

说明书

基于 LDCF 行人检测算法的优化

5130309612_李远

技术领域

本创新工作涉及的是一种行人检测领域的方法，具体是一种基于 LDCF 行人检测算法的优化和改进方法。

背景技术

行人检测是目标检测中的一项热点和重要内容，目标是在灰度、彩色、深度等图像或视频中找到直立行走的行人，在汽车辅助驾驶、智能交通管理、智能视频监控等方面有着广泛的应用。

目前，行人检测算法大体上有三类：DPM 类、决策森林类和深度学习类。其中决策森林类算法更具优势。LDCF 算法可以从训练集中提取颜色、梯度方向和梯度幅值三种特征信息，训练成分类器。在检测行人时，则采用缩放图像构成尺度金字塔的策略，使用决策树进行多尺度检测。

经过对现有技术的文献检索发现，Piotr Doll'ar 等人在“Pedestrian Detection: An Evaluation of the State of the Art”论文中提到了行人检测的一些几何特征，比如行人的宽度与高度的关系，行人的高度与其距离摄像头之间距离的关系等。可以利用这些几何特征，在行人检测的过程中加入几何约束的条件，从而提高 LDCF 检测的正确率。而 Piotr Doll'ar 等人在另一篇论文“Integral Channel Features - Addendum”中，则提到了利用 Non-Maximal Suppression（非极大抑制），来去除重合的检测框并降低 miss rate（失效率）。然而，非极大抑制也可以运用到大小差异较大、部分重合的检测框的去除中，更好地提高算法的正确率。

创新工作内容

本创新工作针对现有方法存在的上述不足，提出了两种基于 LDCF 行人检测算法的优化方法。

第一种优化方法是进行几何约束。在车载视频的应用场景下，车载摄像头安装的位置、焦距参数、图像分辨率等参数容易得到，而需要检测的目标行人在外形方面也有一些几何约束条件。可以根据这些先验知识将长宽比不符合行人特点的以及目标位置不在监控合理区域的检测框删除。

本创新工作是通过以下技术方案实现的，本创新工作包括以下步骤：

第一步、估算地平线高度。

第二步、计算摄像头参数，并根据训练集数据获得行人的高度与距离摄像头的距离之间的关系。

第三步、根据训练集数据获得行人的高度与宽度的关系。

第四步、将前三步推得的三个几何约束条件运用到检测中，提高检测效率。

本创新工作的原理是，LDCF 行人检测算法是一个误检率较高、漏检率较低的行人检测算法。检测框常见的几何错误有三种：检测框位于地平线上，因而这些检测框大多是错误的（如图 2(a)所示）；检测框的宽度过窄或过宽，不满足行人的动作特点（如图 2(b)所示）；检测框的高度与行人距离摄像头的距离不符合近大远小的规律，如距离摄像头较远的行人高度过高（如图 2(c)所示），或者距离摄像头较近的行人高度过矮（如图 2(d)所示）。通过几何

约束，可以去除掉这一部分不合理的检测盒，从而降低该算法的误检率。

与现有技术相比，本创新工作有效地利用了行人的几何特征，提高了算法的准确度。

第二种优化方法是进行关于 LDCF 上应用的贪心非极大抑制算法的改进。由于 LDCF 使用的策略是如果一个窗口被抑制则它不会再次抑制，如果一个正确窗口被抑制，则相应正确的抑制将会丢失，从而存在有效信息未充分利用的缺点。我们小组主要从这点出发研究了对贪心非极大抑制算法的改进。

本创新工作是通过以下技术方案实现的，本创新工作包括以下步骤：

第一步、估算检测分数比和检测对象尺度比的阈值 **Error! Reference source not found.**,

δ 。

第二步、总结出动态阈值的公式，分别实验得到 **Error! Reference source not found.**, δ 最佳值。得到最后的动态阈值公式。

第三步、修改原 NMS 算法，引入保留外围窗口的非极大抑制策略。

第四步、修改原 NMS 算法，引入激活机制的非极大抑制策略。

第五步，分别将前面两种改进(第三种需要和第二种配合使用)应用到实际检测中，观察效果。并将三种改进的组合改进算法应用到实际检测中，观察检测效果，取效果最好的组合方案。

本创新工作的原理是，(1)Ldcf 原始的非极大抑制过程中容易抑制尺度相近的检测窗口，而难以抑制尺度相差较大的窗口，一些主要来自行人腿、手臂等与行人类似地柱状区域会被误检处理而而无法被抑制(如图 4(a)),引入尺度比信息使原来的固定阈值变为动态阈值可以有效解决这个问题。(2)LDCF 贪心保留检测分数最高的检测窗口，实际情况下由于行人的遮挡、姿态等影响，使得检测分数稍低的检测窗口可能与真实行人更匹配。因此，我们在抑制窗口完全包含于被抑制窗口且二者检测分数相差不大的情况下保留外围窗口。(3)原算法一个窗口被抑制就会被判定为误检窗口，就不会被再次被抑制，我们引进的保留外围窗口策略具有一定的激活功能，因此，我们进一步的取消了被抑制窗口是否已经被抑制的检查，使得所有分数比当前抑制窗口低的初始检测窗口，都能进行保留外围窗口的非极大抑制。

与现有技术相比，本创新工作有效利用了检测窗口的各种信息，提高了算法的精确性。

附图说明

图 1 是本创新工作方法的流程图。

图 2 是实施例一中几种常见的几何错误举例。

图 3 是实施例一中的成像模型图。

图 4 是实施例二的使用前后效果对比图。

具体实施方式

下面对本创新工作的实施例作详细说明，本实施例在以本创新工作技术方案为前提下进行实施，给出了详细的实施方式和具体的操作过程，但本创新工作的保护范围不限于下述的实施例。

实施例一（几何约束）

本实施例包括以下步骤：

第一步、选择一系列地平线上无遮挡的图片，估算出地平线的高度。不同数据集的地平线高度不同，如 Caltech-USA 数据集的地平线高度约为 190，ETH 数据集的地平线高度约为 235。

第二步、计算摄像头参数，并根据训练集数据获得行人的高度与距离摄像头的距离之间的关系。

针孔摄像机成像模型如图 3(a)所示。其中 f 为摄像机焦距， w 为屏幕高度， θ 为垂直视角。根据该模型，可以得到焦距、垂直视角与屏幕高度的关系：

$$\left(\frac{w}{2}\right)/f = \tan\left(\frac{\theta}{2}\right)$$

其中，摄像机焦距为 1000 像素，屏幕高度为 480 像素。根据这些已知条件可以计算出垂直视角约为 27° 。

车载摄像机针孔成像模型如图 3(b)所示。其中 H 为行人实际高度， d 为行人到车载摄像机的实际距离， f 为摄像机的焦距， h 为行人在图像中的像素高度。可以发现，对于同一个行人而言，离摄像机越远，其成像高度 h 越小。而行人距离摄像头越远，代表着行人距离图像底部的距离也越远，所以行人距图像底部的距离和行人像素高度之间存在一定的几何关系。

理想的非畸变的行人的成像关系如图 3(c)所示。其中 f 为焦距， h 为图像行人的像素高度， D 为行人底部到图像底部的距离， H 为行人的真实高度， d 为行人到车载摄像机的实际距离。根据该图，可以得到：

$$\frac{h}{f} = \frac{H}{d}$$

$$\frac{D}{f} = \frac{D'}{d}$$

由以上两个公式可知：

$$\frac{h}{D} = \frac{H}{D'}$$

由图及以上公式可得：

$$\frac{f}{h} = \frac{d}{H} = \frac{S + \frac{D}{\tan \alpha}}{H} = \frac{S}{H} + \frac{D'}{H \tan \alpha} = \frac{S}{H} + \frac{D}{h \tan \alpha}$$

所以，

$$D = f \tan \alpha - \frac{S \tan \alpha}{H} h$$

即，当人的高度一定时，行人到图像底部的距离和行人的像素高度存在正比关系。

通过对 Caltech-USA 训练数据集的数据进行简单的函数拟合，得到 D 和 h 的大致关系：

$$D = -0.6927h + 292.2$$

在实际应用中，由于人的高度 H 会变化，以及检测框在检测时也存在一定的误差，将通过不等式来筛选数据：

$$-0.6936h + 223.68 < D < -0.6916h + 352.95$$

通过对 ETH 训练数据集的数据进行函数拟合，得到 D 和 h 的大致关系：

$$D = -0.5675h + 238.3$$

在实际应用中，通过不等式来筛选数据：

$$-0.8079h + 212.31 < D < -0.5261h + 260.59$$

第三步、根据训练集数据获得行人的高度与宽度的关系。

行人的宽高比会随着行人的动作变化而改变。因此可以通过不等式来筛选数据：

$$r_{\min} < \frac{w}{h} < r_{\max}$$

其中， w 为检测框的宽度， h 为检测框的高度， $r_{\min}$ 、 $r_{\max}$ 分别为宽高比的最小值和最大值。

通过对训练数据集的数据进行分析，可以发现行人宽高比的平均值约为 0.41，行人宽高比的最小值为 0.165，最大值为 1，即不等式为：

$$0.165 < \frac{w}{h} < 1$$

第四步、将前三步推得的三个几何约束条件运用到检测中，提高检测效率。

三个几何约束条件是指：

$$\begin{cases} D < D_{\max} \\ -k_1 \bullet h + d_1 < D < -k_2 \bullet h + d_2 \\ r_{\min} < \frac{w}{h} < r_{\max} \end{cases}$$

对 Caltech-USA 训练数据集来说，几何约束条件为：

$$\begin{cases} D < 290 \\ -0.6936h + 223.68 < D < -0.6916h + 352.95 \\ 0.165 < \frac{w}{h} < 1 \end{cases}$$

对 ETH 训练数据集来说，几何约束条件为：

$$\begin{cases} D < 245 \\ -0.8079h + 212.31 < D < -0.5261h + 260.59 \\ 0.165 < \frac{w}{h} < 1 \end{cases}$$

其中， w 为检测框的宽度， h 为检测框的高度， D 为检测框距图像底端的距离。

实施例二(非极大抑制)

设抑制窗口 BB_u 检测分数为 s_u ，高度为 h_u ，被抑制的窗口 BB_d 的检测分数为 s_d ，高度为 h_d ，则动态重合面积阈值 threshold。

本实施例包括以下步骤：

第一步、我们目的是当被抑制窗口尺度小于抑制窗口时，应该进行更严格地抑制，降低重合面积阈值可以抑制这类误检窗口。重合面积的阈值应与尺度比成正相关关系，因此估算检测分数比 s_d/s_u 和检测对象尺度比 h_d/h_u 的阈值 **Error! Reference source not found.**， δ 。

第二步、总结出动态阈值的公式如下：

Threshold=**Error! Reference source not found.**

其中 0.65 为原算法设定的固定的重合面积阈值。**Error! Reference source not found.** =

0.75 为高度比 h_d/h_u 的转折点, **Error! Reference source not found.**=0.7 为检测分数比 s_d/s_u 的转折点。当 $s_d/s_u < \text{Error! Reference source not found.}$ 且 $h_d/h_u < \text{Error! Reference source not found.}$ 时, 我们认为 BB_d 可能是 BB_u 的腿部、手臂等区域, 应采用动态的更小的阈值抑制。当 $h_d/h_u < \text{Error! Reference source not found.}$ 但 $s_d/s_u > \text{Error! Reference source not found.}$ 时我们认为 BB_d 可能是 BB_u 附近小尺度的行人, 不采用动态阈值抑制。

第三步、修改原 NMS 算法, 引入保留外围窗口的非极大抑制策略。当前抑制窗口 BB_u 的区域完全被被抑制窗口 BB_d 所包围, 且二者的检测分数差异极小: $s_d/s_u > \beta$ (这里 β 取 0.86), 则将 BB_u 作为误检窗口剔除, 而将 BB_d 保留下来, 且将 BB_u 的检测分数赋值给 BB_d , 不过 BB_u 此时仍然一直其他未被抑制的低分窗口, 知道本轮循环结束。

第四步、修改原 NMS 算法, 引入激活机制的非极大抑制策略。第三步引入的保留外围窗口的非极大抑制策略一定程度上具有激活功能, 理论上能够减少漏检和误检窗口, 因此, 我们在这里进一步取消被抑制窗口是否已经被抑制的检查过程, 使得所有分数比当前一直窗口低的初始检测窗口都曝秀保留外围窗口的非极大抑制, 激活一些已经被抑制过但是无法被当前抑制窗口抑制的检测窗口。

第五步、分别将前面两种改进(第三种需要和第二种配合使用)应用到实际检测中, 观察效果。并将三种改进的组合改进算法应用到实际检测中, 观察检测效果, 取效果最好的组合方案。

实施效果

依据上述步骤, 我们对 Caltech-USA 测试数据集进行了测试, 该测试集包含 66 个视频, 不进行几何约束时, 得到了 161731 个检测框。仅对地平线的高度进行约束后, 保留下 121362 个检测框, 比例为 75.04%; 仅对行人的宽高比进行约束后, 保留下 159952 个检测框, 比例为 98.90%; 仅对行人的高度和距摄像头的距离进行约束后, 保留下 114406 个检测框, 比例为 70.74%; 同时使用三个条件进行约束, 保留下 107141 个检测框, 比例为 66.25%, 即约束掉了 33.75% 的检测框。Miss rate (误检率) 从 31.0848% 下降到 22.6487%。

依据上述实验步骤, 我们将对原始非极大抑制的集中改进应用到 toolbox 中的非极大抑制函数中, 在 Caltech-USA 部分数据集进行了测试, 仅引入尺度比信息的非极大抑制算法比原始算法的 miss rate 能够降低 1.1918%, 引入保留外围窗口和激活机制的非极大抑制算法能够比原来 MR 降低 1.2913%, 组合引入尺度比信息、保留外围窗口和激活机制的非极大抑制算法能比原始的 MR 降低 1.7768%。

这一实验表明, 本实施例的行人检测优化算法能有效地降低误检率, 提高检测准确度。

说明书附图

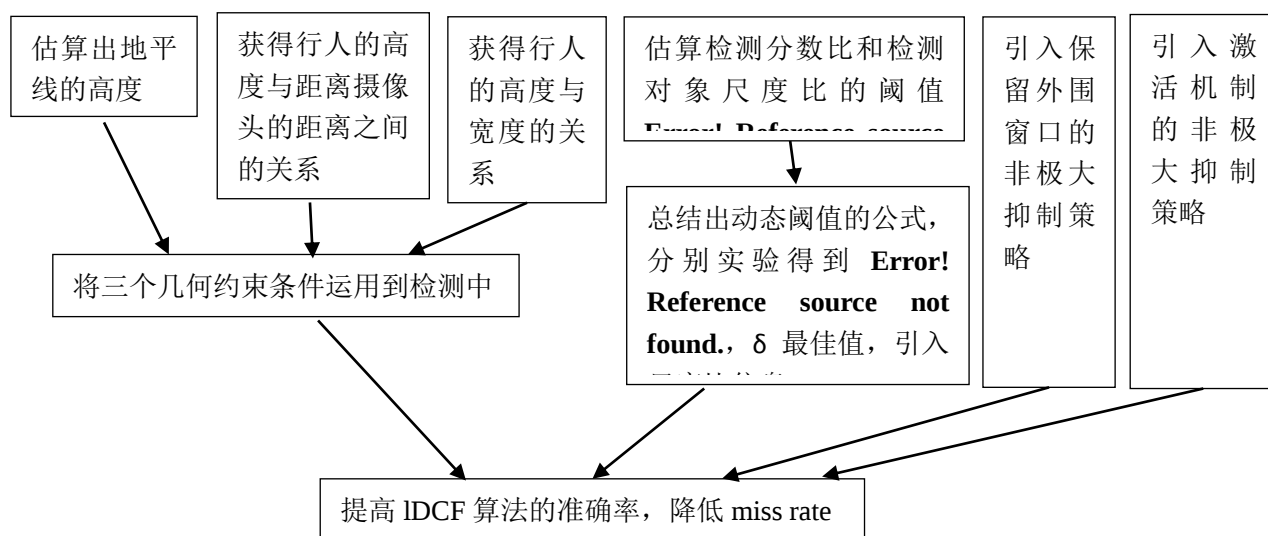


图 1



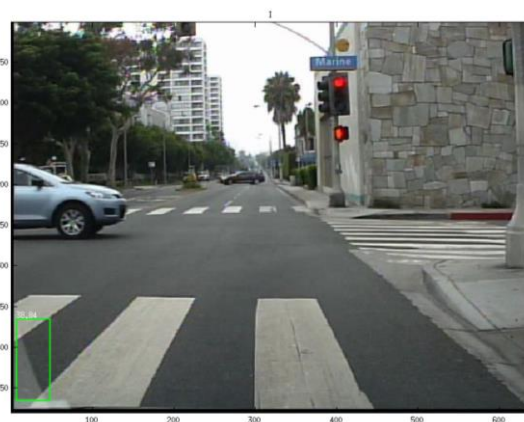
(a)检测框位于地平线上



(b)检测框的宽度过窄

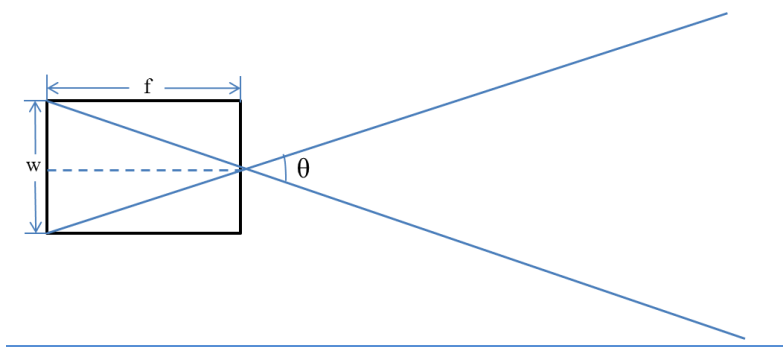


(c)距离摄像头较远的行人高度过高

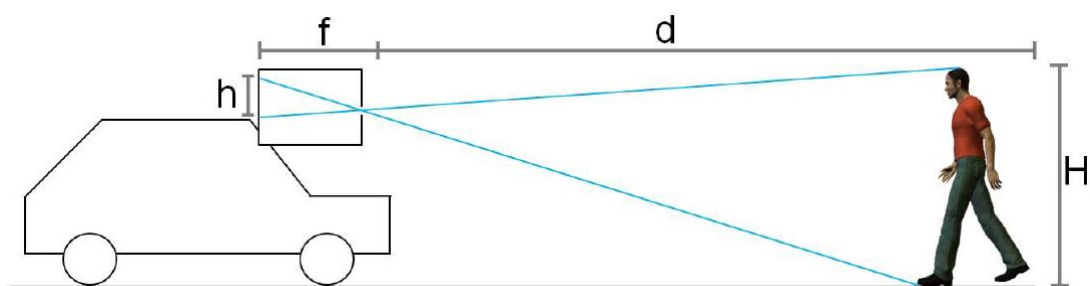


(d)距离摄像头较近的行人高度过矮

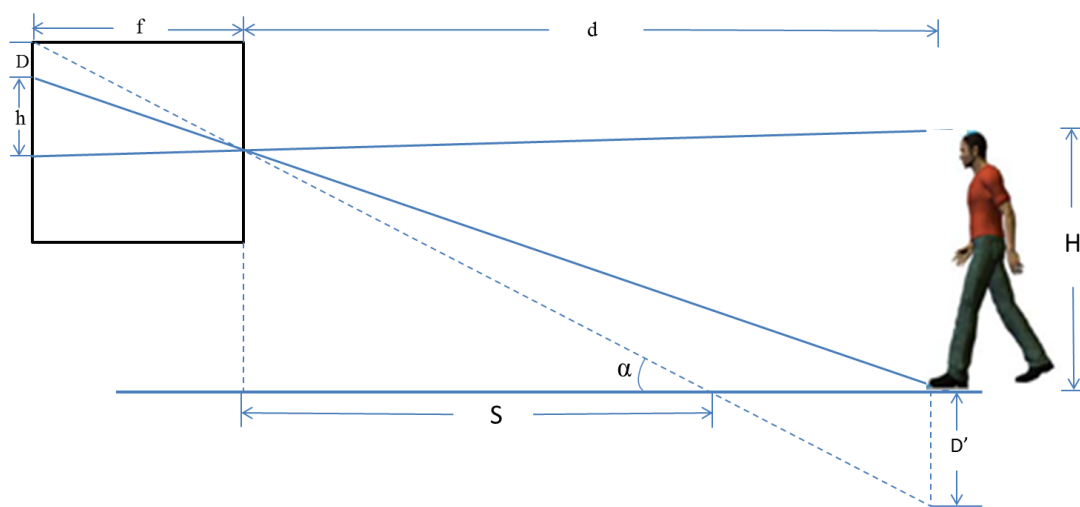
图 2



(a)焦距、视角与屏幕高度的关系

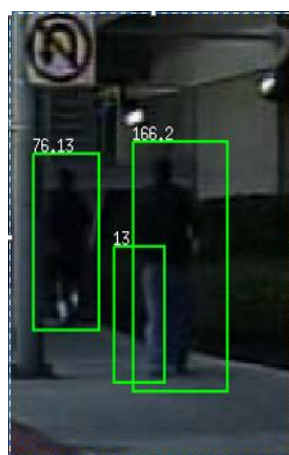


(b)摄像机成像模型



(c)行人几何成像关系图

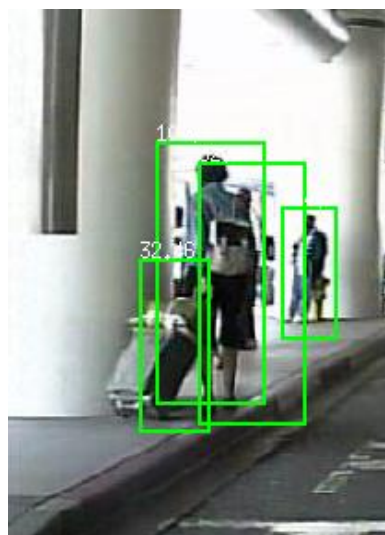
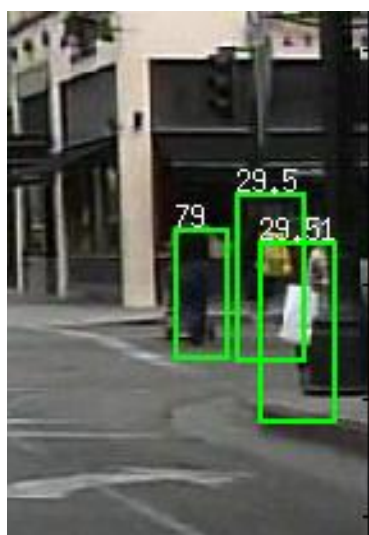
图 3



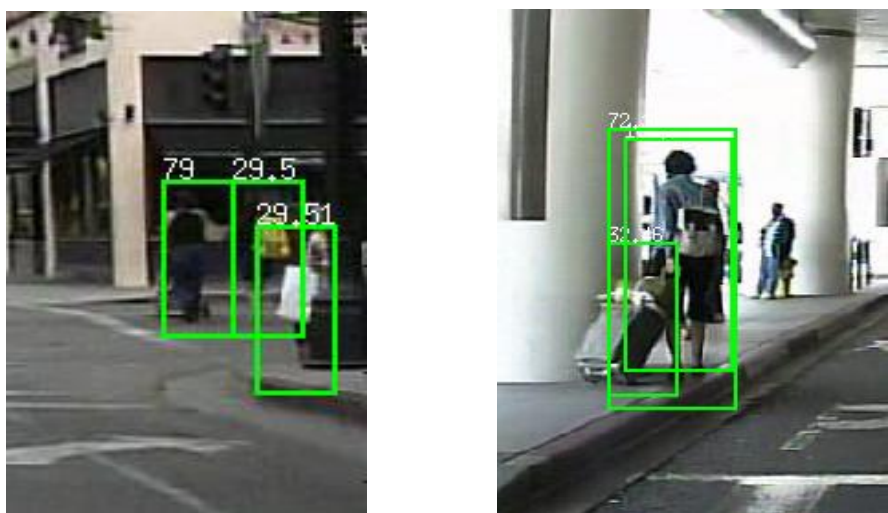
(a)LDCF 原始 NMS 检测效果



(b)引入尺度比信息后的 NMS 检测效果



(c)原始 NMS 检测效果



(d)引入保留外围窗口和激活机制的 NMS 检测效果

图 4

说明书摘要

一种基于 LDCF 行人检测算法的优化和改进方法，首先估算地平线高度、获得行人的高度与距离摄像头的距离之间的关系和行人的高度与宽度的关系。其次，将这三个几何约束条件运用到检测中，提高检测效率。随后，估算检测分数比和检测对象尺度比的阈值 **Error! Reference source not found.**， δ 。接着，总结出动态阈值的公式，分别实验得到 **Error! Reference source not found.**， δ 最佳值，引入尺度比信息。最后，修改原 NMS 算法，引入保留外围窗口的非极大抑制策略和激活机制的非极大抑制策略，并将改进算法应用到实际检测中，观察检测效果。

摘要附图

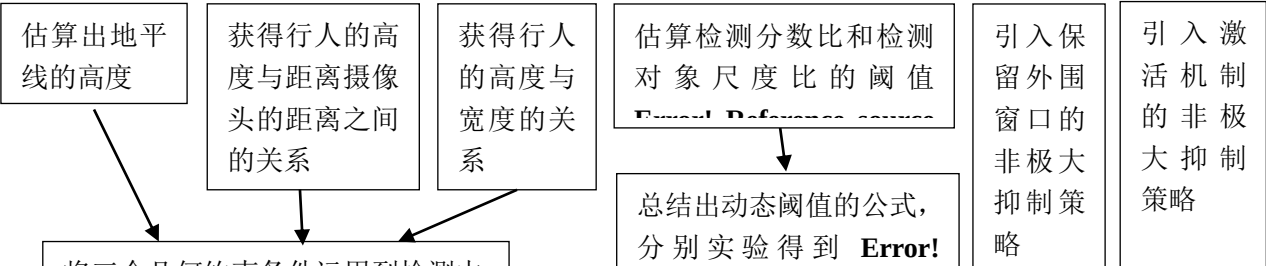


图 5

感谢

首先我要感谢这门课的两任老师以及所有在课堂上传授我们经验的学长学姐，感谢他们对我的谆谆教诲，让我对无线通信与移动网络的相关领域有了更深的理解和认识。其次我要对老师说声对不起，这学期的研究内容没有咨询过两位老师，都是根据自己的感觉和想法自己在做，以至于在最后的课程展示中被老师判定为跑题，对此我有深深愧疚和歉意，同时我也很感动也很感谢班里面的全体同学，他们能够对我的失误进行宽容和理解，在打分方面也没有丝毫的“抱怨”，因此我的最终的展示成绩也不是很差，这个反而增加了我的负罪感，真的是很感谢我的同学。最后，我想说的是这学期我真的做过相关的研究，也对行人检测方面有了一定的了解和认识，也会在今后继续深入的研究。