## Eigenface

Eigenface算法

N\*M矩阵（人脸图像）

x1…xm 为n维向量（人脸图像的每列）

1. 求所有m个n维向量的平均向量

图示

中度可信度描述已自动生成

1. 归一化，让均值为新的0点

图片包含 文本

描述已自动生成

1. 

计算协方差矩阵C =

1. 奇异值分解，计算出C的特征值与特征向量，其中特征向量为特征脸
2. 选定k，k<<N， N-k表示要降的维度数，选前k个最大的特征值对应的特征向量组成一张图，该图就是降维后模糊但保留特征的特征脸图

如何确定k

文本, 信件

描述已自动生成

0.9与0.95代表保留的信息量

特征脸算法缺点：

对于背景与变化（例：两脸重叠）变化敏感

文本

中度可信度描述已自动生成

Fisher face

基于LDA（Linear Discriminant Analysis，线性判别分析）

大部分与pca一样，区别在于，PCA以保留最多信息为前提，既选取一个主元，让其他维度到主元距离尽可能小

LDA选取一个主元，让其他类到主元上的投影能完全被分开

图示

描述已自动生成

## 3D vision

**相机中有四个坐标系，分别为{world}，{camera}，{image}，{pixel}**

{world}为世界坐标系

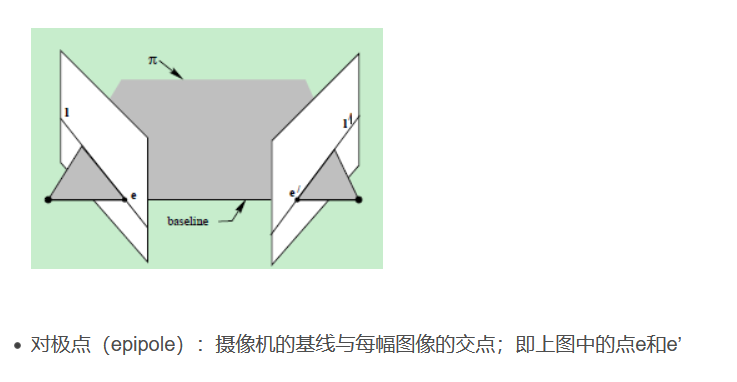
{camera}为相机坐标系，原点位于小孔，z轴与光轴重合

{image}为图像坐标系，原点位于光轴和投影面的交点

{world} -> {image}: x=k[ r | t ]X 图像坐标=内参[旋转|位移]世界坐标

相机定标（calibration）：求内参外参矩阵，内外参合一的矩阵可以将三维世界点投影到二维中，calibration matrix 3\*4 矩阵

Two-view Homography：相机不同角度拍摄同一物体，这两张照片遵循某种投影变换可以互相转换， homography 3\*3 矩阵

对极几何（epipolar geometry）：两视角相机看同一个三维世界点，fundamental矩阵（相机未定标）或essential矩阵（相机已经定标） 均为3\*3矩阵（只与相机外参有关p’Fp=0, p’Ep=0）

1. 对于求essential矩阵(**r+t**)，已知两相机内参矩阵k1,k2,对应点x1,x2

([K2^-1 @ x2).T @ E @ (k1 @ x1](mailto:K2%5e-1@x2).T@E@(k1@x1)) = 0

1. 对于求fundamental矩阵, 已知x1,x2

x2.T @ F @ x1 = 0

1. 三角测距（Triangulation通过图像上的点对算出真实的点的位置X

Stereo Vision（立体视觉）（生成物体到baseline的深度图，可能可以区分出不同物体）：1. 两图片点的对应问题 2. 重构问题

图片包含 游戏机, 电脑

描述已自动生成

光流（optical flow）：表示物体运动的速度，是三维运动场在二维的投影，可用于检测相机运动，物体形状，物体分割

**DLT 算法：**

求相机定标的DLT：

1. 找出六个及以上的点对，世界坐标X与图像坐标x
2. 求Tnorm，Snorm

图示

描述已自动生成

图片包含 文本

描述已自动生成

1. 构造2n\*12的A矩阵，n为点对个数，每个点对按以下构建两行，A矩阵为这些行摞起来

手机屏幕的截图

描述已自动生成

1. 对A做奇异值分解 U \* sigma \* V.T = A，p为V的最后一列
2. 将p reshape 成 3\*4矩阵，最终答案为 Tnorm ^ -1 @ p @ Snorm

分解相机定标矩阵得到k（内参矩阵），R（旋转矩阵），t（相机原点，平移）

1. 先求相机中心c：将p（相机定标矩阵）奇异值分解，p= U \* sigma \* V.T，c为V矩阵最后一列
2. P=[ M|-Mc ], M矩阵为p的前三列，为3\*3矩阵，对M做RQ分解得k矩阵与R矩阵

求homography 的DLT：

1. 找出4个及以上的点对，两张图片的对应点
2. 求Tnorm

黑色的钟表

中度可信度描述已自动生成

1. 把两个图片的点分别标准化

图片包含 图示

描述已自动生成

1. 构造2n\*9的A矩阵，n为点对个数，每个点对按以下构建两行，A矩阵为这些行摞起来

黑色的钟表

中度可信度描述已自动生成

1. 对A做奇异值分解 U \* sigma \* Vt = A，h为V的最后一列
2. 将h reshape 成 3\*3矩阵，最终答案为 Tnorm ^ -1 @ h @ Tnorm

## The epipolar geometry

求fundamental/essential矩阵的dlt，8点算法：

1. 找出8个及以上的点对，两张图片（可能是不同相机拍的）的对应点
2. 求Tnorm1 (), Tnorm2 ()
3. 把两个图片的点分别标准化

图片包含 图示

描述已自动生成

1. 构造n\*9的A矩阵，n为点对个数，A矩阵如下

文本, 信件

描述已自动生成

1. 对A.T @ A 做特征值分解，f为最小的特征值对应的特征向量
2. 把f reshape 成3\*3矩阵F1，对F1做奇异值分解，F1 = U @ sigma @ V.T
3. 将sigma（3\*3的对角矩阵）最后一行全置为0，得到新的sigma1
4. Fundamental矩阵F=Tnorm2.T @ (U @ sigma1 @ V.T) @ Tnorm1

用RANSAC去除离群点用于求fundamental矩阵：

Loop：

选8个最小样本点对

计算F矩阵

确定内点（inliers，要用的点）

Until 选择的内点占样本的比例大于95%或重复了太多次循环

用所有的这些内点计算F矩阵，这样很精确

分解essential矩阵

已知：

黑色的钟表

中度可信度描述已自动生成

钟表的特写

描述已自动生成

可求出tx与R

tx = U @ Z @ U.T, R = U @ W @ V.T 或 U @ W.T @ V.T

三角测距（Triangulation**通过图像上的点对算出真实的点的位置X**）dlt算法：

1. 已知 两相机的定标矩阵, , 照片上的对应点, ,
2. 构建矩阵A如下

电脑屏幕的照片

中度可信度描述已自动生成

1. 对A奇异值分解，A = U @ sigma @ V.T
2. V的最后一列为答案X，3d位置

Stereo Vision 基本算法：

1. Image Acquisition（图像采集）：用一个特别相机坐标（平行的两个位置）获取一个事物的两张图
2. Cameral modelling (相机建模)：知道两个相机的内参，知道两相机之间的关系（外参，旋转，平移）
3. Feature Extraction （特征提取）
4. Image Matching （图像匹配）
5. Depth Interpolation （深度插值）

Disparity Computation (视差计算)

图示

描述已自动生成

Z代表点PL到baseline的距离，d是uL-uR, f是焦距，B是baseline长度

Z = f \* B / (uL-uR)

两图像对应问题（Matching correlation windows，窗口对应，判断这两张图片的两个地方是不是在实际坐标中的同个地方）

因为用了对极几何以及两相机是平行摆放的，所以对应的两个地方必在同一条直线上，且该直线平行于图片坐标系的x坐标

手机屏幕截图

低可信度描述已自动生成

窗口用中心像素点来表示，窗口大小自定（尝试不同大小，选最好的），可能回影响结果

SSD error计算，两窗口中的所有像素差的平方和，选ssd error最小的就是对应点。

光流计算：

图片包含 应用程序

描述已自动生成

灰度恒等方程：

图片包含 图示

描述已自动生成

Ix, Iy, It为对I求的偏导，在计算机视觉里可为差值，u，v是未知数，（u,v）表示运动光流，一个方程两未知数，无法求解。

Lucas–Kanade optical flow algorithm（LK光流算法）：(对噪声不敏感)

假设在两临近的窗口中，运动（u,v）不变，窗口均为5\*5

图示

低可信度描述已自动生成

解方程如下x向量：

文本, 信件

描述已自动生成

A.T @ A 类似哈里斯角点检测里的窗口

Horn–Schunck Optical Flow Algorithm：

文本, 信件

描述已自动生成

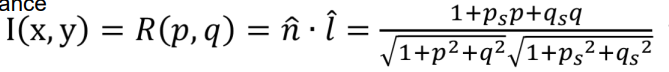
不用窗口，利用全局信息，求出平滑误差与灰度恒等误差，使Ec + lamda\*Es 最小，可求出运动（u,v）

3d重建（shape from x）

1. Shape from shading (阴影)
2. Shape from specular highlights（高光）
3. Shape from texture （纹理）

Lambertian reflection （兰伯特反射）目的求法线，从而3D重建

I = kd N dot L kd是反射系数，N是兰伯特面法线，L是入射光线， N，L为单位向量, I 是像素灰度



商店的招牌

中度可信度描述已自动生成

P,q 代表法线，ps，qs代表光源线

## Neural network

Activation function：sigmoid（映射到0到1之间），ReLU（映射到0，max），tanh（映射到-1到1之间）

手机屏幕截图

中度可信度描述已自动生成sigmoid导数 = s(x)(1-s(x))

文本, 信件

描述已自动生成

图像2分类与多分类，例子均为28\*28图像

2分类，把图像变成1\*784向量x，w是784个像素的权重，为784\*1的向量，b是bias

结果=x@w + b

多分类，还是x向量，w是每个像素每个感知机的权重，为784\*10矩阵，b是每个感知机的bias，10维向量

结果=x@w + b

神经元的输出：图示, 文本

描述已自动生成

权重\*输入的连求和加bias

多分类问题输出，softmax函数，每个数是每个分类的概率，加起来和为1

梯度下降，学习率：

黑色的钟表

低可信度描述已自动生成

CNN识别（recognition alexnet）与检测（detection yolo）的区别：

白板上写着字

描述已自动生成识别只是分类任务，输出是什么东西的概率，检测是输出每个boundingbox 是什么类的概率，且输出boundingbox的中心与这个boundingbox 有多大

反向传播例子