



# 人体行为识别概述



- **VSAM**(Visual Surveillance and Monitoring) 1997
- **IVPL实验室**(The Image and Video Processing Laboratory)
- **AIRVL实验室**(Artificial Intelligence, Robotics, and Vision Laboratory)
- **LPAC实验室**(Laboratory for Perception, Action and Cognition)
- **KNIGHT系统**
- **ISCAPs**(Integrated Surveillance of Crowded Areas for Public Security)
- **REASON**(Robust Methods for Monitoring and Understanding People in Public spaces)
- **CANTATA**(Content Aware Networked systems Towards Advanced and Tailored Assistance)
- 中国科学院自动化所的生物识别与安全技术研究中心**CBSR**
- 微软亚洲研究院

# 行为识别前景



上传图片搜索

框选搜索区域，锁定搜索目标，框线越贴近目标，搜索结果越好

框图示意

正在搜索的图片

商品类目: 箱包 手提包

性别: 女 男

立即

val  
ms  
的事件。

你是不是要找: 女-单肩包 女-斜挎包 女-手拿包 男-手提包

相似商品(3) 相关商品(0) 淘友推荐(0)

从结果中筛选:  相同款  相似款  相同风格  
 服务:  如实描述  7天退换  正品保障  其他  
 颜色:  相同色系

确定

GUCCI/古奇女包/女士包包/帆布包/女包/2010新款/帆布包  
¥ 698.00 广东广州

香港 正品品质 GUCCI大包 古奇新款人气购物包 单肩包  
¥ 1580.00

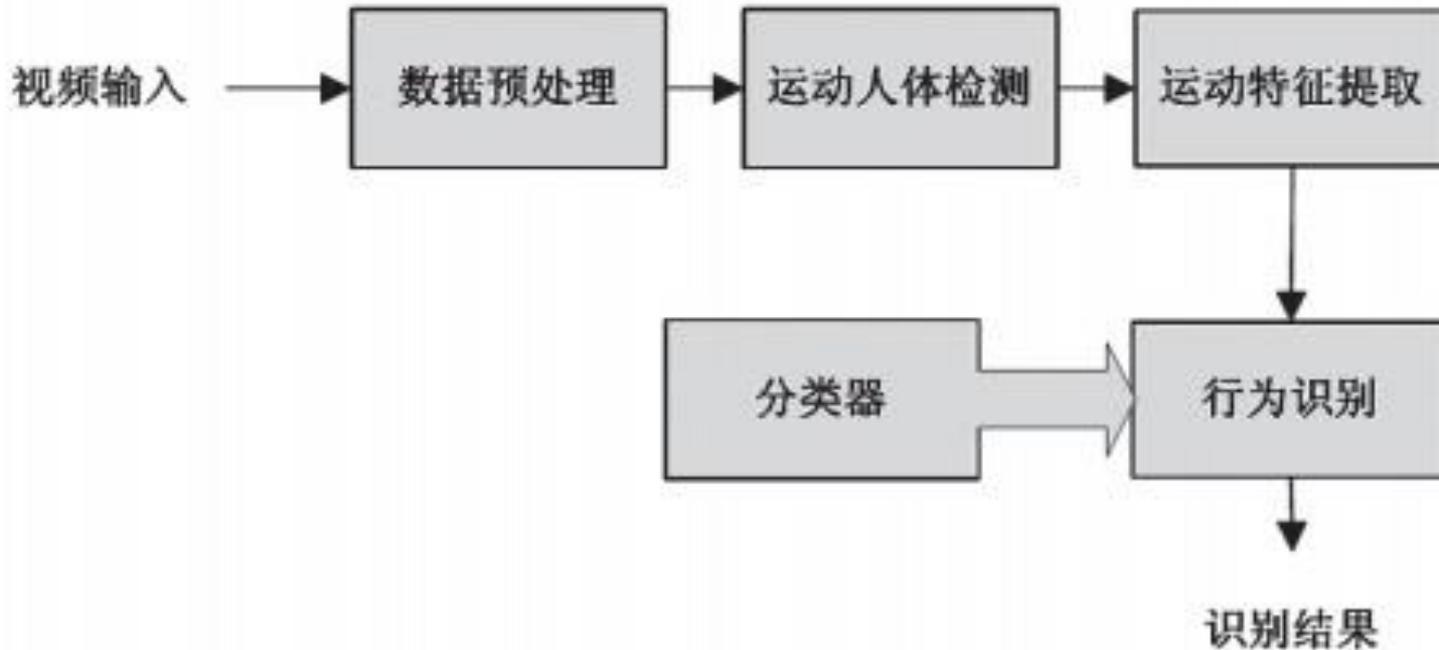
正品品质 新款GUCCI/古奇男包 公文包 斜挎包 手提包  
¥ 698.00 广东广州

## medical area

辅助诊断病人的运动问题。另一

参考文献: Action Recognition in Videos:from Motion Capture Labs to the Web

# 行为识别的流程



- 目标的运动特征，可以用于运动表述，是行为理解等高层部分的基础。
- 常见的运动表述方法有：**运动轨迹**、**时空图表述**和光流。

参考文献：运动人体行为分析  
视频中的人体运动分析及其应用研究

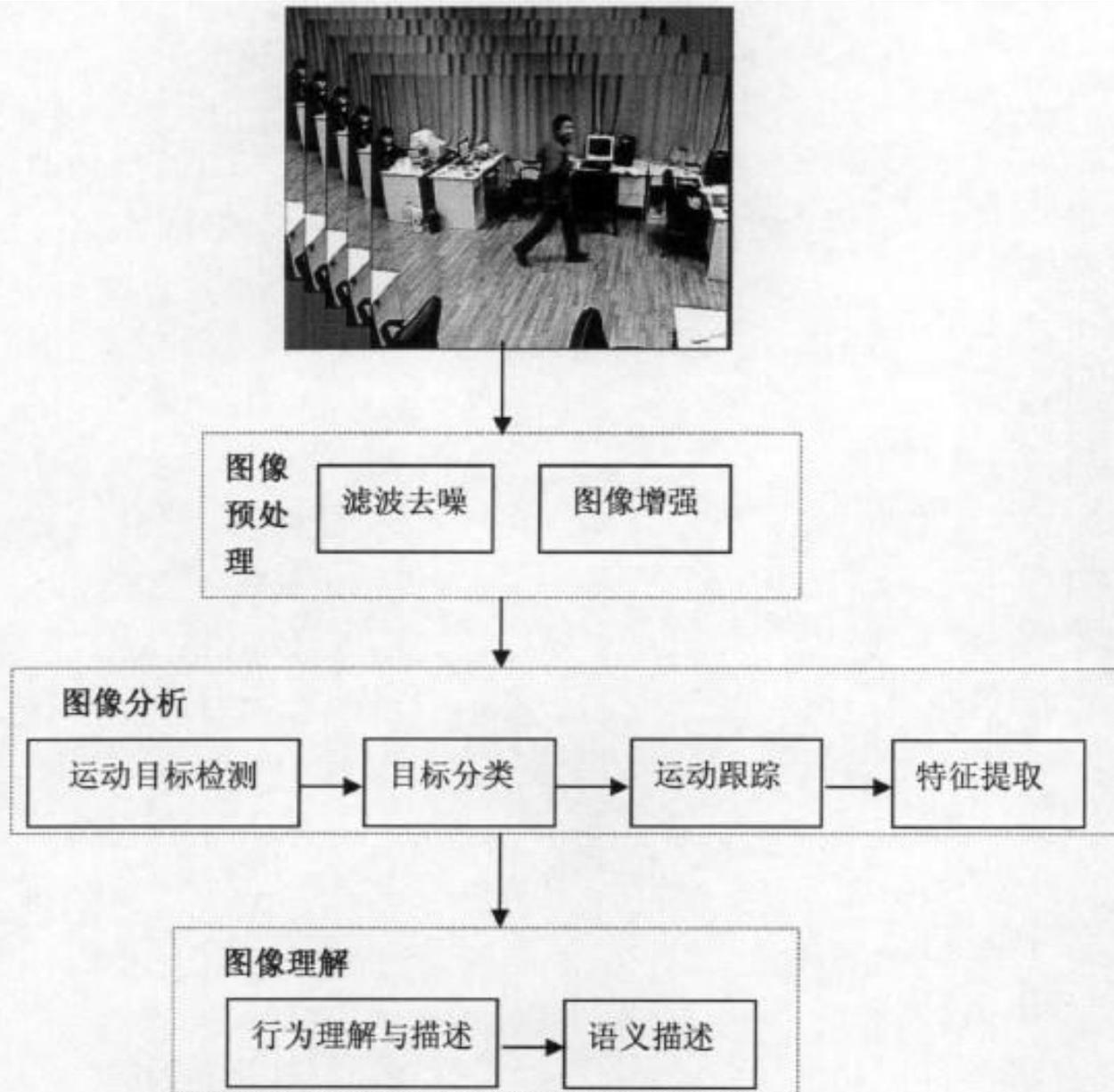


图 1.1 人体运动分析系统的一般性框架



**01**

目标检测

**02**

目标分类

**03**

目标跟踪

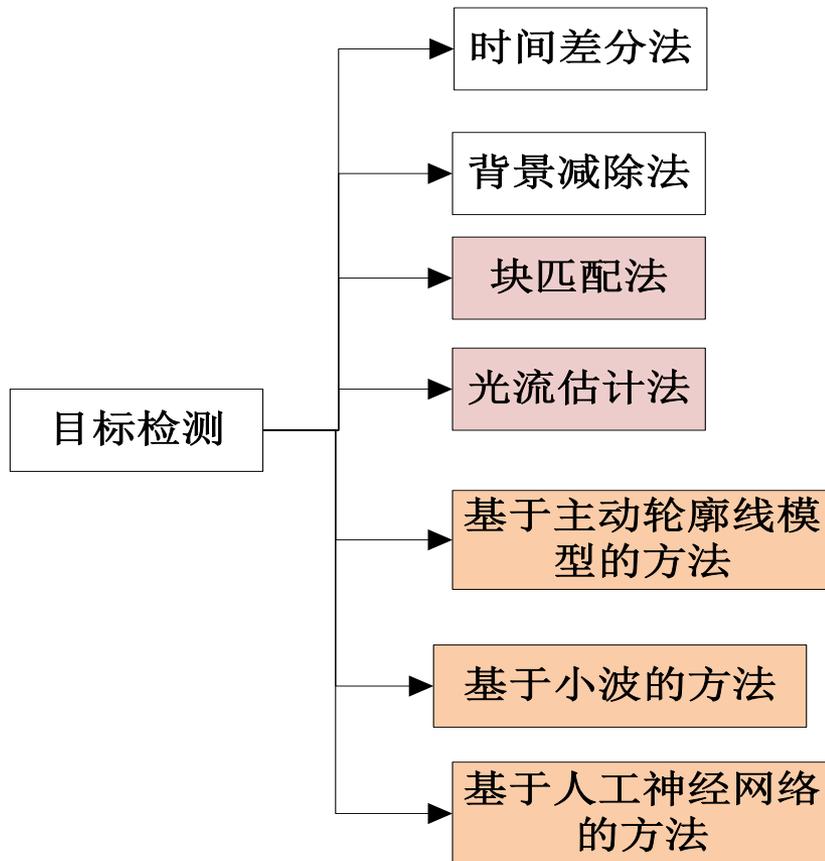
**04**

特征提取

**05**

行为理解

- 对于不依赖先验知识的目标跟踪来讲，运动检测是实现跟踪的第一步
- 运动检测即为从序列图像中将变化区域从背景图像中提取出来。



参考文献：视频中的人体运动分析及其应用研究.nh  
运动目标跟踪算法研究综述.pdf

# 目标检测——时间差分法

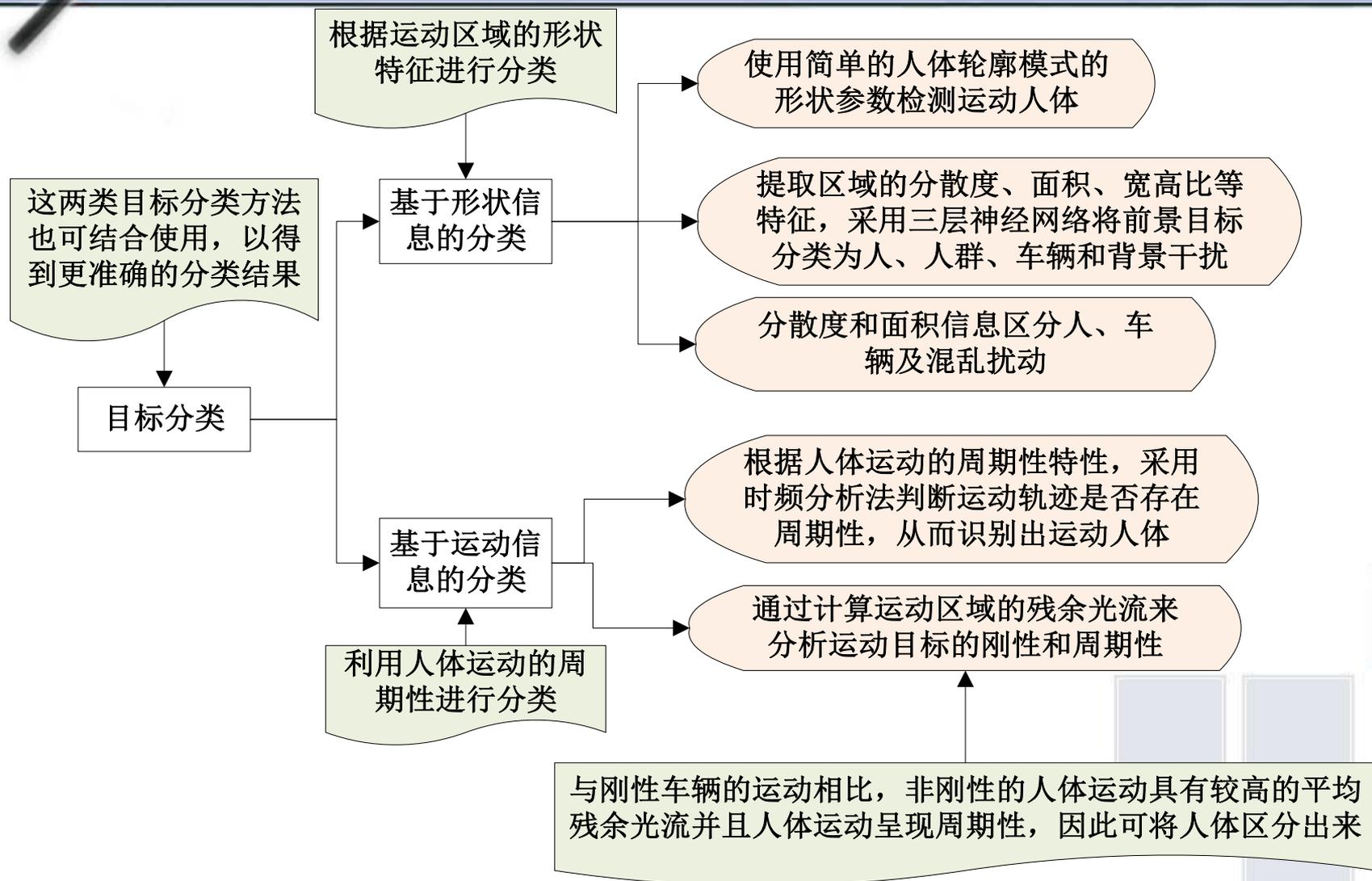
- 在连续的图像序列中**两幅或三幅相邻帧**之间采用**基于像素**的时间差分，并对差分结果进行**阈值化处理**以**提取**图像中的**前景**运动区域。
- **缺点：**
- 前景、背景区域的确定与**阈值的选取**有很大的关系  
当灰度图像序列对比度较低时，由于相邻两帧的差(前景与背景之差)的范围很小，**阈值难以选取**，影响前景目标的分割结果。
- 区域**灰度值**变化较为**平坦时**，容易在人体二值图像内产生**空洞现象**，给后续的目标分类、跟踪和识别造成不便。
- **优点：**
- 对于**动态环境**有**较强的自适应性**

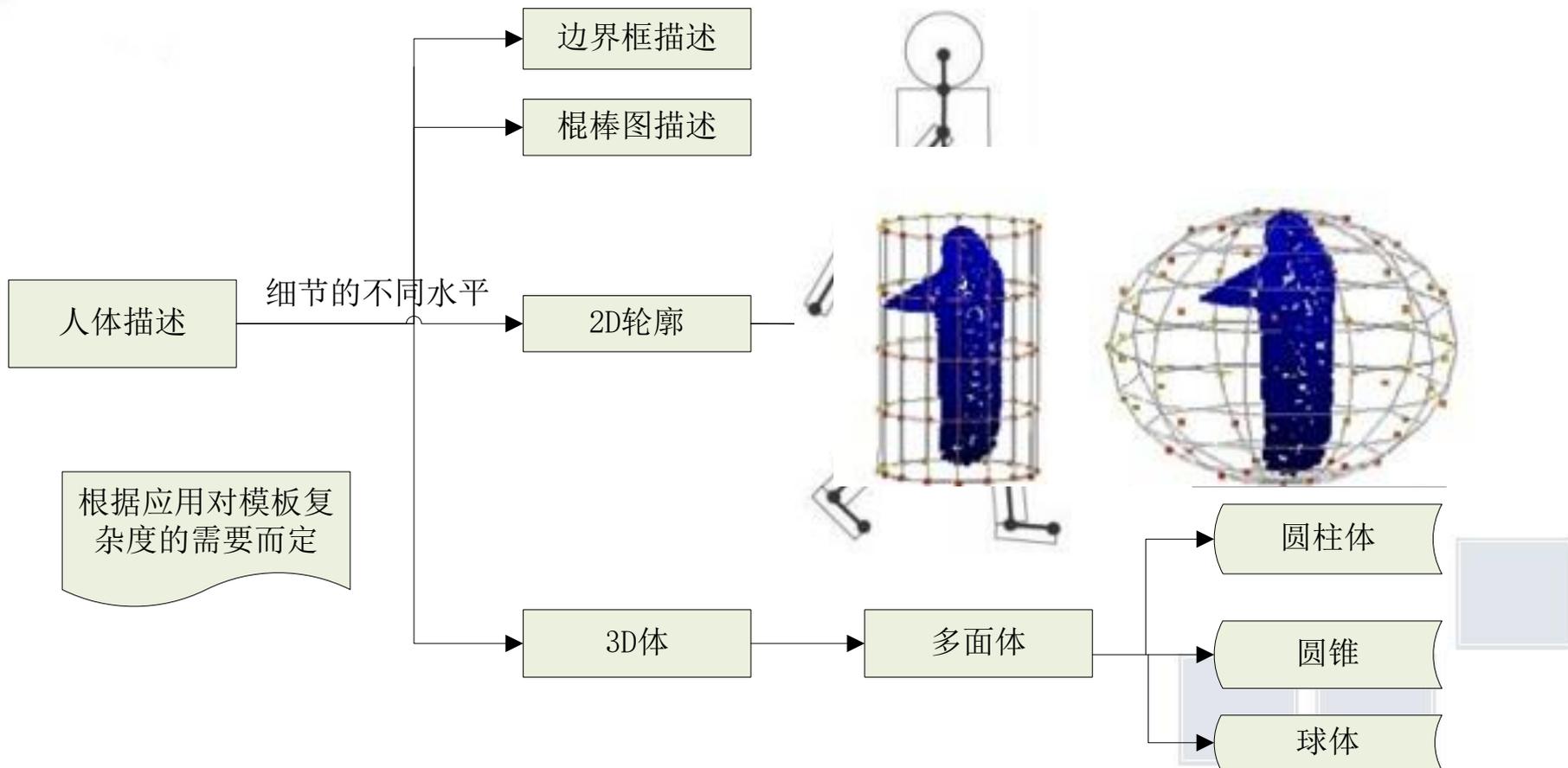


- 基于光流法的运动检测，是利用运动物体随时间变化在图像中表现的光流特性，通过计算位移向量光流场来提取运动目标。
- 光流是空间运动物体在观测成像面上对应像素运动的瞬时速度，是空间物体可见点的三维速度矢量在成像平面上的投影，它携带了丰富的运动和结构信息。
- **优点：**
- 即使在摄像机运动的情况下也能检测出独立的运动目标。
- **缺点：**
- 由于噪声、多光源、阴影和遮挡等原因，计算出的光流场分布并不十分可靠和准确。
- 多数光流法计算复杂、耗时，在实际的系统没有特殊的硬件支持时，很难实现实时检测。



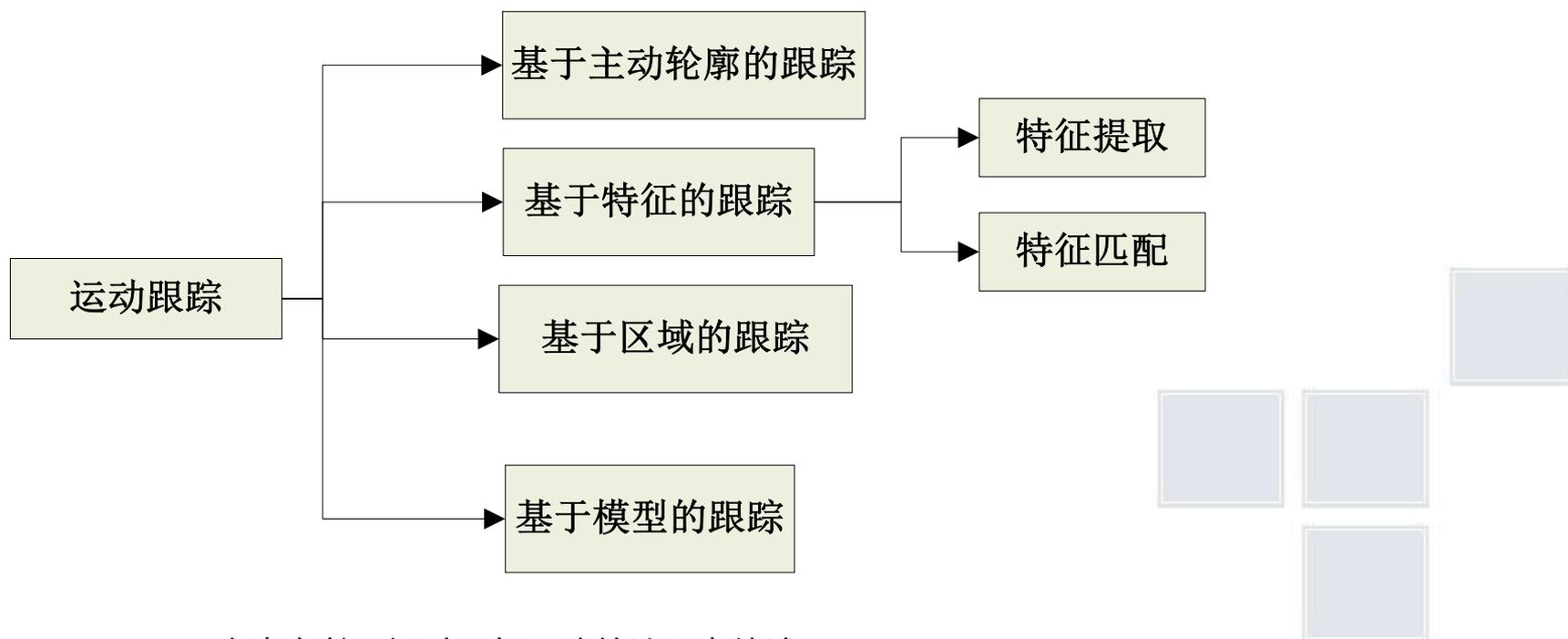
- 最常用且有效的是背景减除法。
- 背景减除法最简单的实现方法是预先选取不含前景运动目标的背景图像，然后将当前图像帧与背景图像相减得到前景目标。
- 背景减除法通常在**摄像机固定的情况下使用**，**关键**是建立随场景变化**不断更新**的背景模型。
- 有**两类**常用的背景更新方法：
  - 1、建立背景模型并采用自适应方法对模型参数进行调整，从而获得新背景图像；
  - 2、从过去的一组观测图像中按一定的假设选择像素值构成当前背景图像。





参考文献: *Human Motion: Modeling and Recognition of Actions and Interactions*

- 运动目标的跟踪，即通过目标的有效表达，在图像序列中寻找与目标模板最相似候选目标区位置的过程。
- 就是在序列图像中为**目标定位**。
- 范例

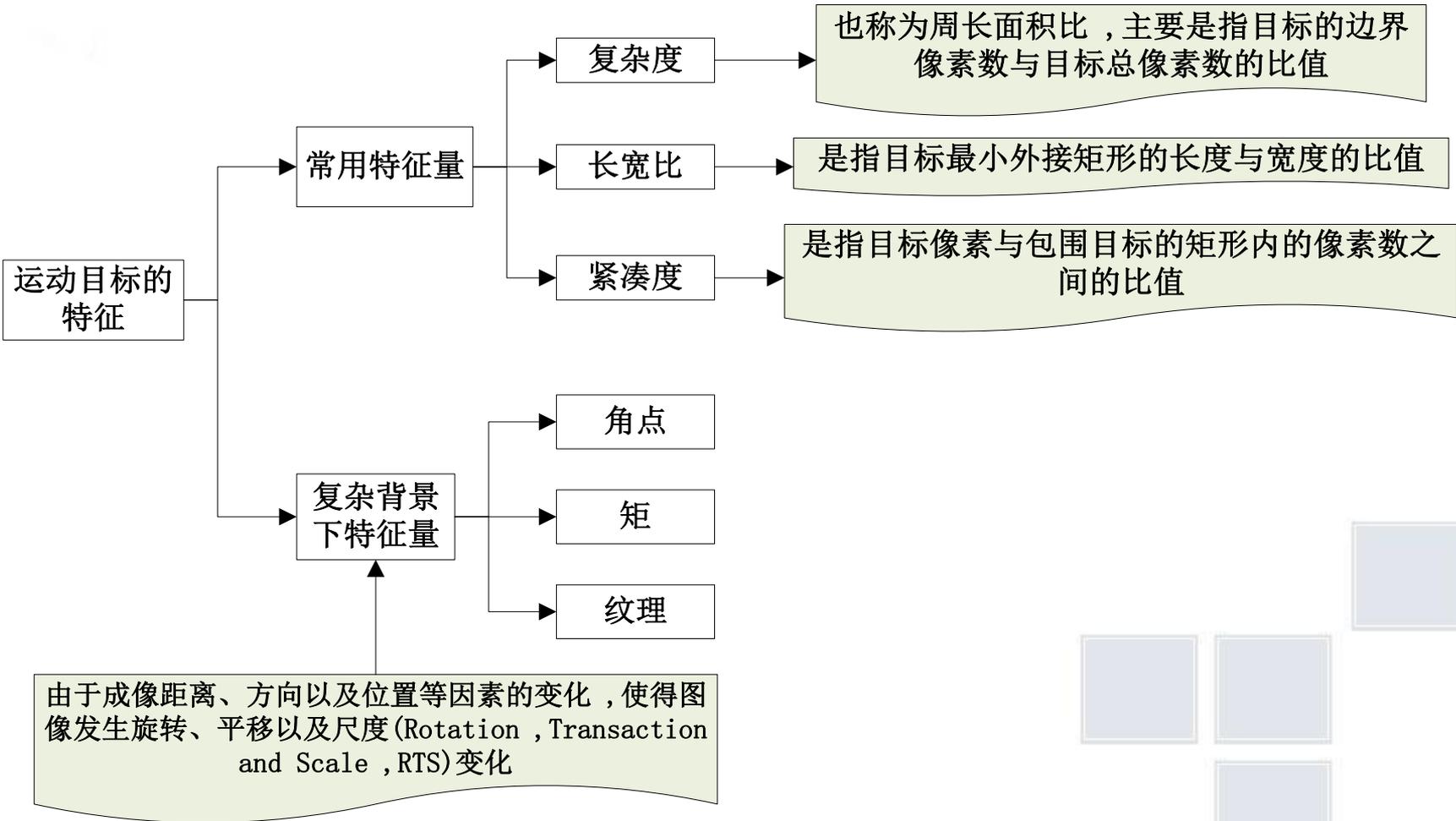


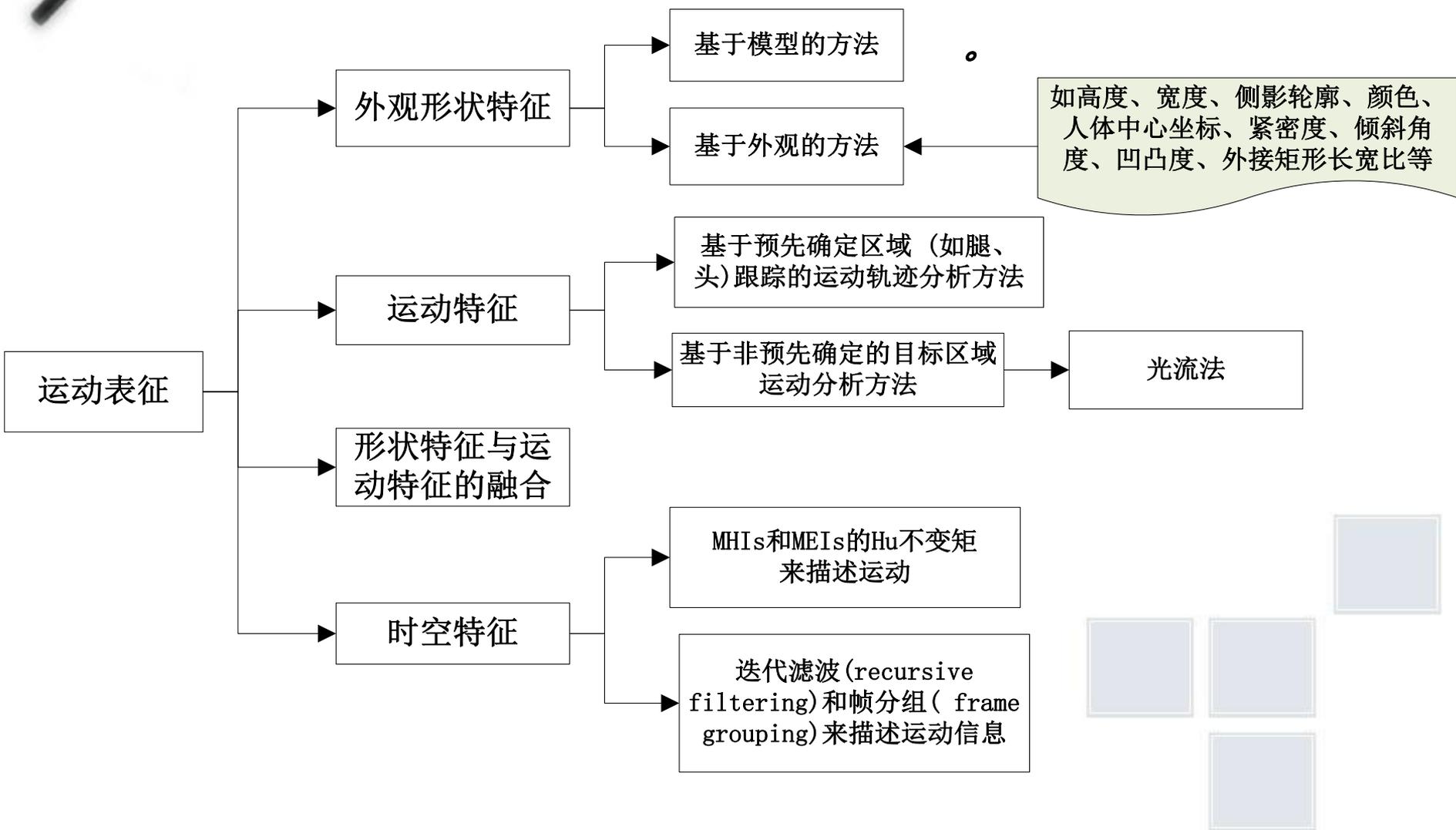


# 四种跟踪方法的比较

方法名	优点	缺点
主动轮廓模型 (snake 模型)	不但考虑来自图像的灰度信息，而且考虑整体轮廓的几何信息，增强了跟踪的可靠性。	1、计算量比较大 2、对于快速运动的物体或者形变较大的情况，跟踪效果不够理想
基于特征的跟踪	对运动目标的尺度、形变和亮度等变化不敏感，与 Kalman 滤波器联合使用，也具有很好的跟踪效果。	1、对于图像模糊、噪声等比较敏感， 2、连续帧间的特征对应关系也较难确定，
基于区域的跟踪	当目标未被遮挡时，跟踪精度非常高、跟踪非常稳定。	1、费时，当搜索区域较大时情况尤其严重； 2、算法要求目标变形不大，且不能有太大遮挡，否则相关精度下降会造成目标的丢失。
基于模型的跟踪	不易受观测视角的影响，具有较强的鲁棒性，模型匹配跟踪精度高，适合于机动目标的各种运动变化，抗干扰能力强，	计算分析复杂、运算速度慢，模型的更新较为复杂，实时性较差。 准确建立运动模型是模型匹配能否成功的关键。

# 运动目标的特征

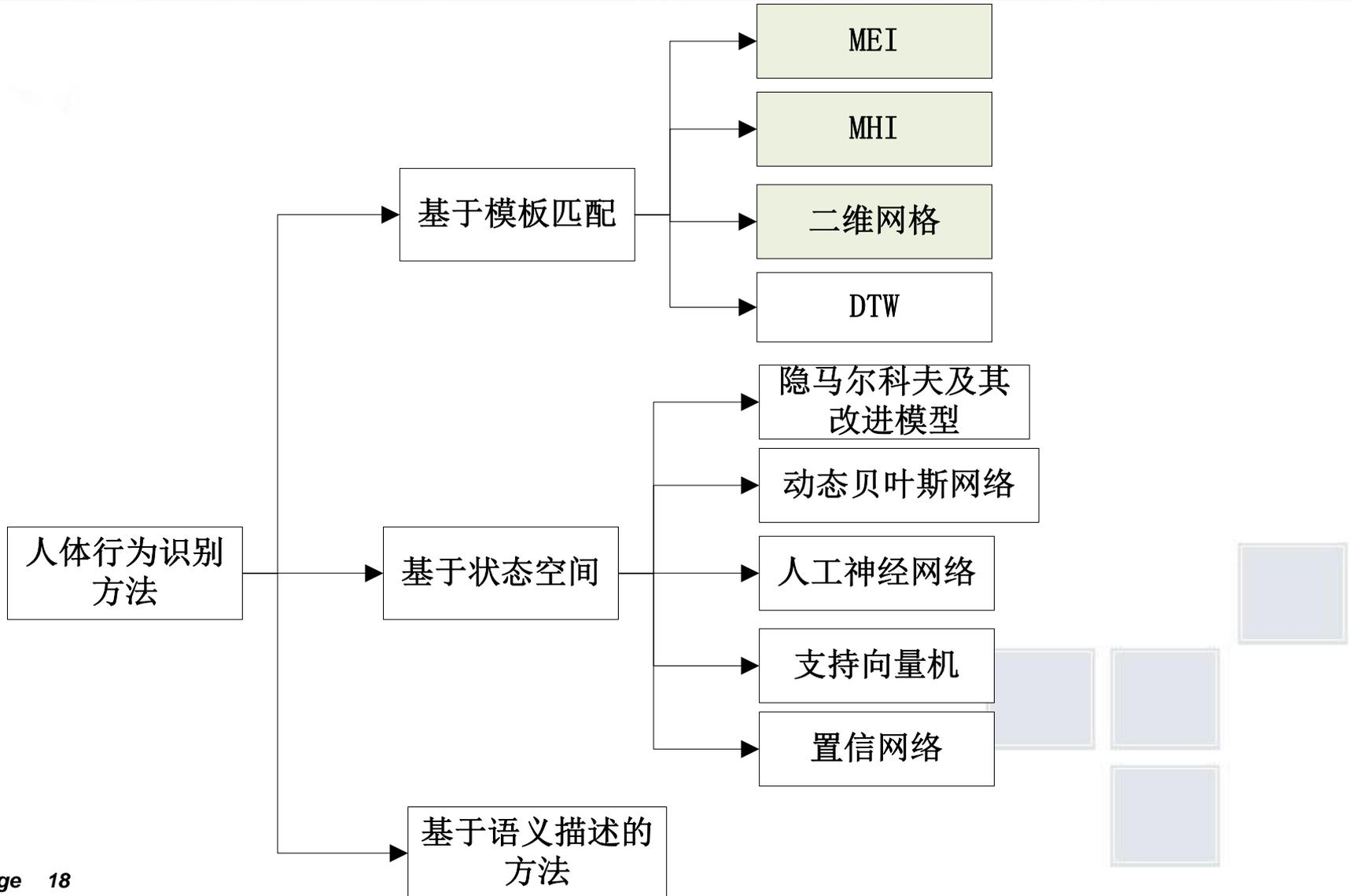






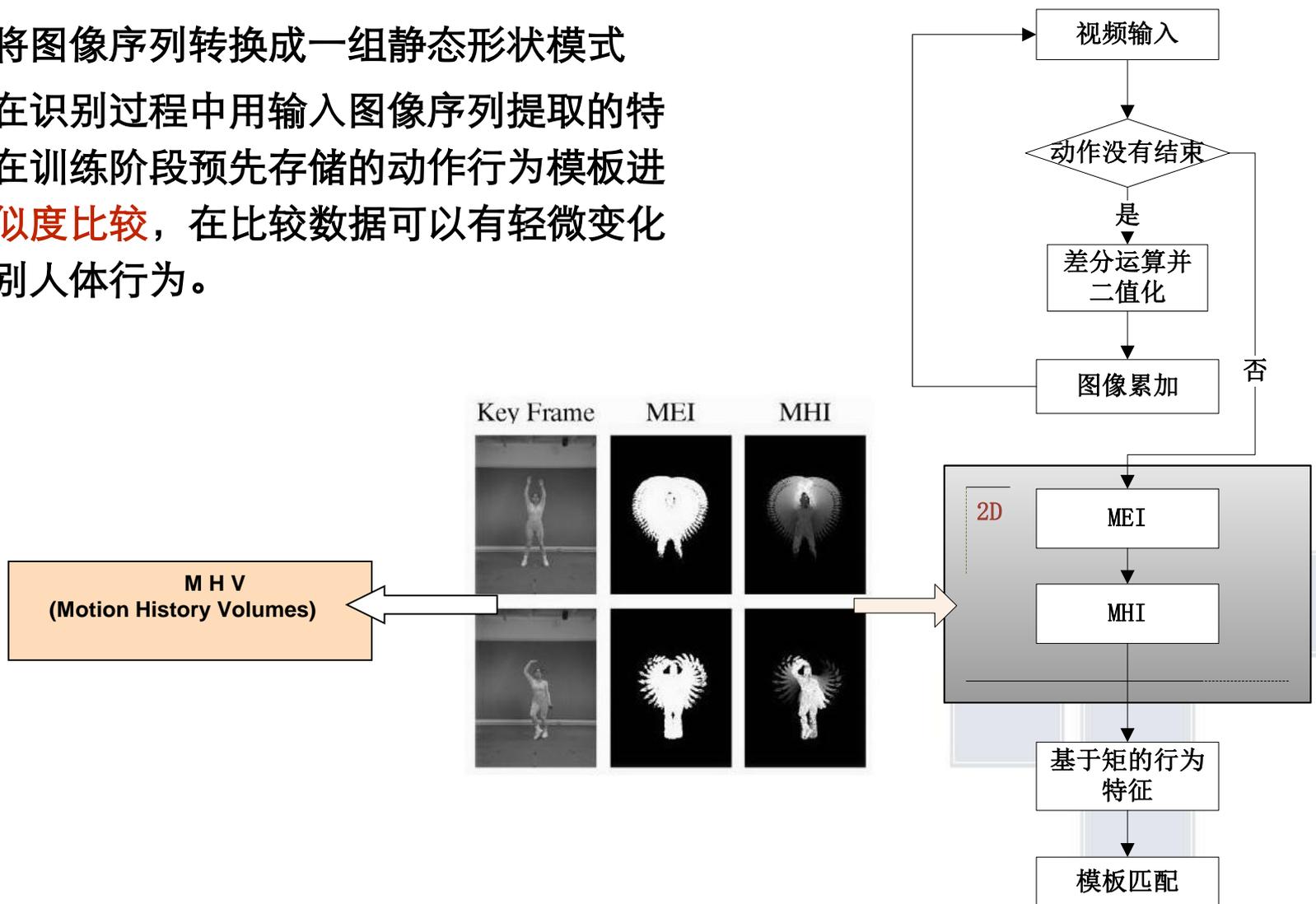
- 人的行为理解与描述是指对人的运动模式进行分析和识别,并用自然语言等加以描述.
- 可以简单地认为是时变数据的分类问题,即将测试序列与预先标定的代表典型行为的参考序列进行匹配.
- 行为理解的**关键问题**是如何从学习样本中**获取参考行为序列**,并且学习和匹配的行为序列能够处理在相似的运动模式类别中空间和时间尺度上轻微的特征变化.





# 基于模板匹配的方法

- 首先将图像序列转换成一组静态形状模式
- 然后在识别过程中用输入图像序列提取的特征与在训练阶段预先存储的动作行为模板进行**相似度比较**，在比较数据可以有轻微变化下识别人体行为。





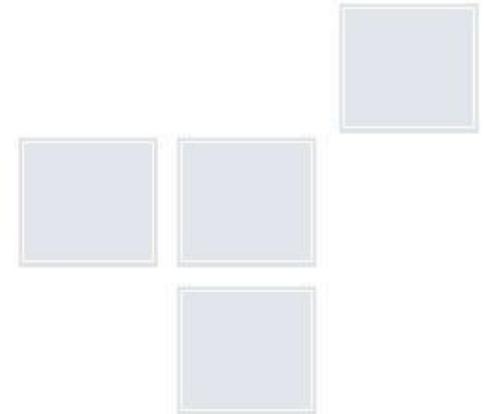
# 基于模板匹配的方法之DTW

- **DTW (dynamic time warping) ——动态时间规整**
- **DTW是一种时变数据序列匹配方法**
- **DTW 优点：概念简单、算法鲁棒，能够对图像序列进行分类。**
- **DTW缺点：算法计算量较大，缺乏考虑相邻时序之间的动态特性，而在实际中，运动序列中相邻序列在时间和空间上有高度的相关性。**

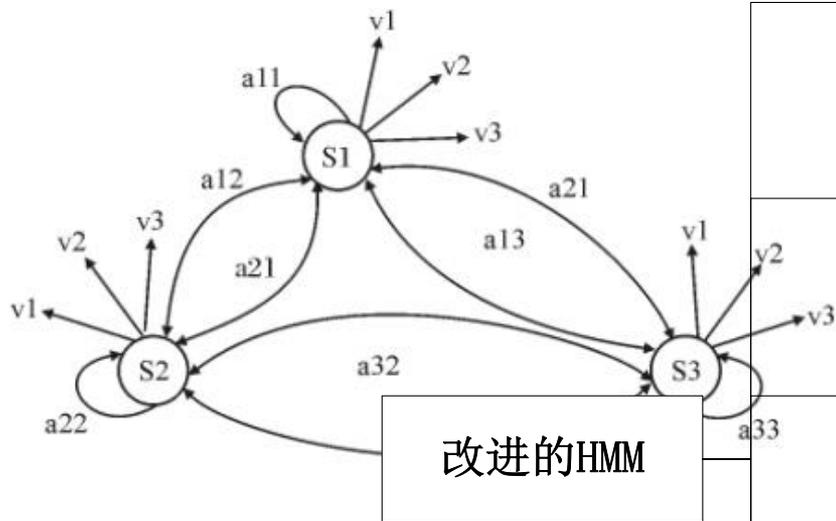




- 将图像序列中的每个**静态姿势或运动状态**作为一个**状态节点**，这些状态节点之间由给定的概率联系起来。
- 任何的**动作序列**可以认为是这些静态动作在不同状态节点中的一次**遍历过程**，计算这个遍历过程的联合概率，**取其最大值**作为分类标准。
- 基于状态空间的方法已经被广泛应用于**预测、估计和检测**时间序列。



# 基于状态空间的方法——HMM



隐马尔可夫模型的基本结构

耦合的HMMs  
(Coupled HMMs)

层级HMMs  
(hierarchical HMMs)

抽象HMMs  
(abstract HMMs)

可变长马尔科夫模型  
(Variable length Markov model)

熵隐马尔科夫模型  
(entropy HMM)

分层HMM  
(layered HMM)



## ■ 动态贝叶斯网络(dynamic Bayesian networks, DBNs)

由于HMMs不能有效处理三个或三个以上独立的过程，作为HMMs的推广方法。

■ 是一种对随机过程描述的有向图解

■ 利用先验知识建立视觉特征之间的因果关系来处理视频处理中固有的不确定性问题。

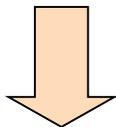
■ **优点：**

■ 可以任意改变拓扑结果或增删变量以反映变量间各种不同的关联关系，而不影响训练算法本身，因此具有**良好的可解释性**，其拓扑结构具有**精确及易于理解的概率语义**。

■ 相比马尔可夫链，贝叶斯网络训练比较复杂。



- 
- 人工神经网络是由人工建立的以有向图为拓扑结构的动态系统，它通过对连续或断续的输入作状态响应而进行信息处理，在分析时变数据时具有很大的优势。
  - 缺点：需要用大量的数据来训练网络

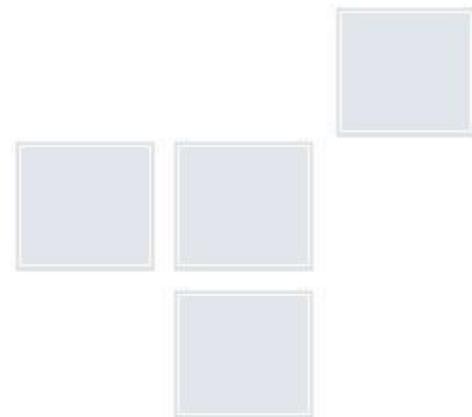


- **TDNN(time-delay neural network)——时延神经网络**
- 是在多层前馈感知器神经网络模型中引入时延单元使得神经网络增加记忆功能；
- 时变序列的前述值被用来预测下一个值。由于大量数据集成为可能，时延神经网络的重点就可以放在时分信息的表达
- 这样由此导出的神经网络模型适合应用于处理序列数据。



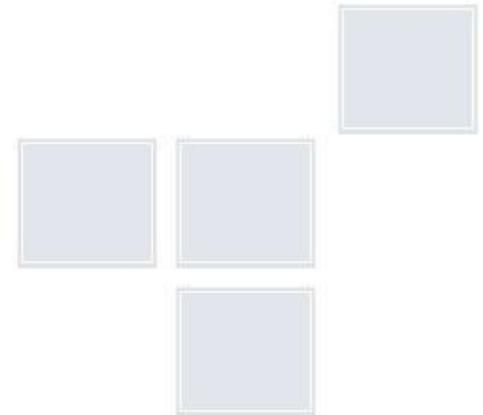
# 基于状态空间的方法——SVM

- 支持向量机(support vector machine, SVM)
- 优点:
- 能够较好地解决常见的非线性分类问题中的小样本、非线性、高维数和局部极小点等实际问题
- 可以避免神经网络结构选择和局部极小点问题



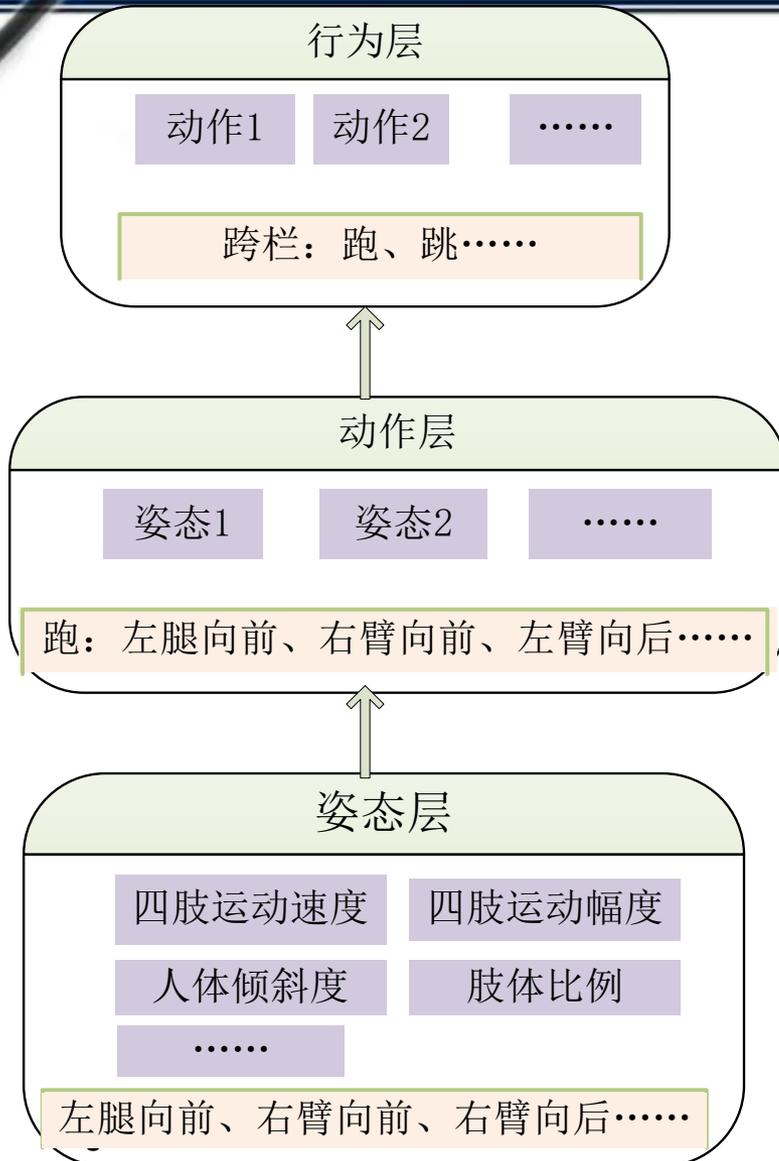


- 模板匹配方法的优点是计算复杂度低、操作和实现简单，但缺乏考虑运动序列中相邻时序之间的动态特性，对于噪声和运动时间间隔的变化相当敏感。
- 状态空间方法虽然能克服模板匹配的缺点，但通常涉及到复杂的迭代运算，算法的步骤较为复杂，难以应用到实际工作中。

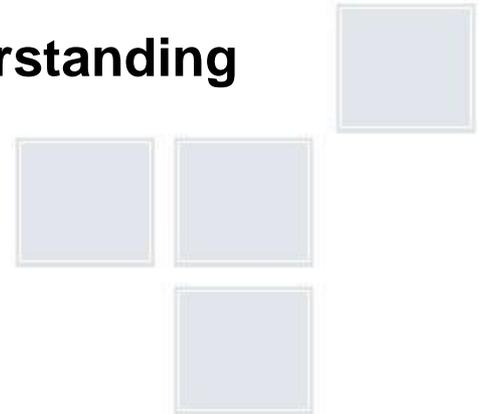


- 用一种形式化的语法格式：主语（人）、谓语（人的动作）、宾语（实物），将场景中人的活动情况用填空的形式填入语法格式中，产生对场景的自然语言描述。
- 基于语义描述的方法是对在一段持续时间内场景内容的分析过程
- 目前还处于对场景中人体行为的简单语义解释
- 对复杂场景中人体复杂行为有效的充分的语义描述还有相当艰巨的工作要做

# 行为的分层模型

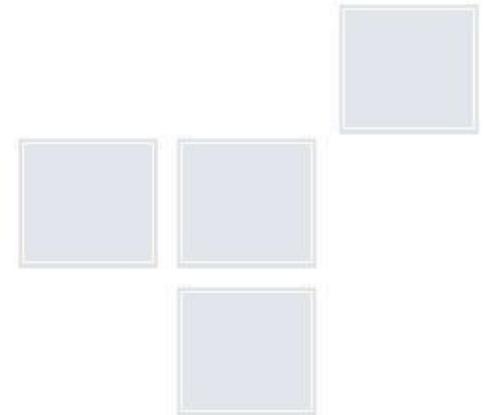


- 常用的概念
- pose、actions、activities
- action/motor primitives
- Human action recognition
- Motion analysis
- Action detection
- Automatic Target Recognition
- Behavior understanding
- visual analysis



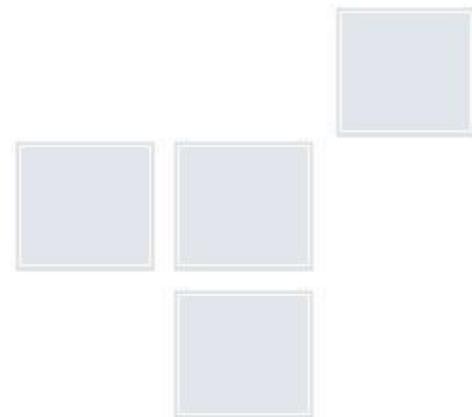


- 运动分割 (Motion Segmentation)
- 人体建模
- 遮挡问题
- 多摄像机的使用
- 运动特征选择与表达
- 行为识别
- 高层行为与场景理解
- 性能评估



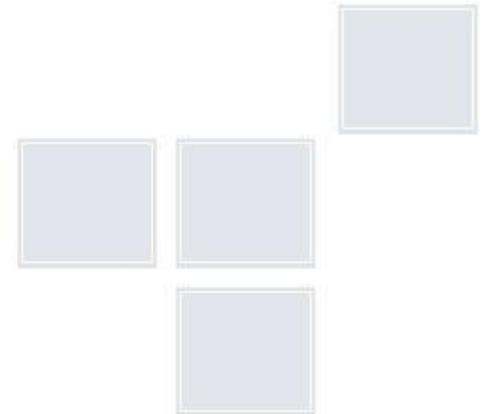


- 音频与视觉相结合的多模态接口
- 行为理解与生物特征识别相结合
- 人的运动分析向行为理解与描述高层处理的转变





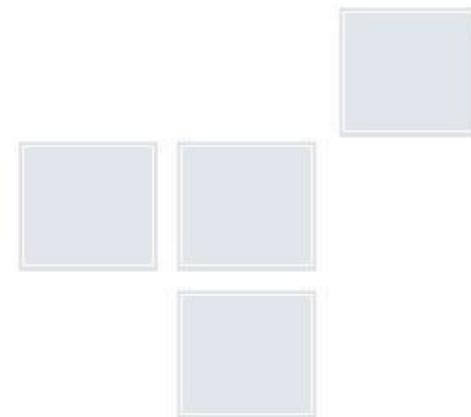
- **KTH human motion dataset**
- **Weizmann human action dataset**
- **INRIA XMAS multi-view dataset**
- **UCF sports action dataset**
- **Hollywood human action dataset**
- **CASIA**



- 
- **ICCV(IEEE International Conference On Computer Vision)**
  - **ICPR(IEEE Conference on Computer Vision and Pattern Recognition)**
  - **ECCV(European Conference on Computer Vision)**
  - **IJCV(International Journal of Computer Vision)**
  - **PAMI(IEEE Transactions on Pattern Analysis and Machine Intelligence)**
  - **IEEE International Workshop Performance Evaluation of Tracking and Surveillance**每年还提供PETS数据库用来评估智能监控系统的性能。
  - **智能视觉监控学术会议**
  - **模式识别会议等**
- 



- **A survey of advances in vision-based human motion capture and analysis**
- **A survey on vision-based human action recognition**
- **Action Recognition in Videos from Motion Capture Labs to the Web**
- 监控视频中的人体异常行为检测研究
- 视频中的人体运动分析及其应用研究
- 自动目标识别与跟踪技术研究综述
- 运动目标跟踪算法研究综述





谢谢