

第3章

机器视觉的工作内容



本章主要介绍机器视觉关于图像的基础理论知识,讲述图像的像素、分辨率、色彩空间,数字图像的获得,以及图像处理的一些相关算法,包括颜色处理、几何变换、图像分割与边缘检测等,为机器人应用打下一定的理论基础。

3.1 图像采集

3.1.1 图像

图像是人对视觉感知的物质再现。这个词汇是视觉技术的基础,在后文中会高频率出现。图像采集、图像处理和图像分析都是与图像相关的话题。

图像可以由光学设备获取,如相机、镜子、望远镜及显微镜等;也可以人为创作,如手工绘画。图像可以记录、保存在纸质媒介、胶片等对光敏感的介质上;而随着数字采集技术和信号处理理论的发展,越来越多的图像以数字形式存储。所以,在如本书这样单纯讨论视觉技术的情况下,“图像”一词实际上指数字图像。

图像分为静态图像(如图片、照片等)和动态图像(如视频等)两种。本节将对数字图像的3个参数——像素、分辨率和色彩空间展开介绍。

1. 像素

像素是图像显示的基本单位,通常被视为图像的最小完整采样。像素的英文单词是 pixel,其中 pix 为英语单词 picture 的常用简写,加上英语单词 element(元素),就得到了 pixel,故“像素”表示“图像元素”,有时也称为 pel(picture element)。每个像素都承载着图像中的信息,单从像素的概念来说,每个这样的信息元素并不能简单地看作一个点或者一个方块,而是一个信息的抽象采样;但是在很多情况下,它们采用点或者方块显示。每个像素的信息可以有各自的颜色值,采用三原色显示时,分成红、绿、蓝3种子像素,如图3-1所示。

一个更加直观的例子是最近兴起的“像素画”,可以将其看作放大的点阵式图像。将图3-2所示的每个小格看作一个像素,这些小格排列整齐,能够轻松构成图像矩阵,且每个

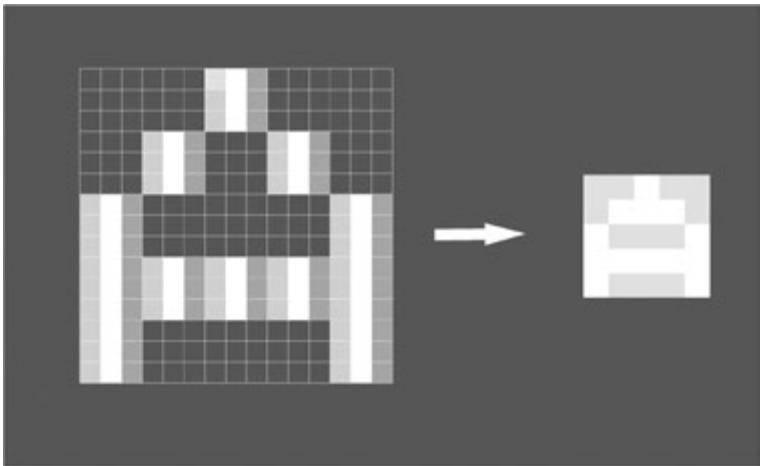


图 3-1 分成红、绿、蓝 3 种子像素显示字母 A

小格中都只能承载一种颜色信息。

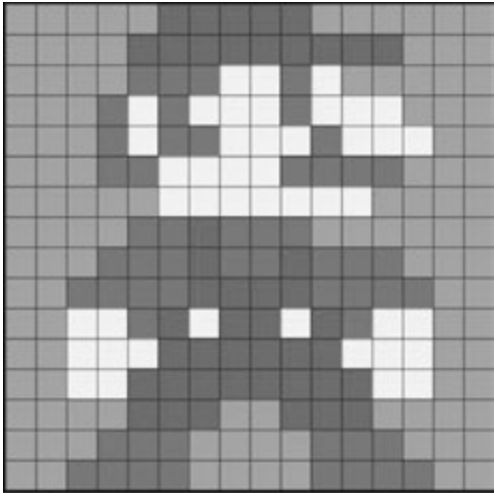


图 3-2 像素画

图像是一个个取样点的集合,单位面积内的像素越多代表分辨率越高,所显示的图像就会越接近真实物体。这也是选购相机时人们关注的相机参数中通常会有分辨率的原因。兆像素是指“一百万像素”,通常用于表示数码相机的分辨率。例如,一个相机可以使用 2048×1536 像素的分辨率,通常称其有“310 万像素”($2048 \times 1536 = 3145728$,通常只取前两位作有效数字)。有关分辨率的内容将在下文详细介绍。

2. 分辨率

分辨率泛指图像或显示系统对细节的分辨能力。日常用语中的分辨率多用于表示图像清晰度,分辨率越高,代表图像质量越好,越能表现出更多的细节。我们经常接触到的有屏幕分辨率和图像分辨率。

屏幕分辨率是屏幕每行的像素点数乘以每列的像素点数,每个屏幕都有自己的分辨率。屏幕分辨率越高,所呈现的色彩越多,清晰度越高。这个概念常常被用在计算机显示器、电

视、投影仪和手机上。如人们选购显示器时所说的分辨率 1920×1080 ，就是指这个显示器能够达到的最大分辨率。

图像分辨率指每英寸图像内的像素点数。分辨率越高，像素的点密度越高，图像越逼真；但相对地，因为记录的信息过多，文件也就会越大。这就是为什么做大幅的喷绘时，要求图片分辨率要尽可能高，就是为了保证每英寸的画面上拥有更多的像素点。图 3-3 所示为不同分辨率的图像差别。

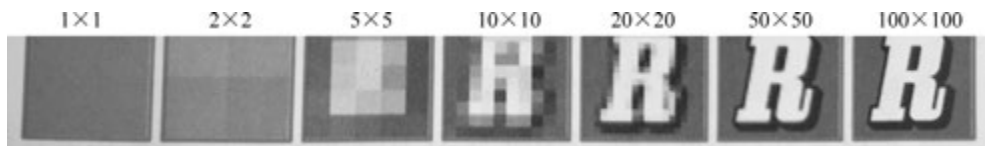


图 3-3 不同分辨率的图像差别(图中的 $N \times N$ 表示的是相邻边的像素数)

对于本书内容而言，更有研究价值的是图像分辨率。描述图像分辨率的单位有 DPI(点每英寸)、LPI(线每英寸)和 PPI(像素每英寸)。LPI 是描述光学分辨率的尺度的，虽然 DPI 和 PPI 也同属于分辨率范畴内的单位，但是它们的含义与 LPI 不同，而且 LPI 与 DPI 无法换算，只能凭经验估算。

另外，PPI 和 DPI 经常会出现混用现象，但是它们所用的领域也存在区别。从技术角度说，“像素”只存在于计算机显示领域，而“点”只出现于打印或印刷领域。很显然，适用于本书内容的分辨率单位是 PPI，其计算公式为

$$PPI = \sqrt{\frac{X^2 + Y^2}{Z}} \quad (3-1)$$

式中：X——长边像素数；

Y——宽边像素数；

Z——图像对角线长度。

从式(3-1)可以看出，如果保持图像的尺寸不变，将图像分辨率提高 1 倍，则图像的像素数变为原来的 4 倍，即图像文件的大小也变为原来的 4 倍。

3. 色彩空间

很多人都知道在绘画时可以使用红色、绿色和蓝色这 3 种原色生成不同的颜色，而这些颜色就定义了一个色彩空间。将红色的量定义为 X 坐标轴、绿色的量定义为 Y 坐标轴、蓝色的量定义为 Z 坐标轴，这样就得到一个三维空间，每种可能的颜色在这个三维空间中都有唯一的一个位置。

但是，这并不是唯一的一个色彩空间。例如，在计算机显示器上显示颜色时，通常使用 RGB(红色、黄色、蓝色)色彩空间进行定义，红色、绿色、蓝色被当作 X、Y 和 Z 坐标轴，如图 3-4 所示。也就是说，任意一种颜色在 RGB 空间中都能找到一个唯一的三维坐标值(R, G, B)。每个坐标值的范围都是 $[0, 255]$ ，由 8 位二进制数保存。

除了 RGB 色彩空间之外，常用的还有 HSB 色彩空间，其采用圆柱坐标系表示色相(Hue)、饱和度(Saturation)和明度(Brightness)，如图 3-5 所示。还有许多其他的色彩空间，可以按照类似方法用三维、更多或者更少维表示，也可能定义更多的颜色，如 Adobe RGB 色彩空间就比 RGB 色彩空间包含更多颜色。

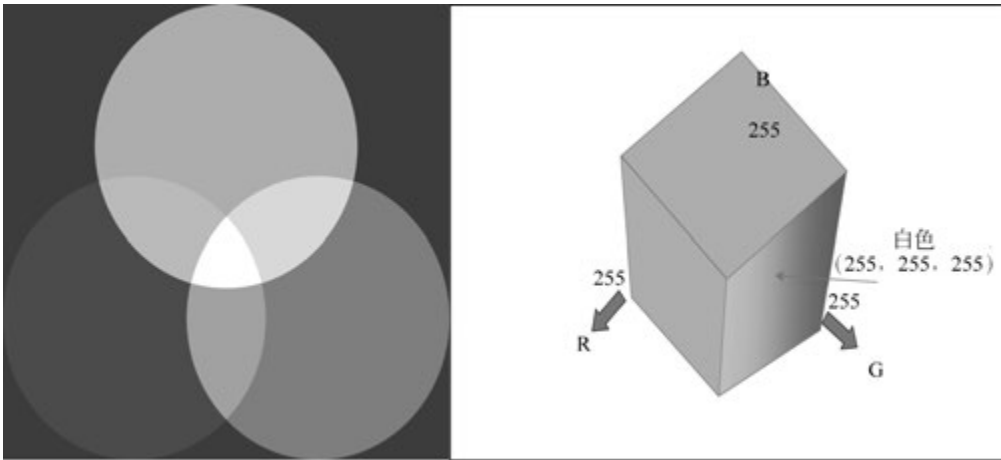


图 3-4 RGB 色彩空间

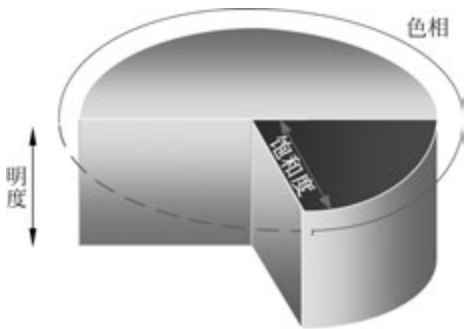


图 3-5 HSB 色彩空间

3.1.2 数字图像的获得

最初的工业相机获得的图像由连续的、有不同色彩及亮度等属性的颜色点组成,以模拟图像的形式存在。如果想要利用数字计算机处理模拟图像,就必须将模拟图像转换为用数字方式表示的数字图像文件,即数字图像。将模拟图像转换成数字图像的过程称为图像数字化过程,如图 3-6 所示。



图 3-6 图像数字化

数字图像的采集过程可以概括为:首先,光经过镜头聚焦在图像传感器上;然后,图像传感器将光信号转换为电信号,传输给 A/D 转换器,由 A/D 转换器完成数字化转换;最后,由图像采集卡输出到图像处理装置。图 3-7 所示为数字图像采集过程示意图。

A/D 转换器的数字化过程又可以细分为采样和量化两个步骤。采样就是将图像在空间上离散化,通常是在二维空间上连续的图像在水平和垂直方向上等间隔地分割成矩形的网状结构,进而得到若干被称为像素的方形小区域。经过采样后,图像在空间上离散为像素,但像素的颜色幅值仍为连续量。将采样所得的各像素的颜色幅值进行离散化的过程,称

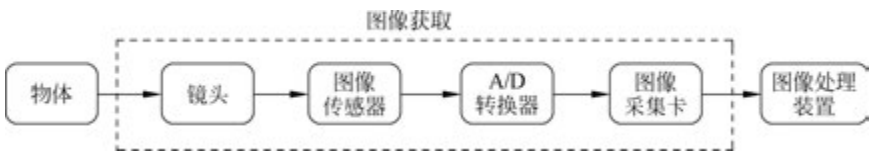


图 3-7 数字图像采集过程示意图

为量化。

经过采样和量化后,每幅图像都被离散化为一个个像素,成为一个由像素组成的二维矩阵,矩阵中的每个元素对应一个像素,如图 3-8 所示。每个像素里保存着该像素的颜色或亮度信息,其值可以为 RGB 图像的向量,也可以为灰度图像的 0~255 的某个数值,或者是二值图像的 0 或 1。

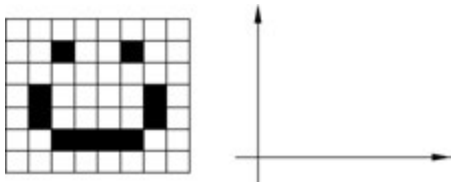


图 3-8 像素矩阵化的图像信息

3.2 图像处理

与其他类型的传感器提供的信息不同,视觉信息丰富多样,图像处理需要从图像中提取数字信息,从而得到用户对场景中感兴趣的目标。图像处理,就是对图像进行分析、加工和处理,使其满足视觉以及其他要求的技术。图像处理是信号处理在图像领域上的一个应用,目前大多数的图像是以数字形式存储的,因而图像处理在多数情况下指数字图像处理。图像处理是信号处理的子类,与计算机科学、人工智能等领域也有密切的关系。很多传统的一维信号处理的方法和概念仍然可以直接应用在图像处理上,如降噪、量化等。然而,图像属于二维信号,和一维信号相比,它有其特殊的一面,处理的方式和角度也有所不同。本节将介绍几种基础而典型的图像处理方式,方便大家理解图像处理都能做些什么,是如何方便后续图像分析的。

3.2.1 颜色处理

1. 灰度变换

灰度图像是每个像素只有一个采样颜色的图像,得到灰度图像的过程称为灰度变换。灰度值为像素光强弱信息的表示,是真实世界图像量化的表现方法。通常,灰度值从最黑到最白为 0~255。光线进入 CCD 感光元件,如果光强达到 CCD 感应的极限,此像素为纯白色,对应于内存中该像素灰度值为 255;如果完全没有光线进入 CDD,此像素为纯黑色,对应于内存中该像素灰度值为 0。

灰度图像的每个像素通常由 8 位二进制数来保存,这样就可以有 256 种灰度表示(0~255)。这种精度刚刚能够避免可见的条带失真,并且非常易于编程。在介绍 RGB 色彩空间

时提到过,每种颜色都可以由一个分别代表红、绿、蓝的三维向量表示,向量中每个分量都需要由 8 位存储,那么每个像素点需要 24 位,因此彩色图像的信息量较大。灰度图像与 RGB 图像相比,后续图像处理的计算量大大减少。

与黑白双色图像不同,灰度图像中的不同像素在黑色与白色之间还有许多级的颜色深度,这些颜色深度经常是在单个电磁波频谐(如可见光)内测量每个像素的亮度得到的。需要注意的是,在数字图像领域之外,“黑白图像”也表示“灰度图像”,就如灰度的照片通常也称为“黑白照片”。

这里介绍一下数字图像处理领域一幅被广泛使用的标准图片——Lena(Lenna),下文中的很多图像对比都将使用此图像作为标准图像。图 3-9 所示为标准图像与灰度图像的对比。



图 3-9 标准图像与灰度图像的对比

在医学图像与遥感图像等技术应用中,经常采用更多的存储级数以充分利用每个采样 10 位或 12 位的传感器精度,并且避免计算时的近似误差。在这样的应用领域流行使用 16 位即 65536 个灰度(或 65536 种颜色)。

2. 二值化

二值化是进行图像分割前经常使用的一种颜色处理方法。在二值化图像时,会把大于某个临界灰度值(阈值)的像素灰度设为灰度极大值(255,即白色),把小于这个值的像素灰度设为灰度极小值(0,即黑色),从而实现二值化。

根据阈值选取的不同,二值化的算法分为固定阈值和自适应阈值。比较常用的二值化方法有双峰法、P 参数法、迭代法和 OTSU 法(最大类间方差法)等。

二值图像是图像二值化的结果,每个像素只有黑色与白色两个可能值而没有中间颜色。也就是说,二值图像是一个对比清晰、非黑即白的单色图像,如图 3-10 所示。人们经常用黑白、B&W、单色图像表示二值图像。但其他领域也用上述这些词汇来表示每个像素只有一个采样值的任何图像,如灰度图像等。

二值图像经常出现在数字图像处理中,使用位图格式存储。一些输入/输出设备,如激光打印机、传真机、单色计算机显示器等都可以处理二值图像。

3.2.2 几何变换

几何变换是指从具有几何结构的集合至其自身或其他此类集合的一种映射。具体来说,几何变换是一个函数。



图 3-10 彩度图像与二值图像的对比

几何变换可以按照其操作集合的维度来分类(可分为平面变换与空间变换),也可依据其保留的性质来分类,如图 3-11 所示。

平移: 保留距离与方向角度。

欧氏: 保留距离与角度。

相似: 保留距离间的比例。

仿射: 保留平行关系。

投影: 保留共线性。

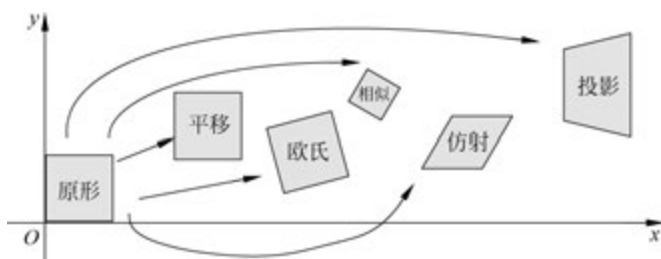


图 3-11 各种几何变换

3.2.3 图像分割

1. 图像分割的定义

多年来,人们提出了很多种图像分割的定义和表述,下面结合集合概念给出图像分割的定义。

假设集合 R 表示整个图像区域,对 R 的分割定义为将 R 分割成满足以下条件的 N 个非空子集 $R_1, R_2, R_3, \dots, R_N$:

- (1) $\bigcup_{i=1}^N R_i = R$;
- (2) 对所有的 $i \neq j$, 有 $R_i \cap R_j = \emptyset$;
- (3) 对 $i=1, 2, \dots, N$, 有 $R_i \cap R_j = \emptyset$;
- (4) 对 $i \neq j$, 有 $P(R_i \cup R_j) = \text{False}$;
- (5) 对 $i=1, 2, \dots, N$, R_i 是连通的区域。

其中, $P(R_i)$ 是对每个集合 R_i 中元素的逻辑谓词, $\emptyset$ 表示空集。

简单地说,图像分割技术是将相似性质的区域按照某种或某些性质分割到一起,性质区别大的被分割到不同的区域。

在进行图像分割时,一定要考虑以下几点。

- (1) 分割后前景的边界是封闭的。
- (2) 有效的分割应该做到分割时去掉无关紧要的细节,保留所关心的特征。
- (3) 有一个合适的计算速度。
- (4) 非阈值化,能否进行自动阈值分割。

对图像分割的研究主要可以分为两方面,一方面是图像分割算法的研究,另一方面是对分割评价方法和评价体系的研究。

2. 图像分割的意义

在图像处理技术中,图像分割是其中一个关键步骤。人们在研究或应用图像时,很可能所感兴趣的仅是图像中的某些部分或区域,这些区域通常称为前景(目标),剩余部分称为背景。前景往往具有一些独特的性质,为了能够更好地辨别和分析前景,最直接的想法就是将它们从图像中分离提取出来,之后才有可能对前景有更进一步的利用和处理。图像分割就是指根据各个区域的特性,如像素的颜色、纹理、灰度等,把图像分成多个区域并将感兴趣的目标提取出来的技术和过程。这里所指的目标可以是多个区域,也可以是单个区域。

3. 图像分割的算法

来自不同领域的图像有很大差别,人们对图像中感兴趣的部分也各不相同,所以图像分割至今没有一个通用的算法,现有的分割算法基本都是针对某些具体问题。

下面详细介绍几种最常用的算法。

1) 阈值化分割算法

阈值化分割算法是较早出现的图像分割算法。

该算法最直接的想法是根据图像的灰度,把图像区分成同实际物体相对应的有意义的区域。每个区域内部从灰度的角度看是均匀的,相邻区域之间是不同的,灰度不同的区域间存在着边界。阈值化分割算法有单阈值算法和多阈值算法。阈值化分割算法最根本的策略是从单阈值化算法开始的,多阈值算法是单阈值算法的发展,单阈值算法是多阈值算法的一种特例。单阈值算法基本策略是预设一个阈值 T ,待处理图像的像素灰度值为 $f(x, y)$,将图像中的像素分成两部分:满足条件 $f(x, y) > T$ 的和 $f(x, y) < T$ 的所有像素,一部分称为前景(目标),另一部分称为背景。这种算法主要应用在指纹识别、印章鉴定、文字识别等领域。多阈值算法是选择多个阈值,将整个灰度范围区分成几部分,每部分的像素自成一类,最终把图像分割成多个不同的区域。

阈值化算法主要有两个步骤:第一步是选择合适的阈值,第二步是通过用图像的每个像素的灰度与选定的阈值做比较确定该像素所属于的类。可见,确定一个合适的阈值是阈值化算法的关键和难点,现有的阈值化算法大多关注最佳阈值的选取。下面介绍几种典型的阈值化分割算法。

(1) 基于直方图的分割算法。

直方图分割算法的基本思想是直接依据图像的原始方图确定阈值。20 世纪 60 年代人们提出了 P -分位数法,该方法是最早的选取阈值的方法,需要一定的先验知识。假设已经获得像素占整幅图像的比例为 P ,用于分割的阈值就应该保证灰度级小于阈值的像素数所

占比率近似等于 P 。 P -分位数法在没有一定的先验知识的情况下往往是不能使用的,这个先验知识一般是指前景点数占有比率。

P -分位数法提出后不久,科学家又提出了 mode 算法。mode 算法的基本思想是在直方图表现出较明显的双峰状时,将双峰中间的谷底对应的灰度级作为分割图像的值。mode 算法的优点是实现起来比较简单;缺点是实际需要分割的图像得到的直方图都比较粗糙,几乎不可能达到理想状态下的光滑度,这样给极大值和极小值的确定带来了极大的困难。

由于直方图的粗糙造成了这些方法的不适用,于是一些学者试图对直方图进行指数平滑处理,目的是解决阻碍辨别极值点的随机干扰点,然后应用一些方法提取出所有可能的极值点,最后再通过相应的评价函数去除不合理的极值点。这种方法优点很明显,缺点同样明显,直方图得到平滑的同时,往往会把一些局部的极值点平滑掉,增加了额外的噪声,影响了图像分割质量。

上面提到的这些直方图算法中,mode 算法使用的频率最高。在实际应用中,并不是每幅图像都满足 mode 算法所需要的直方图特征。因此,一些研究人员提出了一些变换方法,然后通过这些方法将直方图变换成符合 mode 算法要求的直方图,再应用 mode 算法确定阈值。

这些变换方法中使用比较广泛的是四元树直方图变换法,这种方法使得值的选取更容易实现。

以上算法都是单阈值分割算法。在多阈值分割算法中也有基于直方图的分割算法,例如,将二维直方图中每个细化边缘点聚类后得到的平均灰度级作为选取阈值的方法、多阈值选取递归方法以及基于自动分析直方图灰度分布的值化分割算法等。

(2) 基于最大类间方差法。

20 世纪 70 年代末,科学家提出了最大类间方差法。该方法的理论依据是判决分析和最小二乘法。最大类间方差法把通过阈值 T 分成的两类灰度分别进行类内和类间方差计算。这里的阈值 T 是当类间方差最大时确定的灰度级。最大类间方差法的依据是同一个前景内类间的方差较小,但是不同区域的方差最大。这种算法的优点是稳定、实现容易、算法简单。

20 世纪 80 年代初,出现了穷举搜索法、最速上升法等,提高了搜索速度,使得该方法便于实时应用。然而,该方法在类间方差的极大值有多个时计算出的结果不一定是最佳阈值。

(3) 基于最小误差与均匀化误差的分割算法。

为了解决最大类间方法的缺点和不足,出现了最小误差方法,该方法有更为强大的适应能力,解决了前景和背景比例悬殊时遇到的问题。这种方法的具体思路是,由前景和背景像素灰度级组成混合集,用 $P(i)$ 表示该混合集的概率密度函数,直方图可以看作 $P(i)$ 的估计,假设该混合集的每个分量 $P(i|j)$ 服从正态分布,其中均值为 μ_i ,标准差为 σ_j ,然后定义 $J(t)$ 为遵循一定规则的准则函数。在 $J(t)$ 达到极小值时,灰度级 t 就是需要的阈值。20 世纪 80 年代末,出现了经最小误差方法修改后的模型分布方差的偏估计,该算法的提出使得原算法得到了极大的改善,输出效果优势明显,但同时付出了计算复杂度增大的代价。

(4) 基于最大熵的分割算法。

20 世纪 80 年代,应用香农信息论中熵的理论提出了最大后验熵上限法。最大熵算法的基本原理是:在 H_B 加上 H_W 的和取最大值时,前景和背景的一阶灰度的信息量最大,这

时的灰度级就是想要得到的阈值。该方法实现简单,但需要一定的计算量。20 世纪末,该算法有了改进,改进后更加适合计算机处理的算法是通过在重新量化的直方图上取值来确定阈值,最终成功降低了计算量,使得最大熵算法更加适合计算机处理。

2) 基于区域的分割算法

基于区域的分割算法分割图像时,图像的空间信息得到了很好的利用,分割后连续性较好,并且不受图像的分支数影响。主要的分割算法有区域生长和分裂合并。

(1) 区域生长。

区域生长是早期出现的分割算法,是将灰度特征、颜色特征或者纹理特征相似或者相同的像素集合起来构成区域。区域生长法首先在待分割的区域内确定一个种子像素作为生长点,随后让生长点邻域内按照其生长规则的要求与生长点具有相同或者相似特征的像素点合并到生长点所属的类中。在应用区域生长法时,根据不同的初始生长点进行分割时会得到不同的结果,所以初始生长点的选择是区域生长算法的难点。另外,生长规则即生长过程中相似性准则的确定也直接影响图像分割结果,是区域生长算法的另一个难点。

计算复杂度和计算速度是区域生长算法的显著优点,但其同样有明显的缺点,即在缺乏先验知识时,初始生长点选择的随机性和生长规则的随机性往往会产生较差的分割效果。

(2) 分裂合并。

分裂合并算法是一种比较经典的算法。分裂合并算法的基本思想是将待分割的图像先划分成多个不重叠的区域,划分后的区域的大小是任意的,然后按照某分裂合并规则合并或者分裂这些划分后的区域达到分割的要求。分割时,图像区域属性的一致性测度判断准则往往影响着分割的好坏,该准则一般是由待分割图像的一些统计特征来确定的。较常用的分裂合并算法有金字塔分割算法、模糊聚类算法、分水岭算法、RAG 算法、NNG 算法等。其中前三种是分裂算法,后两种是合并算法。

3) 基于边缘检测的算法

边缘是在前景与前景、前景与背景之间、区域与区域之间的灰度变化剧烈、不连续的部分。根据边缘的特性来分割图像是一种重要的分割算法,也就是边缘检测。从人类视觉的角度直接出发,人们发现边缘点往往是图像某些特征变化剧烈的点,这些点通常是一阶导数极大的点或者二阶过零的点,因此提出了一系列的边缘检测算法。

常见的边缘检测算法包括滤波、增强、检测等过程(图 3-12)。滤波过程的目的是对图像进行平滑处理以降低噪声对图像分割的影响,但同时可能会滤去边缘点,因此好的滤波过程应在降低噪声的同时尽量保证边缘的完整。增强过程是为了弥补滤波过程的不足、保证边缘的完整,但同时可能会造成伪边缘的出现。检测过程就是通过某个算法确定边缘点。定位过程是将检测过程中得到的边缘点尽可能精确地定位到实际的边缘。



图 3-12 边缘检测过程

4) 基于 Snakes 模型的分割算法

首先介绍用于图像分割的 Snakes 模型,即

若图像轮廓曲线表示为 $X(s)=(x(s),y(s)),s\in[0,1]$, s 为弧长参数,则该曲线能量

方程为

$$E = \int_0^1 \frac{1}{2} [\alpha |X'(s)|^2 + \beta |x''(s)|^2] + E_{\text{ext}}(X(s)) ds \quad (3-2)$$

其中, α 和 β 为权重系数。式(3-2)中一阶导数和二阶导数项共同构成了轮廓曲线的内部能量。其中一阶导数项为弹性能量, 能量越小, 轮廓点间的距离越小, 由此保证轮廓曲线在外部能量的作用下合理运动和始终连续, 极端情况下轮廓曲线收缩至一个点。二阶导数为刚性能量, 能量越小, 轮廓曲线的局部越趋于直线, 越光滑, 保证轮廓曲线具有形状信息。当 β 为 0 时, 轮廓曲线函数不连续, 可以发展为一个角。第二项 $E_{\text{ext}}(X(s))$ 来自整幅图像, 为外部能量, 负责驱使曲线向图像的边界收敛。

Snakes 模型的适用性较广, 因为 E 采用的是积分的运算方式, 抗噪性能强, 前景的局部模糊敏感度低。该算法的缺点是对初始条件要求高, 收敛速度有待提高。

5) 基于组合优化模型的分割算法

基于组合优化模型的分割算法的主要策略是, 在分割图像时通过定义目标函数, 应用组合优化方法分割图像。目标函数在给定的约束项下的全局最优解就是所求分割的解。

基于组合优化模型的分割算法分割图像的关键是确定合适的区域或边缘的约束条件和对目标函数的估计方法。后来提出的改进算法多是基于以上两点的改进, 如遗传算法、随机退火算法、组合遗传算法等。

6) 基于目标几何和统计模型的分割算法

基于目标几何和统计模型的分割算法是将目标分割和识别统一在一起的算法。该算法的基本策略是把目标的几何知识与统计知识按照一定的规则表示成模型, 使之变成匹配或监督分类。具有代表性的模型分割算法有: 依据目标几何形状特征确定的模板模型分割算法, 采用像概率神经网络、局部线性神经网络等构成目标的基于连接模型的分割算法, 采用传统统计分类策略提取训练图像特征后构成前景的基于特征矢量模型的分割算法。基于目标几何与统计模型的分割算法的优点是能同时解决部分识别任务或者全部识别任务, 效率高; 缺点是在一定条件下误检和漏检的可能性比较大, 在匹配时搜索的速度不够理想。

3.2.4 图像提取

1. 图像边缘类型

边缘是图像像素灰度发生显著变化的部分, 是灰度值不连续的结果, 存在于物体与背景之间、物体与物体之间、区域与区域之间, 一般是由图像中物体的物理特性变化引起的。物理特性变化的不同产生了不同的边缘。图像的边缘具有方向和幅度两个特征。水平与边缘走向像素值变化平缓; 垂直于边缘走向, 像素值变化剧烈, 呈现阶跃状或是斜坡状。因此, 通常将边缘分为阶跃型和屋脊型两种。

阶跃型边缘的两侧灰度值变化较明显, 而屋脊型边缘位于灰度值增减的交界处。从数学角度一般用导数来刻画边缘点的变化, 通常对这两种边缘分别求取一阶、二阶导数即可体现边缘点的变化。阶跃型边缘灰度变化曲线的一阶导数在边缘处达到极大值, 其二阶导数在边缘处与横轴零交叉; 屋脊型边缘灰度变化曲线的一阶导数在边缘处与横轴零交叉, 其二阶导数在边缘处达到负极大值。

2. 图像边缘提取的内容和步骤

图像边缘检测的目的是采用某种算法提取出图像中对象与背景之间的交界线。在最终

提取出图像边缘之前,还需做些前期的图像处理工作,包括滤波去噪、增强边缘部分等,以便提取出真正的、清晰的图像边缘。一般来说,图像的边缘提取包括以下4部分内容。

1) 图像滤波

边缘检测主要是基于图像灰度的导数计算,受噪声影响很大,因此必须采用滤波方法处理受噪声污染的图像,提高边缘检测器的性能。而多数滤波具有低通特性,在去除噪声的同时也会使图像的边缘变得模糊,因此需要选择合适的滤波方法,在保持边缘的同时尽可能地去除噪声,并尽可能地达到增强边缘与降低噪声之间的平衡。

2) 边缘增强

图像增强是指按照特定需要突出图像中的某些信息,处理的结果使图像更适合人的视觉特性和机器的识别系统。图像边缘的增强是确定图像各邻域强度的变化值,显示出邻域或局部强度有显著变化的点,一般是通过计算梯度幅值完成图像边缘的增强。

3) 边缘检测

边缘增强的结果是使图像中灰度有显著变化或者梯度幅值大的点突出显示,但是有些图像中梯度幅值较大的点在某些特定的应用领域中并非边缘点,所以应该用某种算法确定真的边缘点。最简单的是设定梯度幅值作为确定边缘的依据。

4) 边缘定位

根据具体应用领域,结合以上步骤精确确定边缘的位置,提取出图像边缘,得到边缘图像。

在边缘检测算法中,大多数应用仅需指出边缘出现在图像某一像素点的附近,而无须指出边缘的精确位置或方向;然而在某些具体的工程应用中,则需要确定并提取出图像边缘,因此边缘定位也是十分重要的。

边缘提取的基本步骤如图3-13所示。

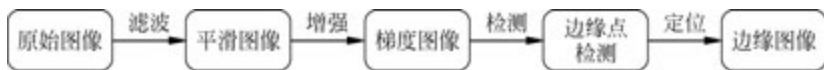


图 3-13 边缘提取的基本步骤

图像边缘提取的基本实现方法如下。

1) 滤波方法

图像在获取、传输和存储过程中不可避免地受到噪声的影响,对图像质量和后续处理通常有较大的影响。图像边缘和噪声均是高频分量,图像降噪和边缘提取是密不可分的,在图像降噪的过程中应尽量保持图像的边缘特征,传统的滤波降噪方法主要有均值滤波法、中值滤波法等。随着对图像质量要求的提高,近年来提出了很多的滤波方法,如自适应滤波方法、小波滤波和形态学滤波等。

(1) 均值滤波法。

均值滤波法是简单的空域处理方法。此方法的基本思想是用一个滤波模板依次滑过整幅图像,然后确定邻域内像素的灰度值,用几个像素灰度的平均值来代替每个像素的灰度,假定有一幅 $N \times N$ 图像 $f(m, n)$,平滑处理后得到图像 $g(x, y)$,像素灰度值为

$$g(x, y) = \frac{1}{M} \sum_{(m, n) \in S} f(m, n), \quad x, y = 0, 1, 2, \dots, N-1 \quad (3-3)$$

其中, S 是以点 (x, y) 为中心的邻域像素点的集合, M 是集合内像素点的总数。也就是说,

平滑化图像 $g(x, y)$ 中每个像素的灰度值均是由包含在 (x, y) 的预定领域中 $f(x, y)$ 的像素灰度值的平均值决定的。

均值滤波法常用的滤波模板有 3×3 、 5×5 和 7×7 ，其优势在于算法简单，计算速度快；但是它在降低图像噪声的同时也会使图像产生模糊，而且它的平滑效果受所使用的模板邻域半径大小的影响，随着邻域半径的加大，图像的模糊程度也愈加严重。为克服这一缺点，可采用阈值减法。当某些点和它邻域内点的灰度平均值之差大于阈值 T 时，就用它们的平均值代替该点的灰度值，否则保留其原灰度值不变。

(2) 中值滤波法。

中值滤波是非线性滤波器的典型代表，主要用于噪声的滤波，在某些条件下可以做到既去除噪声又保护图像边缘。中值滤波基本原理是把数字图像或数字序列中一点的值用该点的一个邻域中各点值的中值代替。中值的定义如下：

一组数 $x_1, x_2, \dots, x_n$ ，把各数按值的大小顺序排列为 $x_{i1} < x_{i2} < \dots < x_{in}$ ，中值 y 计算为

$$y = \text{Med}\{X_1, X_2, \dots, X_n\} = \begin{cases} X_{i[(n+1)/2]}, & n = 1, 3, 5, \dots \\ \frac{1}{2} [X_{i(n/2)} + X_{i(n/2+1)}], & n = 2, 4, 6, \dots \end{cases} \quad (3-4)$$

在一维情况下，中值滤波器是一个含有奇数个像素的滑动窗口，窗口正中间像素的值用窗口内各像素值的中值代替。推广到二维，利用某种形式的二维窗口，即滤波模板，主要有 3×3 和 5×5 模板。中值滤波的优势在于可以克服线性滤波器给图像带来的模糊，能在有效地清除脉冲噪声的同时保护良好的边缘特征，而中值滤波窗口内各点对输出的作用是相同的，如果希望强调中间点或是邻近点的作用，则可以采用加权中值滤波法。

由于不同的滤波器针对不同的噪声，同一幅图像不同区域的滤波强度也应该有所区别，根据这种需求产生了自适应滤波方法。这种方法根据图像的具体情况确定滤波器的系数。图像边缘和图像中的噪声均属于高频分量，但是在不同尺度上又具有不同的特性，于是在不同尺度上提取的边缘在定位精度和抗噪性能上是能够互补的。可以在尺度空间采用滤波方法，即在大尺度下抑制噪声，小尺度下精确定位，以识别可靠边缘，得到边缘的真实位置，从而产生小波滤波方法并应用到图像降噪中。

2) 微分法

(1) 一阶微分方法。

图像的梯度函数在边缘处为局部极大值，即函数梯度变化的速率为局部极大值。通过一阶导数算子或梯度算子估计图像梯度变化的方向，增强图像灰度变化区域，进而判断边缘。

给定函数 $f(x, y)$ ，在点 (x, y) 的 x 方向、 y 方向和 0 方向的一阶方向导数分别为

$$\begin{aligned} f_x(x, y) &= \frac{\partial f}{\partial x} \\ f_y(x, y) &= \frac{\partial f}{\partial y} \\ f_\theta(x, y) &= \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \end{aligned} \quad (3-5)$$

在点 $f(x, y)$ 处的梯度可定义为一个矢量, 即

$$\text{grad}[f(x, y)] = \left[\frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \right]^T \quad (3-6)$$

如果用 $G[f(x, y)]$ 来表示 $\text{grad}[f(x, y)]$ 的幅度, 那么

$$G[f(x, y)] = \max\{\text{grad}[f(x, y)]\} = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{\frac{1}{2}} \quad (3-7)$$

由以上定义可知, 矢量 $\text{grad}[(x, y)]$ 是指向 $f(x, y)$ 最大增长率的方向; $G[(x, y)]$ 是 $\text{grad}[(x, y)]$ 的方向上每单位距离 $f(x, y)$ 的最大增加率, 也就是“梯度的模”。梯度的方向为垂直于边缘的方向, 即

$$\varphi = \arctan \left\{ \frac{\partial f}{\partial y} / \frac{\partial f}{\partial x} \right\} \quad (3-8)$$

在数字图像处理中采用离散形式, 用差分运算代替微分运算。式(3-8)可以用差分式来近似, 即

$$G[f(x, y)] \approx \{ [f(x, y) - f(x+1, y)]^2 + [f(x, y) - f(x, y+1)]^2 \}^{\frac{1}{2}} \quad (3-9)$$

在用计算机计算梯度时, 通常用绝对值运算代替差分运算式, 即

$$G[f(x, y)] \approx |f(x, y) - f(x+1, y)| + |f(x, y) - f(x, y+1)| \quad (3-10)$$

由以上公式可知, 梯度的近似值都和内部像素的灰度差成正比。在一幅图像中, 边缘区梯度值较大, 平滑区梯度值较小, 对于灰度值数值为常数的区域, 梯度值为零, 通过图像与一阶导数算子做卷积, 这些算子体现为模板。先选择多个模板处理结果中的较大值作为在该点的输出值, 然后选择一个合适的阈值做判定, 以得到图像边缘。

(2) 二阶微分方法。

阶跃型边缘对应图像一阶导数的局部极大值, 也对应二阶导数的过零点, 而寻找过零点位置比起寻找极值点更容易且精确, 因此可通过检测图像二阶导数过零点位置来检测图像边缘。

函数 $f(x, y)$ 在点 (x, y) 的 x 方向、 y 方向二阶导数分别为

$$f_{xx}(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} \quad (3-11)$$

$$f_{yy}(x, y) = \frac{\partial^2 f(x, y)}{\partial y^2} \quad (3-12)$$

设 $\vec{n}$ 为梯度方向, 图像 $f(x, y)$ 在点 (x, y) 处沿梯度方向二阶导数为

$$\frac{\partial^2 f(x, y)}{\partial n^2} = \frac{\partial^2 f(x, y)}{\partial x^2} \cos^2 \theta + 2 \frac{\partial^2 f(x, y)}{\partial x \partial y} \sin \theta \cos \theta + \frac{\partial^2 f(x, y)}{\partial y^2} \sin^2 \theta \quad (3-13)$$

由于二阶方向导数算子不具有线性及旋转不变性, 所以通常采用具有旋转对称性和线性且对灰度突变较敏感的算子来实现。

3) 边缘提取

图像边缘在被微分算子增强后, 用边缘提取方法提取出边缘。常用的方法主要有阈值法和零交叉边缘提取方法。

(1) 阈值法。

阈值法的基本方法是通过设定阈值 T , 若微分增强后像素点的数值大于 T , 则认为是边缘点, 否则不是边缘点。合理设置阈值对后续处理非常重要, 如果阈值设置过高, 容易丢

失目标边缘；过低则会增加背景噪声，这样会影响后续处理的精度和实时性。通常边缘提取的阈值是基于边缘增强自适应确定的，根据处理数据的范围分为全局和局部自适应方法。随着理论的发展和技术应用的需要，阈值的选取也产生很多新的方法，主要有迭代阈值法、均匀分布阈值法、一维熵阈值法、模糊阈值法等。

仅有一个阈值并不充分，可能会造成断边缘或是假边缘，因此可以使用双阈值方法。首先利用累计统计直方图得到高阈值，然后再取一个低阈值，如果图像信号的响应大于高阈值，则它一定是边缘；如果低于低阈值则为非边缘；如果在两者之间，则要看它的8个邻接像素有无大于高阈值的值。

(2) 零交叉方法。

零交叉边缘提取法应用也很广泛。使用算子卷积图像，通过判断符号的变化确定零交叉点的位置为边缘点。此外，使用符号结合法定位边缘，也可以取得较好的效果。

3. 图像边缘提取的难点

图像的边缘提取是图像处理和计算机视觉研究的重要课题之一，其提取算法的性能好坏直接决定计算机视觉系统的优劣。好的边缘提取方法应具有定位精度高、抗噪声能力强、不漏掉较真实边缘和不虚报假边缘的性能，但是要完全达到这些要求有很大的困难，主要是定位精度与降噪能力之间的相互矛盾，具体体现在以下几点。

(1) 图像提取中噪声的问题。

在图像的提取、传输、存储过程中都不可避免地增加不同程度和不同性质的噪声，而噪声和边缘均属于图像的高频分量。若考虑噪声的影响，先通过平滑滤除噪声将会失去部分图像边缘信息，若增强边缘的高频分量将会增强该频度内的噪声，即很难较好地达到既能检测到边缘的准确位置，又能够抑制无关的细节和噪声的效果。

(2) 不同尺度选取边缘问题。

图像边缘往往存在于不同的尺度范围内，传统的边缘提取算子为单一尺度，不可能检测出所有正确的边缘。当所选取尺度较大时，抗噪声能力较好但边缘定位不够准确；当尺度较小时，边缘定位精确度提高，但抗噪声能力降低。所以合理选用不同尺度或多尺度方法对图像进行有效的边缘提取也是需要着重考虑的问题。

(3) 图像边缘位置的确定和图像边缘的连续性问题。

检测出的图像边缘的位置会根据算法的不同而不同，并且很多边缘并非单一像素宽，还会因为噪声的影响而导致边缘移位，为真实边缘的确定增加了难度。由于图像几何方面和光学方面的原因，主要是深度的不连续性、颜色和纹理的不同、表面反射、目标物体的阴影等，将致使检测出的边缘具有较低的连续性。

因此，在图像边缘提取过程中，图像的边缘和噪声特性的变化是在不同的尺度空间范围内，清楚地检测到这种特性的变化，对于选择合适的提取方法意义重大，并且不能期望一种检测算子能检测出发生在图像上的所有特性变化，需要考虑多种方法的组合，先找出边缘，再提高边缘的定位精度。

4. 图像边缘提取性能的评价标准

一般来说，对边缘提取方法有如下客观要求。

(1) 要能检测出有效边缘。

(2) 要有高的边缘定位精度。

- (3) 检测的响应最好是单像素的。
- (4) 能对不同尺度的边缘都有较好的响应且能尽量减少漏检。
- (5) 对噪声不敏感。
- (6) 检测的灵敏度受边缘方向的影响小。

第(1)、(2)条要求主要是评价边缘检测算法的性能,第(3)、(4)条判定边缘定位和方向估计算法的性能,最后两条的要求是希望边缘提取算法偏离理想模型的误差范围尽可能地小。这些要求是存在相互矛盾的,使用一个边缘检测器很难达到完美的统一。在评价一个边缘检测器的性能时,主要考察假边缘概率、丢失边缘概率、边缘方向角估计误差、边缘估计值到真实边缘的距离平方值、畸变边缘及其他诸如角点或节点的误差范围等。对图像边缘提取性能的评价主要有主观和客观评价。

由于边缘提取算法的客观评价方法都有其局限性,并且无法证明其有效性,所以目前还是较多地通过人眼主观判断边缘提取方法的优劣,主要是观察边缘的连续性、边缘的光滑性、边缘的细化程度、边缘的定位和定位精度等。

(1) 边缘的连续性。

边缘的连续性要求提取结果的边缘是连续的,需要边缘提取算法受高频信号和噪声的影响小,因为这两者会导致边缘不连续。不连续的算法将会直接导致图像边缘分图断裂或是错误连接其他边缘,致使分割错误。

(2) 边缘的光滑性和细化程度。

边缘的光滑性和细化程度是由算法本身的特性决定的,边缘的光滑性直接影响主观视觉,但对后续处理的影响不大;边缘的细化程度即提取到的图像边缘的宽度。评价图像边缘的这些方面需要使用多幅图像来评价以降低误判率。

(3) 边缘的定位和定位精度。

边缘的定位是考察边缘提取算法能否提取出目标物体的图像边缘,定位错误主要包括漏检和错检。边缘的定位精度考察的是提取出的边缘点与实际边缘点应在的位置的距离,是基于边缘已被定位在真实边缘附近的邻域中。

3.3 图像分析

3.3.1 图像分析的作用

图像分析和图像处理关系密切,两者有一定程度的交叉,但又有所不同。图像处理侧重于信号处理方面的研究,如图像对比度的调节、图像编码、降噪等。而图像分析的侧重点在于研究图像的内容,包括但不局限于使用图像处理的各种技术,它更倾向对图像内容的分析、解释和识别。因而,图像分析和计算机科学领域中的模式识别、计算机视觉关系更密切一些。

图像分析一般利用数学模型并结合图像处理的技术来分析图像的特征和结构,从而提取特定的信息,达到智能化的效果。

3.3.2 模板匹配

模板匹配的本质是基于内容的检索,属于图像分析领域中最常用的技术之一。其目的

是在给定查询图像的前提下,依据内容信息或指定查询标准,在图像数据库中搜索并查找出符合查询条件的相应图像。

互联网上传统的搜索引擎,包括百度、Google 都推出了相应的图片搜索功能,但是这种搜索主要基于图片的文件名建立索引来实现查询功能。这种从查询文字、文件名,最后到图片查询的机制并不是基于图像内容的检索。基于图像内容的检索指的是查询条件本身就是一幅图像,或者是对于图像内容的描述,其建立索引的方式是通过提取图像的底层特征,然后计算比较这些特征和查询条件之间的异同,来决定两幅图片的相似程度。

这里使用一个生活中的实例,以手机 App 中的“拍立淘”功能进行说明,如图 3-14 所示。



图 3-14 模板匹配搜索

在这个实例中,用户想要在网上搜索图 3-14(a)所示的水杯,查询条件就是图 3-14(b)中通过手机后置摄像头扫描到的图像。通过在软件的内部进行图像特征提取,然后在网站的数据库中与众多商品的图像进行查询和对比,即可找到符合特征条件的商品,如图 3-14(c)所示。模板匹配的核心包括以下 3 部分。

1. 特征提取

从广义上讲,图像特征包括基于文本的特征(如关键字、注释等)和视觉特征(如色彩、纹理、形状、对象表面等)两类。视觉特征又可分为通用的视觉特征和领域相关的视觉特征。前者用于描述所有图像共有的特征,与图像的具体类型或内容无关,主要包括色彩、纹理和形状,这些特征又被称为底层特征;后者则建立在对所描述图像内容的某些先验知识(或假设)的基础上,与具体应用紧密相关,如人的面部特征或指纹特征等。

可提取的图像特征可以包括颜色、纹理、平面空间对应关系、外形或者其他统计特征。图像特征的提取与表达是模板匹配的基础。

2. 相似度测量

从图像中提取的特征可以组成一个向量,两幅图像之间可以通过定义一个距离或者相似性的测量度来计算相似度。

3. 模板查询

模板查询有着不同的查询方式,典型的 3 种查询方式如下。

- (1) 按例查询：用户提供一幅查询图像，在数据库中搜索相似图像。
- (2) 按绘查询：用户在类似画笔的接口上进行简单的绘画，以此为标准进行查询。
- (3) 按描述查询：例如，指定条件可以是 30% 的黄色、70% 的蓝色等。

3.3.3 其他可实现的功能

1. 识别物体几何中心位置

在制造装配等应用中，物体的位置起着十分重要的作用。工业应用中，物体通常出现在已知表面(如工作台面)上，而且相机相对台面的位置也是已知的。在这种情况下，图像中物体的位置决定了其空间位置。确定物体位置的方法很多，如采用物体的外接矩形、物体矩心(区域质心)等来表示物体的位置。

1) 区域外接矩形

当物体的边界已知时，可用其外接矩形的尺寸来刻画它的基本形状，如图 3-15 所示。求物体在坐标系方向上的外接矩形，只需计算物体边界点的最大和最小坐标值，就可以得到物体的水平和垂直跨度。

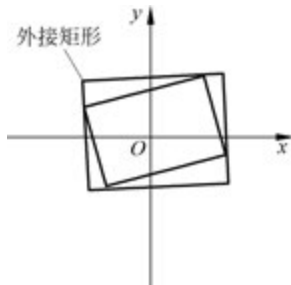


图 3-15 区域外接矩形

2) 区域质心坐标

在工件的分拣操作中，工件的位置信息通常采用质心坐标来表示，从而方便机械手臂抓取目标工件。一般情况下，图像中的物体并不是一个点，因此采用物体或区域面积的中心点作为物体的位置。区域质心是通过将图像进行“全局”运算得到的一个点，因此它对图像中的噪声相对来说是不敏感的。对于二值图像，物体的中心位置与物体的质心位置相同。

2. 识别物体颜色

颜色是彩色图像的直观特征之一，其描述方法与所采用的颜色模型有关，常用的颜色模型有 RGB 颜色空间和 HIS(色相、饱和度、亮度)颜色空间。

为了更好地根据颜色特征实时检测和识别生产线上的工件，可以先采用图像处理软件的检测工具将工件图像的颜色信息记录下来。通常在同一背景下对工件进行判别。当相机定位好后，调用示教模块先记录下工件的颜色信息，然后在工作时调用颜色识别模块对工件进行判别，示教的目的是让程序记住工件的局部颜色信息以及工件的位置。在识别工件时，可以只对已经示教过的区域进行求解，采用对比的检测方法，以灰度或者色彩对比的方式获取当前图像和模型的相似程度，但采用这种方式时产品色彩细微的差异比较容易导致相似度的下降。

3. 识别物体形状

形状识别是指采用边界轮廓对比的方式(边缘检测)获取当前图像和模型的相似程度，产品色彩细微的差异不容易导致相似度的下降。

根据采用的方式不同，识别物体形状大致可以分为以下几类：基于局部图像函数的算法、图像滤波、多尺度算法、基于反应-扩散方程的方法、多分辨率方法、基于边界曲线拟合的方法、主动轮廓线方法等。

在实践时，由于形状检测处理的数据量较小，因此可以高速地进行检测。但若选中的模型(模板)太大，则检测速度会下降。

4. 识别旋转角度

角点是最适合计算工件旋转角度的特征。角点检测算法主要分为基于图像边缘和基于图像灰度两类。前者在图像分割的基础上提取边界构成链码,并将边界上转折较大的点看作角点。该算法过于复杂,且计算量较大,而后者则避开了这些缺陷。

3.4 检测结果的输出

3.4.1 输出对象和输出内容

1. 输出对象

视觉检测的结果由视觉系统的存储单元发出,结果输出对象通常分为两种:显示设备(如显示屏)和上位机。显示屏显示的内容往往是软件界面、相机捕捉画面等内容,以方便用户操作和监控视觉系统。通常应用在工业机器人视觉系统中的上位机有 PLC、PC 和机器人控制器。

2. 输出内容

视觉系统的输出内容见表 3-1。

表 3-1 视觉系统的输出内容

种 类	内 容
状态信号	由视觉控制器确认控制信号或命令输入并开始测量后,利用状态信号(如常见的可检测/正在检测等)向外部装置通知传感器的状态
判定结果	判定结果虽然是个布尔值,只有“是”或“否”(通过/不通过)两种可能,却可以代表多种内容,如目标形状是否匹配、颜色是否匹配、数量是否匹配等。在包含多个处理项目的判定结果中,只要有一个项目判定为不通过,就会输出不通过的结果
测量值	测量值可根据用户自己的设定,输出不同的数值,如目标的位置信息(X 轴坐标、Y 轴坐标、旋转角度等)、数量信息、测量的长度等
字符串输出	字符串输出可以用于输出从条形码、二维码中读取的字符串或数字

3.4.2 视觉系统与上位机的通信

对视觉系统而言,通信非常重要,它是共享数据、支持决策和实现高效率一体化流程的一种方式。视觉控制系统的上位机通常是 PC、PLC 或工业机器人控制器。联网后,视觉系统可以向 PC 传输检测结果以进行进一步分析。工业中更常见的是直接传输给集成过程控制系统的 PLC、机器人和其他工厂自动化设备。

不同品牌的视觉控制系统有其支持的不同通信方式,不同品牌的 PLC 及工业机器人控制器也有不同的接口。要把视觉系统集成到工厂 PLC、机器人或其他自动化装置上,需要找到一种二者相互支持的通信方式或协议。常见的通信方式和协议如下。

1. 并行通信

通过并行接口,在视觉系统和外部装置之间进行通信。

2. 串行通信

通过 RS-232 或 RS-485 串行接口,可以与绝大多数的机器人控制器通信。

3. 工业以太网协议

允许通过以太网网线连接 PLC 和其他装置,无须复杂的接线方案和价格高昂的网络网

关。常见的以太网协议有以下几种。

1) 客户机/服务器

客户机/服务器让视觉系统能够轻易通过以太网与其他系统和控制装置共享数据,无须开发代码。

2) 现场总线网络

通常需要一个协议网关附件将视觉系统添加到现场总线网络,往往不同品牌的产品会使用不同的网络通信协议,但其本质都是以太网通信,如西门子常用的 PROFINET、欧姆龙常用的 Ether CAT、施耐德的 Modbus TCP/IP 等,可根据现场实际情况来选用合适的品牌。

(1) Ether CAT(Ethernet Control Automation Technology)以以太网系统为基础,是一种可实现更高速、更高效通信的高性能工业用网络系统。各节点可高速传输以太网帧,因此可有效缩短通信周期。虽然 Ether CAT 是独立的通信协议,但物理层还是采用标准的以太网技术,因此有良好的通用性,可以使用以太网电缆等。Ether CAT 不仅适用于对处理速度、系统兼容性有高要求的大型控制系统,在中小控制系统中也能充分发挥其作用。

(2) Ether Net/IP 是一种使用以太网的工业用多厂商网络。其规格为开放标准,由 ODVA(Open Device Net Vendor Association)管理,可应用在程序控制及其他自动化的应用中,是通用工业协议(CIP)中的一部分。Ether NAT/IP 采用标准的以太网技术,因此可以与各种以太网设备一起使用。

(3) PROFINET 是一个开放式的工业以太网通信协议,由 PROFIBUS&PROFINET 国际协会提出。PROFINET 应用 TCP/IP 及信息科技的相关需求选择不同层面的功能,各功能的差异在于为了满足高速通信的需求而对应资料交换种类的不同。

视觉控制器与外部上位机根据接口的不同通过不同的通信电缆连接,并根据各种通信协议进行通信。常用的几种通信连接如图 3-16 所示,不同通信方式的优缺点见表 3-2。

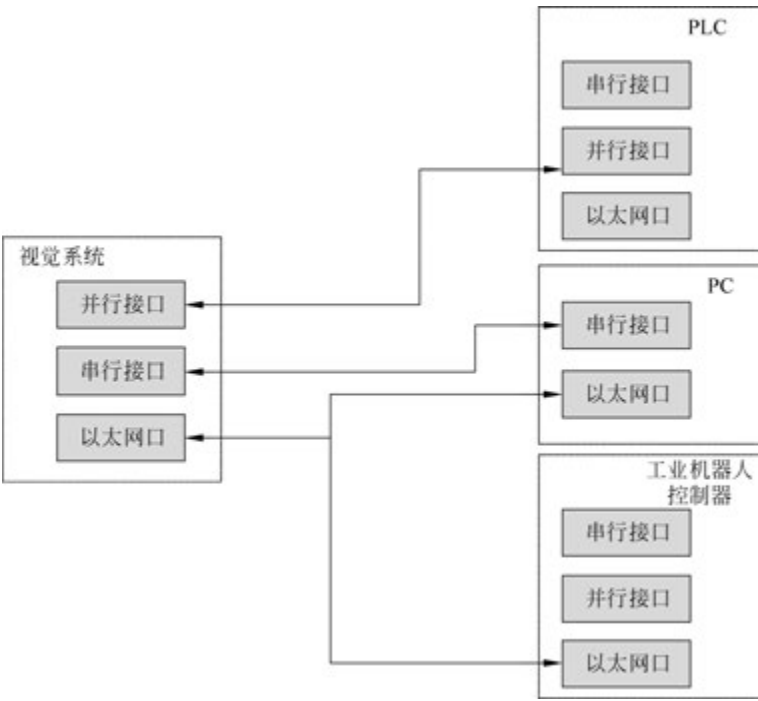


图 3-16 3 种通信方式的连接

表 3-2 视觉系统不同通信方式的优缺点

通信方式	并行通信	串行通信	工业以太网通信
优点	因为可以多位数据一起传输，所以传输速率很快	使用的数据线少，在远距离通信中可以节约通信成本。 在高速传输状态下，串行只有一根数据线，不存在信号线之间的串扰，而且串行通信还可以采用低压差分信号，大大提高抗干扰性，实现更高的传输速率	实时性强。一定的时间内发送一个指令一定要被处理，否则通信就会失败
缺点	内存有多少位，就要用多少数据线，所以需要大量的数据线，成本很高。 在高速传输状态下，并行接口的几根数据线之间存在串扰；而且并行接口需要信号同时发送、同时接收，任何一根数据线的延迟都会引起问题	因为每次只能传输一位数据，所以传输速率比较低	区别于其他的运行环境，工业以太网对温度、干扰要求会更高

3.5 摄像机标定

3.5.1 摄像机标定技术概述

计算机视觉的基本任务之一是从摄像机获取的图像信息出发，计算三维空间中物体的几何信息，并由此重建和识别物体；而空间物体表面某点的三维几何位置与其在图像中对应点之间的相互关系是由摄像机成像的几何模型决定的，这些几何模型参数就是摄像机参数。在大多数条件下，这些参数必须通过实验与计算才能得到，这个过程称为摄像机标定（或称为标定）。标定过程就是确定摄像机的几何和光学参数，以及摄像机相对于世界坐标系的方位。标定精度的大小，直接影响着计算机视觉（机器视觉）的精度。迄今为止，对于摄像机标定已提出了很多方法，本节讨论如何针对具体的实际应用问题，采用特定的简便、实用、快速、准确的标定方法。

摄像机成像模型图像是空间物体通过成像系统在像平面上的投影。图像上每个像素点的灰度反映了空间物体表面某点的反射光的强度，而该点在图像上的位置则与空间物体表面对应点的几何位置有关。这些位置的相互关系，由摄像机成像系统的几何投影模型决定。理想的投影成像模型是光学中的中心投影，也称为针孔成像模型，如图 3-17 所示。

针孔成像模型假设物体表面的反射光都经过一个针孔而投影到像平面上，即满足光的直线传播条件。针孔模型主要由光心（投影中心）、成像面和光轴组成。小孔成像由于透光量太小，因此需要很长的曝光时间，并且很难得到清晰的图像。实际摄像系统通常都由透镜或者透镜组组成。两种模型具有相同的成像关系，即像点是物点和光心的连线与图像平面

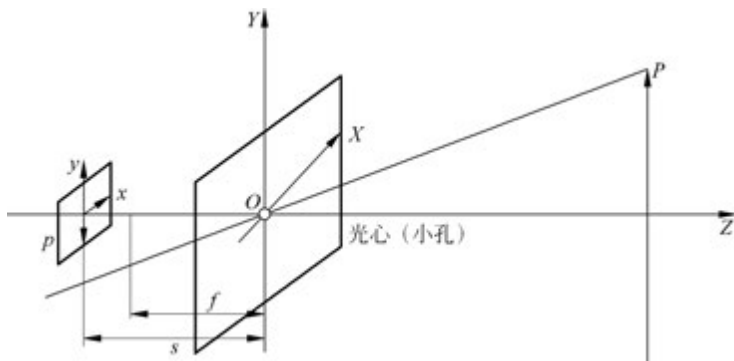


图 3-17 针孔成像模型

的交点。因此,可以用针孔模型作为摄像机成像模型。当然,由于透镜设计的复杂性和工艺水平等因素的影响,实际透镜成像系统不可能严格满足针孔模型,会产生所谓的镜头畸变,常见的如径向畸变、切向畸变、薄棱镜畸变等,因而在远离图像中心处会有较大的畸变,在精密视觉测量等应用方面,应该尽量采用非线性模型来描述成像关系。

3.5.2 摄像机透视投影模型

摄像机通过成像透镜将三维场景投影到摄像机二维像平面上,这个投影可用成像变换(即摄像机成像模型)来描述。摄像机成像模型分为线性模型和非线性模型。针孔成像模型就属于线性摄像机模型,本节就讨论在这种模型下,某空间点与其图像投影点在各种坐标系下的变换关系。图 3-18 所示为 3 个不同层次的坐标系在针孔成像模型下的关系,其中 (X_w, Y_w, Z_w) 为世界坐标系; (x, y, z) 为摄像机坐标系, o 为光心;像面 $X_f O_f Y_f$ 表示的是视野平面,其到光心的距离即 oO 为 f (镜头焦距); $X_f O_f Y_f$ 为以像素为单位的图像坐标系, XOY 为以毫米为单位的图像坐标系。

1) 3 个层次的坐标系统

(1) 世界坐标系 (X_w, Y_w, Z_w) : 也称真实或现实世界坐标系,或全局坐标系。它是客观世界的绝对坐标,一般的三维场景都用这个坐标系来表示。

(2) 摄像机坐标系 (x, y, z) : 以小孔摄像机模型的聚焦中心为原点,以摄像机光轴为 z 轴建立的三维直角坐标系。 x 、 y 一般与图像物理坐标系的 X_f 、 Y_f 平行,且采取前投影模型。

(3) 图像坐标系分为图像物理坐标系和图像像素坐标系两种。图像物理坐标系,其原点为透镜光轴与成像平面的交点, X 与 Y 轴分别平行于摄像机坐标系的 x 与 y 轴,是平面直角坐标系,单位为毫米。图像像素坐标系,即固定在图像上的以像素为单位的平面直角坐标系,其原点位于图像左上角, X_f 、 Y_f 平行于图像物理坐标系的 X 和 Y 轴,对于数字图像,分别为行、列方向。

空间某点 P 到其像点 p 的坐标转换过程主要是通过这几套坐标系的三次转换实现的,首先将世界坐标系进行平移和转换得到摄像机坐标系,然后根据三角几何变换得到图像物理坐标系,最后根据像素和公制单位的比率得到图像像素坐标系。(实际的应用过程是这个过程的逆过程,即由像素长度获知实际的长度。)

通过摄像头的标定,可以得到视野平面上的 mm/pix 分辨率,对于视野平面以外的物体

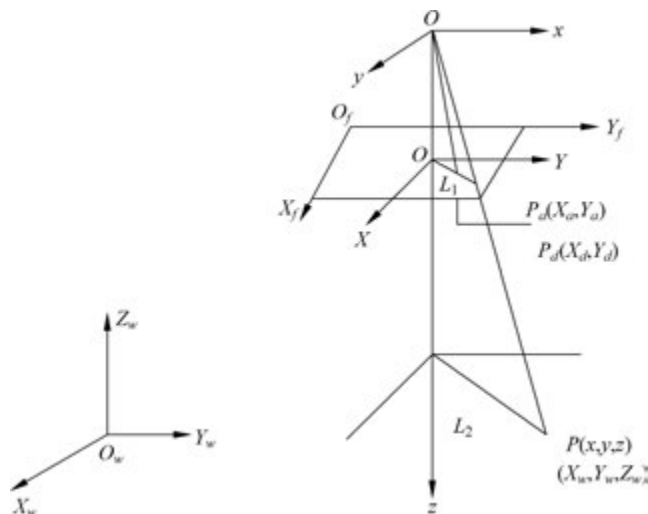


图 3-18 标定系统的坐标系

还是需要通过坐标转换到视野平面上。

2) 摄像机标定参数

摄像机标定是确定摄像机内部参数或外部参数的过程。内部参数是指摄像机内部几何和光学特性,外部参数是指摄像机相对世界坐标系原点的平移和旋转位置。

摄像机内部参数是由摄像机内部几何和光学特性决定的,主要包括以下参数。

(1) 主点 (u_0, v_0) : 图像平面原点的计算机图像像素坐标。

(2) 有效焦距 f : 图像平面到投影中心(光心)距离。

(3) 透镜畸变系数 k : 畸变包括径向畸变和切向畸变,由于一般切向畸变较小,对映射关系影响不大,因此不予考虑。

(4) 轴方向的尺度因子 d_x, d_y : 表示单位像素的实际尺寸。

摄像机外部参数是指从世界坐标系到摄像机坐标系的平移向量和旋转变换矩阵。

3) 坐标系变换关系

定义了上述各种空间坐标系后,就可以建立两两不同坐标变换之间的关系。

(1) 世界坐标系与摄像机坐标系变换关系。世界坐标系中的点到摄像机坐标系的变换可由一个正交变换矩阵 $\mathbf{R}$ 和一个平移变换矩阵 $\mathbf{T}$ 表示。

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{T} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{T} \quad (3-14)$$

齐次坐标表示为

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3-15)$$

式中, $\mathbf{T} = [t_x, t_y, t_z]^T$ 是世界坐标系原点在摄像机坐标系中的坐标; 矩阵 $\mathbf{R}$ 是正交旋转

矩阵,其矩阵元素满足

$$\begin{cases} r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \\ r_{21}^2 + r_{22}^2 + r_{23}^2 = 1 \\ r_{31}^2 + r_{32}^2 + r_{33}^2 = 1 \end{cases} \quad (3-16)$$

正交旋转矩阵实际上只含有 3 个独立变量,再加上 t_x 、 t_y 和 t_z ,总共有 6 个参数决定了摄像机光轴在世界坐标系中空间位置,因此这 6 个参数称为摄像机外部参数。

(2) 图像坐标系与摄像机坐标系变换关系。如图 3-18 所示,摄像机坐标系中的物点 P 在图像坐标系中像点 P_u 的坐标为

$$\begin{cases} X = f_x / z \\ Y = f_y / z \end{cases} \quad (3-17)$$

齐次坐标表示为

$$z \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3-18)$$

将式(3-18)的图像坐标系进一步转换为图像坐标系

$$\begin{cases} u - u_0 = X/d_x = s_x X \\ v - v_0 = Y/d_y = s_y Y \end{cases} \quad (3-19)$$

齐次坐标表示为

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & u_0 \\ 0 & s_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3-20)$$

其中, u_0 、 v_0 是图像中心(光轴与图像平面的交点)坐标, d_x 、 d_y 分别为一个像素在 X 与 Y 方向上的物理尺寸, $s_x = 1/d_x$ 、 $s_y = 1/d_y$ 分别为 X 与 Y 方向上的采样频率,即单位长度的像素个数。

由此可得,物点 P 与图像像素坐标系中像点 P 的变换关系

$$\begin{cases} u - u_0 = f s_x x / z = f_x x / z \\ v - v_0 = f s_y y / z = f_y y / z \end{cases} \quad (3-21)$$

其中, $f_x = f s_x$ 、 $f_y = f s_y$ 分别定义为 X 和 Y 方向的等效焦距。 f_x 、 f_y 、 u_0 、 v_0 参数只与摄像机内部结构有关,因此称为摄像机内部参数。

(3) 世界坐标系与图像坐标系变换关系(共线方程)为

$$\begin{cases} \frac{X}{f} = \frac{u - u_0}{f_x} = \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \\ \frac{Y}{f} = \frac{v - v_0}{f_y} = \frac{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y}{r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z} \end{cases} \quad (3-22)$$

齐次坐标表示为

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{M}_1 \mathbf{M}_2 \mathbf{X} = \mathbf{M} \mathbf{X} \quad (3-23)$$

式(3-23)是图像测量中最基本的共线方程,说明物点、光心和像点这三点必须在同一条直线上。这是针孔模型或者中心投影的数学表达式。根据共线方程,在摄像机内部参数确定的条件下,利用若干已知的特征点和相应的像点坐标,就可以求解出摄像机的6个外部参数,即摄像机的光心坐标和光轴方位的信息。

3.5.3 摄像机镜头的畸变

由于摄像机光学系统并不是精确地按理想化的小孔成像原理工作,存在透镜畸变,物体点在摄像机成像面上实际所成的像与理想成像之间存在光学畸变误差。主要的畸变误差分为三类:径向畸变、偏心畸变和薄棱镜畸变。第一类只产生径向位置的偏差,后两类则既产生径向偏差,又产生切向偏差。图3-19为无畸变理想图像点位置与有畸变实际图像点位置之间的关系。

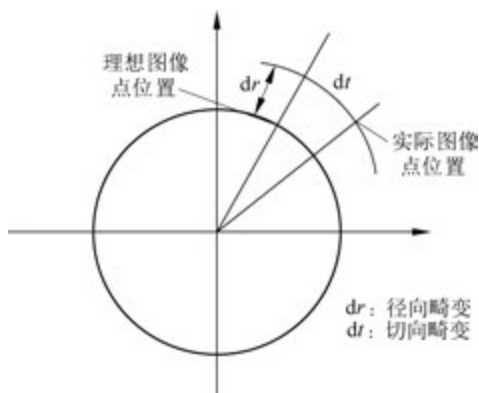


图 3-19 理想图像点与实际图像点

1. 径向变形(径向畸变)

光学镜头径向曲率的变化是引起径向变形的主要原因。这种变形会引起图像点沿径向移动,离中心点越远,其变形量越大。正的径向变形量会引起点向远离图像中心的方向移动,其比例系数增大;负的径向变形量会引起点向靠近图像中心的方向移动,其比例系数减小。如图3-20所示,数学模型如下:

$$\begin{cases} \delta X_r = k_1 X_d (X_d^2 + Y_d^2) + O[(X_d, Y_d)^5] \\ \delta Y_r = k_1 Y_d (X_d^2 + Y_d^2) + O[(X_d, Y_d)^5] \end{cases} \quad (3-24)$$

2. 偏心变形(切向畸变)

由于装配误差,组成光学系统的多个光学镜头的光轴不可能完全共线,从而引起偏心变形。这种变形是由径向变形分量和切向变形分量共同构成,如图3-21所示,其数学模型如下:

$$\begin{cases} \delta X_d = p_1 X_d (3X_d^2 + Y_d^2) + 2p_2 X_d Y_d + O[(X_d, Y_d)^4] \\ \delta Y_d = 2p_1 X_d Y_d + p_2 Y_d (X_d^2 + 3Y_d^2) + O[(X_d, Y_d)^4] \end{cases} \quad (3-25)$$

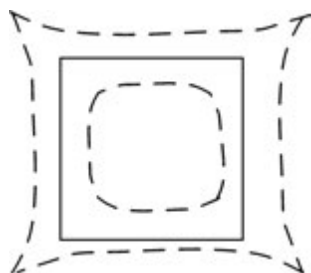


图 3-20 径向畸变

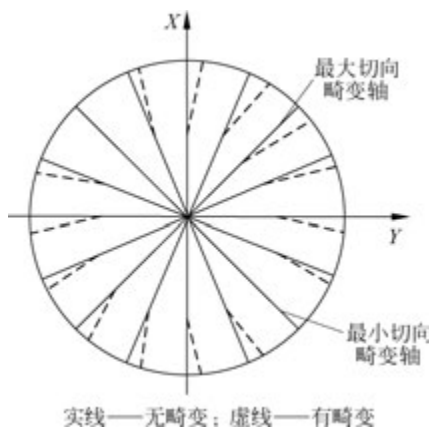


图 3-21 切向畸变

3. 薄棱镜变形

薄棱镜变形指由光学镜头制造误差和成像敏感阵列制作误差引起的图像变形。这种变形是由径向变形分量和切向变形分量共同构成的,其数学模型为

$$\begin{cases} \delta X_p = s_1 (X_d^2 + Y_d^2) + O[(X_d, Y_d)^4] \\ \delta Y_p = s_2 (X_d^2 + Y_d^2) + O[(X_d, Y_d)^4] \end{cases} \quad (3-26)$$

所谓摄像机的非线性模型,是指成像过程不服从小孔成像模型。线性模型为小孔成像模型,不能准确地描述成像几何关系,在远离图像中心处会有较大的畸变,使得所量测的像点坐标产生误差。摄像机的非线性模型可使用式(3-27)表示:

$$\begin{cases} X_u = X_d + \delta_X(X, Y) = X_d + \delta X_r + \delta X_d + \delta X_p \\ Y_u = Y_d + \delta_Y(X, Y) = Y_d + \delta Y_r + \delta Y_d + \delta Y_p \end{cases} \quad (3-27)$$

其中, (X_u, Y_u) 为线性模型下图像点的理想坐标; (X_d, Y_d) 是实际投影点的图像坐标; δ_x, δ_y 是非线性畸变值,它与图像点在图像中的位置有关,可用式(3-28)表示:

$$\begin{cases} \delta_X(X, Y) = k_1 X (X^2 + Y^2) + (p_1 (3X^2 + Y^2) + 2p_2 XY) + s_1 (X^2 + Y^2) \\ \delta_Y(X, Y) = k_2 Y (X^2 + Y^2) + (p_2 (X^2 + 3Y^2) + 2p_1 XY) + s_2 (X^2 + Y^2) \end{cases} \quad (3-28)$$

式中, $k_1, k_2, p_1, p_2, s_1, s_2$ 称为摄像机的非线性畸变参数。镜头的畸变像差与透视畸变并不是同一个概念。镜头的畸变是镜头成像造成的,在设计镜头时可以采取各种手段(如非球面镜)来减小畸变。透视畸变是由视点、视角、镜头指向(俯仰)等因素决定的,这是透视的规律。无论是何种镜头,如果视点相同、视角相同、镜头指向相同,产生的透视畸变是相同的。

3.5.4 经典标定方法简介

1. 摄像机标定方法概述

现有的摄像机标定方法按是否需要参照标定物,大体可以分为两类:传统的摄像机标

定方法和摄像机自标定方法。

1) 传统的摄像机标定方法

传统的摄像机标定方法是使用最为普遍的标定方法,利用一个形状尺寸已知的物体作为标定物,用摄像机拍摄若干幅标定物的图片,通过计算二维图像点与三维空间点之间的关系来完成标定,通过对标定物的合理设计就可以得到高精度的结果。按照标定参照物与算法思路,该方法可以分成基于三维立体靶标的摄像机标定、基于二维平面靶标的摄像机标定以及基于径向约束的摄像机标定等。

(1) 基于三维立体靶标的摄像机标定。

基于三维立体靶标进行摄像机标定是将一个三维立体靶标放置在摄像机前,靶标上每个小方块的顶点均可作为特征点。每个特征点相对于世界坐标系的位置在制作时应精确测定。摄像机获得靶标上特征点的图像后,由于表现三维空间坐标系与二维图像坐标系关系的方程是摄像机内部参数和外部参数的非线性方程,如果忽略摄像机镜头的非线性畸变并把透视变换矩阵中的元素作为未知数,给定一组三维控制点和对应的图像点,那么,就可以利用直接线性变换法求解透视变换矩阵中的各个元素。所以,由靶标上特征点的世界坐标和图像坐标,即可计算出摄像机的内外参数。

(2) 基于二维平面靶标的摄像机标定。

该方法又称为张正友标定法,是一种适合应用的新型灵活方法。该方法要求摄像机在两个以上不同的方位拍摄一个平面靶标,摄像机和二维平面靶标都可以自由移动,且内部参数始终不变。假定二维平面靶标在世界坐标系中的 $Z=0$,那么,通过线性模型分析就可计算出摄像机参数的优化解,然后用基于最大似然法进行非线性求解。在这个过程中得出考虑镜头畸变的目标函数后就可以求出所需的摄像机内、外部参数。这种标定方法既具有较好的鲁棒性,又不需昂贵的精密标定块,实用性很强。但是,张正友方法在进行线性内外参数估计时,由于假定模板图像上的直线经透视投影后仍然为直线,进而进行图像处理,这样实际上会引入误差,所以该方法在广角镜畸变比较大的情况误差较大。

(3) 基于径向约束的摄像机标定。

Tsai(1986)给出了一种基于径向约束的两步法标定方法,该方法的核心是先利用 RAC(Radial Alignment Constraint,径向一致约束)条件用最小二乘法解超定线性方程,以求出除 t_z (摄像机光轴方向的平移)外的其他摄像机外参数,然后再在摄像机有和无透镜畸变两种情况下求解摄像机的其他参数。Tsai 方法的精度比较高,适用于精密测量,但它对设备的要求也很高,不适用于简单的标定。这种方法的精度是以设备的精度和复杂度为代价的。

2) 摄像机自标定方法

摄像机自标定方法,是一种不依赖标定参照物,仅利用摄像机在运动过程中周围环境图像与图像之间的对应关系对摄像机进行标定的方法。目前已有的自标定技术大致可以分为基于主动视觉的摄像机自标定技术、直接求解 Kruppa 方程的摄像机自标定方法、分层逐步标定法、基于二次曲面的自标定方法等。

(1) 基于主动视觉的自标定方法。

所谓主动视觉系统,是指摄像机被固定在一个可以精确控制的平台上,且平台的参数可以从计算机精确读出,只需控制摄像机作特殊的运动来获得多幅图像,然后利用图像和已知的摄像机运动参数来确定摄像机的内外参数。

(2) 基于 Kruppa 方程的自标定方法。

该方法是直接基于求解 Kruppa 方程的一种方法,利用绝对二次曲线像和极线变换的概念推导出 Kruppa 方程。基于 Kruppa 方程的自标定方法不需要对图像序列做射影重建,而是对两图像之间建立方程。这个方法在某些很难将所有图像统一到一致的射影框架场合会比分层逐步标定法更具优势,但代价是无法保证无穷远平面在所有图像对确定的射影空间里的一致性。当图像序列较长时,基于 Kruppa 方程的自标定方法可能不稳定,且其鲁棒性依赖给定的初值。

(3) 分层逐步标定法。

近年来,分层逐步标定法已成为自标定研究中的热点,并在实际应用中逐渐取代了直接求解 Kruppa 方程的方法。分层逐步标定法首先要求对图像序列做射影重建,再通过绝对二次曲线(面)施加约束,最后定出仿射参数(即无穷远平面方程)和摄像机内部参数。分层逐步标定法的特点是在射影标定的基础上,以某一幅图像为基准做射影对齐,从而将未知数数量缩减,再通过非线性优化算法同时解出所有未知数。其不足之处在于非线性优化算法的初值只能通过预估得到,而不能保证其收敛性。由于射影重建时都是以某参考图像为基准,所以参考图像的选取不同,标定的结果也不相同。

(4) 基于二次曲面的自标定方法。

这种自标定方法与基于 Kruppa 方程的方法在本质上是相同的,它们都利用绝对二次曲线在欧氏变换下的不变性。但在输入多幅图像并能得到一致射影重建的情况下,基于二次曲面的自标定方法会更好一些,其根源在于二次曲面包含了无穷远平面和绝对二次曲线的所有信息,且基于二次曲面的自标定方法又是在对所有图像做射影重建的基础上计算二次曲面的,因此,该方法保证了无穷远平面对所有图像的一致性。

2. 基于三维立体靶标进行摄像机标定的技术

$$z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (3-29)$$

式中, $(x_w, y_w, z_w, 1)$ 是空间三维点的世界坐标, $(u, v, 1)$ 为相应的图像坐标, m_{ij} 为透视变换矩阵 $\mathbf{M}$ 的元素。整理,消去 z 后,得到如下关于 m_{ij} 的线性方程:

$$\begin{cases} m_{11}x_w + m_{12}y_w + m_{13}z_w + m_{14} - uz_w m_{31} - uy_w m_{32} - uz_w m_{33} = um_{34} \\ m_{21}x_w + m_{22}y_w + m_{23}z_w + m_{24} - vz_w m_{31} - vy_w m_{32} - vz_w m_{33} = vm_{34} \end{cases} \quad (3-30)$$

该方程描述了三维坐标点 $(x_w, y_w, z_w, 1)$ 与相应图像点 $(u, v, 1)$ 之间的关系。如果已知三维世界坐标和相应的图像坐标,将变换矩阵看作未知数,则共有 12 个未知数。对于每个物体点,都有如上的两个方程。因此,取 6 个物体点,就可以得到 12 个方程,从而求得变换矩阵 $\mathbf{M}$ 的系数。

3. 基于径向排列约束(RAC)的摄像机标定技术

基于 RAC 的摄像机标定方法属于两步法,两步法是近年来较为成功的摄像机标定方法,而两步法中以 Tsai 的标定算法最为典型。其标定过程是先忽略镜头的误差,利用中间变量将标定方程化为线性方程求解出摄像机的外部参数;然后根据已经求得的外部参数考

虑实际的透镜模型,利用非线性优化的方法求取径向畸变系数 k 、有效焦距 f 以及平移分量 t_z 。

径向排列约束就是矢量 $\mathbf{L}_1$ 和矢量 $\mathbf{L}_2$ 具有相同的方向,即方向($\mathbf{L}_1$)=方向($\mathbf{L}_2$)。由成像模型可知,径向畸变不改变 $\mathbf{L}_1$ 的方向,因此,无论有无透镜畸变都不影响以上等式。有效焦距的变化也不影响这个等式,因为焦距的变化只会影响 $\mathbf{L}_1$ 的程度而不影响其方向。

$$\begin{cases} x = r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x \\ y = r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y \\ z = r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z \end{cases} \quad (3-31)$$

RAC 意味着存在

$$\frac{x}{y} = \frac{X_d}{Y_d} = \frac{r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x}{r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y} \quad (3-32)$$

整理得

$$x_w Y_d r_{11} + y_w Y_d r_{12} + z_w Y_d r_{13} + Y_d t_x - x_w X_d r_{21} - y_w X_d r_{22} - z_w X_d r_{23} = X_d t_y \quad (3-33)$$

上式两边同除以 t_y ,得

$$x_w Y_d \frac{r_{11}}{t_y} + y_w Y_d \frac{r_{12}}{t_y} + z_w Y_d \frac{r_{13}}{t_y} + Y_d \frac{t_x}{t_y} - x_w X_d \frac{r_{21}}{t_y} - y_w X_d \frac{r_{22}}{t_y} - z_w X_d \frac{r_{23}}{t_y} = X_d \quad (3-34)$$

再将式(3-34)表示成矢量形式,即

$$\begin{bmatrix} x_w Y_d & y_w Y_d & z_w Y_d & Y_d & -x_w X_d & -y_w X_d & -z_w X_d \end{bmatrix} \begin{bmatrix} r_{11}/t_y \\ r_{12}/t_y \\ r_{13}/t_y \\ r_x/t_y \\ r_{21}/t_y \\ r_{22}/t_y \\ r_{23}/t_y \end{bmatrix} = X_d \quad (3-35)$$

其中,行矢量

$$\begin{bmatrix} x_w Y_d & y_w Y_d & z_w Y_d & Y_d & -x_w X_d & -y_w X_d & -z_w X_d \end{bmatrix}$$

是已知的,而列矢量

$$\begin{bmatrix} r_{11}/t_y & r_{12}/t_y & r_{13}/t_y & t_x/t_y & r_{21}/t_y & r_{22}/t_y & r_{23}/t_y \end{bmatrix}$$

是待求的参数。

对每个物体点,已知其 x_w 、 y_w 、 X_d 、 Y_d ,就可以写出式(3-35),选取合适的 7 个点就可以解出列矢量中 7 个分量。用同一平面上的点来作标定,并选取世界坐标系,使 $z_w = 0$,这样,式(3-35)可以简化为

$$\begin{bmatrix} x_w Y_d & y_w Y_d & Y_d & -x_w X_d & -y_w X_d \end{bmatrix} \begin{bmatrix} r_{11}/t_y \\ r_{12}/t_y \\ t_x/t_y \\ r_{21}/t_y \\ r_{22}/t_y \end{bmatrix} = X_d \quad (3-36)$$

利用最小二乘法求解这个方程组。计算有效焦距 f 、平移分量 t_z 和透镜畸变系数 k 时,先不考虑透镜畸变,可线性求出有效焦距 f 和平移矢量 $\mathbf{T}$ 的 t_z 分量,然后利用这些值作初始值,考虑有畸变系数 k 的模型,利用优化算法求解非线性方程组得到 f 、 t_z 、 k 的精确值。

利用式(3-36)和旋转矩阵为正交阵的特点,可以确定旋转矩阵 $\mathbf{R}$ 和平移分量 t_x 、 t_y 。

利用 RAC 方法将外部参数分离出来,并用求解线性方程的方法求解外部参数。

特别地,可将世界坐标系和摄像机坐标系重合,这样标定时只求内部参数,从而简化标定。

4. 张正友标定法

1) 对每幅图像得到一个映射矩阵 $\mathbf{H}$

一个二维点可以表示为 $\mathbf{m} = [u, v]^T$ 。一个三维点可以表示为 $\mathbf{M} = [X, Y, Z]^T$ 。用来表示其增广矩阵: $\mathbf{m} = [u, v]^T$ 以及 $\mathbf{M} = [X, Y, Z]^T$ 。摄像机通常被简化为一个针孔模型,则一个三维点 $\mathbf{M}$ 与其投影图像点之间的关系可以表示为

$$s\tilde{\mathbf{m}} = \mathbf{A}[\mathbf{R}, t]\tilde{\mathbf{M}}$$

其中, s 是任意标准矢量; $\mathbf{R}$ 、 t 为外部参数,是世界坐标系与摄像机坐标系间的旋转与平移关系; $\mathbf{A}$ 矩阵为摄像机内部参数,表示为

$$\mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3-37)$$

式中, (u_0, v_0) 是坐标系上的原点; α 和 β 是图像上 u 和 v 坐标轴的标准矢量; γ 表示两坐标轴的垂直度。

为了不失一般性,假定模板平面在世界坐标系 $Z=0$ 的平面上,则由式(3-37)可以得到

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3-38)$$

式中, $\mathbf{A}$ 为摄像机内部参数矩阵; $\tilde{\mathbf{M}} = [X, Y, 1]^T$ 为模板平面上的点投影到图像平面上对应点的齐次坐标; $[\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$ 和 t 分别是摄像机坐标系相对于世界坐标系的旋转矩阵和平移矢量。

此时,可以得到一个 3×3 的矩阵 $\mathbf{H}$:

$$\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3] = \lambda \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & t \end{bmatrix} \quad (3-39)$$

利用映射矩阵可得到内部参数 $\mathbf{A}$ 的约束条件

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0$$

2) 利用约束条件线性求解内部参数 $\mathbf{A}$

假设

$$\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta} + \frac{1}{\beta^2} & -\frac{\gamma(v_0\gamma - u_0\beta)}{\alpha^2\beta} - \frac{v_0}{\beta^2} \\ \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} & -\frac{\gamma(v_0\gamma - u_0\beta)}{\alpha^2\beta} - \frac{v_0}{\beta^2} & \frac{(v_0\gamma - u_0\beta)^2}{\alpha^2\beta} + \frac{v_0}{\beta^2} + 1 \end{bmatrix} \quad (3-40)$$

其中, $\mathbf{B}$ 是对称矩阵, 可以表示为六维矢量 $\mathbf{b} = [B_{11}, B_{12}, B_{13}, B_{22}, B_{23}, B_{33}]^T$ 。基于绝对二次曲线原理求出 $\mathbf{B}$ 以后, 再对 $\mathbf{B}$ 矩阵求逆, 利用 Choleski 分解, 便可从 $\mathbf{B}$ 中导出内部参数矩阵 $\mathbf{A}$ 。再由 $\mathbf{A}$ 和映射矩阵 $\mathbf{H}$ 计算每幅图像相对于平面模板的外部参数旋转矩阵 $\mathbf{R}$ 和平移矢量 $\mathbf{t}$:

$$\begin{cases} \mathbf{r}_1 = \lambda \mathbf{A}^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 = \lambda \mathbf{A}^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} = \lambda \mathbf{A}^{-1} \mathbf{h}_3 \end{cases} \quad (3-41)$$

3) 最大似然估计

采用最大似然准则(Maximum Likelihood Estimation)对上述参数进行优化。假设有 n 幅关于模板平面的图像, 而模板平面上有 m 个标定点, 那么最大似然估计值就可以通过式(3-42)最小化得到(即制定评价函数):

$$\sum_{i=1}^n \sum_{j=1}^m \| m_{ij} - m(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j) \|^2 \quad (3-42)$$

其中, m_{ij} 为第 j 个点在第 i 幅图像中的像点; $\mathbf{R}_i$ 为第 i 幅图像旋转矩阵; $\mathbf{t}_i$ 为第 i 幅图像的平移矢量; $\mathbf{M}_j$ 为第 j 个点的空间坐标; 初始估计值利用上面线性求解的结果, 畸变系数 k_1 和 k_2 初始值为 0。

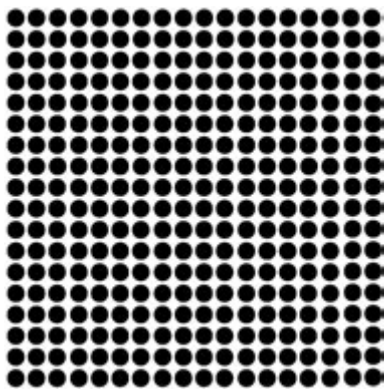


图 3-22 标定用模板

5. 举例

用计算机产生一张有 18×18 个圆形平面图样, 且圆的大小和圆与圆之间的间距都是相同的。然后以 1:1 的比例打印出来, 贴在一块光洁水平的玻璃板上作为标定模板, 如图 3-22 所示。取每个圆形的圆心作为特征点, 则共有 324 个特征点, 选定世界坐标系(如以第一个角点为原点), 便可以得出各特征点的世界坐标。

对标定板图样进行采集。目标图样的采集过程就是通过 CCD 摄像机和图像采集卡等数字图像采集系统将其获取至计算机中成为灰度图像的过程。图 3-23 所示为获得 5 张不同位姿的标定板图像。图 3-23 采集到的

的标定图像利用边缘检测算子, 对图像进行边缘检测, 得到处理后的图像, 如图 3-24 所示。图 3-24 边缘检测图像再取出每个圆圈上的点, 计算其中心位置, 最后对得到的中心点进行排序。将这些提取出的中心点作为参考点, 利用张正友标定法进行标定, 计算出矩阵 $\mathbf{A}$, 以及 k_1 和 k_2 。

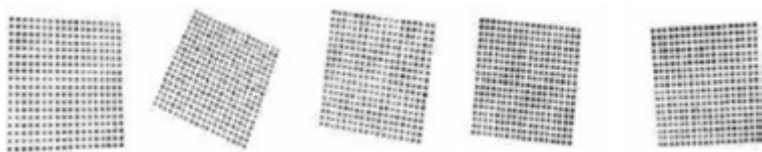


图 3-23 采集到的标定图像

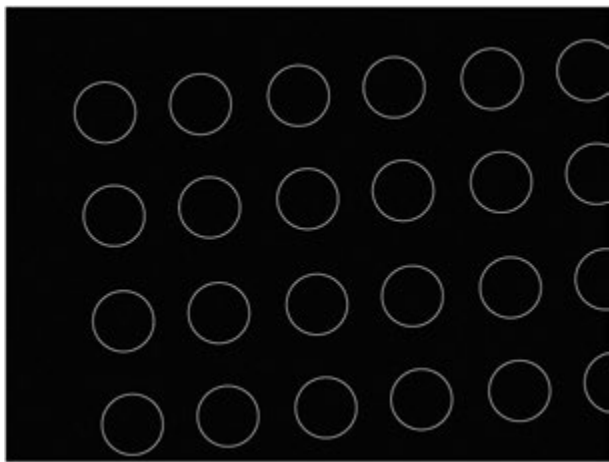


图 3-24 边缘检测图像

将摄像机标定技术应用在机器视觉中,用来检测工业流水线上产品是否为次品。首先是相机标定,有以下两种方法。

(1) 把相机绝对水平放置,然后对一个标准圆拍照,保证成像也是标准圆,记下两个圆的比例关系。以后就可以对成像施以比例变换得到物体的真实面貌。

(2) 相机任意位置放置,对一平面物体图像成像,在相片上取 4 个和物体平面图像对应的点,列方程组解出平面物体成像时的 8 个成像系统的未知常量,然后可以根据变换公式变形复原拍到的真实物体面貌。

第二步即为检测真实物体面貌的主轴。最后,标定任意形状物体主轴后,把拍照的物体主轴和标准物体图像的主轴对齐,计算两幅图的差异,如果差异为 0,那么该产品完全标准;如果差异超过某一值,就可认为是次品。