视觉注意预测模型

视知觉处理机制中包含两种处理方式

* 自底向上
  + 当用户没有任何任务和目标时
  + 倾向于受视觉内容本身的激励来进行观察注意
  + 符合Von Helmholtz的哪里假设和Broadbent的选择性滤波理论
* 自顶向下
  + 用户有任务、担心、希望或者习惯
  + 引入大脑的认知处理过程，有策略地分析场景
  + 符合Deutsch J A和Deutsch D的权重理论

Bottom-UP视觉注意模型

* 从视觉激励开始的信息处理过程
  + 先对较小的知觉单元进行分析，然后转向较大的知觉单元，经过一系列连续阶段的信息处理而达到对视觉激励的解释
    - 视觉显著性计算 Saliency Map
* 用于视觉显著性计算的特征
  + 颜色
  + 亮度梯度
  + 方向梯度
  + 深度
  + 闪烁
  + 运动
* 自底向上的视觉信息处理
  + 又称为数据驱动的处理（Data-driven processing）
  + 特点
    - 数据驱动，自动加工
    - 可以并行处理（由于只处理局部信息）
    - 处理速度快（不进行语义等高级分析）
* 图像显著度
  + 简单定义：一个像素与其他像素的颜色对比度
  + 要精确地描述人眼的注意的还不够
* Itti等人提出基于显著度的视觉注意模型
  + 用于自然场景图像的场景分析
  + 基于局部对比度计算显著性
    - 颜色、强度梯度、方向梯度
    - 闪烁、运动
    - 视觉抑制
  + Center-Surround交叉尺度混合，符合视网膜成像机制
* Feng的基于LEB的RGB-D视觉注意模型
  + 直接从深度信息中测量出显著结构
  + 显著对象在深度上一般都比它的周边区域更靠前

Top-down视觉注意预测模型

* 从相关知觉对象的经验知识开始
  + 有知识经验形成对知觉对象的期望或假设
  + 这种期望或假设制约着视觉认知过程中各个阶段的信息处理过程
  + 又称为任务驱动的处理（Task driven processing）
  + Top-down control如同一个“显著性激励”函数重新加权和整合输入图像的视觉刺激
* Torralba的上下文引导的眼动和视觉注意模型
  + 分为两部分
  + 在图像显著度的基础上加上任务驱动的部分
* Kanan等人提出的SUN模型
  + 也是基于贝叶斯Bayesian机制的外观显著度视觉注意模型
  + 计算模型与Torralba模型计算的信息完全不同但是预测结果相似
  + 可以得到与人类相当的视觉搜索效率
  + 更接近人类的视觉
  + 计算机生成了可选文字:
    —log丆一．0+log丆=fzlC=1)+logp()=1=0
    Self-information:
    Bottom-upsaliency
    Loglikelihood:
    Top-downknowledge
    appearance
    Lexation《丌
    Top-downknowledge
    oftarget'sl'，口m
  + 包含三部分
    - 该点本身的显著度
    - 该点是否属于目标物
    - 该点是否处于目标位置
* Learning based top-down control
  + 采用训练的方法会更接近人类的视觉，因此建立在传统方法基础上的基于学习的方法取得了更还得预测效果
  + 引入了两个问题
    - 每个场景有不同的显眼目标，因此不能采用统一的视觉特征来处理不同的场景，要分别训练
    - 给定不同的bottom-up成分和top-down成分，用怎样的混合策略将二者结合
  + 例子
    - Jia Li等人提出的基于学习的多任务视频注意预测模型
  + Deep Visual Attention Prediction
    - 基于深度学习的视觉注意模型架构
      * 单流架构
      * 多流架构——多尺度输入
      * 旁路架构——防止信号随层数增加而衰减
      * bottom-up与top-down结合
      * 基于旁路的，在单个网络中整合多尺度显著性信息

视觉注意预测模型分类

* 认知模型
* 贝叶斯模型
* 决策理论模型
* 信息理论模型
* 图模型
* 光谱分析模型
* 模式分类模型

视觉注意预测模型的意义

* 复杂场景中的视觉信息过滤
  + 较少视觉干扰
* 视域范围特征区域的自动增强
  + 引导用户注意、重要区域的视觉提示
* 自动场景特征发现
  + 辅助目标识别分类
* 视频中的动态特征发现
  + 自动报警，辅助跟踪识别
* 目标快速搜索识别以及注意力引导
  + 辅助驾驶、辅助作战
* 多目标并行搜索识别
  + 帮助用户高效搜索多个注视目标
* 符合用户任务目标的态势识别，重点区域提示
  + 火灾救援人员的视觉辅助增强
* 提高实时交互效率
  + 辅助提高眼动跟踪效率
  + 提高图像渲染效率

思考：计算视觉显著度low-level feature选取

|  |  |
| --- | --- |
| 亮度梯度 | 有比直接求RGB的平均值更合理的方法吗  人眼对不同颜色光的敏感程度不同，不能简单的平均 |