

基于 opencv 的前沿项目论文进展

2018.11.06 方建勇

提示：采用手机 safari 微软翻译技术

1. 公路交通视频车辆分割背景建模方法的比较

作者:[l. a. marcomini](#), [a. l. cunha](#)

文摘: 本文的目的是比较三种背景建模算法在公路交通视频车辆分割和检测中的性能。所有算法都在 **opencv** 中可用, 并且都在 python 中进行了编码。我们分析了七个视频, 总共录制了 2 个小时。为了比较这些算法, 我们创建了 35 张地面真相图像, 每个视频中创建了 5 个, 我们使用了三个不同的指标: 准确率、精确率和处理时间。通过使用准确性和精确性, 我们的目标是确定算法在检测和分割方面的性能, 同时利用处理时间来评估对计算系统的影响。结果表明, 这三种算法的准确率均在 90% 以上, 而平均准确率为 80%。该算法对处理时间的影响最小, 允许每秒计算 60 帧。少

2018 年 10 月 5 日提交;最初宣布 2018 年 10 月。

2. 用于虹膜识别的开源表示攻击检测基线

作者:[joseph mcgrath](#), [kevin w. bowyer](#), [adam ccajka](#)

摘要: 本文提出了第一个开源表示攻击检测 (pad) 解决方案, 以区分真实的虹膜图像 (可能佩戴清晰的隐形眼镜) 和带有纹理隐形眼镜的虹膜。该软件可以作为各种 pad 评估的基线, 也可以作为一个开源平台, 为虹膜 pad 提供最新的参考方法。该软件是用 c++ 编写的, 仅使用开源资源, 如 **opencv**。该方法不需要虹膜图像分割, 并使用二进制统计图像特征 (bsif) 提取与 pad 相关的特征, 这些特征由支持向量机分类器组合进行分类。附加到当前软件的支持向量机模型已通过 ndcldd-15 数据库进行了培训, 正确识别率超过 98%。但是, 该软件实现了对分类器进行再培训的功能, 具有任何真实和攻击图像的数据库。少

2018 年 9 月 26 日提交;最初宣布 2018 年 9 月。

3. 基于缩小空间差估计的计算高效密集运动目标检测

作者: [goran popović](#) , [antea hadviger](#) , [ivan marković](#) , [ivan petrović](#)

摘要: 从立体相机进行的测量效率高, 运动目标检测和深度估计是许多计算机视觉应用 (包括机器人和自动驾驶) 的一个非常有用的工具。本文介绍了利用一种利用以前帧中的信息提高立体匹配复杂度和准确性的算法来估计视差, 从而可以对运动物体进行密集检测。这种方法背后的主要思想是, 通过使用上一帧的自我运动估计和视差图, 我们可以设置一个以前的基础, 使我们能够降低当前帧视差估计的复杂性, 然后还可以检测运动对象。对于

每个像素，我们运行一个卡尔曼滤波器，递归融合视差预测和减少空间半全局匹配（sgm）测量。利用流单指令多数据指令集和多线程处理，实现并优化了该算法。此外，为了尽可能可靠地估计过程和测量噪声，我们使用三维激光测距传感器获得的地面真相在 kititi 套件上进行了广泛的实验。在差距估计方面，与 **opencv** sgm 的实现相比，该方法在速度和准确性方面都改进了 kititi 数据集序列。少

2018 年 9 月 21 日提交;最初宣布 2018 年 9 月。

4. 面部分析的实时系统

作者: [janne tommola](#), [pedram ghazi](#), [biswo adhikari](#) , [heikki hhtunen](#)

摘要: 本文描述了实时面部分析系统的解剖结构。该系统识别年龄, 性别和面部表情从用户出现在镜头前。所有组件都是基于卷积神经网络的, 我们研究了卷积神经网络在常用训练和评价集上的准确性。这项工作的一个关键贡献是描述了帧抓取、人脸检测和三种类型识别的处理线程之间的相互作用。执行该系统的巨蟒代码使用普通的库——**kerasensorlox**、**opencv** 和 **dlib**——and 可供下载。少

2018 年 9 月 14 日提交;最初宣布 2018 年 9 月。

5. aaui 多模态注释工具箱: 图像和视频中的注释对象

作者: [chris h. bahnsen](#), [andrias mögelmoose](#), [thomas b. moeslund](#)

摘要: 此技术报告介绍了两个注释工具箱，它们可以创建像素和基于多边形的遮罩以及感兴趣对象周围的边界框。这两个工具箱都支持 rgb 和热模式中顺序图像的注释。为每个带注释的对象分配一个分类标记、一个唯一 id 和一个或多个可选的元数据标记。这些工具箱是用 **opencv** 和 qt 库用 c++ 编写的，并使用可视界面和广泛的键盘快捷键进行操作。预构建的二进制文件可用于 windows 和 macos，这些工具也可以从 linux 下的源代码构建。到目前为止，已经使用工具箱对成千上万的帧进行了注释。少

2018 年 9 月 10 日提交;最初宣布 2018 年 9 月。

6. 图像处理和识别, 实现高效的停车位检测和导航辅助

作者: [chetan sai tutika](#), [charan valapaneni](#), [karthikeyan b](#)

文摘: 本文旨在利用 **opencv** 的图像处理能力，开发一种鲁棒、灵活的空置停车位检测算法。它消除了对独立传感器检测汽车的需求，而是使用来自各种来源和服务器的实时图像一起考虑一组插槽。这大大减少了设计高效停车系统所需的费用，并增加了操作的灵活性。该方法包括使用具有识别算法的便携式处理系统，并可选择提取图像并将其导入到指定的服务器。可以在自定义网站

上查看结果, 并可选择保留特定的空插槽和 gps 导航到选定的插槽。少

2018 年 7 月 6 日提交;最初宣布 2018 年 7 月。

7. 用于 adas 的差异图像分割

作者:[viktor mukha](#), [inon sharony](#)

摘要: 我们提出了一个简单的解决方案, 用于使用现有的连接组件标签 (ccl) 算法 (通常应用于二进制图像) 分割灰度图像, 该算法的效率足以在约束 (嵌入式汽车) 中实现建筑。我们的解决方案定制了区域增长和合并方法, 主要针对更接近的对象具有更多相关性的立体视差图像。我们为一些基本案例提供标准 **opencv** 实现的结果, 并提供筑波立体对数据集的图像。少

2018 年 6 月 27 日提交;最初宣布 2018 年 6 月。

8. 基于神经网络的快速探索视频查询

作者:[daniel kang](#) , [peter bailis](#) , [matei zaharia](#)

摘要: 随着视频量的增长, 分析人员越来越多地转向深度学习来处理视觉数据。虽然这些深度网络提供了令人印象深刻的精度水平, 但它们的执行速度比实时 (3 fps) 慢 10 倍 (3 fps), 而 gpu 在规模上是不可行的。此外, 部署这些方法需要使用许多低级库

(如 `opencv`、`mxnet`) 编写复杂的命令式代码, 这通常是一个临时且耗时的过程, 忽略了跨操作员优化的机会。为了大规模地解决视频分析的计算和可用性挑战, 我们引入了 `blazeit` 系统, 该系统优化视频查询以获取对象的时空信息。`blazeit` 通过 `frameql` 接受查询, `frameql` 是一种用于探索性视频分析的声明性语言, 可实现特定于视频的查询优化。我们提出了新的查询优化技术, 特别适合以前工作不支持的视频分析。首先, 我们将控制变量调整为视频分析, 并为聚合查询提供专业化的进步。其次, 我们使用专门的 `nn` 对重要采样进行传统限制的视频搜索 (即擦洗查询)。第三, 我们展示了如何为基于内容的选择推断新的筛选器类。通过结合这些优化, `blazeit` 可以比最近有关视频处理的文献提供三个数量级以上的加速。少

2018 年 5 月 2 日提交;最初宣布 2018 年 5 月。

9. 由轻量级 `cnn` 支持的实时人类检测作为边缘服务

作者 :[seyed yahya nikouei](#), [yu chen](#), [sejunsong](#), [滕华](#), [baek-young choi](#), [timothy r. faughnan](#)

摘要: 边缘计算允许在网络边缘的分散节点上执行更多的计算任务。如今, 许多具有延迟敏感性、关键任务的应用程序可以利用这些边缘设备来减少延迟, 甚至通过其现场存在实现实时在线决策。智能监控中的人体检测、行为识别和预测属于这一类, 大量视频流数据的转换可能需要宝贵的时间, 并给通信网络带来沉重

的压力。人们普遍认识到, 视频处理和对象检测计算密集, 成本太高, 无法由资源有限的边缘设备处理。本文以深度可分离卷积和单射击多盒检测器 (ssd) 为研究对象, 介绍了一种轻量级卷积神经网络 (lcnn)。通过缩小分类器的搜索空间以聚焦监控视频帧中的人体, 提出的 lcn 算法能够检测到具有经济实惠的计算工作量的行人到边缘设备。利用 **opencv** 库在边缘节点 (raspberrypi 3) 上实现了一个原型, 并利用真实世界的监控视频流实现了令人满意的性能。实验研究验证了 lcnn 的设计, 表明它是计算边缘密集型应用的一种很有前途的方法。少

2018 年 4 月 24 日提交;最初宣布 2018 年 5 月。

10. 使用 mmlspark 灵活、可扩展的深度学习

作者 : [mark hamilton](#), [sudarshan raghunathan](#), [akshaya annavajhala](#), [danil kirsanov](#), [eduardo de leon](#), [eli barzilay](#), [ilya matiach](#), [joedavison](#), [maureen busch](#), [miruna Opreescu](#), [ratan sur](#), [roope astala](#), [tong wen](#), [changyoung park](#)

摘要: 在这项工作中, 我们详细介绍了一个新的开源库, 称为 mmlspark, 它结合了灵活的深度学习库认知工具包, 与分布式计算框架 apache spark。为了实现这一目标, 我们向认知工具包提供了 java 语言绑定, 并在 spark 生态系统中添加了几个新组件。此外, 我们还将流行的图像处理库 **opencv** 与 spark 集成在一起, 并提供了一个工具, 用于从任何 sparkml 估计器自动生成 pyspark 包装, 并使用此工具将所有工作暴露给 pyspark 生态

系统。最后, 我们提供了一个大型的工具库, 用于在 spark 生态系统内工作和开发。我们将这项工作应用于从相机陷阱图像中对雪豹进行自动分类, 并为非营利保护组织雪豹信托提供了端到端解决方案。少

2018 年 4 月 11 日提交;最初宣布 2018 年 4 月。

11. 智能手机屏幕截图中的文本提取和检索: 构建媒体中的生活存储库

作者: [agnese chiatti](#), [mu jung cho](#), [anupriya gagneja](#), [xiang, miriam brinberg](#), [katie roehrick](#), [sagnik ray choudhury](#), [nilam ram](#), [byron reeves](#), c. 李·贾尔斯

摘要: 日常参与生活体验越来越多地与移动设备的使用交织在一起。在行为研究中使用秒级的屏幕捕捉, 并实施 "及时" 的卫生干预措施。数字信息的心理广度的不断增加将继续使人们看到的实际屏幕成为生活经历的首选甚至需要的数据来源。有效和高效的信息提取和数字屏幕截图检索是成功使用屏幕数据的关键先决条件。在本文中, 我们提出了实验工作流程, 我们利用: (i) 预处理一个独特的屏幕截图集合, (ii) 提取嵌入在图像中的非结构化文本, (iii) 基于结构化模式组织图像文本和元数据, (iv) 索引生成的文档集合, 以及 (v) 允许通过专用的垂直搜索引擎应用程序进行图像检索。所采用的程序集成了传统图像处理、光学字符识别 (ocr) 和图像检索的不同开源库。我们的目标是评估是否以及如何将最

先进的方法应用于这一新颖的数据集。我们展示了**如何将基于 opencv 的预处理模块与基于长期短期内存 (lstm) 的 Tesseract ocr 版本相结合**, 而无需经过临时培训, 从而获得了 74% 的字符级精度。此外, 我们还将处理后的存储库作为专用图像检索系统的基线, 用于行为和预防科学家的即时使用和应用。我们讨论了文本信息提取和检索的问题, 这些问题是截图图像案例所特有的, 并提出了今后的重要工作。少

2018 年 1 月 4 日提交;最初宣布 2018 年 1 月。

12. 利用图像处理生成人类行为的分析见解

作者: [namit juneja](#), [rajesh kumar muthu](#)

摘要: 本文提出了一种在物理空间中跟踪人类数字的方法, 然后利用这些数据生成几个数据点, 如足迹分布、人口分析、热图以及性别分布。拟议的框架旨在建立这一点, 同时利用最低限度的计算资源, 同时保持实时。它往往是有用的, 如什么样的人访问某个地方或什么小时一天经历最大的活动, 这样的分析可以用来提高销售, 管理大量的人, 以及预测未来的行为。建议的框架的设计方式使其能够从 ip 摄像机中获取输入流, 并使用这些输入流使用 **opencv** 和 **raspberry** 等开源工具生成相关数据点。少

2017 年 11 月 21 日提交;最初宣布 2017 年 11 月。

13. dock 过来面: 一个易于安装和使用更快的 r-inn 人脸检测器在 docker 容器

作者:[nataniel ruiz](#), [james m. rehg](#)

摘要: 人脸检测是面部地标检测、姿态估计、情感分析和人脸识别等应用中一项非常重要的任务,也是必要的预处理步骤。人脸检测不仅是计算机视觉应用中的一个重要预处理步骤,而且在计算心理学、行为成像和其他领域也是一个重要的预处理步骤,在这些领域,研究人员可能不会在计算机视觉框架和最先进的领域中被启动检测应用。现有研究的很大一部分包括人脸检测作为预处理步骤,使用现有的开箱即用的探测器,如基于 hog 的 dlib 和 opencv haar 人脸探测器,这些探测器已不再是最先进的——它们主要用于因为它们易于使用和方便。我们介绍 dockerface, 一个非常精确的更快的 r-cnn 人脸探测器在一个 docker 容器,不需要培训,易于安装和使用。少

2018 年 4 月 5 日提交;v1 于 2017 年 8 月 14 日提交;**最初宣布** 2017 年 8 月。

14. 车辆速度检测应用程序

作者:[伊托罗·伊孔](#)

摘要: 该报告介绍了使用智能手机摄像头测量车辆速度的情况。利用 `opencv` api 的 `lbp` 级联分类器检测车辆在相机机架上的位置, 完成车辆的速度测量, 利用检测到的车辆随时间的位移来计算速度。转换系数确定为将像素位移映射到实际车辆距离。测量的速度与地面真实速度成正比。少

2017 年 3 月 24 日提交;v1 于 2017 年 2 月 17 日提交;**最初宣布** 2017 年 2 月。

15. 图像处理多维约简方法中的高性能软件及其在古代手稿中的应用

作者:[corneliu t. c. arsene](#), [stephen church](#), [mark dickinson](#)

摘要: 多光谱成像是提高文字或印刷文本的可读性的一项重要技术, 这些文字的字母已经褪色, 或者是由于故意擦除, 或者仅仅是由于时间的破坏。通常, 文本可以简单地通过查看单个波长来阅读, 但在某些情况下, 图像需要进一步增强, 以最大限度地增加阅读文本的机会。有许多可能的增强技术, 本文对图像处理的一组扩展维数约简方法进行了评估和比较。我们在两个不同的手稿中评估 15 个降维方法。这一评估既是主观上进行的, 也是询问手稿中使用的语言专家的意见, 以及他们喜欢的技术, 也通过使用 `davis-bouldin` 和 `dunn` 指数来评估生成的图像群集。我们发现, 规范变量分析 (cva) 方法, 这是使用 `matlab` 实现, 我们以前使用来增强多光谱图像, 它确实优于所有其他测试的方法。然

而，它是非常有可能的其他方法将更适合在特定情况下，所以我们仍然建议尝试一系列这些技术。特别是，cva 是一种受监督的聚类技术，因此与非监督技术（如更常用的主成分分析方法（pca）相比，它需要更多的用户时间和精力。如果来自 pca 的结果足以允许读取文本，那么 pva 所需的额外工作量可能是没有道理的。为了比较计算时间和图像结果，还使用 gnu (gnu not unix) 科学库 (gsl) 和 **opencv** (open 源计算机视觉) 计算机实现了 c 编程语言中的 cva 方法。视觉编程库。少

2018年7月18日提交;v1 于 2016 年 12 月 19 日提交;**最初宣布** 2016 年 12 月。

16. 基于图像处理的对象检测

作者: [j 捷德·法雷斯](#), [ilia voronkov](#)

文摘: 无人阿里尔飞行器 (uav) 在军队中对边境安全的重要性更大。本文的主要目的是使用 haar 级联算法开发用于对象和人脸检测的 **opencv-python** 代码。目前，无人机用于探测和攻击浸润的地面目标。这种类型的无人机的主要缺点是，有时没有正确检测到物体，从而导致物体撞上无人机。该项目旨在避免无人机的这种不必要的碰撞和损害。无人机还用于使用 **voila-jones** 算法探测和跟踪人类的监视。该算法采用级联目标检测功能和视觉训练函数来训练算法。此代码的主要优点是减少了处理时间。利

用现有的视频和图像数据库对 python 代码进行了测试, 验证了输出。少

2016 年 11 月 23 日提交;最初宣布 2016 年 11 月。

17. 一种可扩展、稳健的智能实时视频监控框架

作者:shreenath dutt, ankita kalra

文摘: 本文提出了一种利用阿帕奇风暴和 opencv 实现智能、可靠、高效存储的视频监控系统。作为 storm 拓扑, 我们添加了多个信息提取模块, 这些模块只将重要内容写入磁盘。我们的拓扑是可扩展的, 能够根据用例添加新的算法, 而不会影响现有的算法, 因为所有的处理都是相互独立的。此框架还具有高度的可扩展性和容错能力, 这使得它成为需要监控大型监控摄像机网络的组织的最佳选择。少

2016 年 10 月 29 日提交;最初宣布 2016 年 10 月。

18. 用早期 atlas 公共数据测量带电粒子的倍数

作者:gözde 乌斯图恩, erol barut, erhan bektas , v. erkcan zcan

摘要: 我们研究了 atlas 实验记录并为外联目的公开的 100 张早期 lhc 碰撞图像, 并提取带电粒子多样性作为在 $\sqrt{s} = 7 \text{ tev}$ 处质子-质子碰撞的动量函数。由于我们研究的碰撞已经由 atlas

协作机构进行了预处理, 这些轨道是可见的, 但只有低分辨率位图的形式向公众开放。我们采用了两种独立的图像处理方法, 一种基于行业标准的 **opencv** 库和 `c++`, 另一种基于 `python` 中自行开发的算法。我们提出了通过这两种方法获得的粒子的横向动量和方位角分布, 与文献一致。少

2016 年 10 月 8 日提交;最初宣布 2016 年 10 月。

19. 加密 img: 对加密图像的隐私保护处理

作者:[m. tarek ibn ziad](#), [amr alanwar](#) , [moustafa alzantot](#), [mani srivastava](#)

摘要: 云计算服务为图像和多媒体文件的存储和处理提供了可扩展的解决方案。但是, 对隐私风险的担忧会阻止用户与第三方服务共享其个人图像。本文介绍了基于加密图像的模块化隐私保护文信息处理库 `cryptoimg` 的设计与实现。通过使用同态加密, `cryptoimg` 允许用户将其图像处理操作委派给远程服务器, 而无需考虑任何隐私问题。目前, `cryptimg` 支持最常用的图像处理操作的子集, 如图像调整、空间滤波、边缘锐化、直方图均衡等。我们实施了我们的图书馆, 作为一个扩展, 以流行的计算机视觉图书馆 **opencv**.加密 `img` 可以从移动或桌面客户端使用。我们的实验结果表明, `cryptimg` 在加密图像上执行操作时效率很高, 在支持的平台上的误差可以忽略不计, 时间开销也很合理

2016 年 9 月 3 日提交;最初宣布 2016 年 9 月。

20. 空中文字书写

作者:saira beg, m. fahad khan, faisal baig

摘要: 本文提出了一种基于视频的实时指向方法, 允许在移动摄像头前的空中绘制和书写英语文本。提出的方法主要有两个任务: 首先跟踪视频帧中的彩色指尖, 然后将英语 ocr 应用于绘制的图像上, 以识别书面字符。此外, 所提出的方法提供了一种自然的人机交互, 不需要键盘、手写笔、笔或手套等进行字符输入。对于实验, 我们开发了一个使用 java 语言的 **opencv** 的应用程序。我们在三星 galaxy3 android 手机上测试了该方法。结果表明, 该算法在不同形状字母的测试中, 平均精度为 92.083。在这里, 使用了 3000 多个不同的磁性 3d 形状的字符 [ref:http://learnrnd.com/news.php?id=Magnetic_3D_Bio_Printing]。我们提出的系统是基于软件的方法和相关的非常简单, 快速和容易。它不需要传感器或任何硬件, 而不需要相机和繁文缛节。此外, 拟议的方法可以适用于所有断开连接的语言, 但有一个问题, 即它对颜色敏感, 因为在开始字符书写之前, 背景中存在任何红色都可能导致错误的结果。少

2016 年 4 月 27 日提交;最初宣布 2016 年 4 月。

21. 模块化跟踪框架: 基于注册的跟踪的统一方法

作者: [abhineet singh](#), [martin jagersand](#)

摘要: 本文提出了一种模块化、可扩展且高效的基于注册的跟踪开源框架, 称为模块化跟踪框架 (mtf)。它以机器人应用为目标, 完全以 c++ 实现, 并从一开始就设计, 以便轻松地与支持多个主要视觉和机器人库 (包括 **opencv**、ros、visp 和 eigen) 的系统集成。它实现了比其他现有系统更多的方法, 更快、更精确。此外, 其设计的理论基础是一种将基于注册的跟踪器概念化的新方法, 将其分解为三个组成子模块——搜索方法 (sm)、外观模型 (am) 和状态空间模型 (ssm)。在这一过程中, 我们整合了在 2004 年贝克 \ & amp; matthews 的地标性作品之后发表的许多重要进展。除了是快速和高精度跟踪的实用解决方案外, mtf 还可以作为一个有用的研究工具, 允许对任何子模块进行更好的研究。当对其中一种方法引入新的方法时, 这种分解可以帮助实验找到其他最佳方法的组合。通过广泛使用泛型编程, mtf 可以很容易地插入任何子模块的新方法, 以便它不仅可以使用现有方法进行全面测试, 而且还可以立即在使用框架。mtf 已经实施了 16 个 ams、11 个 sms 和 13 个 ssm, 提供了 2000 多个不同的单层跟踪器。它还允许将其中的两个或两个以上的跟踪器以几种方式组合在一起, 以创建几乎无限的各种新型多层跟踪器。少

2018 年 5 月 17 日提交;v1 于 2016 年 2 月 29 日提交;**最初宣布** 2016 年 2 月。

22. 用于多模式深度学习的 c++ 库

作者: [金健](#)

摘要: 一个支持多个模型的深度学习框架, 本文解释了它的理念和功能。mdl 在 linux、mac 和 unix 平台上运行。这取决于 opencv。

2016 年 4 月 12 日提交;v1 于 2015 年 12 月 21 日提交;**最初宣布** 2015 年 12 月。

23. 快速引导式过滤器

作者: [何凯明](#), [孙健](#)

摘要: 引导滤波器是一种边缘感知图像滤波技术。由于其良好的视觉质量, 速度快, 易于实现, 引导滤波器见证了在实际产品中的各种应用, 如手机中的图像编辑应用程序和立体声重建, 并已被纳入官方 matlab 和 **公开简历**。在本说明中, 我们提醒, 引导滤波器可以简单地从 $O(n)$ 时间加快到 $O(N/s^2)$ 时间的子采样比 s 。在各种应用中, 这导致了 $>10\times$ 的加速, 几乎没有明显的退化。我们希望这种加速将提高当前应用的性能, 并进一步推广这种滤波器。代码已释放。少

2015 年 5 月 5 日提交;**最初宣布** 2015 年 5 月。

24. 坚固和实时检测具有驾驶辅助车辆和自主车辆所需坡度的曲线 (曲线)

作者:[amartansh dubey](#), [k. m. bhurchandi](#)

摘要: 道路事故的最大原因之一是曲线车道和盲目转弯。即使是新自主车辆最大的障碍之一，也是检测曲线车道、多条车道和有很多不连续性和噪音的车道。本文提出了一种非常有效和先进的检测具有所需坡度的曲线（特别是实时检测曲线车道）和检测噪声、不连续性和干扰大的曲线（车道）的算法。总体目标是为这一任务制定稳健的方法，即使在不利的条件下也适用。即使在一些最著名和最有用的库，如 **opencv** 和 **matlab**，也没有任何功能可用于检测具有所需坡度、形状和不连续性的曲线。只有几个预定义的形状，如圆，椭圆等，可以检测到使用目前可用的函数。提出的算法不仅可以检测出不连续、噪声大、所需斜率的曲线，还可以对不同曲线进行阴影和光照校正以及检测/判别。少

2015 年 1 月 13 日提交;最初宣布 2015 年 1 月。

25. 飞行时间图像中校准网格的自动检测

作者:[迈尔斯·汉萨德](#)、[拉杜·霍罗](#)、[米歇尔·阿玛特](#)、[乔治·埃万格利迪斯](#)

摘要: 使用棋盘图案的图像, 通过既定的方法校准飞行时间相机是方便的。然而, 振幅图像的低分辨率使得电路板难以可靠地检测到。基于连接的图像组件的启发式检测方法在这些数据上的性能非常差。本文介绍了一种基于 hough 变换的另一种几何原则方法。棋盘的投影由两支铅笔表示, 这两支铅笔在图像的梯度数据中被标识为定向集群。在轴对齐坐标下, 将投影 hough 变换应用于两个簇中的每一个。每个变换的范围都有适当的边界, 因为相应的梯度向量近似平行。这两个变换中的每一个都包含一系列的共线峰; 在给定的铅笔中的每一行一个。通过扫描双线通过转换, 很容易检测到这种模式。将所提出的基于 Hough-based 的方法与标准 **opencv** 检测例程进行了比较, 并将其应用于数百张飞行时间图像。结果表明, 新方法可在更多的姿态上检测到更多的校准板, 而不会造成任何整体精度损失。这一结论是基于对几何误差和光度误差的分析。少

2014 年 1 月 24 日提交;最初宣布 2014 年 1 月。

26. 团队合作行动识别的 gpu 框架

作者: [mohamed elhoseiny](#), [hossam faheem](#), [taymour nazmy](#), [eman shaaban](#)

摘要: 团队行动识别的实时处理是一个挑战, 因为复杂的计算模型实现了较高的系统性能。因此, 本文提出了一个基于图形处理单元 (gpu) 的框架, 以实现基于角色的活动识别工作的显著提

高。该框架可应用于各个领域，特别是运动和军事应用。此外，该框架可以为许多操作识别应用进行定制。本文介绍了 gpu 是性能改进的主要工具的框架的各个阶段。通过在 gpu 上执行视频处理和机器学习算法，实现了加速。视频处理和机器学习算法涵盖了我们的框架中涉及的所有计算。视频处理任务涉及 gpu 的运动检测、分割和目标跟踪算法的实现。此外，我们的框架还与 opencv 集成，gpu 版本的 **opencv** 功能。我们的框架支持机器学习任务，支持向量机 (svm) 的 gpu 实现用于对象分类和特征离散化，隐藏 markov 模型 (hmm) 用于活动识别阶段，id3 算法用于角色团队成员的认可。该系统根据 uc-team-war-war-wat-war-wat-wat—在 nvidia 9500gt 显卡 (32 500mhz 处理器) 上实现了 20x 的加速。少

2013 年 10 月 11 日提交;最初宣布 2013 年 10 月。

27. 子模块间高阶能函数的结构化学习

作者:亚历山大修复,索尔斯滕·约阿奇姆,萨姆·公园, 拉明·扎比

摘要: 子模块化函数可以在多项式时间内精确最小化，而图形切割与最大流的特殊情况 \ 全市 {pami04} 对计算机视觉产生了重大影响 \ citevz {bvz: pami01. 1. kwatra:sigra:sigra:sother:gratcht04。本文讨论了一类重要的子模块 (sos) 函数 \ 西托拉: eccm12 kromorov:dami12}, 它可以

通过称为亚模块化流的最大流变量—[cecleedmond:adm77](#)} 有效地最小化。 sos 函数可以自然地表达涉及的高阶原点, 例如, 局部图像补丁;然而, 由于它们有如此多的参数, 很难充分利用它们的表现力。我们没有试图将现有的高阶先验表述为 sos 函数, 而是采用了一种判别学习的方法, 有效地搜索了 sos 函数的空间, 以获得在我们的训练集中表现良好的高阶先验。采用结构支持向量机方法 ([joachims/](#)), 并根据二次规划方法制定训练问题;因此, 我们可以通过扩展的切割平面算法有效地搜索 sos 素数的空间。我们还展示了如何修改最先进的视觉问题最大流量方法 \ [cite{Goldberg:ESA11}](#), 以有效地解决子模块化流问题。与 `opencv` 实现的 `opencv` 进行了实验比较, 该技术采用手动调谐参数代替机器学习。在标准数据集—[cice:cvpr10](#)} 上, 我们的方法学习具有数百个参数值的高阶原值, 并产生明显更好的分割。虽然我们的重点是二进制标签问题, 但我们表明, 我们的技术可以自然地推广到处理两个以上的标签。少

2013年9月30日提交;[v1](#) 于 2013年9月28日提交;**最初宣布** 2013年9月。

28. 从二维平面图中实现伪三维建筑生成的三相识别方法

作者:[raj kishen moloo](#), [muhammad ajmal sheik dawood](#), [abu salmaan auleear](#)

文摘: 如今, 三维建筑可视化已成为建筑行业建筑产品概念化、设计和展示的有力工具, 提供了现实的工程互动和演练产品。传统的 3d 模型实现方法包括使用专门的 3d 创作工具以及熟练的 3d 设计人员, 其中包含模型的蓝图, 这是一个缓慢而费力的过程。本文的目的是通过简单地分析蓝图文档并自动生成 3d 场景来实现这一过程的自动化。为此, 我们设计了一种从 2d 平面图生成伪 3d 建筑的三相识别方法, 并相应地开发了一个软件。我们的 3 相 3d 构建系统已实现使用 c, c++ 和 **opencv** 库 [24] 的图像处理模块;保存模块生成了一个 xml 文件, 用于存储已处理的平面图对象属性;而 Irrlicht [14] 游戏引擎被用来实现交互式 3d 模块。虽然我们提出的制度仍处于起步阶段, 但取得了值得称道的成果。我们在 6 个平面图上测试了我们的系统, 从低到高的复杂性, 结果似乎很有希望, 平均处理时间约为 3, 3s 生成在 4。此外, 该系统还提供了交互式步行方式, 并允许用户修改组件。少

2011 年 7 月 19 日提交;最初宣布 2011 年 7 月。

29. 增强现实在主流应用中的实现方法

作者: david prochazka, tomas kubek

文摘: 增强现实已成为从空间探索到军事应用的许多领域的有用工具。尽管使用过的理论原理已经众所周知了近 10 年, 但增强现实几乎完全用于具有特殊硬件的高预算解决方案。然而, 在过去

几年里, 我们可以看到许多专注于在不同移动设备上部署增强现实的项目越来越受欢迎。我们的文章针对的是那些考虑为主流市场开发增强现实应用的开发者。这样的开发商将被迫将应用价格, 因此也是开发价格, 保持在合理的水平。使用现有的图像处理软件库可以显著降低开发成本。本文的理论部分概述了增强现实应用结构。此外, 还介绍了选择适当图书馆的方法以及对这一领域重点关注的现有软件库的审查。本文的最后一部分概述了我们使用 **opencv** 库实现增强现实应用的关键部分的情况。少

2011 年 6 月 28 日提交;最初宣布 2011 年 6 月。