

基于粒子滤波的驾驶员视线自动校准算法

付先平¹⁾ 王亚飞^{1),2)} 袁国良¹⁾ 管 潇³⁾ PELI Eli⁴⁾ 罗 罡⁴⁾

¹⁾(大连海事大学 信息科学技术学院 辽宁 大连 116026)

²⁾(大连理工大学 物理与光电工程学院 辽宁 大连 116024)

³⁾(杜兰大学数学系 新奥尔良 美国 70118)

⁴⁾(哈佛大学医学院 波士顿 美国 02114)

摘 要 虽然驾驶员的注视点定位对研究驾驶行为和提高驾驶安全有重要作用,但是光线不均匀和驾驶员头部的大角度转动造成了传统视线跟踪算法必需进行的校准过程在真实驾驶环境下无法实现.文中根据车辆环境和驾驶员注视行为特点,构建驾驶员注视视线区域和头部姿态的映射关系,以实现视线自动校准.文中利用稳健的面部特征定位及三维几何结构估计头部姿态,利用粒子滤波对头部姿态进行学习 and 跟踪.文中使用辅助采样的方法解决粒子滤波在给定的状态空间上的粒子贫化和权值分配问题.文中的自动校准算法误差较小,可以满足驾驶安全和驾驶行为研究的需要.

关键词 视线跟踪;头部姿态;校准;粒子滤波;驾驶行为

中图法分类号 TP319 **DOI号** 10.11897/SP.J.1016.2015.02390

Automatic Calibration Algorithm of Driver Gaze Based on Particle Filter

FU Xian-Ping¹⁾ WANG Ya-Fei^{1),2)} YUAN Guo-Liang¹⁾ GUAN Xiao³⁾ PELI Eli⁴⁾ LUO Gang⁴⁾

¹⁾(Information Science and Technology College, Dalian Maritime University, Dalian, Liaoning 116026)

²⁾(School of Physics and Optoelectronic Engineering, Dalian University of Technology, Dalian, Liaoning 116024)

³⁾(Department of Mathematics, Tulane University, New Orleans 70118, USA)

⁴⁾(Medical School, Harvard University, Boston 02114, USA)

Abstract Driver's eye gaze tracking is very important for driver behavior study and driving safety. But non-uniform light and big rotation angle of head movement cannot make traditional eye gaze tracking algorithm be used in real driving environment. The proposed method based on vehicle environment and driver's eye gaze behavior characteristic is to implement automatic calibration by building the mapping between head pose and gaze zone. The head pose is estimated by robust facial landmarks detection and three-dimensional geometry configuration. And the head movement in the given state space is tracked and learned by particle filter. In order to solve the particle impoverish and weight assignation problem of particle filter, auxiliary sampling method is used. The experimental results show the error rate is low enough to meet the need of driver safety and driving behavior study.

Keywords eye gaze tracking; head pose; calibration; particle filter; driver behavior

收稿日期:2014-04-19;在线出版日期:2014-11-28. 本课题得到国家自然科学基金(61272368,61370142)、辽宁省高等学校优秀人才支持计划(LR2013021)、交通运输部应用基础研究项目(2015329225300)资助. 付先平,男,1971年生,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为工程系统自然场景感知、多媒体技术、图像与视频处理. E-mail: fxp@dlmu.edu.cn. 王亚飞,男,1988年生,博士研究生,主要研究方向为视线估计和机器学习. 袁国良,男,1982年生,博士研究生,主要研究方向为多媒体技术、图像处理与视频处理. 管 潇,男,1989年生,博士研究生,主要研究方向为符号计算. PELI Eli,男,1951年生,博士,教授,主要研究领域为图像处理与理解. 罗 罡,男,1968年生,博士,助理教授,主要研究方向为视觉科学.

1 引言

视线跟踪技术在人机交互、医疗和驾驶行为研究方面有重要的作用. 目前成熟的精度较高的商业视线跟踪仪一般会利用腮托或头戴等方法限制头部运动, 通过基于红外光线的瞳孔-角膜反射视线跟踪方法获取视线参数. 在头部运动受限制的情况下, 头部和眼睛位置估计可以得到较好的准确度^[1-2]. 如果不固定头部运动, 就需要采用运动传感器或多相机对头部运动进行补偿, 但这种视线追踪系统精度不高.

真实驾驶环境下的基于视频的视线追踪方法是智能交通领域的研究热点^[3-4]. 但受驾驶环境和自然环境因素影响, 传统视线追踪方法几乎不可用. 首先, 在自然环境下, 光线比较复杂, 红外光源在户外受阳光影响后不稳定, 传统的瞳孔-角膜反射方法可用性差. 再次, 驾驶员的头部转动幅度很大, 转动角度有时大于 180° , 并且在自然环境下也不利于采用固定或限制头部运动的方法, 因此自然场景下的注视点定位问题一直难以实现. 其中, 视线校准是视线追踪或注视点估计的必要环节, 虽然在简单环境下有专门的校准过程, 校准的算法也比较多^[5], 但在复杂场景下由于条件限制, 一般很难进行校准. 而传统视线跟踪方法必须进行校准, 这也说明了如果要实现注视点定位的自动校准就必须采用新的注视点定位方法.

对于较大的头部旋转角度, 许多研究者通过头部朝向来估计视线方向^[6], 将视线估计问题转换为头部姿态估计问题. 因为人眼的有效视场是有限的, 所以正常情况下头部朝向可以代表 70% 的视线方向, 而驾驶员在注视某个对象时总是转动头部面向该对象. 本文也将驾驶员的视线方向用头部的朝向来表示. 驾驶员的头部姿态估计方法利用脸部特征点的位置、形状、外观或左右瞳孔的距离等参数估计头部的朝向, 头部位置估计通常需要多摄像机和复杂的面部模型来进行初始化. 如 Murphy-Chutorian 等人^[7]提取驾驶员脸部的外观特征, 构建梯度方向直方图, 通过支持向量回归估计视线方向和头部朝向. Tawari 等人^[8]则在驾驶视频图像中定位驾驶员的眼角、鼻角等多个特征点位置, 使用三维模型分析其几何结构来估计头部姿态. 头部姿态的 3 个自由度如图 1 所示.

同时, 交通领域的研究人员利用眼动仪研究汽车驾驶员注视点分布规律等问题, 得到了部分驾驶员视觉搜索行为的统计特性. 如袁伟等人^[9]基于此提出基于马尔可夫链的驾驶员视觉转移特征, 运用

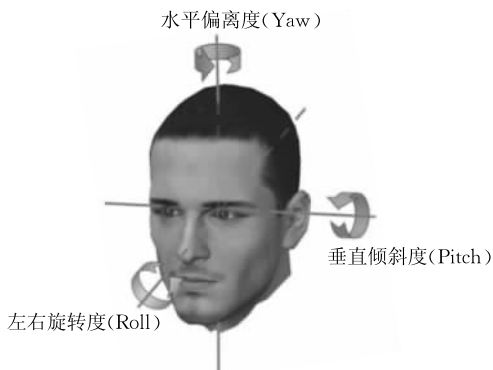


图 1 文献[6]中的头部姿态自由度

动态聚类方法, 对驾驶员注视点的视野平面解析坐标进行聚类, 通过对注视行为平稳分布特征分析得出: 长时间驾驶后各区域的注视点分布的概率会趋于稳定, 并主要集中在若干个区域. 因此可以在驾驶环境中利用头部方向估计代替视线方向估计.

根据驾驶习惯和驾驶员头部方向统计结果^[10], 驾驶员大多数时间都是朝前看, 并且以相对固定的头部转动角度观察左右后视镜、车内后视镜和挡风玻璃的不同区域, 因此头部方向可以采用追踪和学习的方法被实时估计出来, 以实现自动的校准^[11]. 本文算法基于以下两个条件, 一是当驾驶员调整好座椅时, 驾驶员的头部朝向和后视镜等驾驶环境部件的相对位置是固定的, 车内的这些部件可以用来进行视线的校准; 二是驾驶员通常具有一定的驾驶习惯, 例如如何观看后视镜, 打开转向灯前后如何观察后视镜等.

本文算法是在文献[11]的方法上进行了改进, 文献[11]的主要贡献: 一是针对不同的头部姿势和视线方向的自动分类方法提出了一种新的算法; 二是将自动人脸检测、样本集学习算法和粒子滤波追踪估计合并为一个循环结构, 从而使跟踪系统能够自动进行. 本文的改进之处在于: 利用较少形变的几何特征定位方法提高头部姿态的精确度, 同时使用稳健的稀疏人脸追踪方法加快特征点搜索速度; 在状态空间模型中, 对头部姿态进行粒子滤波追踪和学习, 使用辅助采样的方法解决粒子贫化和权值分配问题. 本文不仅侧重于头部姿态的粒子滤波学习, 包括了水平偏移度、垂直倾斜度和左右旋转度(即 Yaw、Pitch 和 Roll)共 3 个自由度; 还基于注视区域的聚类分析优化了注视区域的划分, 使其更符合自然驾驶状态的注视特性.

本文第 2 节介绍驾驶员视线自动校准算法; 第 3 节主要介绍驾驶员状态模型及粒子滤波算法在头部姿态的学习和估计方法中的应用; 第 4 节给出本

文算法的实验结果,并将本文算法与真实数据进行多组对比;第 5 节对本文工作进行总结,并指出进一步的工作。

2 驾驶员视线自动校准算法

我们提出的驾驶员视线自动校准算法由 3 部分构成,包括头部区域确定、基本帧的学习和视线区域的估计,如图 2 所示。确定头部区域简化了人脸特征提取,再使用基于粒子滤波框架的驾驶员状态模型进行基本帧学习,通过度量不同区域的图像距离,确定头部姿态与视线区域的映射关系,从而实现驾驶员视线的自动校准。其中,视线区域是根据驾驶员注视特性和视觉搜索行为特征预先划定的。

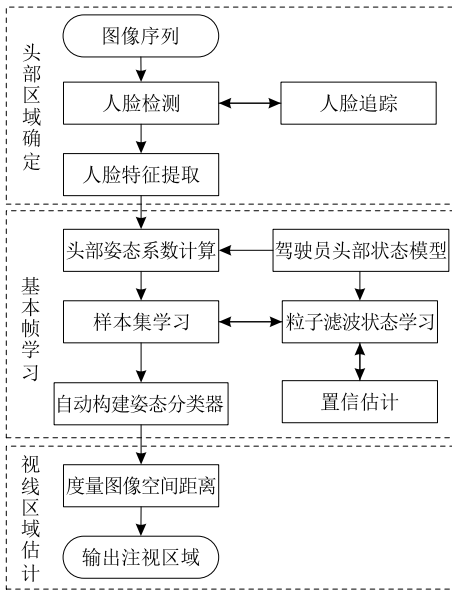


图 2 算法的流程

首先,在头部区域确定阶段,脸部区域在我们接收的整幅图像中被确定,其他区域即被抑制。受光照及驾驶环境影响,人脸检测效率不高,必要的人脸追踪算法被采用以提高人脸区域确定的速度和准确度。本文采用文献[12]的人脸稀疏追踪结合改进的Viola方法^[13]检测驾驶员的脸部,增加了对侧脸等不同姿态人脸的检测,使用追踪结合检测的方法主要是为提高特征点选取的速度和准确度。当系统识别脸部失败,基本帧的学习可能导致最后的相应视线区域的样本集有误,因此在初始化时在固定的时间内接收的图像均进行脸部检测。脸部检测算法在一定时间周期内会自动重复。同时,我们增加遮挡因子及面积因子以降低人脸追踪误差和人脸深度信息缺失带来的影响,两因子均根据统计特性学习得到。对于遮挡因子较大的人脸,即人脸区域被遮挡较多

的人脸在基本帧学习时降低其置信水平。对于面积因子较大的人脸,即尺度发生变化的人脸在基本帧学习时暂时不接受。

在图像序列中利用改进的文献[14]中的方法提取人脸特征,并自动对感兴趣的特征进行特殊标注。通常情况下,遮挡、光线条件、头部大幅度转动和其他影响因素限制了许多已有算法的应用。混合树模型把人脸每个特征点作为一部分,通过全局的混合捕捉人脸不同姿态下的特征点变化。而AAM(Active Appearance Model)模型易错估人脸整体轮廓,CLM(Constrained Local Model)模型则会牺牲鼻子特征点的精确度。由于本系统可以自动进行鼻子中心点的标注,因此系统不需要对一个指定的人进行特别的训练以获得相应的参数,因此不需要驾驶者的特别合作;另外,算法对于驾驶者戴眼镜和大幅度的移动头部具有很好的鲁棒性。不同驾驶员头部大幅度转动条件下的特征提取过程如图 3 所示。其中“×”为自动标注的鼻子顶点位置。



图 3 人脸特征点

其次,在基本帧学习阶段,关键帧表示为与我们划分的区域相对应的头部水平偏离度(Yaw)、垂直倾斜度(Pitch)和左右旋转度(Roll)。我们通过对驾驶员状态模型进行粒子滤波学习,基于图像上定义的距离把所有的样本分为与事先划定的视线区域相同的集合类,并最终用每个集合的中心来代表相对应的视线区域。驾驶员状态模型及粒子滤波学习算法会在本文第 3 节详细阐述。学习算法需要在每一时间段结束后进行重复,使得样本集中的元素能够以更高的置信概率定位视线区域。本文采用与文献[15]相似的学习算法的框架。基于对驾驶者大部分时间会正视前方这个假设,我们选取定位鼻子顶点作为兴趣点并进行加权平均,最后得到的值被认为驾驶者直视前方的中心坐标,取为原点,同时设定相应的

视线区域编号. 中心的选取需要在特定的时间周期内进行重复使得其可信度增高. 视线区域的头部姿态角度关系在系统初始化之前需根据头部传感器给出的数据进行校准对应.

最后, 在视线区域估计阶段, 驾驶员头部的中心由事先确定的校准点进行加权平均算出, 转化为相对应的头部姿态, 最后在集合类中选取距离最近的集合指标作为其估计的视线区域. 头部姿态由兴趣点在三维空间中根据几何关系计算角度得到. 并用此基于粒子滤波算法的估计结果在图像的估计状态和样本学习结果的距离空间度量视线区域结果.

其中, 我们加入了估计的置信映射来帮助了解估计的准确性. 关键帧的集合为 $\mathbf{M} = \{\mathbf{M}_j | 1 \leq j \leq k\}$, 其中 k 是关键帧的数量, 这一数量由事先确定的视线区域决定. 为了建立最优的基本帧的集合, 在每一帧中的图像元素都会被分析, 进而找到其相对应的置信概率最高的视线区域. 当我们确定样本集后, 新的图像帧就可以与样本集的中心进行比较, 这里我们用条件分布 $p(\mathbf{M} | \mathbf{M}_j)$ 来表示在新的图像帧的状态已知的情况下推测其属于样本 j 的概率, 在求出相应的离散分布后我们选择概率最大的样本的指数作为当下驾驶者的视线区域号. 为了计算简便, 样本集的基数会尽可能地小. 此外, 头部姿态的估计模型还保留了随机变量的相关矩阵 $\mathbf{A}_x = E(\mathbf{M} - \boldsymbol{\mu})(\mathbf{M} - \boldsymbol{\mu})'$, 此为头部姿态的协方差矩阵, 并以此为依据估计残差.

3 粒子滤波学习过程

粒子滤波方法或序贯蒙特卡罗方法是一组通过直接实现贝叶斯递归方程来估计状态空间后验密度的在线后验密度估计算法. 它提供了一种产生样本的行之有效的办法, 且无需假设状态空间模型或状态分布. 状态空间模型可以是非线性的, 初始状态和噪声分布可以采取任何形式要求. 概率分布用一组粒子的集合来表示, 通过寻找一组在状态空间中传播的随机样本来近似地表示概率密度函数, 获得系统状态的最小方差估计.

我们假定驾驶员头部为刚性物体, 可以通过帧与帧之前的特征点跟踪计算运动参数, 即头部姿态参数. 将人脸单一特征点映射至三维图像空间, 利用粒子滤波的追踪方法估计和学习驾驶员头部姿态, 头部姿态角度由三维空间几何关系计算得到.

与通常在三维空间内进行所有可能的全局校准不同, 我们在目标平面自动锁定事先选择的已知点,

然后对驾驶员移动头部位置的情况进行已知点的匹配和插值. 目标的识别通过基于动态的重要性采样和重采样过程的算法来实现. 在最初校准的过程中, 我们可以通过标准化的模式平移向量来帮助在新的图像中定位已知的校准点. 因此, 当前帧的头部姿态可以基于这些校准点集合加上系统刚开始进行时所施加的学习算法-粒子滤波算法更加准确的估算出来.

3.1 状态空间模型

在状态空间模型中考虑非线性非高斯的动力学模型, 其中 $\mathbf{X}_{1:t} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_t\}$ 作为感兴趣的不可观测状态随时间演变, 而在连续的时间点上也有部分可观测状态 $\mathbf{Y}_{1:t} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_t\}$. 同时, 假定状态序列为马尔科夫链. 在此情况下, 与文献[16]类似, 状态空间模型可以表示时间序列的过程, 其主要构成为:

$$\begin{aligned} \text{状态转换模型: } \mathbf{X}_t &= F_t(\mathbf{X}_{t-1}, \mathbf{U}_t) \leftrightarrow \overbrace{f_t(\mathbf{X}_t | \mathbf{X}_{t-1})}^{\text{转换密度}}; \\ \text{观测模型: } \mathbf{Y}_t &= G_t(\mathbf{X}_t, \mathbf{V}_t) \leftrightarrow \overbrace{g_t(\mathbf{Y}_t | \mathbf{X}_t)}^{\text{观测密度}}. \end{aligned}$$

我们定义驾驶员头部姿态状态 $\mathbf{X}_t = [\mathbf{A}_t \mathbf{v}_t]^\top$, 其中 $\mathbf{A}_t = (T_x, T_y, T_z, \alpha_t, \beta_t, \gamma_t)$ 是一个六维向量, 由平面坐标的 x, y, z 轴的位移和头部的水平偏离度 α_t 、垂直倾斜度 β_t 、左右旋转度 γ_t 组成. 这里 $\mathbf{v}_t$ 为线速度和角速度组成的向量. 在对头部运动进行追踪时, 我们假定驾驶者的视线在大多数时间专注于一个方向, 这与真实驾驶行为一致, 故我们在初始化时锁定这一视线区域的参数. 当驾驶者的头部有很大偏移时, 我们的模型能计算驾驶者当前的头部中心位置和实现锁定的视线区域, 进而基于事先设定的视线区域标号对当前的视线区域进行最近距离的估计. 我们用文献[7]中提到的混合状态 $\mathbf{X}_t^*$, 其中 $\mathbf{X}_t = (1 - \tau_t)\mathbf{X}_t^{(1)} + \tau_t\mathbf{X}_t^{(2)}$ 来代表当下的状态:

$$\mathbf{X}_t^* = (\mathbf{X}_t, \tau_t) \quad (1)$$

其中 τ_t 为二值随机变量, 取值范围 0 和 1, 且 $\mathbf{X}_t^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{X}_{t-1}^{(1)} + \begin{pmatrix} u_t^{(1)} \\ 0 \end{pmatrix}$ 代表时间点 t 速度几乎为零的状

态, $\mathbf{X}_t^{(2)} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \mathbf{X}_{t-1}^{(2)} + \begin{pmatrix} 0 \\ u_t^{(2)} \end{pmatrix}$ 代表速度为常量, 在这

里我们忽略了对加速度的考虑, 这是因为头部在转动的时候更趋向于常速度连续的转动. 其中, $\mathbf{u}_t$ 是微小的头部瞬间转动. 因此其实现为 $\mathbf{Y}_t = \mathbf{T}_t \mathbf{X}_t + \mathbf{V}_t$, $\mathbf{T}_t$ 是两个空间的转换, $\mathbf{V}_t$ 是 t 时刻的噪声. τ_t 的分布基于物体的转动速度, 当 $\tau_t = 1$ 时满足 $d(\mathbf{Y}_t, \mathbf{Y}_{t-1}) > \epsilon_1$, $d(\mathbf{A}_t, \mathbf{A}_{t-1}) > \epsilon_2$. 其中 ϵ_1, ϵ_2 为转动速度阈值, 即两个

时间点脸部特征的移动超过这个值我们认为头部在转动,否则 $\tau_t=0$.

3.2 粒子滤波算法

接下来我们介绍适用于时间序列模型即状态空间模型的粒子滤波算法.在贝叶斯框架中,依赖后验概率密度进行推理,状态 $\mathbf{X}_t$ 和观测 $\mathbf{Y}_t$ 的联合概率密度为

$$p_{0:T,0:T}(\mathbf{X}_{0:T}, \mathbf{Y}_{0:T}) =$$

$$p_0(\mathbf{X}_0)g(\mathbf{Y}_0|\mathbf{X}_0) \times \prod_{t=1}^T f(\mathbf{X}_t|\mathbf{X}_{t-1})g(\mathbf{Y}_t|\mathbf{X}_t) \quad (2)$$

其中 p_0 是 $\mathbf{X}_0$ 的初始概率密度.因此,依据状态转换模型和观测模型,状态和数据从状态空间模型中随机采样的过程中,样本路径 $(\tilde{\mathbf{x}}_0, \dots, \tilde{\mathbf{x}}_T)$ 取初值 $\tilde{\mathbf{x}}_0 \sim p_0(\mathbf{X}_0)$,其余 $\tilde{\mathbf{x}}_t \sim f_t(\mathbf{x}_t|\mathbf{X}_t=\tilde{\mathbf{x}}_{t-1})$;而对应的 $(\tilde{\mathbf{y}}_0, \dots, \tilde{\mathbf{y}}_T)$ 取初值 $\tilde{\mathbf{y}}_0 \sim g(\mathbf{Y}_0|\mathbf{X}_0)$,其余的 $\tilde{\mathbf{y}}_t \sim g_t(\mathbf{y}_t|\mathbf{X}_t=\tilde{\mathbf{x}}_t)$.

在粒子滤波算法中,对转换密度和观测密度函数的分布估计是至关重要的.事实上对于当前状态分布 $p(\mathbf{X})$ 我们不知道其准确的趋势,故用任一经标准化的重要性分布 $q(\mathbf{X})$ 替代,然后根据前一时刻的状态和观测值对当前的样本数据赋予权重.

对第 i 个样本权重 $w_t^{(i)}$ 有 $w_t^{(i)} = \frac{p_t(\mathbf{x}_t^{(i)})}{q_t(\mathbf{x}_t^{(i)})}$. 由

$$p(\mathbf{X}_t=\mathbf{x}_t|\mathbf{Y}_{1:t}=\mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{y}_{1:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{1:t-1})}, \text{ 设}$$

$$\sigma_{t|0:t-1} = \int_{\mathbb{R}^n} g_t(\mathbf{Y}_t=\mathbf{y}_t|\mathbf{x}) p_{t|0:t-1}(\mathbf{X}_t=\mathbf{x}|\mathbf{y}_{0:t-1}) d\mathbf{x},$$

则有

$$p_{0:t|0:t}(\mathbf{X}_{0:t}=\mathbf{x}_{0:t}|\mathbf{Y}_{0:t}=\mathbf{y}_{0:t}) = \frac{g_t(\mathbf{x}_t|\mathbf{Y}_t)p_{0:t|0:t-1}(\mathbf{x}_{0:t}|\mathbf{y}_{0:t})}{\sigma_{t|0:t-1}(\mathbf{y}_t|\mathbf{y}_{0:t-1})} \quad (3)$$

由于我们基本不可能依据密度函数 $p_{0:t|0:t}(\mathbf{X}_{0:t}|\mathbf{Y}_{0:t})$ 进行取样,故我们根据概率密度 $q_{(0:t)}(\mathbf{X}_{0:t}|\mathbf{Y}_{0:t})$ 选取 N 个样本 $\tilde{\mathbf{x}}_{0:t}^{(i)}, i=1,2,\dots,N$,重要性权重为

$$\tilde{w}_t^{(i)} = \frac{p_{0:t|0:t}(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{0:t})}{q_{0:t}(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{0:t})}, i=1,\dots,N \quad (4)$$

这些权重经过标准化被映射到 $[0,1]$ 区间,其中

$$w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}}. \text{ 基于下面的公式}$$

$$q_{0:t}(\mathbf{X}_{0:t}=\mathbf{x}_{0:t}|\mathbf{Y}_{0:t}=\mathbf{y}_{0:t}) = \underbrace{q_{0:t-1}(\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t-1})}_{\text{存在路径}} \underbrace{q_t(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)}_{\text{扩展路径}} \quad (5)$$

我们找到了 $w_t^{(i)}$ 和 $w_{t-1}^{(i)}$ 的关系,由公式表示得

$$\tilde{w}_t^{(i)} = \frac{p_{0:t|0:t}(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{0:t})}{q_{0:t}(\tilde{\mathbf{x}}_{0:t}^{(i)}|\mathbf{y}_{0:t})}$$

$$= \tilde{w}_{t-1}^{(i)} \frac{f_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})g_t(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})\sigma_{t|0:t-1}(\mathbf{y}_t|\mathbf{y}_{0:t-1})}$$

$$= c w_{t-1}^{(i)} \frac{f_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)})g_t(\mathbf{y}_t|\tilde{\mathbf{x}}_t^{(i)})}{q_t(\tilde{\mathbf{x}}_t^{(i)}|\tilde{\mathbf{x}}_{t-1}^{(i)}, \mathbf{y}_t)\sigma_{t|0:t-1}(\mathbf{y}_t|\mathbf{y}_{0:t-1})} \quad (6)$$

其中 c 为常数, $i=1,2,\dots,N$. 值 $\sigma_{t|0:t-1}(\mathbf{y}_t|\mathbf{y}_{0:t-1})$ 由于会最终进行标准化所以在计算中不起作用.由于时间过长后权重的作用会渐渐失效,故我们要在每一步计算完权重后进行重采样.

粒子滤波方法用于原则上提升时间点 t 的新状态的概率分布,我们在标准重采样过程上进行了改进,使用辅助采样的方法使得在下一时刻有利于粒子的存活.

首先,假设在时间点 $t-1$ 联合后验概率函数可

以很好的被 $\hat{p}_{0:t-1|0:t-1}(d\mathbf{x}_{0:t}) = \sum_{i=1}^N w_t^{(i)} \delta_{\mathbf{x}_{0:t}^{(i)}}(d\mathbf{x}_{0:t})$ 一致的近似,其中 $\delta_{\mathbf{x}_{0:t}^{(i)}}$ 是 $\mathbf{x}_{0:t}^{(i)}$ 点的狄拉克测度.这时,时间点 t 的联合后验概率可以近似为

$$p_{0:t|0:t}(d\mathbf{x}_{0:t}|\mathbf{y}_{0:t}) \approx \frac{1}{Z} \sum_{i=1}^N w_{t-1}^{(i)} \delta_{\mathbf{x}_{0:t-1}^{(i)}}(d\mathbf{x}_{0:t-1}) g(\mathbf{y}_t|\mathbf{x}_t) f(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}) d\mathbf{x}_t \quad (7)$$

其中归一化因子 Z 定义为

$$Z = \sum_{j=1}^N w_{t-1}^{(j)} \int f(\mathbf{x}|\mathbf{x}_{t-1}^{(j)}) g(\mathbf{y}_t|\mathbf{x}) d\mathbf{x}.$$

因此新粒子 $\mathbf{x}_{0:t}^{(i)}$ 的整体路径的联合分布是

$$q_t(d\mathbf{x}_{0:t}) = q_{0:t-1}(d\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t}) q_t(d\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{y}_t)$$

$$= \left(\sum_{i=1}^N v_{t-1}^{(i)} \delta_{\mathbf{x}_{0:t-1}^{(i)}}(d\mathbf{x}_{0:t-1}) \right) \times (q_t(d\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}, \mathbf{y}_t)) \quad (8)$$

其中 $\sum_{i=1}^N v_{t-1}^{(i)} = 1$ 并且 $v_{t-1}^{(i)} > 0$. 公式由两部分组成,过程 $\mathbf{x}_{0:t-1}$ 过去路径的边缘分布 $q_{0:t-1}$ 和新状态 $\mathbf{x}_t$ 的条件分布 q_t . 第一部分在老粒子路径 $\{\mathbf{x}_{0:t-1}^{(i)}\}$ 中心离散分布.权重函数 $v_{t-1}^{(i)}$ 依赖于数据,对于新数据点 $\mathbf{y}_t$ 可以预先选择适应性好的粒子.

取状态某一点的值 $\boldsymbol{\mu}_t^{(i)}$ 作为 $f(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)})$ 的平均,计算权值函数作为该点的似然函数, $v_{t-1}^{(i)} = g(\mathbf{y}_t|\boldsymbol{\mu}_t^{(i)})$,如果时间点 $t-1$ 的粒子被权值化,则 $v_{t-1}^{(i)} = w_{t-1}^{(i)} g(\mathbf{y}_t|\boldsymbol{\mu}_t^{(i)})$. 因为,理想情况下对于时间点 t 的滤波,从边缘条件分布 $p_{0:t-1|0:t}$ 可以得到过去的路径 $\mathbf{x}_{0:t-1}$,使用时间点 $t-1$ 的粒子近似可以得到

$$p_{0:t-1|0:t}(d\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t})$$

$$\propto \int p_{0:t-1|0:t-1}(d\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t-1}) f(\mathbf{x}_t|\mathbf{x}_{t-1}) g(\mathbf{y}_t|\mathbf{x}_t) d\mathbf{x}_t$$

$$\approx \sum_{i=1}^N w_{t-1}^{(i)} \delta_{\mathbf{x}_{0:t-1}^{(i)}}(d\mathbf{x}_{0:t-1}) \int f(\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)}) g(\mathbf{y}_t|\mathbf{x}_t) d\mathbf{x}_t \quad (9)$$

在蒙特卡罗方法中,使用粗略的近似函数来处理,令 $f(\mathrm{d}\mathbf{x}_t|\mathbf{x}_{t-1}^{(i)})\approx\delta_{\mu_t^{(i)}}(\mathrm{d}\mathbf{x}_t)$,则有

$$p_{0:t-1|0:t}(\mathrm{d}\mathbf{x}_{0:t-1}|\mathbf{y}_{0:t})\approx$$
$$\sum_{i=1}^Ng(\mathbf{y}_t|\boldsymbol{\mu}_t^{(i)})w_{t-1}^{(i)}\delta_{\mathbf{x}_{0:t-1}^{(i)}}(\mathrm{d}\mathbf{x}_{0:t-1}) \quad (10)$$

我们使用上述机制定义粒子的广义重要性比率,即

$$\tilde{w}_t^{(i)}=\frac{\tilde{w}_{t-1}^{(i)}}{v_{t-1}^{(i)}}\times\frac{g(\mathbf{y}_t|\mathbf{x}_t^{(i)})f(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)})}{q_t(\mathbf{x}_t^{(i)}|\mathbf{x}_{t-1}^{(i)},\mathbf{y}_t)}$$
 (11)

与标准的序贯重要性采样相比,我们的采样方法在计算偏置时修正重要权值等于 $1/v_{t-1}^{(i)}$,同时第一层权值为比率 $w_{t-1}^{(i)}/v_{t-1}^{(i)}$. 这样在采样之前的重采样过程中,可以利用原来时刻预测的似然度大的粒子扩展到该时刻,增加了粒子的多样性的同时减少了重要性权值的方差,可以产生更精确的估计.

我们能够基于这个状态模型和观测密度函数 $g_t(\mathbf{Y}_t|\mathbf{X}_t)\propto(1/B)\exp(-\lambda d(\mathbf{Y}_t,\mathbf{T}_\in\mathbf{X}_t+\mathbf{V}_t))$ 来估计转移密度函数 $f_t(\mathbf{X}_t|\mathbf{X}_{t-1})$,其中 B 为标准化常数. 利用粒子滤波方法估计当前的状态 $\mathbf{X}_t$,其由样本点的加权平均值 $\hat{\mathbf{X}}_t=\sum_{n=1}^Nw_t^{(n)}\bar{\mathbf{x}}_t^{(n)}$ 得到.

粒子滤波方法作为一种基于贝叶斯估计的非线性滤波算法,在处理非高斯非线性时变系统的参数估计和状态滤波问题方面有独到的优势. 由上述方法可知,在系统运行之前,设定所有要求的参数,则系统可以通过粒子滤波对头部姿态的追踪和估计,实现视线区域的自动校准,完全不需要驾驶人特殊的合作.

4 实验结果

4.1 实验环境及数据来源

在系统初始化前,我们用驾驶员头部运动传感器实时记录头部姿态的水平偏离度、垂直倾斜度和左右旋转度,并与设定好的视线区域编号进行对应,确定相应的头部姿态角度和区域编号的映射. 具体的注视区域划分详见 4.2 节. 然后使用摄像头实时记录驾驶员的图像,对每帧图像进行人脸检测和追踪,提取感兴趣的特征点,并估计驾驶员的头部姿态角度. 为保证头部姿态角度的正确性,会使用部分数据初始化头部姿态零点. 最后,施加样本集学习和粒子滤波算法对头部姿态更新,根据置信概率确定状态可信性,并输出当前的注视频率区域,完成对注视频率的校准.

为了验证本文的自动校准算法效果,我们采集

了真实环境的图像数据进行测试实验. 其中,实验环境为 Intel Core i5 2.67 GHz CPU 和 4 GB RAM 的电脑. 图像数据来源于同一个 USB 摄像头,图像分辨率为 640×480 像素. 摄像头固定在靠近方向盘的仪表盘中心,距驾驶员有一定角度. 我们在捕捉驾驶员脸部特征的摄像头里加装了能够同时适应白天和黑夜的红外光源及光照转换装置,增强实验装置抗弱光线条件的鲁棒性. 驾驶者测试时头戴 Trivisio Colibri 惯性位置追踪器,以接收头部姿态的实际数据.

我们随机选取 8 个个体,共收集 16 万帧的连续图像数据,来测试我们提出的算法适应不同个体的驾驶习惯的能力. 其中,部分被试驾驶员情况信息如表 1 所示. 由于不同的驾驶员的视觉搜索模式(注视持续时间、注视点分布、注视频角等)具有差异性,因此表中列出驾驶员的性别、年龄、身高、驾龄等情况,并在之后的部分对驾驶员的误差进行了对比. 我们将驾驶员划分为熟练驾驶员和非熟练驾驶员:其中熟练驾驶员为驾龄 5 年及以上的驾驶员,即表 1 中前 4 位驾驶员;而非熟练驾驶员则为驾龄 5 年以下的驾驶员,即表 1 中后 4 位驾驶员.

表 1 被试驾驶员情况信息

编号	性别	年龄	身高	驾龄
001	48	女	160	7 年
002	32	男	176	5 年
003	42	男	186	9 年
004	60	男	170	16 年
005	31	女	172	0.5 年
006	25	女	163	1 年
007	40	女	156	2 年
008	23	男	167	0.3 年

4.2 视线区域的表示

在一般情况下,驾驶员以头部中心为轴,转向其注视的物体. 因此,驾驶者头部姿态决定了视线区域. 我们根据驾驶环境及注视习惯划分了不同的视线区域. 本文所划分的 9 个视线区域参照了文献[11]中所描述的方法. 文献[11]中的部分视线区域在正常行驶过程中注视概率几乎为零,并且不同驾驶员朝向后视镜的注视概率并不稳定,因此我们使用 K-mean 聚类的方法将原有 12 个视线区域优化为 9 个视线区域.

如图 4 所示,视线区域的划分是基于真实的驾驶环境. 由于朝向后视镜头部特征几乎无法提取,我们采用带有倒车影像的车辆规避倒车时头部转向. 因此中央控制台区域的注视频率实际也包含了朝向后视镜的注视. 在朝向风挡玻璃即区域 3、区域 5 及

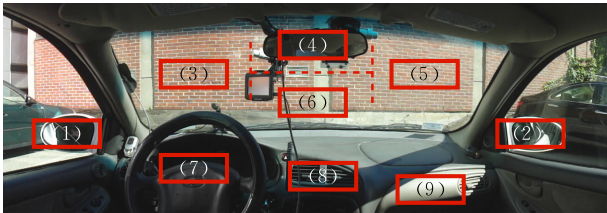


图 4 用于自动校准的视线区域的划分

区域 6 时,驾驶员通常会结合头部转动和眼球运动,但只对同一视野的近距离、中距离和远距离的车前状况,有较多的眼球运动观察.

4.3 视线区域的估计

视线区域的估计步骤需在平面上水平偏离度、垂直倾斜度和左右旋转度零点位置已知的情况下进行初始化.通常我们认为直视视线区域 6 的中心位置为零点.在此实验中,我们用 500 帧的图像初始化零点,进而认为其欧拉角为零(水平偏离度、垂直倾斜度和左右旋转度均为零). $R(\theta)=(\alpha,\beta,\gamma)$ 表示在 XYZ 坐标系下的欧拉角,其中 α 为水平偏离度、 β 为垂直倾斜度、 γ 为左右旋转度,如图 5 所示.这里使用单目摄像头的深度估计建立鼻子顶点 P 到坐标轴中心点 O 的位移和偏移欧拉角.其中坐标轴中心点即头部中心点.

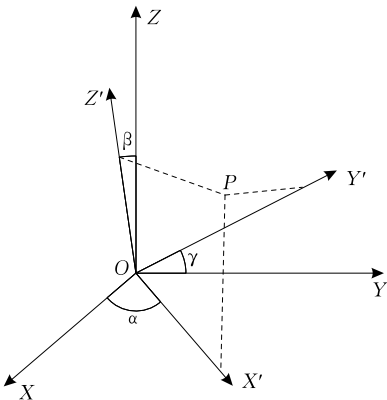


图 5 欧拉角示意

之后,我们对图像序列进行训练以寻找基本帧,进而在形成样本集之后我们能用样本集的元素来估计当前驾驶者所处的状态,进而求出其视线区域.当学习算法对 50 000 帧的图像经过两轮学习过后,我们得到表 2,即每个视线区域中心的水平偏离度、垂直倾斜度与左右旋转度.当追踪器得到每一帧图像后,算法会针对当前状态在样本空间生成概率分布,最后概率值最大的样本集指标为驾驶者当前视线区域的标号.

对于同一个图像提取装置和同一段图像序列,由于我们采取了随机结构用于驾驶者脸部的识别,

表 2 视线区域中心对应的头部姿态

视线区域	临近视线区域	头部姿态
		(水平偏离度,垂直倾斜度,左右旋转度)
1	3,7	(32.131,0.319,-0.352)
2	5,9	(-17.462,1.543,2.615)
3	1,4,6,7	(23.342,1.466,-0.208)
4	3,5,6	(-6.014,-5.037,0.103)
5	2,4,6,8,9	(-11.535,1.463,2.234)
6	3,4,5,7,8	(0.000,0.000,0.000)
7	1,3,6,8	(25.653,-1.578,-0.217)
8	5,6,7,9	(-9.426,-2.445,0.128)
9	2,5,8	(-14.095,2.973,2.301)

故我们重复实验所得到的数据会有区别于表 2,但不影响对视线区域的估计结果.

我们对全部图像帧进行水平偏离度、垂直倾斜度和左右旋转度的计算后发现,其平均值最后稳定到 0.024、-0.005 和 0.013.这表示驾驶者大部分时间都注视前方,若与此结果相悖,则很有可能在这段时间驾驶者处于比较疲劳的阶段,则能够通过此结果对驾驶员进行警示或者采取刹车处理.

在真实驾驶环境中,部分驾驶员的置信概率平均值曲线如图 6 所示.基于当前状态的置信评估来判断头部姿态估计的准确性,初期置信概率较不稳定,随着帧数的增加,置信概率的平均值逐步上升,这说明对头部姿态的估计数值逐渐更为准确.以 001 驾驶员为例,在真实驾驶环境中,我们得到至 9000 帧时,置信概率的平均值为 0.6743,标准差为 0.0963,基本趋于稳定.而其他驾驶员的置信概率也分别稳定至 0.6682 和 0.7052 之前.

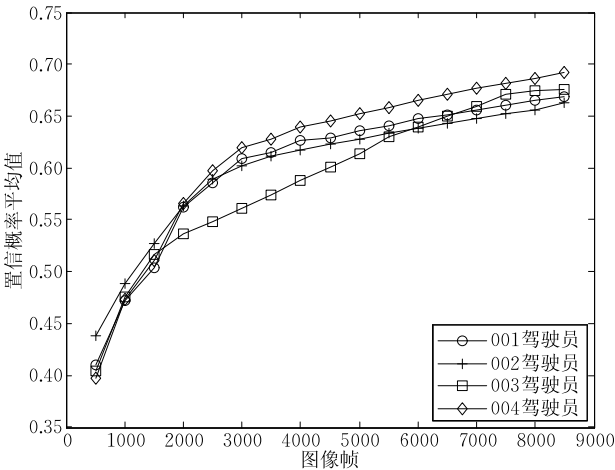


图 6 头部姿态估计的置信概率平均值曲线

除此之外,式(1)中的 τ_i 需进行准确取值,在我们的实验中其当前值的确定依赖于之前一个时刻头部摆动速度的值, $\tau_i=1$ 必有 v_i 超出了我们事先设定

的阈值. 经过对初始化前传感器的样本进行统计我们得到平均角速度大约为 1.5° , 故我们取 1.4° 为阈值, 这是因为在统计过程中计算的方差较大, 取这个值可以大致符合实际情况. 此阈值会在对头部姿态判断的过程中进行不断修改使得其相对准确.

4.4 视线区域注视的频率统计

我们对所有数据帧进行视线区域注视概率的统计得到了表 3, 其中对于熟练驾驶员和非熟练驾驶员分别进行了统计.

表 3 视线区域注视频率统计

视线区域	熟练驾驶员 平均注视频率	非熟练驾驶员 平均注视频率
1	0.0425	0.0575
2	0.0467	0.0614
3	0.0455	0.0273
4	0.0501	0.0609
5	0.0396	0.0253
6	0.6733	0.6654
7	0.0034	0.0056
8	0.0910	0.0921
9	0.0079	0.0045

由表 3 可知, 区域 6 即正前方区域最频繁被使用. 除了区域 6 以外, 区域 1、区域 2 及区域 4 也经常是驾驶员注视的目标, 即左视镜, 右视镜和后视镜, 这符合大多数人的开车习惯. 非熟练驾驶员比熟练驾驶员更注意区域 1、区域 2 及区域 4, 即左视镜, 右视镜和后视镜. 而熟练驾驶员对前方的区域注视频率比非熟练驾驶员更为平均一些.

4.5 与标准数据进行对比

我们对每个被试驾驾驶员所收集的数据帧中选取帧 500 到帧 7500 进行测试, 并对比我们的方法估计的头部姿态和头部运动传感器的真实头部姿态, 这时估计值是乘以扩张系数的最终值. 我们最终得到的水平偏离度、垂直倾斜度和左右旋转度的误差曲线如图 7 和图 8 所示. 由于有预先的视线区域的划分, 在本文中我们设定水平偏移角和垂直倾斜角的容忍误差可接收率由划分的视线区域块的最小块决定, 其中水平偏离度 2° , 垂直倾斜度和左右旋转度

均为 1.5° . 由图 7 和图 8 所示, 水平偏移角和垂直倾斜角的平均绝对误差曲线是收敛的, 并且极限值均在 1.5° 左右, 数值能够保证大部分时间内视线区域的准确性. 当然对于倾斜度而言, 估计误差稍大. 这些数据保证了在长时间驾驶中此自动校准算法能够对驾驶人的视线区域有相当准确的判断.

综合图 7、图 8 和表 4 结果可知, 熟练驾驶员的绝对平均误差曲线收敛速度略快于非熟练驾驶员.

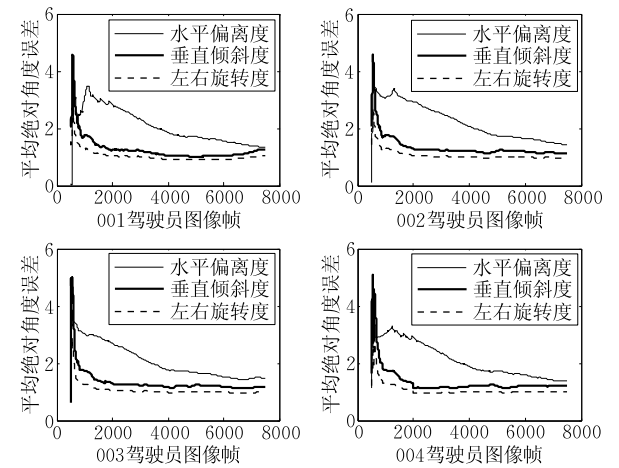


图 7 熟练驾驶员头部姿态的帧 500 到帧 7500 平均绝对误差曲线

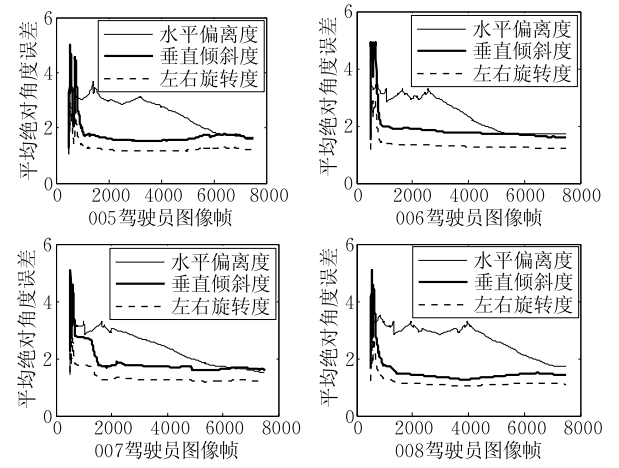


图 8 非熟练驾驶员头部姿态的帧 500 到帧 7500 平均绝对误差曲线

表 4 绝对误差的统计特性

编号	水平偏离度(Yaw)			垂直倾斜度(Pitch)			左右旋转度(Roll)		
	平均绝对 误差/($^{\circ}$)	绝对误差 方差	可接收率/%	平均绝对 误差/($^{\circ}$)	绝对误差 方差	可接收率/%	平均绝对 误差/($^{\circ}$)	绝对误差 方差	可接收率/%
001	1.326	2.585	97.94	1.244	1.262	99.33	1.022	1.112	99.77
002	1.418	2.504	97.58	1.132	1.134	99.75	0.966	1.136	99.76
003	1.470	2.639	97.86	1.183	1.212	99.66	0.992	1.167	99.86
004	1.374	2.521	97.35	1.213	1.236	99.43	1.006	1.171	99.75
005	1.634	2.954	97.25	1.604	1.398	97.85	1.202	1.102	99.85
006	1.708	2.989	97.06	1.610	1.323	98.04	1.205	1.105	99.74
007	1.498	2.750	97.33	1.619	1.401	97.78	1.209	1.173	99.74
008	1.707	2.978	97.13	1.408	1.365	98.99	1.104	1.123	99.76

而非熟练驾驶员的视线区域的可接收率略小于熟练驾驶员. 但总体而言, 所有被试驾驶员的平均绝对误差相差并不大, 并且误差曲线均收敛, 同时绝对误差方差较小, 而且具有较高的可接收率. 因此, 该结果表明, 本文算法受不同个体驾驶习惯影响较小.

4.6 自动校准算法的应用

我们的方法无需驾驶员主动配合, 便可自动校准驾驶员的注视视线区域, 并稳定的输出驾驶员的头部姿态角度和当前注视. 由结果可知, 基于头部姿态的注视视线估计的误差低, 可用性高. 同时, 关联头部姿态和注视视线区域的自动校准方法受不同驾驶员及驾驶习惯的影响小.

通过追踪和估计驾驶员的注视视线及视线区域, 可以分析驾驶员的注视持续时间、注视点分布、注重视角等特性, 实时掌握驾驶员的状况, 感知驾驶员对路况及交通信息的注意程度, 以便搭建新型的车联网智能交通安全驾驶系统.

5 结 论

本文提出了一个在真实驾驶环境中对驾驶员视线方向进行自动校准和估计的算法. 系统的运行基于嵌入的样本集学习算法和粒子滤波算法. 两个算法融合在一起可以估计驾驶员头部运动, 也可以在其他方面有广泛应用.

这种算法的优势在于其能够使系统进行自动估计而不需要驾驶员的合作, 系统运行之前会有初始化的过程, 进行标准数据预设参数和视线区域的划分. 在算法执行之前我们加入脸部识别算法以确定头部姿态零点, 利用稳定的人脸特征提取方法提取特征. 之后, 学习算法在图像中对不同的视线区域选取基本帧构成不同的集合类, 即视线区域的商集, 进而对粒子滤波算法提供基本的样本集, 经过粒子滤波算法对当前驾驶员的状态进行取样, 计算加权平均值, 我们得到估计的水平偏离度、垂直倾斜度和左右旋转度, 进而构造其在样本集空间的条件概率分布, 选取概率最大的样本指标作为视线区域的编号. 本文算法能够在连续时间内自动地对驾驶员的视线区域进行估计, 因此, 此算法具有自适应性, 可以嵌入到硬件设备中, 自动对驾驶员视线校准并对注视区域进行准确估计.

参 考 文 献

- [1] Zhang Chuang, Chi Jian-Nan, Zhang Zhao-Hui, Wang Zhi-Liang. A novel eye gaze tracking technique based on pupil center cornea reflection technique. *Chinese Journal of Computers*, 2010, 33(7): 1272-1282(in Chinese)
(张闯, 迟健男, 张朝晖, 王志良. 一种新的基于瞳孔-角膜反射技术的视线追踪方法. *计算机学报*, 2010, 33(7): 1272-1282)
- [2] Ebisawa Y, Fukumoto K. Head-Free, remote eye-gaze detection system based on pupil-corneal reflection method with easy calibration using two stereo-calibrated video cameras. *IEEE Transactions on Biomedical Engineering*, 2013, 60(10): 2952-2960
- [3] Lee S J, et al. Real-time gaze estimator based on driver's head orientation for forward collision warning system. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12(1): 254-267
- [4] Yang H, et al. Recent advances and trends in visual tracking: A review. *Neurocomputing*, 2011, 74(18): 3823-3831
- [5] Hansen D W, Ji Q. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(3): 478-500
- [6] Murphy-Chutorian E, Trivedi M M. Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(4): 607-626
- [7] Murphy-Chutorian E, Trivedi M M. Head pose estimation and augmented reality tracking: An integrated system and evaluation for monitoring driver awareness. *IEEE Transactions on Intelligent Transportation Systems*, 2010, 11(2): 300-311
- [8] Tawari A, Martin S, Trivedi M M. Continuous head movement estimator for driver assistance: Issues, algorithms, and on-road evaluations. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(2): 818-830
- [9] Yuan Wei, Fu Rui, Guo Ying-Shi, Xue Chuan-Liang, Ma Yong. Driver's visual transition characteristics based on the Markov chain. *Journal of Chang'an University (Natural Science Edition)*, 2012, 32(6): 88-93(in Chinese)
(袁伟, 付锐, 郭应时, 薛传亮, 马勇. 基于马尔可夫链的驾驶人视觉转移特征. *长安大学学报(自然科学版)*, 2012, 32(6): 88-93)
- [10] Dong Y, Hu Z, Uchimura K, et al. Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems*, 2011, 12(2): 596-614
- [11] Fu Xianping, Guan Xiao, Peli Eli, et al. Automatic calibration method for driver's head orientation in natural driving environment. *IEEE Transactions on Intelligent Transportation Systems*, 2013, 14(1): 303-312
- [12] Zhang Kaihua, Zhang Lei, Yang Ming-Hsuan. Real-time compressive tracking//*Proceedings of the 12th European Conference on Computer Vision*. Firenze, Italy, 2012: 864-877
- [13] Viola P, Jones M. Robust real-time face detection. *International Journal of Computer Vision*, 2004, 57: 137-154
- [14] Zhu X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild//*Proceedings of the 25th IEEE Conference on Computer Vision and Pattern Recognition*. Providence, USA, 2012: 2879-2886

[15] Toyama K, Blake A. Probabilistic tracking in a metric space//Proceedings of the 8th International Conference on Computer Vision. Vancouver, Canada, 2001, 2: 50-59

[16] Cappe O, Godsill S J, Moulines E. An overview of existing methods and recent advances in sequential Monte Carlo. Proceedings of the IEEE, 2007, 95(5): 899-924



FU Xian-Ping, born in 1971, Ph.D., professor. His research interests include perception of natural scenes in engineering systems, multimedia technology and image/video processing.

WANG Ya-Fei, born in 1988, Ph.D. candidate. His research interests include gaze estimation and machine learning.

YUAN Guo-Liang, born in 1982, Ph.D. candidate. His research interests include multimedia technology and image/video processing.

GUAN Xiao, born in 1989, Ph.D. candidate. His research interest is symbolic computation.

PELI Eli, born in 1951, Ph.D., professor. His research interests include image processing and image understanding.

LUO Gang, born in 1968, Ph.D., assistant professor. His research interest is vision science.

Background

Driver’s gaze estimation plays a significant role in driver intelligent assist system and vehicle active safety system. On account of not need to install other devices on user’s body, non-intrusive video-based gaze estimation method is becoming the main trend. But traditional methods like infrared based method, appearance based method and shape based method meet great challenges in natural driving environment.

Head pose can be used to express sight and focus of driver’s attention. Most recent systems demonstrated a hybrid approach to tracking and estimate the head pose in the study of driver’s gaze and attention. But there are still many problems when in the face of dramatically illumination changes and large head rotation.

Our system assumes the driver’s head pose maintains a certain degree of stability when looking at the fixed component inside the car and the eye gaze point doesn’t fall outside

the field-of-view in defined head pose. Thus, the head pose corresponding to the given gaze zone can be calculated by the rotation angle of yaw, pitch and roll. According to driving habits and statistics characteristics of head pose, estimation and learning algorithms can be implemented, so the system can automatically calibrate.

Our group has been focusing on the research of driver’s gaze estimation on Safety Driving Assist System for a long time. Many outstanding works have been published in world-wide conferences and transactions. The research of the paper is supported by the National Natural Science Foundation of China under Grant Nos. 61272368, 61370142, and Excellent Talent Supporting Project of Higher Education Institution in Liaoning Province under Grant No. LR2013021, and Applying Basic Research program of ministry of Transport under Grant No. 015329225300.