一、填空题

1. 图像灰度均值、方差

   - 图像的灰度平均值是平指灰度的平均水平。

   - 平均方差是衡量一个样本波动大小的量，对图像来说，平均方差反应的是图像高频部分的大小。方差小，则图片看着较暗；方差大，则图片看着较亮。

   - 灰度的均值代表图像的亮度。 方差代表其对比度

2. 采样间距、图像数据量、质量

   - 图像的取样率：单位距离的取样数目（在两个空间方向上）
   - 图像中可分辨的最小细节，主要由采样间隔值决定
   - 采样间隔值越小，空间分辨率越高

   > ◆ 一般来说，采样间隔越大，所得图像像素数越少，空间分辨率低，质量差，严重时出现马赛克效应；采样间隔越小，所得图像像素数越多，空间分辨率高，图像质量好，但数据量大。

3. 检测图像边缘的2种数学方法

   一阶微分算子和二阶微分算子。

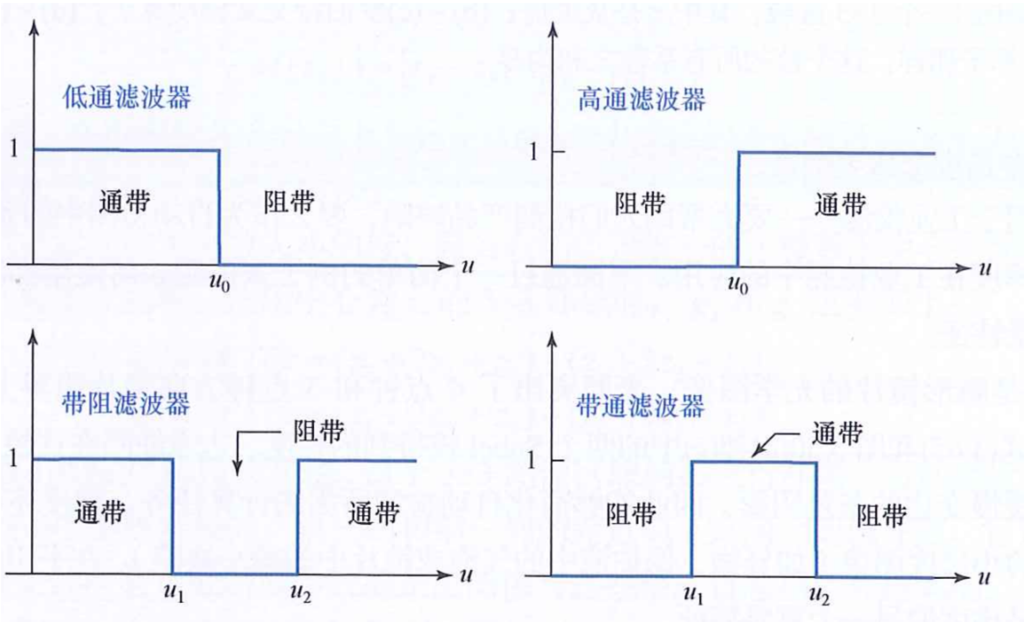
   > 一阶：Roberts算子以及在此基础上发展的Sobel算子[9,10]、Isotropic Sobel算子, Prewitt算子、Kirsch算子和Robinson算子等
   > 二阶：Laplacian算子, LOG算子, DOG算子和Canny算子等

   > 边缘检测：识别图像中的突然变化(不连续：表面法向量、深度、表面颜色、照明)
   > 边缘是图像强度函数变化迅速的地方
   > 梯度指向强度增长最快的方向

   > 补充学习： 一阶微分和二阶微分的区别:
   > (1)一阶微分处理通常会产生较宽的边缘
   > (2)二阶微分处理对细节有较强的响应,如细线和孤立点
   > (3)一阶微分处理一般对灰度阶梯有较强的响应
   > (4)二阶微分处理对灰度级阶梯变化产生双响应
   > (5)二阶微分在图像中灰度值变化相似时,对线的响应要比对阶梯强,且点比线强.
   > 　　大多数应用中,对图像增强来说.二阶微分处理比一阶微分好,因为形成细节的能力强. 而一阶微分处理主要用于提取边缘.

4. (低通、高通)滤波器

   都是线性空间域滤波器（非线性：中值、最大值、最小值）



> 平滑滤波器（低通）
> 　　主要用途：钝化图像、去除噪音
> 锐化滤波器（高通）
> 　　主要用途：边缘增强、边缘提取
> 带通滤波器
> 　　主要用途：删除特定频率、增强中很少用

5. 灰度变换函数

   s = T(r)
   三种基本类型

   - ①线性的(正比或反比)
   - ②对数的(对数和反对数的)
   - ③幂次的(n次幂和n次方根变换)

   反转变换：$s = L - 1 - r$

   　适于处理增强嵌入于图像暗色区域的白色或灰色细节,特别是当黑色面积占主导地位时.

   对数变换：$s = c \log (1 + r)$

   　使一窄带低灰度输入图像映射为一宽带输出值.可以用于扩展图像中的暗像素.

   幂次变换：$s = c r ^ { r }$

   　幂次曲线中的γ值决定了是把输入窄带暗值映射到宽带输出值还是把输入窄带亮值映射到宽带输出.

6. 图像量化

   - 数字化坐标值称为取样
   - 数字化幅度值称为量化。

图像量化（和采样）分为均匀量化和非均匀量化。

> ◆ 量化等级越多，所得图像层次越丰富，灰度分辨率高，图像质量好，但数据量大；量化等级越少，图像层次欠丰富，灰度分辨率低，会出现假轮廓现象，图像质量变差，但数据量小。

7. 梯度下降法
    - 梯度下降法的计算过程就是沿梯度下降的方向求解极小值
    - 过梯度下降法并不能保证全局最优解，主要由目标函数的非凸性造成的,容易陷入局部最优解

8. SLAM

    SLAM:Simultaneous localization and mapping缩写,意为"同步定位与建图"

    主要用于解决机器人在未知环境运动时的定位与地图构建问题

    根据传感器不同分为：视觉SLAM、激光SLAM

9. ORB特征
    - 特征点的信息
        - 位置、大小、方向、评分等——关键点
        - 特征点周围的图像信息——描述子（Descriptor）

> 补充1：特征匹配 :通过描述子的差异判断哪些特征为同一个点

> 补充2：特征匹配之后，得到了特征点之间的对应关系
> 　　如果只有两个单目图像，得到2D-2D间的关系 ——对极几何
> 　　如果匹配的是帧和地图，得到3D-2D间的关系 ——PnP
> 　　如果匹配的是RGB-D图，得到3D-3D间的关系 ——ICP

10. 外参

    对于世界坐标系的旋转和平移。

> 两像机间的相对位置和方向。
> 旋转矩阵R和平移矢量T

11. VO（Visual Odometry：视觉里程计）算法通常分为两类
    - 特征点法

- 直接法

12. SLAM后端优化

- 以EKF为代表的滤波方法
- 以图优化为代表的非线性优化方法

13. 畸变

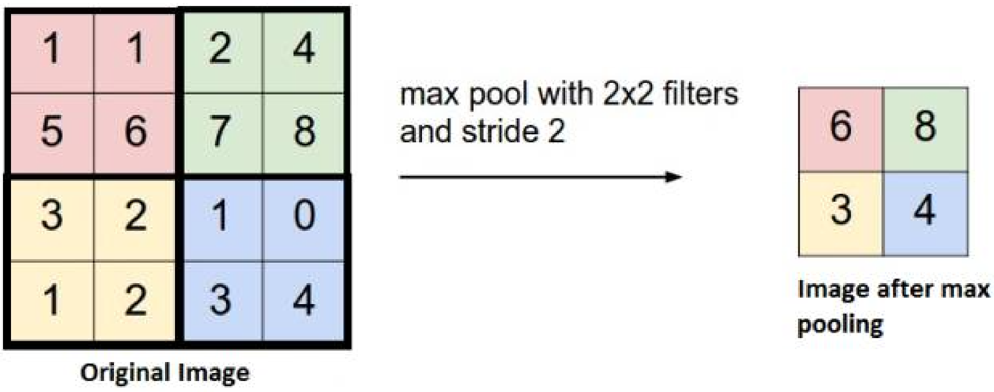
- 径向畸变：由透镜形状引起的畸变称为径向畸变，径向畸变主要分为桶形畸变和枕型畸变
- 切向畸变：在相机的组装过程中由于不能使透镜严格和成像平面平行，会引入切向畸变

14. 神经网络前向传播、反向传播

- 神经网络从输入到输出的计算过程叫做前向传播(Forward propagation)
- 利用误差反向传播算法进行反向计算的过程也叫反向传播(Backward propagation)。

15. 池化（汇聚）

- 最大池化
- 平均池化



16. 正则化

正则化

通过限制网络参数的稀疏性，可以来约束网络的实际容量

这种约束一般通过在损失函数上添加额外的参数稀疏性惩罚项实现

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \cdots + \beta_n x^n + \varepsilon.$$

多项式的项数越多，表达能力越强，增加包含多项式系数β的参数ε可以改变多项式的表达能力

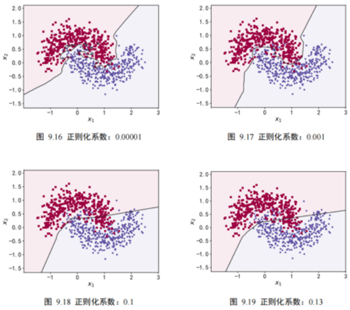


图 9.16 正则化系数：0.00001

图 9.17 正则化系数：0.001

图 9.18 正则化系数：0.1

图 9.19 正则化系数：0.13

$Minimize\ \mathcal{L}(f_\theta(\boldsymbol{x}), y), (\boldsymbol{x}, y) \in \mathbb{D}^{train}$ 原来的损失

$Minimize\ \mathcal{L}(f_\theta(\boldsymbol{x}), y) + \lambda * \Omega(\theta), (\boldsymbol{x}, y) \in \mathbb{D}^{train}$ 添加惩罚项后

其中Ω (θ)表示对网络参数θ 的稀疏性约束函数，一般，参数θ 的稀疏性约束通过约束参数θ (w,b)的L 范数实现

17. Dropout

    Dropout假设网络间连接越少越好

    我们在前向传播的时候，让某个神经元的激活值以一定的概率p停止工作，这样可以使模型泛化性更强，因为它不会太依赖某些局部的特征

    训练时，减少每次实际参与计算的模型的参数量

    测试时， Dropout会恢复所有的连接，保证模型测试时获得最好的性能。

    步骤

    1. 随机断掉网络部分隐藏神经元
    2. 修改后的网络前项传播，损失结果通过修改的网络反向传播(wb)
    3. 恢复被删删掉的神经元，重复这一过程

二、简答题

1. 图像锐化与图像平滑

    - 平滑处理：降低了图像的“尖锐”变化。

    > 均值滤波器、中值滤波器

    - 锐化处理：突出图像中的细节或者增强被模糊了的细节.

    > 空间微分。图像微分增强了边缘和其他突变(如噪声)而消弱了灰度变化缓慢的区域.

2. 电磁波谱成像(举例)

光 —— 可以被人眼感知的电磁波。（一般人的眼睛可以感知的电磁波的波长在400～760nm之间，但还有一些人能够感知到波长大约在380～780nm之间的电磁波）电磁波是能量的一种，任何有能量的物体，都会释放电磁波。人从物体感受的颜色由物体反射光决定。

- 若所有反射的可见光波长均衡，则物体显示白色
- 有颜色的物体是因为物体吸收了其他波长的大部分能量，从而反射某段波长范围的光。
- 没有颜色的光叫单色光或消色，灰度级通常用来描述单色光的强度，其范围从黑到灰，最后到白。
- 在原理上，如果可以开发出一种传感器，能够检测由一种电磁波谱发射的能量，就可以在那一段波长上对感兴趣的物体成像。

> 无线电波、红外线、可见光、紫外线、X射线、γ射线都是电磁波
> 举例：
> γ射线:核医学和天文观测。
> x射线：X光片、CAT扫描
> 紫外光成像：平板印刷术、工业检测、显微镜方法、激光、生物成像和天文观测等。
> 可见光及红外波段成像：最常见的方式。
> 微波成像：雷达
> 无线电波成像：核磁共振成像（MRI

3. 统计排列滤波器

属于空间平滑滤波器。

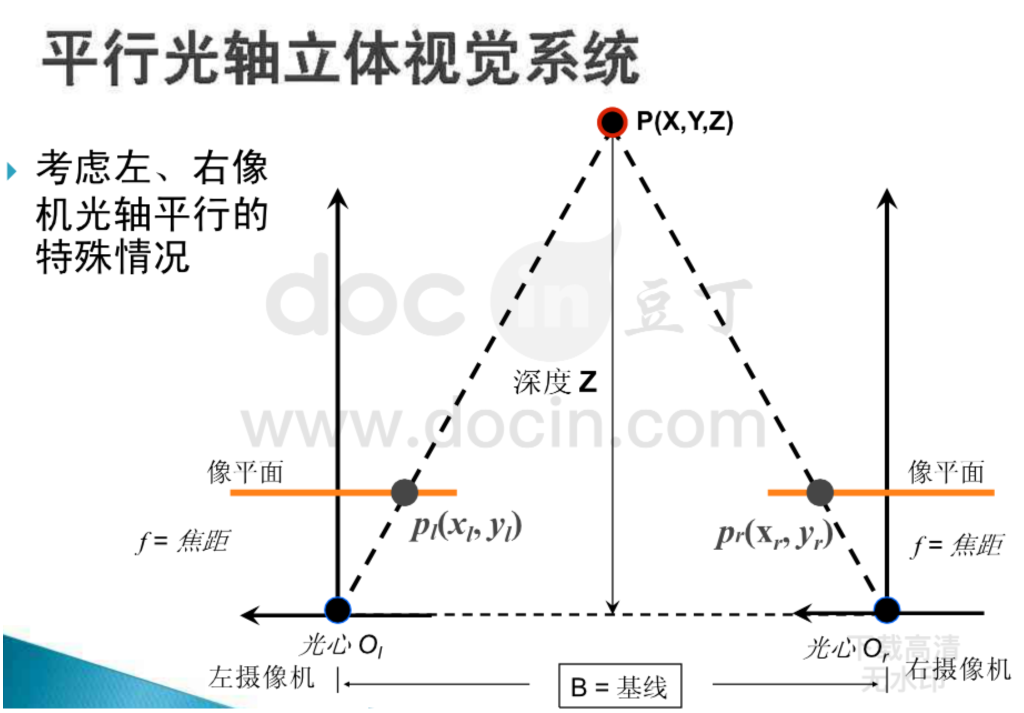
一种非线性滤波器,它的响应基于图像滤波器包围的图像区域中像素的排序,然后由统计排序结果决定的值代替中心像素的值. 最常见的是中值滤波器 。

中值滤波器:

- 算法：先将掩模内欲求的像素及其领域的像素值排序（升序或降序），确定出中值,并将中值赋予该像素点.
- 主要功能：使拥有不同灰度的点看起来更接近于它的邻近值。
- 主要用途：去除“椒盐”噪声（消除孤立的亮点、暗点）。可以较好地保留图像的细节。

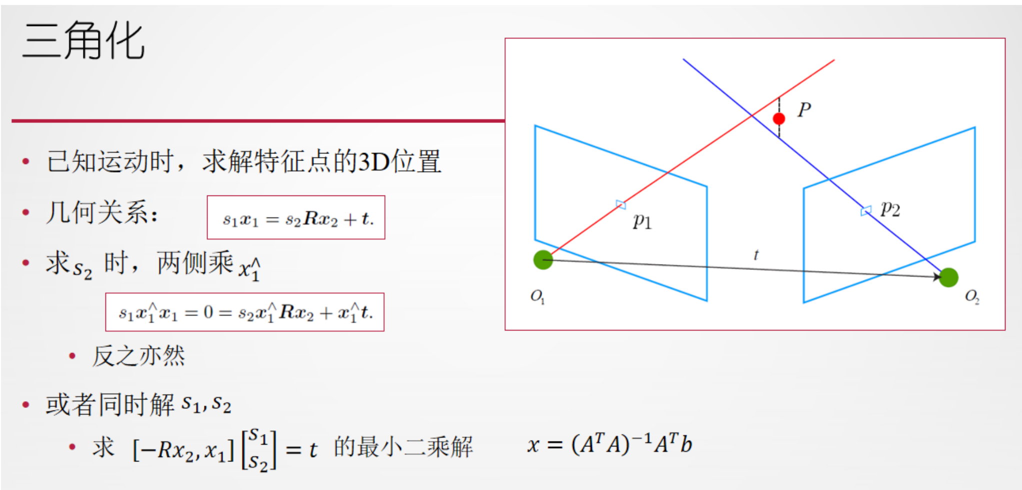
> 注：二维中值滤波的窗口形状和尺寸对滤波效果影响较大，不同的图像内容和不同的应用要求，往往采用不同的窗口形状和尺寸。常用的二维中值滤波窗口有线状、方形、圆形、十字形以及圆环形等。窗口尺寸一般先用3X3,再取5×5逐渐增大，直到滤波效果满意为止。就一般经验来讲，对于有缓变的较长轮廓线物体的图像，采用方形或圆形窗口为宜。对于包含有尖顶物体的图像，用十字形窗口，而窗口大小则以不超过图像中最小有效物体的尺寸为宜。如果图像中点、线、尖角细节较多，则不宜采用中值滤波

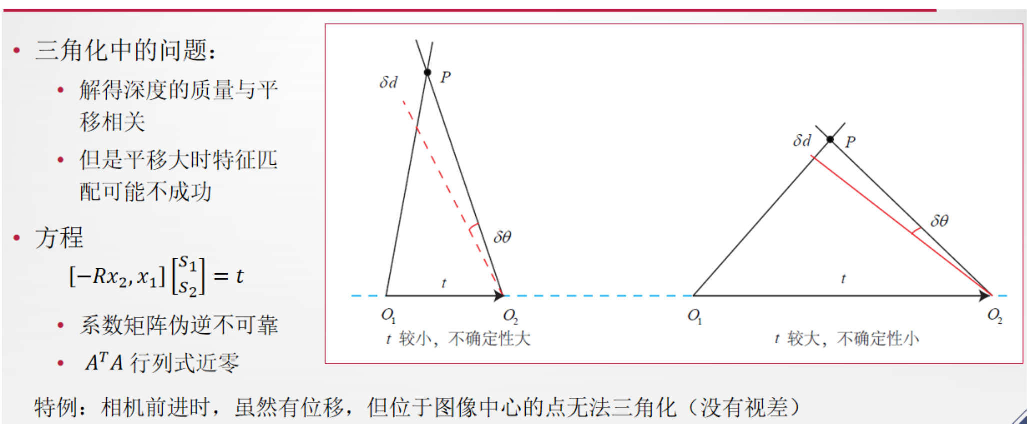
4. 立体视觉三角化（SLAM中，下面第二、三张图）



$$z = \frac { B \cdot f } { d }$$

其中：$d = x - x ^ { \prime }$为视差。





5. 迭代最近点ICP

ICP算法能够使不同的坐标下的点云数据合并到同一个坐标系统中，首先是找到一个可用的变换，配准操作实际是要找到从坐标系1到坐标系2的一个刚性变换。

ICP算法本质上是基于最小二乘法的最优配准方法。该算法重复进行选择对应关系点对， 计算最优刚体变换，直到满足正确配准的收敛精度要求。

ICP 算法的目的是要找到待配准点云数据与参考云数据之间的旋转参数R和平移参数 T，使得两点数据之间满足某种度量准则下的最优匹配

假设给两个三维点集 X1 和 X2，ICP方法的配准步骤如下：

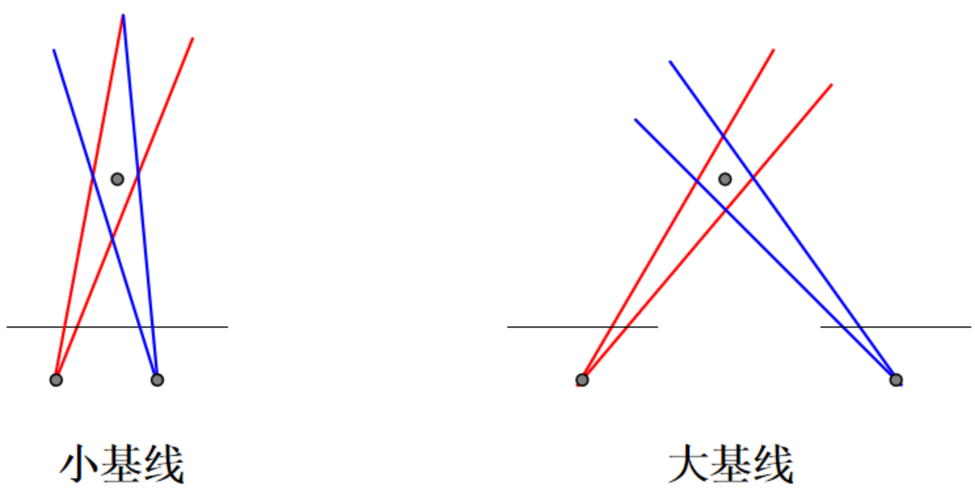
第一步，计算X2中的每一个点在X1 点集中的对应近点；

第二步，求得使上述对应点对平均距离最小的刚体变换，求得平移参数和旋转参数；

第三步，对X2使用上一步求得的平移和旋转参数，得到新的变换点集；

第四步， 如果新的变换点集与参考点集满足两点集的平均距离小于某一给定阈值，则停止迭代计算，否则新的变换点集作为新的X2继续迭代，直到达到目标函数的要求。

6. 立体视觉中大基线和小基线(短基线和长基线)
   - 基线：连接两个摄像机中心的线



短基线:

较大的公共视野区域;

深度误差（不确定区域）较大。

长基线

公共视野区域较小;

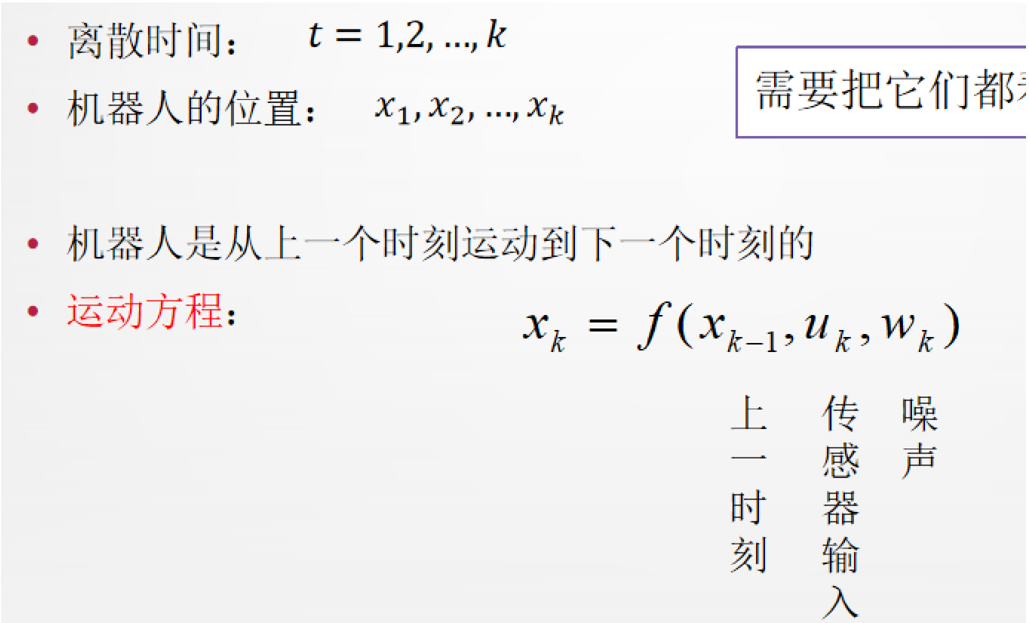
深度误差（不确定区域）较小。

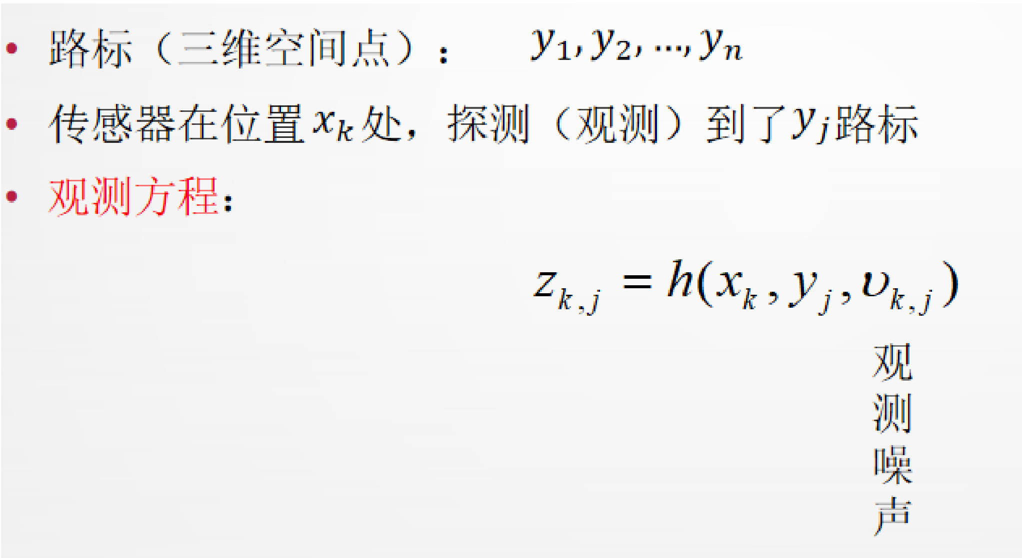
7. SLAM方程

   两个基本方程：

   运动方程：$x_{k}=f(x_{k-1},u_{k},w_{k})$

   观测方程：$z_{k,j}=h(x_{k},y_{j},v_{k,j})$

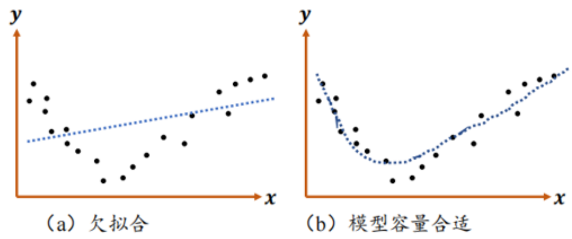




8. 过拟合或者欠拟合（神经网络）

欠拟合(Underfitting)

当模型的**容量过小**时，模型不能够很好的学习到训练集数据的模态，导致训练集上表现不佳，同时在未见的样本上表现也不佳。



具体表现

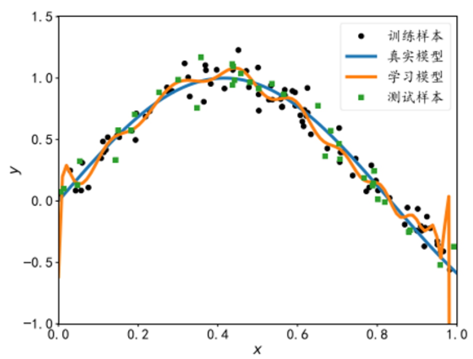
学习到的模型在训练集上的误差(如均方差)较大，同时在测试集上面的误差也较大。

解决办法

通过增加神经网络的层数、增大中间维度

过拟合(Overfitting)

当模型的容量过大时，网络模型除了学习到训练集数据的模态之外，还把额外的观测误差也学习进来，
导致学习的模型在训练集上面表现较好，但是在未见的样本上表现不佳，也就是泛化能力偏弱。



过拟合

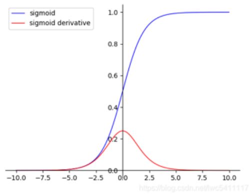
> 防止过拟合：正则化、添加动量、学习率衰减、早停、DropOut、数据增强

9. 激活函数（sigmoid、tanh优缺点）

激活函数：帮助网络去理解、学习非常复杂的非线性函数，输入产生怎样的响应

激活函数

sigmoid $f(x)=\frac{1}{1+e^{-x}}$



导数值：(0,0.25]

把$x \in R$ 的输入“压缩”到$x \in$ [0,1]区间

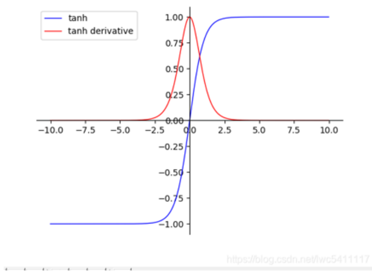
概率分布 [0,1]区间的输出和**概率**的分布范围契合，可以通过 Sigmoid 函数将输出转译为概率输出。

Sigmoid 函数**连续可导**，相对于阶跃函数，可以直接利用梯度下降算法优化网络参数，应用的非常广泛

缺点 Sigmoid 函数在输入值较大或较小时容易出现梯度值接近于 0 的现象，
称为**梯度弥散现象**，网络参数长时间得不到更新，很难训练较深层次的网络模型。

双曲正切函数
Tanh

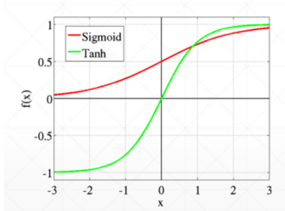
$$tanh(x)=\frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})}$$
$$=2*sigmoid(2x)-1$$



Tanh 函数能够将$x \in R$ 的输入“压缩”到[−1,1]区间

可以看到 tanh 激活函数可通过 Sigmoid 函数缩放平移后实现

导数值(0,1]



tanh(x) = sinh(x)/cosh(x) = (eˣ - e⁻ˣ)/(eˣ + e⁻ˣ)

Relu Sigmoid和Tanh易出现梯度弥散的现象

在 2001 年，神经科学家 Dayan 和 Abott 模拟得出更加精确的脑神经元激活模型，如下图所示

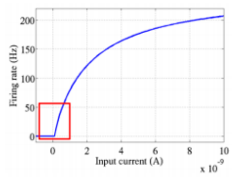


图 6.10 人脑激活模型 (Glorot, Bordes, & Bengio, 2011)

Relu

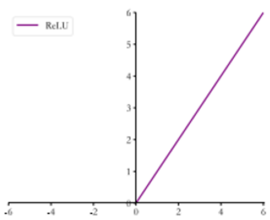
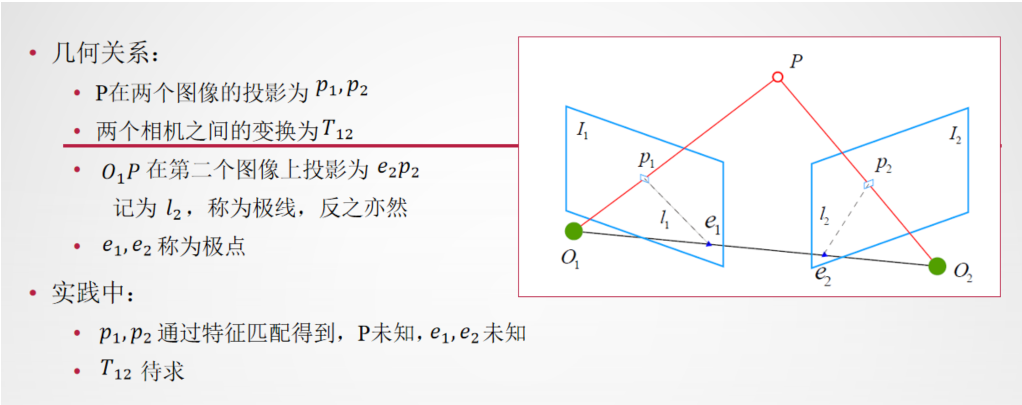


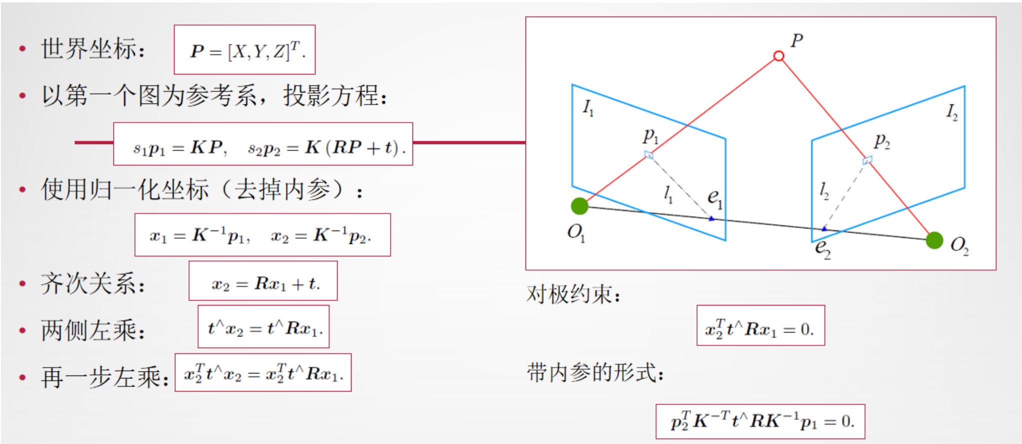
图 6.9 ReLU 激活函数

ReLU 对小于 0 的值全部抑制为 0；对于正数则直接输出，梯度恒为1

三、计算题

1. 推导对极约束方程





2. SLAM流程图（每一部分都是干什么）

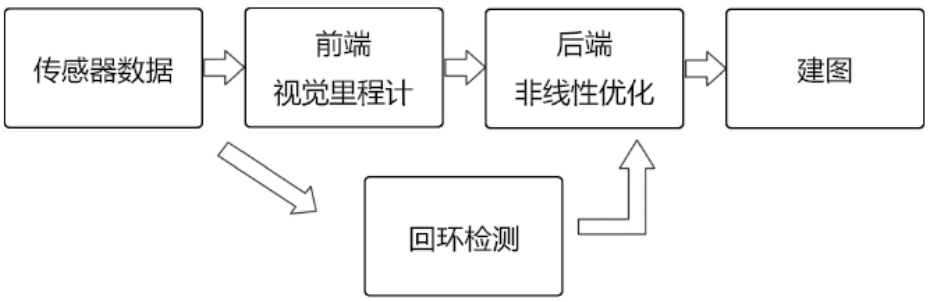


图 2-7 整体视觉 SLAM 流程图。

传感器数据读取：

- 在视觉SLAM中主要为相机信息的读取和预处理：获取图片或视频序列

视觉里程计：

- 相邻图像估计相机运动
- 基本形式：通过两张图像计算运动和结构
- 不可避免地有累积漂移

后端优化：

- 从带有噪声的数据中优化轨迹和地图 状态估计问题

回环检测：

- 检测机器人是否回到早先位置
- 识别到达过的场景
- 计算图像间的相似性
- 减小累积误差

建图：

- 用于导航、规划、通讯、可视化、交互等
- 度量地图 vs 拓扑地图
- 稀疏地图 稠密地图

3. SLAM非线性优化的图优化模型

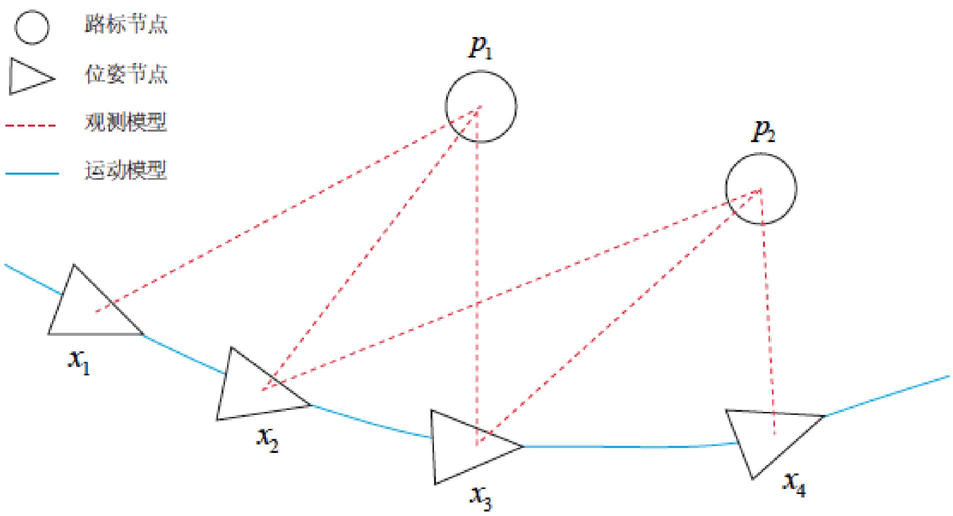
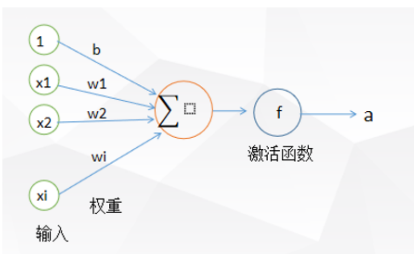


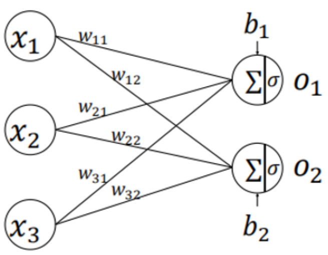
图 6-2 图优化的例子。

4. 简单全连接神经网络结构和参数量（3层）



$out = \boldsymbol{f}\ (X\ @W + b\ )$

两个输入$[x^1_1, x^1_2, x^1_3]$，$[x^2_1, x^2_2, x^2_3]$



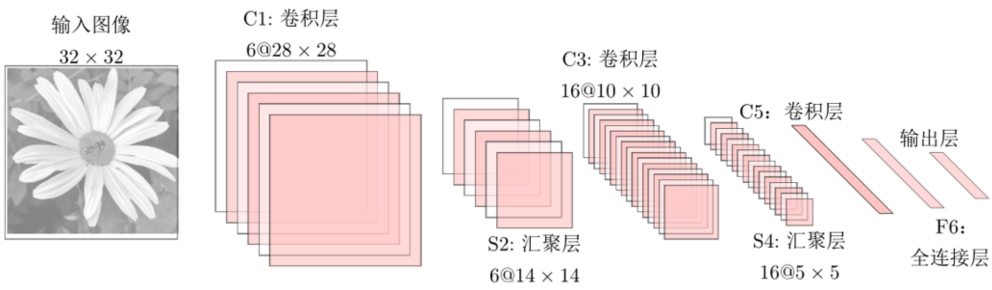
$$\begin{bmatrix} o_1^1 & o_2^1 \\ o_1^2 & o_2^2 \end{bmatrix} = \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 \\ x_1^2 & x_2^2 & x_3^2 \end{bmatrix} @ \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} + [b_0 \quad b_1]$$

我们把这种由**神经元**构成的网络叫做神经网络

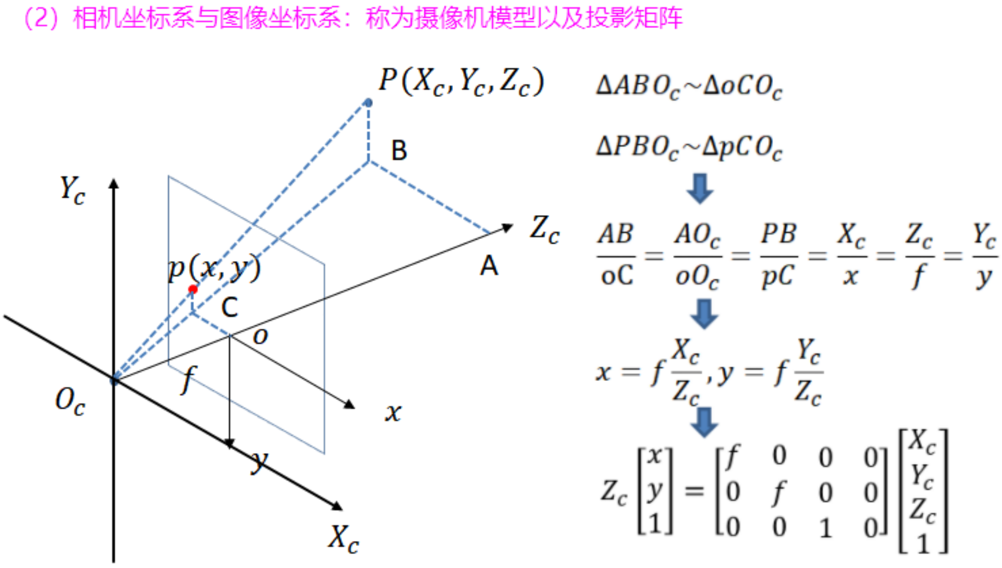
每个**输出**节点与每个**输入**节点连接，这种网络层称为**全连接层**(Fully-connected Layer)

5. LeNet-5

   典型的7层卷积神经网络。3个卷积核。卷——池——卷——池——卷——连——出



6. 投影方程（相机）



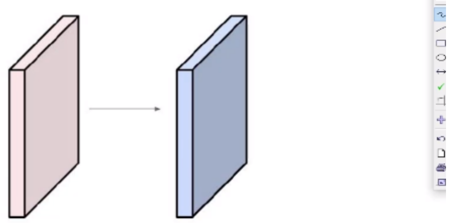
7. 卷积尺寸计算（输入输出尺寸）

   卷积结果计算公式：

   长度： $H_2 = \frac{H_1 - F_H + 2P}{S} + 1$

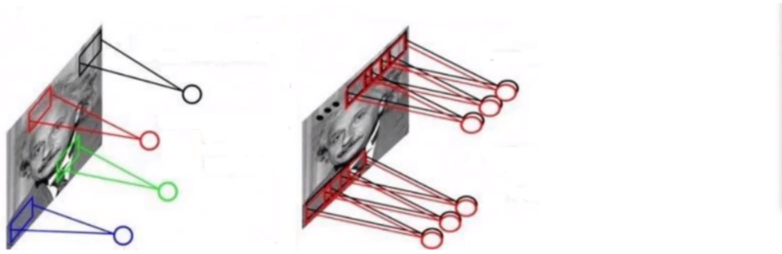
   宽度： $W_2 = \frac{W_1 - F_W + 2P}{S} + 1$

   其中W1、H1表示输入的宽度、长度；W2、H2表示输出特征图的宽度、长度；F表示卷积核长和宽的大小；S表示滑动窗口的步长;P表示边界填充(加几圈0)。



   我的补充：卷积参数共享

卷积参数共享：



数据依旧是32*32*3的图像，继续用10个5*5*3的filter来进行卷积操作，所需的权重参数有多少个呢？

5*5*3 = 75，表示每一个卷积核只需要75个参数，此时有10个不同的卷积核，就需要10*75 = 750个卷积核参数，不要忘记还有b参数，每个卷积核都有一个对应的偏置参数，最终只需要750+10=760个权重参数。

8. 简单卷积计算（矩阵、卷积核、步长→结果）