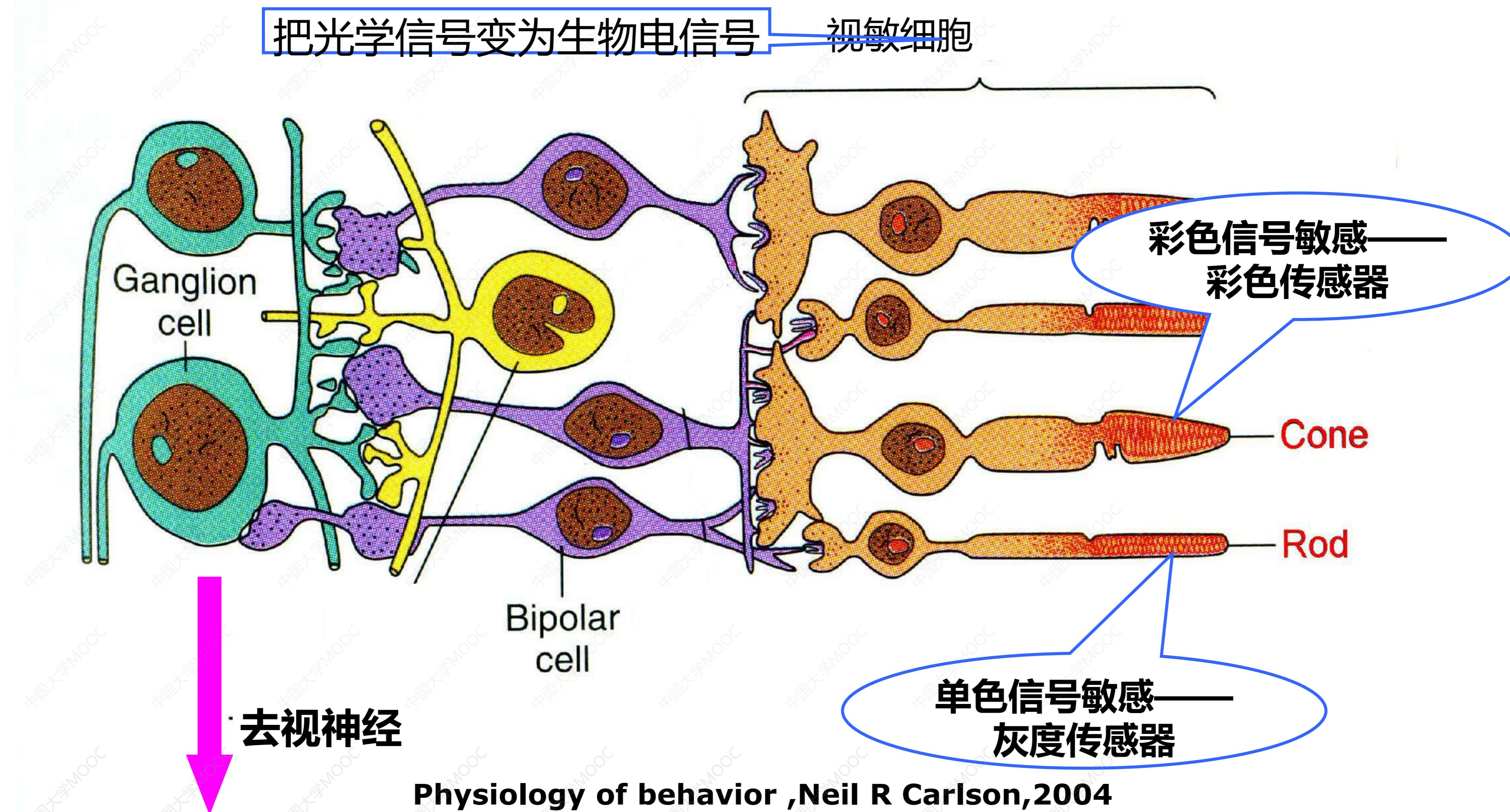


Theory of lens



2.1 人类视觉系统

视网膜处理



■ 2.1 人类视觉系统



锥状细胞起作用



杆状细胞起作用

■ 2.1 人类视觉系统

杆状细胞工作的光通量范围： 10^{-6} - 10^0 ML

锥状细胞工作的光通量范围： 10^{-3} - 10^4 ML

1毫朗伯 (ML) = 10流明 (lm)

照度单位：勒克斯 (lx) = 流明每平方米 (lm/m^2)

晴天室外环境： 10^3 - 10^5 lx——锥状细胞工作；

晴天室内环境： 10^2 - 10^3 lx——锥状细胞工作；

阴天室外环境：50-500 lx——锥状细胞工作；

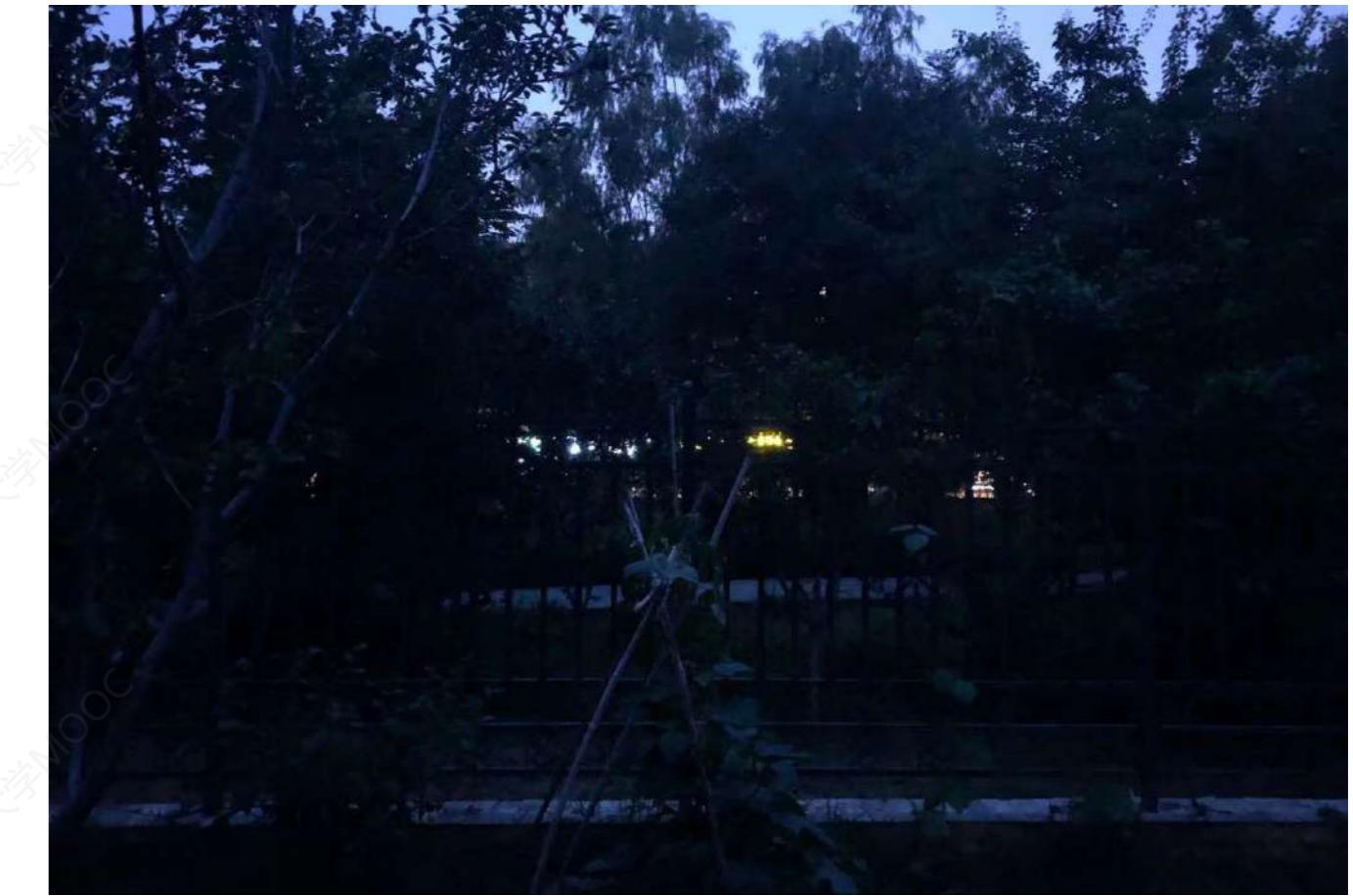
阴天室内环境：5-50 lx——锥状细胞工作；

满月夜室外环境：0.03-0.3 lx——锥状细胞和杆状细胞同时工作；

无月夜室外环境：0.001-0.02 lx——杆状细胞工作；

一般室内环境：50-100 lx——锥状细胞工作；

阅读环境：300 lx——锥状细胞工作。



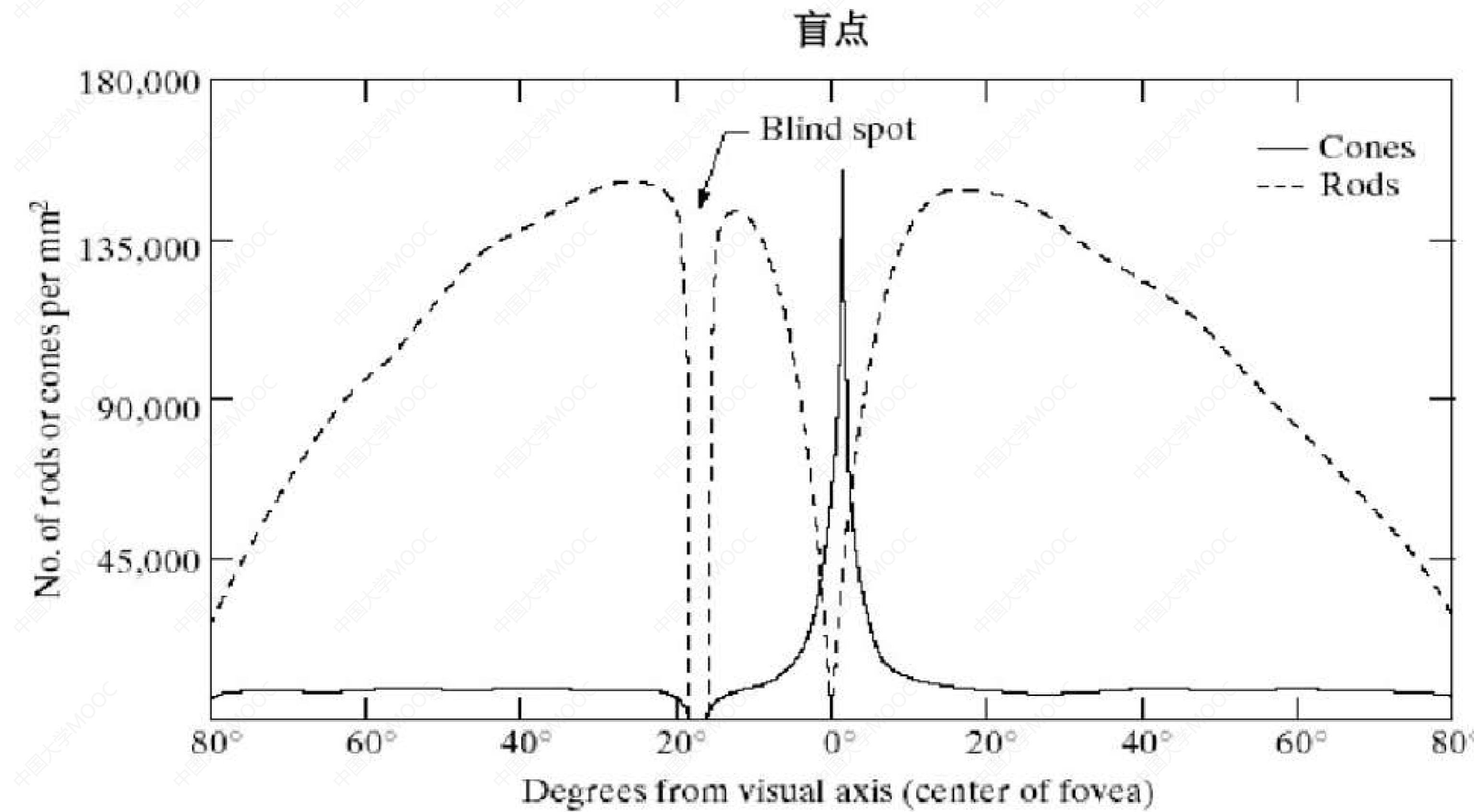
杆状细胞工作场景示例



锥状细胞工作场景示例



2.1 人类视觉系统



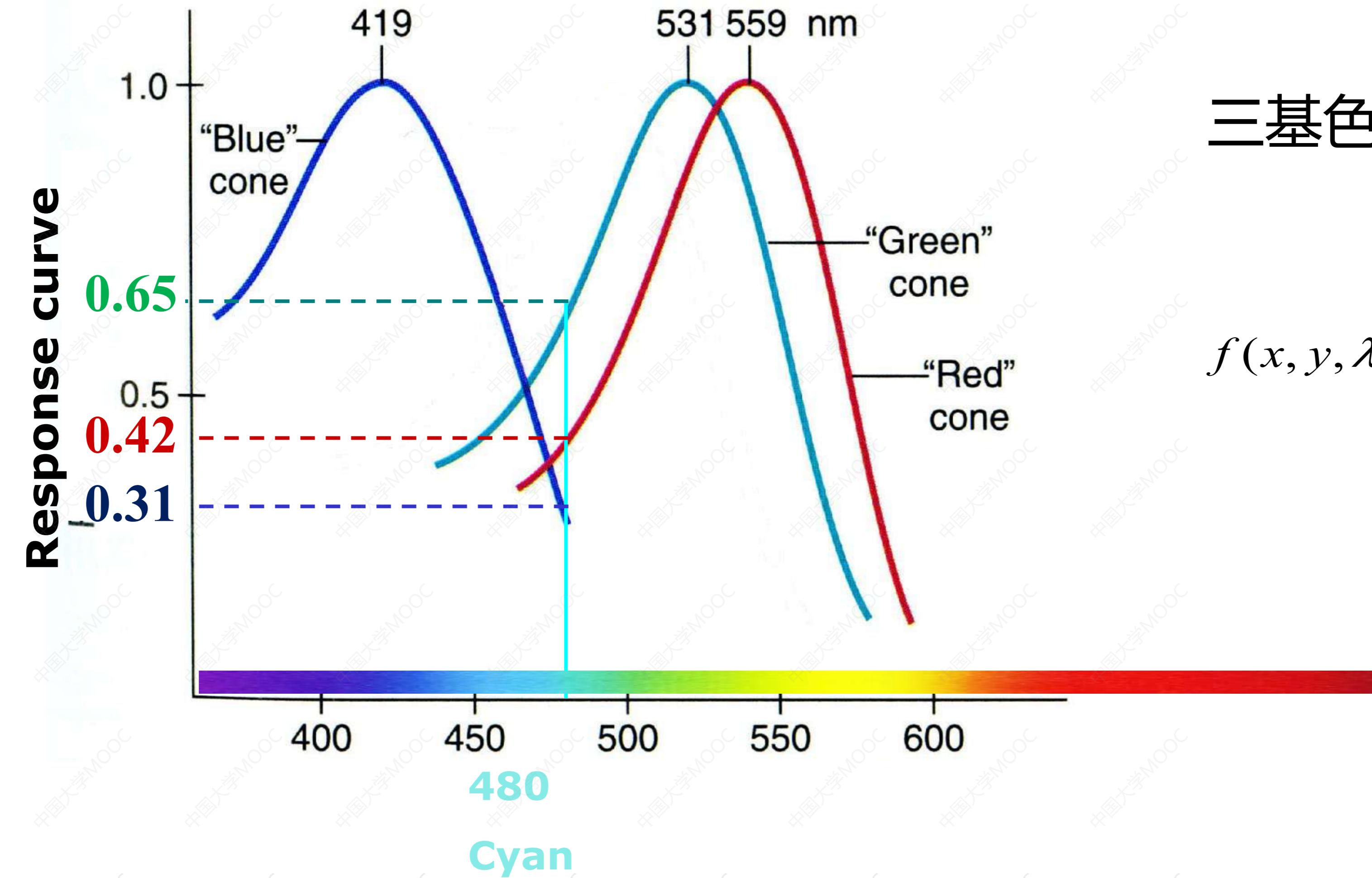
杆状细胞和锥状细胞在视网膜上的分布

2.1 人类视觉系统

表 2.2 杆状细胞与锥状细胞特性比较

	锥状细胞	杆状细胞
细胞数量 (万)	600-700	7600-15000
敏感因素	彩色	亮度
视觉	明视觉	暗视觉
分辨力	高	低
分布位置	黄斑区	整个视网膜

2.1 人类视觉系统



三基色原理

$$f(x, y, \lambda) \Leftrightarrow \begin{cases} R(x, y) \\ G(x, y) \\ B(x, y) \end{cases}$$

光度学基本知识

- 《数字图像处理学》，阮秋琦，电子工业出版社
 - 图像 (image) : 以某一技术手段被再现于二维画面上的视觉信息
 - 静止的、平面的、彩色的图像表示为 $f(x, y, \lambda)$

光度学基本知识

➤ 相对视敏函数

- ✓ 在辐射功率相同的情况下，不同的光不仅给人以不同的色彩感觉，而且给人以不同的亮度感觉。
- ✓ 在获得相同的亮度感觉的前提下，测量不同波长的辐射功率为 $P(\lambda)$ ，则视敏度函数 $k(\lambda)=1/P(\lambda)$

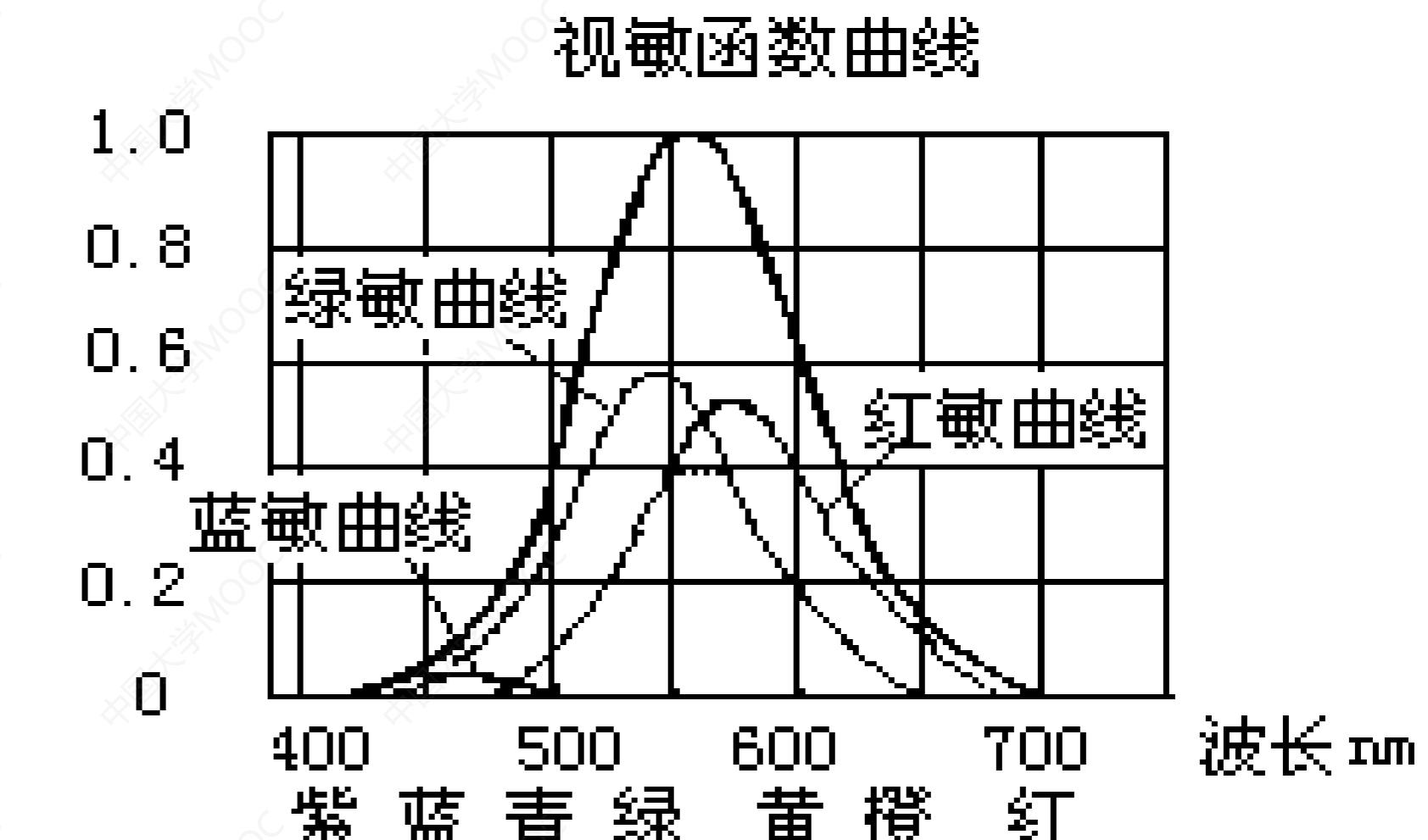
光度学基本知识

$k(\lambda)$ 越大，人眼对 λ 光越敏感

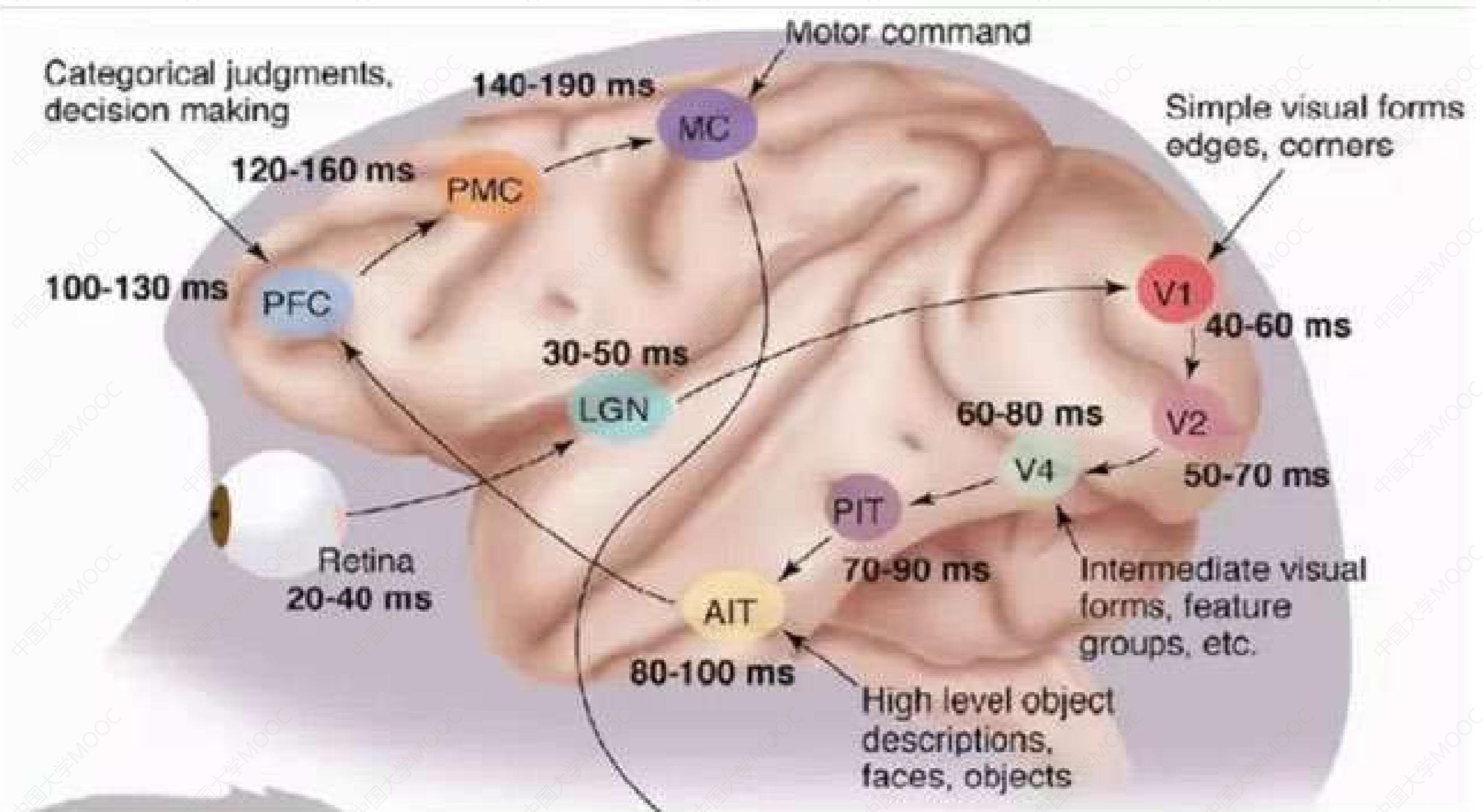
$$k_{\max} = k(555)$$

相对视敏函数

$$V(\lambda) = k(\lambda) / k(555)$$

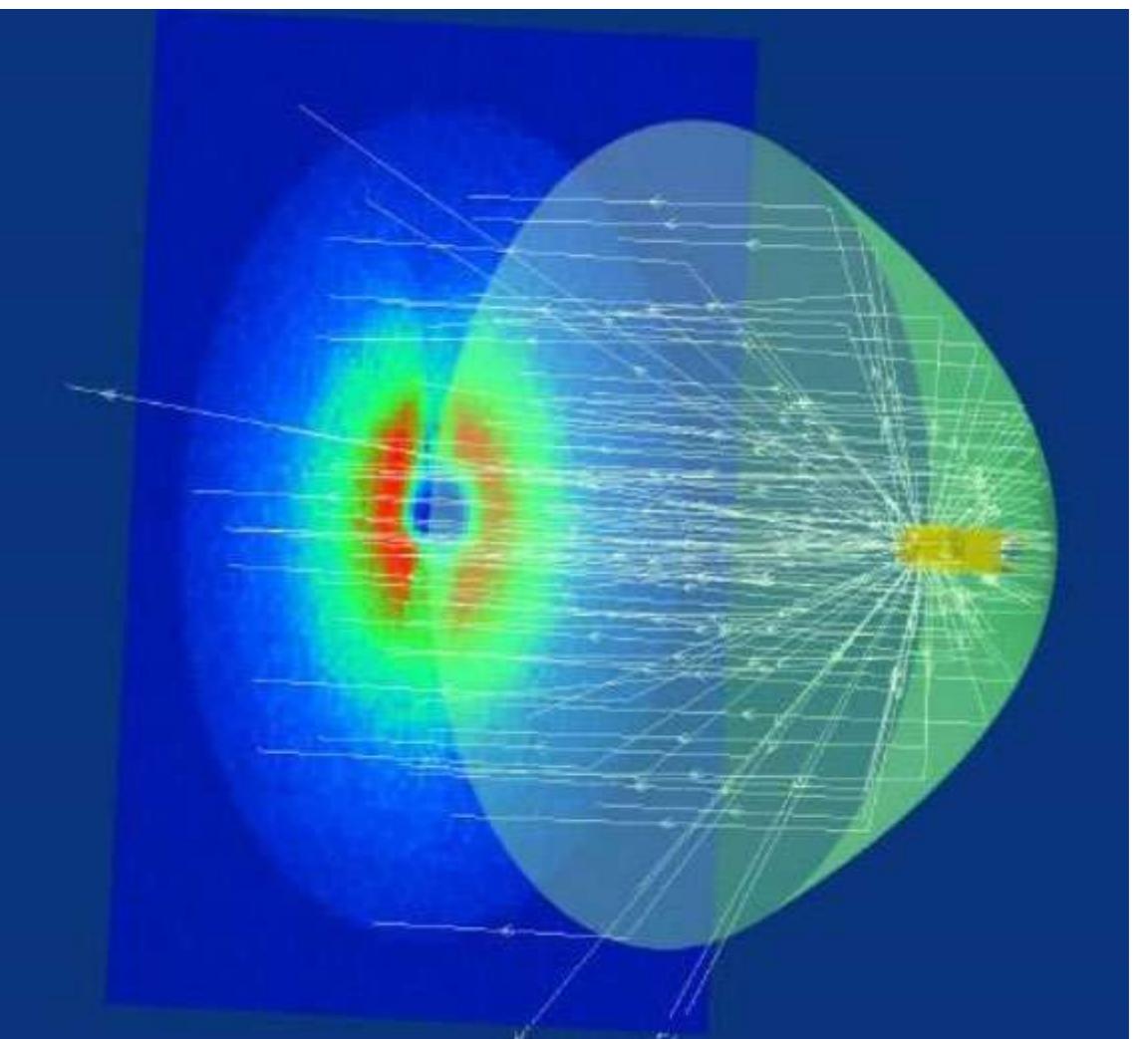
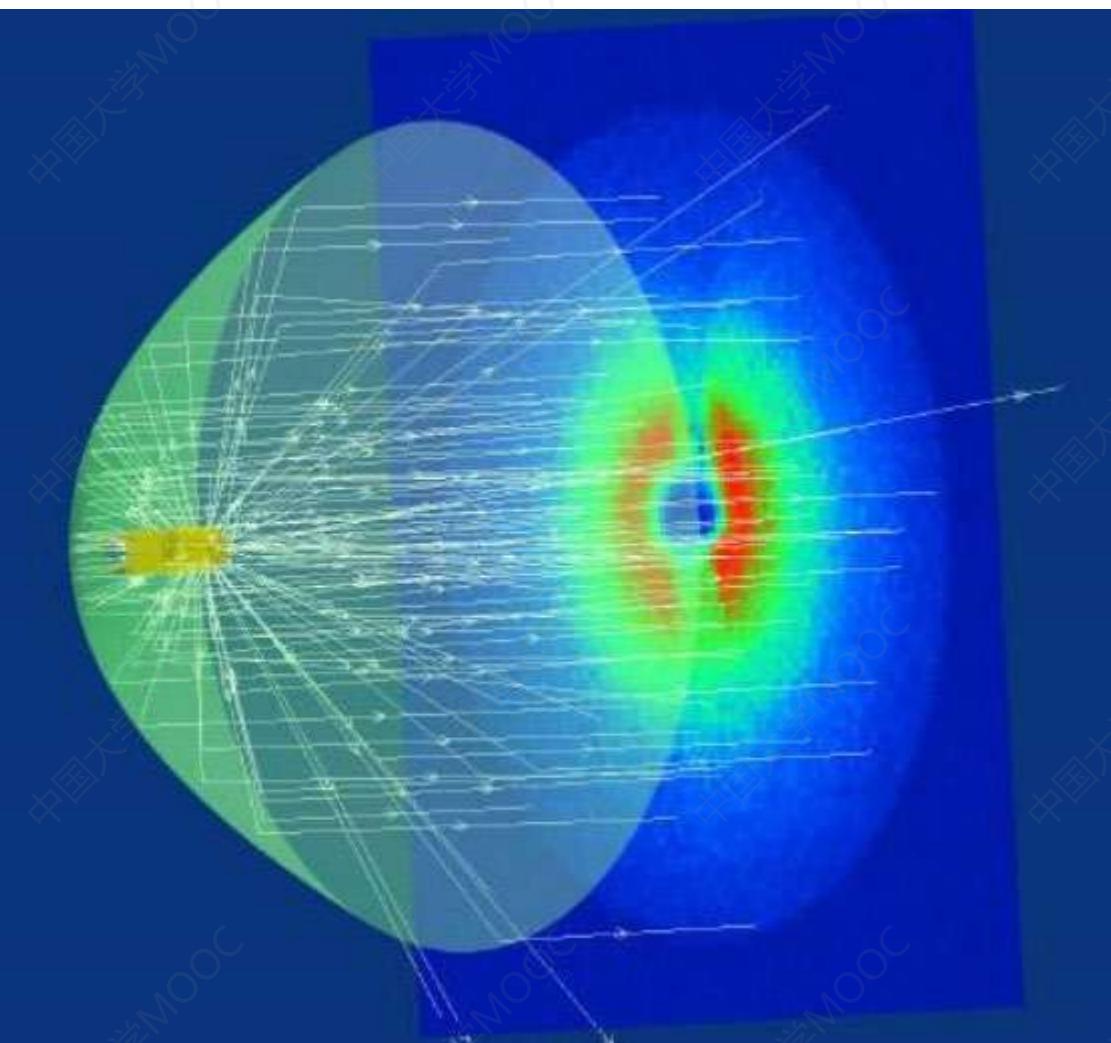


人类视觉组成



- 1981 年的诺贝尔医学奖，颁发给了 David Hubel (出生于加拿大的美国神经生物学家) 和 Torsten Wiesel, 以及 Roger Sperry。前两位的主要贡献是“发现了视觉系统的信息处理”：可视皮层是分级的

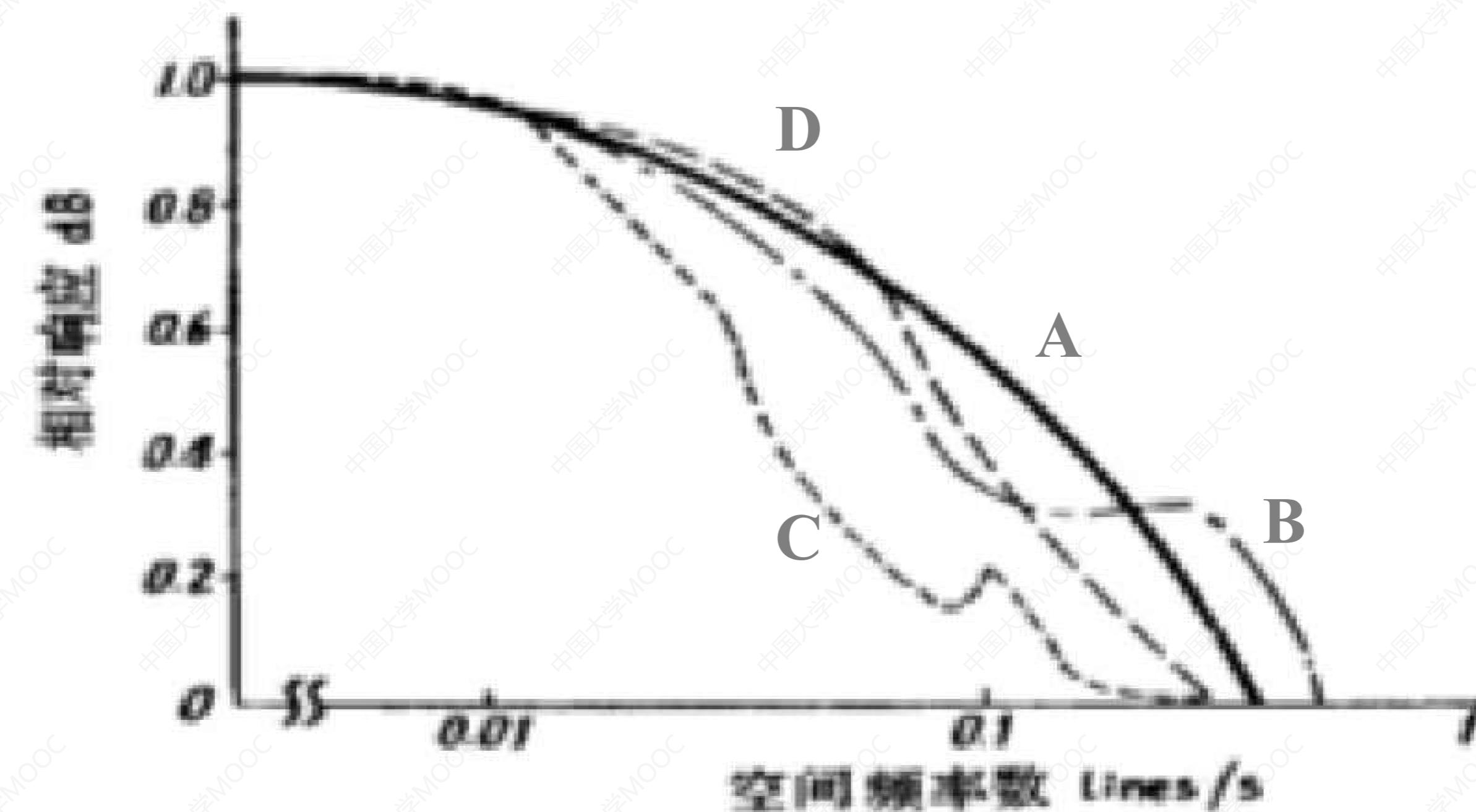
2.2 视觉特性



■ 2.2 视觉特性

1. 空间频率特性

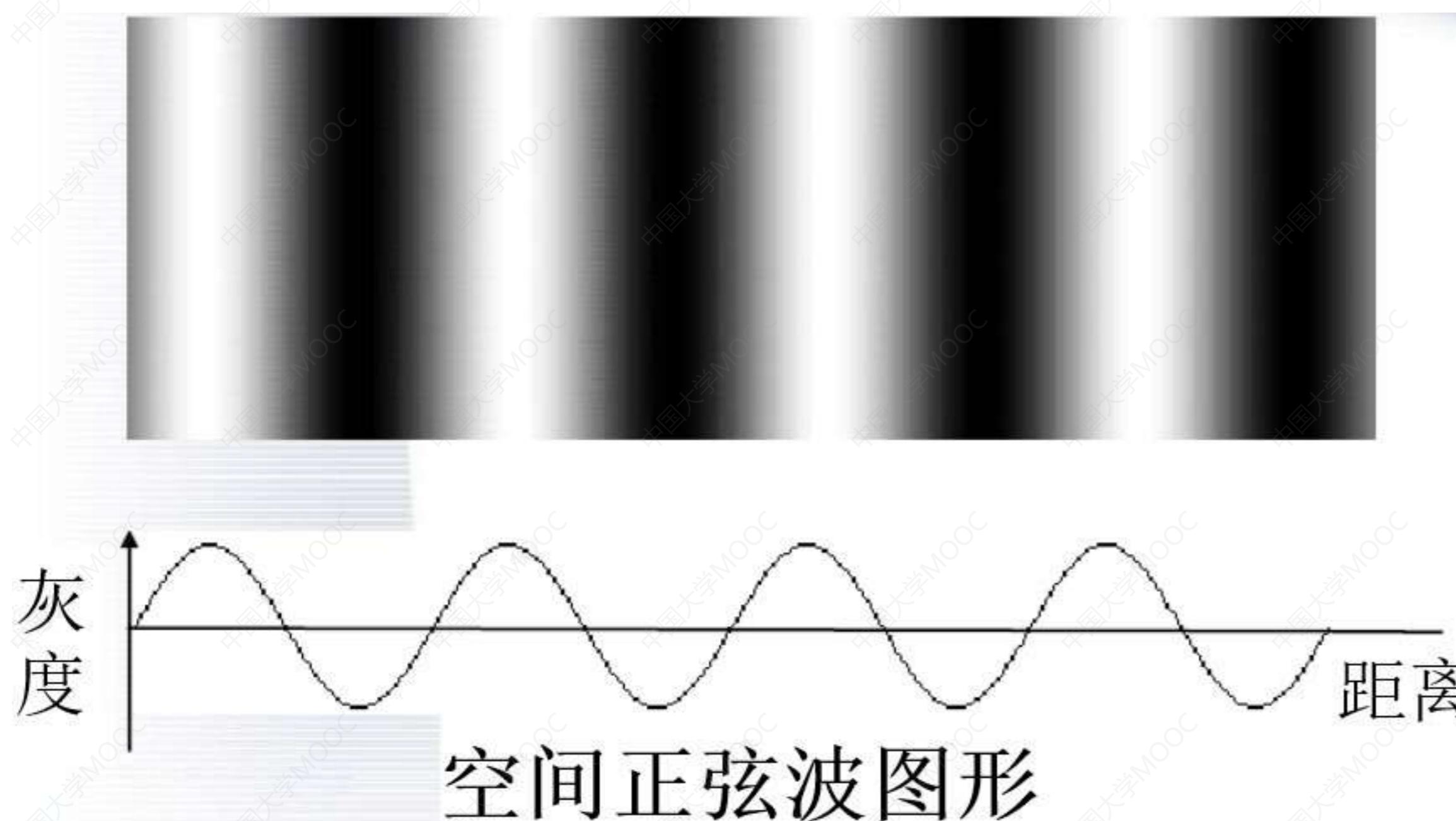
□ 眼睛光学系统的空间频率特性



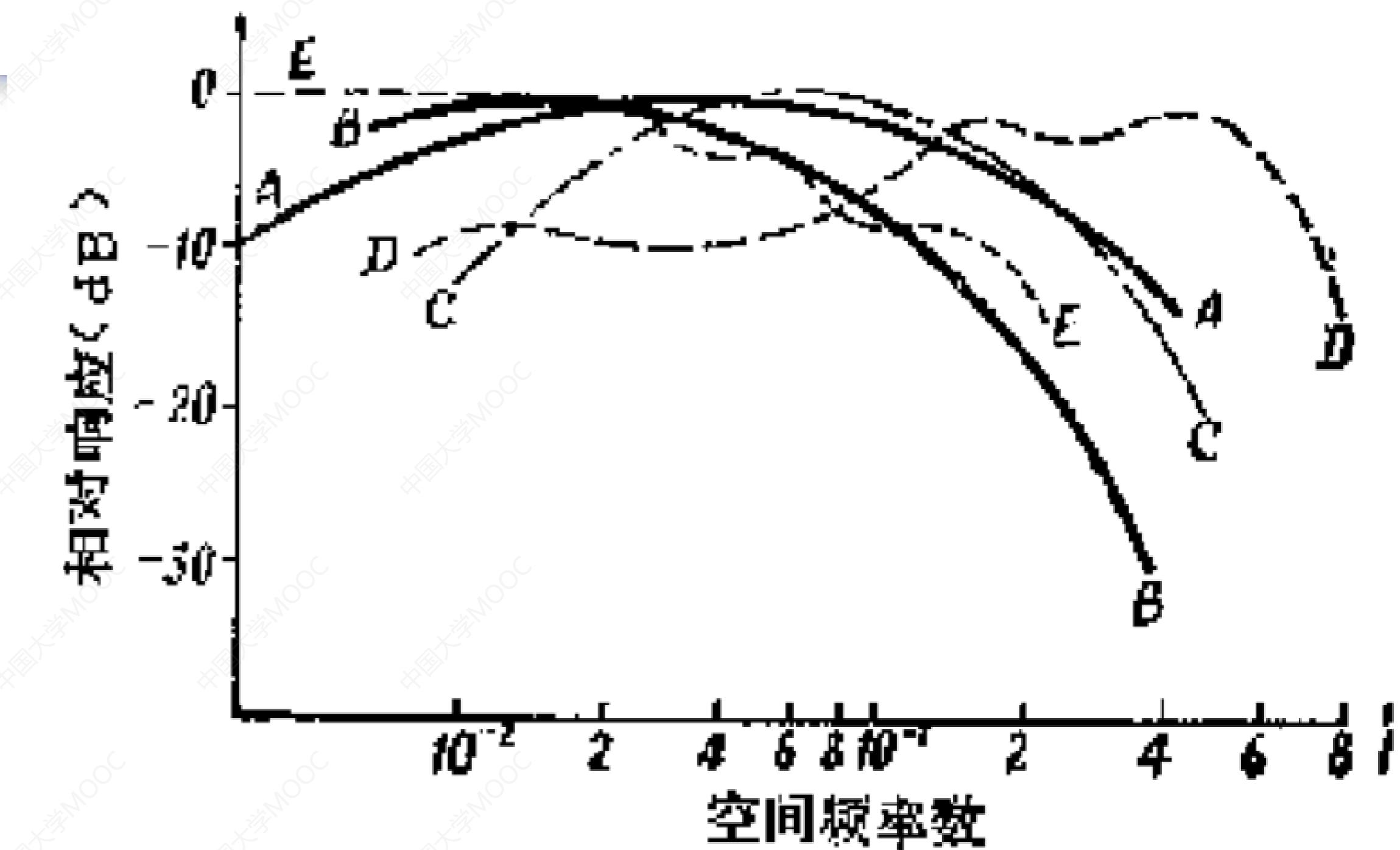
曲线A：克劳斯科普夫测定
曲线B：罗尔勒测定；
曲线C：迪莫特测定
曲线D：拉曼特测定

2.2 视觉特性

1. 视觉的空间频率特性



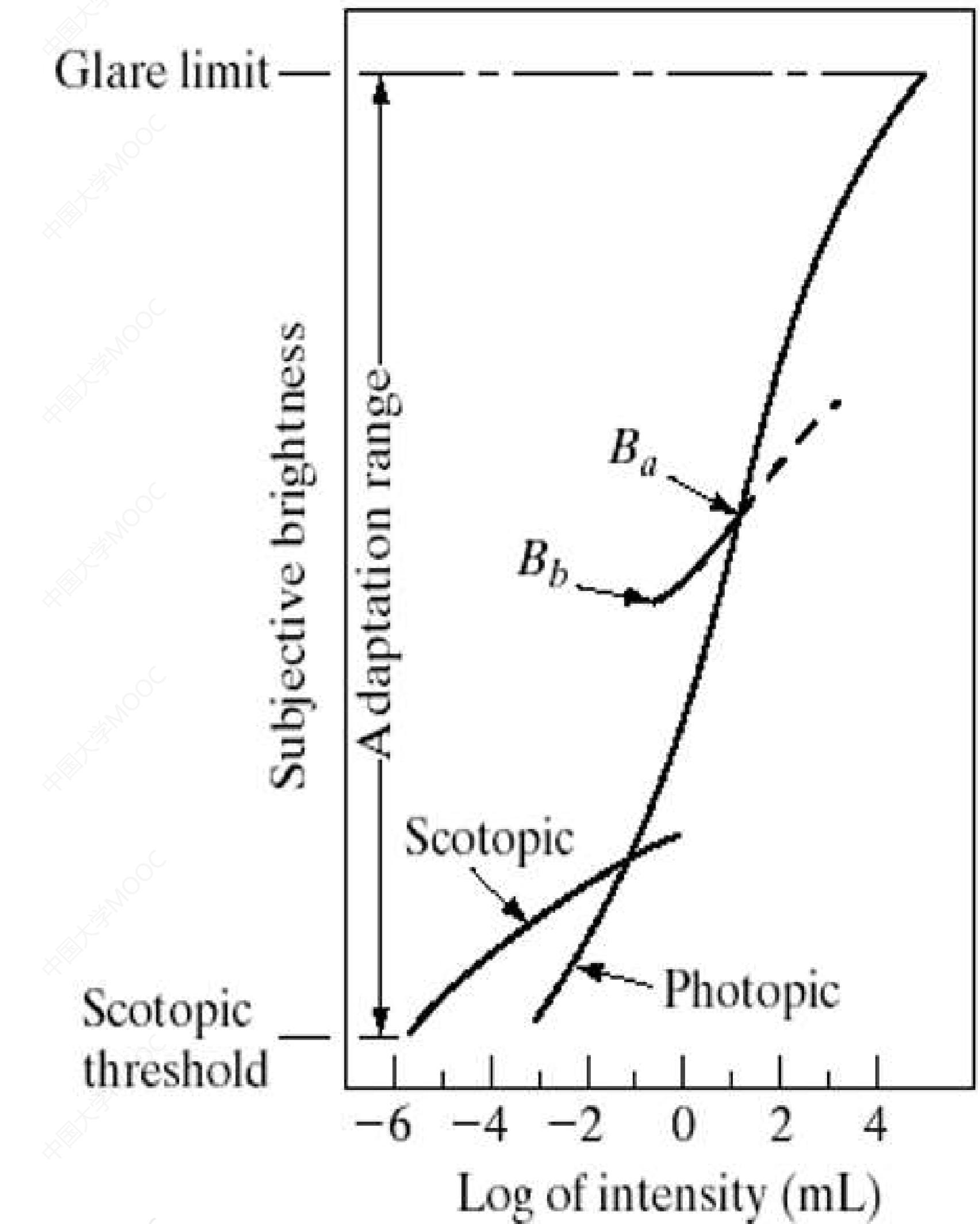
曲线A: 日本NHK测定
曲线B: 大上 (日本) 测定;
曲线C: 谢德 (Schade) 测定



■ 2.2 视觉特性

2. 亮度适应性

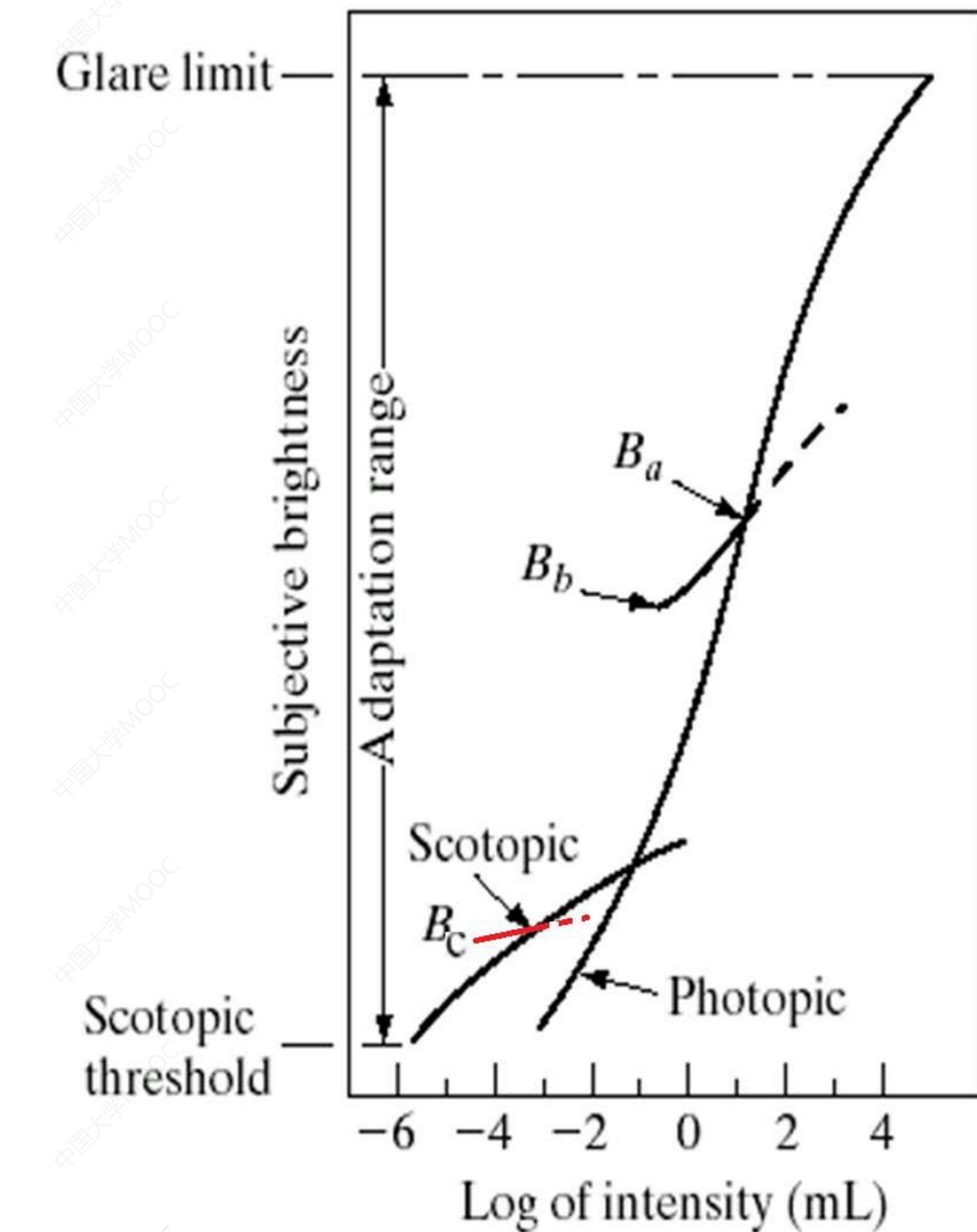
- 客观亮度 (*intensity*) : 外界输入到人的视觉系统的光源强度 (流明)
- 主观亮度 (*subjective brightness*) : 视觉系统感知的外界亮度;
- 主观亮度近似为客观亮度的对数 ;



2.2 视觉特性

2. 亮度适应性

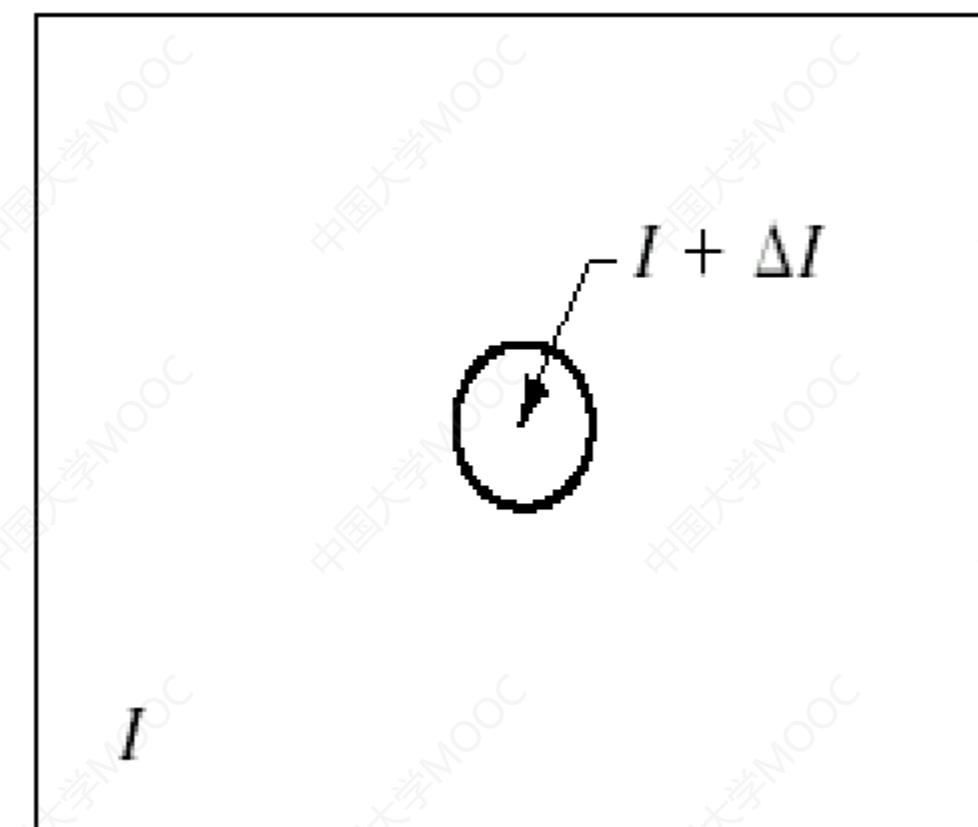
- 人的视觉系统能够感知的亮度范围 10^{-6} - 10^4 毫朗伯
- 视觉系统不可能同时感知如此大的范围，在特定亮度下形成特定的亮度响应曲线，感知一定亮度范围的变化，换一个场景，则可以自动调整为与亮度环境相适应的亮度相应曲线，感知相应范围的亮度变化。这就是视觉系统的亮度适应性。



2.2 视觉特性

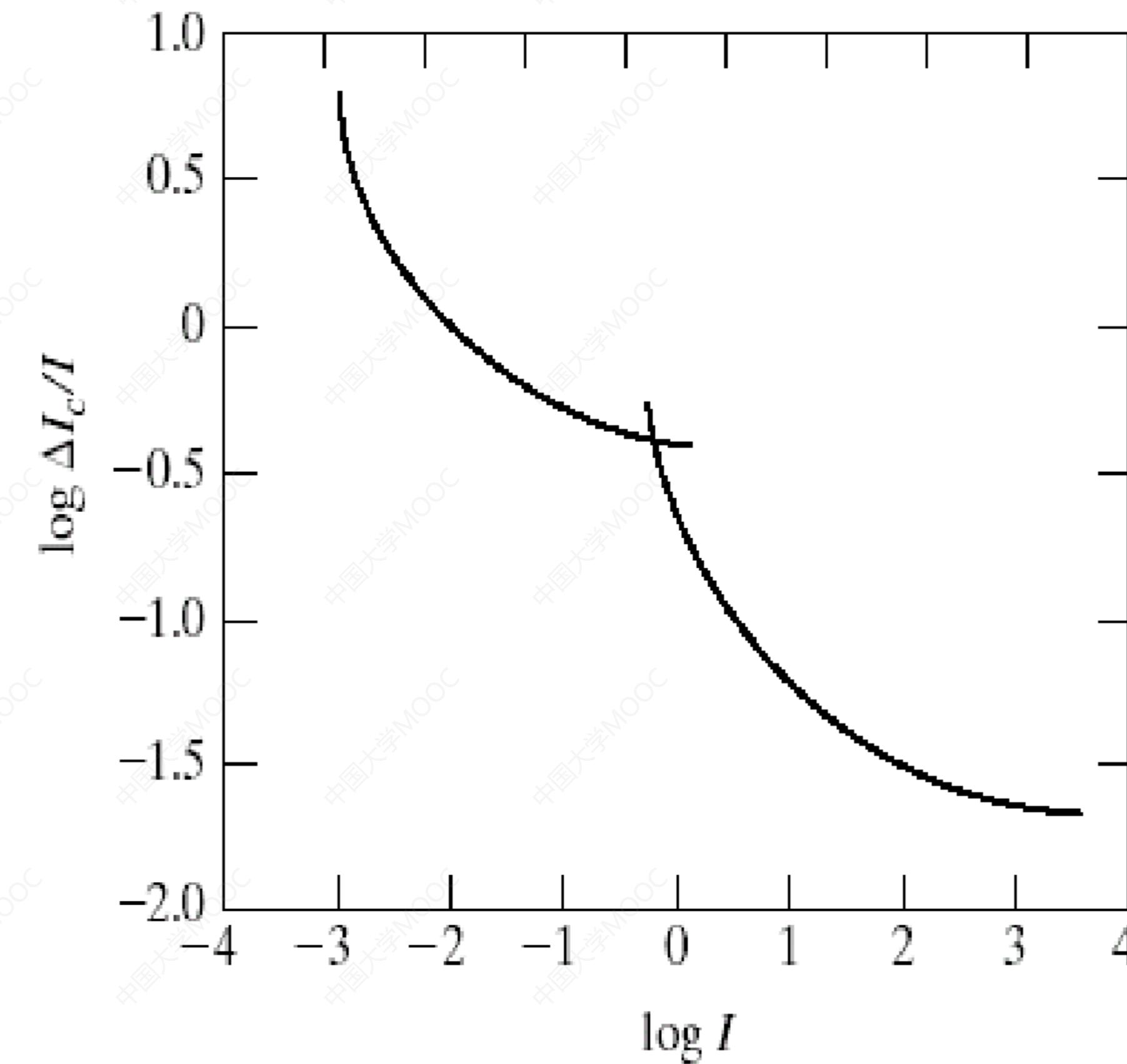
3. 视觉分辨力

► 经典实验——韦伯实验



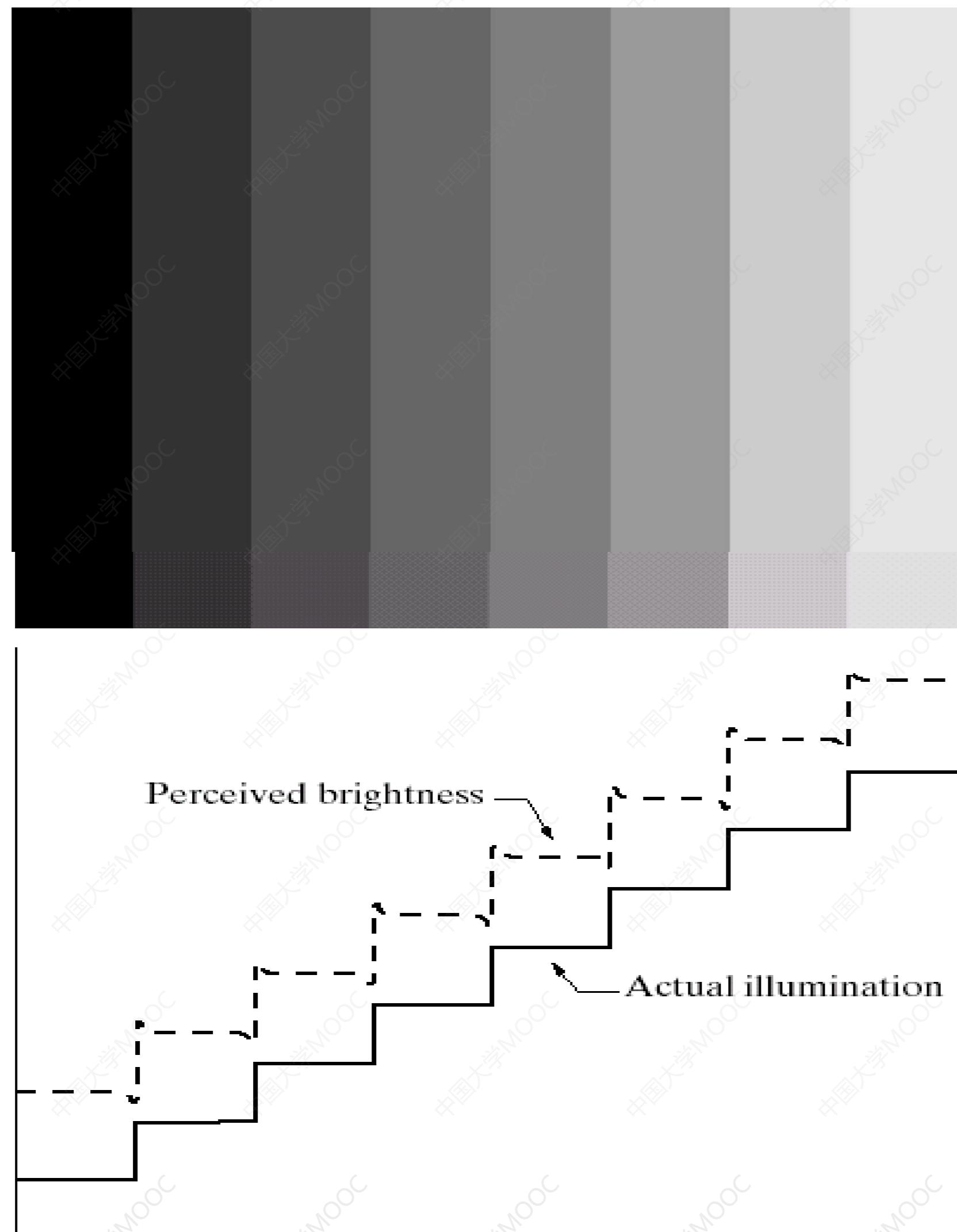
$\Delta I_c / I$ 韦伯比

ΔI_c 以 $I + \Delta I_c$ 的亮度在中心区域闪烁，视觉系统可以感受到50%的闪烁时对应的亮度增加值。 ΔI_c 越小，说明相应亮度 I 下的亮度变化越容易分辨，视觉的亮度分辨率越高。



2.2 视觉特性

4. 两个视觉现象——马赫带效应



a
b

FIGURE 2.7

(a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

■ 2.2 视觉特性

4. 两个视觉现象——同时对比效应

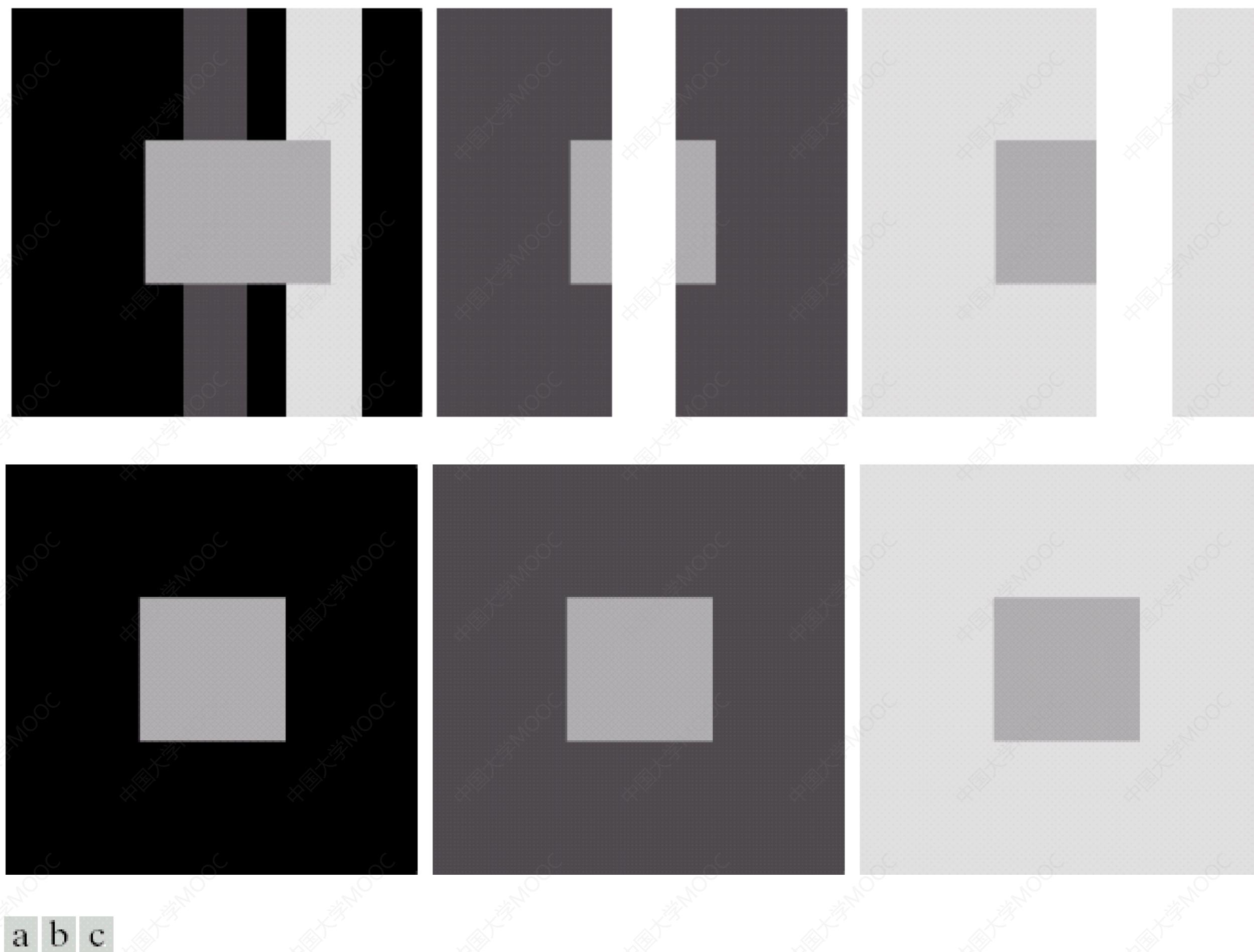
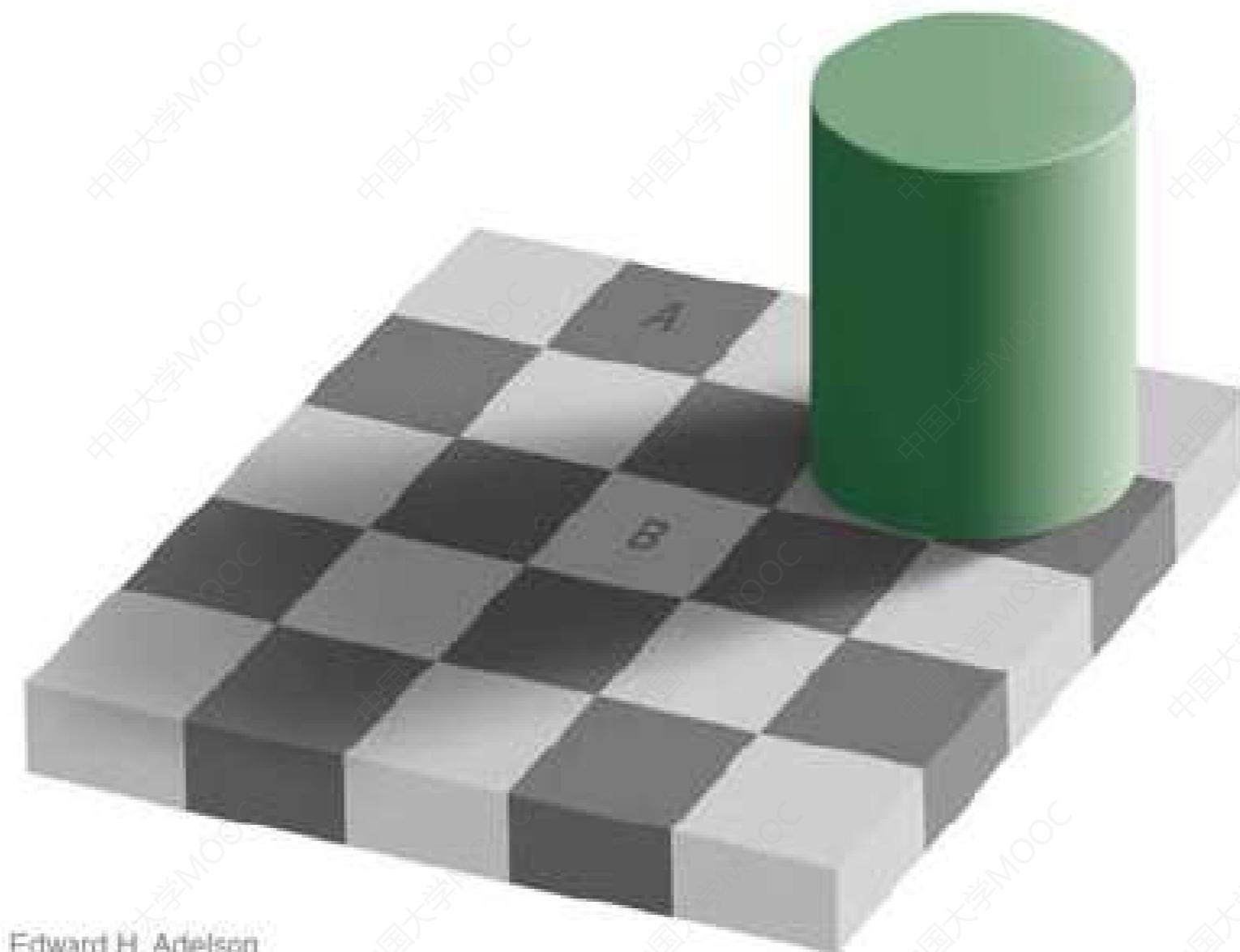


FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

■ 2.2 视觉特性

4. 两个视觉现象——同时对比效应

Brightness: Measurement vs. Perception



Edward H. Adelson

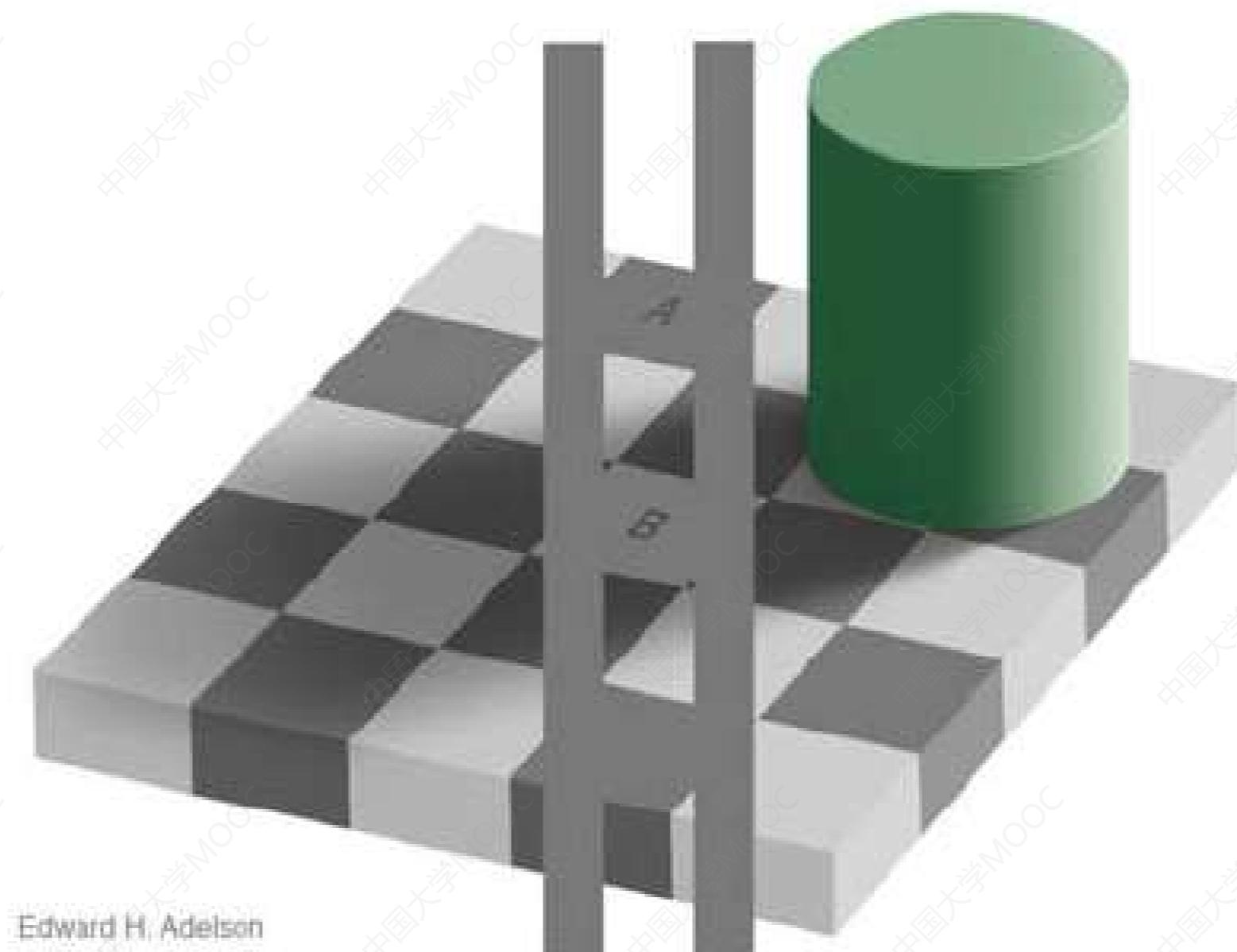
Do squares A and B have the same brightness?

Kodak

■ 2.2 视觉特性

4. 两个视觉现象——同时对比效应

Brightness: Measurement vs. Perception



Squares A and B have the same **measured** brightness but a different **perceived** brightness!

Kodak

■ 2.2 视觉特性

➤ 5 视觉惰性

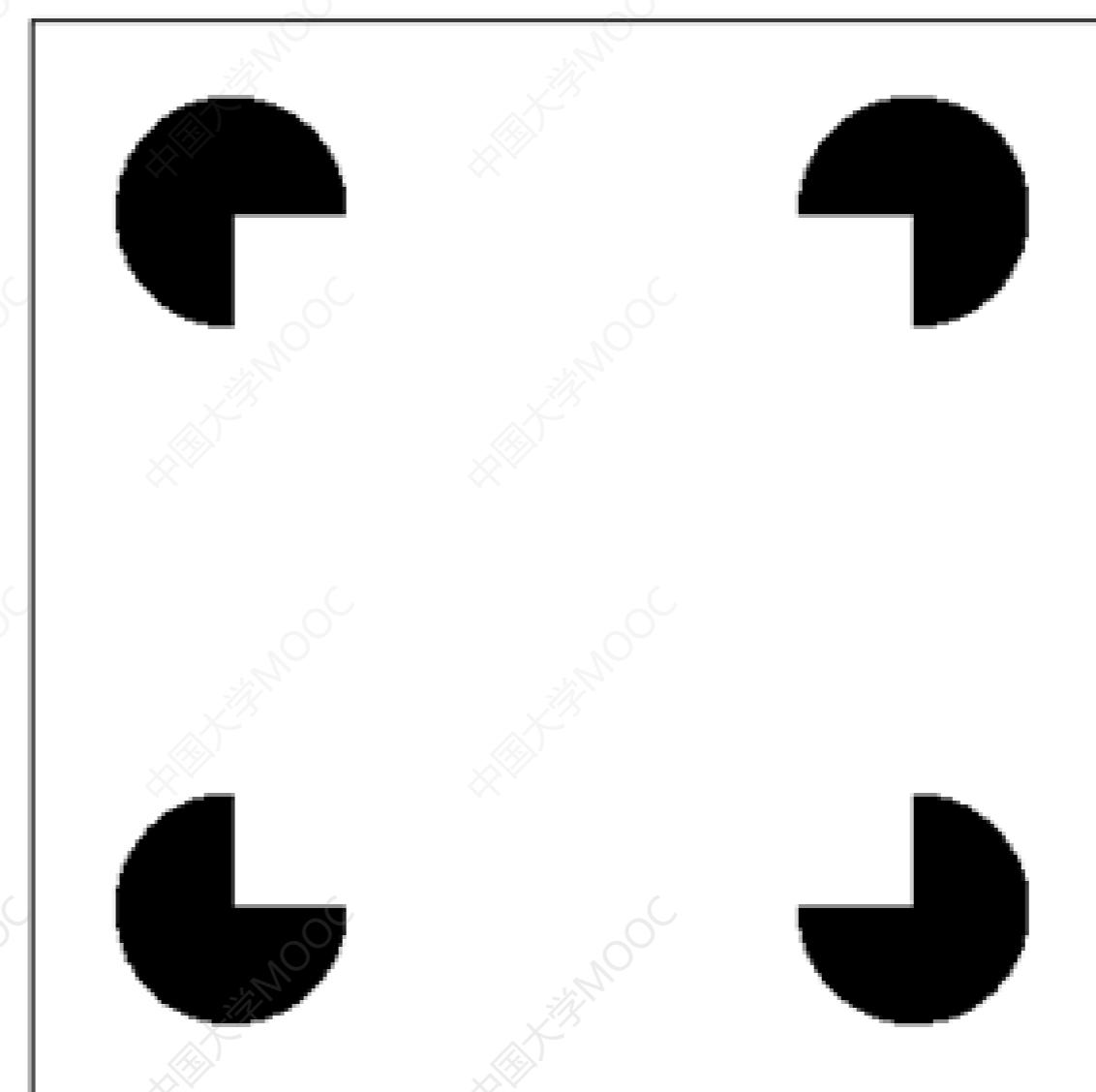
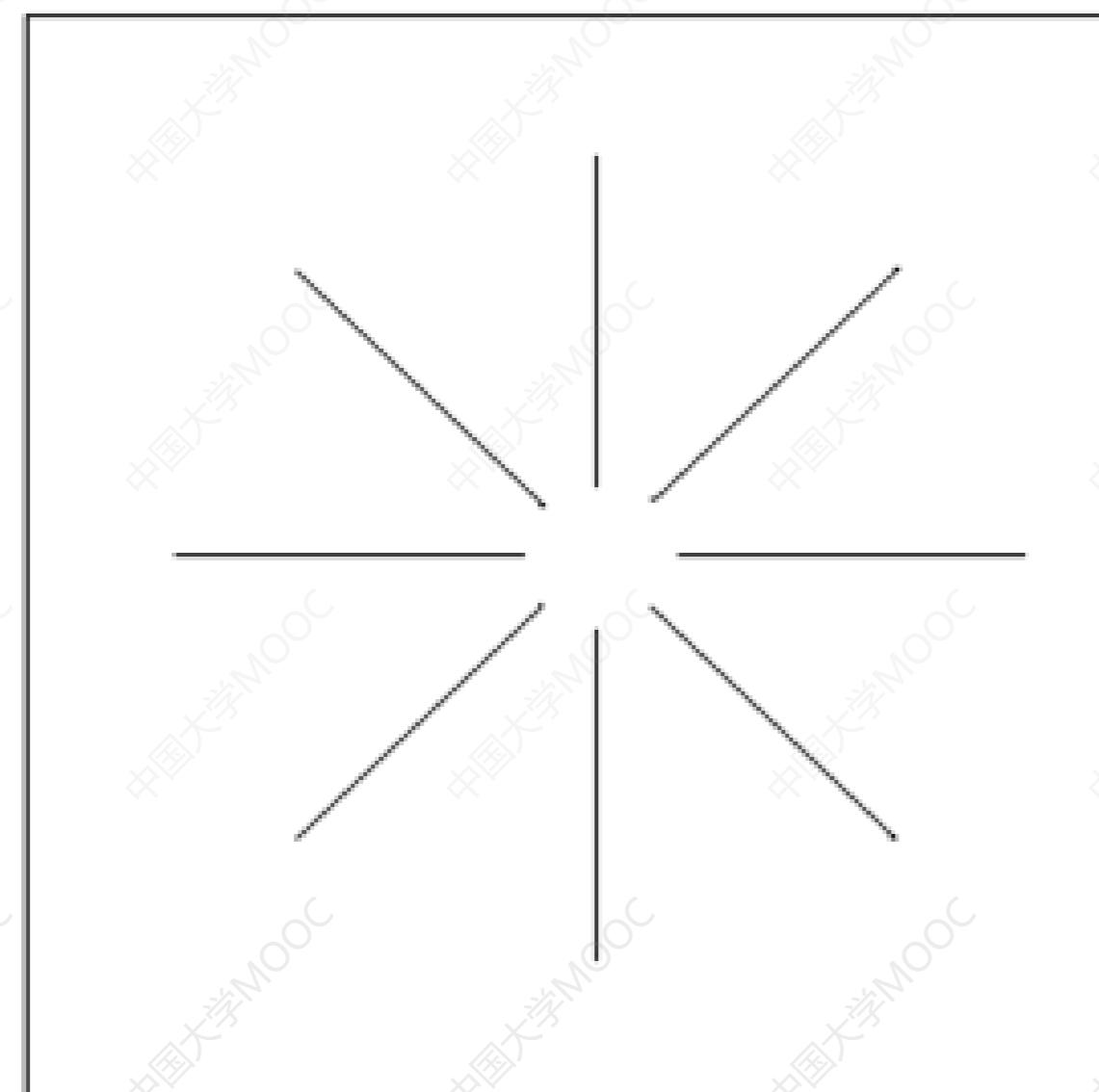
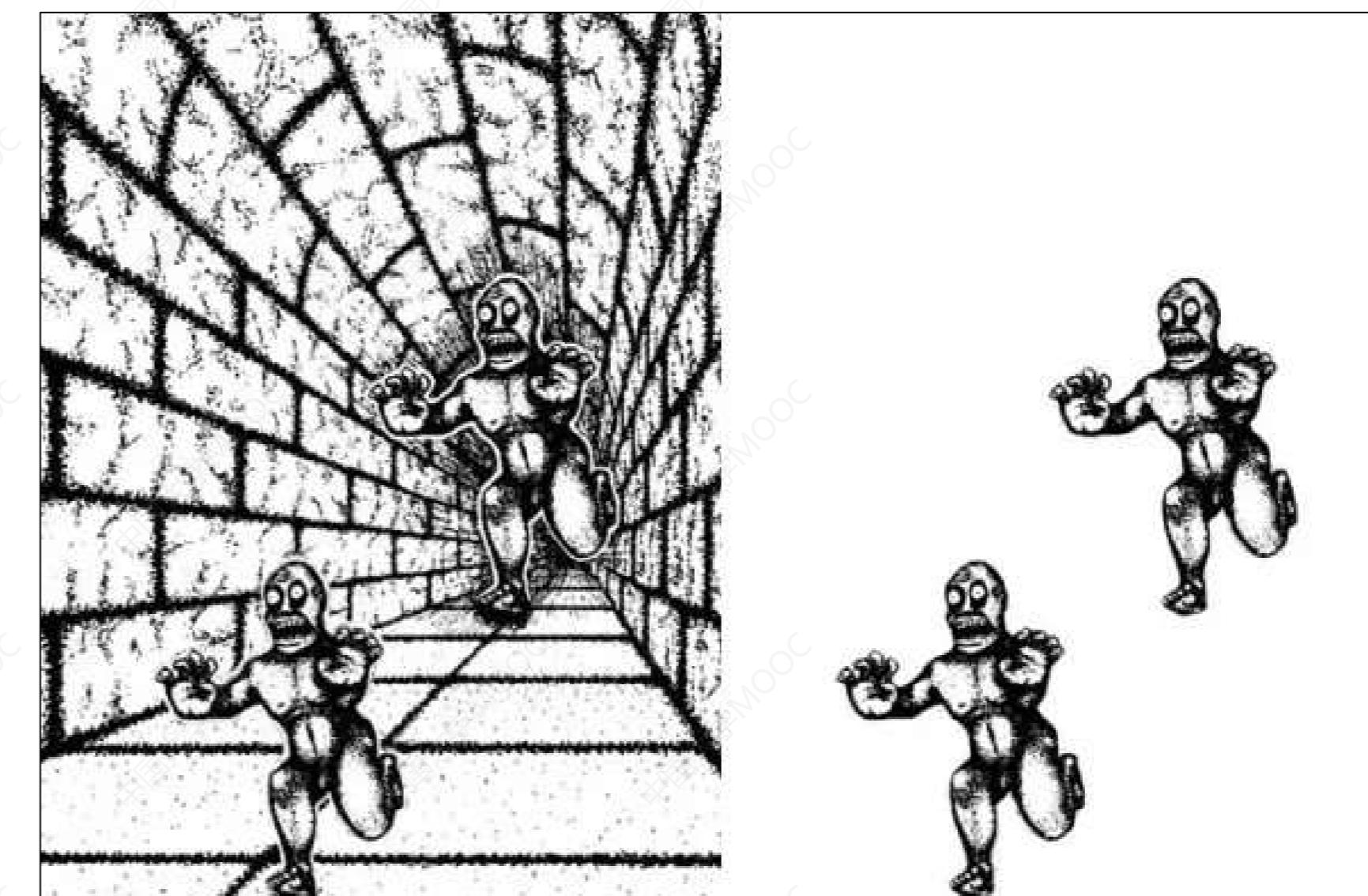
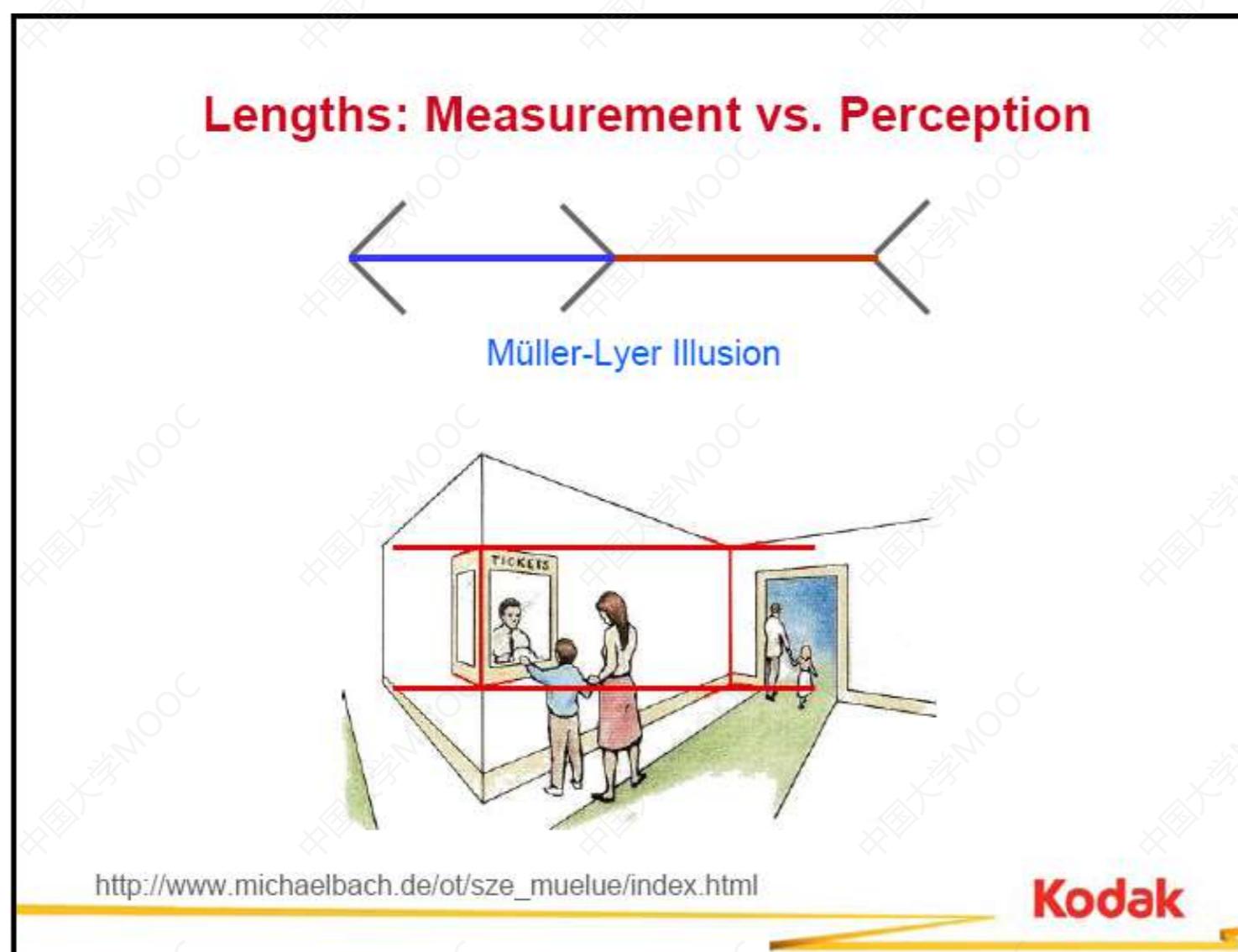
即光学图像一旦在视网膜上形成，视觉将会对这个图像的感觉维持一个有限的时间，这种生理现象也叫做视觉暂留。对于中等亮度的光刺激，视觉暂留时间为0.05至0.2秒

➤ 6 彩色视觉

光谱上380 ~ 760纳米(nm)波长的辐射能量作用于人的视觉器官所产生的颜色感觉，又称色觉，彩色视觉包括亮度、色调和色饱和度。

2.2 视觉特性

➤ 7 幻觉 (Illusion) ——人类视觉系统的主观感受和像素级表达不一致



■ 2.2 视觉特性

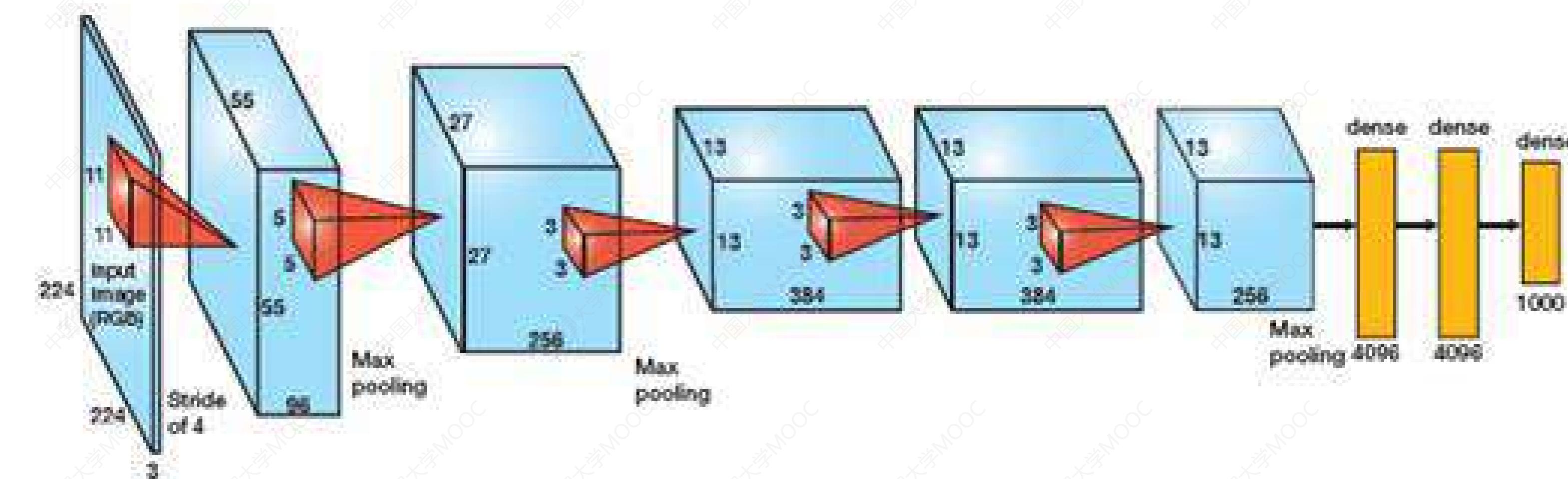
➤ 8 视觉系统模型

➤ 1 人眼—>视敏细胞—>视神经细胞—>视觉中枢

- 低通—>对数 —>高通
- 带通—>对数
- 对数—>带通
- 神经网络模型

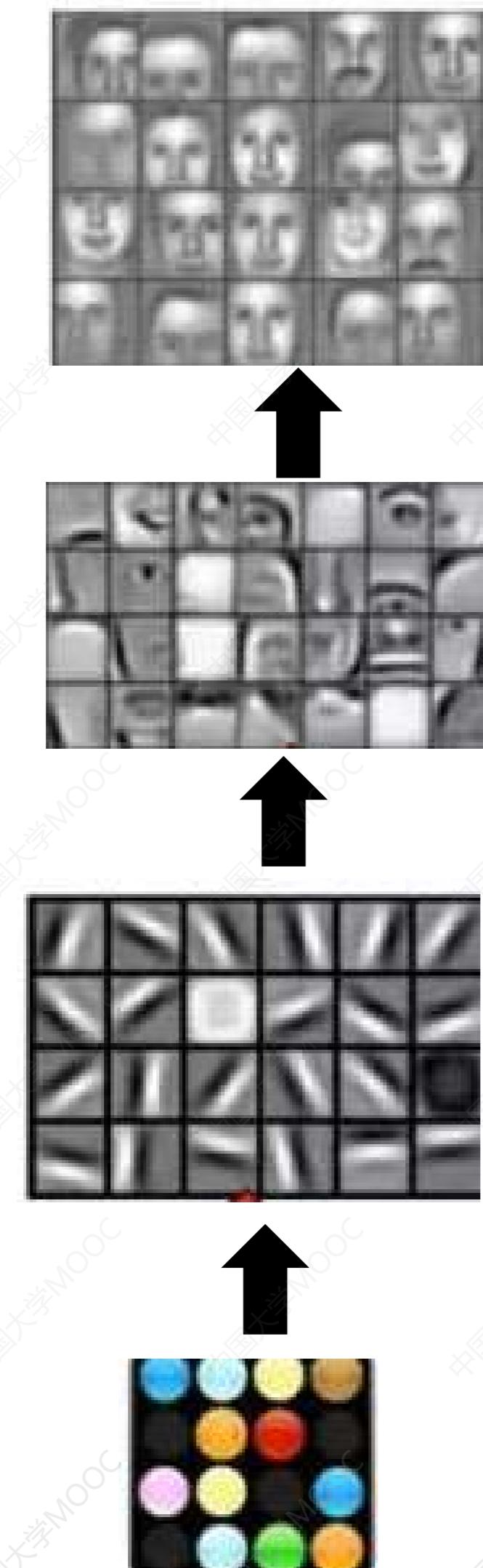
■ 2.2 视觉特性

- 2012年Geoffrey Hinton和他的学生Alex Krizhevsky 在2012年的imagenet上的图像分类challenge上，提出的Alex-net网络结构模型赢得了冠军，前五个分类准确率 (0.15315) , 第二名 (0.26172)



2.2 视觉特性

7 视觉系统模型

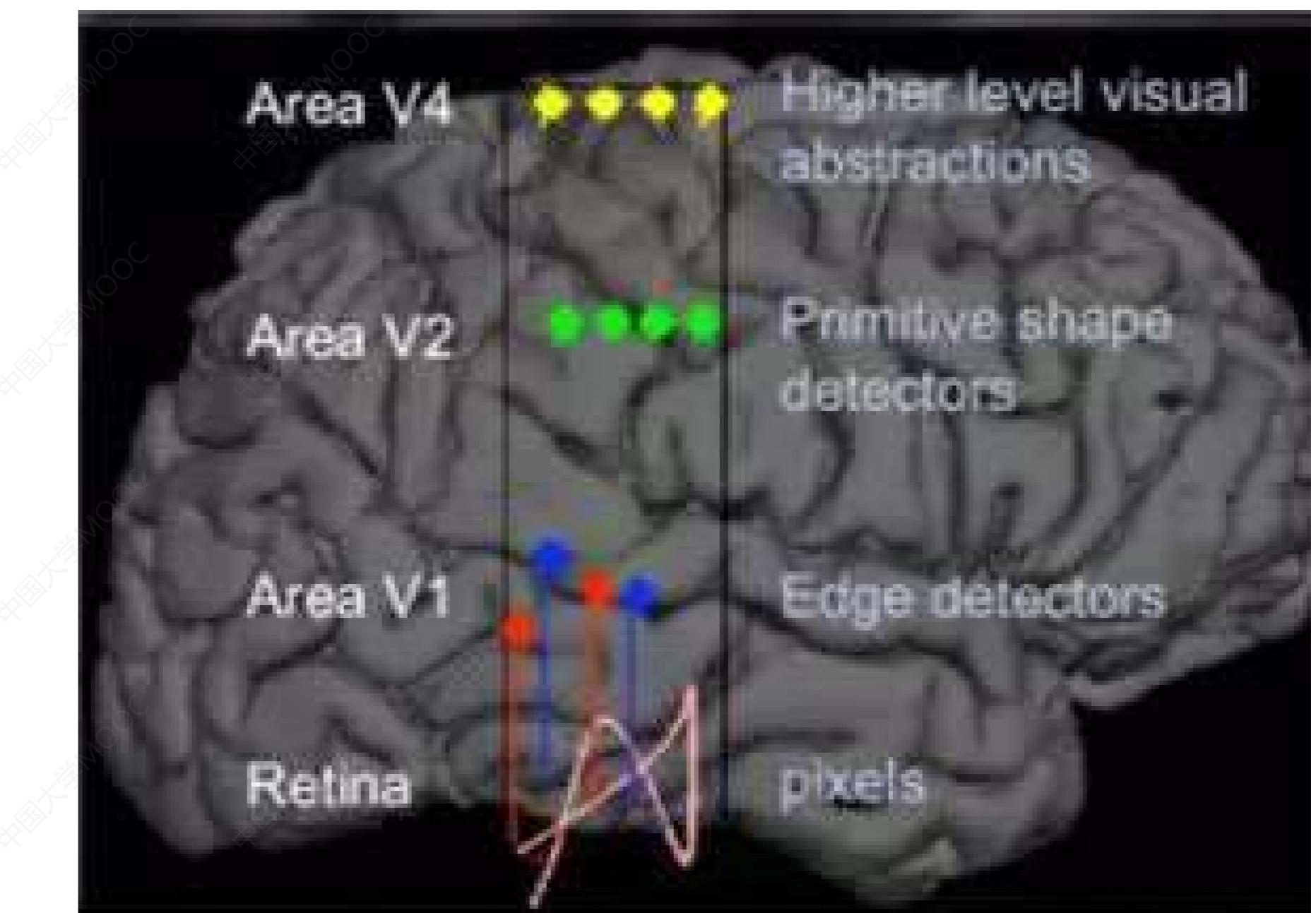


object models

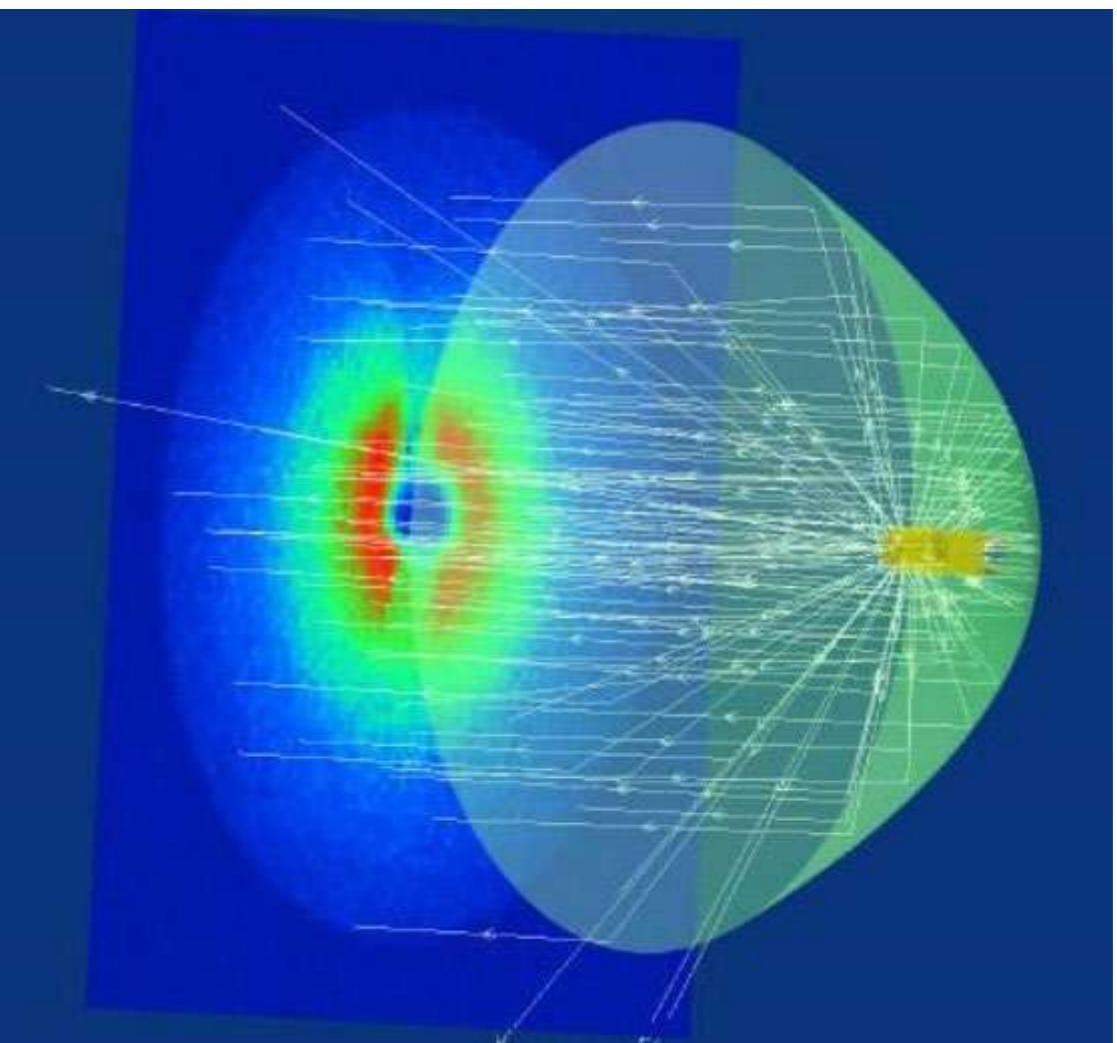
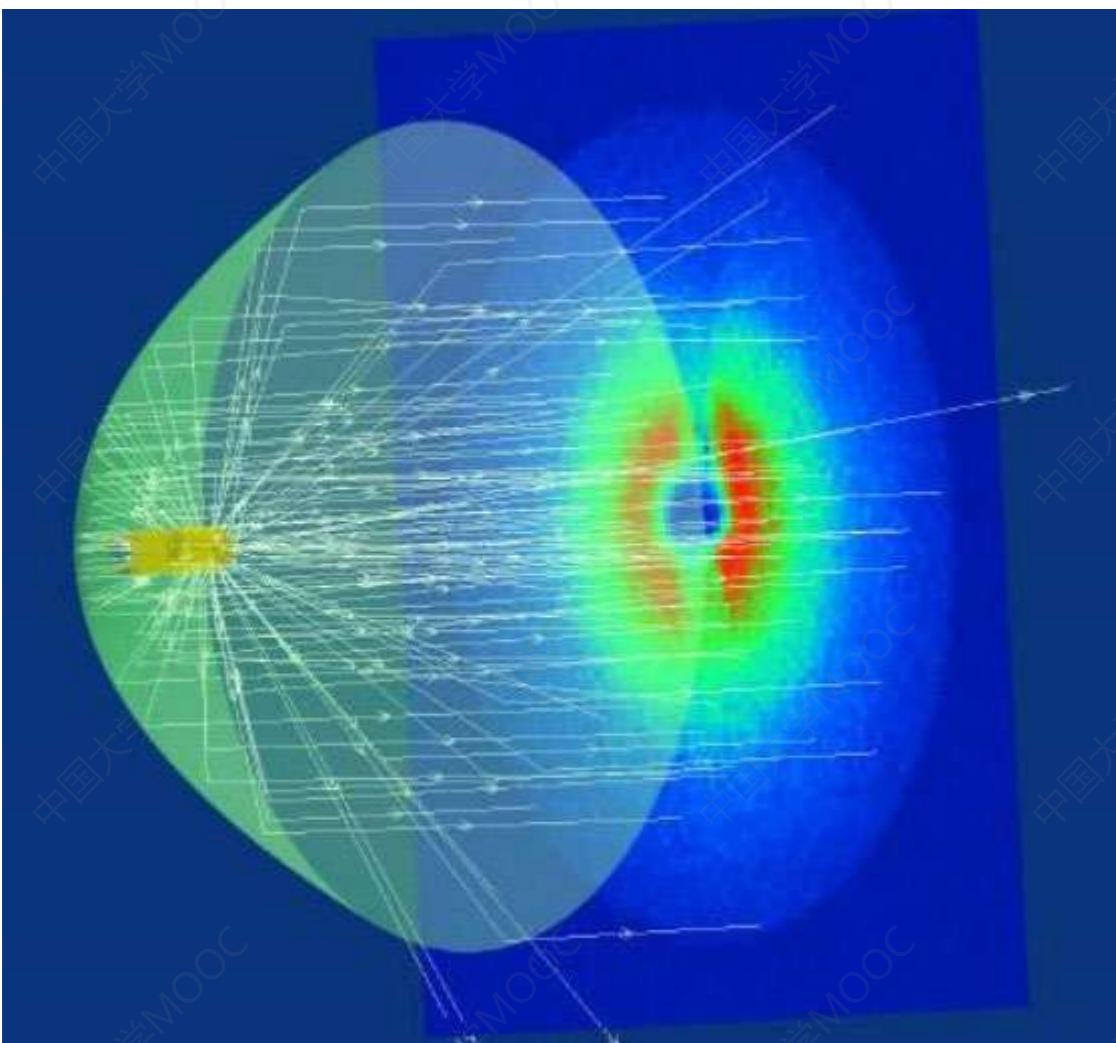
Object parts
(combination
Of edges)

edges

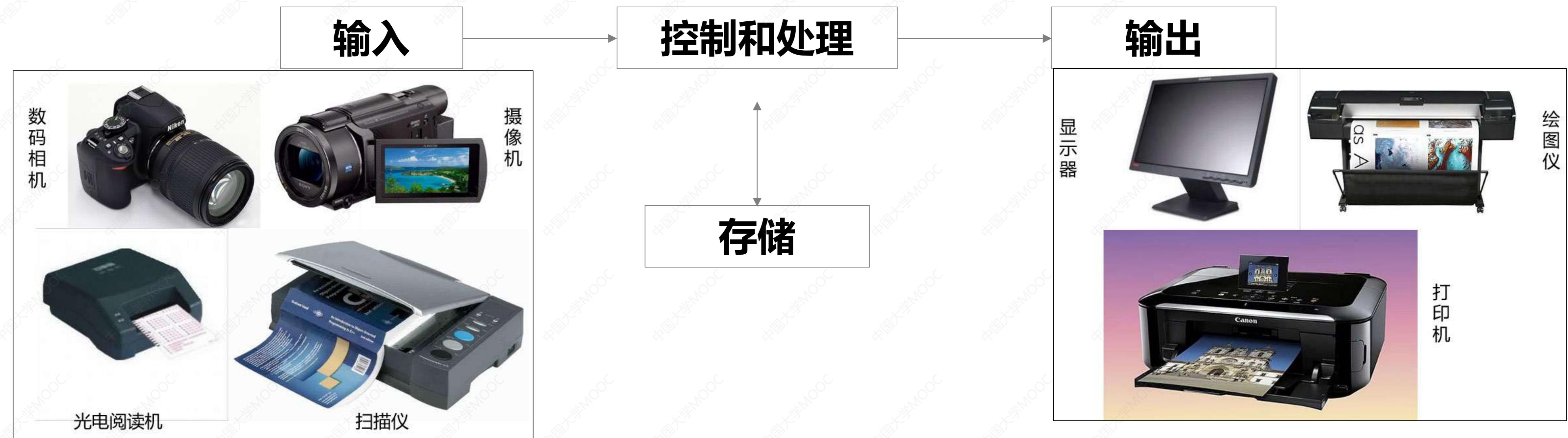
pixels



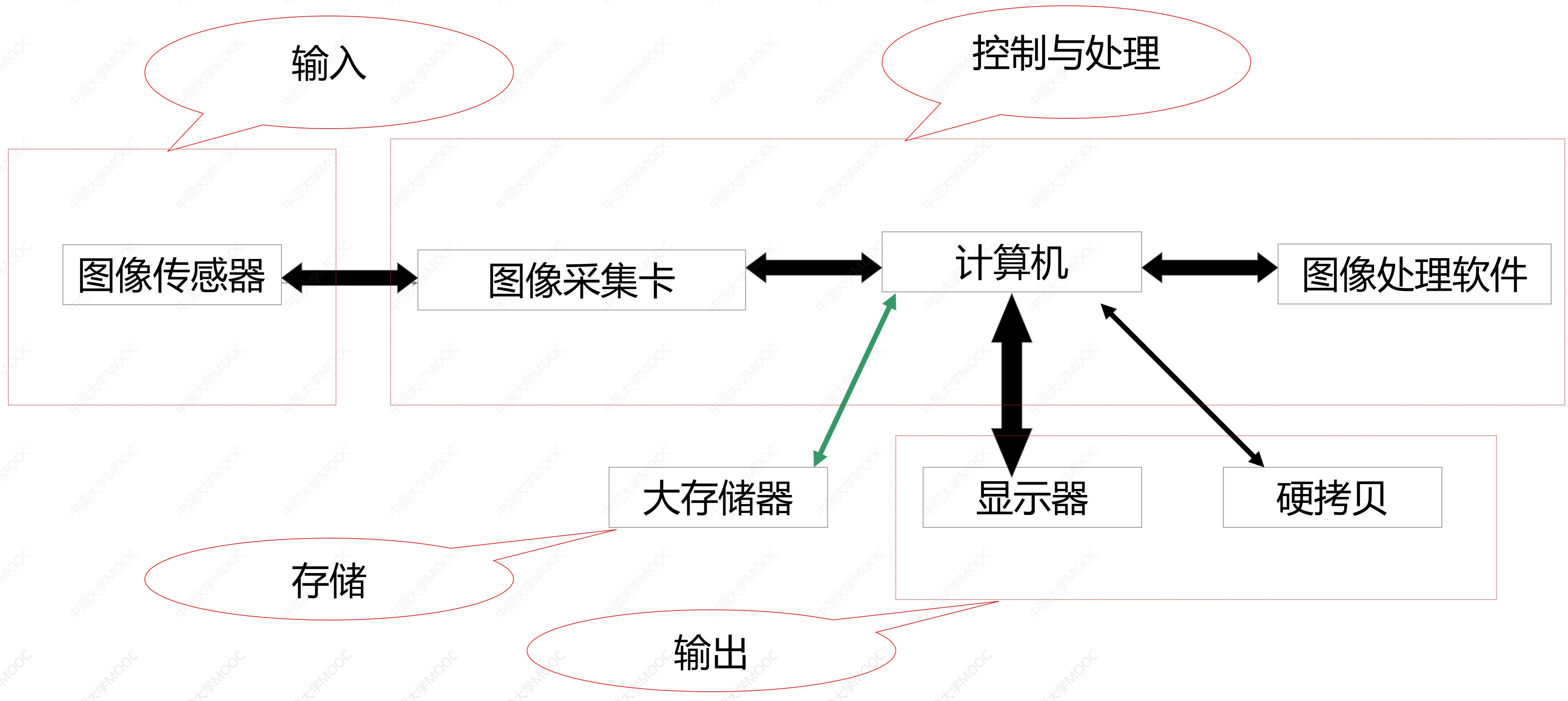
2.3 成像与数字化



2.3.1 图像处理系统组成



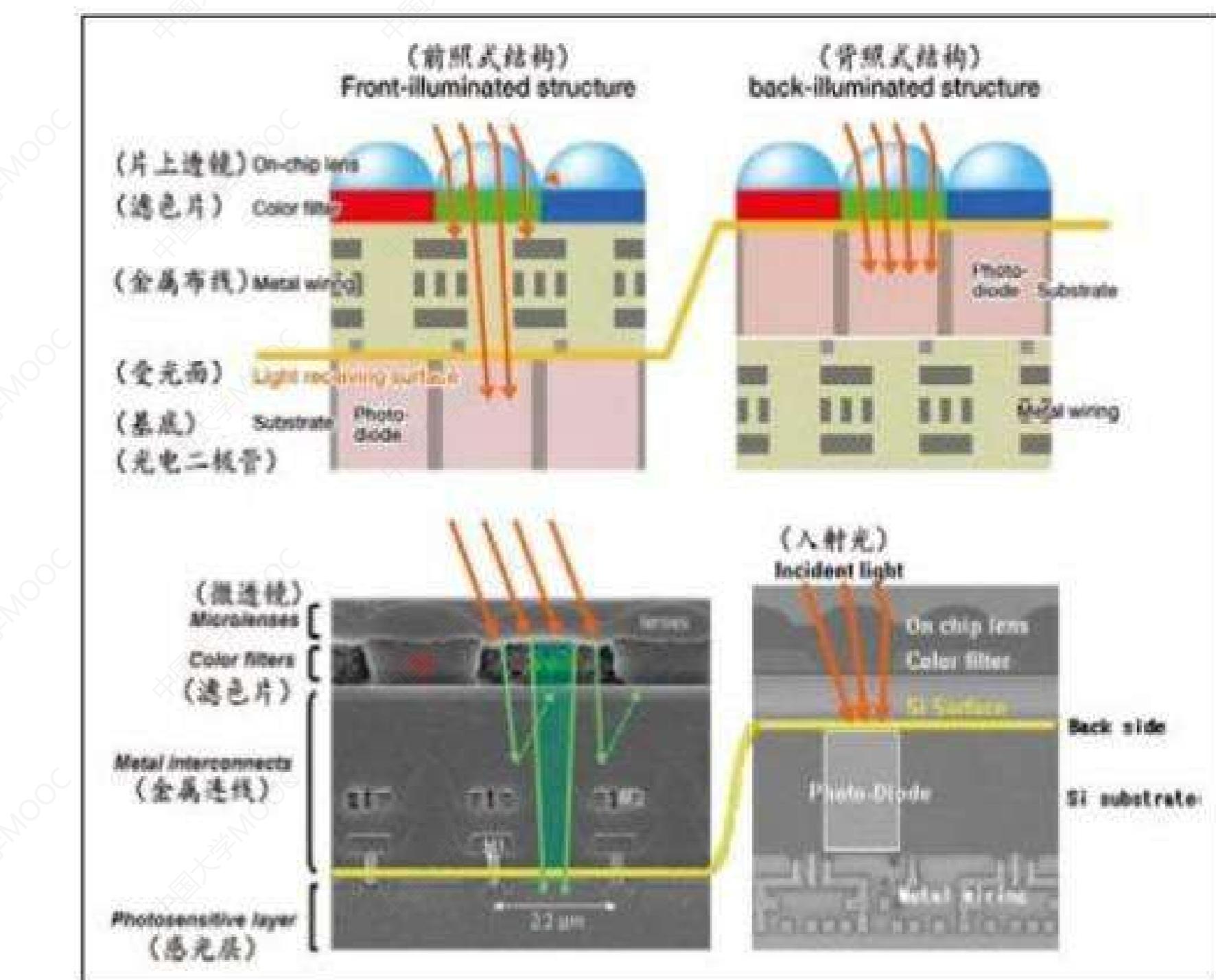
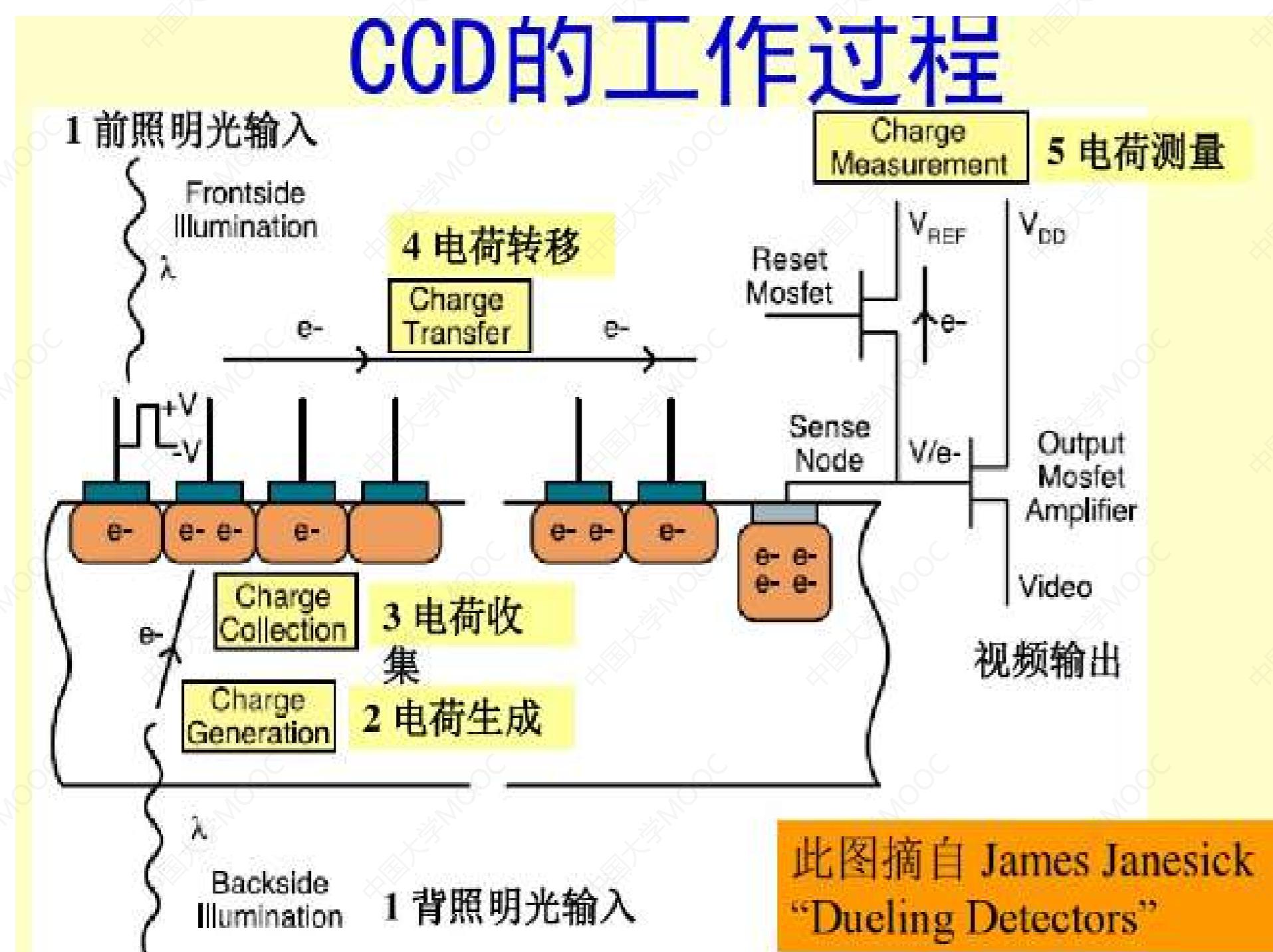
图像处理系统组成



图像处理系统组成

• 传感器类型

- CMOS: 灵敏度低、成本低、分辨率低、噪音大、功耗低
- CCD: 灵敏度高、成本高、分辨率高，噪音小、功耗高



图像数字化

➤采样

- 采样定理：设x,y方向的采样间隔分别为 Δx 和 Δy ，其采样频率分别为：

$$\Delta u = \frac{2\pi}{\Delta x} \quad \Delta v = \frac{2\pi}{\Delta y}$$

- 设图像在x,y方向的最高频率分别为 u_c 和 v_c ，则不失真采样应满足

$$\Delta u \geq 2u_c \quad \Delta v \geq 2v_c$$

■ 图像数字化

- 量化：采样点灰度值的离散化过程
- 采样点数与量化级数的选取
 - ✓ 图像大小： $M \times N$, 量化灰度级数Q。
 - ✓ 1. Q一般为2的整数次幂, $Q = 2^b$
 - 一般情况下: $b=5-8$, 人眼的灰度分辨率为32-64灰度级。
 - 遥感图像, 医学图像: $b=8-12$
 - ✓ 2. M,N的选取应满足采样定理。

■ 图像的空间分辨率和灰度分辨率

- 采样决定了图像的空间分辨率
 - 空间分辨率即图像中的最小可分辨细节，可以用单位距离上的线对数量来表示；如每英寸点数 (dpi) 来表示。
- 量化决定了灰度分辨率
 - 灰度分辨率即图像中的最小可分辨灰度变化，测量可分辨的灰度变化比较主观；
 - 当实际的物理分辨率度量没有必要时，通常一幅图像的空间分辨率用其像素点数来表示，灰度分辨率用其灰度级数来表示。

图像数字化



a



b



c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

■ 图像数字化

- 256 gray level



■ 图像数字化

- 128 gray level



■ 图像数字化

- 64 gray level



■ 图像数字化

- 32 gray level



■ 图像数字化

- 16 gray level



■ 图像数字化

- 8 gray level



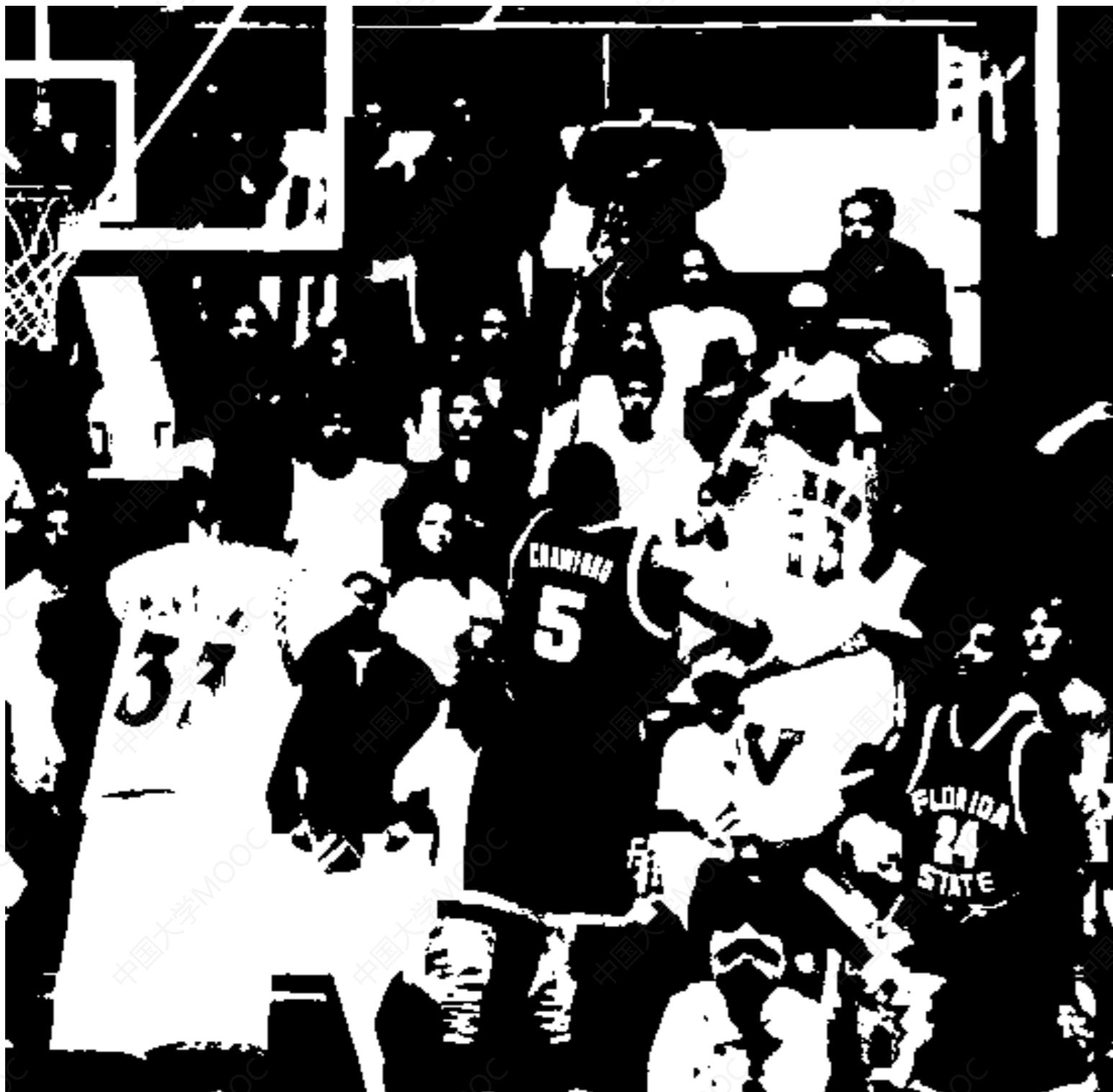
■ 图像数字化

- 4 gray-level



■ 图像数字化

- 2 gray-level



■ 图像数字化

- $\frac{1}{2}$ Sub-sample



■ 图像数字化

- $\frac{1}{4}$ Sub-sampling



■ 图像数字化

- 1/8 Sub-sampling



■ 图像数字化

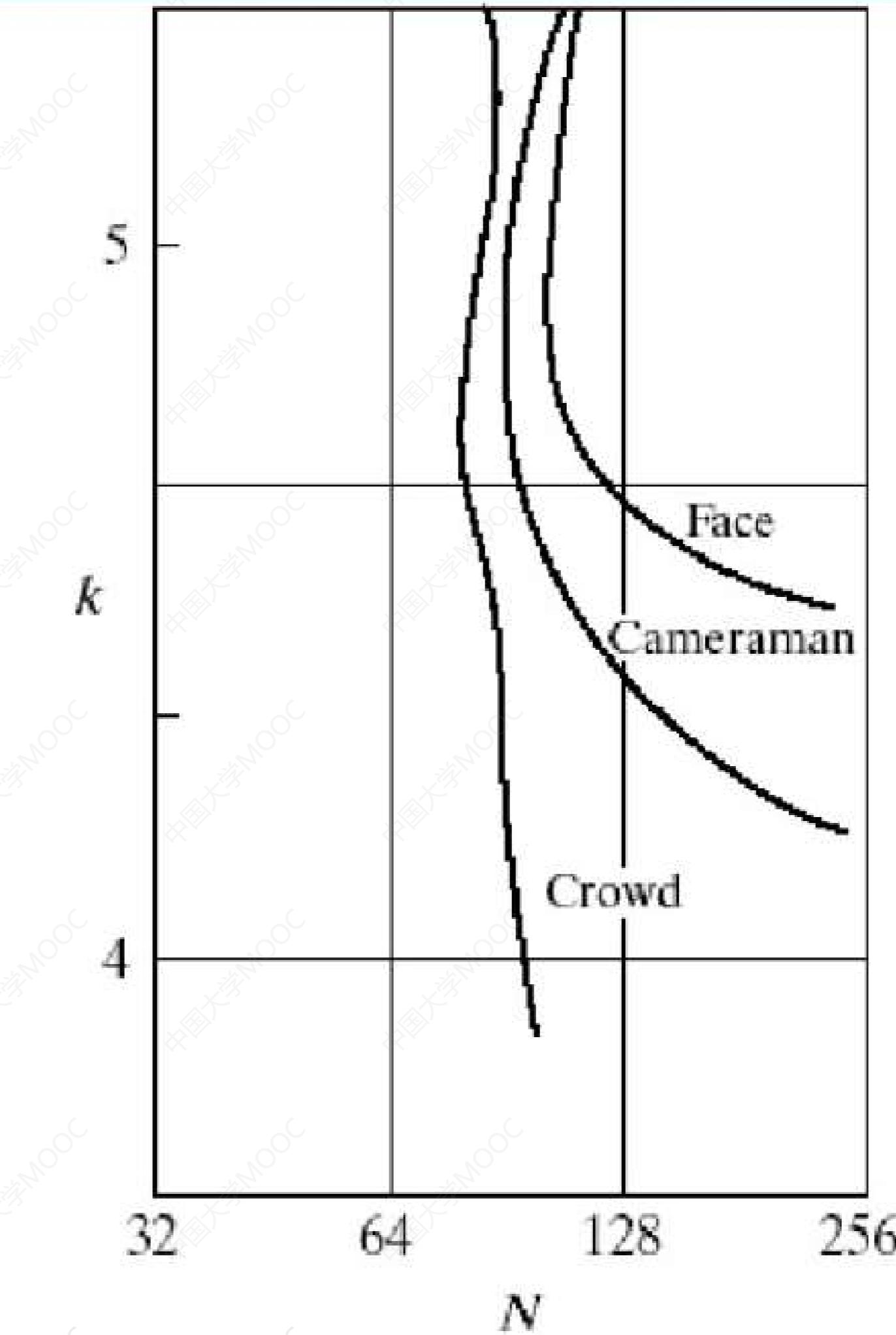
✓ 3. 降低分辨率和灰度级数对图像质量的影响

- a) 对于低复杂度图像，图像中有大面积灰度变化平缓的区域，降低分辨率不会引起细节损失，因而图像质量保持保持较好；减小灰度级数，容易导致灰度平滑区域出现假轮廓，引起图像质量下降；
- b) 对于高复杂度图像，图像中细节较多，如群众场面，降低分辨率容易引起细节损失，导致图像质量下降；减小灰度级数，不致丢失图像细节，因而图像质量保持较好。

等偏爱曲线

FIGURE 2.23
Representative
isopreference
curves for the
three types of
images in
Fig. 2.22.

等偏爱曲线
(Isopreferencecurve): 在**N-k**
平面内对应于主观感觉质量相等
的曲线；**•实验结果：**当图像中
细节增加时，等偏爱曲线趋于更
接近**N-k**平面内的垂直直线。



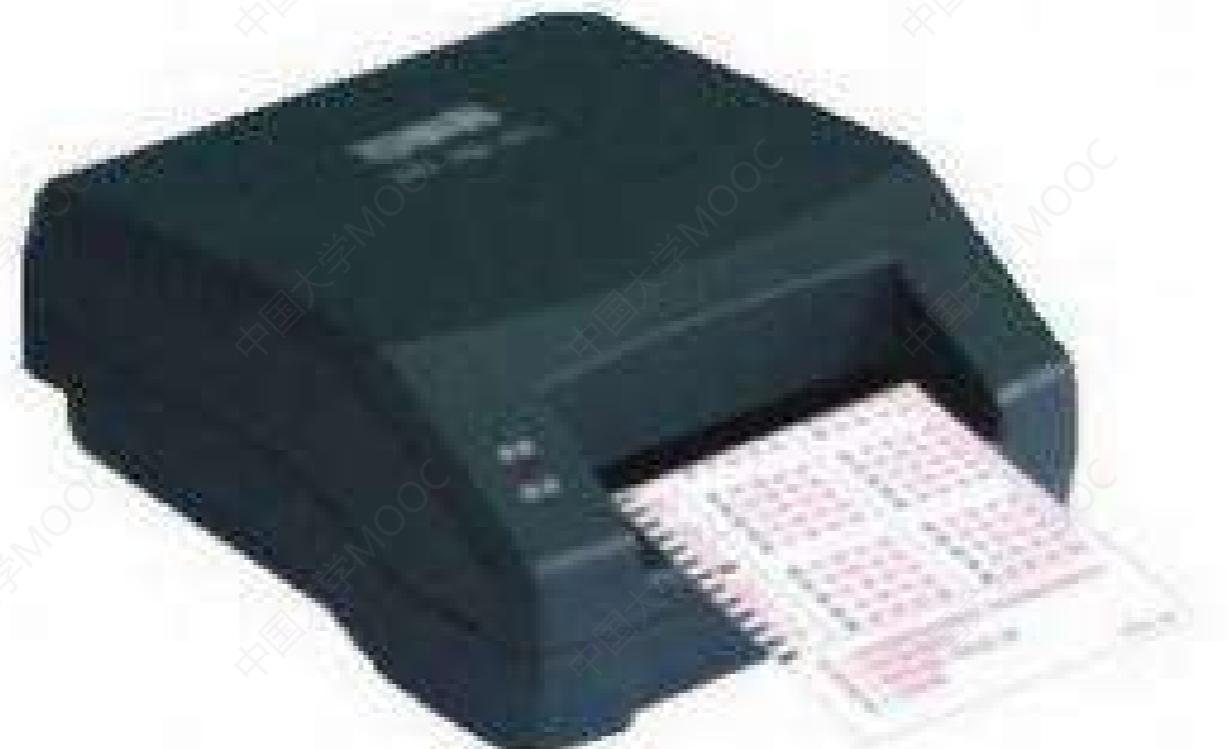
图像处理系统组成

- 图像输入设备
 - 数码相机
 - 摄像机
 - 扫描仪
 - 光电阅读机、

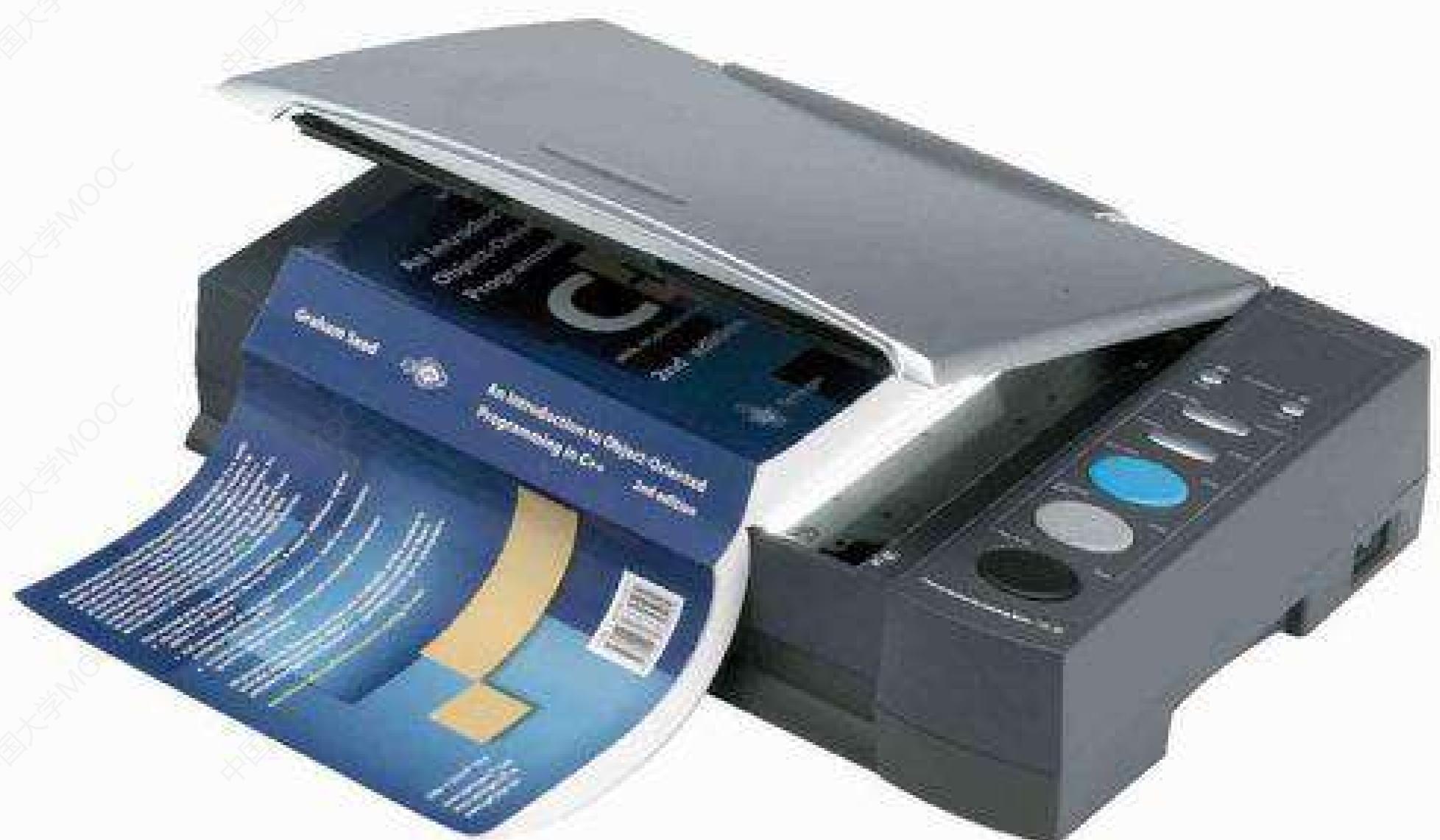
数码相机



摄像机



光电阅读机



扫描仪

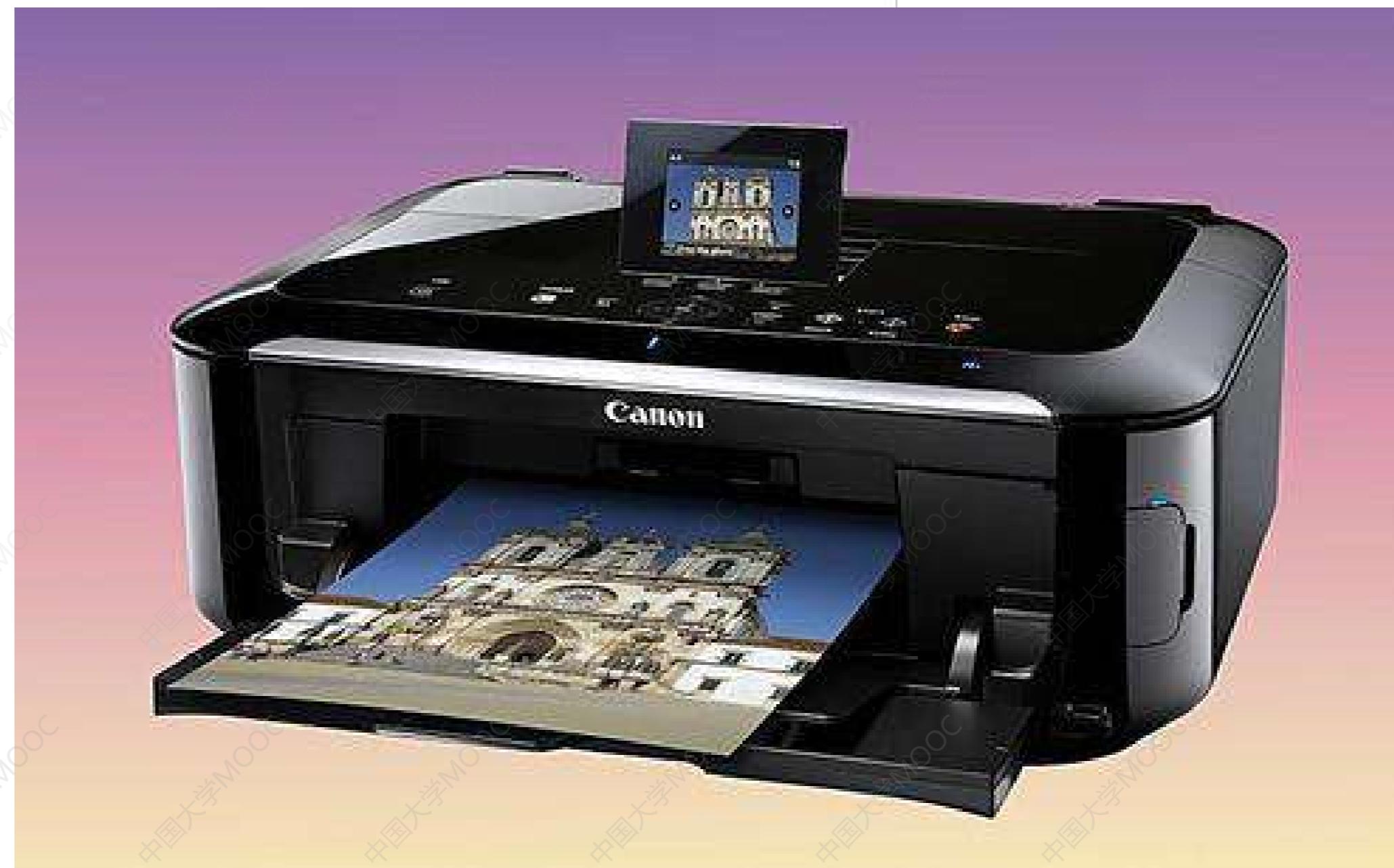
图像处理系统组成

- 图像输出设备
 - 显示器
 - 打印机
 - 绘图仪

显示器

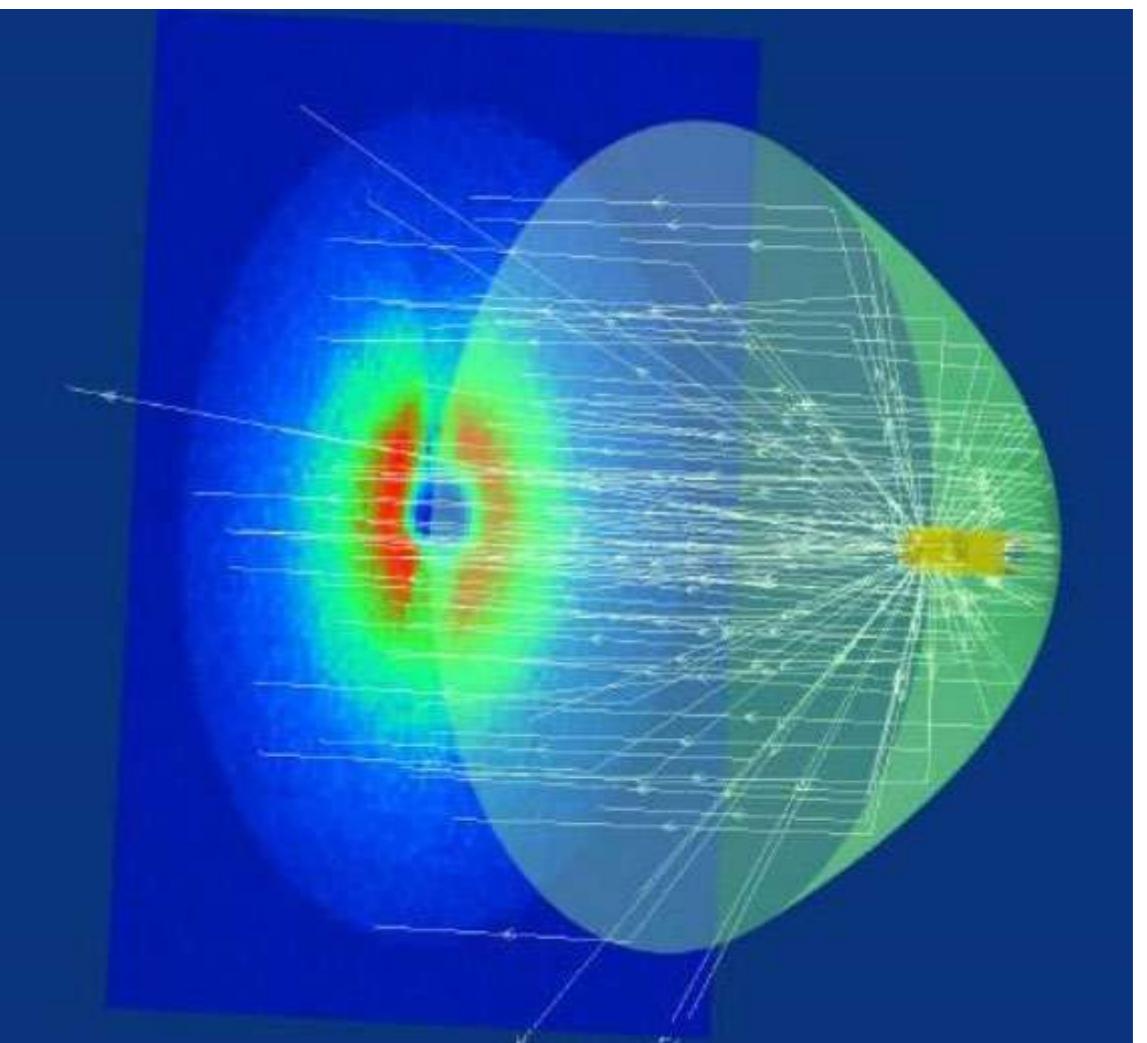
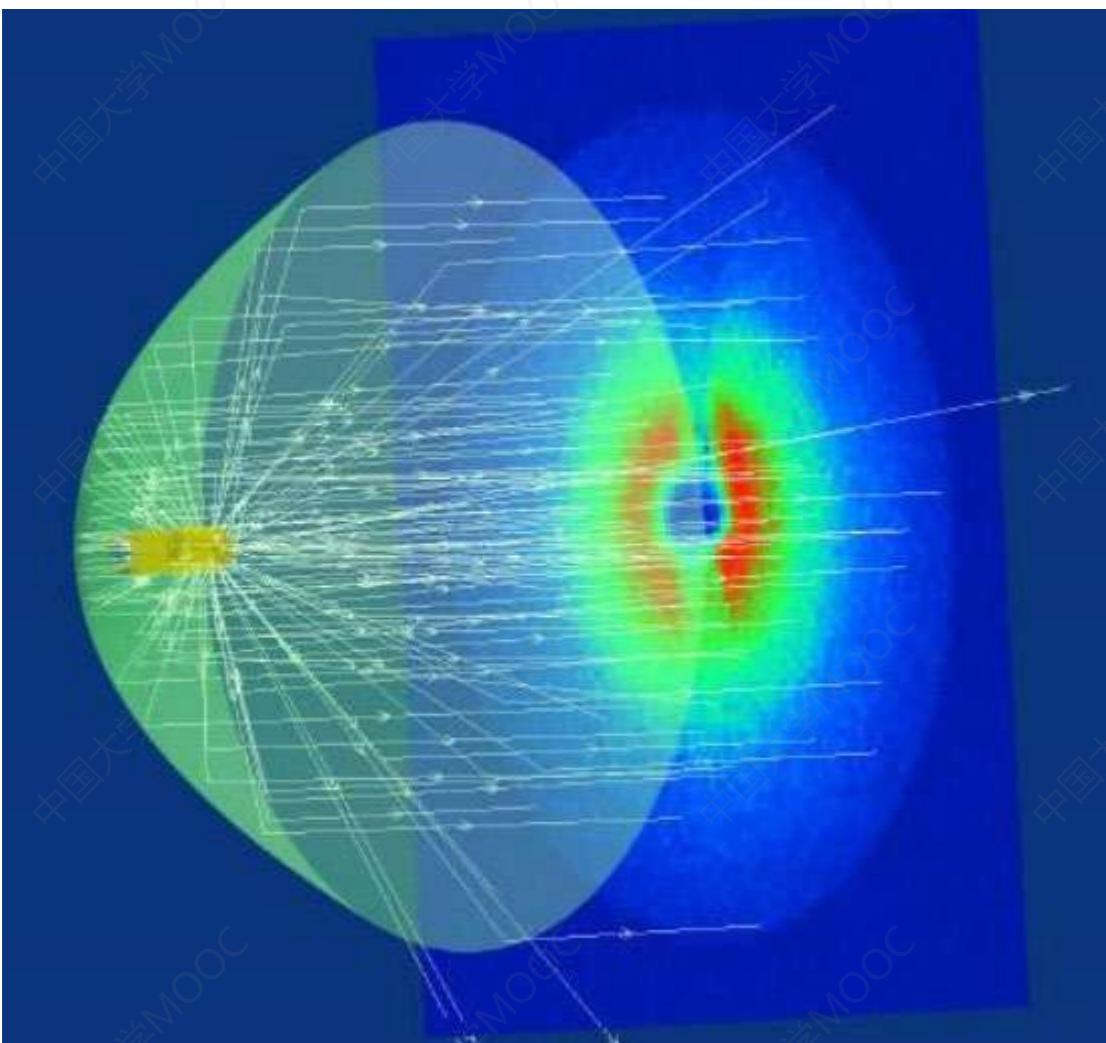


打印机



绘图仪

2.4 图像处理中的常用 术语



■ 常用术语

1. 邻域 (Neighbors)

✓ 4邻域 $N_4(p)$

✓ 对角邻域 $N_D(P)$

✓ 8邻域

$$\bullet N_8(P) = N_D(P) \cup N_4(p)$$



■ 常用术语

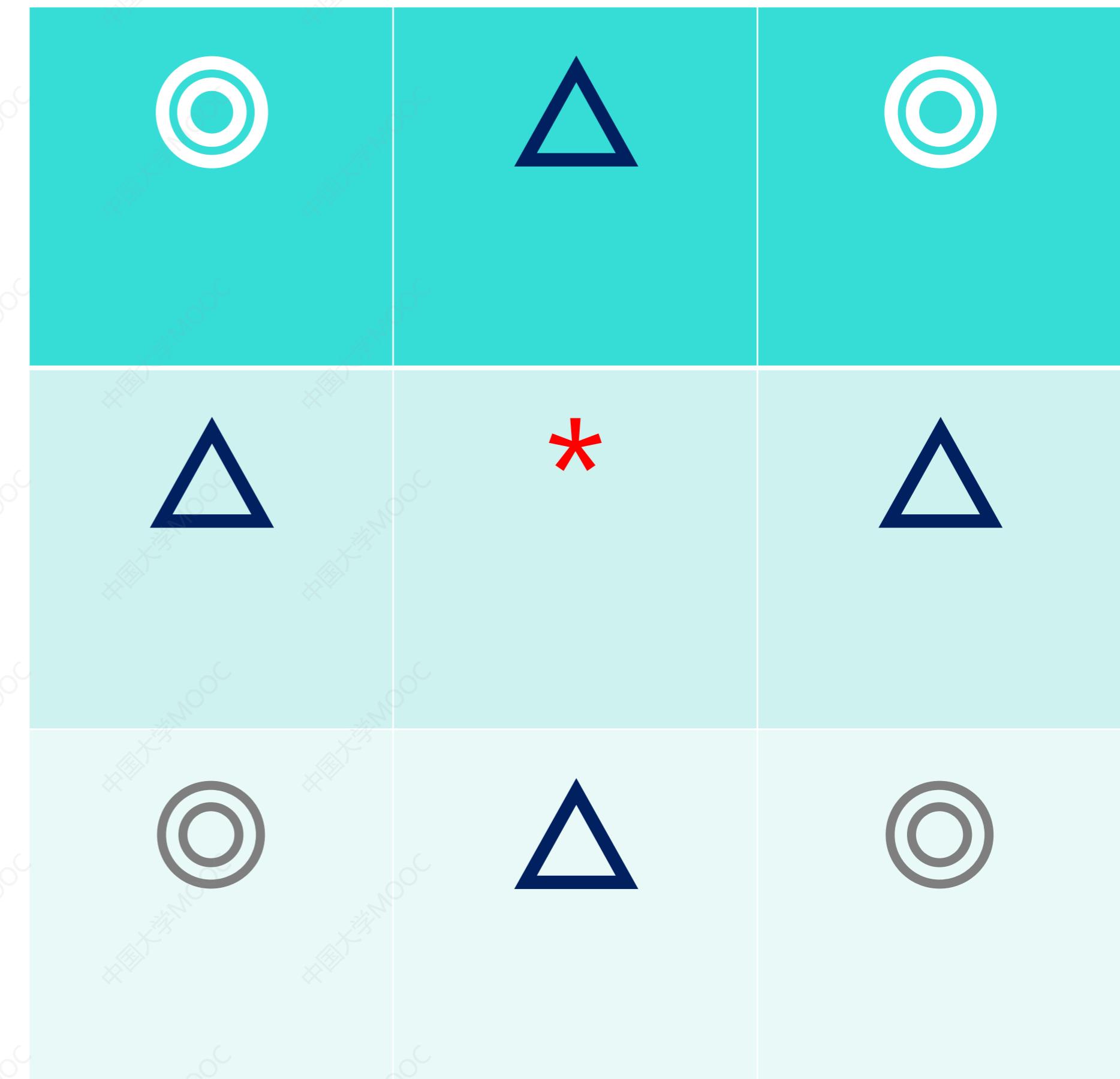
1. 邻域 (Neighbors)

✓ 4邻域 $N_4(p)$

✓ 对角邻域 $N_D(P)$

✓ 8邻域

$$\bullet N_8(P) = N_D(P) \cup N_4(p)$$



常用术语

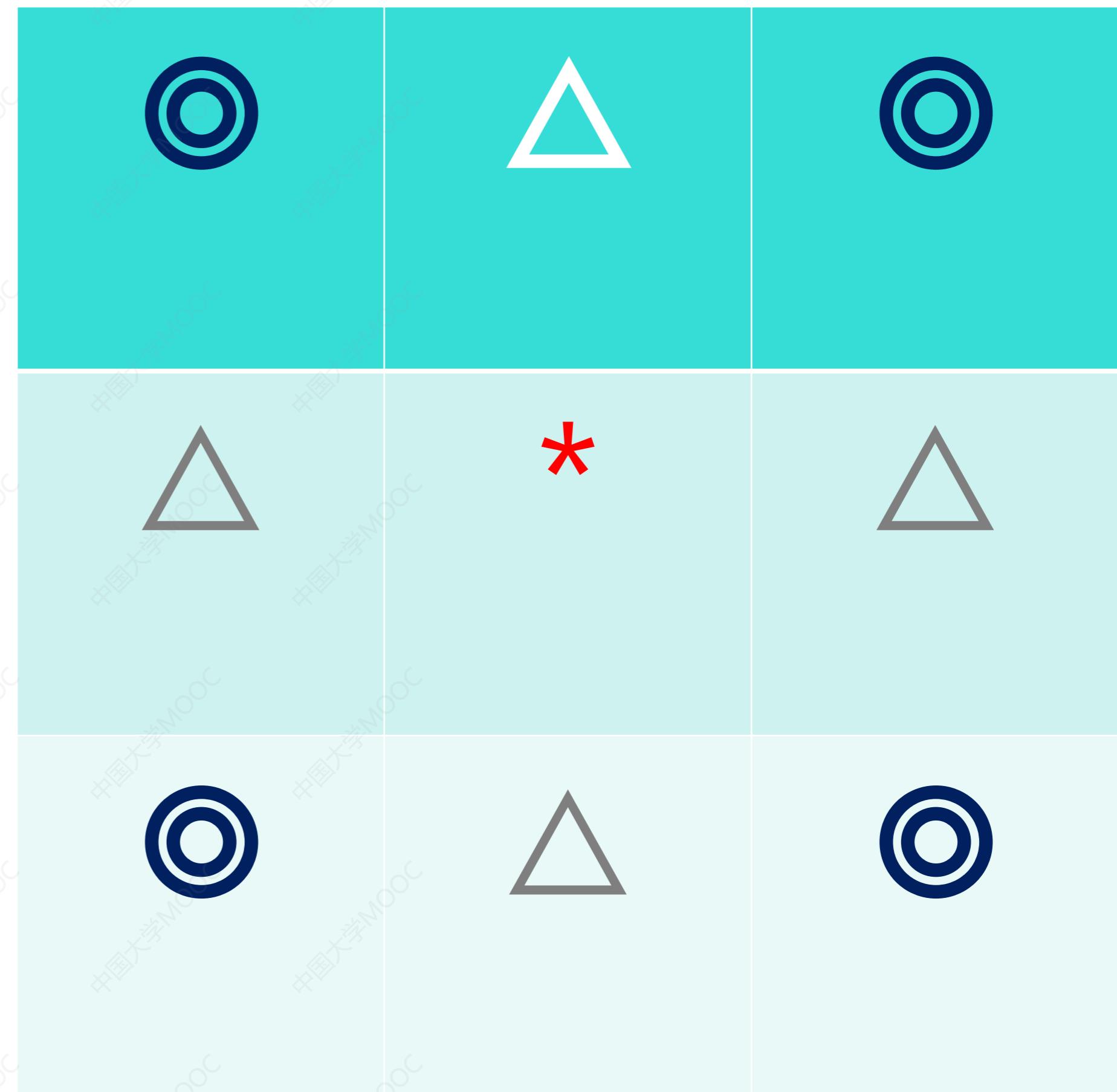
1. 邻域 (Neighbors)

✓ 4邻域 $N_4(p)$

✓ 对角邻域 $N_D(P)$

✓ 8邻域

$$\bullet N_8(P) = N_D(P) \cup N_4(p)$$



■ 常用术语

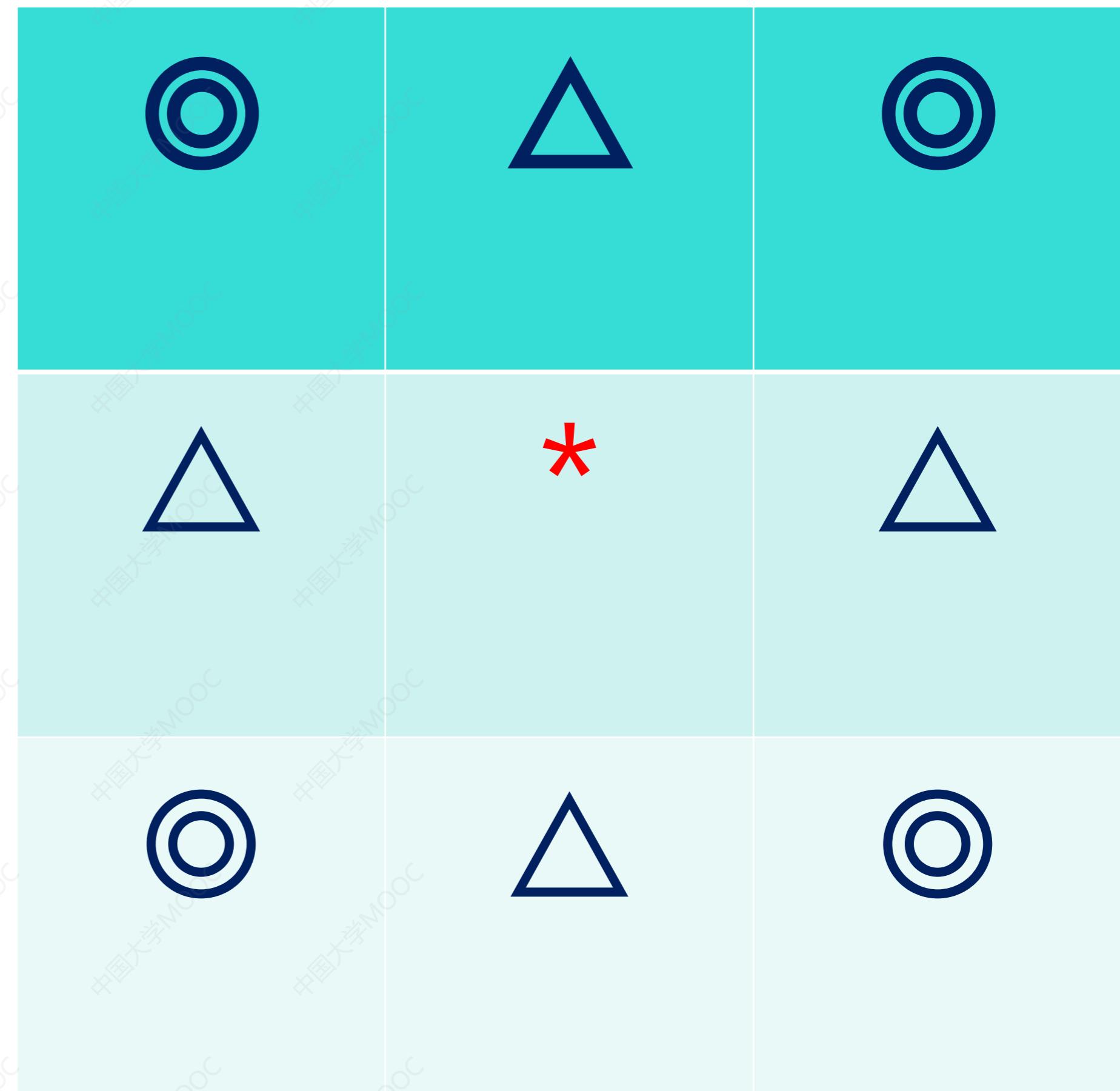
1. 邻域 (Neighbors)

✓ 4邻域 $N_4(p)$

✓ 对角邻域 $N_D(P)$

✓ 8邻域

$$\bullet N_8(P) = N_D(P) \cup N_4(p)$$



■像素间的基本关系

• 2. 邻接 (连通) (Adjacency)

✓ 相似度准则 V 是用于定义邻接性的灰度 (颜色) 值集合

- 在二值图像中, $V=\{1\}$ 或 $V=\{0\}$
- 在灰度图像中, V 可以是所有灰度值的任意子集
- 在彩色图像中, V 可以是所有颜色值的任意子集

✓ 4邻接 (连通)

- p 和 q 的灰度值 (颜色值) 均属于集合 V
- p 和 q 互为4邻域

✓ 8邻接 (连通)

- p 和 q 的灰度值 (颜色值) 均属于集合 V
- p 和 q 互为8邻域

■ 像素间的基本关系

• 3. 通路 (路径) (Path)

✓ 令：

- 像素 p 坐标为 (x, y)
- 像素 q 坐标为 (s, t)

✓ 像素 p 到像素 q 的通路是特定的像素序列

✓ $(x_0, y_0), (x_1, y_1), \dots (x_i, y_i) \dots (x_n, y_n)$

- $(x_0, y_0) = (x, y)$
- $(x_n, y_n) = (s, t)$
- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是邻接的

✓ 通路长度为 n

■像素间的基本关系

✓ 跟通路相关的几个概念

- 闭合通路

- $(x, y) = (s, t)$

- 4通路

- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是4邻接的

- 8通路

- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是8邻接的

0	1	1	q
0	1	0	
0	0	1	p

相似度准则 $V=\{1\}$

0	1	1	q	0	1	-1
0	1	0	0	1	0	
0	0	1	p	0	0	1

2条8路径长度不同，容易引起歧义

没有4通路 2条8通路

■像素间的基本关系

✓ M (Mixed) 邻接 (混合邻接/混合连通)

- p 和 q 的灰度值 (颜色值) 均属于集合V;
- p 和 q 互为4邻域, 或者 p 和 q 互为对角邻域且 p 的4邻域与 q 的4邻域的交集中的像素点的灰度值 (颜色值) 均不属于集合V;
- 目的:
 - 为了消除使用8邻接时经常出现的二义性

■像素间的基本关系

✓ 跟通路相关的几个概念

- 闭合通路

- $(x_i, y_i) = (s, t)$

- 4通路

- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是4邻接

- 8通路

- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是8邻接的

- m通路

- (x_i, y_i) 和 (x_{i-1}, y_{i-1}) 是m邻接的

✓ 跟8通路不同，m通路没有歧义性

相似度准则 $V=\{1\}$

0	1	1	q	0	1	1	0	1	1	q
0	1	0		0	1	0	0	1	0	
0	0	1	p	0	0	1	0	0	1	p

没有4通路 2条8通路 m通路

歧义

■像素间的基本关系

• 4. 连通性 (Connectivity)

- ✓ 令 S 是图像中的一个像素子集， p 和 q 是 S 中的像素
- ✓ p 和 q 是连通的
 - p 和 q 两者之间在 S 中存在通路
- ✓ 连通分量
 - 对于 S 中的任何像素 p ， S 中连通到该像素的像素集称为连通分量
- ✓ 连通集
 - 若 S 只有一个连通分量，则集合 S 称为连通集

■像素间的基本关系

• 5. 区域 (Region)

✓ 令 R 是图像的一个像素子集，如果 R 为连通集，则 R 为该图像的一个区域

• 6. 边界/轮廓 (Boundary/Contour)

✓ 令 R 为图像中的一个区域

✓ 边界 B

- B 是 R 的一个子集

- B 中每个元素的一个或多个邻域像素点不属于 R ;

✓ 如果 R 是一个完整的图像

- 它的边界 B 是图像的第一行 (列) 和最后一行 (列) 的所有像素点组成的集合

✓ 区域的边界形成闭合路径

■像素间的基本关系

• 7. 边缘 (Edge)

✓ 边缘也是一组像素点的集合，这些像素点的导数值超过预设阈值，表示灰度不连续性；

✓ 边缘和边界的区别

- 边界：一个“全局”的定义，与区域相对应；

- 边缘

- 一个局部的定义；

- 表达某一像素点与其相邻像素点灰度变化不连续；

- 在二值图像中

- 边缘和边界通常对应

- 在灰度图像中

- 边缘和边界不完全等价，但二者有很大相关性

■ 距离

✓ 对于像素点 p, q , 和 z , 其坐标分别为 $(x, y), (s, t), (v, w)$, D 如果满足以下条件, 则 D 是一个距离函数 (度量函数)

- (a) $D(p, q) \geq 0$ (如果 $p=q$, $D(p, q) = 0$)
- (b) $D(p, q) = D(q, p)$
- (c) $D(p, z) \leq D(p, q) + D(q, z)$.

■ 像素间的基本关系

- 1. 欧几里得距离 (Euclidean distance)

$$D_e(p, q) = \sqrt{(x - s)^2 + (y - t)^2}$$

- 与像素点 (x, y) 的欧氏距离小于等于 r 的像素点，是所有包含在以 (x, y) 为圆心、半径 r 的圆盘中的像素点

- 2. D_4 距离 (城区距离) (D4 distance/City-block distance)

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

- 与像素点 (x, y) 的 D_4 距离小于等于某个值 r 的像素点，形成以 (x, y) 为中心的菱形区域

- 3. D_8 距离 (棋盘距离) (D8 distance/ Chessboard distance)

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

- 与像素点 (x, y) 的 D_8 距离小于等于某个值 r 的像素点，形成一个以 (x, y) 为中心的正方形区域

■像素间的基本关系

	2			
2	1	2		
2	1	0	1	2
2	1	2		
2				

$$D_4 \leq 2$$

菱形区域

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

$$D_8 \leq 2$$

正方形区域

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

$$D_e \leq 3$$

圆形区域

■ 像素间的基本关系

• 4. 路径长度 (Length of the path) 和距离 (Distance)

- 两点间的 D_m 距离定义为两点间的最短 m 通路

D_m 距离取决于通路上像素的值以及它们邻域像素的值

0 1 1 (q)

1 1

1 (p)

$$D_m=4$$

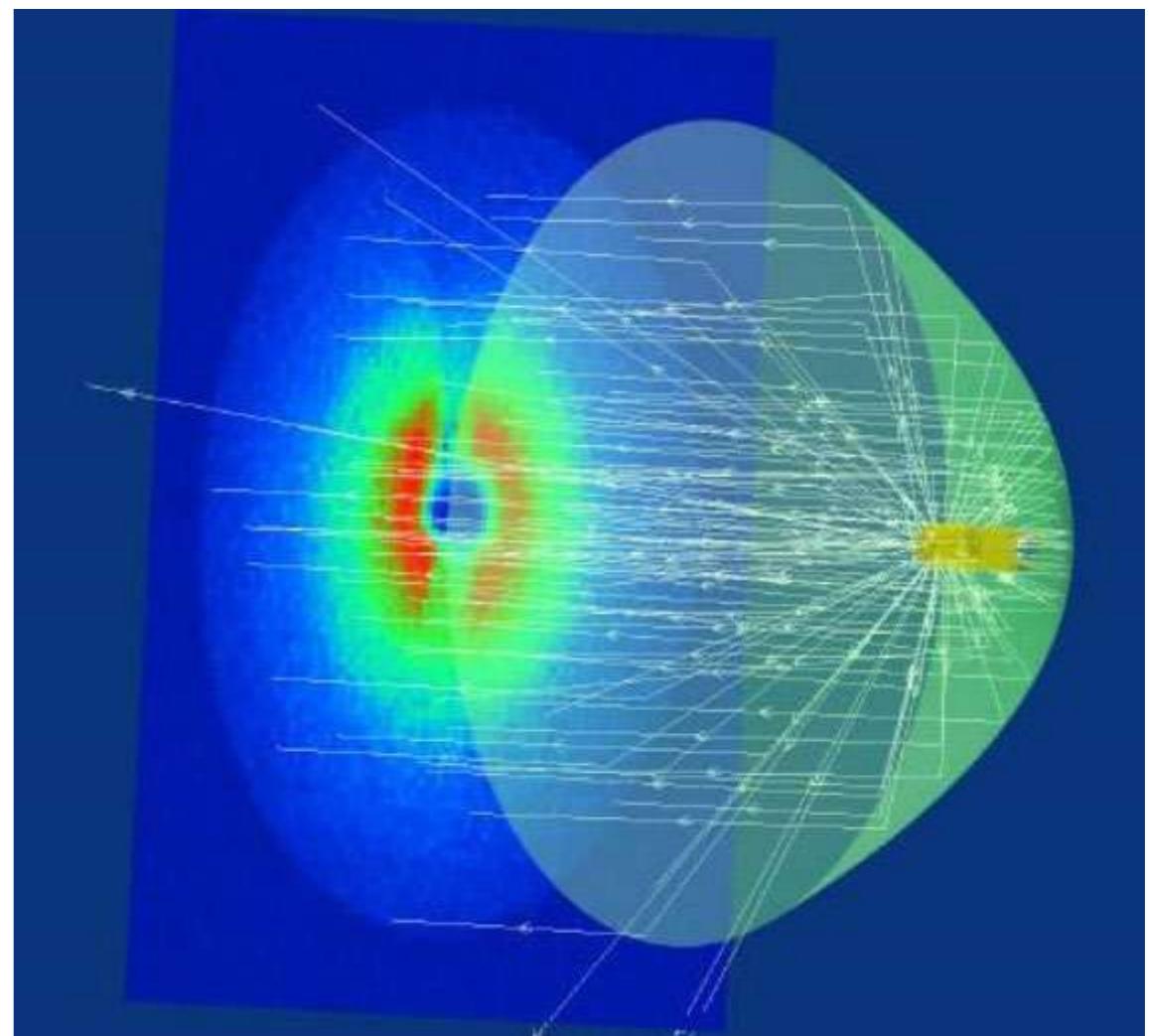
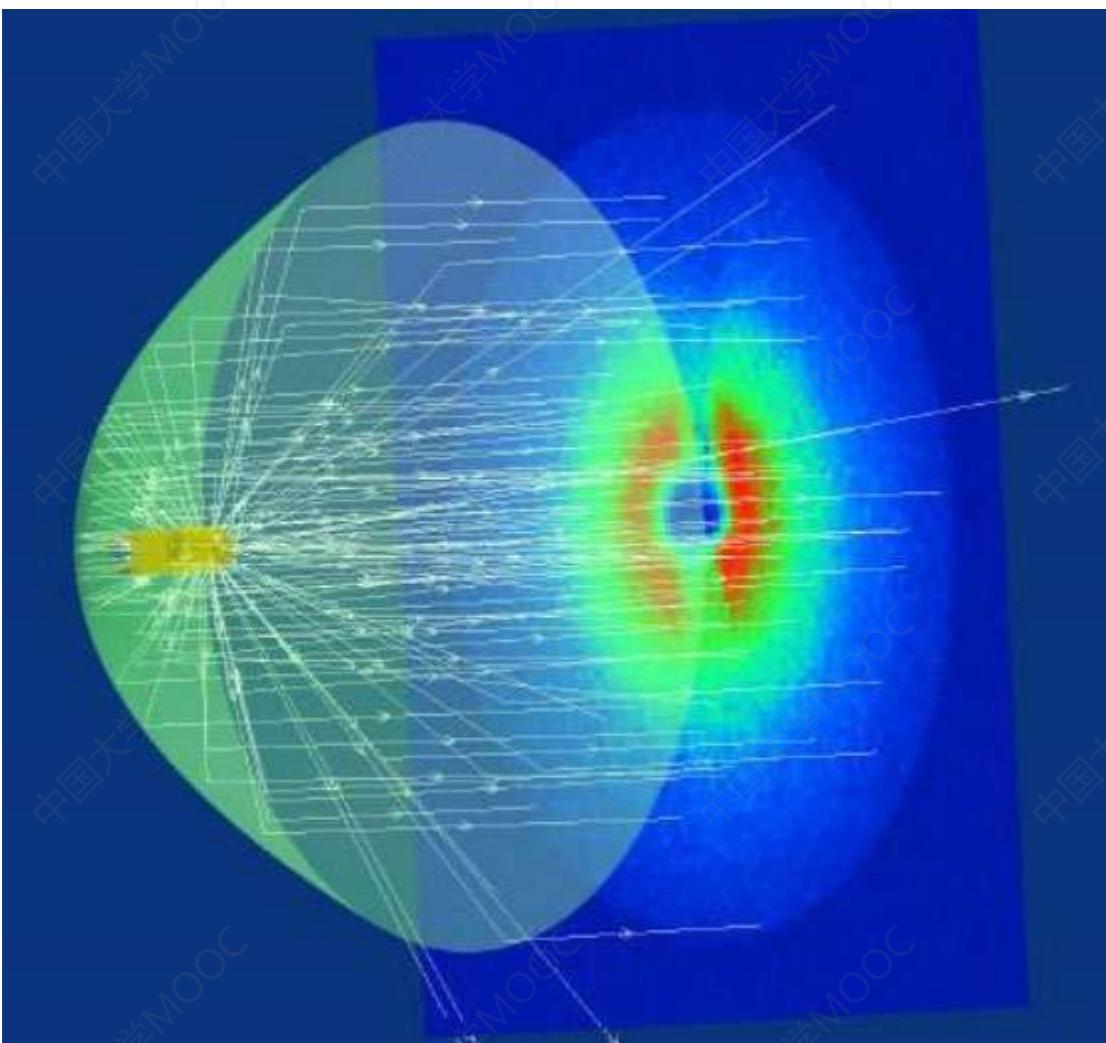
0 0 1 (q)

0 1

1 (p)

$$D_m=2$$

2.5 图像质量评价



图像质量评价

➤ 逼真度(fidelity): 设计能精确、自动感知图像质量的计算

模型，获得图像质量的定量表示。客观评价方式。

- 全参考图像质量评价
- 半参考图像质量评价
- 无参考图像质量评价

图像质量评价

- 可懂度(intelligibility): 以人为观察者, 对图像质量进行主观定性评价。主观评价方式。
- 绝对评价方式: 观察者根据自己的知识和理解, 参照原始图像对待定图像按照某些特定评价性能对图像的绝对好坏进行评价。
 - 相对评价方式: 没有原始图像作为参考, 观察者对一批待评价图像进行相互比较, 从而判断出每个图像的优劣顺序, 并给出相应的评价值;
 - 评价指标: 平均主观得分 (MOS: Mean Opinion Score)

图像质量评价

- 全参考图像质量评价 (FR: Full Reference)
 - 评价一幅图像和标准图像的偏离程度；
- 半参考图像质量评价 (RR: Reduced Reference)
 - 只提取标准图像的部分信息作为参考，对失真图像质量进行评价，该方法的关键性能指标是特征提取和特征比较；
- 无参考图像质量评价 (NR: No Reference)
 - 在没有标准图像的情况下，直接根据参考质量评价指对失真图像质量进行评价；

■ 逼真度评价方式(1)

- 归一化均方误差(NMSE) Normalized Mean Square Error

$$NMSE = \frac{\sum \sum (f(x,y) - f'(x,y))^2}{\sum \sum (f(x,y))^2}$$

- 归一化信噪比(NSNR)

$$NSNR = -10 \log_{10} (NMSE)$$

■ 逼真度评价方式(2)

➤ 峰值均方误差(PMSE) Peak Mean Square Error

$$PMSE = \frac{\frac{1}{MN} \sum \sum (f(x,y) - f'(x,y))^2}{A^2}$$

- 其中A为 $f(x,y)$ 的最大取值， M,N分别为图像的宽度和高度。
- 峰值信噪比

$$PSNR = -10 \log_{10}(PMSE) \text{ dB}$$

逼真度评价方式(3)

- 结构相似度评价方法SSIM(Structure Similarity)
- 结构相似性，设两幅图像x,y

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

亮度相似度

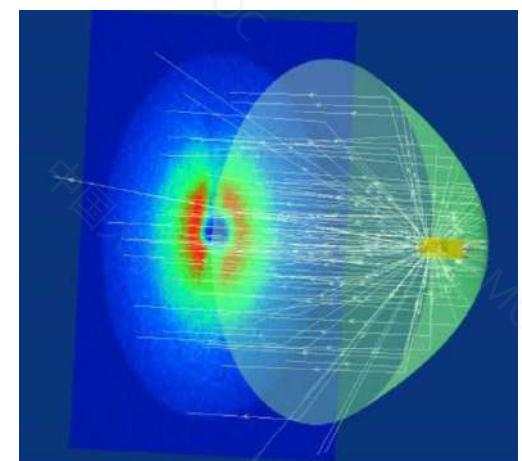
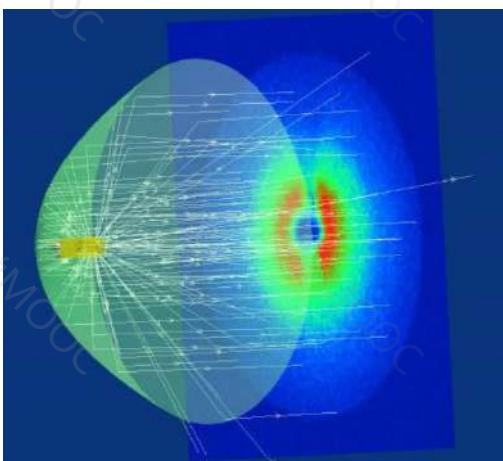
对比度相似度

结构相似度

$$\begin{aligned} SSIM(x, y) &= [l(x, y) \cdot c(x, y) \cdot s(x, y)] \\ &= \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \end{aligned}$$

- μ_x 为 x 的均值 • σ_x^2 为 x 的方差
- μ_y 为 y 的均值 • σ_y^2 为 y 的方差
- σ_{xy} 为 x 和 y 的协方差
- $c_1 = (k_1 L)^2, c_2 = (k_2 L)^2$ 为常数
- L 为像素值的范围, $2^B - 1$
- $k_1 = 0.01, k_2 = 0.03$ 为默认值

第三章 图像空间域 增强



■ 图像增强

- 面向某一目标对图像感兴趣的部分进行处理，以改善图像的视觉质量。
- 面向目标，不考虑降质原因（基于人类主观偏好）设计方法
- 特点
 - 方法设计比较主观
 - 突出图像中感兴趣的部分,达到局部最优；
 - 方法相对简单，不需要估计图像降质过程，研究相对活跃；

图像增强

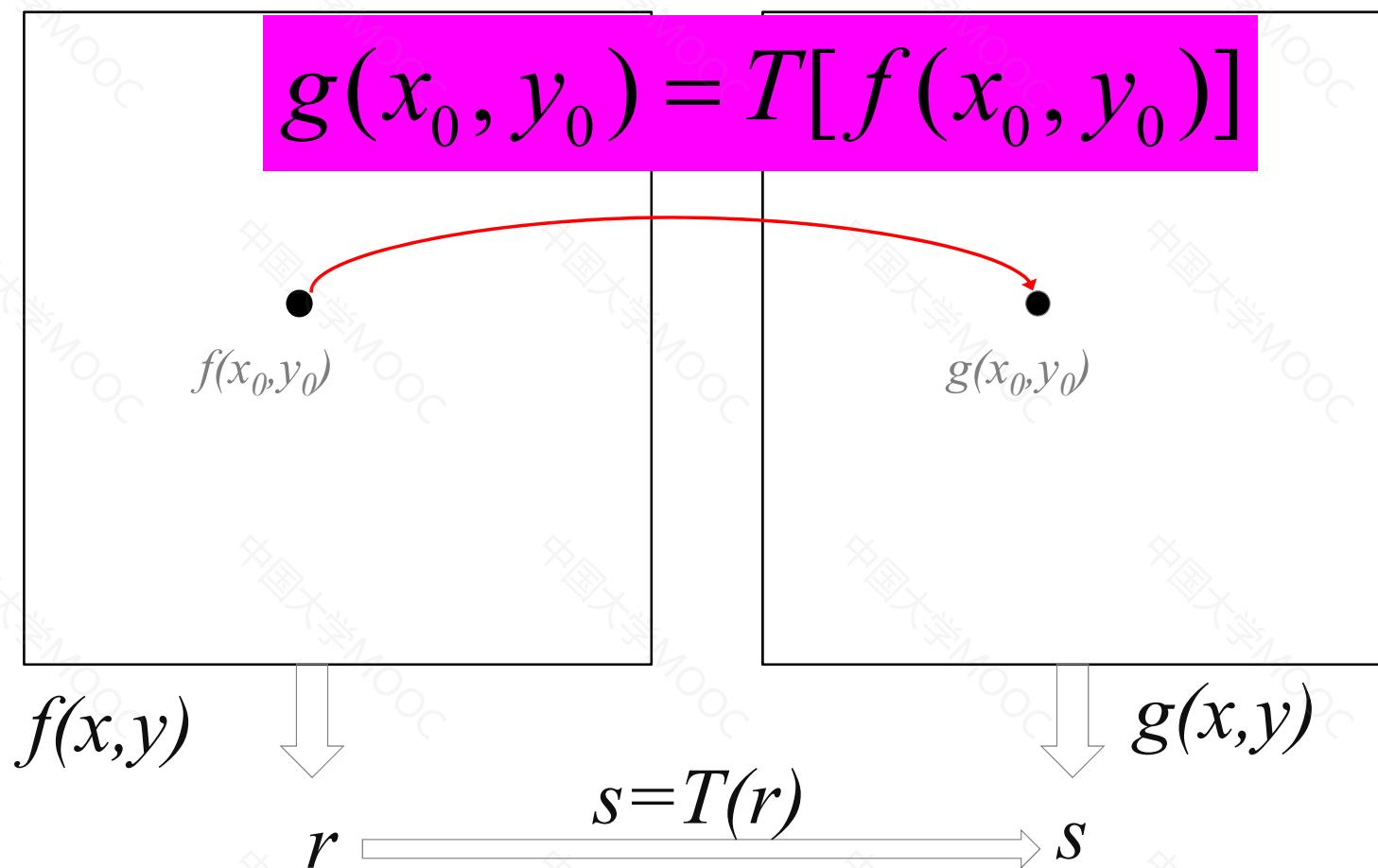
➤ 图像增强的方法：

- 空间运算：基于图像像素点灰度值进行增强处理
- 变换运算：从空间域变换到新的域里进行处理，处理结束后将处理结果再映射回空间域
- 伪彩色增强：把灰度图像映射到彩色空间，常用于遥感图像处理、医学图像处理

■ 引言

- 3.1 点运算
- 3.2 图像滤波
- 3.3 伪彩色增强

■3.1 点运算



■ 3.1 点运算

点运算的约束条件

- 1) 对于 $0 \leq r \leq 255$, 有 $0 \leq s \leq 255$
- 2) 在 $0 \leq r \leq 255$ 区间内, $s = T(r)$ 单值单调增长。



■3.1 点运算

- 3.1.1 简单的灰度变换
- 3.1.2 直方图均衡
- 3.1.3 其他直方图处理方法

■ 3.1.1 简单的灰度映射

简单的灰度映射，基于已有函数，将其特定化为问题适合的映射函数；

1. 对数变换 (\log)

函数形式： $s=c\log(a+r)$, a,c 为常量待定

令： $r=0,s=0;$

$r=1,s=1;$

用这两个约束条件，即得到两个公式

$$0=c\log(a)$$

$$1=c\log(a+1)$$

计算得到：

$$a=1;$$

$$c=1/\log(2);$$

映射公式 $s=1/\log(2)*\log(1+r)$

■ 3.1.1 简单的灰度映射

⑩ 2. 指数变换 (Inverse log)

$$s = c(e^r - a)$$

$$r=0 \rightarrow s=0, a=1$$

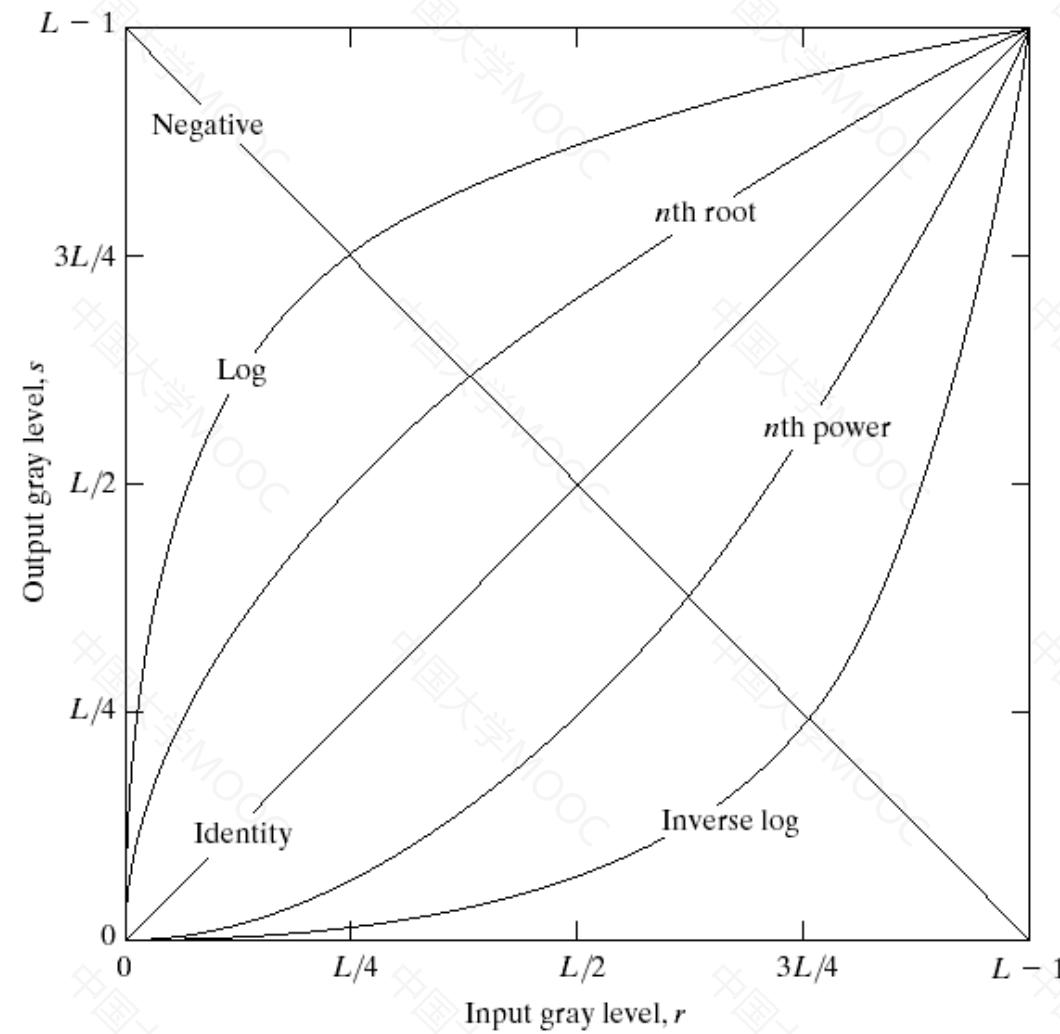
$$r=1 \rightarrow s=1, c = \frac{1}{e-1}$$

⑩ 3. 幂指数变换 (Power law)

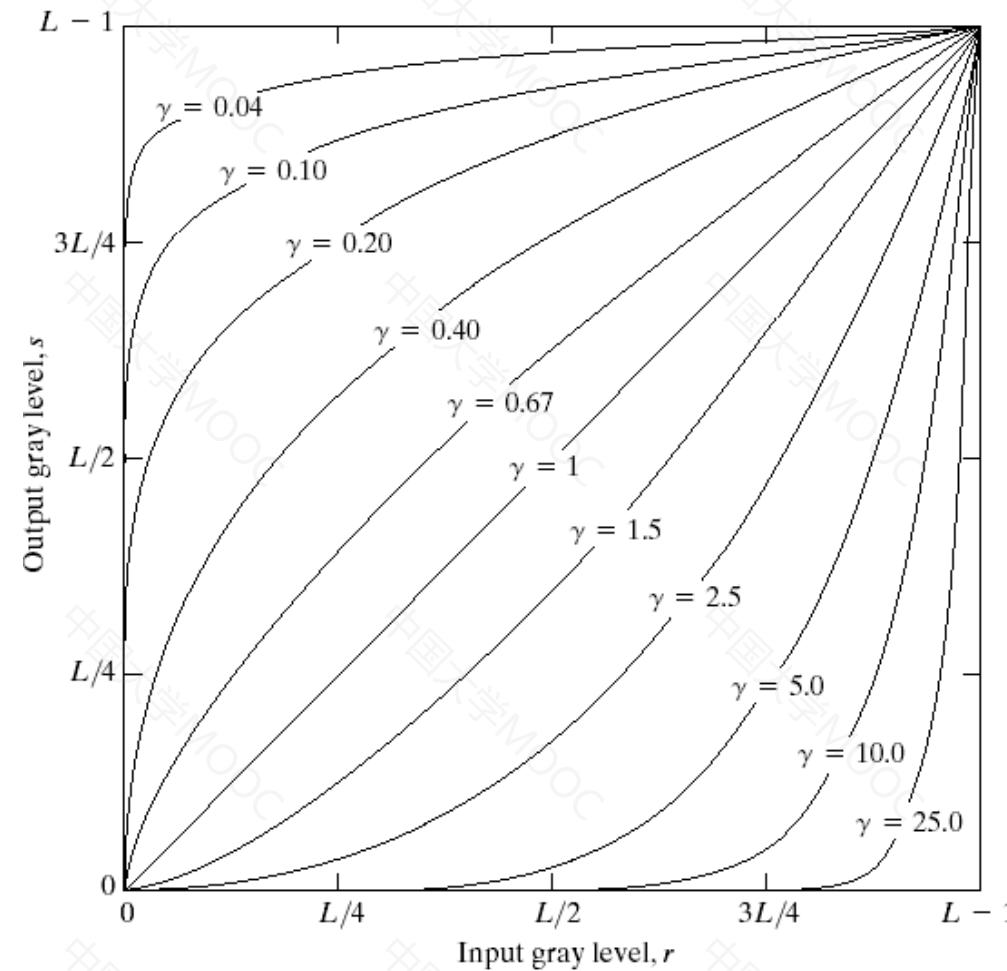
$$s = cr^n$$

$$r=1 \rightarrow s=1, c=1$$

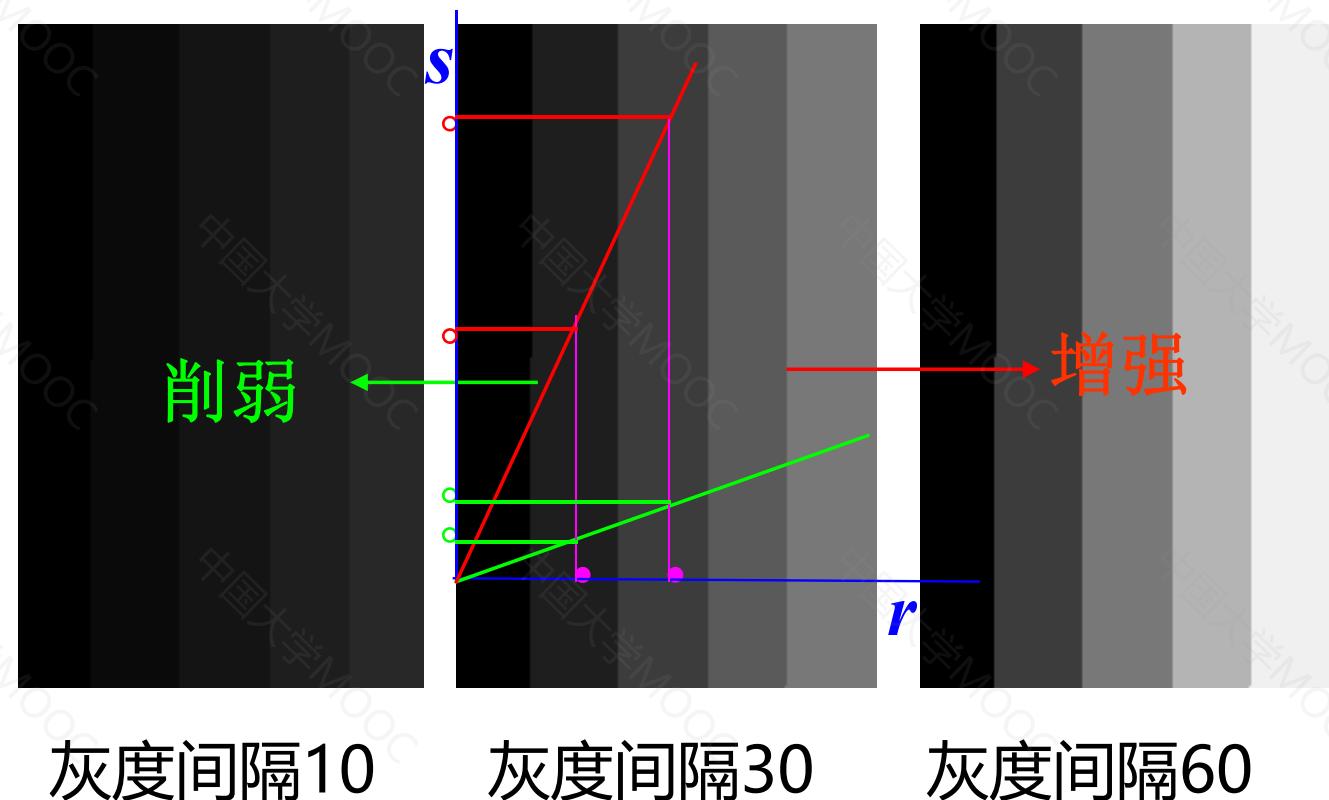
■ 3.1.1 简单的灰度映射



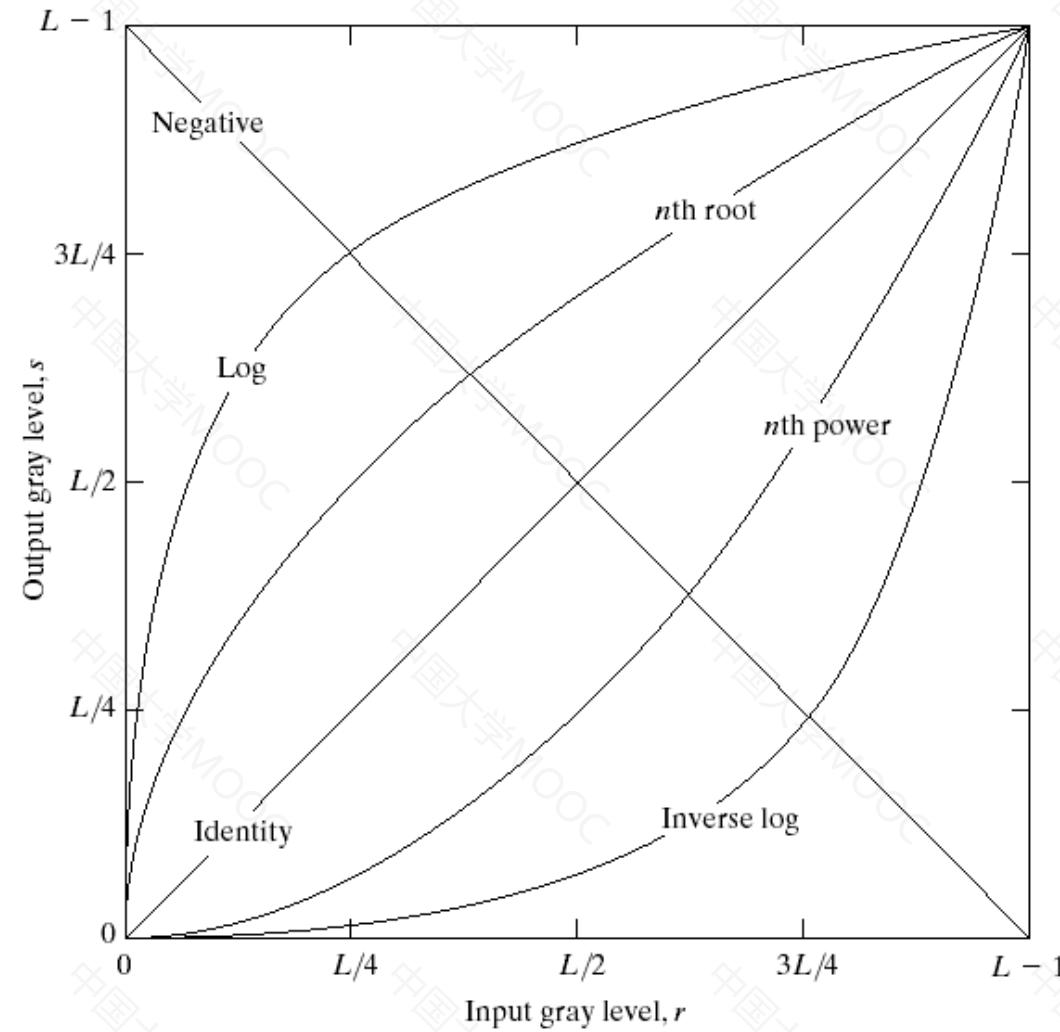
■3.1.1简单的灰度映射



■ 点运算中增强和削弱的概念

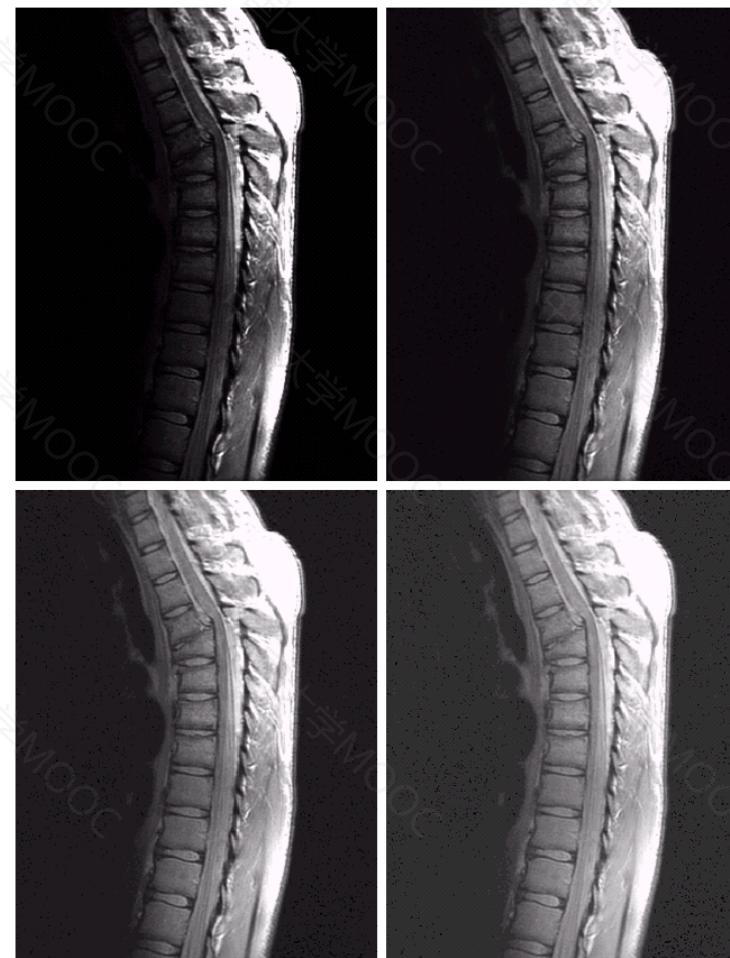


■ 3.1.1 简单的灰度映射



■ 类对数增强举例

输入
图像



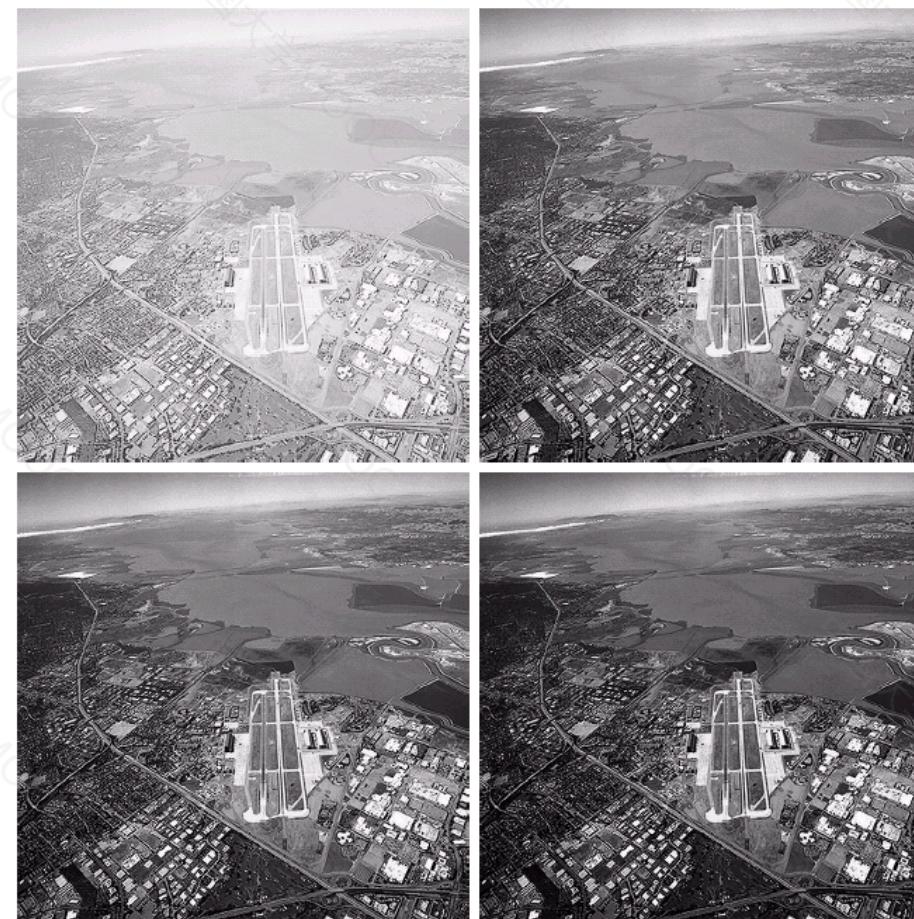
$c=1$
 $\gamma=0.6$

$c=1$
 $\gamma=0.4$

$c=1$
 $\gamma=0.3$

■类指数增强举例

输入
图像



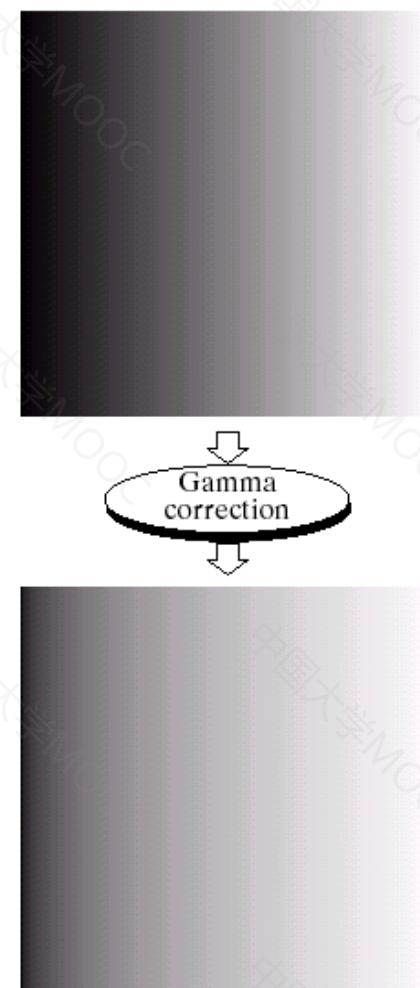
$c=1$
 $\gamma=3.0$

$c=1$
 $\gamma=4.0$

$c=1$
 $\gamma=5.0$

■伽马校正

输入
图像



伽马校
正曲线

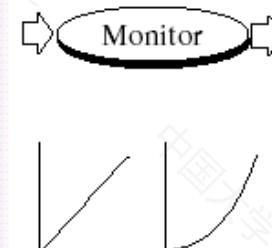
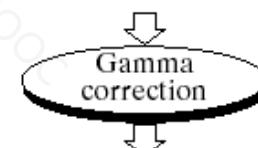
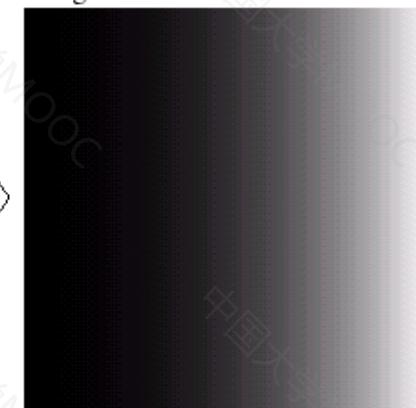


Image as viewed on monitor



显示器
直接显
示结果

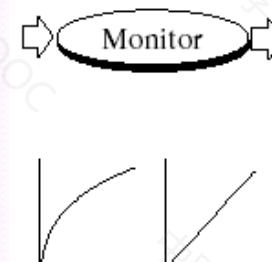
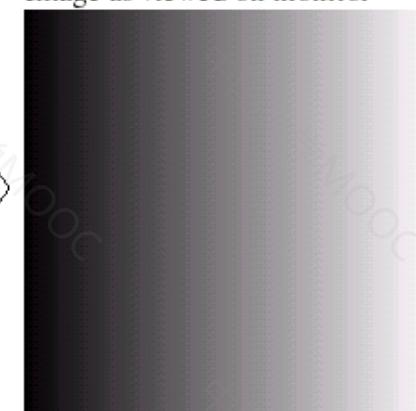


Image as viewed on monitor



伽马校
正后显
示结果

■ 3.1.2 直方图均衡

1. 直方图

2. 直方图均衡

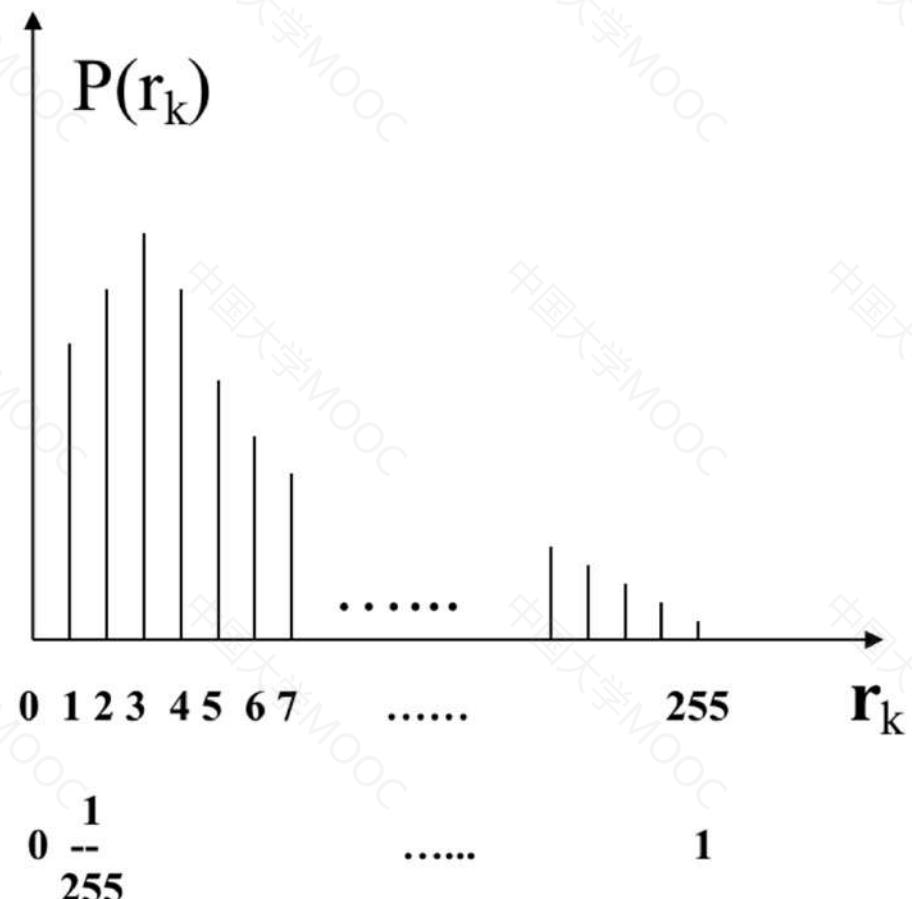
■ 3.2 直方图均衡

► 1. 直方图

图像直方图表示图像中不同灰度级象素出现的相对频率，是图像灰度分布概率的估值。

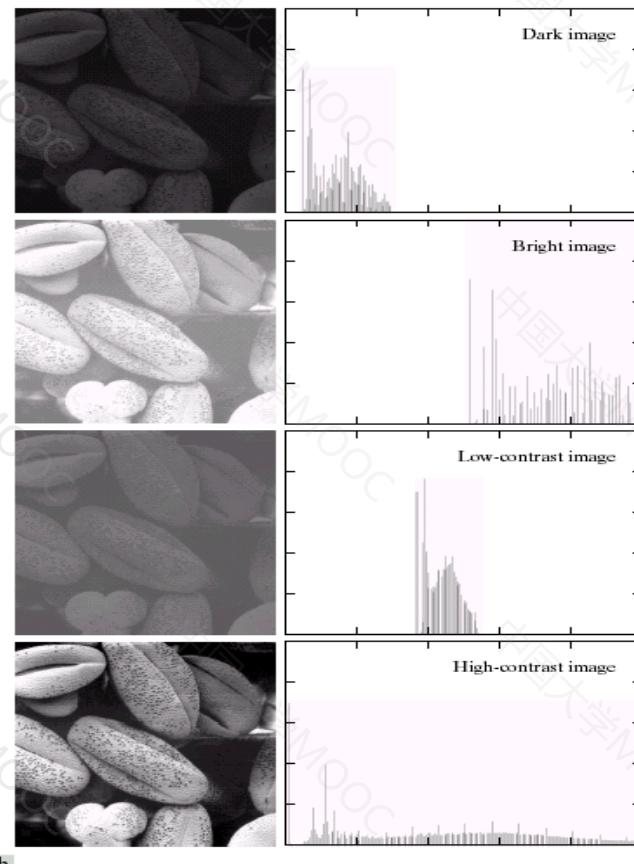
$$P(r_k) = n_k/n$$

$$\sum_{k=0}^{L-1} p(r_k) = 1$$



直方图和归一化直方图

图像变换



a b

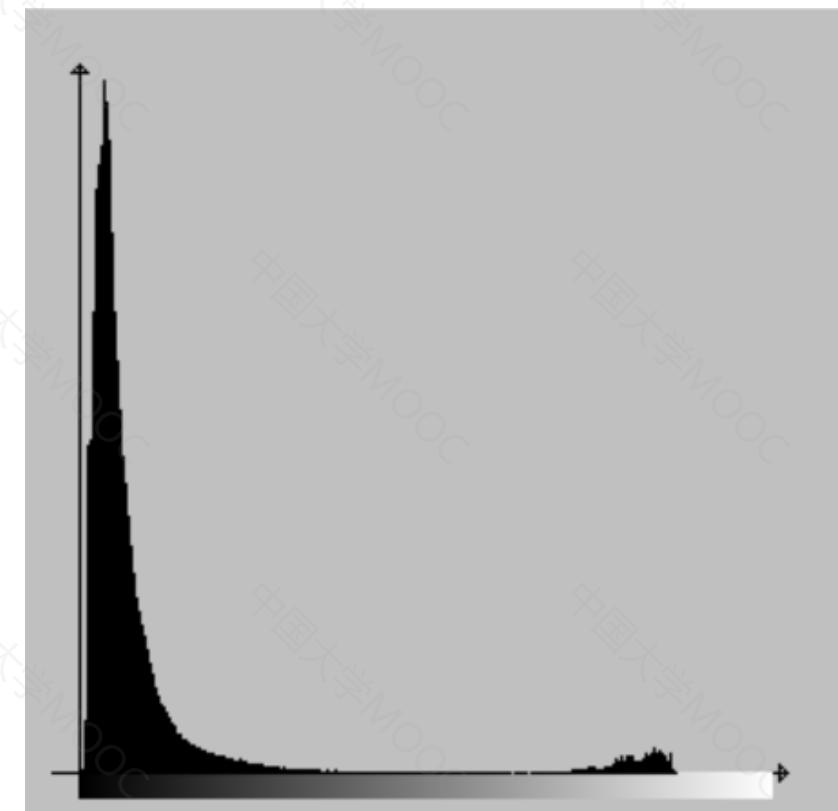
FIGURE 3.15 Four basic image types: dark, light, low contrast, high contrast, and their corresponding histograms. (Original image courtesy of Dr. Roger Heady, Research School of Biological Sciences, Australian National University, Canberra, Australia.)

- 图像偏暗 \leftrightarrow 直方图集中在低端
- 图像偏亮 \leftrightarrow 直方图集中在高端
- 图像对比度低 \leftrightarrow 直方图分布集中
- 质量很好的图像 \leftrightarrow 直方图分布在整个灰度范围，而且趋近于均匀分布

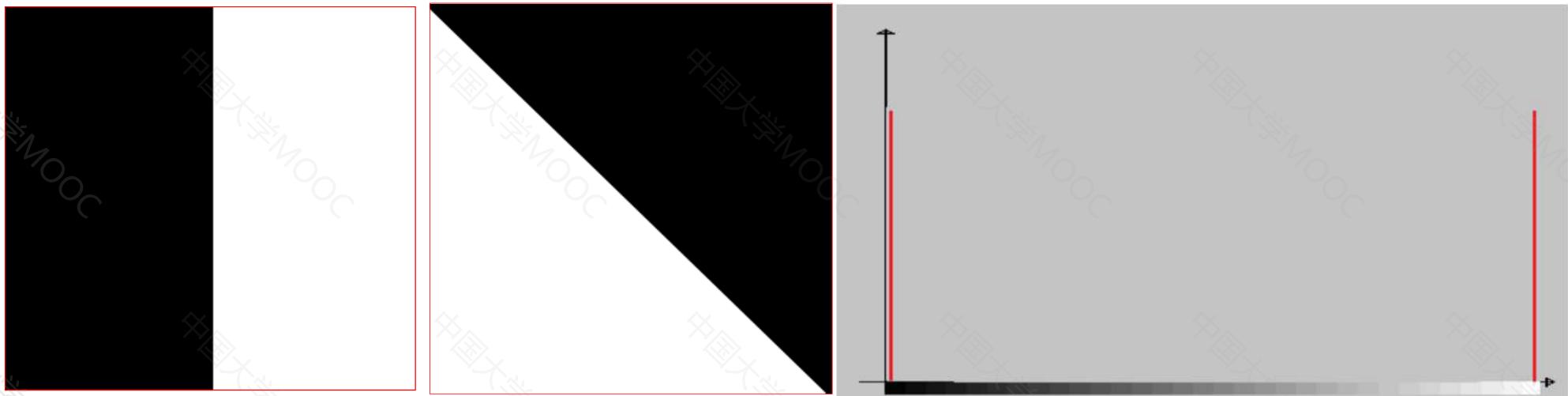
直方图是图像灰度分布统计特性的反映



一幅整体偏暗的图像

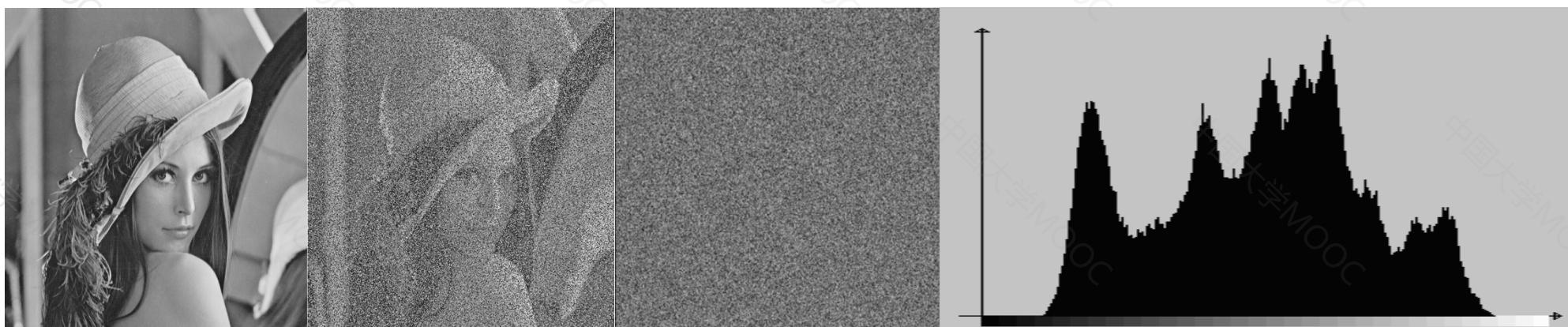


左边图像对应的直方图



两个图像各50%的像素点灰度为0和255

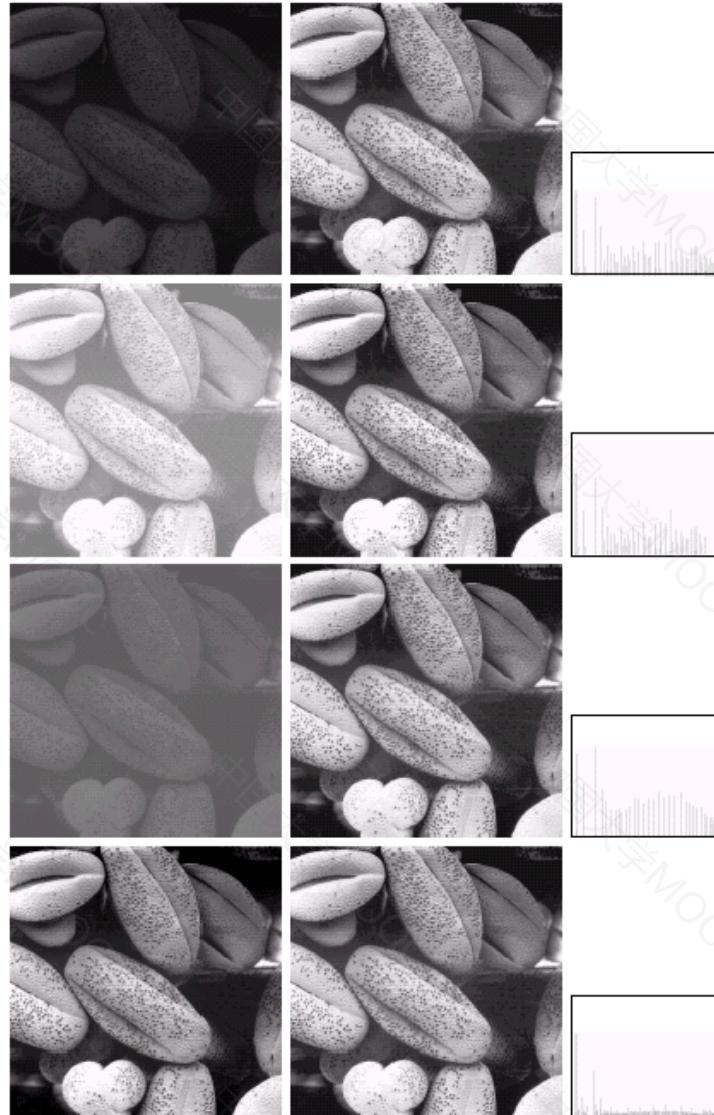
左侧两个图像的直方图



Lena图像

70%像素点互换位置 100%像素点互换位置

左侧三个图像的直方图



■直方图均衡(Histogram Equalization)

直方图均衡是由一幅灰度概率分布已知的图像，生成一幅灰度概率分布均匀的新图像。

随机变量r的概率密度函数 $p_r(r)$ ，随机变量s是r的函数 $s=T(r)$ ，请问随机变量s的概率密度函数如何计算？

$$p_s(s) = [p_r(r) \frac{dr}{ds}]_{r=T^{-1}(s)}$$

直方图均衡问题表达：已知一幅图像 $f(x,y)$ ，他的灰度分布概率 $p_r(r)$ 已知，通过 r 到 s 的映射， $s=T(r)$ ，生成一个新图像 $g(x,y)$ ，他的灰度分布概率 $P_s(s)$ 等于常量。

■直方图均衡的基本原理

$$p_s(s) = [p_r(r) \frac{dr}{ds}]_{r=T^{-1}(s)} = \text{常数}$$

$$\frac{dr}{ds} = \frac{1}{p_r(r)}$$

■直方图均衡的基本原理

$$p_s(s) = [p_r(r) \frac{dr}{ds}]_{r=T^{-1}(s)} = 1$$

$$\frac{dr}{ds} = \frac{1}{p_r(r)}$$

■直方图均衡的基本原理

$$p_s(s) = [p_r(r) \frac{dr}{ds}]_{r=T^{-1}(s)} = 1$$

$$\frac{dr}{ds} = \frac{1}{p_r(r)}$$

$$\frac{ds}{dr} = p_r(r)$$

$$s = T(r) = \int_0^r p_r(t) dt$$

■直方图均衡的基本原理

➤ 灰度映射公式

$$s = T(r) = \int_0^r p_r(t) dt$$

➤ 实际计算公式

$$s_k = \sum_{n=0}^k p_r(r_n)$$

$$s_0 = p_r(r_0)$$

$$s_1 = p_r(r_0) + p_r(r_1)$$

$$s_k = p_r(r_0) + p_r(r_1) + p_r(r_2)$$

$$s_k = p_r(r_0) + p_r(r_1) + \dots + p_r(r_k)$$

■ 2 计算举例(1)

例1:设图像灰度分布概率如下：

r	0	1	2	3	4	5	6	7
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

试进行直方图均衡，给出灰度映射关系以及均衡后图象的直方图分布。

■ 2 计算举例(1)

➤ 1. 对直方图进行归一化

r	0	1/7	2/7	3/7	4/7	5/7	6/7	1
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

➤ 2. 灰度映射

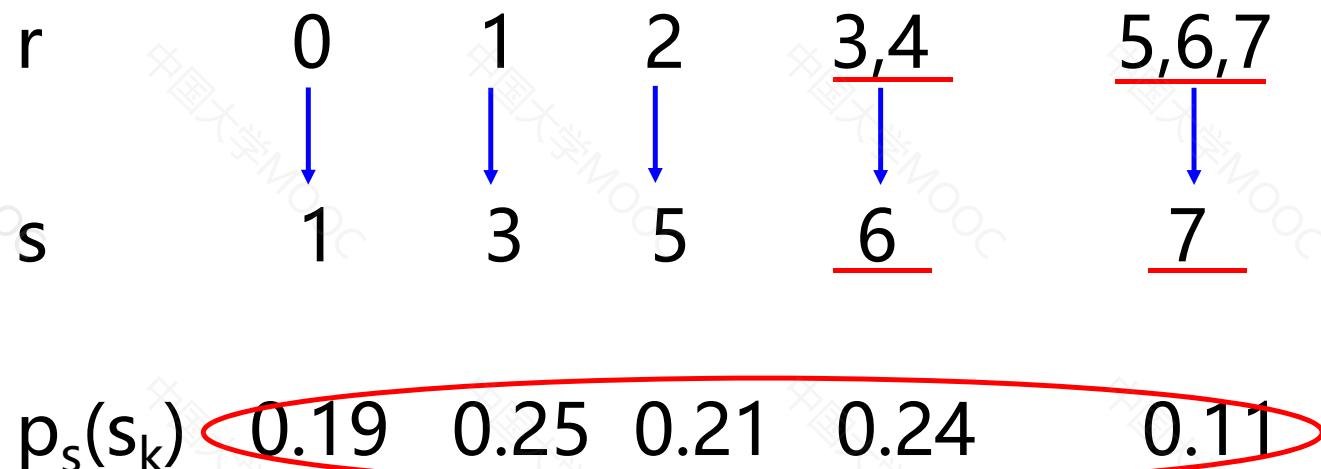
s'	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
------	------	------	------	------	------	------	------	------

➤ 3. 量化

s	1/7	3/7	5/7	6/7	6/7	1	1	1
-----	-----	-----	-----	-----	-----	---	---	---

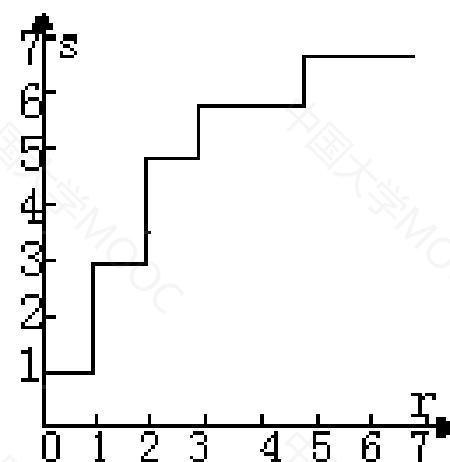
■ 2 计算举例(1)

► 4. r -- s 映射关系及 s 的灰度分布概率

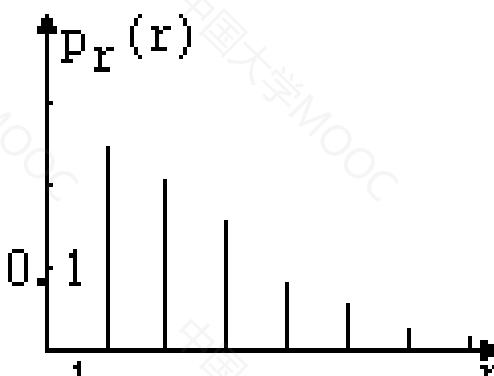


■ 2 计算举例(1)

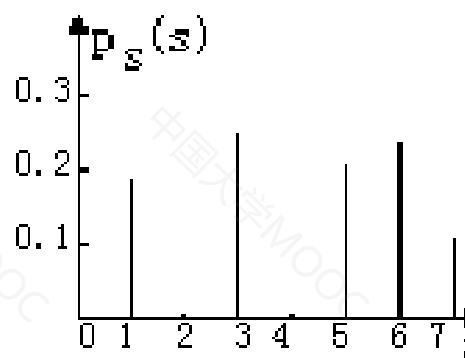
➤ 5. 图形显示



r--s 映射关系图



原图直方图分布



均衡后直方图分布

■ 3 实例演示

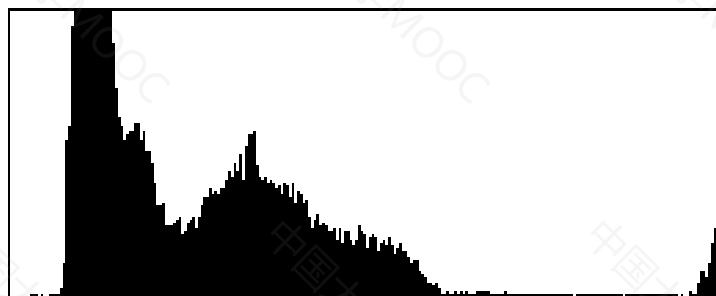


原始图象



均衡后图象

■ 3 实例演示



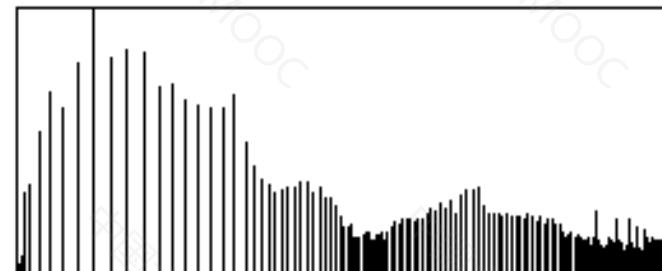
平均值: 67.83

非零灰度数: 241

最小灰度数: 2 (灰度7)

最大灰度数: 929 (灰度26)

最大最小灰度数比: 464.5



平均值: 127.75

非零灰度数: 132

最小灰度数: 34 (灰度0)

最大灰度数: 929 (灰度29)

最大最小灰度数比: 27.32

■直方图均衡

总结

- 直方图均衡实现高概率增强、低概率削弱
- 直方图均衡存在灰度级简并现象
- 均衡后图象直方图不是严格均匀分布，只是近似均匀分布
- 直方图均衡重复做有意义吗？
- 点运算中的局部优化效果讨论

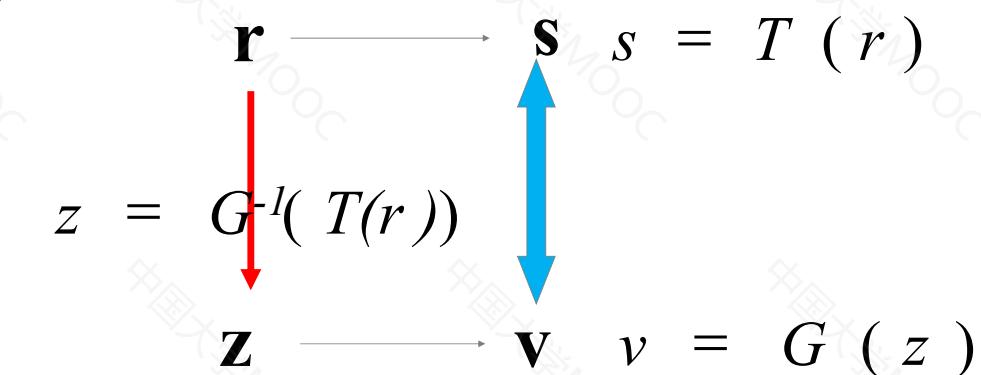
■ 3.1.3 其他直方图处理方法

直方图规定化处理 (histogram matching/specification)

已知一幅图像 $f(x,y)$, 它的灰度分布概率 $p_r(r)$ 已知, 通过 r 到 z 的映射, $z=T(r)$, 生成一个新图像 $w(x,y)$, 他的灰度分布概率 $P_z(z)$ 符合规定的形状。

$P_r(r)$: 输入图像的概率分布

$P_z(z)$: 输出图像的目标概率分布



■计算举例 (1)

例1:设输入图像灰度分布概率如下:

r	0	1	2	3	4	5	6	7
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

输出图像的目标灰度概率分布如下

z	0	1	2	3	4	5	6	7
$p_z(z_k)$	0	0	0	0.15	0.20	0.30	0.20	0.15

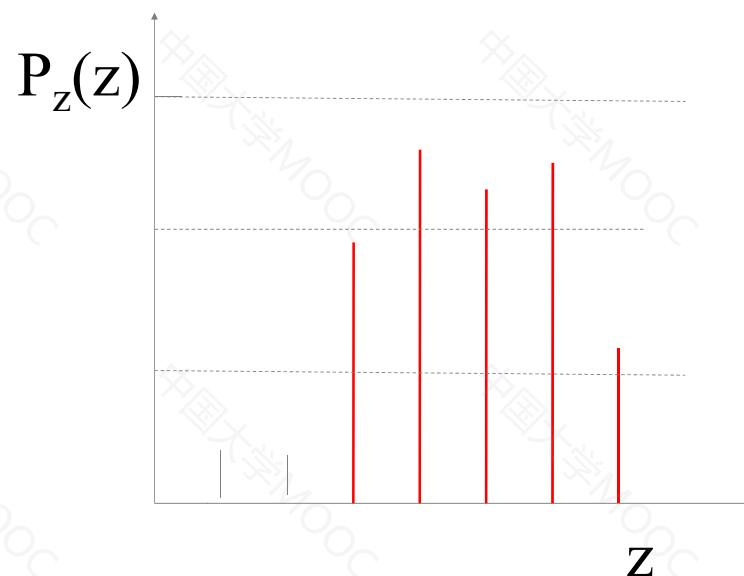
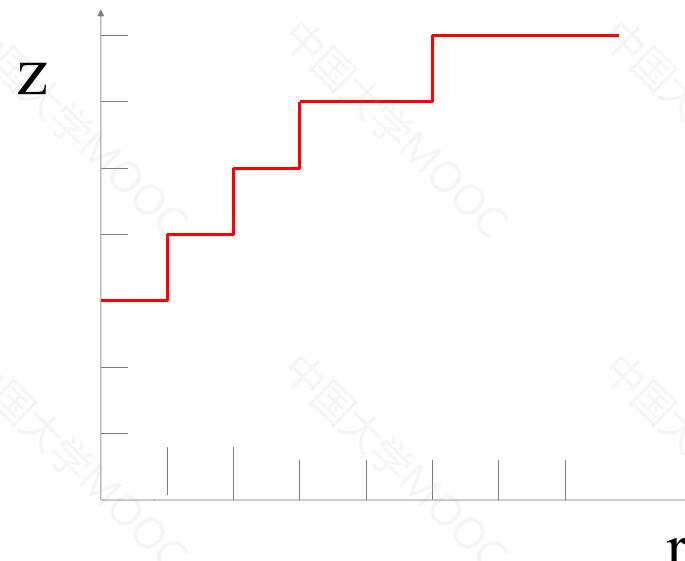
试进行直方图规定化处理，给出灰度映射关系以及输出图象的直方图分布。

■ 3.1.3 其他直方图处理方法

r	1/7	2/7	3/7	4/7	5/7	6/7	1	
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
v	0	0	0	0.15	0.35	0.65	0.85	1.0
$p_z(z_k)$	0	0	0	0.15	0.20	0.30	0.20	0.15
z	0	1/7	2/7	3/7	4/7	5/7	6/7	1

■ 3.1.3 其他直方图处理方法

• r	0	1	2	3,4	5,6,7
• z	3	4	5	6	7
$P_z(z)$	0.15	0.25	0.21	0.24	0.11



-输出图像直方图近似于规定直方图
-G 满足严格单值单调递增

■ 3.1.3 其他直方图处理方法



原始图像



直方图均衡结果图像



直方图规范化处理结果图像

■直方图均衡的标准化处理

➤直方图均衡的问题

- 均衡后的图像直方图不完全均匀

■ 直方图均衡的标准化处理

➤例:设图像灰度分布概率如下:

r	0	1	2	3	4	5	6	7
$p_r(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

试进行直方图均衡的标准化处理，给出灰度映射关系以及均衡后图象的直方图分布。

■直方图均衡的标准化处理

- 一个8灰度图像灰度分布概率理想均匀的情况下应该为 $1/8=0.125$

	$p_s(s_0)$	$p_s(s_1)$	$p_s(s_2)$	$p_s(s_3)$	$p_s(s_4)$	$p_s(s_5)$	$p_s(s_6)$	$p_s(s_7)$
$P_r(r_0)=0.19$	0.125	0.065						
$P_r(r_1)=0.25$		0.06	0.125	0.065				
$P_r(r_2)=0.21$				0.06	0.125	0.025		
$P_r(r_3)=0.16$						0.10	0.06	
$P_r(r_4)=0.08$							0.065	0.015
$P_r(r_5)=0.06$								0.06
$P_r(r_6)=0.03$								0.03
$P_r(r_7)=0.02$								0.02
	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125

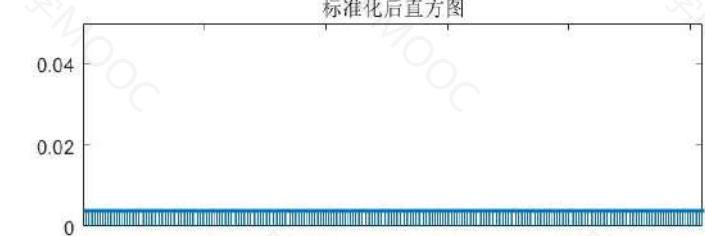
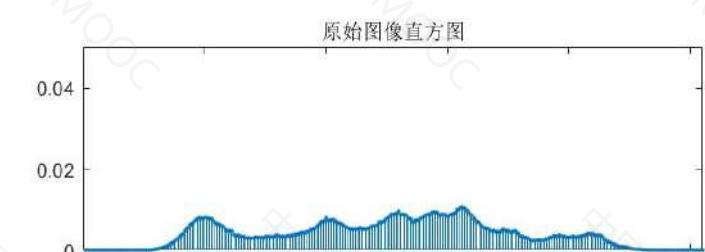
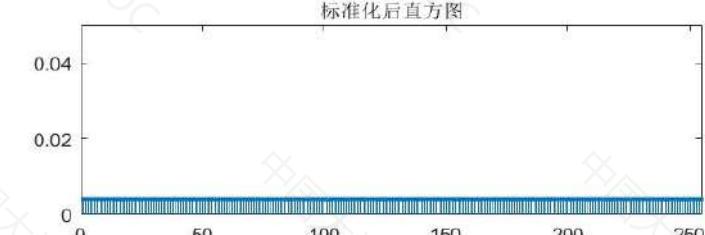
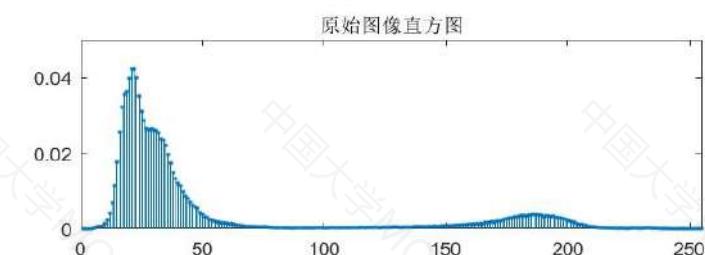
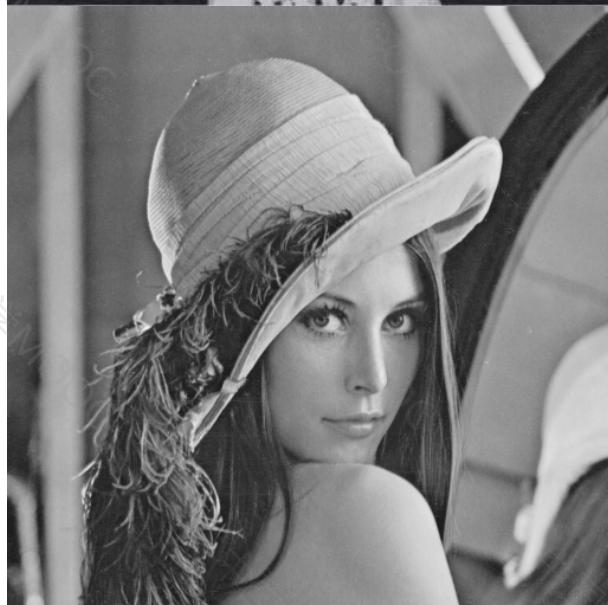
$r=r_o$ 时， $s=?$

$$s=s_o$$

$$[0, 0.658)$$

$$s=s_I$$

$$[0.658, 1.0]$$



■§3.2 空间运算

目标图像中任一点的取值,不仅取决于原图像中相应像素点,而且取决于原图像中其邻域点

$$g(x_0, y_0) = R\{f(x_0, y_0), f(x_0-1, y_0), f(x_0, y_0-1), f(x_0+1, y_0), (x_0, y_0+1)\dots\}$$

■ 3.2.1 掩膜及掩膜运算

掩膜为一组用矩阵表示的系数，如下为一组 3×3 的掩膜

$$\begin{bmatrix} w(-1, -1) & w(-1, 0) & w(-1, 1) \\ w(0, -1) & w(0, 0) & w(0, 1) \\ w(1, -1) & w(1, 0) & w(1, 1) \end{bmatrix}$$

■掩膜运算

$f(0,0)$	$f(0,1)$	$f(0,2)$	$f(0,3)$	$f(0,4)$
$f(1,0)$	$f(1,1)$	$f(1,2)$	$f(1,3)$	$f(1,4)$
$f(2,0)$	$f(2,1)$	$f(2,2)$	$f(2,3)$	$f(2,4)$
$f(3,0)$	$f(3,1)$	$f(3,2)$	$f(3,3)$	$f(3,4)$
$f(4,0)$	$f(4,1)$	$f(4,2)$	$f(4,3)$	$f(4,4)$

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

$g(1,1)$

$$\begin{aligned} g(1,1) = & f(0,0)*w(-1,-1) + f(0,1)*w(-1,0) \\ & + f(0,2)*w(-1,1) + f(1,0)*w(0,-1) \\ & + f(1,1)*w(0,0) + f(1,2)*w(0,1) \\ & + f(2,0)*w(1,-1) + f(2,1)*w(1,0) \\ & + f(2,2)*w(1,1) \end{aligned}$$

■掩膜运算

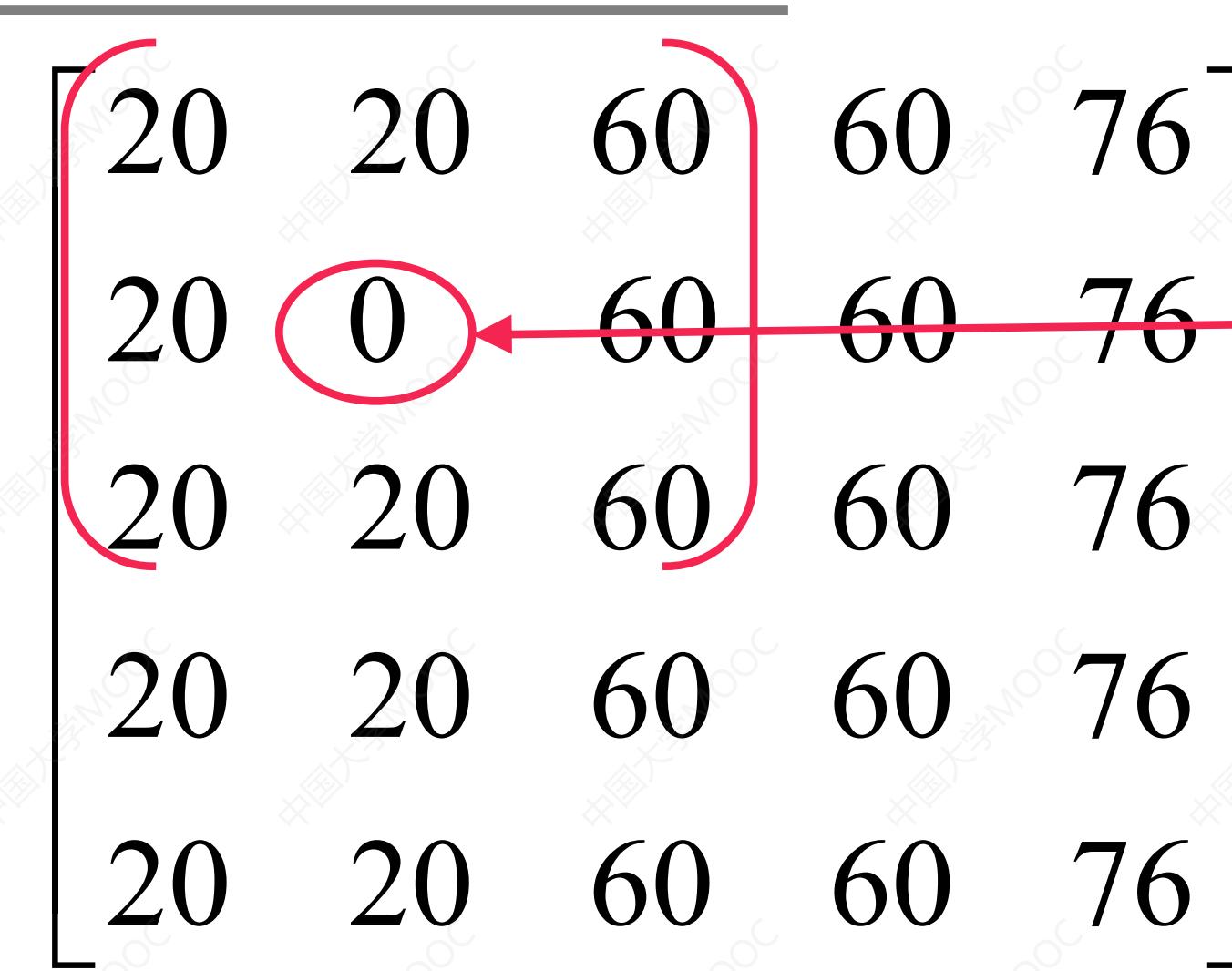
$f(0,0)$	$f(0,1)$	$f(0,2)$	$f(0,3)$	$f(0,4)$
$f(1,0)$	$f(1,1)$	$f(1,2)$	$f(1,3)$	$f(1,4)$
$f(2,0)$	$f(2,1)$	$f(2,2)$	$f(2,3)$	$f(2,4)$
$f(3,0)$	$f(3,1)$	$f(3,2)$	$f(3,3)$	$f(3,4)$
$f(4,0)$	$f(4,1)$	$f(4,2)$	$f(4,3)$	$f(4,4)$

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

$g(1,1)$	$g(1,2)$
----------	----------

$$\begin{aligned} g(1,2) = & f(0,1)*w(-1,-1) + f(0,2)*w(-1,0) \\ & + f(0,3)*w(-1,1) + f(1,1)*w(0,-1) \\ & + f(1,2)*w(0,0) + f(1,3)*w(0,1) \\ & + f(2,1)*w(1,-1) + f(2,2)*w(1,0) \\ & + f(2,3)*w(1,1) \end{aligned}$$

■掩膜运算



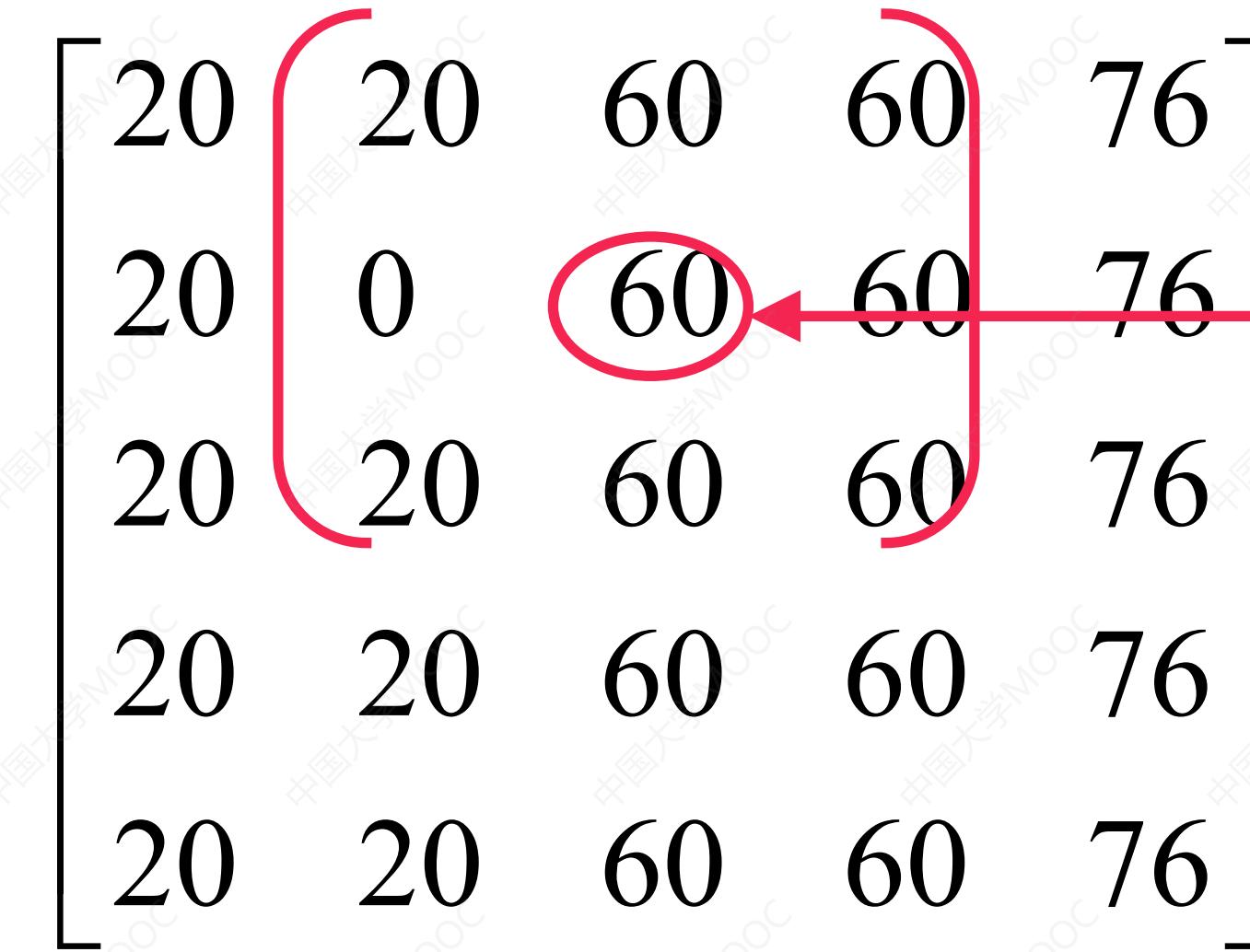
A 3x3 kernel matrix with all elements set to $1/9$. The central element is circled in red.

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$



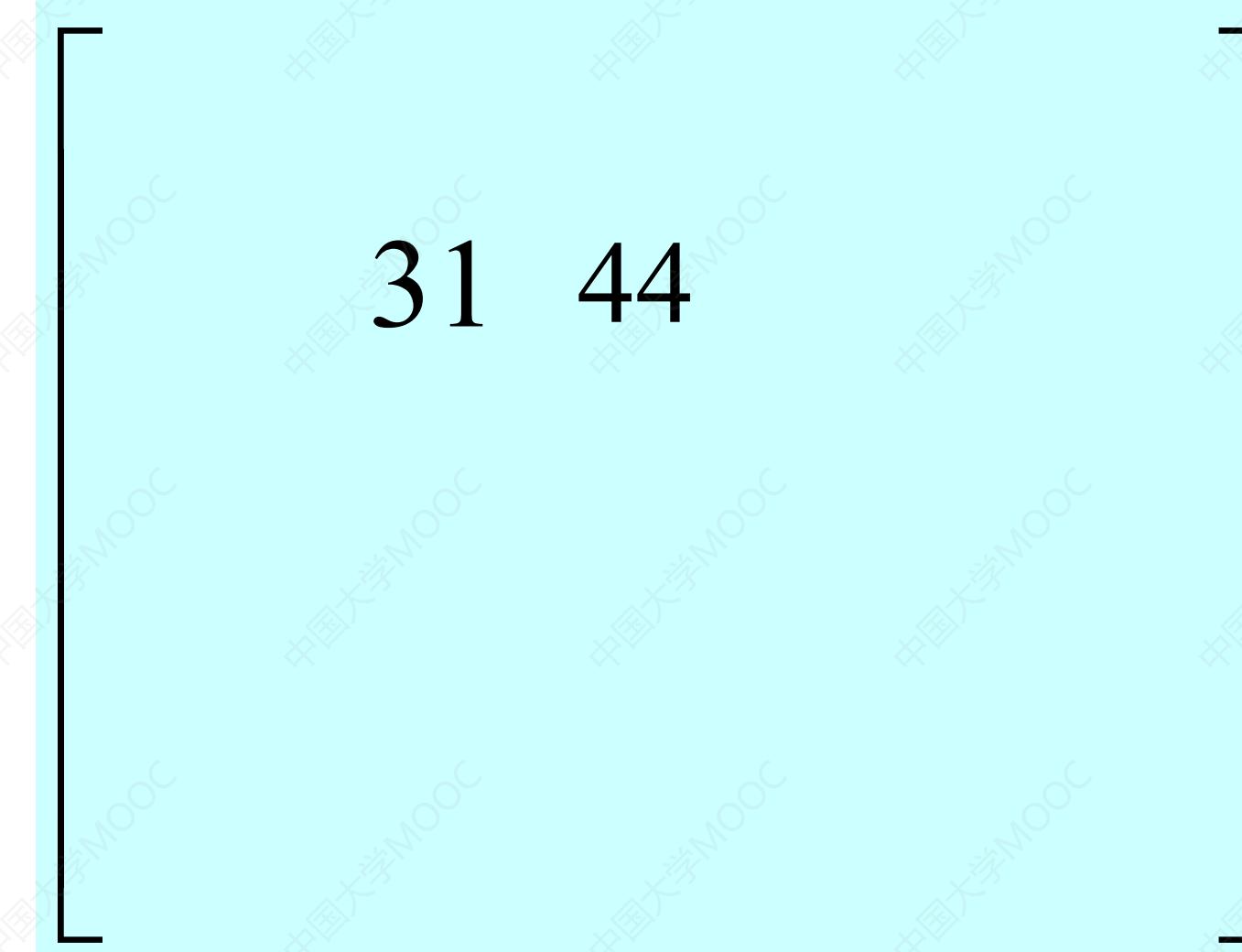
$$\begin{aligned} g(1,1) &= 20 * 1/9 + 20 * 1/9 + 60 * 1/9 \\ &\quad + 20 * 1/9 + 0 * 1/9 + 60 * 1/9 \\ &\quad + 20 * 1/9 + 20 * 1/9 + 60 * 1/9 \\ &= 31.111 \approx 31 \end{aligned}$$

■掩膜运算



A 3x3 mask matrix:

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & \textcircled{1/9} & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$



$$\begin{aligned} g(1,2) &= 20 * 1/9 + 60 * 1/9 + 60 * 1/9 \\ &\quad + 0 * 1/9 + 60 * 1/9 + 60 * 1/9 \\ &\quad + 20 * 1/9 + 60 * 1/9 + 60 * 1/9 \\ &= 44 \end{aligned}$$

■掩膜运算

$$\begin{bmatrix} 20 & 20 & 60 & 60 & 76 \\ 20 & 0 & 60 & 60 & 76 \\ 20 & 20 & 60 & 60 & 76 \\ 20 & 20 & 60 & 60 & 76 \\ 20 & 20 & 60 & 60 & 76 \end{bmatrix}$$

$$\begin{bmatrix} 31 & 40 & 65 \\ 31 & 44 & 65 \\ 33 & 47 & 65 \end{bmatrix}$$

$$\begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix}$$

■掩膜与卷积的关系

卷积运算

$$g(x, y) = f(x, y) * h(x, y) = \sum_{m=-l}^l \sum_{n=-l}^l f(x-m, y-n)h(m, n)$$

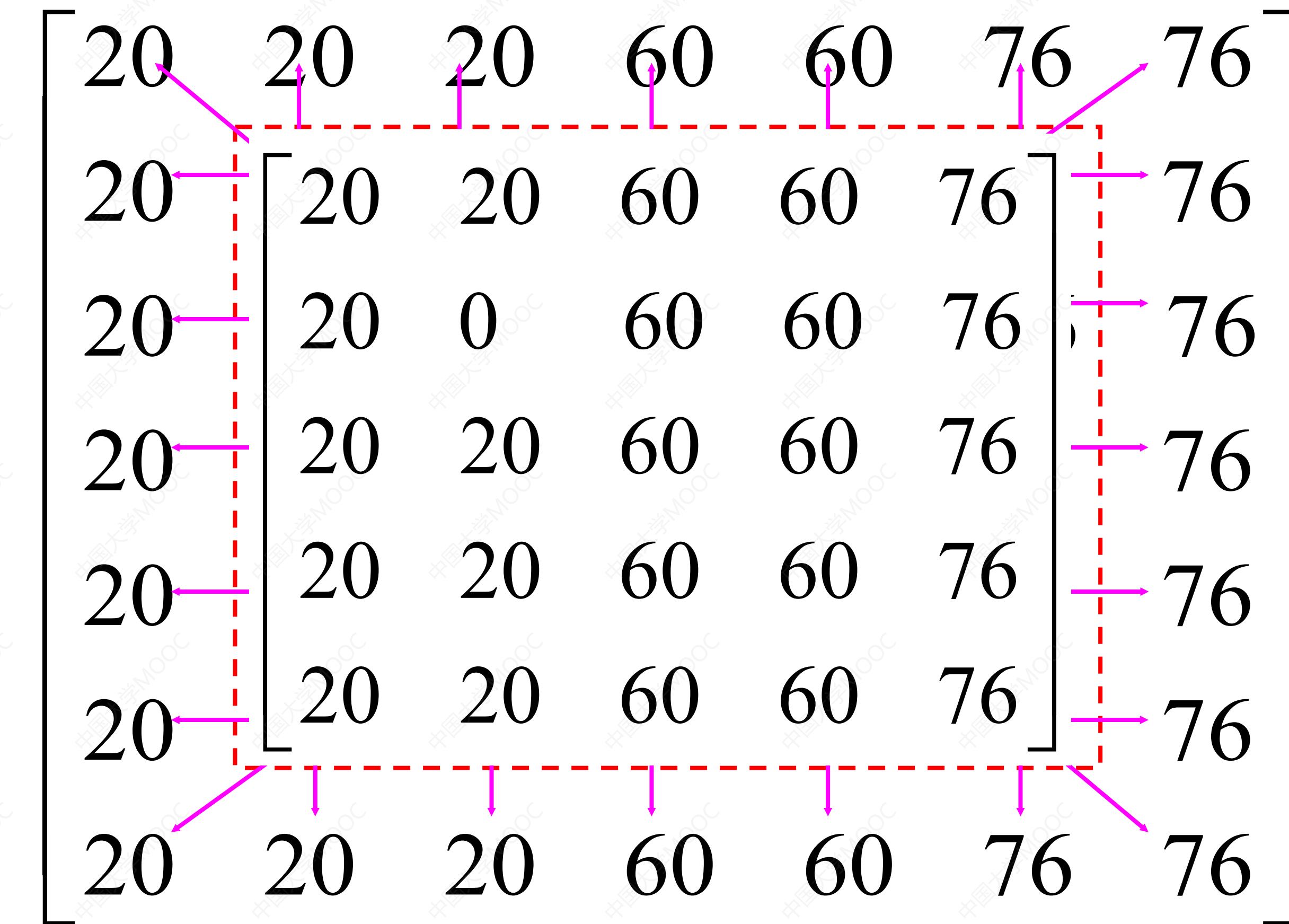
掩膜运算

$$\begin{aligned} g(1,1) &= f(0,0) * w(-1,-1) + f(0,1) * w(-1,0) \\ &\quad + f(0,2) * w(-1,1) + f(1,0) * w(0,-1) \\ &\quad + f(1,1) * w(0,0) + f(1,2) * w(0,1) \\ &\quad + f(2,0) * w(1,-1) + f(2,1) * w(1,0) \\ &\quad + f(2,2) * w(1,1) \end{aligned}$$

掩膜是卷积的特例

■边缘的处理

- 边界延拓



■平滑掩膜

$$H1 = \frac{1}{9} * \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad H2 = \frac{1}{5} * \begin{vmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$
$$H3 = \frac{1}{8} * \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{vmatrix} \quad H4 = \frac{1}{4} * \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$

■常用方式

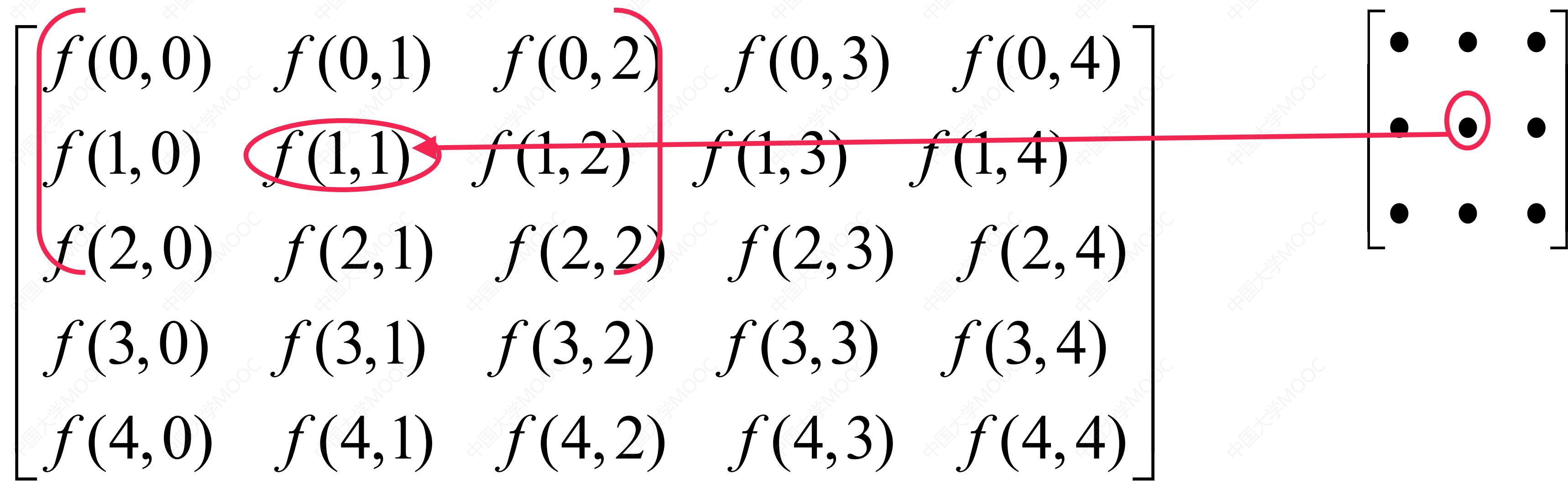
$$g(x, y) = \begin{cases} f'(x, y) & |f(x, y) - f'(x, y)| > C_{th} \\ f(x, y) & |f(x, y) - f'(x, y)| \leq C_{th} \end{cases}$$

$f'(x, y)$ 是用上述平滑掩膜H3处理的结果

C_{th} 为门限，主要目的是去除随机白噪音

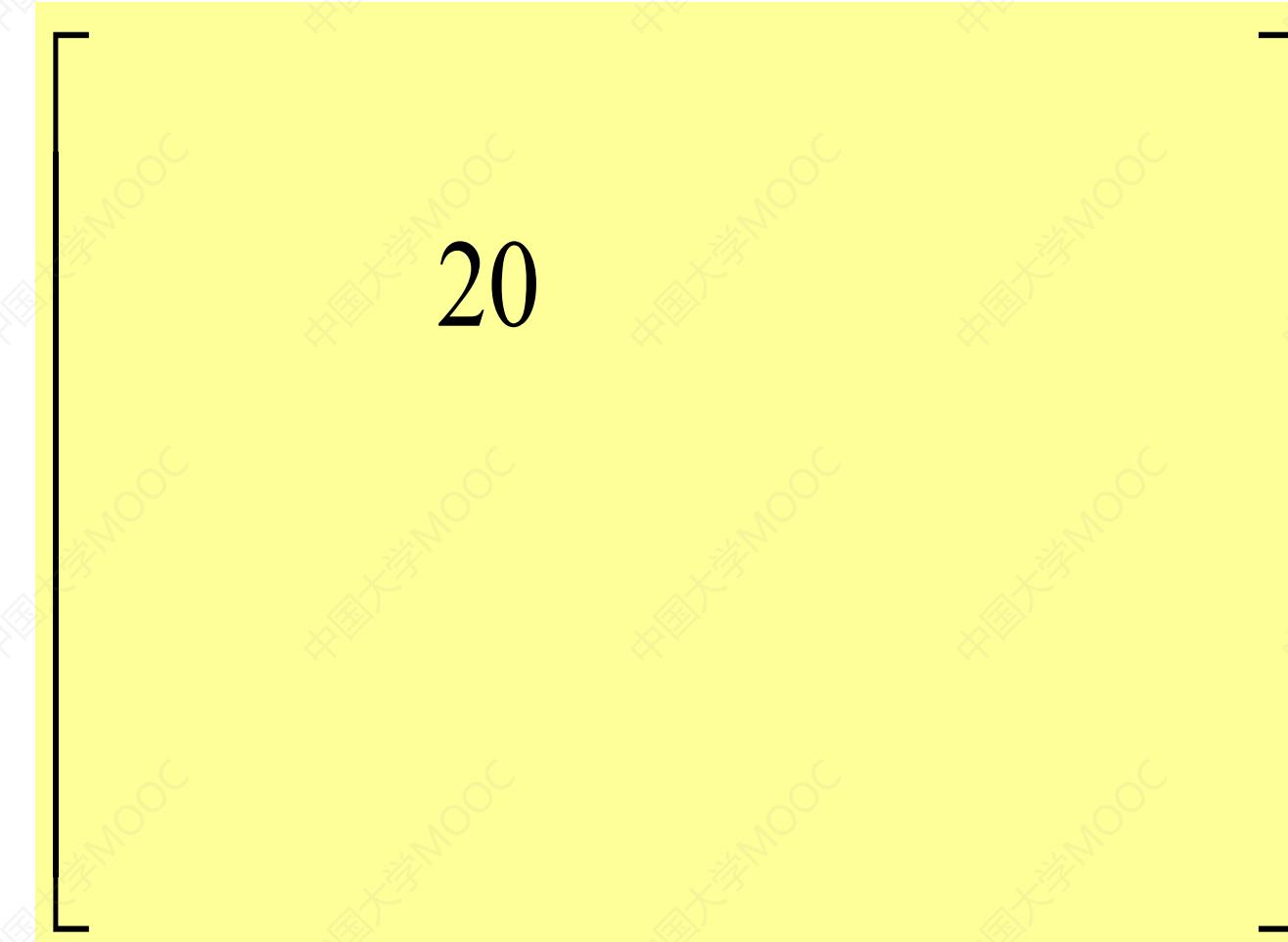
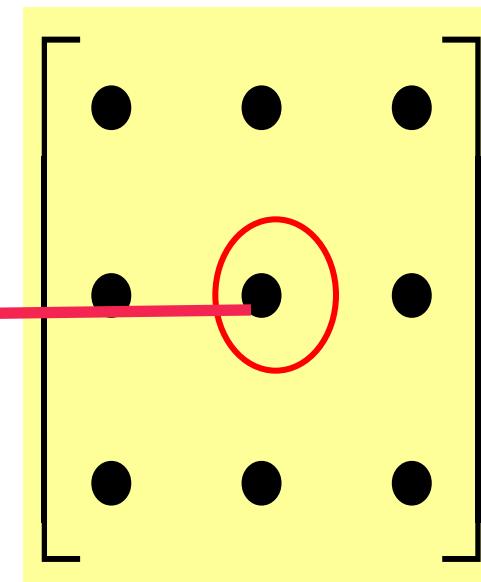
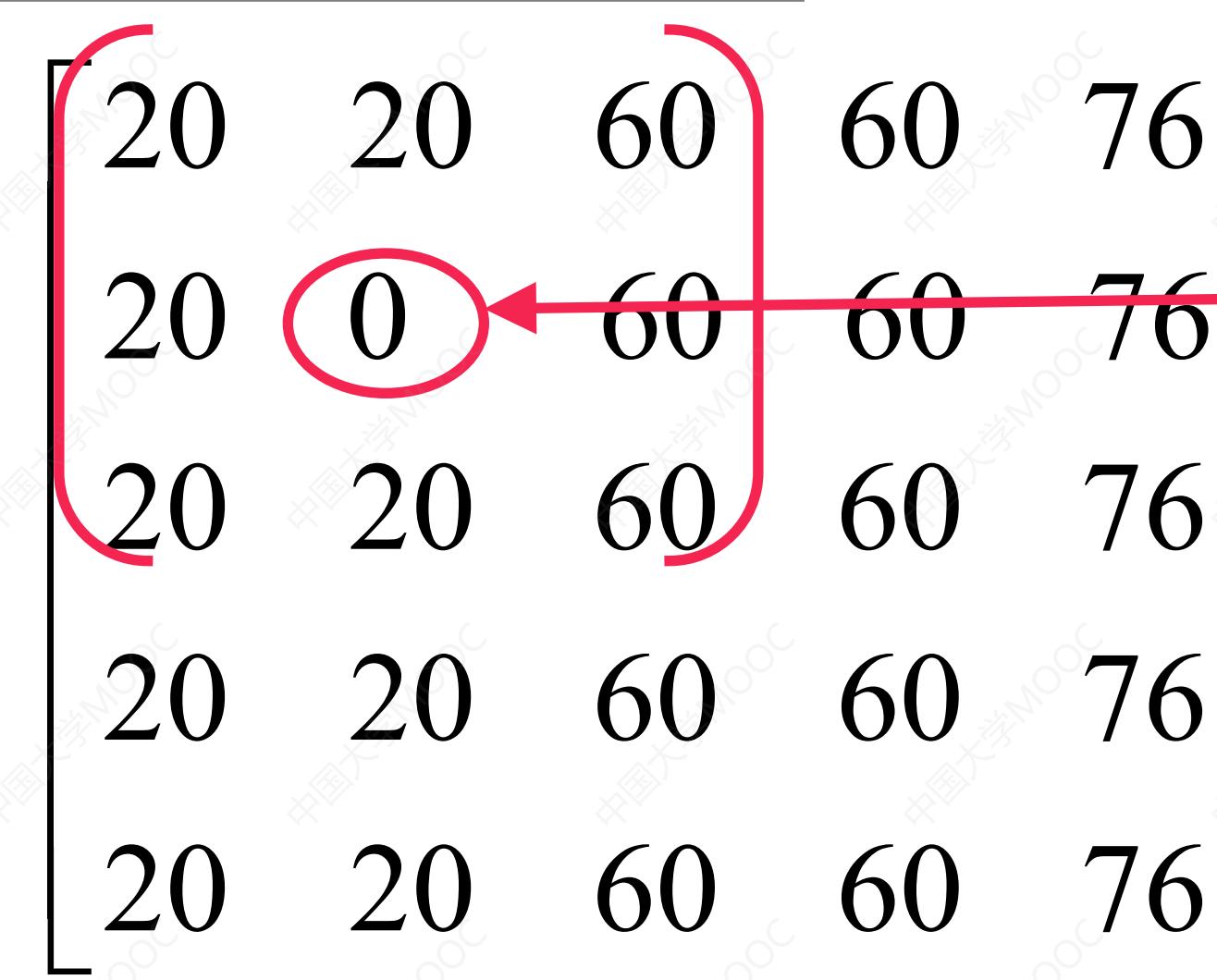
二 中值滤波

排序统计滤波器



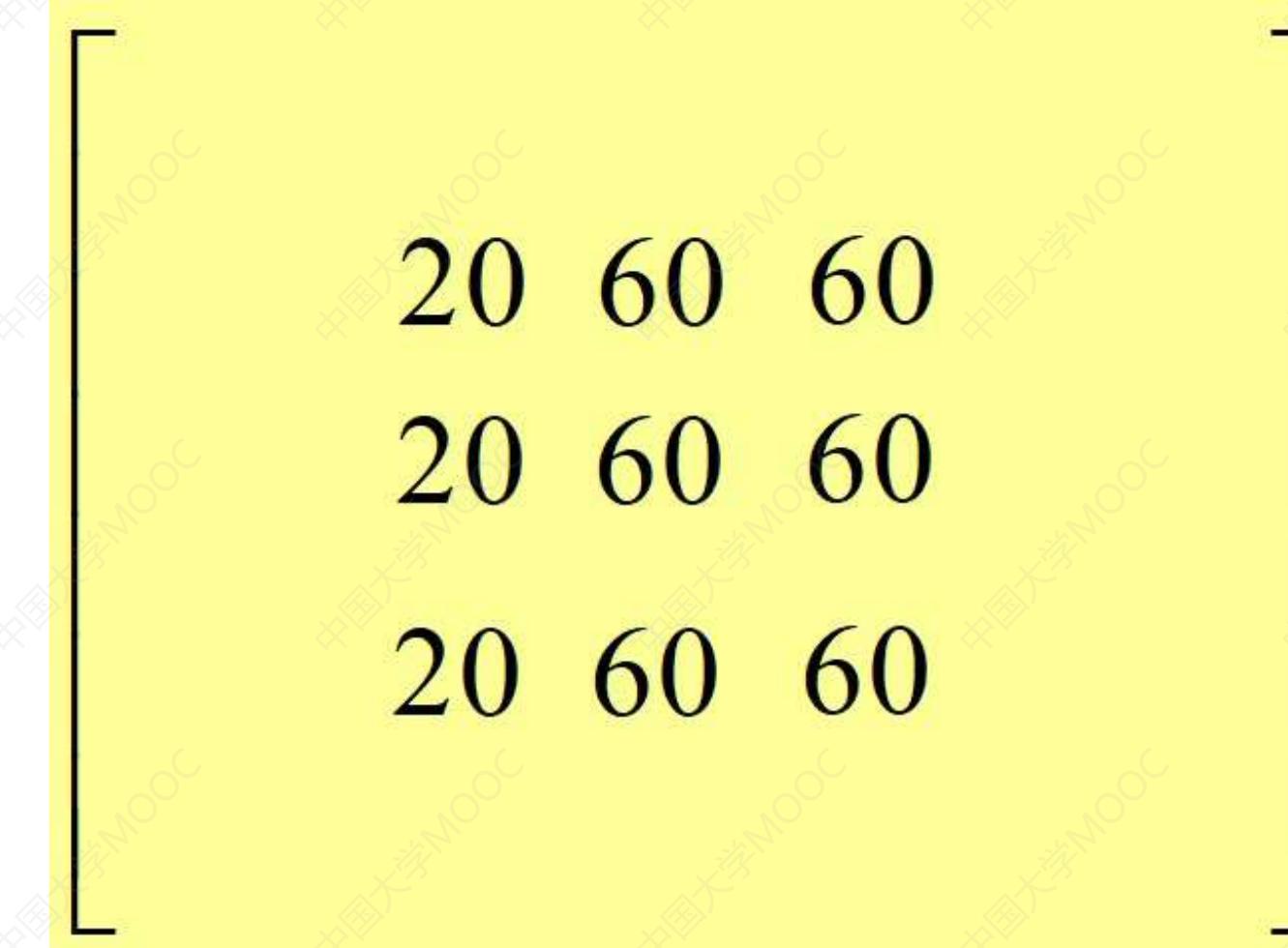
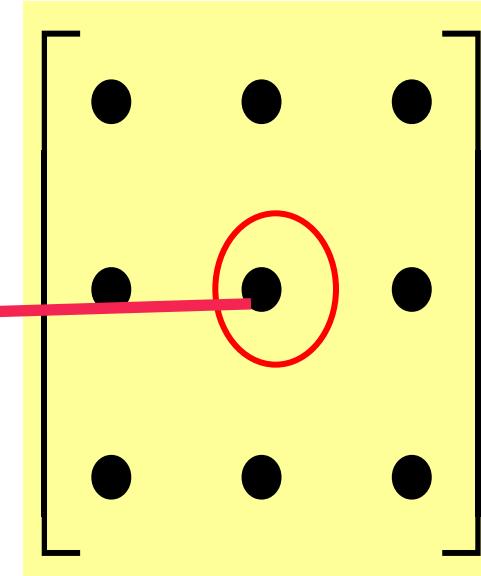
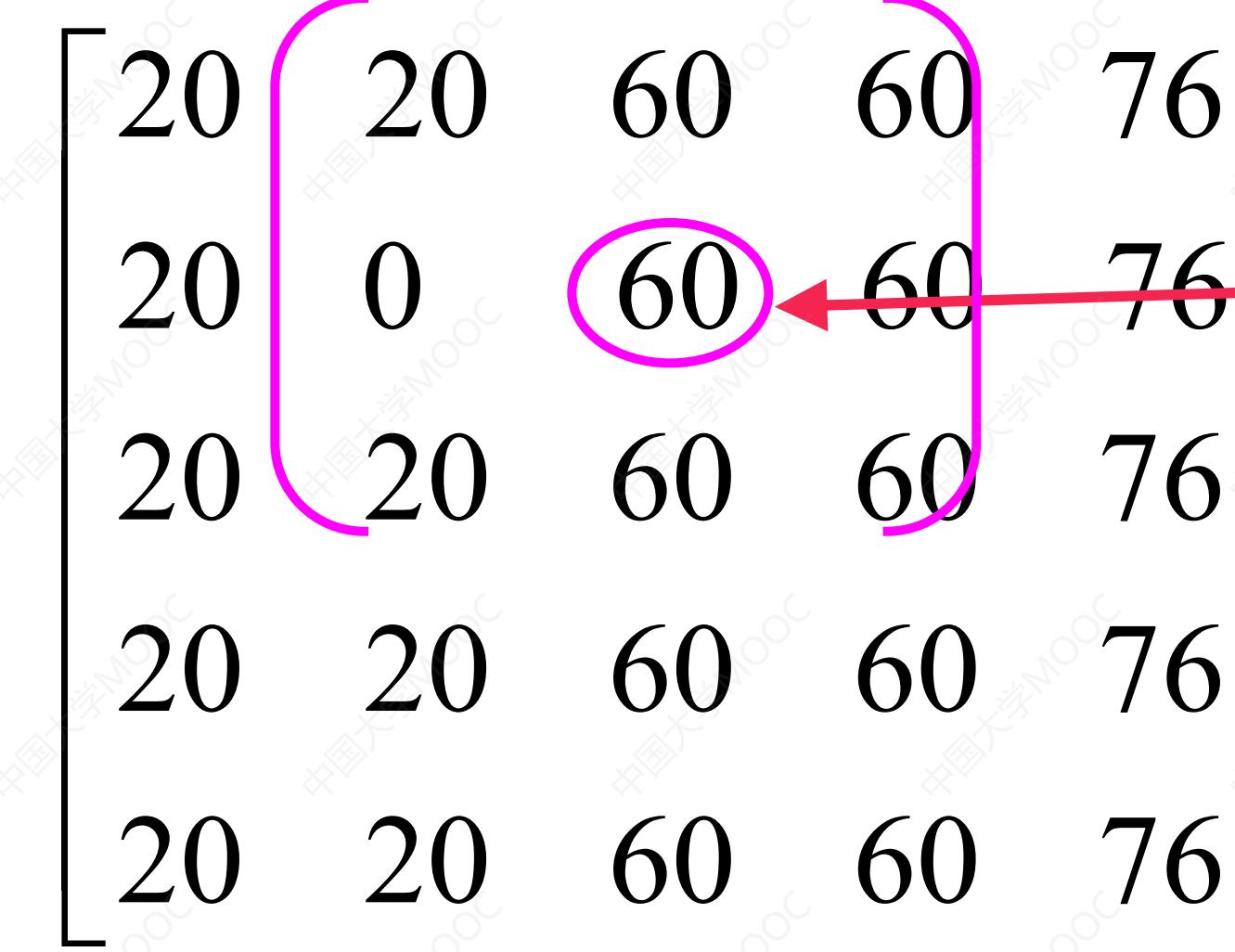
将 $\{f(0,0), f(0,1), f(0,2), f(1,0), f(1,1), f(1,2), f(2,0), f(2,1), f(2,2)\}$ 进行排序，取排在第N的值作为输出。常见：
取最大值——最大滤波
取最小值——最小滤波
取中间值——中值滤波

二 中值滤波



$$\begin{aligned} & \text{Rank}\{20, 20, 60, 20, 0, 60, 20, 20, 60\} \\ &= \{60, 60, 60, 20, 20, 20, 20, 0\} \end{aligned}$$

二 中值滤波



$$\begin{aligned} &\text{rank}\{20, 60, 60, 0, 60, 60, 20, 60, 60\} \\ &= \{60, 60, 60, 60, \textcircled{60}, 60, 20, 20, 20\} \end{aligned}$$

■ 二 中值滤波

➤ 基本原理

$$g(x, y) = \text{Median}\{f(m, n), (m, n) \in W_{(x, y)}\}$$

➤ 窗口类型

线形，方形，十字形，圆形，环形等

■ 中值滤波举例

➤ 例 1：试用 3×3 的十字窗口对下图像进行中值滤波，给出输出图像。

0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0	0
0	0	1	0	0	0	1	0	0
0	0	1	0	0	1	0	0	0
0	0	0	0	1	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	0
0	0	1	0	0	0	0	0	0
0	0	1	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	0	0	0

● 中值滤波很容易滤除图像里的线状细节。

中值滤波举例

➤例2：试用 3×3 的十字窗口和方形窗口分别对下图像进行中值滤波，给出输出图像。

中值滤波举例

十
字
窗

■中值滤波举例

方形窗口

.	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	5	5	5	5	5	1	1	1
1	1	5	5	5	5	5	5	5	1	1	1
1	1	5	5	5	5	5	5	5	1	1	1
1	1	5	5	5	5	5	5	5	5	1	1
1	1	5	5	5	5	5	5	5	5	1	1
1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1

- 用十字窗可以很好地保留图像中的边角细节，而用方形窗很容易滤除这些细节。
- 窗口大小大于图像中的细节时，细节容易被滤除

■中值滤波举例

例3：试用 3×3 的十字窗口分别对下图进行中值滤波和平滑掩膜，给出输出图像。

•	1	1	1	1	5	5	5	5
1	1	1	1	1	5	5	0	5
1	1	1	1	1	5	5	5	5
1	0	1	1	1	5	5	5	5
1	1	1	1	1	0	5	5	5
1	1	1	1	1	5	5	5	5

■ 中值滤波举例

例 3：用 3×3 的方形窗口分别对下图像进行中值滤波和平滑，给出输出图像。

•	1	1	1	1	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5
1	1	1	1	5	5	5	5	5

中值滤波结果

- 中值滤波在滤除噪音的同时，很好地保持了图像边缘。

■ 中值滤波举例

例 3：试用 3×3 的方形窗口分别对下图像进行中值滤波和平滑，给出输出图像。

•	1	1	1	2	4	5	4	5
1	1	1	2	4	4	4	4	4
1	1	1	2	4	5	4	5	5
1	1	1	2	3	5	5	5	5
1	1	1	1	3	4	5	5	5
1	1	1	2	3	5	5	5	5
1	1	1	2	3	5	5	5	5

- 平滑掩膜滤除了一个噪音点，另外两个噪音点被放大了变模糊了；
- 平滑掩膜使图像边缘变模糊。

平滑结果

■3中值滤波的特点

- a) 中值滤波是非线性滤波
- b) 滤除噪声的同时可以很好的保护轮廓边缘
- c) 窗口大小不超过图像中最小物体的有效尺寸为宜
- d) 窗口中噪声像素比例过高，则中值滤波失效，即当噪声不严重时采用中值滤波效果较好，反之，噪声严重时采用中值滤波，将会使图像质量下降，不宜采用。

■ 二 中值滤波

中值滤波对图像质量的影响，分四种情况讨论：

Case1：当前像素点是噪音，并且噪音比例比较低，**图像质量大概率改善**；

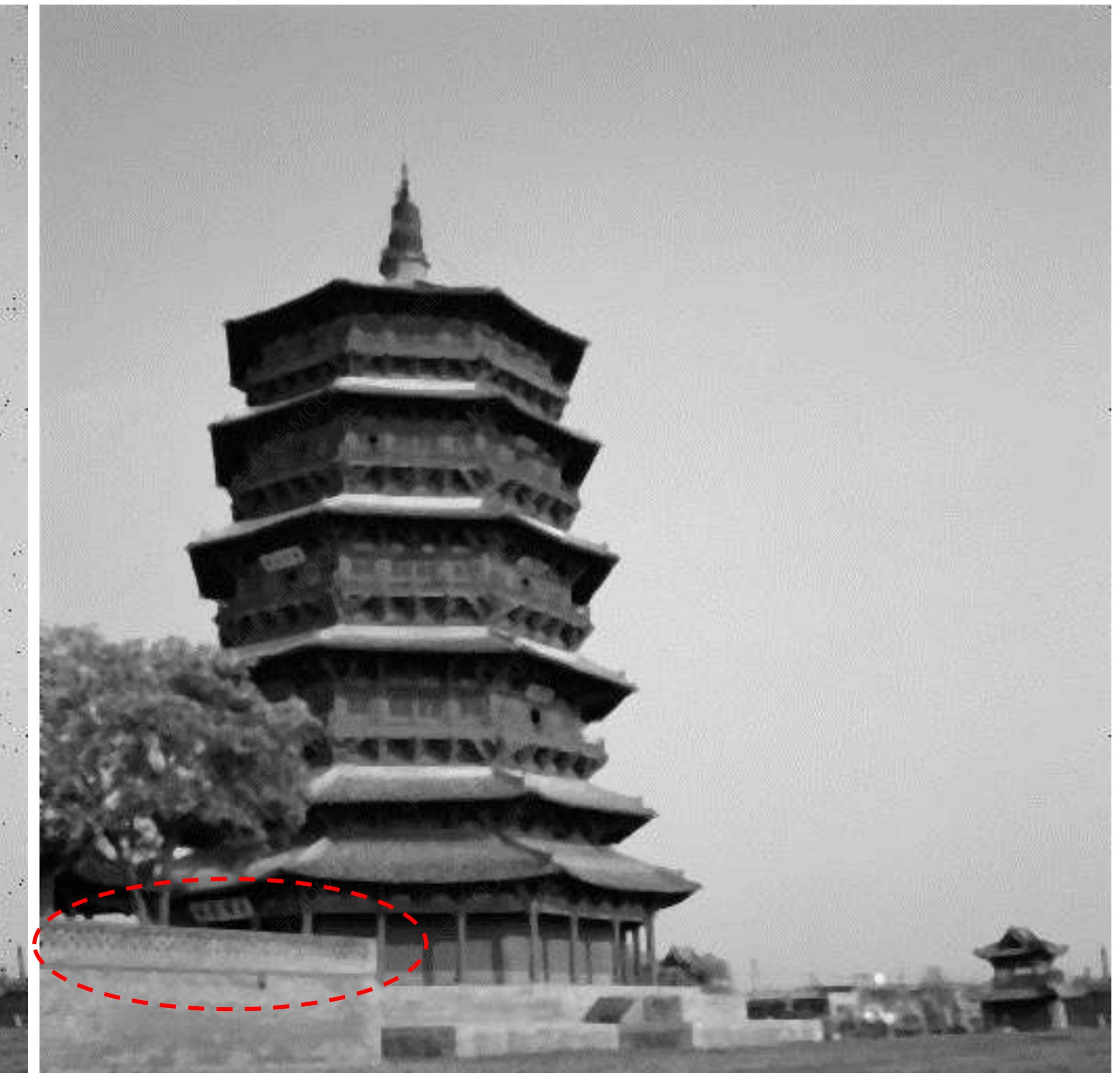
Case2：当前像素点是噪音，并且噪音比例比较高，**图像质量大概率不会改善**；

Case3：当前像素点不是噪音，而且排在中间的值是当前像素点的邻域像素点的值，**图像质量可能会下降**；

Case4：当前像素点不是噪音，而且排在中间的值是当前像素点的值，**图像质量不会下降**；



有1%椒盐噪音的图像



中值滤波结果图像



有30%椒盐噪音的图像



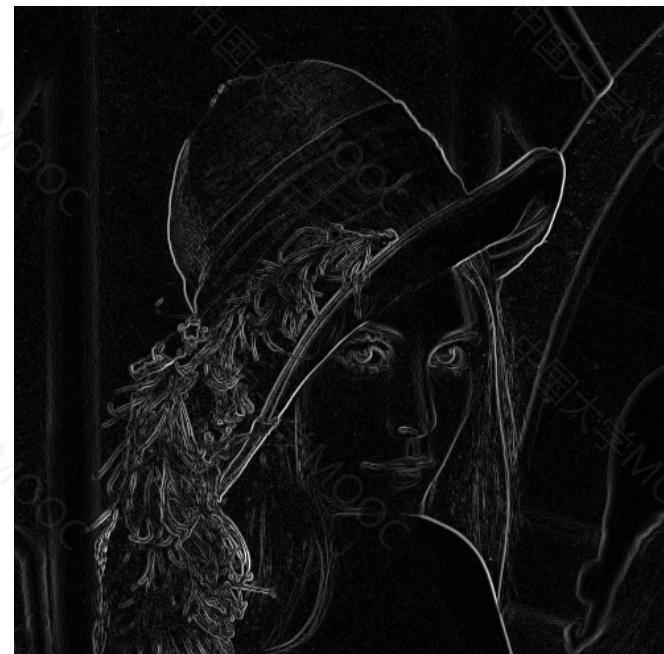
中值滤波结果图像

3.2.3 图像锐化——增强或提取图像中的边缘

- 景物的边缘是指图像中存在灰度突变的地方，可用梯度运算和差分运算把边缘提取出来。



输入图像



边缘检测



边缘增强

2 Laplace算子

• 定义

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \nabla_x^2 f + \nabla_y^2 f$$

$$\nabla_y f(x, y) = [f(x, y + 1) - f(x, y)]$$

$$\nabla_x f(x, y) = [f(x + 1, y) - f(x, y)]$$

$$\begin{aligned}\nabla_x^2 f(x, y) &= [\nabla_x f]_x = [f(x, y+1) - f(x, y)] - [f(x, y) - f(x, y-1)] \\ &= f(x, y+1) + f(x, y-1) - 2f(x, y)\end{aligned}$$

$$\begin{aligned}\nabla_y^2 f(x, y) &= [\nabla_y f]_y = [f(x+1, y) - f(x, y)] - [f(x, y) - f(x-1, y)] \\ &= f(x+1, y) + f(x-1, y) - 2f(x, y)\end{aligned}$$

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

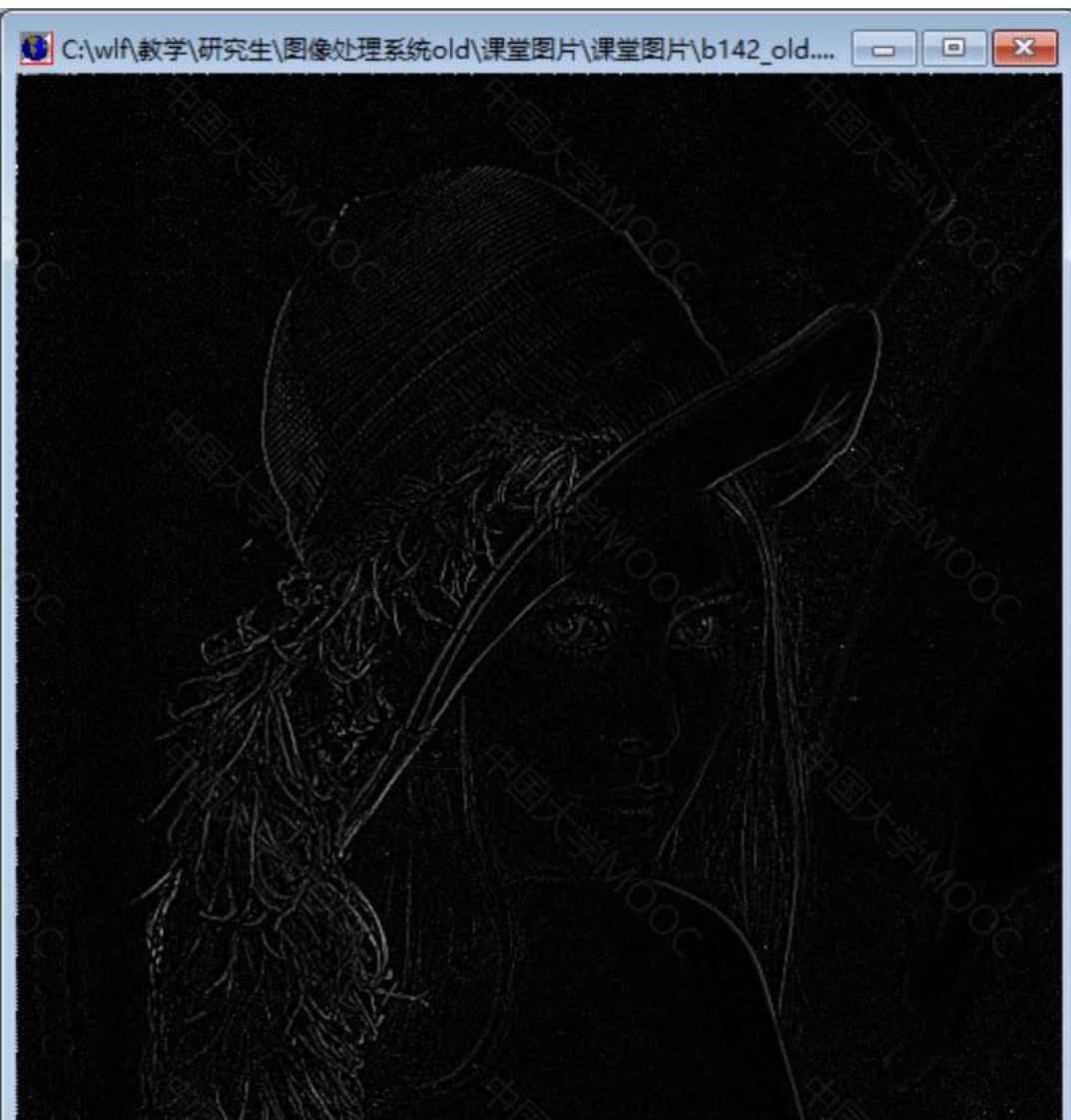
对应于

$$H_I = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix} \quad H_2 = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}$$

(a) 边缘检测

(b) 边缘检测

中国大学MOOC



$$\begin{aligned}\nabla_x^2 f(x, y) &= [\nabla_x f]_x = [f(x, y+1) - f(x, y)] - [f(x, y) - f(x, y-1)] \\ &= f(x, y+1) + f(x, y-1) - 2f(x, y)\end{aligned}$$

$$\begin{aligned}\nabla_y^2 f(x, y) &= [\nabla_y f]_y = [f(x+1, y) - f(x, y)] - [f(x, y) - f(x-1, y)] \\ &= f(x+1, y) + f(x-1, y) - 2f(x, y)\end{aligned}$$

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

对应于

$$H_I = \begin{vmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{vmatrix}$$

2021/11/26

(a) 边缘检测

$$H_2 = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}$$

(b) 边缘检测

$$H_3 = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{vmatrix}$$

(c) 边缘增强

- 1 梯度法

- **梯度（梯度向量）定义：**

设二元函数 $f(x,y)$ 在平面区域D上具有一阶连续偏导数，则函数在点P(x,y)的梯度定义为

-

$$\vec{G}[f(x,y)] = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

- 梯度幅度

$$G[f(x, y)] = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}$$

- 梯度幅角

$$\theta_M = \arctg \frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

- 图像里梯度一般指的是梯度幅度

梯度幅度的简化运算：取绝对值来代替平方和开方

$$G[f(x, y)] = |f(x+1, y) - f(x, y)| + |f(x, y+1) - f(x, y)|$$

简化后的梯度运算对应的掩膜

$$H_1 = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}$$

Roberts梯度：交叉梯度

$$G[f(x, y)] = |f(x + 1, y + 1) - f(x, y)| + |f(x + 1, y) - f(x, y + 1)|$$

$$H_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Sobel算子

$$H_1 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Prewitt算子

$$H_1 = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad H_2 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

锐化掩膜特性

$W(x, y) > 0 \text{ or } < 0$

$\sum W(x, y) = 0$ (边缘检测)
 $= 1$ (边缘增强)

更多扩展掩膜

$$H_4 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad H_5 = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \quad H_6 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$



Lena图像



Canny算子边缘检测结果图像

■彩色图像处理

➤ 3.4.1. 颜色模型

- RGB
- CMYK
- HSI
- Lab

➤ 3.4.2. 伪彩色增强

- 密度分割（灰度分层）
- 灰度级彩色转换

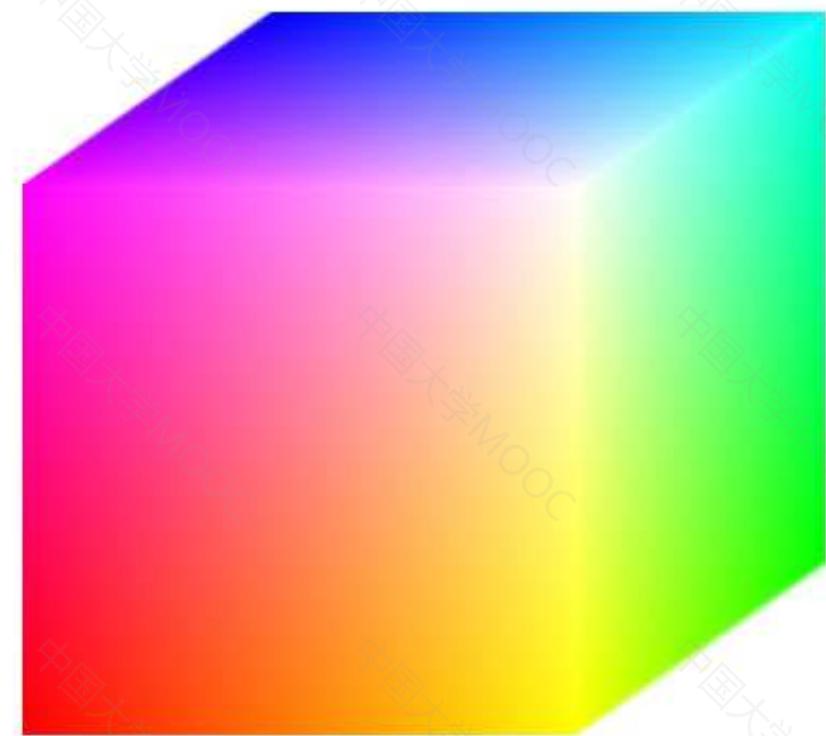
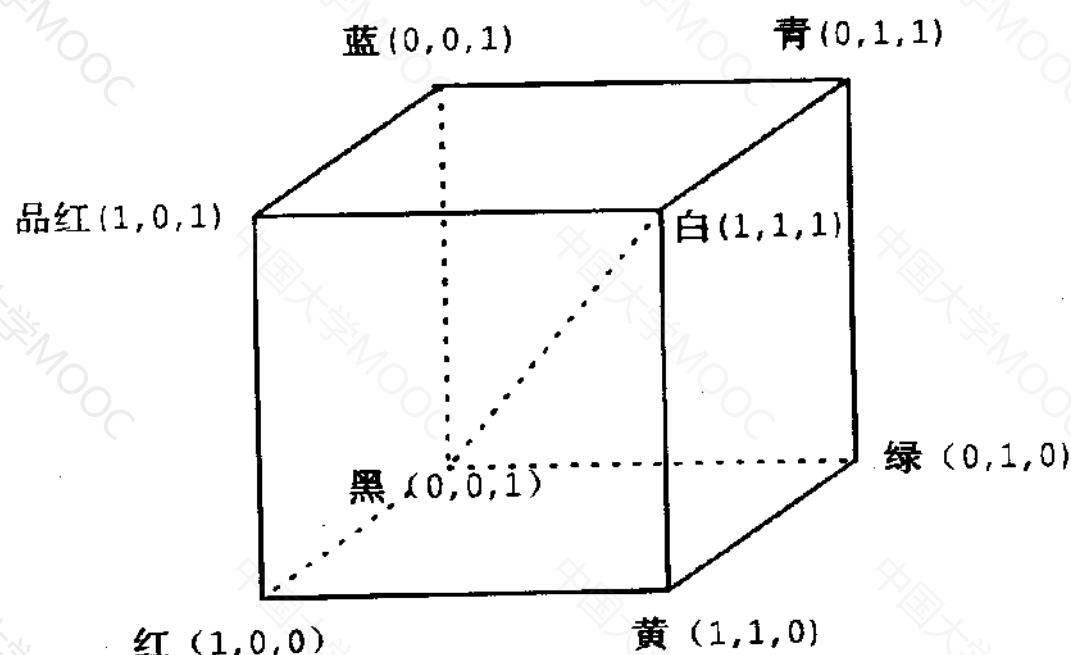
➤ 3.4.3. 彩色直方图均衡

■彩色图像处理

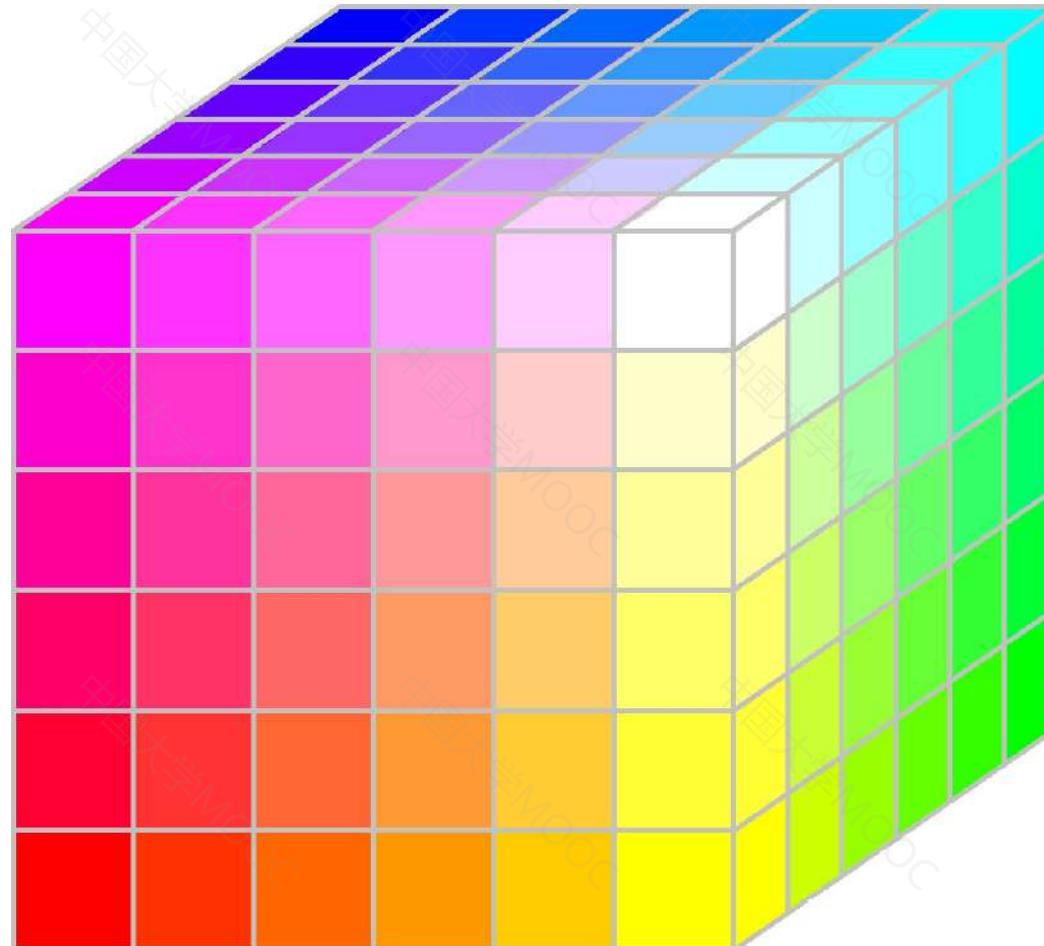
➤ 3.4.1. 颜色模型

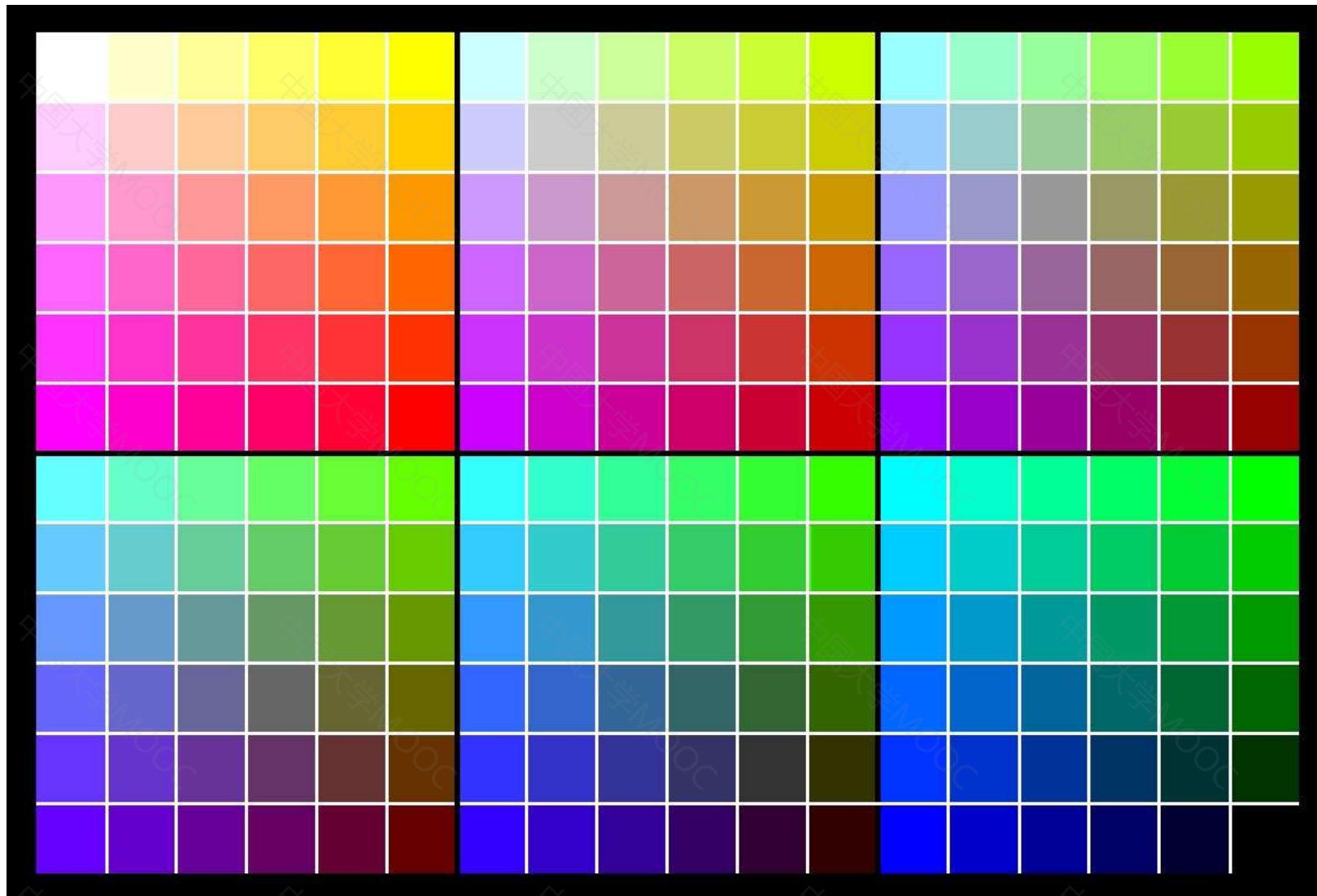
- RGB
- CMYK
- HSI
- Lab

■颜色模型

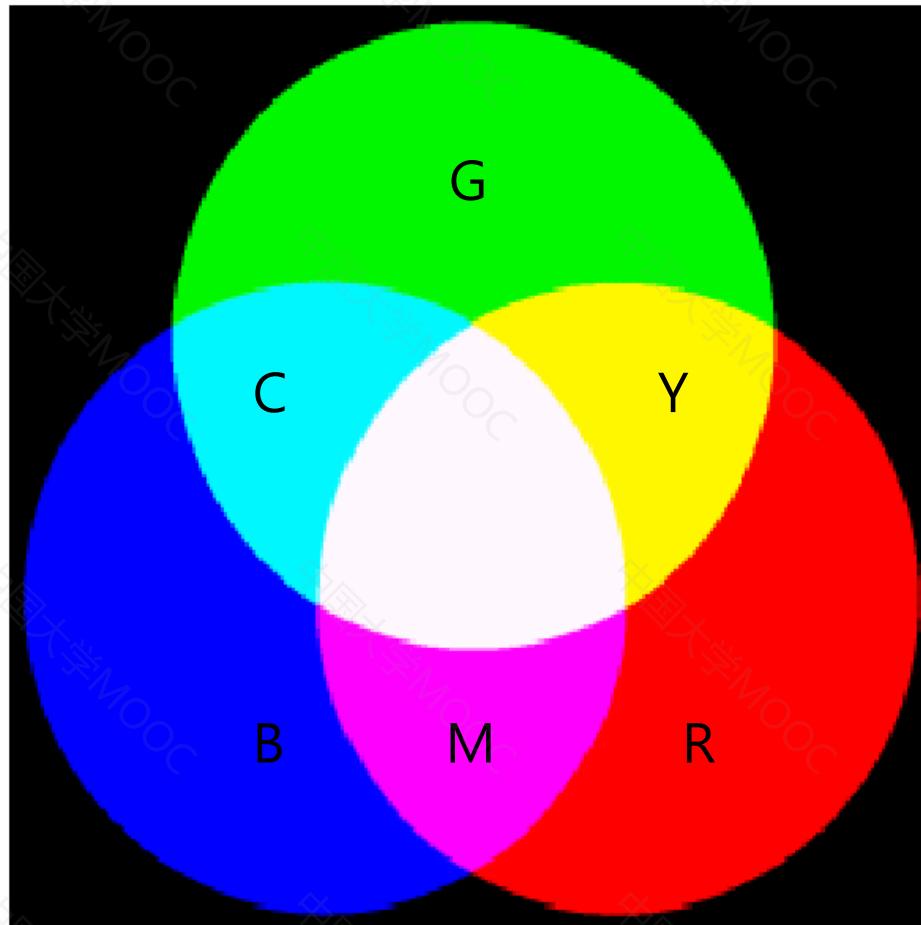


- 24位全彩色图像 可显示颜色 $(2^8)^3 = 16777216$





- CMYK颜色空间



R: Red (红色)

G: Green (绿色)

B: Blue (蓝色)

C: Cyan (青色)

M: Magenta (品红色)

Y: Yellow (黄色)

K: black (黑色)

HSI颜色空间

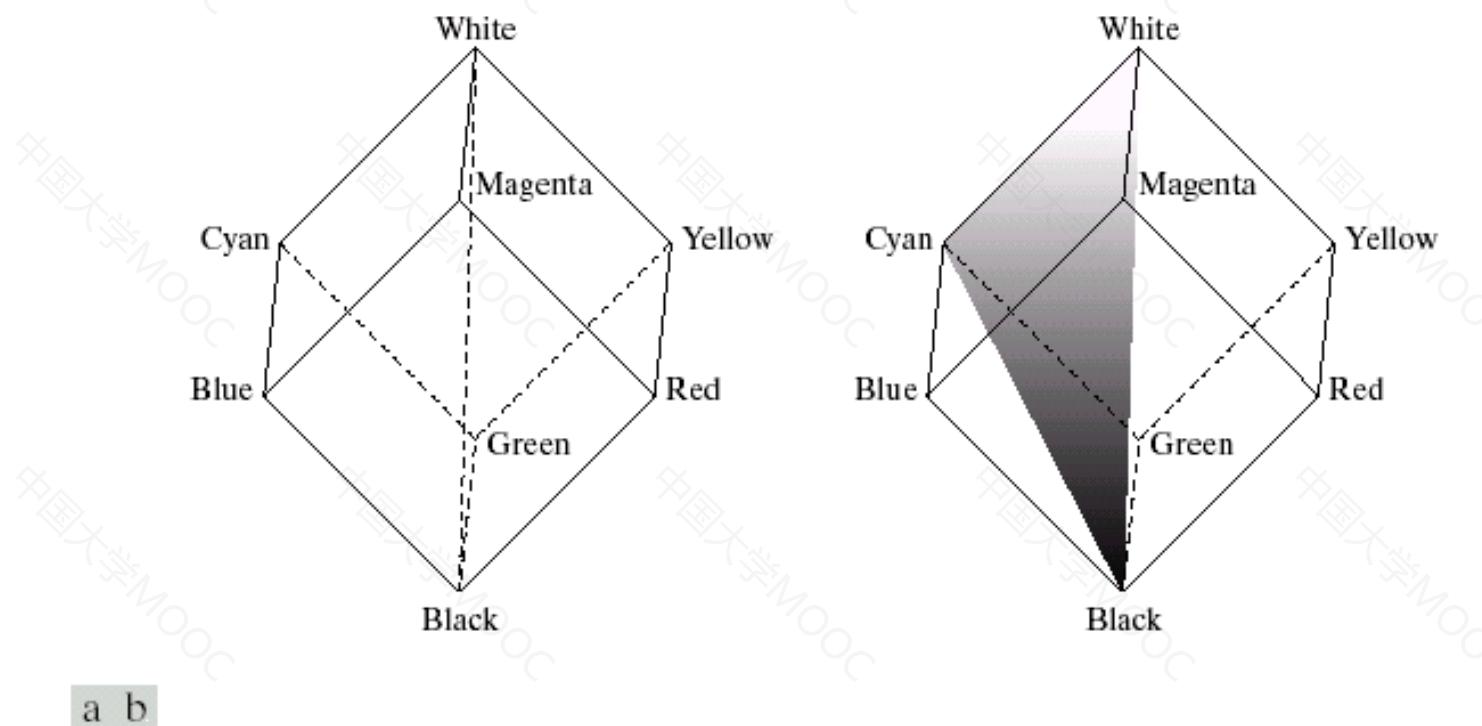
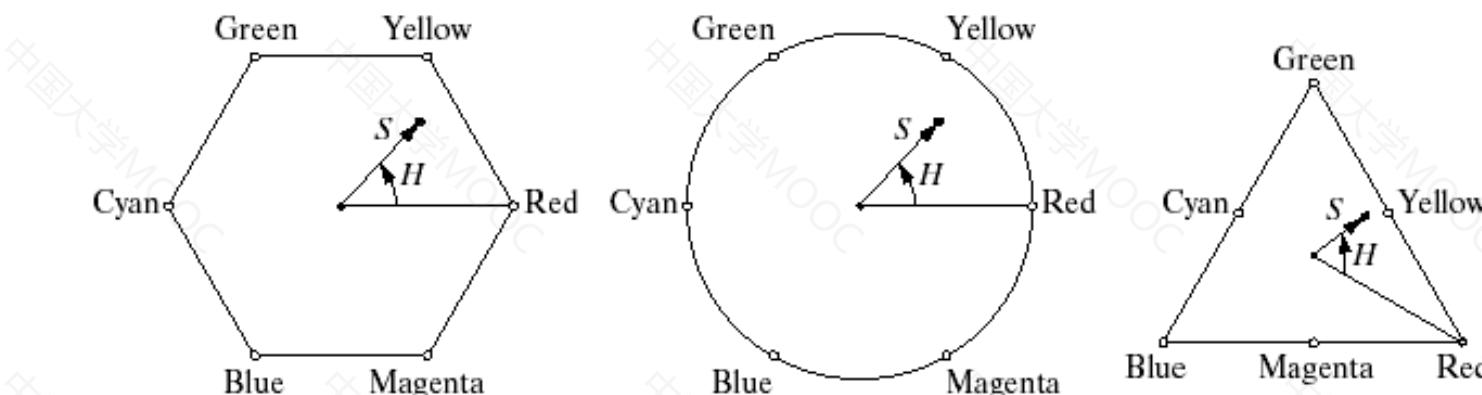
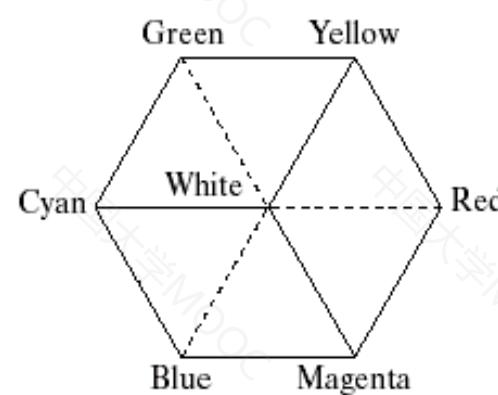
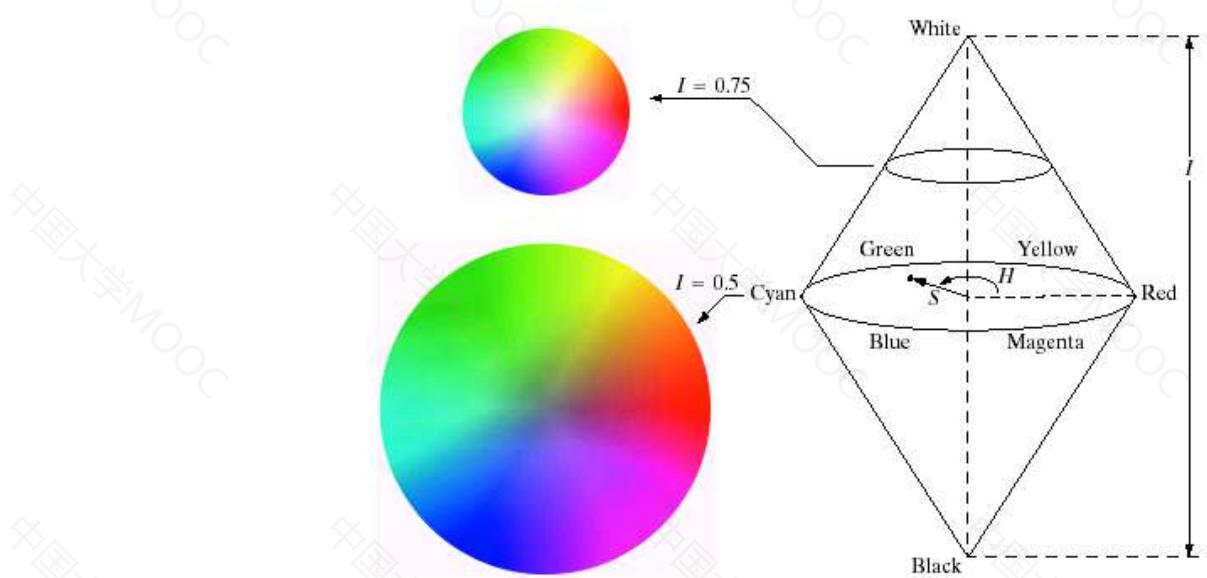
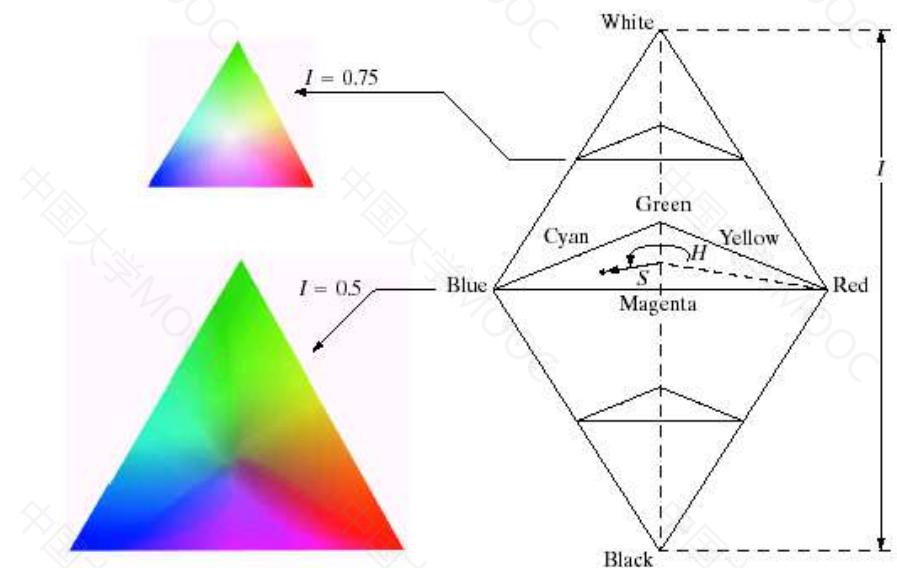


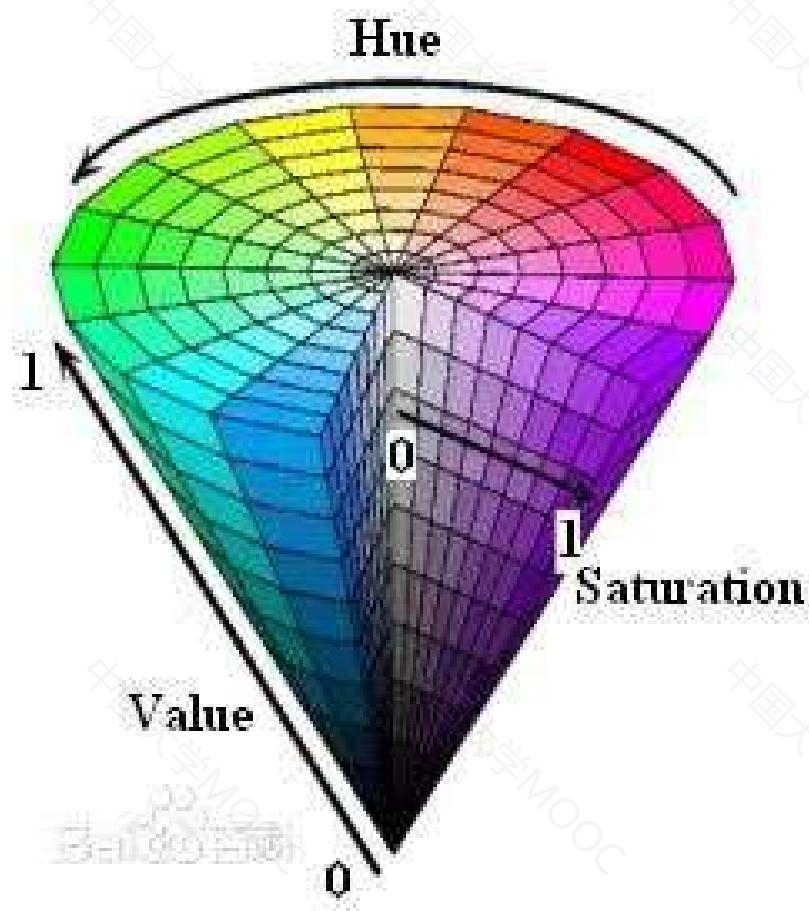
FIGURE 6.12 Conceptual relationships between the RGB and HSI color models.



a
b c d

HSI彩色模型





■ RGB → HSI

$$\theta = \cos^{-1} \left\{ \frac{[(R-G)+(R-B)]/2}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}$$

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases}$$

$$S = 1 - \frac{3 \cdot \min(R, G, B)}{R + G + B}$$

$$I = (R + G + B) / 3$$

**RGB无法直接转换成LAB，需要先转换成XYZ再转换成LAB，
即：RGB——XYZ——LAB**

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9508 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$f(t) = \begin{cases} t^{\frac{1}{3}} & t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3} \left(\frac{29}{6}\right)^2 t + \left(\frac{6}{29}\right) & \text{otherwise} \end{cases}$$

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16$$

$$a^* = 500[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)]$$

$$b^* = 200[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)]$$

$$X_n = 0.9505, Y_n = 1.0, Z_n = 1.0893$$

RGB--→YCbCr

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

$$Cb = -0.1687 R - 0.3313 G + 0.5 B + 128$$

$$Cr = 0.5 R - 0.4187 G - 0.0813 B + 128$$

YCbCr--→RGB

$$R = Y + 1.402 (Cr-128)$$

$$G = Y - 0.34414 (Cb-128) - 0.71414 (Cr-128)$$

$$B = Y + 1.772 (Cb-128)$$

■ 伪彩色增强

- 人类视觉可分辨的灰度32-64，可分辨的颜色几百甚至上千种，因此，将会度图像映射为彩色图像，可能提高图像的可分辨质量。

■ 伪彩色增强

1. 密度分割法

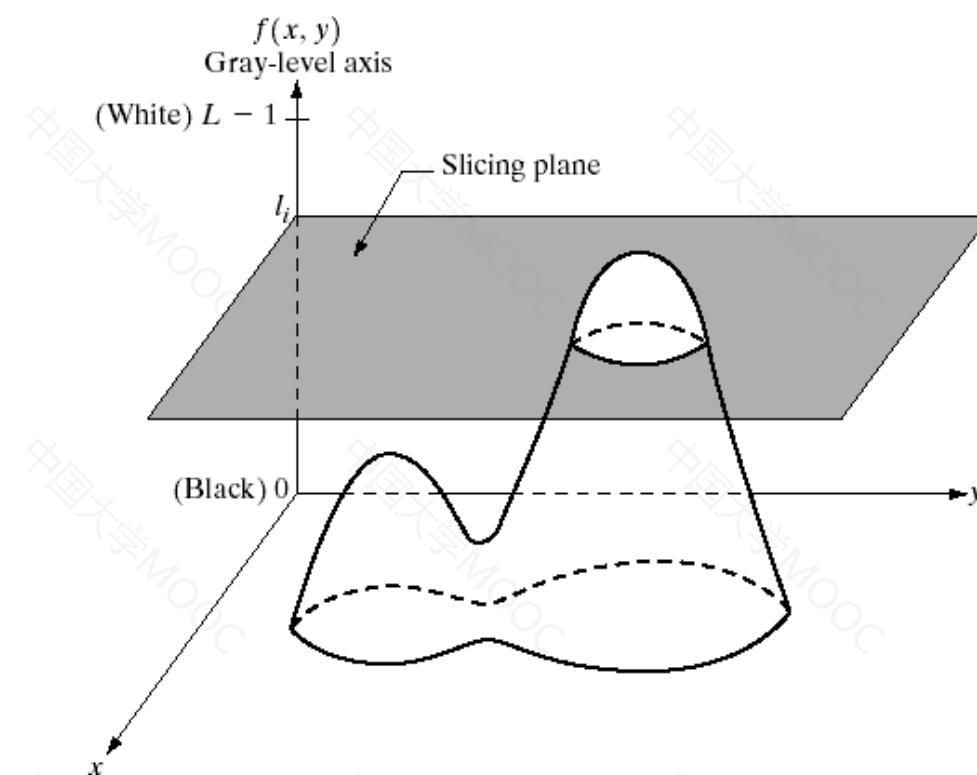


FIGURE 6.18 Geometric interpretation of the intensity-slicing technique.

■ 3.4.2 伪彩色增强

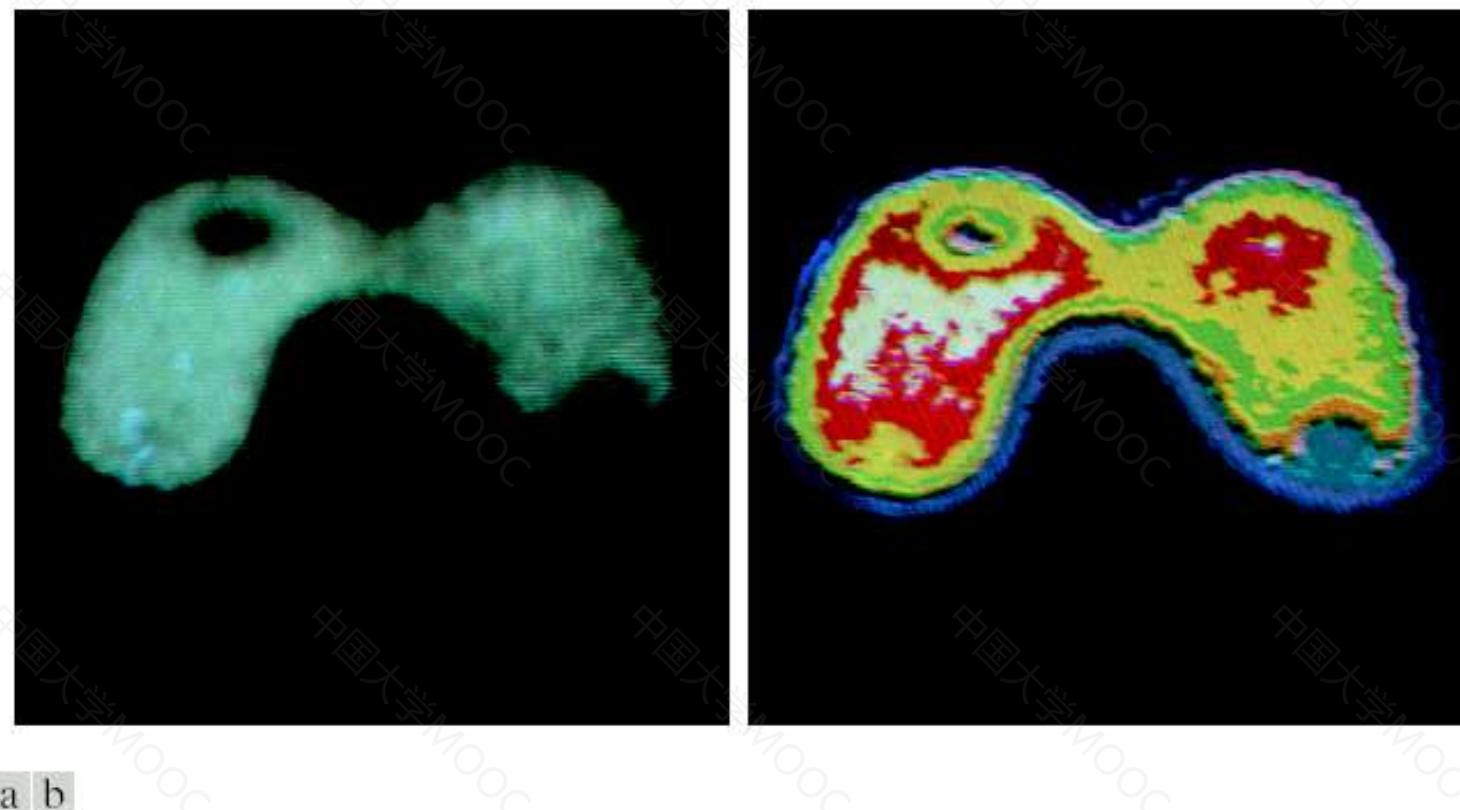
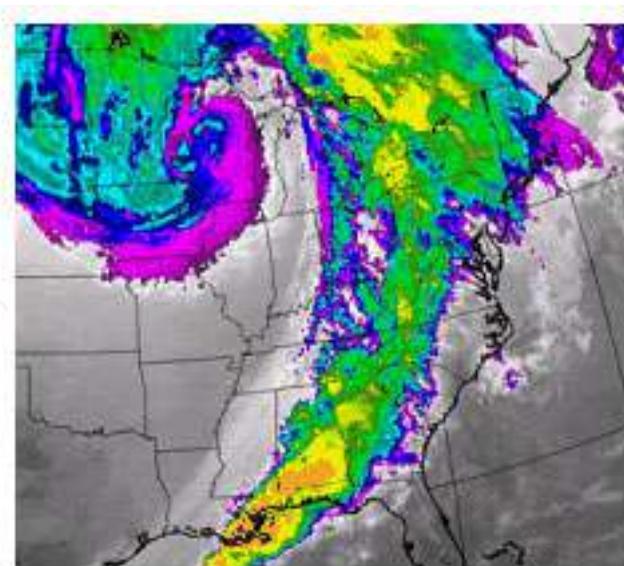


FIGURE 6.20 (a) Monochrome image of the Picker Thyroid Phantom. (b) Result of density slicing into eight colors. (Courtesy of Dr. J. L. Blankenship, Instrumentation and Controls Division, Oak Ridge National Laboratory.)

■ 伪彩色增强



灰度图像



密度切割法结果

■ 伪彩色增强

• 2. 灰度级彩色变换

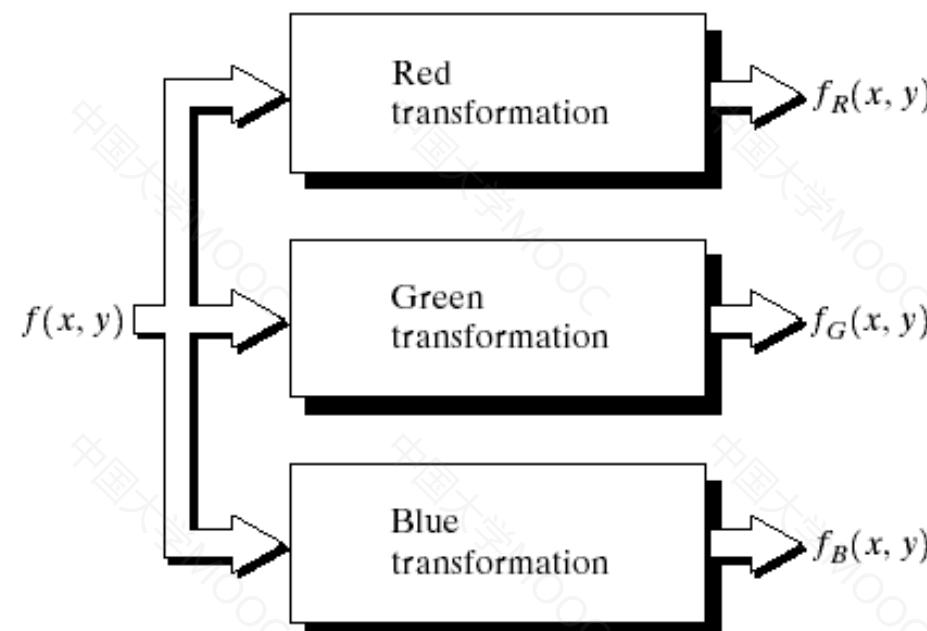
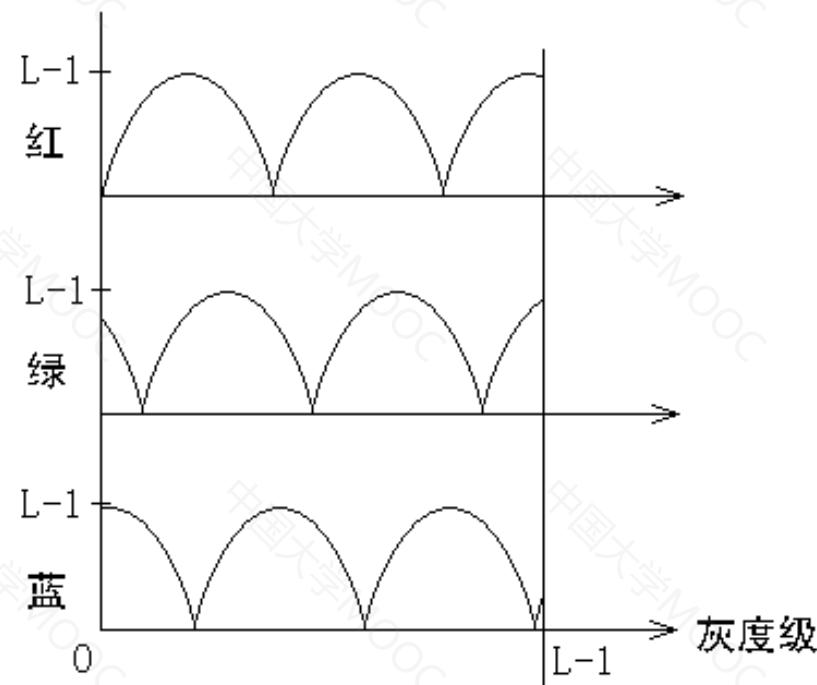


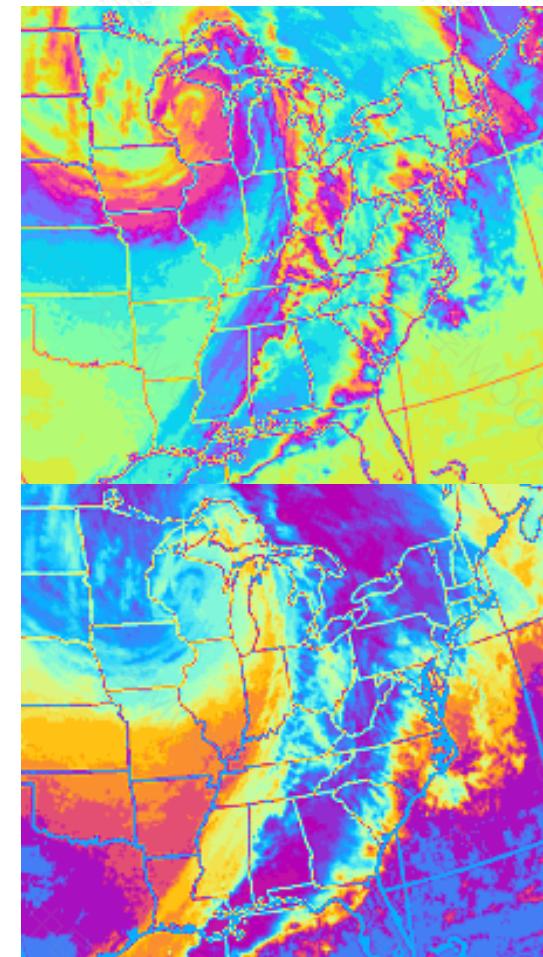
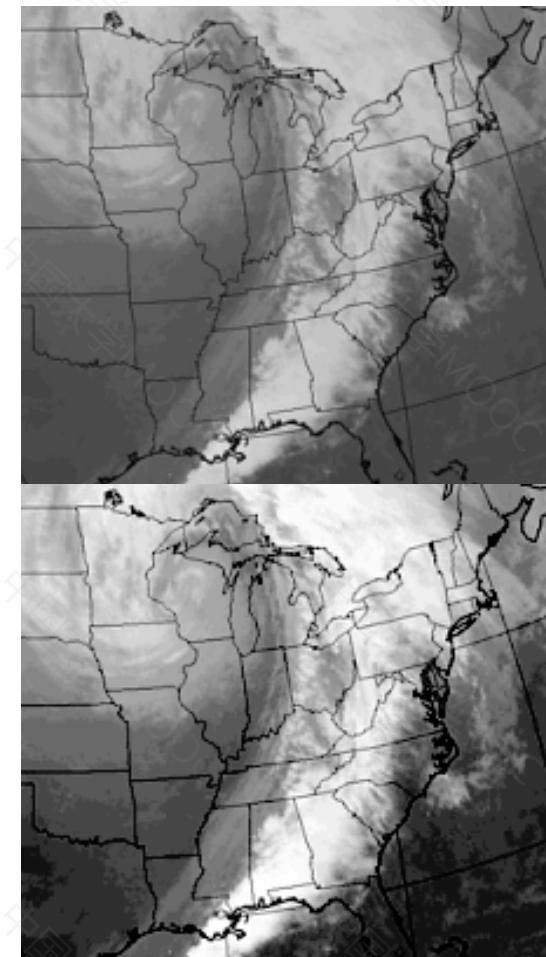
FIGURE 6.23 Functional block diagram for pseudocolor image processing. f_R , f_G , and f_B are fed into the corresponding red, green, and blue inputs of an RGB color monitor.

■ 伪彩色增强

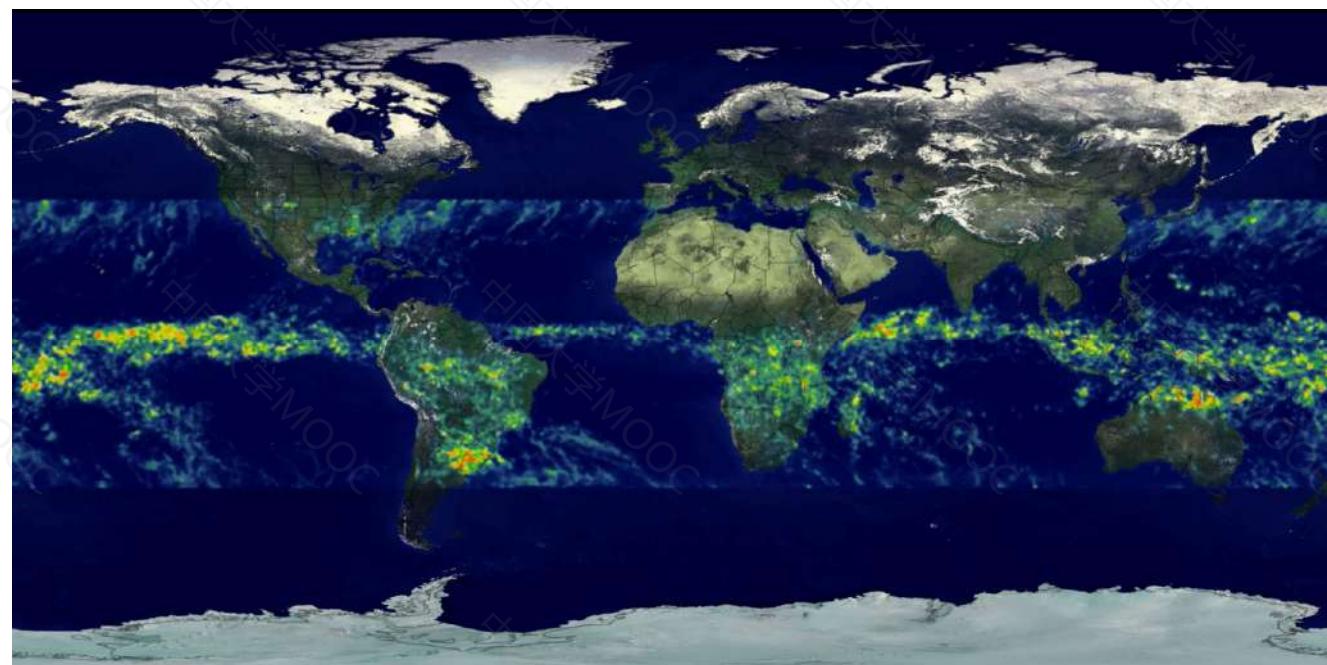
2. 灰度级彩色变换



■ 伪彩色增强

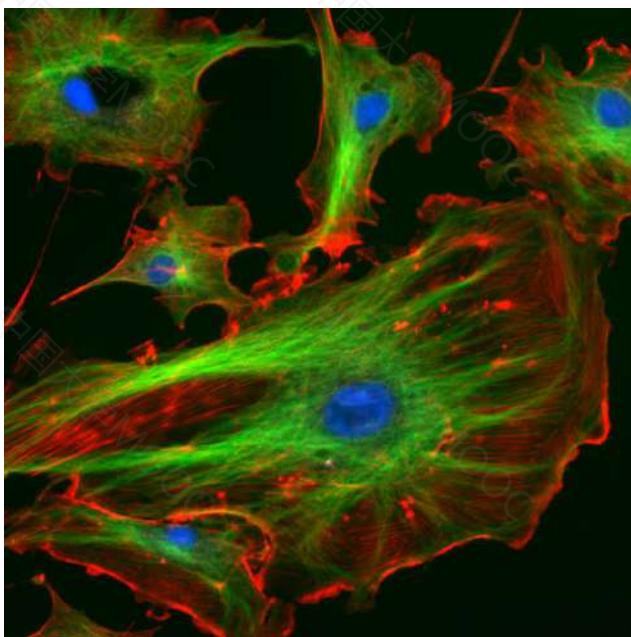


伪彩色增强

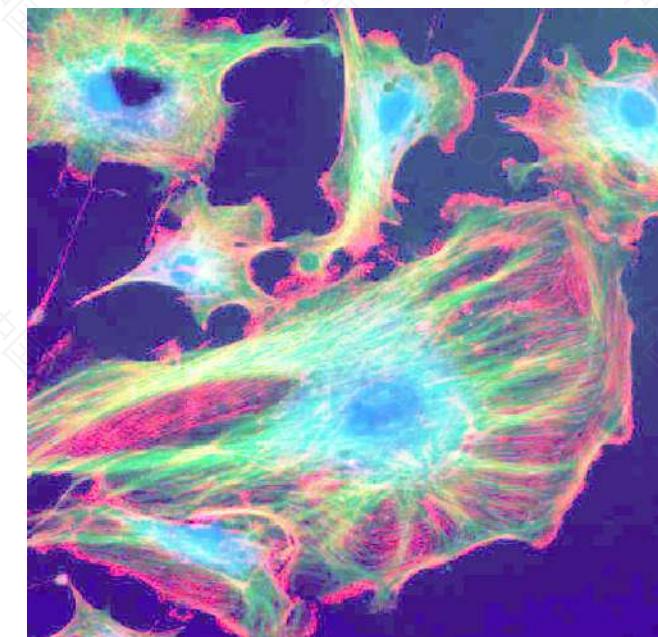


灰度级彩色变换法举例

■ 彩色直方图均衡效果



输入图像



RGB三分量分别做直方图均衡结果图像

■彩色直方图均衡方法

1) HSI空间

直接对I分量进行直方图均衡 $I \rightarrow I'$

由 (I', H, S) 重构原始图像

2) YCbCr空间

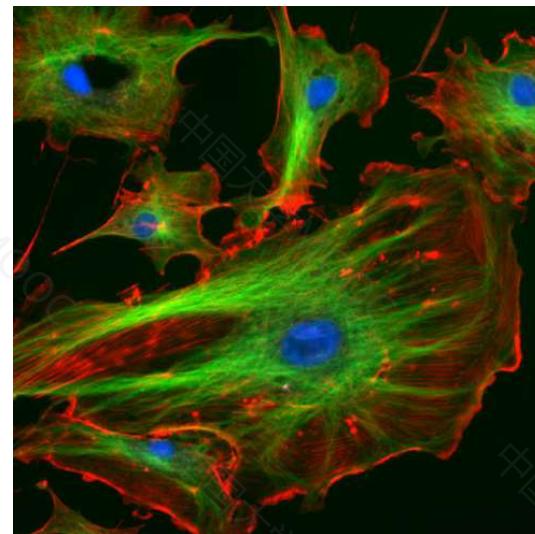
直接对Y分量进行直方图均衡 $Y \rightarrow Y'$

由 (Y', Cb, Cr) 重构原始图像

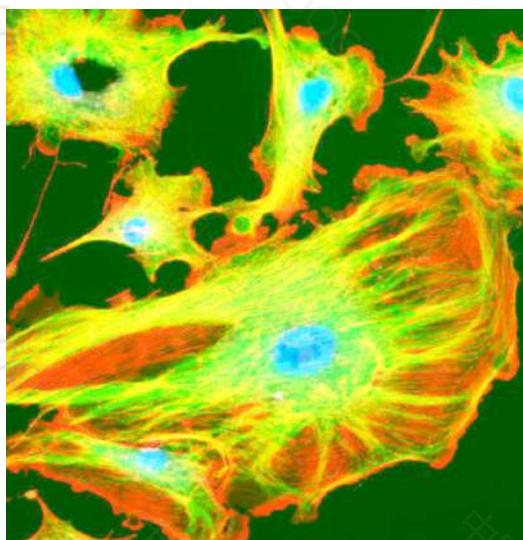
3) RGB空间

每个像素点分别计算R, G, B三个分量的映射比例 rR, rG, rB , 找到最大比例

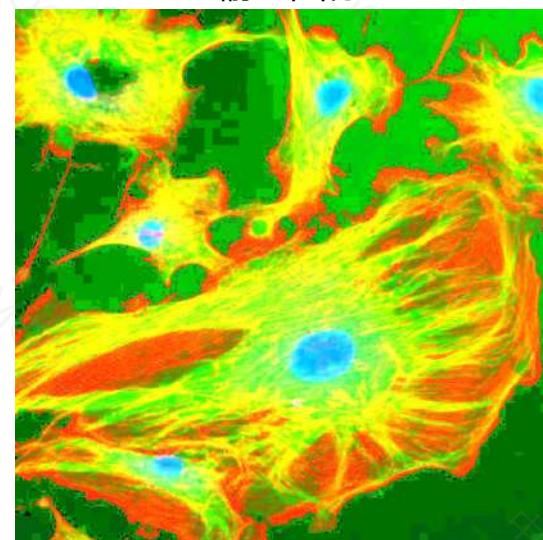
$r_{max} = \max\{rR, rG, rB\}$, 用该比例估计映射后的 R', G', B' , 如所有分量均小于等于255, 则输出相应像素点。否则取第二大比例重复上述过程, 直到三个颜色分量同比例映射后的所有分量均小于255。如所有比例都不合适, 则输出原始像素点的值。



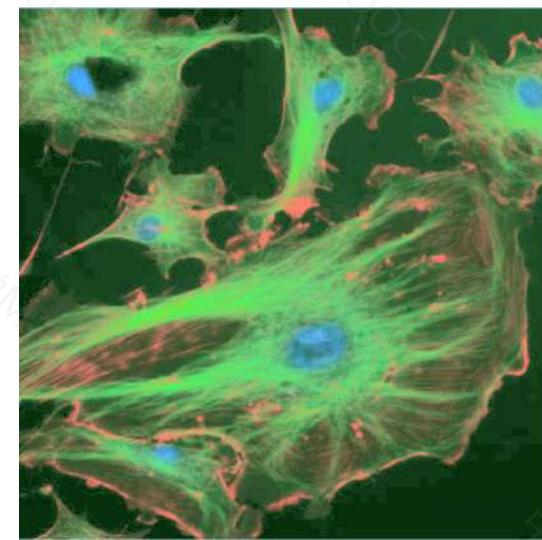
输入图像



基于RGB的直方图均衡结果



基于HSI的直方图均衡结果



基于YCbCr的直方图均衡结果



输入图像



基于RGB的直方图均衡结果

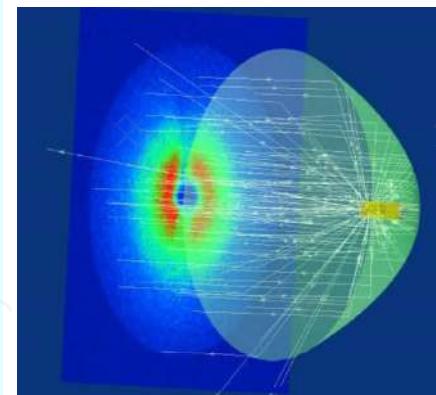
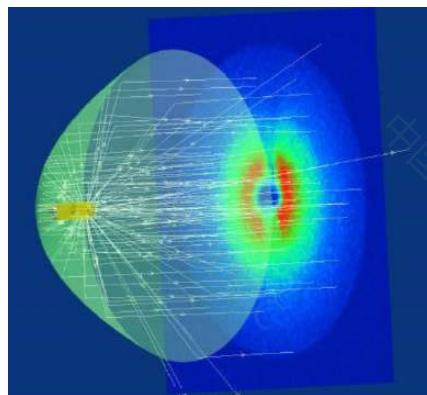


基于HSI的直方图均衡结果



基于YUV的直方图均衡结果

第四章 图像变换域 增强



引言

- 面向特定目标对图像进行处理，以改善图像的视觉质量。
- 举例：直方图均衡，边缘增强
- 特点
 - 方法设计比较主观
 - 突出图像中感兴趣的部分，达到局部最优
 - 面向目标

引言

- 4. 1 离散傅里叶变换
- 4. 2 通带滤波
- 4. 3 同态滤波

4.1 离散傅立叶变换 (Discrete Fourier Transform)

一. 离散傅立叶变换定义

$$F(u, v) = A \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})}$$

$$f(x, y) = B \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})}$$

$$A * B = \frac{1}{MN} \begin{matrix} & A & \frac{1}{MN} & \frac{1}{\sqrt{MN}} & 1 \\ & B & 1 & \frac{1}{\sqrt{MN}} & \frac{1}{MN} \end{matrix}$$

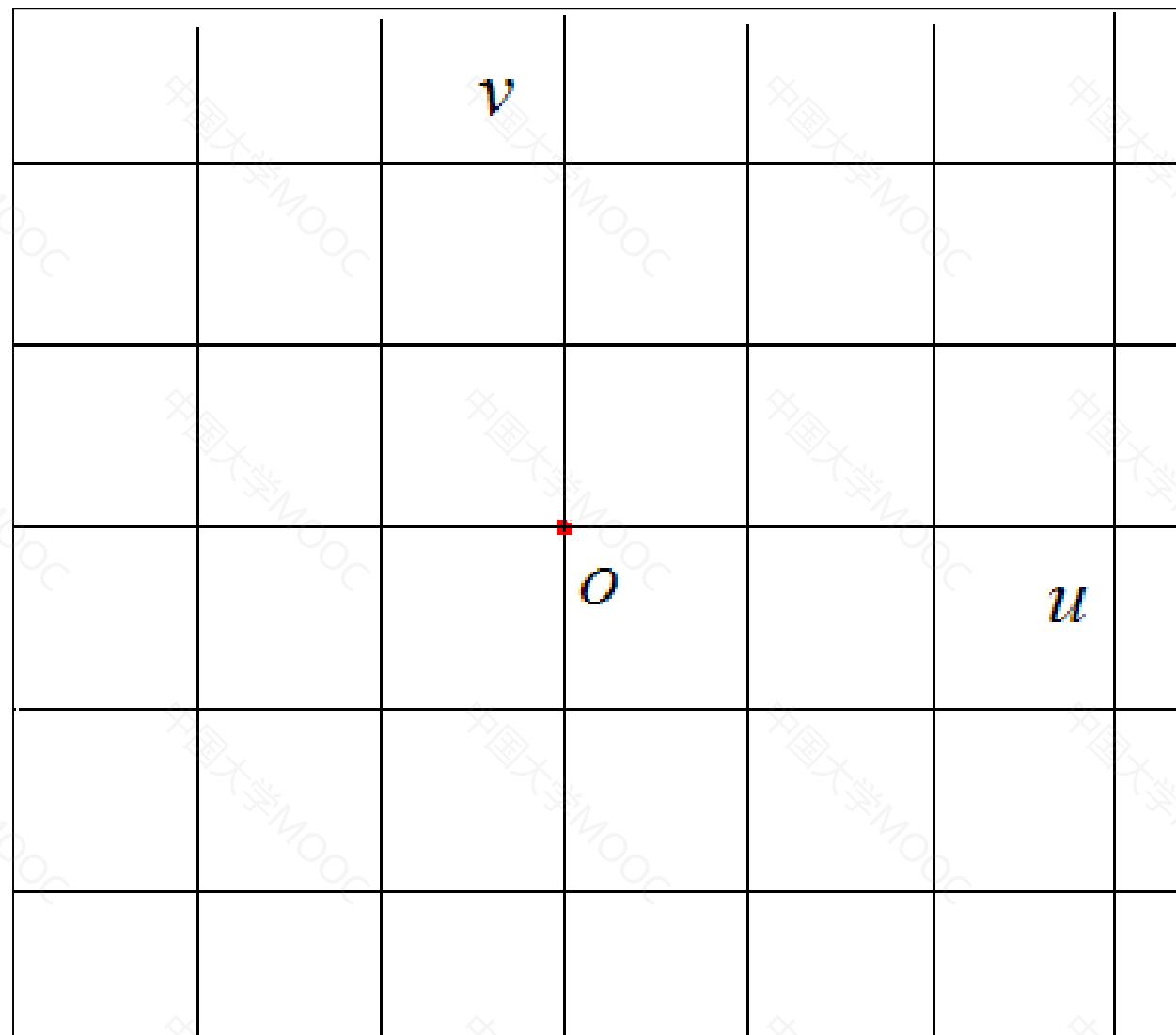
■ 二. 离散傅立叶变换性质

■ 1. 线性

$$f_1(x, y) \longleftrightarrow F_1(u, v)$$

$$f_2(x, y) \longleftrightarrow F_2(u, v)$$

$$af_1(x, y) + bf_2(x, y) \longleftrightarrow aF_1(u, v) + bF_2(u, v)$$



2. 周期性

$$F(u, v) = F(u + aM, v + bN)$$

$$\begin{aligned} F(u + aM, v + bN) &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j[\frac{2\pi}{M}(u+aM)x + \frac{2\pi}{N}(v+bN)y]} \\ &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi[\frac{ux}{M} + \frac{vy}{N}]} e^{-j2\pi[ax+by]} \\ &= F(u, v) \end{aligned}$$

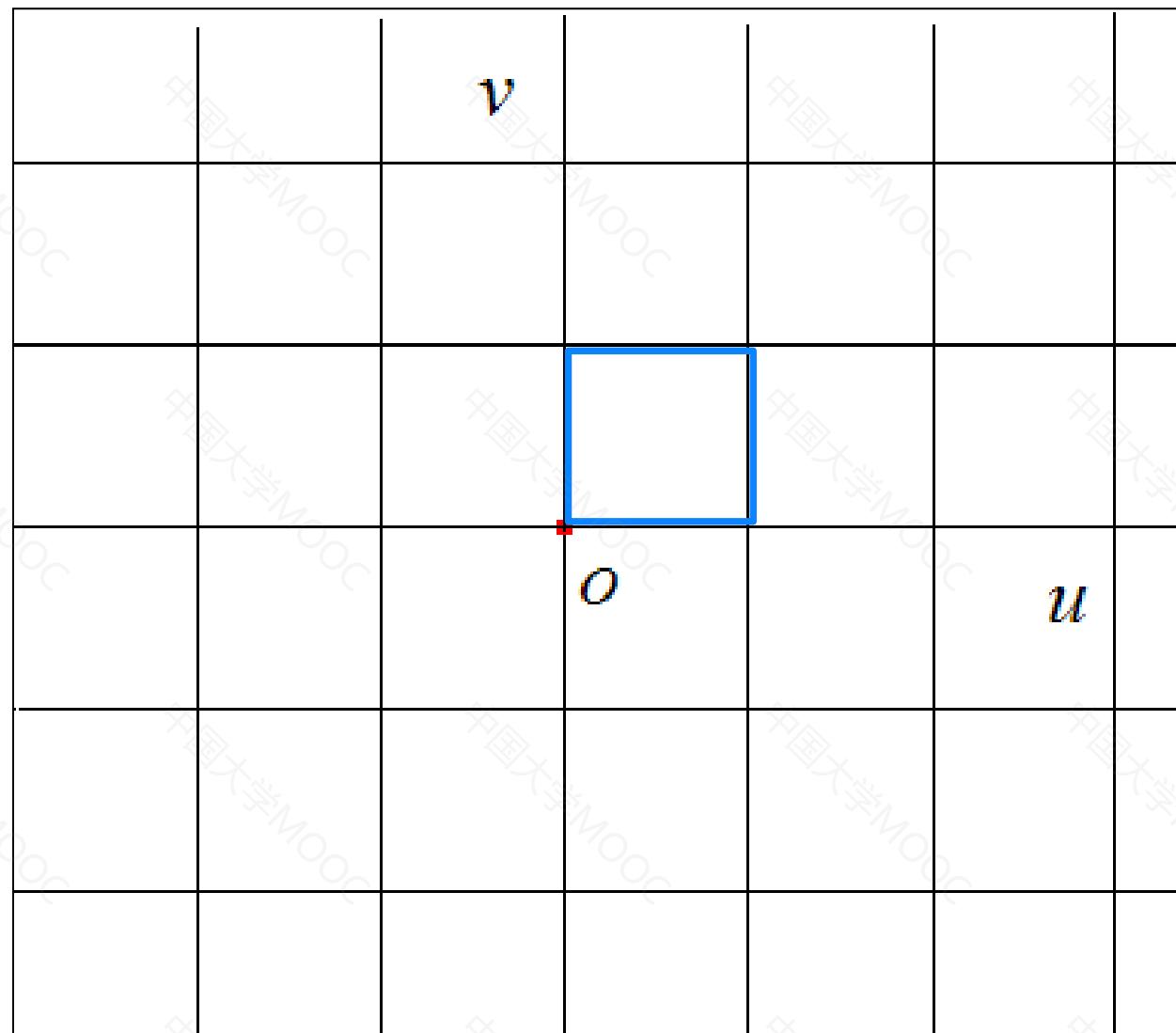
一. 离散傅立叶变换定义

$$F(u, v) = A \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} d$$

$u=0, 1, 2, \dots, M-1, v=0, 1, 2, \dots, N-1$

$$f(x, y) = B \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{ux}{M} + \frac{vy}{N})} d_l$$

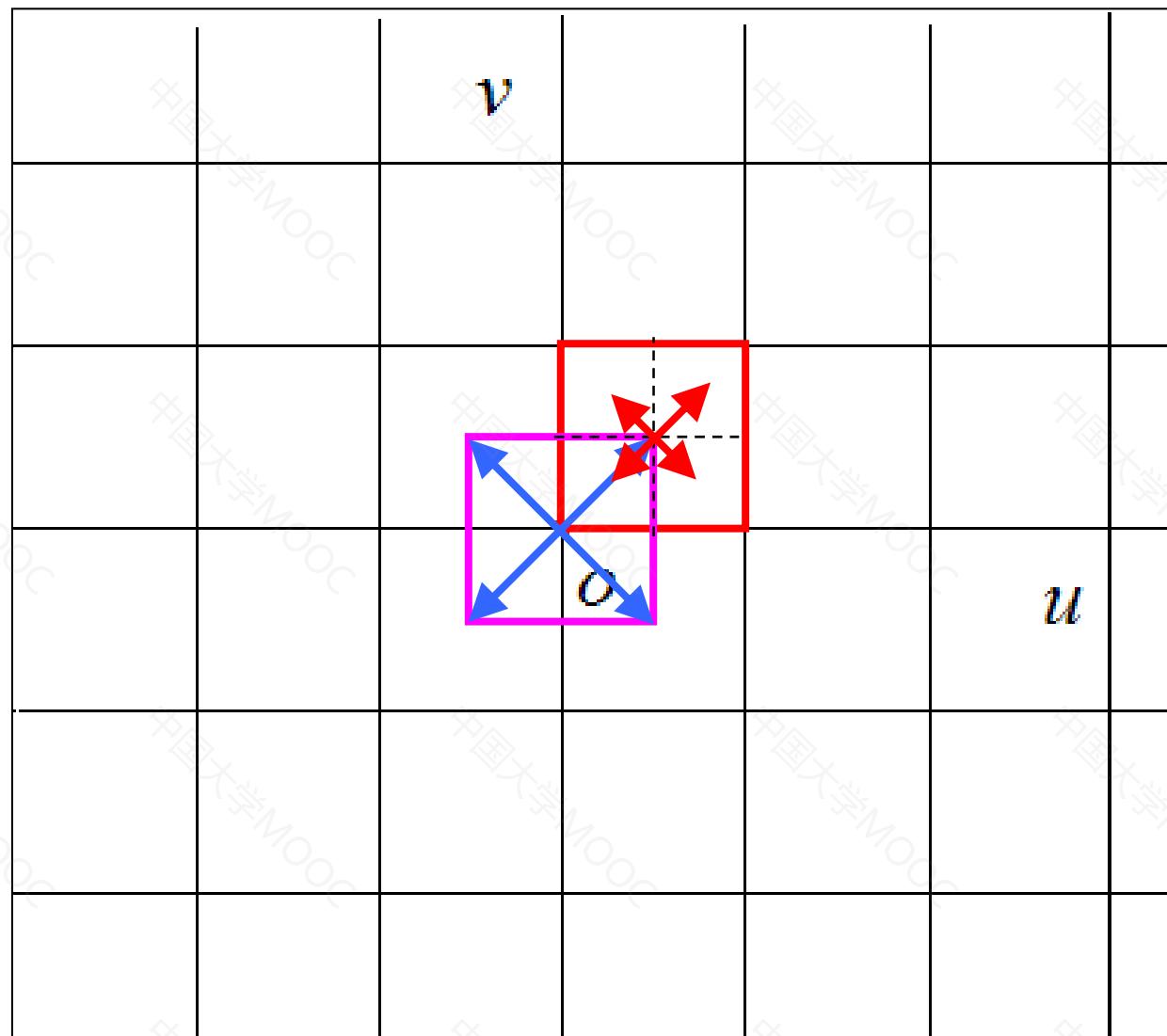
$x=0, 1, 2, \dots, M-1, y=0, 1, 2, \dots, N-1$



3. 对称性

$$F(u, v) = F^*(-u, -v)$$

$$\begin{aligned} F * (-u, -v) &= \\ &\left[\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j\left[\frac{2\pi}{M}(-u)x + \frac{2\pi}{N}(-v)y\right]} \right]^* \\ &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f^*(x, y) e^{-j2\pi\left[\frac{ux}{M} + \frac{vy}{N}\right]} \\ &= \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi\left[\frac{ux}{M} + \frac{vy}{N}\right]} \\ &= F(u, v) \end{aligned}$$



4. 平移特性

空域（频域）中点的平移在频域（空域）并没有产生幅度上的改变，只是有附加相移产生。

$$f(x, y)e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} \Leftrightarrow F(u - u_0, v - v_0)$$

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)e^{-j2\pi(\frac{ux_0}{M} + \frac{vy_0}{N})}$$

为了获得完整信号频谱

$$u_0 = \frac{M}{2}, \quad v_0 = \frac{N}{2}$$

$$f(x, y)e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} = f(x, y)(-1)^{x+y} \Leftrightarrow F(u - \frac{M}{2}, v - \frac{N}{2})$$

图像空域平移示例

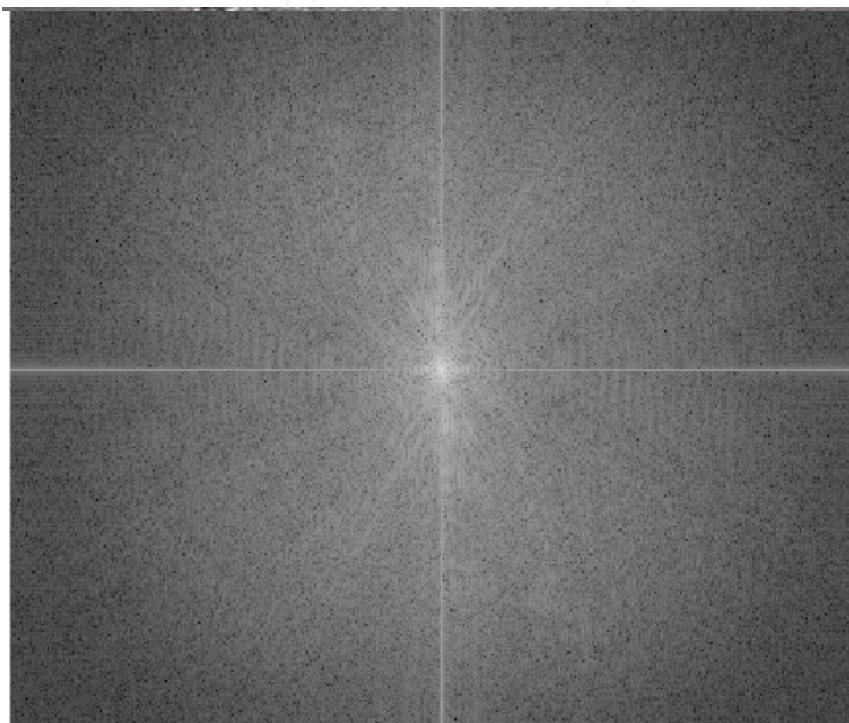


注册人脸图像

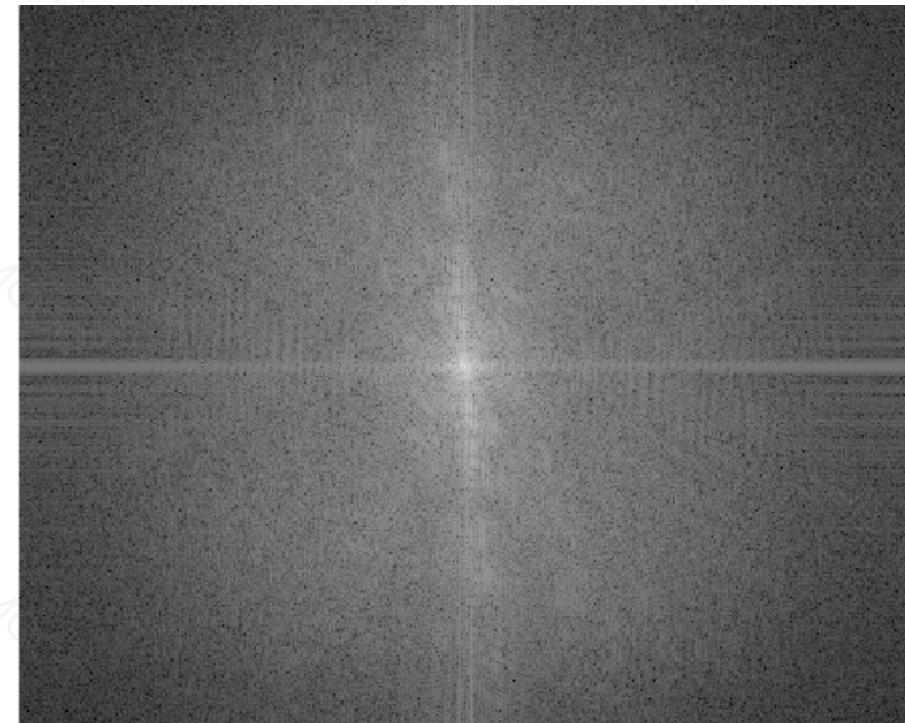


测试人脸图像

图像空域平移示例

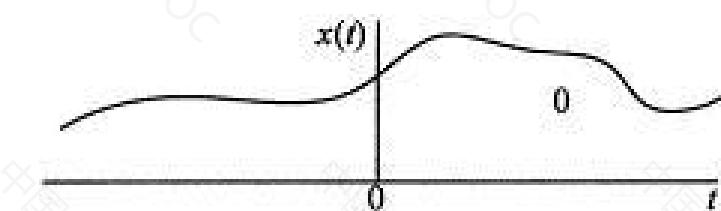


注册人脸图像幅度谱

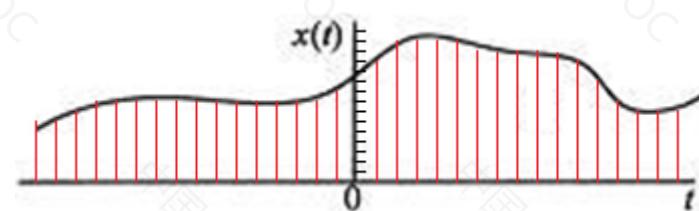


测试人脸图像幅度谱

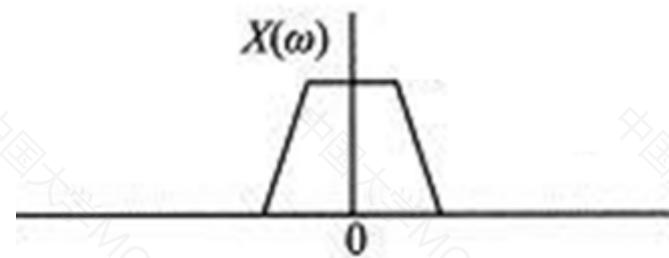
两幅人脸图像的幅度谱很相似



模拟信号

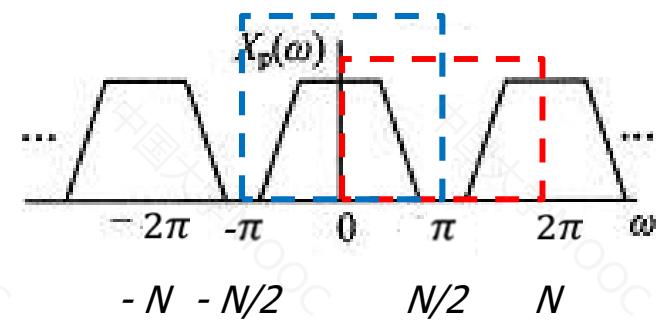


数字信号

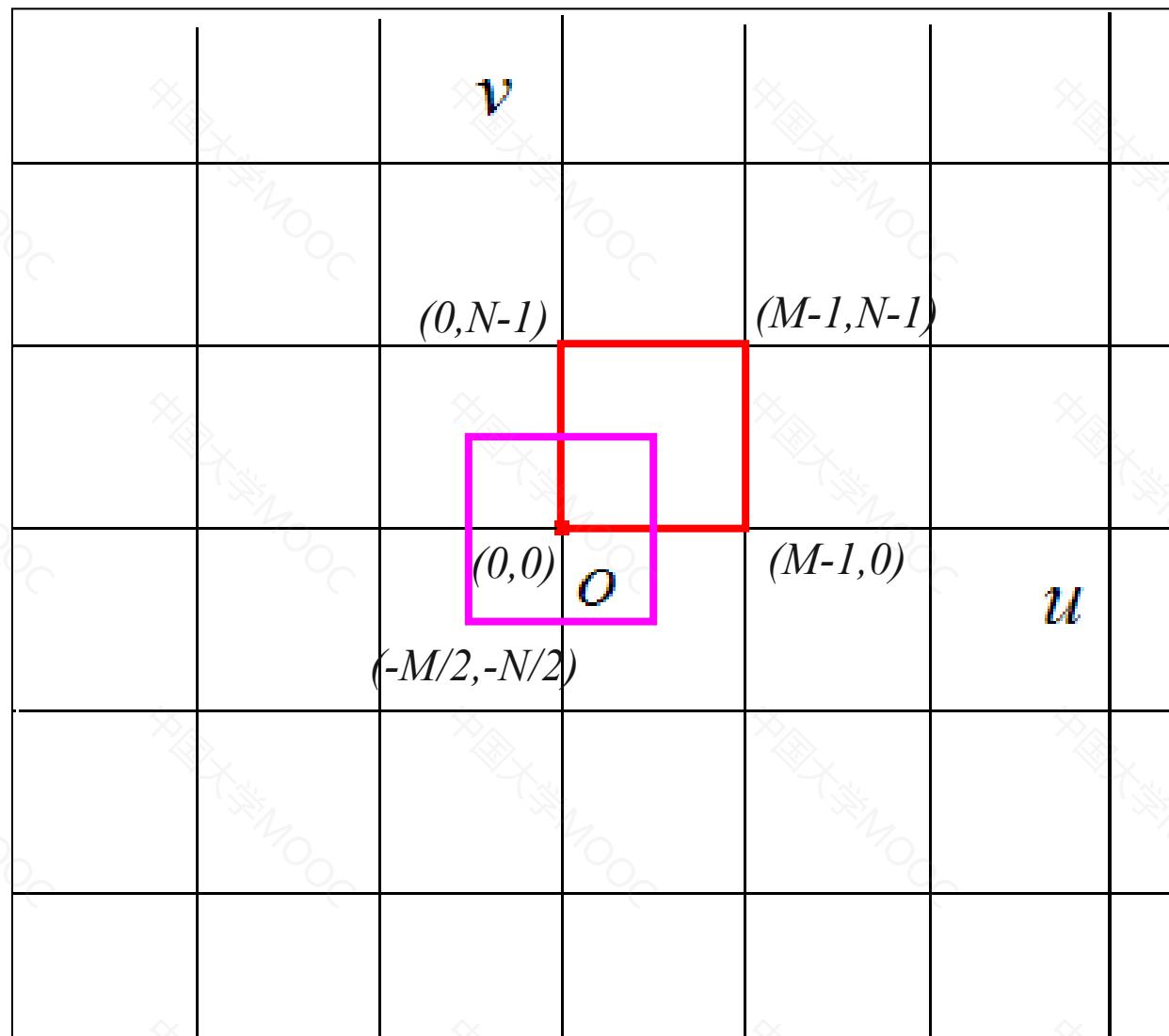


高频 低频 高频

模拟信号频谱



数字信号频谱



4. 平移特性

空域（频域）中点的平移在频域（空域）并没有产生幅度上的改变，只是有附加相移产生。

$$f(x, y)e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} \Leftrightarrow F(u - u_0, v - v_0)$$

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)e^{-j2\pi(\frac{ux_0}{M} + \frac{vy_0}{N})}$$

为了获得完整信号频谱

$$u_0 = \frac{M}{2}, \quad v_0 = \frac{N}{2}$$

$$f(x, y)e^{j2\pi(\frac{u_0x}{M} + \frac{v_0y}{N})} = f(x, y)(-1)^{x+y} \Leftrightarrow F(u - \frac{M}{2}, v - \frac{N}{2})$$

5. 可分离性

$$\begin{aligned} F(u, v) &= \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi (\frac{ux}{M} + \frac{vy}{N})} \\ &= \sum_{x=0}^{M-1} e^{-j2\pi \frac{ux}{M}} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \frac{vy}{N}} \\ &= \sum_{x=0}^{M-1} e^{-j2\pi \frac{ux}{M}} F(x, v) = F(u, v) \end{aligned}$$

6. 旋转特性

$$f(x, y) \Leftrightarrow F(u, v)$$

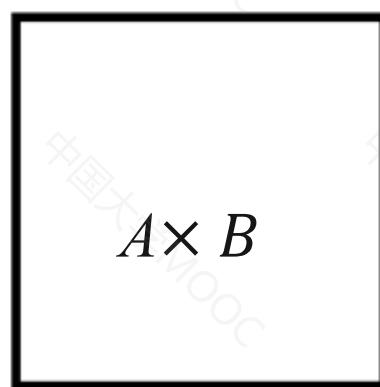
$$\begin{aligned} & f(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta) \\ & \Leftrightarrow F(u \cos \theta + v \sin \theta, -u \sin \theta + v \cos \theta) \end{aligned}$$

图像在空域旋转 θ 角，其频谱在频域旋转 θ 角

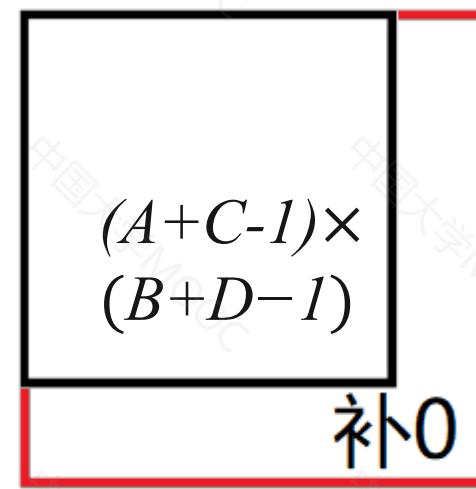
7. 卷积定理

$$f(x, y) * h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n)h(x - m, y - n)$$

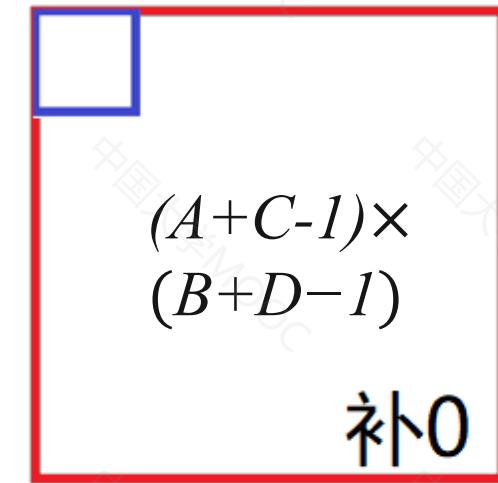
$$f(x, y) * h(x, y) \Leftrightarrow F_e(u, v) \bullet H_e(u, v)$$



$f(x, y)$



$f_e(x, y)$

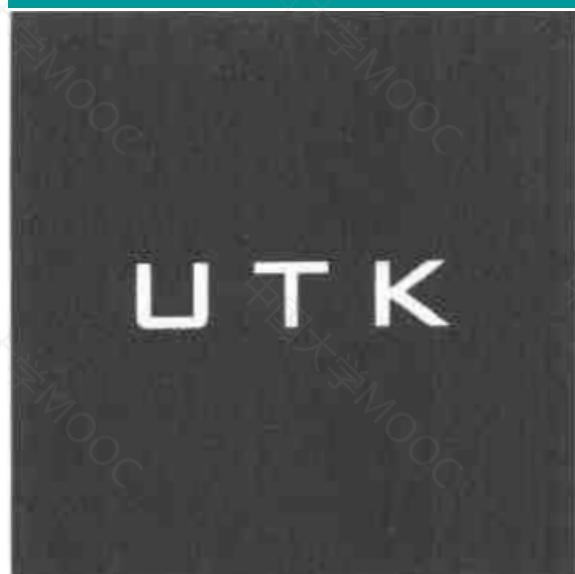


$h_e(x, y)$

8.相关定理

$$f(x, y) \circ h(x, y) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) h(x + m, y + n)$$

$$f(x, y) \circ h(x, y) \Leftrightarrow F_e(u, v) \bullet H^*_e(u, v)$$



$f(x, y)$



$h(x, y)$



$g(x, y) = f(x, y) \circ h(x, y)$


$$f(x,y)$$
T
$$h(x,y)$$


$$g(x,y) = IDFT(F_e(u,v)H^*_e(u,v))$$

T
$$f_e(x,y)$$
$$h_e(x,y)$$


2021/12/2

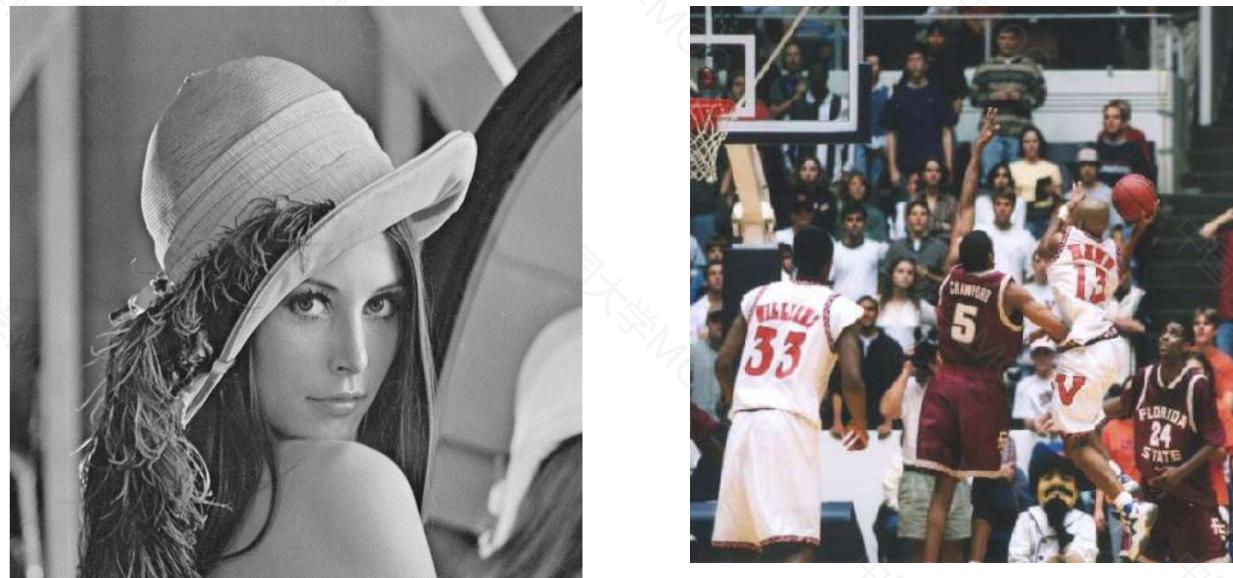
■ 9 微分

$$DFT(f(x,y)) = F(u,v)$$

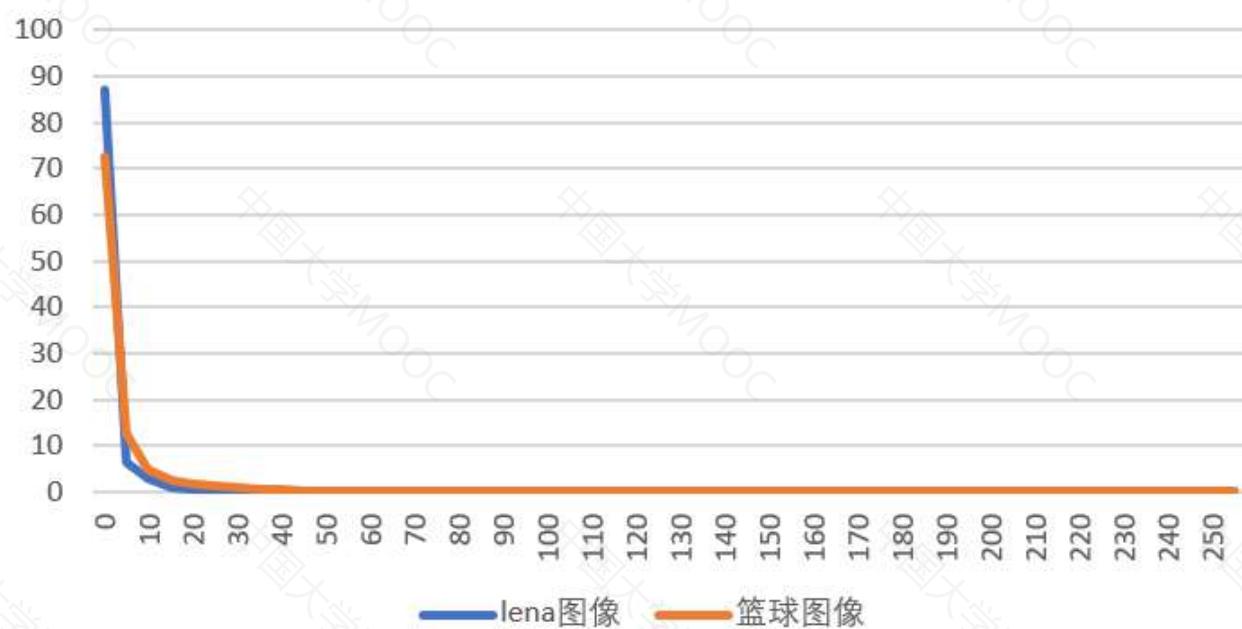
$$DFT\left(\frac{\partial^n f(x,y)}{\partial x^n}\right) = (ju)^n F(u,v)$$

$$DFT\left(\frac{\partial^n f(x,y)}{\partial y^n}\right) = (jv)^n F(u,v)$$

- ### 三. 图像频谱分析
- 1 图像纹理特征
- 图像频谱是图像纹理特征的反映，
如果图像具有比较平坦的纹理，则
图像低频分量丰富，反之，若图像
细节较多，则图像高频分量丰富。



频谱分布对比



3 相位信息的重要性

$$F(u, v) = R(u, v) + jI(u, v)$$

幅度谱

$$|F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)}$$

相位谱

$$\theta(u, v) = \arctan\left(\frac{R(u, v)}{I(u, v)}\right)$$

三种方式重建图像



幅度谱

相位谱

■ 方式1：

$$|F(u, v)|$$

$$\theta(u, v)$$

■ 方式2

$$|F(u, v)|$$

$$\theta(u, v) = 0$$

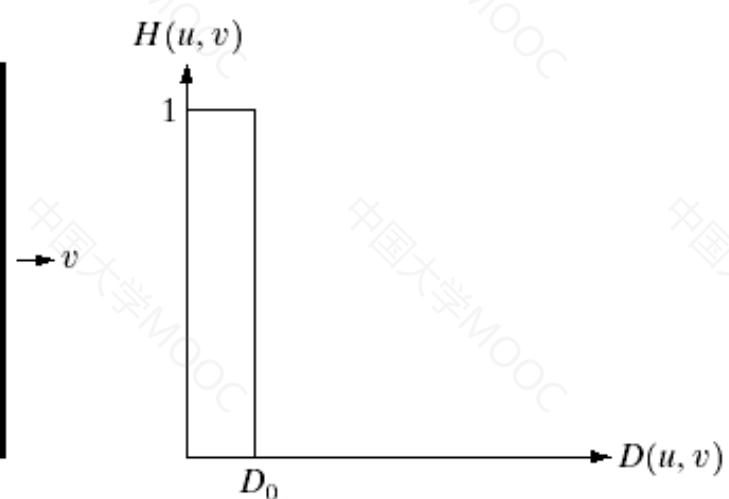
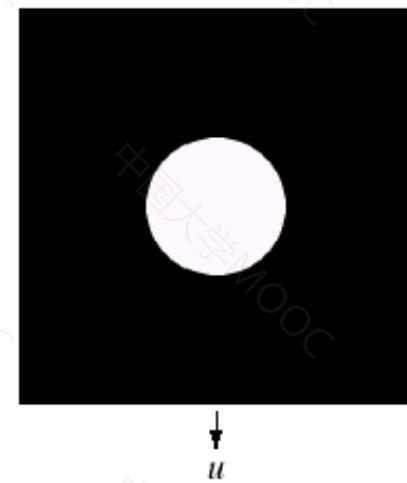
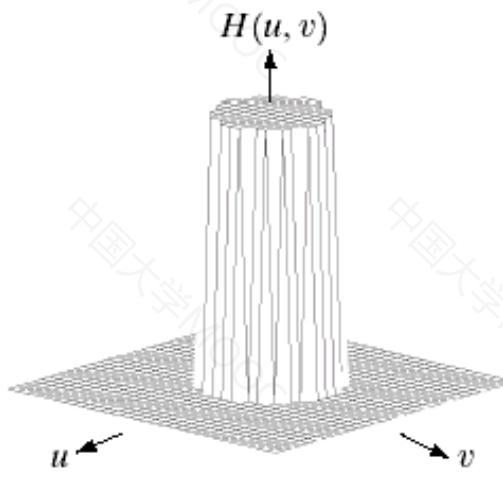
■ 方式3：

$$|F(u, v)| = A$$

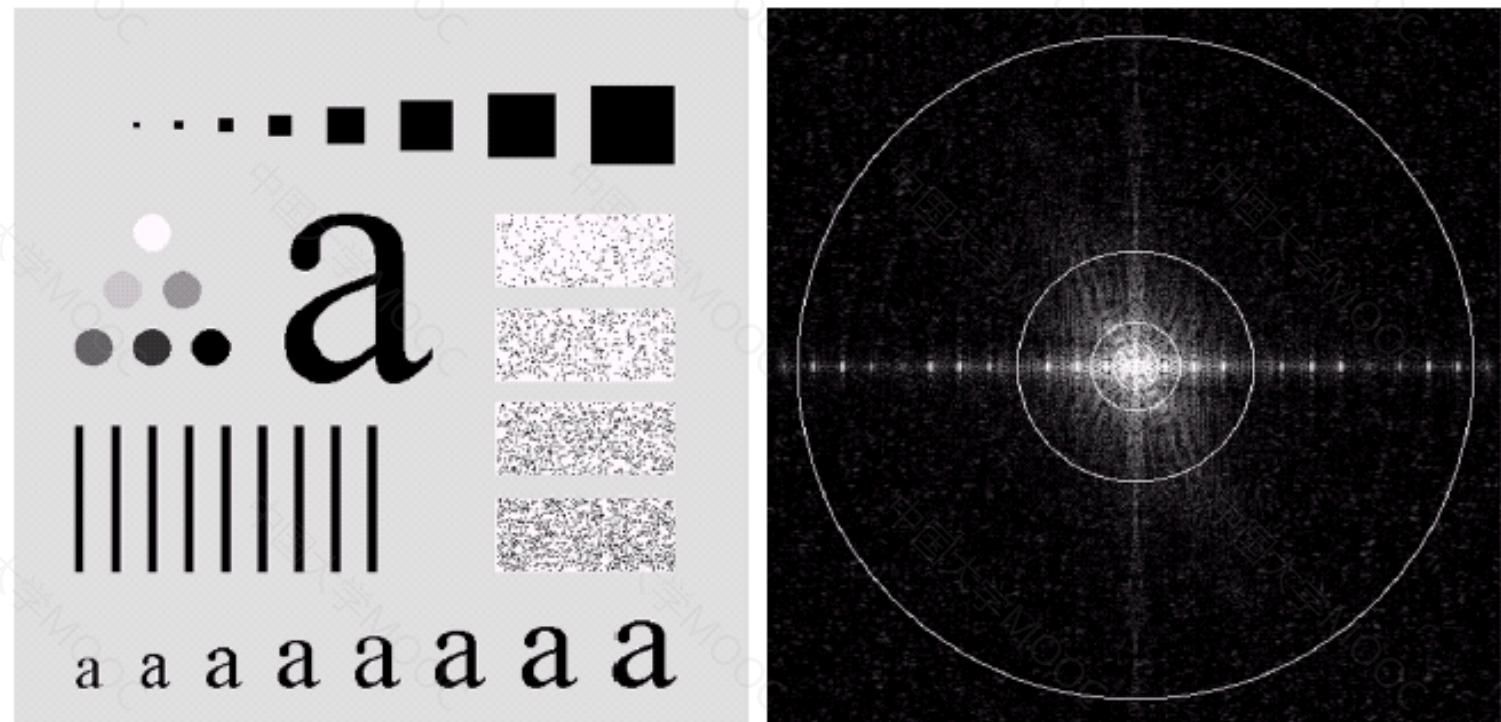
$$\theta(u, v)$$

一、通带滤波





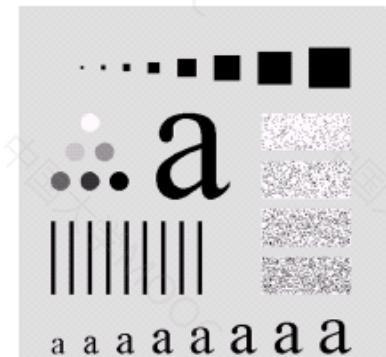
理想低通滤波器



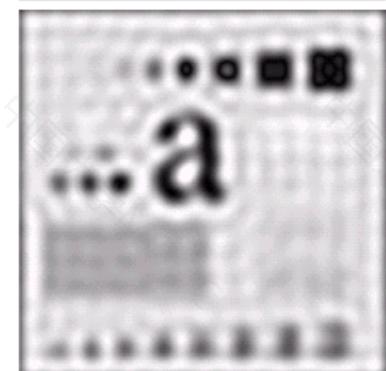
频谱特性分析

半径 (像素) : 5 15 30 80 230
功率占比: 92.0% 94.6% 96.4% 98.0% 99.5%

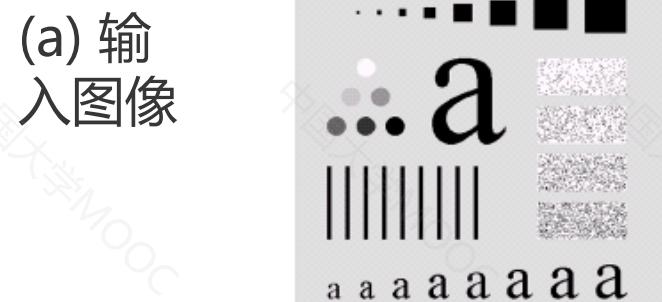
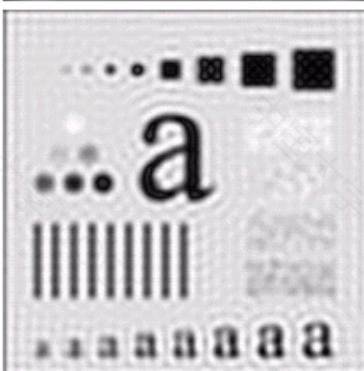
(a) 输入图像



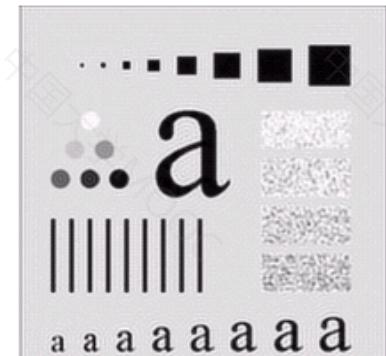
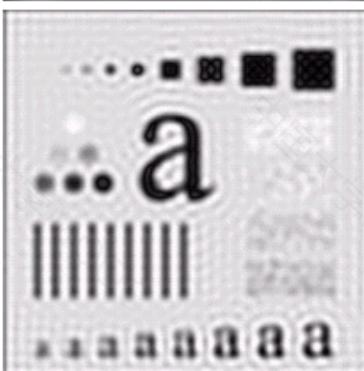
(b) 截止频率为15的滤波结果



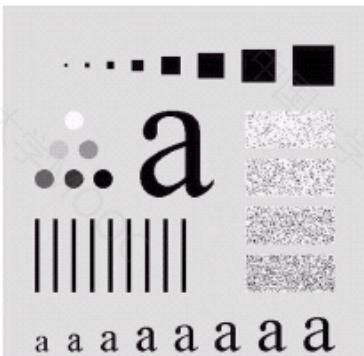
(c) 截止频率为30的滤波结果



(d) 截止频率为30的滤波结果



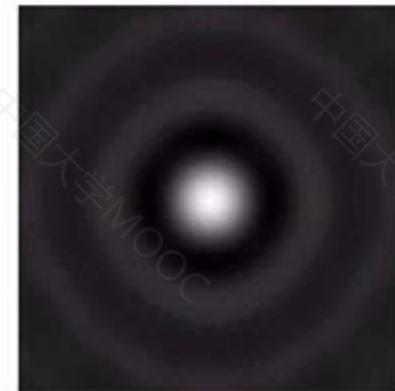
(e) 截止频率为80的滤波结果



理想低通滤波结果图像



(a) 半径为5 的理想低通滤波器的频域表达



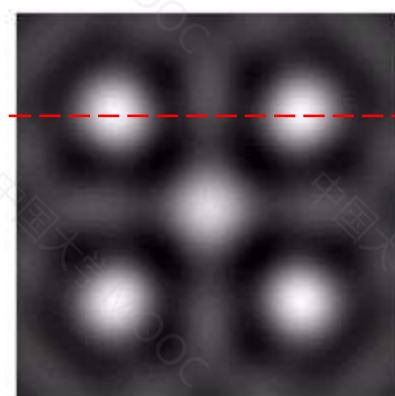
(b) 半径为5 的理想低通滤波器的空域表达



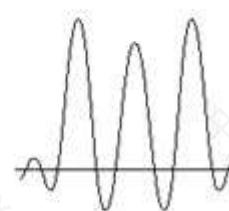
(c) 图(b)中心线上的值



(d) 输入图像在5的像素点有值

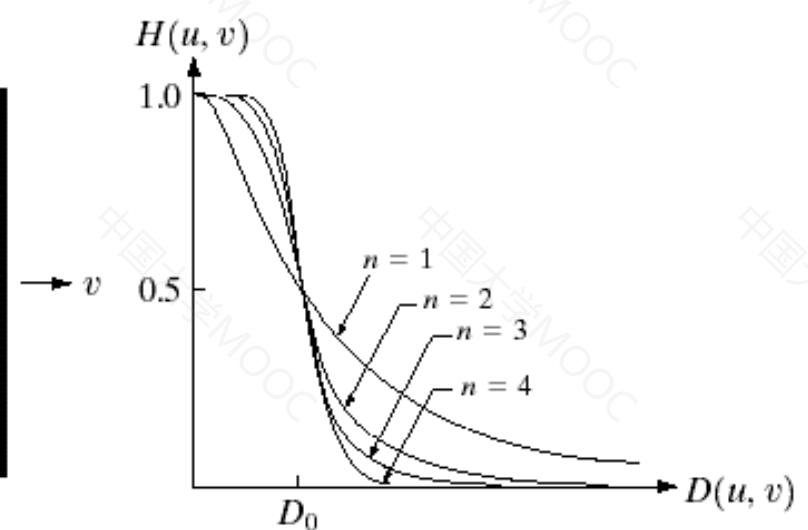
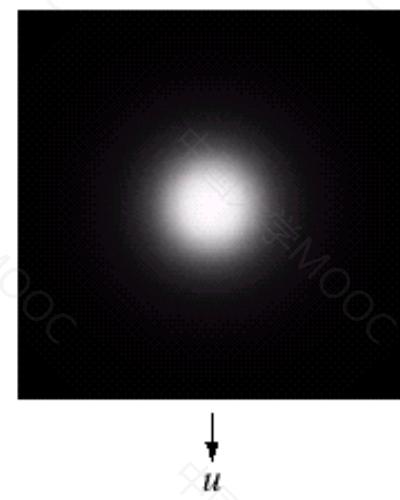
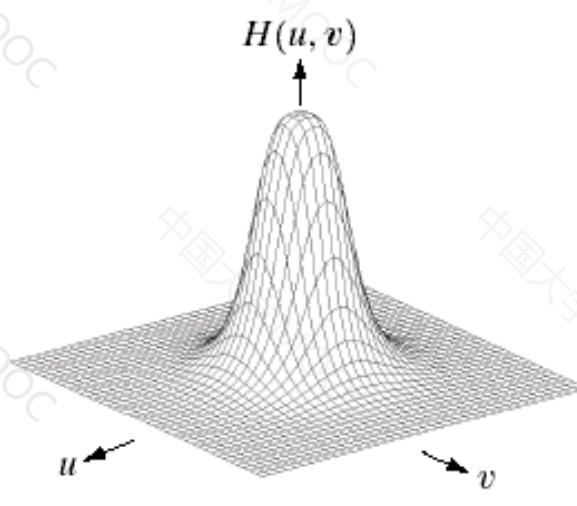


(e) 图(b) 和(d)卷积结果图像

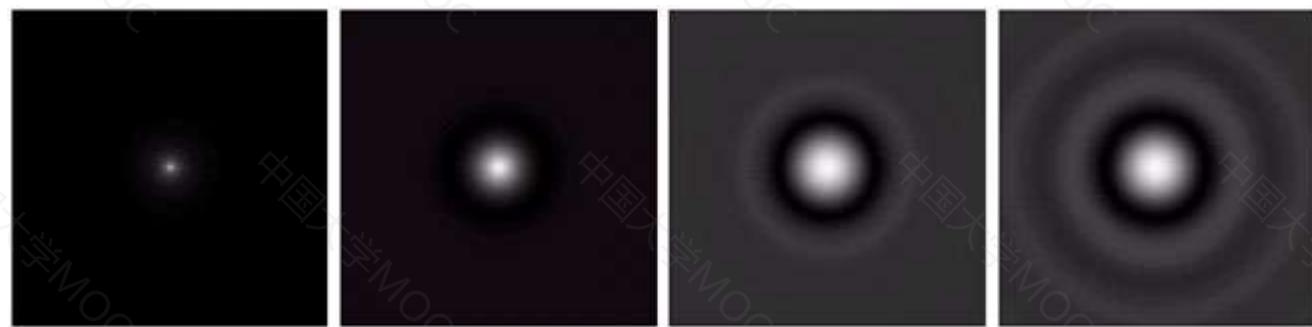


(c) 图(e)红色虚线上的值

振铃效应：就是指输出图像的灰度剧烈变化处产生的震荡



Butterworth 低通滤波器



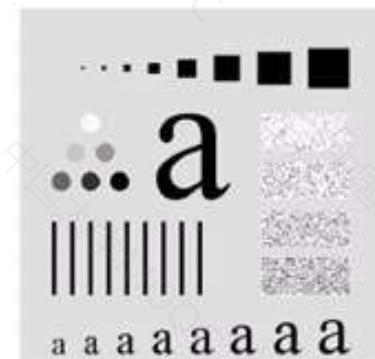
$n=1$

$n=2$

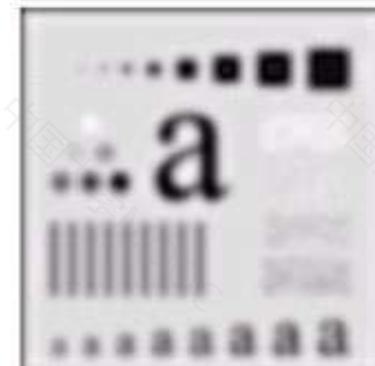
$n=5$

$n=20$

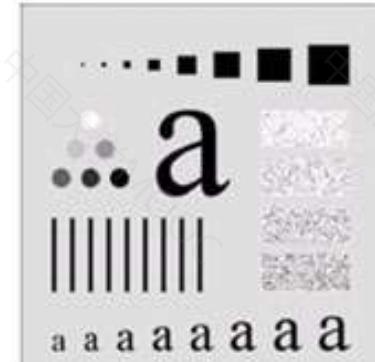
(a) 输入图像



(c) 截止频率为15的滤波结果



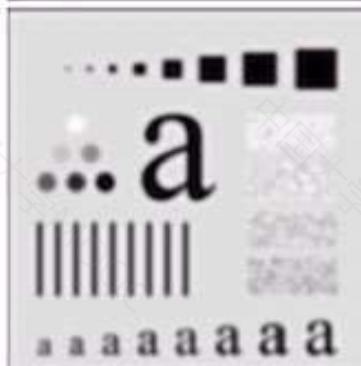
(e) 截止频率为80的滤波结果



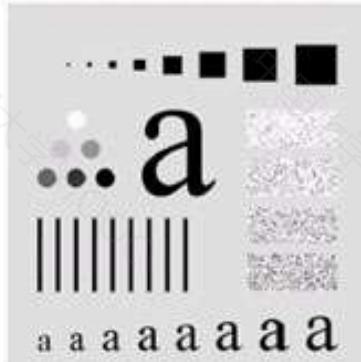
(b) 截止频率为5的滤波结果



(d) 截止频率为30的滤波结果

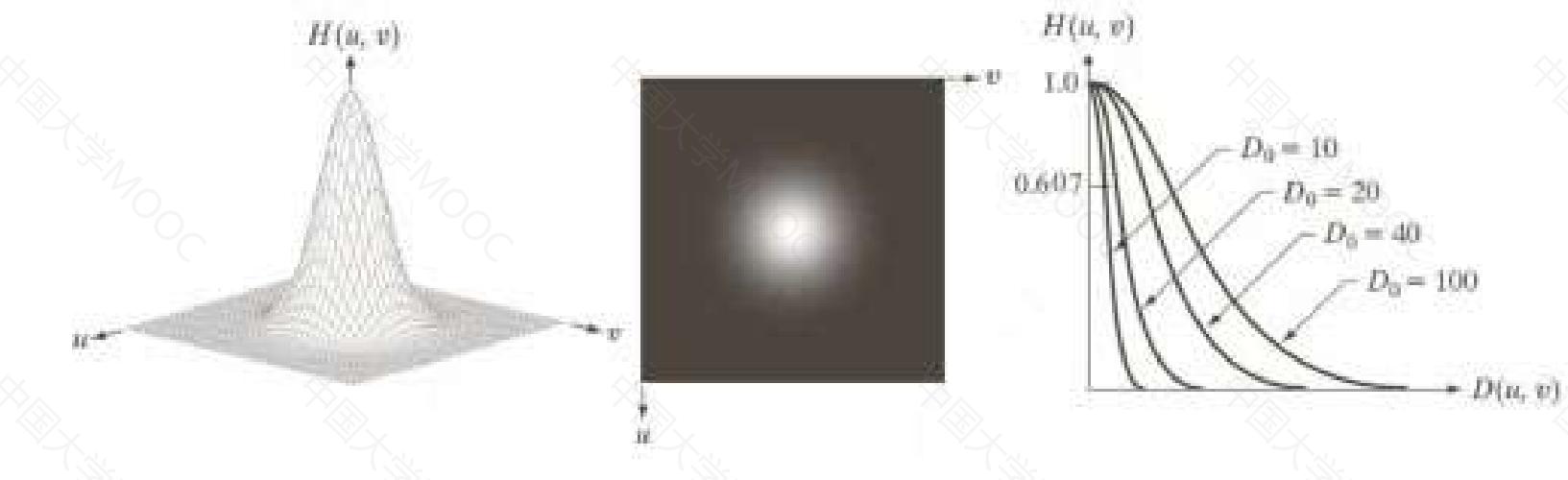


(f) 截止频率为230的滤波结果



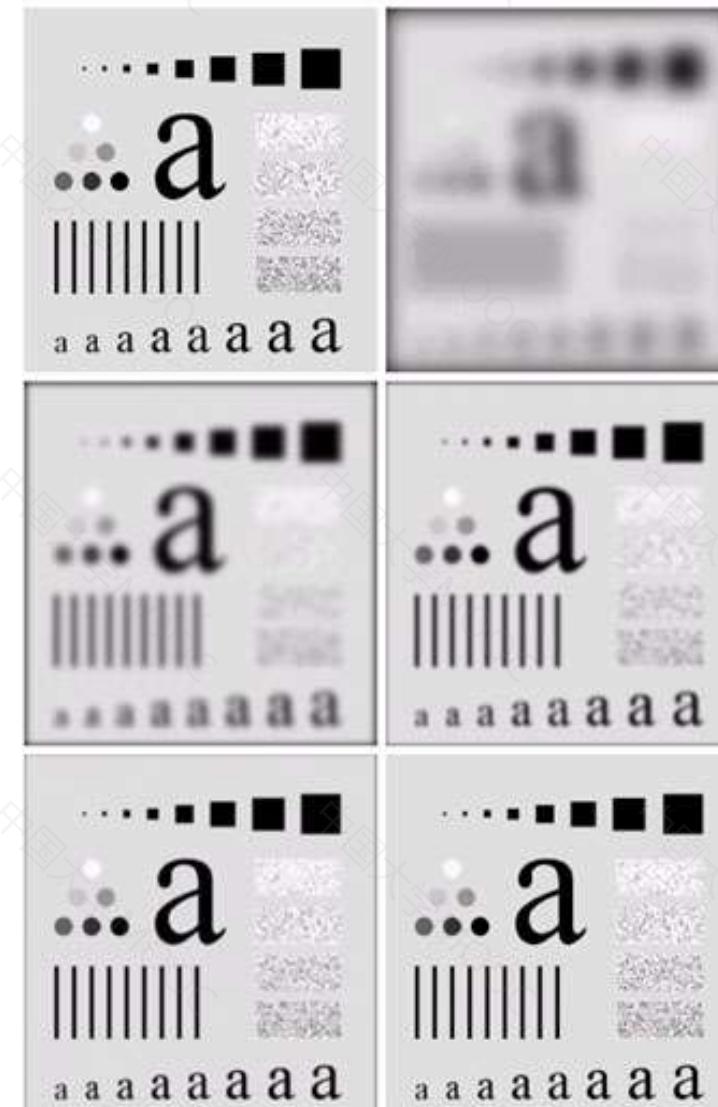
Butterworth低通滤波结果图像

$$H(u, v) = e^{-D^2(u, v)/2D_0^2}$$



高斯低通濾波器

(a) 输入图像



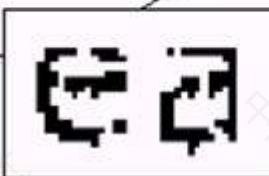
(b) 截止频率为5的滤波结果

(d) 截止频率为30的滤波结果

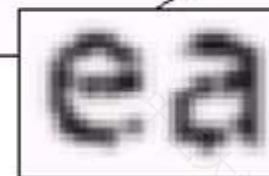
(f) 截止频率为230的滤波结果

高斯低通滤波结果图像

Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.



Historically, certain computer programs were written using only two digits rather than four to define the applicable year. Accordingly, the company's software may recognize a date using "00" as 1900 rather than the year 2000.



文本扫描图像文字不连续，低通滤波
处理后改善很多

低通滤波应用举例一



(a)

(b)

(c)

(a) 原始图像 (b) 截止频率100低通滤波后, 眼角皱纹有所改善,
, (c)截止频率80低通滤波后, 眼角皱纹进一步改善。

低通滤波应用举例二

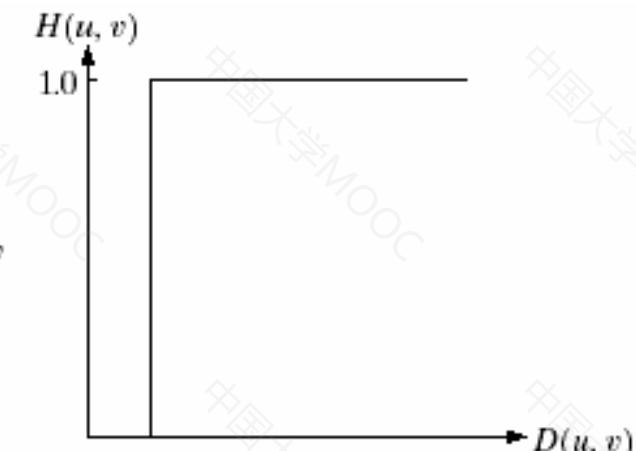
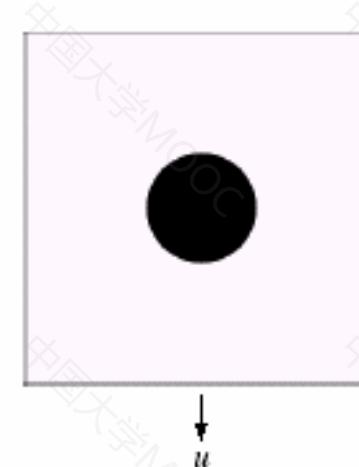
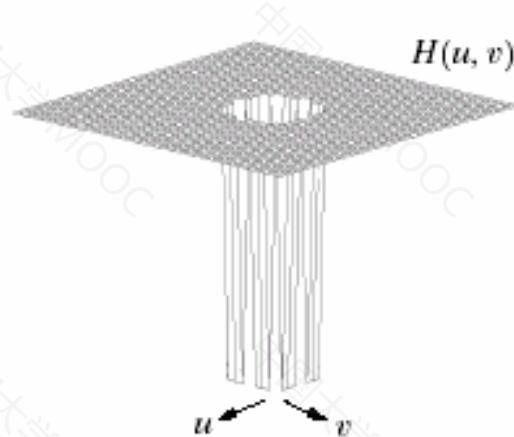
高通濾波器

頻域表达

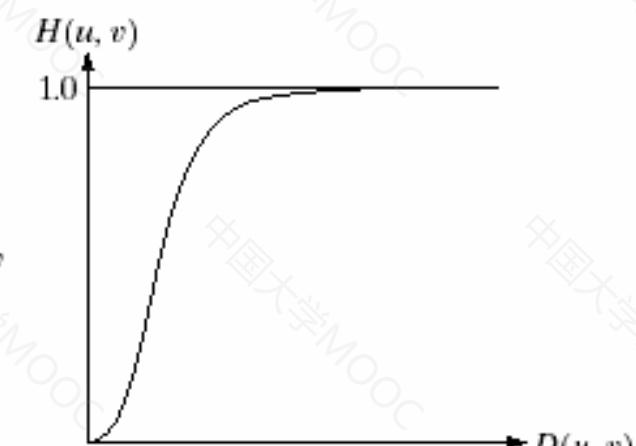
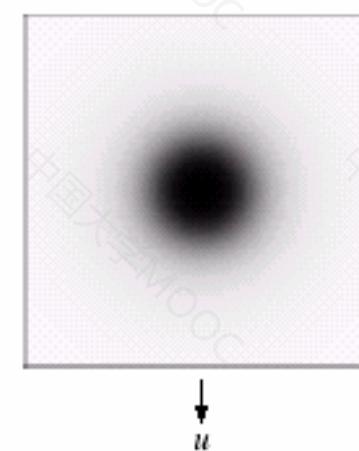
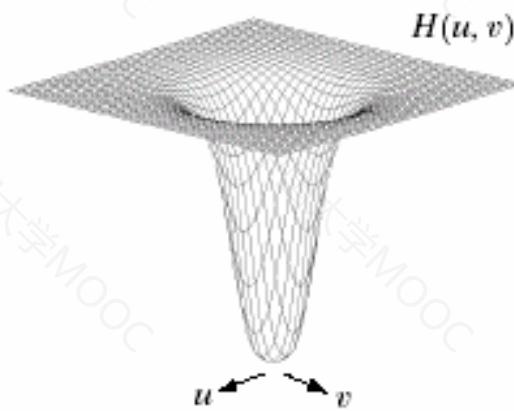
$$H_{hp}(u, v) = 1 - H_{lp}(u, v)$$

高通滤波器的频域形态

理想
高通
滤波
器



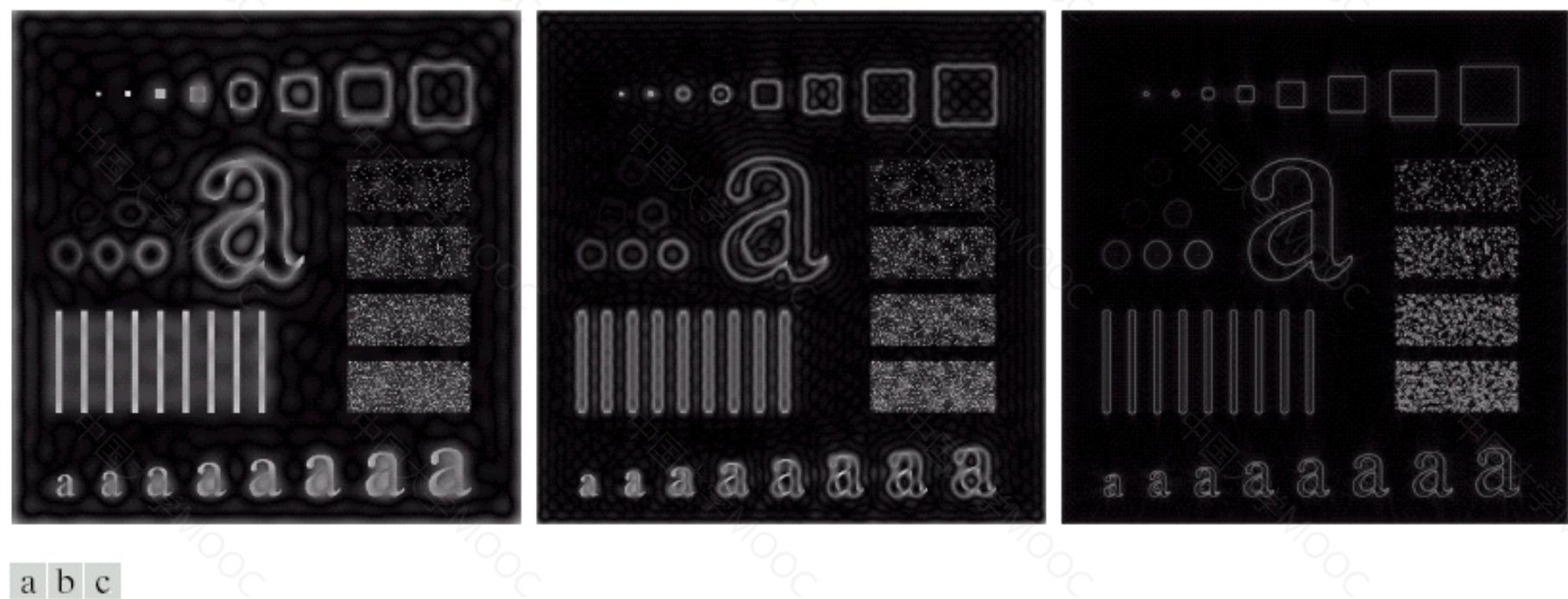
巴特
沃斯
高通
滤波
器



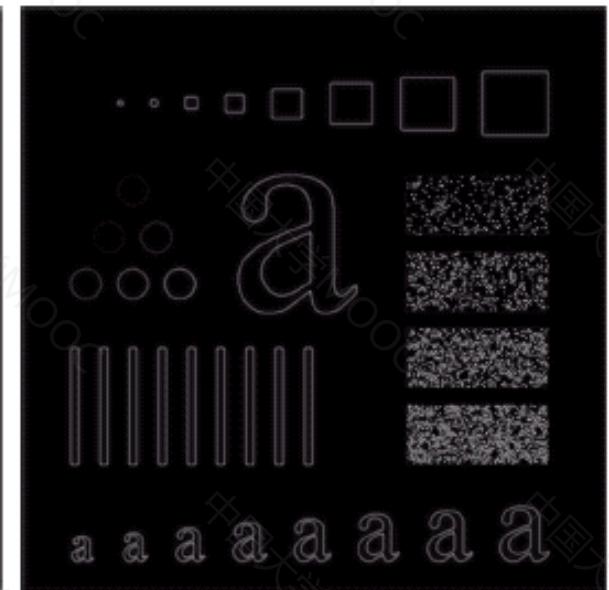
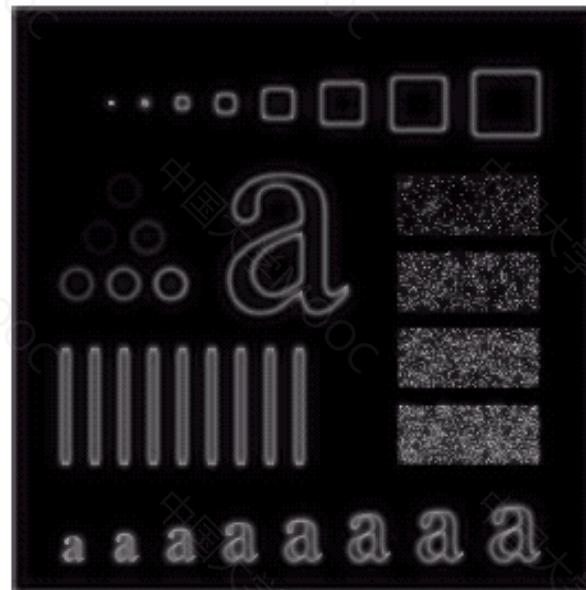
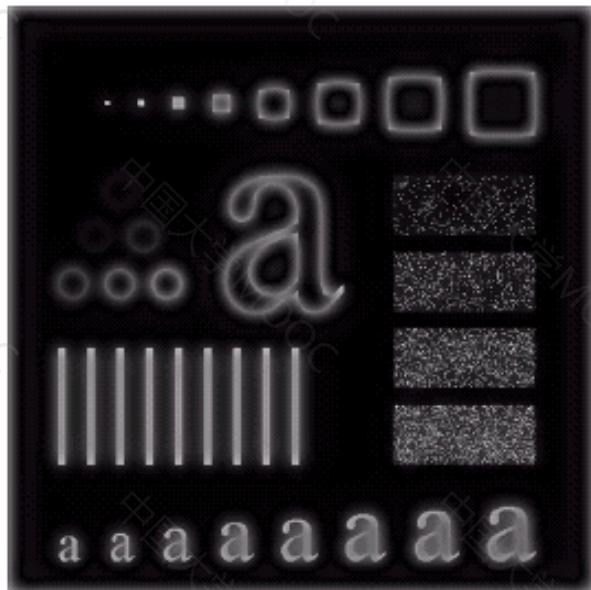
透视图

图像表示

横截面

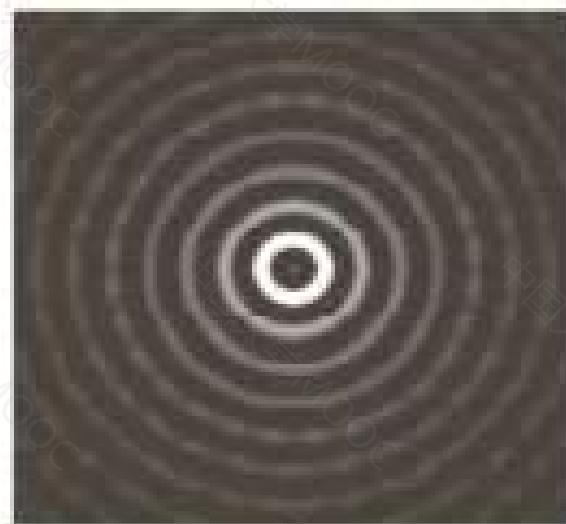


理想高通滤波结果

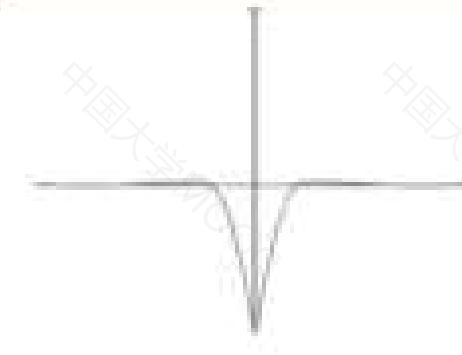


Butterworth高通滤波结果

高通滤波器的空域形态



理想高通滤波器



巴特沃斯高通滤波器

高通滤波器

频域表达

$$H_{hp}(u, v) = 1 - H_{lp}(u, v)$$

空域表达

$$h_{hp}(x, y) = \delta(x, y) - h_{lp}(x, y)$$

张元、汪庆宝、方穗明. 数字信号处理. 北京工业大学出版社. 1995.11

频域通带滤波与空域 掩膜运算的关系？

Laplace算子

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \nabla_x^2 f + \nabla_y^2 f$$

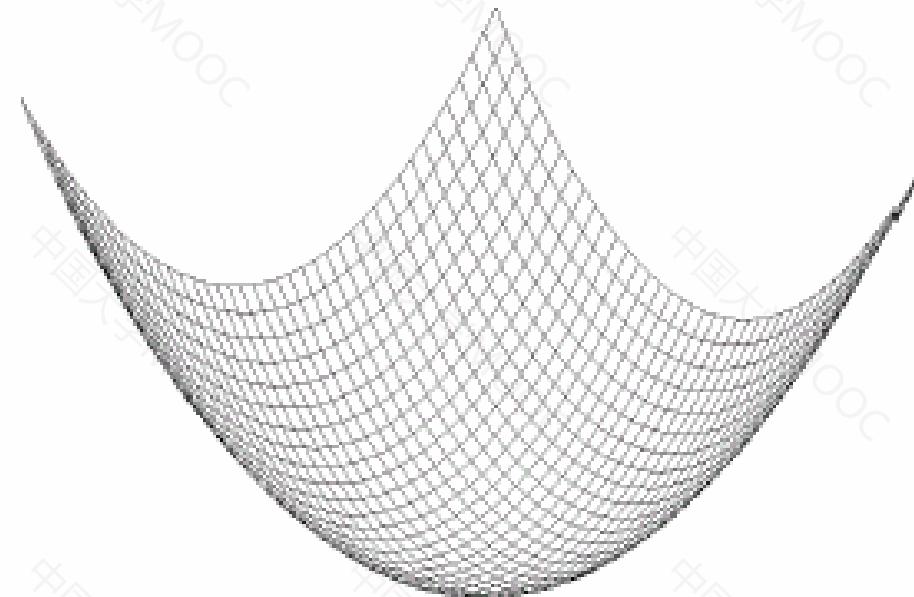
$$g(x, y) = \nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = \nabla_x^2 f + \nabla_y^2 f$$

$$DFT\left(\frac{\partial^n f}{\partial x^n}\right) = (ju)^n F(u, v)$$

$$DFT\left(\frac{\partial^n f}{\partial y^n}\right) = (jv)^n F(u, v)$$

$$\begin{aligned} DFT(g(x, y)) &= (ju)^2 F(u, v) + (jv)^2 F(u, v) \\ &= -(u^2 + v^2) F(u, v) \\ &= H(u, v) F(u, v) \end{aligned}$$

$$|H(u, v)| = u^2 + v^2$$



$H(u, v)$ 等价于一个高通滤波器

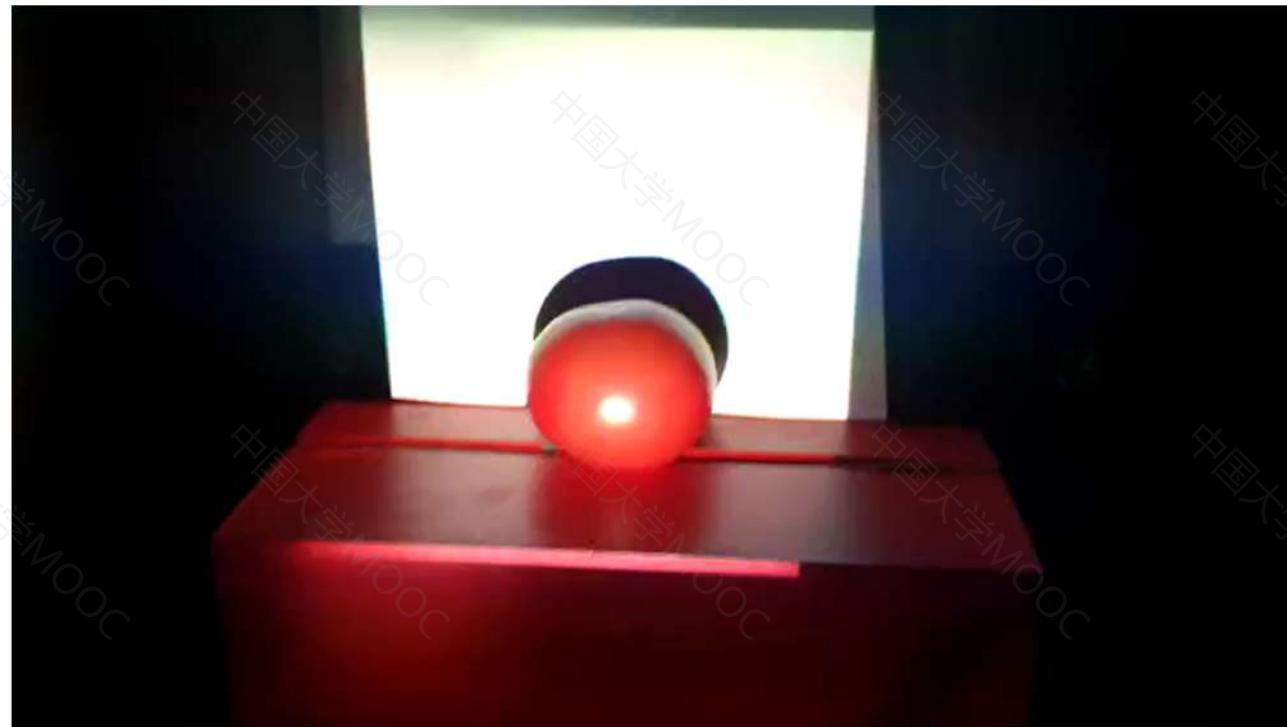
平滑等价于频域低通滤波
锐化等价于频域高通滤波

二、同态滤波

反射成像原理

$$f(x, y) = i(x, y) * r(x, y)$$

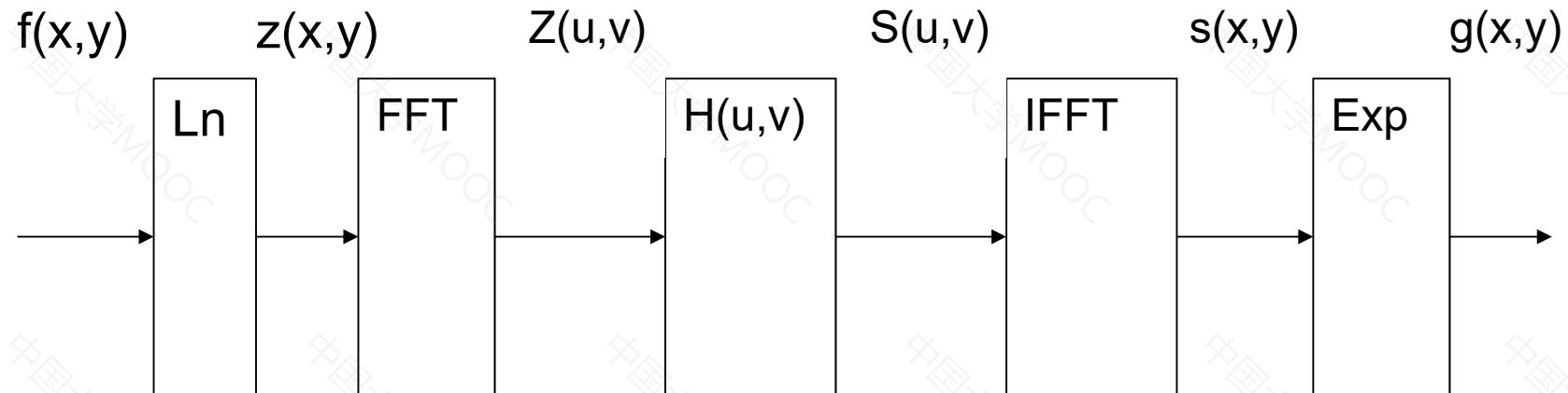
应用背景



**西红柿放在蓝光下，会是什么颜色？
答案是：黑色**

原因：入射光中蓝色份量为1，红色分量为0，而西红柿把蓝色光都吸收了。

■ 同态滤波算法



$$\begin{aligned} z(x,y) &= \ln(f(x,y)) = \ln(i(x,y)) + \ln(r(x,y)) \\ &= i'(x,y) + r'(x,y) \end{aligned}$$

$$Z(u,v) = I'(u,v) + R'(u,v)$$

$$\begin{aligned} S(u,v) &= H(u,v) * Z(u,v) = H(u,v) * I'(u,v) + H(u,v) * R'(u,v) \\ &= H(u,v) * R'(u,v) = R'(u,v) \end{aligned}$$

$$s(x,y) = r'(x,y)$$

$$g(x,y) = \exp(s(x,y)) = \exp(r'(x,y)) = r(x,y)$$

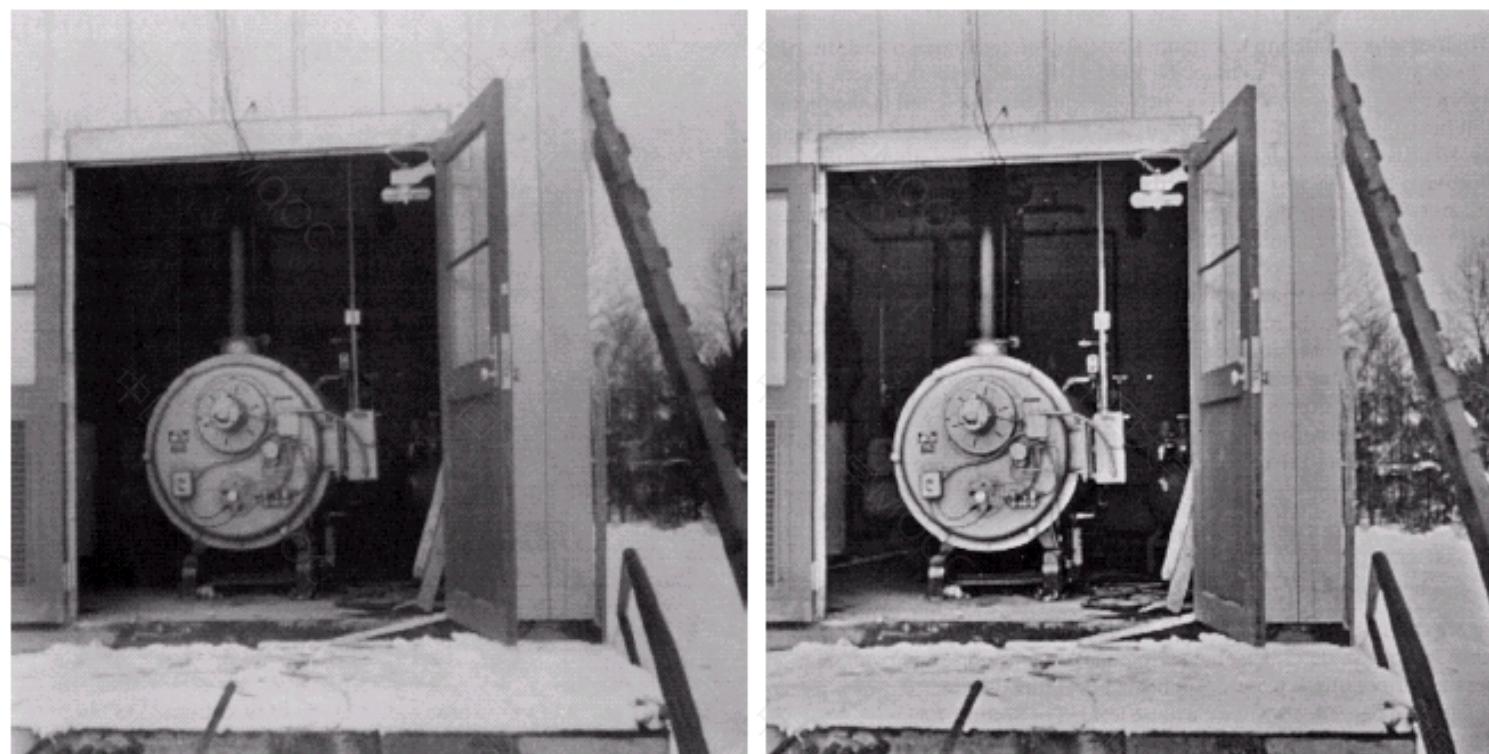
同态滤波应用

- 去除光照变化的影响
- 滤除乘性噪音

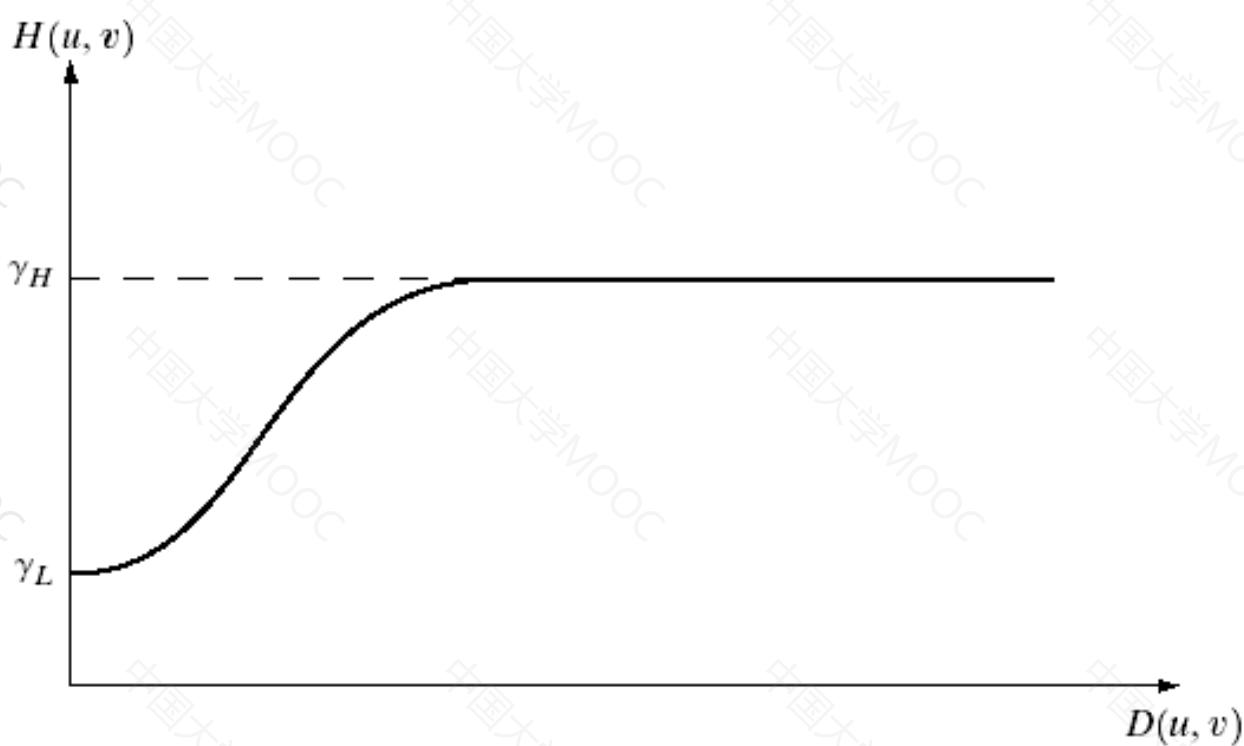
a b

FIGURE 4.33

(a) Original image. (b) Image processed by homomorphic filtering (note details inside shelter). (Stockham.)



■ 压缩动态范围，同时增强对比度



$$\gamma_L < 1$$

$$\gamma_H > 1$$

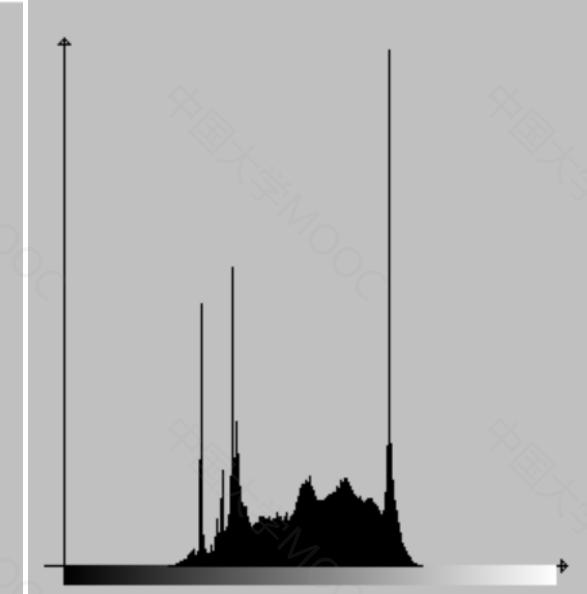
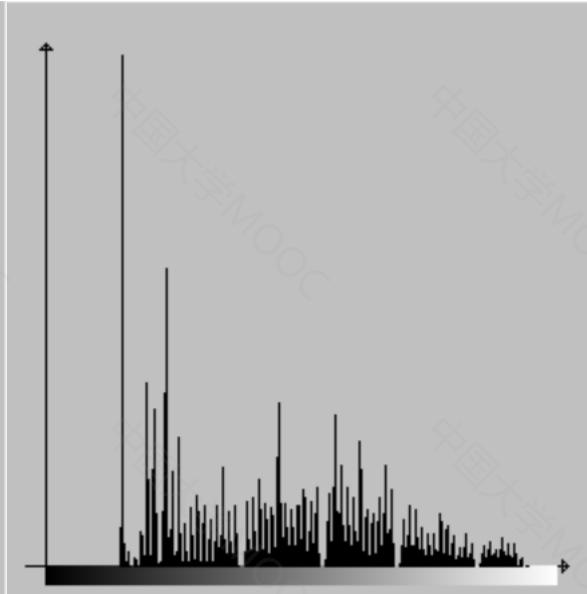
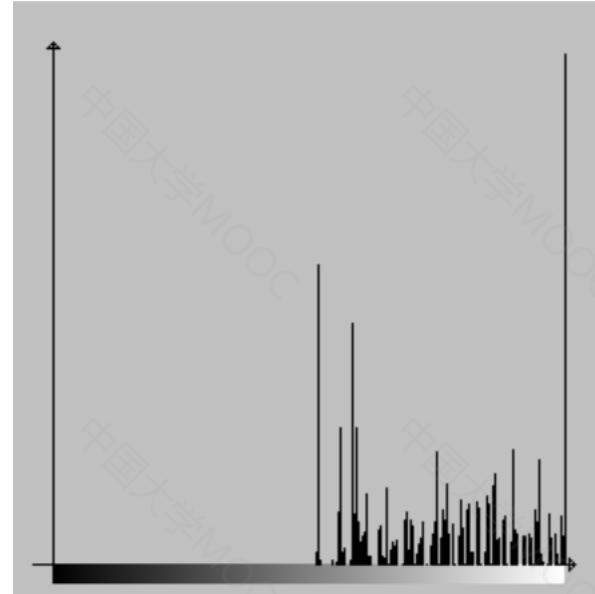
输入图像



同态滤波结果



线性映射结果



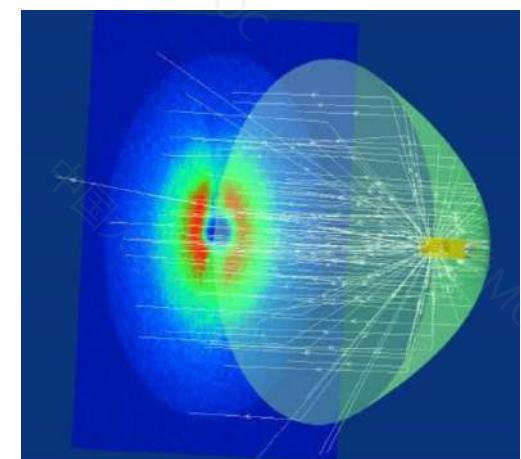
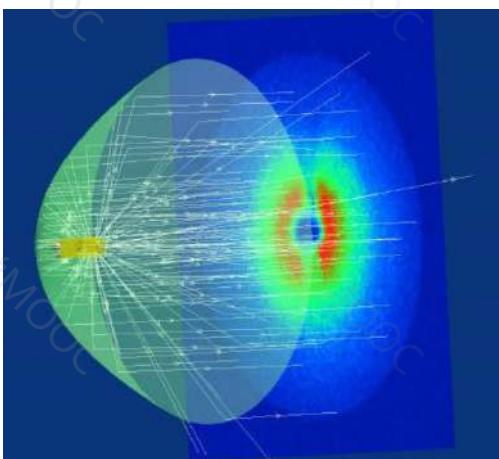
第五章 图像复原 (Image Restoration)

- 图像恢复/复原 (Restoration)：对一幅降质图像，找到降质原因，估计降质模型，通过求逆运算恢复图像的本来面目。
- 一幅模糊图像，处理过程首先弄清楚图像降质的原因，建立相应的数学模型，并沿着使图像质量下降的逆过程来恢复图像。
- 特点
 - ✓ “达到图像质量的整体最优”
 - ✓ 客观方式进行评价
 - ✓ 面向原因

■第五章 图像复原 (Image Restoration)

- 5.1 图像降质原因和降质模型
 - 5.1.1 图像降质原因
 - 5.1.2 图像降质模型
 - 5.1.3 噪音模型估计
- 5.2 图像降质函数估计
- 5.3 降质抑制 (去除)
 - 5.3.1 噪音抑制方法
 - 5.3.2 逆滤波
 - 5.3.3 维纳滤波

5.1 图像降质原因和 降质模型



■ 图像降质原因和降质模型

一、图像降质原因

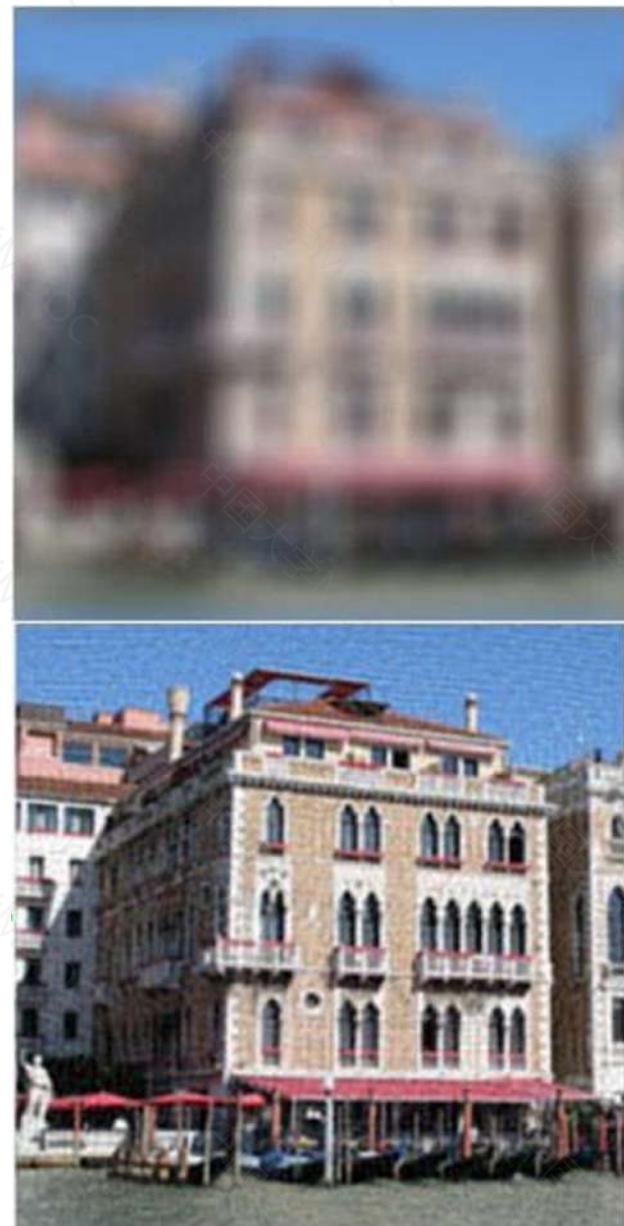
- 运动模糊；
- 散焦模糊；
- 雨、雾等天气影响；
- 云层干扰；
- 系统噪音；
- 光照影响；
- 几何失真；
- 等等

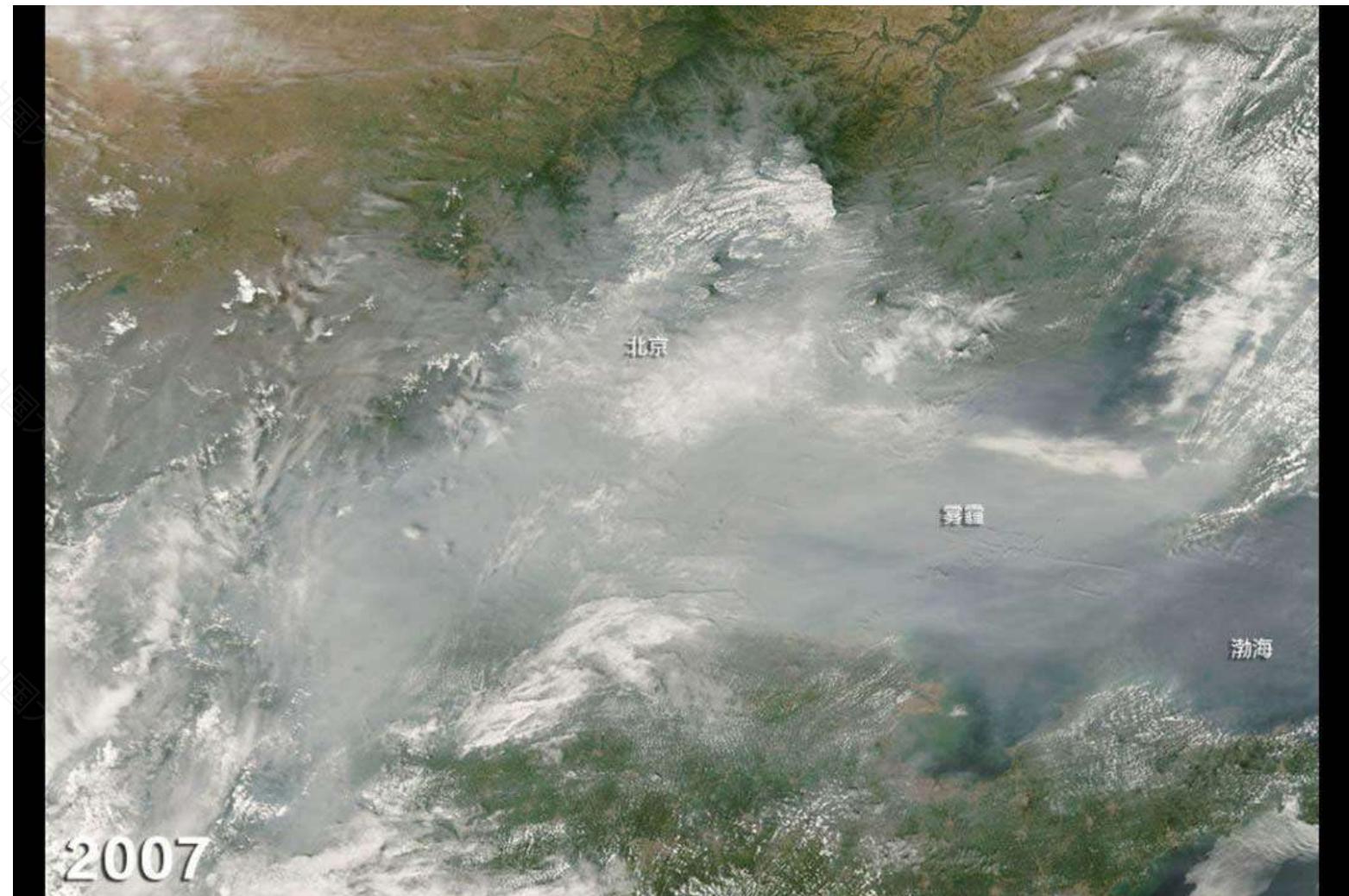
不同降质原因对应不同的降质模型

散 焦 模 糊 图 像



运动模糊图像

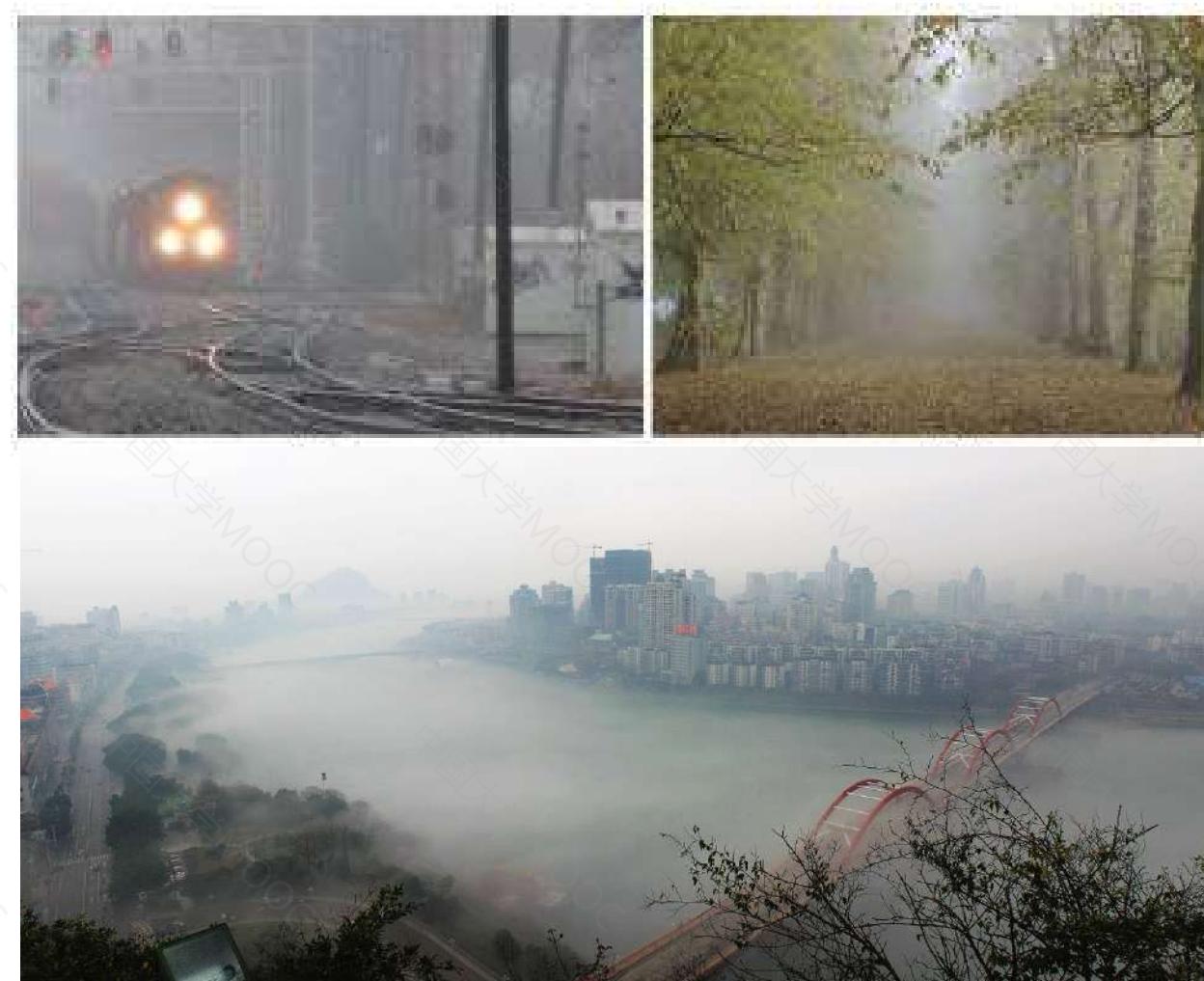




云层遮挡的图像



有雨的图像



有雾的图像

加入高斯噪声的图像



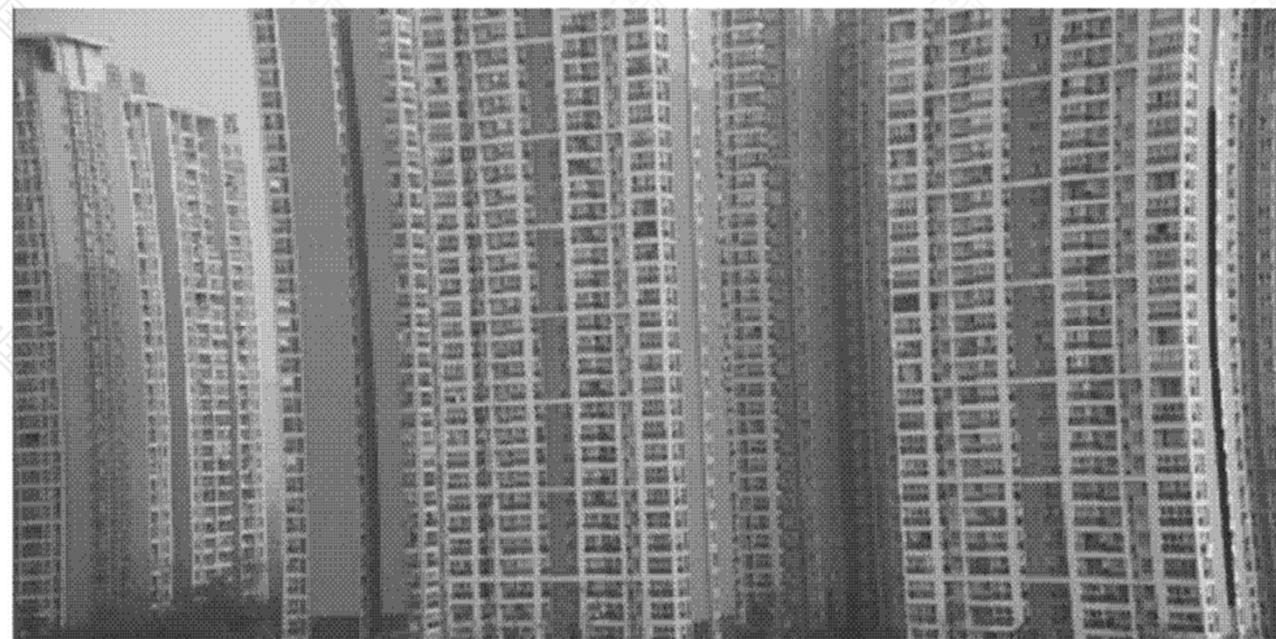
加入椒盐噪声的图像



有噪音的图像



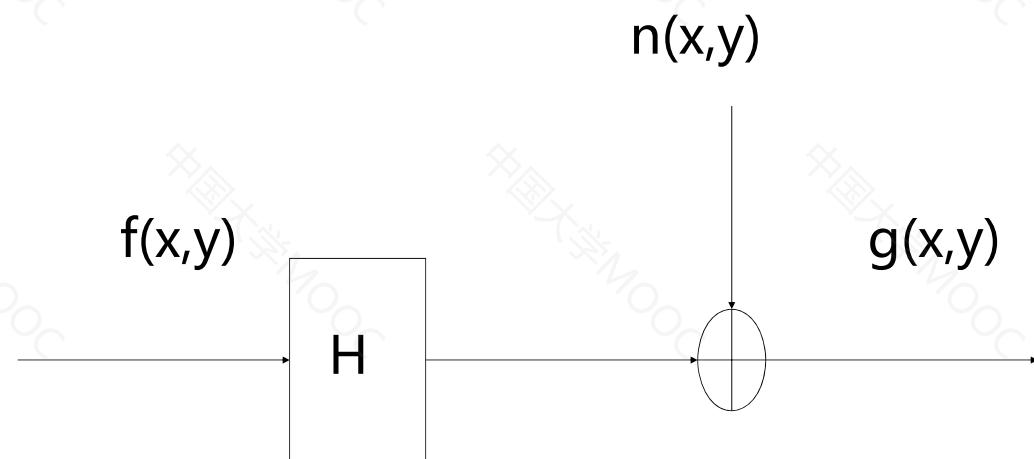
有光照变化的图片



有几何畸变的图像

■ 图像降质原因和降质模型

一、图像降质模型



- 如果系统满足线性移不变特性

$$g(x, y) = f(x, y) * h(x, y) + n(x, y)$$

$$G(u, v) = F(u, v)H(u, v) + N(u, v)$$

■ 噪音模型估计

随机噪音

- ✓ 随机产生的不能预测的噪音；它是一种前后独立的平稳随机过程，在任何时刻它的幅度、波形及相位都是随机的。但每一种噪声还是服从于一定的统计分布规律。

随机噪音分类

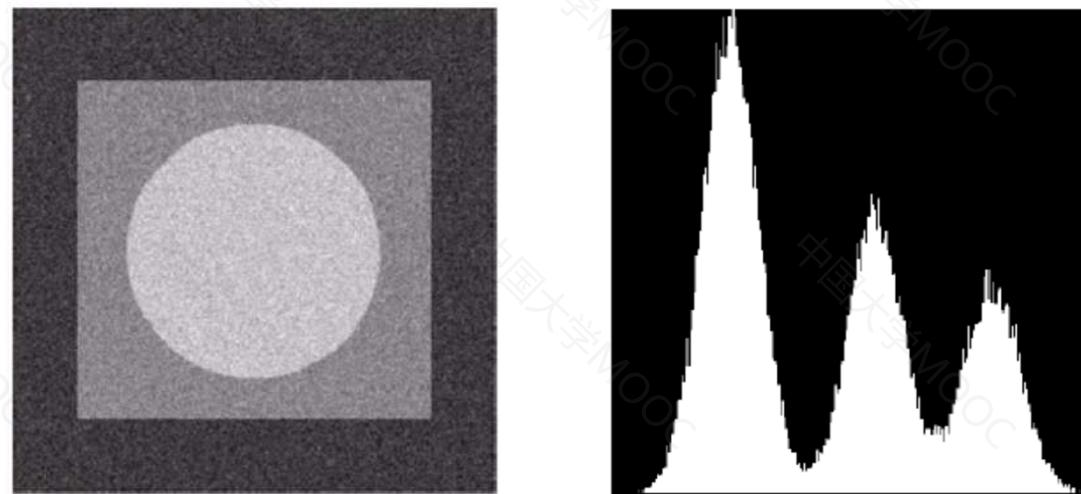
➤ 按功率谱密度

- ✓ 白噪音：一种功率频谱密度为常数的随机信号或随机过程
- ✓ 有色噪音：功率频谱密度不为常数的随机信号或随机过程

■ 噪音模型估计

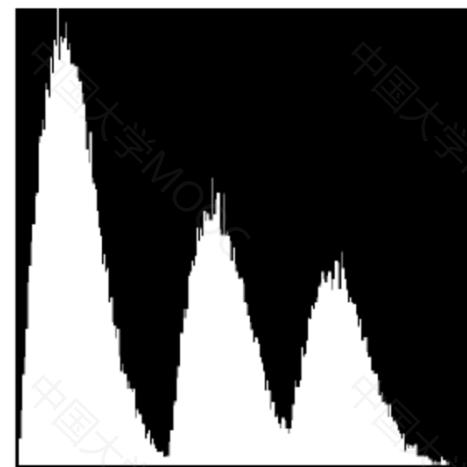
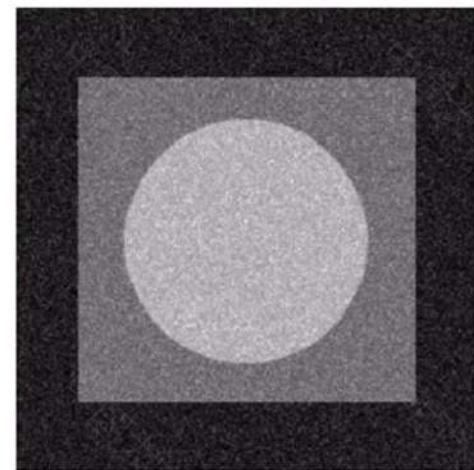
➤ 按概率密度函数分布

- ✓ 高斯噪音
- ✓ 瑞利噪音
- ✓ 伽马噪音
- ✓ 指数噪音
- ✓ 均匀噪音
- ✓ 椒盐噪音
- ✓ 等等



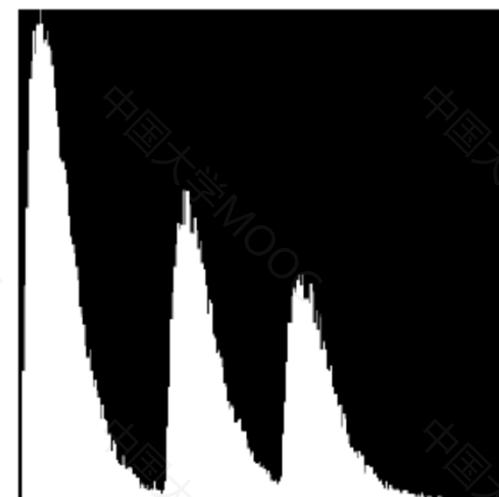
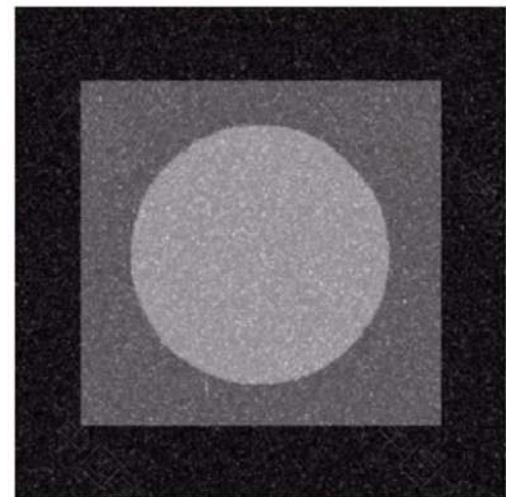
高斯噪音图像

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2}$$



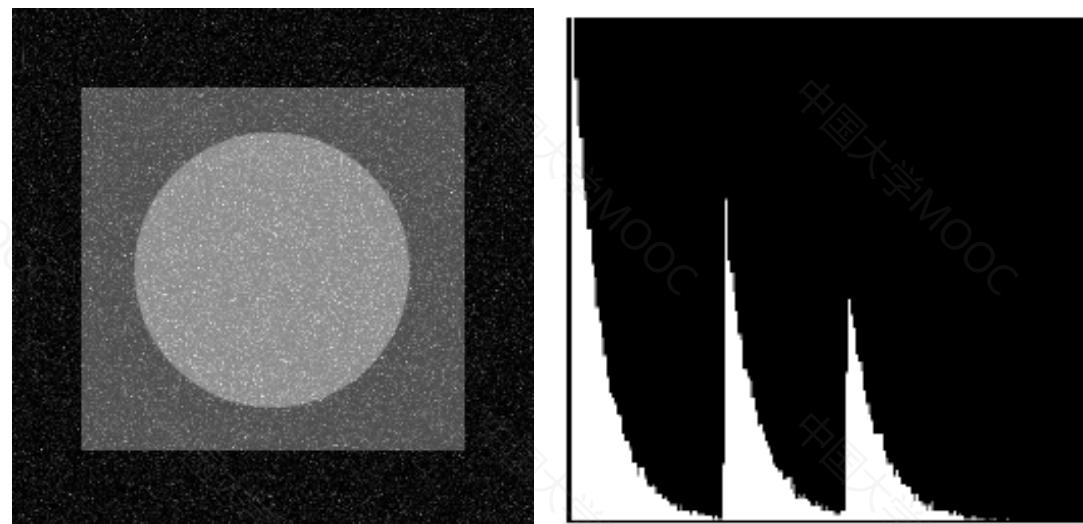
瑞利噪音图像

$$p(z) = \begin{cases} \frac{2}{b}(z-a)e^{-(z-a)^2/b} & z \geq a \\ 0 & z < a \end{cases}$$



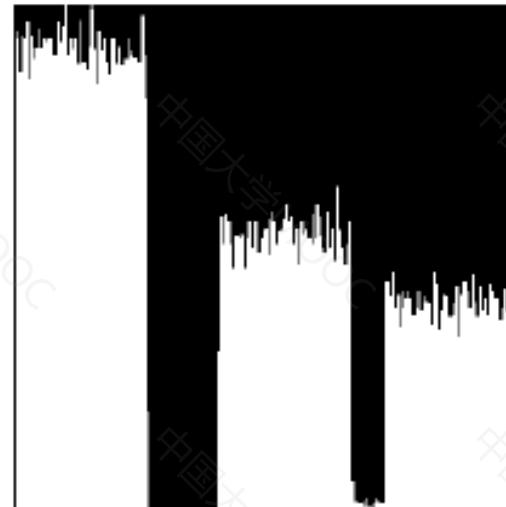
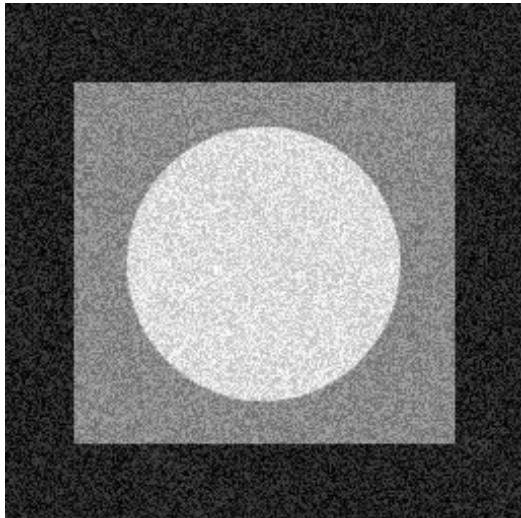
伽马噪音图像

$$p(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az} & z \geq 0 \\ 0 & z < 0 \end{cases}$$



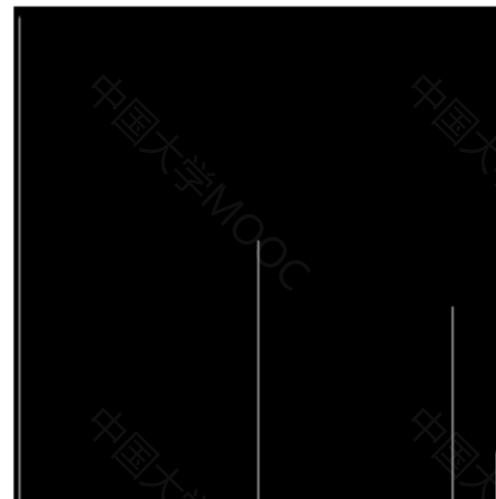
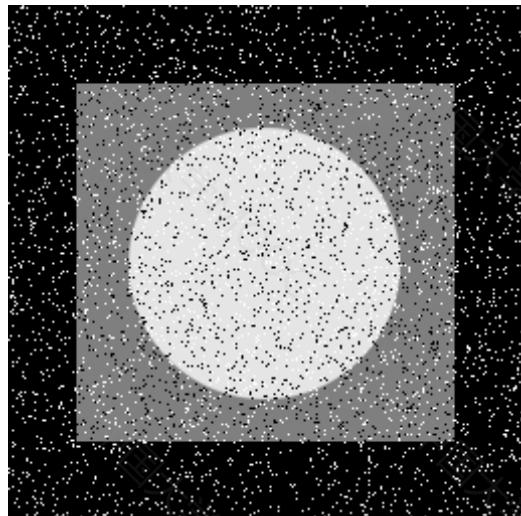
指数噪音图像

$$p(z) = \begin{cases} ae^{-az} & z \geq 0 \\ 0 & z > 0 \end{cases}$$



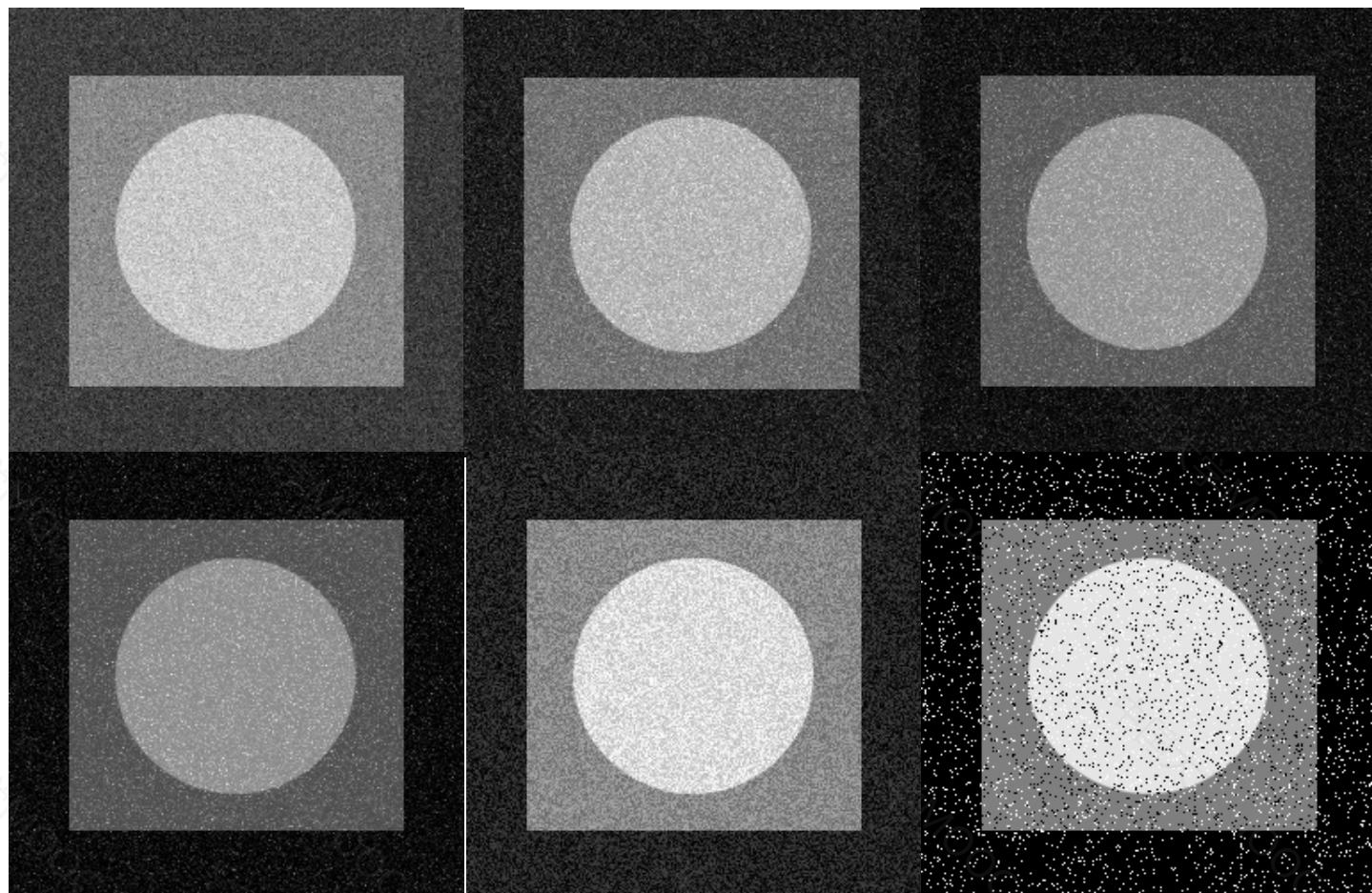
均匀噪音图像

$$p(z) = \begin{cases} \frac{1}{b-a} & a \leq z \leq b \\ 0 & otherwise \end{cases}$$

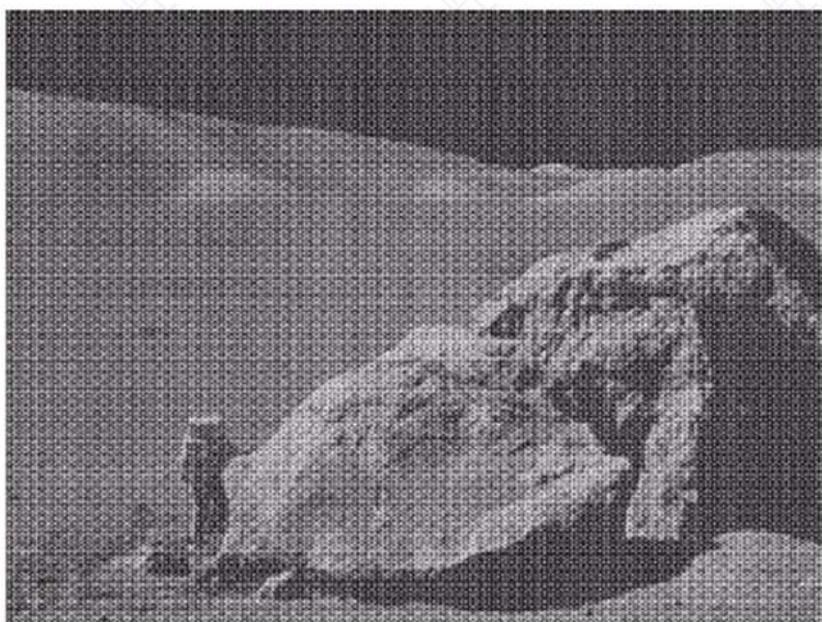


椒盐噪音图像

$$p(z) = \begin{cases} P_a & z = a \\ P_b & z = b \\ 0 & otherwise \end{cases}$$



六种噪音图像对比



有周期噪音的图像



左图的频谱图

■ 噪音模型估计

周期噪音

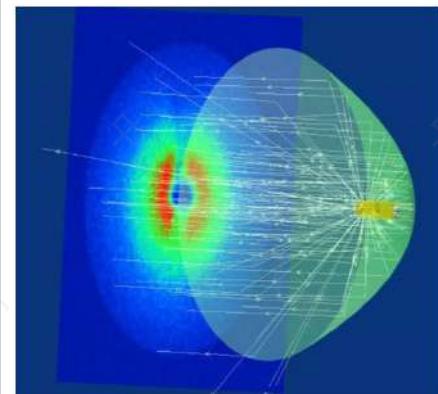
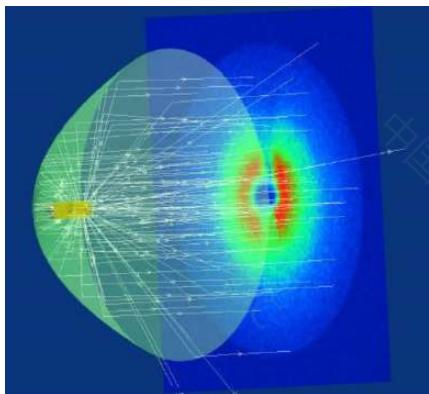
离散傅里叶变换 (DFT)

随机噪音

由成像环境推导

实验获得

5.2 图像降质函数估计

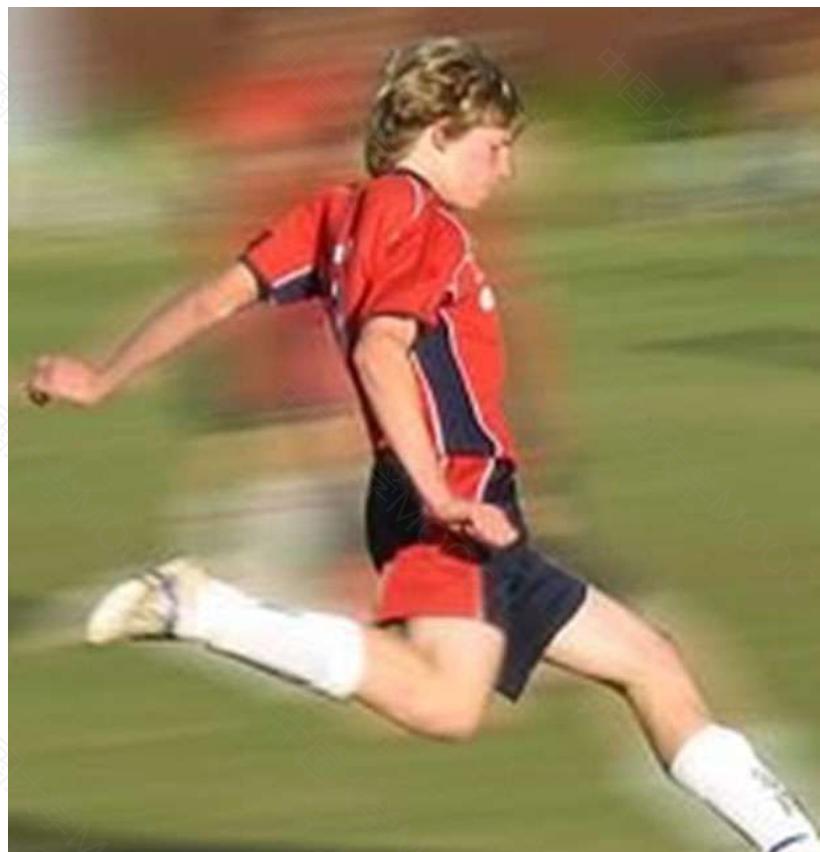


图像降质函数估计方法

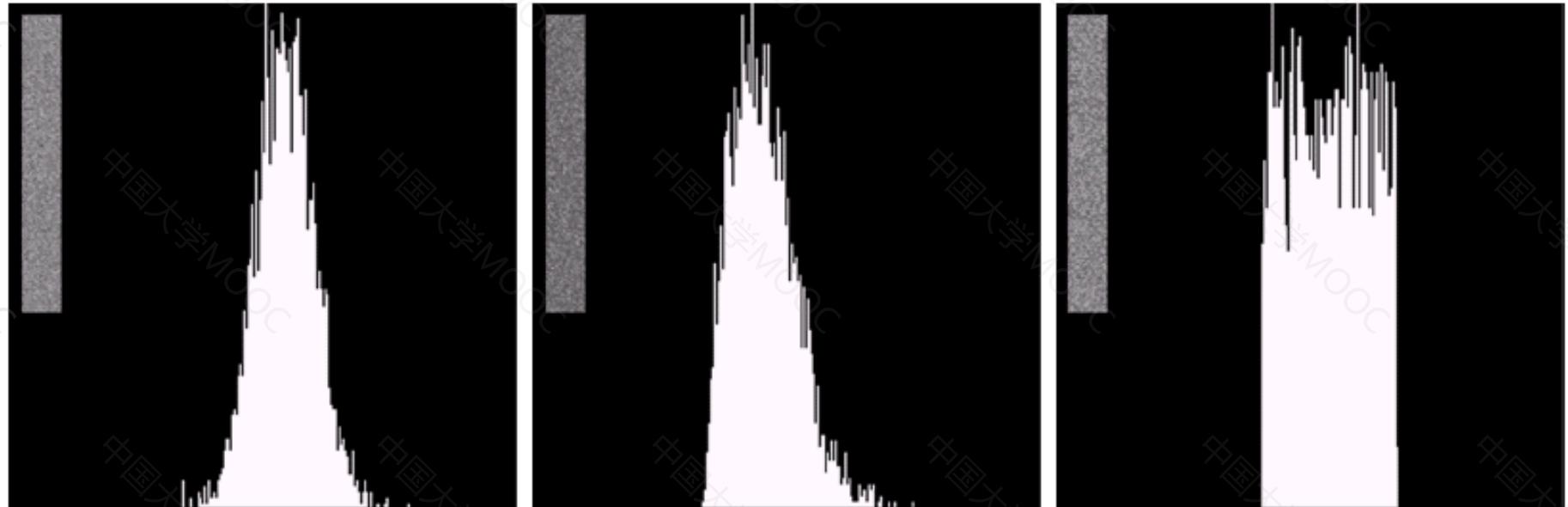
- 1) 观察法
- 2) 重复实验法
- 3) 理论建模法

一、观察法

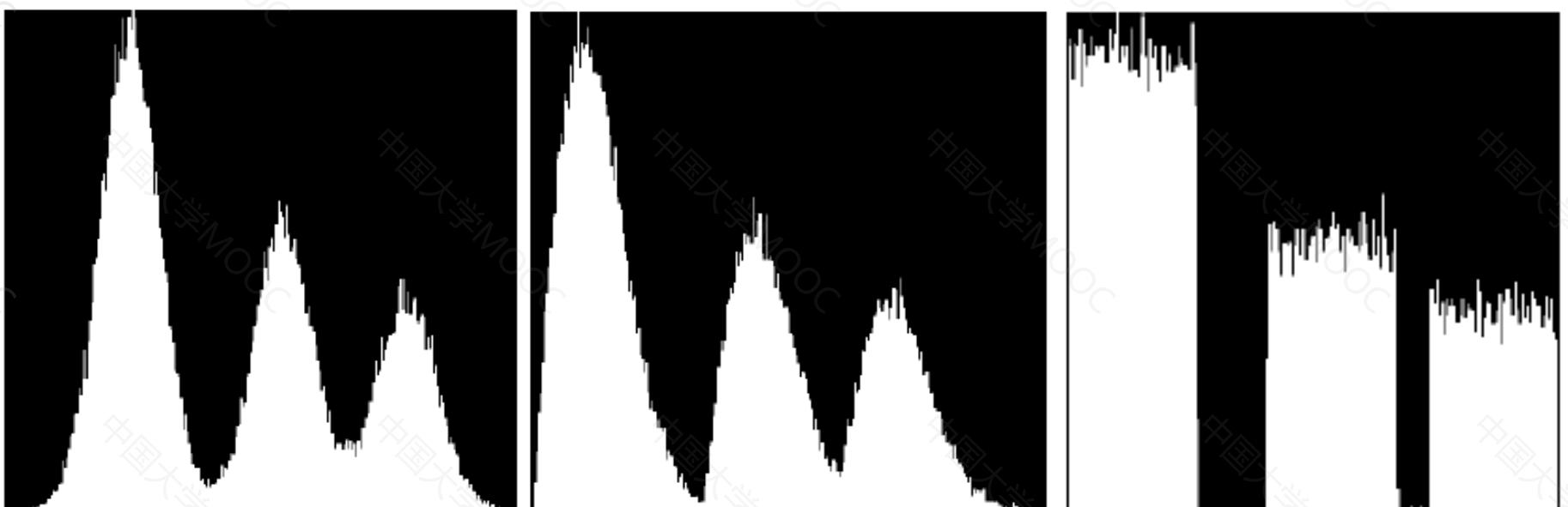
- 从包含简单结构的图像中收集信息



$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}$$



局部图像区域及噪音分布



4.1节中对应的原始图像的噪音分布

- 利用样本灰度，我们可以构造出与观察子图像大小和特征相同的无模糊图像
- 观测到的子图 $g_s(x, y)$,
- 构造的子图 $f_s(x, y)$
- 退化函数 $h_s(x, y)$

$$H_s(u, v) = \frac{G_s(u, v)}{\hat{F}_s(u, v)}$$

- 根据移不变特性

$$H(u, v) = H_s(u, v)$$

- 2. 重复实验法

- 如果设备可用 (重复降质过程, 获取降质图像)
- 降质系统的点扩展函数 (PSF) 通过使用相同系统成像对点光源来获得;
- 如果该系统是一个线性移不变系统, 则可以得到其脉冲响应。脉冲的傅里叶变换是常数A
- 可以得到脉冲响应(点扩展函数)
$$H(u, v) = \frac{G(u, v)}{A}$$

3. 理论建模法

- 结合降质原因，对降质过程进行理论分析，估计降质模型；
- 理论建模方法考虑导致降质的环境条件，如果能够考虑的因素全面，还是比较有效的。
- 举例：Hufnagel和Stanley提出的云层干扰图像的降质模型是基于大气湍流的物理特性分析建立的。

这种模型有一种常见的形式：

$$H(u, v) = e^{-k(u^2 + v^2)^{5/6}}$$

- 其中k是常数，除了指数的5/6次方外，这个方程的形式与高斯函数的形式相同

a
b
c
d

FIGURE 5.25

Illustration of the atmospheric turbulence model.

(a) Negligible turbulence.
(b) Severe turbulence,
 $k = 0.0025$.

(c) Mild turbulence,
 $k = 0.001$.

(d) Low turbulence,
 $k = 0.00025$.

(Original image courtesy of NASA.)



- 匀速直线运动模糊图像降质函数估计
 - 什么是运动模糊?
 - 如果景物与相机之间存在相对运动，则造成图像模糊，称为运动模糊。
 - 相机曝光时间T
 - $x_0(t)$ 和 $y_0(t)$ 分别是x和y方向上运动的时变分量
- $$g(x, y) = \int_0^T f(x - x_0(t), y - y_0(t)) dt$$
- 对应成像像素点的总曝光量是曝光时间内不同时刻的瞬时曝光量的积分。

- 匀速直线运动模糊图像降质函数估计
- 什么是运动模糊？

$$g(x, y) = \int_0^T f(x - x_0(t), y - y_0(t)) dt$$

– 如果曝光时间内相机和景物之间不存在相对运

动，则 $x_0(t)=0, y_0(t)=0$

$$g(x, y) = \int f(x, y) dt$$

$$= f(x, y) \int dt$$

$$= T f(x, y)$$

- 匀速直线运动模糊图像降质函数估计
- 什么是运动模糊?
 - 如果相机与景物之间的相对运动是匀速直线运动，则产生的模糊图像是匀速直线运动模糊图像

$$g(x, y) = \int_0^T f(x - x_0(t), y - y_0(t)) dt$$

$$\begin{aligned}x_0(t) &= at / T \\y_0(t) &= bt / T\end{aligned}\quad \int_0^T f\left(x - \frac{at}{T}, y - \frac{bt}{T}\right) dt$$



相机和景物存在**45度**方向相对运动



相机存在垂直方向运动



相机存在水平方向运动

$$G(u, v)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) e^{-j2\pi(ux+vy)} dx dy$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[\int_0^T f(x - x_0(t), y - y_0(t)) dt \right] e^{-j2\pi(ux+vy)} dx dy$$

$$= \int_0^T \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x - x_0(t), y - y_0(t)) e^{-j2\pi(ux+vy)} dx dy \right] dt$$

$$\begin{aligned} & x' = x - x_0(t) \\ & y' = y - y_0(t) \end{aligned}$$
$$= \int_0^T \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x', y') e^{-j2\pi(ux'+vy')} dx' dy' \right] e^{-j2\pi(ux_0(t)+vy_0(t))} dt$$

$$= F(u, v) \int_0^T e^{-j2\pi(ux_0(t)+vy_0(t))} dt$$

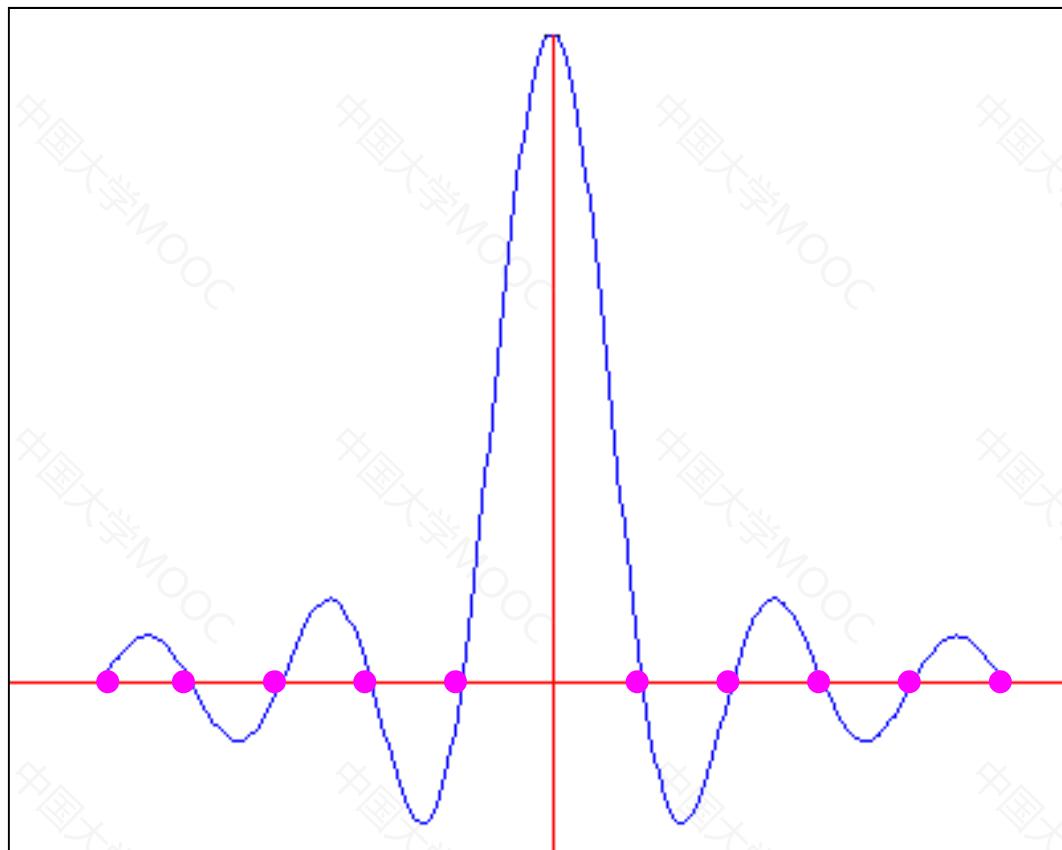
F (u,v)

$$= F(u, v) H(u, v)$$

$$\begin{aligned}
H(u, v) &= \int_0^T e^{-j2\pi(ux_0(t)+vy_0(t))} dt \\
&\stackrel{x_0(t)=at/T}{=} \int_0^T e^{-j2\pi(ua+vb)t/T} dt \\
&\stackrel{y_0(t)=bt/T}{=} \int_0^T e^{-j2\pi(ua+vb)t/T} dt \\
&\stackrel{t'=-j2\pi(ua+vb)t/T}{=} \int_0^{-j2\pi ua} e^{t'} \left(-\frac{T}{2j\pi(ua+vb)} dt' \right) \\
&= \frac{T}{-2j\pi(ua+vb)} \int_0^{-j2\pi(ua+vb)} e^{t'} dt'
\end{aligned}$$

$$\begin{aligned} H(u, v) &= \frac{T}{-2j\pi(ua+vb)} e^{jt'} \Big|_0^{-j2\pi(ua+vb)} \\ &= \frac{T(e^{-j2\pi(ua+vb)} - 1)}{-2j\pi(ua+vb)} \\ &= \frac{T}{\pi(ua+vb)} \frac{e^{-j\pi(ua+vb)} - e^{j\pi(ua+vb)}}{-2j} e^{-j\pi(ua+vb)} \\ &= T \frac{\sin(\pi(ua+vb))}{\pi(ua+vb)} e^{-j2\pi(ua+vb)} \end{aligned}$$

$$|H(u, v)| = T \left| \frac{\sin(\pi(ua + vb))}{\pi(ua + vb)} \right|$$



当 $ua+vb=0$ 时， $H(u, v)=T$

当 $ua+vb=\pm 1$ 时， $H(u, v)=0$

当 $ua+vb=\pm 2$ 时， $H(u, v)=0$

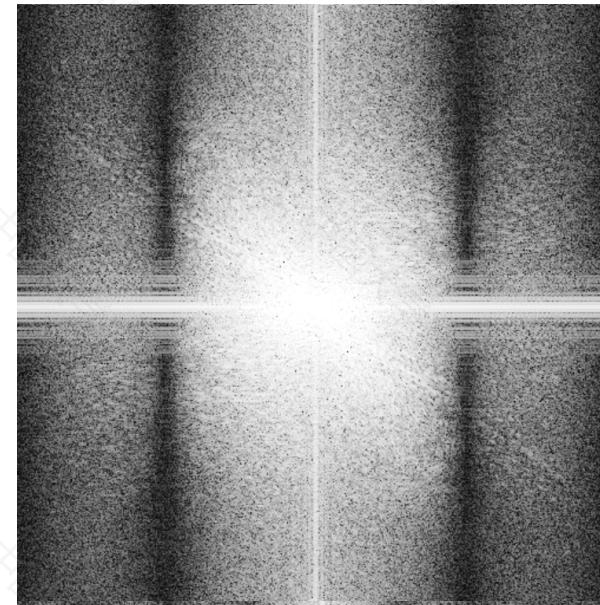
当 $ua+vb=\pm k$ （ k 为整数）时， $H(u, v)=0$

- 水平方向运动模糊

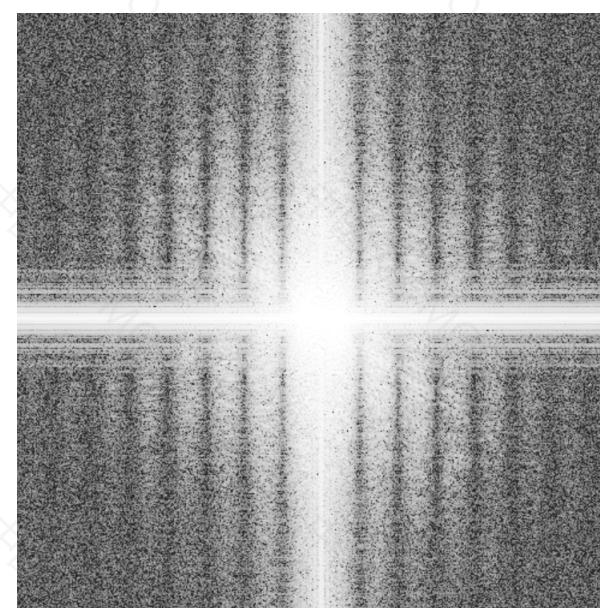
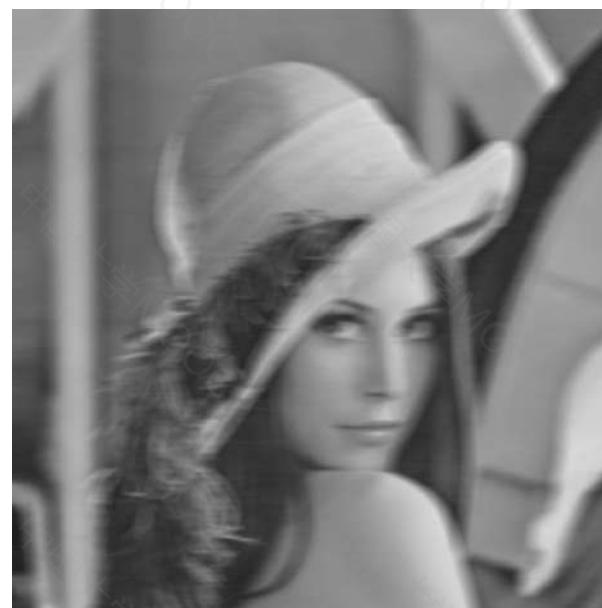
$$x_0(t) = at / T$$

$$y_0(t) = 0$$

$$\begin{aligned} H(u, v) &= T \frac{\sin(\pi(ua + vb))}{\pi(ua + vb)} e^{-j2\pi(ua+vb)} \\ &= T \frac{\sin \pi ua}{\pi ua} e^{-j2\pi ua} \end{aligned}$$

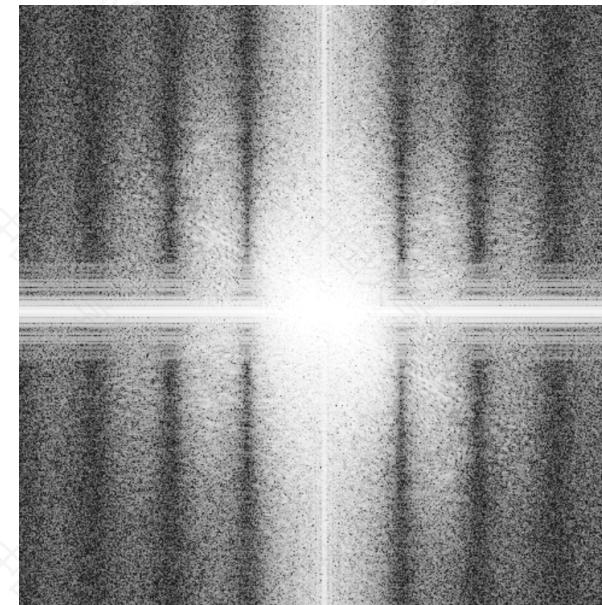
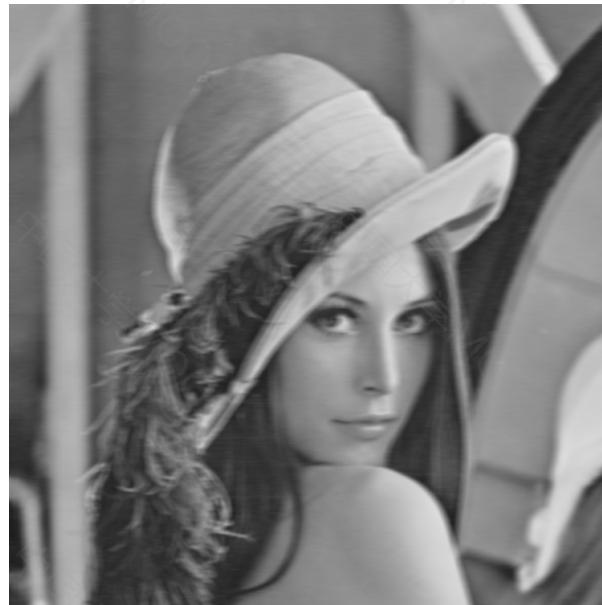


水平方向运动幅度为4的图像和幅度谱

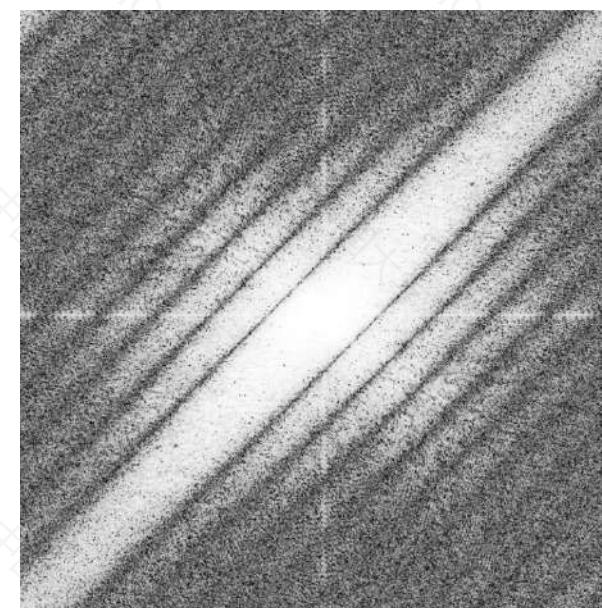


水平方向运动幅度为16的图像和幅度谱

匀速直线运动模糊图像频谱上有平行暗条；
暗条方向和运动方向垂直；
暗条间隔和运动幅度呈反比；



水平方向运动幅度为8的图像和幅度谱



45度方向运动幅度为16的图像和幅度谱

- 匀速直线运动模糊图像的频谱中出现明显的相互平行的暗条，暗条垂直于运动方向。举例来讲，若存在水平方向运动，则暗条为垂直竖条。运动幅度越大，暗条间隔越小
- 可根据频谱中出现的平行暗条方向及间距来判断运动的方向和幅度。



相机和景物之间存在非匀速非直线运动时，利用重力加速度传感器获取的相机运动轨迹示意图

5.3 降质抑制（去除）

- 5.3.1 噪音抑制
- 5.3.2 逆滤波
- 5.3.3 维纳滤波

5.3 降质抑制（去除）

➤ 噪音抑制

- 随机噪音
 - 椒盐噪音抑制
 - 中值滤波
 - 加权中值滤波：模板不同位置对影像素点的排列次数不一致；
 - 自适应中值滤波
 - 高斯噪音抑制
 - 总变差
 - 稀疏编码等

3. 加权中值滤波

- 中值滤波时，各像素点赋以不同的权值，决定其排序次数。
- 例如：窗口

1 1 1

1 1 1

1 1 1

中值滤波模板

1 2 1

2 5 2

1 2 1

加权中值滤波模板1

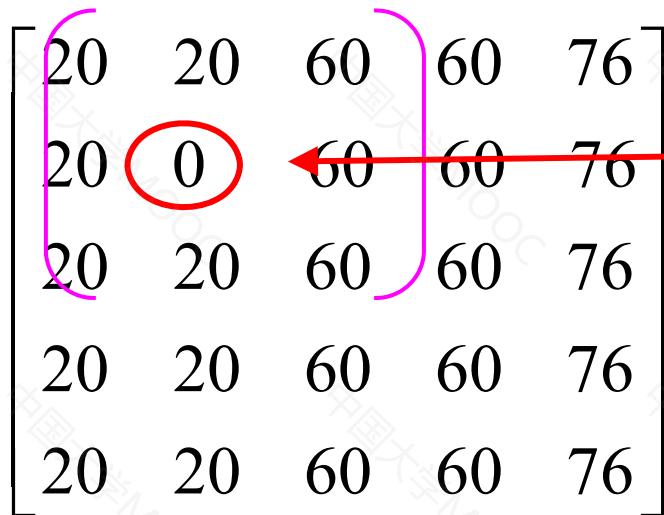
1 1 3

1 3 3

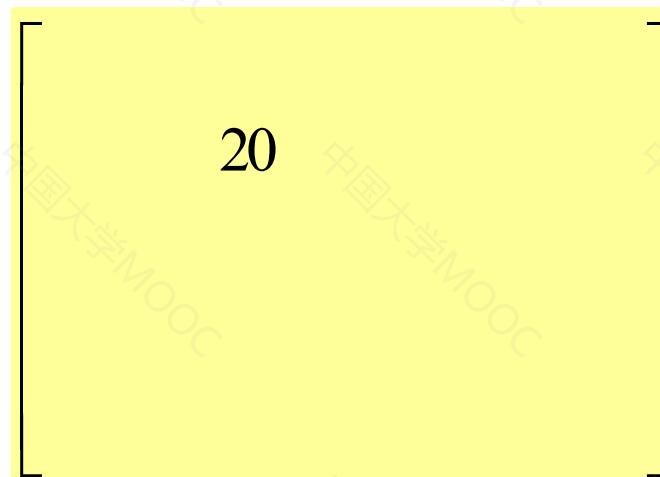
1 1 3

加权中值滤波模板2

中值滤波



$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



$$nk\{20, 20, 60, 20, 0, 60, 20, 20, 60, 50, 60, 60, 20, 20, 20, 20, 0\}$$

*

20	20	60	60	76
20	0	60	60	76
20	20	60	60	76
20	20	60	60	76
20	20	60	60	76

60 → 1

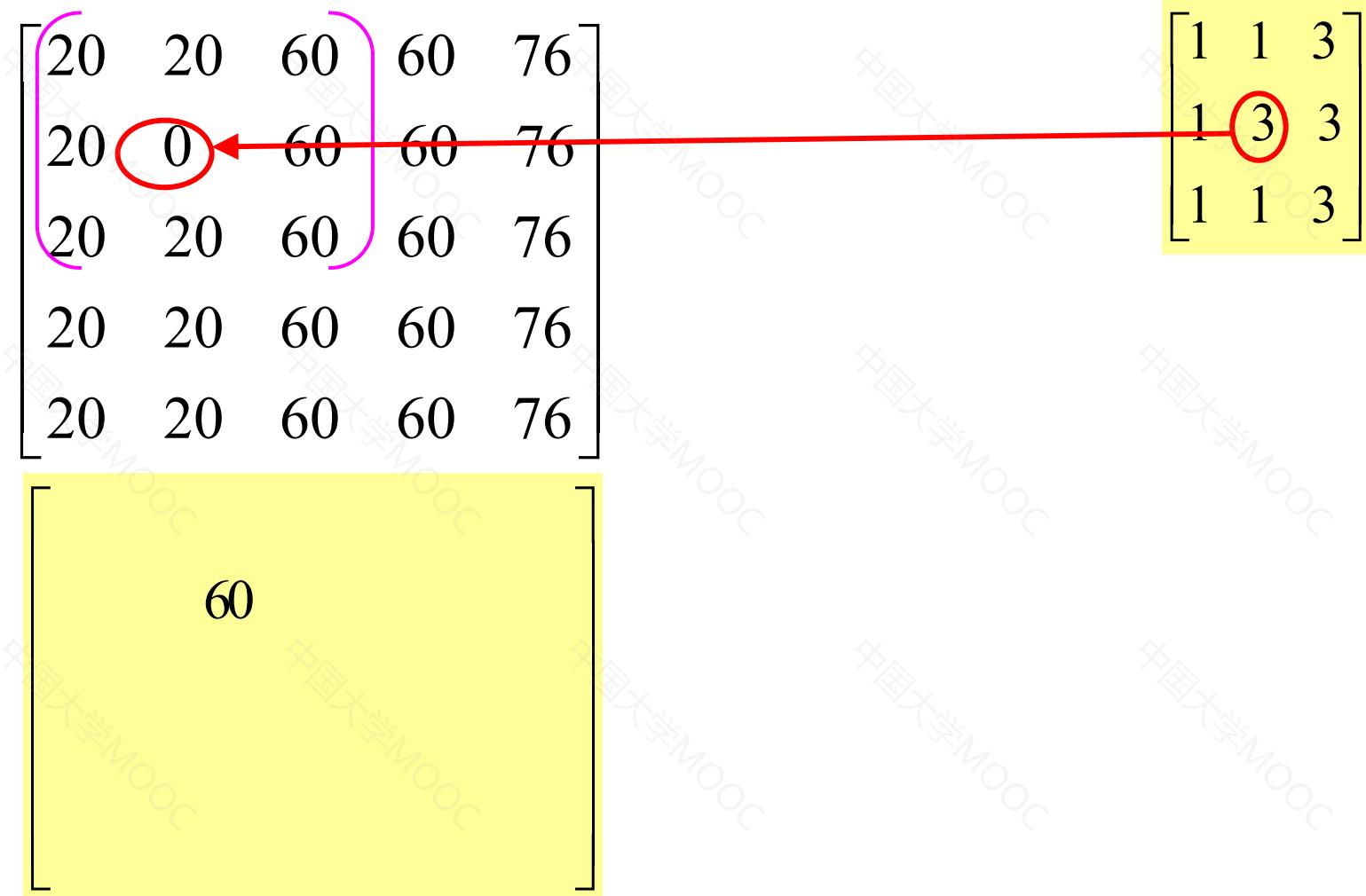
1	1	1
1	1	1
1	1	1

20	60
----	----

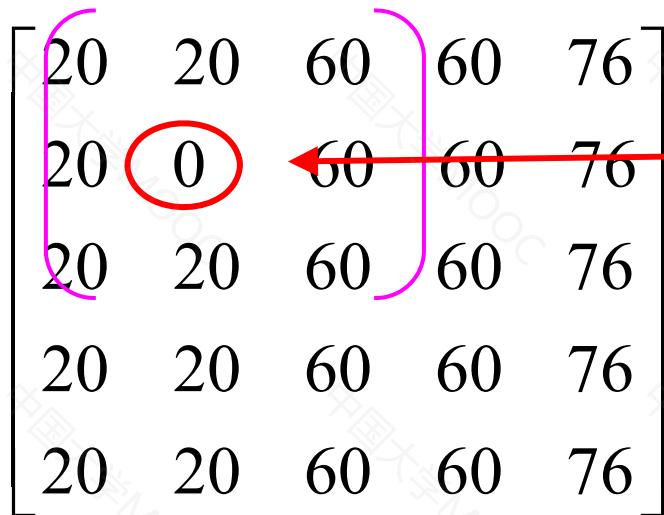
$$\begin{aligned}
 & rank\{20, 60, 60, 0, 60, 60, 20, 60, 60\} \\
 & = \{60, 60, 60, 60, 60, 60, 20, 20, 20\}
 \end{aligned}$$

*

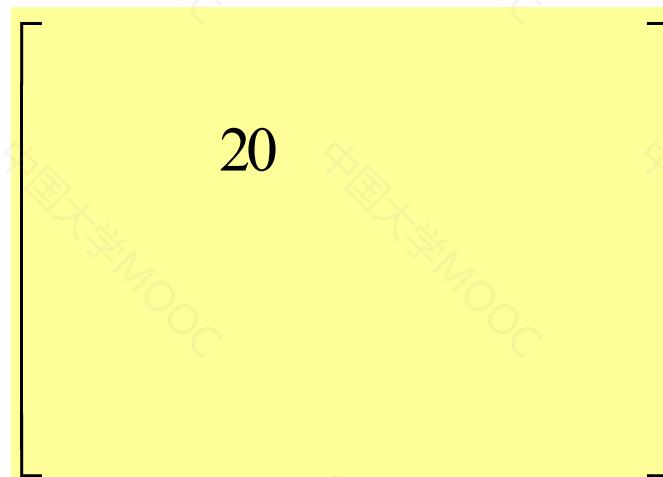
加权中值濾波



中值滤波

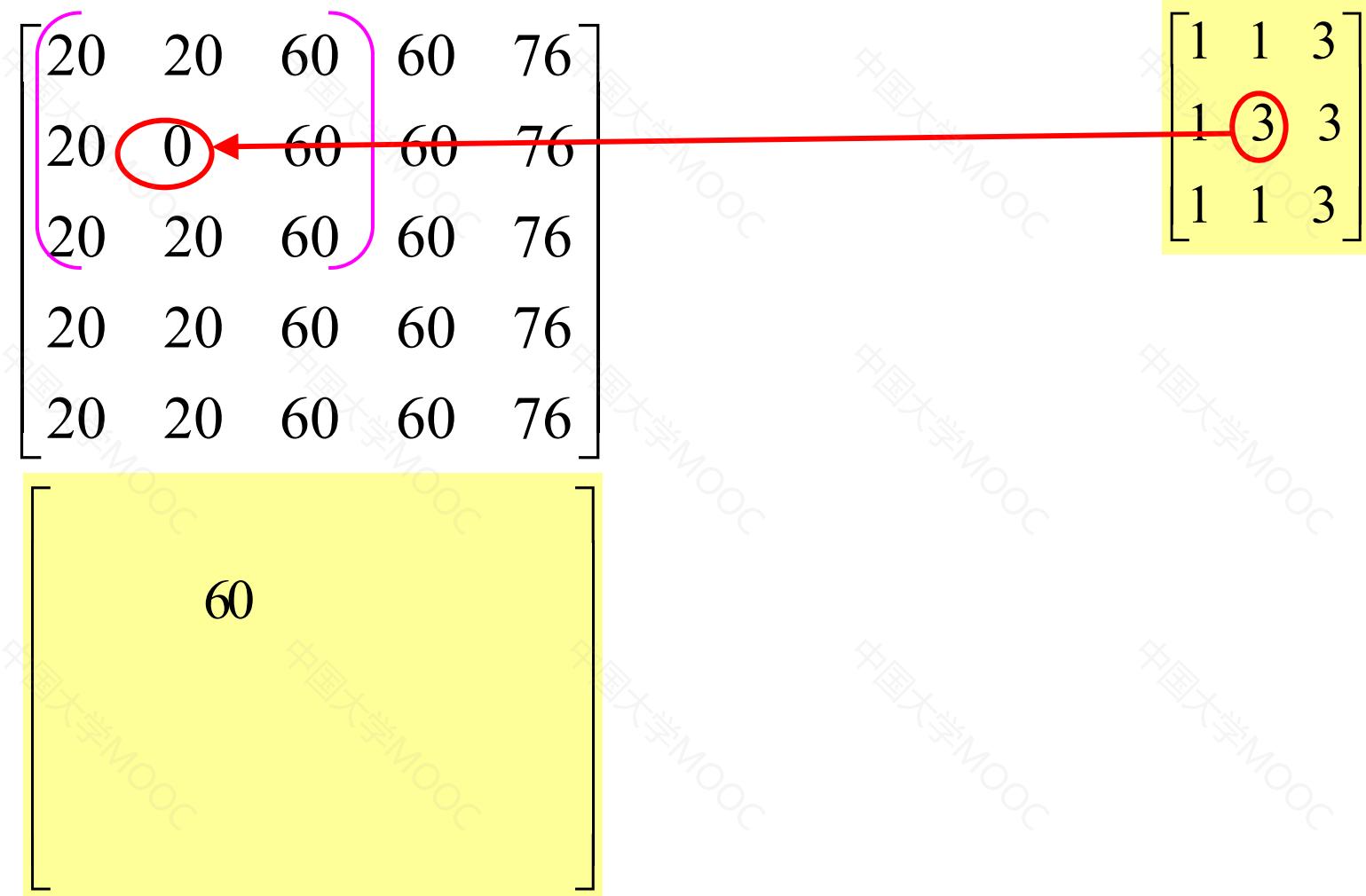


$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & \textcircled{1} & 1 \\ 1 & 1 & 1 \end{bmatrix}$$



$nk\{20, 20, 60, 20, 0, 60, 20, 20, 60, 50, 60, 60, 20, \textcircled{20}, 20, 20, 20, 0\}$

加权中值濾波



20	20	60	60	76
20	0	60	60	76
20	20	60	60	76
20	20	60	60	76
20	20	60	60	76

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

20

$$Rank\{20, 20, 60, 20, 0, 0, 0, 0, 60, 20, 20, 60\} \\ = \{60, 60, 60, 20, 20, 20, 20, 0, 0, 0\}$$

*

4. 自适应中值滤波

- 设中值滤波窗口区域 S_{xy}
- 输入图像 $f(x, y)$ ，输出图像： $g(x, y)$
- $f_{\min}, f_{\max}, f_{\text{med}}$ 分别为 S_{xy} 中的最小灰度、最大灰度和中值
- S_{\max} 为 S_{xy} 的最大可能取值

Step1：寻找真正的中值，如找不到，输出图像直接赋原值

$$A1 = f_{med} - f_{min}$$

$$A2 = f_{med} - f_{max}$$

若 $A1 > 0$ 且 $A2 < 0$, 转 step2

否则 $S_{xy} = S_{xy} + 1$

若 $S_{xy} < S_{max}$ 转 step1

否则, $g(x, y) = f(x, y)$

Step2：估计当前像素点是否噪音点，如果是，输出图像赋中值，否则直接赋原值

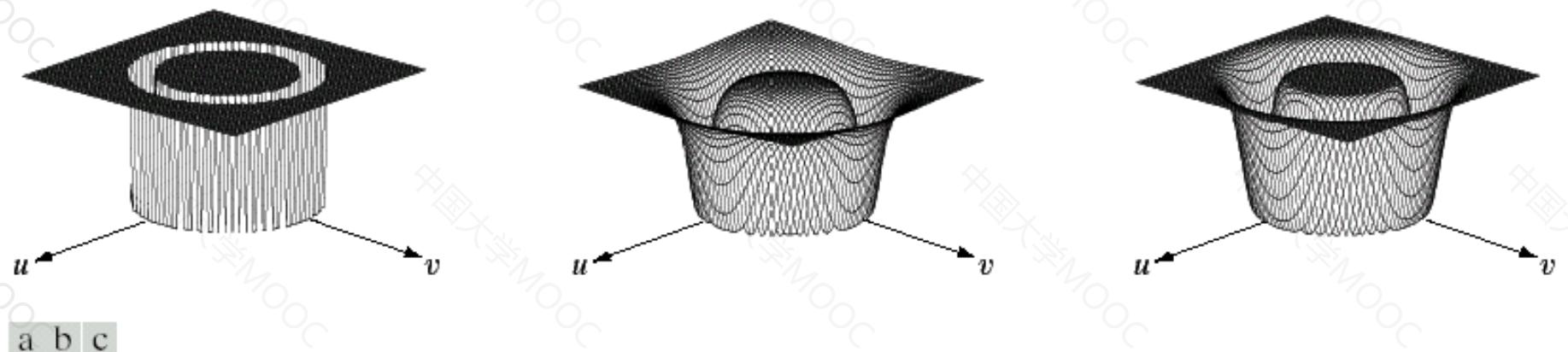
$$B1 = f(x,y) - f_{min}$$

$$B2 = f(x,y) - f_{max}$$

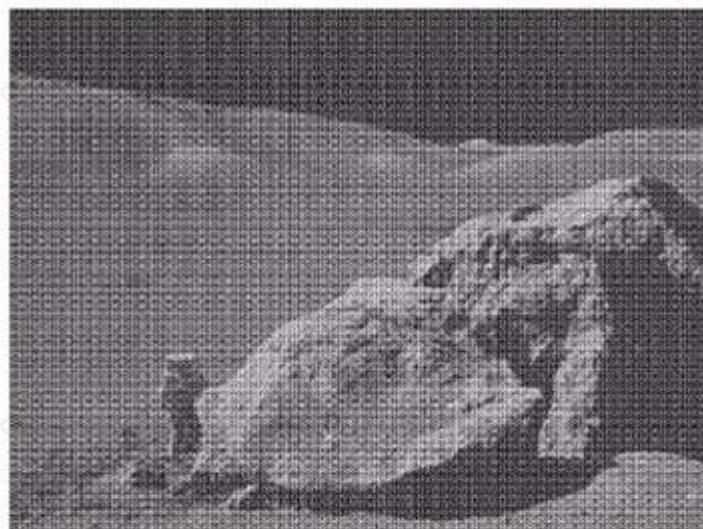
若 $B1 > 0$ 且 $B2 < 0$, $g(x, y) = f(x, y)$

否则 $g(x, y) = f_{med}$

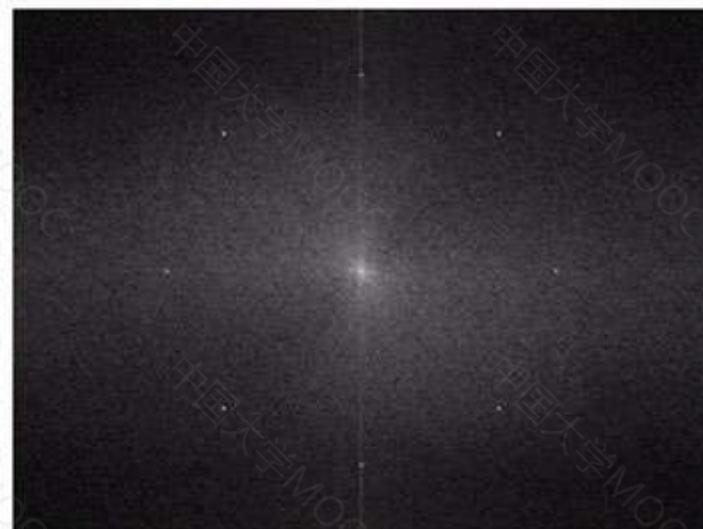
周期噪音抑制 一带阻滤波器



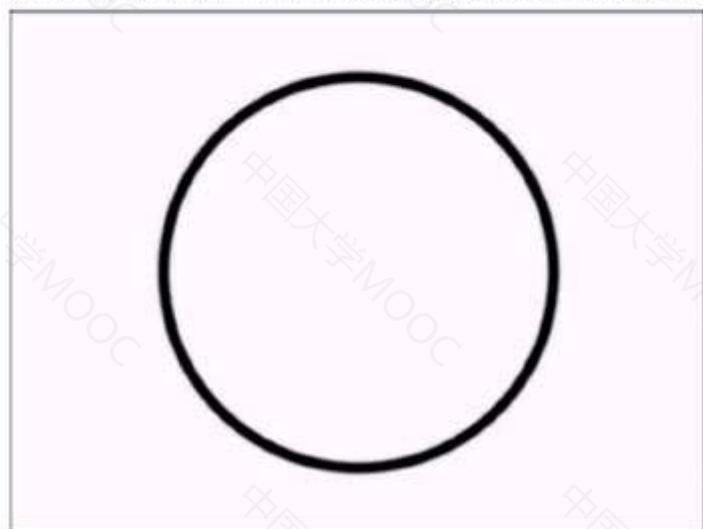
输入
图像



左图
的频
谱图

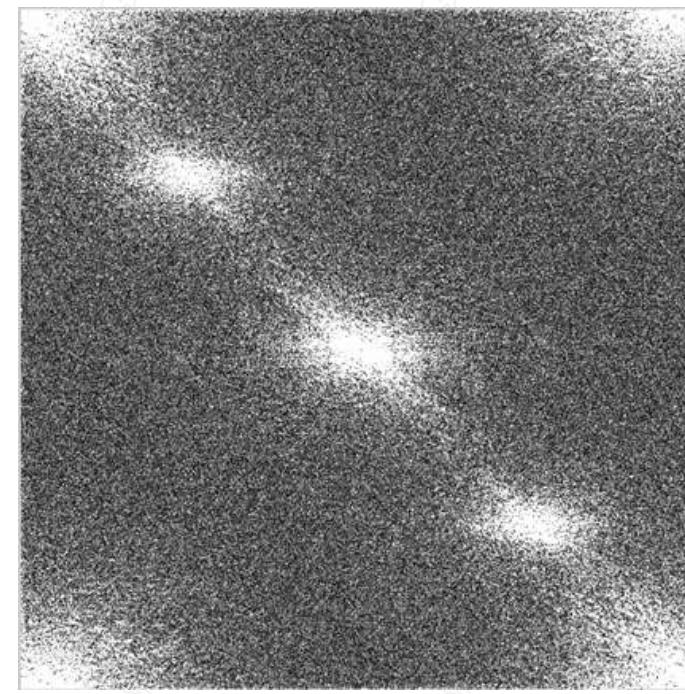


Buttwo
rth带阻
滤波器

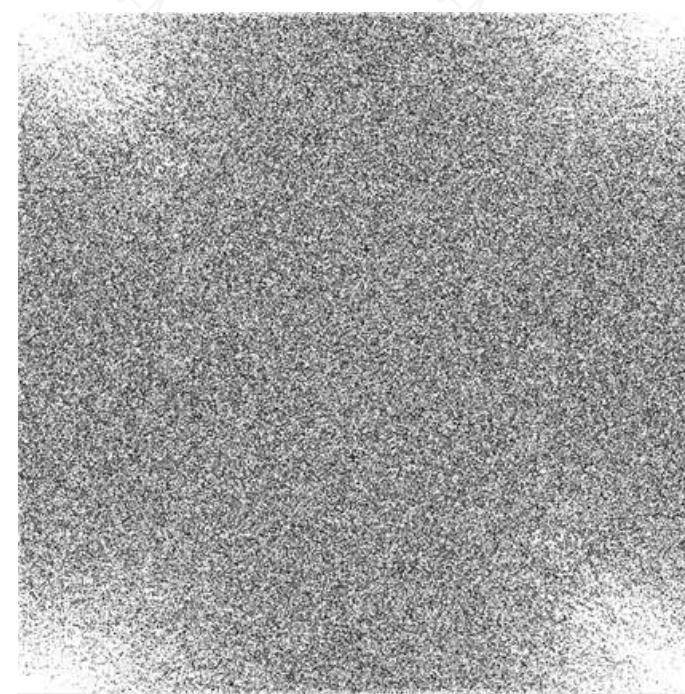


滤波
结果
图像

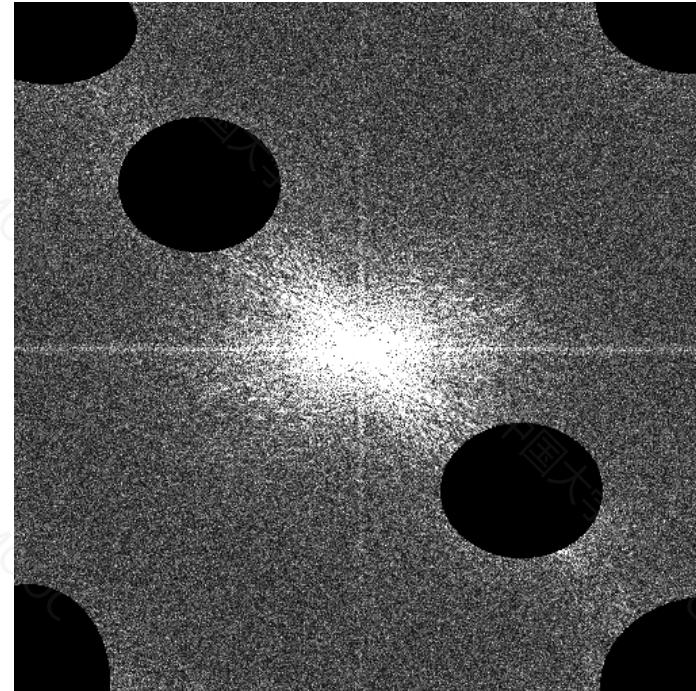




含有周期噪
声和椒盐噪
音的图像和
频谱



仅含有椒盐
噪声的原图
和频谱



带阻滤波器滤除噪音频率



加权中值滤波滤除随机噪音

$$\begin{matrix} 1 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{matrix}$$

5.4 图像复原方法

1. 逆滤波

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)}$$

$$\begin{aligned}\hat{F}(u, v) &= \frac{F(u, v) H(u, v) + N(u, v)}{H(u, v)} \\ &= F(u, v) + \frac{N(u, v)}{H(u, v)}\end{aligned}$$

$F(u, v) \gg N(u, v)$ 时

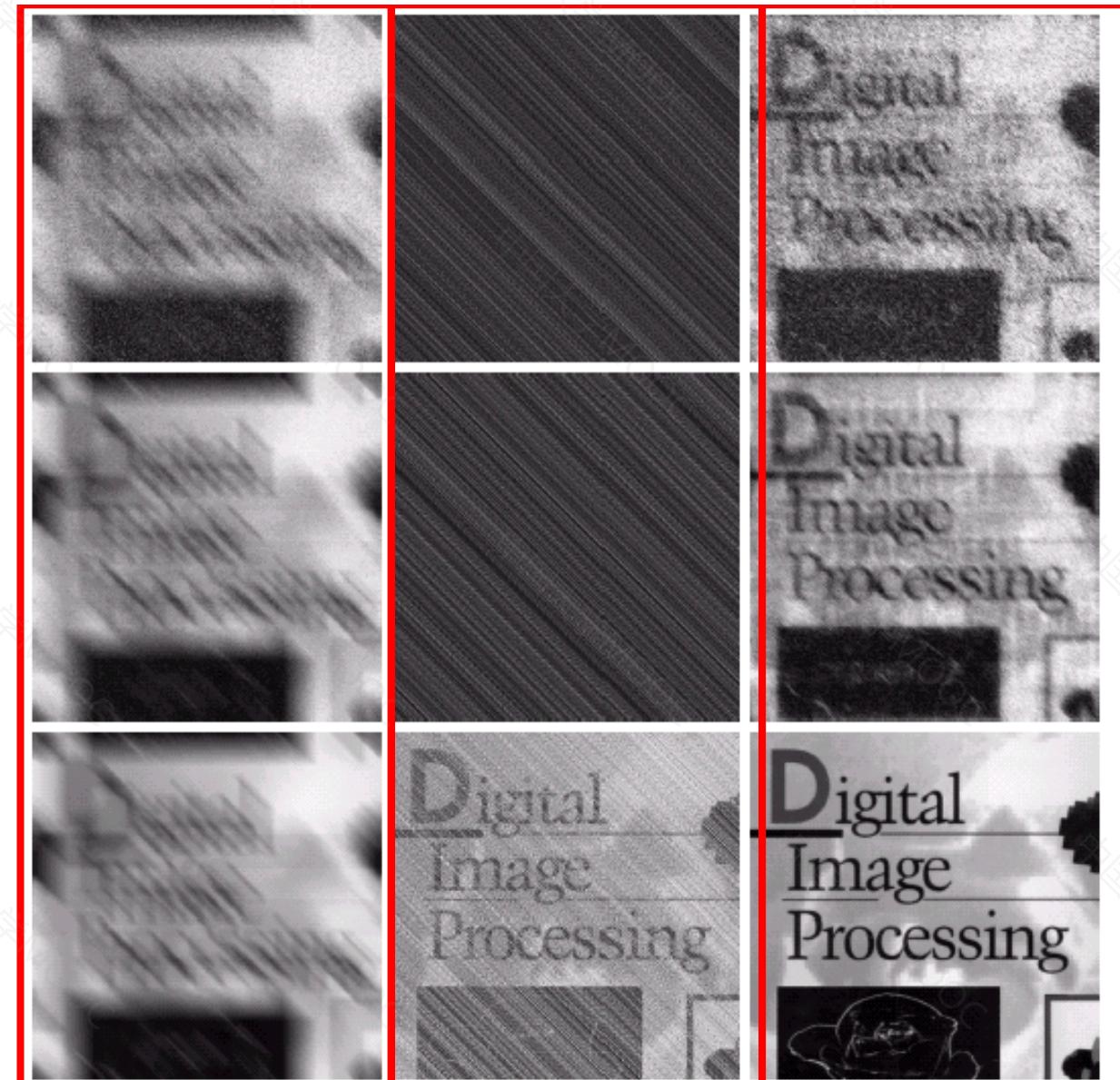
$$\hat{F}(u, v) \approx F(u, v)$$

问题：

零点问题

当 $H(u, v) = 0$ 时，由于 $N(u, v)$ 不可能等于 0，导致 $\hat{F}(u, v) \rightarrow \infty$ ，无法恢复原始图像。

高频噪音放大问题



左：降质图像

中：逆滤波复原结果图像

右：维纳滤波复原结果图像

上：运动
模糊+加性
噪音 10^5

中：运动
模糊+加性
噪音 10^2

下：运动
模糊

二、存在问题

1 零点问题：

$$H(u, v) = 0 \text{ 时} \quad F(u, v) \rightarrow \infty$$

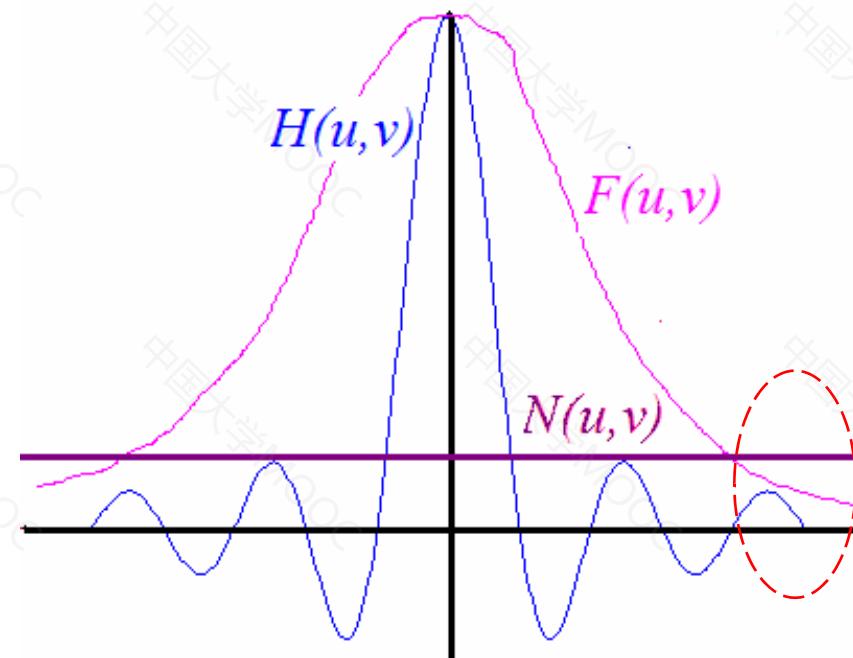
2 高频噪音放大，频率较高时

$$H(u, v) \quad \downarrow$$

$N(u, v)$ 为常量

$$F(u, v) \quad \downarrow$$

$$F'(u, v) - F(u, v) = E(u, v) = \frac{N(u, v)}{H(u, v)} \uparrow$$

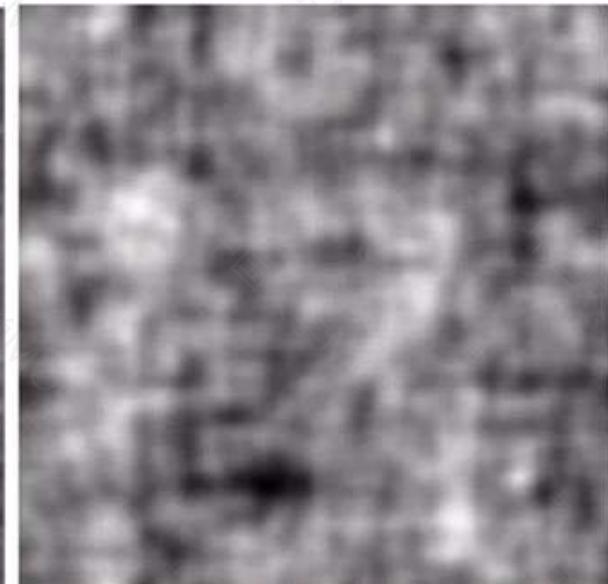




(a) 原始图像



(b) 降质图像



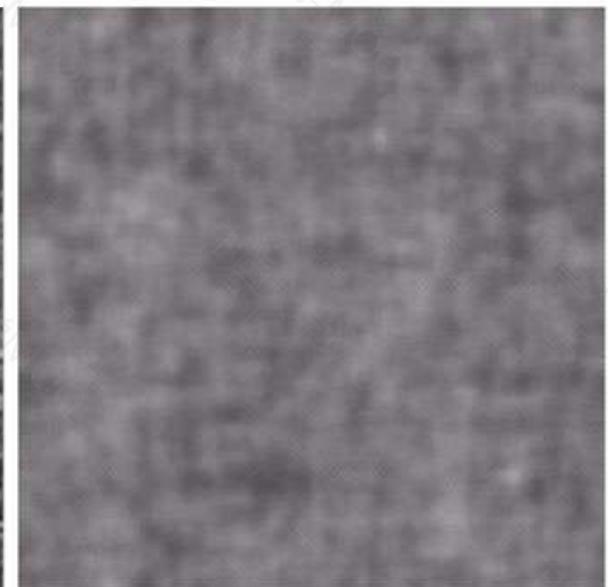
(c) 全频带逆滤波结果图像



(d) 截止频率为40的逆
滤波结果图像；



(e) 截止频率为70的逆
滤波结果图像；



(f) 截止频率为85的逆
滤波结果图像

2. 维纳滤波器

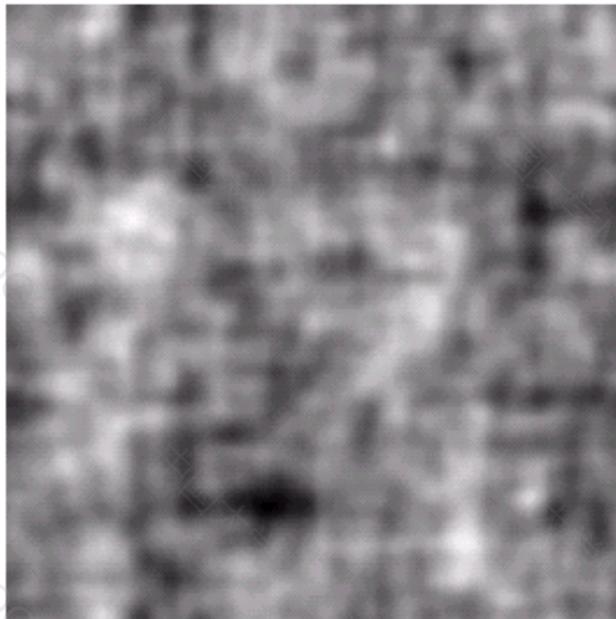
$$\begin{aligned}\hat{F}(u, v) &= \left[\frac{H^*(u, v)}{|H(u, v)|^2 + S_\eta(u, v) / S_f(u, v)} \right] G(u, v) \\ &= \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + S_\eta(u, v) / S_f(u, v)} \right] G(u, v)\end{aligned}$$

$H(u, v)$ 退化函数

$S_\eta(u, v)$ 噪声功率谱

$S_f(u, v)$ 未退化图像的功率谱

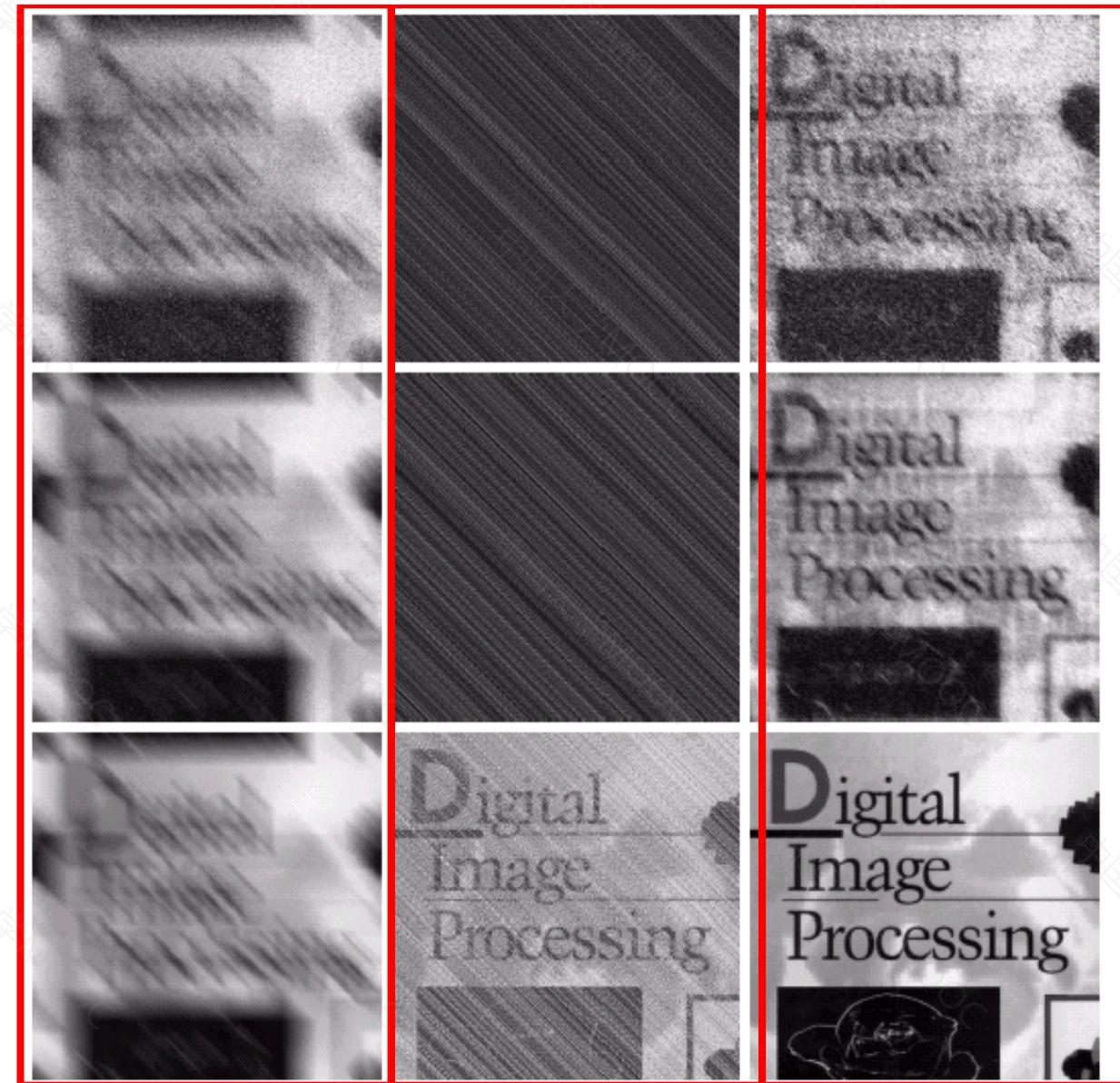
$$\hat{F}(u, v) \approx \left[\frac{1}{H(u, v)} \frac{|H(u, v)|^2}{|H(u, v)|^2 + K} \right] G(u, v)$$



左：全频带逆滤波结果图像；

中：截止频率为70的逆滤波结果图像；

右：维纳滤波结果图像



左：降质图像

中：逆滤波复原结果图像

右：维纳滤波复原结果图像

上：运动
模糊+加性
噪音 10^5

中：运动
模糊+加性
噪音 10^2

下：运动
模糊

■ 图像压缩(Image Compression)

- 6.1 图像压缩基础知识
- 6.2 信息熵编码
- 6.3 基于预测的编码
- 6.4 其他编码：变换编码 矢量量化编码
- 6.5 JPEG标准

■ 6.1 图像压缩基础知识

- 1. 数据和信息
- 2. 数据冗余
- 3. 图象压缩的必要性
- 4. 图像压缩的可行性

■ 6.1 图像压缩基础知识

• 1. 数据和信息

- **数据≠信息**
- **数据是信息的载体**
- **信息是数据要传送的内容**

■ 6.1 图像压缩基础知识

2. 数据冗余

你的妻子，Helen，将于明天晚上6点零5分在上海的虹桥机场接你。

($23*2+10=56$ 个半角字符)

你的妻子将于明天晚上6点零5分在虹桥机场接你。

($20*2+3=43$ 个半角字符)

Helen明晚6点在虹桥接你。

($8*2+6=22$ 个半角字符)

**图象压缩是
数据压缩**

结论：只要接收端不会产生误解，就可以减少承载信息的数据量。

■6.1 图像压缩基础知识

- 3. 图象压缩的必要性——图象数据量巨大，图像压缩非常有必要
 - 标清电视：720*576
 - 每帧数据量： $720*576*3=1,244,160$ 字节= $1.24M$ 字节
 - 每秒数据量： $25*1244160=31,104,000$ 字节= $29.66M$ 字节
 - 20GU盘：容纳690秒=11.5分钟
 - 1.5小时的电影需要占用156G的硬盘空间

■6.1 图像压缩基础知识

4. 图像压缩的可行性——图像数据存在冗余



描述语言

- 1) “这是一幅 $2*2$ 的图像，图像的第一个像素是红的，第二个像素是红的，第三个像素是红的，第四个像素是红的”。
- 2) “这是一幅 $2*2$ 的图 像，所有像素点都是红色的”。

由此我们知道，改变图像的描述方法可以达到压缩的目的。

■ 6.1 图像压缩基础知识

4. 图像压缩的可行性——图像数据存在冗余

图像数据存在冗余——图像数据有损压缩的原理

36	35	34	34	34
34	34	32	34	34
33	37	30	34	34
34	34	34	34	34
34	35	34	34	31



34	34	34	34	34
34	34	34	34	34
34	34	34	34	34
34	34	34	34	34
34	34	34	34	34



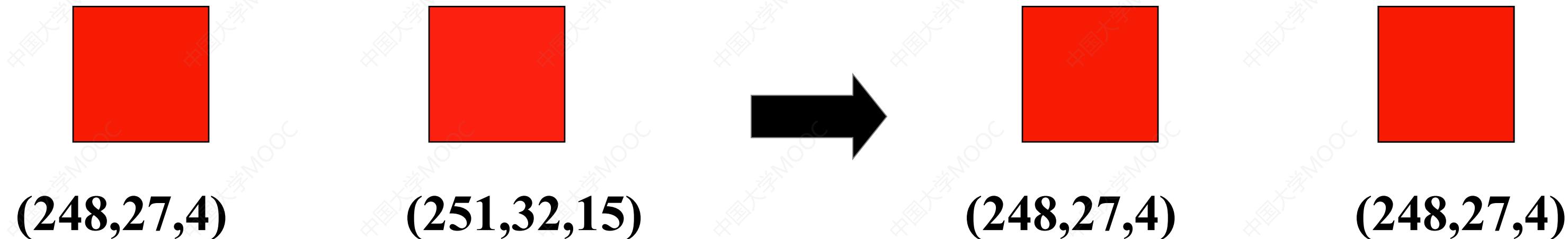
25	34
----	----

■ 6.1 图像压缩基础知识

4. 图像压缩的可行性——图像数据存在冗余 图象的视觉冗余（彩色）

$\rightarrow 2^8 * 2^8 * 2^8 = 2^{24}$

$$2^{24} = 16,777,216$$



■图像数据存在冗余

4. 图像压缩的可行性——图像数据存在冗余

图象的视觉冗余（灰度图像）



4. 图像压缩的可行性——图像数据存在冗余

图像数据冗余分析结论

由于一幅图像存在数据冗余和主观视觉冗余，我们的压缩方式就是从这两方面着手来开展的。

- 1) 因为有数据冗余，当我们把图像的描述方式改变之后，可以去除部分数据冗余。
- 2) 因为有主观视觉冗余，当我们忽略一些视觉感知不太明显的微小差异，可以进行所谓的“有损”压缩。

■ 6.1.2 图像数据冗余分析

图像数据存在各类冗余

- ✓ 空间冗余
- ✓ 时间冗余
- ✓ 统计冗余
- ✓ 视觉心理冗余

■ 6.1.2 图像数据冗余分析

空间冗余

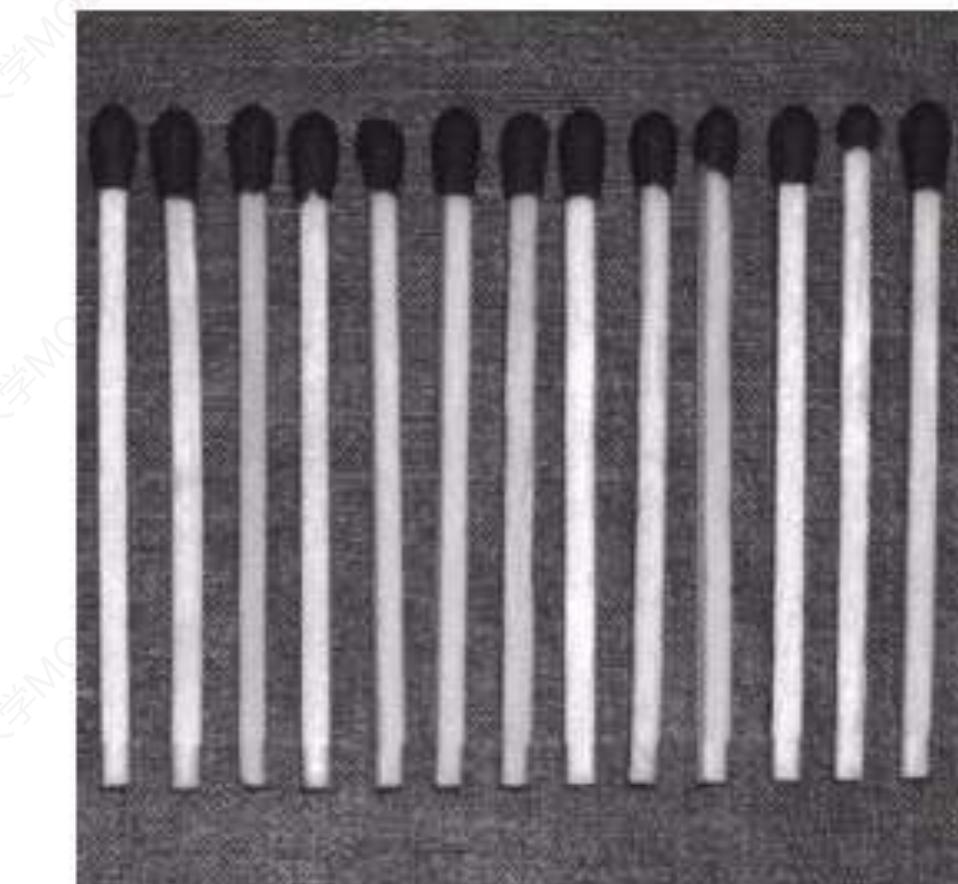
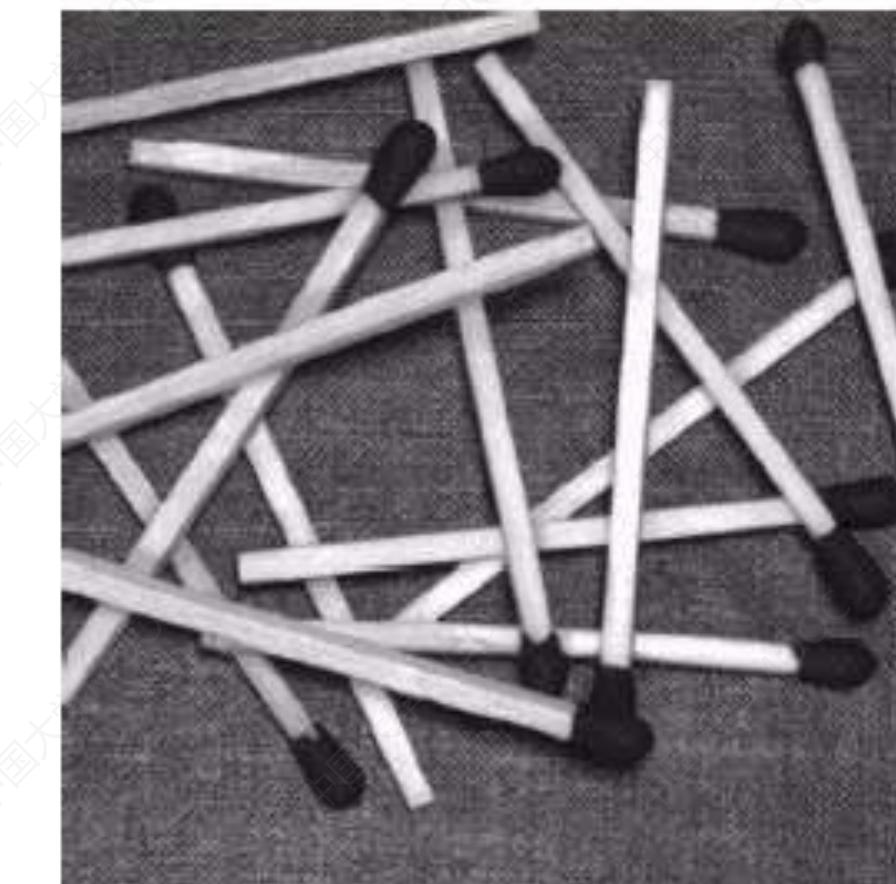
相邻像素点之间存在较强的相关性

$$A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{N-1-\Delta n} f(x, y)f(x, y + \Delta n)$$

$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)}$$

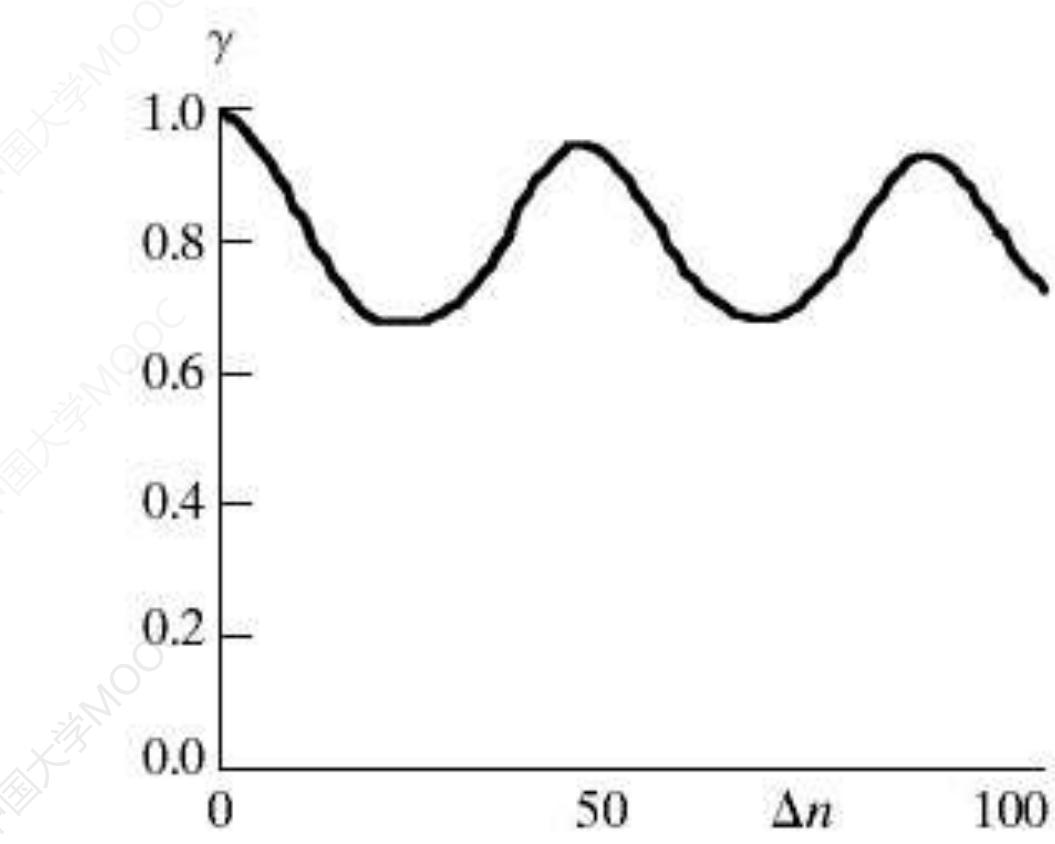
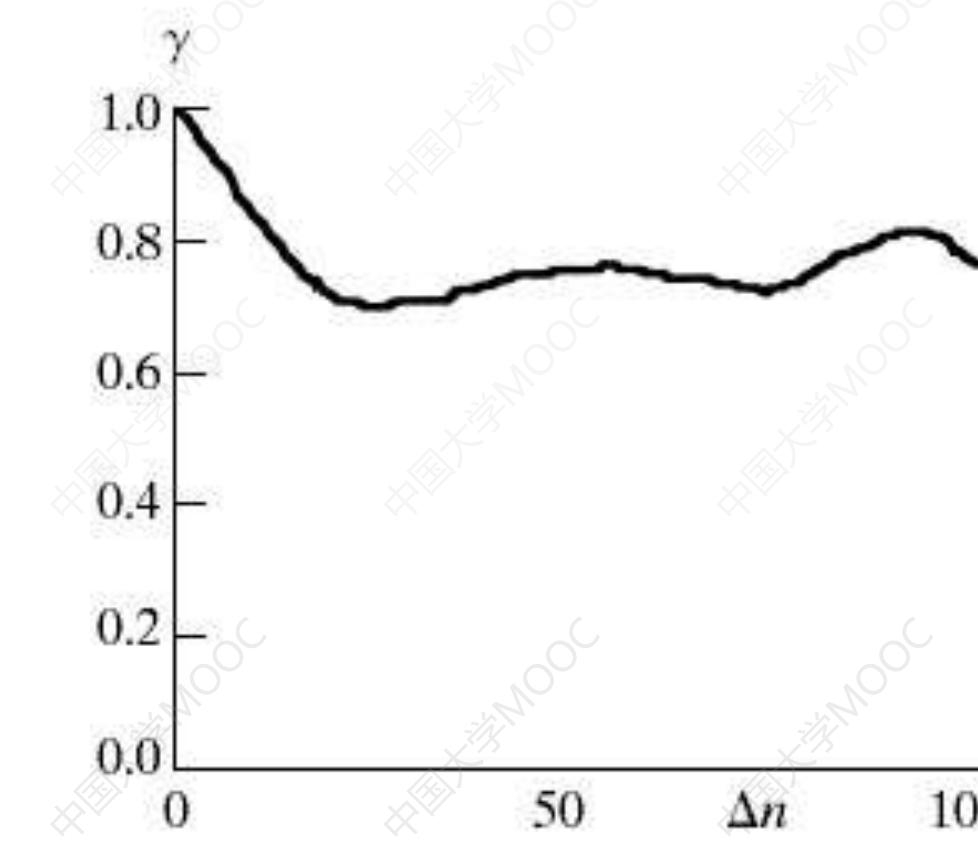
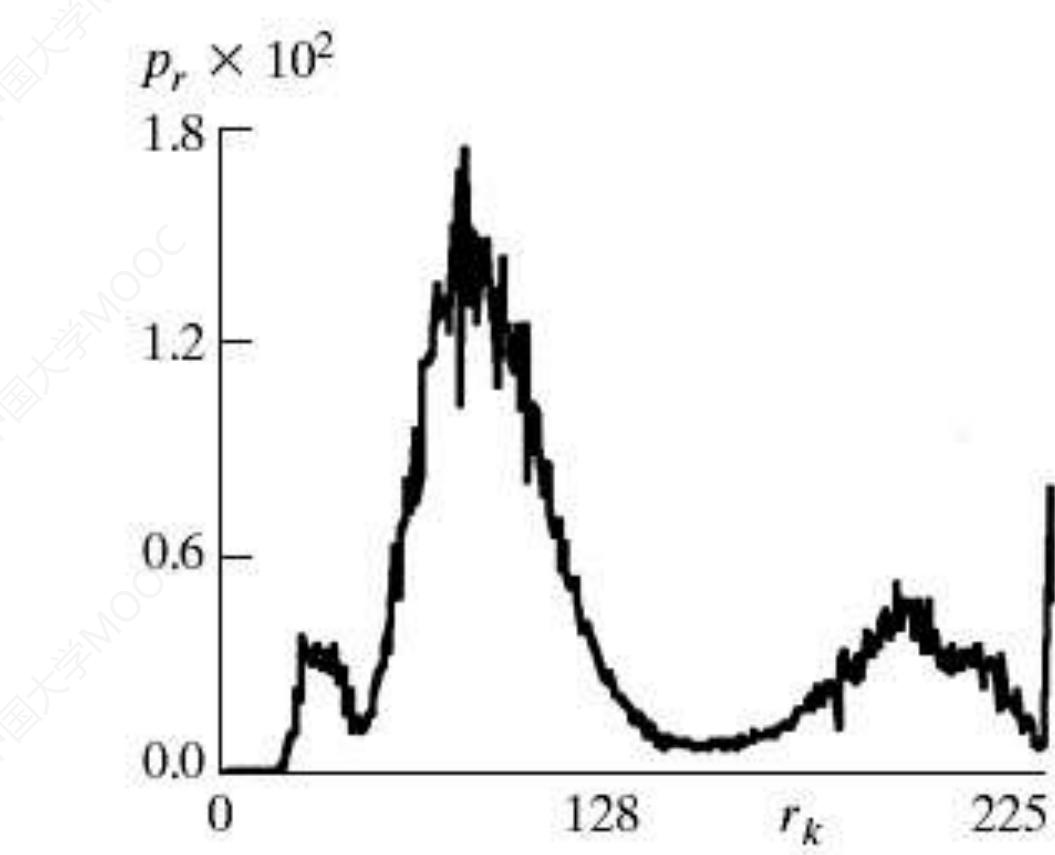
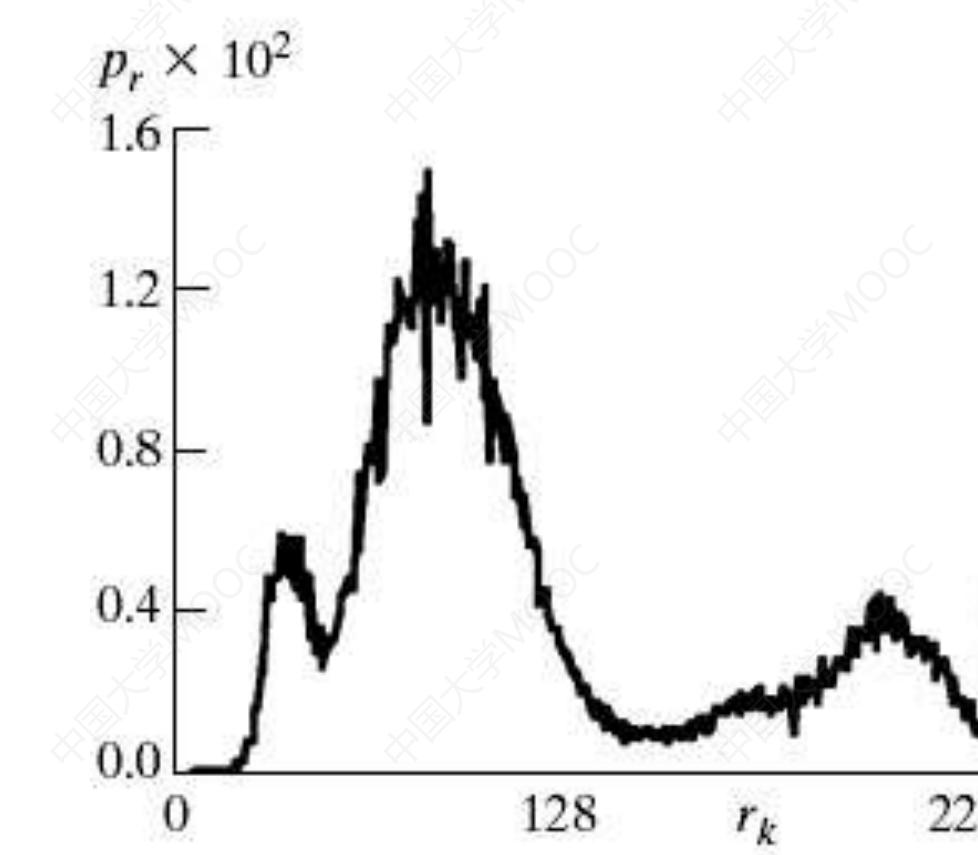
$\Delta n = 1$ 时，

γ 分别为 0.9922 和 0.9928



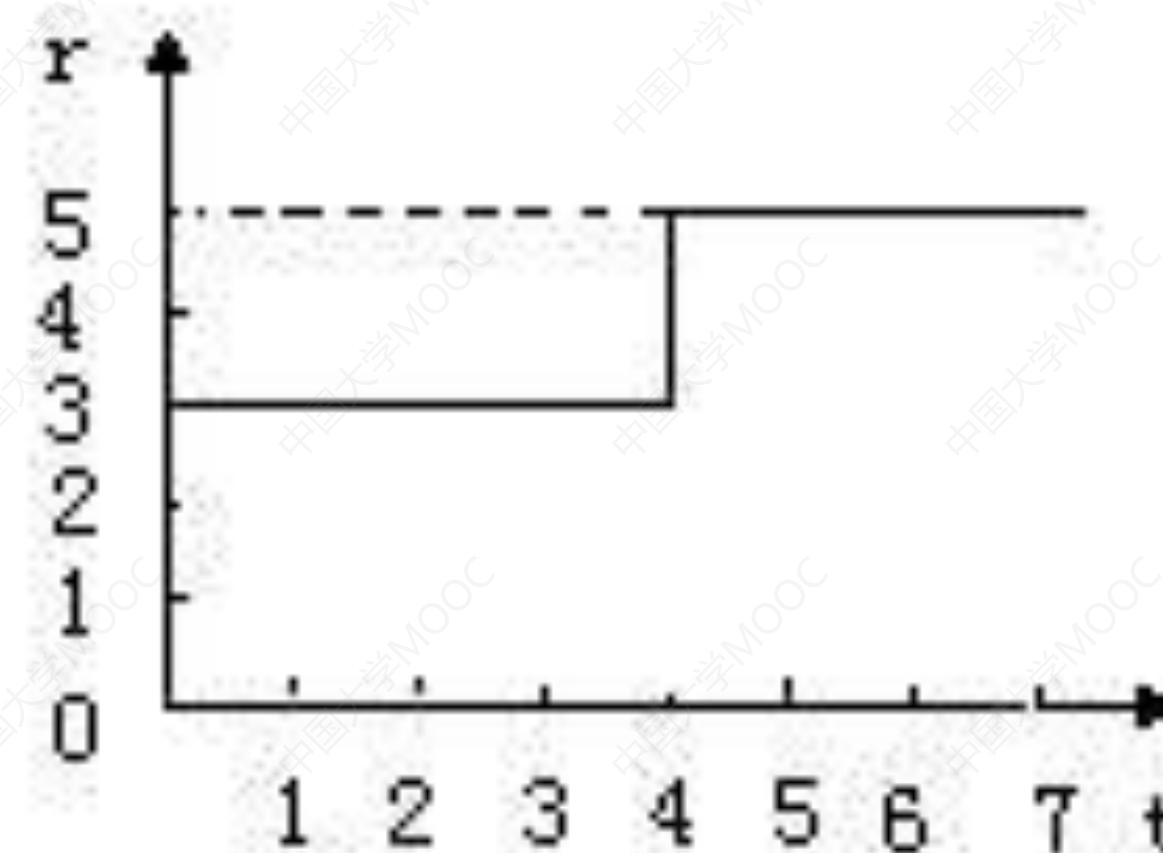
a
b
c
d
e
f

FIGURE 8.2 Two images and their gray-level histograms and normalized autocorrelation coefficients along one line.

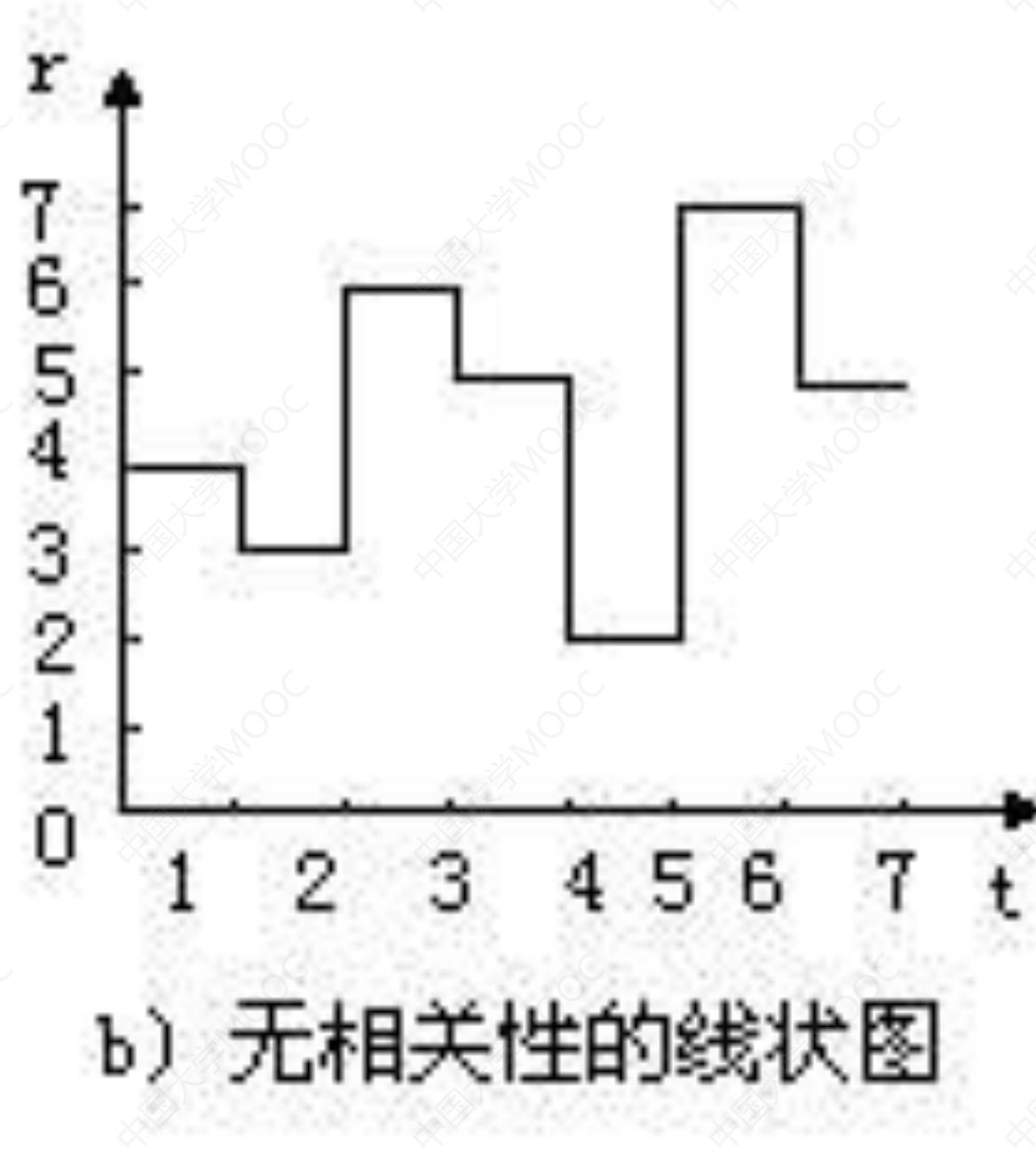


■ 6.1.2 图像数据冗余分析

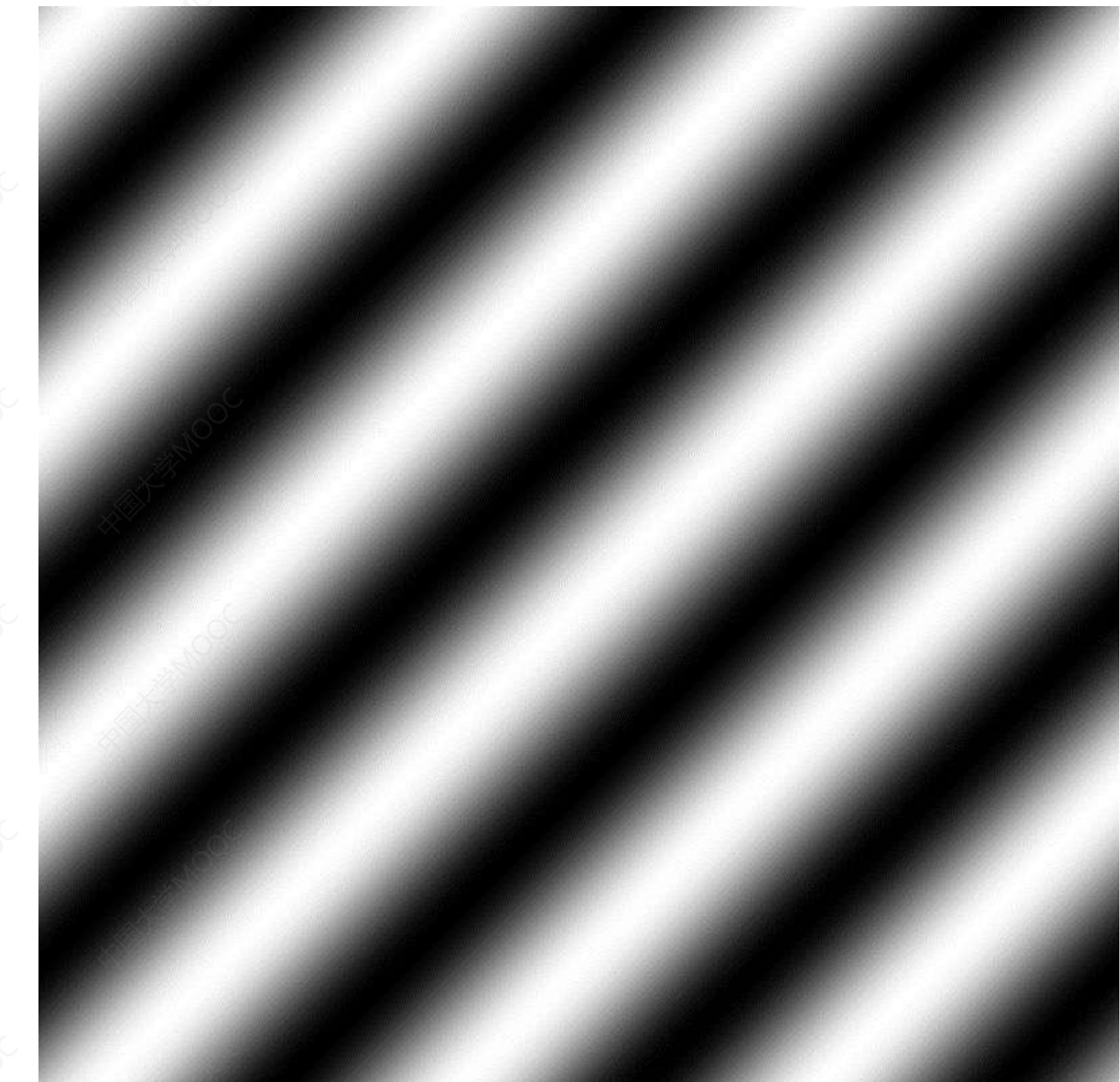
空间冗余



a) 相关性较强的线状图



b) 无相关性的线状图



$$A_0=10, u_0=1, v_0=1, \varphi=\pi/6$$

$$f(x, y) = A_0 \sin(u_0 x + v_0 y + \varphi)$$

■ 6.1.2 图像数据冗余分析

时间冗余

- ✓ 序列图像相邻帧之间存在冗余。
- ✓ 去除时间和空间冗余：预测，变换



■ 6.1.2 图像数据冗余分析

统计冗余（信息熵冗余）

熵的定义

- 随机变量 $X=\{x_1, x_2, x_3 \dots x_n\}$, 其中 x_n 称为消息。消息 x_n 的自信息量与其出现的概率成反比。

$$I(x_n) = -\log(p_n)$$

- 熵是平均信息量，是各消息自信息量的加权平均，

$$H = - \sum p_i \log(p_i)$$

■ 6.1.2 图像数据冗余分析

对数的底数

信息量单位

2:

bit 比特

e:

nat 奈特

10:

hart 哈特

熵与概率分布的关系

➤ 设 $X = \{x_1, x_2, x_3, x_4\}$

1. $p_1 = p_2 = p_3 = p_4 = 1/4 \quad H = ? \quad 2$

2. $p_1 = 1/2; p_2 = 1/4; p_3 = 1/8; p_4 = 1/8 \quad H = ? \quad 7/4$

均匀分布熵最大，等于等字长编码的平均码长

概率分布越不均匀，熵越小。

■ 6.1.2 图像数据冗余分析

实际图像的灰度分布概率均匀吗？

$H <$ 均匀分布概率对应的熵 H_{max} = 等字长编码的平均码长 L_{avg}

统计冗余 = $L_{avg} - H > 0$

统计冗余是由于统计分布概率的不均匀造成的

去除统计冗余——熵编码/统计编码/符号编码

■ 6.1.2 图像数据冗余分析

视觉心理冗余

- 人眼感受到某一点的区域亮度取决于
 - ✓ 入射光和反射系数
 - ✓ 视觉系统反应
- 人类视觉系统对所有视觉信息的敏感不是相等的，在正常的视觉处理过程中，各种信息的相对重要程度不同，那些不十分重要的信息称作视觉心理冗余。
- 去除这些冗余基本不会削弱图像的感知质量

■ 6.1.2 图像数据冗余分析 ■ 视觉心理冗余的去除——量化

视觉心理冗余去除——量化

IGS (Improved Gray-Scale) 量化：

- 利用眼睛对边缘固有的敏感性，通过将一个伪随机数加到每个象素上将这些边缘拆散。
- 伪随机数在对结果进行量化之前根据相邻像素灰度级的低位生成
- 由于低位时完全随机的所以，这样做增加了与伪轮廓相关的人工边缘随机性的灰度级

■ 6.1.2 图像数据冗余分析

视觉心理冗余去除——量化

IGS (Improved Gray-Scale) 量化：

	像素	灰度级	和	IGS量化	普通量化
	$i-1$	N/A	00000000	N/A	N/A
	i	01101100	01101100	0110	0110
	$i+1$	10001011	10010111	1001	1000
	$i+2$	10000111	10001110	1000	1000
	$i+3$	11110100	11110100	1111	1111

■ 6.1.2 图像数据冗余分析

视觉心理冗余去除——量化

IGS (Improved Gray-Scale) 量化：



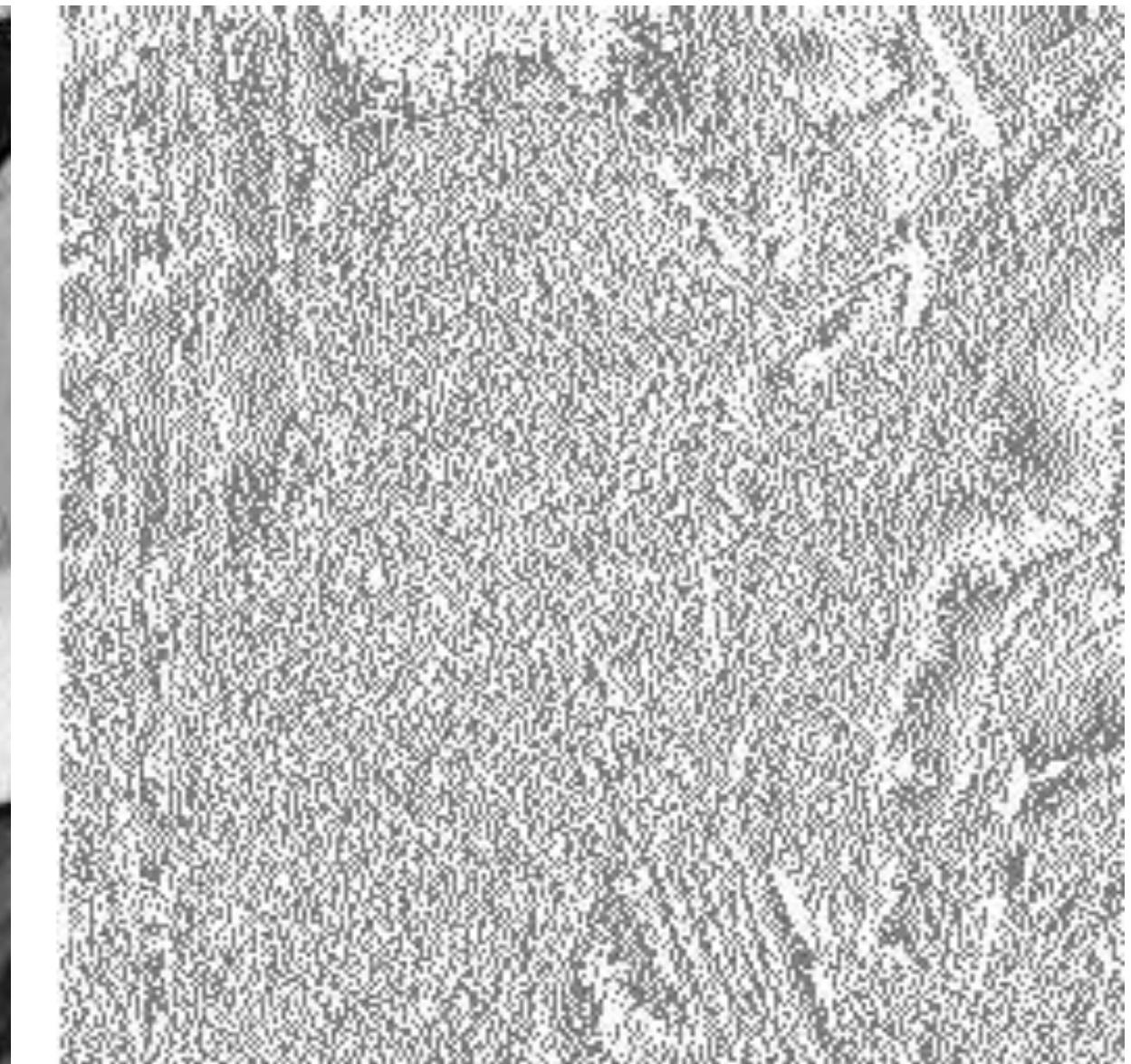
原图(8比特)



普通量化 (4比特)



IGS量化 (4比特)



IGS量化 - 普通量化

普通量化和IGS量化相同像素点达53.9%

■ 6.1.2 图像数据冗余分析

图像压缩编码框架

1. 基本编码框架



2. 去相关技术——去除像素间冗余

基于预测 (Predictive)

基于变换 (Transform)

基于矢量量化 (Vector Quantization)

基于子带变换 (Subband Transform)

■ 6.1.2 图像数据冗余分析

3. 信息熵编码 (Entropy Coding) / 统计编码：利用图像灰度分布统计特性非均匀的特点，采用不等字长进行编码

Huffman 编码

算术编码

Shannon-Fano 编码

B1 码

B2 码

等等

■ 6.1.2 图像数据冗余分析

数据压缩方法分类

- ✓ 无失真（无损）压缩（编码）：Lossless compression/ coding；可逆编码 (reversible coding)；信息保持编码(Imformation preserving coding)：压缩结果无失真，可以完全重构原始图像
- ✓ 有失真（有损）压缩（编码）：Lossy compression/ coding；不可逆编码 (irreversible coding)：压缩结果有失真，不可完全重构原始图像。

■ 6.2 信息熵编码

信息熵编码又称熵编码、统计编码。

主要目的

去除图像灰度分布不均匀带来的信息熵冗余。

■ 6.2 信息熵编码

最佳变长编码定理

在变字长编码中，对于出现概率大的信息符号编以短码字，对于出现概率小的符号编以长码字，如果码字长度严格按照符号概率大小的相
反顺序排列，则平均码字长度一定
小于任何其他排列方式得到的码字
长度。

大概率分配短码字

小概率分配长码字

主要方法

1. *Huffman* 编码
2. *Shannon-Fano* 编码
3. *B1* 码
4. *B2* 码
5. 算术编码

■ 6.2 信息熵编码

*Huffman*编码

1951年，哈夫曼和他在MIT信息论的同学需要选择是完成学期报告还是期末考试。导师Robert M. Fano给他们的学期报告的题目是：寻找最有效的二进制编码。由于无法证明哪个已有编码是最有效的，哈夫曼放弃对已有编码的研究，转向新的探索，最终发现了基于有序频率二叉树编码的想法，并很快证明了这个方法是最有效的。这个算法使用自底向上的方法构建二叉树，避免了他的导师和信息论创立者香农共同提出的Shannon-Fano编码的最大弊端——自顶向下构建树；

1952年，David A. Huffman在麻省理工攻读博士时发表了“*A Method for the Construction of Minimum-Redundancy Codes*”，它就叫做Huffman编码

这个故事告诉我们不要害怕挑战权威！

■ 6.2 信息熵编码

*Huffman*编码

已知一个随机变量X的6种消息取值及取值概率

	x1	x2	x3	x4	x5	x6
p(xi)	0.4	0.06	0.1	0.04	0.3	0.1

试进行*Huffman*编码，对编码结果进行分析，并给出符号序列x1x2x5x6x3x1x6的输出码流。

■ 6.2 信息熵编码

*Huffman*编码——概率重组

$p(x_i)$

$x_1 \quad 0.4$

$x_5 \quad 0.3$

$x_3 \quad 0.1$

$x_6 \quad 0.1$

$x_2 \quad 0.06$

$x_4 \quad 0.04$

■ 6.2 信息熵编码

*Huffman*编码——概率重组

$p(x_i)$

$x_1 \quad 0.4$ — 0.4

$x_5 \quad 0.3$ — 0.3

$x_3 \quad 0.1$ — 0.1

$x_6 \quad 0.1$ — 0.1

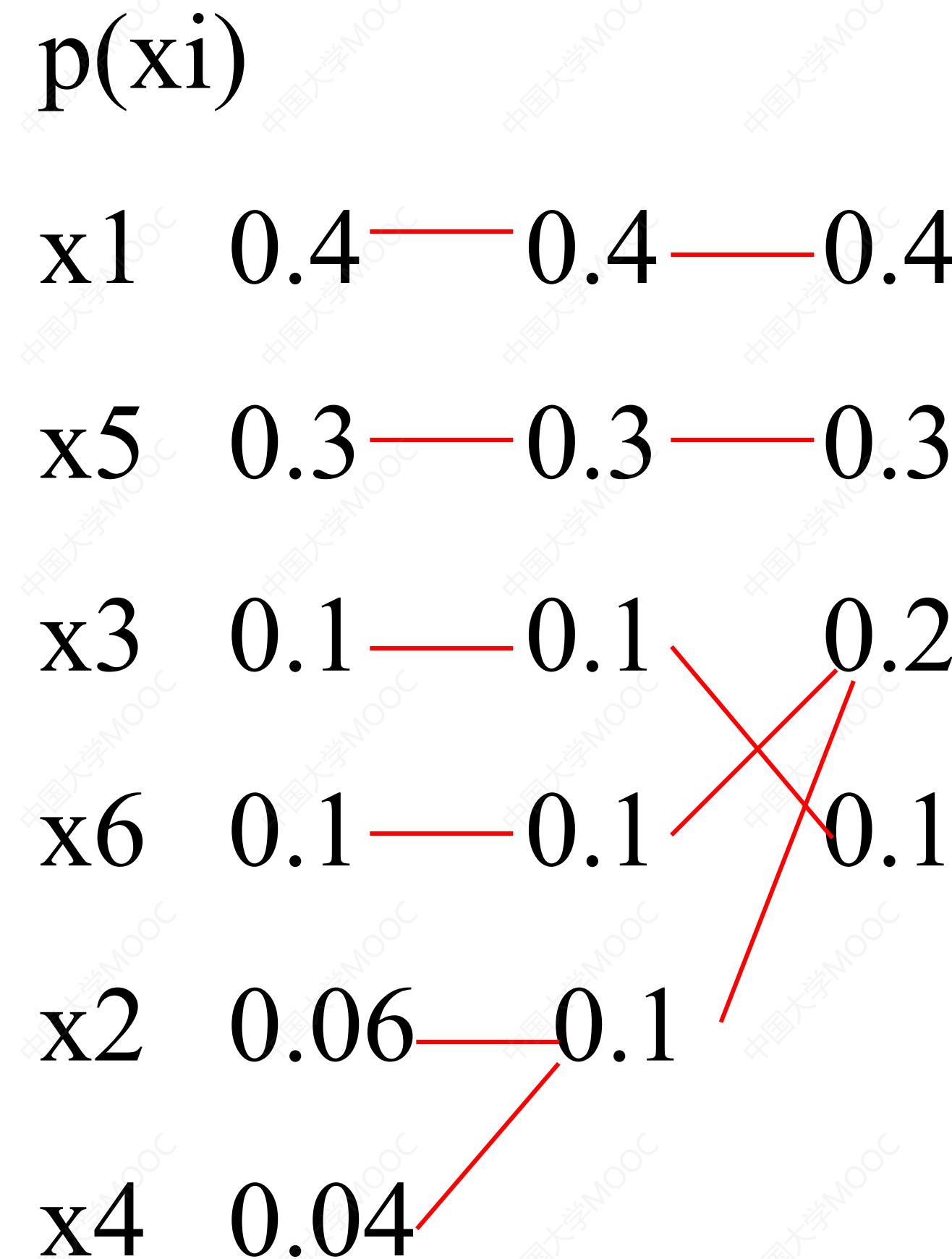
$x_2 \quad 0.06$ — 0.1

$x_4 \quad 0.04$



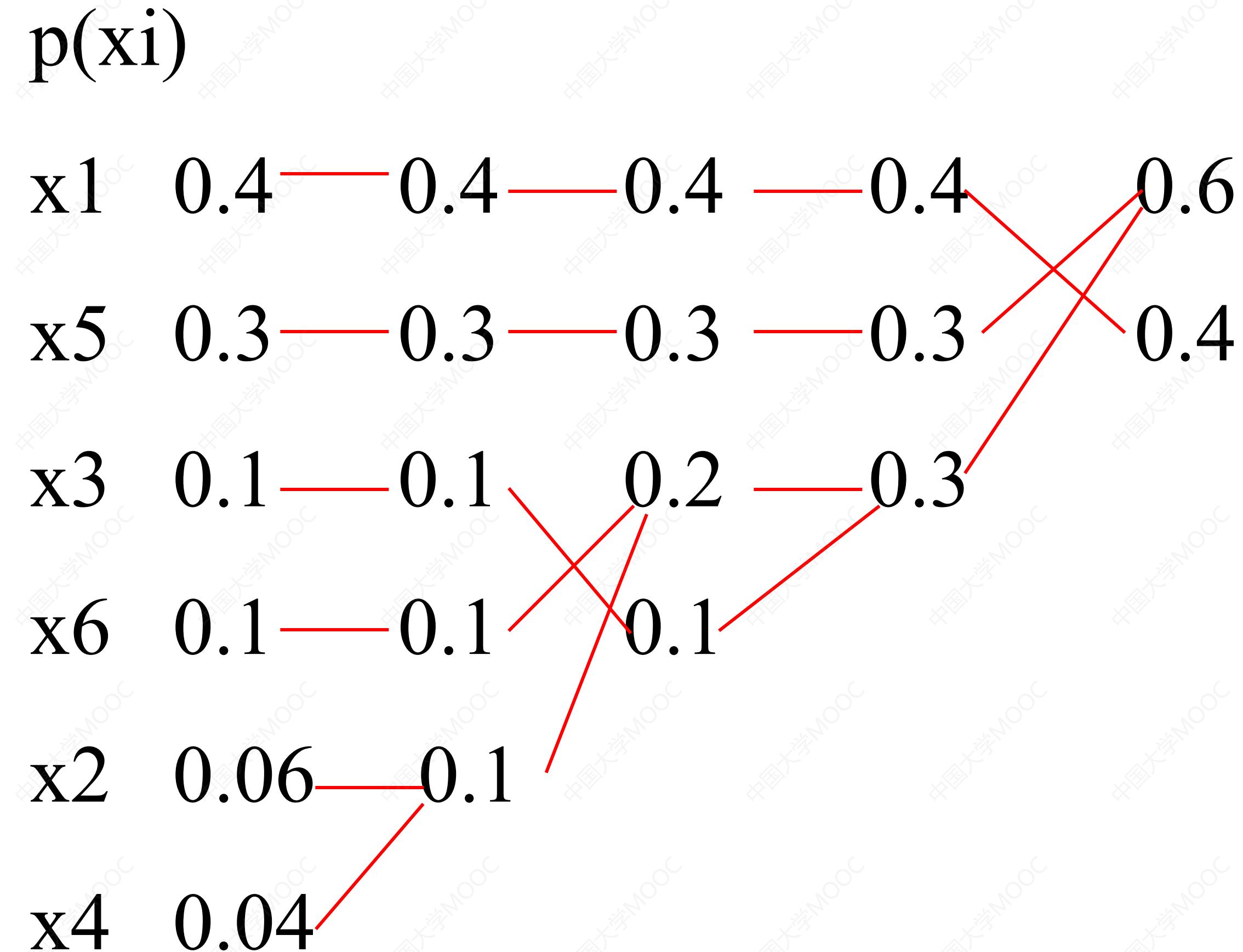
■ 6.2 信息熵编码

*Huffman*编码——概率重组



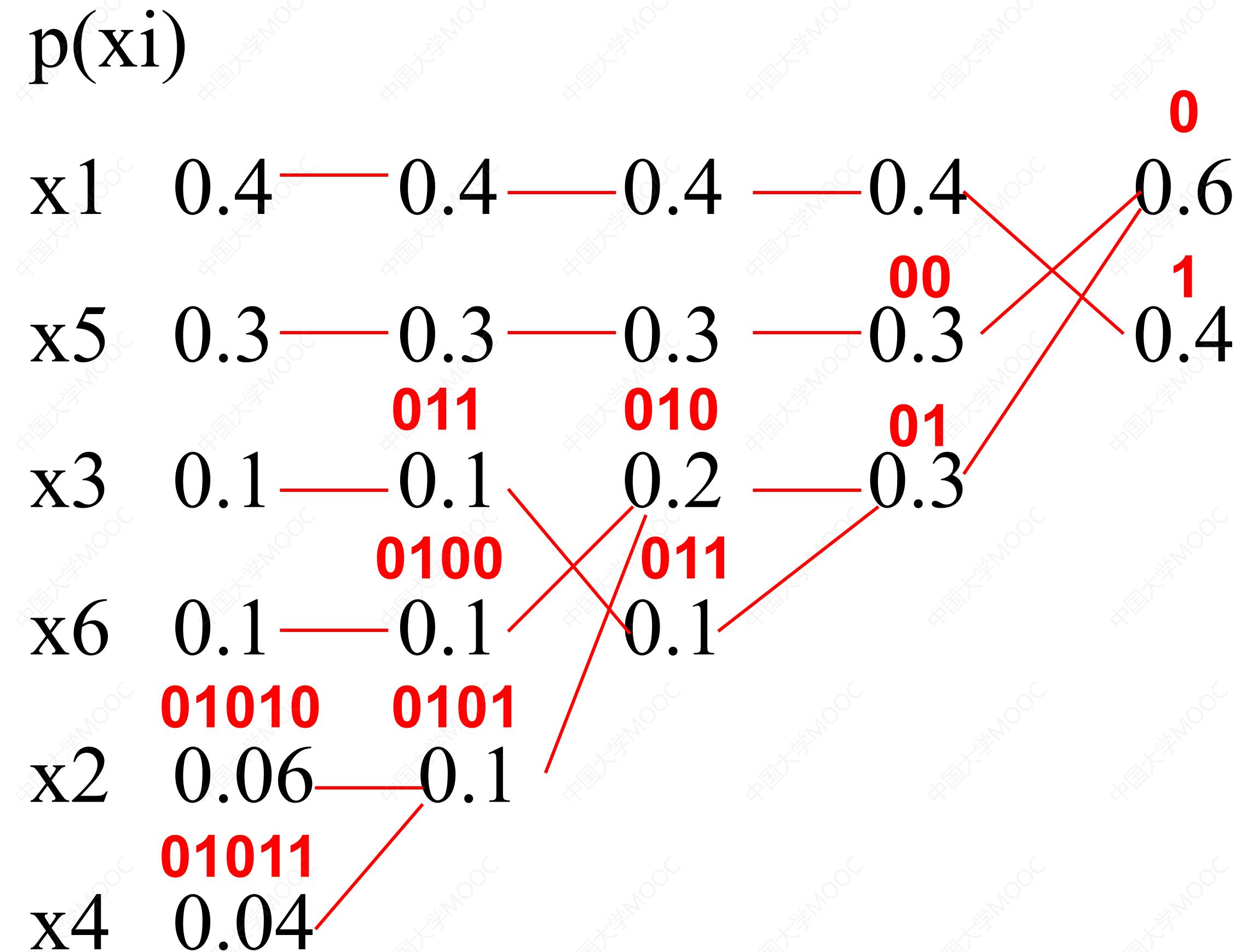
6.2 信息熵编码

Huffman 编码——概率重组



■ 6.2 信息熵编码

*Huffman*编码——码字分配



■ 6.2 信息熵编码

*Huffman*编码——编码结果

x1	x2	x3	x4	x5	x6
0.4	0.06	0.1	0.04	0.3	0.1
1	01010	011	01011	00	0100

平均字长/码字

$$R = 0.4 + 0.06 * 5 + 0.1 * 3 + 0.04 * 5 + 0.3 * 2 + 0.1 * 4 = 2.2 \text{ 比特}$$

熵 $H = 2.14 \text{ 比特}$

等字长编码码长 $A = 3 \text{ 比特}$

压缩比 $= A/R = 1.36$ 编码效率 $= H/R = 97.27\%$

■ 6.2 信息熵编码

*Huffman*编码——编码举例

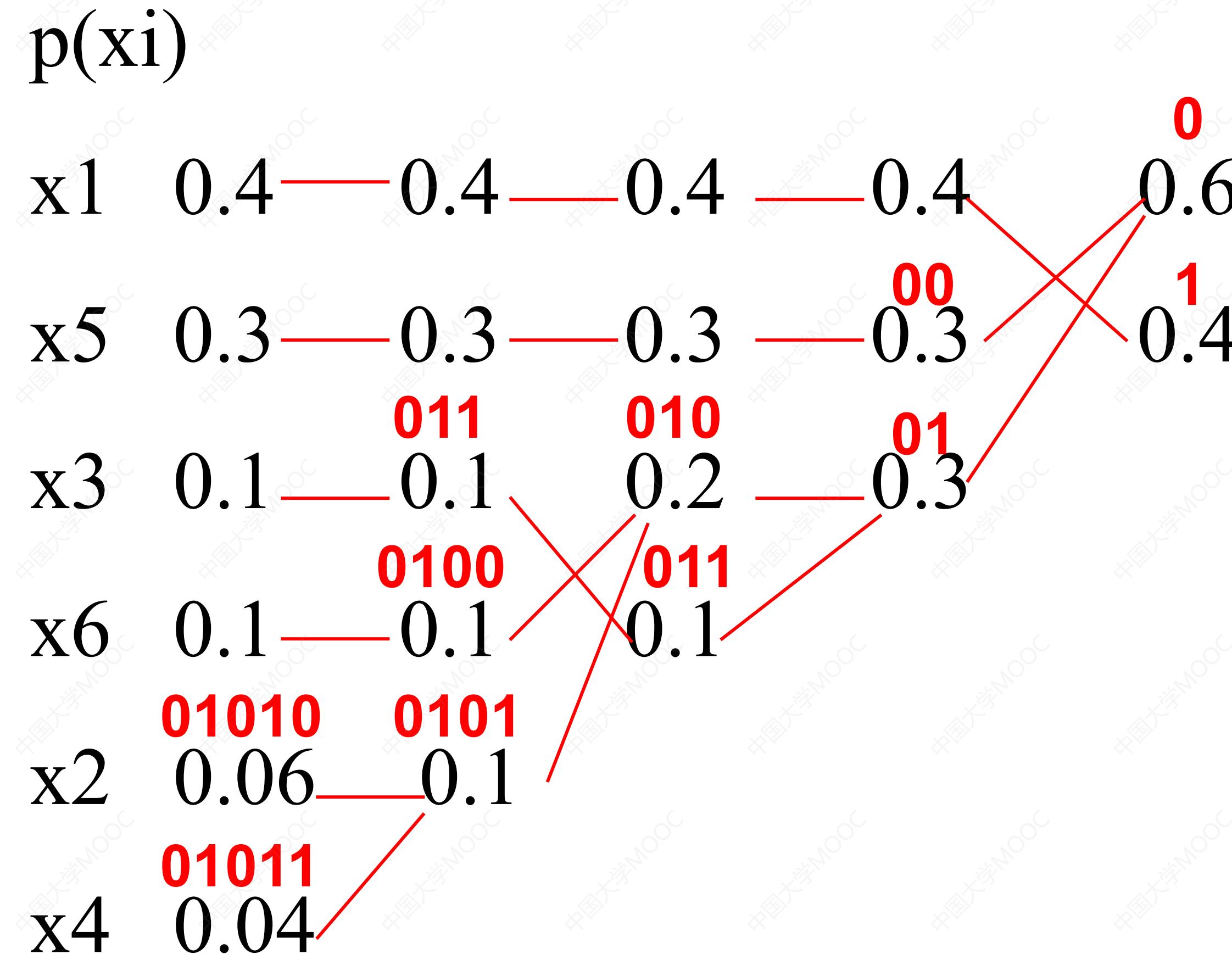
对符号序列x₁x₂x₅x₆x₃x₁x₆的编码结果

1 01010 00 0100 011 1 0100

问题1：等概率时，换位是否影响最终的编码结果？是否影响编码性能？

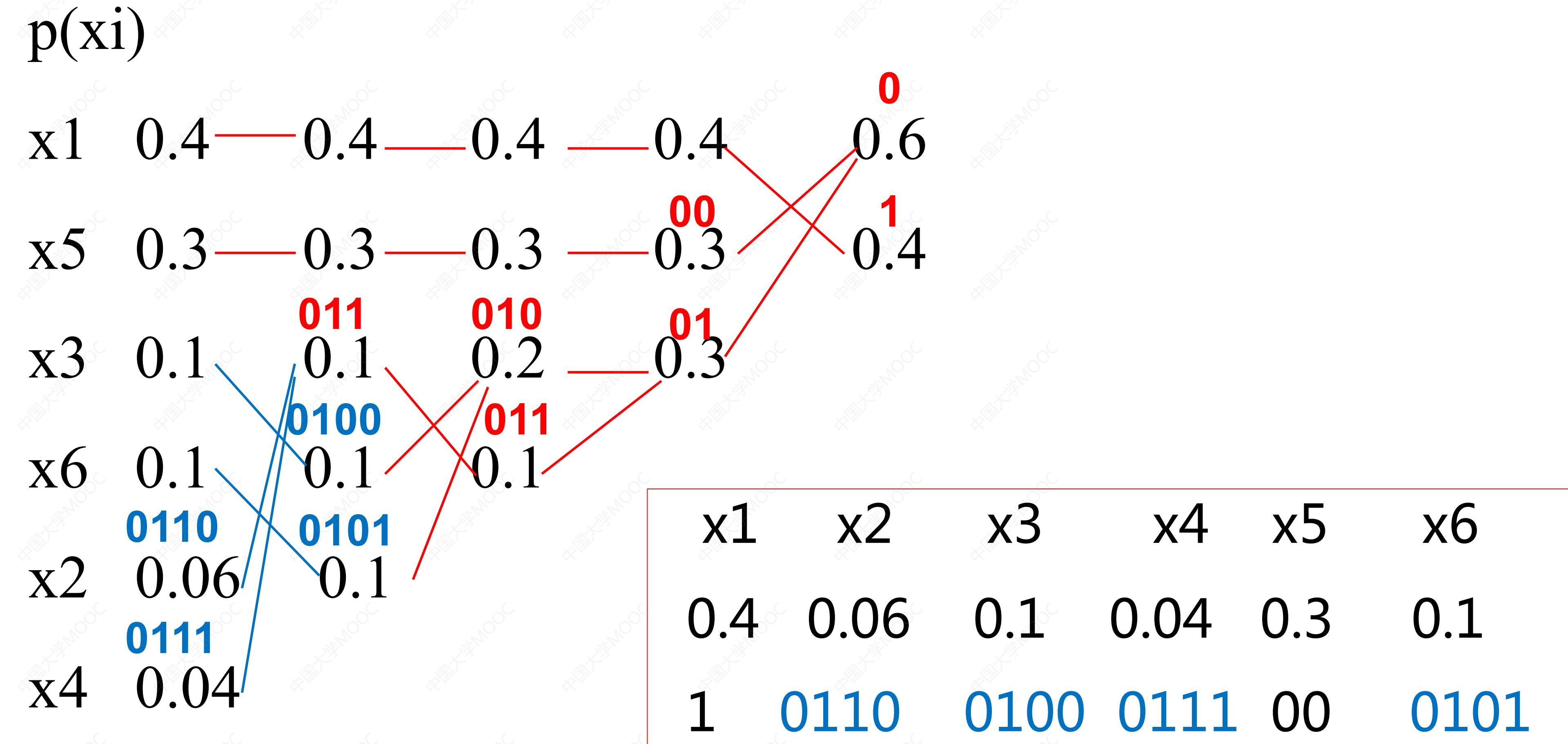
■ 6.2 信息熵编码

等概率时换位对编码结果的影响



■ 6.2 信息熵编码

等概率时换位对编码结果的影响



平均字长/码字

$$R = 0.4 + 0.06 * 4 + 0.1 * 4 + 0.04 * 4 + 0.3 * 2 + 0.1 * 4 = 2.2 \text{ 比特}$$

■ 6.2 信息熵编码

*Huffman*编码——编码序列举例

对符号序列x₁x₂x₅x₆x₃x₁x₆的编码结果

1 01010 00 0100 011 1 0100

问题1：等概率时，换位是否影响最终的编码结果？是否影响编码性能？

问题2：码字赋值0,1交换，对编码结果有何影响？

Huffman编码的局限性：编码结果不唯一，编解码必须依据相同的码表

■ 6.2 信息熵编码

编码序列举例：x₁x₂x₅x₆x₃x₁x₆

1 01010 00 0100 011 1 0100

问题1：等概率时，换位是否影响最终的编码结果？是否影响编码性能？

问题2：码字赋值0,1交换，对编码结果有何影响？

Huffman编码的局限性：编码结果不唯一，编解码必须依据相同的码表

■ 6.2 信息熵编码

Huffman 编码的软件实现

最优二叉树：每个非叶节点有且仅有两个子节点。

叶节点数： n **总节点数**： $2n-1$

Struct nodetype{ weight, parent, lch,rch}

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1	0.4	0	0	0
2	0.06	0	0	0
3	0.1	0	0	0
4	0.04	0	0	0
5	0.3	0	0	0
6	0.1	0	0	0

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1	0.4	0	0	0
2	0.06	7	0	0
3	0.1	0	0	0
4	0.04	7	0	0
5	0.3	0	0	0
6	0.1	0	0	0
7	0.1	0	2	4

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1	0.4	0	0	0
2	0.06	7	0	0
3	0.1	8	0	0
4	0.04	7	0	0
5	0.3	0	0	0
6	0.1	8	0	0
7	0.1	0	2	4
8	0.2	0	3	6

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1 **0.4** 0 0 0

2 **0.06** 7 0 0

3 **0.1** 8 0 0

4 **0.04** 7 0 0

5 **0.3** 0 0 0

6 **0.1** 8 0 0

7 **0.1** 9 2 4

8 **0.2** 9 3 6

9 **0.3** 0 8 7

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1 **0.4** 0 0 0

2 **0.06** 7 0 0

3 **0.1** 8 0 0

4 **0.04** 7 0 0

5 **0.3** 10 0 0

6 **0.1** 8 0 0

7 **0.1** 9 2 4

8 **0.2** 9 3 6

9 **0.3** 10 8 7

10 **0.6** 0 5 9

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1 **0.4** 11 0 0

2 **0.06** 7 0 0

3 **0.1** 8 0 0

4 **0.04** 7 0 0

5 **0.3** 10 0 0

6 **0.1** 8 0 0

7 **0.1** 9 2 4

8 **0.2** 9 3 6

9 **0.3** 10 8 7

10 **0.6** 11 5 9

11 **1.0** 0 10 1

■ 6.2 信息熵编码

Num Weight Par Lch Rch

1 0.4 11 0 0

2 0.06 7 0 0

3 0.1 8 0 0

4 0.04 7 0 0

5 0.3 10 0 0

6 0.1 8 0 0

7 0.1 9 2 4

8 0.2 9 3 6

9 0.3 10 8 7

10 0.6 11 5 9

11 1.0 0 10 1

1

0110

0100

0111

00

0101

2 *Shannon-Fano*编码

x1 0.4

x5 0.3

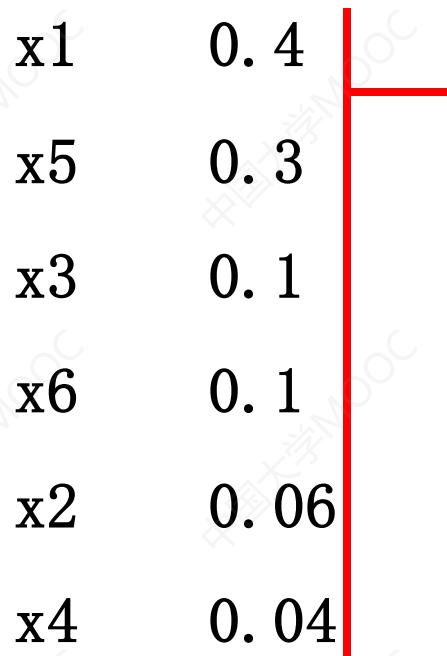
x3 0.1

x6 0.1

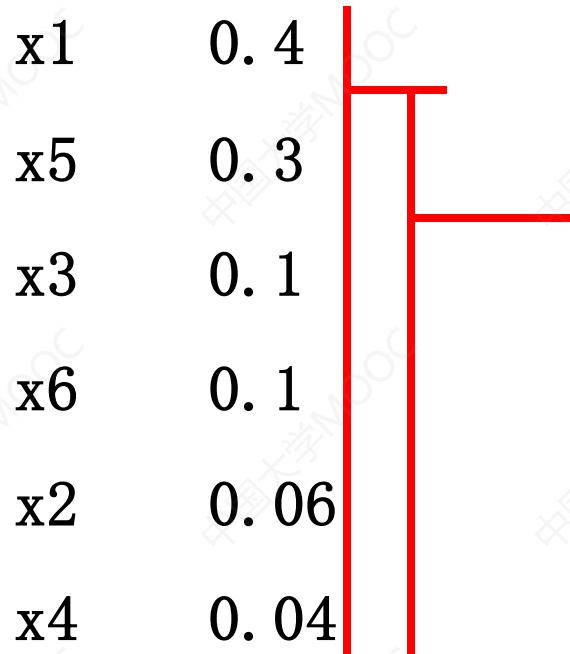
x2 0.06

x4 0.04

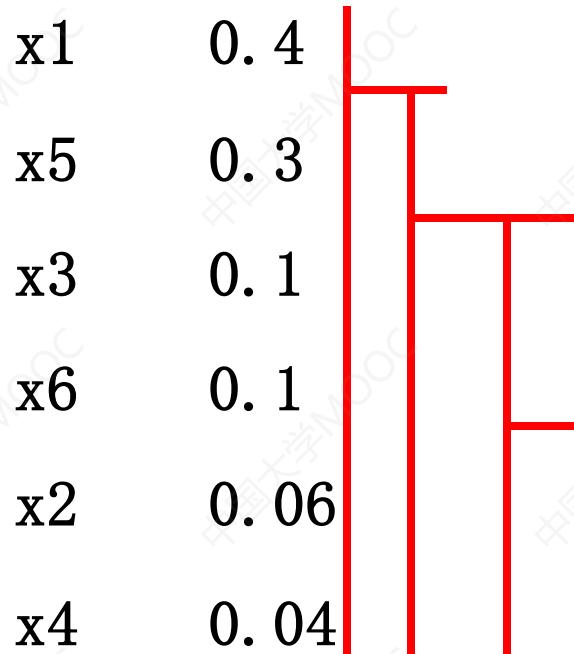
2 Shannon-Fano 编码



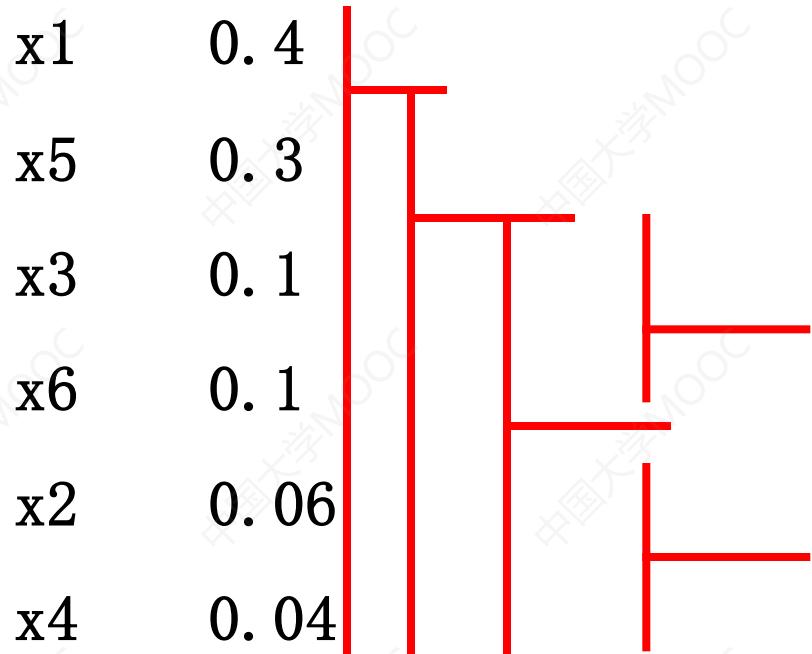
2 Shannon-Fano 编码



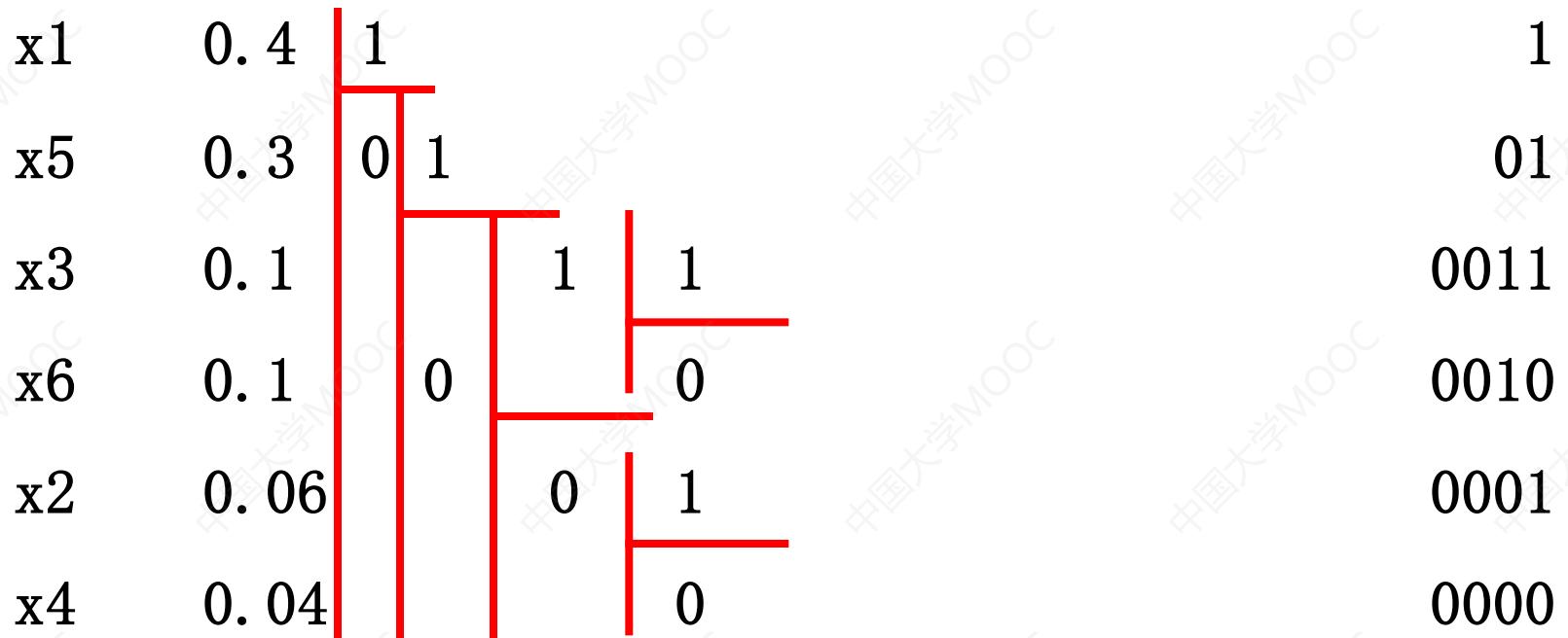
2 Shannon-Fano 编码



2 Shannon-Fano 编码



2 Shannon-Fano 编码



符号序列和输出码流：

符号序列： $x_1 x_2 x_5 x_6 x_3 x_1 x_6$:

输出码流： 1 0001 01 0010 0011 1 0010

3. B1码

延续比特+信息比特

x_1	x_2	x_3	x_4	x_5	x_6
0.4	0.06	0.1	0.04	0.3	0.1

3. B1码

延续比特+信息比特

	x1	x5	x3	x6	x2	x4
按概率大小排序	0.4	0.3	0.1	0.1	0.06	0.04
分配信息比特	0	1	00	01	10	11
组成B1码	c0	c1	c0c0	c0c1	c1c0	c1c1

符合序列和编码结果:

符号序列: 延续比特待定

编码结果: 00 | 1110 | 01 | 1011 | 0000 | 10 | 0001

4. B2码

延续比特+信息比特

	x1	x5	x3	x6	x2	x4
按概率大小排序	0. 4	0. 3	0. 1	0. 1	0. 06	0. 04
分配信息比特	00	01	10	11	0000	0001
组成B1码	c00	c01	c10	c11	c00c00	c00c01

符号序列和编码结果

符号序列: $x_1x_2x_5x_6x_3x_1x_6$:

编码结果: 000 100100 001 111 010 100 011

■ 5. 算术编码

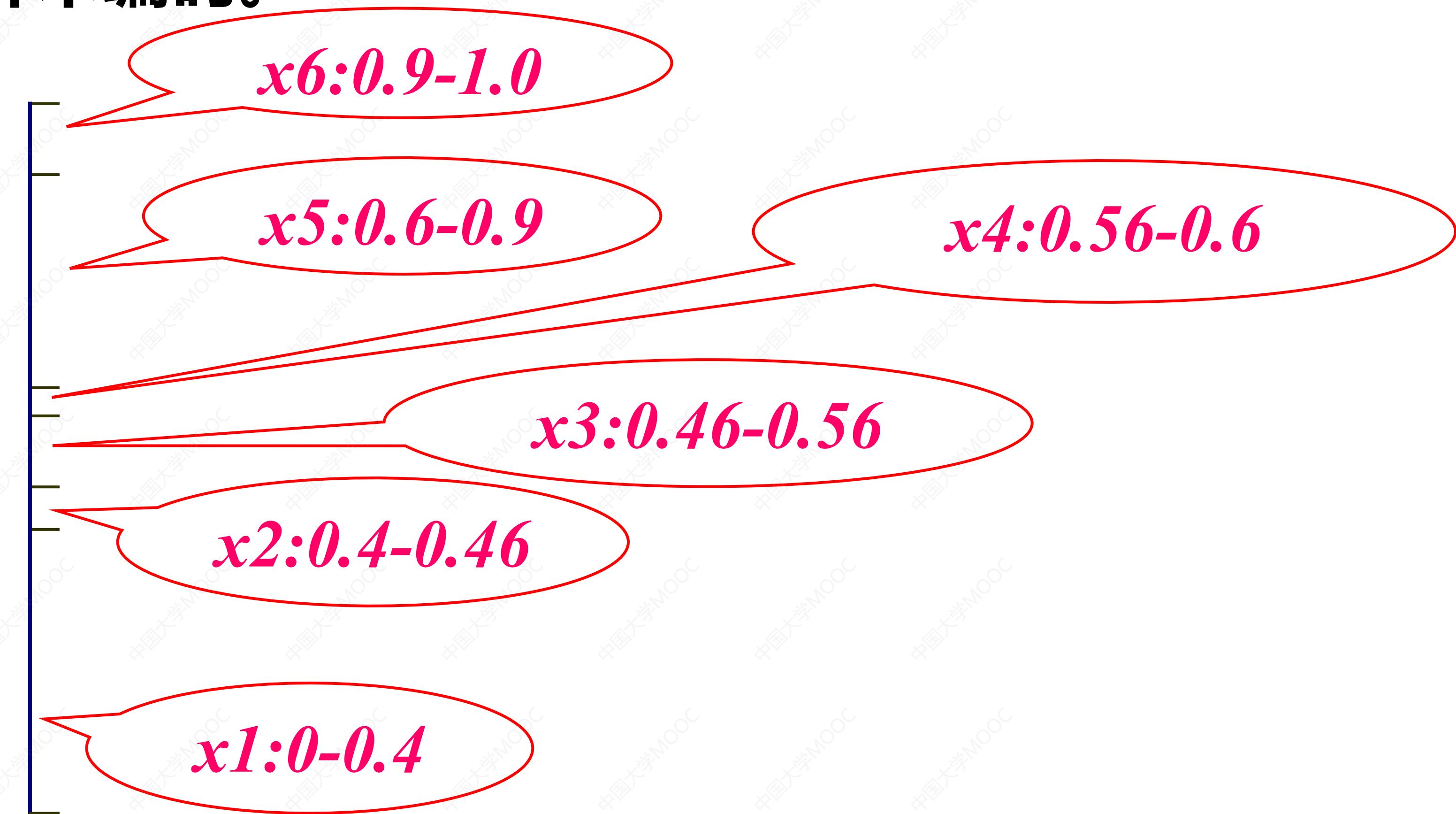
算术编码是给整个信源符号序列分配一个单一的算术码字，这个码字本身定义了一个介于0和1之间的实数间隔。没有码表。

- ✓ 每个符号根据其出现概率减小间隔。
- ✓ 概率越小，则间隔越小，表示间隔所需的信息单元数目变多；概率越大，则间隔越大，表示间隔所需的信息单元数目变少。
- ✓ 随着符号数目的增加，实数间隔变小，则表示间隔所需的信息单元数目变多。

已知

	$x1$	$x2$	$x3$	$x4$	$x5$	$x6$
$p(xi)$	0.4	0.06	0.1	0.04	0.3	0.1

试进行算术编码。



■ 算术编码过程

实数间隔: $Range$,

实数间隔的下边界: $Range_{low}$,

实数间隔的上边界: $Range_{high}$

新符号的上边界: New_{high}

新符号的下边界: New_{low}

算术编码的过程就是一个不断修正实数间隔上 下边界的过程

- 起始实数间隔范围：

$$Range_{low} = 0, Range_{high} = 1, Range = 1$$

- 实数间隔修正公式

$$Range_{low} = Range_{low} + New_{low} \times Range$$

$$Range_{high} = Range_{low} + New_{high} \times Range$$

设符号序列：

$x_1x_2x_5x_6x_3:$

符号 $x_1:0-0.4$

$$New_{low} = 0, New_{high} = 0.4$$

$$Range_{low} = 0 + 0 \times 1 = 0$$

$$Range_{high} = 0 + 0.4 \times 1 = 0.4$$

$$Range = 0.4 - 0 = 0.4$$

设符号序列：

$x_1x_2x_5x_6x_3:$

符号 $x_2: 0.4-0.46$

$$New_{low} = 0.4, New_{high} = 0.46$$

$$Range_{low} = 0 + 0.4 \times 0.4 = 0.16$$

$$Range_{high} = 0 + 0.46 \times 0.4 = 0.184$$

$$Range = 0.184 - 0.16 = 0.024$$

设符号序列：

$x_1x_2x_5x_6x_3$:

符号 $x_5: 0.6-0.9$

$$New_{low} = 0.6, New_{high} = 0.9$$

$$Range_{low} = 0.16 + 0.6 * 0.024 = 0.1744$$

$$Range_{high} = 0.16 + 0.9 * 0.024 = 0.1816$$

$$Range = 0.1816 - 0.1744 = 0.0072$$

设符号序列：

$x_1x_2x_5x_6x_3:$

符号 $x_6:0.9-1.0$

$$New_{low} = 0.9, New_{high} = 1.0$$

$$Range_{low} = 0.1744 + 0.9 * 0.0072 = 0.18088$$

$$Range_{high} = 0.1744 + 1.0 * 0.0072 = 0.1816$$

$$Range = 0.1816 - 0.18088 = 0.00072$$

设符号序列：

$x_1x_2x_5x_6x_3:$

符号 $x_3:0.46-0.56$

$$New_{low} = 0.46, New_{high} = 0.56$$

$$range_{low} = 0.18088 + 0.46 * 0.00072 = 0.1812112$$

$$range_{high} = 0.18088 + 0.56 * 0.00072 = 0.1812832$$

$$Range = 0.1812832 - 0.1812112 = 0.000072$$

编码结果对应实数区间：

$0.1812112 - 0.1812832$

代表字符： $Value = 0.18122$

解码过程是不断修正代表字符的过程：

$Value \in$ 符号范围

$Value = (Value - New_{low}) / New_{range}$

$$Value = 0.18122 \in [0, 0.4)$$

解出符号: $x1$

$$New_{low} = 0, New_{range} = 0.4$$

$$Value = (0.18122 - 0) / 0.4 = 0.45305$$

$$Value = 0.45305 \in [0.4, 0.46)$$

解出符号: $x2$

$$New_{low} = 0.4, New_{range} = 0.06$$

$$Value = (0.45305 - 0.4) / 0.06 = 0.88416$$

$Value=0.88416 \in [0.6,0.9)$

解出符号: $x5$

$New_{low}=0.6, New_{range}=0.3$

$Value=(0.88416-0.6)/0.3=0.9472$

$Value=0.9472 \in [0.9,1.0)$

解出符号: $x6$

$New_{low}=0.9, New_{range}=0.1$

$Value=(0.9472-0.9)/0.1=0.472$

$Value=0.472 \in [0.46,0.56)$

解出符号: x_3

$New_{low}=0.46, New_{range}=0.1$

$Value=(0.472-0.46)/0.1=0.12$

算术编码必须设定终止符

- 需要生成码表
 - Huffman编码
 - Shannon-Fano编码
 - B1码
 - B2码
- 不需要生成码表，但需要生成实数区间
 - 算术编码
- 常用熵编码方法
 - Huffman编码
 - 算术编码

6.3 去相关编码

针对空间和时间冗余去除设计的编码方法

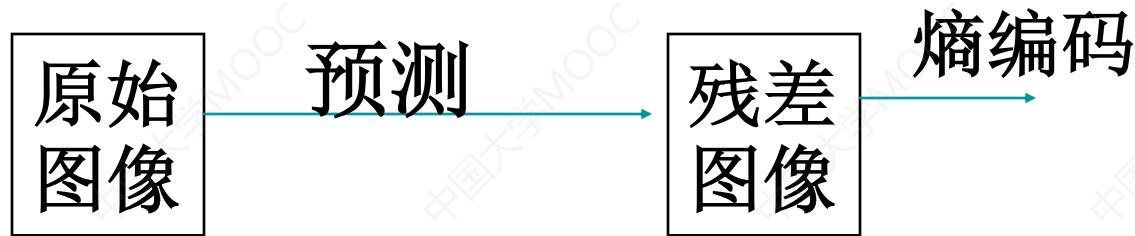
一、预测编码

二、变换编码

三、矢量量化编码

预测编码

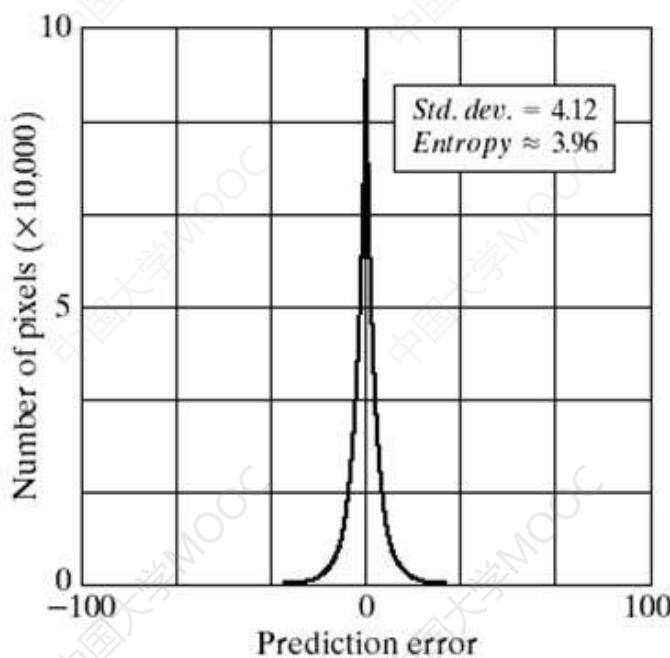
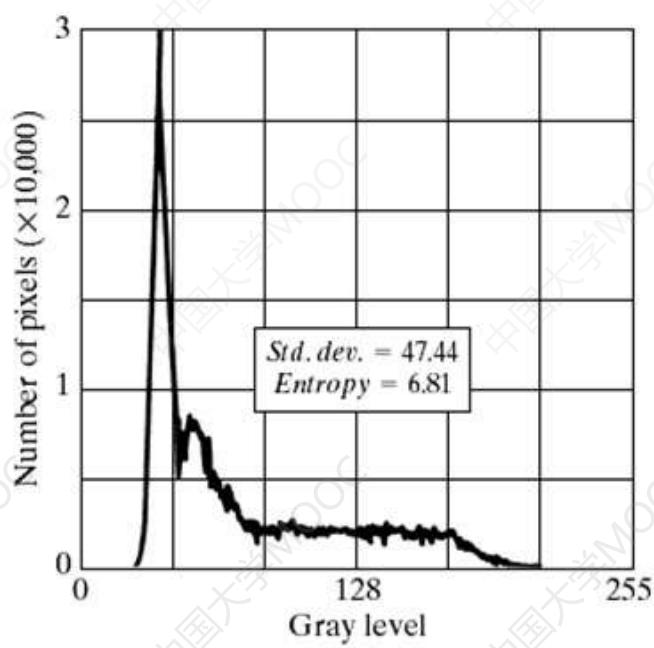
一 基本原理



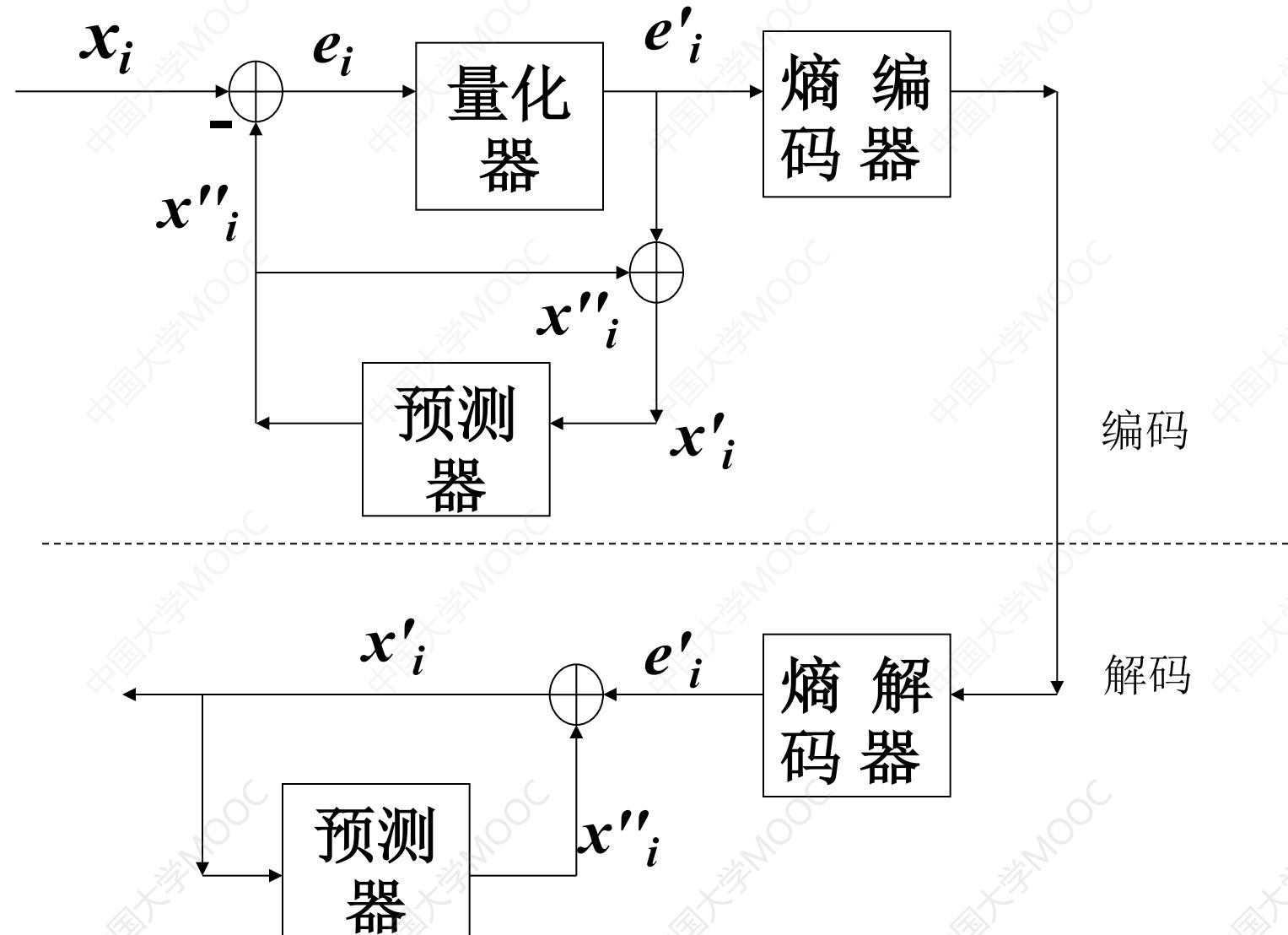
预测编码的好坏直接取决于预测模型。

预测编码分为两种：线性预测和非线性预测。

差分脉冲编码调制DPCM (Differential Pulse Code Modulation),



预测编码原理图



式中: $x_i - x''_i = e_i$

$$x'_i = x''_i + e'_i$$

无量化器: $e'_i = e_i$ 则 $x'_i = x_i$ 无损压缩

有量化器: $e'_i \neq e_i$ 则 $x'_i \neq x_i$ 有损压缩

例下面为一行数据，试进行有损DPCM编码。

编码预测器 $x''_i = x_{i-1}$

编码量化器

$$Q(e) = \begin{cases} -64 & e \leq -64 \\ e & -64 < e < 63 \\ 63 & e \geq 63 \end{cases}$$

给出编码结果序列（误差序列）和解码重构序列

$$X(n) = \{95, 93, 18, 27, 111, 11, 45, 186, 110, 155\}$$

$X(i)$	95	93	18	27	111	11	45	186	110	155
$X'(i)$	0	63	93	29	27	90	26	45	108	110
$e(i)$	95	30	-75	-2	84	-79	19	141	2	45
$e'(i)$	63	30	-64	-2	63	-64	19	63	2	45
$X'(i)$	63	93	29	27	90	26	45	108	110	155

二. 最佳线性预测器设计

常用三阶预测器预测：

$$\hat{f}(x, y) = a_1 f(x, y-1) + a_2 f(x-1, y) + a_3 f(x-1, y-1)$$

最简预测模式： $a_1=1, a_2=a_3=0$ 不一定最优。

何时最优？ 残差图像的分布期望值越小越好，

即 $e(x, y) = f(x, y) - \hat{f}(x, y)$

以 $E\{e^2(x, y)\} = E\{(f(x, y) - \hat{f}(x, y))^2\}$

最小为目标求 a_i

令 $\frac{\partial E\{e^2(x, y)\}}{\partial a_i} = 0$

得到3个独立方程组成的方程组

求解方程组，求出 a_i

三 其他预测方法

自适应预测:同一幅图像采用不同的预测模型

非线性预测

基于模型的预测

变换编码

先用正交基对图象进行变换，然后对变换域系数进行量化和编码

变换的目的

去除空间信号的相关性（空间冗余）

将空间信号的能量集中到频域的一小部分系数

上

能量小的系数可通过量化去除，而不会严重影响重构图像的质量

基于变换的编码模式



K-L变换 (Karhunen–Loeve 变换)

对于向量x用确定的完备正交归一化向量基展开

$$\mathbf{x} = \sum_{j=1}^{\infty} y_j \mathbf{u}_j$$

得到一组向量的协方差矩阵 $R=XX^T$

计算R的特征值和特征向量

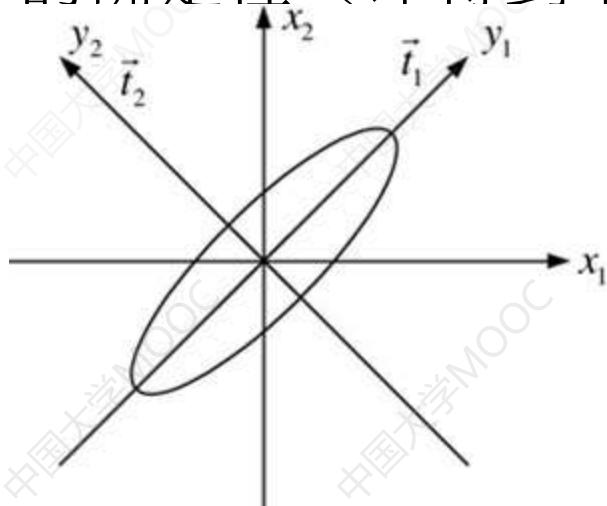
$$RU=AU$$

即可得到基函数向量U

K-L变换消除了原有x向量之间的相关性

K-L变换的特点

- 变换后的特征分量不相关。KL变换后协方差矩阵变为对角阵，消除了原有向量空间的各分量之间的相关性，从而可能去掉带有较少信息的基函数，达到降维的目的
- 最佳逼近性。采用同样维数向量表示原始信号，KL变换的结果与原始信号误差最小
- 能量集中效率最高。KL变换后能量向某些分量相对集中，增强了总体的确定性（即得到主成分）



DCT变换

- DCT变换（离散余弦变换）

- ✓ 二维DCT变换及反变换

$$T(u, v) = a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$u, v = 0, 1, \dots, N - 1$$

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} a(u)a(v) T(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$x, y = 0, 1, \dots, N - 1$$

$$\frac{1}{2} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \cos \frac{\pi}{16} & \cos \frac{3\pi}{16} & \cos \frac{5\pi}{16} & \cos \frac{7\pi}{16} & \cos \frac{9\pi}{16} & \cos \frac{11\pi}{16} & \cos \frac{13\pi}{16} & \cos \frac{15\pi}{16} \\ \cos \frac{2\pi}{16} & \cos \frac{6\pi}{16} & \cos \frac{10\pi}{16} & \cos \frac{14\pi}{16} & \cos \frac{18\pi}{16} & \cos \frac{22\pi}{16} & \cos \frac{26\pi}{16} & \cos \frac{30\pi}{16} \\ \cos \frac{3\pi}{16} & \cos \frac{9\pi}{16} & \cos \frac{15\pi}{16} & \cos \frac{21\pi}{16} & \cos \frac{27\pi}{16} & \cos \frac{33\pi}{16} & \cos \frac{39\pi}{16} & \cos \frac{45\pi}{16} \\ \cos \frac{4\pi}{16} & \cos \frac{12\pi}{16} & \cos \frac{20\pi}{16} & \cos \frac{28\pi}{16} & \cos \frac{36\pi}{16} & \cos \frac{44\pi}{16} & \cos \frac{52\pi}{16} & \cos \frac{60\pi}{16} \\ \cos \frac{5\pi}{16} & \cos \frac{15\pi}{16} & \cos \frac{25\pi}{16} & \cos \frac{35\pi}{16} & \cos \frac{45\pi}{16} & \cos \frac{55\pi}{16} & \cos \frac{65\pi}{16} & \cos \frac{75\pi}{16} \\ \cos \frac{6\pi}{16} & \cos \frac{18\pi}{16} & \cos \frac{30\pi}{16} & \cos \frac{42\pi}{16} & \cos \frac{54\pi}{16} & \cos \frac{66\pi}{16} & \cos \frac{78\pi}{16} & \cos \frac{90\pi}{16} \\ \cos \frac{7\pi}{16} & \cos \frac{21\pi}{16} & \cos \frac{35\pi}{16} & \cos \frac{49\pi}{16} & \cos \frac{63\pi}{16} & \cos \frac{77\pi}{16} & \cos \frac{91\pi}{16} & \cos \frac{105\pi}{16} \end{bmatrix}$$

$$\begin{bmatrix} 0.35356 & 0.35356 & 0.35356 & 0.35356 & 3.35356 & 0.35356 & 0.35356 & 0.35356 \\ 0.49039 & 0.41573 & 0.27779 & 0.09755 & -0.09755 & -0.27779 & -0.41573 & -0.49039 \\ 0.46194 & 0.19134 & -0.19134 & -0.46194 & -0.46194 & -0.19134 & 0.19134 & 0.46194 \\ 0.41573 & -0.09755 & -0.49039 & -0.27779 & 0.27779 & 0.49039 & 0.09755 & -0.41573 \\ 0.35356 & -0.35356 & -0.35356 & 0.35356 & 0.35356 & -0.35356 & -0.35356 & 0.35356 \\ 0.27779 & -0.49039 & 0.09755 & 0.41573 & -0.41573 & -0.09755 & 0.49039 & -0.27779 \\ 0.19134 & -0.46194 & 0.46194 & -0.19134 & -0.19134 & 0.46194 & -0.46194 & 0.19134 \\ 0.09755 & -0.27779 & 0.41573 & -0.49039 & 0.49039 & -0.41573 & 0.27779 & -0.09755 \end{bmatrix}$$

矢量量化编码

1 理论基础——数据相关的信息熵定理

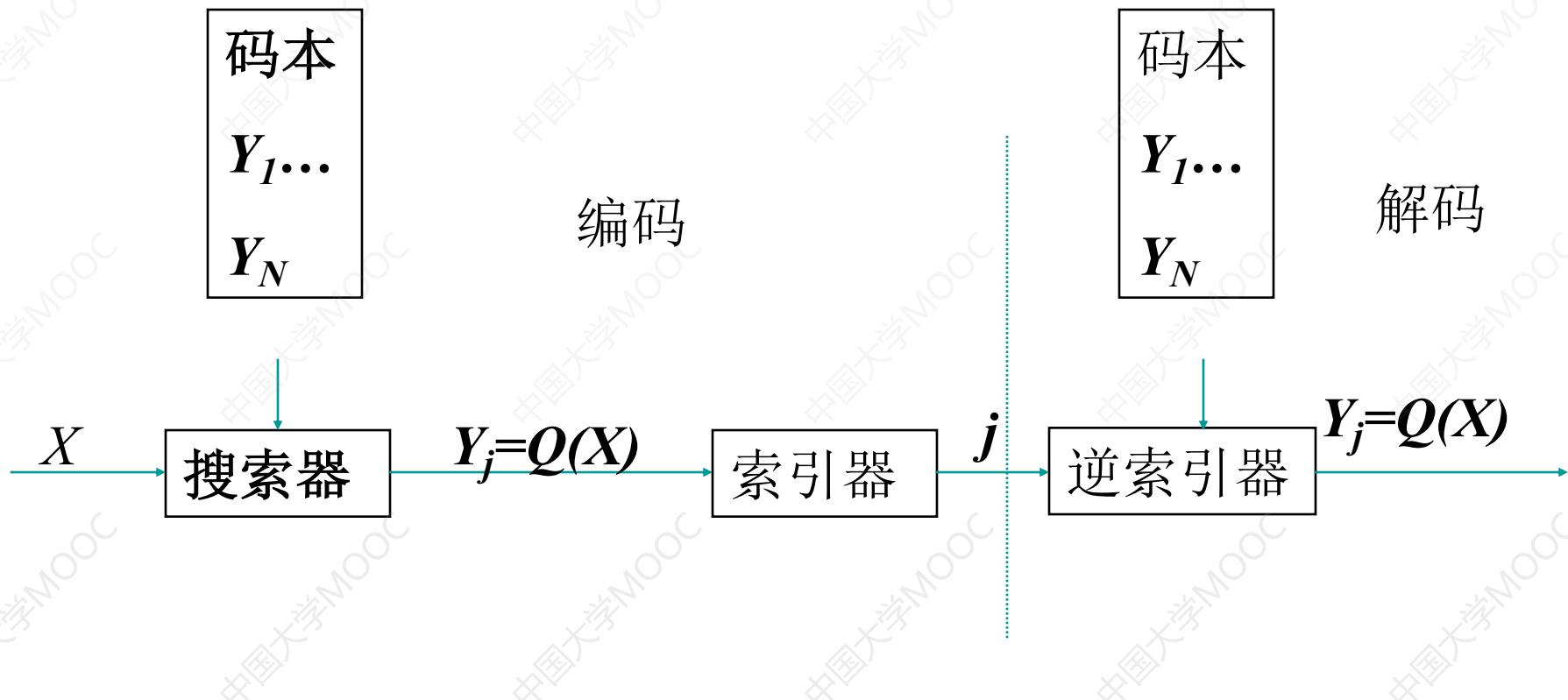
- ✓ 相关阶数K，表明信源中K个元素相关，
 $K=1$ 时不相关， $K=2$ 时相邻两个元素相关。

- ✓ 矢量编码维数n<K时

$$H_n < H_{n-1} \dots < H_2 < H(X)$$

2. 矢量量化编码系统组成

- ✓ 码本：存储N个标准矢量
- ✓ 搜索器：寻找与输入矢量X最接近的码本矢量
- ✓ 索引器：由找到序号j，并输出
- ✓ 逆索引器：由序号j找到码本矢量，并输出



3. 矢量量化的关键问题

- 码本的生成
- 搜索器性能的优化

6.4 图像压缩标准

1 相关压缩标准

JPEG

JPEG-LS

JPEG2000

ISO/ITU

MPEG1

MPEG2

MPEG4 AVC (H.264)

ISO/IEC

AVS1

AVS2

AVS3

中国AVS工作组

H.261

H.263

H.263+

ITU

二JPEG编码标准

JPEG (Joint Photographic Experts Group) 是一个由 ISO和CCITT两个组织机构联合组成的一个图像专家小组。JPEG是一个适用范围很广的静态图像数据压缩标准，既可用于灰度图像又可用于彩色图像。JPEG不仅适于静止图像的压缩，电视图像的帧内图像的压缩编码，也常采用此算法。

标准主要采用了两种基本的压缩算法：

- 采用以离散余弦变换(DCT)为基础的有损压缩算法。(压缩比 10：1以上)
- 采用以预测技术为基础的DPCM无损压缩算法。 (2:1)



图象编码



JPEG数据文件

01011101110101000111

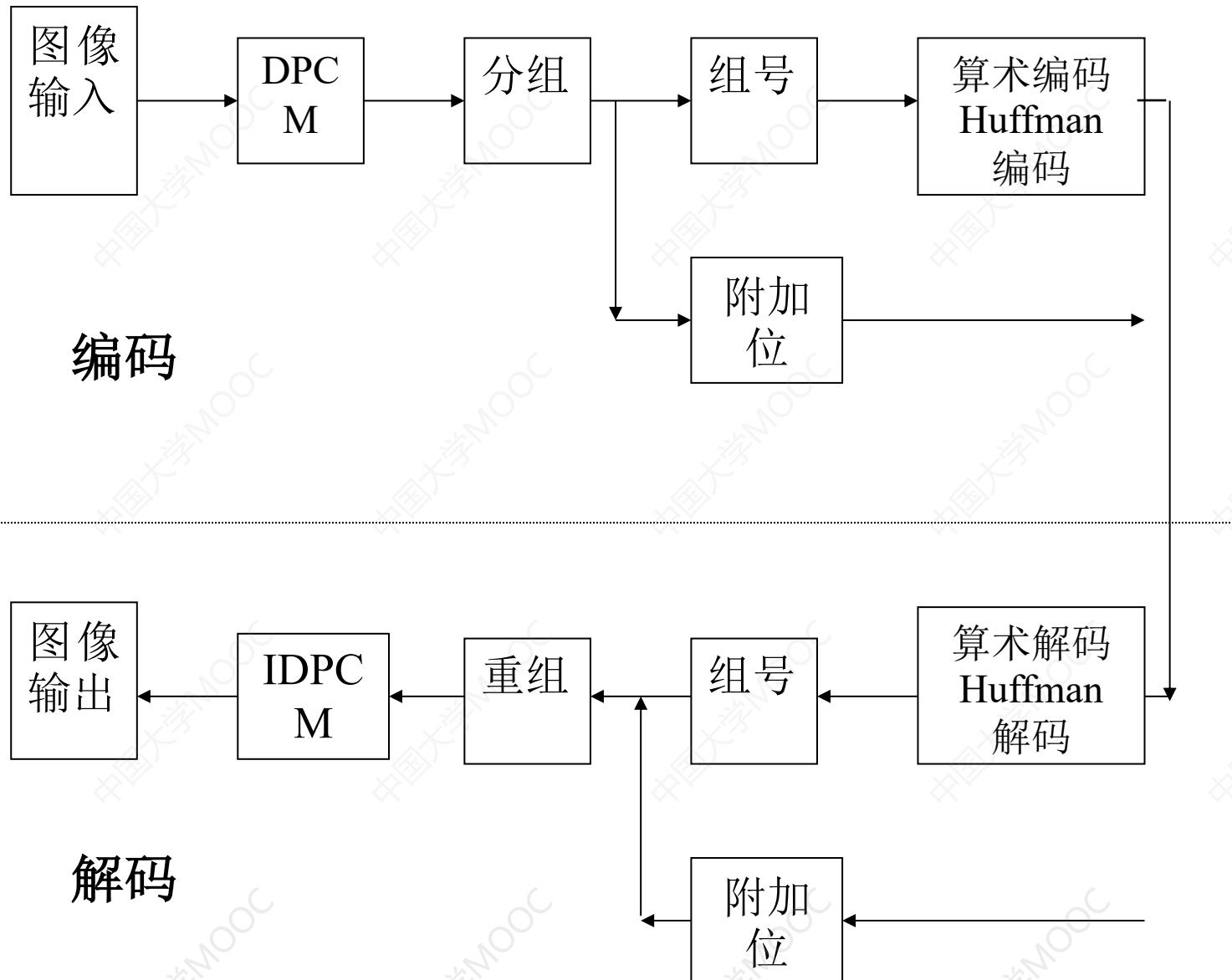
2.JPEG无损压缩标准

核心技术

DPCM

算术编码/Huffman编码

分组技术



DPCM预测器

- 0 无预测
- 1 A
- 2 B
- 3 C
- 4 $A+B-C$
- 5 $A+(B-C)/2$
- 6 $B+(A-C)/2$
- 7 $(A+B)/2$

C B
A Y

第一行固定采用
Y-A预测，第一列
固定采用Y-B预测，
第一行且第一列
点无预测

分组编码

组号	差分值	附加位位数
0	0	0bit
1	-1, 1	1bit: -1:0, 1:1
2	-3, -2, 2, 3	2bit: -3:00, -2:01, 2:10, 3:11
3	-7∽-4, 4∽7	3bit: -7:000, -6:001, -5:010, -4:011 4:100 5: 101 6: 110 7: 111
4	-15∽-8, 8∽15	4bit
5	-31∽-16, 16∽31	5bit
6	-63∽-32, 32∽63	6bit
7	-127∽64, 64∽127	7bit
8	-255∽-128, 128∽255	8bit

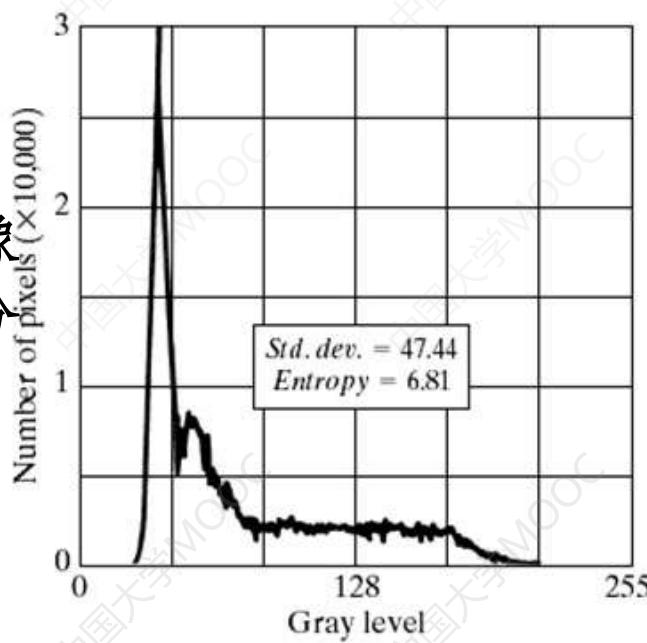
码字=组号+附加位

分组编码

原始
图像



原始图像
的灰度分
布概率



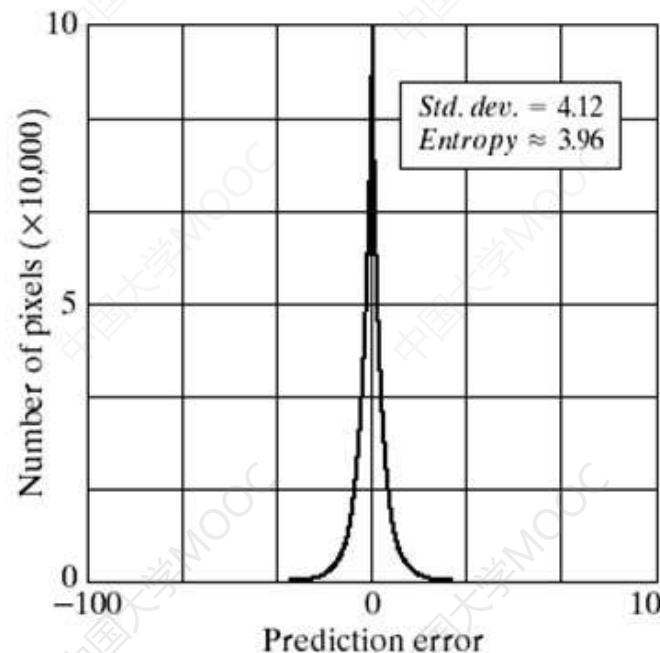
预测模型

$$\hat{f}(x, y) = f(x, y-1)$$

残差
图像



残差图像
的灰度分
布概率



分组编码

组号	差分值	附加位位数
0	0	0bit
1	-1, 1	1bit: -1:0, 1:1
2	-3, -2, 2, 3	2bit: -3:00, -2:01, 2:10, 3:11
3	-7∽-4, 4∽7	3bit: -7:000, -6:001, -5:010, -4:011 4:100 5: 101 6: 110 7: 111
4	-15∽-8, 8∽15	4bit
5	-31∽-16, 16∽31	5bit
6	-63∽-32, 32∽63	6bit
7	-127∽64, 64∽127	7bit
8	-255∽-128, 128∽255	8bit

码字=组号+附加位

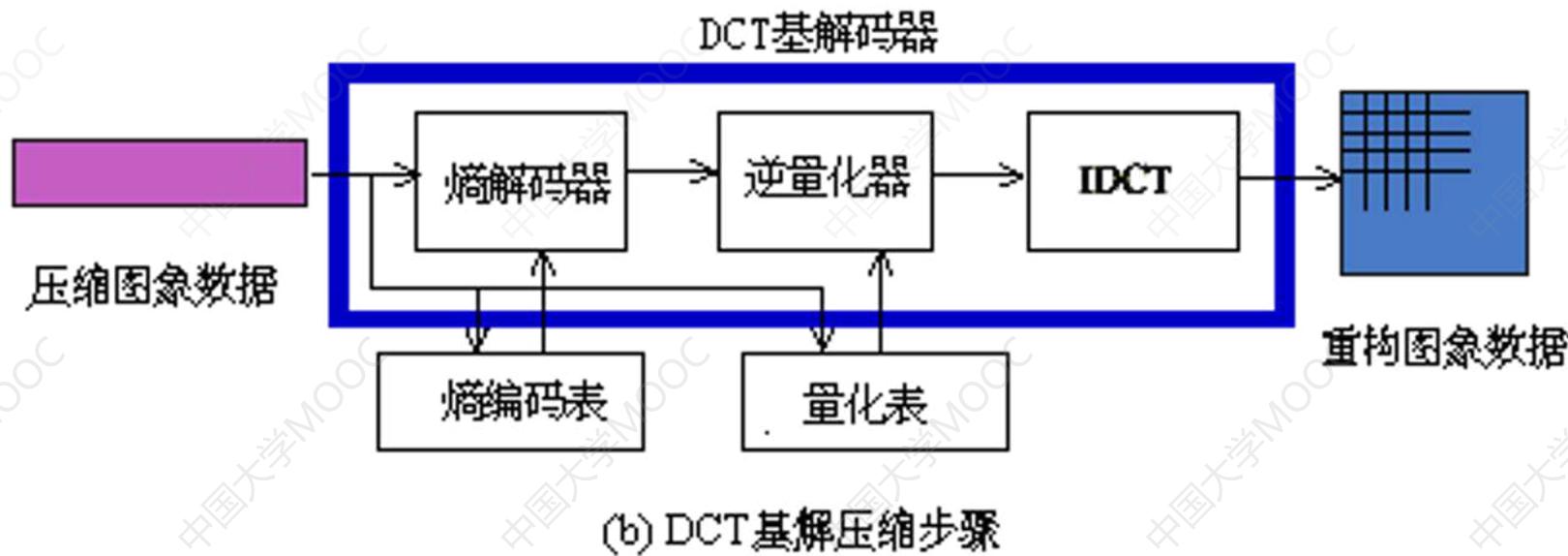
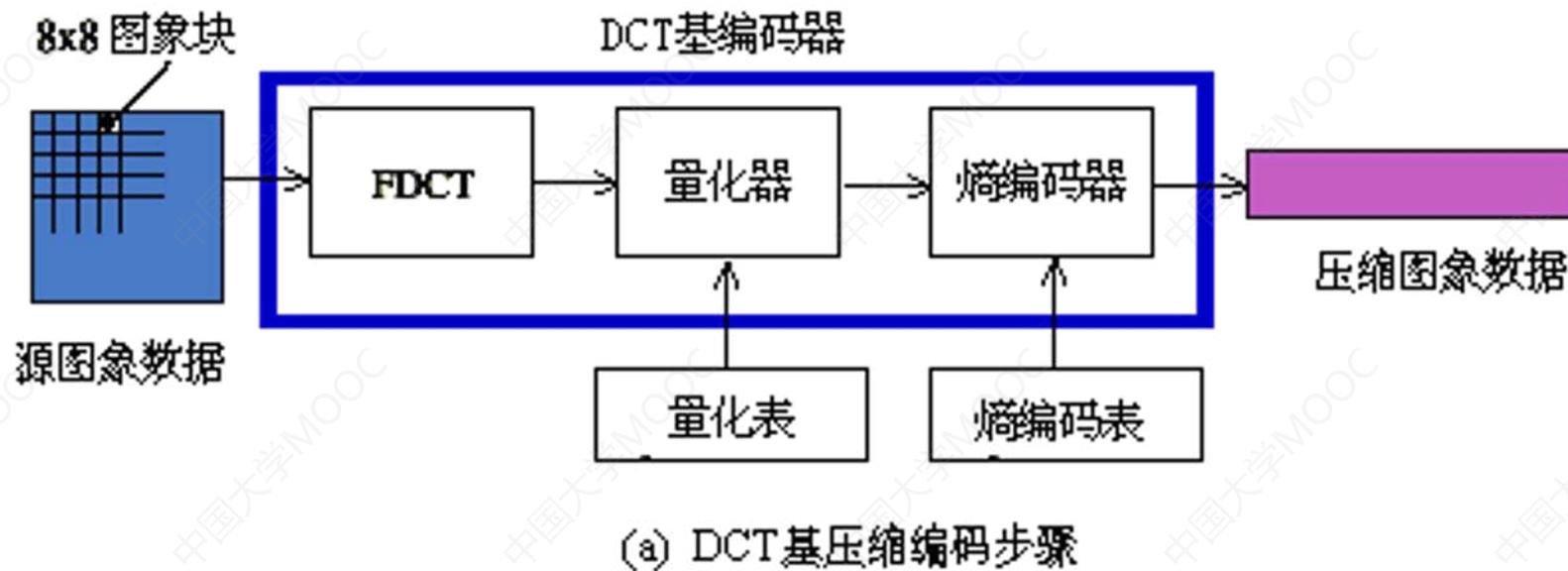
3.JPEG有损压缩标准

› 将彩色由RGB空间转换到YCbCr空间。

› 关键技术

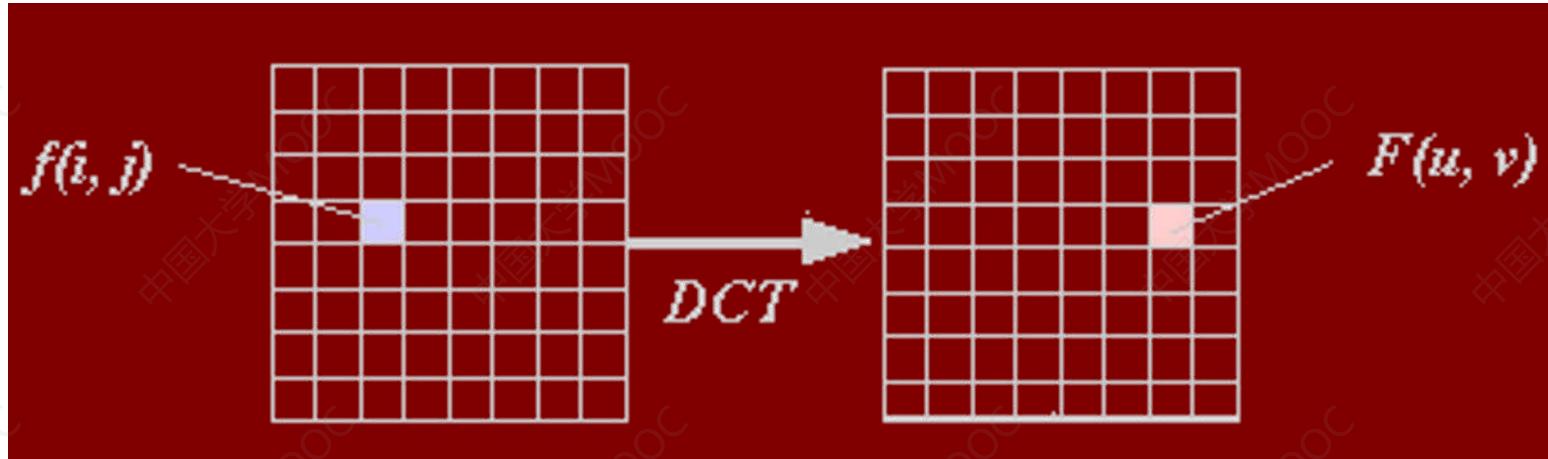
- DCT：分块
- 量化参：参照量化表
- 游程长+分组
- Huffman编码/算术编码

JPEG有损编解码步骤



正向离散余弦变换

对每个单独的彩色图像分量，把整个分量图像分成 8×8 的图像块，如图所示，并作为二维离散余弦变换DCT的输入。通过DCT变换，把能量集中在少数几个系数上。



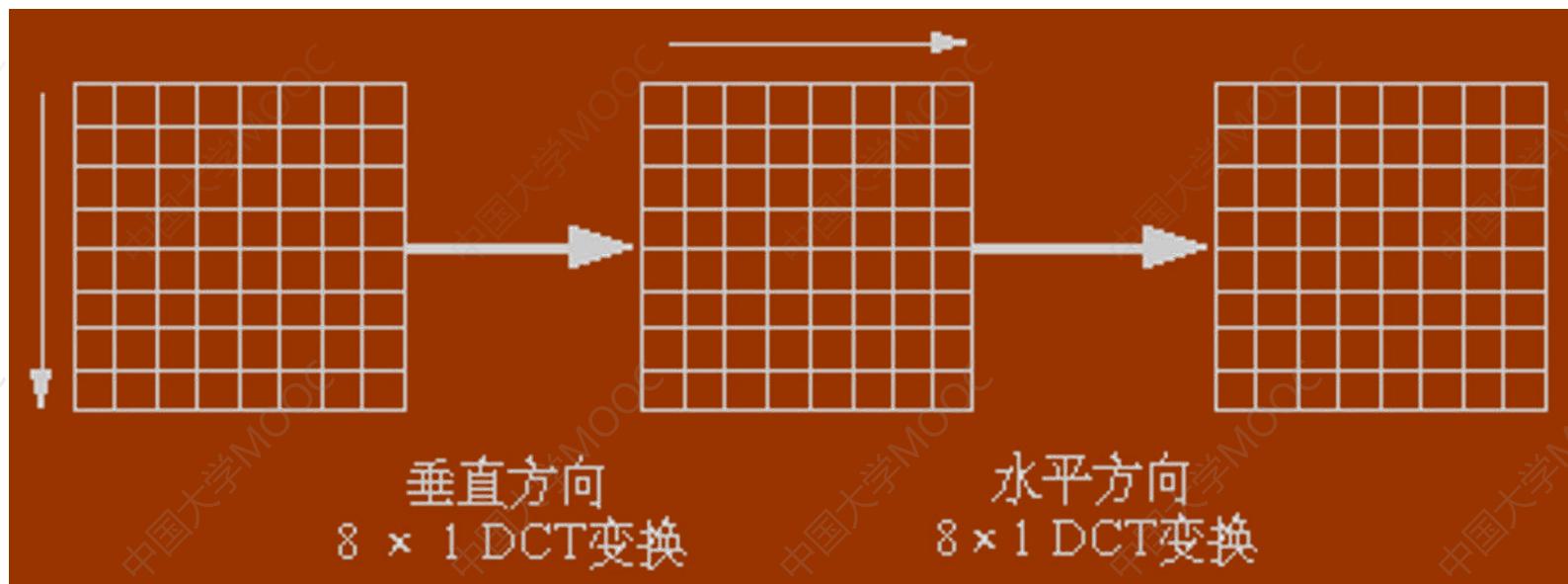
DCT变换使用下式计算

$$F(u, v) = \frac{1}{4} C(u)C(v) \left[\sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \right]$$

在计算两维的DCT变换时，可使用下面的计算式把两维的DCT变换变成一维的DCT变换。

$$F(u, v) = \frac{1}{2} C(u) \left[\sum_{i=0}^7 G(i, v) \cos \frac{(2i+1)u\pi}{16} \right]$$

$$G(i, v) = \frac{1}{2} C(v) \left[\sum_{j=0}^7 f(i, j) \cos \frac{(2j+1)v\pi}{16} \right]$$



量化

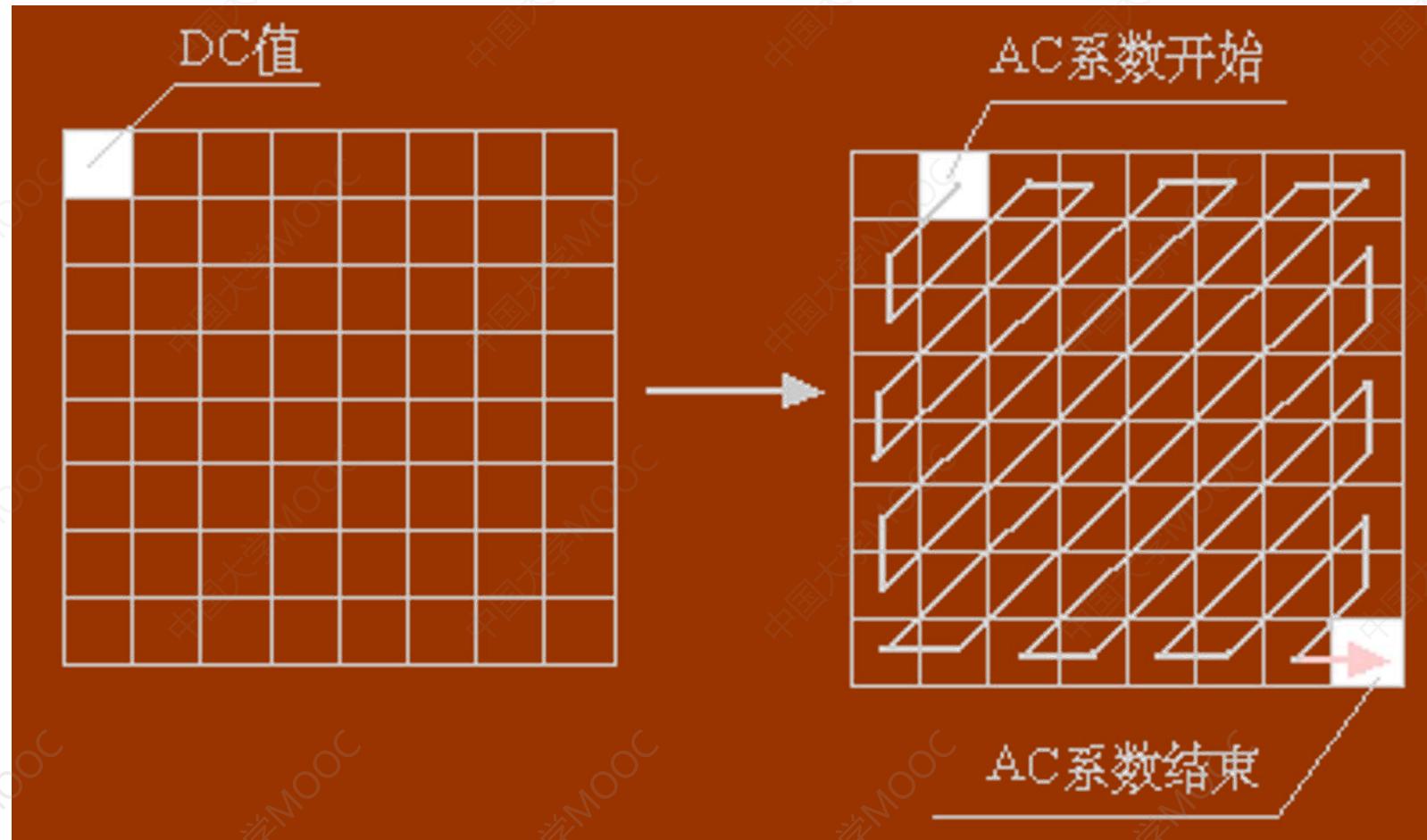
量化是对经过FDCT变换后的频率系数进行量化。量化的目的是减小非“0”系数的幅度以及增加“0”值系数的数目。

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Z字型编码

对图像数据进行展开,为后面的行程编码做准备



Z字编码后的数据顺序

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

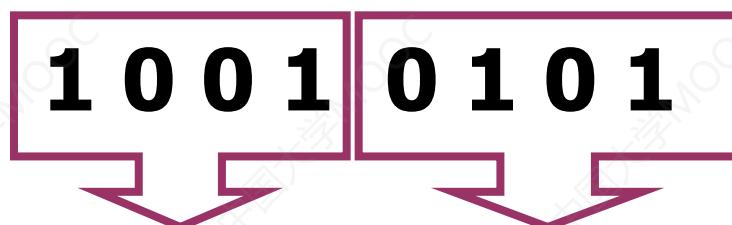
量化DCT系数的序号

直流系数的编码

直流分量数值较大且变化不是很大, **JPEG**算法使用了差分脉冲调制编码(**DPCM**)技术, 对相邻图像块之间量化DC系数的差值进行编码。

交流系数的编码:游程长+分组

在量化后高频分量产生了大量的连零串, 因此使用非常简单和直观的游程长度编码(**RLE**)对它们进行编码。



连续“0”的个数 下一个非“0”系数对应的组号

交流系数的编码:游程长+分组

在量化后高频分量产生了大量的连零串，因此使用非常简单和直观的游程长度编码(RLE)对它们进行编码。

举例，下面一个序列进行游程长+分组

序列：6，5，3，2，0，0，1，0，0，0，0，0，0，0，
0，0，0，0，0，0，0，0，0，0，0，1

编码结果：03，03，02，02，21，f0，31

熵编码

使信息量更接近信息熵，使用霍夫曼编码器的理由是可以使用很简单的查表(lookup table)方法进行编码。

- 直流系数表

0x0002, 0x0023, 0x0033, 0x0043, 0x0053, 0x0063, 0x00e4, 0x01e5, 0x03e6, 0x07e7, 0x0fe8, 0x1fe9

- 交流系数码表

0x000a4,0x00002,0x00012,0x00043,0x000b4,0x001a5,00787,0x00f88,0x03f6a,0x0ff82,0x0ff83
{0x00000,0x000c4,0x001b5,0x00797,0x01f69,0x07f6b,0x0ff84,0x0ff85,0x0ff86,0x0ff87,0x0ff8
{0x00000,0x001c5,0x00f98,0x03f7a,0x0ff4c,0x0ff89,0x0ff8a,0x0ff8b,0x0ff8c,0x0ff8d,0x0ff8e}
{0x00000,0x003a6,0x01f79,0x0ff5c,0x0ff8f,0x0ff90,0x0ff91,0x0ff92,0x0ff93,0x0ff94,0x0ff95}
{0x00000,0x003b6,0x03f8a,0x0ff96,0x0ff97,0x0ff98,0x0ff99,0x0ff9a,0x0ff9b,0x0ff9c,0x0ff9d}
{0x00000,0x007a7,0x07f7b,0x0ff9e,0x0ff9f,0x0ffa0,0x0ffa1,0x0ffa2,0x0ffa3,0x0ffa4,0x0ffa5},
{0x00000,0x007b7,0x0ff6c,0x0ffa6,0x0ffa7,0x0ffa8,0x0ffa9,0x0ffaa,0x0ffab,0x0ffac,0x0ffad}
{0x00000,0x00fa8,0x0ff7c,0x0ffae,0x0ffaf,0x0ffb0,0x0ffb1,0x0ffb2,0x0ffb3,0x0ffb4,0x0ffb5},
{0x00000,0x01f89,0x0ff81,0x0ffb6,0x0ffb7,0x0ffb8,0x0ffb9,0x0ffba,0x0ffbb,0x0ffbc,0x0ffbd},
{0x00000,0x01f99,0x0ffbe,0x0ffb0,0x0ffc0,0x0ffc1,0x0ffc2,0x0ffc3,0x0ffc4,0x0ffc5,0x0ffc6},
{0x00000,0x01fa9,0x0ffc7,0x0ffc8,0x0ffc9,0x0ffca,0x0ffcb,0x0ffcc,0x0ffcd,0x0ffce,0x0ffcf},
{0x00000,0x03f9a,0x0ffd0,0x0ffd1,0x0ffd2,0x0ffd3,0x0ffd4,0x0ffd5,0x0ffd6,0x0ffd7,0x0ffd8},
{0x00000,0x03faa,0x0ffd9,0x0ffda,0x0ffdb,0x0ffdc,0x0ffdd,0x0ffde,0x0ffd0,0x0ffe0,0x0ffe1},
{0x00000,0x07f8b,0x0ffe2,0x0ffe3,0x0ffe4,0x0ffe5,0x0ffe6,0x0ffe7,0x0ffe8,0x0ffe9,0x0ffea},
{0x00000,0x0ffeb,0x0ffec,0x0ffed,0x0ffee,0x0ffef,0x0fff0,0x0fff1,0x0fff2,0x0fff3,0x0fff4}
{0x07f9b,0x0fff5,0x0fff6,0x0fff7,0x0fff8,0x0fff9,0x0ffffa,0x0ffffb,0x0ffffc,0x0ffffd,0x0ffffe} };

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	160	159	159	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

源图象样本

235.6	-1.0	-12.1	-5.20	2.1	-1.7	-2.7	1.3
-22.6	-18.5	-6.2	-3.2	-2.9	-0.1	0.4	-1.2
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3
1.8	-0.2	-1.6	-0.3	-0.8	1.5	1.0	-1.0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8
-2.6	1.6	-3.8	-1.8	1.9	1.2	-0.6	-0.4

FDCT 系数

15, 0, -2, -1, -1, -1, 0, 0, -1, 0, 0

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	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
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

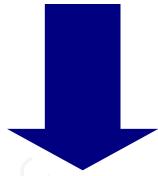
量化表

15, 0, -2, -1, -1, -1, 0, 0, -1, 0, 0

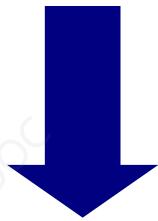
DC (3) (5), (1, 2) (-2),
(0, 1) (-1), (0, 1) (-1),
(0, 1) (-1), (2, 1) (-1),
(EOD)

DPCM和游程编码

其中一种较特殊的情况是出现大于**16**个的连零的情况这时用**(15, 0)**来表示



查熵编码表



写入**JPEG**流



生成**JPEG**文件

JPEG图像文件格式

JEPG文件大体上可以分成两个部分：标记码（tag）和压缩数据。标记码部分给出了JPEG图像的所有信息，如图像的宽、高、Huffman表、量化表等等。

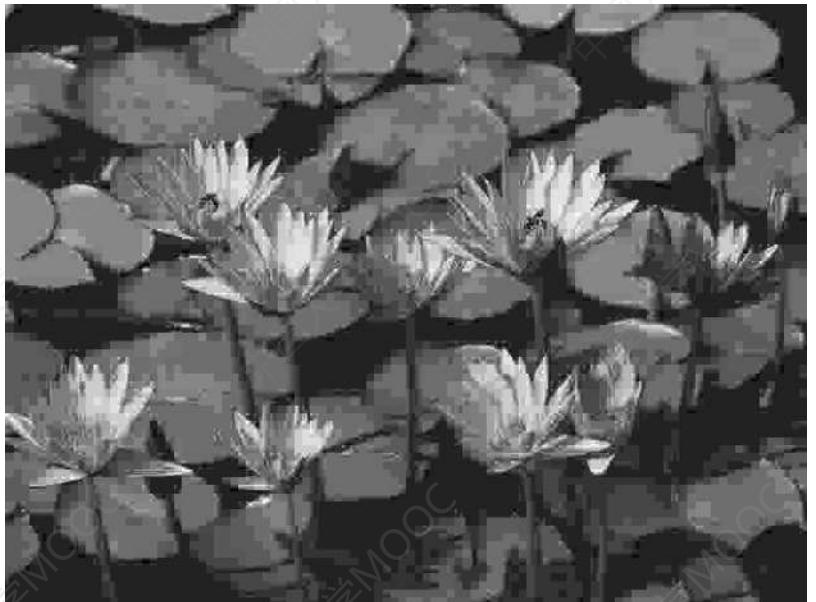
SOI	0xD8	图像开始
APP0	0xE0	JFIF应用数据块
APPn	0xE1 - 0xEF	其他的应用数据块(n, 1~15)
DQT	0xDB	量化表
SOF0	0xC0	帧开始
DHT	0xC4	霍夫曼(Huffman)表
SOS	0xDA	扫描线开始
EOI	0xD9	图像结束



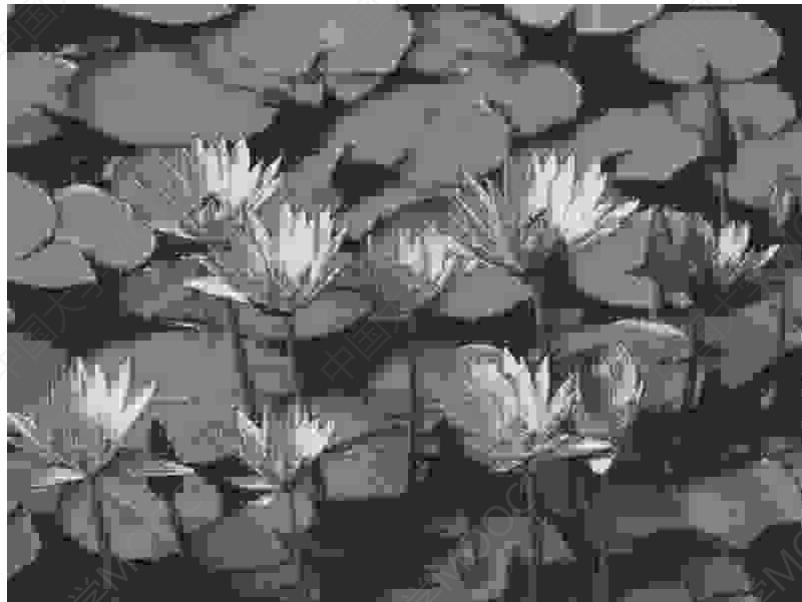
原始图象



10:1



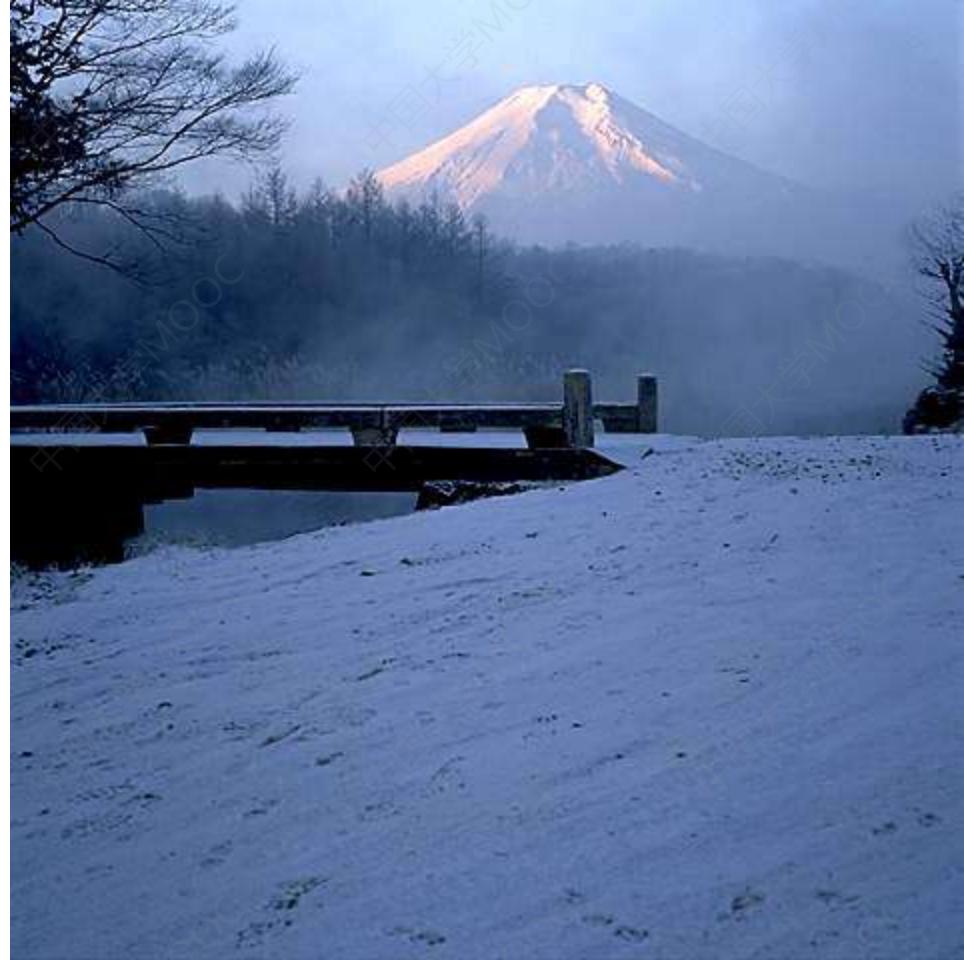
25:1



32:1



原始图象



压缩后的图象，26:1



原始图象



压缩后的图象， 7.8:1

像

■ 引言

图像分割：把图像分成若干个特定的、具有独特性质的区域的技术。

图像分割是由狭义图像处理到图像分析理解的关键步骤。



■ 引言

➤ 三类图像分割技术

➤ 图像分割

- 将分属不同物体的像素区域分开
- 如前景与背景，狗的区域、猫的区域、背景等

➤ 语义分割

- 在普通分割的基础上，分类出每一块区域的语义（即这块区域是什么物体）
- 如把画面中的所有物体都指出它们各自的类别

➤ 实例分割

- 在语义分割的基础上，给每个物体编号
- 如这个是该画面中的狗A，那个是画面中的狗B

■ 引言

五类图像分割方法：

基于阈值的分割方法

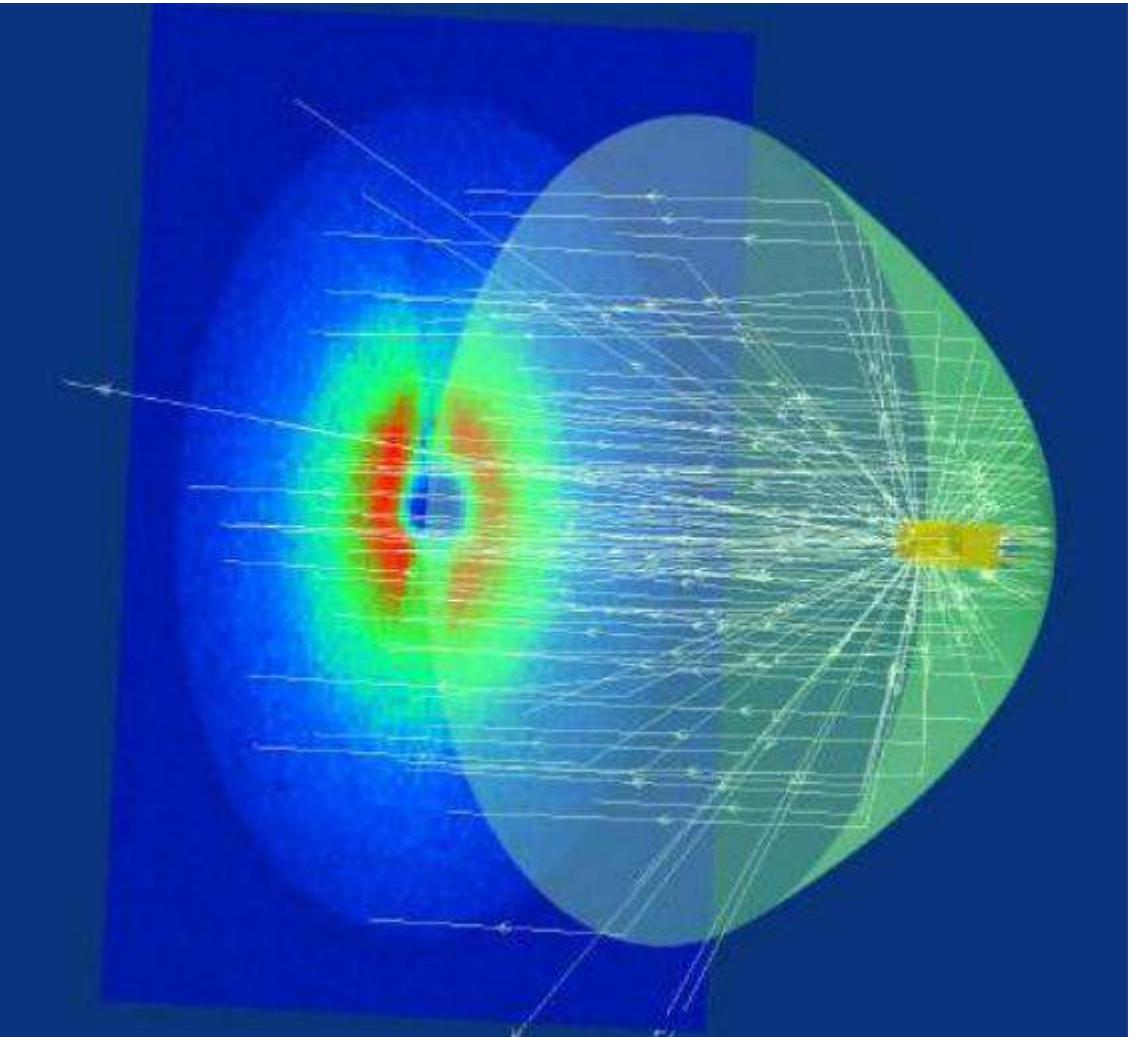
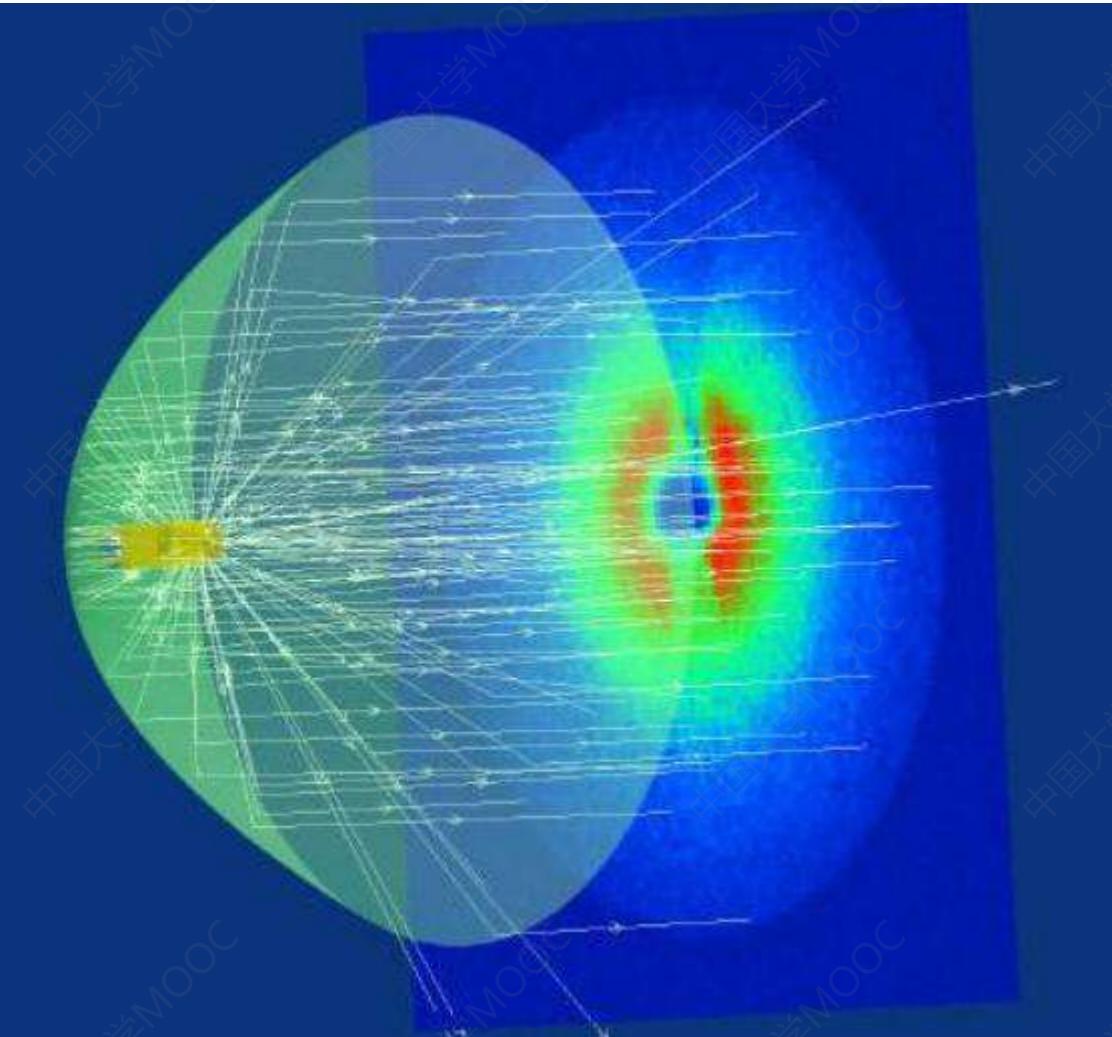
基于边缘的分割方法

基于区域的分割方法

基于特定理论的分割方法

基于深度学习的分割方法

7.1 基于阈值的分割方法



■ 7.1 基于阈值的分割方法

➤ 基本原理

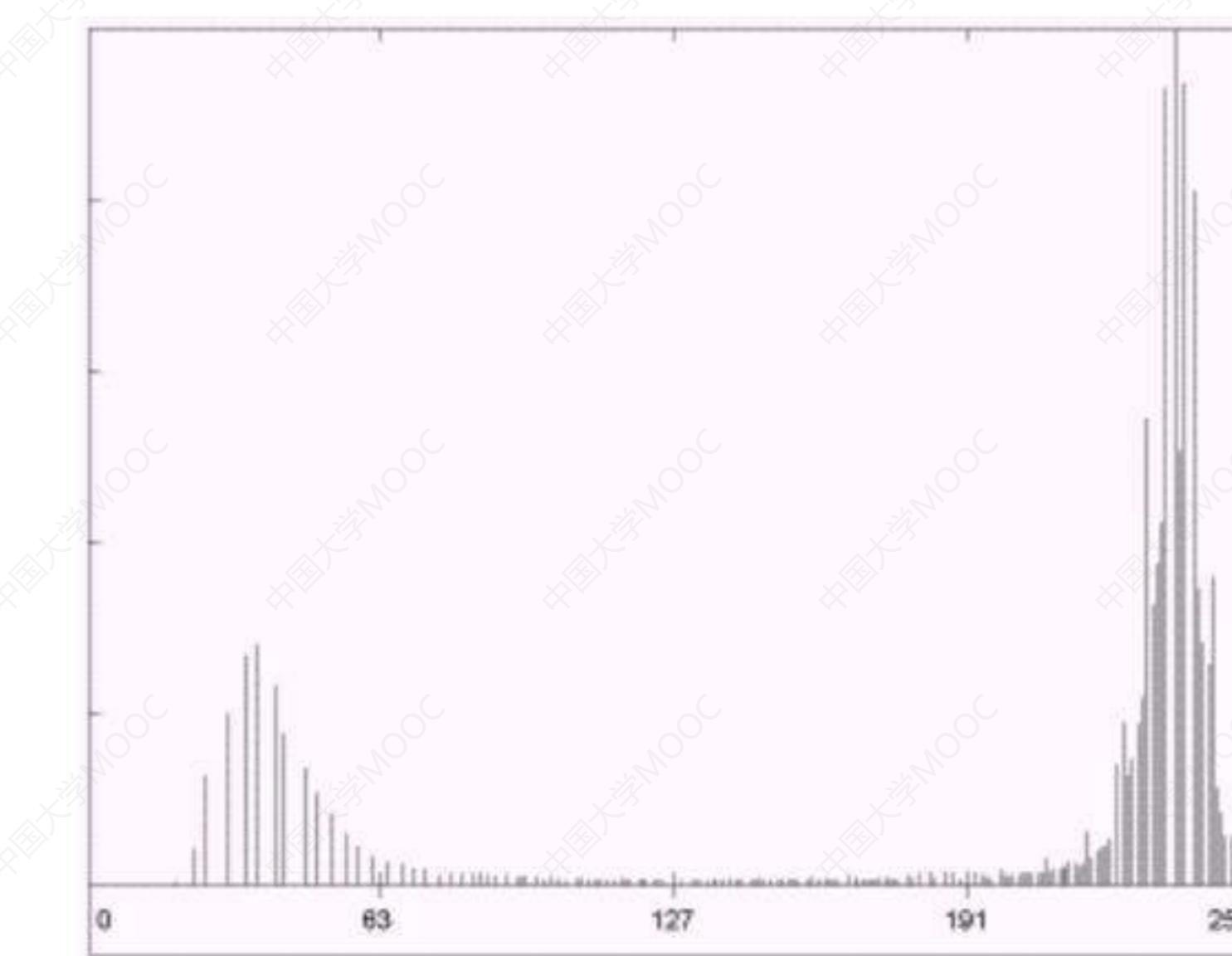
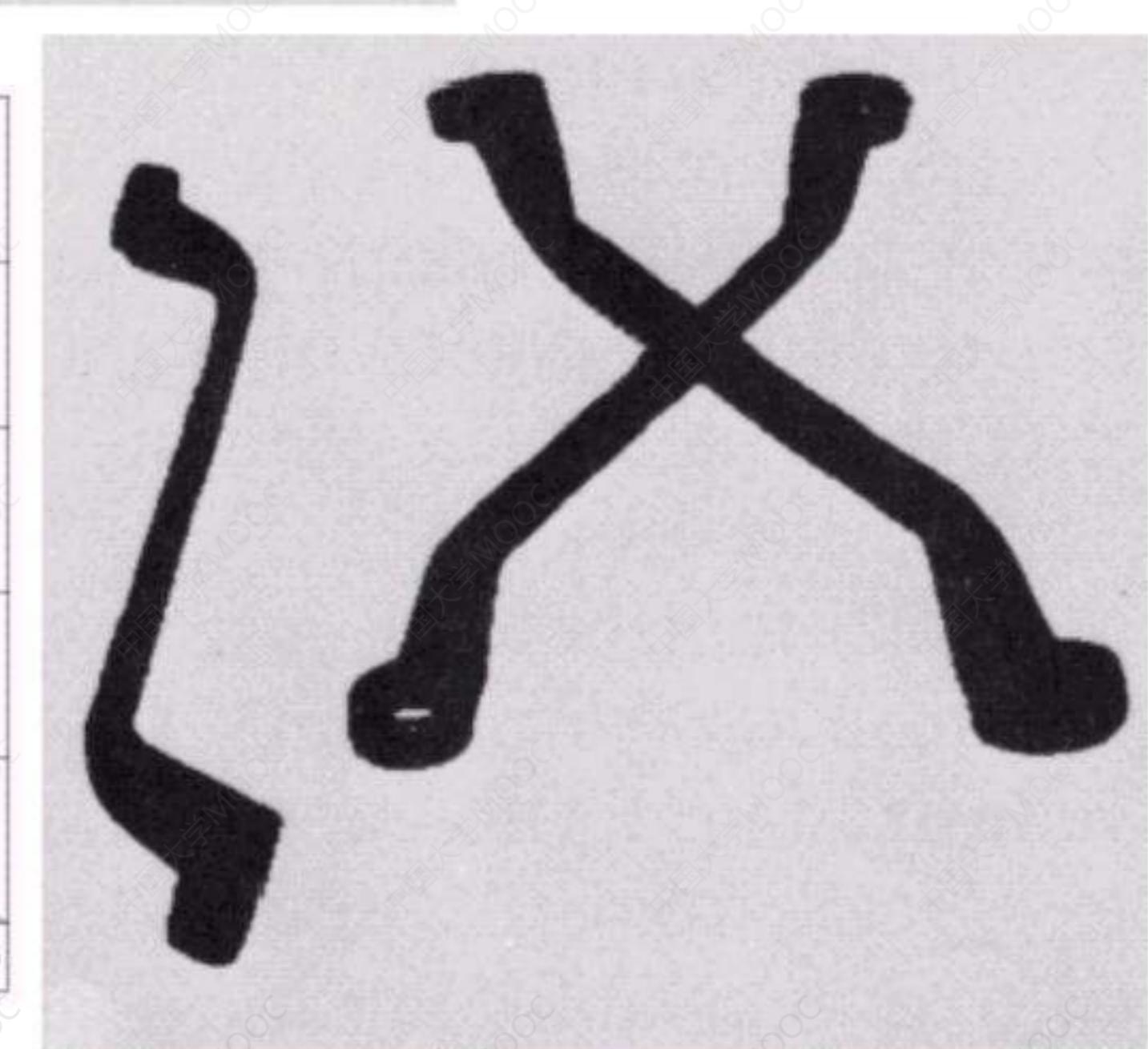
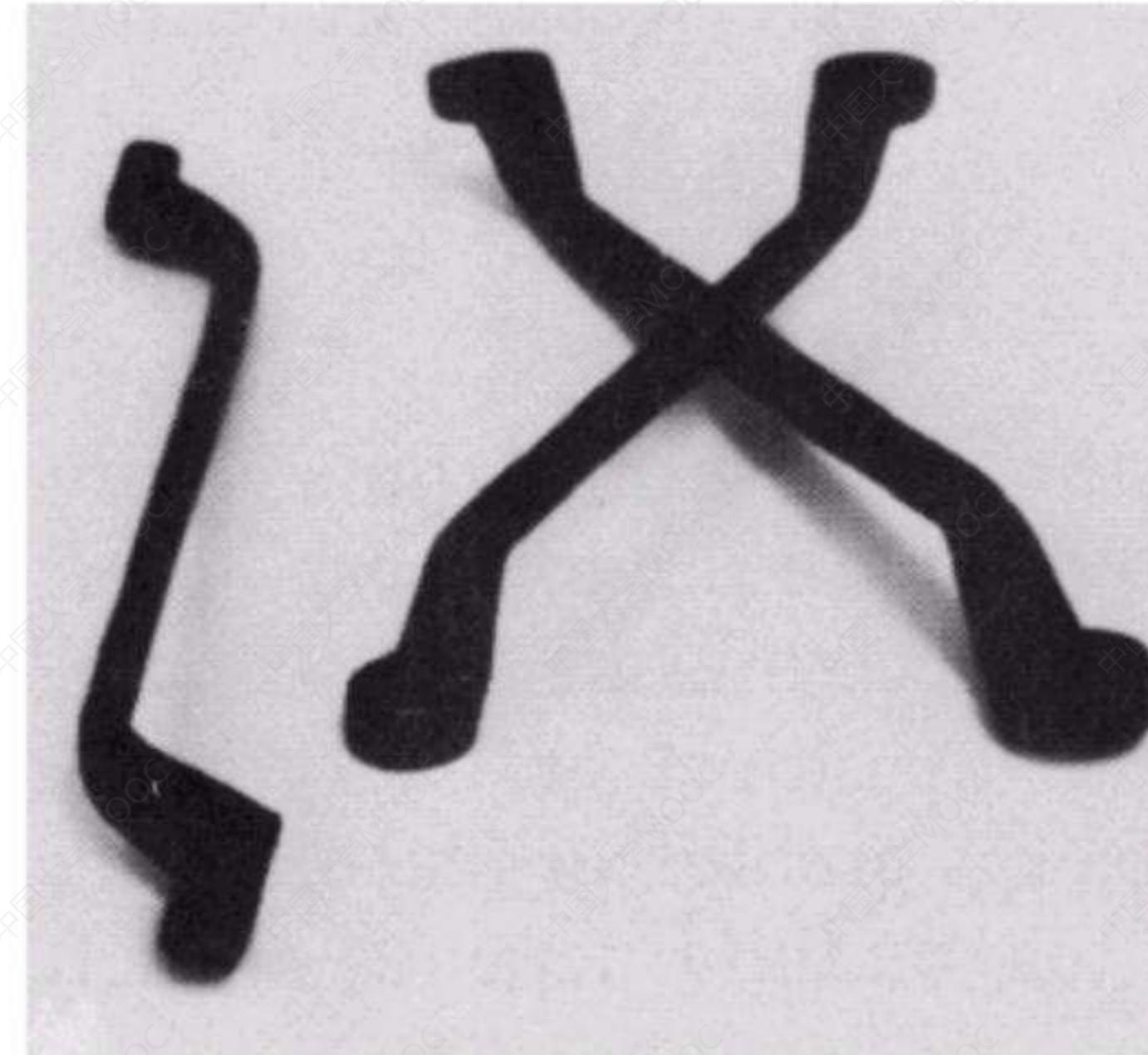
- 前景和背景亮度差别较大，故采用阈值法

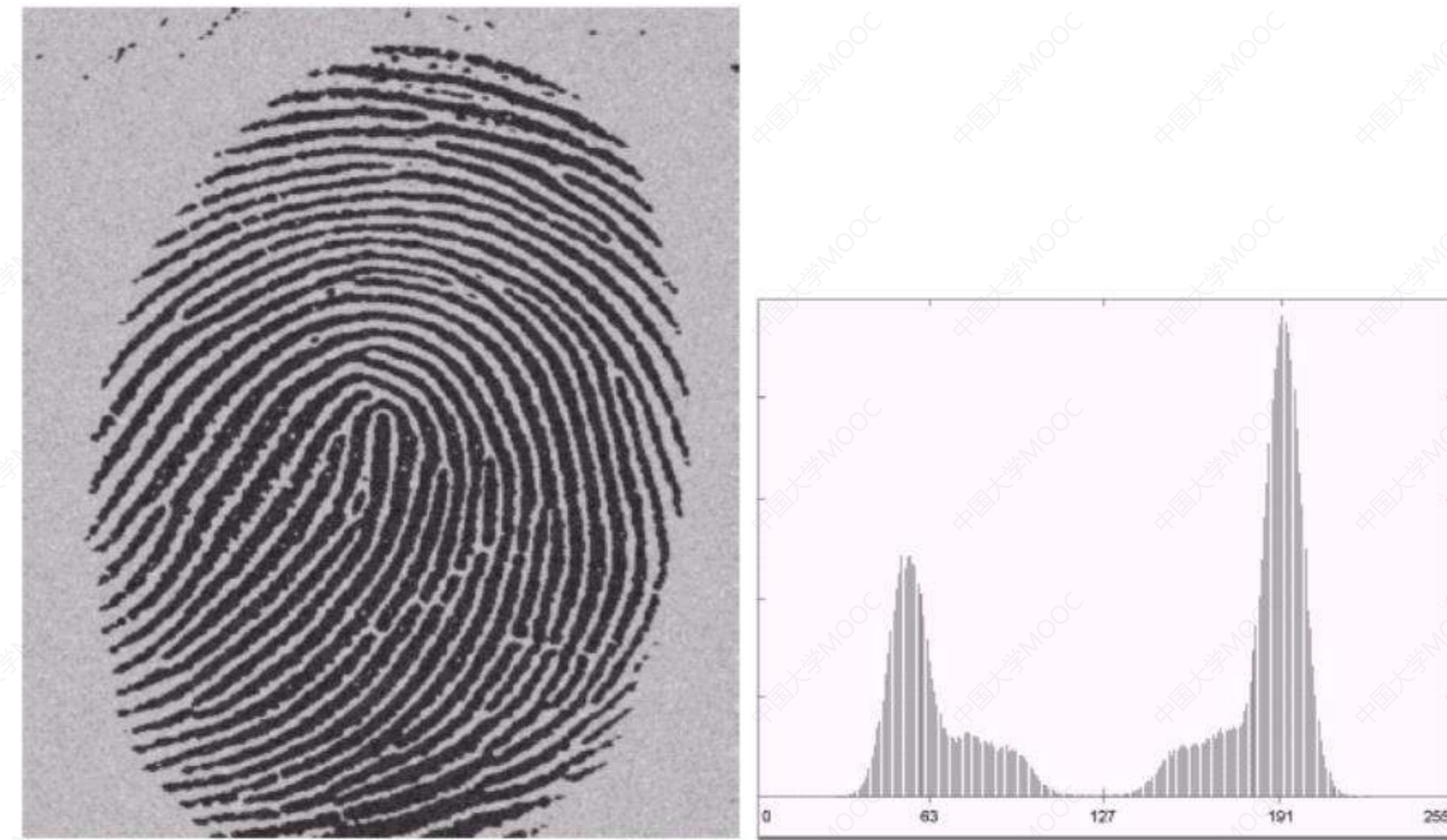
➤ 阈值的选择

- 全局阈值
- 自适应阈值
- 最优化阈值

a
b c

FIGURE 10.28
(a) Original image. (b) Image histogram.
(c) Result of global thresholding with T midway between the maximum and minimum gray levels.



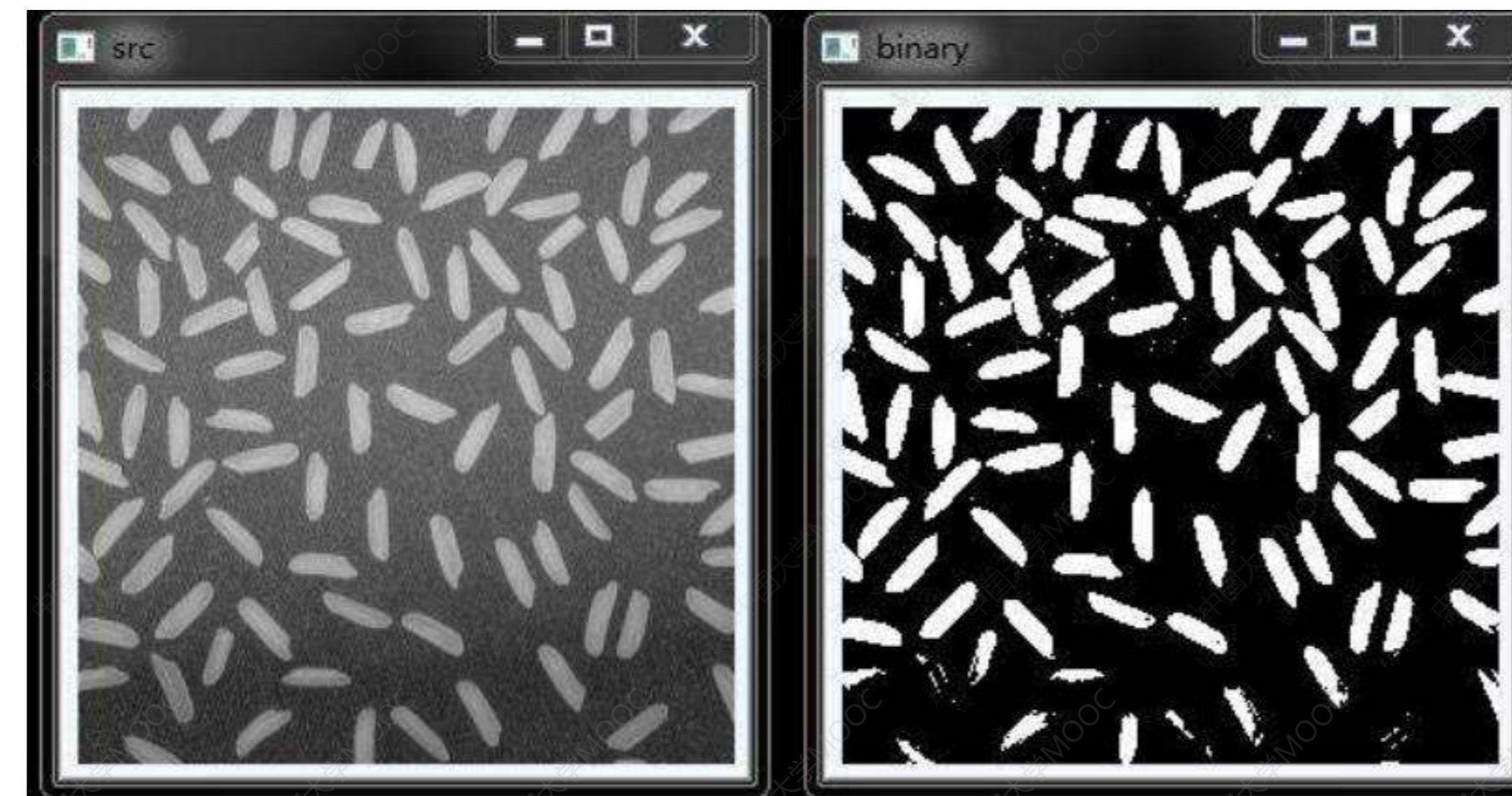
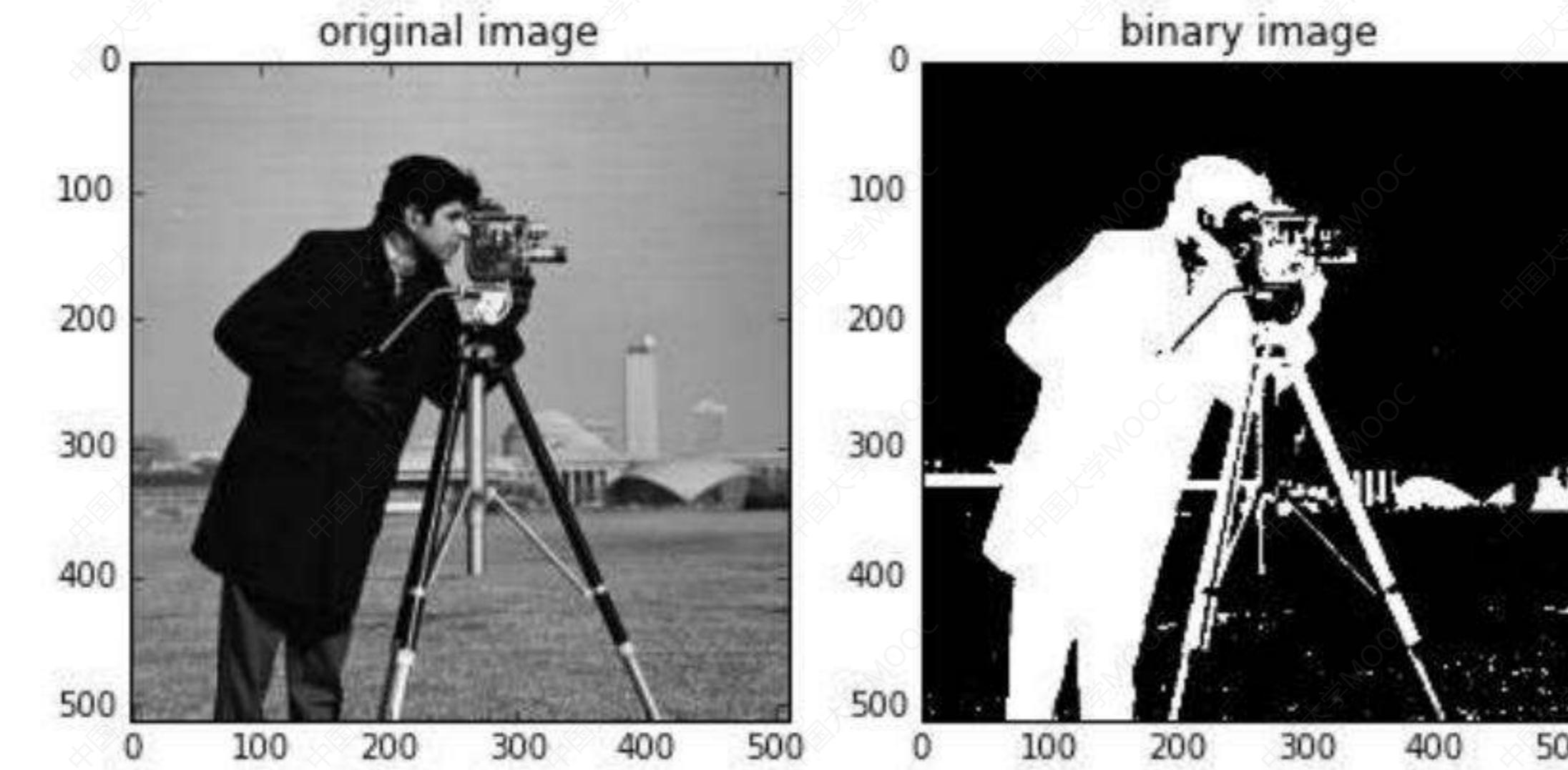


a
b
c

FIGURE 10.29

(a) Original image. (b) Image histogram.
(c) Result of segmentation with the threshold estimated by iteration.
(Original courtesy of the National Institute of Standards and Technology.)

7.1 基于阈值的分割方法



■ 7.1 基于阈值的分割方法

➤ 自适应阈值确定方法——OSTU法（大津法）

Ostu假设图像是由前景区域和背景区域两部分组成的，通过遍历计算不同阈值（通常为[0 255]区间范围内）下分割结果中前景区域和背景区域的灰度直方图，然后比较两者之间的方差，使得方差最大化的那个灰度阈值即为所求二值化阈值。

■ 7.1 基于阈值的分割方法

w_0 前景点所占比例

$w_1 = 1 - w_0$ 背景点所占比例

u_0 = 前景灰度均值

u_1 = 背景灰度均值

$u = w_0 * u_0 + w_1 * u_1$ 全局灰度均值

$g = w_0(u_0 - u)^*(u_0 - u) + w_1(u_1 - u)^*(u_1 - u)$

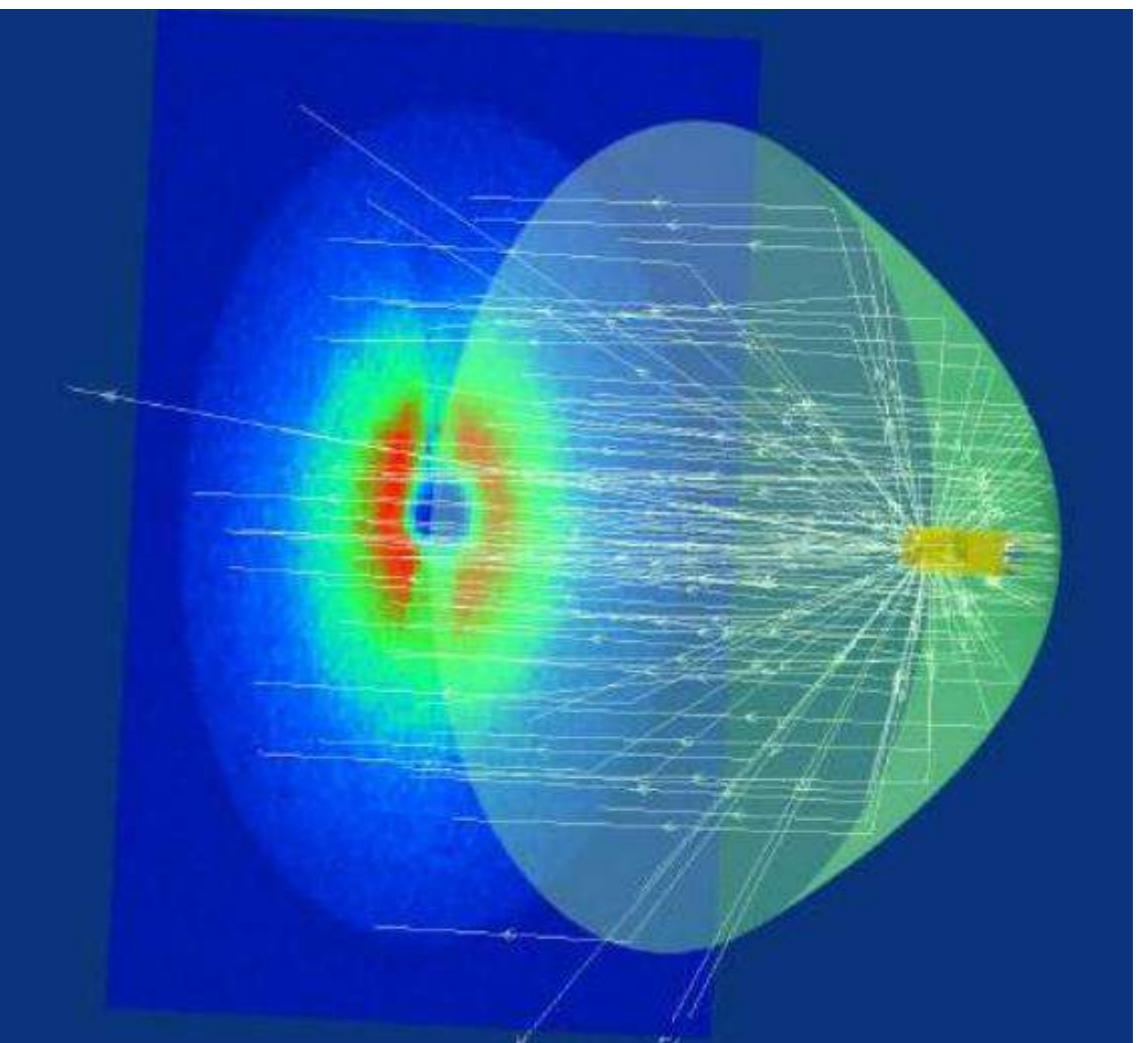
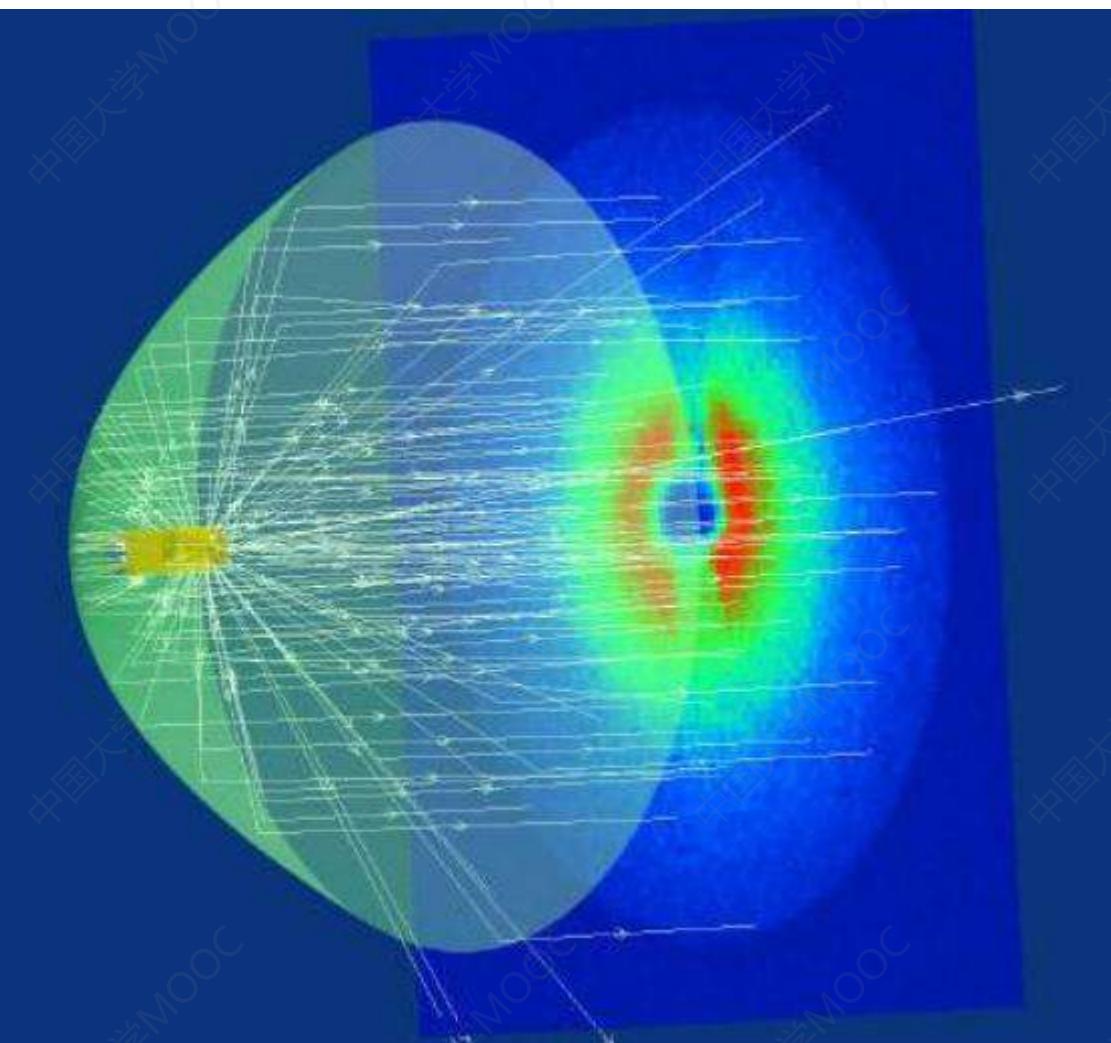
目标函数为 g , 反映了前景和背景的差值, 目标是 g 越大越好。

■ 7.1 基于阈值的分割方法

优点：速度快

缺点： 1) 对噪音和目标大小十分敏感，它仅对类间方差为单峰的图像产生较好的分割效果
2) 当目标与背景的大小比例悬殊时效果不好

7.2 基于边缘的分割方法



■ 7.2 基于边缘的分割方法

- 利用前景背景之间的不连续性，灰度变化位置就是前景轮廓。
- 首先通过边缘检测获取前景和背景之间的边缘，进一步由边缘获取轮廓。



■ 7.2 基于边缘的分割方法

➤ 局部处理

- 梯度强度 + 梯度方向

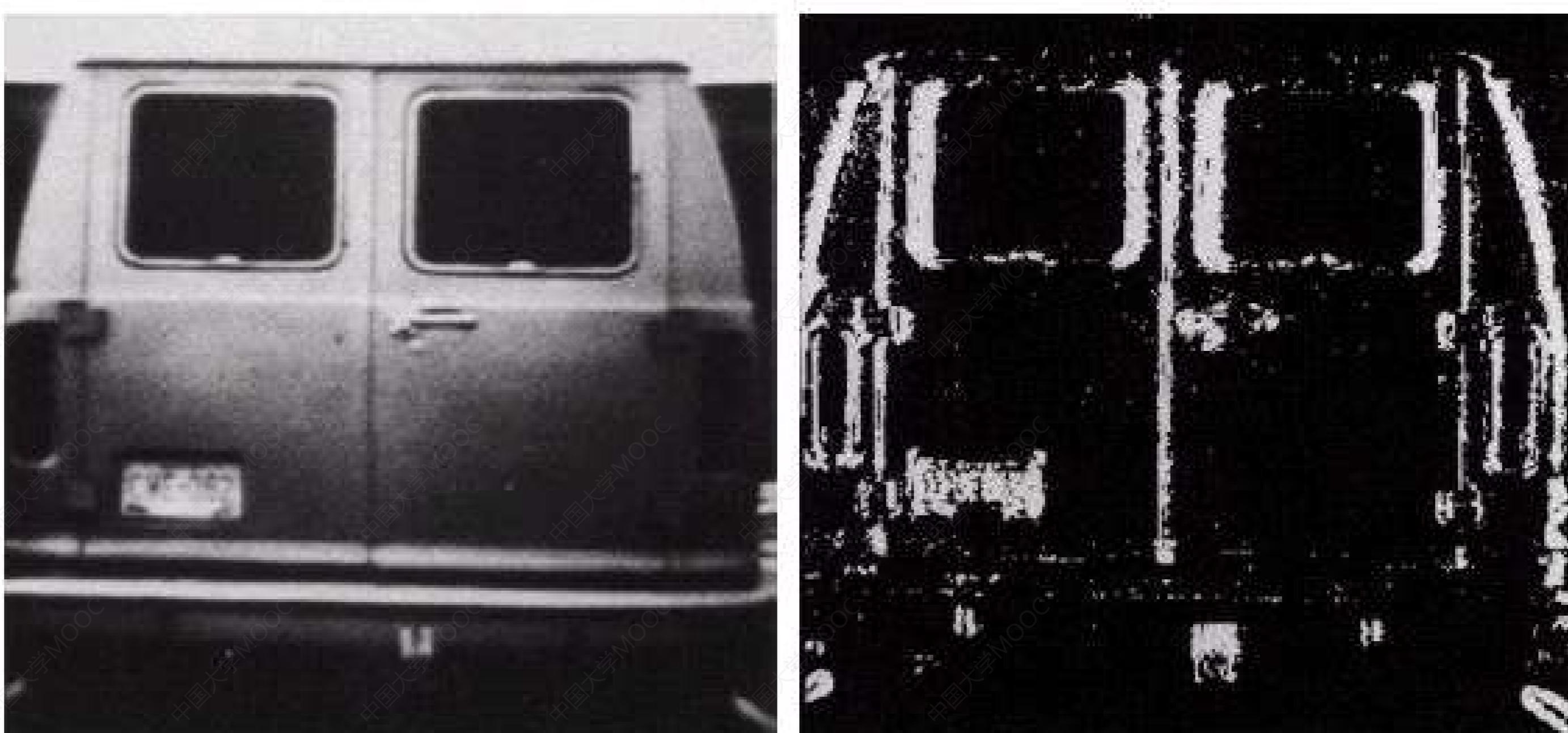
➤ 全局处理

- 基于Hough 变换的方法
- 基于曲线拟合的方法
- 其他方法

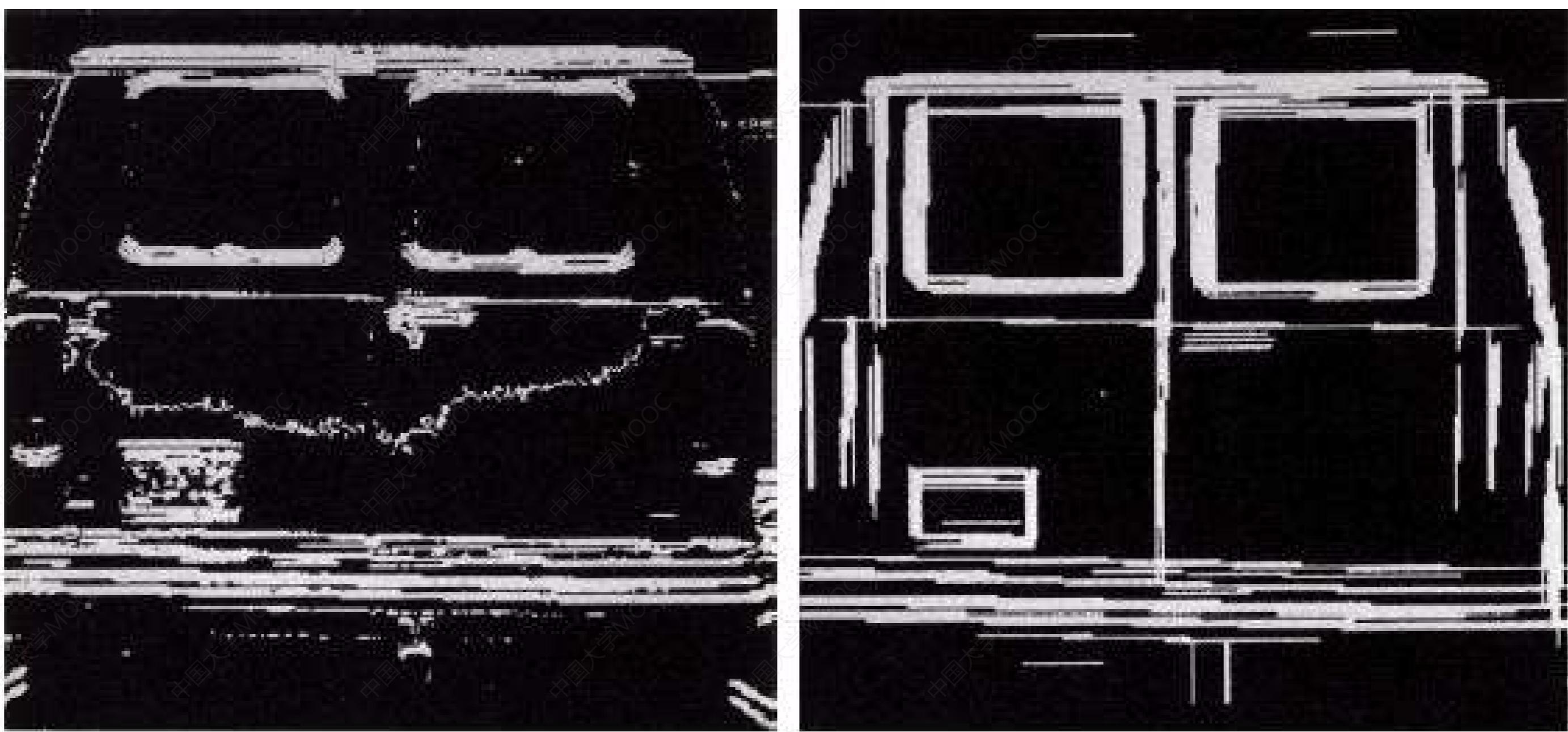
a b
c d

FIGURE 10.16

- (a) Input image.
- (b) G_y component of the gradient.
- (c) G_x component of the gradient.
- (d) Result of edge linking. (Courtesy of Perceptics Corporation.)



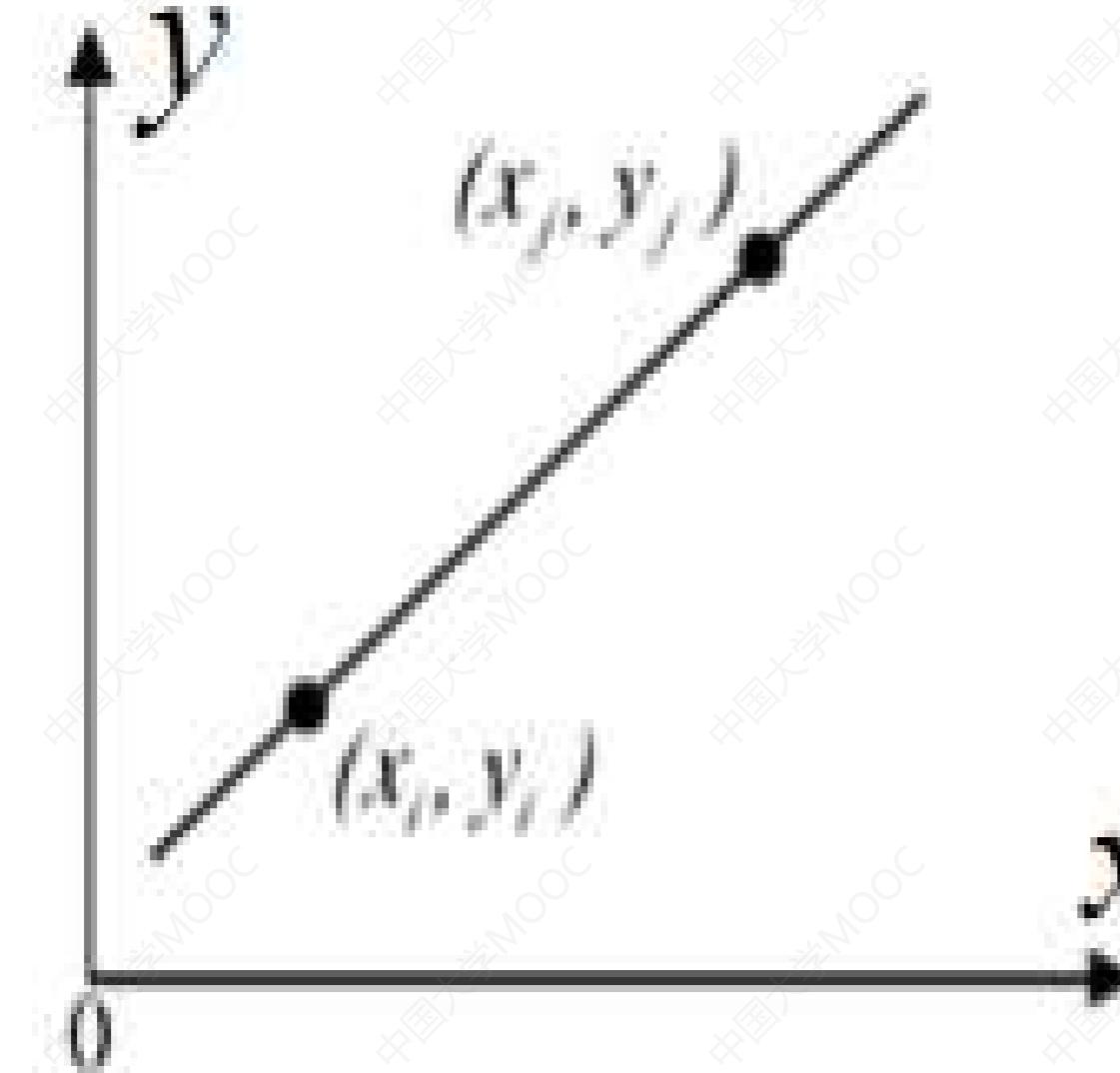
幅度大于25,
方向偏差不大
于15度



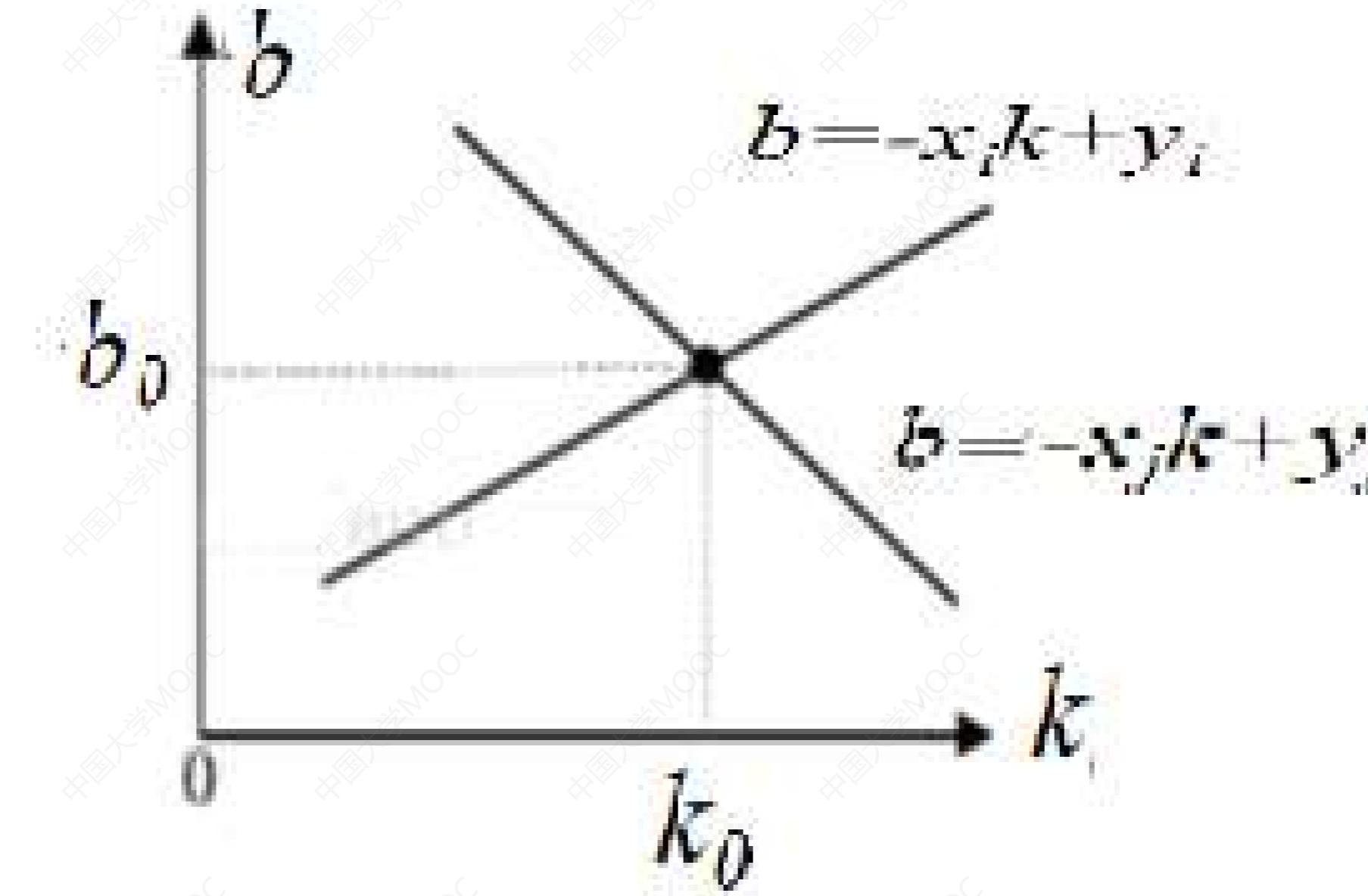
■ 7.2 基于边缘的分割方法

- 全局处理——Hough变换
- **基本原理：**通过参数空间映射提取图像中的可以用参数表示的曲线；
- 设直线方程： $y=k_0x+b_0$
- 直线上的每一点 (x_i, y_i) 均对应于参数空间 (k, b) 的一条直线 $b = -x_i k + y_i$ ；
- 且每条直线都满足方程： $y_i = k_0 x_i + b_0$ ，所以每条直线都经过点 (k_0, b_0)

7.2 基于边缘的分割方法



(a) 图像空间



(b) 参数空间

直线解析式一： $y = k \cdot x + b,$

参数空间映射 $(x, y) \rightarrow (k, b)$

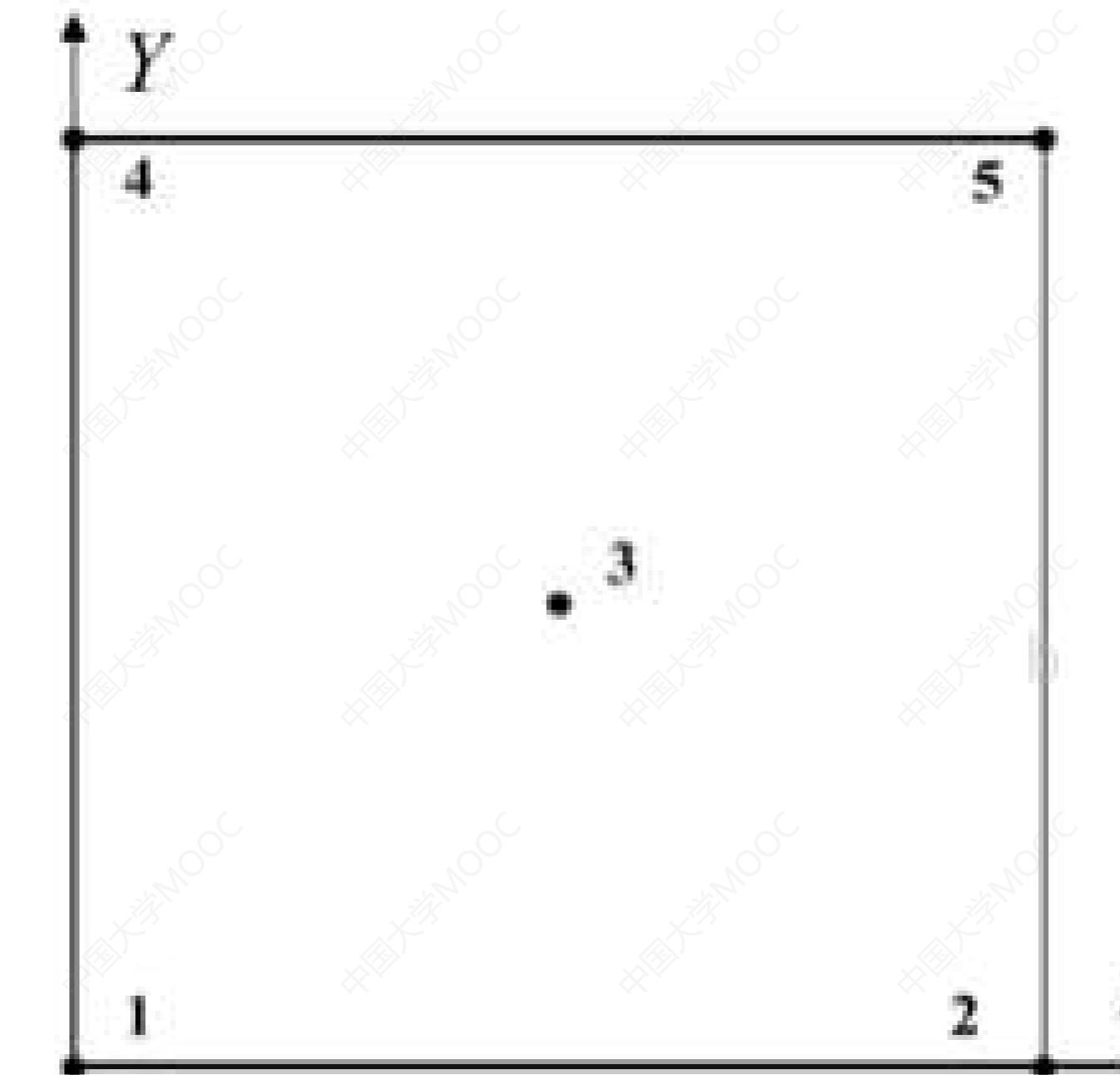
如果 $x=C$, 则 k 值接近无穷, 直线上的点在参数空间无法表示

7.2 基于边缘的分割方法

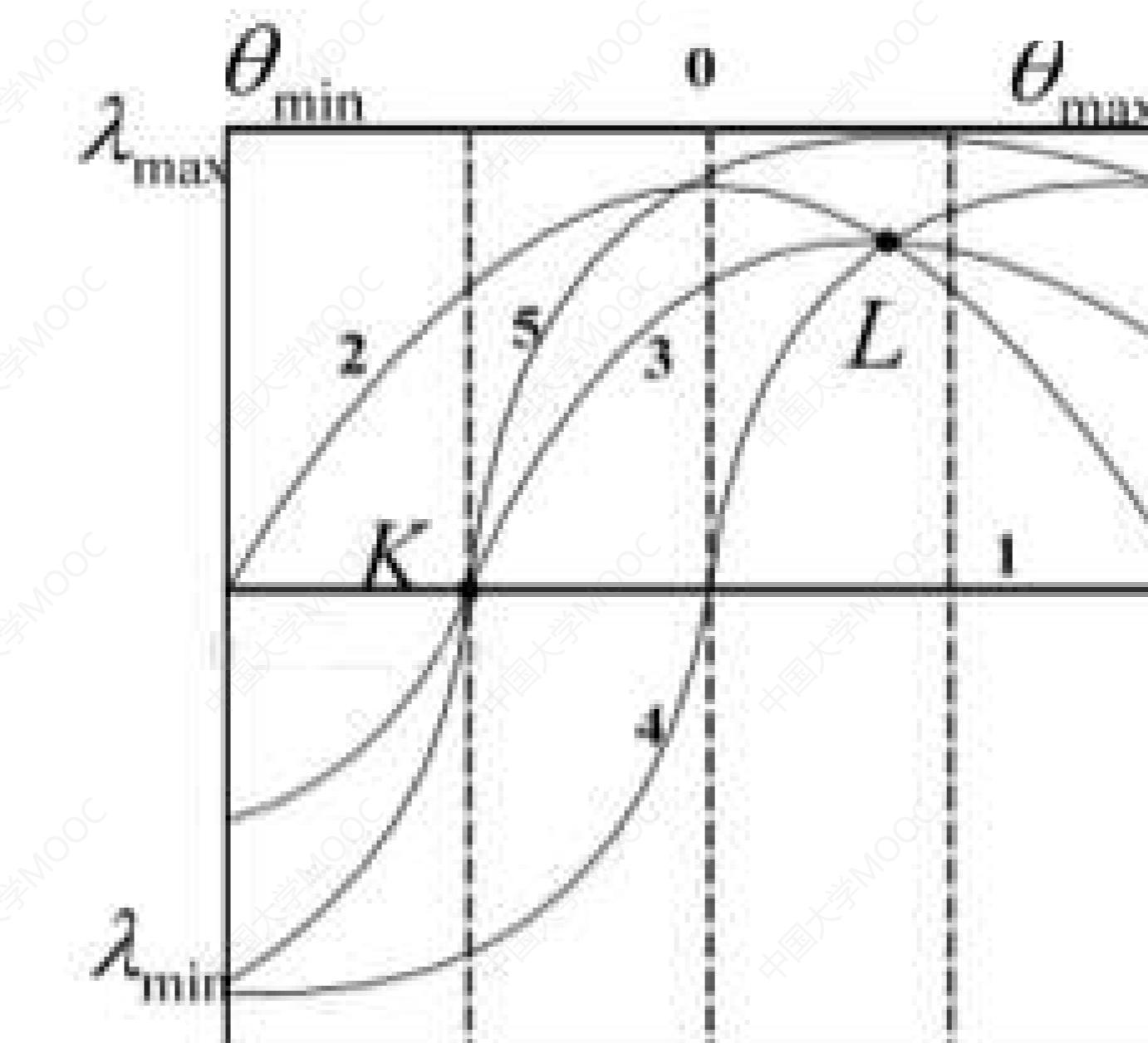
Hough变换的改进方案——将直角坐标改为极坐标

直线解析式二： $\rho = x \cdot \cos \theta + y \cdot \sin \theta$

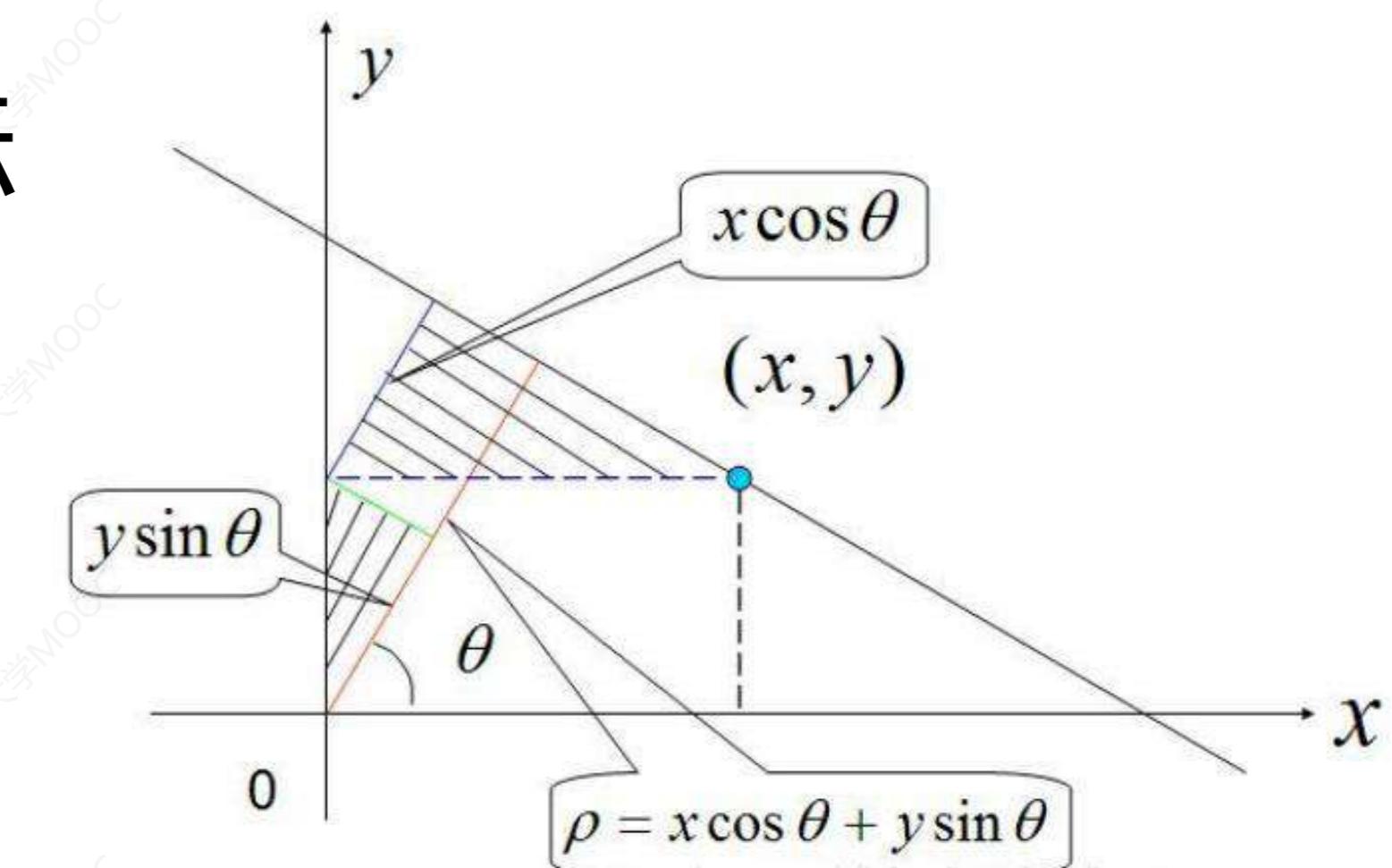
参数空间映射 $(x, y) -> (\rho, \theta)$



(a) 图像空间

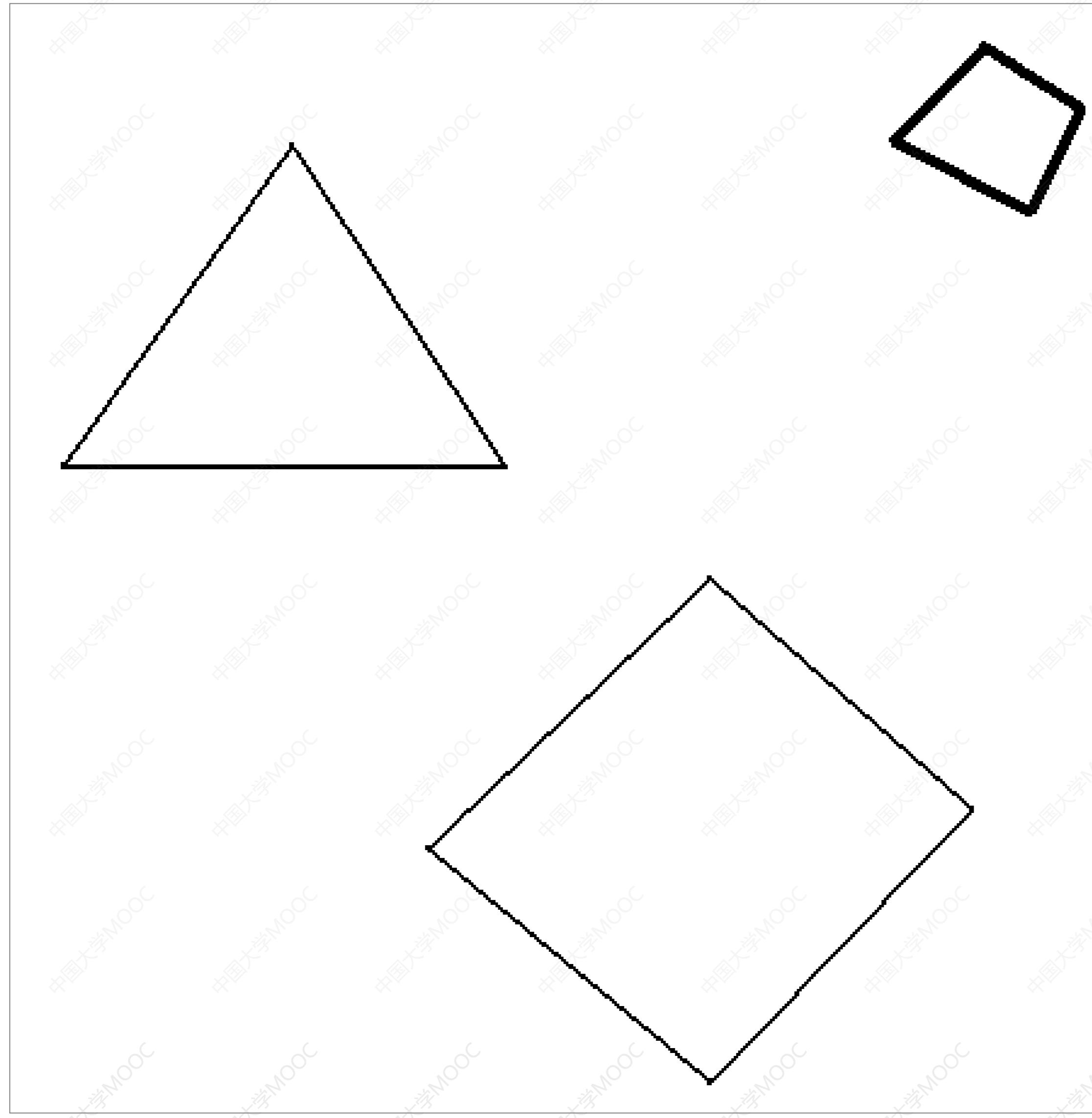


(b) 参数空间



■ 7.2 基于边缘的分割方法

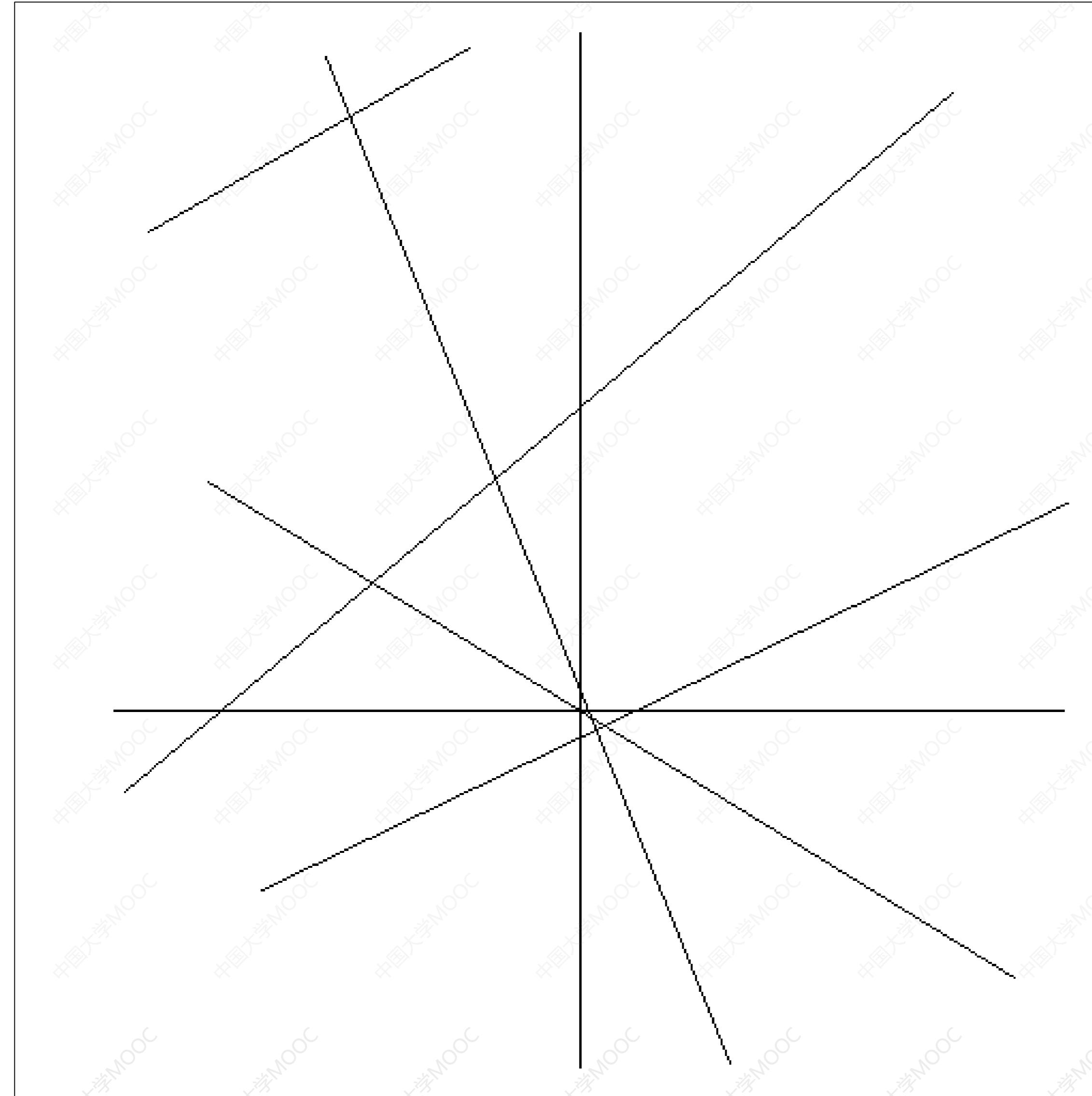
输入
图像



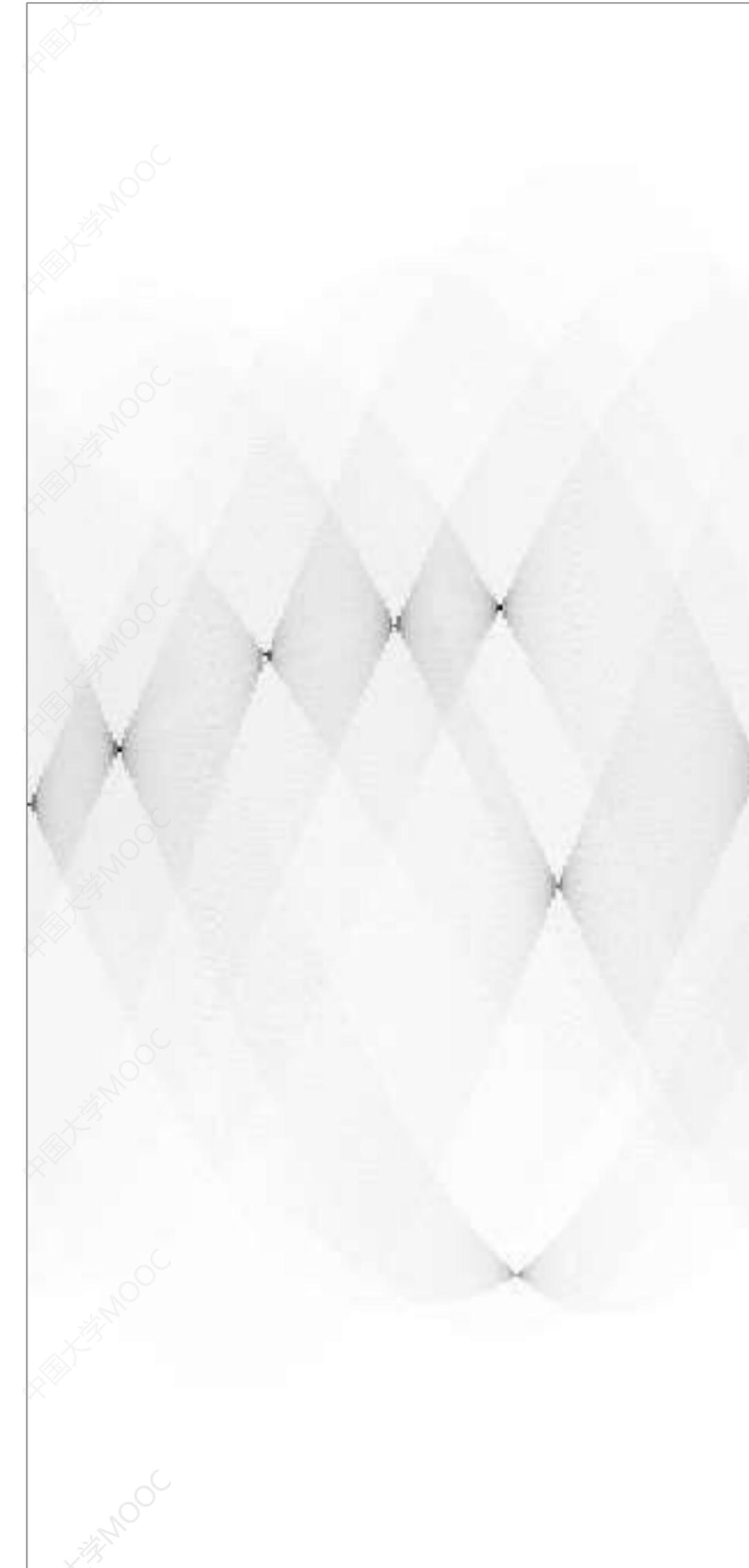
Hough
变换结
果图像

■ 7.2 基于边缘的分割方法

输入
图像

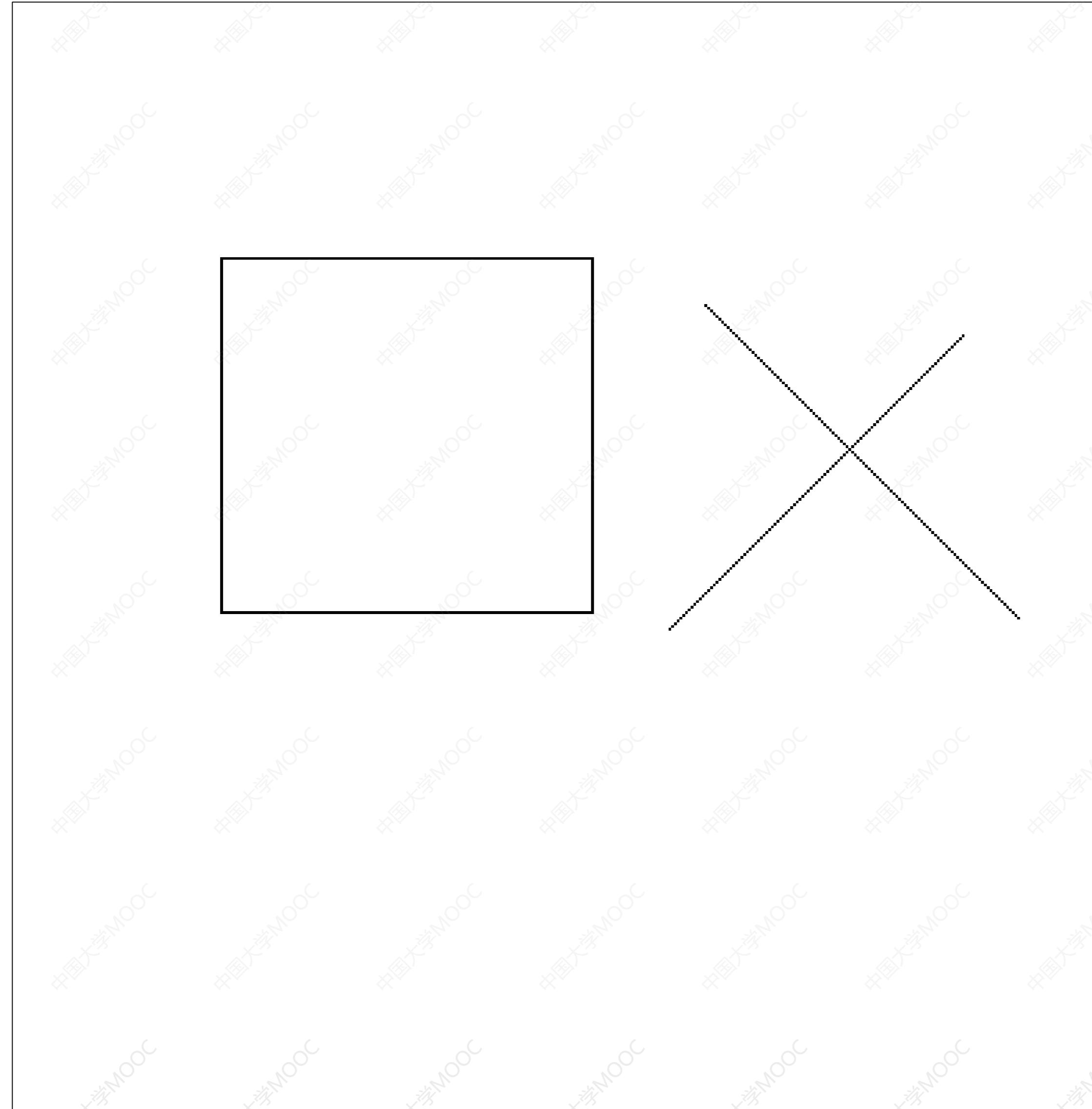


Hough
变换结
果图像



7.2 基于边缘的分割方法

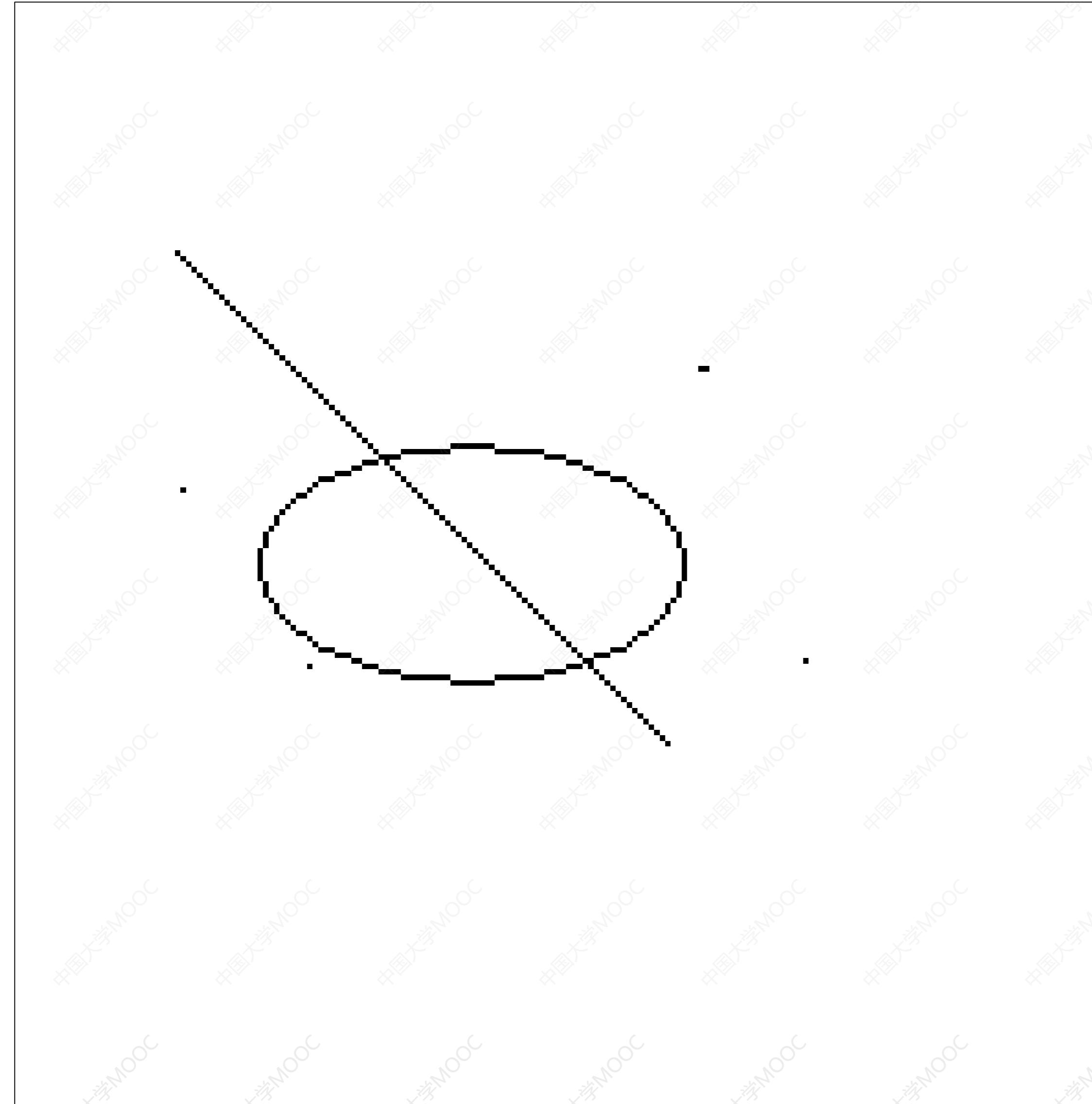
输入
图像



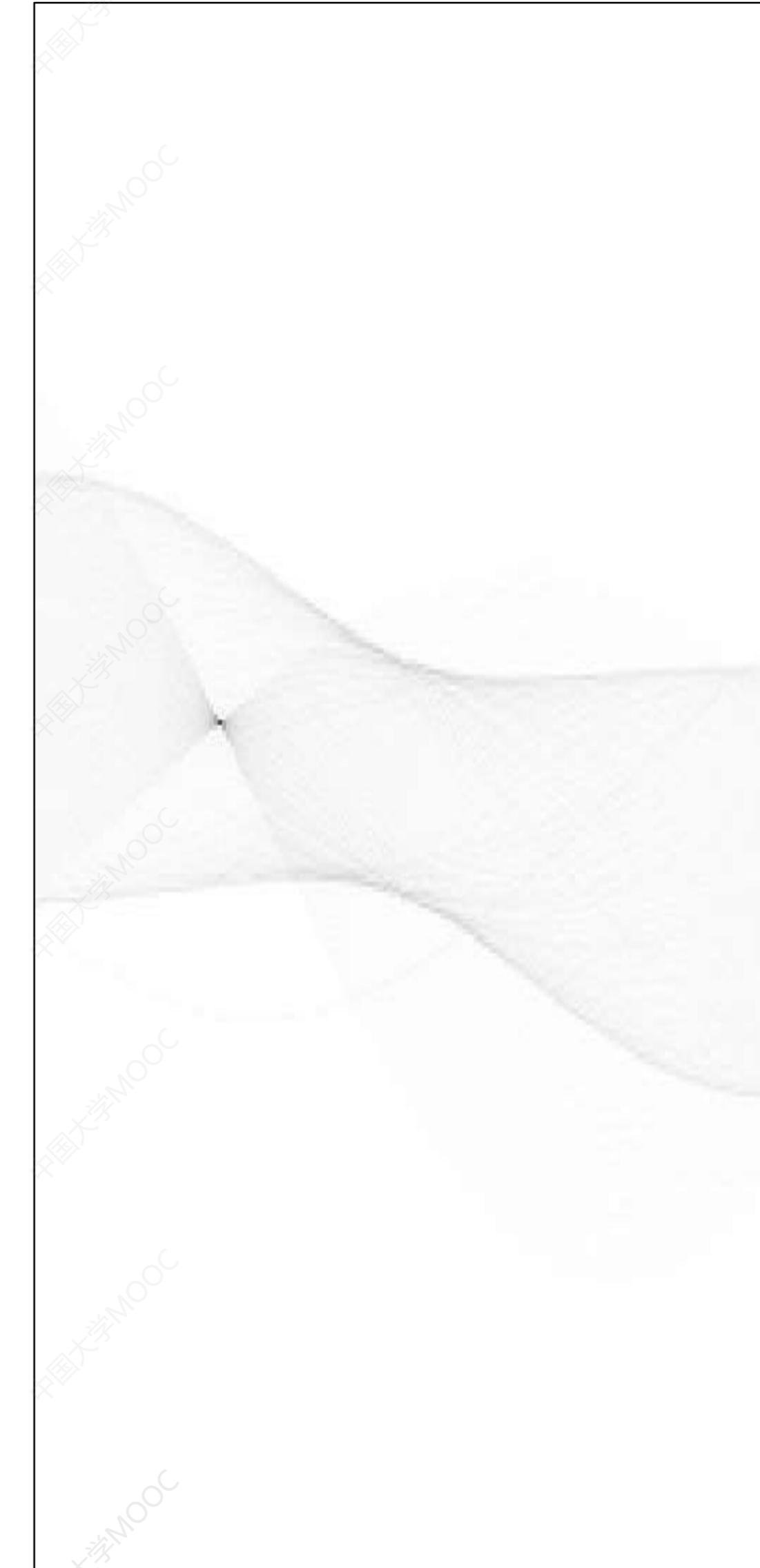
Hough
变换结
果图像

■ 7.2 基于边缘的分割方法

输入
图像



Hough
变换结
果图像



■ 7.2 基于边缘的分割方法

➤ Hough变换的扩展——提取任意参数曲线

- 圆心为 (x_0, y_0) , 半径为 r 的圆的方程表达为

$$(x - x_0)^2 + (y - y_0)^2 = r^2$$

- 圆上任意一点 (x, y) 对应于 (x_0, y_0, r) 参数空间上一个在经过 (x, y) 点的圆锥

$$(x_0 - x)^2 + (y_0 - y)^2 - r^2 = 0$$

■ 7.2 基于边缘的分割方法

➤ 基于Hough变换及改进算法的直线和参数曲线的提取

- RHT (Random Hough Transform)
- CHT (Circle Hough Transform)

等等

■ 7.2 基于边缘的分割方法

- 基于离散直线段 (Discrete Line Segments, DLS) 合并的方法，
 - 跟踪边缘后再分割成很多短的近似位于同一直线上的 DLS，然后通过一定的规则将这些 DLS 进行合并。
 - 这种方法的特点是局部性好能够保持良好的边缘特性，但传统合并方法很难保证结果的全局最优。

■ 7.2 基于边缘的分割方法

➤ 基于卡尔曼滤波的边缘跟踪方法

- 提取并细化边缘
- 然后利用卡尔曼滤波器跟踪边缘点，获得直线支持区
- 最后设计一种稳健的方法来拟合直线支持区
- 实验结果表明，这种算法具有良好的性能，适应于较广范围的实际图像

■ 7.2 基于边缘的分割方法

➤ 基于链码跟踪的直线提取

- 先对目标边界进行链码跟踪，
- 然后在得到的链码串集合中进行直线段提取。
- 这类方法的优点是计算量小，并且能同时得到直线段的位置、长度、方向等信息。
- 不足之处是算法性能受目标边界的跟踪算法制约。

■ 7.2 基于边缘的分割方法

➤ 基于Zernike矩的直线提取算法

- 首先利用ZMs在局部坐标系中计算边缘点所在直线方程
- 然后通过坐标变换将直线方程变换到图像坐标系中，从而建立了边缘点和所在直线参数之间的一一映射关系
- 最后对于所求出的直线参数在参数空间对参数单元投票取阈值提取直线
- 实验表明，该算法速度快、精度高和抗干扰性强

■ 7.2 基于边缘的分割方法

➤ 基于参数控制的直线提取算法

- 根据最终目标确定一些控制参数
- 在计算过程中利用参数尽早淘汰图像中与所需结果无关的部分，减轻后续复杂操作的负担，使计算量有数量级程度地减少
- 本算法在处理像素、使用先验知识、拟合直线的方式以及固有的时间性能等方面有突出的表现

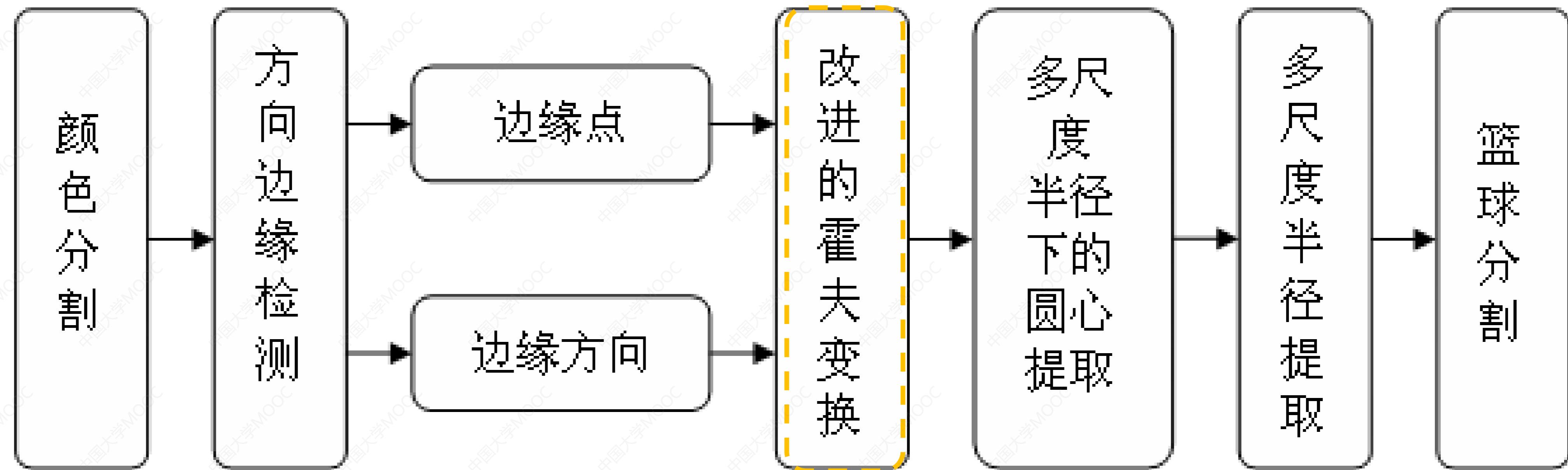
■ 7.2 基于边缘的分割方法

Hough变换方法应用——基于改进Hough变换提取圆方法



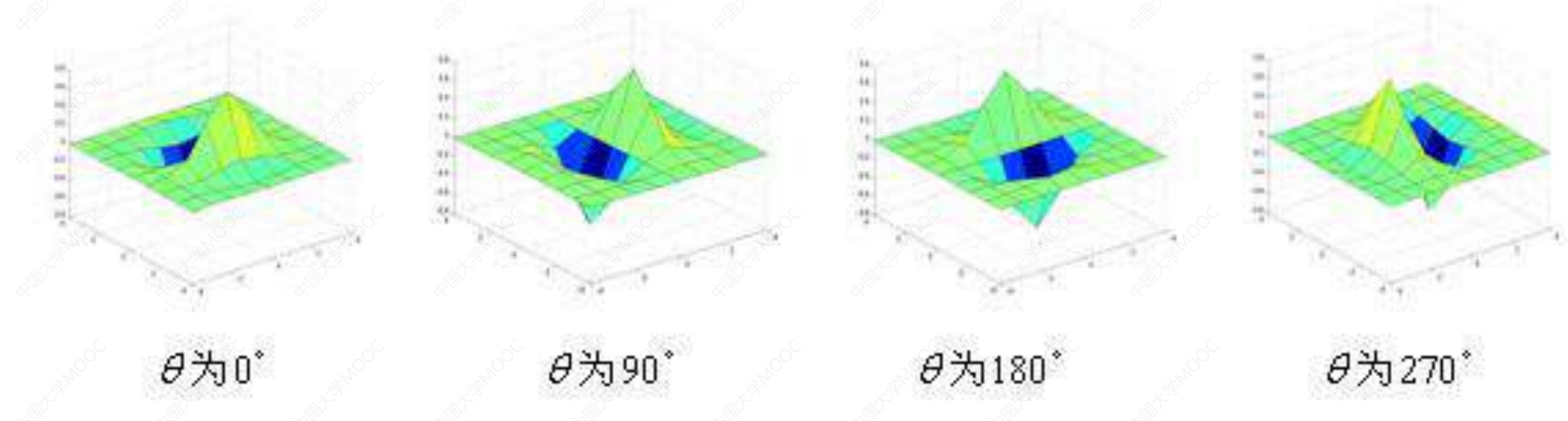
■ 基于改进Hough提取圆方法

• 算法流程



■ 基于改进Hough提取圆方法

方向可变滤波器

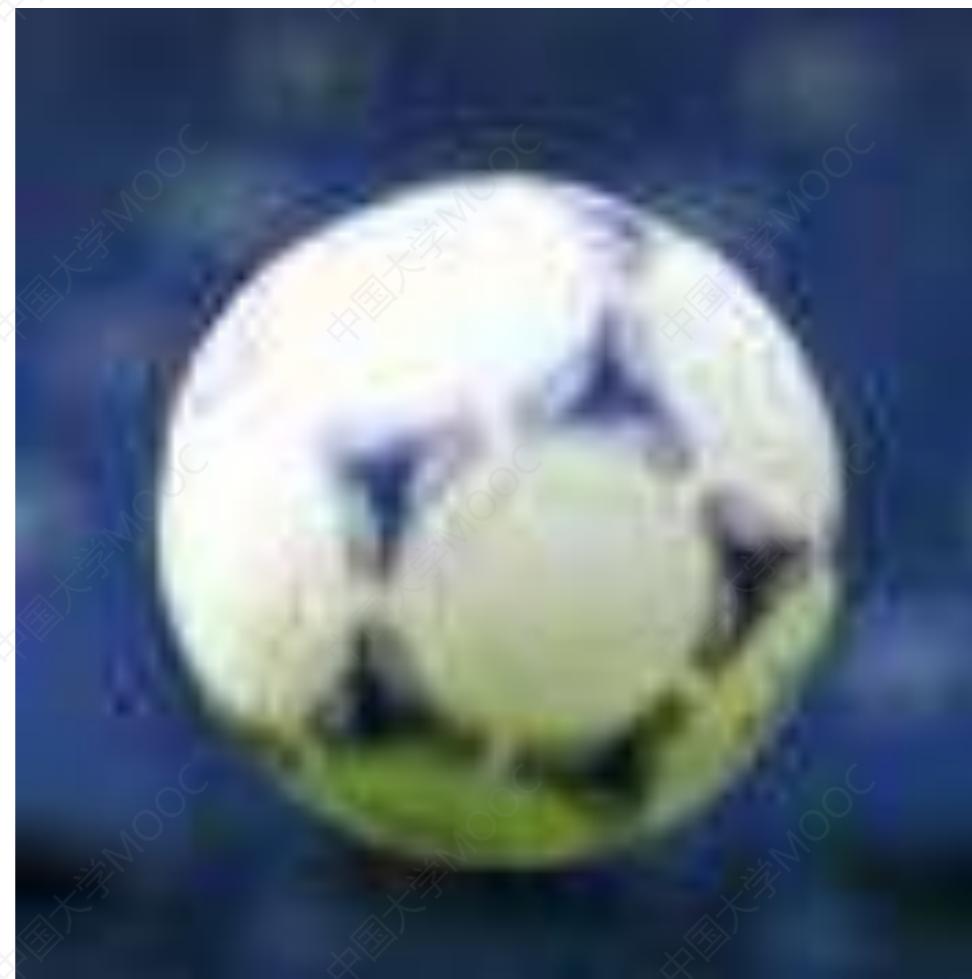


不同方向 θ 下的高斯函数二阶Hilbert
方向可变滤波器 $H_2^\theta(x, y)$ 结构

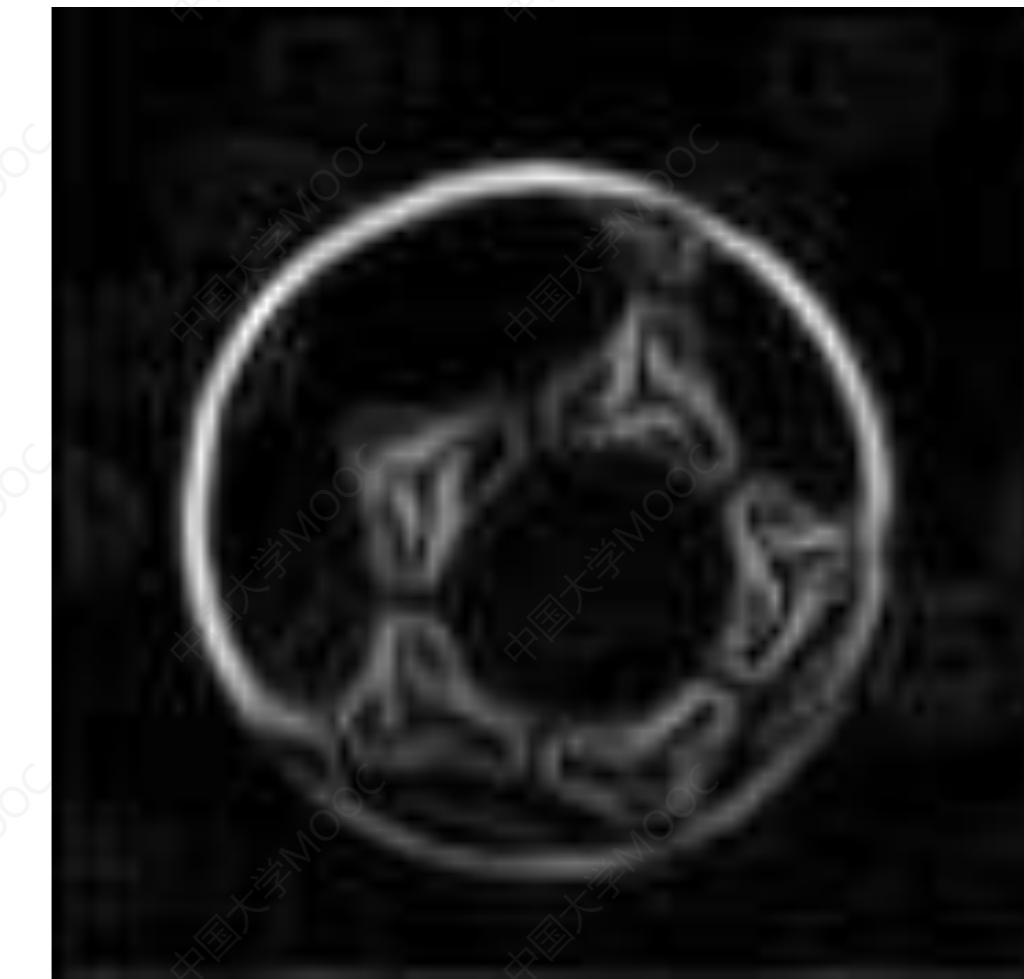
■ 基于改进Hough提取圆方法

方向边缘检测

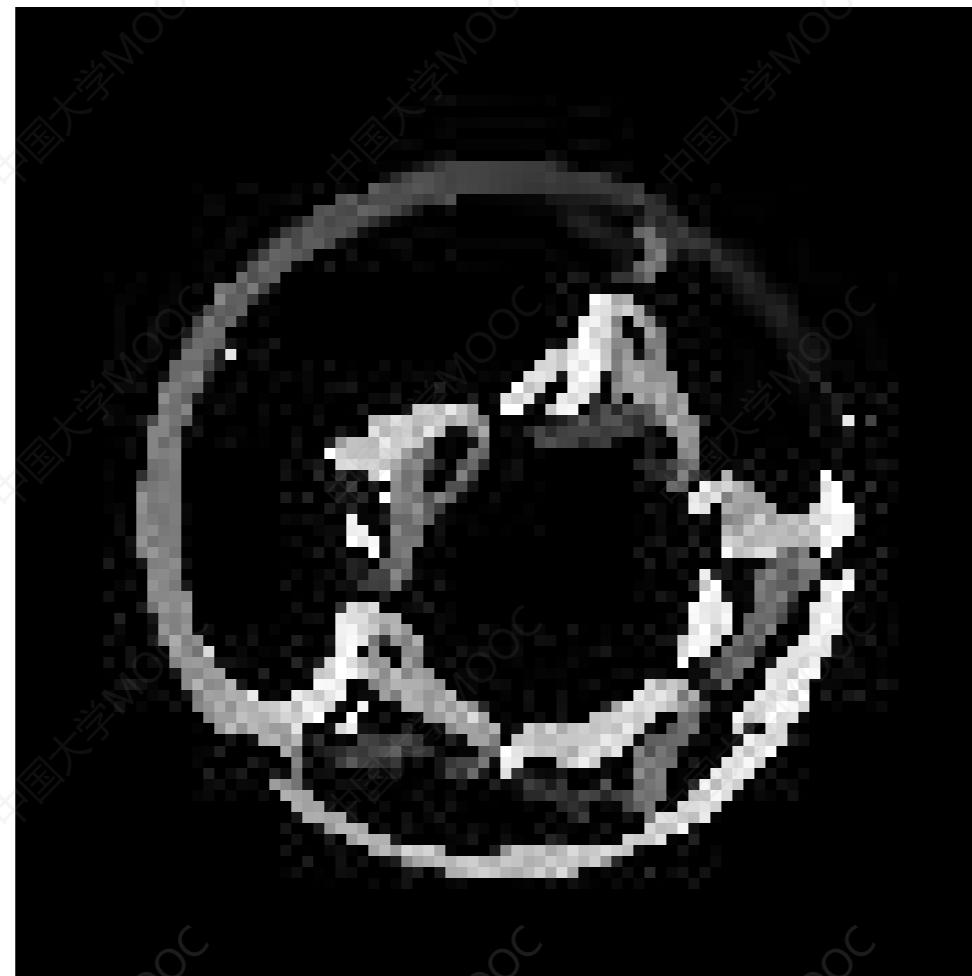
原始
图像



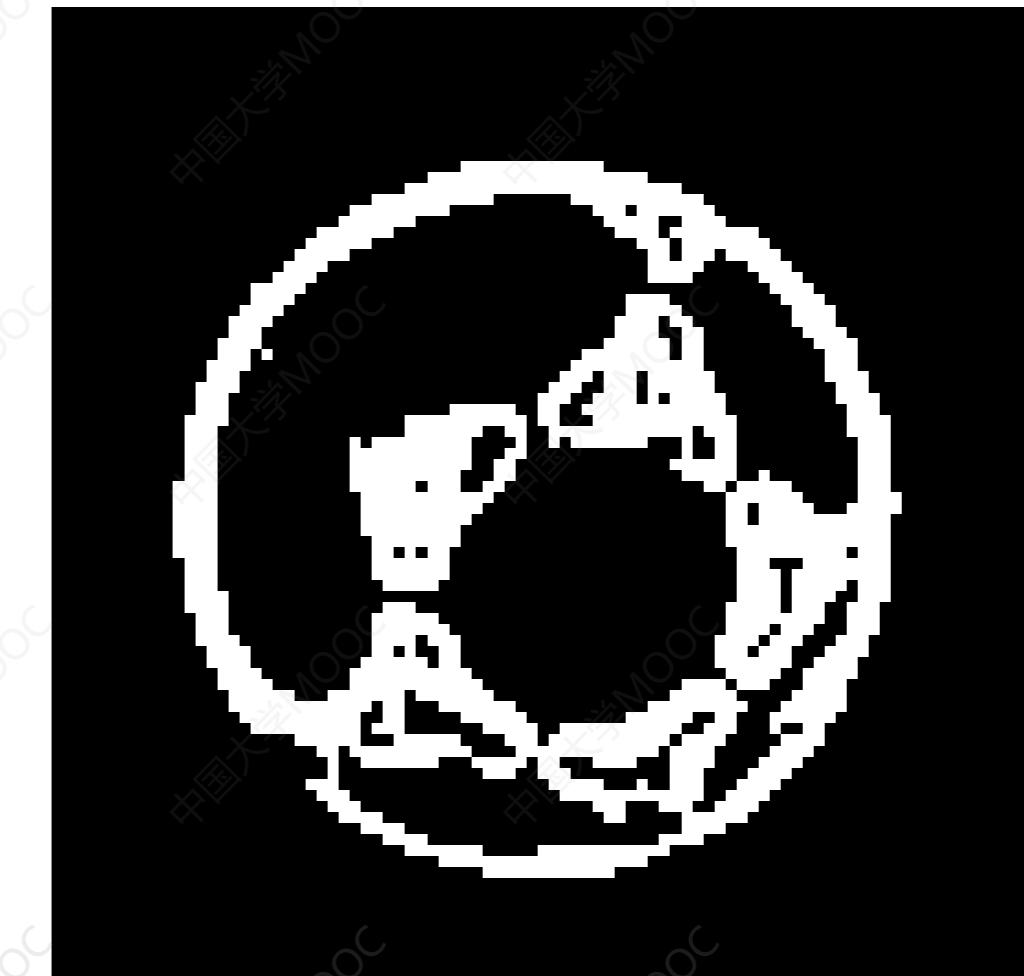
边缘能
量图



边缘方
向图



边缘像
素点图



刘勋, 毋立芳等. 一种改进的Hough变换提取圆方法. Apr. 2003 20(6):623-627.

■ 基于改进Hough提取圆方法

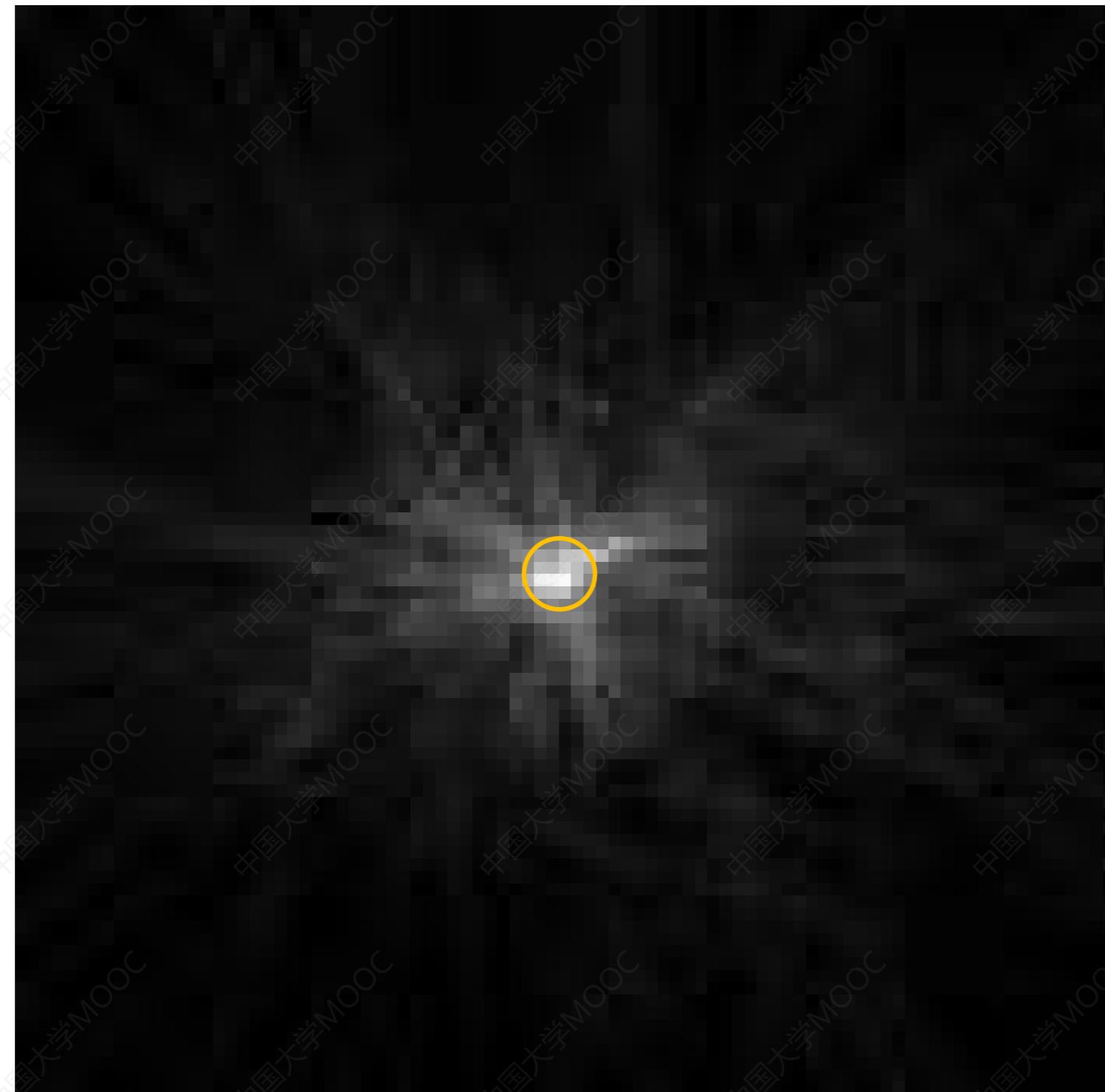
方向可变滤波器可以
准确提取边缘方向



- 沿着圆形轮廓，
边缘法线方向逐
次递增
- 范围[0, 359]

■ 基于改进Hough提取圆方法

圆心提取



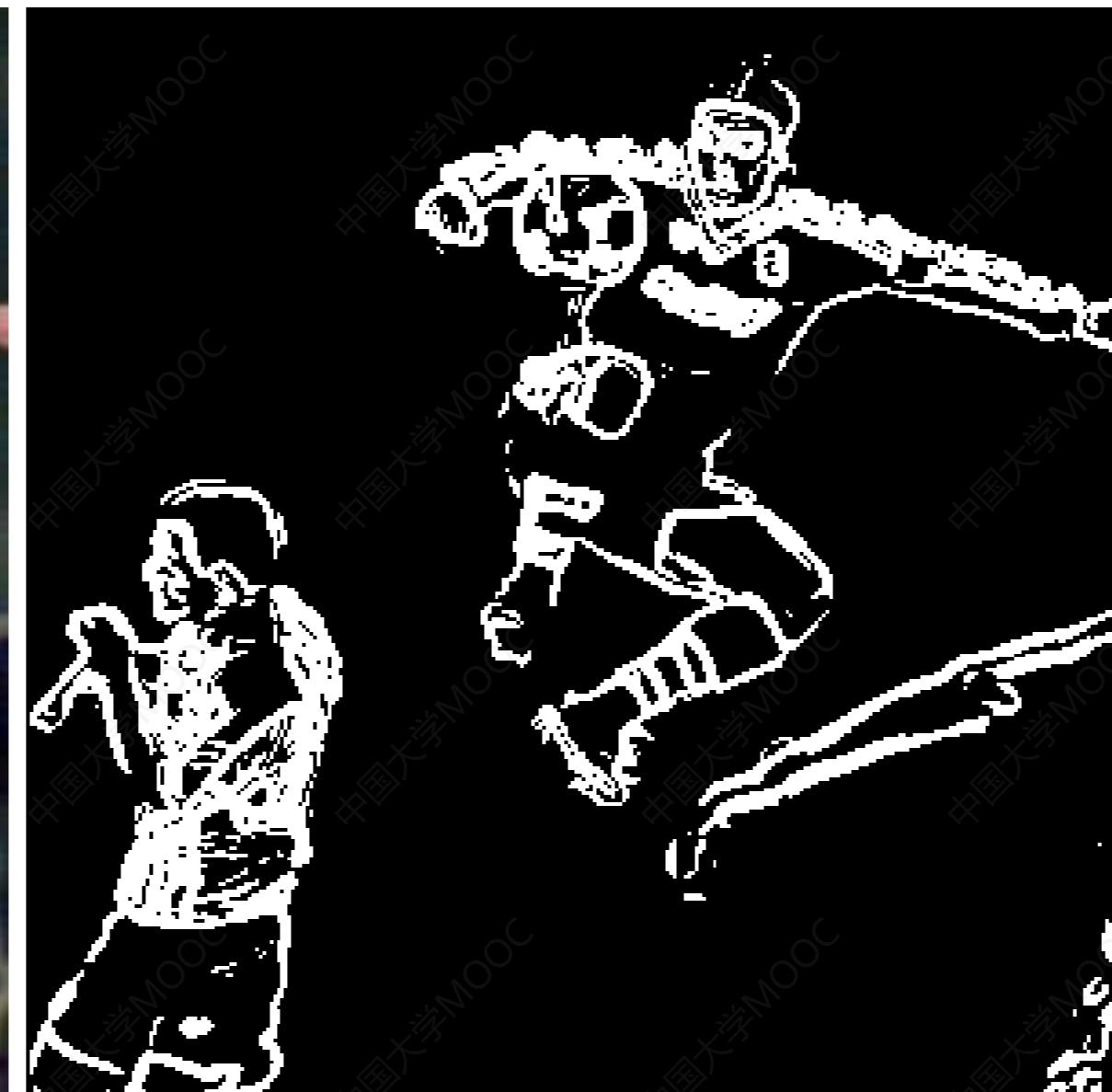
- 特征 - 圆形边缘点法线交于圆心
- 建立圆心参数平面
- 作边缘点法线
- 统计圆心参数平面的灰度峰值

■ 基于改进Hough提取圆方法

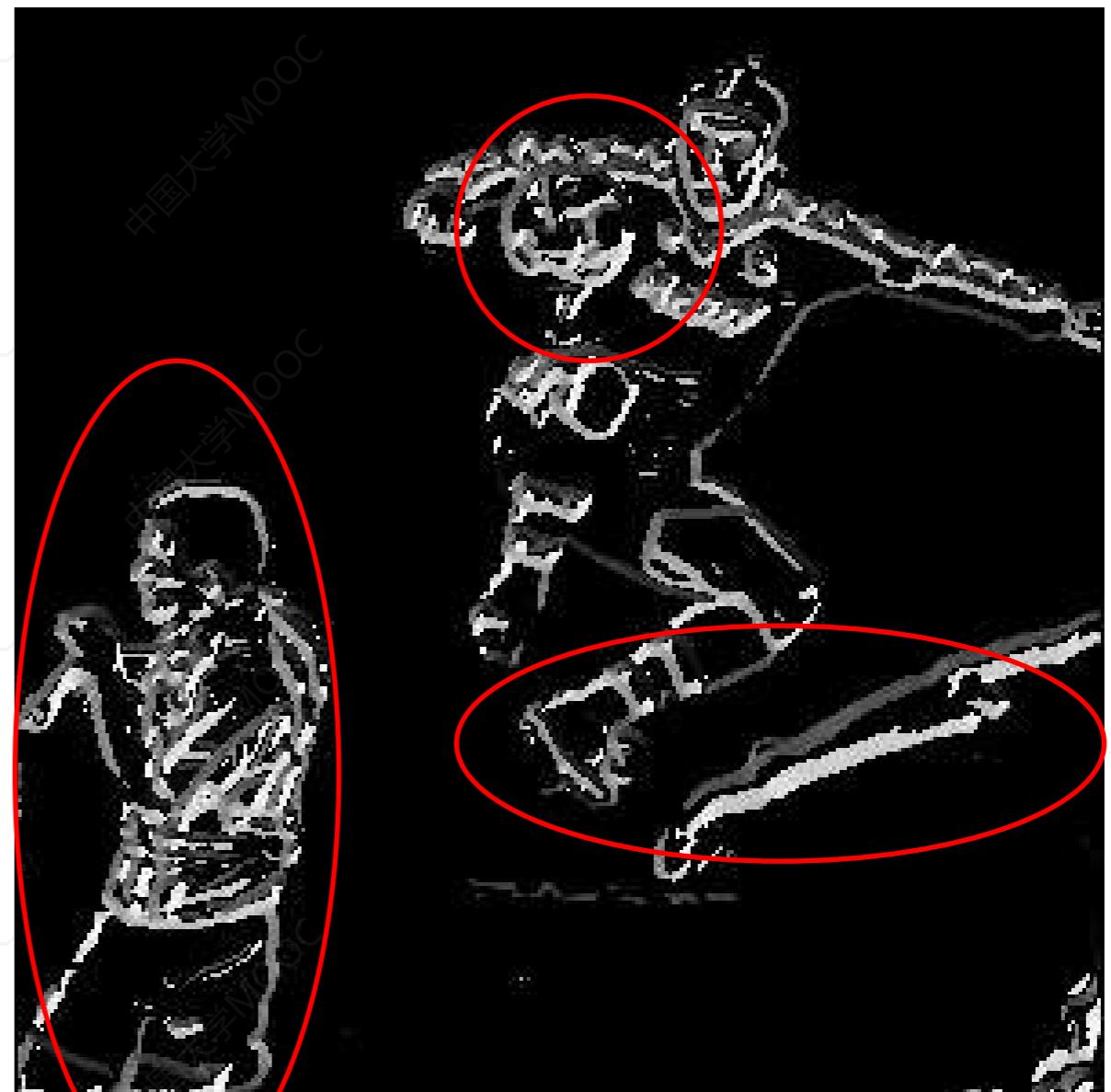
复杂场景的影响



原始图像



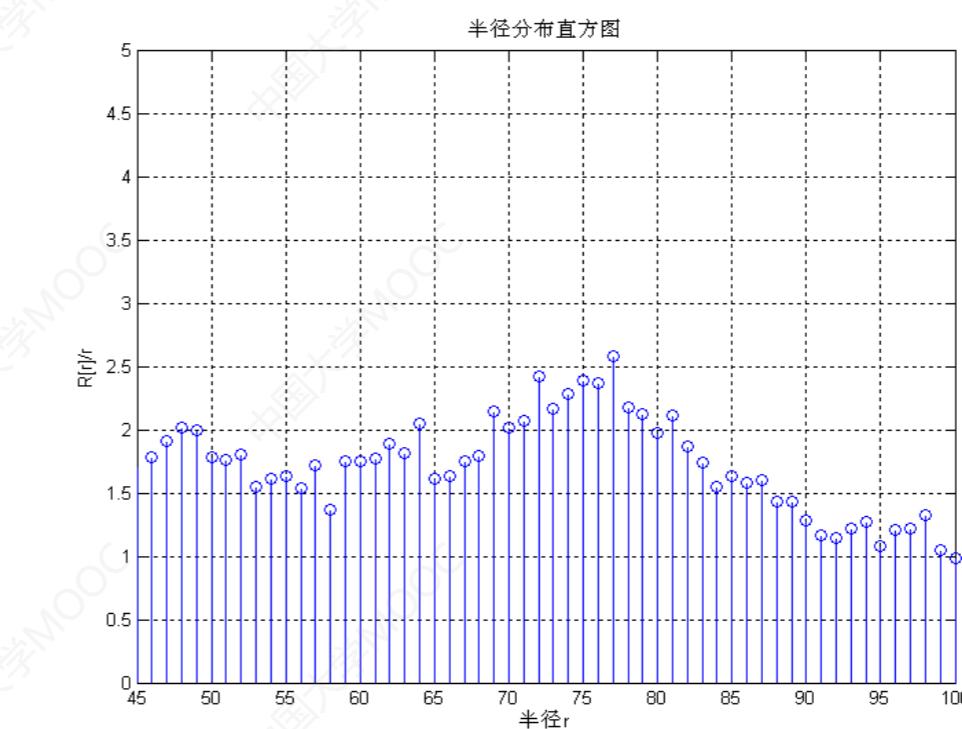
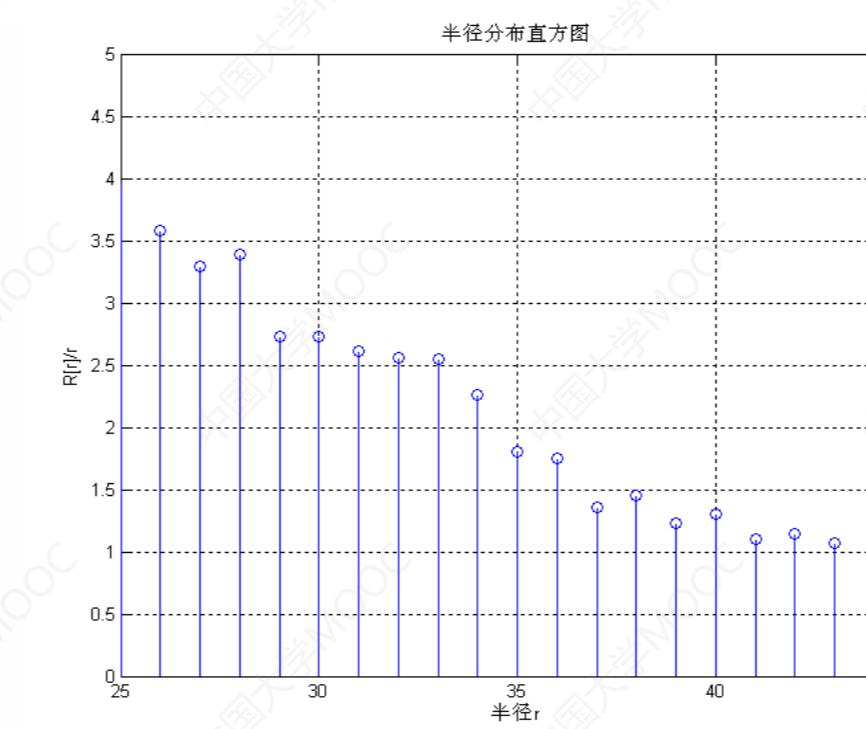
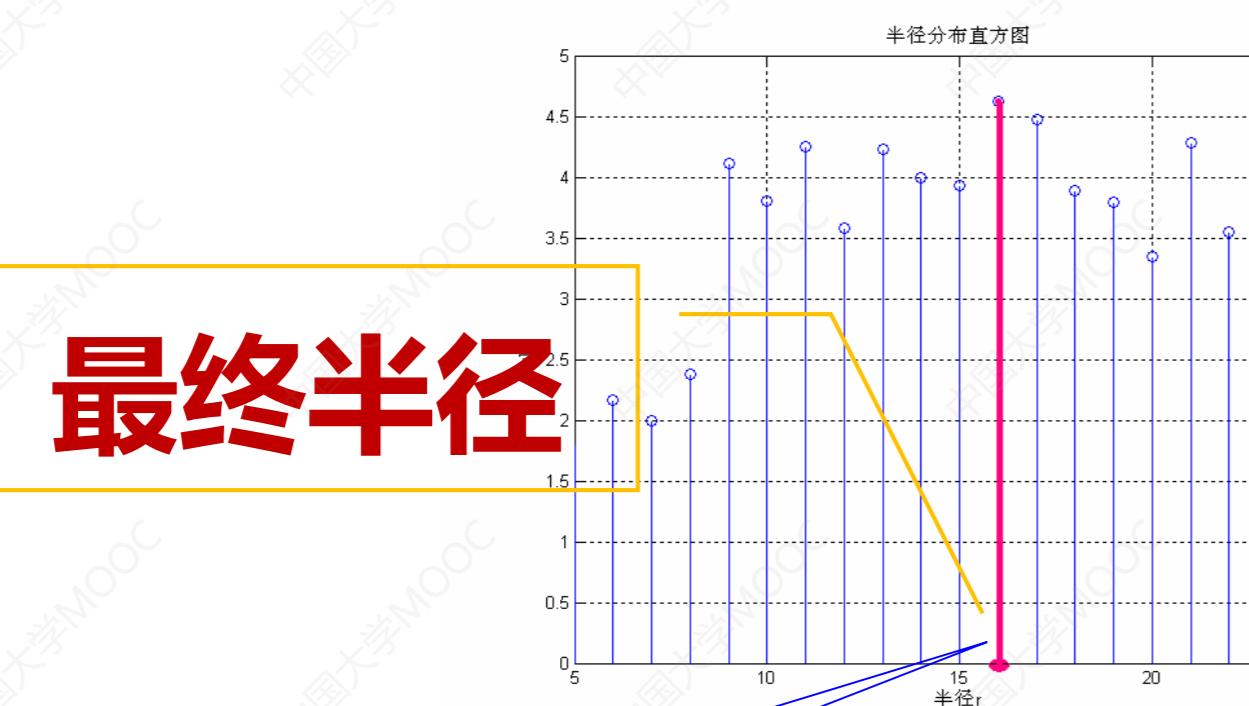
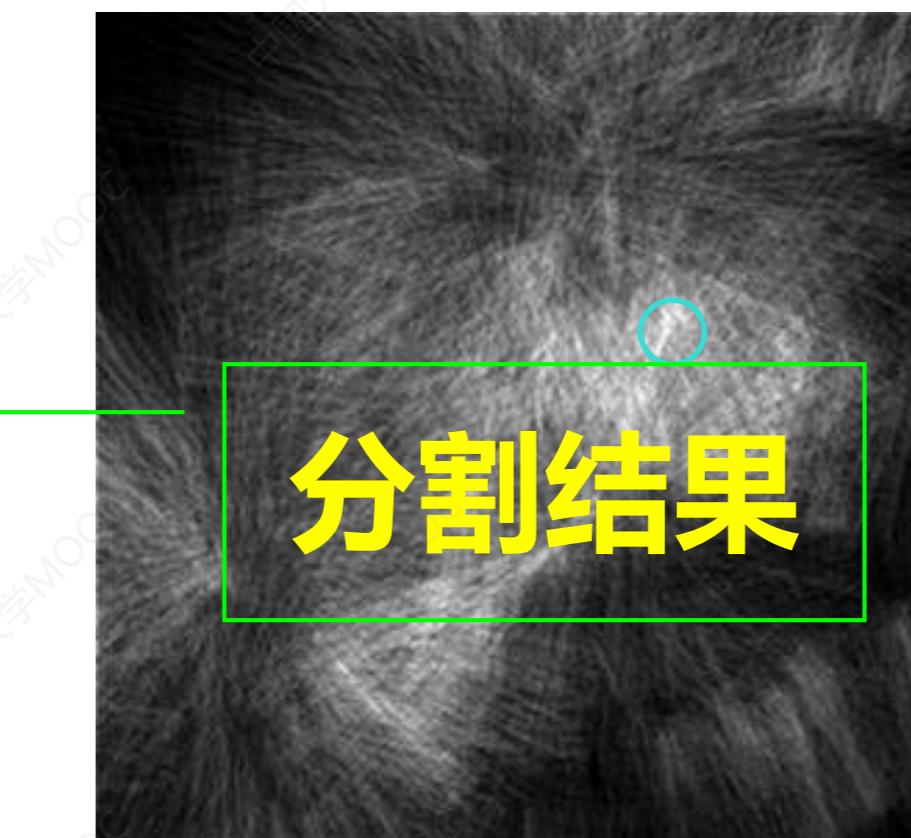
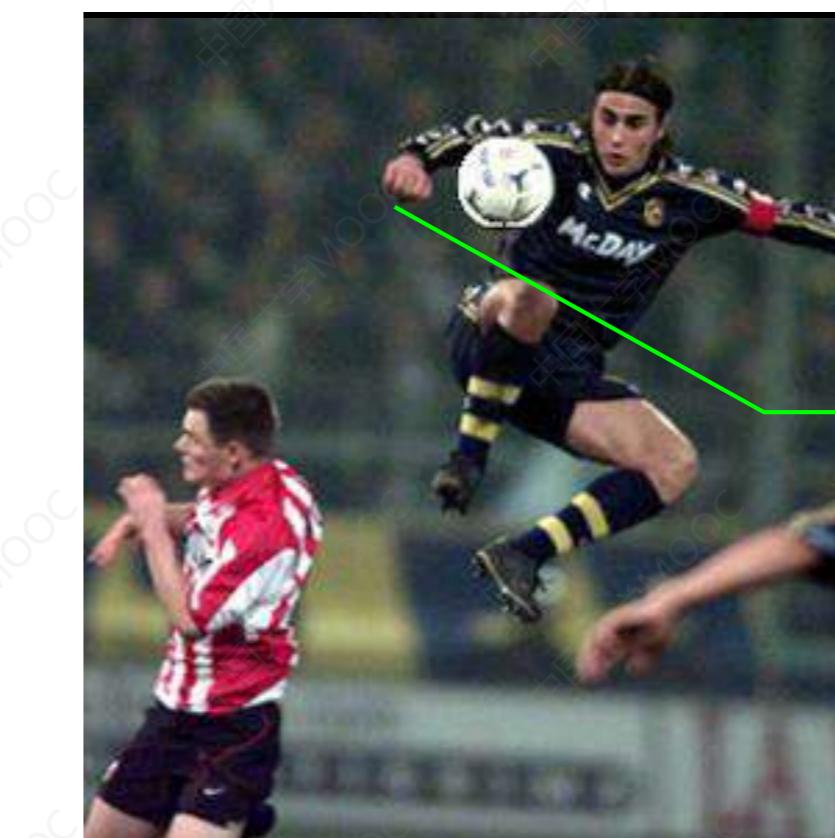
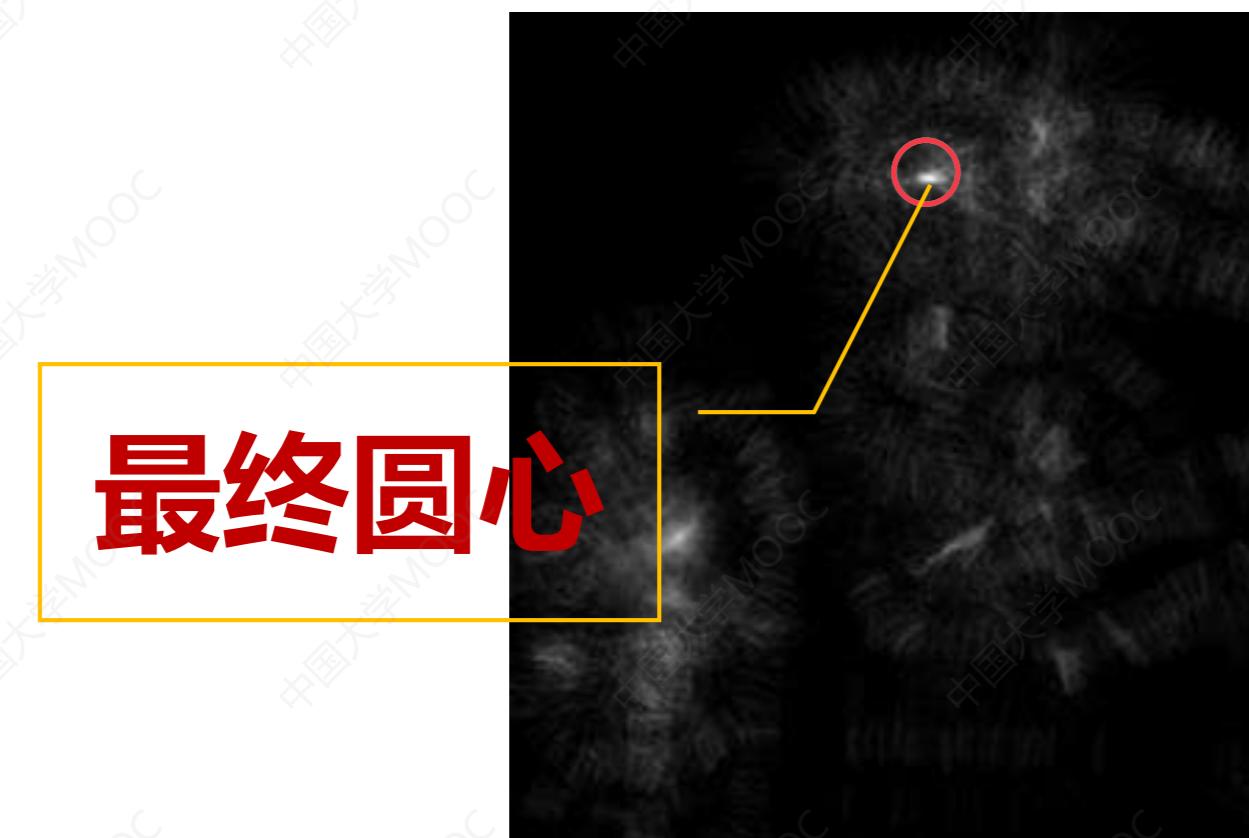
边缘像素点



最大边缘方向图

■ 基于改进Hough提取圆方法

多尺度半径下圆心和半径提取



$R=16$

$R=5-25$

$R=25-45$

$R=45$ 以上

刘勋, 毋立芳等. 一种改进的Hough变换提取圆方法. Apr. 2003 20(6):623-627.

■ 基于改进Hough提取圆方法

■ 球类分割结果

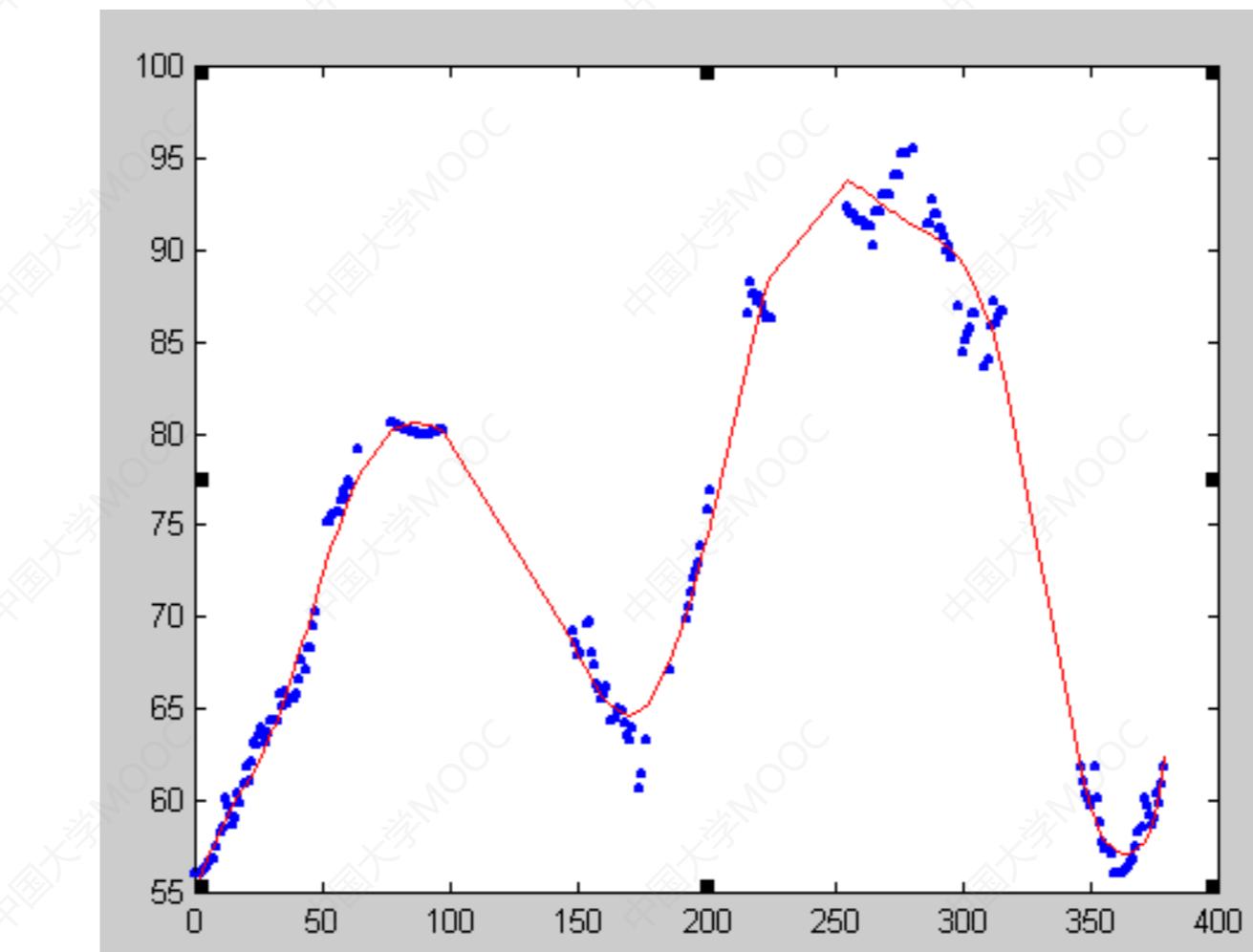
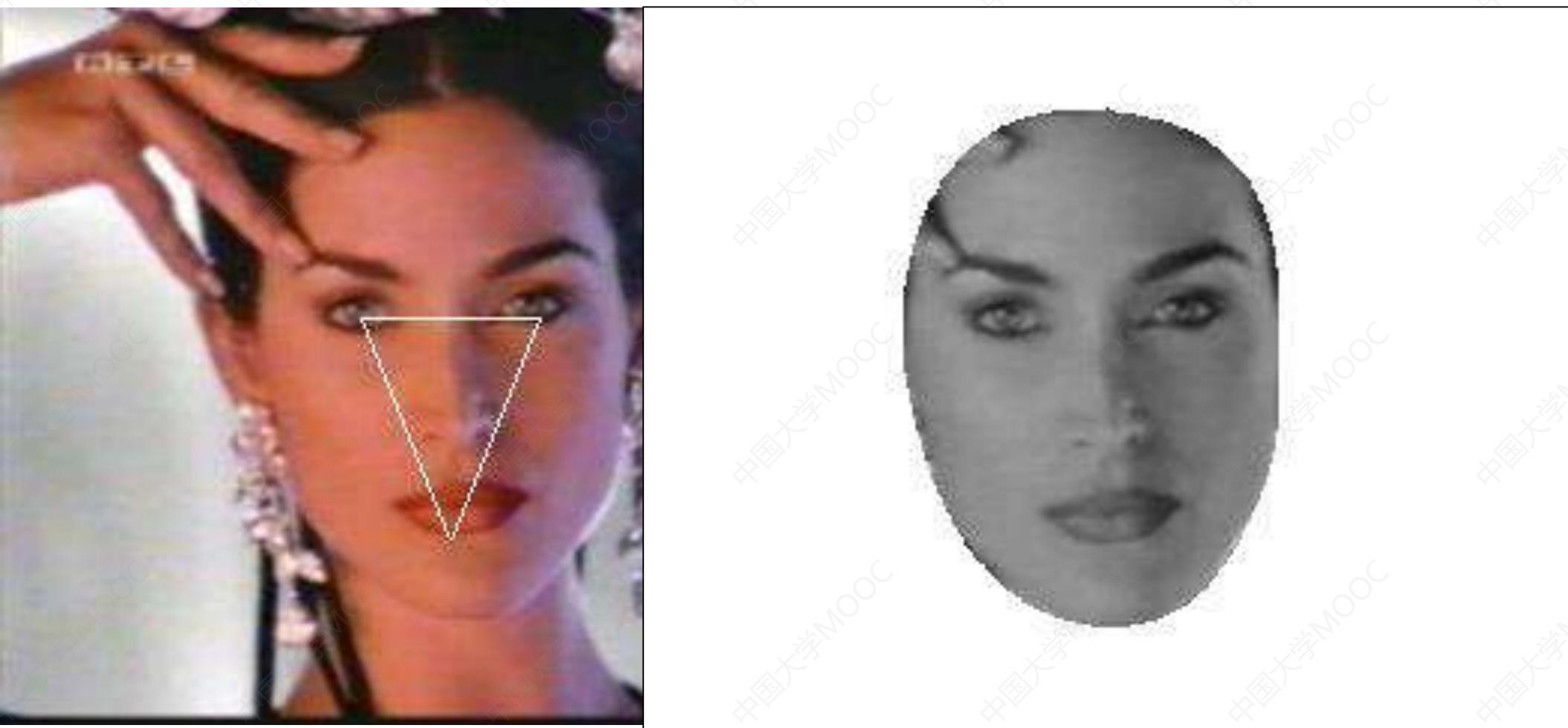


- 对部分遮挡和运动模糊有较好的鲁棒性；
- 同样的方法对排球和足球也有较好的鲁棒性

■ 7.2 基于边缘的分割方法

基于参数曲线拟合的人脸分割

- 提取人脸轮廓，保留人脸区域



搜索范围：
内边界：椭圆
长轴: $0.8 * D_{\text{eye-eye}}$
短轴: $D_{\text{eye-mouth}}$
外边界：矩形
宽度: $3 * D_{\text{eye-eye}}$,
高度: $2.25 * D_{\text{eye-mouth}}$

■ 7.3 基于区域的方法

- 基于区域的分割是以直接寻找区域为基础的分割技术，利用了对象区域与背景区域灰度分布的相似性。
- 基于区域的图像分割方法可以分为两大类
 - 区域生长法
 - 区域分裂与合并

■ 7.3 基于区域的方法

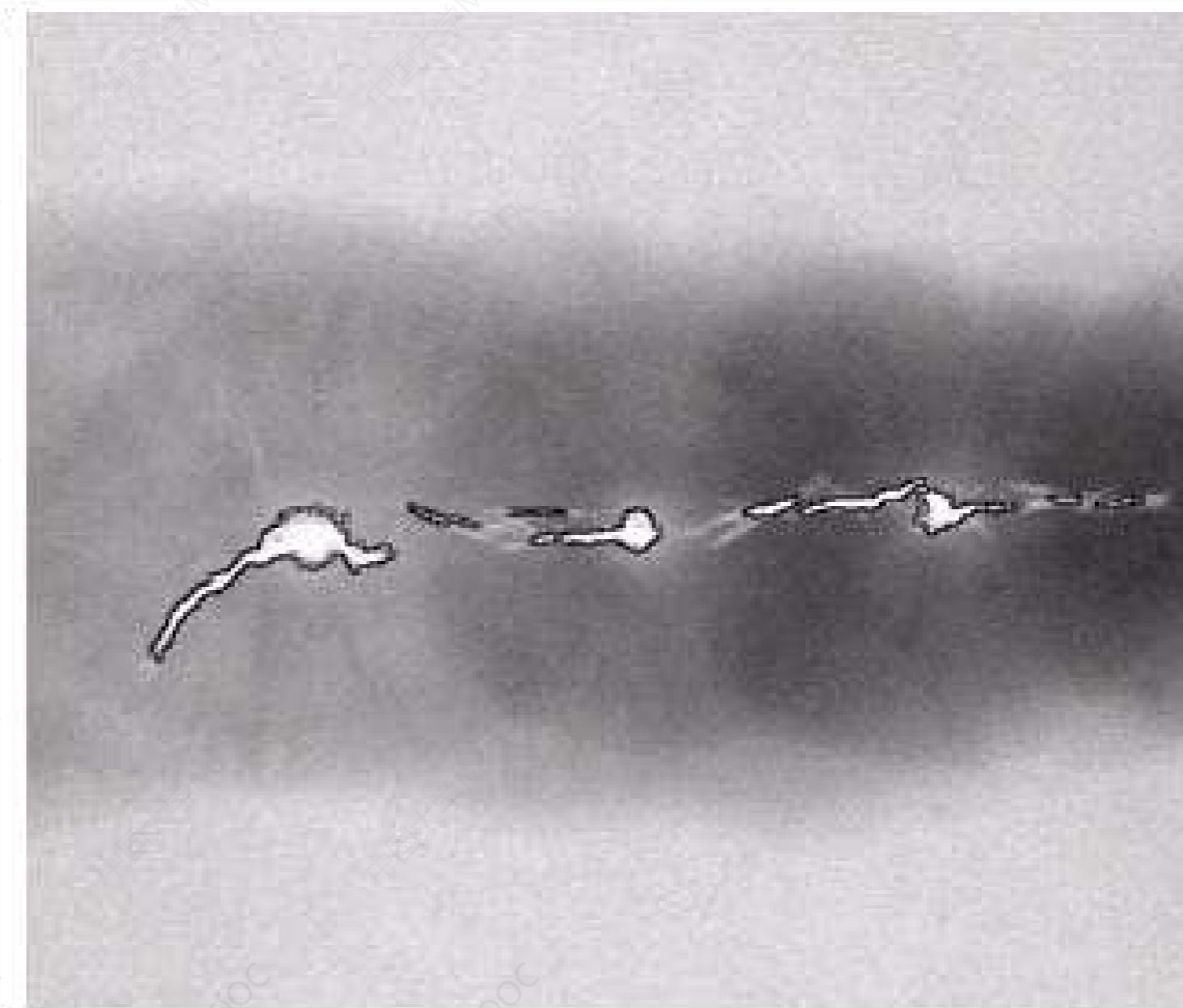
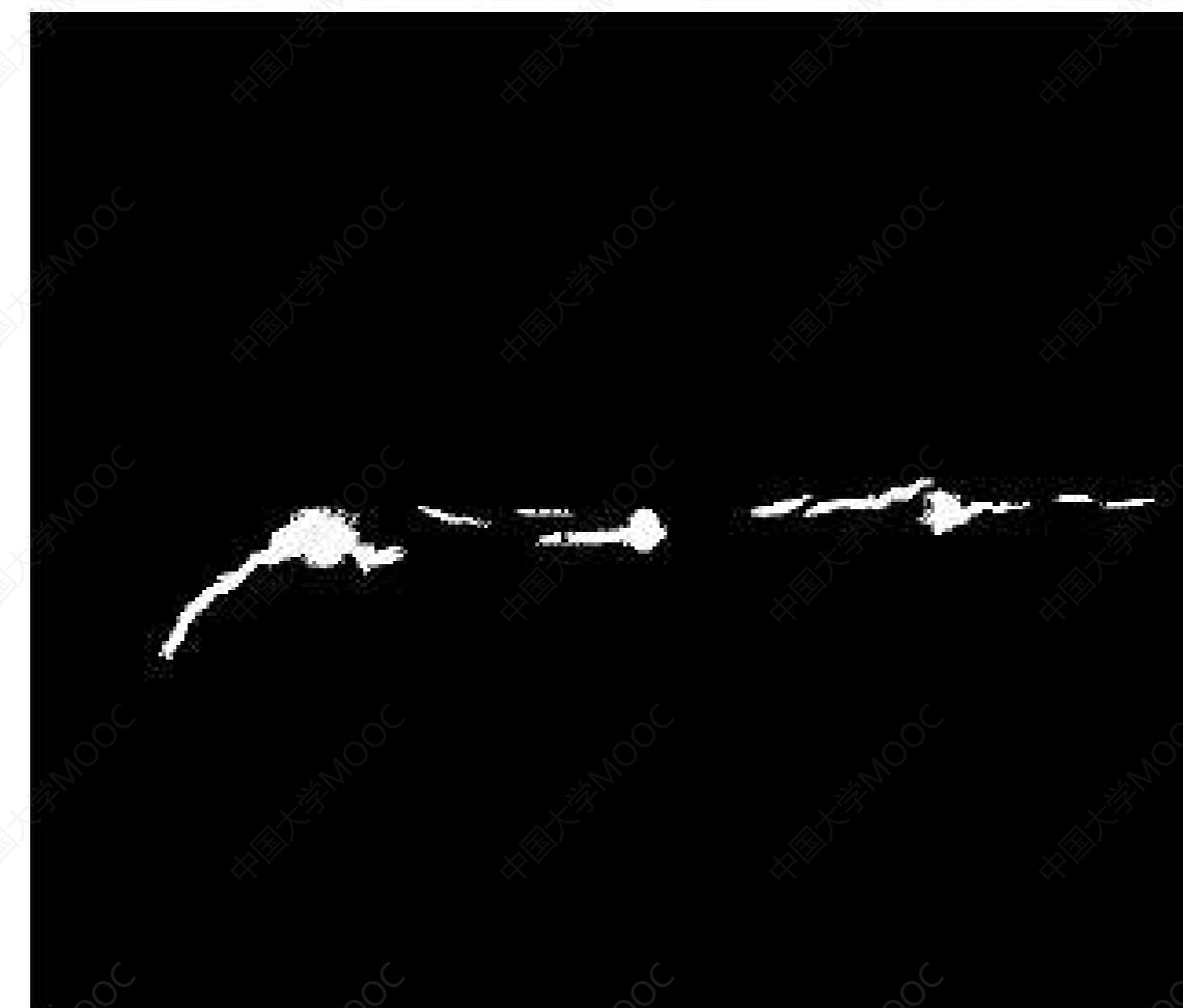
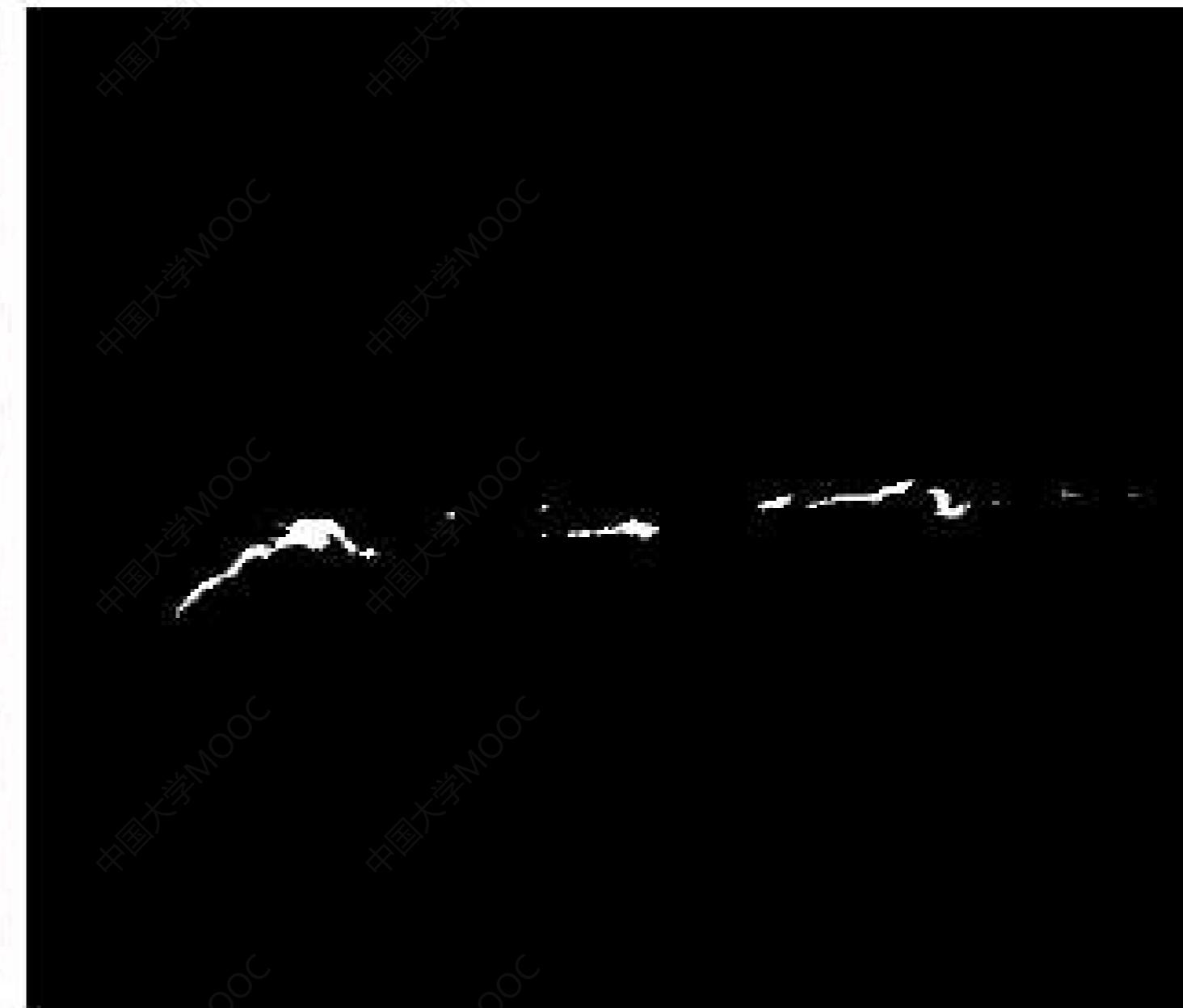
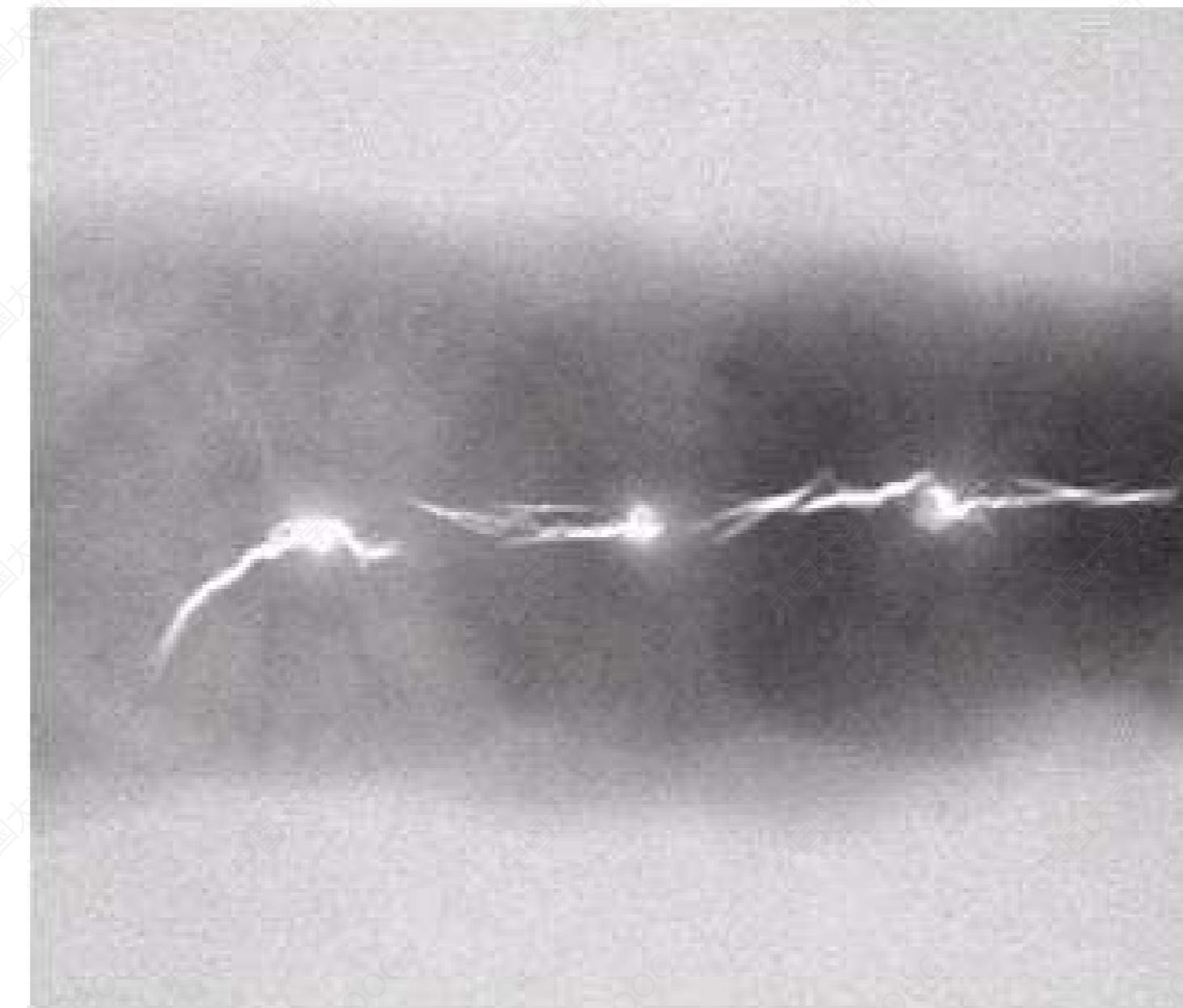
➤ 区域生长法

- 从某个像素（种子点）出发，按照一定的准则，逐步加入邻近像素，当满足一定的条件时，区域生长终止。
- 三个要素
 - 种子点、生长准则、终止准则
- 种子点：区域生长的起始点；
- 生长准则：加入邻近像素点原则，大部分区域生长准则使用图像的局部性质，不同的生长准则会影响区域生长的过程、结果；
- 终止准则：区域生长过程停止的原则。

a b
c d

FIGURE 10.40

(a) Image showing defective welds. (b) Seed points. (c) Result of region growing. (d) Boundaries of segmented defective welds (in black). (Original image courtesy of X-TEK Systems, Ltd.).



■ 7.3 基于区域的方法

➤ 区域分裂与合并算法

- 基本思路类似于无穷分割，然后将分割后满足相似度准则的区域进行合并。
- 基本算法
 - ✓ 对图像中灰度级不同的区域，均分为4个子区域
 - ✓ 若相邻的子区域所有像素的灰度级相同，则将其合并
 - ✓ 重复1~2步骤，直到不再有新的分裂与合并为止

■ 7.4 基于形态学分水岭的方法

分水岭分割方法，是一种基于拓扑理论的数学形态学的分割方法，其基本思想是把图像看作是测地学上的拓扑地貌，图像中每一点像素的灰度值表示该点的海拔高度，每一个局部极小值及其影响区域称为集水盆，而集水盆的边界则形成分水岭。

■ 7.4 基于形态学分水岭的方法

➤ 图像中的三类点

- 局部极小点
- 只能往一个极小点逼近的点
- 等概率的向多个极小点逼近的点

➤ 满足第三个条件的点为“分水线”，图像分割的目的是找到“分水线”

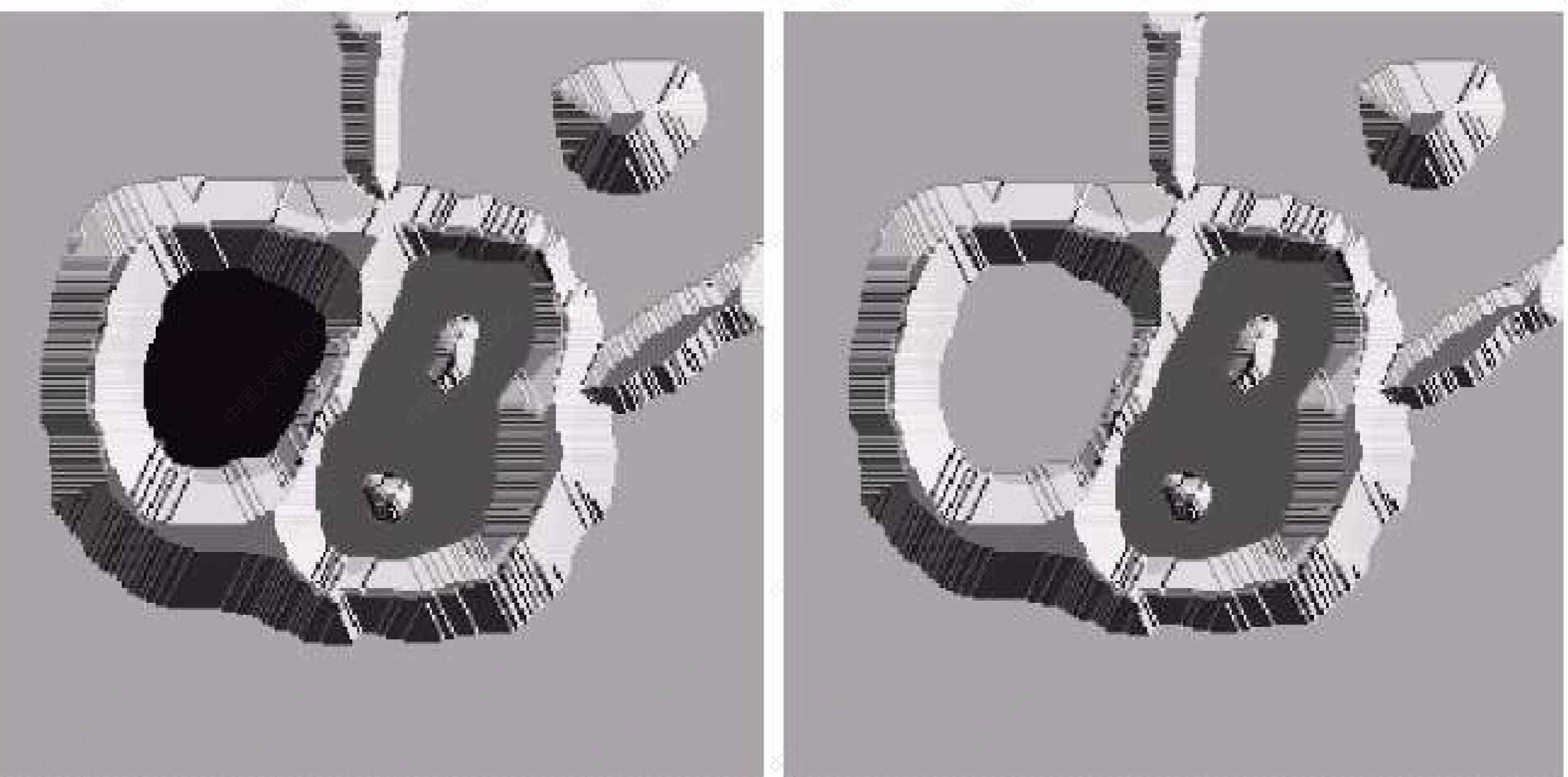
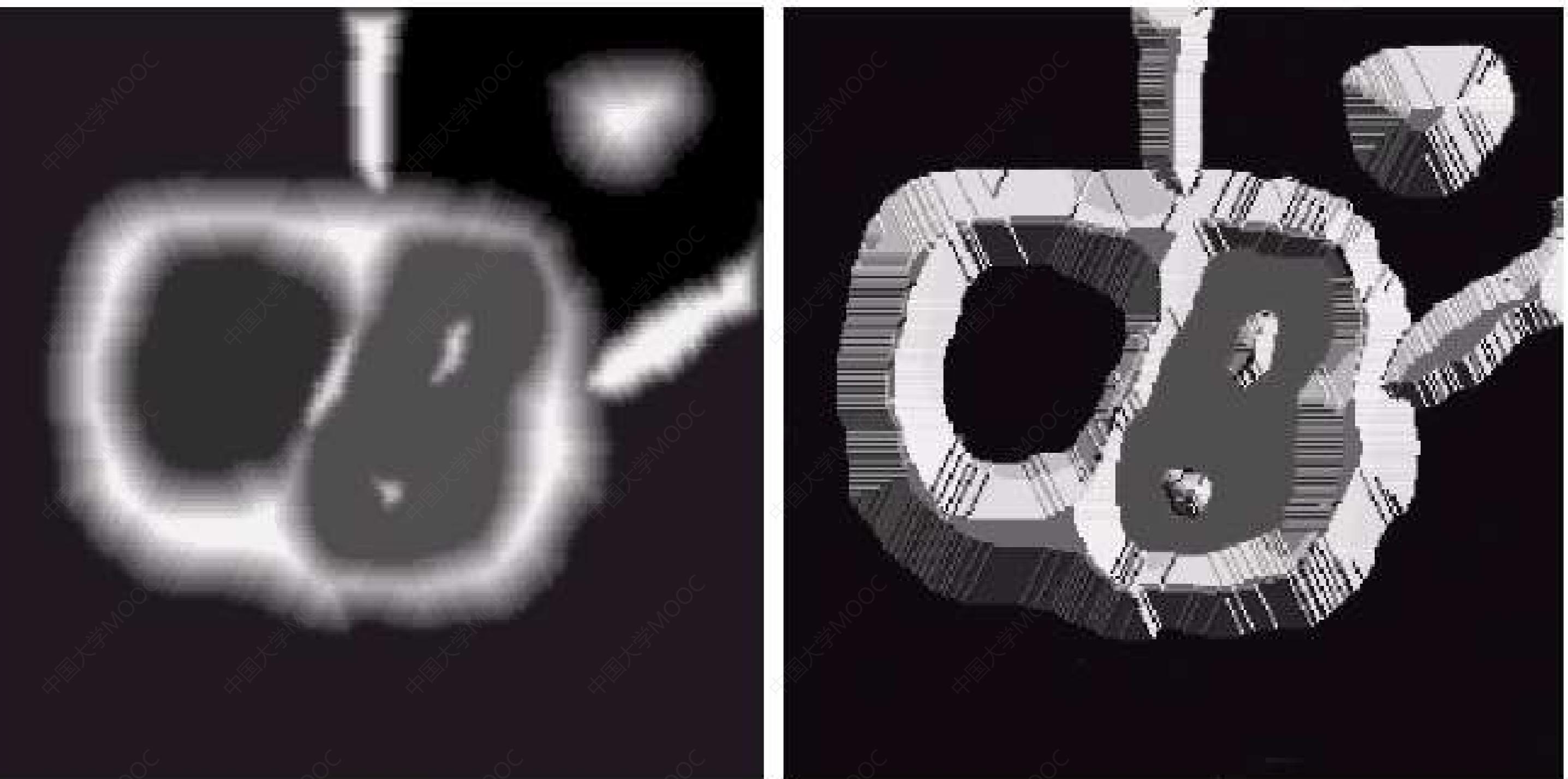
■ 7.4 基于形态学分水岭的方法

- 基于分水岭的图像分割算法的主要目标是找出分水线。
- 基本思想
 - 假设在每个区域最小值的位置上打一个洞并且让水以均匀的上升速率从洞中涌出，从低到高淹没整个地形。当处在不同的汇聚盆地中的水将要聚在一起时，修建的大坝将阻止聚合。水将只能到达大坝的顶部处于水线之上的程度。这些大坝的边界对应于分水岭的分割线。所以，它们是由分水岭算法提取出来的(连续的)边界线。

a
b
c
d

FIGURE 10.44

- (a) Original image.
(b) Topographic view. (c)–(d) Two stages of flooding.

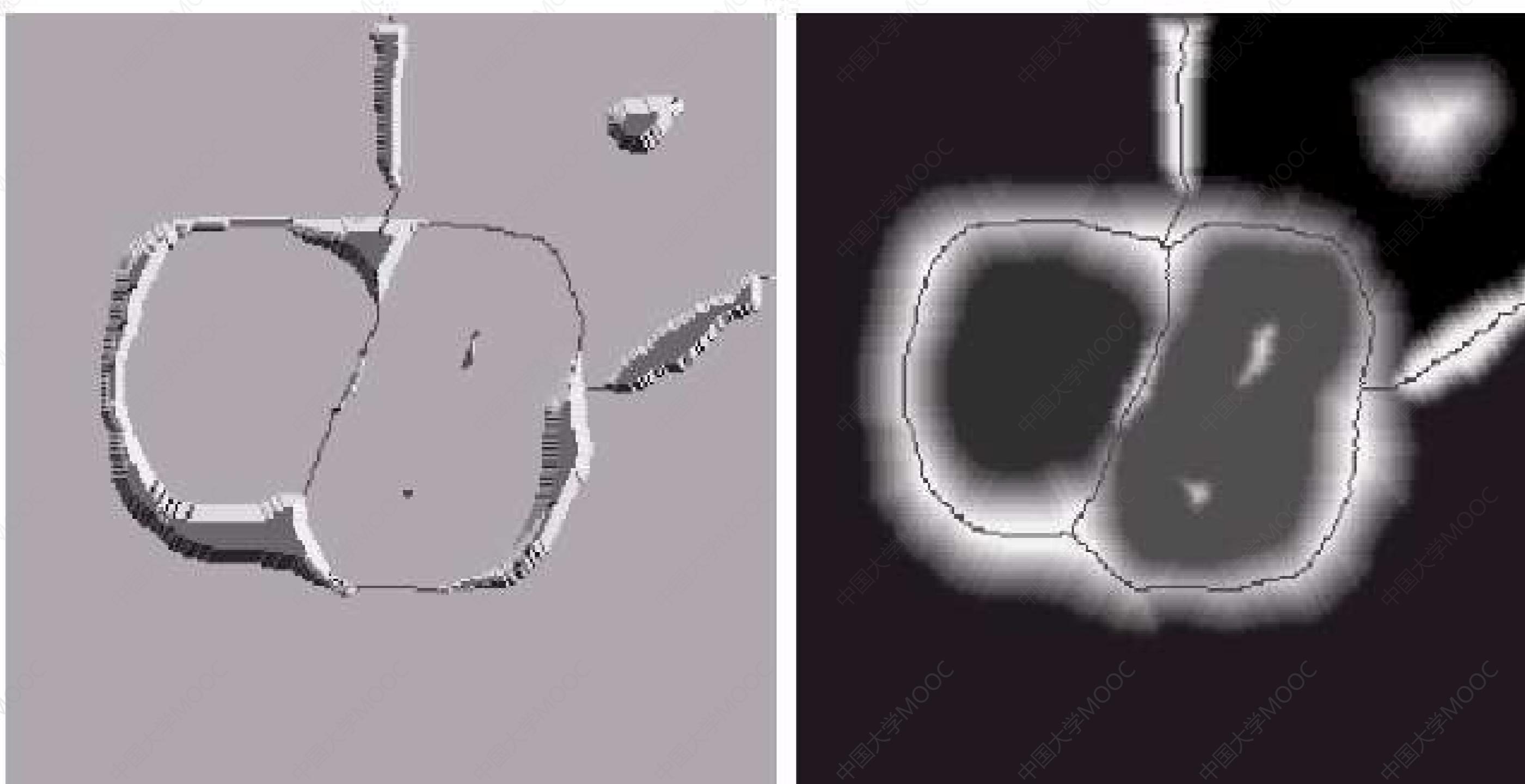
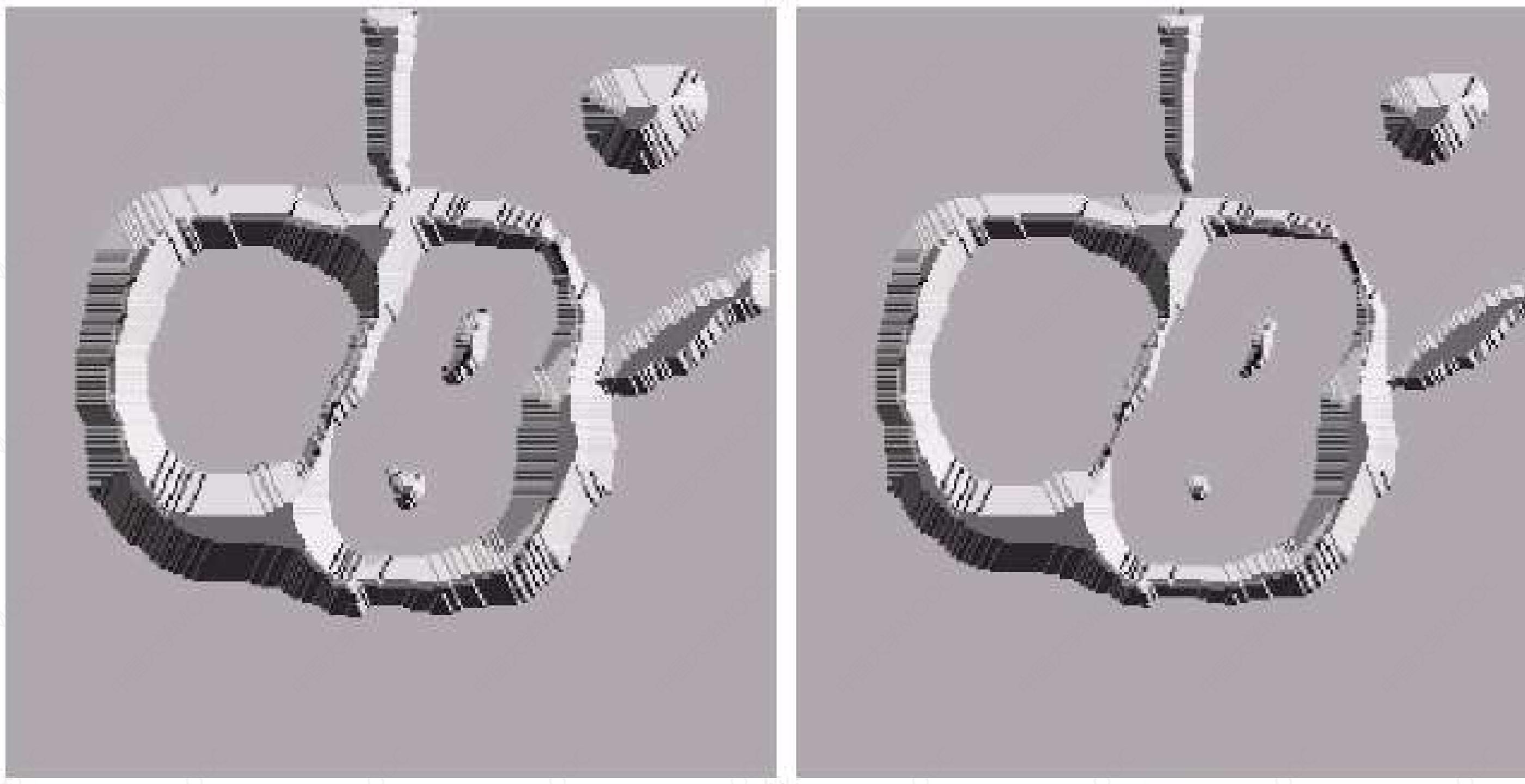


e f
g h

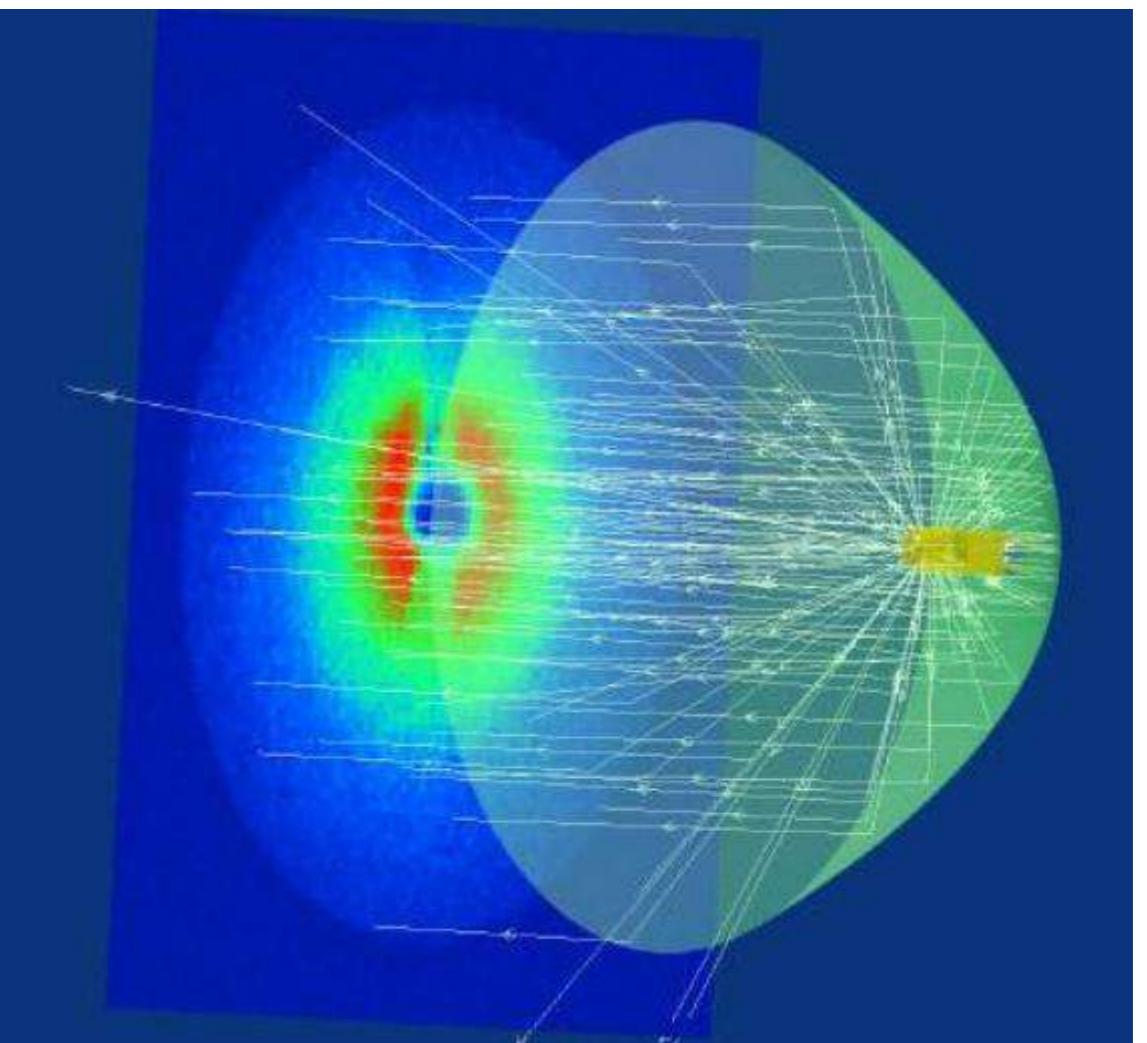
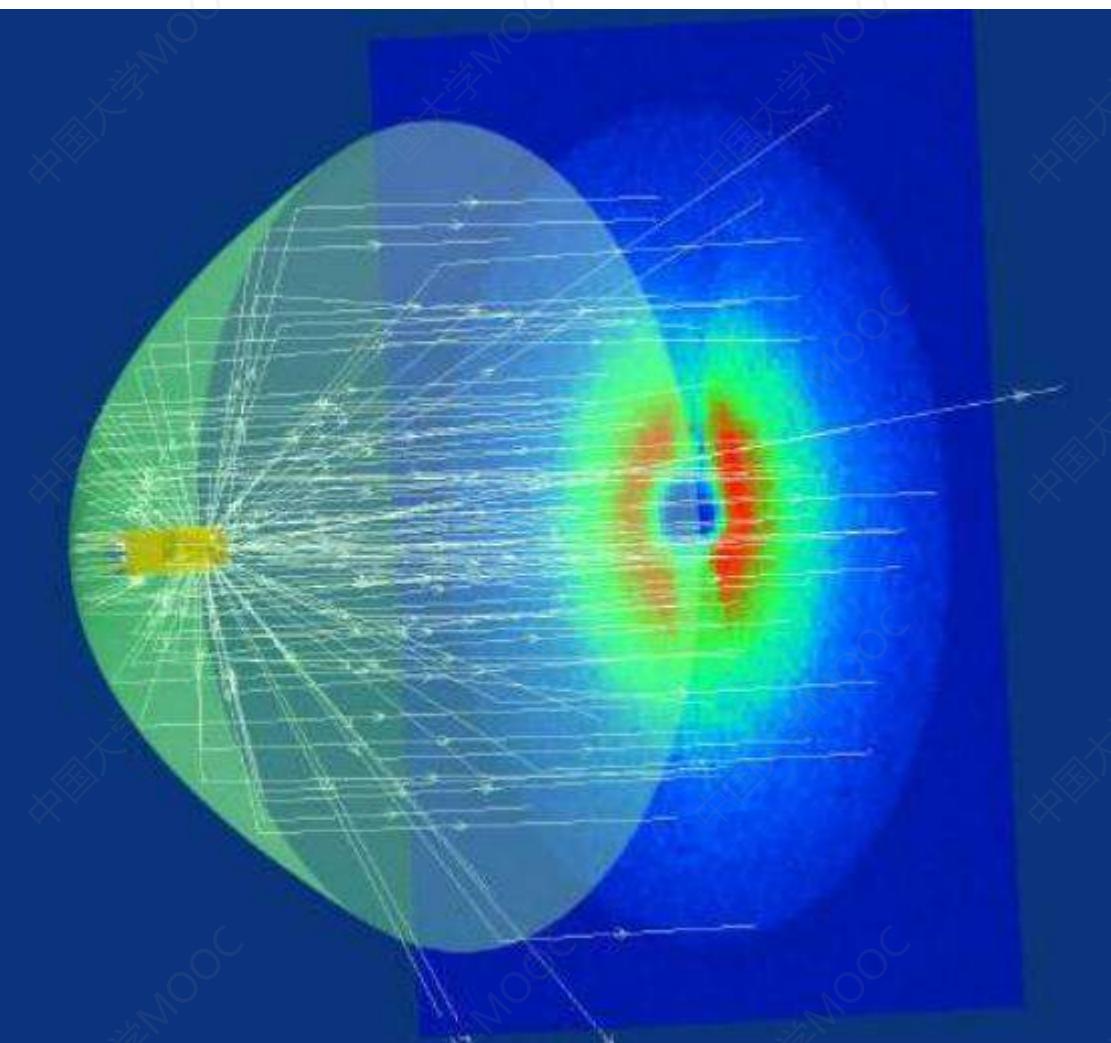
FIGURE 10.44

(Continued)

(e) Result of further flooding.
(f) Beginning of merging of water from two catchment basins (a short dam was built between them). (g) Longer dams. (h) Final watershed (segmentation) lines. (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)



7.5 基于学习的图像分割方法



■ 7.5 基于学习的图像分割方法

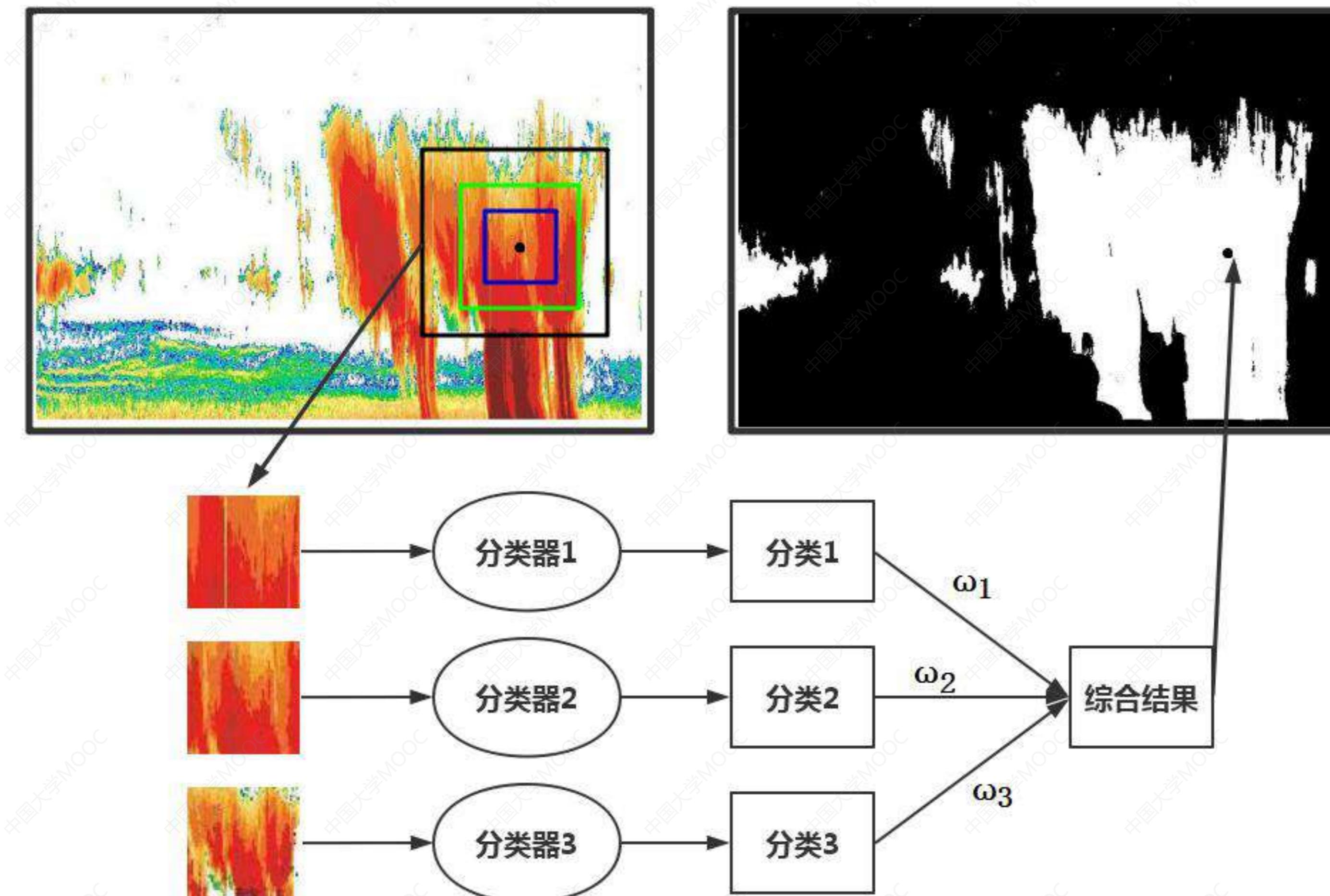
➤ 图像分割的本质——像素点属性分类

- 前景1
- 前景2
- 前景3.....
- 前景N
- 背景

➤ 由像素点获得多尺度图像，把像素点属性分类任务转化为多尺度图像分类任务

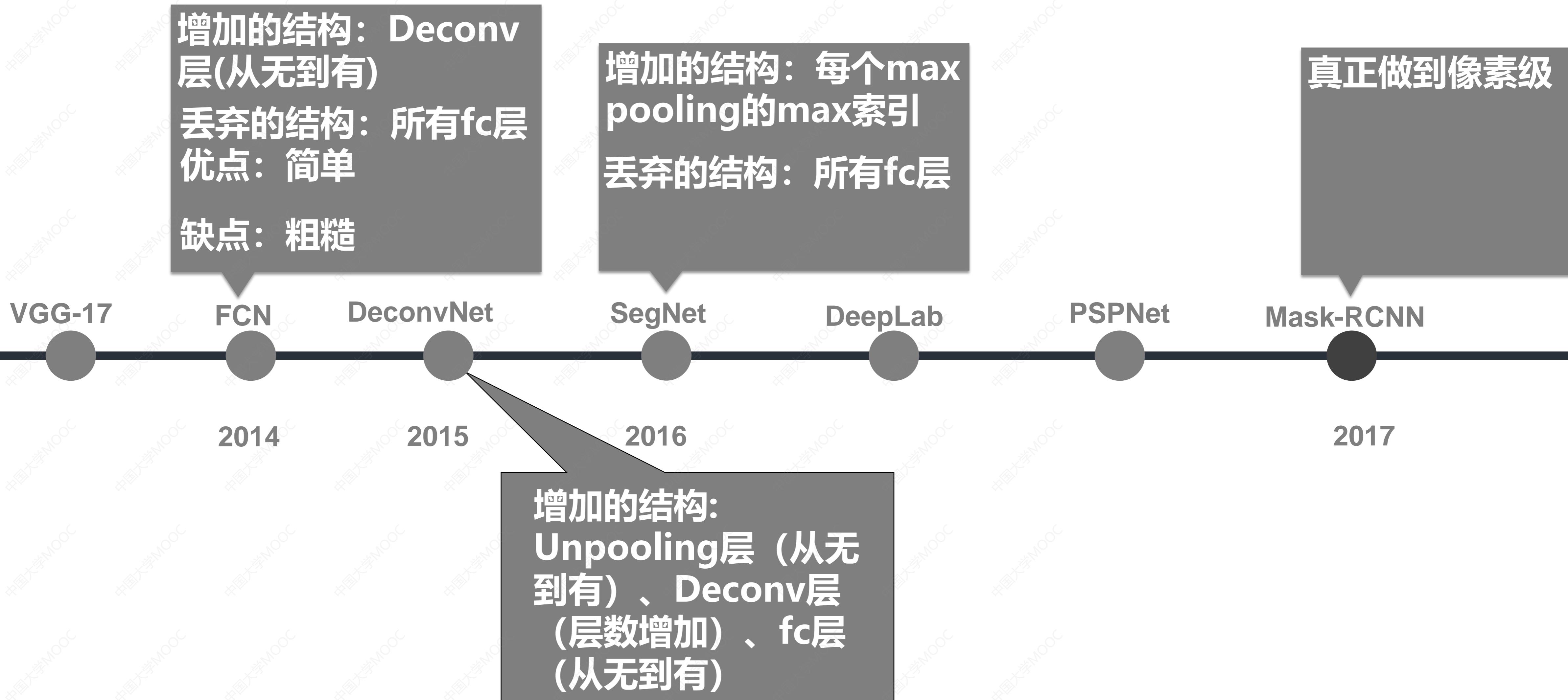
- 基于MR-CNN的分类方法
- 基于SP-CNN的分类方法
- FCN-CNN相结合的分类方法

7.5 基于学习的图像分割方法

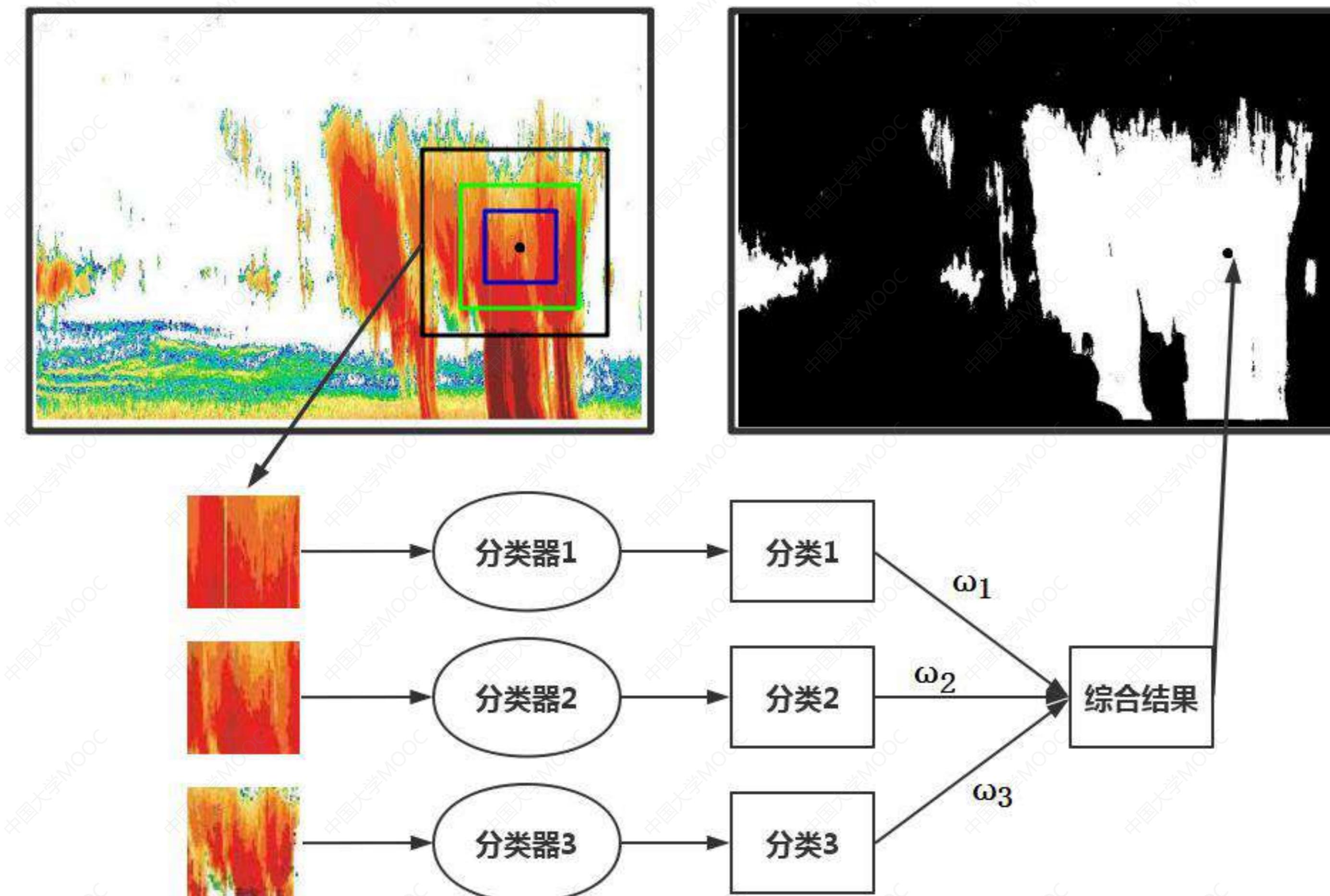


MR-CNN (多分辨率CNN)

7.5 基于学习的图像分割方法

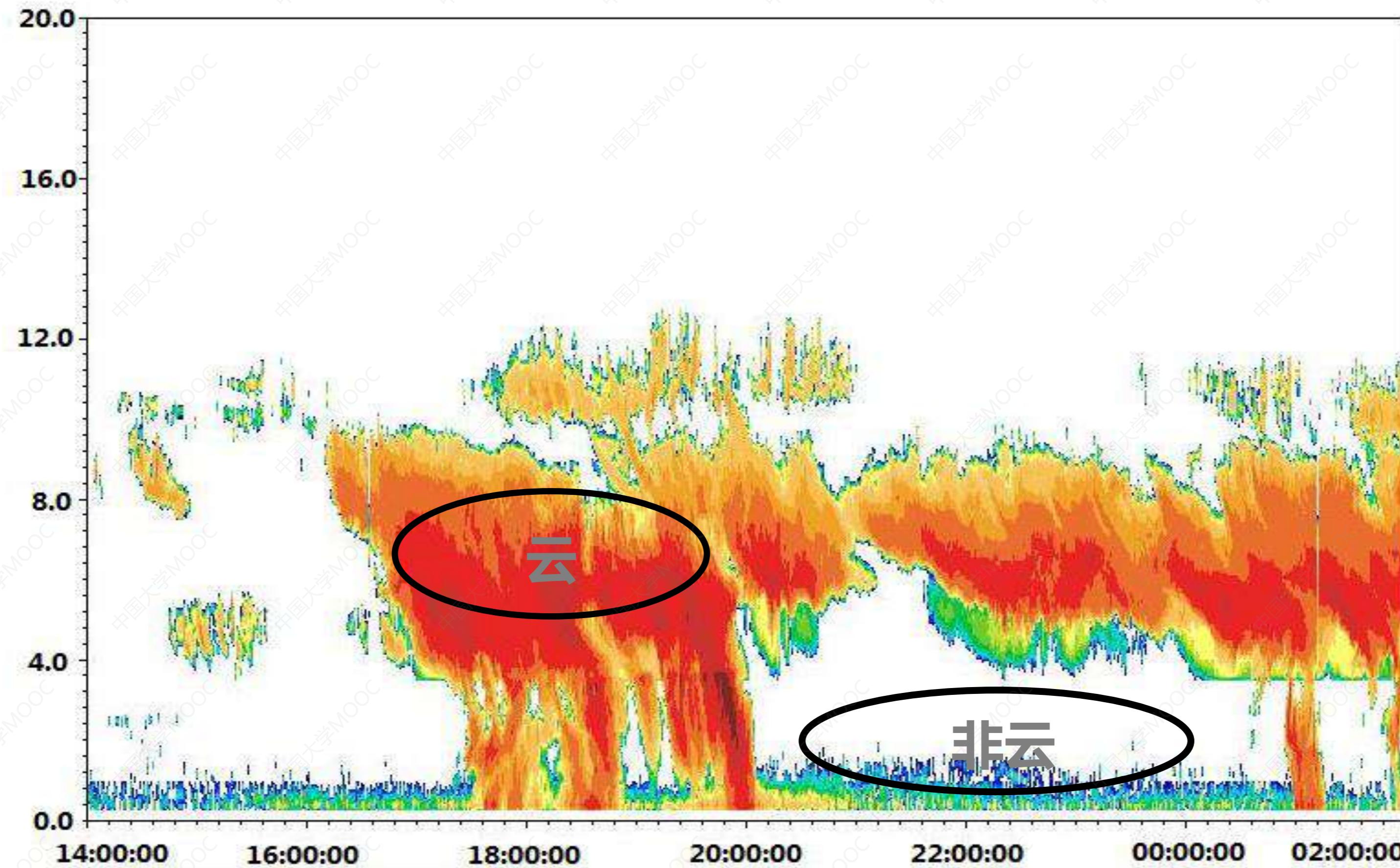


7.5 基于学习的图像分割方法

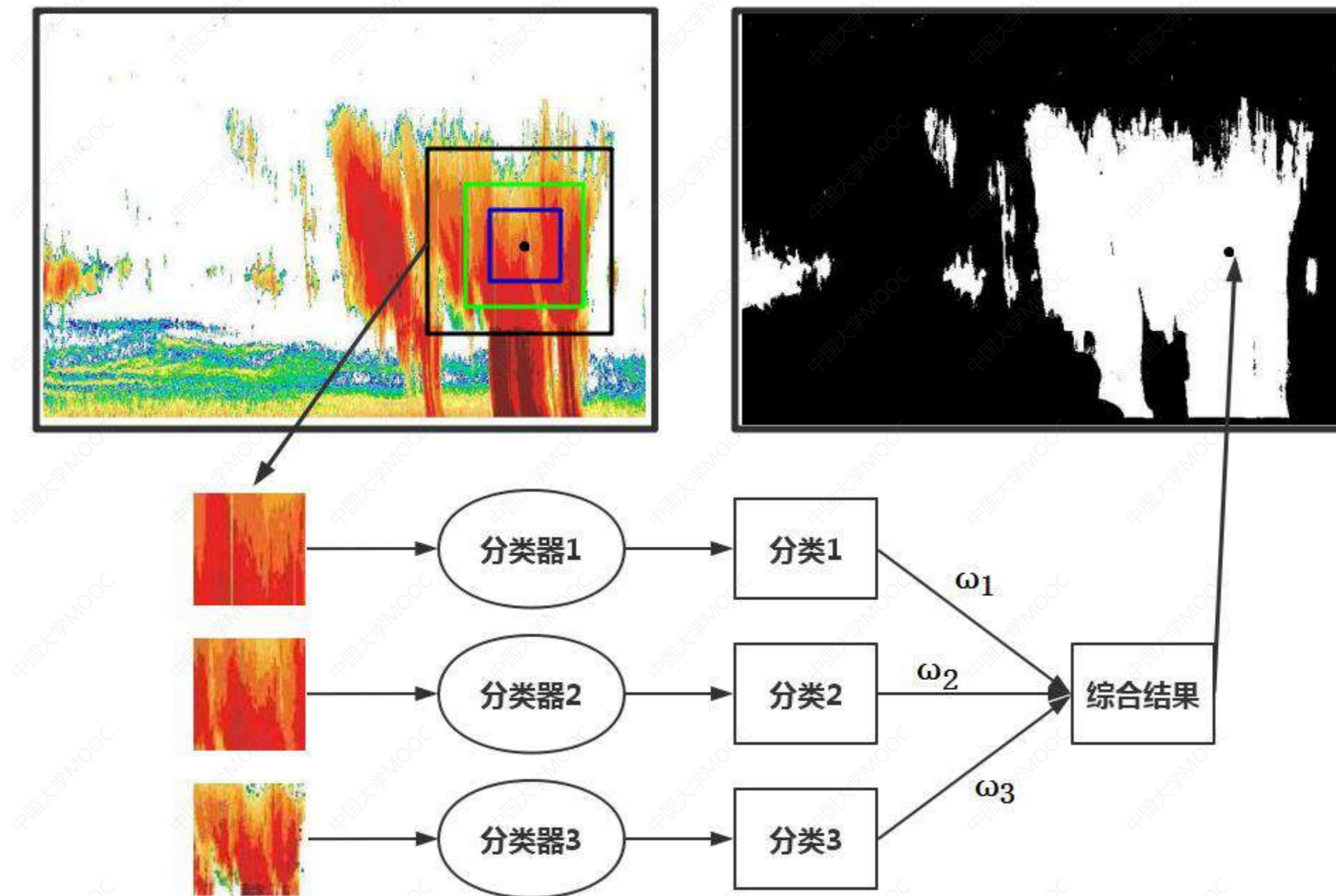


MR-CNN (多分辨率CNN)

7.5 基于学习的图像分割方法



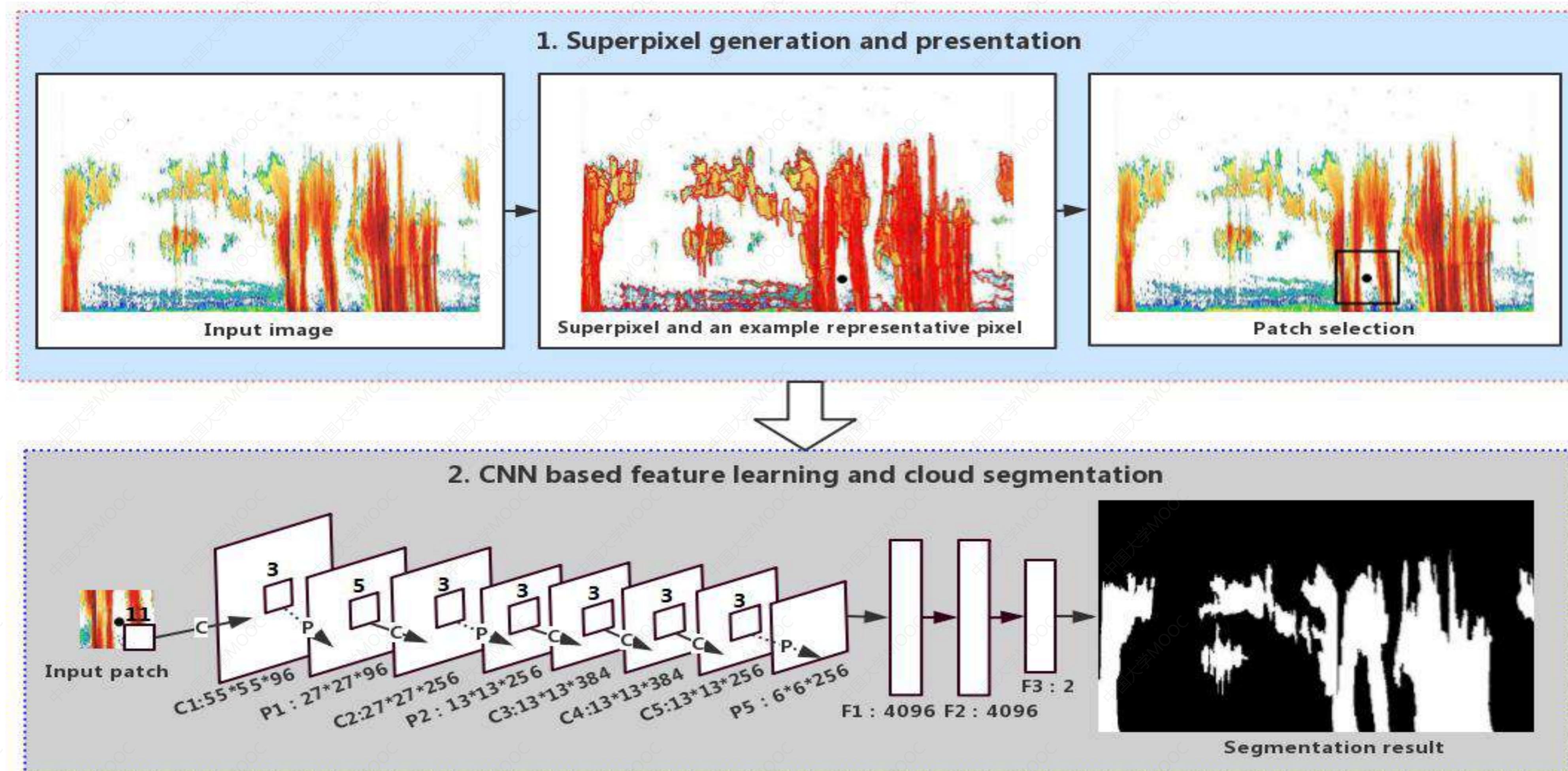
7.5 基于学习的图像分割方法



MR-CNN

Wu L, He J, Jian M, et al. Fast cloud image segmentation with superpixel analysis based convolutional networks[C]// International Conference on Systems, Signals and Image Processing. IEEE, 2017:1-5. DOI:[10.1109/IWSSIP.2017.7965591](https://doi.org/10.1109/IWSSIP.2017.7965591)

7.5 基于学习的图像分割方法



SP-CNN

Wu L, He J, Jian M, et al. Fast cloud image segmentation with superpixel analysis based convolutional networks[C]// International Conference on Systems, Signals and Image Processing. IEEE, 2017:1-5. DOI:[10.1109/IWSSIP.2017.7965591](https://doi.org/10.1109/IWSSIP.2017.7965591)

7.5 基于学习的图像分割方法

SP-CNN 过程中随机与等间隔选取关键点的性能比较

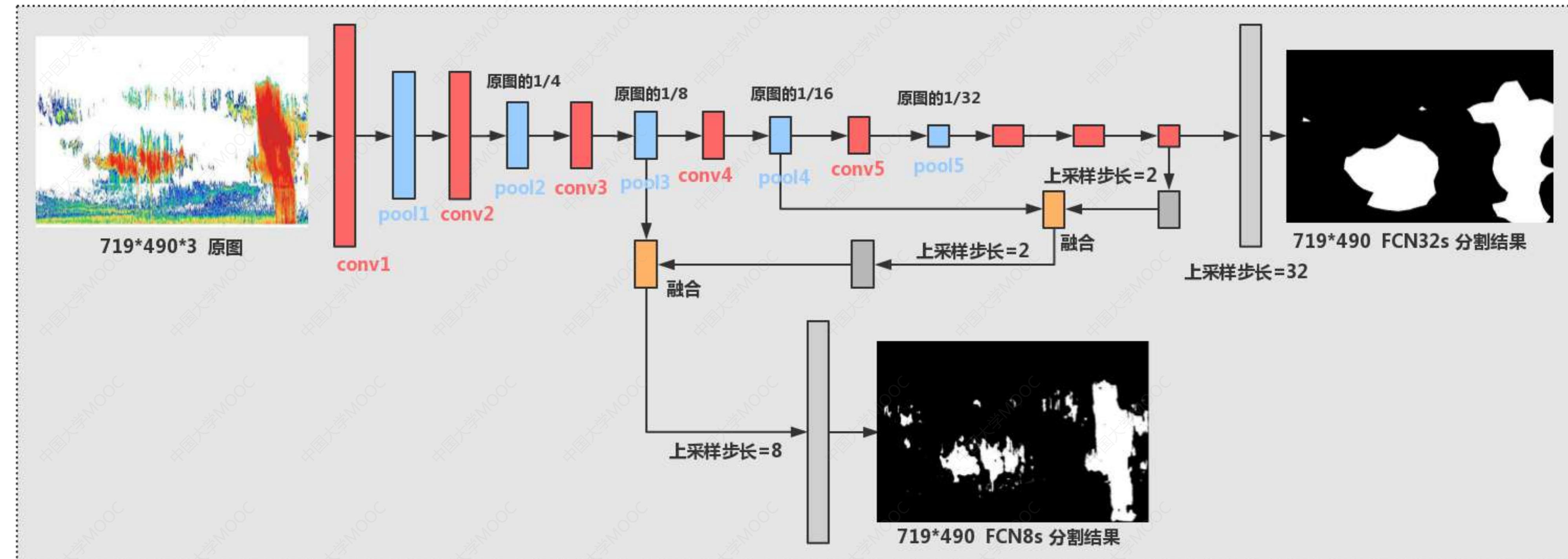
	随机选取关键点			等间隔选取关键点		
	$n = 1$	$n = 3$	$n = 5$	$n = 1$	$n = 3$	$n = 5$
准确度(%)	99.171 ± 0.027	99.422 ± 0.011	99.433 ± 0.003	98.821 ± 0.047	99.476 ± 0.002	99.543 ± 0.003
时间	9.12	21.12	33.12	9.12	21.12	33.12

云图分割 SP-CNN 方法与其他方法的性能比较

	阈值法[11]	MR-CNN[15]	FCN32s[17]	FCN8s[17]	SP-CNN
准确度 (%)	96.751 ± 0.003	99.360 ± 0.005	94.441 ± 1.456	98.630 ± 0.002	99.543 ± 0.003
交并比 IOU (%)	60.253 ± 3.346	83.744 ± 0.492	65.553 ± 3.717	55.884 ± 2.692	80.132 ± 0.512
时间 (秒)	-	17615	4.23	4.94	33.12

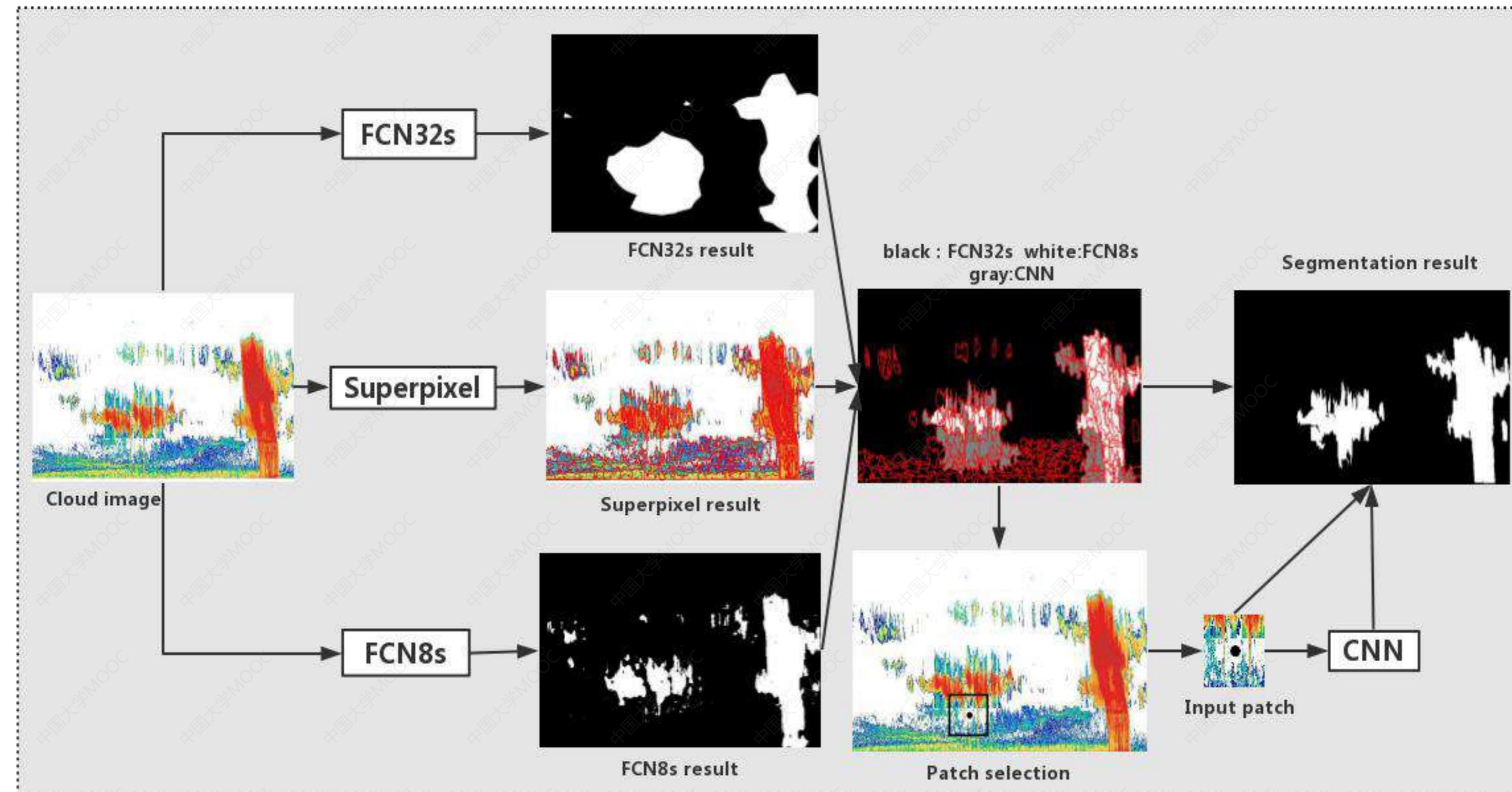
Wu L, He J, Jian M, et al. Fast cloud image segmentation with superpixel analysis based convolutional networks[C]// International Conference on Systems, Signals and Image Processing. IEEE, 2017:1-5. DOI:[10.1109/IWSSIP.2017.7965591](https://doi.org/10.1109/IWSSIP.2017.7965591)

7.5 基于学习的图像分割方法



FCN结果示例

7.5 基于学习的图像分割方法



FCN-CNN

7.5 基于学习的图像分割方法

CNN 处理的平均超像素数量及时间统计

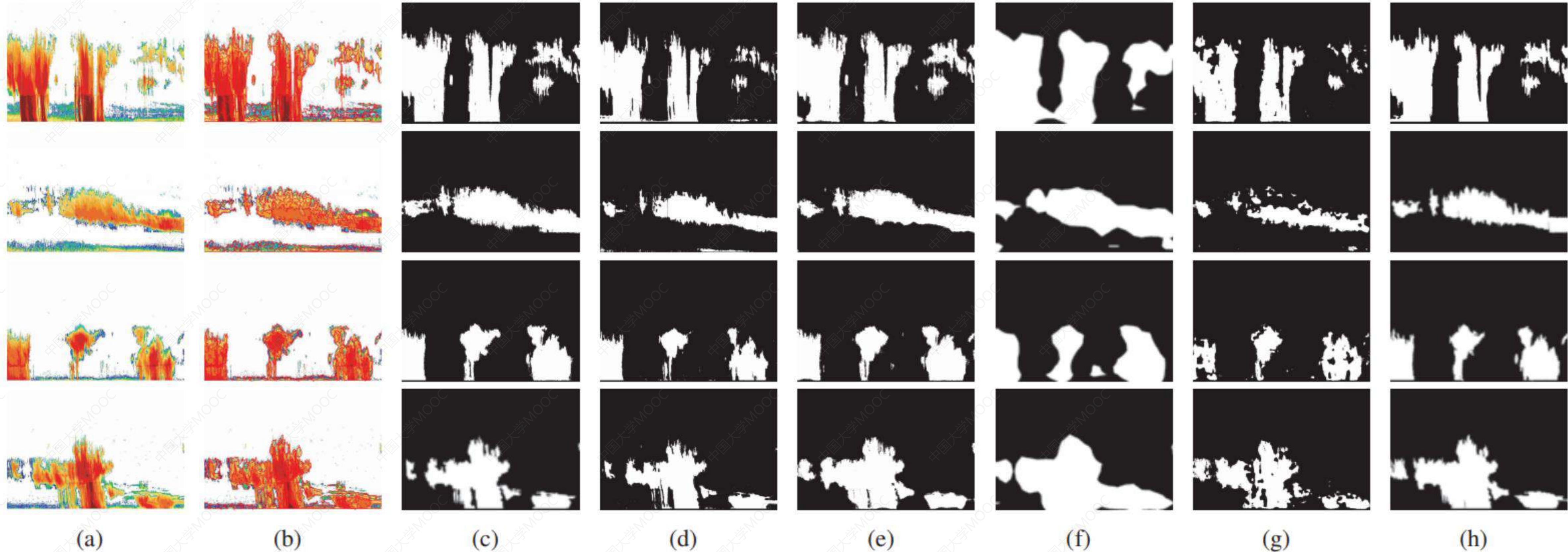
方法	SP-CNN	FCN8s-CNN	FCN32s-CNN	FCN-CNN
需要 CNN 处理的超像素(个)	150	84	103	38.47
CNN 处理所需时间 (秒)	30	16.8	20.6	7.69

本文的云图分割 FCN-CNN 方法与其他方法的性能比较

方法	准确度 (%)	交并比 IoU (%)	时间 (秒)
阈值 ^[13]	96.751 ± 0.003	60.253 ± 3.346	—
MR-CNN ^[15]	99.500 ± 0.005	83.744 ± 0.492	17615
FCN8s ^[17]	98.630 ± 0.002	55.884 ± 2.692	4.23
FCN32s ^[17]	94.441 ± 1.456	65.553 ± 3.717	4.94
SP-CNN	99.543 ± 0.003	80.132 ± 0.512	33.12
FCN8s-CNN	99.517 ± 0.012	79.564 ± 0.976	24.86
FCN32s-CNN	99.543 ± 0.004	80.782 ± 0.785	27.95
FCN-CNN	99.564 ± 0.003	81.235 ± 0.506	19.98

毋立芳,贺娇瑜,简萌,邹蕴真,赵铁松. 局部聚类分析的FCN-CNN云图分割方法. 软件学报, 2018,29(4):1049-1059.

7.5 基于学习的图像分割方法



(a) 原始图像 (b) 超像素结果图像 (c) 人工标注的真实结果 (d) 基于阈值的分割结果 (e) 多分辨率CNN分割结果 (f) FCN32S分割结果 (g) FCN8S分割结果 (h) SP-CNN分割结果

第八章 特征提取和识别

毋立芳

2020.10.28

■ 引言

图像分割技术将图像分割成一个一个独立的区域，这一章的任务是将每个区域映射成为一个语义概念，基本的思路包括特征提取和模式识别两步。

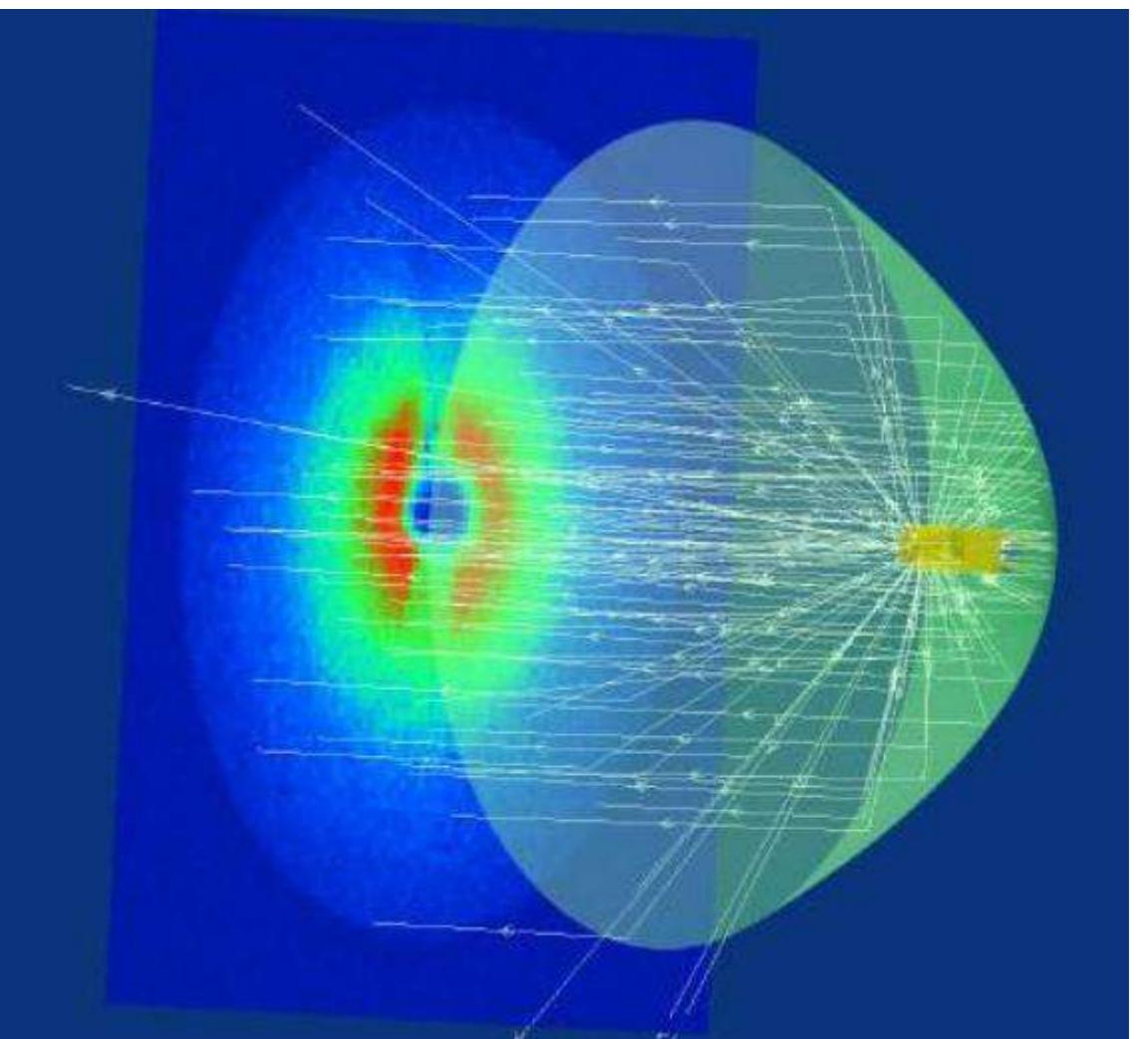
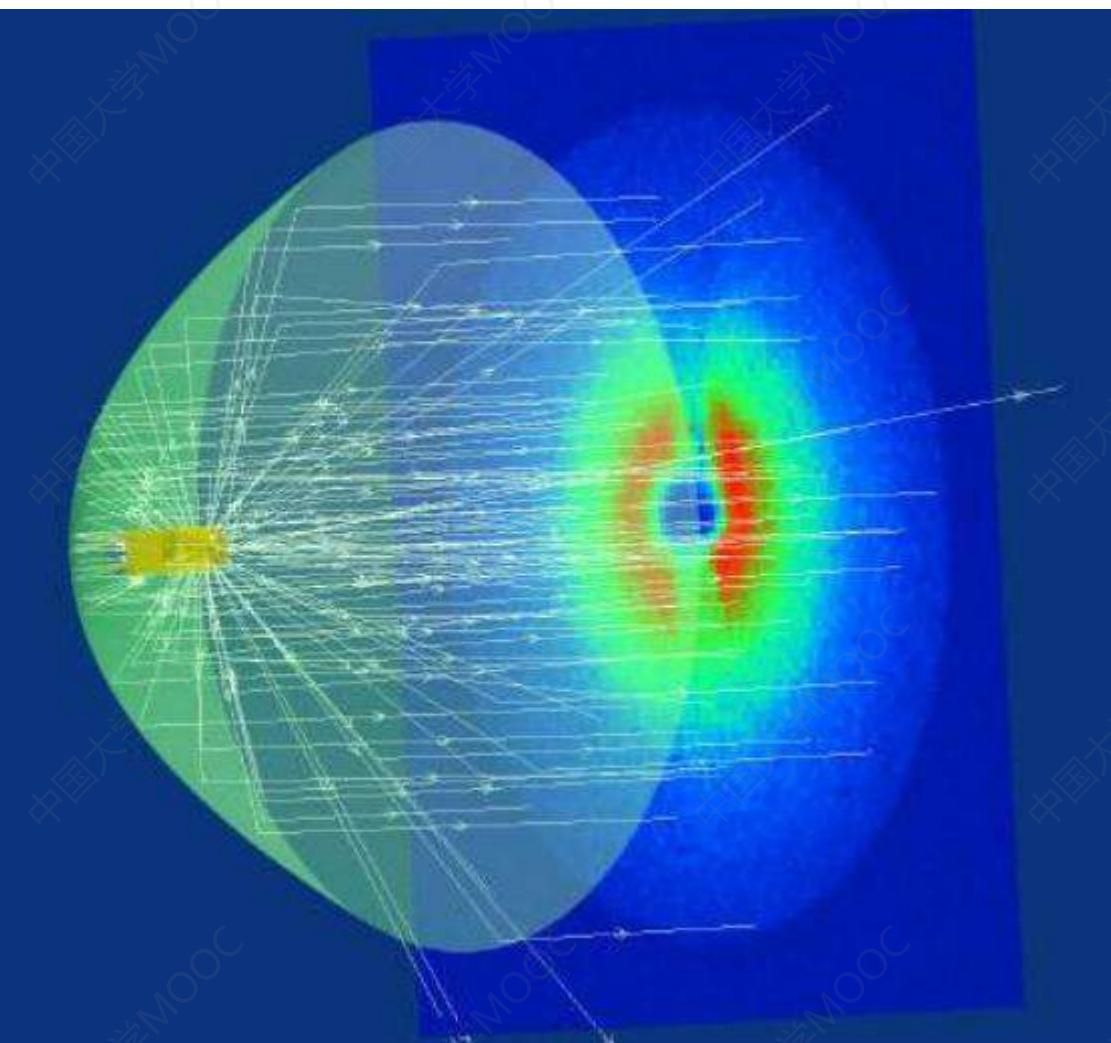
特征提取：将对应区域表达为一个（组）特征向量，用于模式识别、图像检索等任务

模式识别：根据特征向量和语义对象之间的映射关系，输出对应区域的语义概念。

■ 引言

- 8.1 数据级特征表达
- 8.2 形状与颜色描述子
- 8.3 区域描述子
- 8.4 模式识别

8.1 数据级特征表达



■ 8.1 数据级特征表达

- 8.1.1 方向链码

- 8.1.2 多边形近似

- 8.1.3 标记图法

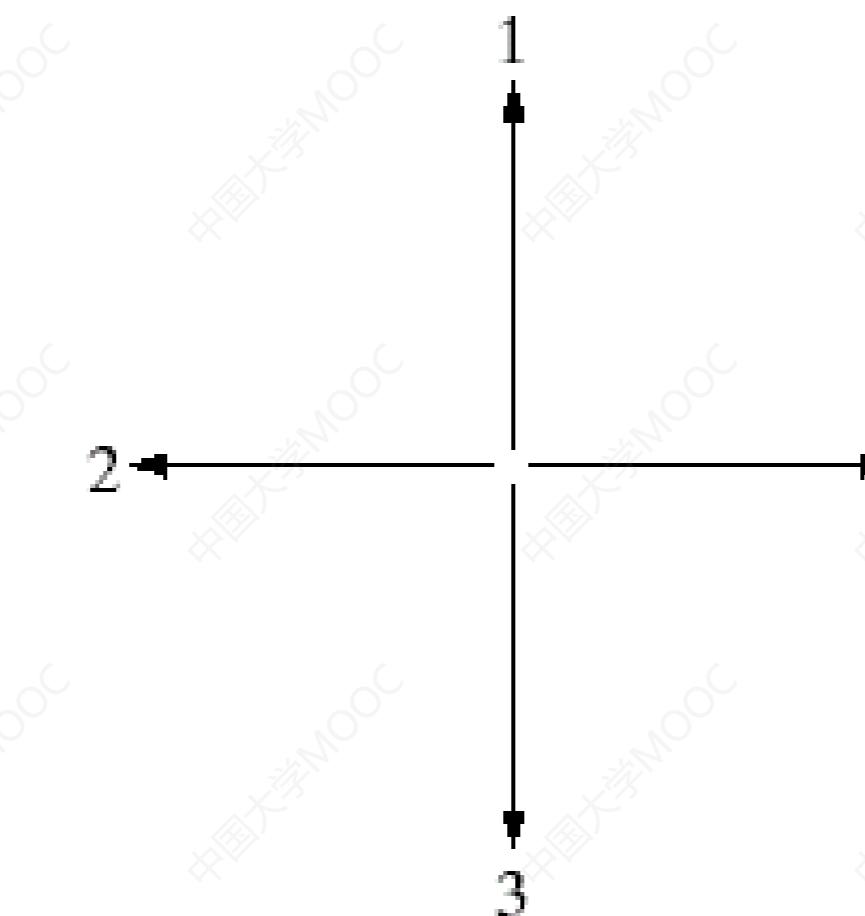
- 8.1.4 边界分段

- 8.1.5 骨架法

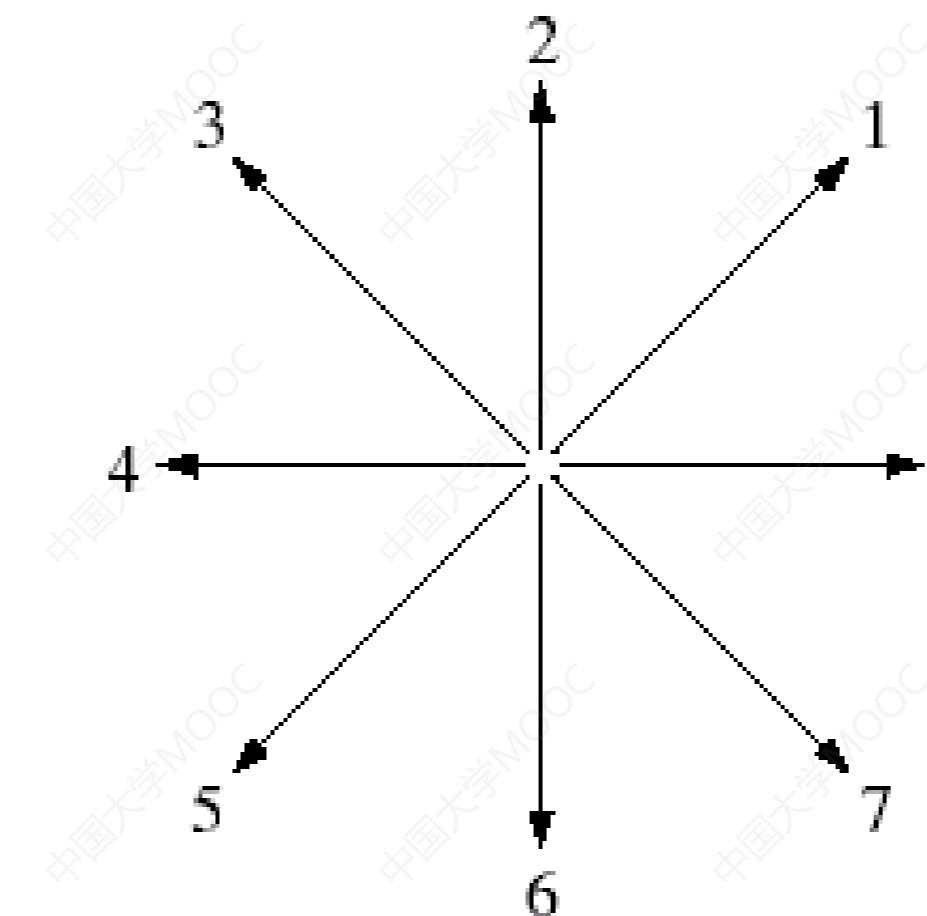
■ 8.1 数据级特征表达

➤ 方向链码

- 把轮廓上相邻像素点的转换关系表达为一种方向码。



四方向编码

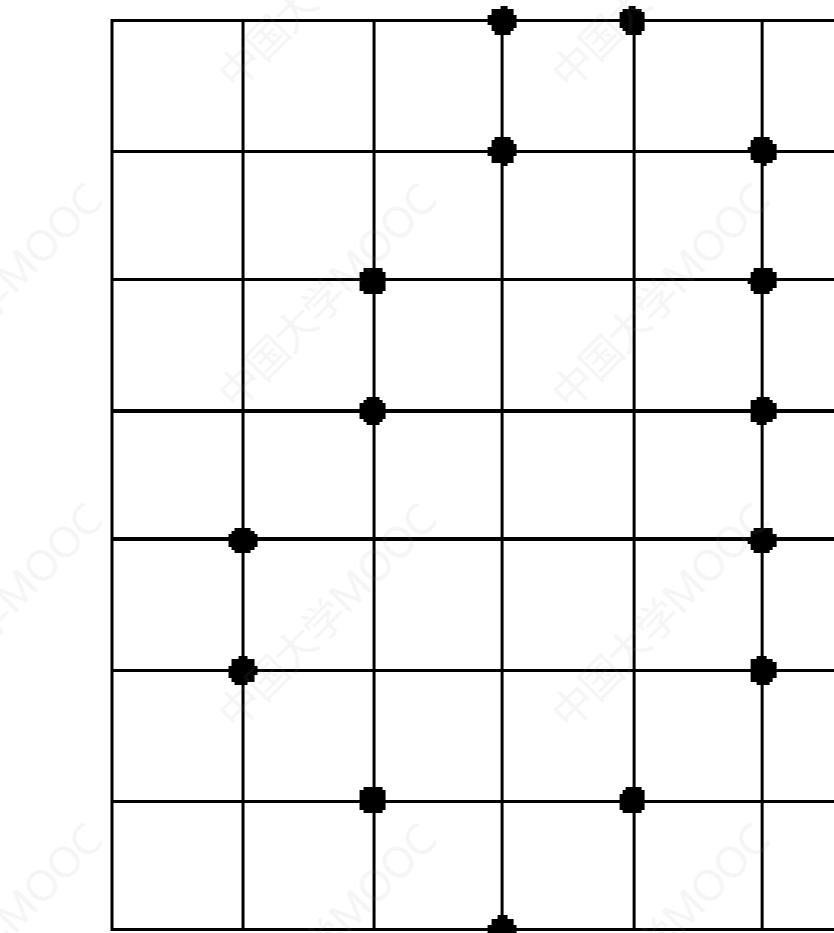
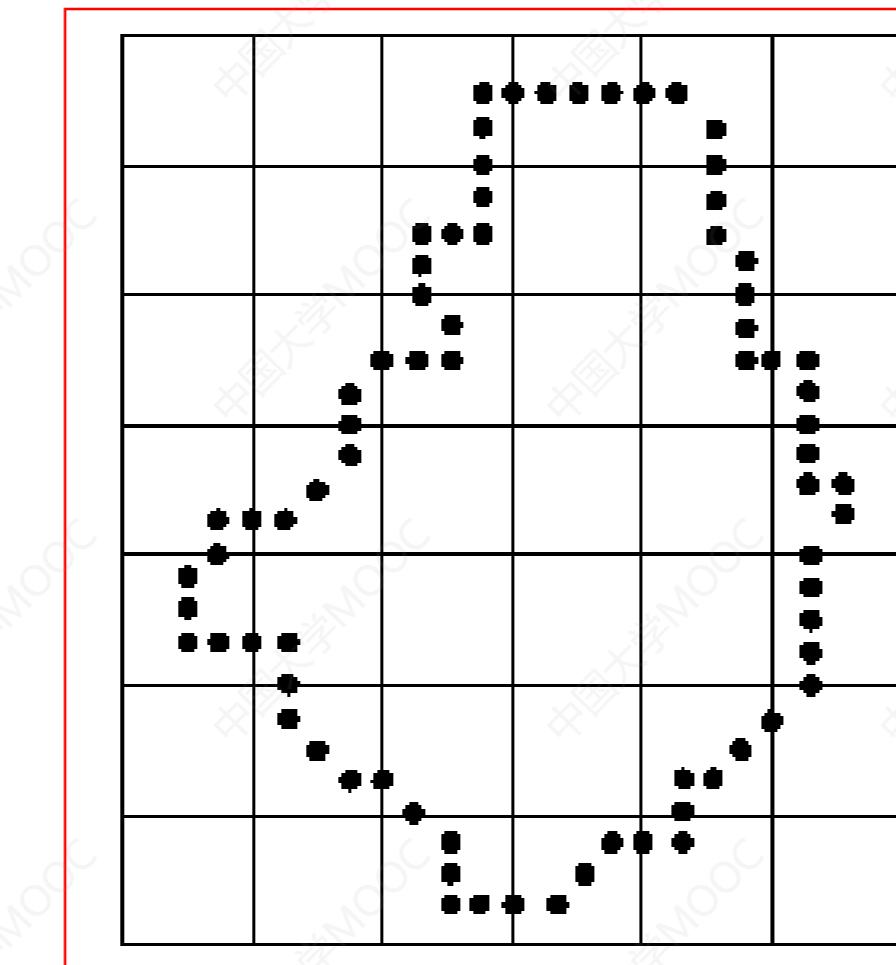


八方向编码

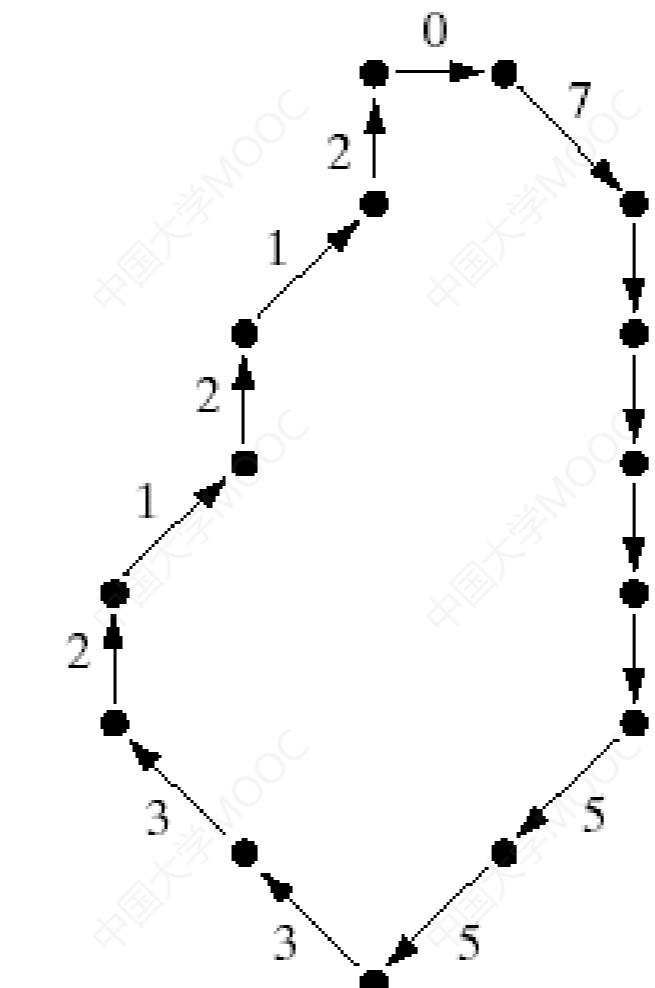
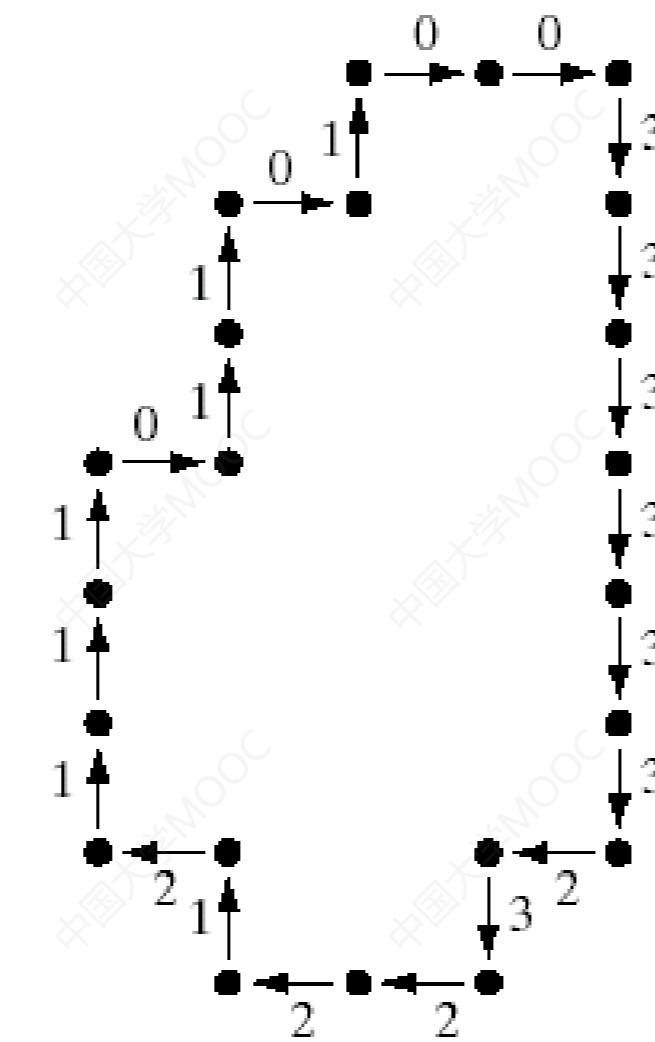
■ 8.1 数据级特征表达

➤ 方向链码

输入
图像



四方向
链码



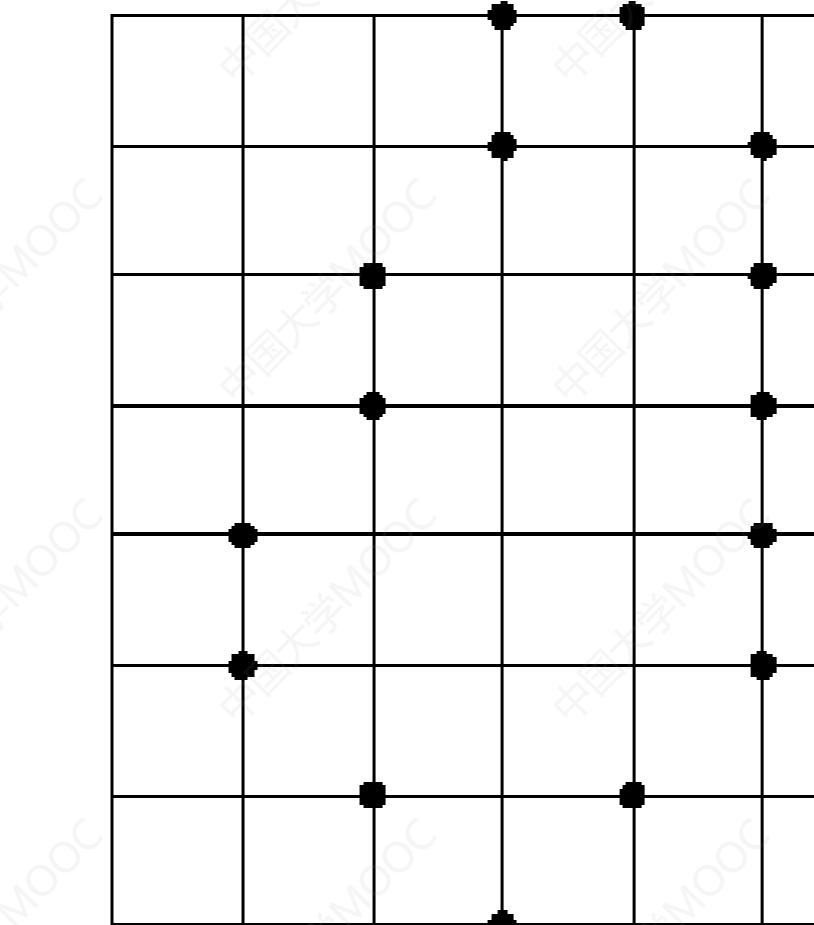
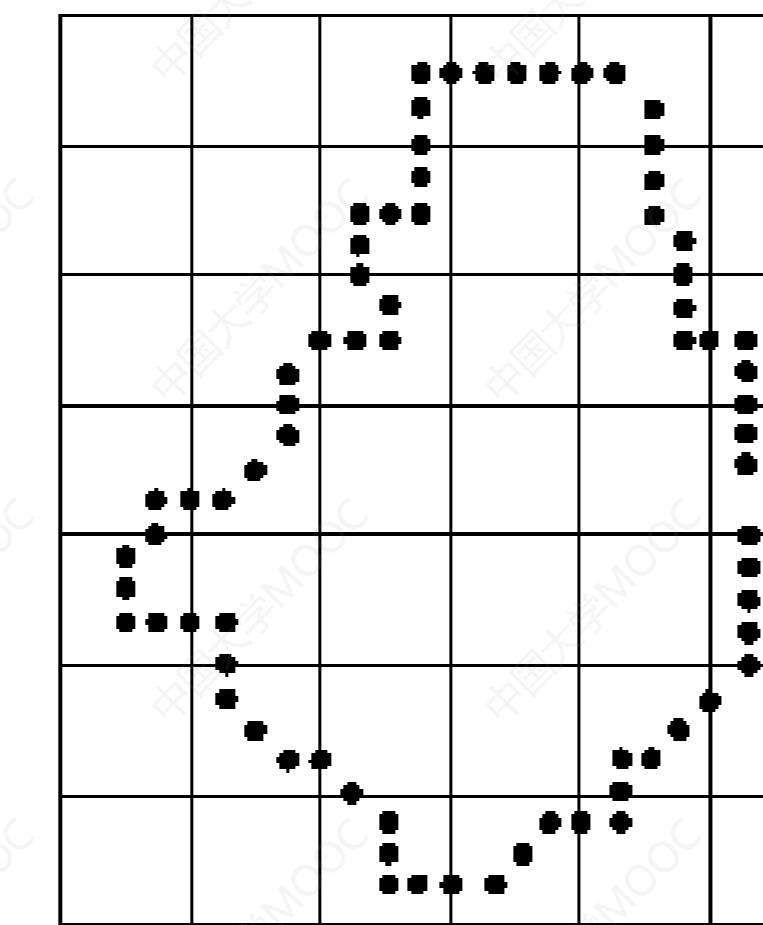
八方向
链码

陈廷标, 夏良正编著, 数字图像处理, 人民邮电出版社, 1990年。

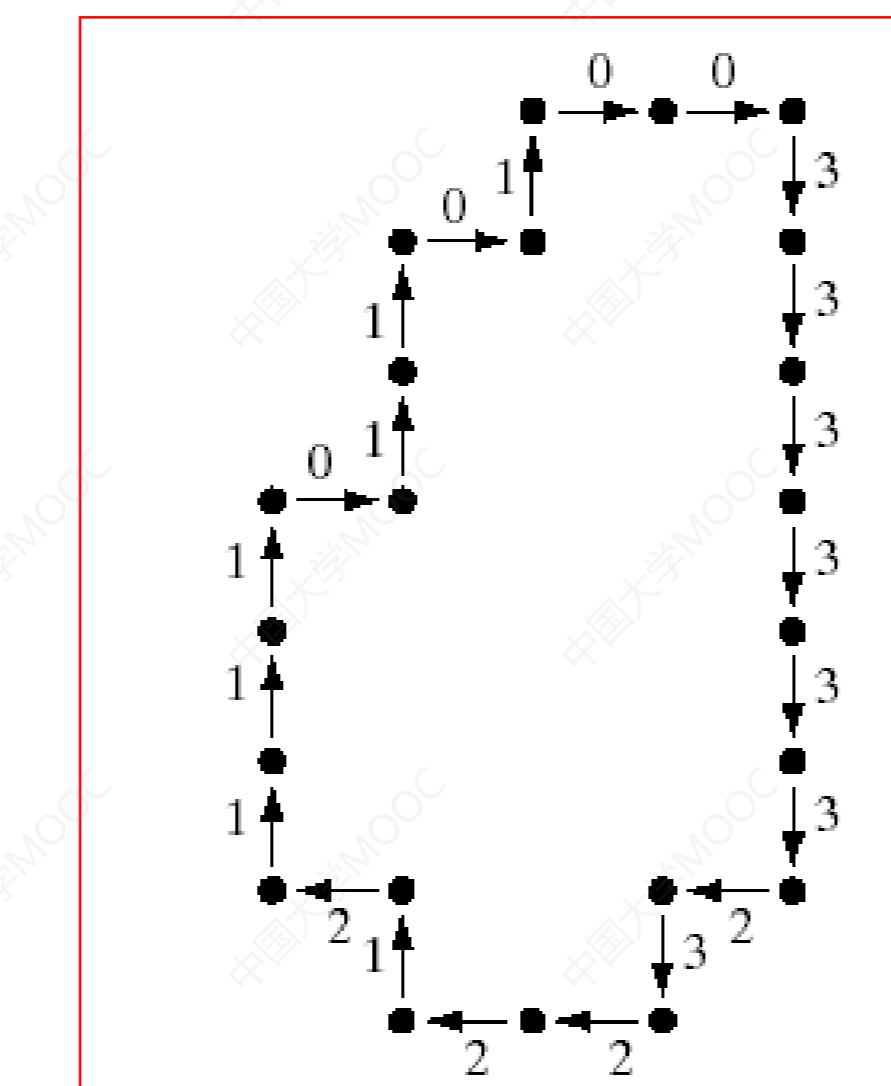
■ 8.1 数据级特征表达

➤ 方向链码

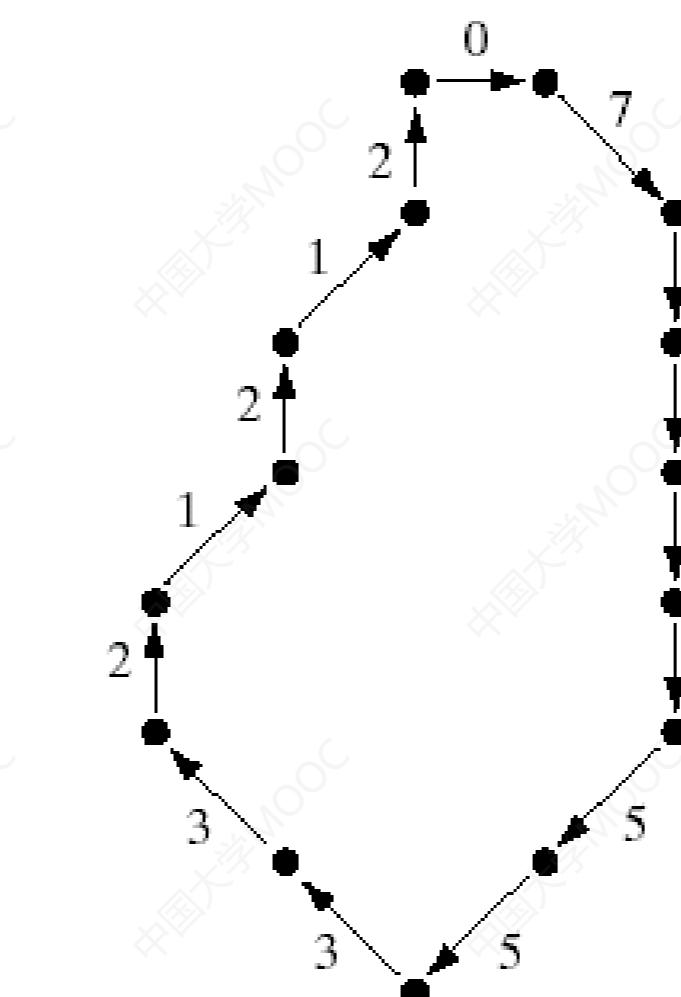
输入
图像



四方向
链码



八方向
链码

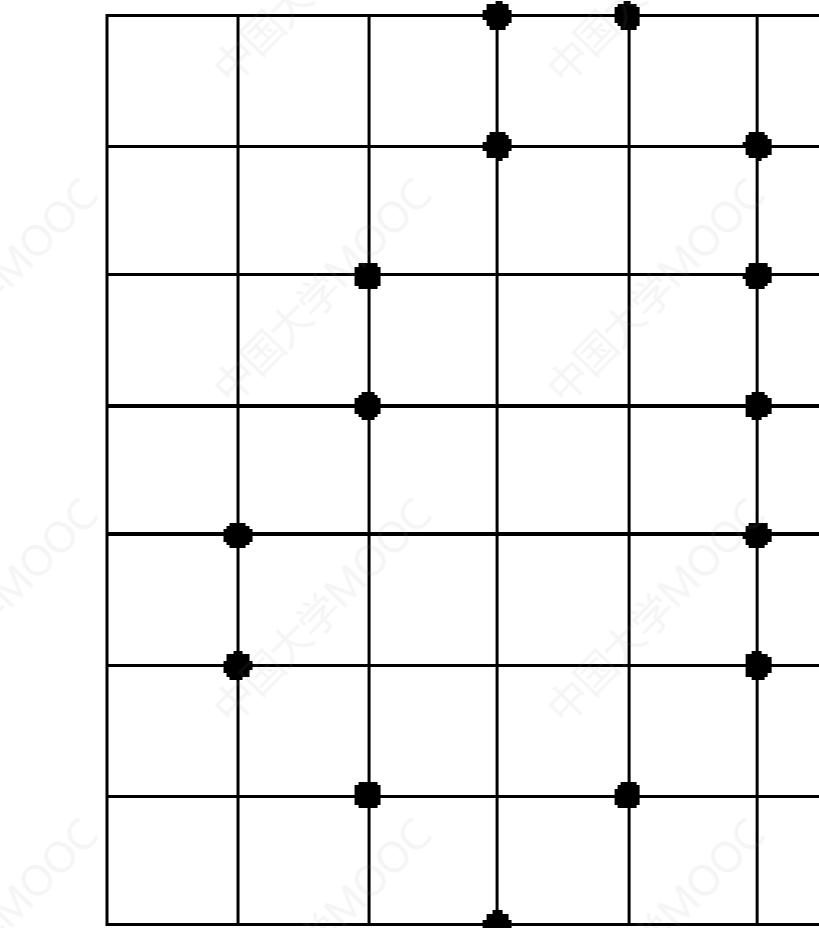
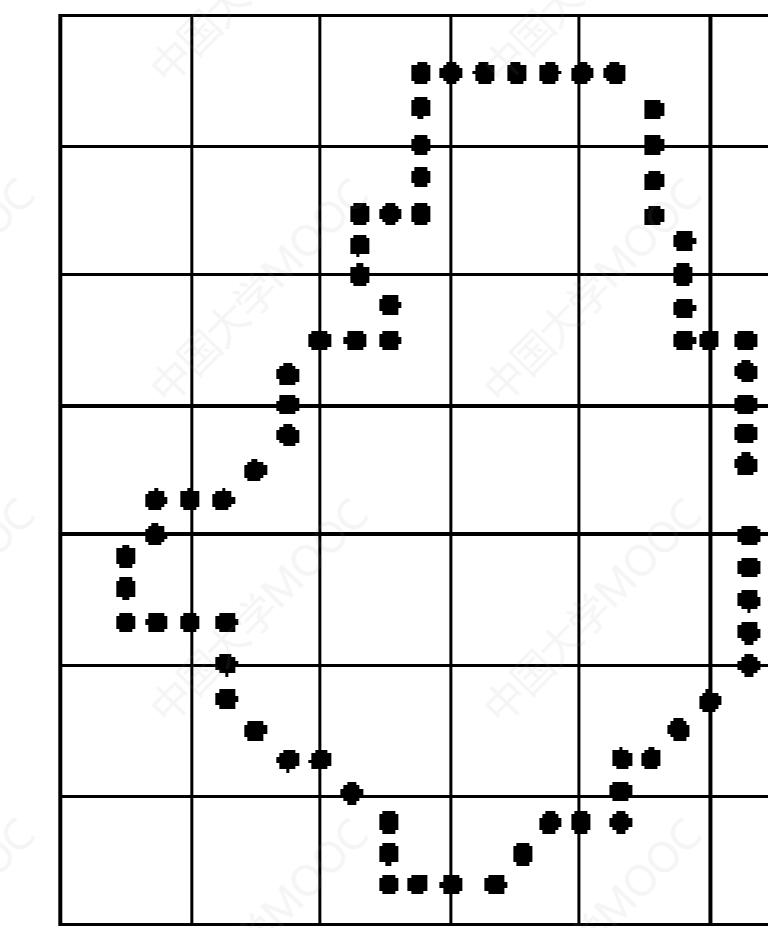


陈廷标, 夏良正编著, 数字图像处理, 人民邮电出版社, 1990年。

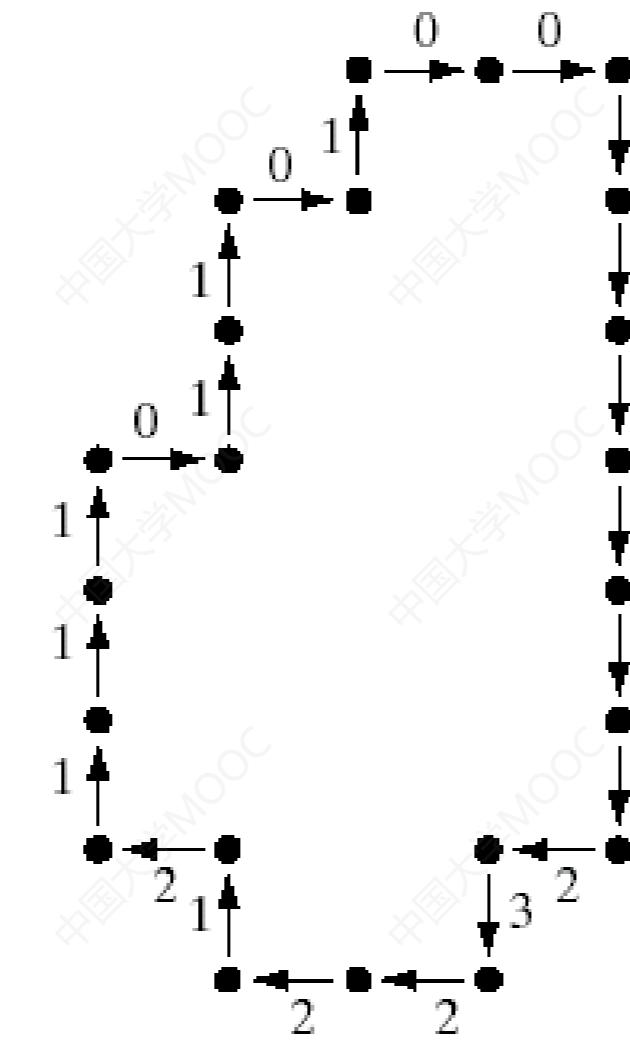
■ 8.1 数据级特征表达

➤ 方向链码

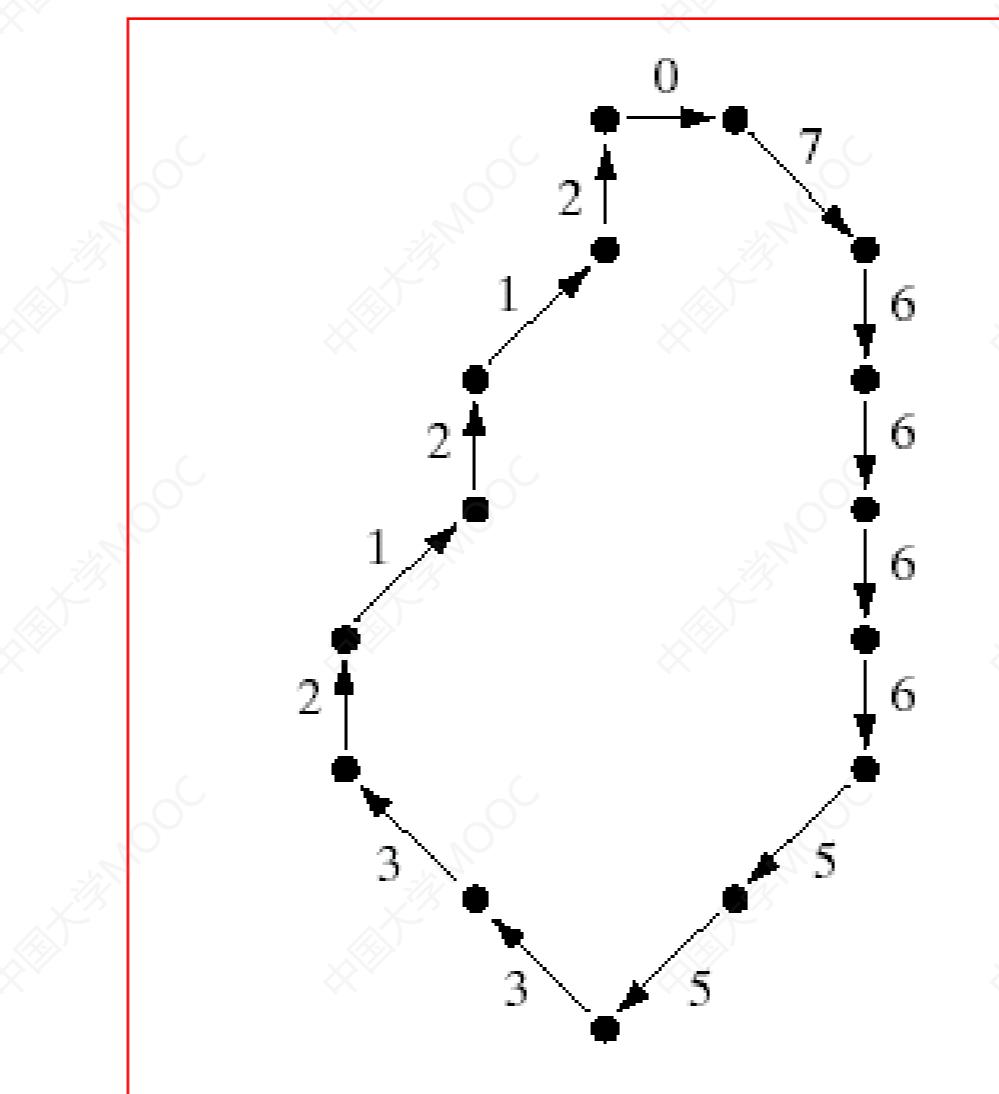
输入
图像



四方向
链码



八方向
链码



陈廷标, 夏良正编著, 数字图像处理, 人民邮电出版社, 1990年。

■ 8.1 数据级特征表达

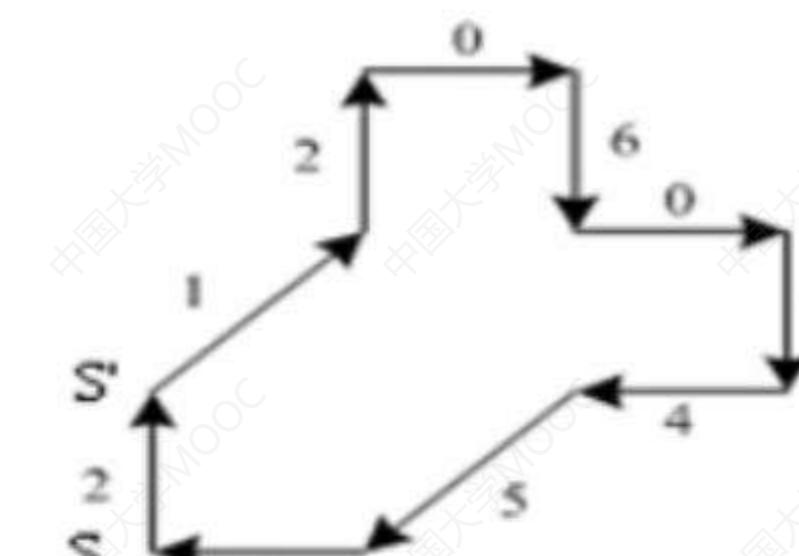
➤ 归一化方向链码

➤ 解决图像旋转引起的方向链码起始点不同的问题；

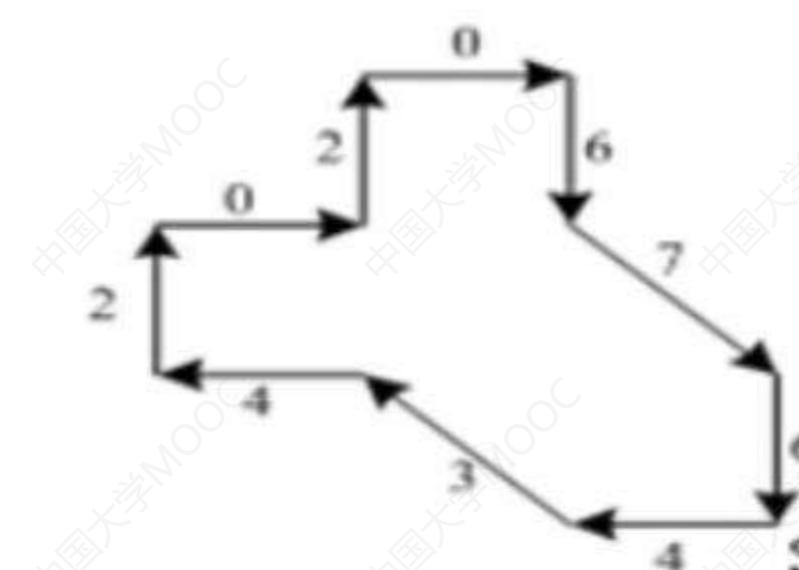
➤ 对于闭合边界，任选一起点S得到原链码，将链码看作由各方向数构成的n位自然数，将该码按一个方向循环，使其构成的n位自然数最小，此时就形成起点唯一的链码，称为归一化链码，也称为规格化链码。我们将这样转换后所对应的链码起点作为这个边界的归一化链码的起点。

8.1 数据级特征表达

➤ 归一化差分链码——具有旋转不变性



(a) 原始目标的区域



(b) 逆时针旋转 90° 后的区域

原链码: (4) 2 1 2 0 6 0 6 4 5

差分码： 6 7 1 6 6 2 6 6 1

(c) 旋转面

归一化链码

$$\overline{M}_8 = 0606454212$$

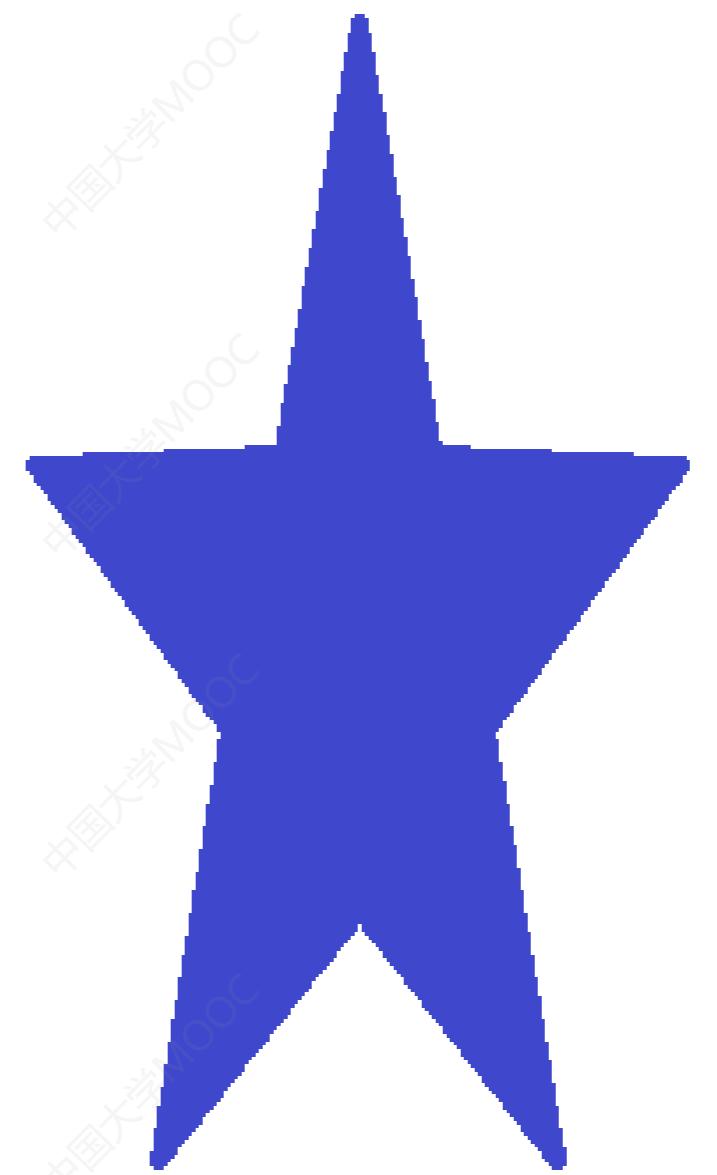
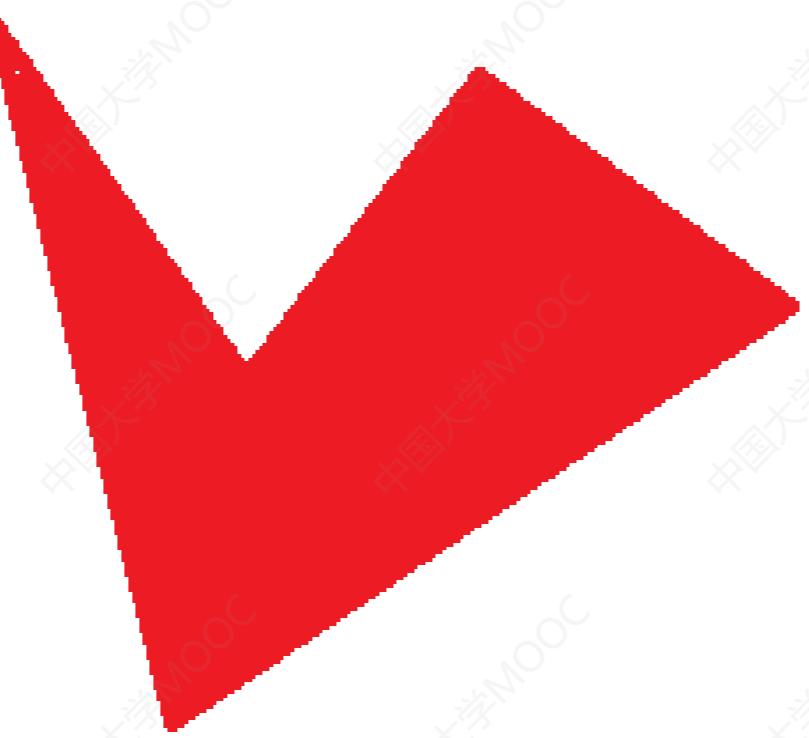
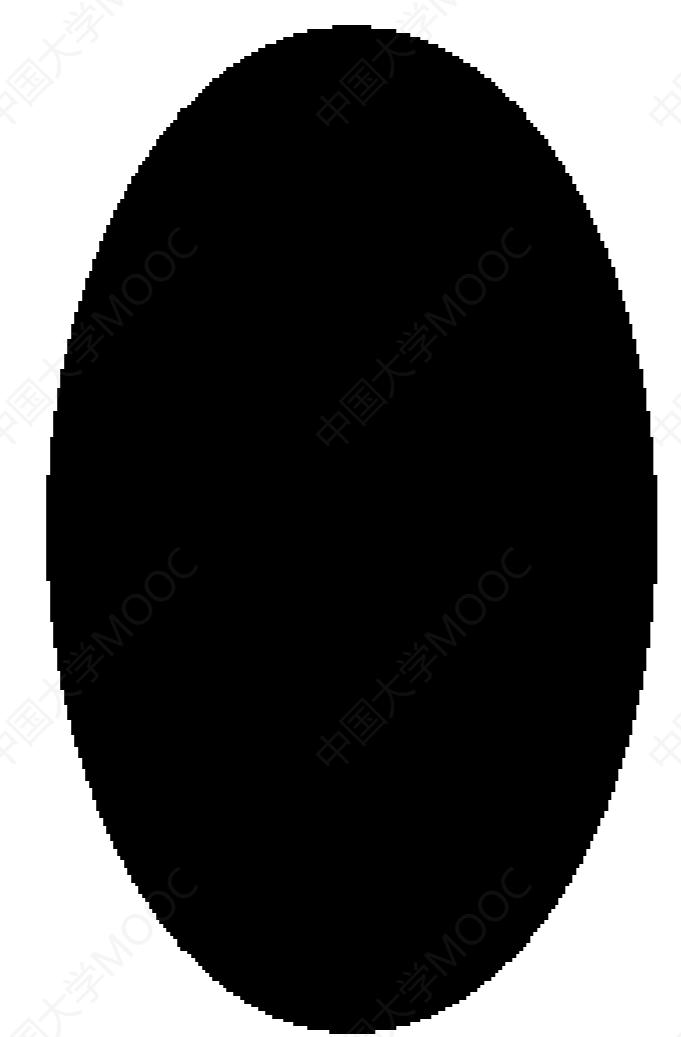
原链码: (6) 4 3 4 2 0 2 0 6 7

差分码: 6 7 1 6 6 2 6 6 1

(d) 旋转

归一化差分码

$$\bar{M}_s' = 1662661767$$



■ 8.1 数据级特征表达

➤ 多边形近似——区域边界可以由多边形以任意精度来逼近

两种方法：

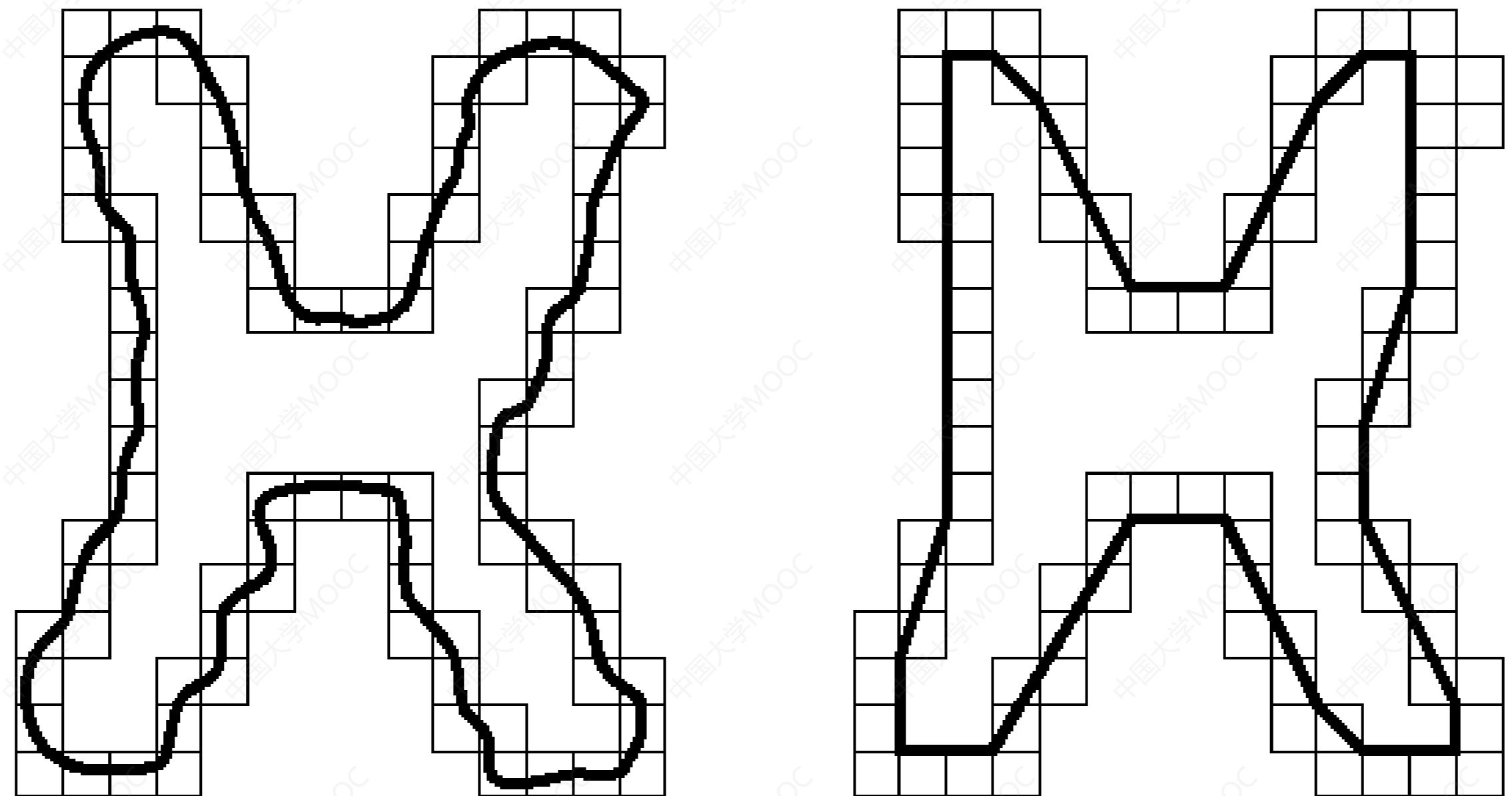
1) 最小周长多边形

2) 分割法

■ 8.1 数据级特征表达

➤ 最小周长多边形方法

用一系列彼此连接的单元将轮廓包围

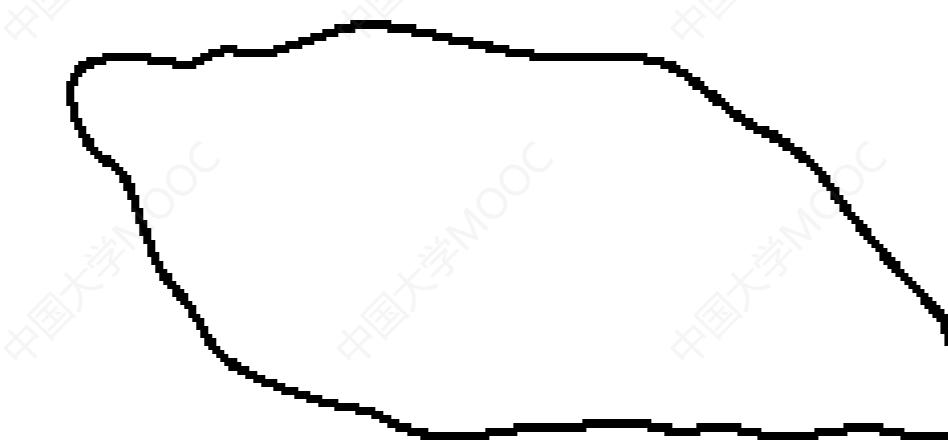


■ 8.1 数据级特征表达

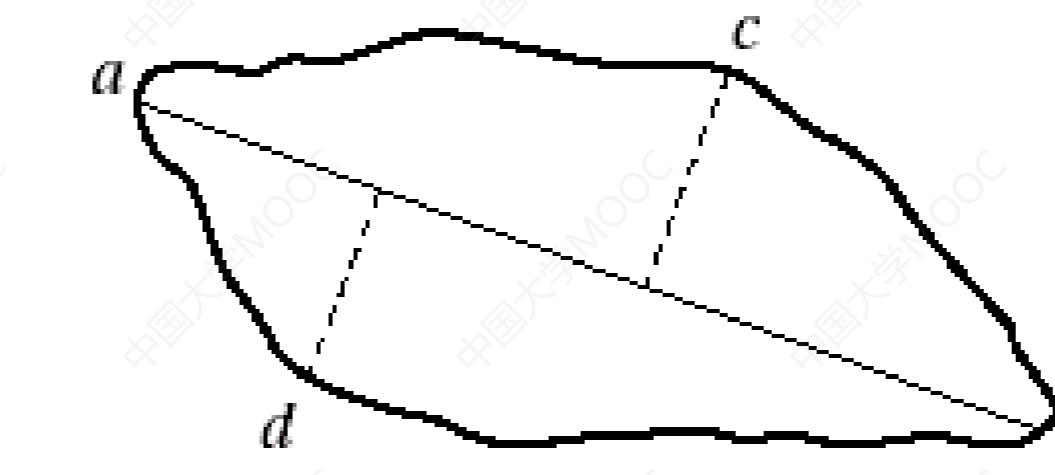
➤ 分割法

道格拉斯-普克算法(将一个分割段依次细分为两部分，直到满足规定的条件)

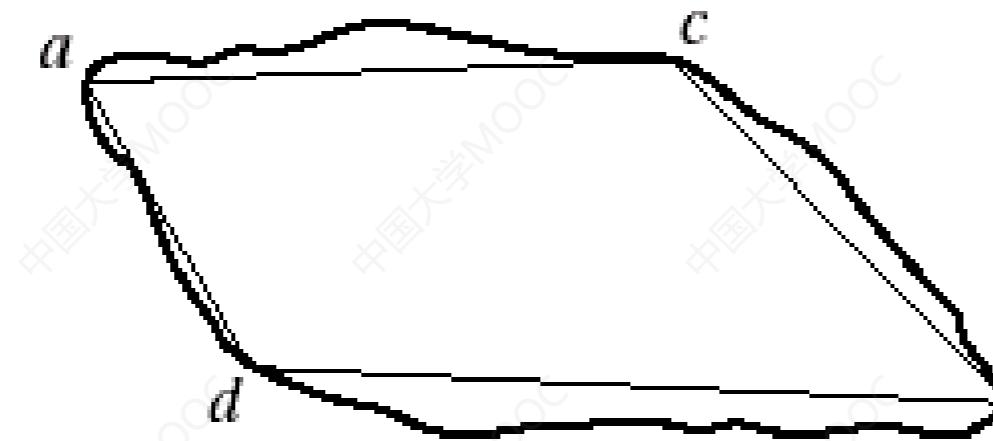
输入
图像



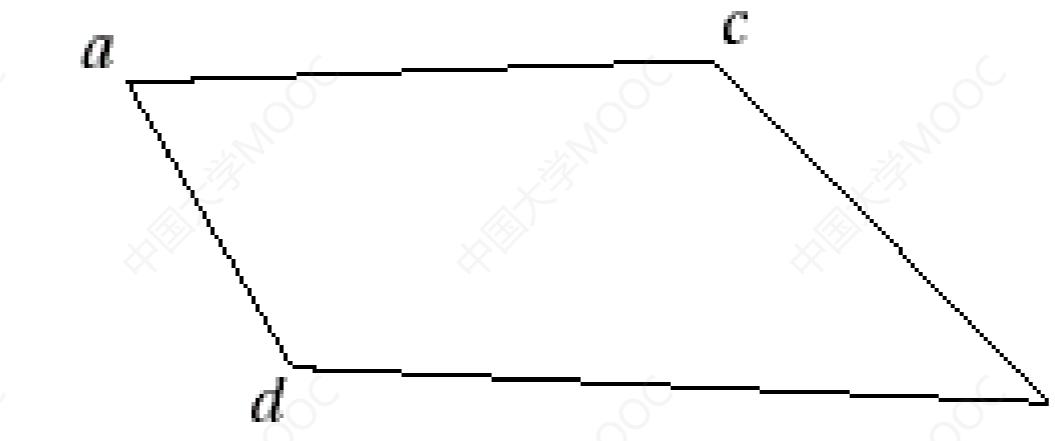
第一次
分割



第二次
分割



多边形
近似

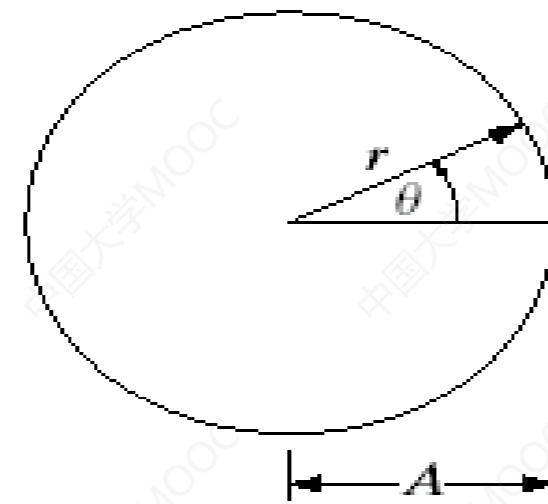


■ 8.1 数据级特征表达

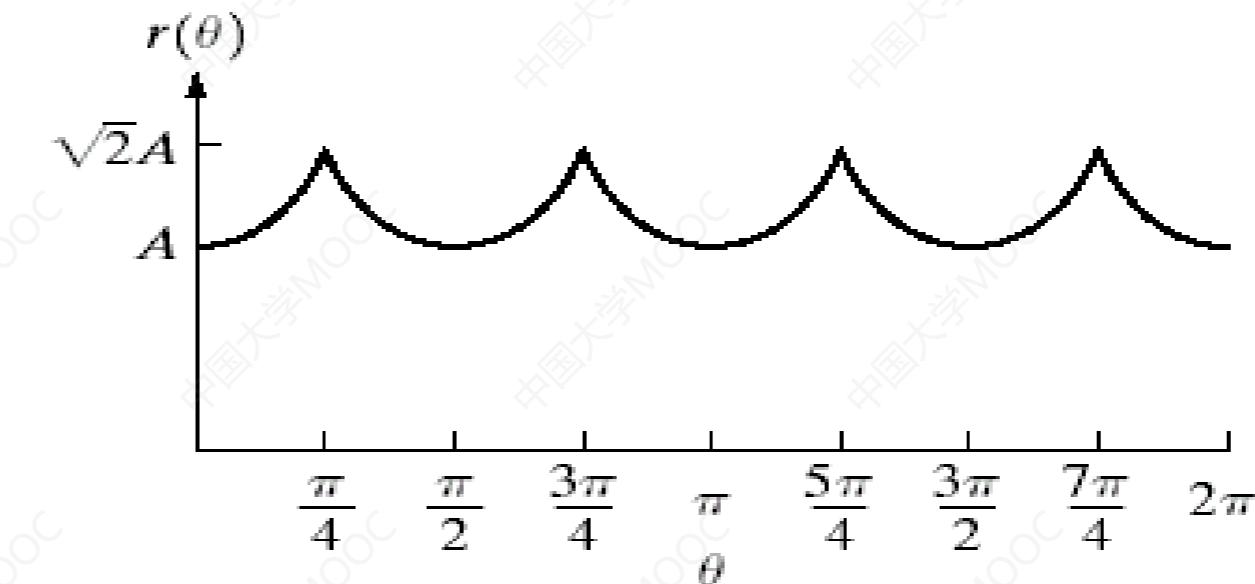
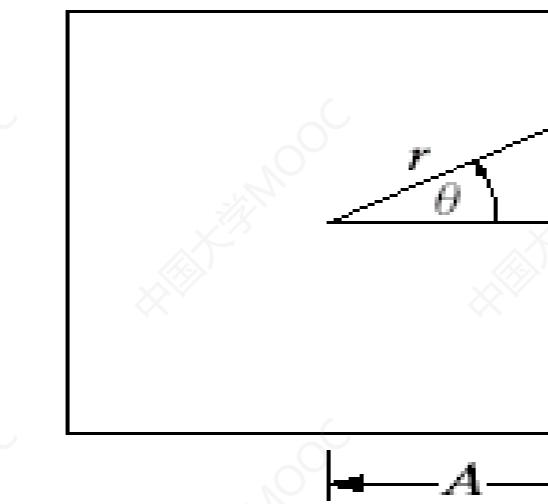
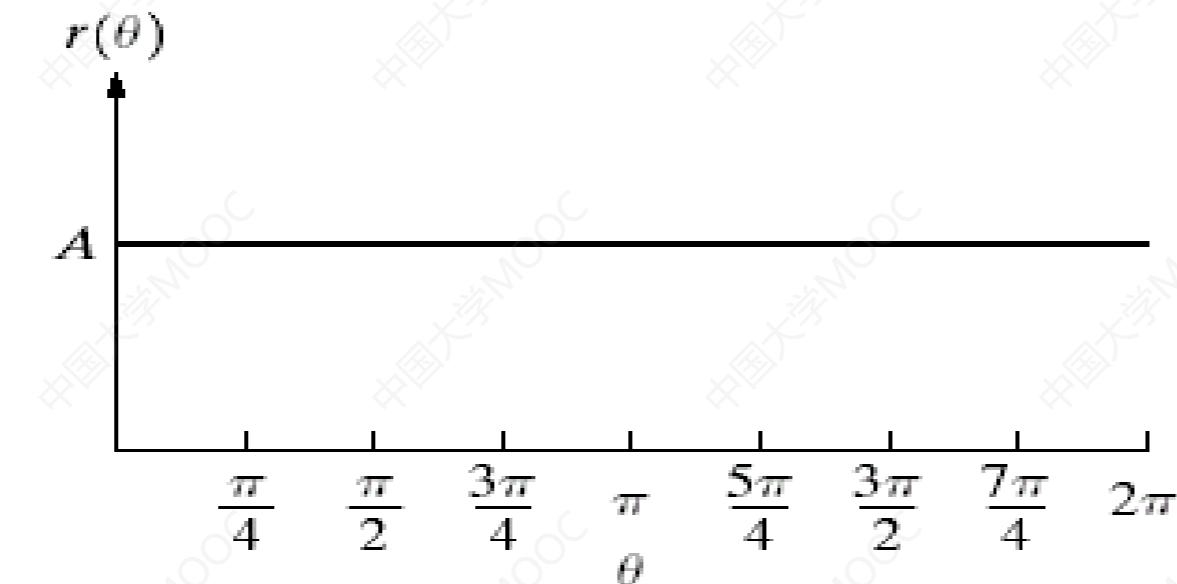
➤ 标记图法——将区域边界表达为一维函数

标记图法有很多种，可以通过各种方式生成。最简单的方法之一是将中心到边界像素点的距离绘制为角度的函数，

输入
图像



标记图
表示

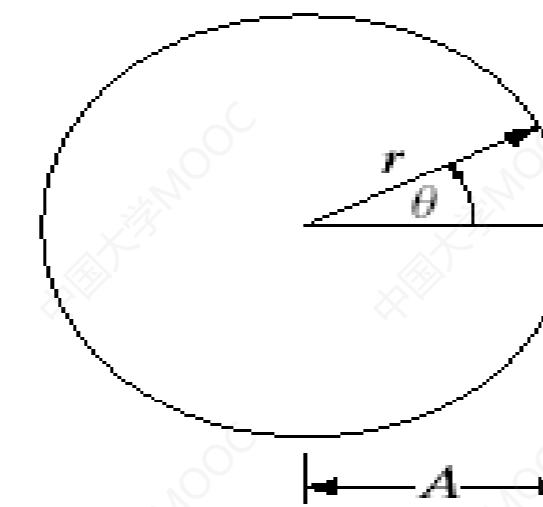


■ 8.1 数据级特征表达

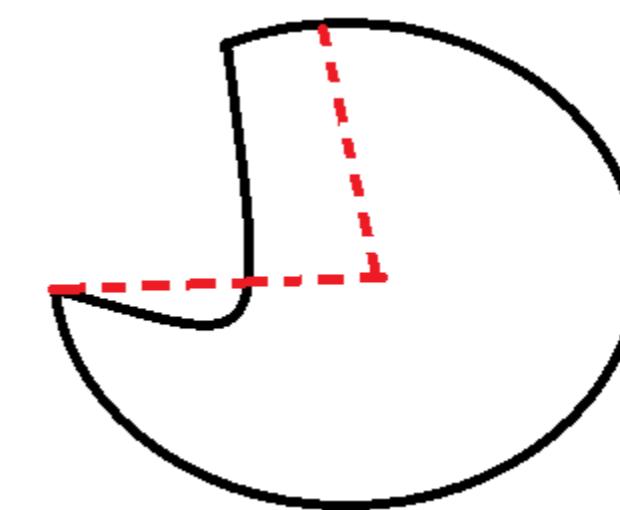
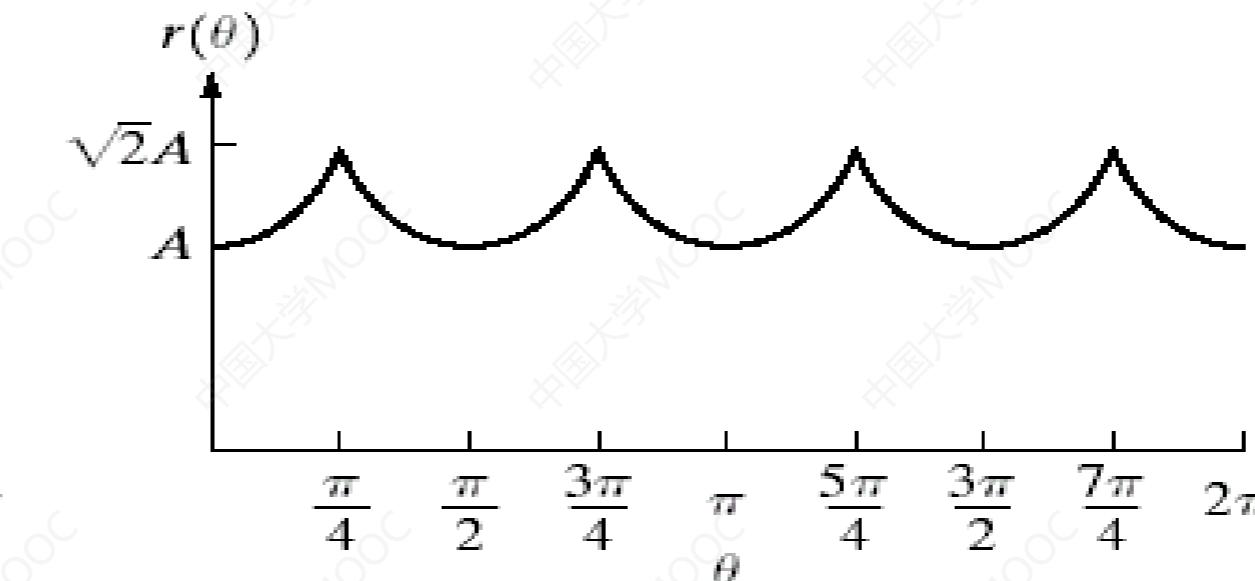
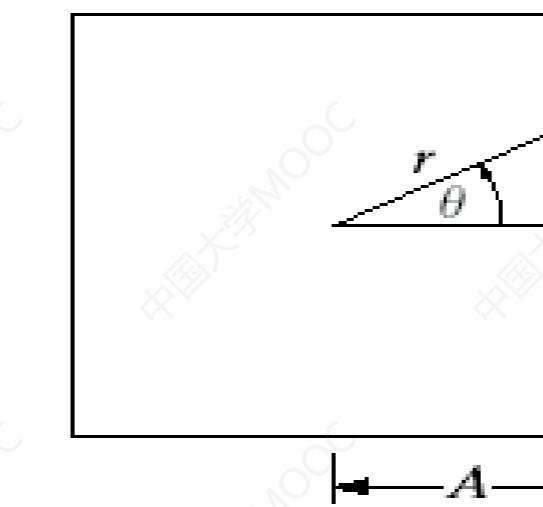
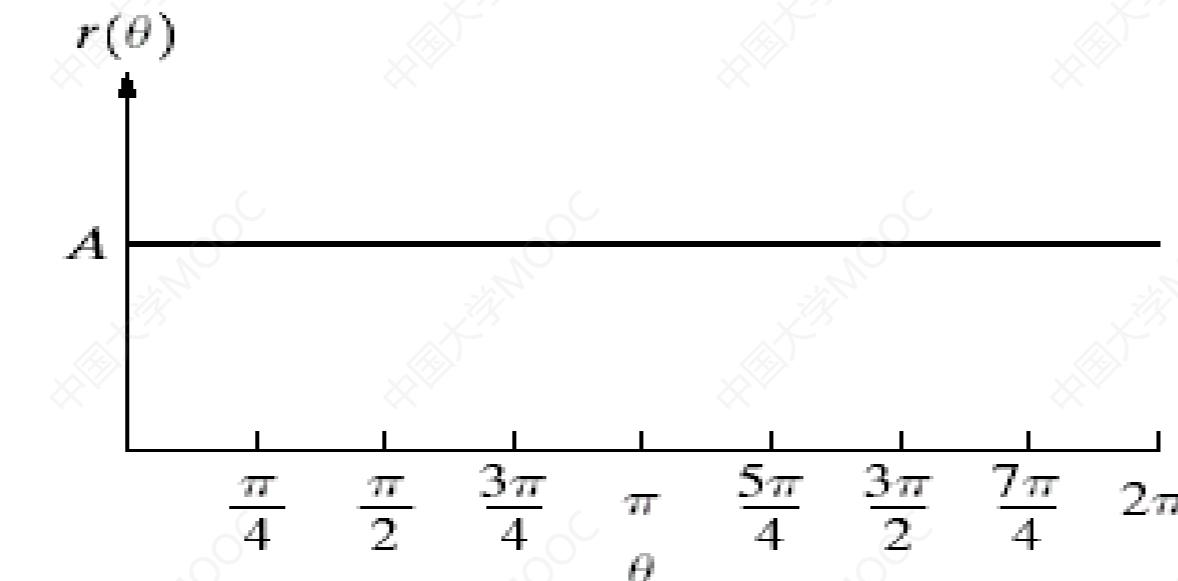
➤ 标记图法——将区域边界表达为一维函数

标记图法有很多种，可以通过各种方式生成。最简单的方法之一是将中心到边界像素点的距离绘制为角度的函数，

输入
图像



标记图
表示



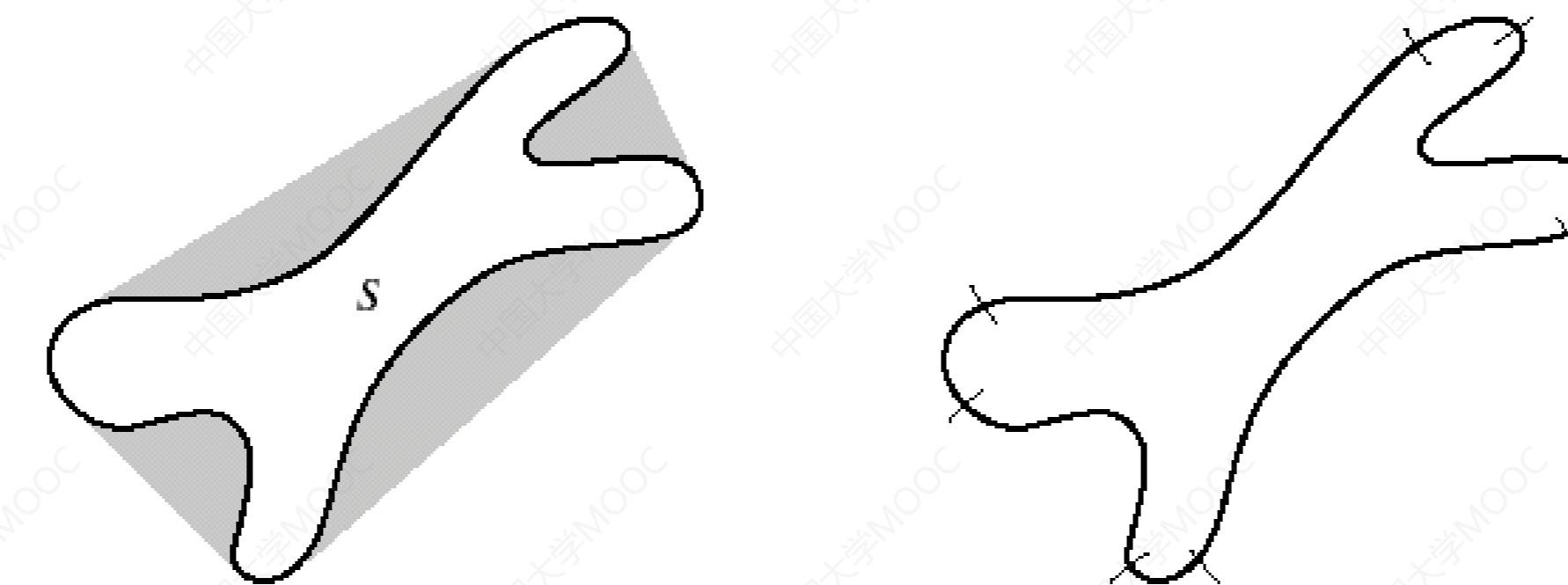
非凸区域示例

■ 8.1 数据级特征表达

➤ 边界分段

将区域边界分解为多个片段，通过分解降低了边界的复杂性，从而简化了描述过程。

当边界包含一个或多个包含形状信息的重要凹面时，这种方法特别有吸引力。在这种情况下，使用边界所包围的凸包区域是实现边界鲁棒定位的有力工具。



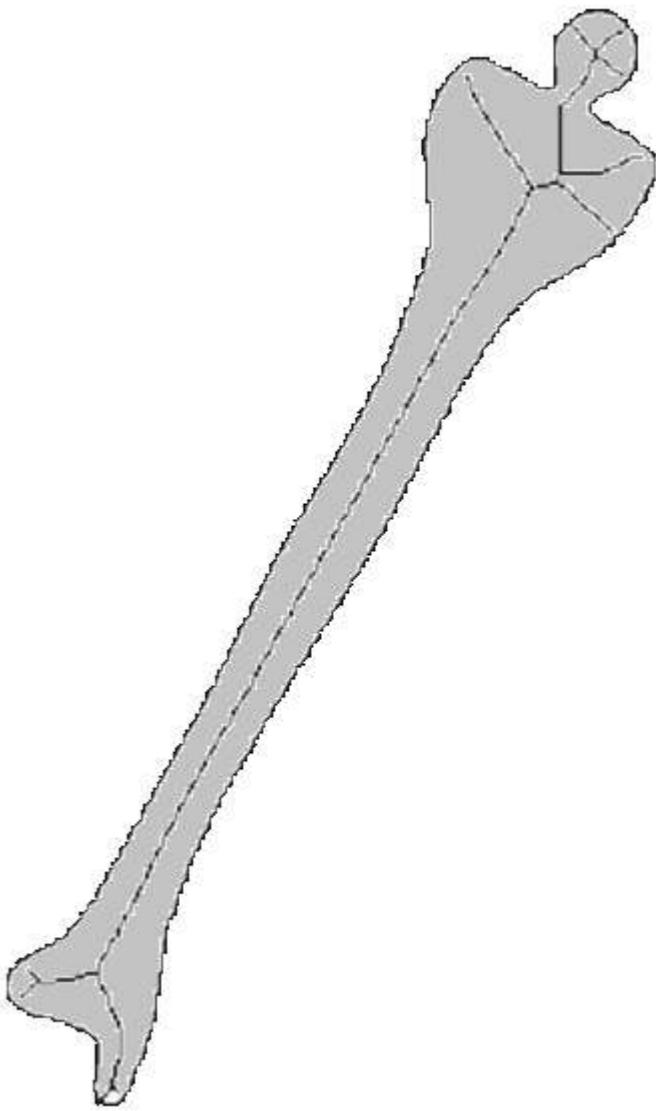
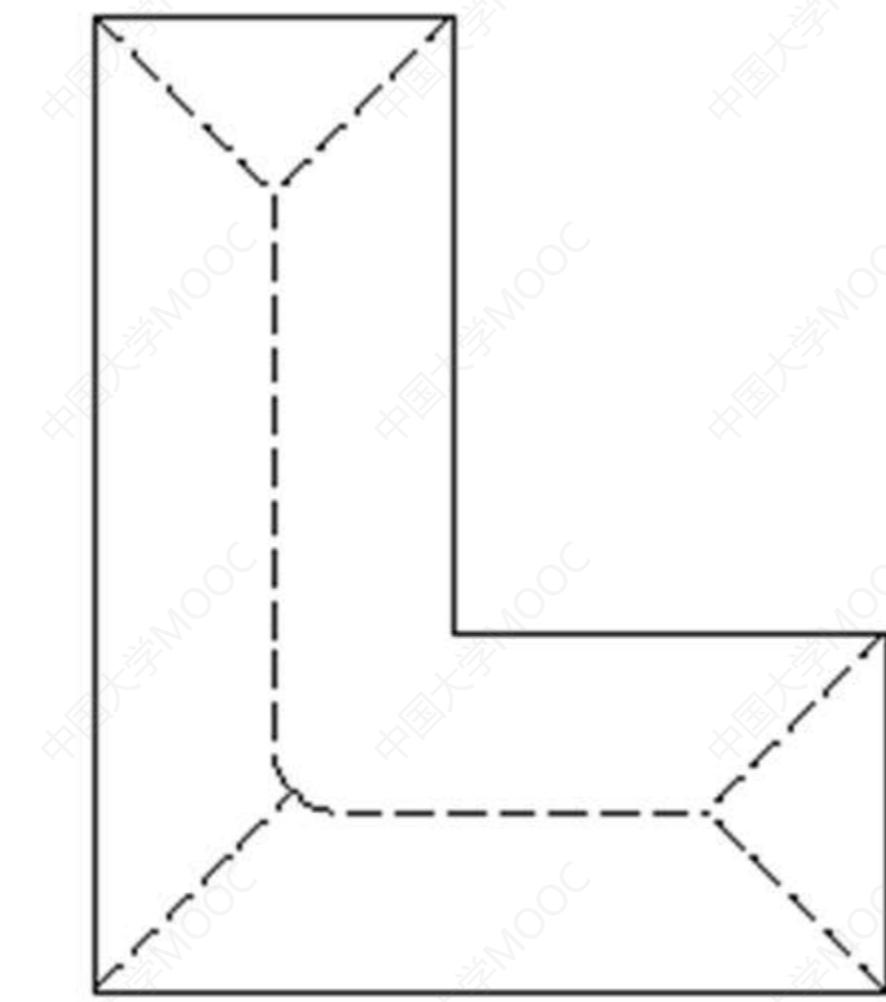
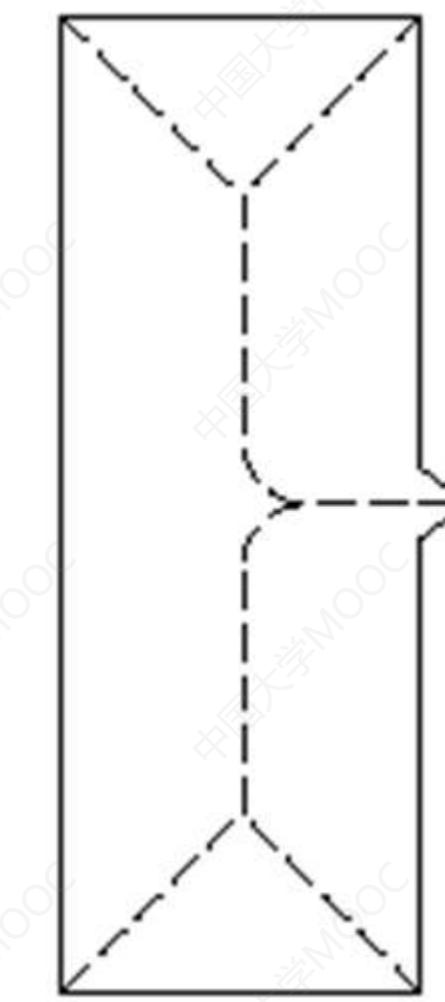
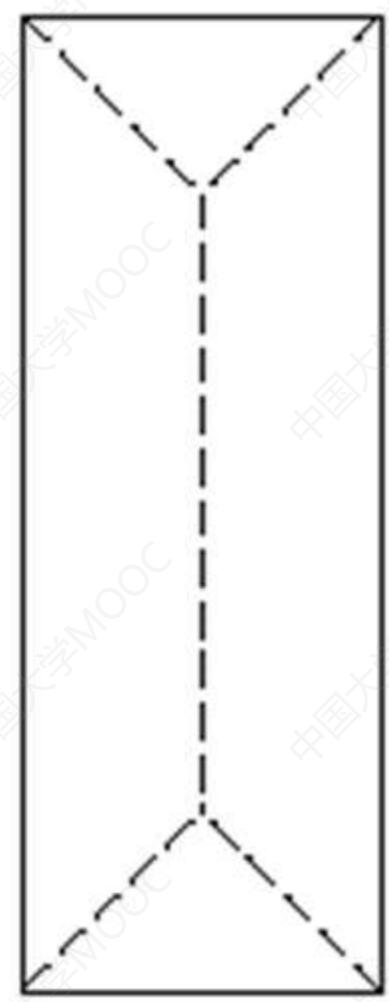
输入图像

边界分段示意图

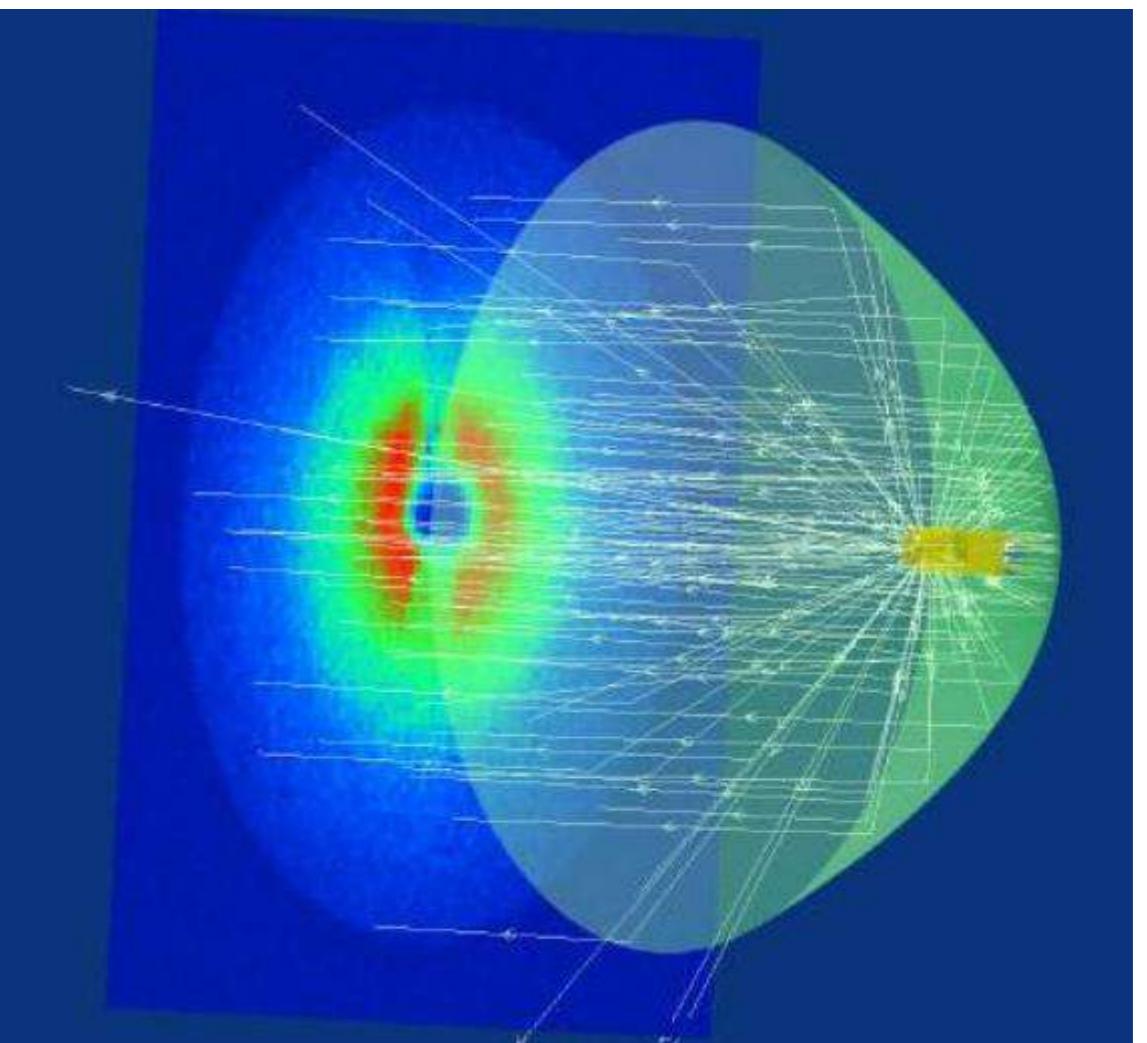
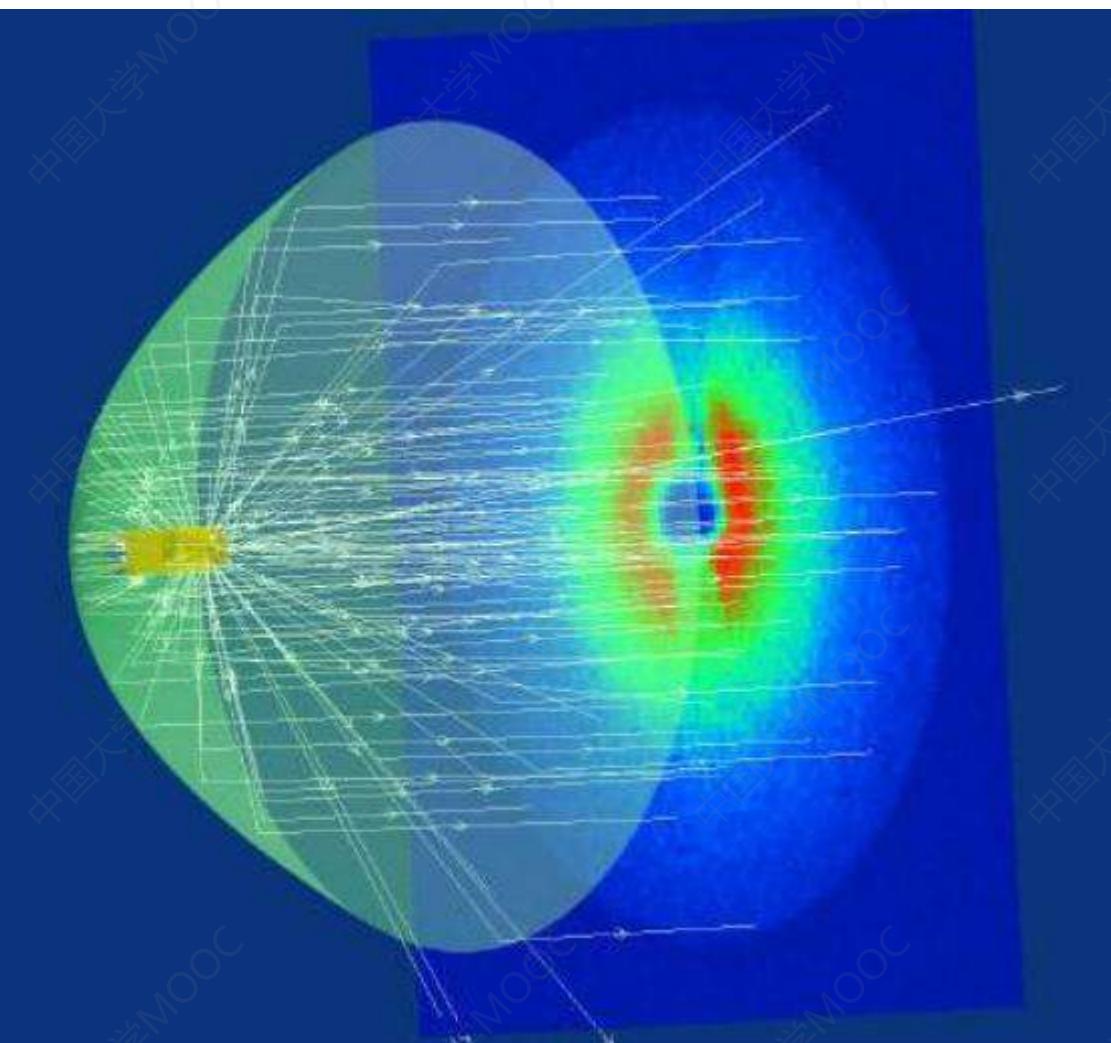
■ 8.1 数据级特征表达

➤ 骨架法

通过区域细化算法(也称为骨架化算法)获得区域的骨架来完成。



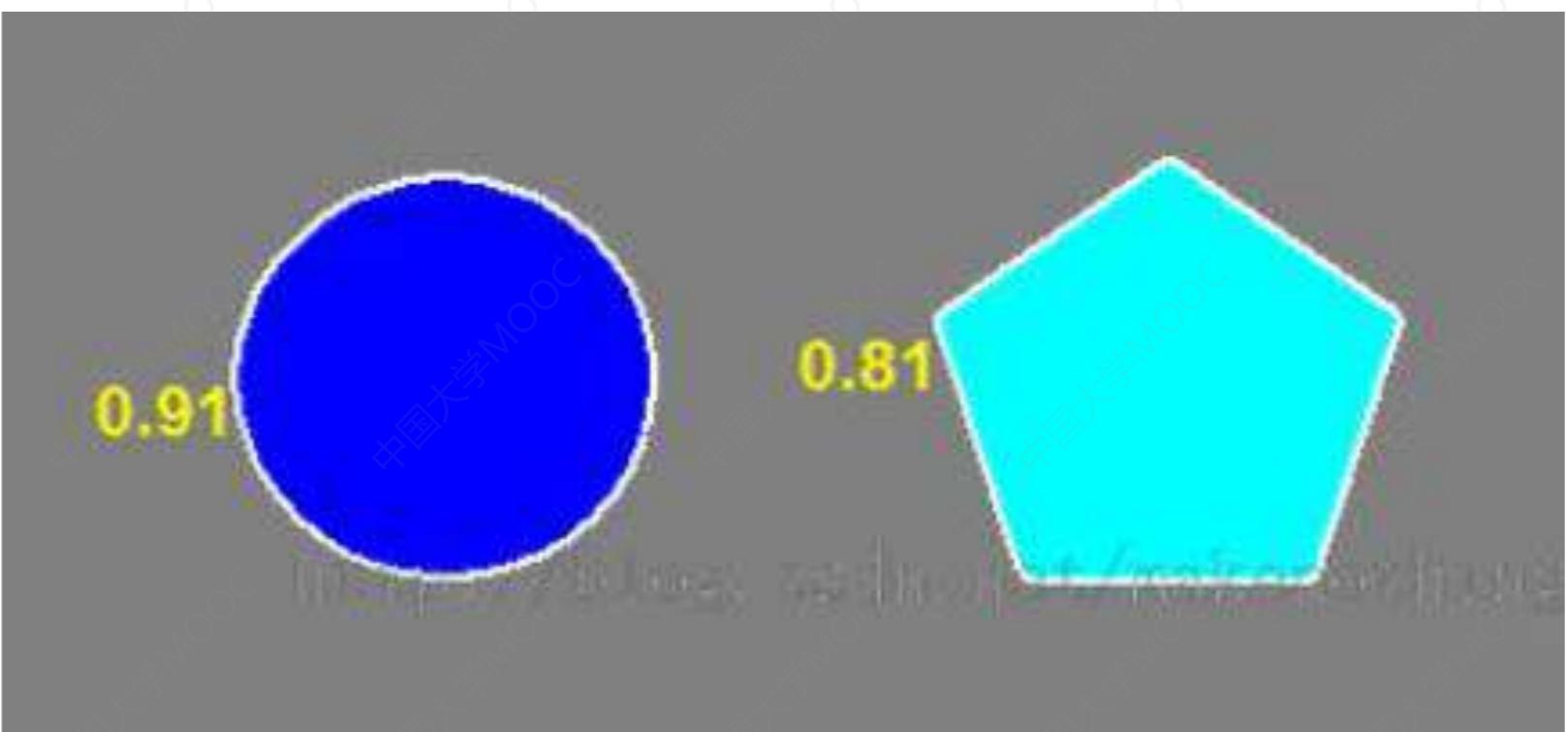
8.2 形状与颜色描述 子



■ 8.2 形状与颜色描述子

➤ 一些简单的描述子

- 长度
- 直径
- 主轴
- 曲率
- 圆度: $4\pi \times \text{面积} / (\text{周长})^2$
- 形状数



■ 8.2 形状与颜色描述子

➤ 形状数

Order 4

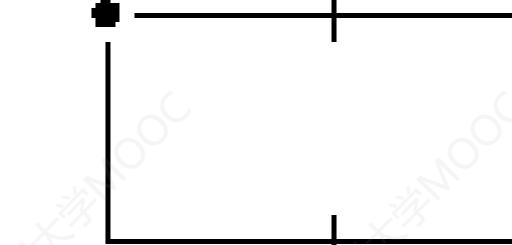


Chain code: 0 3 2 1

Difference: 3 3 3 3

Shape no.: 3 3 3 3

Order 6

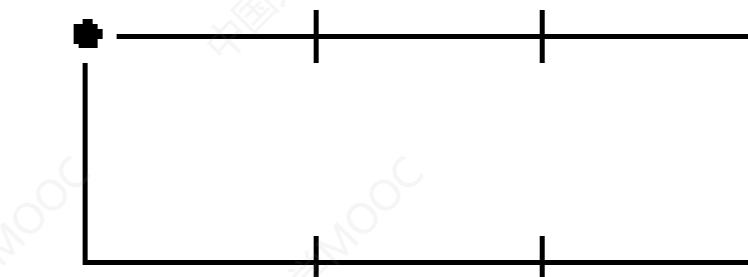
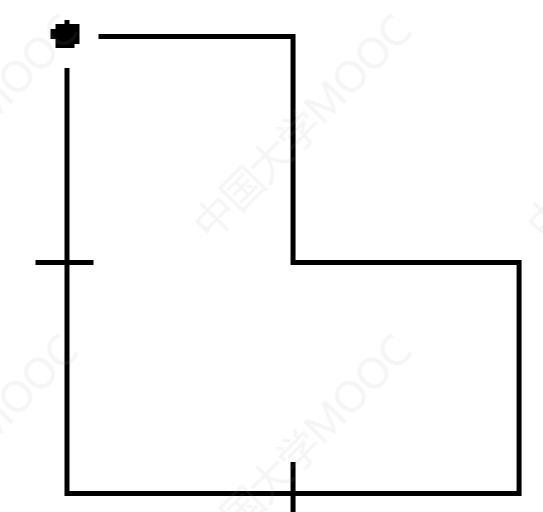
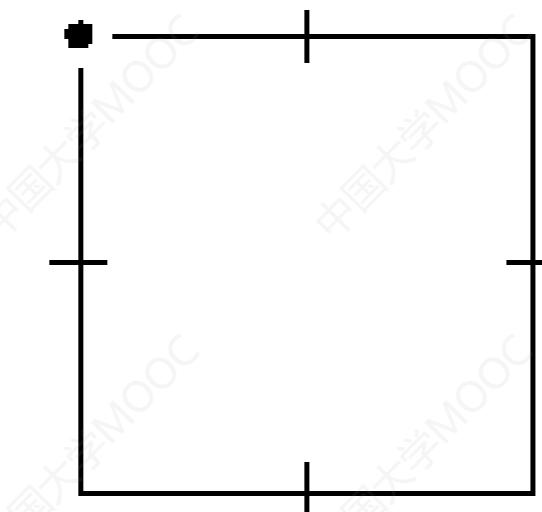


Chain code: 0 0 3 2 2 1

Difference: 3 0 3 3 0 3

Shape no.: 0 3 3 0 3 3

Order 8



Chain code: 0 0 3 3 2 2 1 1

Difference: 3 0 3 0 3 0 3 0

Shape no.: 0 3 0 3 0 3 0 3

Chain code: 0 3 0 3 2 2 1 1

Difference: 3 3 1 3 3 0 3 0

Shape no.: 0 3 0 3 3 1 3 3

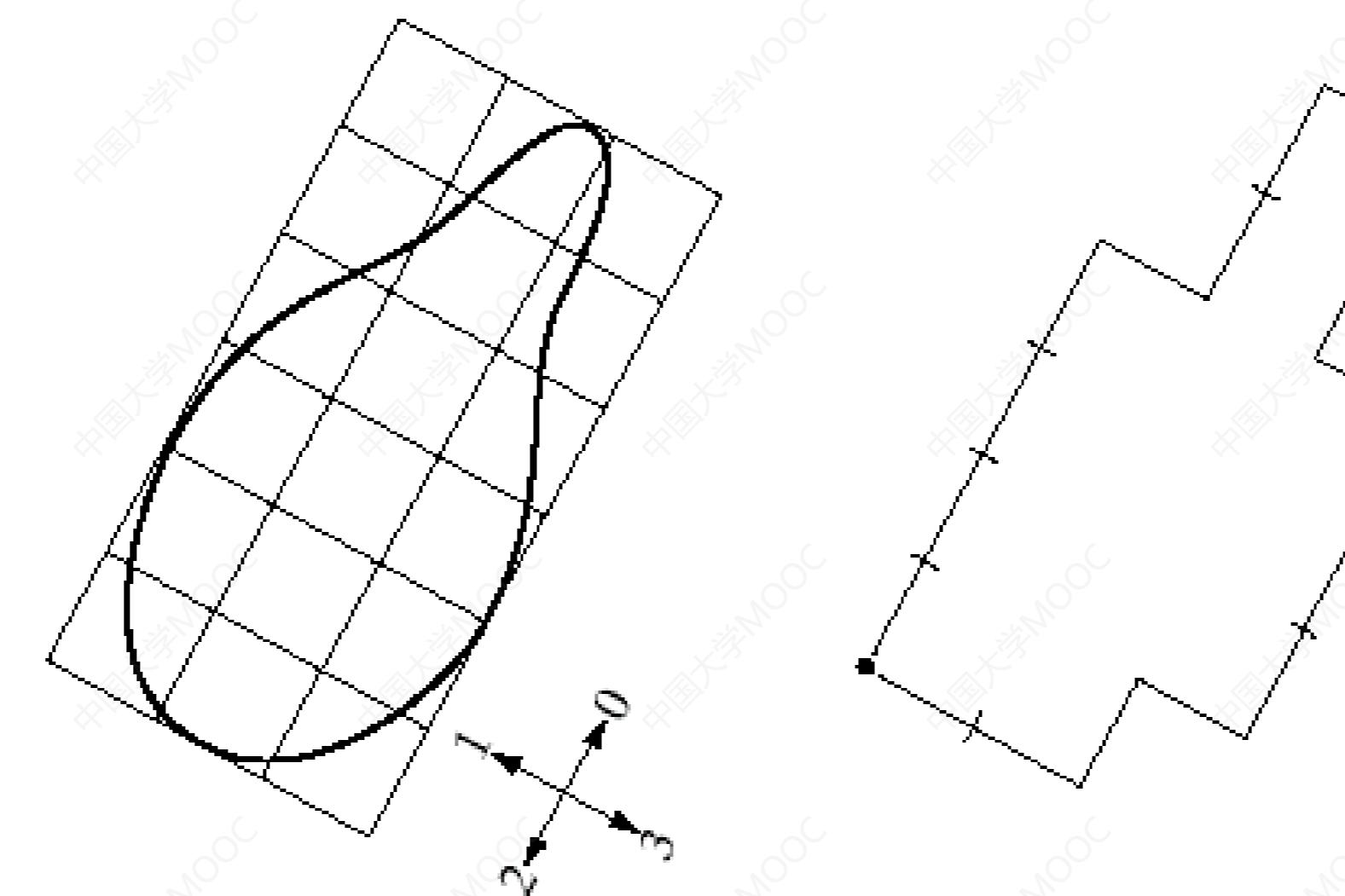
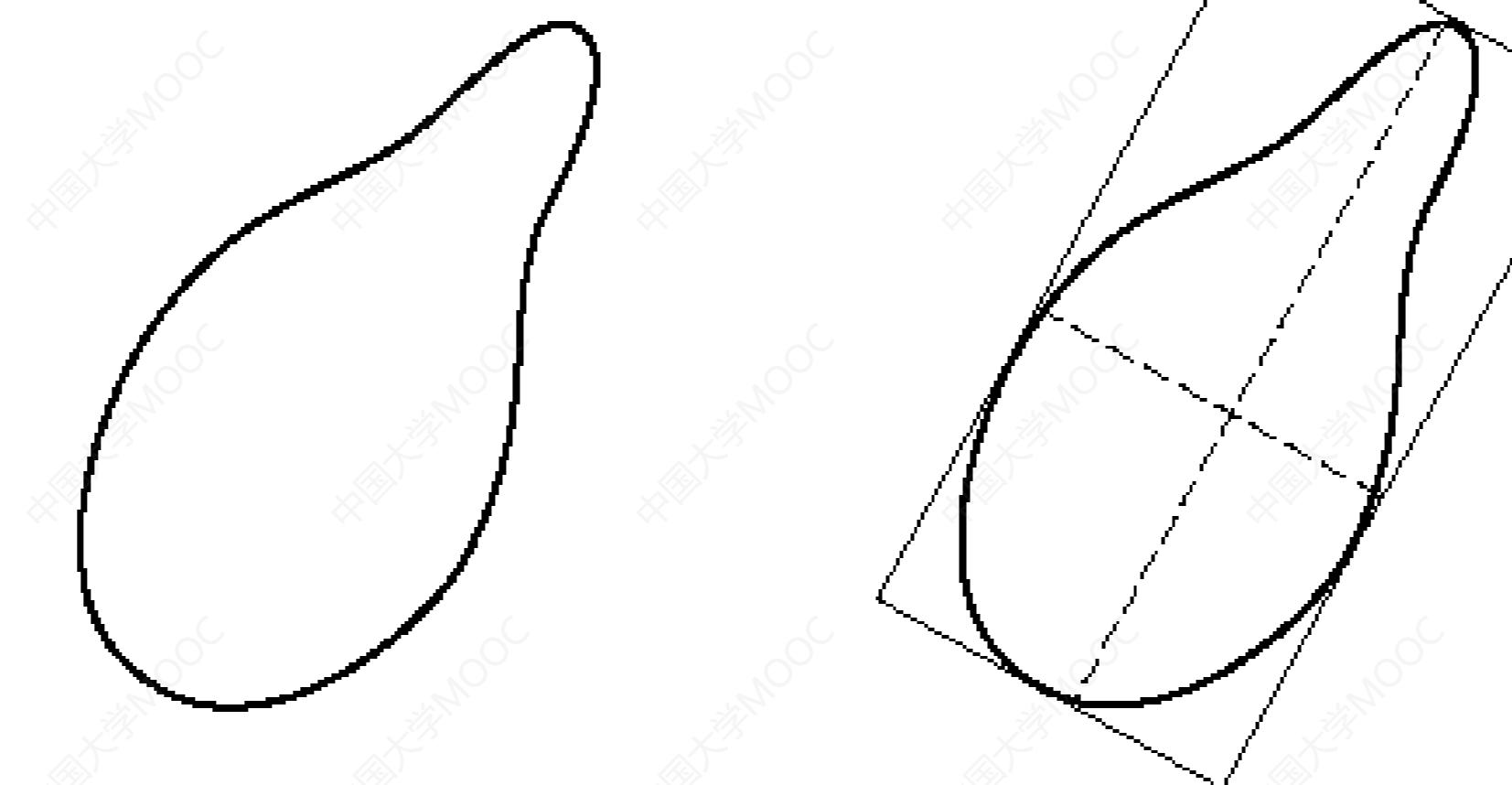
Chain code: 0 0 0 3 2 2 2 1

Difference: 3 0 0 3 3 0 0 3

Shape no.: 0 0 3 3 0 0 3 3

■ 8.2 形状与颜色描述子

➤ 形状数



Chain code: 0 0 0 0 3 0 0 3 2 2 3 2 2 2 1 2 1 1

Difference: 3 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 0

Shape no.: 0 0 0 3 1 0 3 3 0 1 3 0 0 3 1 3 0 3

■ 8.2 形状与颜色描述子

➤ 傅里叶描述子

傅里叶描述子的基本思想是：假定物体的形状是一条封闭的曲线，边界曲线上点 $P(i)$ 的 x ， y 坐标变化 $x(i) + jy(i)$ 是一个以形状边界周长为周期的函数，这个周期函数可以用傅里叶级数展开表示，傅里叶级数中的一系列系数 $z(k)$ 是直接与边界曲线的形状有关的，称为傅里叶描述子。

■ 8.2 形状与颜色描述子

➤ 傅里叶描述子

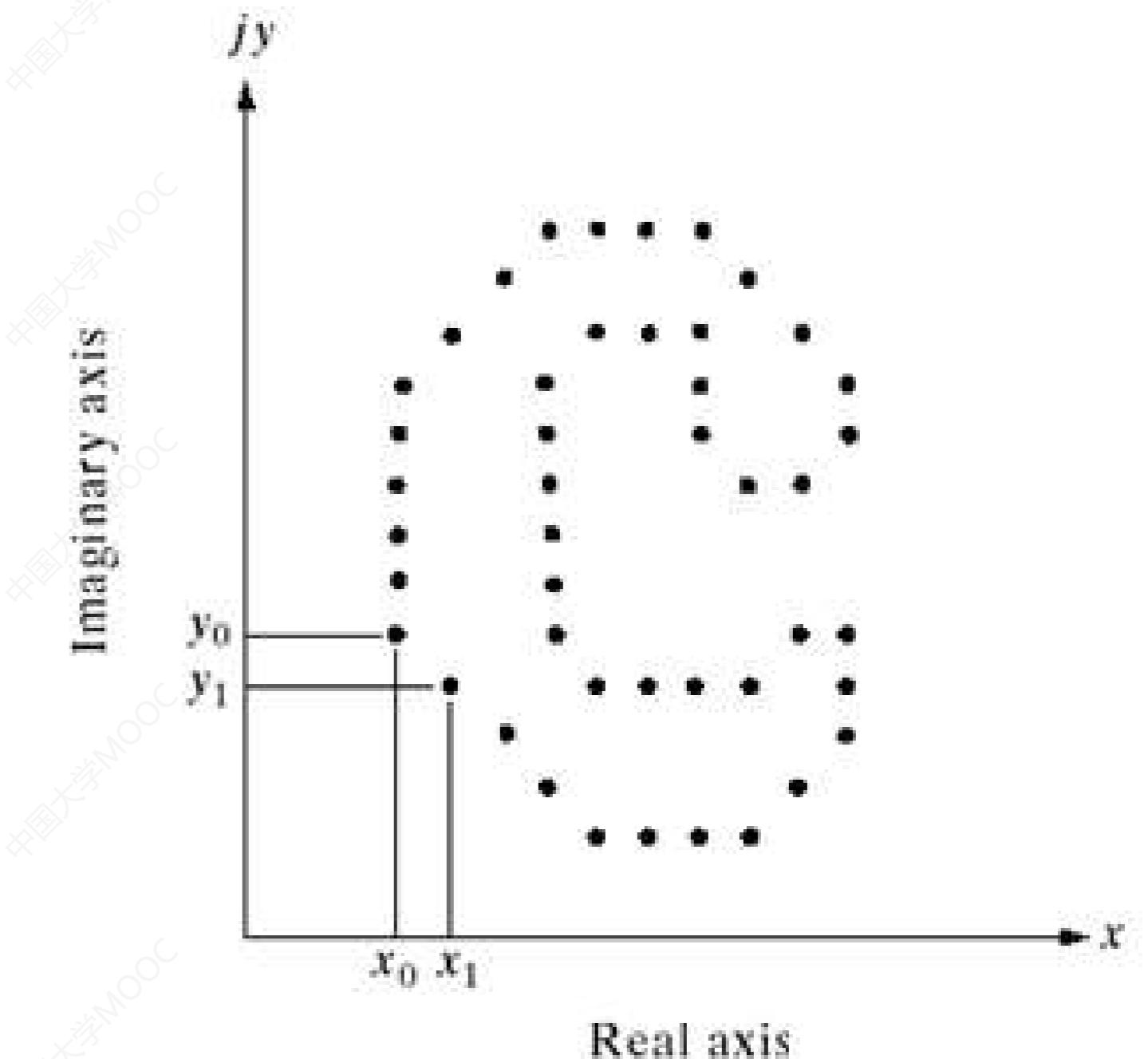
图中显示了x-y平面上的k点数字边界。从任意点 (x_0, y_0) 开始。

假设沿逆时针方向穿越边界时，会遇到坐标对 (x_0, y_0) 、

(x_1, y_1) 、 (x_2, y_2) 、 \dots (x_{K-1}, y_{K-1}) 。

边界本身可以表示为一系列坐标 $(x_k, y_k) \ k = 0, 1, 2, K - 1$ 。此

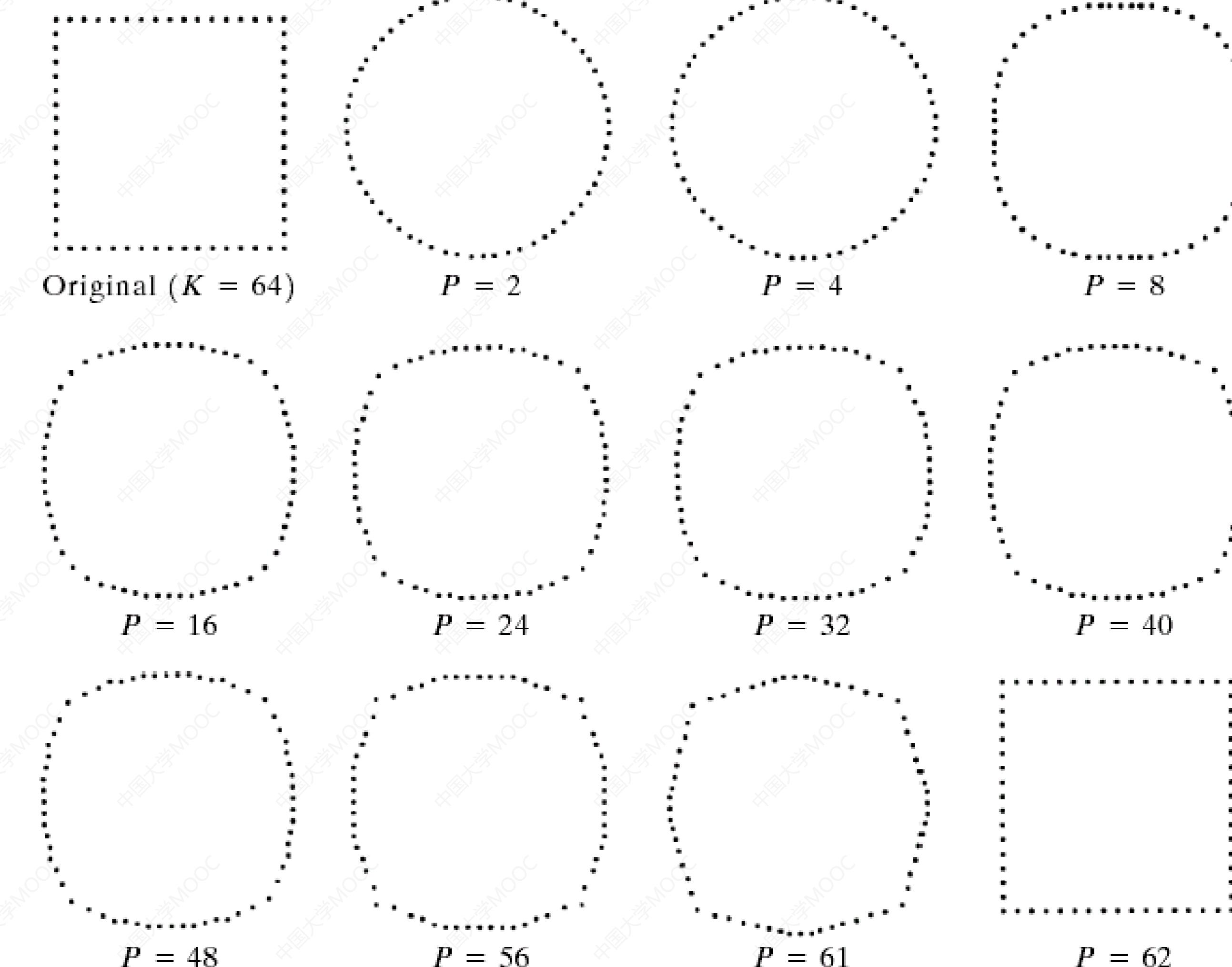
外，每个坐标对可以视为一个复数，即 $s_k = x_k + j y_k$



■ 8.2 形状与颜色描述子

➤ 傅里叶描述子

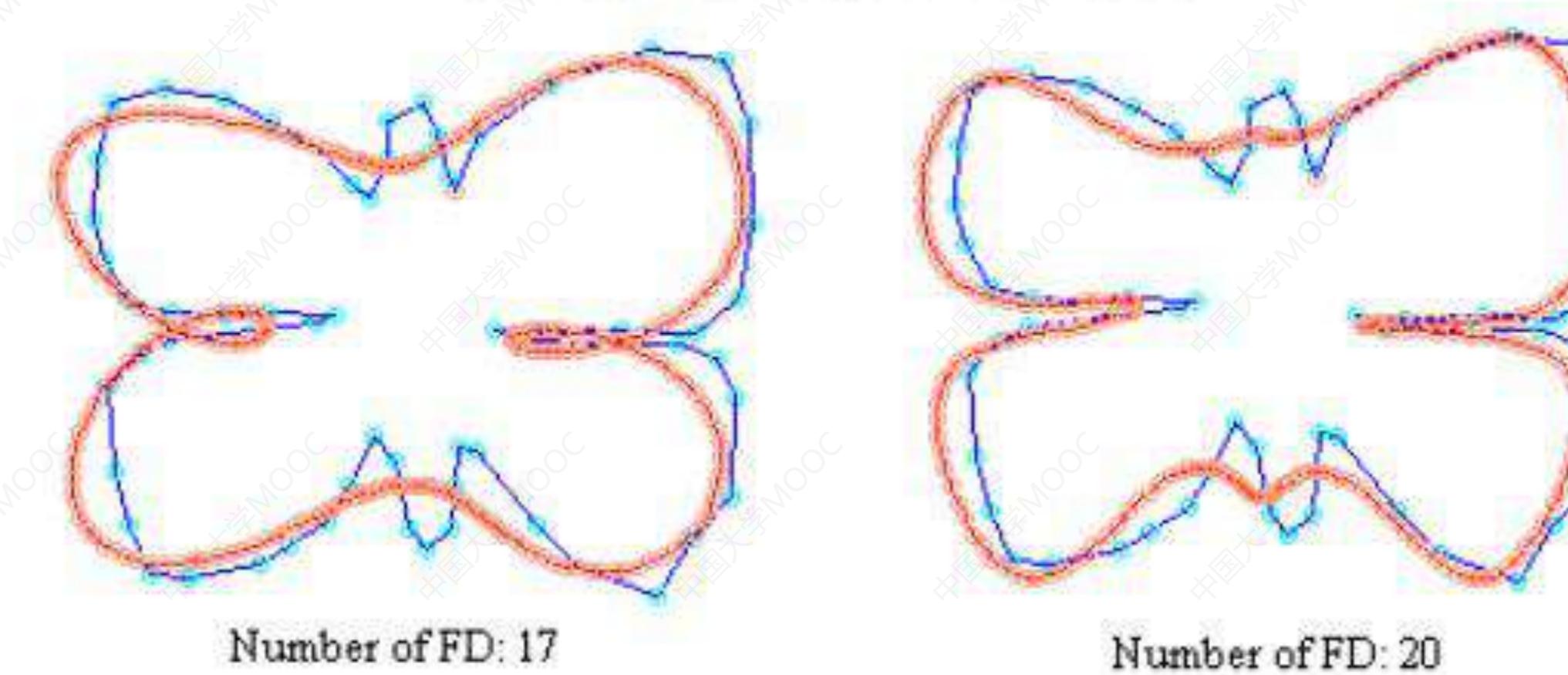
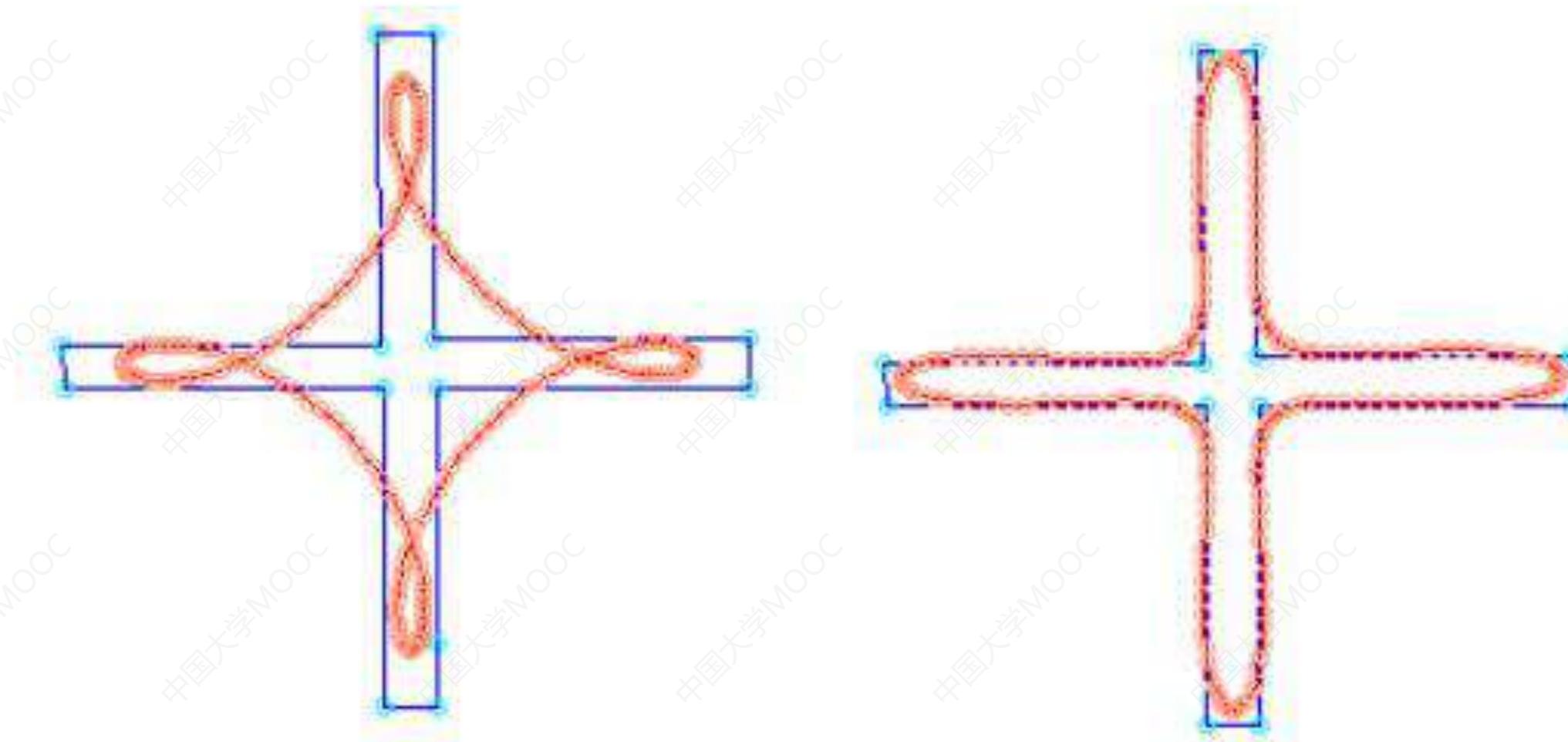
正方形形状用傅里叶描述子表示，二阶描述子表达为圆形，随着引入更多阶数，表达的形状越来越接近正方形。



■ 8.2 形状与颜色描述子

➤ 傅里叶描述子

更多分解实例。



■ 8.2 形状与颜色描述子

➤ 图像的矩 (Image Moments)

矩用来抽取图像(块)的特征。图像中，低阶矩反映低频(主要的)信息，高阶矩反映高频(细节)信息。从一幅图像计算出来的矩集，不仅可以描述图像形状的全局特征，还可以提供大量关于该图像不同的几何特征信息，如大小、位置、方向和形状等。矩函数在图像分析中有广泛应用，如模式识别、目标分类、目标识别与方位估计等。

像素点 $f(x,y)$ 的 $(p+q)$ 阶矩定义为：

$$m_{pq} = \sum_{x=1}^c \sum_{y=1}^R x^p y^q f(x, y), p, q = 0, 1, 2 \dots$$

■ 8.2 形状与颜色描述子

➤ 图像的矩 (Image Moments)

- 各阶矩的物理意义与计算:
- 0阶矩(m^{00}): 目标区域的面积(质量);
- 1阶矩((m^{01}, m^{10})): 目标区域的质心;
- 2阶矩(m^{20}, m^{02}, m^{11}): 目标区域的旋转半径;
- 3阶矩($(m^{30}, m^{03}, m^{12}, m^{21})$): 目标区域的方位和斜度, 反映目标的扭曲。

■ 8.2 形状与颜色描述子

➤ 图像的中心矩——构造平移不变性

- 以目标区域的质心为中心构建中心矩，那么矩的计算永远是目标区域中的点相对于目标区域的质心，而与目标区域的位置无关，即具备了平移不变性。

$$x_0 = \frac{m_{10}}{m_{00}}, y_0 = \frac{m_{01}}{m_{00}}$$

- 根据求得的质心坐标，可以构造出中心距

$$u_{pq} = \sum_{x=1}^C \sum_{y=1}^R (x - x_0)^p (y - y_0)^q f(x, y), p, q = 0, 1, 2 \dots$$

■ 8.2 形状与颜色描述子

➤ 图像的归一化中心矩——构造尺度不变性

- 为抵消尺度变化对中心矩的影响，利用零阶中心矩 u_{00} 对各阶中心矩进行归一化处理，得到归一化中心矩：

$$\eta_{pq} = \frac{u_{pq}}{u_{00}^r}, r = \frac{q + p}{2} + 1$$

- 零阶矩表示目标区域的质量（面积），如果目标区域的尺度发生变化，其零阶中心距也会相应发生改变，使得矩具备尺度不变性。

■ 8.2 形状与颜色描述子

➤ 图像的Hu矩——构造旋转不变性

- 利用二阶和三阶中心距可推导出以下7个不变矩组($\Phi_1 \sim \Phi_7$),他们在图像平移、旋转和比例变化时保持不变。

$$\Phi_1 = \eta_{20} + \eta_{02}$$

$$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\Phi_3 = (\eta_{20} - 3\eta_{12})^2 + 3(\eta_{21} - \eta_{03})^2$$

$$\Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\begin{aligned}\Phi_5 = & (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

$$\Phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

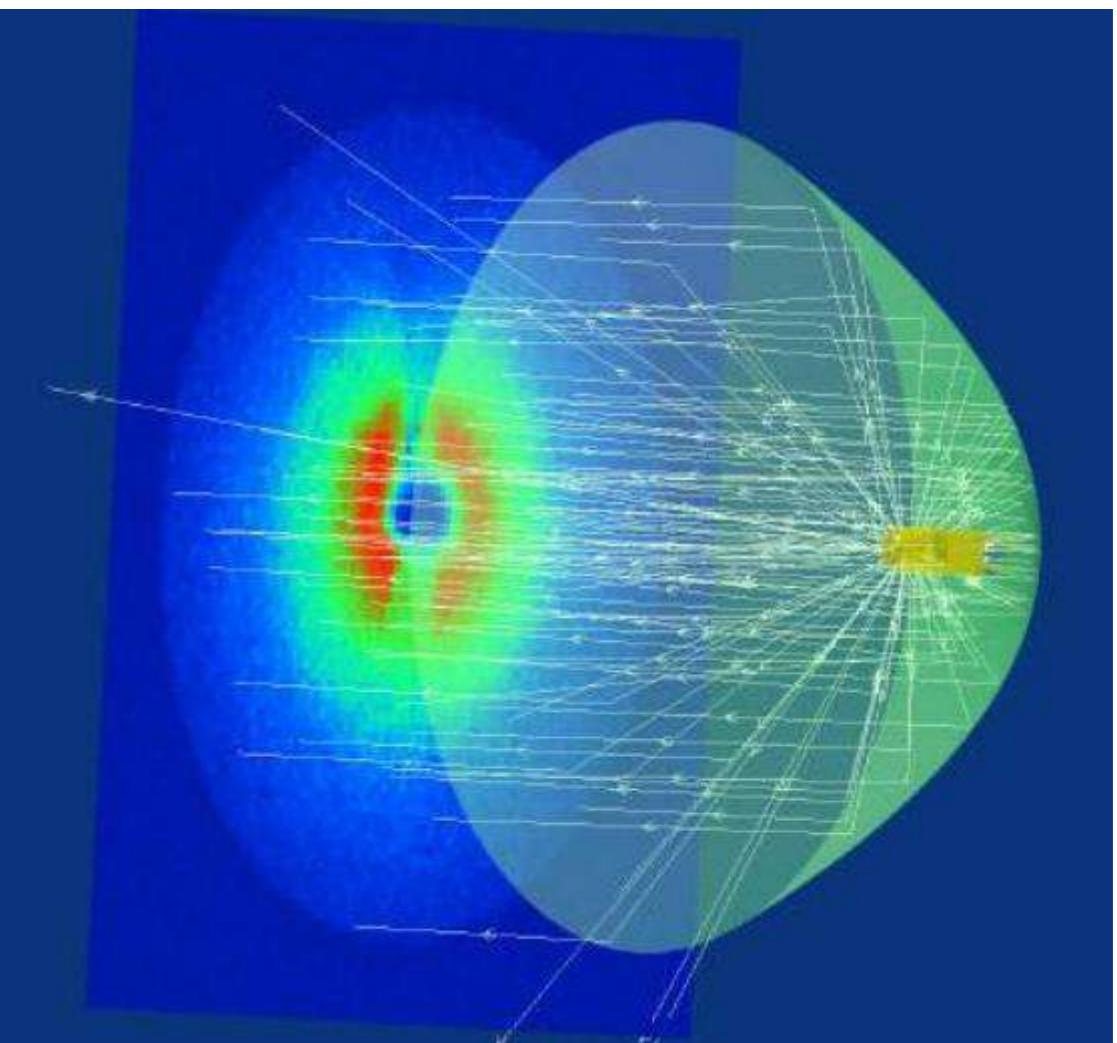
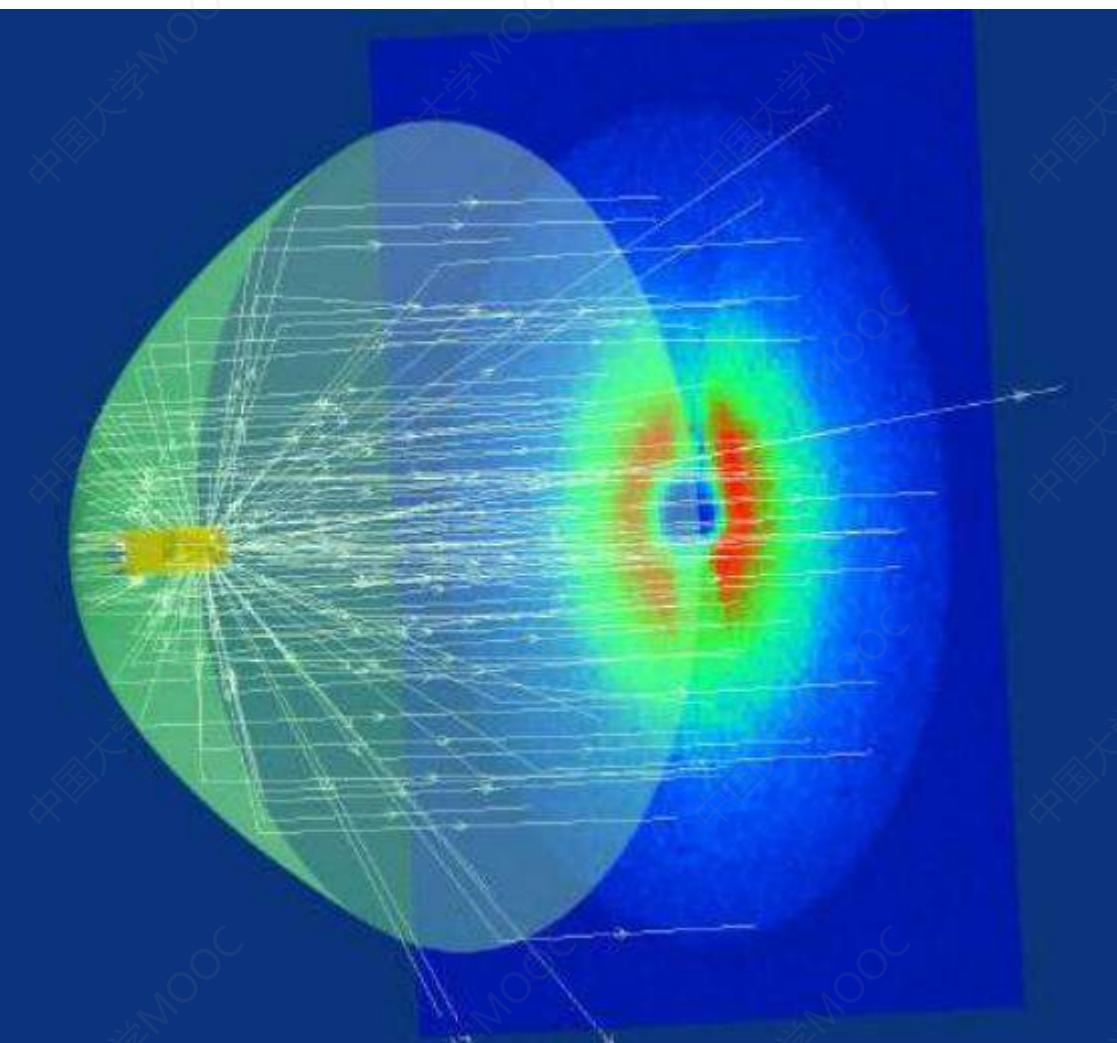
$$\begin{aligned}\Phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\end{aligned}$$

■ 8.2 形状与颜色描述子

➤ 颜色描述子——图像中的颜色分布特征

- 颜色直方图：颜色直方图用以反映图像颜色的组成分布，即各种颜色出现的频率。
- 主颜色：图像或区域中的出现像素点最多的颜色。
- 颜色矩：利用线性代数中矩的概念，将图像中的颜色分布用其矩表示。利用颜色一阶矩（平均值Average）、颜色二阶矩（方差Variance）和颜色三阶矩（偏斜度Skewness）来描述颜色分布。

8.3 区域描述子



■ 8.3 区域描述子

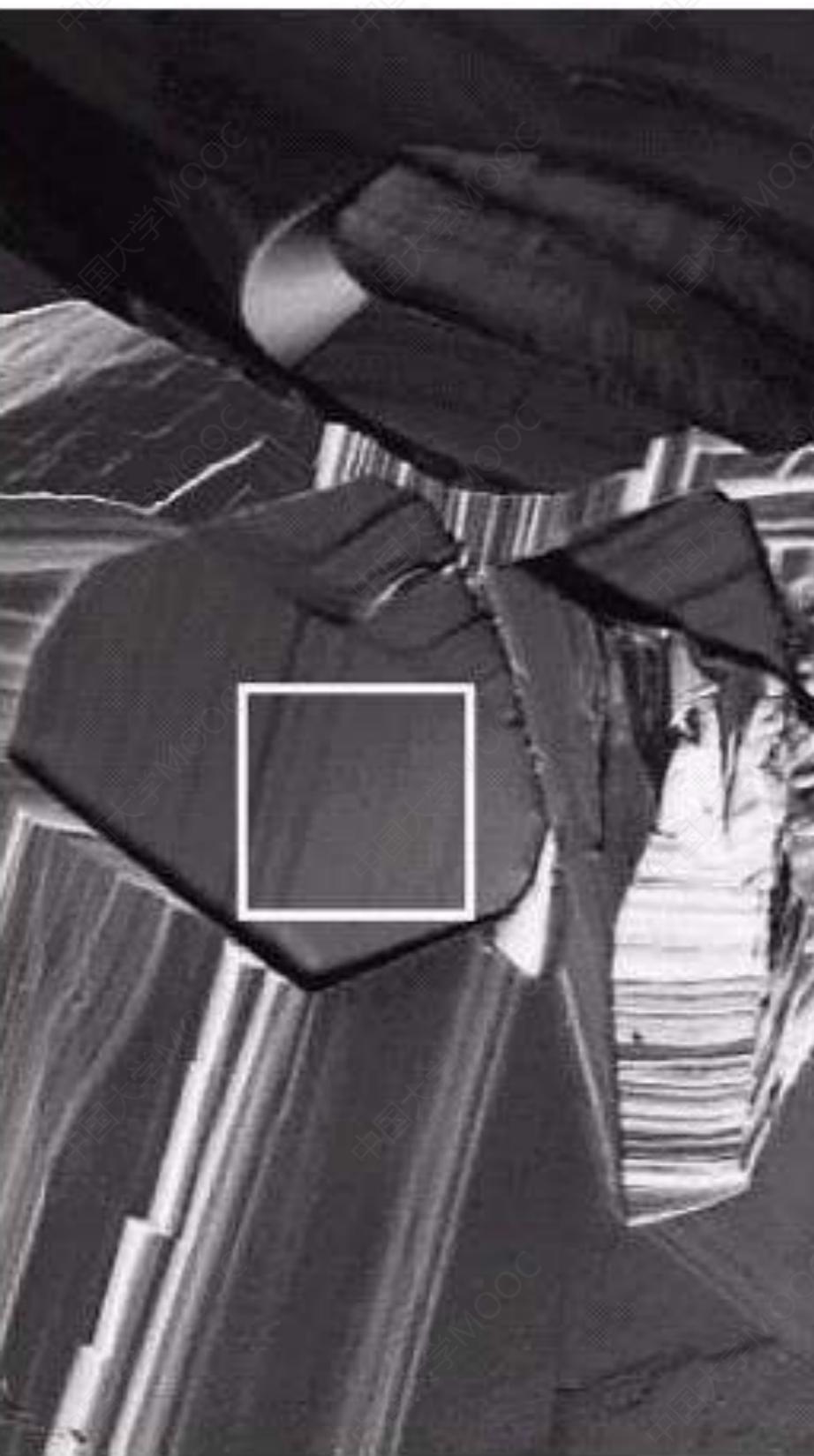
➤ 简单的描述子

- 紧致度= $(\text{周长})^2/\text{面积}$

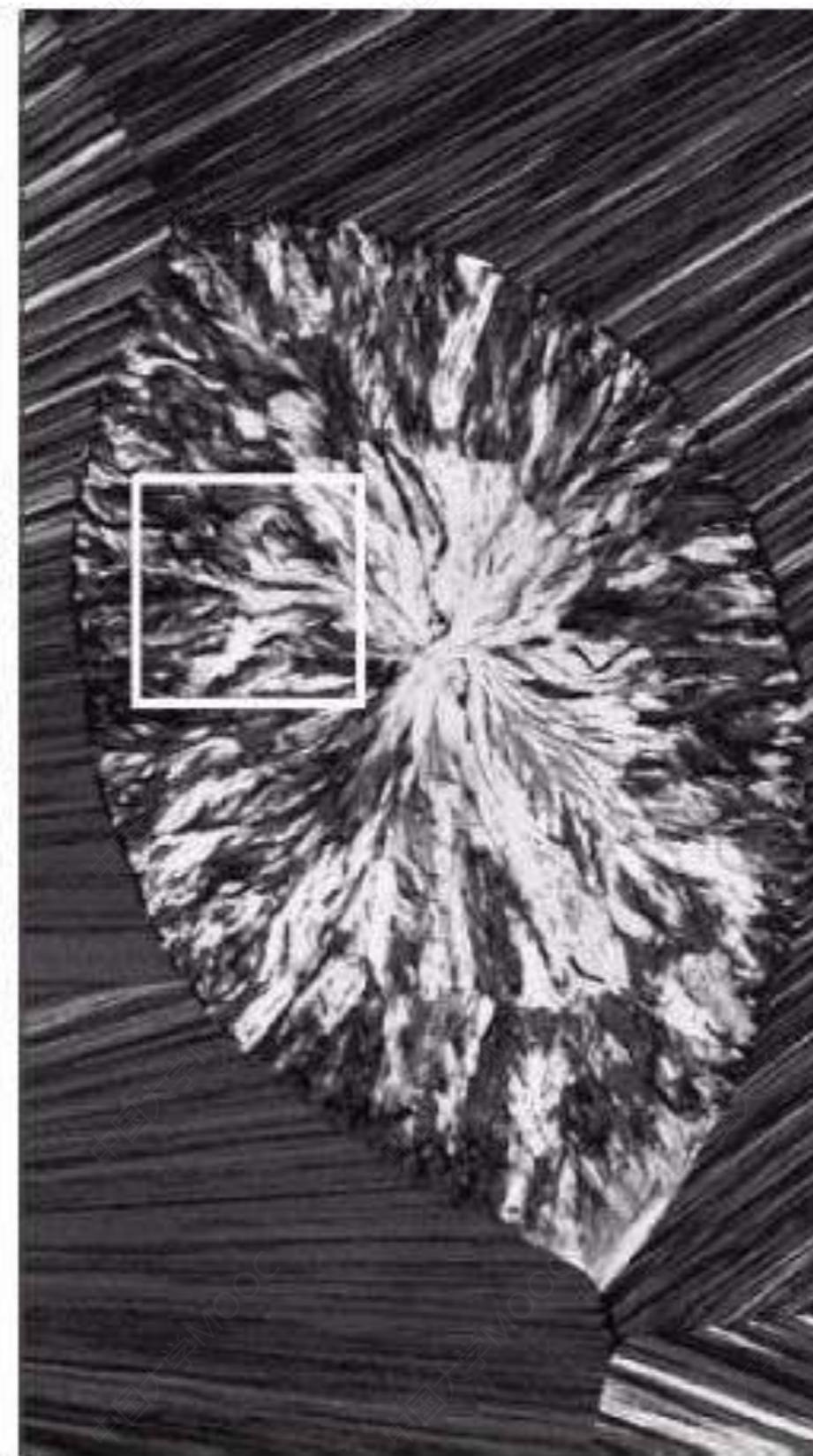
— 纹理

- 区域描述的一类重要方法是对其纹理内容进行量化。尽管没有纹理的正式定义，但直观地说，这个描述子提供了一些属性的度量，比如平滑度、粗糙度、规律性等
- 图像处理中描述区域纹理的三种主要方法是统计、结构和颗粒化。

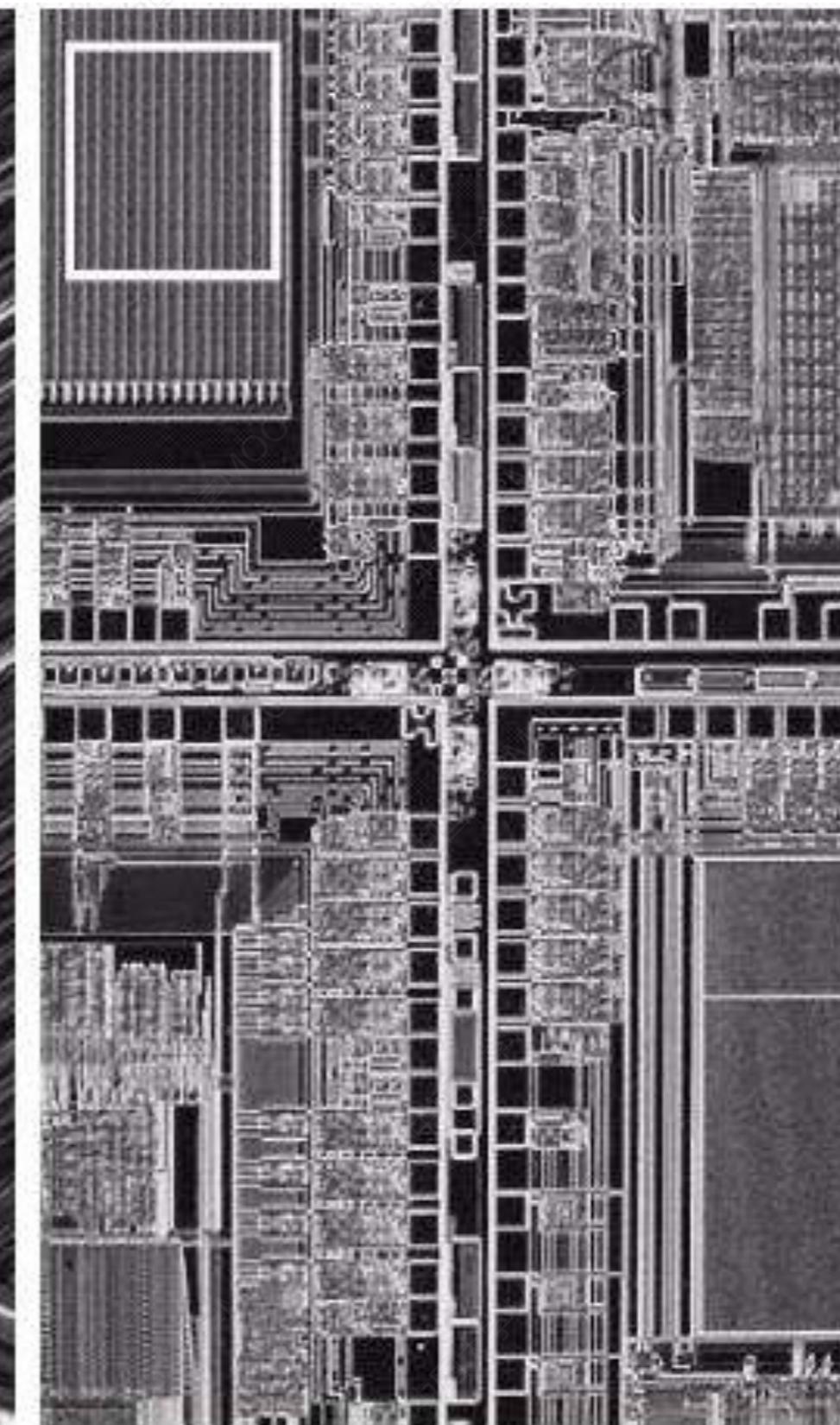
■ 8.3 区域描述子



纹理平滑区域



纹理丰富区域

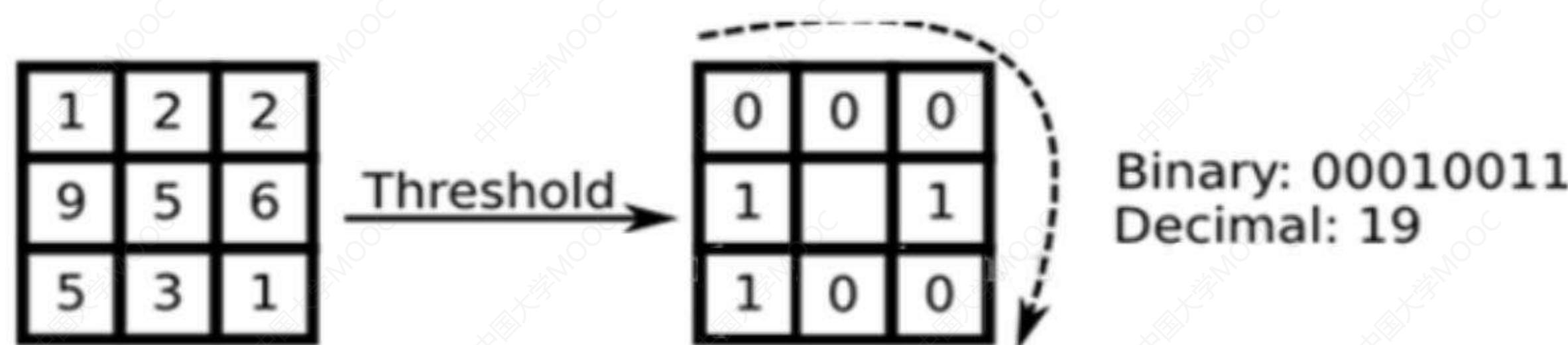
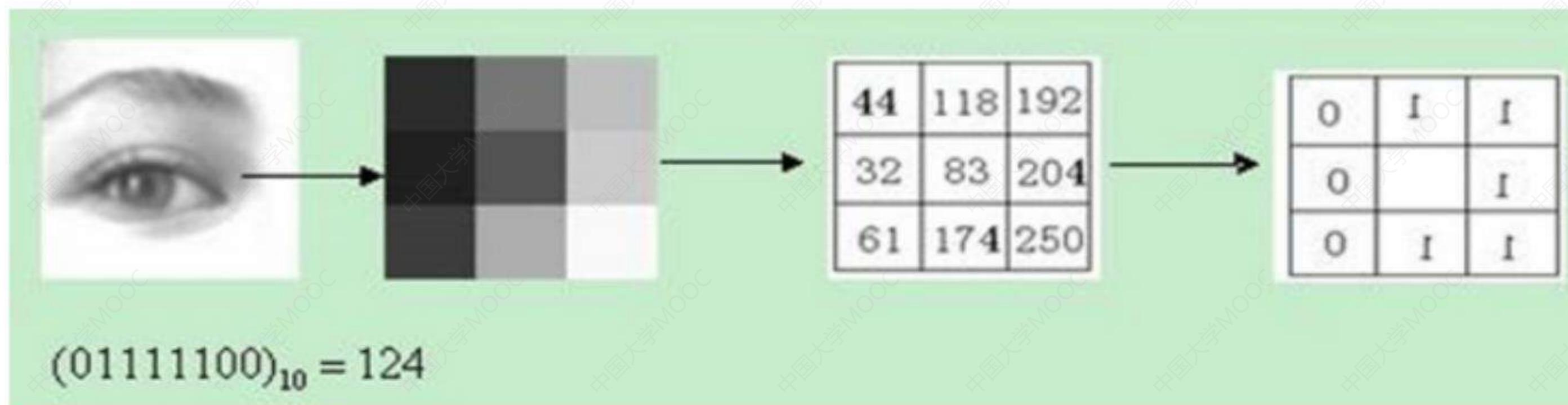


纹理规则区域

三种不同类型的纹理区域示例

■ 8.3 区域描述子

➤ LBP (local Binary pattern) 局部二值模式



用公式表示就是：

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c)$$

■ 8.3 区域描述子

➤ LBP (local Binary pattern) 局部二值模式



输入图像



LBP图像

■ 8.3 区域描述子

➤ 局部二值模式 (LBP)

- LBP (local Binary pattern) 局部二值模式:
- 旋转不变LBP
- 均匀模式LBP
- MBLBP (Multi-Block LBP)
- LBPH (Local Binary Patterns Histogram)
- LGBP (Local Gabor Binary Patterns)

■ 8.3 区域描述子

➤ 灰度共生矩阵

首先对于一幅图像定义一个方向 (orientation) 和一个以pixel为单位的步长(step)，灰度共生矩阵T ($N \times N$)，则定义 $M(i,j)$ 为灰度级为i和j的像素同时出现在一个点和沿所定义的方向跨度步长的点上的频率。其中N是灰度级划分数目。由于共生矩阵有方向和步长的组合定义，而决定频率的一个因素是对矩阵有贡献的像素数目，而这个数目要比总共数目少，且随着步长的增加而减少。因此所得到的共生矩阵是一个稀疏矩阵，所以灰度级划分N常常减少到8级。如在水平方向上计算左右方向上像素的共生矩阵，则为对称共生矩阵。

■ 8.3 区域描述子

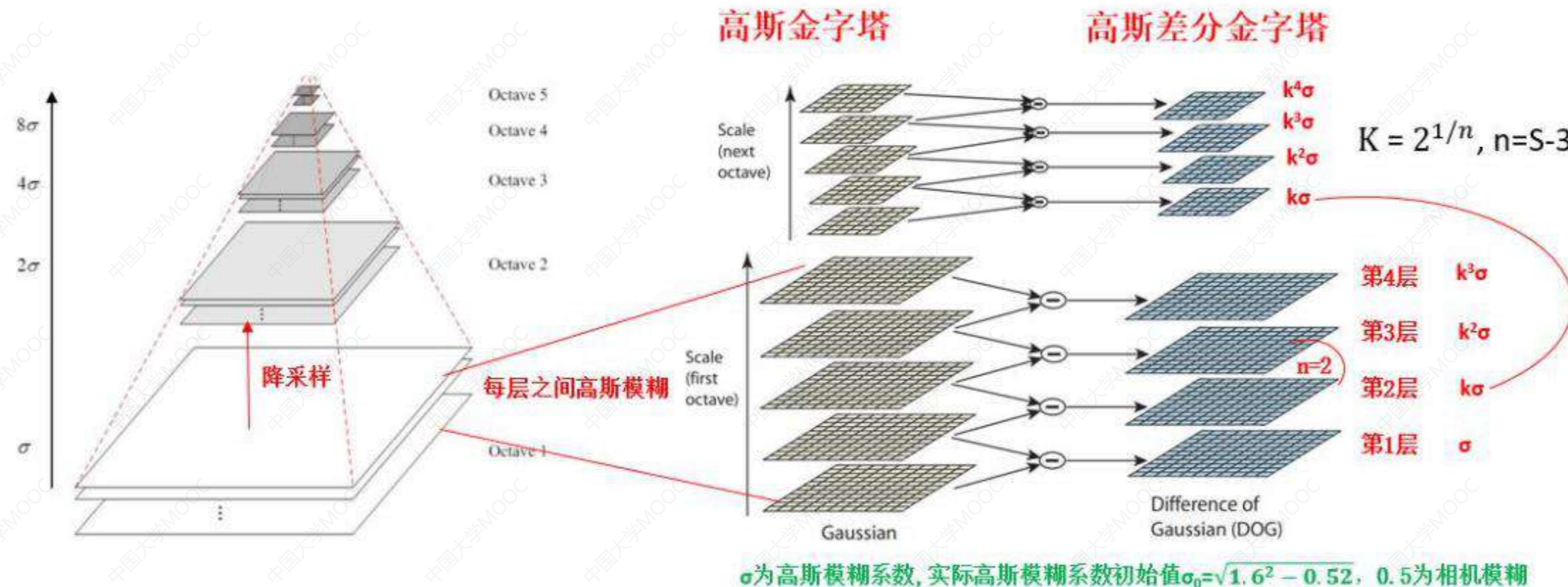
➤ SIFT (Scale-invariantfeature transform)

- SIFT (Scale-invariantfeature transform) : 尺度不变特征变换，在不同的尺度空间上查找关键点(特征点)，并计算出关键点的方向。计算步骤：
 - 1) 多尺度空间极值点检测；
 - 2) 关键点的主方向计算；
 - 3) 描述子的构造；
 - 4) 特征向量

Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60(2): 91-110

8.3 区域描述子

➤ SIFT (Scale-invariantfeature transform)

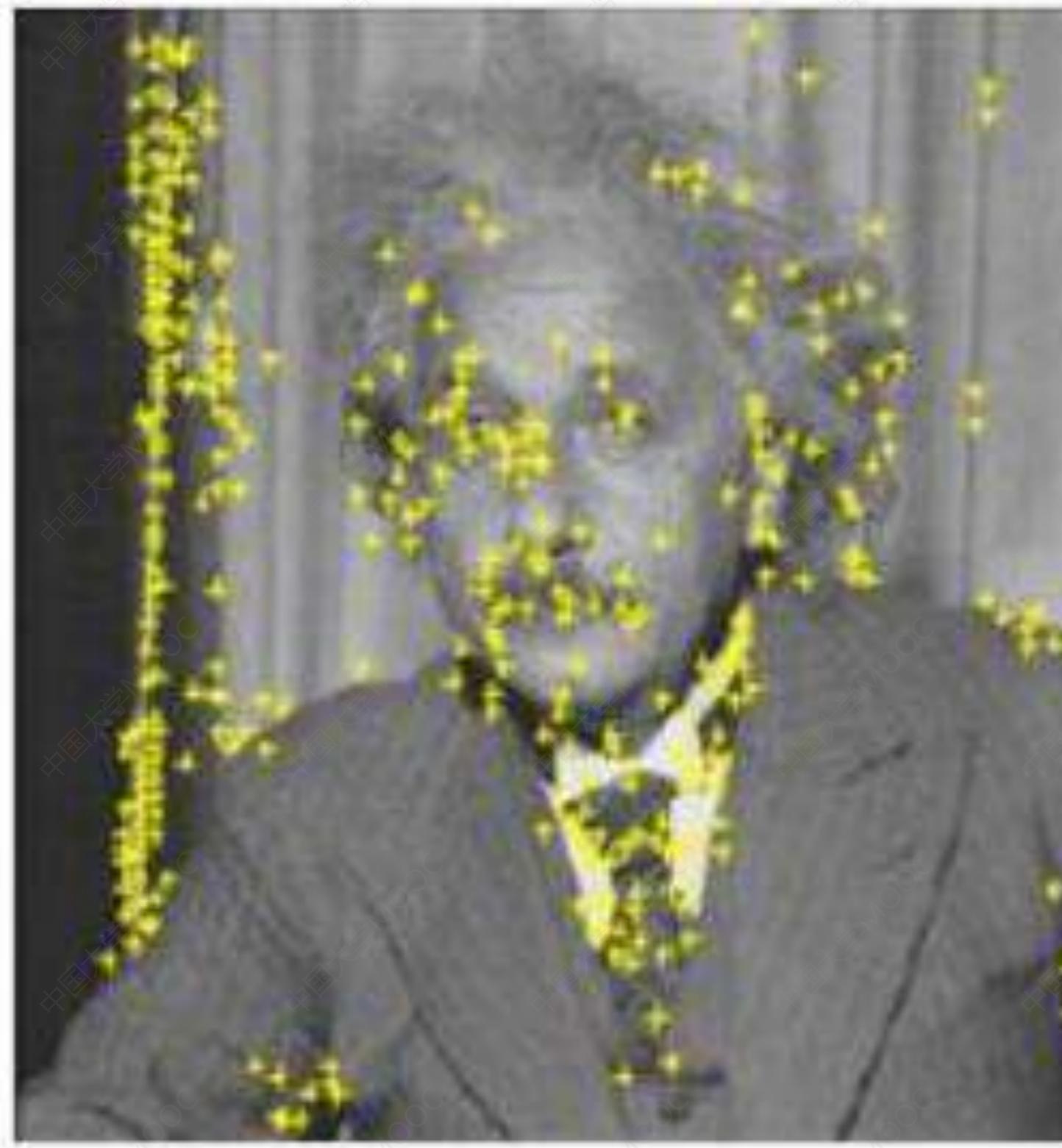


高斯差分金字塔

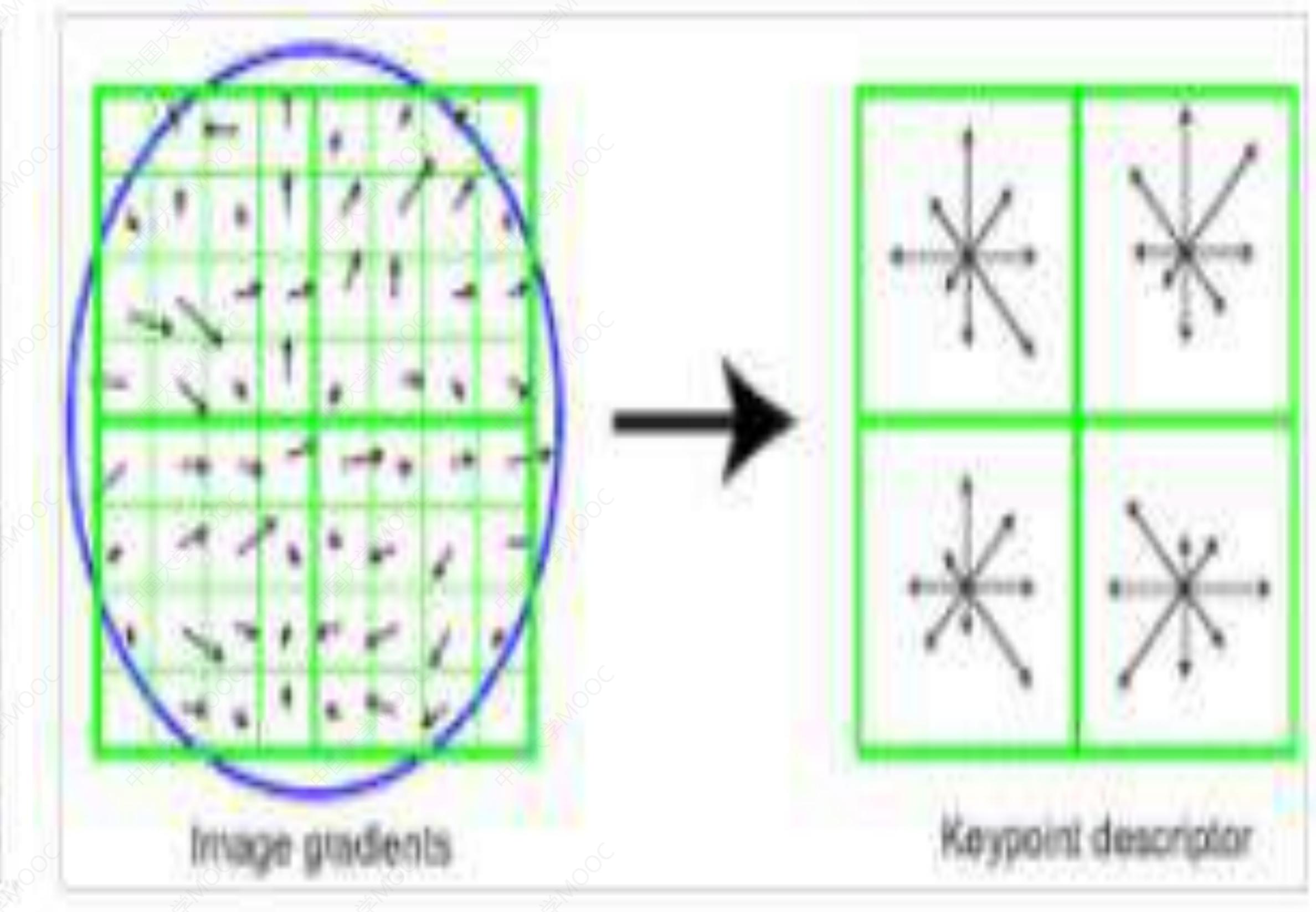
DOG空间极值点检测

■ 8.3 区域描述子

➤ SIFT (Scale-invariantfeature transform)



极值点



每个极值点的特征描述子

■ 8.3 区域描述子

➤ SIFT (Scale-invariantfeature transform)

- SIFT所找到的关键点是一些十分突出，不会因光照，仿射变换和噪音等因素而变化的点，鲁棒性较好。
- 由于 SIFT 对尺度、旋转以及一定视角和光照变化等图像变化都具有不变性，并且 SIFT 具有很强的可区分性，自它提出以来，很快在物体识别、宽基线图像匹配、三维重建、图像检索中得到了应用，局部图像特征描述子在计算机视觉领域内也得到了更加广泛的关注。

■ 8.3 区域描述子

➤ SURF (Speeded Up Robust Features)

- SURF (Speeded Up Robust Features) 是对 SIFT 的改进版本，它利用 Haar 小波来近似 SIFT 方法中的梯度操作，同时利用积分图技术进行快速计算，SURF 的速度是 SIFT 的 3-7 倍，大部分情况下它和 SIFT 的性能相当，因此它在很多应用中得到了应用，尤其是对运行时间要求高的场合。

Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008

■ 8.3 区域描述子

➤ ASIFT (Affine SIFT)

- ASIFT (Affine SIFT) 通过模拟所有成像视角下得到的图像进行特征匹配，可以很好地处理视角变化的情况，尤其是大视角变化下的图像匹配。

➤ DAISY

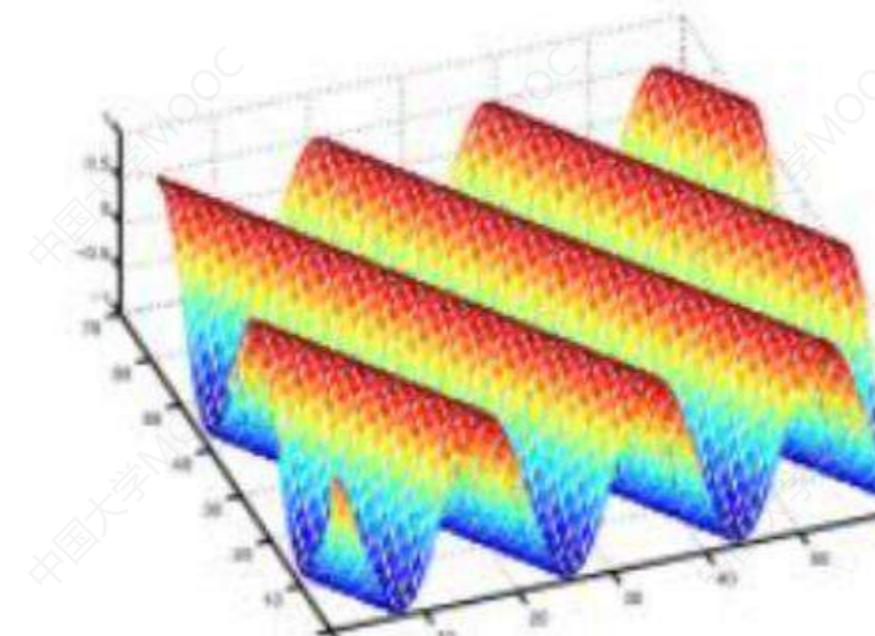
- 面向稠密特征提取的可快速计算的局部图像特征描述子，本质思想和SIFT是一样的。

Tola E , Lepetit V , Fua P . DAISY: an efficient dense descriptor applied to wide-baseline stereo.[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010. 32(5):815-830

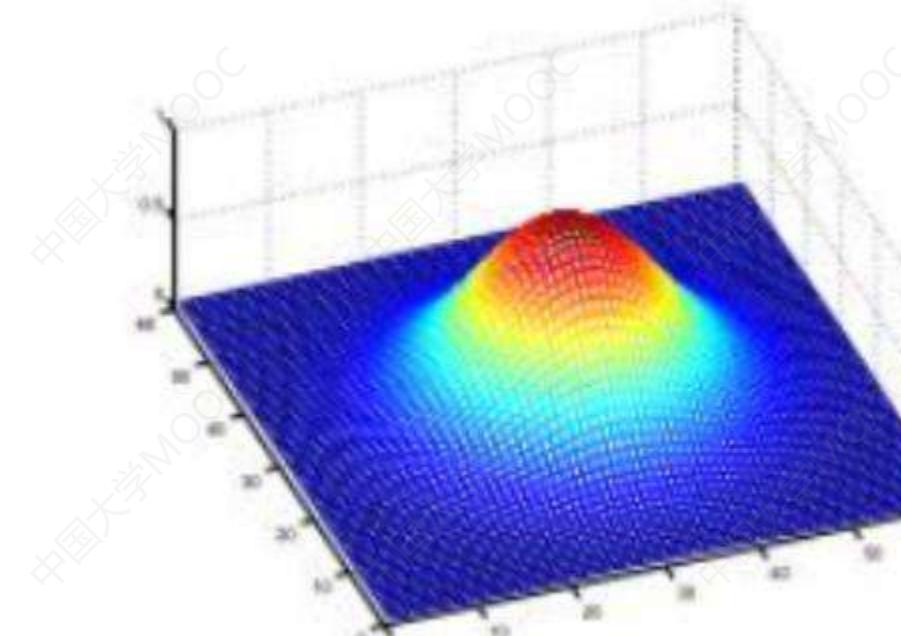
■ 8.3 区域描述子

➤ Gabor特征

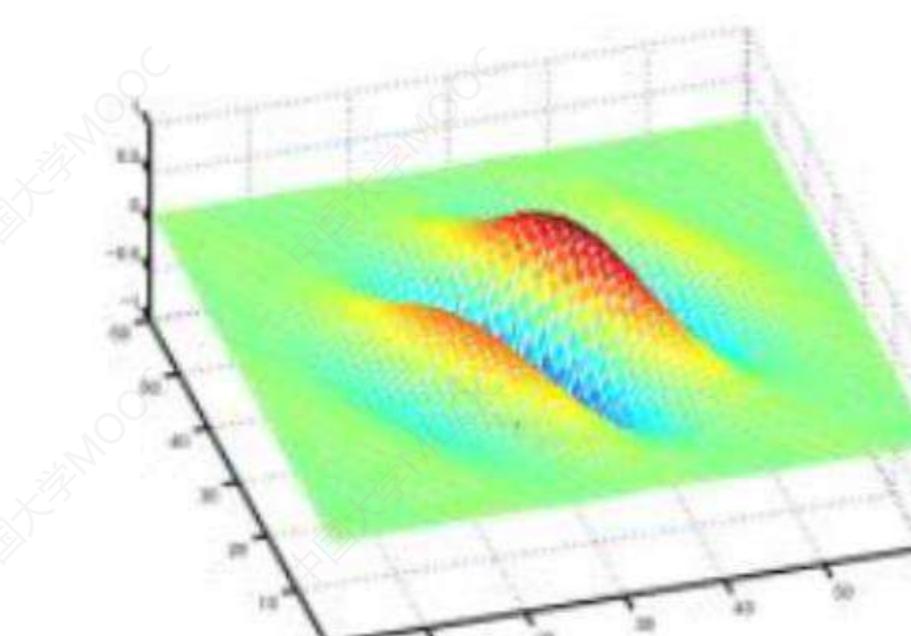
- Gabor特征是一种可以用来描述图像纹理信息的特征，Gabor滤波器的频率和方向与人类的视觉系统类似，特别适合于纹理表示与判别，Gabor特征主要依靠Gabor核在频率域上对信号进行加窗，从而能描述信号的局部频率信息。
- Gabor核本质上是高斯核与正弦波调制的结果。
- 如下图所示，分别为正弦函数，高斯函数，调制后的Gabor核函数。通过频率参数和高斯函数参数的选择，Gabor变换可以选取很多纹理特征，但是Gabor是非正交的，不同特征分量之间有冗余。



正弦函数



高斯函数



Gabor核函数

■ 8.3 区域描述子

➤ Gabor特征

- 从傅里叶变换的角度来看，Gabor变换就是窗函数取高斯窗时的短时傅里叶变换。

- 复数
$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i(2\pi \frac{x'}{\lambda} + \psi)\right)$$

- 实部
$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right)$$

- 虚部
$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi \frac{x'}{\lambda} + \psi\right)$$

- 这里 $x' = x \cos \theta + y \sin \theta$ $y' = -x \sin \theta + y \cos \theta$

- λ 为正弦函数的波长， θ 为核函数方向， ψ 为相位偏移， σ 为高斯标准差， γ 为 x 、 y 两个方向的纵横比
(指定了Gabor函数的椭圆率)

■ 8.3 区域描述子

➤ Gabor特征

- 一个Gabor核能获取到图像某个频率邻域的响应情况，这个响应结果可以看做是图像的一个特征。如果用多个不同方向、不同频率的Gabor核去获取图像相应频率和方向的响应情况，这个特征就可以描述图像的频率信息。

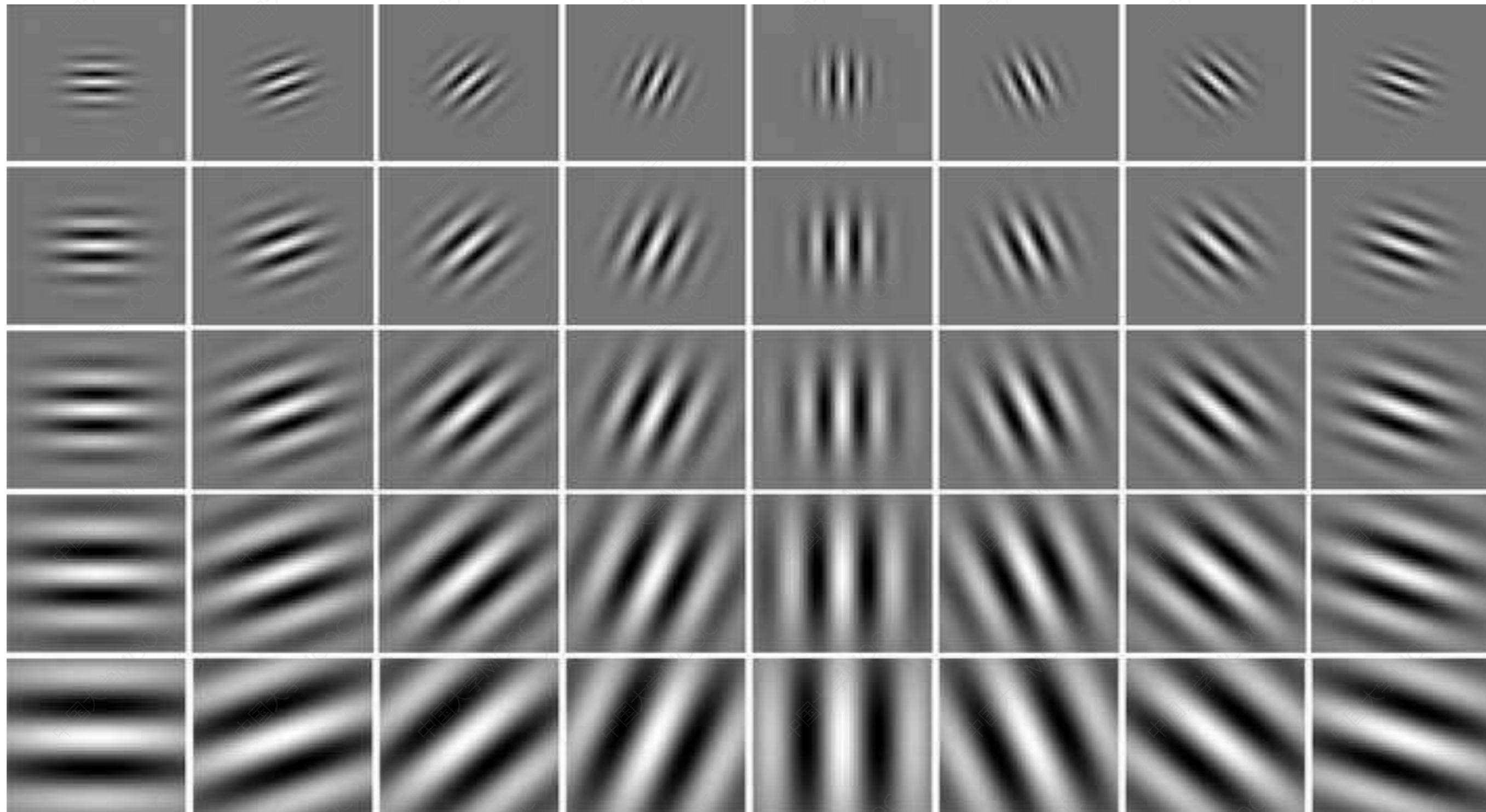
$$Gabor(x_0, y_0, \theta, \sigma_x, \sigma_y, u_0, v_0) = s(x, y) \omega_r(x, y)$$

$$= K \exp(-\pi((x - x_0)^2 / \sigma_x^2 + (y - y_0)^2 / \sigma_y^2)) \exp(i2\pi(u_0 x + v_0 y))$$

- (x_0, y_0) 高斯核的中心点, θ 高斯核的旋转方向 (顺时针), (σ_x, σ_y) 高斯核两个方向上的尺度, (u_0, v_0) 频域坐标, K 高斯核的幅度比例。

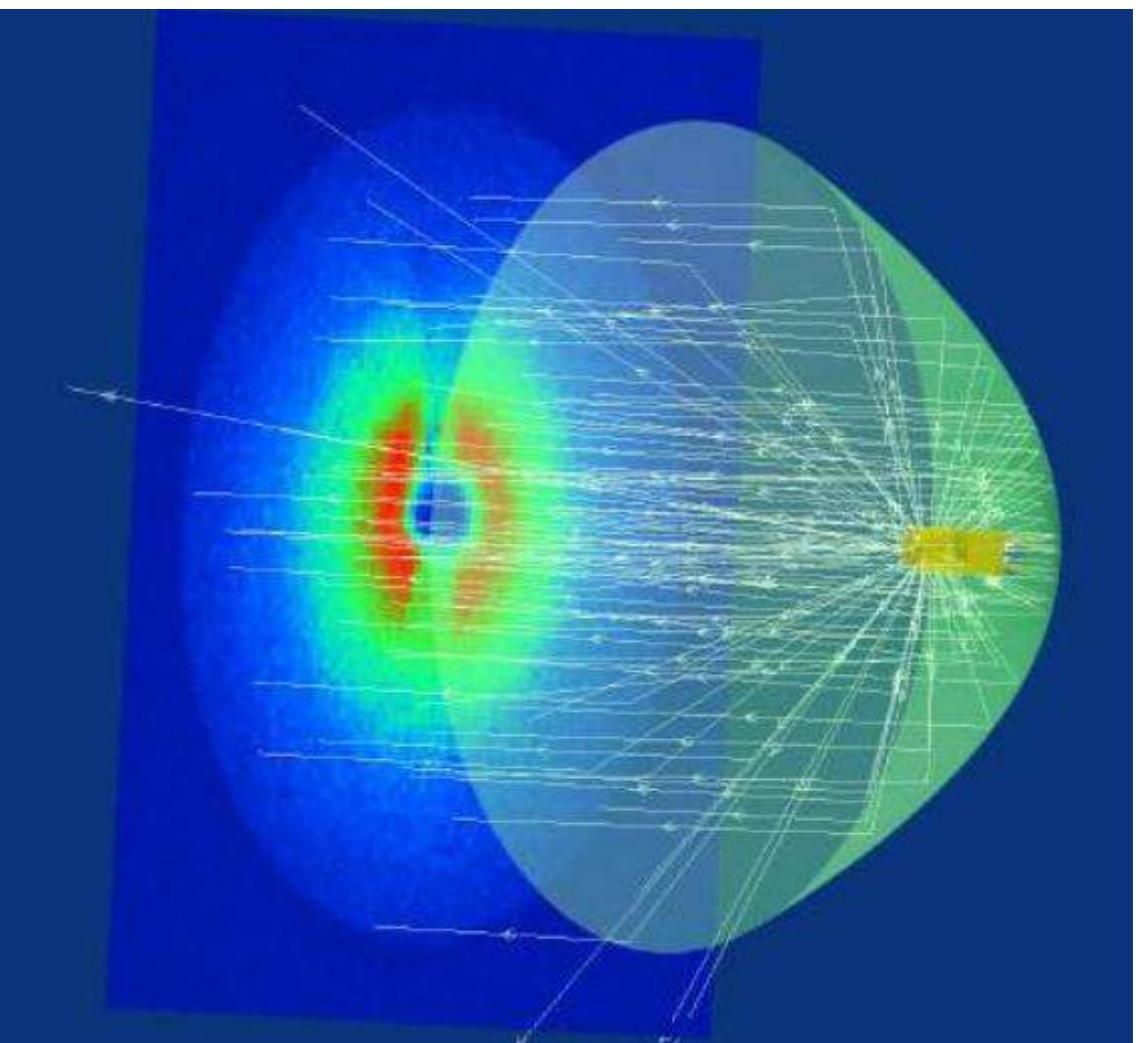
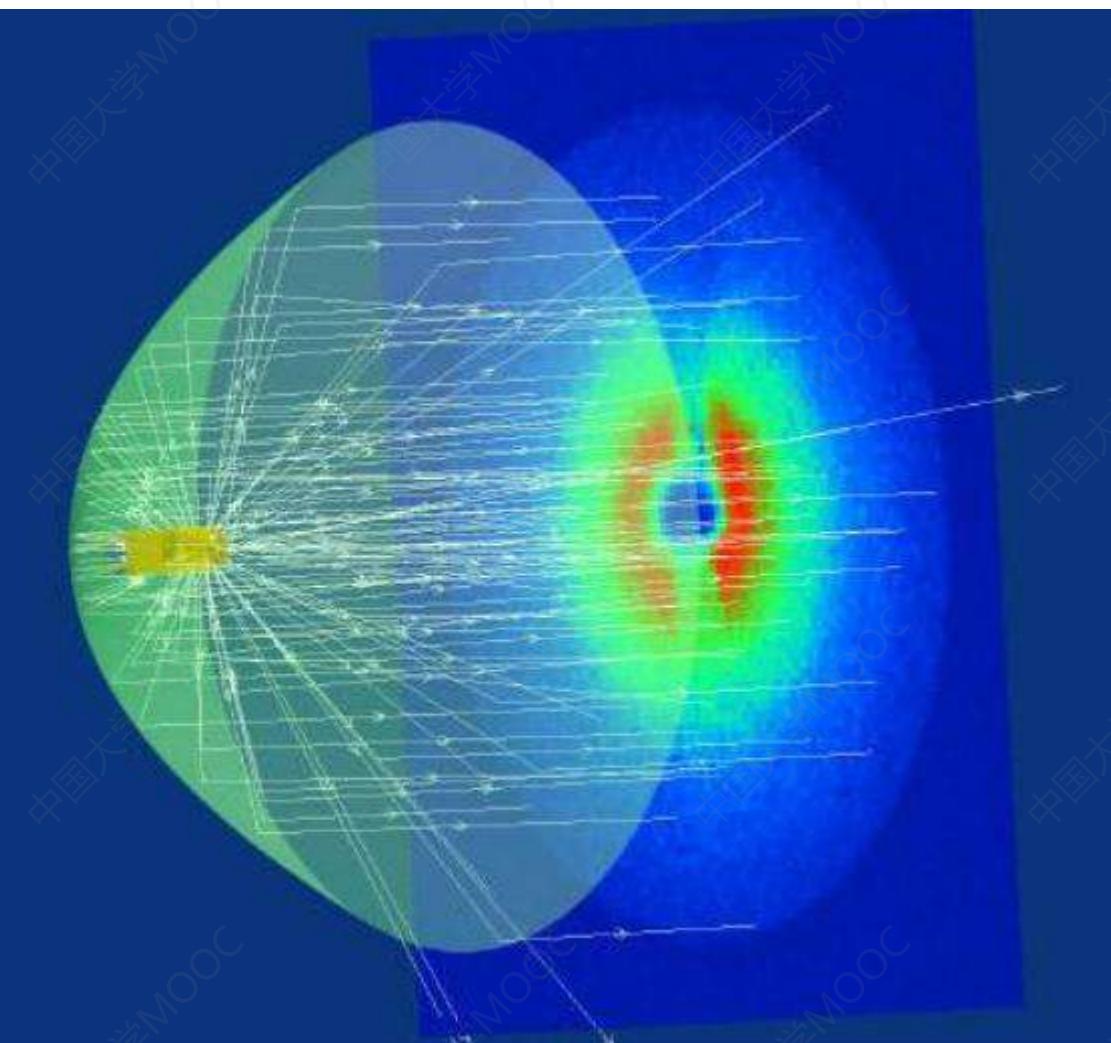
■ 8.2 区域描述子

➤ Gabor特征



一组Gabor核函数

8.4 模式识别/分类



■ 8.4 模式识别/分类

- 模式和模式类
- 最小距离分类器
- 贝叶斯分类器
- 神经网络分类器
- SVM分类器

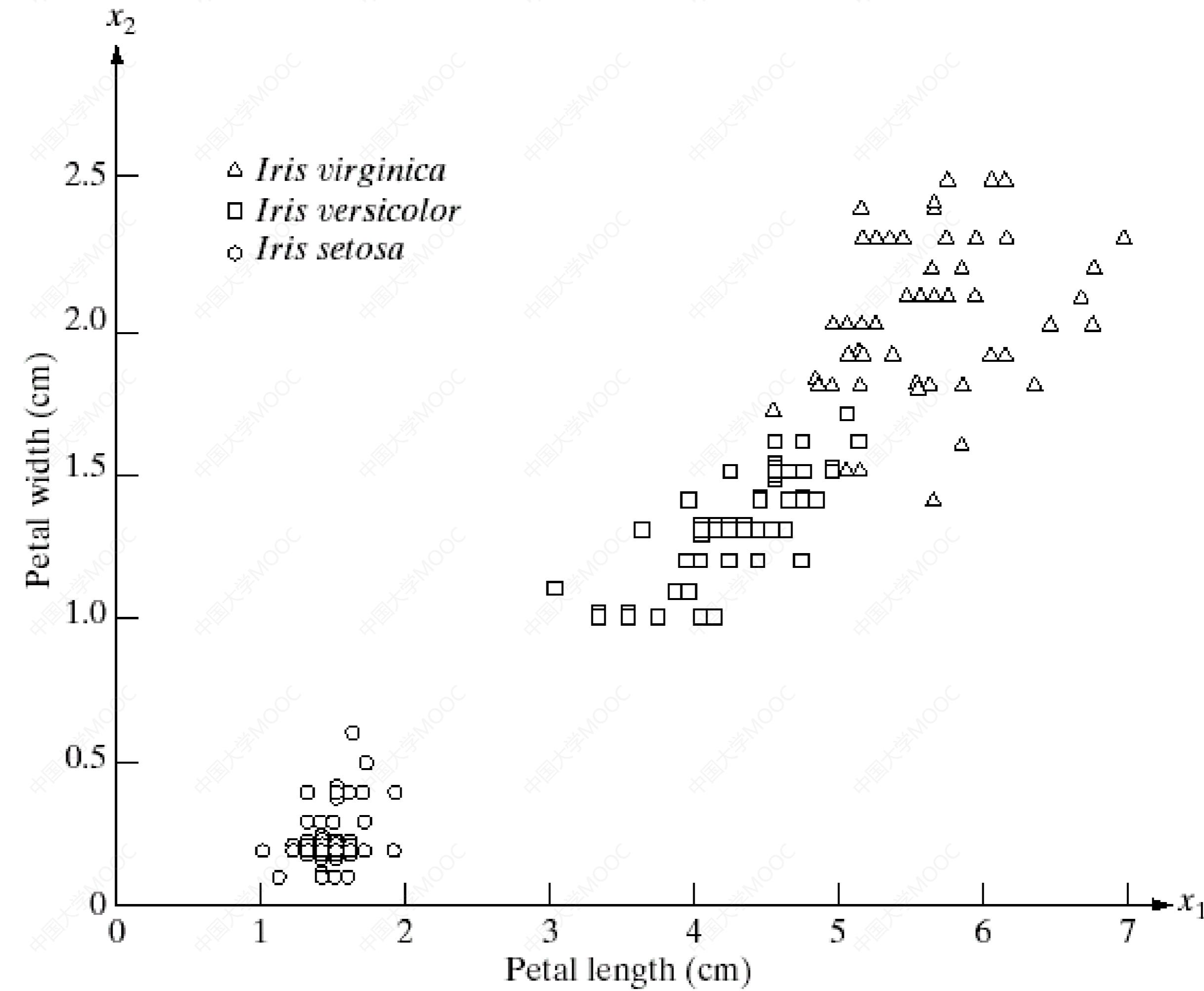
■ 8.4 模式识别/分类

➤ 模式和模式类

- 一个模式是一组特征向量, $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, n为特征向量维数;
- 模式类是对应于多个类标签的模式, 不同簇的模式对应于不同的类标签 $\omega_1, \omega_2, \dots, \omega_w$, w 是标签类别数。

8.4 模式识别/分类

➤ 模式和模式类



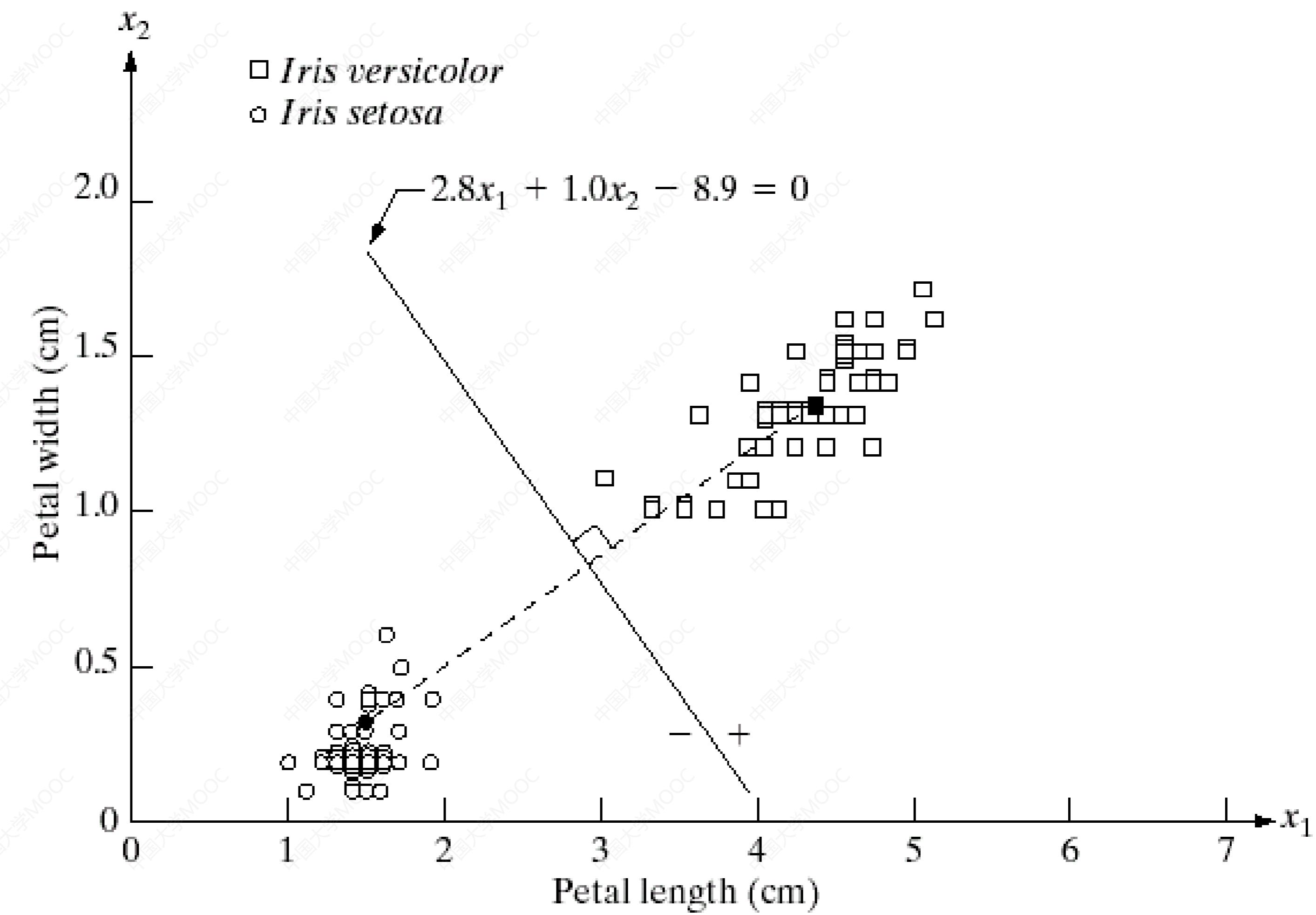
■ 8.4 模式识别/分类

➤ 最小距离分类器

- 假设 $x = [x_1, x_2, \dots, x_n]^T$ 是一个 n 维模式向量，有 W 个模式类 $\omega_1, \omega_2, \dots, \omega_W$ 。
- 模式识别的任务就是给一个模式向量分配类标签的过程。
- 模式识别都需要首先进行学习，学习的模式类别是固定的，所以又称模式分类。模式识别（分类）的本质就是寻找模式空间的最优分类边界或者判决函数 $d_i(x)$, $i=1, 2, \dots, W$
- 对于一个未知模式 x , 如果 $d_i(x)$ 最大（最小）则认为 x 属于第 i 类
- 分类边界意味着在其上的模式向量到两类的判决函数值相等 $d_i(x) = d_j(x)$ 或者 $d_i(x) - d_j(x) = 0$

8.4 模式识别/分类

➤ 最小距离分类器



■ 8.4 模式识别/分类

➤ 贝叶斯分类器

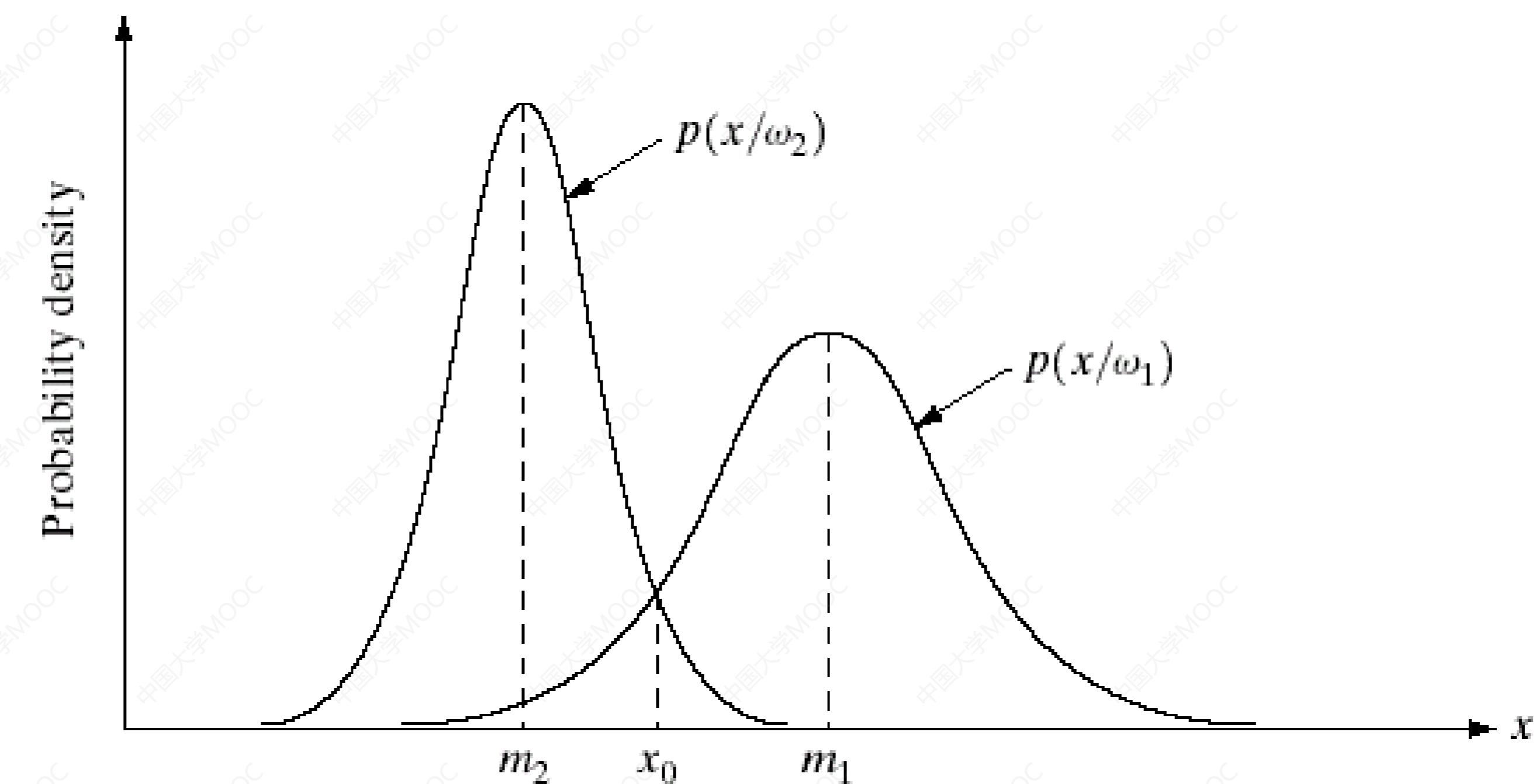
- 概率统计的方法，分类边界的划分保证错误分类的样

本数最少。

$$P(\omega_i|x) = \frac{P(x|\omega_i)P(\omega_i)}{P(x)} \quad i = 1, 2, \dots, W$$

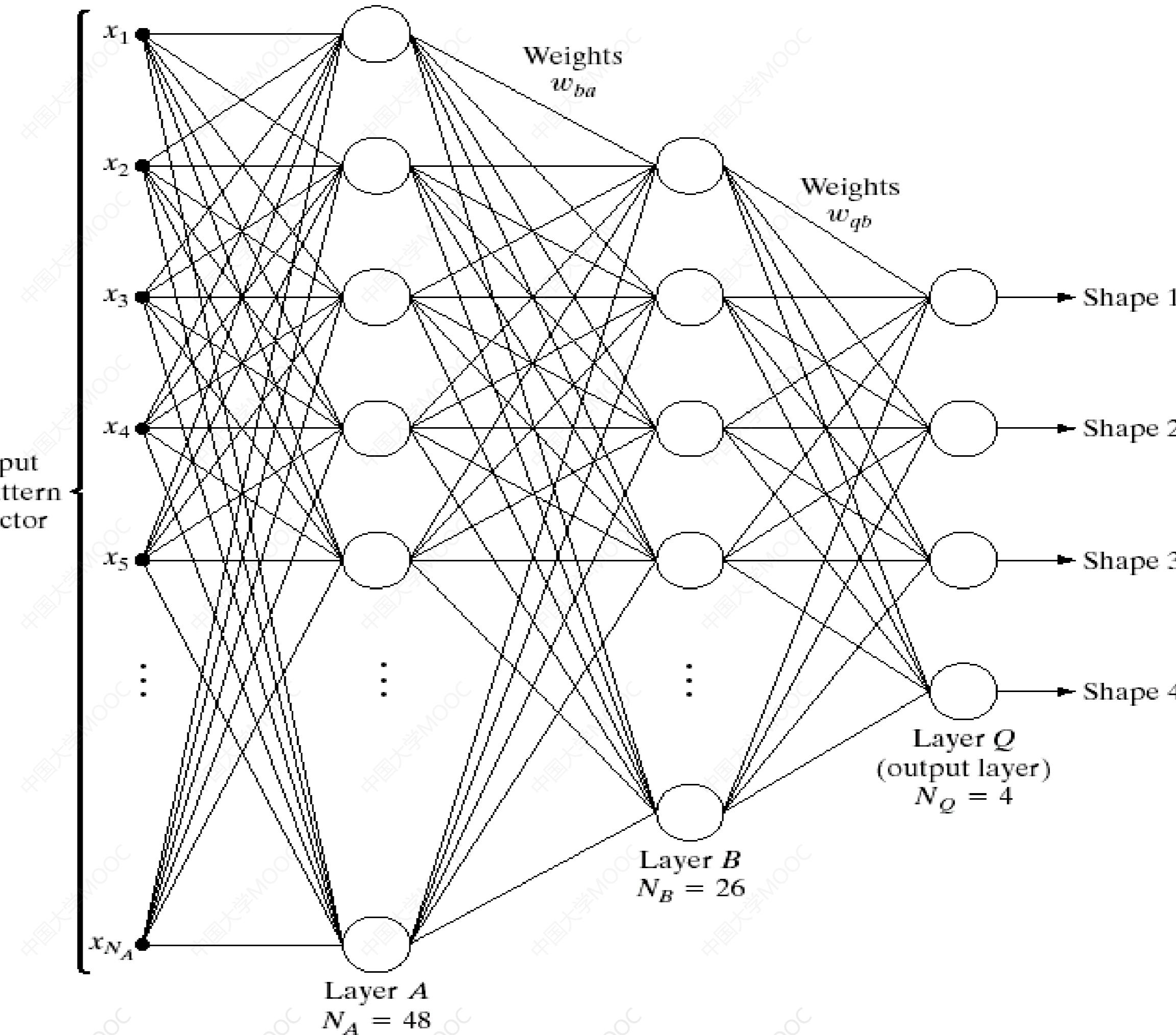
8.4 模式识别/分类

➤ 贝叶斯分类器



■ 8.4 模式识别/分类

➤ 神经网络分类器



■ 8.4 模式识别/分类

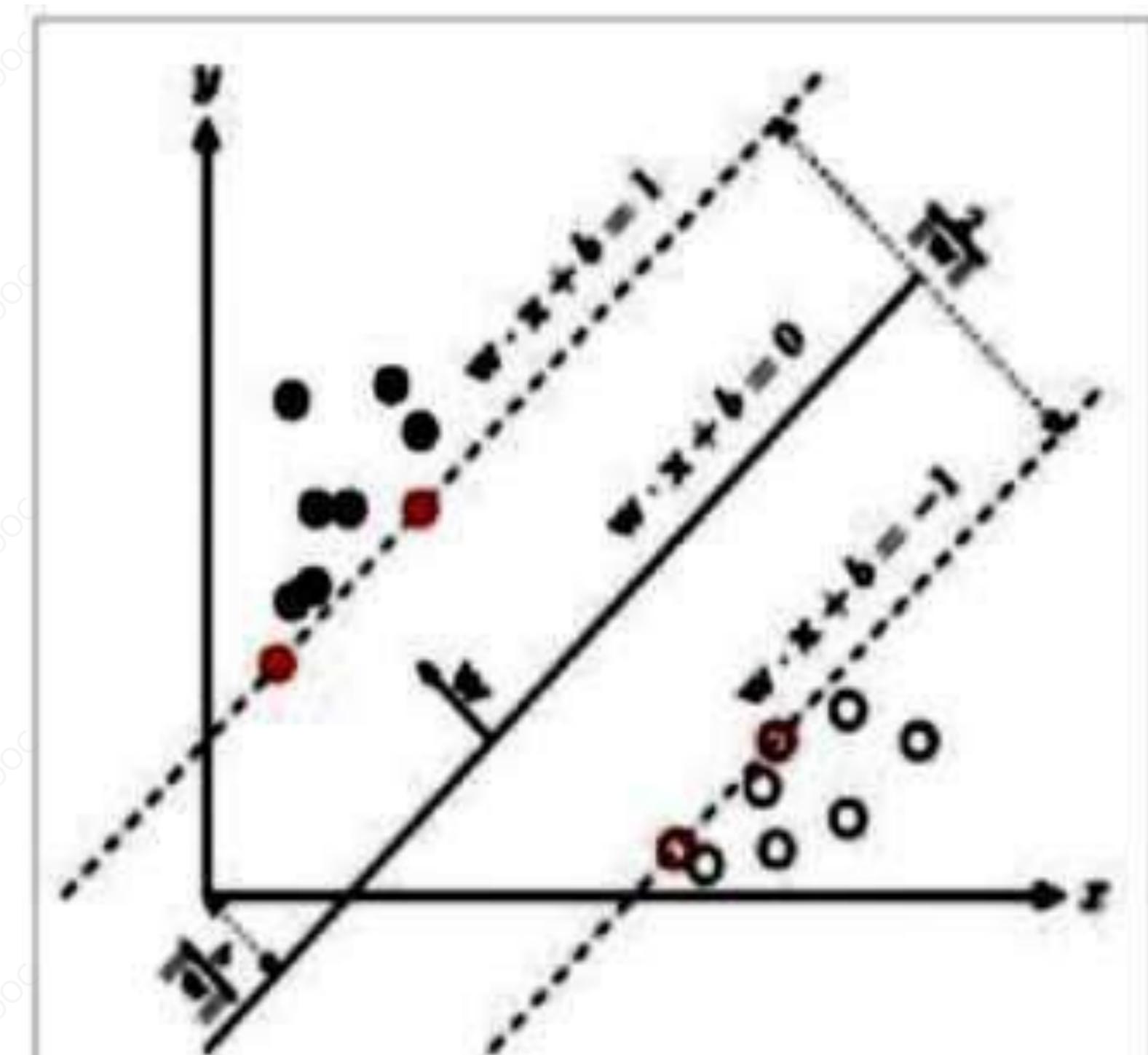
➤ SVM分类器

- SVM又称支持向量机（Support Vector Machine），基本模型是定义在特征空间上间隔最大的线性分类器。是一种二分类模型。
- 超平面——分类的决策边界。在SVM中，希望找到离分隔超平面最近的点（称为支持向量），确保它们离分隔超平面的距离尽可能的远。通过最大化支持向量到分隔超平面的距离来训练SVM。

■ 8.4 模式识别/分类

➤ SVM分类器

- 在分类问题中给定输入数据 $X = \{X_1, \dots, X_N\}$ 和学习目标 $Y = \{y_1, \dots, y_N\}$ ，其中输入数据的每个样本都包含多个特征并由此构成特征空间(feature space)，而学习目标为二元变{-1,1}表示负类(negative class)和正类(positive class)。
- 若输入数据所在的特征空间存在作为决策边界(decision boundary)的超平面将学习目标按正类和负类分开，并使任意样本的点到平面距离大于等于1。
- 决策边界 $w^T X + b = 0$
- 点到平面的距离 $y_i(w^T X_i + b) \geq 1$
- 则称该问题具有线性可分性，参数 w, b 分别为超平面的法向量和截距。



线性可分：决策边界（实），间隔边界（虚），支持向量（红点）
[2]

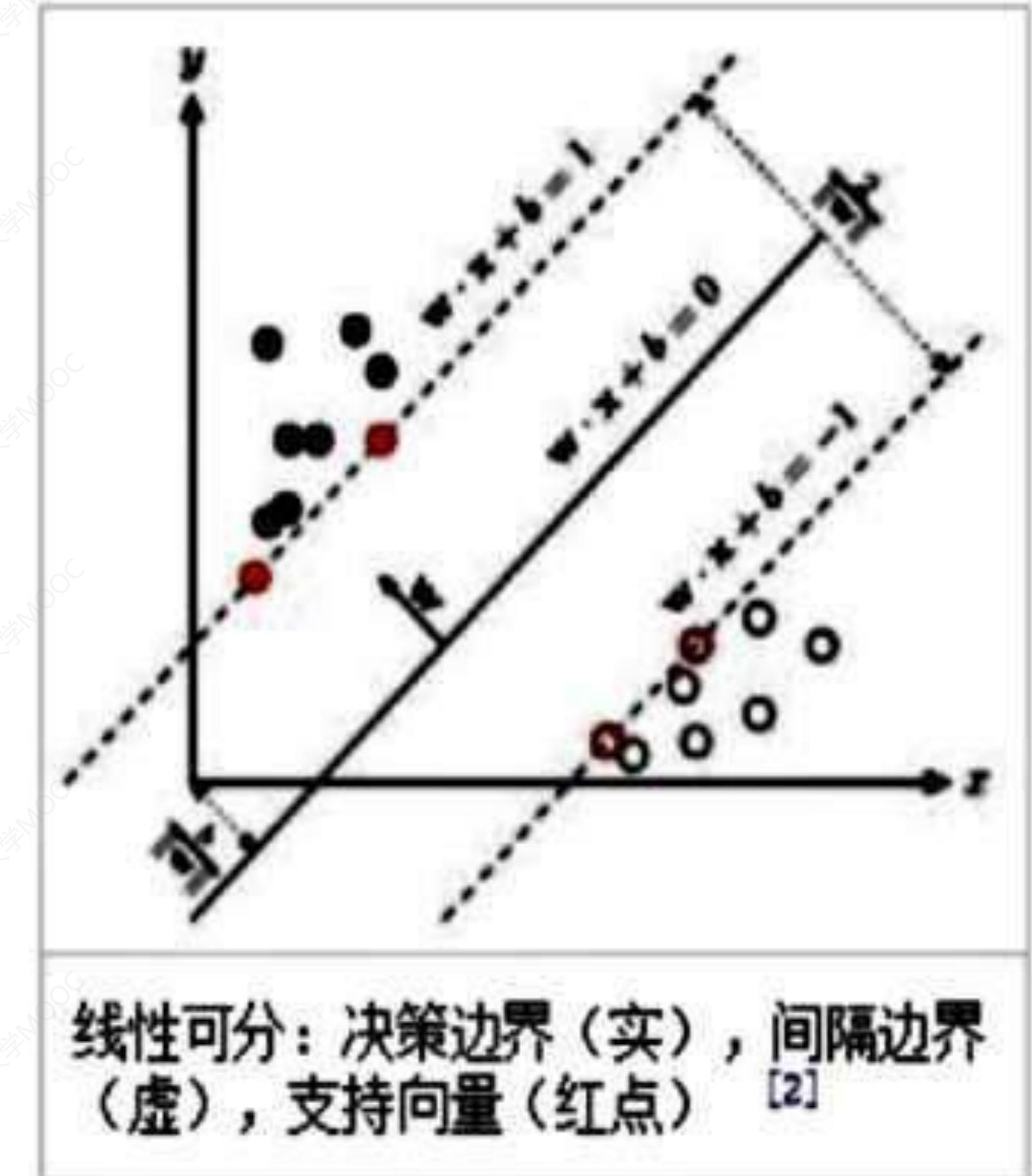
■ 8.4 模式识别/分类

➤ SVM分类器

- 满足该条件的决策边界实际上构造了2个平行的超平面作为间隔边界以判别样本的分类：

- $w^T X + b \geq +1 \Rightarrow y_i = +1$
- $w^T X + b \leq -1 \Rightarrow y_i = -1$

- 上间隔边界上方的所有样本属于正类，下间隔边界下方的样本属于负类。两个间隔边界的距离于间隔边界上的正类和负类样本为支持向量 (support vector)。
$$d = \frac{2}{\|w\|}$$
 (margin) 位



■ 8.4 模式识别/分类

➤ SVM分类器

- 对特征空间划分的最优超平面是SVM的目标,最大化分类边际的思想是SVM方法的核心
- 支持向量是SVM的训练结果,在SVM分类决策中起决定作用的是支持向量
- 在小样本训练集上能够得到比其它算法好很多的结果。