Here's a breakdown to address each part of your question:

以下是解决您问题的每个部分的细分：

---

3 (a) Differences between One-Stage and Two-Stage Detectors

3 (a) 一级检测器和两级检测器之间的差异

One-stage and two-stage detectors are both used in object detection tasks, but they differ in how they approach detection:

一级和两级检测器都用于物体检测任务，但它们的检测方式有所不同：

1. **One-Stage Detectors**: These detectors (e.g., YOLO, SSD) perform object classification and localization in a single step. They apply a dense set of predictions across the image and are typically faster because they directly predict the bounding boxes and class scores, making them more suitable for real-time applications.

   单级检测器：这些检测器（例如 YOLO、SSD）一步执行对象分类和定位。它们在图像上应用一组密集的预测，并且通常速度更快，因为它们直接预测边界框和类别分数，使它们更适合实时应用。

2. **Two-Stage Detectors**: These detectors (e.g., Faster R-CNN) work in two phases. In the first stage, they generate a set of region proposals likely to contain objects. In the second stage, these regions are classified and refined. Two-stage detectors are often more accurate due to this selective processing but are generally slower than one-stage detectors.

   两阶段检测器：这些检测器（例如 Faster R-CNN）分两个阶段工作。在第一阶段，他们生成一组可能包含对象的区域提案。第二阶段，对这些区域进行分类和细化。由于这种选择性处理，两级检测器通常更准确，但通常比一级检测器慢。

---

3 (b) Two-Stream Network for Human Action Recognition in Videos

3(b) 用于视频中人类动作识别的双流网络

A two-stream network for human action recognition processes video data through two separate pathways, commonly using spatial and temporal streams:

用于人类动作识别的双流网络通过两个独立的路径处理视频数据，通常使用空间流和时间流：

- **Spatial Stream**: This pathway takes a single frame (or a stack of frames) to capture spatial information (appearance). It often uses a convolutional neural network (CNN) to extract features related to the objects and scenes within individual frames.

  空间流：该路径采用单个帧（或一堆帧）来捕获空间信息（外观）。它通常使用卷积神经网络（CNN）来提取与各个帧内的对象和场景相关的特征。

- **Temporal Stream**: This pathway uses optical flow between consecutive frames to capture motion information. By analyzing the temporal changes, it helps recognize actions based on movement patterns over time.

  时间流：该路径使用连续帧之间的光流来捕获运动信息。通过分析时间变化，它有助于根据一段时间内的运动模式识别动作。

Diagram:

[Draw a simple diagram with two streams (arrows) labeled "Spatial Stream" and "Temporal Stream," each converging into a final "Fusion" node for classification.]

[绘制一个简单的图表，其中两个流（箭头）标记为"空间流"和"时间流"，每个流汇聚成最终的"融合"节点以进行分类。]

- **Fusion Layer**: The outputs of the two streams are combined, usually by fusion methods like averaging or concatenation, to produce the final action prediction.

  融合层：两个流的输出通常通过平均或串联等融合方法进行组合，以产生最终的动作预测。

---

3 (c) Approaches to Stereo Matching3 (c) 立体匹配方法

Stereo matching techniques generally fall into two categories:立体匹配技术通常分为两类：

1. **Local (Window-Based) Methods**: These methods use a small window around each pixel to compare pixels between stereo images, minimizing a cost function to find matches. They are computationally efficient and suitable for real-time applications. However, they struggle in low-texture or repetitive regions.

   局部（基于窗口）方法：这些方法使用每个像素周围的小窗口来比较立体图像之间的像素，从而最小化寻找匹配的成本函数。它们计算效率高，适合实时应用。然而，它们在低纹理或重复区域中挣扎。

2. **Global Methods**: These methods consider the entire image when matching, optimizing an energy function globally across all pixels. They produce smoother results and handle ambiguous regions better, but are computationally expensive, making them slower than local methods.

   全局方法：这些方法在匹配时考虑整个图像，在所有像素上全局优化能量函数。它们产生更平滑的结果并更好地处理模糊区域，但计算量大，使得它们比本地方法慢。

---

3 (d) Derivation of the Brightness Constancy Equation

3(d) 亮度恒常性方程的推导

Given:

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x, y, t)$$

1. **Expand Using Taylor Series**: Assuming Δx, Δy, and Δt are small, apply the Taylor series expansion to the left side:

   使用泰勒级数展开：假设 Δx , Δy ， 和 Δt 很小，将泰勒级数展开到左侧：

$$I(x+\Delta x, y+\Delta y, t+\Delta t) \approx I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t$$

2. **Subtract** $I(x, y, t)$ from both sides:**减去** $I(x, y, t)$ 从双方来看：

$$0 \approx \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t$$

3. **Divide by** Δt and let $\Delta x/\Delta t = u$ and $\Delta y/\Delta t = v$ (the optical flow components):

   除以 Δt 并让 $\Delta x/\Delta t = u$ 和 $\Delta y/\Delta t = v$ （光流组件）：

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0$$

This gives the **brightness constancy equation**:这给出了**亮度恒定方程**：

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0$$

This equation forms the basis for optical flow estimation in computer vision.

该方程构成了计算机视觉中光流估计的基础。