

上海交通大学， 2019-2020 学年第 1 学期

《计算机视觉》课程报告

题目： 组装线上电脑主板的框定

学生姓名： 李涛

专业： 自动化

班级： F1603201

学号： 516030910010

任课教师： 赵旭

日期： 2019-12-14

目 录

一、项目介绍.....	1
二、算法介绍.....	1
2.1 主板待匹配区域确定	1
2.2 特征点匹配及改进.....	4
2.2.1 sift+RANSAC 算法 ^[1]	4
2.2.2 refine 算法	5
2.3 边界框定	6
2.4 多块主板的处理	7
三、实验结果.....	8
3.1 sift+RANSAC 算法	8
3.2 refine 算法	10
四、项目总结.....	11
参考文献.....	12

一、项目介绍

在实际的电脑组装产线上，需要实时检测视频中产线上的电脑主板，用来计算工人的工序执行情况和操作时间。本项目的目的是利用主板的模板图像，设计算法检测出视频截取图像帧中的电脑主板，并利用边界框出。

在本项目中，我们提出了一种通过边缘检测和形态学操作得到电路板待匹配区域的算法，使用比较成熟的 sift+RANSAC 算法进行图片的配准，针对某些图片框定不准确的情况提出了一种 refine 算法，并处理了多个电路板的情况，使得在所有视频帧中的工人手头的

所有主板都得到了精准的框定。

二、算法介绍

我们在本项目中使用的算法主要由以下几个部分组成：主板待匹配区域确定、特征点匹配及改进、边界框定和多块主板的处理。

2.1 主板待匹配区域确定

原图尺寸很大，通过 sift 得到的特征点极多，直接拿整张图同模板做特征点匹配显然是不合适的，因为这样会造成大量的计算浪费同时也会出现很多错配造成干扰。我们希望截取出只包含主板的小区域同模板进行匹配，这样比较高效和准确。我们观察到主板上一个明显的特征是上面密密麻麻分布着很多电路元件，考虑通过边缘检测响应应该比较大，边缘检测的效果如下：



图 1 Canny 边缘检测的响应

可以观察到通过 Canny 边缘检测电路板区域响应很密集而且区域最大。我们接下来将采用形态学操作来抠取出电路板区域：首先我们希望将那些稠密的响应相互连通形成一整块区域，故我们采用 3×3 的全 1 模板对得到的边缘响依次应进行膨胀和闭操作：



图 2 对得到的边缘响应进行膨胀



图 3 对膨胀后的结果进行形态学闭操作

可以观察到电路板区域基本上已经粘合成一个整体了，于是我们可以采用形态学开操作把背景复杂的线条去除（由于背景的各种线条要么水平的宽度很小，要么竖直的宽度很小，而电路板区域水平、竖直的宽度都很大，故我们采用 1×15 和 15×1 两种模板进行形态学开操作）：

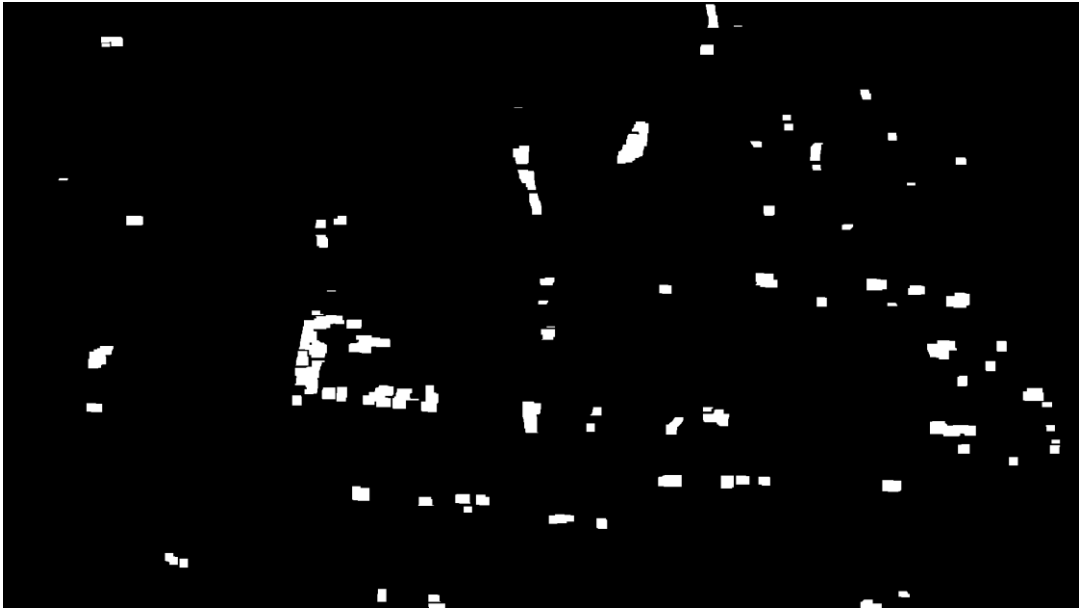


图 4 对形态学闭操作后的结果进行形态学开操作以消除背景

此时整张图片中的人、生产线等都已经消失，可以明显的观察到电路板区域有比较集中的白色块存留，面积最大。我们采用形态学闭操作消除块间的缝隙（采用较大的模板 35×35 ）：



图 5 采用形态学闭操作消除电路板区域的缝隙

此时我们可以观察到电路板主体区域已经联通，且面积最大。我们使用 opencv 的 `findContours` 函数得到所有联通区域的轮廓，选择面积最大的轮廓即为主板电路区域，最后我们使用外接矩形框定：

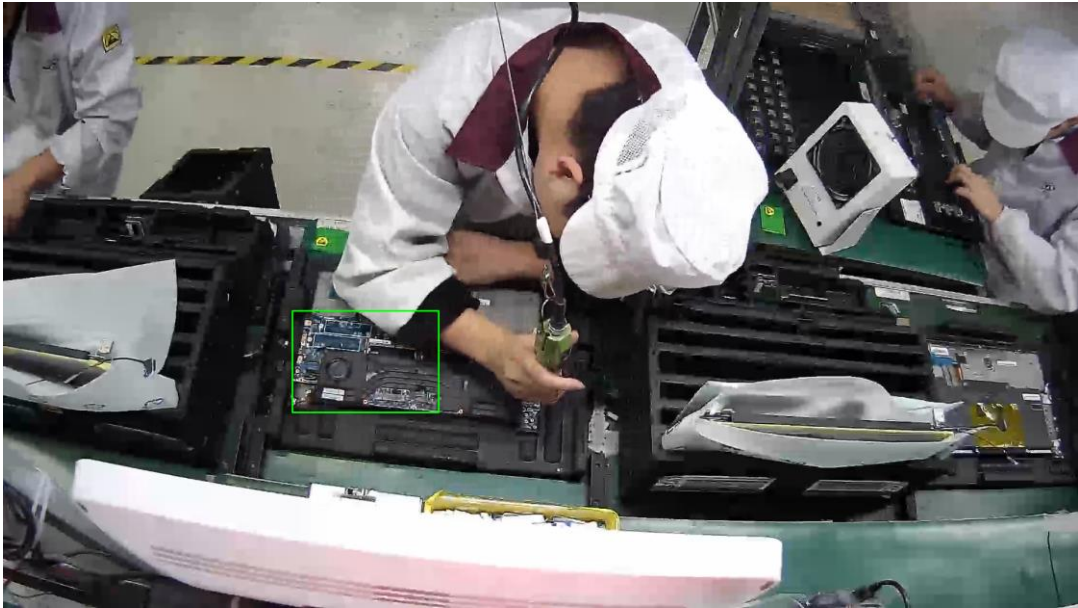


图 6 根据最大联通区域外接矩形框定结果

上图中绿色的矩形即为我们选择的待匹配区域。虽然由于手遮挡和主板上半部分比较光滑响应较小等原因不能将整块主板都圈定，但这一小块区域已经足够同模板进行匹配了，而且另外一个好处是该区域绝大部分都是真正的主板区域，而且是主要有效的主板区域，没有背景物体的干扰，有利于减少背景区域错误匹配带来的干扰。

总结一下我们确定主板待匹配区域主要有以下几个步骤：

- 1) 对原图使用 Canny 边缘检测
- 2) 膨胀加粗电路板元件的响应
- 3) 使用形态学闭操作使得电路板元件部分融合
- 4) 使用形态学开操作消除背景线条
- 5) 使用较大模板进行形态学闭操作使得电路板完全融合
- 6) 求取出最大的轮廓，使用外接矩形进行框定，得到最终待匹配区域

2.2 特征点匹配及改进

接下来我们要对待匹配区域和模板进行配准，以获得从模板到视频帧图片中真实电路板的变换矩阵，从而将视频帧图片中的主板框定出来。我们将分为基础的 sift+RANSAC 算法和在此基础上的 refine 算法进行介绍。

2.2.1 sift+RANSAC 算法^[1]

我们首先对待匹配的图片 and 模板图片分别使用 sift 算法获得特征点和相应的特征向量，然后使用 opencv 自带的 kdtree 算法对每一个待匹配的图片中的特征点在模板图片中找到特征向量最近的两个特征点，如果最近的距离和次近的距离之比小于一个比例系数，那么我们认为最近的点和该特征点的匹配是可靠的，下面展示的是系数取 0.8 时候的结果：

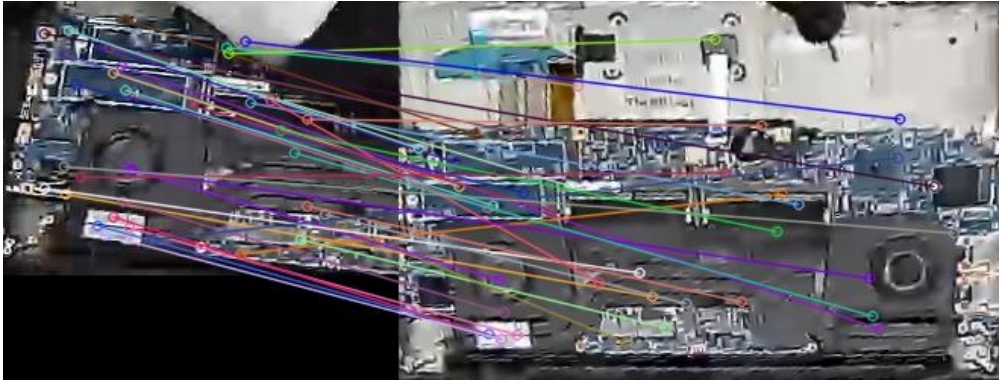


图 7 待匹配区域同模板的配准（系数取 0.8）

我们观察到上图中大部分的点匹配是比较一致的，但也有很多匹配错误的点（如白色的手套和模板上面白色的主板部分容易错配）。直接通过所有匹配点对拟合变换矩阵由于很容易受匹配错误的点对的影响因而偏差很大。我们采用 RANSAC (random sample consensus 随机一致性采样) 来剔除这些误匹配点对，过程如下，选取四对匹配点（因为最少四个点才能描述变换过程，四个点有八个自由度，而射影变换也有八个自由度）得到一个变换矩阵 M ，用 M 和一个误差度量函数与计算当前满足这个变换矩阵的一致集的数量，如果一致集超过了给定的数量，就停止迭代。

2.2.2 refine 算法

在实际测试后我们发现对于遮挡比较复杂的情况下（如双手遮挡、黄色工具箱遮挡等情况）使用上述 sift+RANSAC 算法边界不能够精准的定位。在经过反复的实验、观察和总结后我们对上述算法做了一定的 refine 使得在定位不准确的四张图片上也得到了令人满意的效果。

我们注意到假设把 sift 可靠匹配的比例系数改小，那么得到的匹配就会很可靠。如下图取比例系数为 0.6，得到的 6 个匹配都是正确的。但是如果使用这 6 个匹配来拟合出变换矩阵（实际上四个匹配理论上就足够了），得到的结果却很糟糕，因为这些看似正确的匹配实际上是有微小的误差的。如下图：



图 8 全部正确的匹配（系数取 0.6）



图 9 用全部正确的匹配拟合出来的结果

由此我们可以得出一个结论:一个准确的匹配需要大量的可能不是那么准确的匹配来共同拟合,光靠少量的完全正确的匹配是得不到精确的框定的。我们基于此观点,首先使用较小的比例系数来拟合出一个比较“粗糙”的模型。然后使用较大的比例系数,此时具有很多很多的良莠不齐的匹配,我们根据之前得到的模型粗略地筛选掉一些错误的匹配,最后在剩下的匹配中重复 RANSAC 算法提炼出一个 refine 的模型。这样的好处是可供我们拟合的匹配数目大大增加了,从而能拟合出更准确的框定。

假设两对匹配是正确的,那么当待匹配图和模板分开摆放时,对应点和对应点之间的连线应该是近似平行的。我们首先计算出在“粗糙”的模型中对应点连线倾斜角的平均值,然后根据这个平均值设定一个阈值来筛选掉错误的匹配。

具体的对比将在实验结果一节详细介绍。

2.3 边界框定

我们得到模板图片到待匹配区域的变换矩阵 M 后,将模板图片四个顶点的坐标在射影坐标系下分别与矩阵 M 相乘,再加上截取的待匹配区域在原图中的坐标,即可得到原图中主板四个顶点的坐标,并将其在原图中框定出来,效果如下:

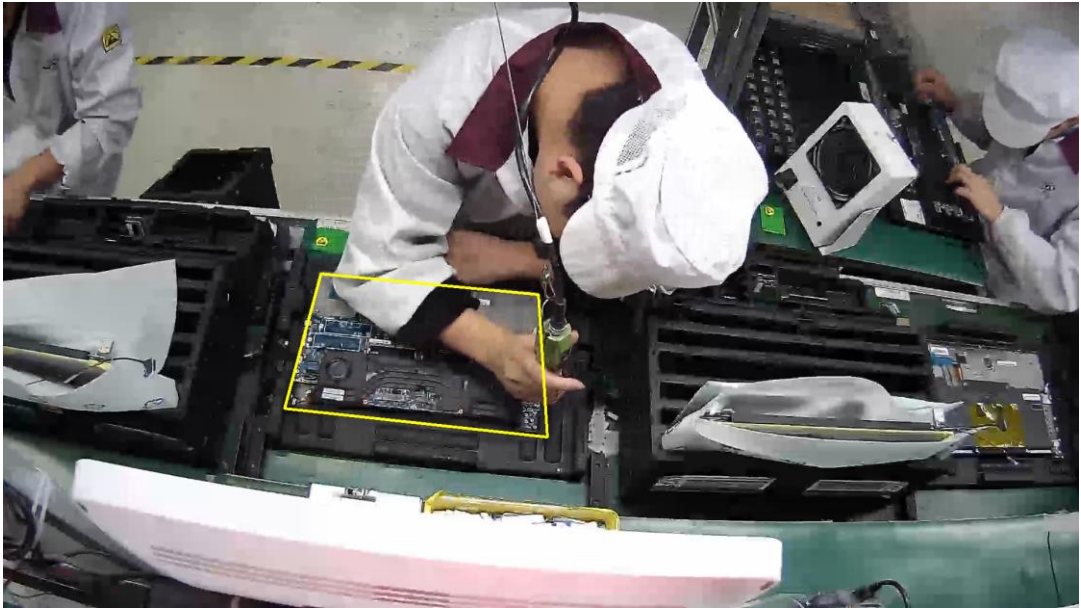


图 10 最终框定效果图

可以看到虽然有手的遮挡，但框定的结果还是相当的准确的，这也说明只使用部分的电路板图片也足够恢复出整个的变换矩阵。

2.4 多块主板的处理

最后处理视频帧中可能存在多块主板的情况。由于之前框定主板待匹配区域是根据最大轮廓面积的外接矩形来确定的，在框定完前一个主板后，我们可以方便地将二值图中该矩形区域置成全黑，然后继续寻找下一个最大面积轮廓。只要设定轮廓面积的一个阈值，我们就可以重复上述框定过程，轻松地框出所有轮廓。

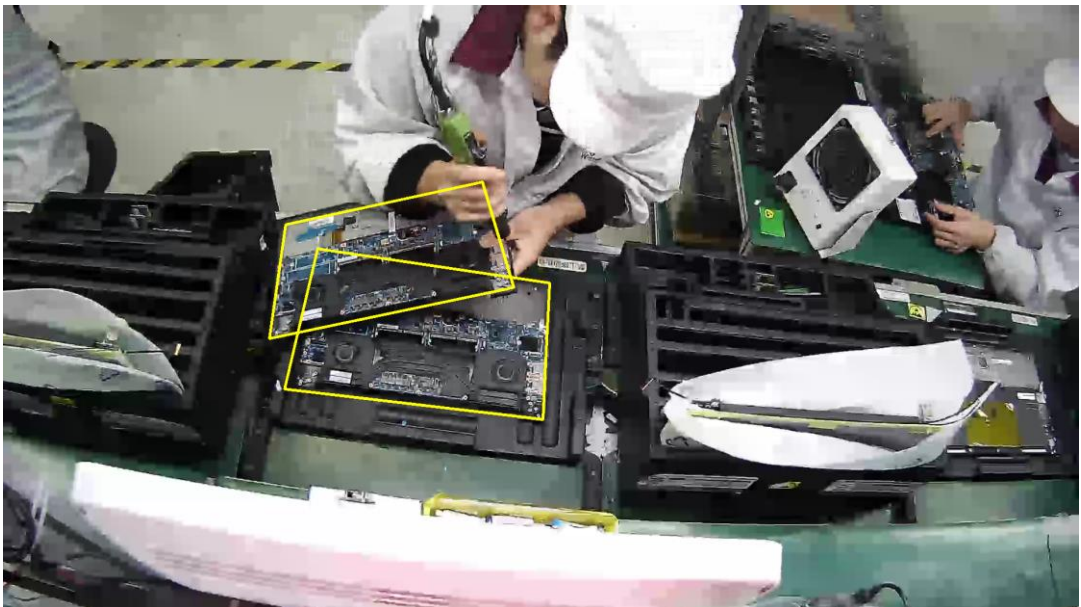


图 11 同时框定图片中多块主板

三、实验结果

首先由于给定的主板模板和视频帧中装配线上实际的主板图片清晰度、细节等差异很大，直接使用该模板进行匹配效果不好，故我们从给定的视频帧图片中挑选了一张遮挡较小的（8 螺丝枪轻微遮挡.png），通过 opencv 中的透视变换函数将图片中的主板截取出来，获得了一张与实际接近的模板图片（如下图），在之后我们将使用该图作为我们匹配的模板。

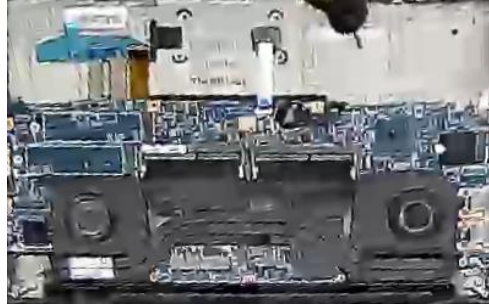


图 12 重新制作的主板模板

3.1 sift+RANSAC 算法

我们首先使用基本的 sift+RANSAC 算法在**所有参数固定**（比例系数取 0.8，RANSAC 的 ransacReprojThreshold 取 10）的情况下对所有图片进行测试，在前八张图片有 6 张图片得到了边界精准的定位，



图 13 sift+RANSAC 算法前 8 张图精确定位

前八张图中另外两张图边界框定有轻微的偏差：



图 14 sift+RANSAC 算法前 8 张图轻微偏差

对于后面五张图匹配有三张图片框定失败，两张图可以精准框定，通过观察发现主要原因是后边图片中的电路板由于已经完成了一部分工序，部分电路元件的颜色相较于模板发生了改变，即视频帧中的电路板和实际模板不同，如下图我们可以观察到左边图中两个标签的颜色（白色）和右图模板中的颜色（蓝色）不同：

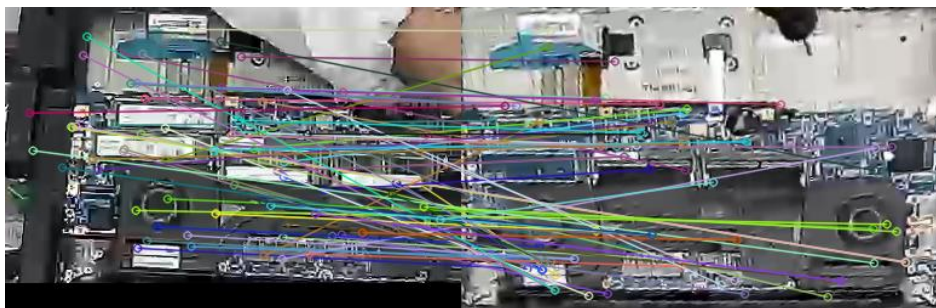


图 15 实际电路板和模板电路板不同（标签颜色）



图 16 sift+RANSAC 算法后 5 张与模板不一致的图片中精确框定的 2 张

在我们重新更换模板后再使用原来代码进行测试：



图 17 后 5 张图片使用的新的模板



图 18 更换模板重新测试（1 张精准匹配，2 张有些偏差）

我们发现之前匹配失败的最后 3 张图片中有一张图片得到了精确定位，另外两张图片定位有一些偏差。综上所述我们可以看出，sift+RANSAC 算法在图片存在遮挡、透视变换等复杂情况下具有令人满意的效果，在测试的 13 张图片中，电路板基本能够匹配成功，其中 9 张图片得到了非常精确的框定，4 张图片的框定存在些微的偏差。

3.2 refine 算法

我们首先在比例系数为 0.7, ransacReprojThreshold 为 4 的基础上通过 RANSAC 得到了一个不准确的匹配，计算对应点的斜率并设定斜率误差系数为 0.1，然后在设定比例系数为 0.9，并根据斜率进行初步筛选，得到的匹配如下：



图 19 进行初步筛选的匹配

注意到通过这个筛选，把绝大多数错误的匹配过滤了，但仍然存在错误的匹配。我们在此基础上再次进行 RANSAC 算法。我们在之前匹配不准确的四张图片上进行测试：



图 20 refine 算法改进效果对比图

可以观察到改进后四张图片都得到了令人满意的框定效果。至此我们完成了所有图片所有电路板的准确框定。

四、项目总结

在本次项目中我们主要要实现利用已知主板模板,对生产线视频帧中工人手头的主板进行框定。我们首先通过边缘检测和形态学操作等算法对视频帧中有效的待匹配区域进行选定,然后通过 sift+RANSAC 算法实现了对图片中的主板和模板的配准,针对遮挡较严重配准不准确的情况,我们给出了一个 refine 算法:利用较少的可靠匹配点对得到一个“粗糙”的模型对较多的点对进行初步筛选,再在较多的点对中重新使用 RANSAC 提炼出一个准确的模型。最后我们处理了图片中多块主板的情况,完成了所有图片中工人手头所有主板的准确框定。

本项目中不足之处有虽然 refine 算法在原先匹配不准确的四张图片中取得了优异的效果,但由于使用了多次 RANSAC 和斜率阈值参数选择等导致参数数量大大增加,模型的鲁棒性相较于原先有所下降,对于不同的图可能需要调整参数来取得最佳的效果。这也反映了准确性和鲁棒性之间存在一定的矛盾。另外,在筛选掉错误的匹配中本项目只使用了斜率要接

近的标准，这在某些情况是不可靠甚至错误的，可以进一步完善和优化。另外针对 refine 算法对参数更加敏感的问题也值得在下一步进一步思考和改进。

参考文献

[1] Lowe D G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110.