

DOI: 10.3724/SP.J.1218.2014.00076

基于车载单目图像的 3 维地平面估计

向文辉¹, 刘羽¹, 曹洋¹, 汪增福^{1,2}

(1. 中国科学技术大学自动化系, 安徽 合肥 230027; 2. 中国科学院合肥智能机械研究所, 安徽 合肥 230031)

摘要: 提出了一种由车载摄像头获取的单目图像估计场景的 3 维地平面以及深度信息的算法。该算法首先融合图像的散焦信息、饱和度信息以及暗通道先验, 得到场景的相对深度图。然后在基于地平线分段平滑的假设下, 进行双边中值滤波推断出 3 维地平面。最后在此基础上利用成像几何原理计算出绝对深度图。为了验证算法的有效性, 不仅在离线计算机上进行了大量的对比实验, 而且还将该算法应用于机器人小车的室外自主避障。实验结果表明本文算法可以较好地估计出 3 维地平面和场景深度, 机器人小车可以利用这些信息成功检测并躲避障碍物。

关键词: 机器人小车; 单目图像; 图像融合; 3 维地平面估计; 深度估计

中图分类号: TP242

文献标识码: A

文章编号: 1002-0446(2014)-01-0076-07

3D Ground Plane Estimation from a Monocular Vehicle-borne Image

XIANG Wenhui¹, LIU Yu¹, CAO Yang¹, WANG Zengfu^{1,2}

(1. Department of Automation, University of Science and Technology of China, Hefei 230027, China;

2. Institute of Intelligent Machine, Chinese Academy of Sciences, Hefei 230031, China)

Abstract: An algorithm is proposed to estimate the 3D ground plane region and scene depth information from a monocular image captured by a vehicle-borne camera. Firstly, information about image defocus, image saturation and dark channel prior are fused to estimate a relative depth map of the scene. Then, the 3D ground plane can be inferred by using a bilateral median filter based on the assumption that the horizon is piecewise smooth. Finally, absolute depth map can be obtained by using the principle of imaging geometry. To verify the effectiveness of the proposed algorithm, not only numerous comparative experiments are performed on an offline computer, but also it is applied to the outdoor autonomous obstacle avoidance of a robot vehicle. Experiment results demonstrate that both the 3D ground plane and scene depth information can be well estimated by the proposed algorithm, with which the robot vehicle can successfully detect and avoid the obstacles.

Keywords: robot vehicle; monocular image; image fusion; 3D ground plane estimation; depth estimation

1 引言 (Introduction)

移动机器人导航和 SLAM 技术在制造业、空间开发和服务机器人等领域都具有广阔的前景, 而移动机器人实现自主导航的前提是对自身的精确定位。传统方法中, 基于激光扫描仪或双目立体视觉的方法常被用于距离信息的测量, 但这些方法存在造价昂贵或标定困难等问题, 实用性受到很大的限制。随着计算机视觉技术的发展, 基于单目视觉估计场景深度信息的方法因其低成本及易于操作等优势受到了广泛的关注, 成为该领域近年来的研究热点之一。

由于场景在成像过程中会丢失大量信息, 因此基于单幅图像的场景深度恢复是十分困难的, 是一

个病态反问题, 有无穷多个解, 往往需要加入较强的先验知识才能恢复出场景的 3 维信息。现有的一些基于单幅图像估计深度的方法包括从图像的明暗信息恢复形状 (shape from shading)^[1]、从图像的散焦信息恢复场景深度 (depth from defocus)^[2] 等, 但这些方法有着单一光源、静态背景等比较严格的先验假设, 难以实用。Loh 等^[3] 利用图像的纹理梯度求得场景深度值, 但这一方法容易受到噪声的干扰。Delage 等^[4] 提出了基于动态贝叶斯模型的学习方法, 但该方法只适用于室内图像。Saxena 等^[5-7] 提出了一种基于单幅图像来获得深度的算法, 该算法对地面上的阴影及纹理变化具有一定的鲁棒性, 能检测到不连续区域, 可以用于机器人的自主

避障, 不过该算法效率较低, 无法在实际应用中达到实时性的要求。

快速准确的 3 维地平面估计是解决移动机器人自主导航问题的一个关键点, 它有助于场景中地标的定位。Hoiem 等^[8]采用一种基于图的纹理分割算法^[9]将图像场景分割成地面、垂直于地面的物体以及天空, 但是这种方法获得的结果更接近于图像分类, 无法用于机器人的精确定位。Cherian 等^[10]在 Make3D 结果^[11]的基础上提出了一种利用图像分割来进行地面估计的算法。这个方法能比较准确地估计出地面区域, 但不能获得精确的深度信息, 而且同样有计算代价过高的问题。

3 维地平面估计的另一个主要作用是可以获得场景的深度信息。本文提出了一种由车载单目摄像头获取的单幅图像估计 3 维地平面及深度信息的算法。该算法综合利用了图像中的散焦信息^[12]、暗通道先验信息^[13]和图像饱和度信息, 融合得到比较精确的相对深度图, 然后在相对深度图上搜索出地平面边界, 求得准确的地平面区域, 最后进一步利用摄像头的成像原理求得绝对深度图。

在离线计算机上的实验结果表明, 本文提出的算法有很高效率, 处理一幅分辨率为 320×240 的图像所用时间不超过 0.2s。此外, 还将该算法应用于轮式移动机器人在室外环境中的自主避障, 结果也验证了本文算法的快速性和准确性。

2 算法框架 (Algorithm framework)

整个算法的流程如图 1 所示, 该算法主要包括计算相对深度图和计算绝对深度图 2 个方面。

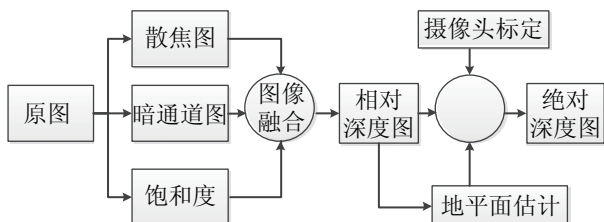


图 1 算法流程图

Fig.1 Flow chart of the proposed algorithm

计算相对深度图的基本思路是将图像中的散焦信息、暗通道先验信息以及图像饱和度信息综合起来, 通过图像融合技术求解获得。首先, 包含图像中物体散焦信息的散焦图可以用来表示场景相对深度, 但实际情况中的散焦信息并不完全准确, 这是由于在焦平面之外可能会存在锐利的边界, 而在焦平面上也可能会出现模糊的边缘。这些情况将会导致相对深度图的不准确, 进而影响地平面的估计。

其次, He 等^[13]基于大气散射模型提出了暗通道先验的概念, 并利用其实现了单幅图像的去雾工作。暗通道先验信息指出对于清晰的户外无雾图像, 非天空区域的大部分局部图像块中的像素至少有一个颜色通道具有很低的像素值。但在实际场景中, 雾、霾等情况是始终存在的, 且对远处的物体影响大于近处的物体, 因此暗通道图的亮度可以为场景深度的估计提供线索。但是, 由于暗像素在低饱和度区域 (如天空) 存在较少, 所以由暗通道图估计出的深度也会不可避免地存在错误。最后, 图像在 HSV 空间中的饱和度信息可用于纠正暗通道先验估计错误的区域, 提高结果的可靠性。

但是, 相对深度图并不能直接用来指导机器人的自主导航, 因此第 3 节的工作就是通过相对深度图求绝对深度。该工作基于这样一个假设: 场景中所有的物体都垂直于地平面, 而没有悬挂在空中的物体, 一幅图像可以用地平面和垂直地平面的物体来表示。这样, 物体的绝对深度可以根据其与地面的接触点到摄像头的距离求得。在该假设之下, 地平面和垂直物体之间的边界应该非常明显并且局部平滑。本文首先计算相对深度图在垂直方向上的梯度, 然后用双边中值滤波器滤掉细小纹理所带来的噪声, 最后在该垂直梯度图上采用一种新型的搜索曲线策略来获得需要的地平面边界。在获取地平面区域后, 利用摄像头焦距和成像几何关系, 并结合相对深度图即可计算出绝对深度图。

3 相对深度图的估计 (Estimation of the relative depth map)

3.1 计算散焦图

生成散焦图的过程包括: 首先使用薄棱镜模型^[14]估计图像阶跃边缘的宽度进而获取图像的散焦信息, 然后利用引导图像滤波将边缘信息扩散到整幅图像。

采用文 [12] 中局部散焦估计的方法获得粗糙的散焦图, 图 2 是该方法在 1 维情况下的原理图。用 2 个标准差不同的高斯函数对图像进行 2 次重滤波, 则原阶跃边缘幅值 E 与重滤波后幅值 E_1 、 E_2 的差在位置 x 处的比值为

$$k(x) = \frac{E(x) - E_1(x)}{E_1(x) - E_2(x)} \quad (1)$$

该比值最大的地方就是边缘的位置, 且利用该最大比值即可求得该边缘位置的散焦量。在实验中, 高斯函数的标准差分别取为 $\sigma_1 = 2$, $\sigma_2 = 4$ 。

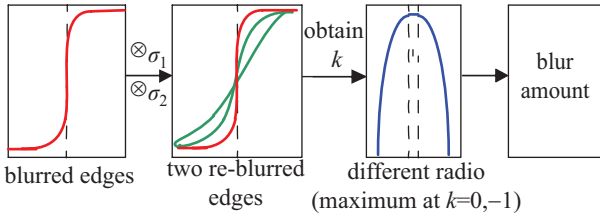


图2 局部散焦估计方法

Fig.2 Local blur estimation approach

能够证明, 式(1)中的比值在边缘位置 $x = -1$ 和 $x = 0$ 处取到最大. 当 $\sigma_1 \gg \sigma$, $\sigma_2 \gg \sigma$ 时, 能够估计边缘的半径:

$$\sigma \approx \frac{\sigma_1 \cdot \sigma_2}{(\sigma_2 - \sigma_1) \cdot k_{\max}(x) + \sigma_2} \quad (2)$$

具体到2维图像散焦信息的计算, 采用的是各向同性的2维高斯函数, 由于其各个方向都是1维高斯函数, 因此计算方法与上面1维的情形一样, 实验得到的粗糙散焦图如图3(b)所示.

该粗糙散焦图需要经过细化才能在后续步骤中使用, 按照文[15]提出来的算法, 此细化问题可以转化为最小化代价函数:

$$\mathbf{E}(\mathbf{t}) = \mathbf{t}^T \mathbf{L} \mathbf{t} + \lambda (\mathbf{t} - \hat{\mathbf{t}})^T \mathbf{A} (\mathbf{t} - \hat{\mathbf{t}}) \quad (3)$$

式中, \mathbf{t} 为细化后的散焦图, $\hat{\mathbf{t}}$ 是稀疏散焦图, \mathbf{A} 是对角矩阵, 当像素 (i, j) 位于边缘位置时值为1, 否则取为0, λ 是平滑系数, 用于控制第1项和第2项的权重, \mathbf{L} 是拉普拉斯矩阵 (matting Laplacian matrix), \mathbf{L} 中的元素值 $\mathbf{L}(i, j)$ 定义为

$$\sum_{k|(i \in W_k, j \in W_k)} \left\{ \delta_{ij} - \frac{1}{|W_k|} [1 + (\mathbf{I}_i - \boldsymbol{\mu}_k)^T \times (\boldsymbol{\sigma}_k^2 + \frac{\varepsilon}{|W_k|} \mathbf{U}_3)^{-1} (\mathbf{I}_j - \boldsymbol{\mu}_k)] \right\} \quad (4)$$

其中, δ_{ij} 为克罗内克符号, $\boldsymbol{\mu}_k$ 和 $\boldsymbol{\sigma}_k$ 分别是以像素 k 为中心的窗口 W_k 内所有像素 RGB 三通道亮度值的均值向量和协方差矩阵, \mathbf{I}_i 和 \mathbf{I}_j 为输入图像 \mathbf{I} 在像素 i 和 j 处的 RGB 值, ε 是正则化参数, $|W_k|$ 是窗口 W_k 的大小, \mathbf{U}_3 是3阶单位阵.

在式(3)中, 由代价函数对 \mathbf{t} 求偏导不难得出, 该优化问题可以转化为求解稀疏线性方程组:

$$(\mathbf{L} + \lambda \mathbf{A}) \mathbf{t} = \lambda \mathbf{A} \hat{\mathbf{t}} \quad (5)$$

由于直接求解式(4)非常耗时, 因此本文利用 He 等人在文[16]提出的引导图像滤波 (guided image filtering) 的方法来计算拉普拉斯矩阵, 这样可以在稍微牺牲图像质量的前提下较大程度地提

高算法效率. 图3(c)所示为粗糙散焦图细化后的结果.

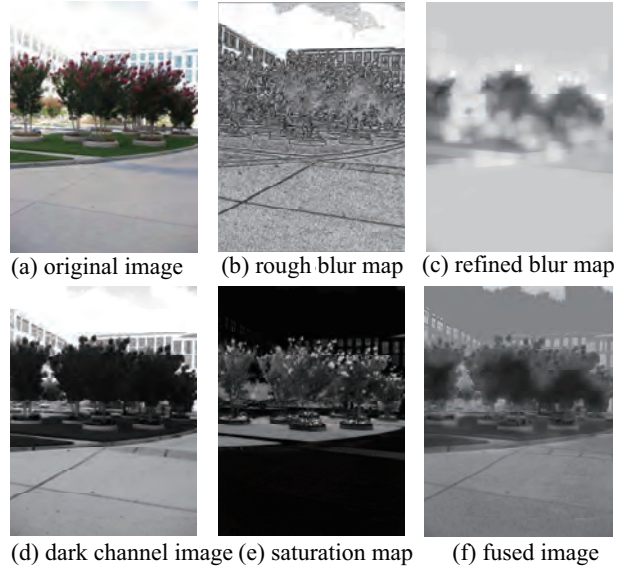


图3 相对深度图的计算

Fig.3 Calculation of the relative depth map

3.2 计算暗通道图

前已述及, 暗通道图的亮度能够为场景深度的估计提供线索. 一幅图像的暗通道可以定义为^[13]

$$J^{\text{dark}}(x) = \min_{c \in \{r, g, b\}} \min_{y \in \Omega(x)} J^c(y) \quad (6)$$

其中, c 为图像 RGB 三通道, J^c 是图像 J 的对应彩色通道, $\Omega(x)$ 是以 x 为中心的小区域. 图3(d)即为图3(a)求得的暗通道图.

暗通道图作为深度信息依据的主要缺陷在于暗像素在低饱和度区域可能不存在. 如图3中, 由于天空和地面区域的饱和度很低, 所以暗通道图在对应的区域取值很大. 为此, 使用图像 HSV 空间中的饱和度信息来对暗通道图中的深度错误进行纠正. 图3中原图的饱和度图如图3(e)所示.

3.3 图像融合

为了获得更为精确的深度信息, 本文将散焦图、暗通道图和饱和度图中的冗余互补信息进行融合, 图像融合的公式为

$$d_i = w_1 b_i J_i + w_2 J_i S_i^{(1 - \text{sigmoid}(y))} + w_3 b_i (1 - S_i)^{(1 - \text{sigmoid}(y))} \quad (7)$$

其中, b_i 是散焦图, J_i 是暗通道图, S_i 是饱和度图, d_i 是融合结果, 即相对深度图, w_1 、 w_2 和 w_3 是归一化权重系数, y 是当前像素 i 的归一化垂直坐标, sigmoid 函数为 $\text{sigmoid}(y) = 1/(1 + e^{-y})$.

式(7)中共有3项, 第1项 $b_i J_i$ 将散焦图与暗通道图进行融合, 第2项 $J_i S_i^{(1 - \text{sigmoid}(y))}$ 在 y 较小时

(天空区域的位置) 取值较小, 这样可以一定程度上抑制暗通道图在低饱和度区域的错误结果, 第 3 项 $b_i(1 - S_i)^{1-\text{sigmoid}(y)}$ 用以对第 2 项的抑制效果进行补偿. 图 4 给出了在不同权值组合下的融合结果, 由于所求的是相对深度图, 只要 w_1 、 w_2 和 w_3 设定在一个较宽松的范围内 (均保持在 0.25 到 0.5 之间), 最终得到的结果都很准确. 在实验中, w_1 、 w_2 和 w_3 根据实验获得的经验分别取为 0.4、0.3 和 0.3. 图 3(f) 为最终的融合结果, 即为原图的相对深度图.

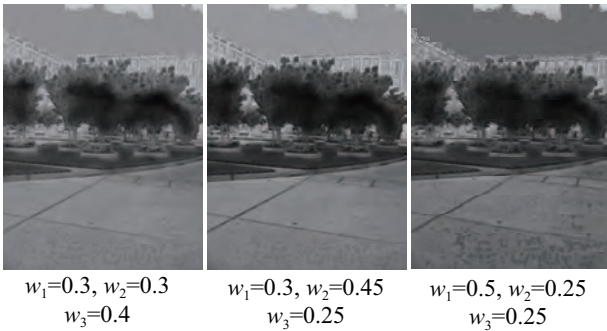


图 4 不同权重融合结果
Fig.4 Results of different weigh

4 3 维地平面与绝对深度图的估计 (Estimation of the 3D ground plane and absolute depth map)

4.1 估计 3 维地平面

方法基于这样一个合理假设, 即场景中所有的物体都垂直于地平面, 地面区域处在图像的底部并且局部连续. 在这个假设下, 需要准确估计出地平面和垂直物体的边界, 为此首先计算了相对深度图在垂直方向的梯度, 并使用阈值截断和中值滤波的方法去除了不显著的以及孤立的点, 这样求得的垂直方向梯度图如图 5(b) 所示. 这是一幅二值图像, 提取出每一列位置最低的非零点构成以下点集:

$$\Omega = \{G(i_x, j_x)\}, \quad x = 1, 2, \dots, N \quad (8)$$

其中, i_x 是图像水平方向坐标, j_x 是图像垂直方向坐标, x 是该集合点的索引, N 是该集合中元素的个数, 易见 N 的值最大不超过图像的列数. 这些点都分布在笛卡尔坐标系中, 图像的原点设在左上角.

因此, 基于前面提到的假设, 地平面的边界检测问题可以转化为在该垂直方向梯度图中利用集合 Ω 中的点拟合出一条尽量平滑的曲线, 当然还需要

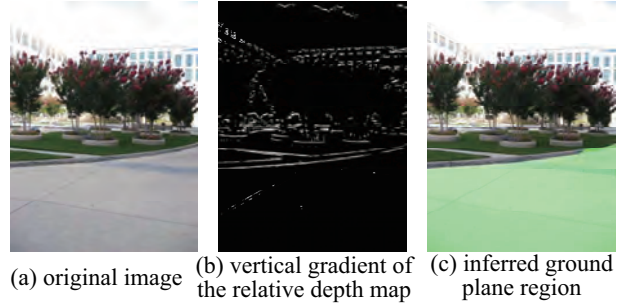


图 5 地平面的估计过程

Fig.5 Procedure of 3D ground plane estimation

附加一些限制条件. 该拟合问题可以表示为

$$\arg \max_{\tau} \int_x \tau(i_x) - \lambda (\|\nabla \tau(i_x)\|^2) \quad (9)$$

限制条件为

$$0 \leq \tau(i_x) \leq e j_x \quad (10)$$

$$b_{\min}(i_x) \leq T(i_x) \leq b_{\max}(i_x) \quad (11)$$

其中, $\tau(i_x)$ 为最终地面边界 τ 在 Ω 中索引为 x 的点所在列上的位置, 参数 λ 控制解的平滑程度, $T(i_x)$ 是 $\tau(i_x)$ 的一阶微分, $b_{\min}(i_x)$ 和 $b_{\max}(i_x)$ 分别表示 $\tau(i_x)$ 一阶微分的最高和最低范围. 求解该最优化问题非常耗时, 为提高算法效率, 采用 Tarel 等^[17] 提出的双边中值滤波器对地平面边界进行求解, 相应的算法步骤为

$$A(i_x) = \text{median}_{S_w}(j_x) \quad (12)$$

$$B(i_x) = A(i_x) - \text{median}_{S_w}(|j_x - A(i_x)|) \quad (13)$$

$$C(i_x) = B(i_{x-1}) + \max(\min(B(i_x) - B(i_{x-1}), b_{\max}(i_x)), b_{\min}(i_x)) \quad (14)$$

$$\tau(i_x) = \max(\min(C(i_x), j_x), 0) \quad (15)$$

其中, S_w 是窗口的大小, 图 4(c) 所示为 S_w 为 5×5 时求得地平面区域的结果, 其中绿色区域为检测到的地平面.

4.2 计算绝对深度图

在确定图像中的地平面区域之后, 采用文 [10] 中的方法, 利用摄像头的成像几何原理^[18] 获得真实的绝对深度信息, 其成像原理如图 6 所示.

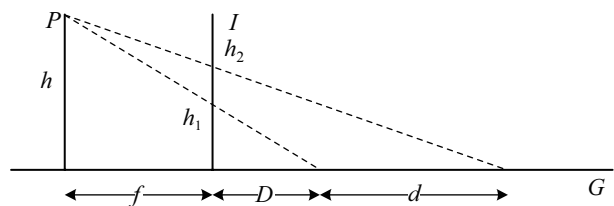


图 6 摄像头成像原理图

Fig.6 Schematic of camera imaging

图6中, P 为光学中心, h 为摄像头的高度, h_1 和 h_2 代表该物体投影到图像平面上的高度, f 为摄像头的焦距, I 为图像平面, G 为地平面, D 为图像中某一聚焦物体与地平面的交界点到图像平面的距离, 即相对深度. d 为图像中某物体到该聚焦物体的水平距离, 即为所求的绝对深度, 利用图6中的相似三角形关系容易推导得到下式:

$$d = f \frac{h(h_2 - h_1)}{(h - h_1)(h - h_2)} = D \frac{h(h_2 - h_1)}{h_1(h - h_2)} \quad (16)$$

由于已假设场景中所有的物体都垂直于地平面, 因此对于地面上的物体, 它的深度就设定为其与地面接触点的深度. 这样, 一旦地平面被提取出来, 即可利用式(16)计算出场景的绝对深度信息.

5 实验及结果 (Experiments and results)

5.1 离线实验

为了快速验证本文算法的有效性, 首先在离线计算机上用 MATLAB 进行编程实验, 实验在平台为 AMD® Athlon™ II X4 640 CPU, 3.12GB RAM.

实验中使用的是 Make3D 数据集^[19], 该数据集包含了 534 幅分辨率为 1704×2272 的室外图像以及相对应由激光扫描仪测得的分辨率为 86×107 的真实深度图.



图7 本文算法的地平面估计结果

Fig.7 Ground plane estimation results of the proposed algorithm

图7所示为本文算法检测3维地平面的部分实验结果, 其中第1行为原始图像, 第2行图像中的绿色区域(包括阴影部分对应的深色区域)为估计得到的地平面. 可以看出, 该算法对各种室外场景都有比较好的鲁棒性, 求得的地平面边界平滑连续, 不会受到阴影、地面纹理变化等干扰. 同时, 算法对图像中的障碍物能很好地识别, 如第1幅图像中的自行车、第2幅图像中的阶梯等.

另一方面, 将该算法最终获得的绝对深度图与 Cherian 等^[10]的结果进行比较, 结果如图8所示.

其中, 第1列为原始图像, 第2列为 Cherian 等人的结果, 第3列为本文实验结果, 第4列为激光扫描获得的结果.

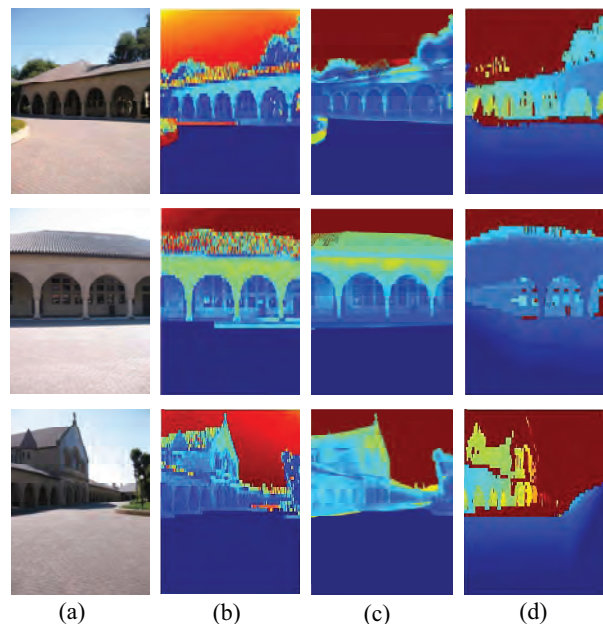


图8 绝对深度图结果比较

Fig.8 Comparison of absolute depth maps

可以看出, Cherian 等的算法和本文提出的算法都能较好地恢复场景的深度信息. 相比之下, 本文算法获得结果的分辨率更高, 而且在一些细节上比 Cherian 等的结果要好, 比如第2行图中的屋顶区域本文算法的深度结果更为平整. 同时, 将本文算法以及 Make3D 算法估计得到的深度图分别与网站^[19]提供的真实深度图进行了比较, 从 Make3D 数据集中选取 30 幅分辨率为 341×455 的图像, 平均运行时间和平均相对误差的结果如表1所示. 深度图的相对误差用 $|d - \hat{d}|/d$ 计算, 其中 d 为求得的深度值, \hat{d} 为真实深度值. 由表1可以看出, 虽然本文算法估计的深度图在精度上稍差于 Make3D 算法, 但在效率上远高于 Make3D 算法, 实用价值更大.

表1 本文算法与 Make3D 算法比较

Tab.1 Comparison between the proposed method and Make 3D

算法	所用时间 /s	平均相对误差
Make3D	52.32	0.8137
本文算法	0.28	0.8194

为了进一步验证本文算法的准确性与高效性, 还与 Hoiem 等^[8]的 Photo pop-up 算法进行了比较. 表2所示为两种算法对分辨率为 341×455 的图像的平均运行时间, 从中可以看出本文算法的运算效率远远高于 Photo pop-up 算法. 图9为两种算法运

行结果的比较, 其中第 1 行为原始图像, 第 2 行为 Hoiem 等的结果, 第 3 行为本文实验结果. 可以看出, 本文算法鲁棒性更高, 可以较好地排除阴影和纹理的干扰, 如第 1 列图像中的阴影. 同时, 本文算法也不会将垂直地面的物体检测为地平面, 如第 2 列图像中的自行车和第 4 列图像中的建筑物.

表 2 本文算法与 Photo pop-up 算法比较

Tab.2 Comparison between the proposed method and Photo pop-up

算法	所用时间 /s
Photo pop-up	35.45
本文算法	0.26



图 9 地平面估计结果比较

Fig.9 Comparison of estimated ground planes

5.2 实际环境中的机器人实验

为了验证本文算法在实际环境中的准确性与可行性, 还将该算法应用于轮式移动机器人的室外自主避障实验中.



图 10 轮式小车

Fig.10 The wheeled vehicle

5.2.1 机器人平台

实验中, 使用的是四轮驱动轮式小车, 如图

10 所示. 小车车顶安装了 1 部安卓智能手机, 它拥有 1GHz 双核 Cortex™-A9 处理器和 1GB 内存, 以及一个 500 万像素摄像头. 对该手机进行编程, 使其能够采集分辨率为 320×240 的图像, 然后利用本文提出的算法对采集到的图像进行实时处理, 生成控制指令, 并将控制指令通过蓝牙串口模块发送到小车控制板, 控制小车的动作.

5.2.2 小车自主避障实验

小车的自主避障实验在室外进行, 利用安装在车顶上的智能手机拍摄图像并进行处理, 得益于本文算法的复杂度比较低, 在 320×240 的图像分辨率下能达到 3 帧/秒左右的处理速度, 这样在低速情况下能基本满足地面机器人控制的实时性要求. 图 11 所示为室外场景中, 利用本文算法估计出来的 3 维地平面以及绝对深度图. 其中, 第 1 列为原始图像, 第 2 列为地平面区域估计结果, 第 3 列为最终求得的场景深度图.

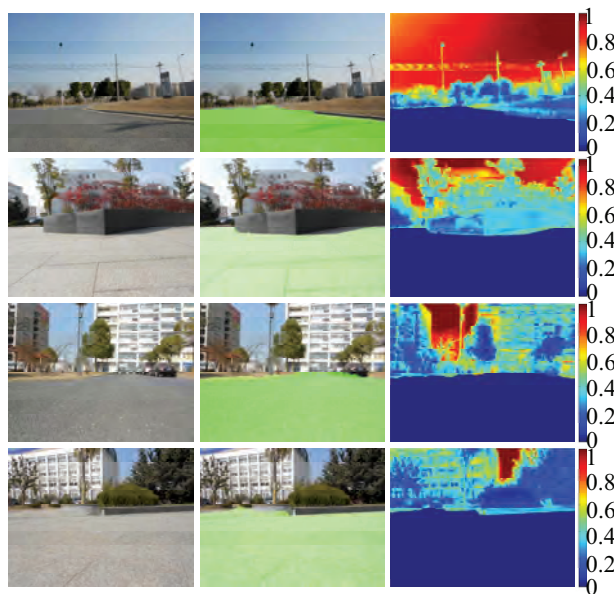


图 11 室外场景实验结果

Fig.11 Experimental results of outdoor scene

为了更好地说明算法用于小车自主避障的有效性, 还用手持摄像机录制了小车的一个避障运动全过程, 如图 12 所示. 其中, 第 1 列是从手持摄像机拍摄的视频中按时间顺序抽取的几帧图像 (人的角度), 第 2 列是从小车上的手机拍摄的图像序列中抽取与第 1 列图像大致相对应的几帧 (小车的角度), 第 3 列是利用本文算法处理第 2 列图像得到的地平面区域, 可以看出地平面可以被很好地检测到. 在实验中, 小车的速度保持稳定, 只需考虑运动方向的控制问题, 而且小车不去识别障碍物的类型, 只搜索最安全的路径, 最大程度地远离障碍

物, 保证自身的安全. 具体来说, 小车在运动过程中对获得的地平面边界进行搜索, 选择距离小车最远的边界点作为当前目标, 并根据新获得的数据不断更新前进方向, 达到避障的目的. 在图 12 描述的场景中, 小车根据检测到的地平面边界不断调整方向, 逐渐右转, 最终成功避开了前方的障碍.

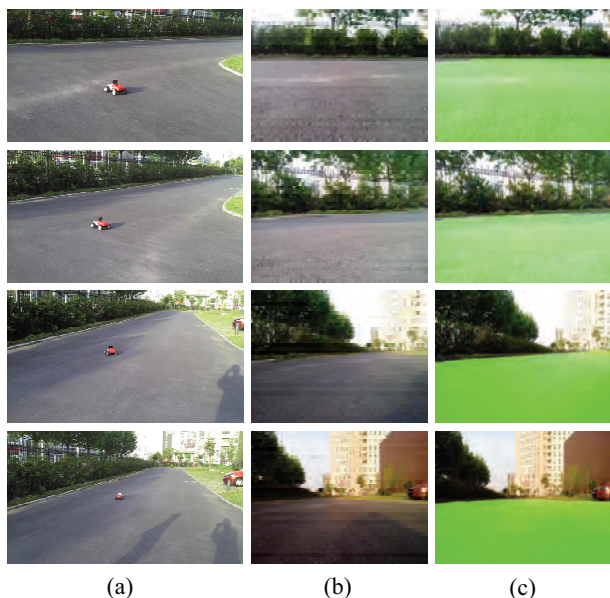


图 12 小车自主避障实验

Fig.12 Autonomous obstacle avoidance of the vehicle

6 结论 (Conclusion)

提出了一种基于车载单目图像的 3 维地面估计及场景深度估计算法. 该算法首先结合了图像的散焦信息、暗通道先验信息和饱和度信息, 估计出一幅比较精确的相对深度图, 然后基于场景中所有的物体都垂直于地平面的合理假设, 在相对深度图上估计出场景地平面的边界, 进而获得 3 维地平面区域, 最后利用摄像头成像原理求得场景的绝对深度图. 实验结果表明, 本文算法鲁棒性较高, 可以准确地估计出地平面和场景深度, 而且还具有很高的效率. 同时, 该算法能够实时处理车载智能手机拍摄得到的分辨率为 320×240 的图像, 并获得比较准确的场景信息, 指导轮式移动机器人的自主避障.

参考文献 (References)

- [1] Zhang R, Tsai P, Cryer J E, et al. Shape from shading: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(8): 690-706.
- [2] Chaudhuri S, Rajagopalan A N. Depth from defocus: A real aperture imaging approach[M]. Berlin, Germany: Springer-Verlag, 1999.
- [3] Loh A M, Kovese P. Estimation of surface normal of a curved surface using texture[C]//Proceedings of the Digital Image Computing: Techniques and Applications Conference. Piscataway, USA: IEEE, 2003.

- [4] Delage E, Lee H, Ng A Y. A dynamic Bayesian network model for autonomous 3D reconstruction from a single indoor image[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2006: 2418-2428.
- [5] Sun M, Saxena A, Ng A Y. Learning 3-D scene structure from a single still image[C]//11th Conference on Computer Vision. Piscataway, USA: IEEE, 2007: 1-8.
- [6] Saxena A, Chung S H, Ng A Y. 3-D depth reconstruction from a single still image[J]. International Journal of Computer Vision, 2008, 76(1): 53-69.
- [7] Saxena A, Chung S H, Ng A. Learning depth from single monocular images[C]//Advances in Neural Information Processing Systems. 2005: 1161-1168.
- [8] Hoiem D, Efros A A, Hebert M. Automatic photo pop-up[J]. ACM Transactions on Graphics, 2005, 24(3): 577-584.
- [9] Huttenlocher D, Felzenszalb F. Efficient graph based image segmentation[J]. International Journal of Computer Vision, 2004, 59(2): 167-181.
- [10] Cherian A, Morellas V, Papanikolopoulos N. Accurate 3D ground plane estimation from a single image[C]//International Conference on Robotics and Automation. Piscataway, USA: IEEE, 2009: 2243-2249.
- [11] Saxena A, Sun M, Ng A Y. Make3D: Learning 3D scene structure from a single still image[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(5): 824-840.
- [12] Hu H, Haan G. Adaptive image restoration based on local robust blur estimation[C]//9th International Conference on Advanced Concepts for Intelligent Vision Systems. Berlin, Germany: Springer-Verlag, 2007: 461-472.
- [13] He K, Sun J, Tang X. Single image haze removal using dark channel prior[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2009: 1956-1963.
- [14] Hecht E. Optics[M]. 4th ed. Hoboken, USA: John Wiley & Sons Inc., 2001.
- [15] Levin A, Lischinski D, Weiss Y. A closed form solution to natural image matting[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2006: 61-68.
- [16] He K, Sun J, Tang X. Guided image filtering[C]//11th European Conference on Computer Vision. Berlin, Germany: Springer-Verlag, 2010: 1-14.
- [17] Tarel J P, Hautiere N. Fast visibility restoration from a single color or gray level image[C]//11th Conference on Computer Vision. Piscataway, USA: IEEE, 2009: 2201-2208.
- [18] Zhang Z Y. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- [19] Make3D: Convert your still image into 3D model[EB/OL]. [2013-01-14]. <http://make3d.cs.cornell.edu/>.

作者简介:

向文辉 (1988-), 男, 硕士生. 研究领域: 机器人技术, 计算机视觉.

刘羽 (1988-), 男, 博士生. 研究领域: 图像处理, 计算机视觉, 模式识别.

曹洋 (1980-), 男, 博士, 讲师. 研究领域: 机器人技术, 计算机视觉, 模式识别.