



北京航空航天大学

Beijing University of Aeronautics and Astronautics

数字图像处理

图像的识别与解释



■ 内容回顾

■ 视觉理论与信息系统

■ 运动目标检测跟踪



- **特征点匹配**
- **特征内容匹配**



• 特征点匹配

- 图像匹配的计算量很大，而且受图像采集时光照影响很大
- 可以先取得图像中一些特征点（根据灰度特点、几何位置特点等），借助特征点进行匹配
- 常用特征点：边缘点、角点、拐点



• 主要步骤

- Step1: 特征点的选取
- Step2: 匹配特征点, 计算变换模型
- Step3: 对结果进行插值





- Harris角点

- 根据点的梯度值来判断是否为特征点，时间复杂度较小，准确度较高
- 具有旋转不变(Rotation Invariant)和光照不变性
- 步骤：
 - Step1:对每一像素点计算自相关矩阵M

$$M = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} = \begin{pmatrix} A & D \\ C & B \end{pmatrix}$$

- Step2:计算每像素点的Harris角点响应

$$R = (AB - C^2) - k(A + B)^2$$

- Step3:在 $w \times w$ 范围内寻找极大值点，若Harris角点响应大于阈值，则视为角点

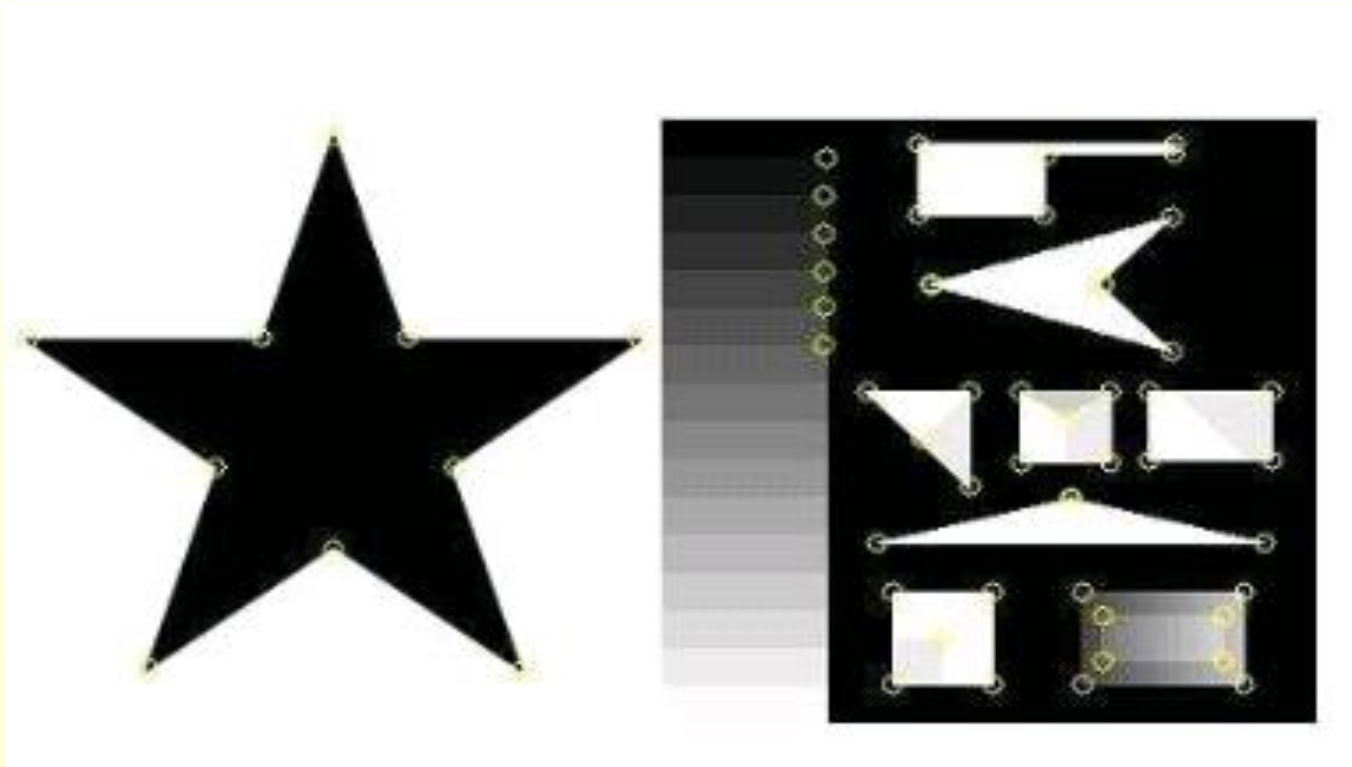


Harris算子是一种有效的点特征提取算子

优点:

- 计算简单，只用到灰度的一阶差分及滤波
- 提取的特征均匀且合理，在纹理丰富区域可以提取出大量有用的特征点，在纹理较少区域，提取的特征点较少
- 稳定，对图像旋转、灰度变化、噪声影响和视点变化不敏感

局限性: 对尺度敏感，提取的角点是像素级的



Harris算子对简单图像的响应



- SIFT特征点匹配

- **S**cale **I**nvariant **F**eature **T**ransform, 即尺度不变特征变换

- SIFT特征点匹配基本思想:

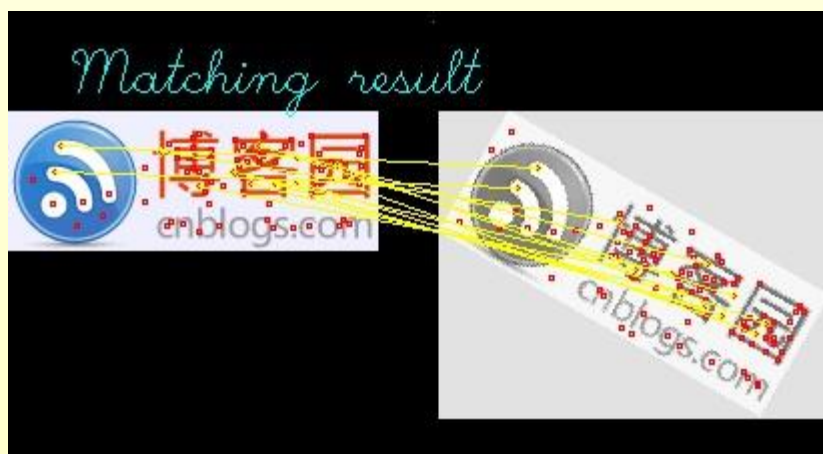
- 首先在尺度空间进行特征检测, 并确定关键点(Keypoints)的**位置**和关键点所处的**尺度**: 在图像二维平面空间和高斯差分尺度空间同时检测局部极值, 即一个点如果在DOG尺度空间本层以及上下层的26个邻域中是最大或最小值时, 就认为该点是图像在该尺度下的一个特征点。
 - 然后使用关键点邻域梯度的主方向作为该点的**方向**特征, 以实现算子对尺度和方向的无关性
 - SIFT特征向量的生成: 对于一个关键点产生128个数据, 即最终形成128维的SIFT特征向量



北京航空航天大学

Beijing University of Aeronautics and Astronautics

SIFT特征点



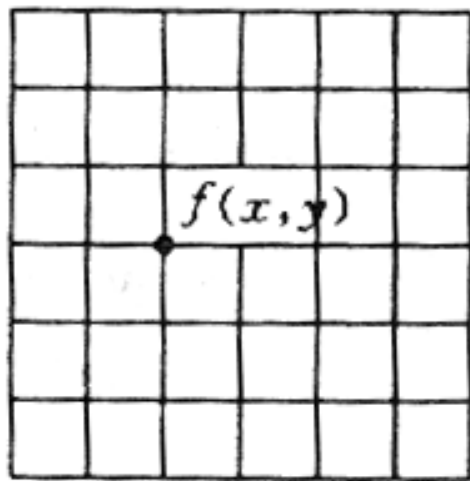
计算机学院



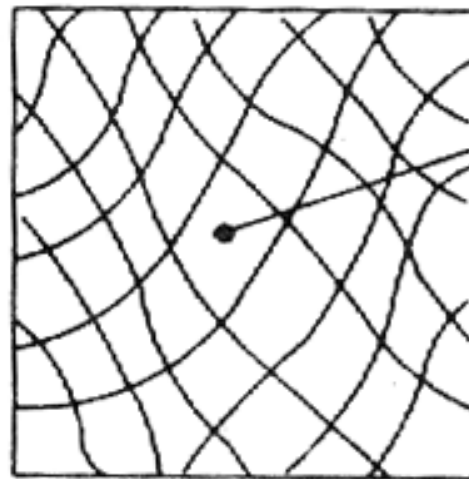
- 特征点的特征描述符应是不变量，以确保最低限度地受摄像机运动或光照变化等因素的影响。
- 匹配特征点时，根据特征向量的相似性来进行匹配，一般采用各种距离函数作为特征的相似性度量，如欧氏距离、马氏距离等。
- 变换：最小二乘、RANSAC



- 插值情况:** 在图像的几何变换中, 比如缩放和旋转, 输出图像上的像素点坐标有可能对应于原图像上几个像素点之间的位置。



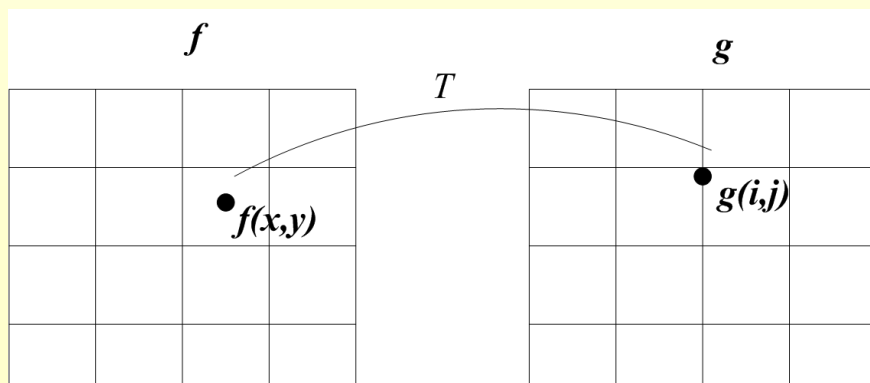
(a)



(b)



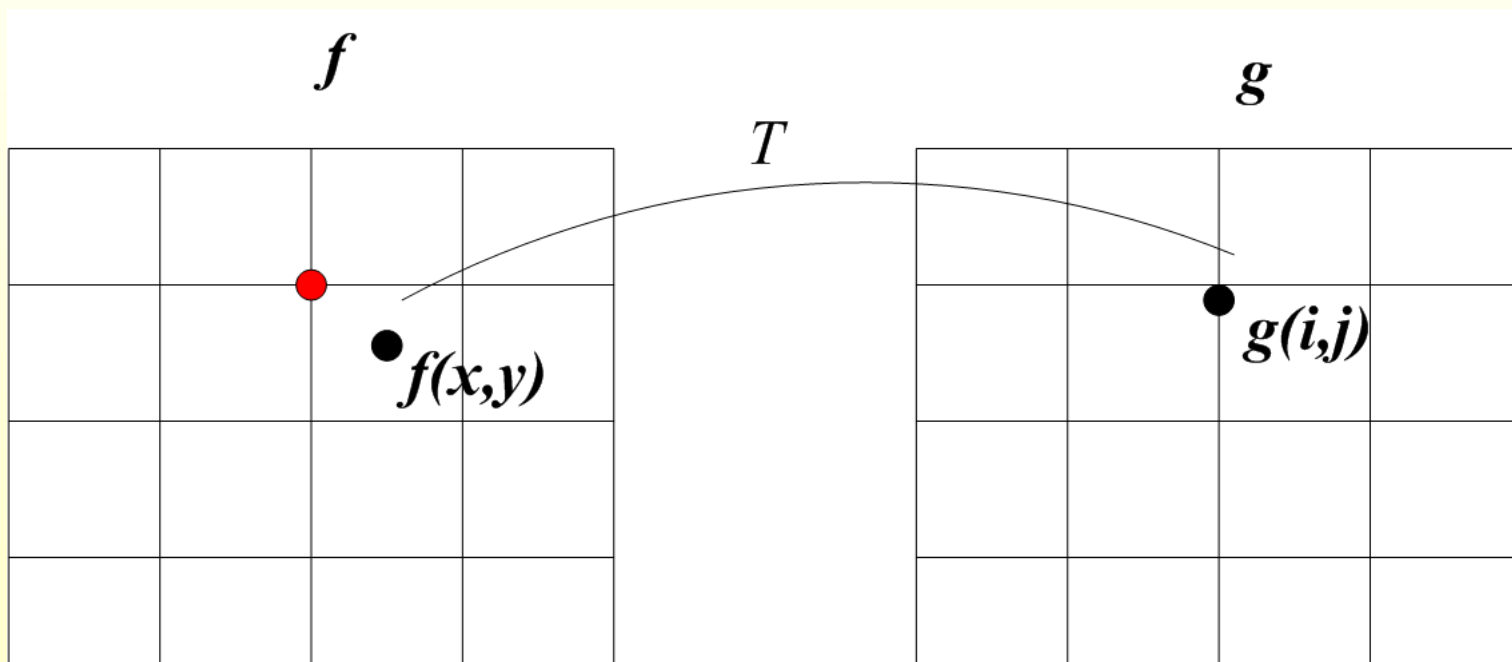
- 设图像 f 为参考图像，图像 g 为待匹配图像。采用 $T:(x,y) \rightarrow (i,j)$ 对图像 g 进行空间变换，得到点阵 g_T 。
- 假设变换 T 是可逆的，其逆变换为 T^{-1} 。对于 g_T 中的坐标点 (i,j) ，其原像为 g 中的坐标点 $(x,y)=T^{-1}(i,j)$ 。
- 由于 (x,y) 不一定是整数网格点，从而不能把 $f(x,y)$ 直接赋给 g_T 中坐标点 (i,j) ，作为灰度 $g_T(i,j)$ 。
- 重采样就是根据图像 f 中与 (x,y) 临近的像素点的灰度，利用逼近的方法得到 $g_T(i,j)$ ，从而得到最终的匹配图像。





- **最邻近法：**令变换后像素的灰度值等于距它最近的输入像素的灰度值

特点：简单快速，但是对于邻近像素点的灰度值有较大改变，细微结构是粗糙的

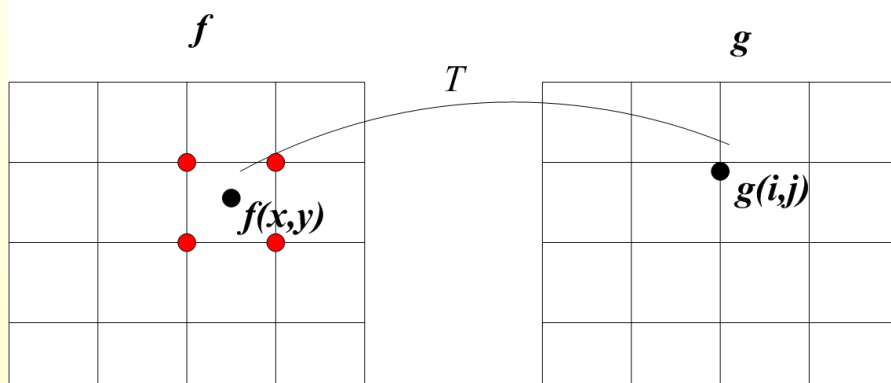




- **双线性法：**根据图像 g 中与 $T^{-1}(i,j)$ 最临近的4个像素点的灰度，利用线性逼近的方法，得到灰度 $g_T(i,j)$

$$g_T(i, j) = (1 - u)(1 - v)g(i_0, j_0) + u(1 - v)g(i_0 + 1, j_0) + (1 - u)v g(i_0, j_0 + 1) + uv g(i_0 + 1, j_0 + 1)$$

其中 $T^{-1}(i, j) = (x, y), \text{int}(x) = i_0, \text{int}(y) = j_0, u = x - i_0, v = y - j_0$



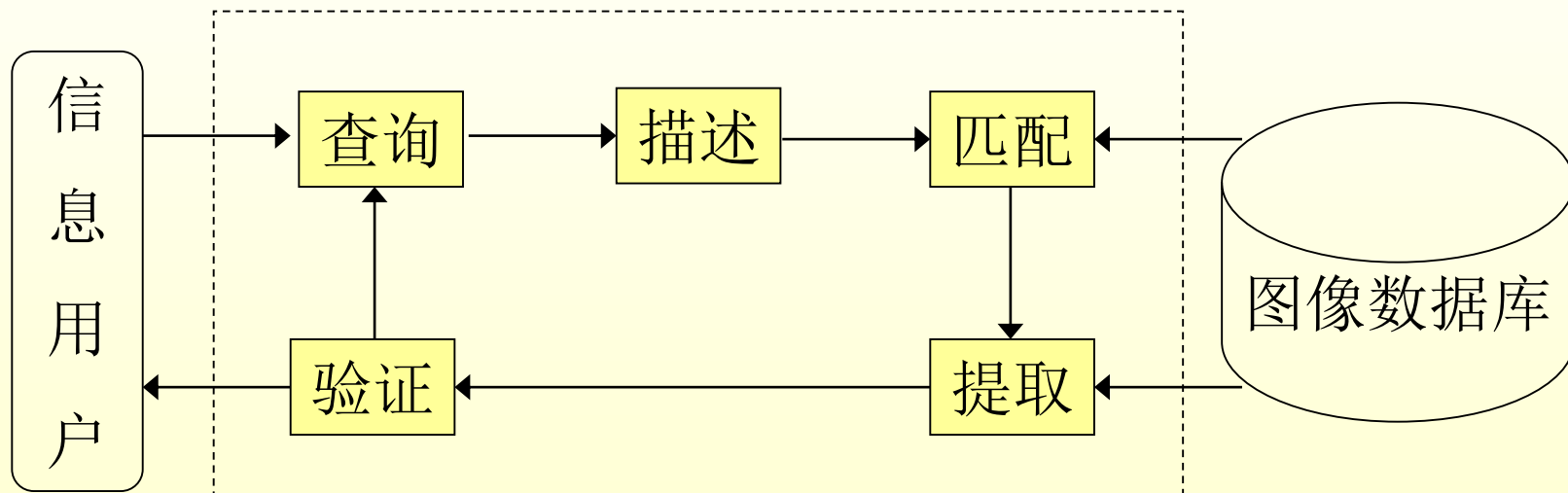
特点：具有低通滤波器特质，使高频信号受损，图像轮廓模糊，但同时边缘处的过度也比较自然。缺点是运算速度比最近邻插值要慢



- **特征内容匹配**
 - 一般框架和系统模块
 - 颜色匹配
 - 纹理匹配
 - 形状匹配
 - 综合特征匹配
 - 语义内容匹配

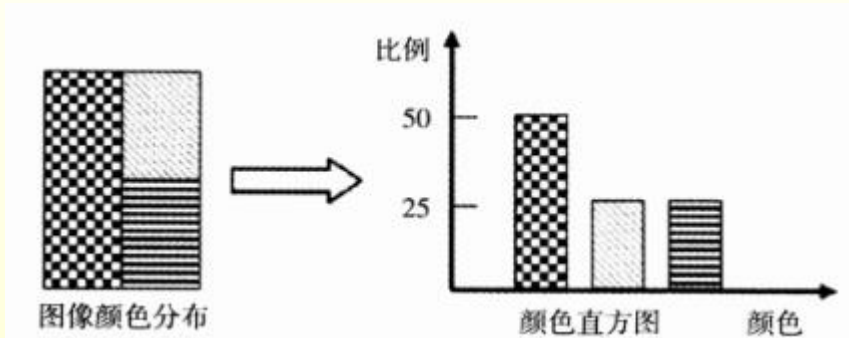



• 一般框架和系统模块



• 颜色匹配

- 颜色空间RGB,HSI,YUV.....
- 统计直方图 $H(k)=n_k/n \quad k=0,1,2,...,L-1$



- 直方图相交法  累积直方图
- 距离法：采用直方图的均值来粗略表达颜色信息
- 参考颜色表法：参考色的数量要比原图少，得到简化的直方图



颜色直方图的相似性



(a)



(b)



(c)



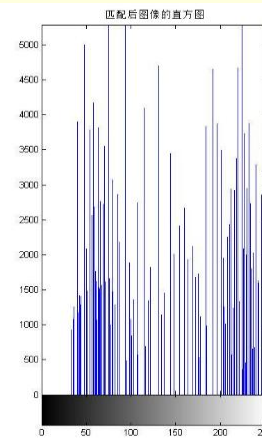
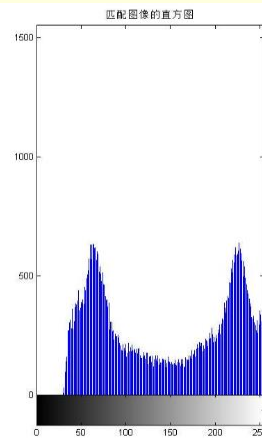
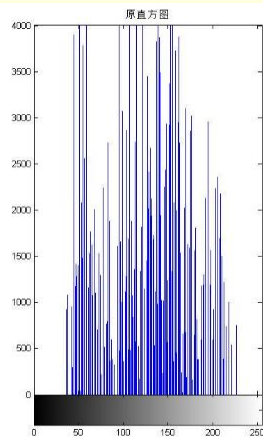
(d)

匹配图像

原图像

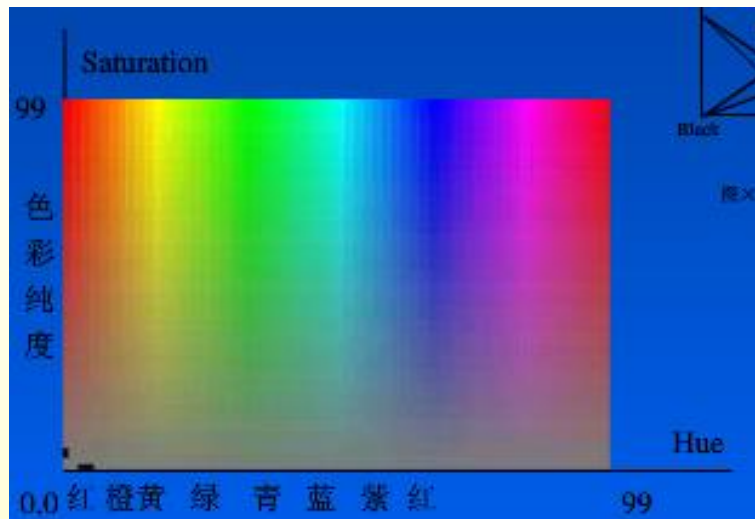


匹配后的图像





HSI/HSV颜色空间

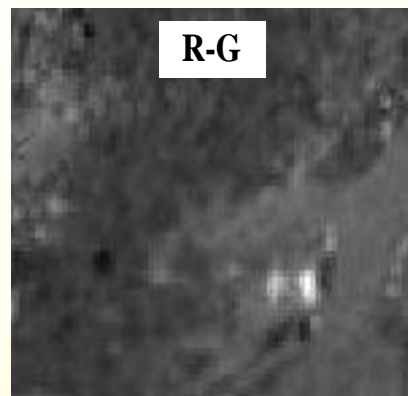




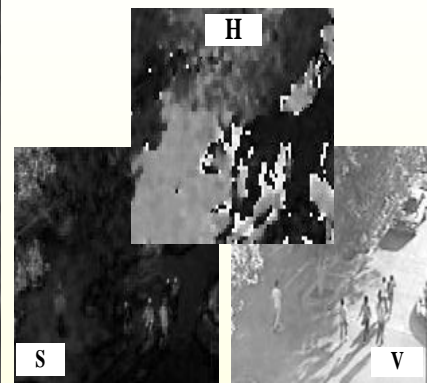
(a)原图



2G-B



R-G

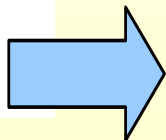


H

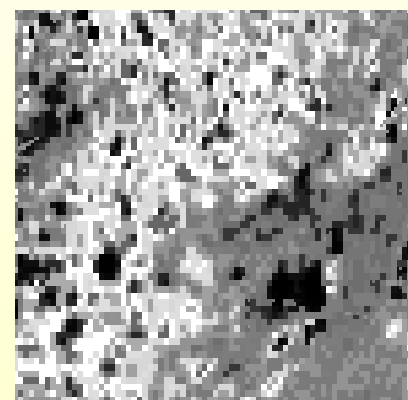
S

V

投影结果



(b)2G-B



(c)R-G



(d)HSV

不同颜色空间下目标置信度图效果比较



• 纹理匹配(Texture)

- 对比度、角度方向二阶矩、熵、平均值
- **统计法**：利用灰度直方图的矩来描述纹理，可分为**灰度差分统计法**和**行程长度统计法**
- **联合概率矩阵法（共生矩阵）**：一幅图象的灰度共生矩阵能反映出图象灰度关于方向、相邻间隔、变化幅度的综合信息，它是分析图象的局部模式和它们排列规则的基础。



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

频度矩阵

10	0	1	2
0	6	4	0
1	0	4	3
2	0	0	3

01	0	1	2
0	6	4	0
1	0	4	3
2	0	0	3

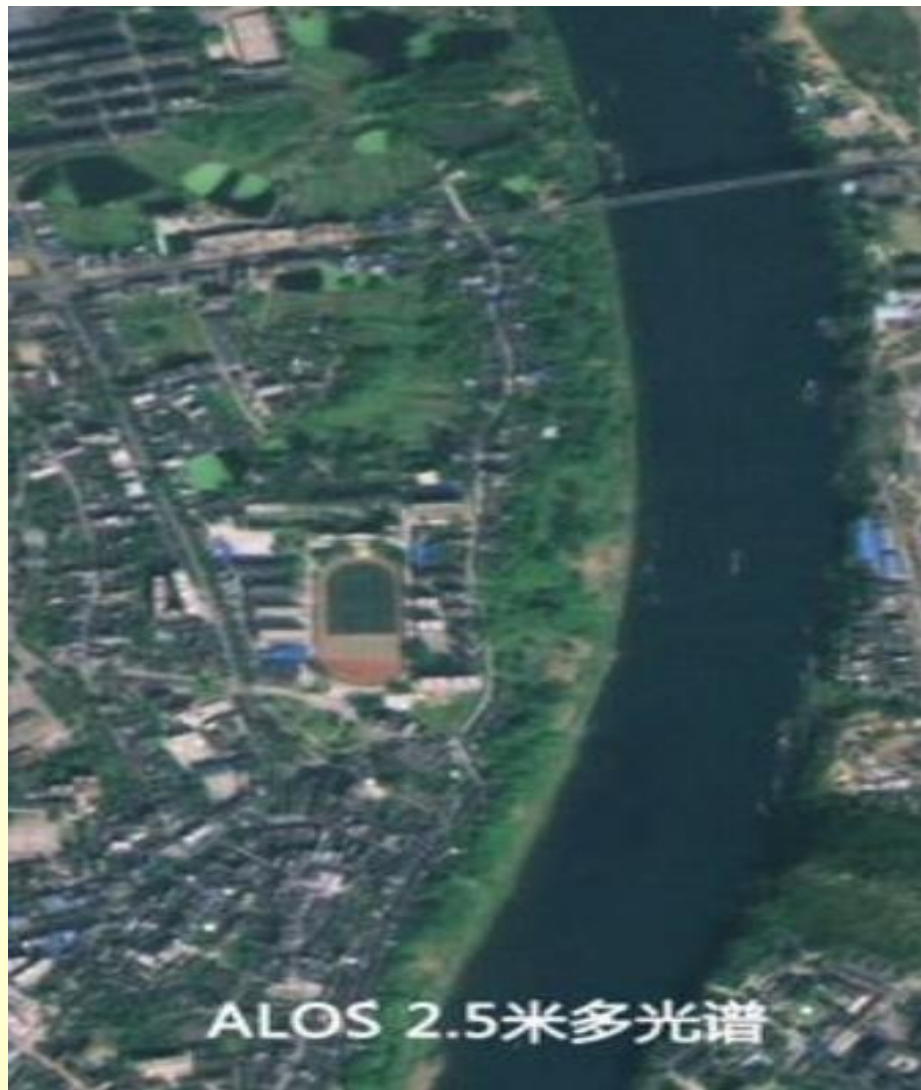
11	0	1	2
0	3	7	0
1	0	0	5
2	0	0	1

位置算子 δ_{10} , δ_{01} 的共生矩阵: $\frac{1}{20} \begin{bmatrix} 6 & 4 & 0 \\ 0 & 4 & 3 \\ 0 & 0 & 3 \end{bmatrix}$

δ_{11} 的共生矩阵 $\frac{1}{16} \begin{bmatrix} 3 & 7 & 0 \\ 0 & 0 & 5 \\ 0 & 0 & 1 \end{bmatrix}$



- 卫星多光谱图像：地物类型分析（森林、树林、草地、城区、小片灌溉区、大片灌溉区和水域，共7类）
- 将**纹理特征**（灰度共生矩阵）和**多光谱灰度特征**组合成 16×1 特征向量，对七类地域分别取训练样本和检验样本，提取特征利用分类器进行分类





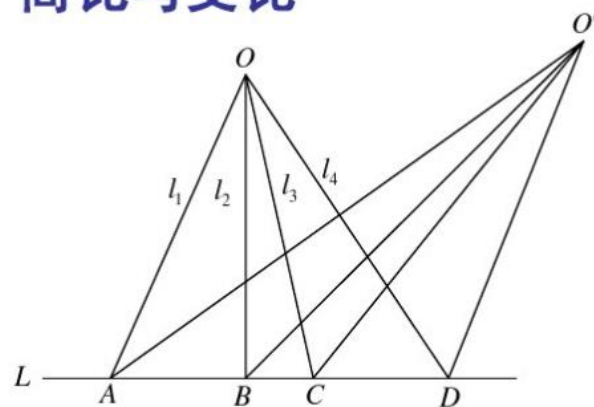
• 形状匹配

链码、游程码、freeman码等、样条(B样条,3次,5次样条)插值、多项式、多边形逼近、特征点检测.....

• 常用的形状描述子.

- 基于几何特征: 紧密度、实心度、偏心率、不规则度等;
- 基于统计特征: 粗糙度、均值、方差等;
- 变换域特征: 矩、Fourier 描绘子、小波描绘子、形态描绘子等;
- 仿射不变量: 简比等;
- 射影不变量: 交比等.

简比与交比



简比(simple ratio):

$$SR(A, B; C) = AC / BC$$

交比(cross ratio):

$$CR(A, B; C, D) = \frac{SR(A, B; C)}{SR(A, B; D)} = \frac{AC / BC}{AD / BD}$$



综合特征匹配

将颜色、纹理、形状以及空间位置等不同特征进行任意组合，从而实现综合特征的图像识别与理解。

• 语义内容匹配

— 关键问题：

- 语义特征具有“模糊”性，要解决特征的模糊化
- 语义特征与人的视觉理解紧密相关，要融入人的视觉感知，在“理解”图像的基础上检索图像



北京航空航天大学

Beijing University of Aeronautics and Astronautics

数字图像处理

图像的识别与解释



- 内容回顾
- 视觉理论与信息系统
- 运动目标检测跟踪



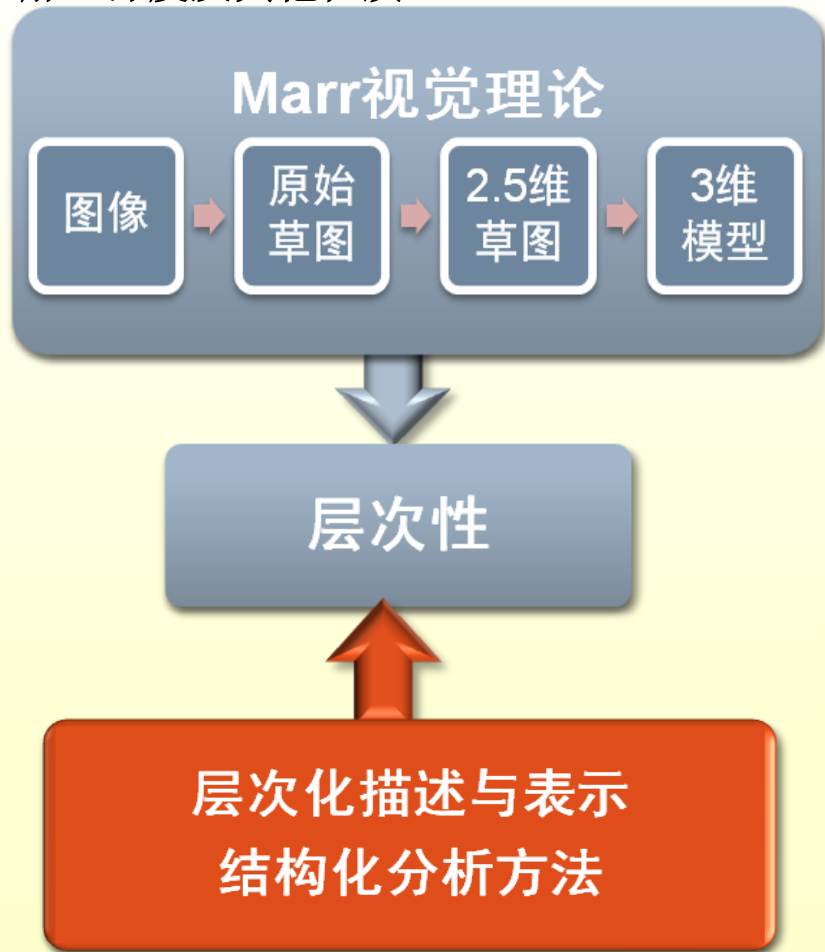
- **视觉信息从广义上包括**
 - 由人类视觉系统感知的信息
 - 由人类发明的各种视觉装置所获取的信息
 - 由这些信息推导出的其他表示形式
 - 从上述这些信息中抽象出来的高级表达和行为规划,以及与这些信息密切相关的知识和处理这些信息所需的经验



- 视觉信息系统是**采集、表达、处理、分析和理解**上述视觉信息的系统
- 研究内容：
 - 研究人类视觉系统是如何表达视觉信息的
 - 研究在视觉信息系统中应该如何表达和组织这些信息

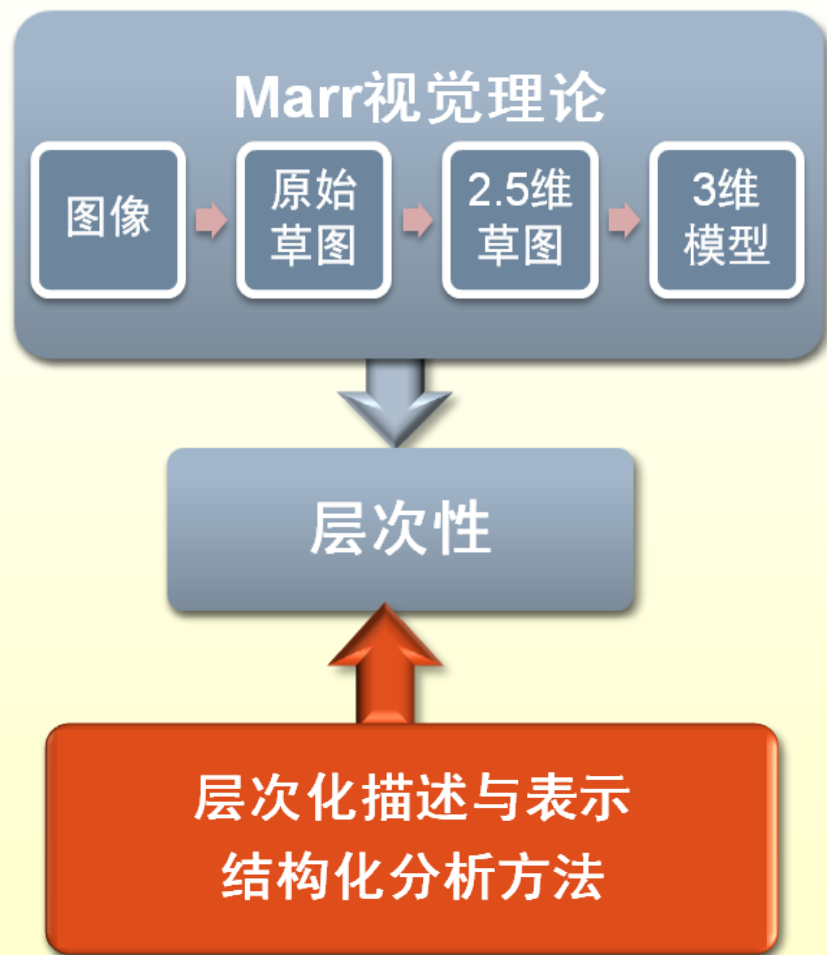


2.5维图(2.5 Dimensional sketch)——以观察者为中心，描述可见表面的方位、轮廓、深度及其他性质



Marr理论的要点:

- 视觉是一个复杂的信息加工过程，要完成视觉任务，需要在若干**不同层次和方面**进行，**人类认知可分为底层特征、几何模型、语义描述等三层，具有鲜明的层次性**；
- 视觉信息加工的三个要素：**计算理论+算法实现+硬件实现**
- 视觉信息的**三级内部表达**
- 视觉信息处理是按照功能模块的形式**组织**起来的
- 计算理论形式化表示必须考虑**约束条件**

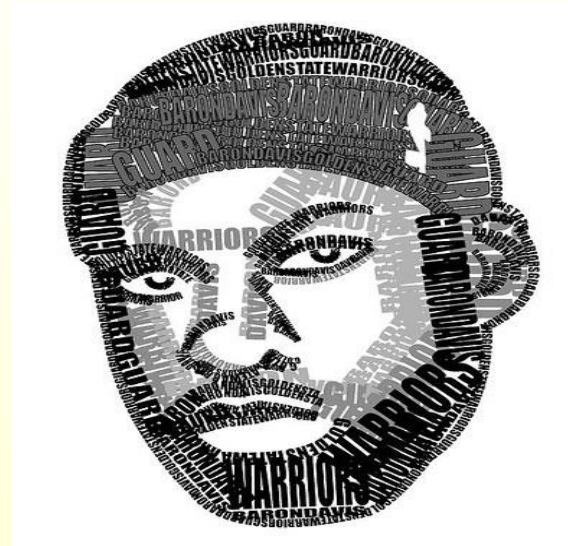
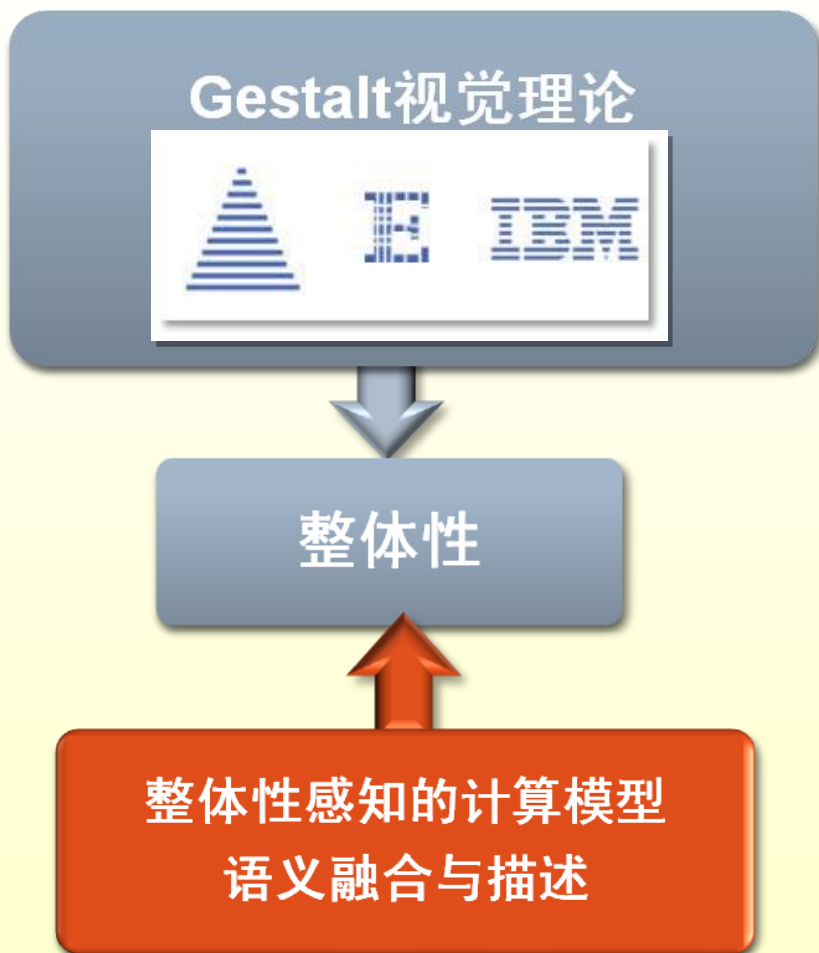


Marr理论的不足:

- 框架中输入是被动的，给什么图像，系统就处理什么图像；
- 框架中加工目的不变，总是恢复场景中物体的位置和形状等；
- 框架缺乏或者未足够重视高层知识的指导作用；
- 整个框架中信息加工过程基本自下而上，单向流动，没有反馈。



Gestalt心理学理论在宏观上强调知觉组织的整体化效应。



神经生物学发现在视觉信息处理通路上存在自上而下的调控，以此实现对初级皮层信息处理过程的调整与结果的整合。



视觉计算理论

Marr视觉理论



层次性

层次化描述与表示
结构化分析方法

Gestalt视觉理论



整体性

整体性感知的计算模型
语义融合与描述



- 一方面，在媒体计算模型的研究中注意吸收利用底层、局部性认知机理方面的成果；
- 另一方面，通过媒体计算模型研究形成有关高层、整体性认知的假说，在认知科学的探索中进行验证。
- 形成认知机理与计算模型研究的双向支持、相互借鉴，促进符合媒体理解层次性和整体性的理论框架的构建。



多层次特征表示

易于提取
树干的轮廓



树干轮廓的最佳尺度

不易于提取
树干的轮廓



树干轮廓的次优尺度



树干轮廓的临界尺度

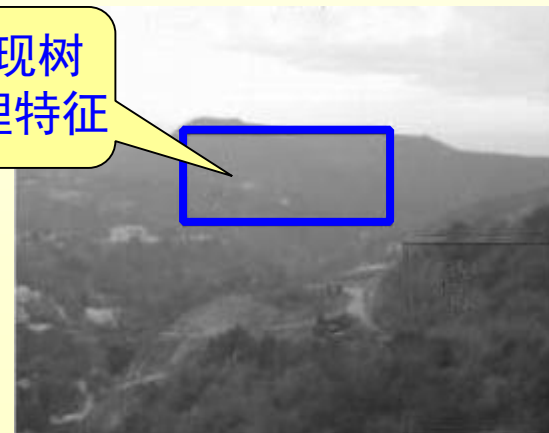


树林纹理的最佳尺度



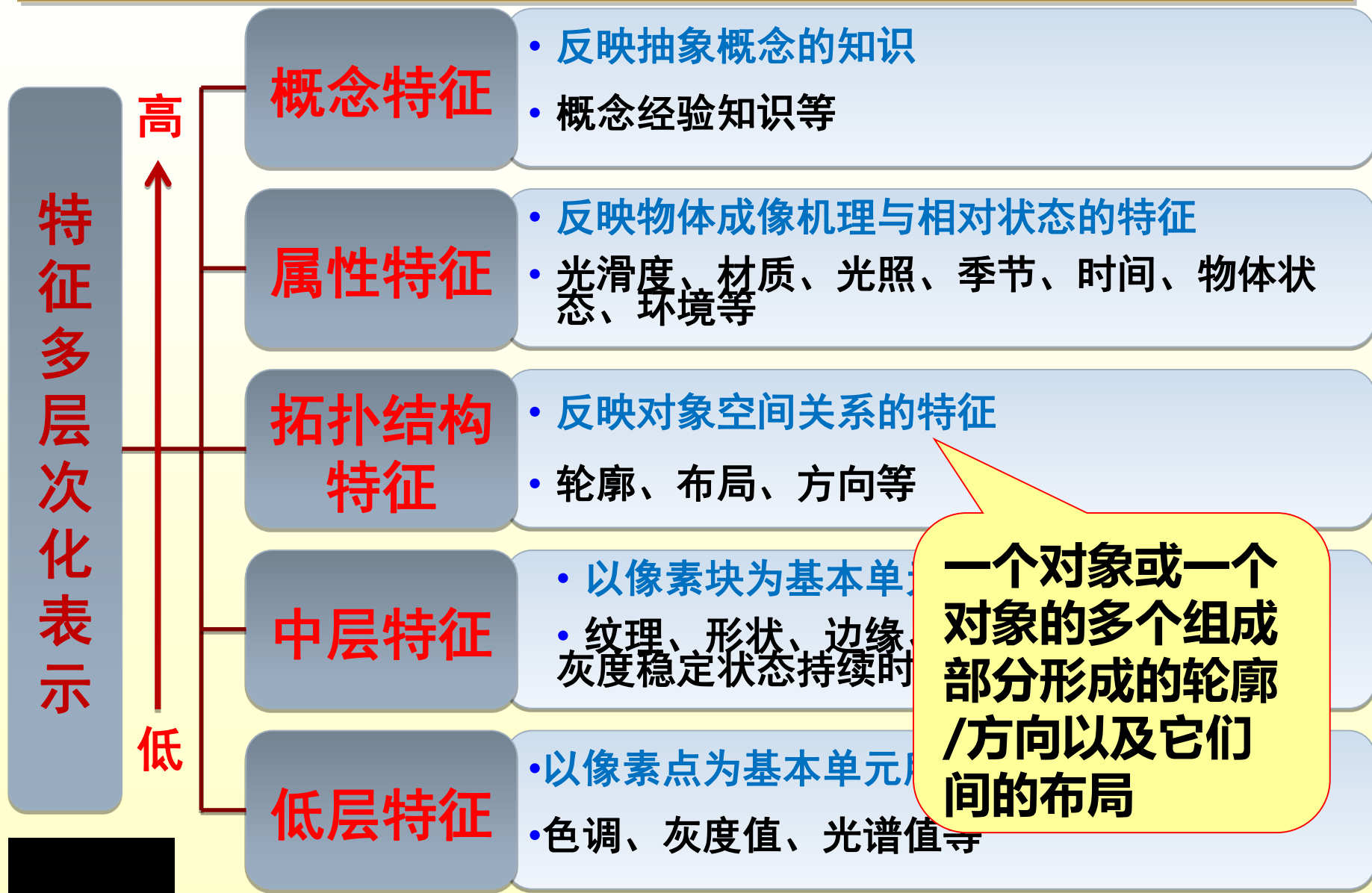
树林纹理的次优尺度

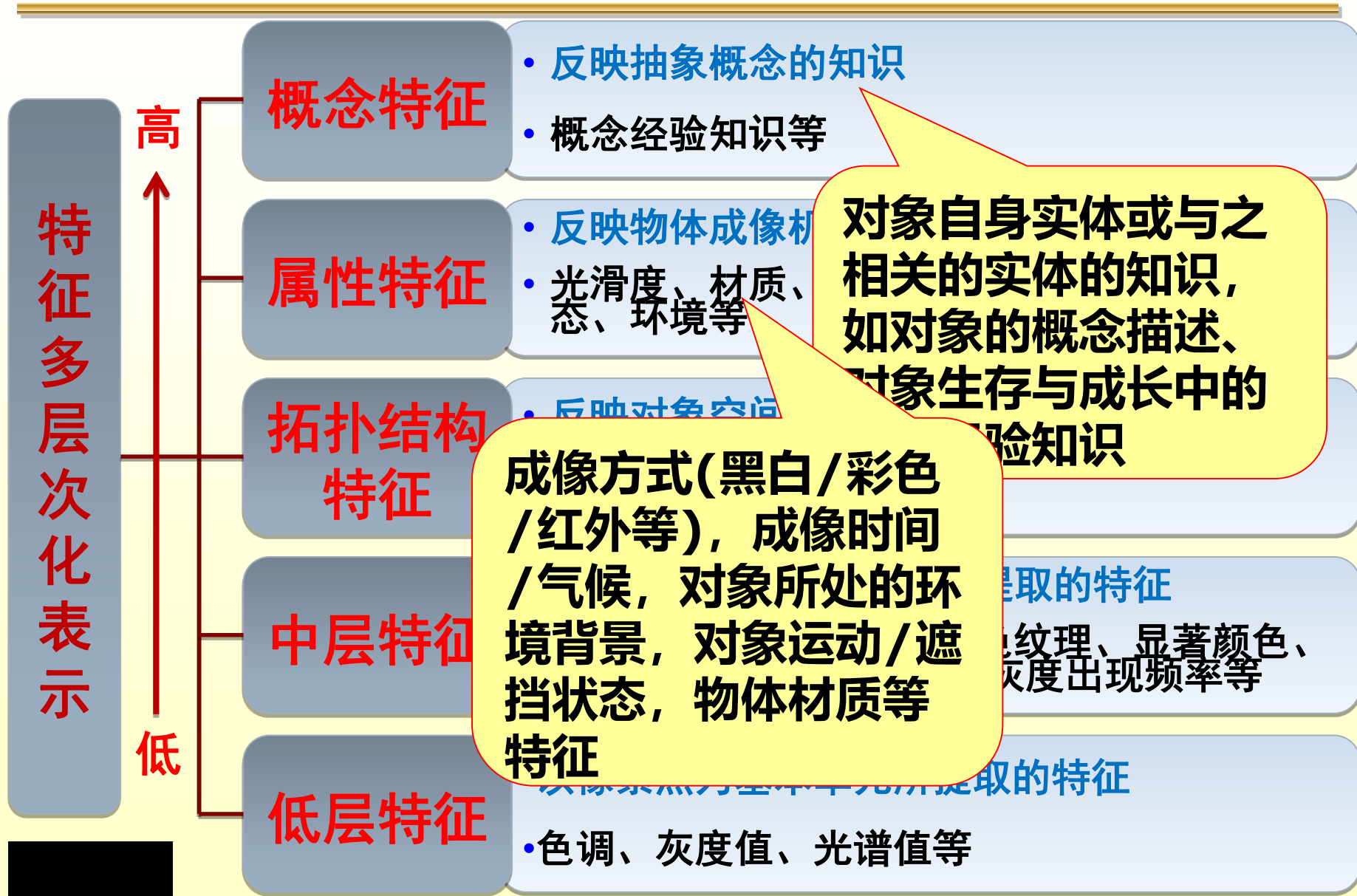
无法展现树
林的纹理特征

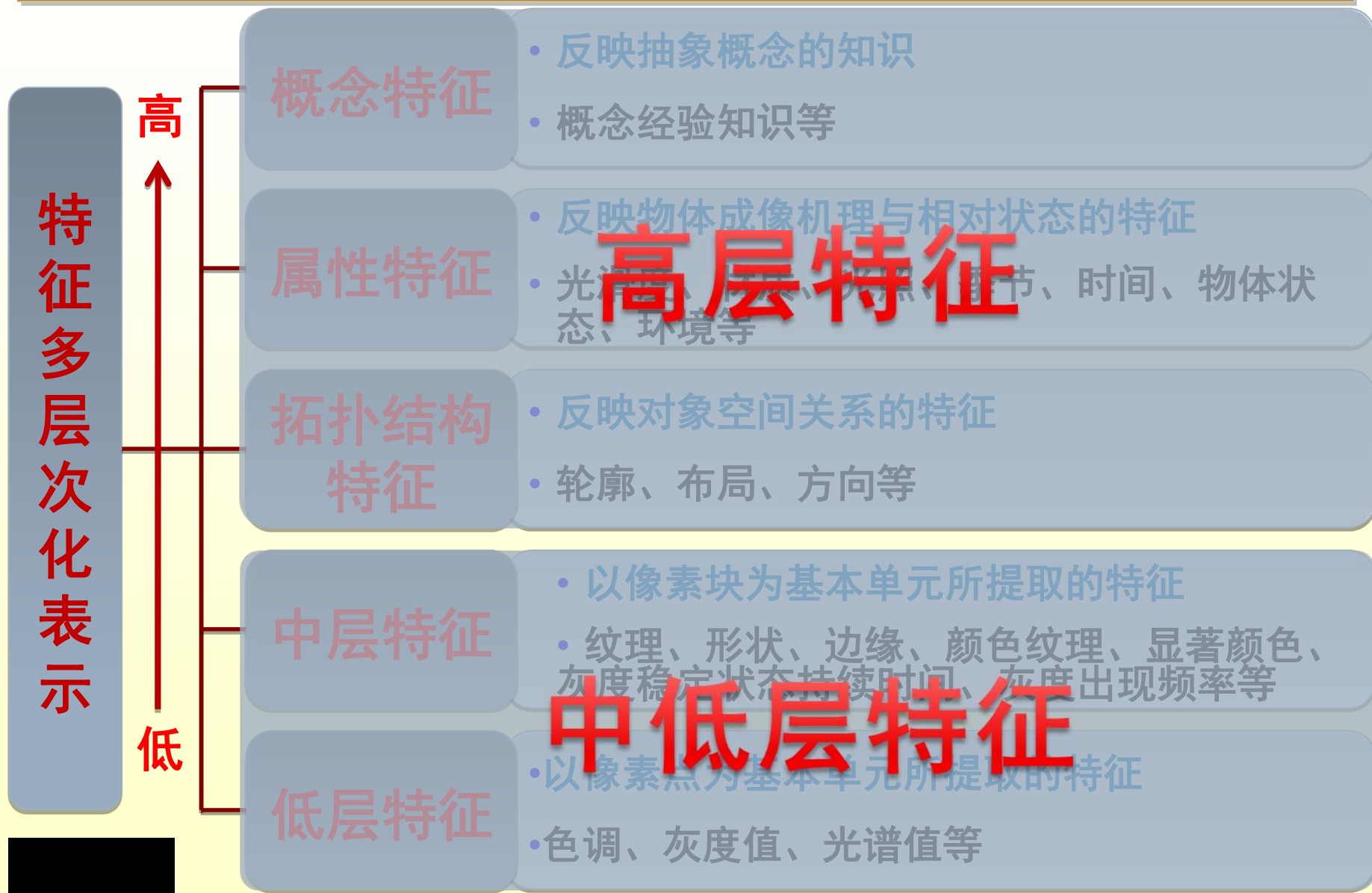


树林纹理的不可见尺度









■ 概念特征/环境属性特征在理解图像对象中的重要性

色调、纹理、
形状相似，同
一物体？

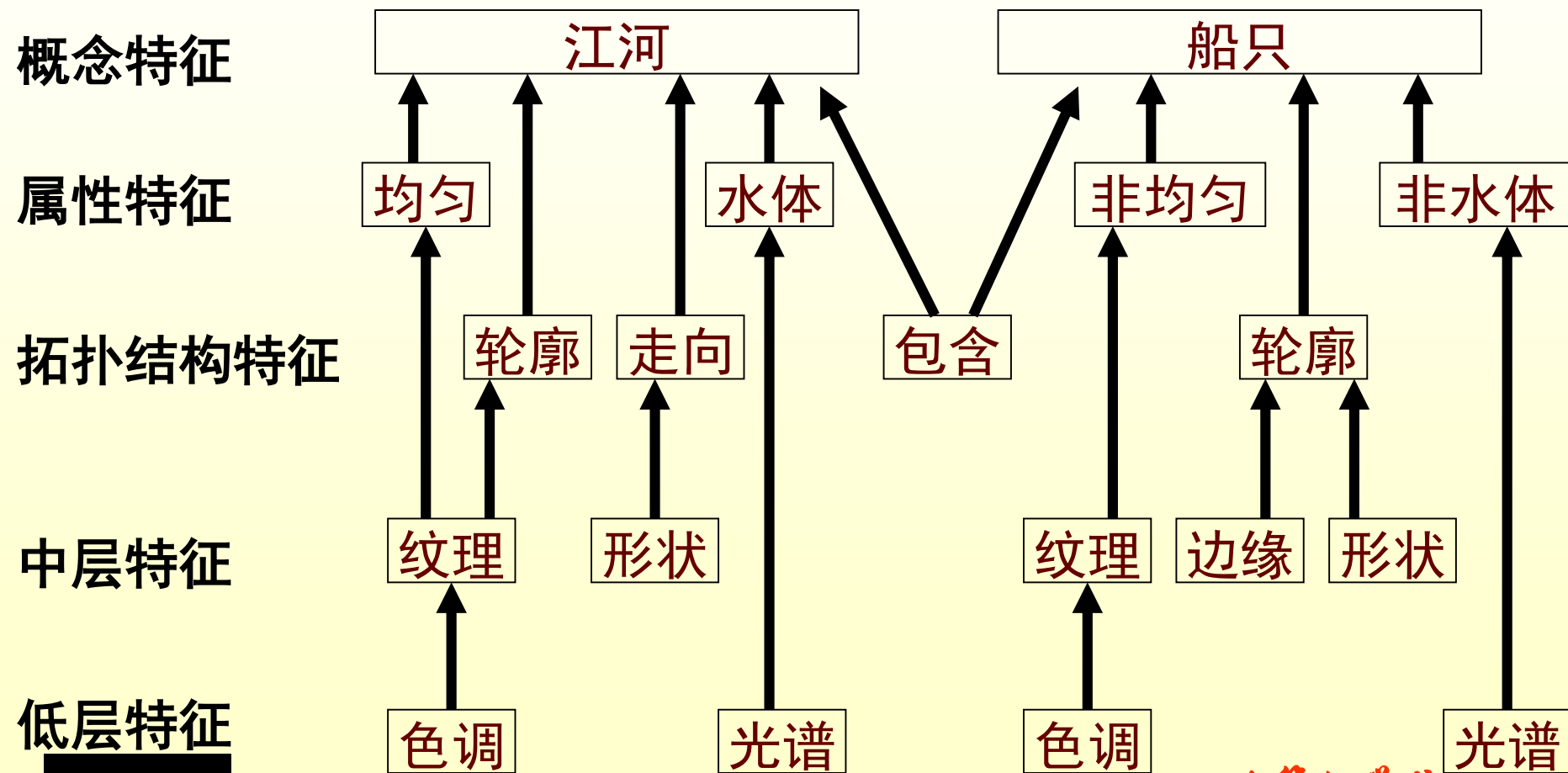
道路上的车

水中的船



• 自底向上逐层抽象

– 利用低层特征获取高层特征





• 自顶向下指导与调控

- 利用高层特征选取部分特征

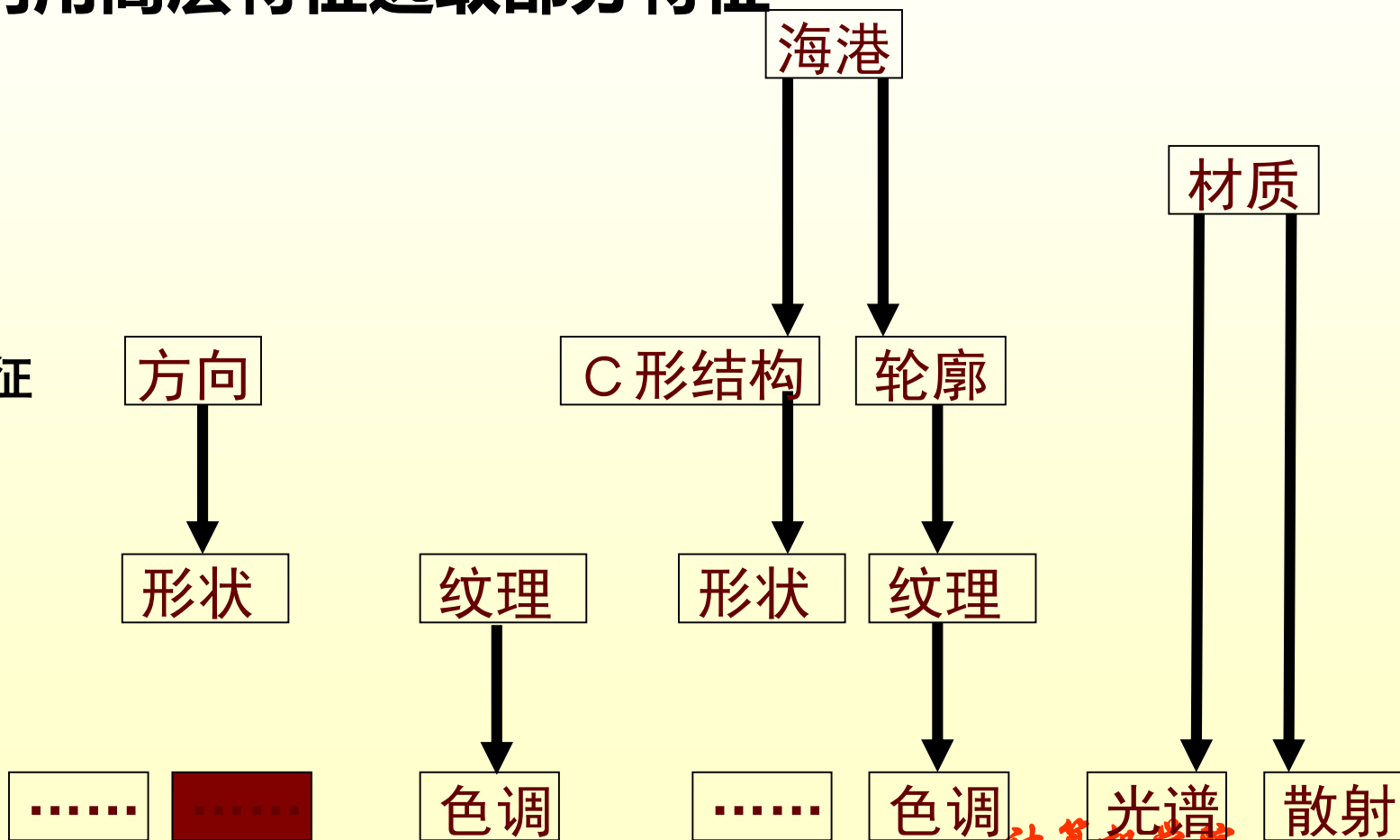
概念特征

属性特征

拓扑结构特征

中层特征

低层特征



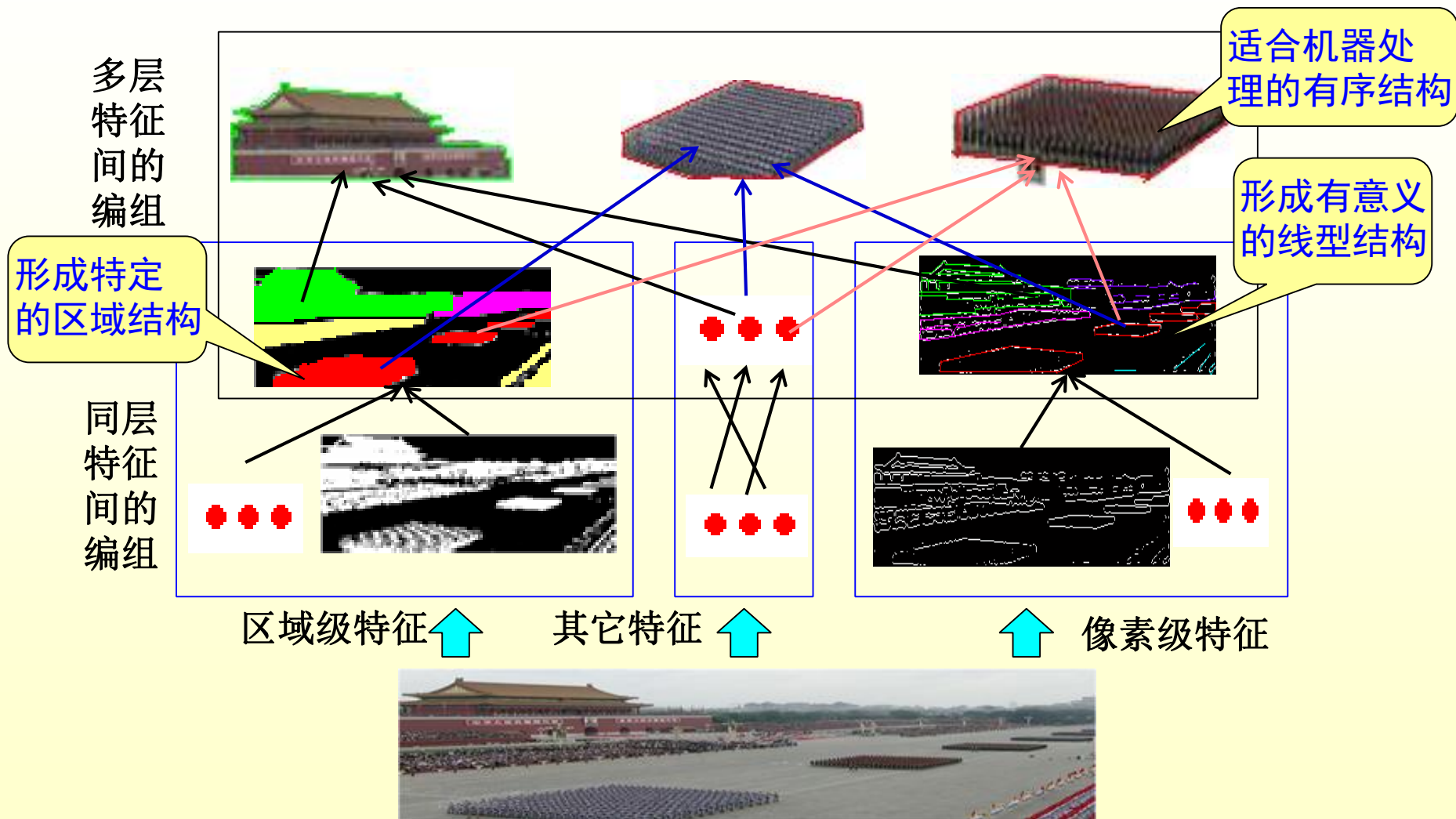


■ 特征的层次化编组

- 从媒体中提取的特征，尤其是低层的像素级特征，是离散的、无序的、非结构化的，通过编组将其组织为有序的、蕴含一定高层语义的结构。



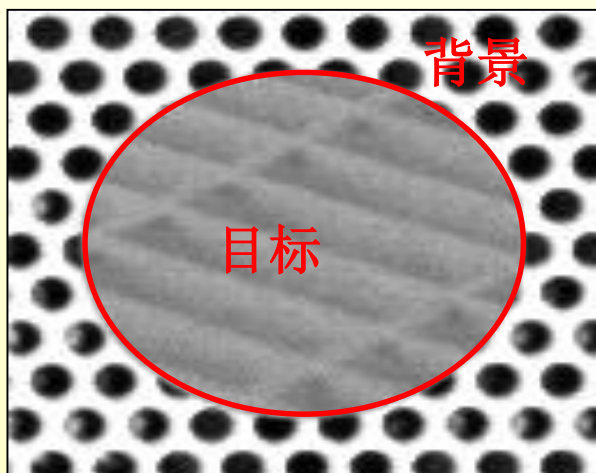
多层次特征表示



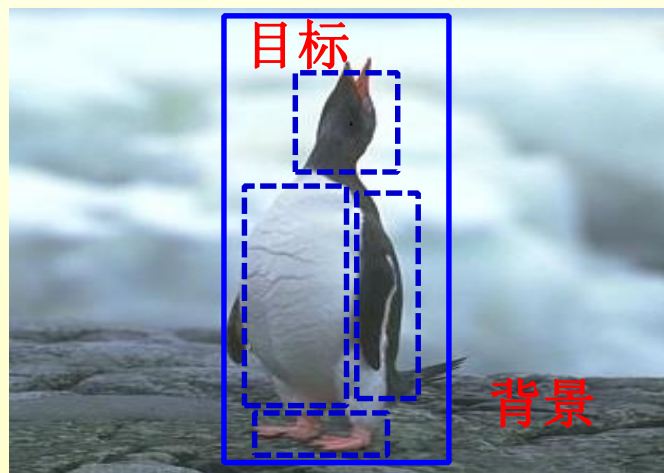


• 轮廓的提取

- 轮廓是目标的闭合空间边界
- 轮廓提取准则涉及两个层面
 - 低层特征的一致性，即目标组成部分“内”的纹理、颜色的一致性
 - 需要先验知识和认知的支持，即目标各组成部分属于一个统一的整体



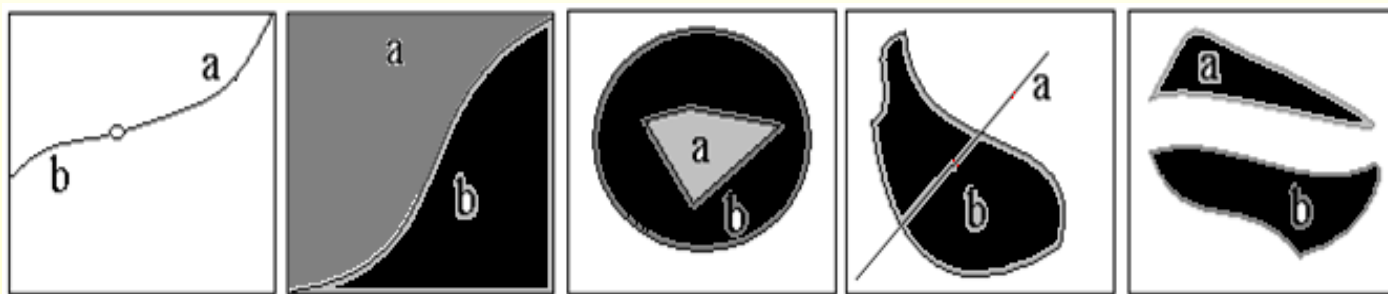
从低层特征看：纹理
与颜色具有一致性



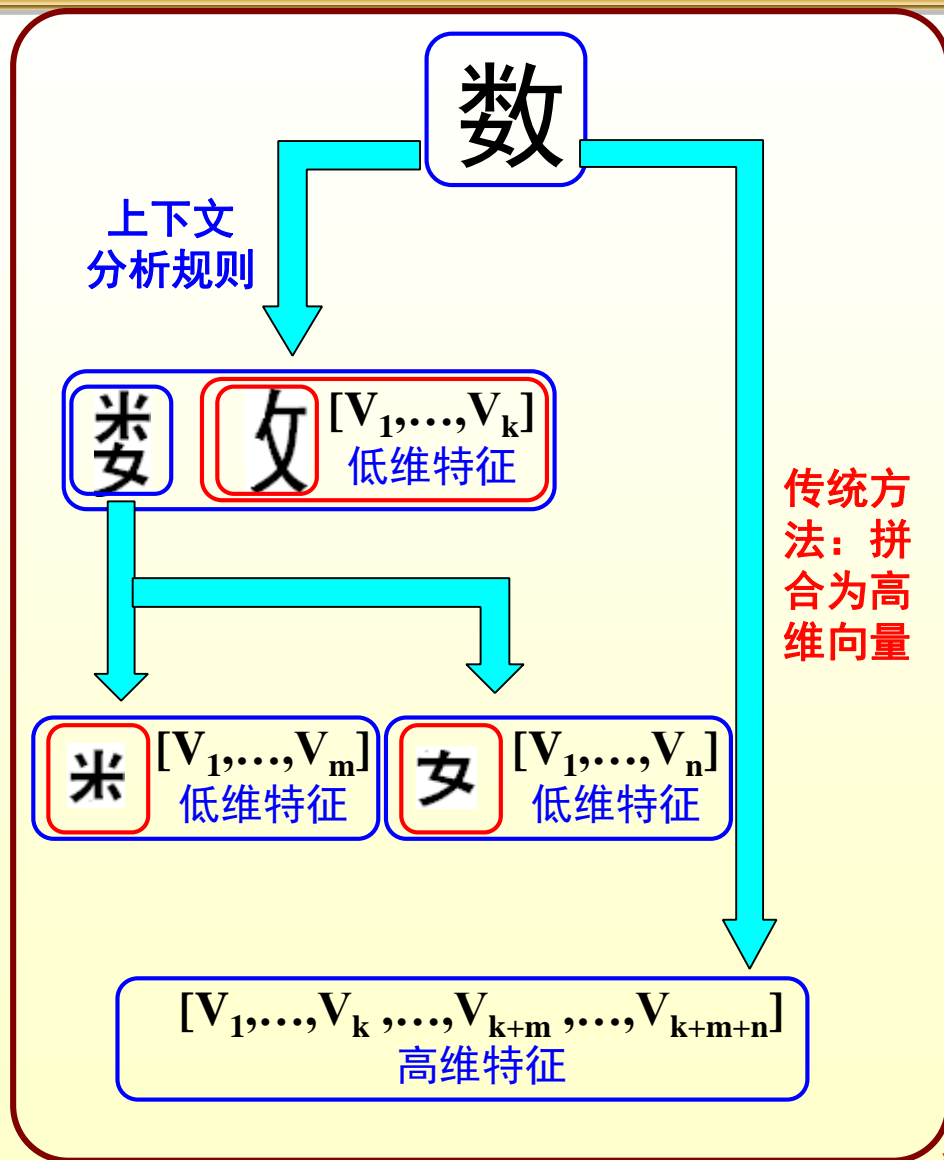
由经验知识得出：企鹅的头、腹、脚的
低层特征不同，但认识上是统一的整体

• 空间结构的提取

— 空间结构分为连接、相邻、包含、穿越和相离等类型



(1) 连接 (2) 相邻 (3) 包含 (4) 穿越 (5) 相离





- 现有特征间的关系及映射采用典型特征关系分析方法：
 - 将获取的特征组成一个高维特征向量或特征矩阵, 通过机器学习的方法分析选择特征
- 特征间关系不清晰, 不利于特征的分析、提取和利
用的效率
- 不同特征之间的差异性
- 同一特征的显著性差异

局部特征

- 信息量最大子块：具有最大信息量的区域，可能是对人类视觉刺激最强烈的显著区域，也可能是感兴趣区域，还有可能是会场自由讨论时的一段声音.....
- 信息量最大子块的提取和应用相关，和上下文相关



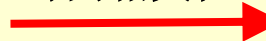
空间相关性



整体级特征：显著颜色（和背景差异大的目标颜色）



时间相关性



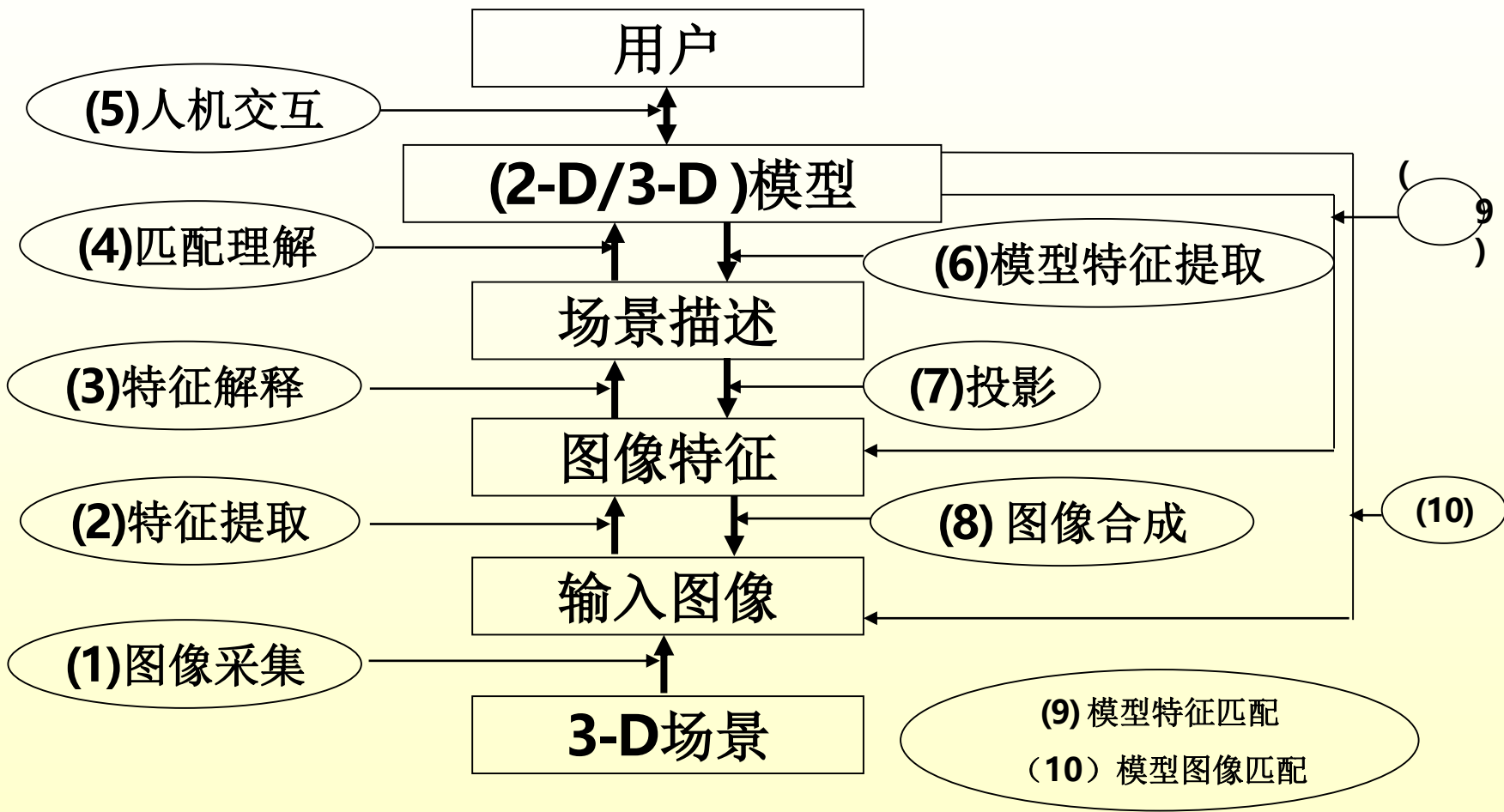
下半身形变大



属性级特征：非刚体

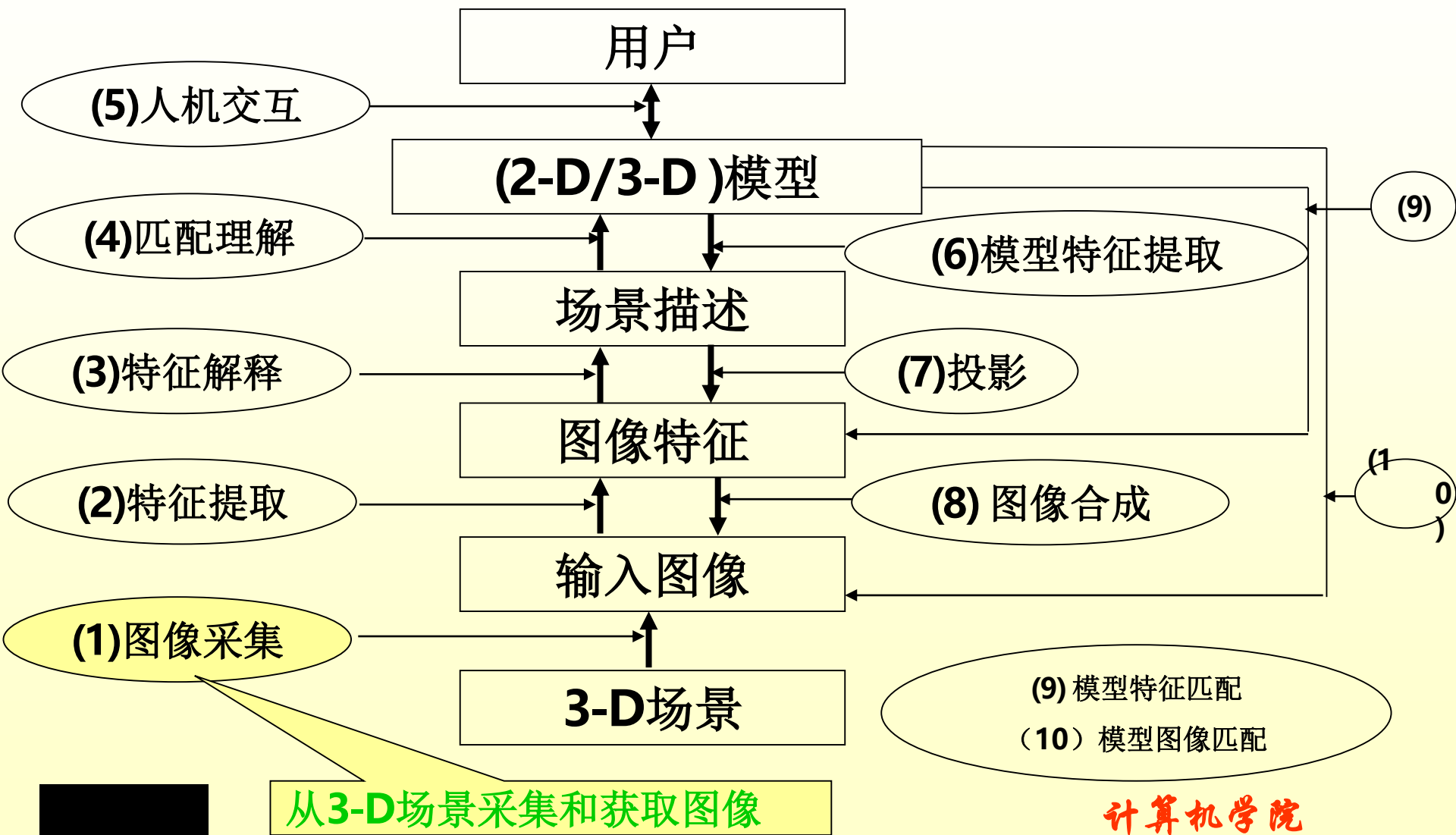


视觉信息系统模型



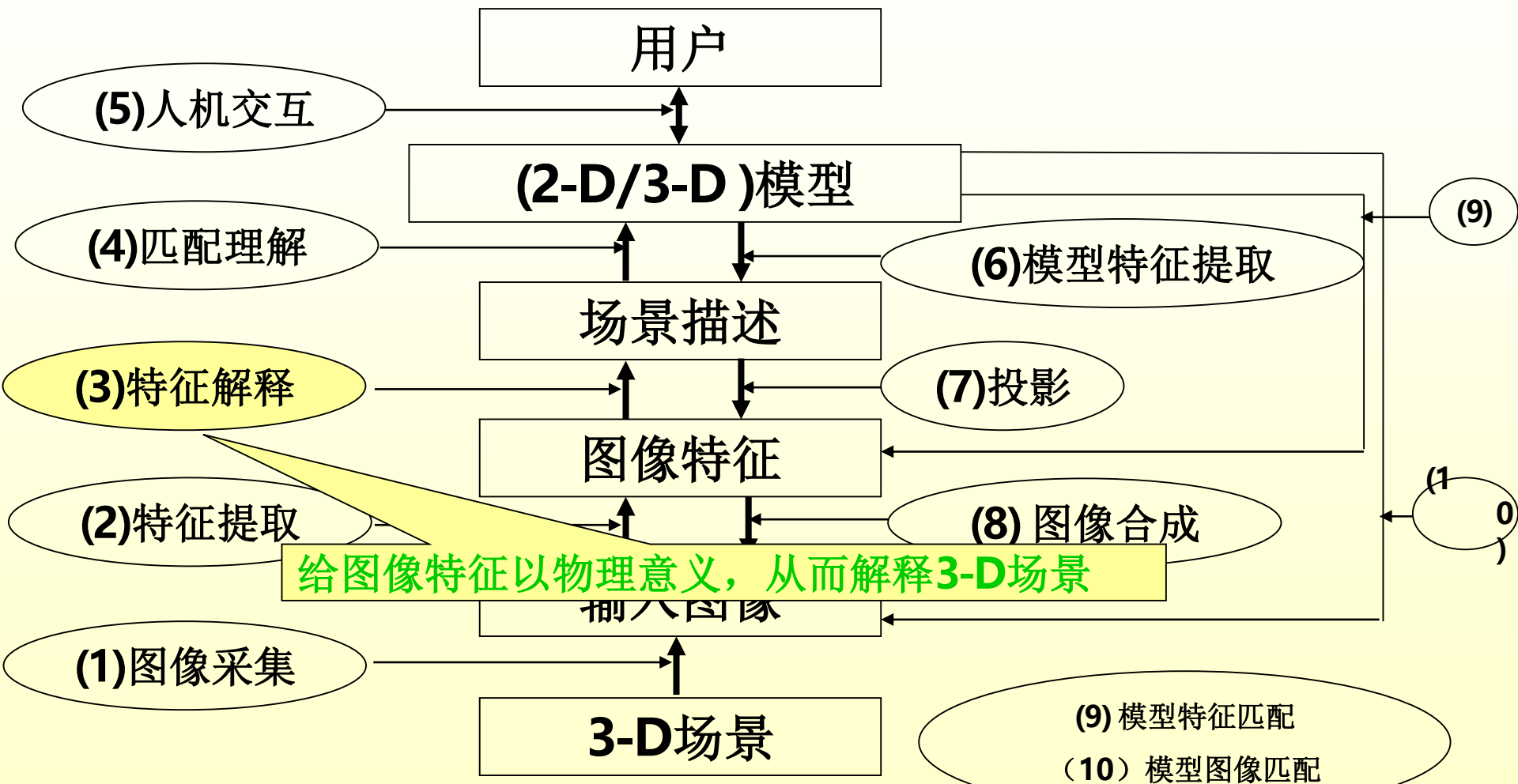


视觉信息系统模型



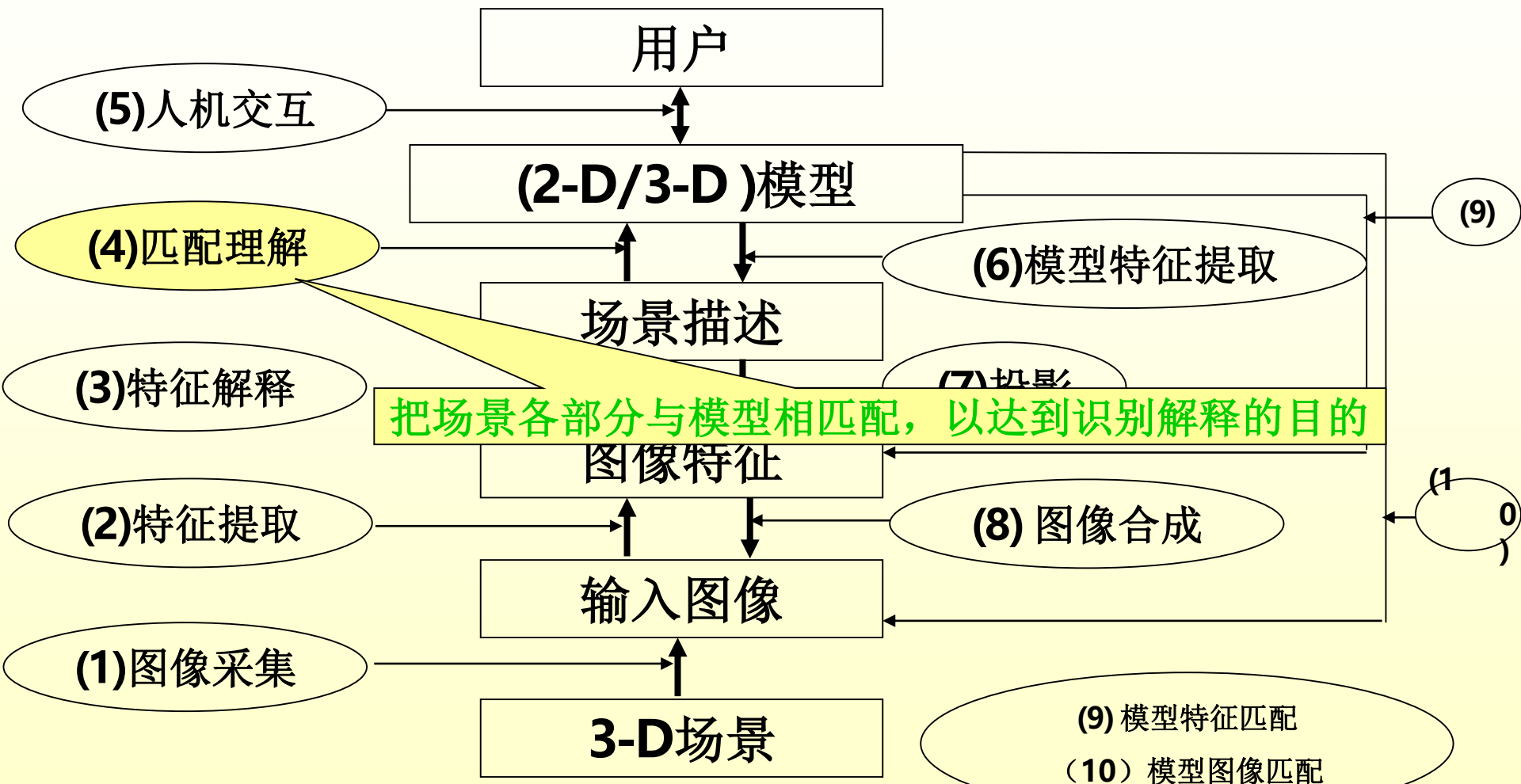


视觉信息系统模型



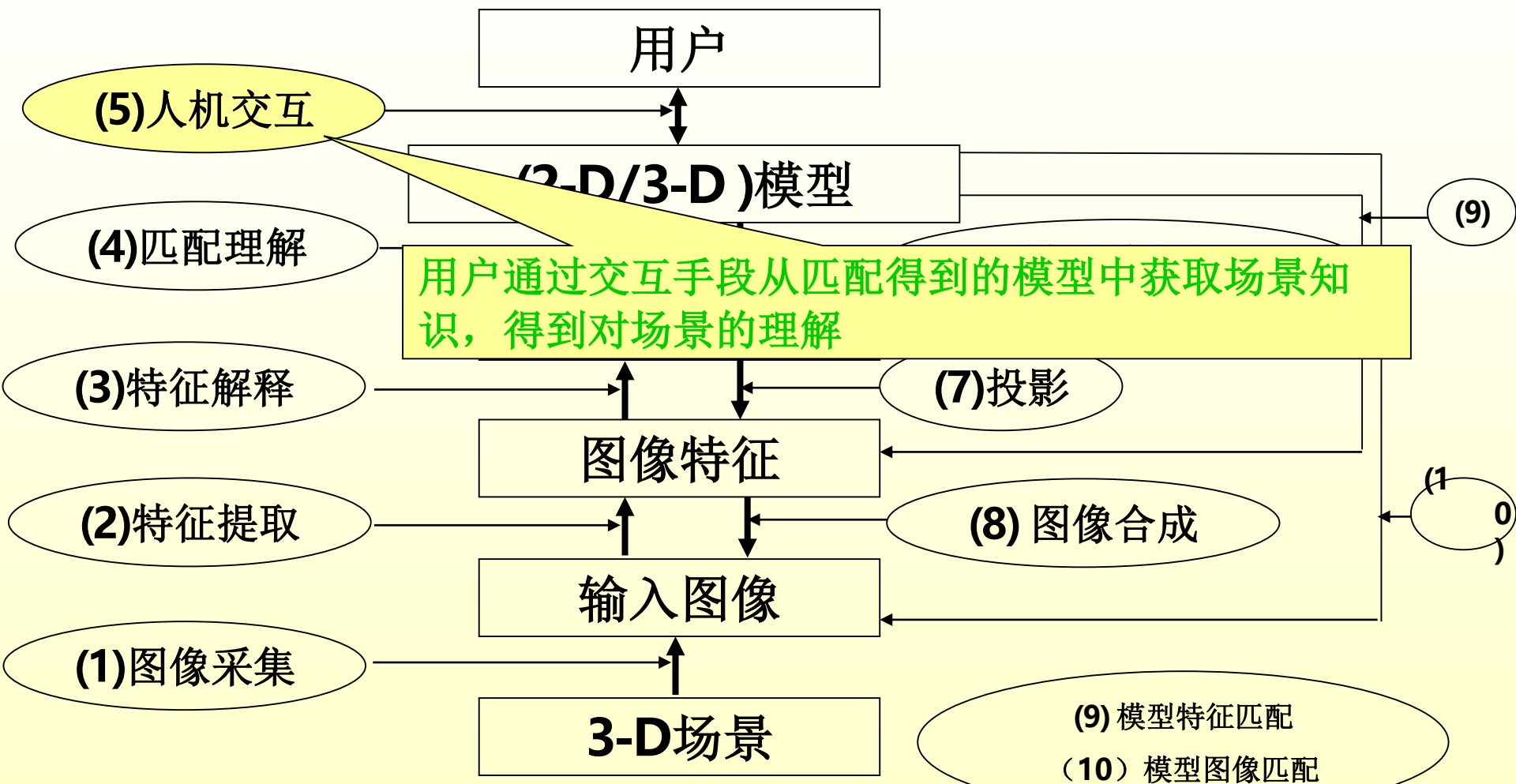


视觉信息系统模型





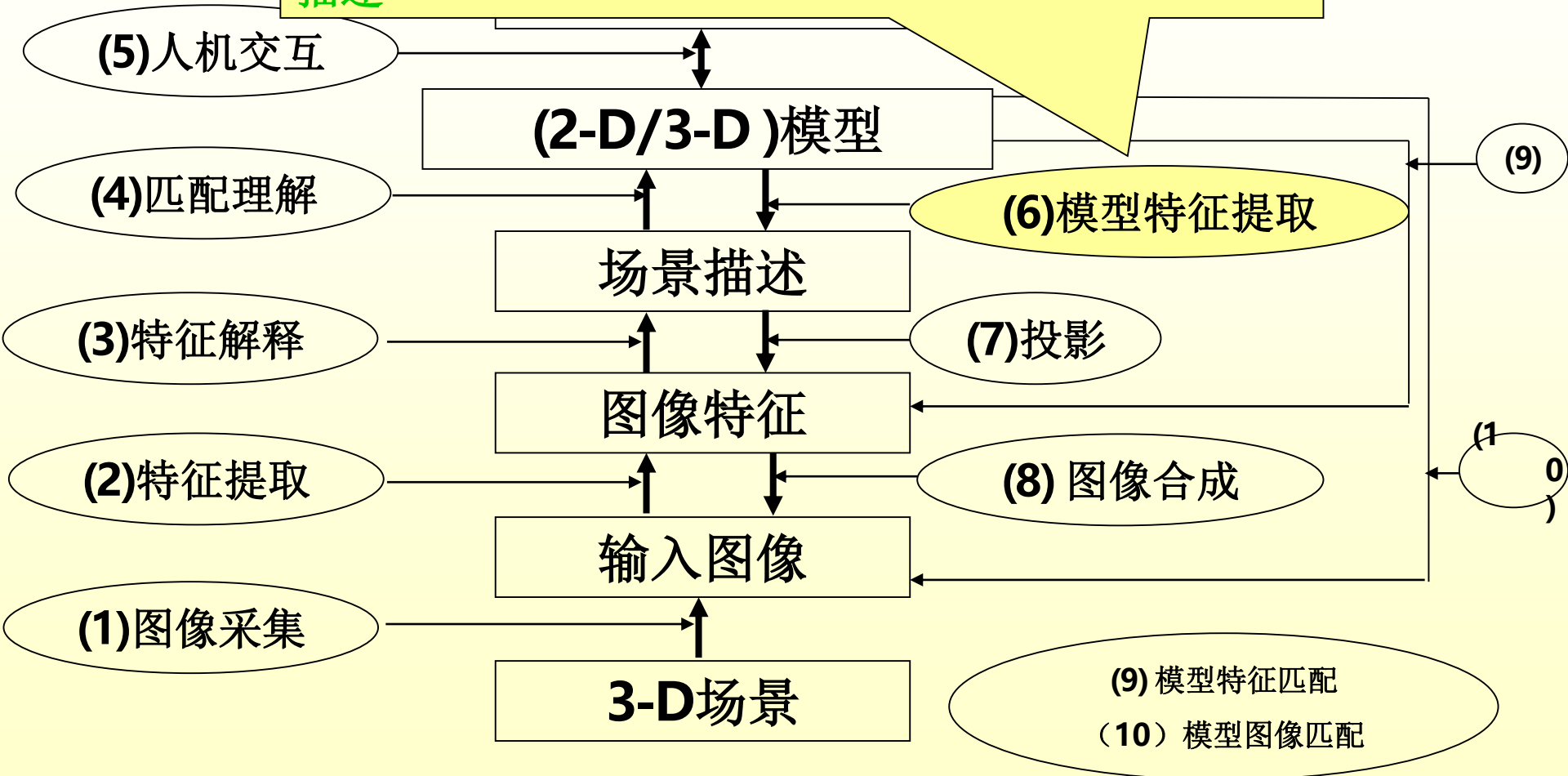
视觉信息系统模型





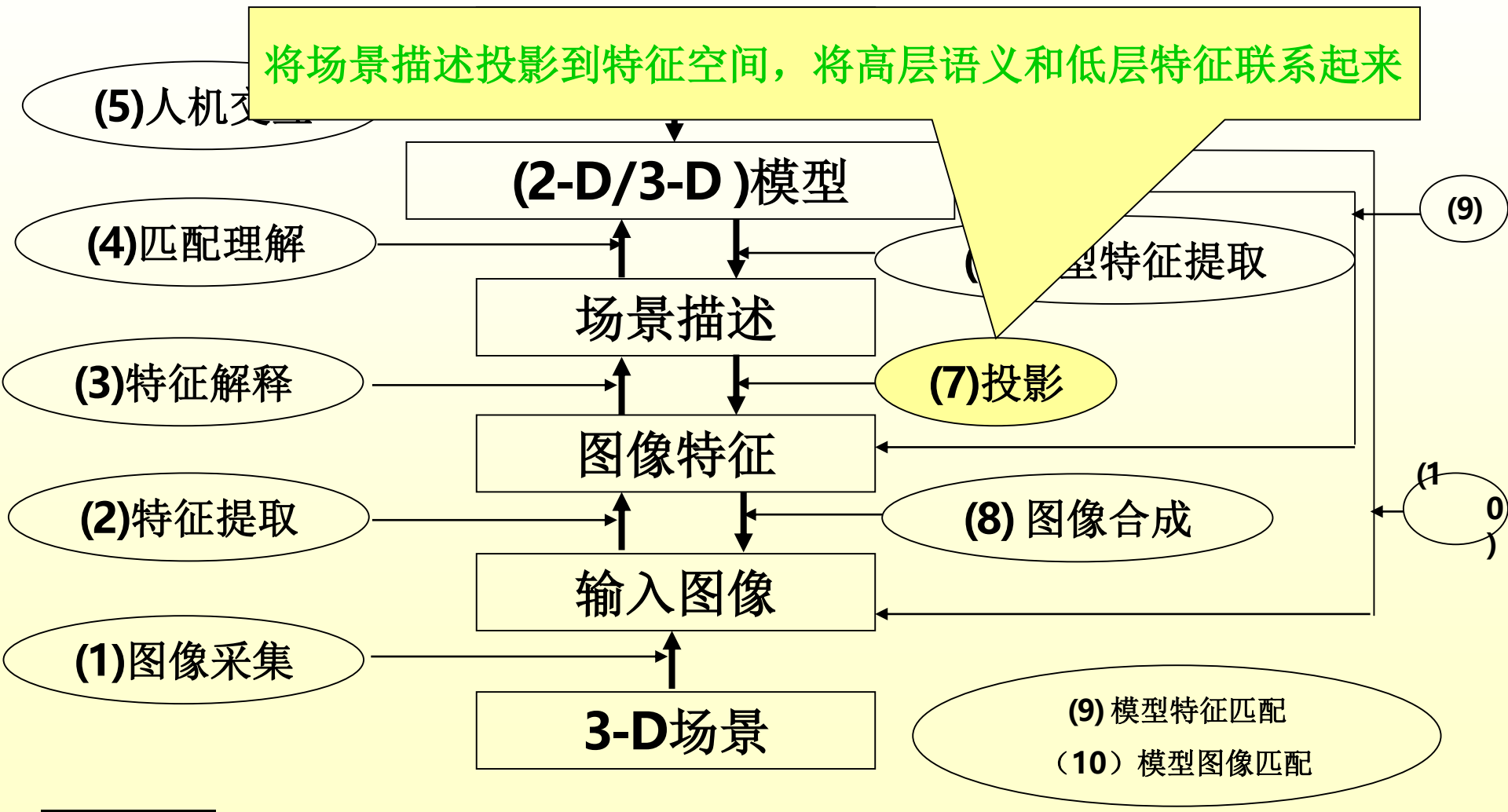
视觉信息系统模型

如果从模型中将典型的特征提出来就有可能恢复场景描述



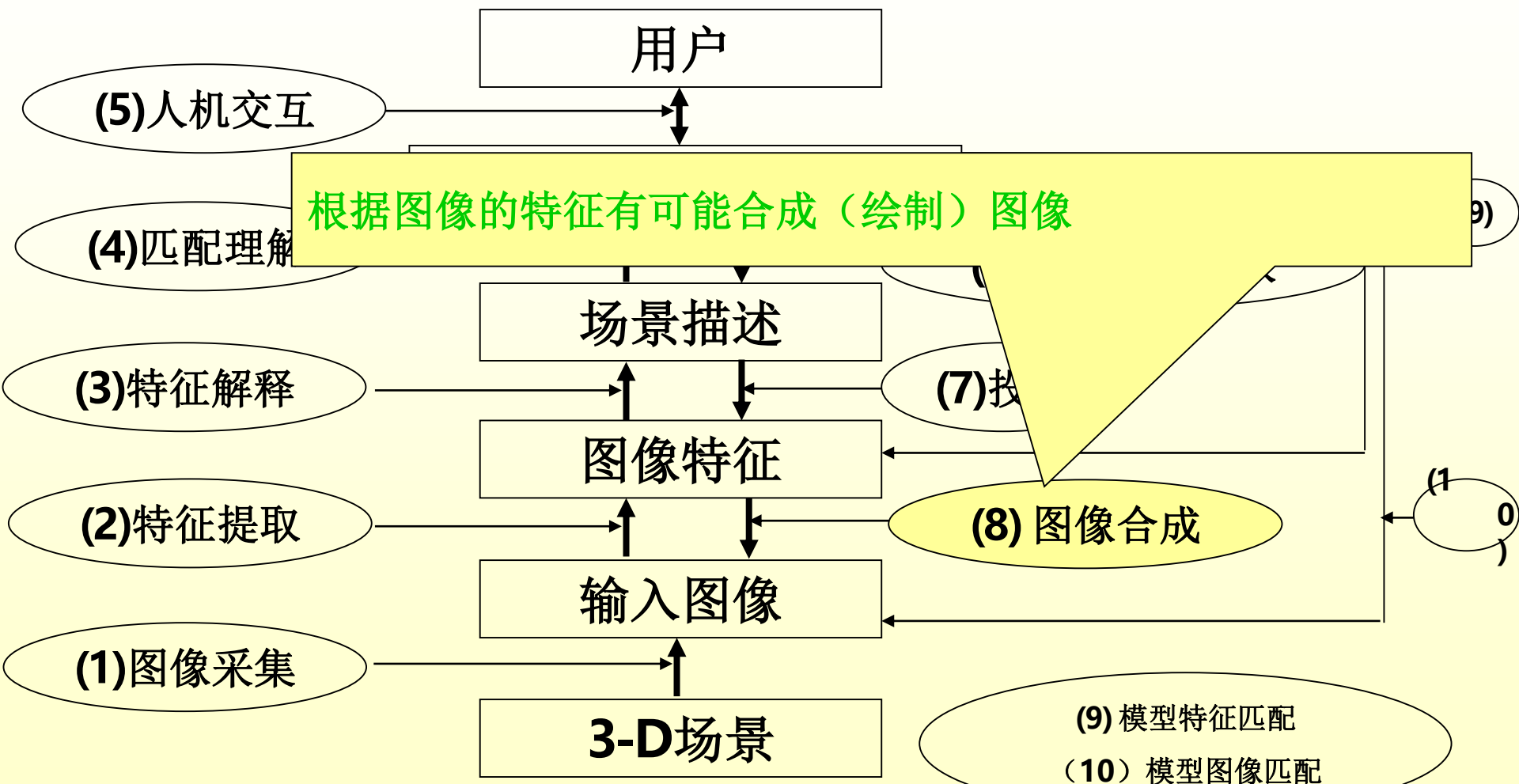


视觉信息系统模型





视觉信息系统模型



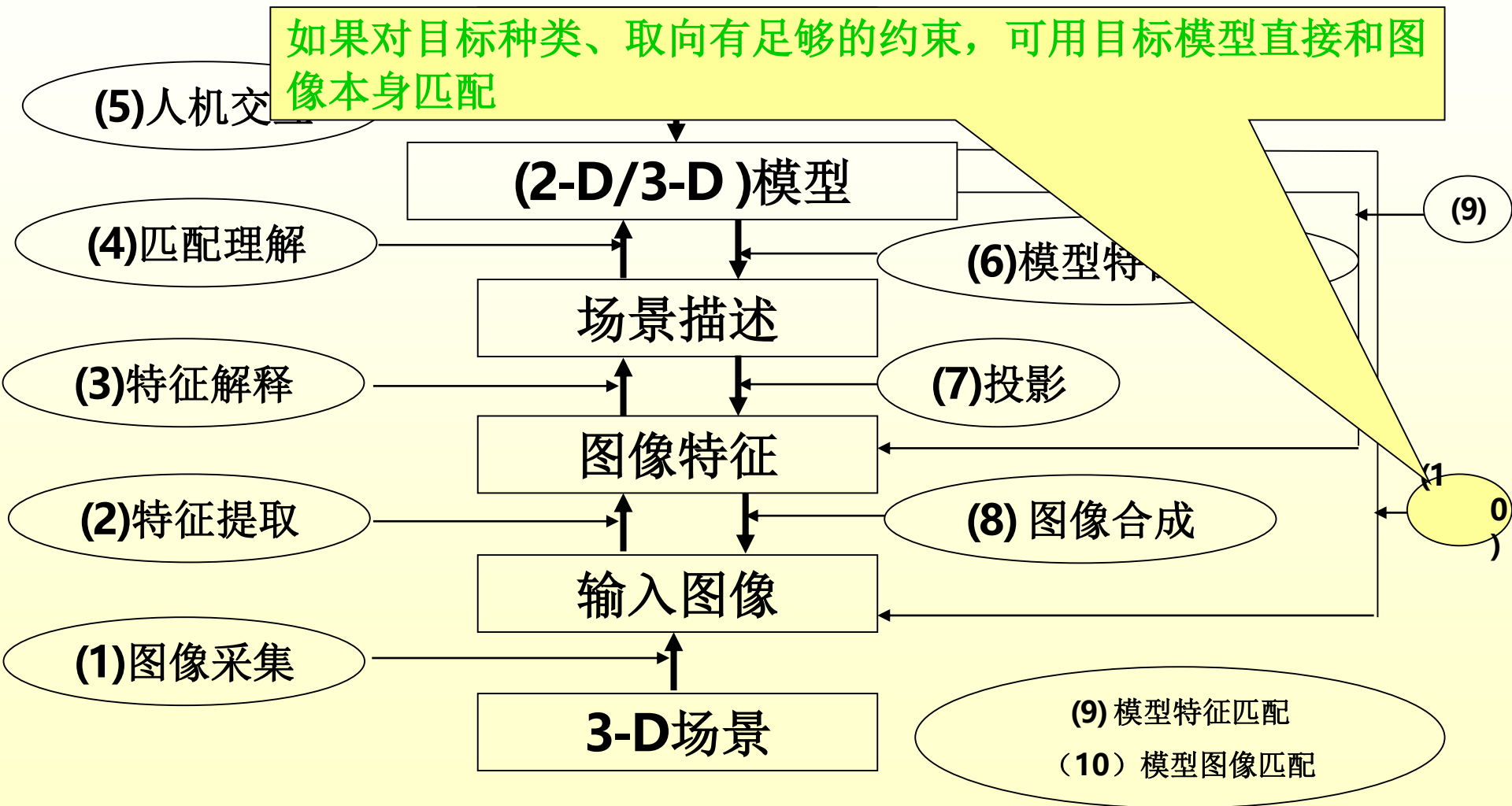
如果对目标种类、取向有足够的约束，可用目标模型直接和图像特征匹配





视觉信息系统模型

如果对目标种类、取向有足够的约束，可用目标模型直接和图像本身匹配





- 内容回顾
- 视觉理论与信息系统
- 运动目标检测跟踪



- **视频运动对象分析**有助于提取事件的发生、经过和结果等信息，是视频内容理解的基础。
- 作为视频运动对象分析核心的**运动对象检测和跟踪**，对提高视频应用的**效率与效果**都很重要，是系统**智能化**的体现。
 - **视频监控**：克服人易疲劳和连续反应速度迟缓等缺陷,避免连续观看丢失目标，实现无人值守
 - **视频存储与传输**：自动存储异常事件视频, 减少存储量，节省传输带宽
 - **视频内容检索与理解**：形成语义描述
 - **交通管理**：道路流量控制
 - **医学生物组织运动分析、气象云图分析预报**
 - **应用于人无法参与的领域**
 -

• 相邻帧差法



当前帧



相邻帧差法

不能适应摄像机运动、相邻帧光照突变情况，受噪声干扰较大，不能检测缓慢运动，易产生空洞。



□ 单分布高斯背景模型

- 对一个背景图像，像素亮度的分布满足高斯分布，即对背景图像B， (x, y) 点的亮度满足：

$$B(x, y) \sim N(\mu, \delta)$$

背景模型的每个像素属性包括两个参数：平均值 μ ，标准差 δ

- 如果满足
$$e^{\left(\frac{-(I(x,y)-\mu(x,y))^2}{2\delta^2}\right)} > T$$

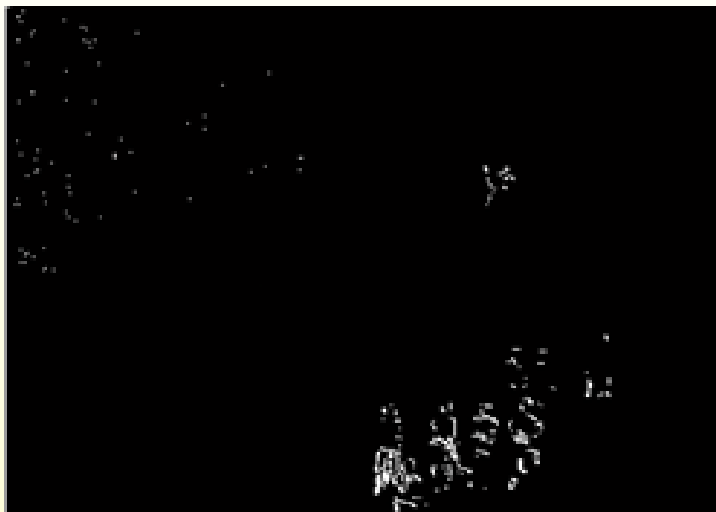
则认为 (x, y) 是背景点，反之是前景点

- 更新以适应背景图像的变换：

$$\mu(t+1, x, y) = \alpha \cdot \mu(t, x, y) + (1-\alpha) \cdot I(x, y)$$

α 为更新参数

- 背景减法



相邻帧差法



背景减法

不能适应摄像机运动、光照变化等情况，受噪声影响大，如果学习阶段本身存在物体运动，检测效果差。



□ 混合高斯模型

- K 个高斯混合模型表示监控场景背景像素点的强度变化，对于每个像素点， $\{I_1, \dots, I_t\}$ 表示已观察到的背景像素点强度，在时刻 t 具有强度 I_t 的某个像素点的概率估计为

$$P(I_t) = \sum_{j=1}^K \omega_{j,t} \cdot \eta(I_t, \mu_{j,t}, \Sigma_{j,t})$$

K 表示高斯分布的数目， K 由计算机的内存和计算能力决定，一般取值3~5即可。

$\omega_{j,t}$, $\mu_{j,t}$ 分别是 t 时刻第 j 个高斯分布的权重、均值， $\Sigma_{j,t}$ 是协方差， η 表示高斯分布函数。

$$\eta(I_t, \mu_{j,t}, \Sigma_{j,t}) = \frac{1}{(2\pi)^{\frac{d}{2}} \left| \Sigma_{j,t} \right|^{\frac{1}{2}}} e^{-\frac{1}{2}(I_t - \mu_{j,t})^T \Sigma_{j,t}^{-1} (I_t - \mu_{j,t})}$$



• 高斯混合模型

相邻帧差法



背景减法



高斯混合模型法



合理的模型可以解决“树叶晃动”、“背景物体移动”等问题。

不能适应摄像机运动、光照突变的情况。

对缓慢运动目标处理不好。



视频运动目标检测



线性预测法

不能适应摄像机运动、光照突变的情况，对光照渐变处理较好，但容易漏检，训练阶段有物体运动也能有较好效果。



特征背景法

不能适应摄像机运动和噪声干扰大的情况，对光照变化具有一定的适应性，结果受训练阶段的影响大。



- 训练阶段先获取 n 帧采样图像，每一帧图像有 P 个像素点，计算均值图像帧为 $back$ 。然后计算协方差矩阵，把 M 个特征向量存储在特征向量矩阵 Φ_{Mb} 中。



$back(352*288)$

```
Cb=cov(back);
```

```
Lb=eig(Cb);
```

```
Lb=sort(Lb,'descend');
```

```
Qm=zeros(row,row);
```

```
for i=1:row;
```

```
    Qm(i,i)=Lb(i);
```

```
end
```

$Cb(352*352)$

$Qm(288*288)$



视频运动目标检测

- 检测阶段时每到一个新帧 I ，把它投影到特征空间中 $I' = \Phi_{Mb}(I - back)$ 。然后把 I' 投影到图像空间 $I'' = \Phi_{Mb}^T I' + bac$ ，满足 $|I - I''| > T$ 的点作为是前景点。

$II = Qm * (I - back);$

$III = Qm' * II + back;$

$Diff = I - III;$

$m = \text{abs}(Diff) > \text{threshold};$



I



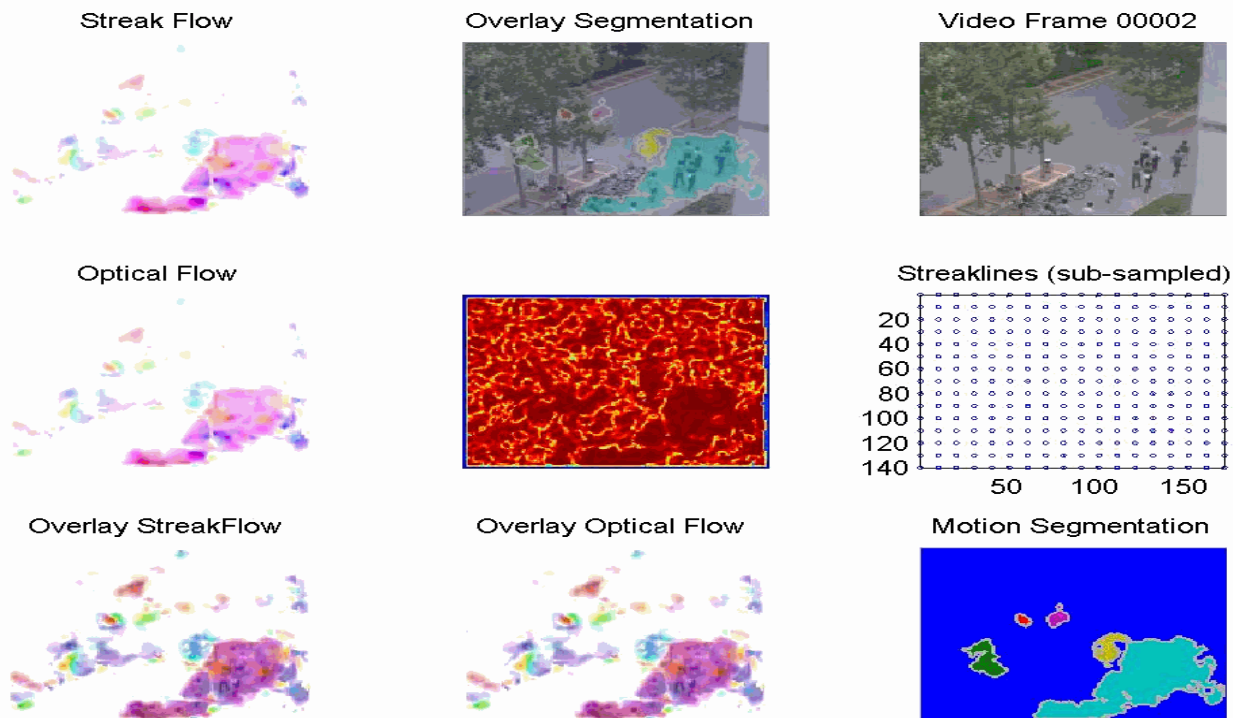
I''



视频运动目标检测

光流：二维可视速度场的表示。光流法的思想是在亮度不变假设下，求取连续图象帧间的相对运动，象素点的运动矢量构成光流场。

匹配法在相邻图象中查找象素（块）之间的对应关系，两幅图象之间对应象素（块）之间的位移即是所求的光流。





感兴趣区域提取

第一步

训练阶段视频亮度序列 Y_1, Y_2, \dots, Y_T

运动掩模 M_m

利用带状算子生成有意义运动掩模 M_{sm}

第二步

彩色背景帧 $(\bar{H}, \bar{S}, \bar{V})$

色彩感知粗分割, 生成粗分割掩模 M_c

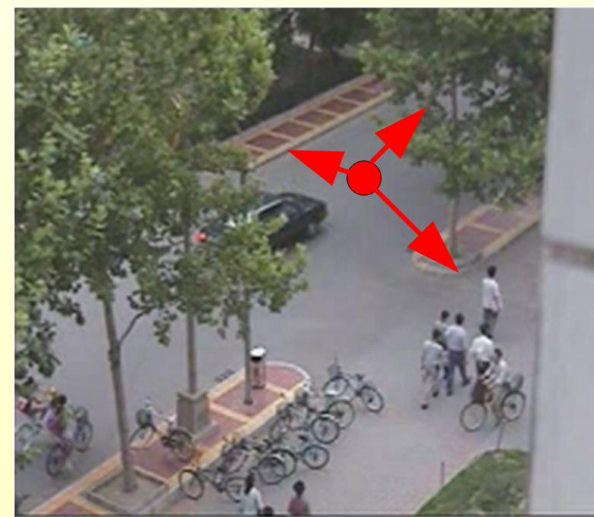
灰色区域亮度精分割; 彩色区域色调精分割; 生成颜色类型图 T_c

第三步

基于 M_{sm} 利用颜色一致进行区域生长

确定干扰对象进行区域退化

感兴趣区域 R



- 构造带状算子提取有意义运动区域

假设训练阶段的 T 帧中，像素点 (i,j) 取得的亮度最大值和最小值之差为 $D_{MN}(i,j)$

$$D(i, j) = \begin{cases} 1 & \text{if } D_{MN}(i, j) > dTh \\ 0 & \text{else} \end{cases}$$





- 构造带状算子提取有意义运动区域



se1 =

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	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
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

se2 =

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

se3 =

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

se4 =

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

se5 =

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



感兴趣区域提取



(a) 第100帧



(b) M_m



(c) M_{sm}



(d) M_d

训练阶段有意义运动区域提取结果



感兴趣区域提取



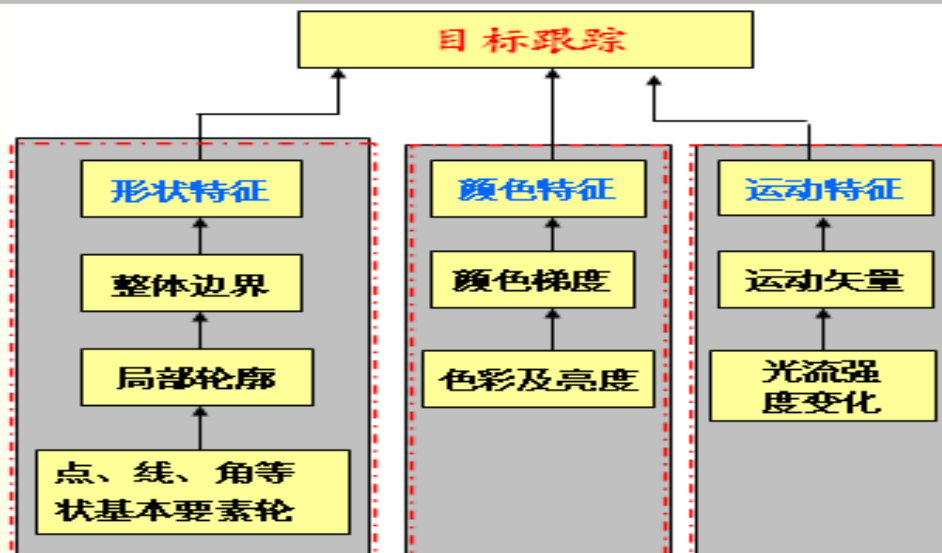


■ 目标跟踪

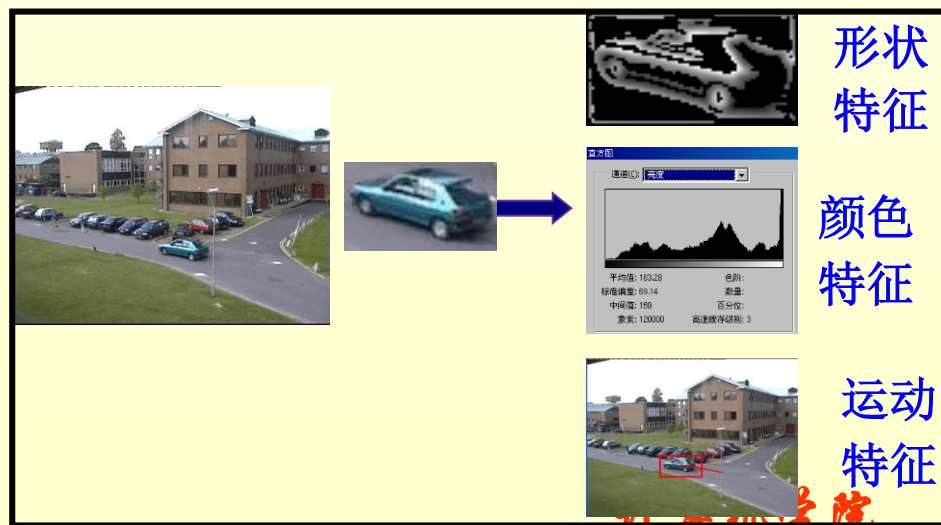
- 目标特征建模：颜色、形状、运动等特征，以及多特征融合
- 数据关联滤波：Kalman滤波，粒子滤波，相关滤波
- 特点：
 - 多特征加权融合
 - 模糊逻辑等方法调节特征权值

人类视觉对目标状态感知时，至少通过以下三方面

- **结构/形状**: 对点、线、角、轮廓等物体形状基本要素的感知
- **颜色**: 对色彩变化、对比度、饱和度等的感知
- **运动**: 通过前景与背景的差异，对物体运动的方向、速度等感知



人类视觉系统的感知





■构造特征提取算子

特征名称	层 次	作 用	提取方法
自适应显著颜色	中层特征	反映人眼敏感性提取显著颜色	非线性变换建立目标和背景颜色分类器提取目标显著颜色
加权灰度直方图	中层特征	刻画目标，进行遮挡判断	以目标区域中心位置为核函数中心，构建核函数加权灰度直方图

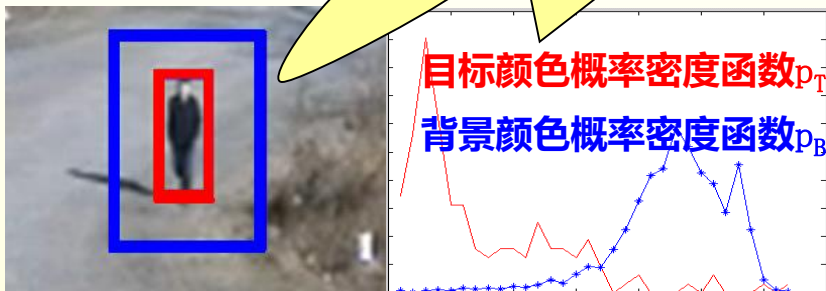
■构造特征提取算子

特征名称	层次	作用	提取方法
角点纹理	中层特征	提取目标显著点 周围纹理	SIFT
边缘轮廓	拓扑结构特征	提取目标轮廓	Canny, 水平集
纹理特征	中层特征	提取图像纹理	HOG

■突出目标特性的显著颜色提取方法

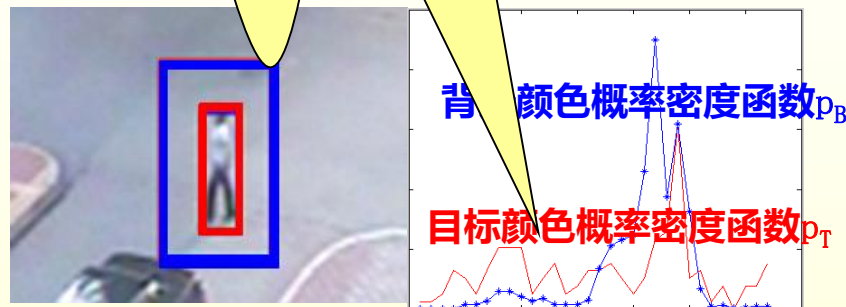
□ 难点:

上衣、裤子颜色
为显著颜色



目标和背景颜色不相似

头、裤子颜色为
显著颜色



目标和背景颜色存在相似

□ 研究思路:

- 根据人类视觉对醒目颜色的注意机制, 提取目标中和背景差异大的颜色作为显著颜色

■ 颜色差异模型 $(i) = \log \frac{\max \{ p_T(i), \delta \}}{\max \{ p_B(i), \delta \}}$

考虑目标和
背景的差异
计算机学院



- 对于灰度图像，一般为了去除噪点，所以会先利用离散高斯平滑模板进行平滑
- 梯度法求图像梯度

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{bmatrix}$$

$$|\nabla f| = \|\nabla f\|_2 = [G_x^2 + G_y^2]^{1/2} = \left[\left(\frac{\delta f}{\delta x} \right)^2 + \left(\frac{\delta f}{\delta y} \right)^2 \right]^{1/2}$$

$$G_h(x, y) = f(x+1, y) - f(x-1, y) \quad \forall x, y$$

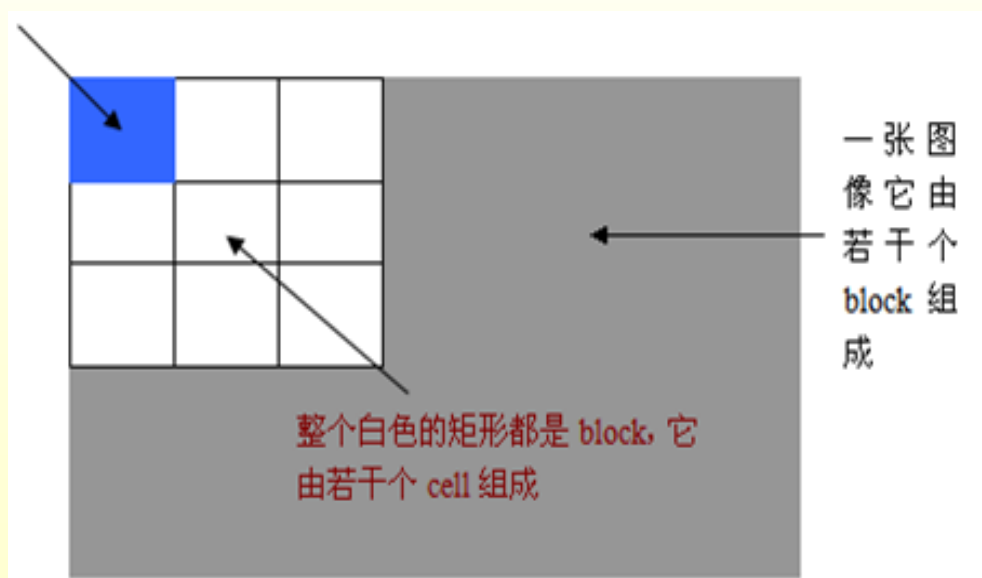
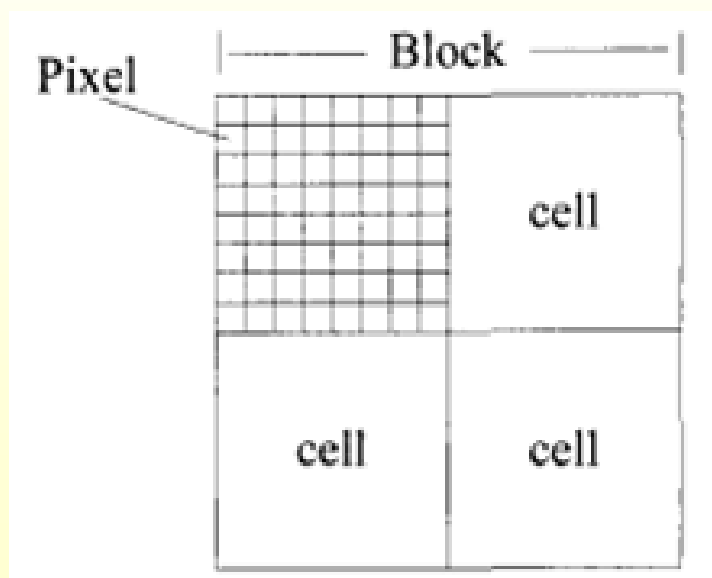
$$G_v(x, y) = f(x, y+1) - f(x, y-1) \quad \forall x, y$$

$$M(x, y) = \sqrt{G_h(x, y)^2 + G_v(x, y)^2} \approx |G_h(x, y)| + |G_v(x, y)|$$

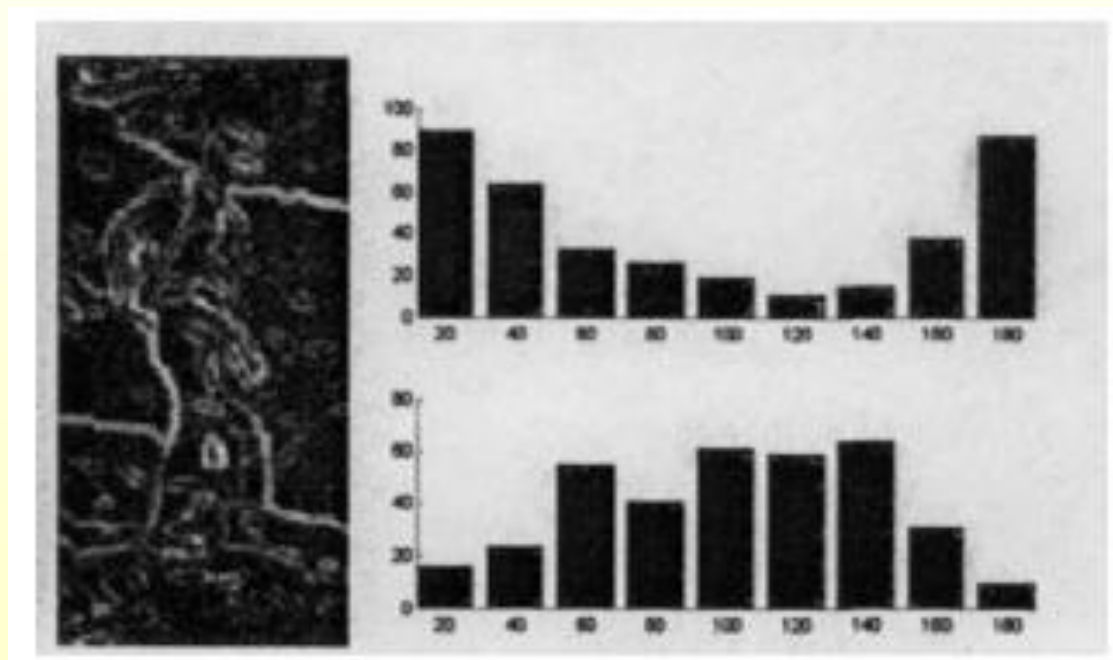
$$\theta(x, y) = \arctan(G_h(x, y) / G_v(x, y))$$

- 矩形HOG块的划分

- 一般一个块 (Block) 都由若干单元 (Cell) 组成, 一个单元都由若干个像素点组成。



- 在每个Cell中独立做梯度方向统计，从而形成以梯度方向为横轴的直方图。梯度分布平均分成数个方向角度（orientation bins），每个方向角度范围都会对应一个直方柱。





- **主要步骤:**
- 将图像划分成小cells (例如 $6*6$ 像素/cell) ;
- 统计每个cell的梯度直方图 (不同梯度的个数) , 即可形成每个cell的descriptor;
- 将每几个cell组成一个block (例如 $3*3$ 个cell/block) , 一个block内所有cell的特征descriptor串联起来便得到该block的HOG特征descriptor。
- 将图像image内的所有block的HOG特征descriptor串联起来就可以得到该image (你要检测的目标) 的HOG特征descriptor了。这个就是最终的可供分类使用的特征向量。



- **HOG与SIFT区别**

- HOG和SIFT都属于描述子，在具体操作上有很多相似的步骤，但是两者在使用目的和具体处理细节上有很大的区别。HOG与SIFT的主要区别如下：
 - ① SIFT是基于关键点特征向量的描述。
 - ② HOG是将图像均匀的分成相邻的小块，然后在所有的小块内统计梯度直方图。
 - ③ SIFT一般有两大步骤，第一个步骤是对图像提取特征点，而HOG不会对图像提取特征点。
 - ④ SIFT需要对图像尺度空间下对像素求极值点，而HOG中不需要。



- **HOG的优点:**
- **HOG表示的是边缘（梯度）的结构特征，因此可以描述局部的形状信息；**
- **而且由于它这种分块分单元的处理方法，也使得图像局部像素点之间的关系可以很好得到的表征；**
- **位置和方向空间的量化一定程度上可以抑制平移和旋转带来的影响；**
- **采取在局部区域归一化直方图，可以部分抵消光照变化带来的影响；**
- **采用梯度直方图，使得图像所需要的表征数据的维度降低了。**

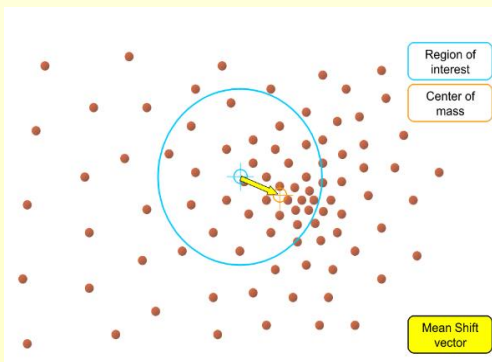
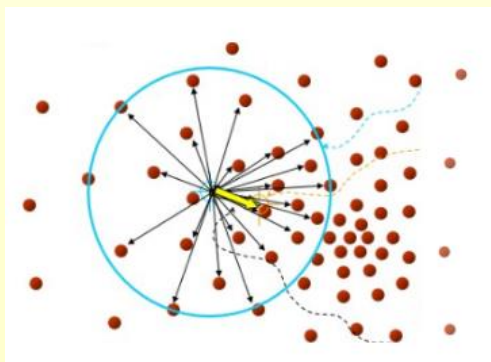


- **HOG的缺点:**
- 描述子生成过程冗长，导致速度慢，实时性差；
- 很难处理遮挡问题；
- 由于梯度的性质，该描述子对噪点相当敏感

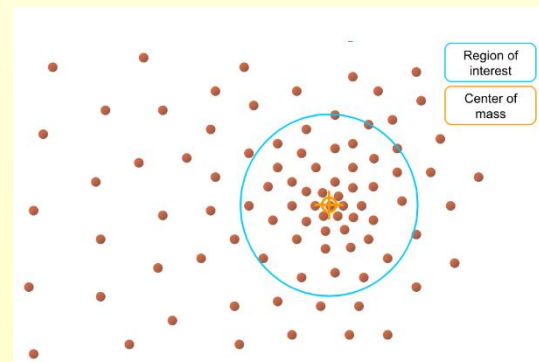


• Kalman滤波、Meanshift、粒子滤波、相关滤波

- Kalman滤波：认为物体的运动模型服从高斯模型，对目标的运动状态进行预测，然后通过与观察模型进行对比，根据误差来更新运动目标的状态
- Meanshift：本质是一个迭代的过程，在一组数据的密度分布中，使用无参密度估计寻找到局部极值



.....





北京航空航天大学

Beijing University of Aeronautics and Astronautics



计算机学院



- **Kalman滤波、Meanshift、粒子滤波、相关滤波**
- **粒子滤波** (PF: Particle Filter) 的思想基于蒙特卡洛方法, 它是利用粒子集来表示概率, 可以用在任何形式的空间状态模型上。其核心思想是通过**从后验概率中抽取的随机状态粒子来表达其分布**, 是一种顺序重要性采样法(Sequential Importance Sampling)。
- 简单来说, 粒子滤波法是指通过**寻找一组在状态空间传播的随机样本对概率密度函数进行近似, 以样本均值代替积分运算, 从而获得状态最小方差分布的过程**。这里的样本即指粒子, 当样本数量 $N \rightarrow \infty$ 时可以逼近任何形式的概率密度分布。



- 尽管算法中的概率分布只是真实分布的一种近似，但由于非参数化的特点，它摆脱了解决非线性滤波问题时随机量必须满足高斯分布的制约，能表达比高斯模型更广泛的分布，也对变量参数的非线性特性有更强的建模能力。因此，粒子滤波能够比较精确地表达基于观测量和控制量的后验概率分布。
- 非高斯、非线性

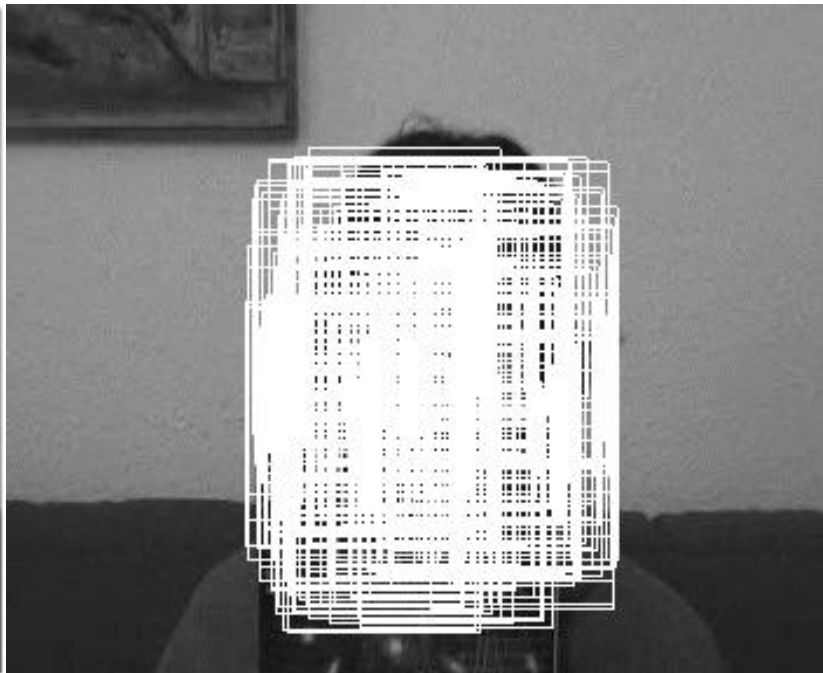
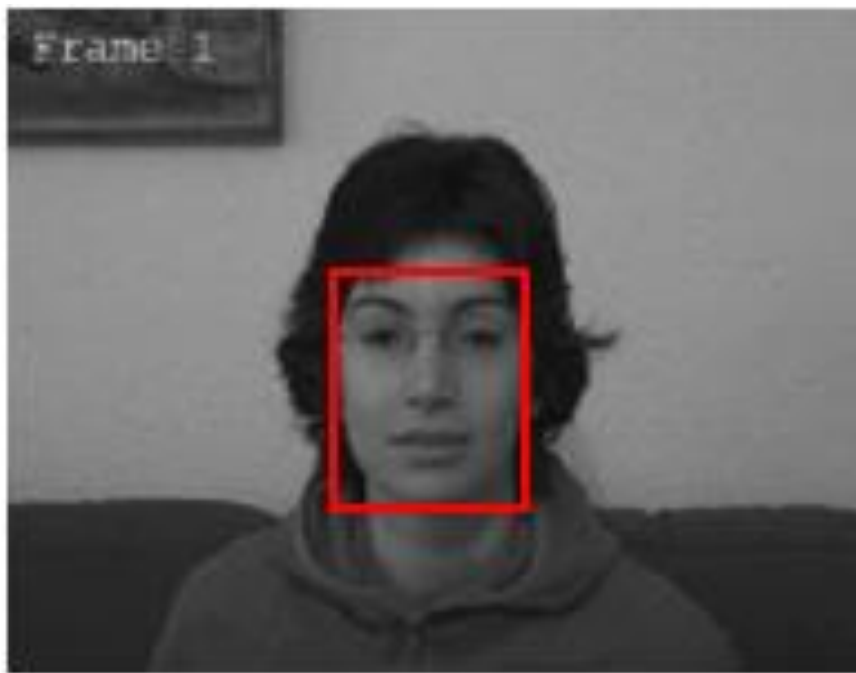


- 初始化所有粒子
- 更新粒子当前位置
- 评估每个粒子的重要性
- 根据粒子重要性重新采样

多特征融合的运动目标跟踪算法

■ 采样

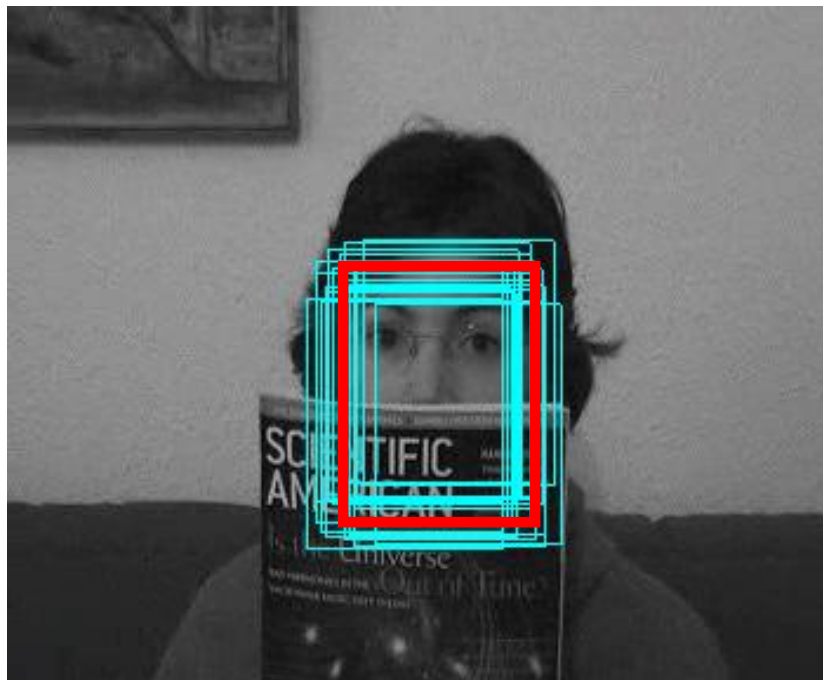
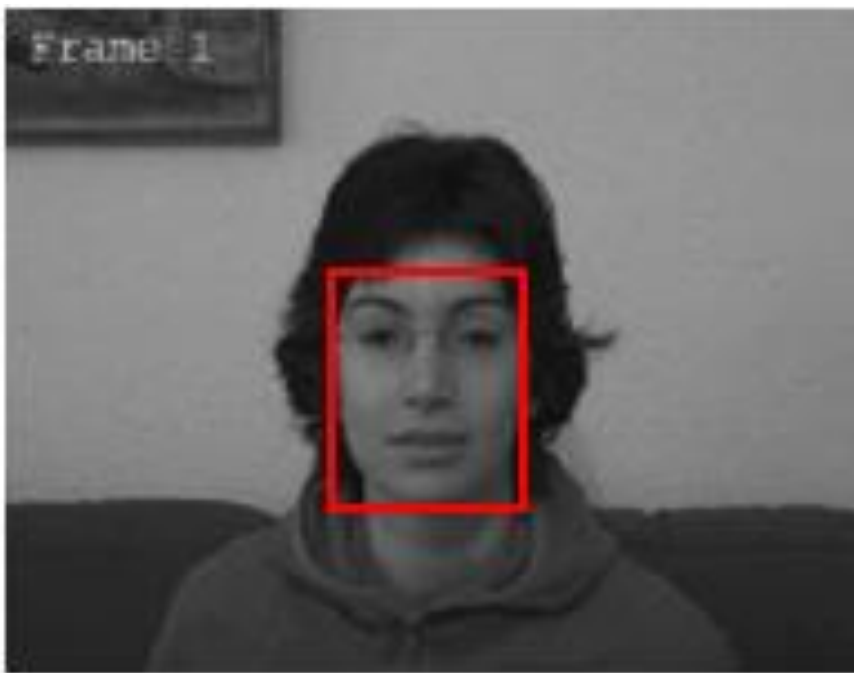
□ 散布粒子，得到候选目标



运动目标跟踪算法

■ 目标定位

□ 相似度前10%的粒子加权



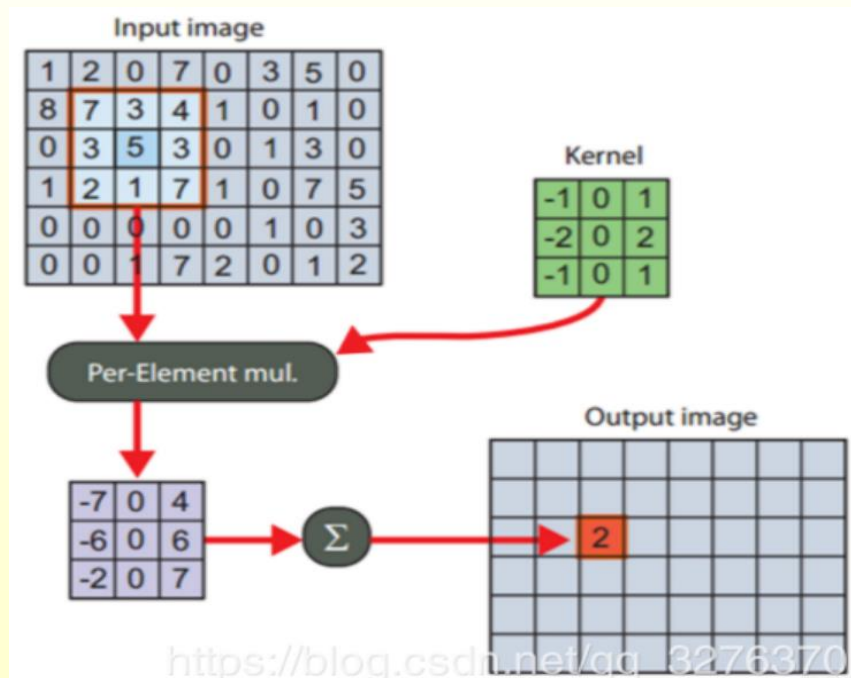


- **主要问题：需要用大量的样本数量才能很好地近似系统的后验概率密度。环境越复杂，描述后验概率分布所需要的样本数量就越多，算法的复杂度就越高。因此，能够有效地减少样本数量的自适应采样策略是该算法的重点。**
- **重采样阶段会造成样本有效性和多样性的损失，导致样本贫化现象。如何保持粒子的有效性和多样性，克服样本贫化，也是该算法研究重点。**



- 相关滤波
- 相关：用来描述两个信号的相似程度

-1	-1	-1
2	2	2
-1	-1	-1



输入信号的信息与核越相似，响应值就越高。只要在下一帧找到响应值最高的位置就可以定位目标，实现跟踪