

基于立体视觉平面单应性的智能车辆 可行驶道路边界检测

郭春钊¹ 山部尚孝¹ 三田诚一¹

摘要 道路检测是智能车辆及先进驾驶辅助系统 (Advanced driver assistance systems, ADAS) 研究的关键问题之一. 本文提出了一种基于立体视觉的可行驶道路区域与非道路区域间边界的检测方法. 该方法基于立体视觉平面单应性建立了一个隐马尔科夫模型 (Hidden Markov model, HMM). 针对该模型, 我们应用 Viterbi 算法, 并提出了一种巧妙的状态序列的观测概率函数, 以寻找道路/非道路边界的最优状态序列. 实验结果证明了该方法在各种典型且复杂的实际道路场景中的有效性和鲁棒性.

关键词 智能车辆, 道路检测, 立体视觉, 单应, 隐马尔科夫模型, 全局优化

引用格式 郭春钊, 山部尚孝, 三田诚一. 基于立体视觉平面单应性的智能车辆可行驶道路边界检测. 自动化学报, 2013, 39(4): 371–380

DOI 10.3724/SP.J.1004.2013.00371

Drivable Road Boundary Detection for Intelligent Vehicles Based on Stereovision with Plane-induced Homography

GUO Chun-Zhao¹ YAMABE Takayuki¹ MITA Seiichi¹

Abstract Road detection is one of the key issues for intelligent vehicles and advanced driver assistance systems (ADAS). In this paper, we present a stereovision-based approach for estimating the boundary between the drivable road region and the non-road region. It is based on a formulation of stereo with the ground-plane-induced homography in a hidden Markov model (HMM). Under this formulation, we employ the Viterbi algorithm and propose a sophisticated measure of the probability of the state sequence to find the most likely road/non-road boundary. Experimental results on a wide variety of typical but challenging real road scenes have substantiated the effectiveness as well as robustness of the proposed approach.

Key words Intelligent vehicles, road detection, stereovision, homography, hidden Markov models, global optimization

Citation Chun-Zhao Guo, Takayuki Yamabe, Seiichi Mita. Drivable road boundary detection for intelligent vehicles based on stereovision with plane-induced homography. *Acta Automatica Sinica*, 2013, 39(4): 371–380

尽管众多学者数十年来一直在进行研究并且取得了显著的进展, 道路检测仍是智能车辆的关键问题之一. 其中, 基于视觉的方法是该领域的一个重要分支. 一些方法使用单目摄像机通过利用道路外观的特征提取道路区域. 例如, Nieto 等^[1] 计算每个像素的纹理方向, 然后通过投票方案寻找道路的消失点 (Vanishing point), 最后利用颜色线索来定位道路边界. Franke 等^[2] 把道路检测定义为最大后验概率 (Maximum a posteriori,

MAP) 估计问题, 基于灰度、颜色及纹理特征对回旋曲线 (Clothoid) 道路模型的参数进行优化. McCall 等^[3] 通过提取人工绘制的车道线来确定道路. Kuhnl 等^[4] 运用慢特征分析 (Slow feature analysis, SFA) 得到具有颜色和纹理特征的低阶特征集, 进而训练基于图像分块的 GentleBoost 分类器. 由于道路图像的纹理随与摄像机之间的距离而变化, 因此一些方法采用颜色分析, 在不用的颜色空间对道路外观进行建模. 例如, 文献 [5] 利用色调亮度饱和度 (Hue-saturation-intensity, HSI) 的颜色特征对道路模式进行建模. 文献 [6] 利用了 RGB 颜色空间中高斯分布以及混合高斯模型. 这些基于外观的方法可以较好地应用于某些特定环境中. 然而在道路并不充分对应于具有先验定义特征的模型的情况下, 上述方法便缺乏有效性. 另外, 预定义的模型也难以包含所有可能的道路图

收稿日期 2012-03-13 录用日期 2012-11-26
Manuscript received March 13, 2012; accepted November 26, 2012

丰田工业大学智能车辆研究中心资助
Supported by Research Center for Smart Vehicles, Toyota Technological Institute, Japan

本文责任编辑 王飞跃
Recommended by Associate Editor WANG Fei-Yue

1. 丰田工业大学智能车辆研究中心, 名古屋 4688511, 日本
1. Toyota Technological Institute, Nagoya 4688511, Japan
该文的英文版同时发表在 *Acta Automatica Sinica*, vol. 39, no. 4, pp. 371–380, 2013.

像. 因此, 在复杂场景中这些方法通常仅能够检测出与其模型一致的部分可行驶道路. 此外, 还有一些方法研究了基于单目摄像机的连续时间图像序并且利用连续图像间的对应目标的位移来检测道路. 例如, 文献 [7] 应用逆向光流技术实现多尺度道路区域的自适应分割. 文献 [8] 利用从运动信息中恢复三维场景结构 (Structure from motion, SFM) 的技术估计基于地图的道路边界模型. 这些基于运动的方法可以实现可行驶道路的一般检测并给出目标位移和场景结构/深度等信息. 然而, 当摄像机不稳定、光流估计不具备较高的鲁棒性时, 以上方法就不能取得令人满意的结果. 最近, 人们又重新开始研究基于单幅静止图像的几何估计方法. 这些方法通常将图像分割, 进而基于纹理梯度、透视、模糊等单目线索推断得出最有可能的三维结构. 尽管这些方法已经取得了显著的成果^[9], 但它们仍不能满足在自动驾驶和先进驾驶辅助系统 (Advanced driver assistance systems, ADAS) 等方面的直接应用.

当提取的特征具有较强的差异性和鲁棒性时, 上述基于单目的方法可以得到良好的结果. 然而, 当特征具有噪声或不断变化时, 这些方法便可能失效. 一般而言, 基于立体视觉的方法较之基于单目的方法具有更好的鲁棒性, 这是因为基于立体视觉的方法可以提供诸如三维三角测量等信息, 并且对于尺度丢失和车辆动态运动具有较强的鲁棒性. 给定一组图像对, 基于立体匹配的方法通过匹配并计算视差来提取场景的三维特征. 文献 [10–11] 阐述了城市三维重建问题. 然而, 由于需要解决每个像素的对应问题, 从而该类基于计算视差图的三维结构重建问题需要相当大的计算量. 另外, 还有一些其他的方法可以利用立体线索而无需计算视差. 文献 [12–13] 利用道路平面单应性来寻找像道路像素的对应关系. 与基于立体匹配的方法相比, 尽管这些方法仅仅适用于平坦道路的假设, 但它们对于自动驾驶应用中可行驶道路检测系统而言更为简便而有效. 本文所提出的算法即属于该类方法.

本文提出了基于立体视觉的可行驶道路区域和非道路区域间边界的检测方法. 该方法基于立体视觉平面单应性建立了一个隐马尔科夫模型 (Hidden Markov model, HMM). 针对该模型, 我们应用 Viterbi 算法, 并提出了一种巧妙的状态序列的观测概率函数, 以寻找道路/非道路边界的最优状态序列. 本文与其他相关工作的区别主要体现在以下几个方面:

1) 本文算法面向复杂场景, 并强调系统的鲁棒性. 在复杂场景中, 道路通常具有不一直的表面, 并且其外观亦随时间和空间的变化而出现较为显著的改变. 因此, 通过构建道路外观描述来检测这样的复杂道路是不可靠的. 为了克服这些困难, 我们应该依

据道路的几何性质而非其外观来进行可形式道路区域的分类, 因为几何性质可以反映车辆可行驶道路的的必要特征. 本文基于平坦道路假设, 将可行驶道路区域定义为地面上车辆前方相对平坦且可以安全通过的区域. 因此, 在本文算法中, 我们利用道路的平坦几何性质作为检测图像中可行驶道路边界的唯一准则, 而视觉特征仅仅用于提供估计边界位置的观测依据.

2) 基于 HMM 的立体表述来定义道路检测问题, 其优越性如下: a) 全局优化. 寻找最有可能边界的每一步并非出于局部的考量, 而是作为全局优化问题的一部分, 以使问题得到最优解决. b) 空间和时间的平滑性. 通过惩罚边界点空间变化的不连续性以及时间预测的偏差保证平滑性.

本文算法是在基于立体视觉的导航系统的框架下进行开发, 该系统集成于实验所用德智能车辆内, 如图 1 所示. 左侧是一辆小型电动汽车, 配有两台计算机、一台立体摄像机、四台单束激光扫描仪、一台 GPS 以及其他传感器. 右侧是一辆混合动力汽车, 配有六台计算机、一台立体摄像机、两台四束激光扫描仪、一台 GPS 以及其他传感器. 另外, 我们之前曾开发了一个基于马尔科夫随机场 (Markov random field, MRF) 的可行驶道路的检测算法, 并且利用置信传播算法 (Belief propagation) 推断整个输入图像中道路/非道路分类的二值标记. 然而, 基于立体视觉平面单应性的道路检测问题, 并不需要检测道路内部像素, 而只许如本文所述的找到匹配与不匹配区域间的边界即可, 这样可以大大减少计算量.



图 1 用于实验的智能车辆

Fig. 1 Our experimental intelligent vehicles

本文结构如下: 第 1 节介绍自动平面单应性的自动估算方法. 第 2 节阐述了基于 HMM 的可行驶道路边界检测算法. 第 3 节给出了不同照明和天候条件下实际复杂道路场景的实验结果. 最后给出本文的结论.

1 平面单应性的自动估计

在两个摄像机和一个空间平面组成的射影几何中, 平面上点的图像通过一个平面单应矩阵对应于第二个视图中相应的图像点, 如图 2(a) 所示. 从点 x 到点 x' 的映射为平面 π 诱导的单应矩阵 H . 在

图像平面间, 有

$$x' = Hx \quad (1)$$

如图 2(b) 所示, 当利用道路平面诱导的单应矩阵对右图像进行变换并与左图像进行比较时, 由式 (1) 可知只有道路平面上的点可以较好地重合, 其他点则不会. 因此, 道路边界检测实际上是立体匹配问题中较为简单的一种形式, 这是因为道路/非道路边界仅仅存在于匹配与不匹配的区域之间. 为了自动第计算地面诱导的单应矩阵, 本文提出一种结合了特征对应与区域优化的估算方法. 该方法总结为如下步骤, 其应用实例如图 3 所示.

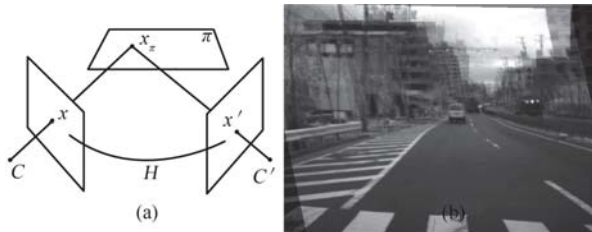


图 2 立体视觉平面单应性 ((a)) 和左基准图像与右变换图像的对照比较 ((b))

Fig.2 The homography induced by a plane ((a)) and comparison between the transformed right image and the original left image ((b))

1) 类道路区域的初步检测. 输入第一个图像对后, 本文算法首先进行车辆前方道路亮度的直方图分析, 所得结果即输入道路外观模型以选择最合适的 ANN 分类器. 该模型利用不同道路表面类型的真实路面数据训练得出. 所选择的分类器随后用于类道路区域的初步检测. 这些区域去除了大部分道路平面以外的特征点, 从而使得单应性估计的初始化具有更高的准确性和鲁棒性.

道路外观模型如图 4 所示. 模型中道路外观类由道路向量 (μ, σ) 定义, 其中 μ 和 σ 分别为直方图的均值和标准差. 每一类具有相应的反向传播 ANN 分类器, 用于类道路区域的初步检测. 该分类器具有 2 层、25 个输入、7 个隐层节点以及 2 个输出. 对学习而言, 目前直方图汇集了 7 类道路向量的 100 张道路图像. 应用这些图像分块, 对直方图进行分析, 所得结果映射到与道路向量对应的道路外观模型. 另外, 模型中每个 ANN 分类器都是在利用独立的监督学习算法使用数十个同一类道路图像进行训练.

2) 特征点图生成. 我们利用 Harris 检测器来检出作为特征点的角点^[16] 从第一个输入图像中独立提取, 它们能以亚像素精度确定角点的位置. 此外, 我们只保留图像对类道路区域内的角点以生成特征点图.

3) 假定对应. 该步的目的在于提供一个初始点对应集. 通过利用式 (2) 中归一化交叉相关 (Nor-

malized cross correlation, NCC) 来匹配角点图上的特征点来实现, 式 (2) 度量了两点及其相邻点间的相似度.

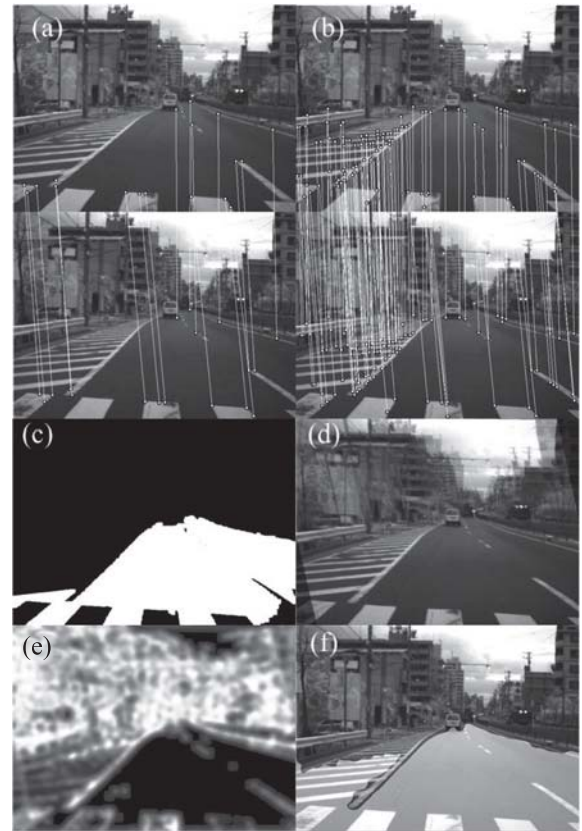


图 3 基于特征的检测 ((a) 具有道路外观模型所选 ANN 分类器的内围层对应; (b) 不具有 ANN 分类器的内围层对应; (c) 近似道路区域; (d) 变换后的右侧图像与原始左侧图像的叠加图像; (e) 平滑化及直方图均衡化的绝对差分图像; (f) 计算区域 (高光区域))

Fig.3 Feature-based detection ((a) Inlier correspondences with selected ANN classifier of road appearance model; (b) Inlier correspondences without ANN classifier; (c) Road-like region; (d) Superimposed image of the transformed right image and the original left image; (e) Smoothed as well as histogram equalized absolute difference image; (f) Computational region (highlighted region))

$$NCC(u, v) = \frac{\sum_{(u,v) \in W} [i_1(u, v) - \bar{i}_1][i_2(u, v) - \bar{i}_2]}{\sqrt{\sum_{(u,v) \in W} [i_1(u, v) - \bar{i}_1]^2 [i_2(u, v) - \bar{i}_2]^2}} \quad (2)$$

式中, W 为计算窗口, $i_k(u, v)$ 为第 k 个图像分块, \bar{i}_k 为图像分块的均值. 利用角点对间最大的 NCC 值对假定对应进行赋值.

4) 利用随机样本一致性 (Random sample consensus, RANSAC) 算法进行单应矩阵计算. 对假定对应集应用 RANSAC 算法估计单应矩阵以及与单应矩阵一致的内围层对应点.

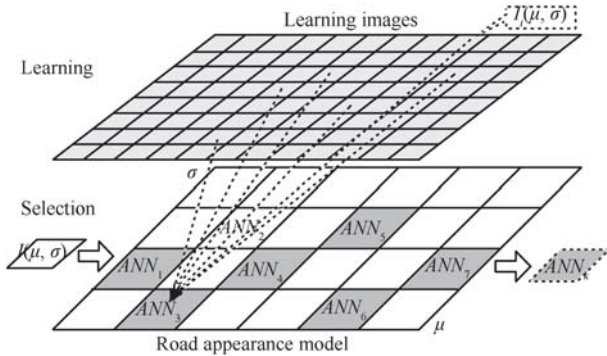


图 4 道路外观模型的学习和选择 (从学习过程中得到预定义的 ANN 分类器. 检测时将选择学习与当前输入具有同一种实际道路外观的模型)

Fig. 4 Learning and selection of the road appearance model (The pre-defined ANN classifiers are obtained from a learning process. For detection, the one, which is learned from the same kind of real road appearance as the current input, will be selected.)

5) 优化. 利用全部内围层对应重新估计 H , 运用 LM 算法 (Levenberg-Marquardt algorithm, LMA) 使总投影误差函数达到最小.

$$\sum_{x_j \in S_i} (\hat{H}x_j - x'_j)^2 \quad (3)$$

式中, S_i 为左侧图像的内围层集, x_j, x'_j 为对应点, \hat{H} 为进行优化的单应矩阵.

6) 基于区域的动态优化. 运用所得单应矩阵对右图像进行变换再与左图象进行差分, 并进行高斯平滑化及直方图均衡化处理. 保留处理后图像中亮度小于道路阈值的像素. 这一过程实际上等价于右变换图像与左基准图像的 SAD 匹配. 接下来进行连通成分分析得到最大的连通区域. 随后利用连通区域所有像素信息对单应矩阵进行优化. 优化过程的迭代步骤可以修正特征点误匹配引起的潜在误差, 使得基于区域的优化实现准确性更高、鲁棒性更强的单应矩阵估计. 这里需要寻找某个单应矩阵 H , 使得式 (4) 中图像对间的总灰度误差函数 $E(H)$ 最小.

$$e(H) = \sum_{x \in Rc} |I_l(x) - I_r(Hx)| \quad (4)$$

式中, I_l, I_r 分别为左、右图像, Rc 为计算区域, 即连通成分分析后的保留区域. 基于 LMA 的非线性优化过程如图 5 所示.

图 3 给出的实例证明了本文提出的单应矩阵自动估计的有效性. 另外, 使用/不使用 ANN 分类器

的特征对应比较如表 1 所示. 从图 3 和表 1 中可以看出, 道路外观模型 ANN 分类器的应用有效减小了特征点图的大小, 降低了对应的数目, 并且去除了道路平面外的界外对应点. 因此, 基于单应矩阵估计并且具有道路外观模型 ANN 分类器的道路检测具有更高的准确性和鲁棒性.

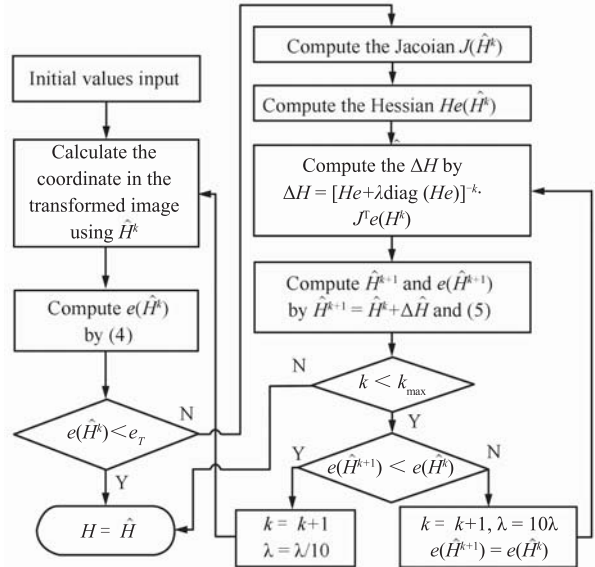


图 5 基于 LMA 的非线性优化过程

Fig. 5 Nonlinear optimization process using LMA

表 1 使用/不使用 ANN 分类器的道路外观模型的特征对应
Table 1 Feature correspondences with/without ANN classifier of road appearance model

| 图像 | 角图上的点 | 假定对应 | 内围层对应 | 道路平面上的内围层 |
|-----|-------|------|-------|-----------|
| (a) | 75 | 39 | 18 | 100 % |
| (b) | 2005 | 233 | 100 | 85 % |

这里所介绍的单应矩阵自动估计算法使得本文算法对于任意输入立体数据均有效, 并且不需要进行事先的校准. 所得单应矩阵为接下来的边界估计步骤揭示出可行驶道路的内近平坦的几何性质. 需要说明的是, 单应矩阵的自动估计需要至少 4 个地面上的特征点, 而大多数实际道路场景可以满足这一要求.

2 可行驶道路边界估计

给出变换图像以及基准图像, 本文所提方法当前的目标是找到匹配与不匹配区域间最有可能的边界点位置, 这可以通过在 HMM 中应用 Viterbi 算法来实现.

2.1 隐马尔科夫模型 (HMM)

HMM 是一种统计马尔科夫模型, 模型中所建模的系统假定为具有未观测 (隐) 状态的马尔科夫过

程. 这里首先简要介绍标准连续时间有限状态马尔科夫模型. 模型包括两个互相关过程: 一个潜在的不可观测马尔科夫链, 具有有限个数状态、一个状态转移概率矩阵以及一个初始状态概率分布; 一个与各状态相关的概率密度函数集. HMM 具有下列特征:

1) 模型中状态的个数 N . 定义状态集 $S = \{S_1, S_2, \dots, S_N\}$, $S_i, i = 1, 2, \dots, N$ 为各单独状态. T 时刻的状态定义为 $q_t \in S, 1 \leq t \leq T$, 其中 T 为观测序列的长度.

2) 不同观测符号的个数 M . 观测符号对应于所建模系统的物理输出. 定义符号集 $V = \{V_1, V_2, \dots, V_M\}$, 其中 $V_i, i = 1, 2, \dots, M$ 为各单独符号.

3) 状态转移概率矩阵 A , 即 $A = \{a_{i,j}\}$, 其中

$$a_{i,j} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i, j \leq N \quad (5)$$

常数

$$0 \leq a_{i,j} \leq 1 \quad (6)$$

并且

$$\sum_{j=1}^N a_{i,j} = 1, \quad 1 \leq i \leq N \quad (7)$$

4) 观测符号概率矩阵 B , 即 $B = \{b_j(k)\}$, 其中

$$b_j(k) = P(O_t = V_k | q_t = S_j), \quad 1 \leq j \leq N, 1 \leq k \leq M \quad (8)$$

并且 O_t 为时刻 t 的观测符号.

5) 初始状态分布 Π , 即 $\Pi = \{\pi_i\}$, 其中

$$\pi_i = P(q_1 = S_i), \quad 1 \leq i \leq N \quad (9)$$

2.2 利用 HMM 进行可行驶道路边界检测

隐马尔科夫模型已经成功应用于时域模式识别中, 如语音、手写、手势识别等. 本文所提方法阐述了 HMM 中可行驶道路边界检测问题, 如图 6 所示. 与数据随时间变化的常规 HMM 不同, 所提 HMM 自左向右随图像坐标的列而变化, 状态为列中图像像素的垂直位置. 观测源于参考图像与变换图像的匹配. 最有可能的隐状态序列, 即道路/非道路边界点, 其存在于匹配与不匹配的区域之间.

运用 Viterbi 算法解决 HMM 中的道路检测问题, 该算法可以选择最佳的状态序列, 使得对于给定观测序列的状态序列的似然度达到最大. 观测序列如下列递归方程所示:

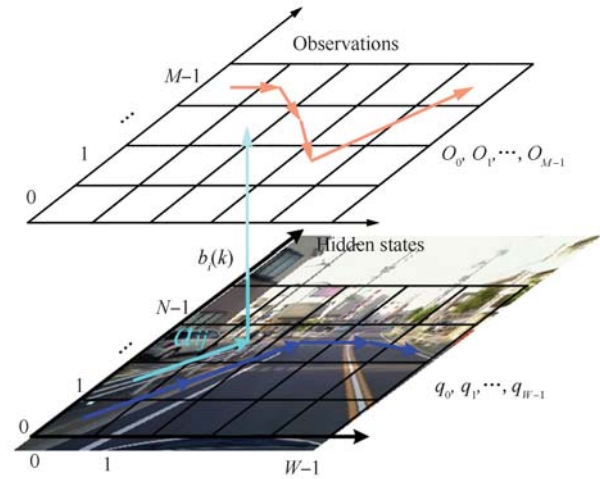


图 6 用于可行驶道路边界检测的 HMM (网格对应于图像像素)

Fig. 6 The HMM for drivable road boundary detection (The grids correspond to the image pixels.)

$$X_{0,j} = b_j(O_0) \cdot \pi_j$$

$$X_{i,j} = b_j(O_i) \max_k (X_{i-1,k} \cdot a_{k,j}) \quad (10)$$

式中, $X_{i,j}$ 为结束于状态 j 的 i 列状态序列的最大概率. 对式 (10) 取对数, 上述方程等价于

$$X_{0,j} = V_{0,j} + P_{0,j}$$

$$X_{i,j} = V_{i,j} + \max_k (X_{i-1,k} + t_{i,j}^{i-1,k} + t_{i,j}^{i,j'}) \quad (11)$$

式中, $V_{i,j}$ 为像素 (i, j) 作为边界点的观测度量. $t_{i,j}^{i-1,k}$ 为像素 $(i-1, k)$ 至像素 (i, j) 的路径成本. $t_{i,j}^{i,j'}$ 为偏离轨迹的路径成本, 而 j' 为通过相邻序列图像的自行运动或单应变换得到的预测. $X_{i,j}$ 为直至第 i 列的最佳评分.

所提阐述中, 边界下方的像素可以良好地进行匹配, 边界上方的像素则不然. 因此, 定义 $V_{i,j}$ 如下:

$$V_{i,j} = \lambda_1 D_{i,j} + \lambda_2 G_{i,j} \quad (12)$$

$$D_{i,j} = \sum_{k=h_1}^j \frac{c(i, k)}{j - h_1} - \sum_{k=j}^{h_2} \frac{c(i, k)}{h_2 - j}$$

$$G_{i,j} = \frac{\sum_{k=j-\Delta w}^j c(i, k) - \sum_{k=j}^{j+\Delta w} c(i, k)}{2\Delta w}$$

$$c(i, k) = \sum_m \gamma_m (\Phi_r^m(i, k) - \Phi_t^m(i, k))^2$$

λ_1, λ_2 为加权参数. $c(i, k)$ 为参考图像与变换图像间像素 (i, k) 的兼容成本. $\Phi_r(\cdot)$ 和 $\Phi_t(\cdot)$ 为所构建的特征向量, γ_m 为第 m 个元素的权值. 对于 RGB

空间中的彩色图像, 向量是 9 维的, 包括 3 个颜色数值以及 1 个 x 和 y 方向上 6 维颜色梯度向量. 对于灰度图像而言, 向量是 3 维的, 包括亮度值以及 1 个二维梯度向量. 使用梯度向量的目的在于通过强化与亮度边缘一致的匹配不连续性改善估计性能. 第 1 项 $D_{i,j}$ 度量了当前扫描线中红色区域与黄色区域间像素平均兼容成本的差异, 如图 8 所示. 第 2 项 $G_{i,j}$ 计算像素 (x, y) 兼容成本的梯度值, 同时通过强化与亮度边缘一致的边界改善估计性能. $[h_1, h_2]$ 给出了扫描线的计算范围, Δw 为梯度计算的常数, 如图中绿色与红色框所示.

$t_{i,j}^{i-1,k}$ 为空间平滑的转移矩阵, $t_{i,j}^{i,j'}$ 为时间平滑的转移矩阵. 由于相邻列间道路/非道路边界不会显著变化, 因此定义 $t_{r,\theta_i}^{k,\theta_{i-1}}$ 如下:

$$t_{i,j}^{i-1,k} = -\min(\tau_s, \kappa_s |k - j|) \quad (13)$$

式中, κ_s 为正的权值参数, τ_s 为两点间距离饱和的截断阈值. 这样可以通过惩罚不连续性保证路径的空间平滑. 同时, 截断阈值的使用可以保留不连续性, 这在某些情况下可能是必要的.

类似地, 定义 $t_{r,\theta_i}^{r',\theta_i}$ 如下:

$$t_{i,j}^{i-1,j'} = -\min(\tau_t, \kappa_t |j' - j|) \quad (14)$$

式中, κ_t 为正的权值参数, τ_t 为当前状态与预测状态间偏差饱和的截断阈值. 状态转移的步骤如图 7 所示.

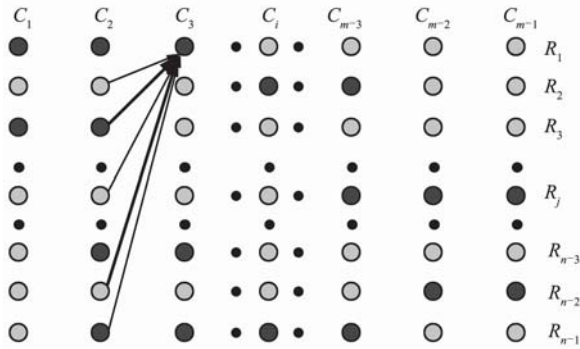


图 7 状态转移 (行表示状态, 列表示每一步的观测. 深灰色表示高观测似然度, 浅灰色表示该步的低观测似然度. 箭头表示相邻步状态间的转移概率. 粗箭头表示高转移概率 (低成本), 细箭头表示低转移概率 (高成本))

Fig. 7 The state transition (The rows are states and the columns are observations at each step. The dark gray indicates high observation likelihood while the gray indicates low observation likelihood at this step. The arrows indicate the transition probabilities between the states in the adjacent steps. Thick arrows represent high transition probabilities (low costs) while thin arrows indicate low transition probabilities (high costs).)

这样, 通过运用所提 Viterbi 算法得到的最优路

径便对应于近似平坦道路的可行驶道路边界. 图 8 给出了一个实例结果. 需要注意的是, 最后的边界是通过用直线连接相邻的边界像素而产生的. 这样的处理可以保证近似垂直边界段的情况下道路边界的封闭性.

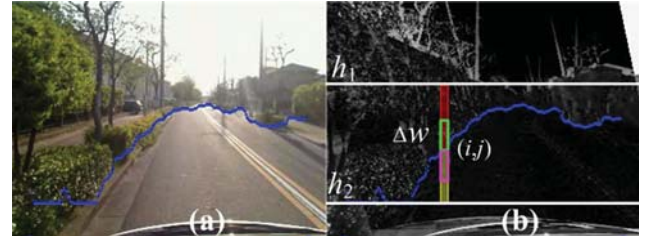


图 8 可行驶道路边界检出例 ((a)) 和隐马尔科夫模型的最优状态序列 ((b))

Fig. 8 An example result of the drivable road boundary detection ((a)) and the most likely sequence of hidden states in the difference image ((b)) (The colored areas and frames indicate the computational regions for the observation measure.)

2.3 无纹理区域的处理

立体视觉算法通常在计算无纹理区域时得到错误的结果. 在本文算法中, 对于与道路区域连接的无纹理区域, 如黑色的汽车、白色的墙壁、蓝色的天空等, 当将图像由一个视图转换为另一视图时, 尽管其像素位置发生改变, 但是其特征向量并不改变, 为此这些区域经常会引起检测错误. 该问题如图 8 所示. T_1, T_2 为参考图像的无纹理区域, T'_1, T'_2 为变换图像相应的区域. 所提方法对所有可行驶道路边界内的匹配像素进行分类. 显然, T'_1 中一部分像素匹配得很好便引起这样的错误. 因此, 无纹理区域的道路分类应该从整个区域而非区域内个别像素决定. 如图 9(c) 所示, 某一缺少纹理区域属于可行驶道路当且仅当其轮廓内部所有像素均良好地对应, 如 T_2, T'_2 . 对于其他无纹理区域, 如 T_1, T'_1 , 轮廓的差异表示它们不遵循道路平面单应性, 因而实际上是在道路平面之外. 在这种情况下, T'_1 中一定有某些像素存在于检测边界之外, 如图 8 中红色部分所示.

在本文方法中, 我们通过以下步骤来处理无纹理区域引起的问题:

- 1) 提取变换图像的无纹理区域. 文献 [19] 将无纹理区域定义为这样的区域, 在该区域内固定大小方形窗口的均方水平亮度梯度小于某一给定阈值. 本文方法计算了四个方向 (水平、垂直及两条对角线方向) 上图像的梯度以避免连接不同的无纹理区域.
- 2) 标记并记录每个相连的无纹理区域.
- 3) 计算每个无纹理区域检测边界外异常像素的比例.
- 4) 如果无纹理区域的异常比例高于某一给定阈

值时, 便将整个区域从检测道路区域中剔除.

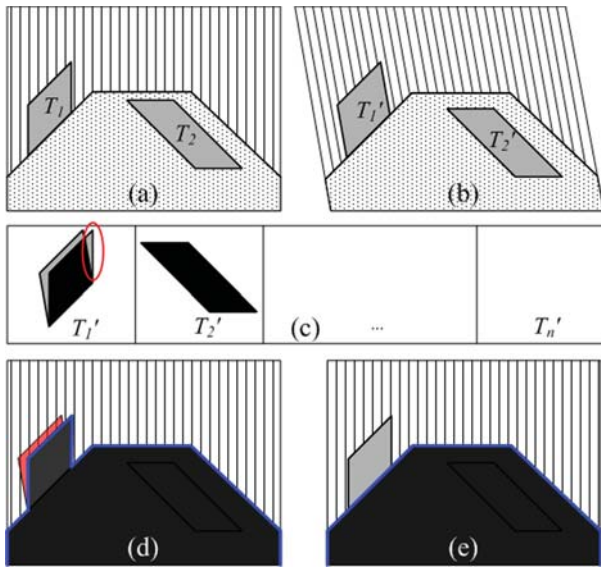


图9 无纹理区域的处理 ((a) 和 (b) 分别为参考图像和变换图像, 其中点表示道路区域, 条纹表示非道路区域, 灰色表示无纹理区域; (c) 为变换图像的无纹理区域列表, 黑色表示具有匹配特征的像素; (d) 和 (e) 分别为无纹理区域处理前后的检测结果)

Fig. 9 Dealing with textureless regions ((a) and (b) are the reference image and the transformed image, respectively, where dots indicate the road region, strips indicate the non-road region and gray indicates the textureless regions. (c) is the list of textureless regions in the transformed image, where black indicates the pixels with matched features. (d) and (e) are the detection results before and after dealing with textureless regions, respectively.)

3 实验结果

实验时, 我们使用本文算法来处理各种典型且复杂的真实场景中. 实验规范如表 2 所示. 计算时间小于 100 ms/帧, 可以满足实时应用. 计算时间可以通过优化代码及利用硬件加速可以进一步的缩短.

由于正常环境下的道路检测已获得较好的验证, 这里我们仅关注复杂场景. 实验总共得到了超过 5 000 帧的实例结果, 所提系统对其中超过 92% 的图像进行了准确的可行道路区域边界检测.

3.1 复杂情境下的表现

不同复杂照明和天候条件下典型的实验结果如图 10~12 所示. 图 10 中, (a)~(e), (i)~(j), (l)~(m), (p)~(t) 及 (u)~(x) 表明所提方法克服噪声以及在强背光、重阴影、复杂道路模式等条件下定位可行道路边界的能力. (k)~(l) 及 (p)~(v) 表明所提方法可以检测作为障碍的汽车、摩托车、行人等. (a)~(j) 及 (m)~(y) 表明所

提方法也可以检测作为非道路区域的路边栅栏、灌木丛、路缘等. 从 (f)~(i)、(p)~(r)、(u)~(v) 及 (w)~(y) 可以看出, 所提方法可以良好地应用于弯路、十字路口和弯曲/倾斜的道路.

对于在雨夜提取数据的可行道路边界检测结果如图 11 所示. 无论是不均匀的道路表面 (k)~(m), 还是雨滴 (a)~(e) 和雨刷器 (r)~(y) 引起的遮挡, 或是下雨及黑夜造成的低能见度, 所提方法仍可以准确地找到可行道路边界.

对于在大雪天提取数据的检测结果如图 12 所示. 由于雪、阴影、车轮轨迹、灯光反射等原因, 道路表面十分复杂, 然而由于应用了基于近似平坦几何性质的排他判断准则, 得到了所有检测可行道路边界内的道路区域. 需要注意的是, 如图 11 (v) 所示, 检测出了处于可行道路边界之外的运动场上的一个雪人.

3.2 讨论

源于激光测距仪 (LRF) 的道路检测通常包括以下步骤. 首先确定道路平面, 该平面将作为点云分类的参考. 二维激光传感器, 如 SICK, 只能检测超过位移阈值突出于地面的障碍, 并且假定余下地面是可行驶的. 改进的高解析度激光传感器, 如 Velodyne, 可以提供更多三维点云, 利用它们可以确定道路表面的地面模型. 这种基于 LRF 的方法可以给出准确而可靠的测距结果. 然而, 其测量是稀疏的并且检测结果取决于目标的大小和反射率, 因此检测概率随距离而减小, 同时有效检测距离相对较短. 计算机视觉可以提供长距离充分的信息, 然而其受亮度变化的约束.

因此很多系统采用了多传感器融合的方法以使的结果更加准确和完整. 一个著名的例子是 Stanley 的道路检测系统, 该自主汽车赢得了 2005 年 DARPA 挑战赛^[6].

表 2 实验规范

Table 2 Experimental specifications

| | |
|----------|----------------------------|
| CPU | Xeon(R) X5550 @2.67 GHz |
| 内存 (RAM) | 12.0 GB |
| 操作系统 | Window Xp Professional Sp2 |
| 编程语言 | C++ |
| 实验数据 | Grayscale 图像 |
| 分辨率 | 320 × 240 |
| 计算时间 | < 100 ms/帧 |

在 Stanley 系统中, 其激光传感器的最大有效距离约为 22 米, 然而其视觉模块通常可以看到前方 70 米. 因此, Stanley 首先从激光分析中提取车辆前方的一个四边形区域, 并将其投射到摄像机图像. 然后基于混合高斯模型的自适应计算机视觉算法, 利用四边形内的图像像素作为可行道路的训练样本, 从而找出激光距离外的道路. 学习以两种可能的方

式与图像相适应: 通过调整之前确定的内部高斯模型至实际图像像素, 以及通过引入新的高斯模型而丢弃旧的模型. 第一种方式使得 Stanley 适应于缓慢变化的照明条件; 第二种方式使其能够迅速适应于新的路面颜色.

作为比较, 本文运用 Stanley 视觉系统以及本文所提方法共同应用于处理不均匀草地区域的图像序列. 实例比较结果如图 13 所示. 由于图像所有区域均是可行驶的, 用下半部分图像定义训练样例的四边形. 图 13 中间一行的序列表明 Stanley 计算机视觉系统的快速适应性, 其中视觉模块成功地由草地极快适应于光秃区域. 然而, 该图像序列同时显示出 Stanley 视觉结果不能用于转弯行为决策, 这是因为图中土路区域显然是可行驶的, 而 Stanley 并不能从远处检测出这一点. 原因在于当激光主要从近距离扫描草地表面时, 对于远距离区域没有训练样例, 因此不能够被检测出.

只有在 Stanley 进入土路区域后并且得到其像素作为训练样例, 分类才会改变. 然而, 本文算法

可以成功在远距离找到可行驶道路边界, 而不受完全不同的道路表面的影响, 这可以体现复杂环境中所提方法的优势. 对于定量的比较, 我们计算了两个系统对于图像序列共 157 帧的检测错误率. 对于 Stanley 系统对于不均匀的道路颜色/纹理场景的总错误率高于 30%, 而所提方法的错误率则低于 2%.

4 结语

本文提出了基于立体视觉的可行驶道路区域和非道路区域间边界的估计方法. 本文的第一个贡献在于提出了基于特征对应及区域优化的立体视觉单应矩阵的自动估计方法; 本文的第二个贡献是构建了基于立体视觉单应性及 HMM 的道路边界检测算法. 本文的第三个贡献则是提出了一个巧妙的状态序列的观测概率函数以寻找最优的道路/非道路边界. 上述贡献提高了在无任何先验知识条件下可行驶道路边界检测的准确性及鲁棒性. 实验结果证明了该方法在各种典型且复杂的实际道路场景中的有效性和鲁棒性.



图 10 基于所提方法的典型城市街道的实例检测结果

Fig. 10 Example detection results of the proposed approach in typical urban streets

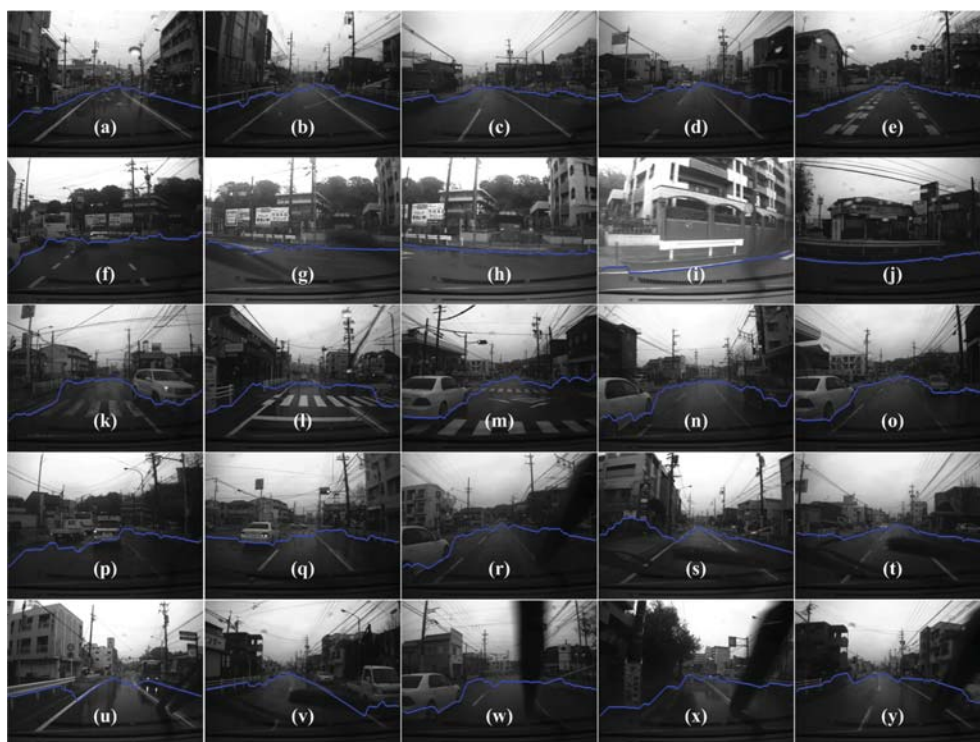


图 11 基于所提方法在下雨黑夜的实例检测结果

Fig. 11 Example detection results of the proposed approach in a dark rainy evening

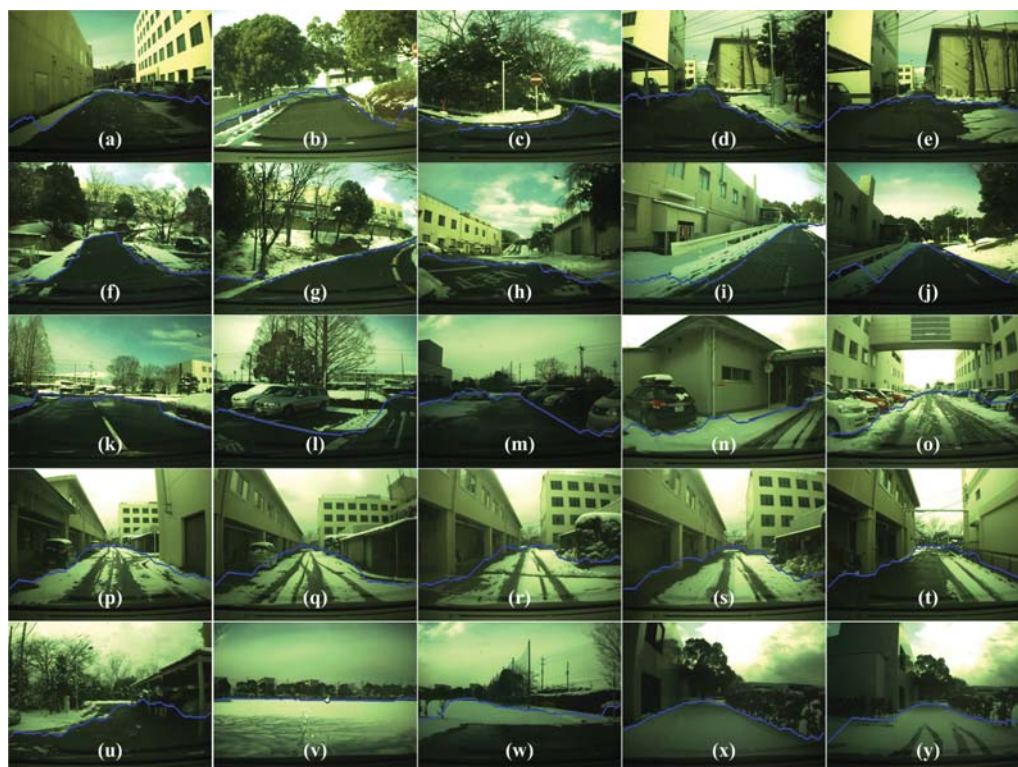


图 12 基于所提方法在大雪天的实例检测结果

Fig. 12 Example detection results of the proposed approach in a heavy snowy day

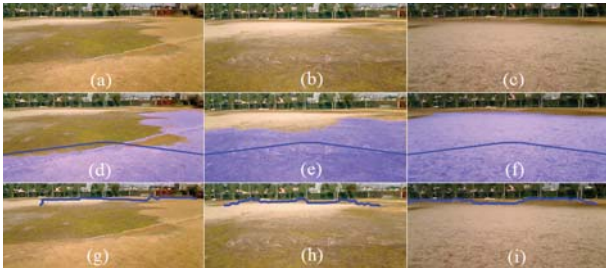


图 13 对于相同数据 (a)~(c) 应用 Stanley 视觉系统 (d)~(f) 和所提系统 (g)~(i) 的实例比较结果

Fig. 13 Example comparison results with the same data (a)~(c) between the Stanley's vision system (d)~(f) and the proposed system (g)~(i)

References

- Nieto M, Salgado L. Real-time vanishing point estimation in road sequences using adaptive steerable filter banks. In: Proceedings of the 9th International Conference on Advanced Concepts for Intelligent Vision Systems. Berlin, Heidelberg: Springer-Verlag, 2007. 840–848
- Franke U, Loose H, Knoppel C. Lane recognition on country roads. In: Proceedings of the 2007 IEEE Intelligent Vehicles Symposium. Istanbul, Turkey: IEEE, 2007. 99–104
- McCall J C, Trivedi M M. Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation. *IEEE Transactions on Intelligent Transportation Systems*, 2006, **7**(1): 20–37
- Kuhnl T, Kummert F, Fritsch J. Monocular road segmentation using slow feature analysis. In: Proceedings of the 2011 IEEE Intelligent Vehicles Symposium. Baden-Baden, Germany: IEEE, 2011. 800–806
- Sotelo M A, Rodriguez F J, Magdalena L. Virtuous: vision-based road transportation for unmanned operation on urban-like scenarios. *IEEE Transactions on Intelligent Transportation Systems*, 2004, **5**(2): 69–83
- Thrun S, Montemerlo M, Dahlkamp H. Stanley: the robot that won the DARPA grand challenge. *Journal of Field Robotics*, 2006, **23**(9): 661–692
- Lookingbill A, Rogers J, Lieb D, Curry J, Thrun S. Reverse optical flow for self-supervised adaptive autonomous robot navigation. *International Journal of Computer Vision*, 2007, **74**(3): 287–302
- Darms M, Komar M, Lueke S. Map based road boundary estimation. In: Proceedings of the 2010 IEEE Intelligent Vehicles Symposium. San Diego, CA: IEEE, 2010. 609–614
- Gupta A, Efros A A, Hebert M. Map based road boundary estimation. In: ECCV. 2010
- Akbarzadeh A, Frahm J M, Mordohai P, Clipp B, Engels C, Gallup D, Merrell P, Phelps M, Sinha S, Talton B, Wang L, Yang Q, Stewenius H, Yang R, Welch G, Towles H, Nister D, Pollefeys M. Towards urban 3D reconstruction from video. In: Proceedings of the 3rd International Symposium on 3D Data Processing, Visualization, and Transmission. Chapel Hill, NC: IEEE, 2006. 1–8
- Gallup D, Frahm J M, Pollefeys M. Piecewise planar and non-planar stereo for urban scene reconstruction. In: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, CA: IEEE, 2010. 1418–1425
- Kubota S, Nakano T, Okamoto Y. A global optimization algorithm for real-time on-board stereo obstacle detection systems. In: Proceedings of the 2007 IEEE Intelligent Vehicles Symposium. Istanbul, Turkey: IEEE, 2007. 7–12
- Seki A, Okutomi M. Robust obstacle detection in general road environment based on road extraction and pose estimation. In: Proceedings of the 2006 IEEE Intelligent Vehicles Symposium. Tokyo, Japan: IEEE, 2006. 437–444

- Guo C, Mita S, McAllester D. Robust road detection and tracking in challenging scenarios based on Markov random fields with unsupervised learning. *IEEE Transactions on Intelligent Transportation Systems*, 2012, **13**(3): 1338–1354
- Hartley R, Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2003
- Harris C, Stephens M. A combined corner and edge detector. In: Proceedings of the 4th Alvey Vision Conference. Manchester, UK: Organising Committee AVC, 1988. 147–151
- Gill P E, Murray W. Algorithms for the solution of the non-linear least-squares problem. *SIAM Journal on Numerical Analysis*, 1978, **15**(5): 977–992
- Rabiner L R. A tutorial on hidden Markov models and selected applications in speech recognition.
- D. Scharstei and R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, in Int. J. of Computation Vision, 2002, 47(1/2/3): 7–42.



郭春钊 丰田工业大学智能车辆研究中心助理教授, 分别于 2002 年和 2007 年在中国科学技术大学自动化系取得学士和博士学位, 主要研究方向为机器学习, 计算机视觉, 智能车辆, 仿生机器人。E-mail: guo@toyota-ti.ac.jp

(GUO Chun-Zhao Assistant professor at Toyota Technological Institute,

Japan. He received the B.S. and Ph.D. degrees from University of Science and Technology of China (USTC) in 2002 and 2007, respectively. His research interest covers machine learning, computer vision, intelligent vehicles, cooperative driving, and biomimetic robots. Corresponding author of this paper.)



山部尚孝 丰田工业大学智能车辆研究中心硕士研究生, 主要研究方向为计算机视觉和智能车辆。

E-mail: sd11434@gmail.com

(YAMABE Takayuki Master student at Toyota Technological Institute, Japan. His research interest covers computer vision and intelligent vehicles.)



三田诚一 丰田工业大学智能车辆研究中心主任, 特任教授, 丰田工业大学芝加哥分校联合教授, 分别于 1969 年, 1971 年和 1989 年在京都大学取得学士, 硕士和博士学位, 主要研究方向为数字信号处理, 机器学习, 智能车辆。E-mail: smita@toyota-ti.ac.jp

(MITA Seiichi Distinguished professor at the Research Center for Smart Vehicles, Toyota Technological Institute, Japan. He received the B.S., M.S., and Ph.D. degrees from Kyoto University in 1969, 1971, and 1989, respectively. He is a joint professor with the Toyota Technological Institute at Chicago. His research interest covers digital signal processing for recording and communication channels and machine learning applications for vehicle environment recognition.)