

# 一种基于机器视觉的智能化芯片和电路板耦合方法、装置、终端设备及存储介质

发明人

## 技术领域

本发明设计图像处理、智能感知及机器视觉等领域，在利用YOLOv7深度学习模型检测芯片目标的基础上，利用边缘检测算法来识别芯片边缘，成功实现基于机器视觉的智能化芯片和电路板耦合，可广泛应用于工业检测、质量监控等场景。

## 技术背景

当代科技的快速发展和应用的历程中，电子制造业在整个科技产业链中都扮演着不可或缺的关键角色。未来，如何进一步提升电子制造的智能化、自动化水平，将成为电子制造业持续发展的重点方向。其中，实现芯片与电路板间的智能化耦合是必须要解决问题，但是由其引起的精确度较低、稳定性较差以及在恶劣环境下鲁棒性降低等一系列挑战都有待解决。

目前传统方法实现芯片的耦合过程通常依赖于人工手动操纵或利用简单的机械装置，但是，不可避免的会遇到难以保持亚毫米级别的精确对准的问题，影响耦合的质量和可靠性。并且，依赖手工操纵不仅效率极其低下，而且难以保证耦合质量的一致性。此外，由于芯片本身脆弱、对环境敏感的特性，恶劣环境情况下使用传统的耦合方法难以满足高可靠性的要求。

随着现代机器视觉技术的发展与进步，我们考虑利用机器视觉技术结合自动化设备来实现智能化的芯片和电路板间的耦合，在机器视觉实现精确的芯片边缘检测的基础上，配合机械手臂、滑轨等自动化设备，可以实现高度自动化、亚毫米级精度的芯片与电路板耦合。这种方式不仅大幅度提高生产效率，而且可以保证所有芯片耦合质量的一致性，满足高可靠性要求。此外，恶劣环境对自动化设备的性能影响较小，系统的鲁棒性较好，解决了传统方法的局限性。

总而言之，该基于机器视觉的智能化芯片电路板耦合技术必然具有广阔的应用前景，对实现电子制造业的智能化、自动化与可持续化发展具有重大意义。

## 发明内容

本发明提出一种基于机器视觉算法的智能化芯片电路板耦合技术，所述方法即通过固定于机台上的相机对芯片和电路板进行实时识别，在利用YOLOv7深度学习模型成功检测芯片目标的大致位置基础上，使用边缘检测算法对芯片的边缘进行检测，通过三维重建进一步得到更精确的位姿，进而控制机械手臂、滑轨等自动化设备实现高度自动化、亚毫米级精度的芯片与电路板耦合。

本发明提出方法的实现主要包括三个模块，其分别为：感知设备转换模块、芯片检测计算模块与相机标定重构模块。感知设备转换模块利用相机等设备获取芯片目标所处区域图像，并将其转换为待处理的数字信号以便于后续进行芯片的位置检测与计算。芯片检测计算模块为由 GPU 和 CPU 组成的小型计算机，首先需利用YOLO深度学习模型快速精确地定位出芯片在图像中所处的大致位置坐标，为后续对芯片的边缘精确检测提供关键信息；其次，利用边缘检测算法对预处理的图像信息进一步处理，进而成功检测得到芯片边缘。相机标定重构模块由标定板和摄像头组成，使用张正友标定法建立真实三维空间与图像像素坐标系之间的映射关系，结合利用成像原理，将二维图像中的像素坐标转换为真实世界的三维坐标，同时对相机拍摄过程中出现的各种光学畸变进行精确建模和校正，使得图像获得更高的几何精度，最终便可以成功实现亚毫米级别的精确度。

本发明提出的智能化芯片和电路板耦合技术，相较于传统利用人工手动操纵的方法，用智能化工厂的形式代替人力的参与，可以大大提高芯片与电路板的耦合效率，且在避免人为操作失误造成的影响同时，保证所有芯片耦合精度的一致性。此发明可降低电子制造业的整体成本，便于推动电子制造业的快速发展。

## 技术核心

基于机器视觉算法的智能化芯片和电路板耦合方法，其核心主要集中于芯片检测计算模块与相机标定重构模块，更准确地说，是集中于利用YOLOv7深度学习模型检测芯片目标、利用边缘检测算法检测芯片边缘与利用张正友标定法实现二维图像中的像素坐标与真实世界的三维坐标之间的相互转换这三个主要部分。

当前的目标检测算法主要分为单阶段模式与双阶段模式，其中单阶段模式算法是指以YOLO深度学习模型为代表，通过一个单独的神经网络便可以直接输出图像中目标的位置及其对应类别，其由于网络简单，易于优化，以及端到端的训练模式可以保证其运算速度。对本项智能化芯片和电路板耦合技术而言，存在

速度响应的需求,故选取单阶段模式算法的代表YOLO深度学习模型作为实现算法,满足项目实时性的要求。而本发明选取YOLO算法的改进、优化版本YOLOv7深度学习模型作为目标检测的算法,引入多尺度注意力机制以便于网络适应不同尺度目标,并通过数据增强和硬负样本挖掘等策略来提高模型的泛化能力。由于其快速的检测速度和较高的检测精度,YOLOv7深度学习模型在实际应用中具备明显优势。故智能化芯片和电路板耦合技术利用该算法模型,在速度、精度和实时性等方面都具备巨大优势。

相机标定算法可以通过确定相机内部、外部参数,从世界坐标坐到相机的像素点需要经过三次坐标转换,为世界坐标系→相机坐标系→图像坐标系→像素坐标系。在本发明中,需要实现其逆过程来实现坐标转换。世界坐标系为在世界中固定的一个三坐标轴的坐标系,其单位为绝对长度。本发明将世界坐标系直接设立在标定板的左上角。相对的,相机坐标系的坐标则建立在相机的光心原点,其Z轴与光轴重合。从世界坐标系转换到相机坐标系需要进行一次刚体变换,即进行一次平移和旋转的变换,故外参是描述相机当前相对位置的一个平移向量和旋转矩阵。相机的内参则表征着每个像素在相机CMOS传感器中的绝对长度和焦距等,是绝对长度变换为像素长度的关键参数。求解相机内外参数具有重要意义。

目前主流的相机标定算法有五类,即:张正友标定法、基于稀疏特征点的标定、基于深度信息的标定、实时在线标定和基于深度学习的标定。其中张正友标定法目前最为广泛运用,智能化实现芯片和电路板的耦合过程中,相加标定的环境并不复杂,相机的坐标也固定。所以,此时使用计算简单、鲁棒性强、标定精度可达亚像素级和目前算法流程较为成熟的张正友标定法最合适。该相机标定算法在预先测量得到的标定板特征点的三维坐标的基础上,采集标定板在多个不同角度和距离的图像,利用非线性最小二乘优化方法来优化目标函数在图像上的重投影误差,最终输出得到相机焦距、主点坐标、畸变系数等相机内参数。其中估算得到的相机镜头畸变参数可以对图像进行校正,世界坐标系中的位置和姿态以便于后续的目标定位与空间重建,为智能化芯片和电路板耦合装配过程提供强有力的技术支持。

边缘检测作为图像分割、纹理分析和图像识别等技术所依赖的重要基础,其需提取周围灰度强度有反差变化的像素几何作为物体边缘。在具体实现上,由于本发明的识别目标在灰度值上存在明显的分层,为了保证运算的简便性和实时性。我们首先对输入图像通过设定一定阈值的二值化,对于种类不同的芯片,可以在识别类别后进行阈值的自调节,从而保证边缘算法识别的稳定性。在进一步寻找轮廓时,本发明采用OpenCV库中的findcontours()函数来进行查找,其功能多样,在进行轮廓提取时,可以在外围轮廓,所有轮廓,具有父子关系等等的轮廓之间进行选择,保证了轮廓选取的灵活性,对于不同型号的芯片都具有很强的适应性,并可以把检测的轮廓全部返回为点的集合,在后续的操作提供了便利。

## 实现方法

针对基于机器视觉的智能化芯片和电路板耦合技术，实现过程主要为感知设备转换模块、芯片检测计算模块与相机标定重构模块三个主要模块的一体化整合，其主要包括：

步骤一：安装相机等感知设备模块，其中相机的安装角度需保证拍摄画面中可以清晰观察到芯片和电路板，并选择合适的拍摄距离防止过近导致的拍摄视野过小或太远导致图像清晰度受到影响。尽量稳定芯片和电路板的位置，并让两者位于画面中央避免边缘畸变的产生。调整室内灯光的照明角度，并检查周围环境的反光现象，尽可能减少直射光源造成的反光问题。

步骤二：确认实验环境符合要求后，[放置标定棋盘，确定世界坐标系，进行相机外参标定，求解出当前相机坐标轴相对于世界坐标轴的位移、旋转向量。](#)之后打通摄像头和后台主机间的通信链路，并执行芯片检测计算模块，利用YOLOv7深度学习模型快速精确地定位出芯片在图像中所处的大致位置坐标，得到芯片的大致位置，然后使用边缘检测算法对目标芯片的边缘进行检测并计算得到芯片的中心点[像素坐标](#)。

步骤三：利用芯片检测计算模块得到的相关芯片坐标信息，通过相机标定重构模块建立真实三维空间与图像像素坐标系之间的映射关系，[进行深度估计，并将二维图像中的像素坐标转换为真实世界的三维坐标](#)，并对相机拍摄过程中出现的各种光学畸变进行精确建模和校正，以便于在控制机械手臂、滑轨等自动化设备时实现亚毫米级别的精确度。

以上步骤中，在利用YOLOv7深度学习模型快速定位芯片位置坐标时，需要利用感知设备转换模块采集大量相机在不同角度和距离拍摄芯片和电路板得到的图像数据，利用人工进行标注，构建符合目标的训练数据集，而数据集的采集和制作过程会对最终YOLOv7深度学习模型进行目标识别的效果产生影响。同时，在获取得到芯片在图像中所处的大致位置坐标后需对图像进行放大处理后，再利用边缘检测算法对目标芯片的边缘进行检测，这样才可以在相机像素精度受限的情况下尽可能实现亚毫米级别的精确度。

## 技术原理

下面以芯片边缘检测为例，阐述本发明的技术原理。该部分将 YOLO 目标检测算法、按相机标定法原理、边缘检测算法与多边形拟合，[深度估计与坐标重建](#)顺序阐述：

### 1、YOLO 目标检测算法原理与网络训练

本发明采用的目标检测算法是 YOLO 目标检测算法。YOLO 目标检测算法的核心思想是将目标检测任务转化为一个回归问题，通过将图像分成网格，并同时预测每个网格中的目标的边界框和类别信息。在本发

明中利用 YOLO 目标检测算法初步估计芯片的位置。

首先，YOLO 目标检测算法需要提取固定摄像头所拍摄的芯片视频图像信息，用于目标的位置预测。之后这些图片要经过单一的前向传播卷积层网络，进行下采样以及特征提取，形成多个不同尺度的特征图。

接着，我们**创新性地**引入多输入堆叠结构进行特征提取，以实现数据增强。首先，我们将不同尺度的特征图置于 SPPCSPC 模块处理，以增加感受野。随后，将特征图进行采样率变化并与不同尺度的特征图进行卷积融合和堆叠融合，实现特征的融合和加强，便于后续预测。

之后，需要将特征层传入 Yolo Head 获得预测结果。在此，我们**创新性地**引入了 RepConv 结构的前端处理。通过引入特殊的残差结构辅助训练，提高了网络的预测性能。接着，将经过 RepConv 网络处理的特征图接入不同输出通道数的 Yolo Head 卷积网络，得到不同尺度的目标检测结果。

进一步的，需要对预测结果进行解码操作。在本发明中，我们首先对三个尺度的输出进行了解码，得到了每个尺度下预测框的位置、置信度和类别概率。之后通过 sigmoid 函数将预测的中心坐标和宽高调整参数转换成实际的相对值。并根据当前网格单元的位置和锚框的大小，计算出预测的边界框坐标。

而对于 Yolo Head 卷积网络需要大量数据集训练，在网络训练过程，我们突破性地利用自适应多正样本匹配模型，通过计算真实框与先验框的 IoU 和种类的 cost 确定正样本和匹配框的种类信息。之后我们根据正样本的匹配情况，构建用于训练的目标。接着，通过对比预测框和真实框的匹配情况，构建分类损失函数，回顾损失函数以及置信度损失函数。联合优化这三个部分的损失，使得模型在训练过程中逐渐提高对目标的分类、定位和置信度预测的准确性，从而提高物体检测的性能。

在实际预测过程中，由于网络需要保证输入图片的大小相同，我们将其像素长宽约束为一个 32 的倍数，对每一个视频帧都进行一次 resize，并使用灰条来填补图像大小的差异部分。

为了保证算法的鲁棒性，我们仅接受置信度  $> 0.6$  的预测框，同时还引入了非极大值抑制，对于视频流中多个可能的预测框，我们只保留了置信度最高的预测框，保证每个芯片的识别框的唯一性。由于 YOLO 识别框对于识别的芯片而言尚有空缺，我们对不同的类别分别进行了不同程度的截取和缩放，故边缘检测结果可以直接对更小的图片进行识别，显著优化了识别速度以及边缘检测效果。

值得一提的是，对于类别的分类可根据业务所需进行定制，比如对于芯片边缘检测而言，根据有无芯片，可以分成两类。也就是说，目标检测算法可以提供芯片是否在视野范围内的状态检测功能，提高算法的鲁棒性以及实用性。

## 2、相机标定方法

为了便于机械臂拾取芯片，需要进行相机标定，将二维的图像坐标  $(u, v)$  转化为世界坐标  $(X_w, Y_w, Z_w)$ ，这样才能真正由机械臂进行抓取。本发明采用了张正友相机标定法，旨在计算相机内参，相机外参以及畸变系数，便于后续直接计算出图像芯片中心点和边缘点在世界坐标系下的位置。

在本发明中，在平面标定板上可以方便的建立世界坐标系，此时可以设置  $Z_w=0$ ，图像坐标系的坐标与世界坐标系的坐标满足以下关系：

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ 1 \end{bmatrix}$$

其中等式右方第一项为内参矩阵，R, T 分别表示旋转矩阵和偏移向量构成外参矩阵，而  $X_w, Y_w$  表示世界坐标系下的坐标。

将内参和外参组成的矩阵设置为单应性矩阵 H，通过求解

在相机内参的求解中，我们利用到了棋盘格标定板，标定板上的空间点需要通过一个单应性映射矩阵与像素点建立关系，通过 cholesky 分解最终得到相机内部参数。在不同角度和距离拍摄图片并在每张图片中选择多个角点，计算其在世界坐标系中的位置，并且通过 Harris 角点检测算法计算对应角点的图像坐标系坐标，我们使用了 OpenCV 中的 CV.findchessboardcorner() 函数来进行棋盘参数的查找，为了提高精度，



我们进一步地选择亚像素角点，保证棋盘的各点标注精确。根据标定原理，当用于标定的图片数量大于等于 3 时，则可以得到唯一的内参矩阵；

接下来，在求解完内参矩阵后，输入标定板角点的世界坐标系坐标，利用内参矩阵与外参矩阵求解该角点在图像坐标系下的位置，并与图像中理想点坐标对比，构建重投影误差函数。并且利用最大似然优化，通过优化内参矩阵和外参矩阵，使得重投影误差取最小值，使得初步求解的内参矩阵和外参矩阵得到优化。

由于场景的限定，相机的位置是固定的。拍摄的照片来自不同的角度和距离，只能求得最小重投影误差的外参，为保证算法的鲁棒性以及适应本场景，我们需要一个稳定的唯一的外参，在开头进行的相机标定得到的外参为拍摄的图片，在固定摄像头后，我们还需要得到目前相机相对于探测区域的外参（位移，旋转等参数）。因而当前使用还需要使用 PnP 算法求解单应性矩阵，计算外参。

PnP 算法是一个求解给定  $n$  个 3D 空间参考点，以及各点在相机图像上对应的成像点，求参考点所在坐标系与相机的空间关系的算法，其基于 Levenberg-Marquardt 优化的迭代方法，使重投影误差最小化。通过代入内参中的畸变系数和内参矩阵。我们将世界坐标设定在标定棋盘上，由于标定棋盘的各个方格距离均等，通过求解深度，我们可以直接得到各个角点的世界坐标  $X_w, Y_w$ ，再通过像素点的  $u, v$  坐标可以求得一个外参矩阵，也就是相机当前对于世界坐标系的位姿。我们得到了相机的外参  $R, T$ ，内参  $K$ ，畸变系数  $dist$ 。至此，完成对相机的标定，完成了对相机内参，畸变系数的计算。

### 3、边缘检测算法与多边形拟合

根据 Yolo 算法返回的值，本发明首先通过划分阈值进行二值化，其结果已经形成了较为准确的边缘。再使用 `cv.findContours()` 函数得到所有边缘点的集合，计算出边缘中心点的  $u, v$  坐标。注意到，如果单纯使用边缘检测会导致边缘产生大量的毛刺和抖动，不利于后续进行信息提取和进一步操作。因而使用 `approxPolyDP` 函数进行四边形的拟合，其原理是对得到的所有相邻边缘点进行链接得到一条直线，计算原本边缘曲线对该直线的最大距离，如果小于阈值，就使用该直线来作为边缘曲线的拟合，将边缘的凸起化为平整的直线。在本实验场景中，我们选择只接受最后返回四个角点，对四个角点进行首尾链接，我们就得到了一个逼近芯片中心轮廓的四边形，观察到其稳定度上升，便于后续的深度计算。

至此，利用目标检测算法和边缘检测算法，完成了对摄像头画面中芯片与电路板的像素位置估计，得到边缘与中心点的  $u, v$  坐标。

### 4、深度估计与坐标重建

在本发明中，可对芯片目前的深度进行估计，将  $Z$  轴坐标反馈于机械臂可以便于机械臂进行贴合操作。

首先，基于相似三角形的变换依赖于边缘的像素长度，由于芯片的四个边缘会因为观测角度的不同，在像素上的长度会有所不同，因而在提取像素坐标之前，需要经历一次仿射变换，将芯片的四个角度转化为较为平整的  $90$  度。若不如此，就只能在靠近相机的中央部分得到比较准确的结果。我们可以使用 PnP 算法返回的结果来进行角度变换，保证长宽对应一致。

我们使用成像原理将相机距离元件的深度进行计算，由小孔成像原理得两个相似三角形：

$$\begin{aligned}x &= \frac{X}{Z_c'} f \\y &= \frac{Y}{Z_c'} f\end{aligned}$$

$Z_c'$  为元件距离元件的深度， $X, Y$  为芯片的绝对尺寸， $x, y$  为芯片在像素坐标上的像素尺寸。由于芯片长度在世界坐标系下和相机坐标系的长度相同，且相机外参给定了相机距离世界坐标的距离  $Z_c$ ，我们可以计算  $Z_c'$ ，元件的实际高度  $Z_w$  为：

$$Z_w = Z_c - Z_c'$$

由于感兴趣区域的像素较低，轻微变化的边缘就可能造成  $Z$  轴的较大扰动，因而我们对元件的四个边缘都

进行一次计算并求其均值，同时为了保证贴合时 Z 轴的稳定程度，对于视频流的每一帧，将每一次的输出结果推入一个容量为 30 帧的队列，对 Z 轴数值进行一次平滑，保证 Z 轴坐标的平稳性。  
因而最后求得的 Z 轴坐标为：

$$Z_W = Zc - average(\sum \frac{X}{x} f_x + \frac{Y}{y} f_y)$$

其中Zc'为物体距离相机的距离，l 为归一化坐标系转换为绝对坐标的系数，其数值取决于使用的标定板的绝对长度，本发明中，其取值为 10mm。

在仅仅已知像素坐标 u,v 的情况下，会求出无穷个世界坐标的解，但如果加入限定Z<sub>W</sub>的条件(已经由上文计算得出)，所以Z<sub>W</sub>已知的条件下，可以代入逆向求解世界坐标的公式：

$$\begin{bmatrix} X_W \\ Y_W \\ Z_W \end{bmatrix}^T = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}^{-1} \left( Zc' \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - l \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} \right)$$

我们仅需解得 x<sub>w</sub> 与 y<sub>w</sub>。至此，我们得到了目前芯片的三个世界坐标，可以将其输入机械臂进行下一步操作。本发明所述的基于机器视觉的智能化芯片和电路板耦合方法的技术原理已阐述完毕。

附图

