

DOI: 10.11992/tis.201801016

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180628.1622.004.html>

# 基于车内外视觉信息的行人碰撞预警方法

杨会成, 朱文博, 童英

(安徽工程大学 电气工程学院, 安徽 芜湖 241000)

**摘 要:** 行人碰撞预警系统通常依据行人检测与碰撞时间判断的方式为驾驶员提供预警信息。为了提供更加可靠的危险判断依据, 本文提出一种同时分析道路状况与驾驶员头部姿态的行人碰撞预警方法, 用两个单目相机分别获取车辆内外环境图像。通道特征检测器用于定位行人, 根据单目视觉距离测量方法估计出行人与自车间的纵向与横向距离。多任务级联卷积网络用于定位驾驶员面部特征点, 通过求解多点透视问题获取头部方向角以反映驾驶员注意状态。结合行人位置信息与驾驶员状态信息, 本文构建模糊推理系统判断碰撞风险等级。在实际路况下的实验结果表明, 根据模糊系统输出的风险等级可以为预防碰撞提供有效的指导。

**关键词:** 碰撞预警; 内外信息; 行人定位; 驾驶员状态; 单目视觉; 通道特征; 多任务级联卷积网络; 模糊推理系统  
**中图分类号:** TP181   **文献标志码:** A   **文章编号:** 1673-4785(2019)04-0752-09

中文引用格式: 杨会成, 朱文博, 童英. 基于车内外视觉信息的行人碰撞预警方法 [J]. 智能系统学报, 2019, 14(4): 752-760.

英文引用格式: YANG Huicheng, ZHU Wenbo, TONG Ying. Pedestrian collision warning system based on looking-in and looking-out visual information analysis[J]. CAAI transactions on intelligent systems, 2019, 14(4): 752-760.

## Pedestrian collision warning system based on looking-in and looking-out visual information analysis

YANG Huicheng, ZHU Wenbo, TONG Ying

(College of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China)

**Abstract:** Pedestrian collision warning systems usually provide early warning for drivers based on the technologies of pedestrian detection and collision time measurement. To provide a more reliable basis for risk assessment, a pedestrian collision warning method that involves analyzing the road condition and driver's head pose simultaneously is proposed in this paper. Two monocular cameras are used to capture vehicle exterior and interior images, and a channel features detector is applied to locate pedestrians. The vertical and horizontal distances between pedestrians and ego-vehicle are estimated based on monocular vision distance measurement. The multi-task cascaded convolutional network is utilized for facial landmark detection. By solving a perspective-n-point (PnP) problem, the estimated head angles can reflect driver's attention states. By combining both pedestrian location information and driver's attention information, we implemented a fuzzy inference system to assess collision risk level. An experiment in real-world driving conditions demonstrated that the risk levels obtained from the fuzzy system are reliable and can provide guidance for collision avoidance.

**Keywords:** collision warning; internal and external information; pedestrian positioning; driver states; monocular vision; channel features; multi-task cascaded convolutional network; fuzzy inference system

高级驾驶辅助系统 (advanced driver assistance system, ADAS) 是目前车辆安全领域的研究热点, 该系统通过车载传感器收集并分析车内外环境数据, 为驾驶员提供辅助信息并对危险进行提醒。

碰撞预警是 ADAS 的一个重要功能, 碰撞预警系统通常根据传感器获取的前方障碍物相对距离和速度计算碰撞时间 (time to collision, TTC), 可靠的预警一般设置为潜在的碰撞前约 2 s 的时间。目前应用的碰撞保护系统大多针对车辆间的碰撞, 然而在人车碰撞事故中, 没有保护装置的行人更

收稿日期: 2018-01-08. 网络出版日期: 2018-06-29.

基金项目: 安徽省高校自然科学研究重点项目 (KJ2018A0122).

通信作者: 朱文博. E-mail: [vembozhu@163.com](mailto:vembozhu@163.com).

容易受到致命的伤害。统计数据表明,在我国城市道路中,每年死于交通事故的人口中有接近50%为身处车外的行人。因此研究行人碰撞预警系统,在可能的碰撞发生前对驾驶员做出提醒,具有重要的现实意义。

行人碰撞预警的实现广泛采用视觉传感器获取行人位置信息,根据相对距离和车速做出判断。由于姿态与外形的多样性,行人检测一直是计算机视觉领域的一个重要研究课题。行人检测的经典方法有 Viola 等<sup>[1]</sup>提出的 Haar-like 特征和级联分类器, Dalal 等<sup>[2]</sup>提出的 HOG 特征结合 SVM 分类器,以及 Felzenszwalb 等<sup>[3]</sup>提出的可变形部件模型 (deformable part model, DPM)。卷积神经网络 (convolutional neural networks, CNN) 由于其强大的特征学习和表达能力,在计算机视觉领域得到了广泛的应用。深度学习目标检测方法著名的有 Girshick 等<sup>[4-5]</sup>提出的基于区域建议的卷积神经网络模型 (region CNN, R-CNN), 通过 CNN 完成特征提取与目标定位。

根据美国高速交通安全管理局 (NHTSA) 的统计<sup>[6]</sup>,有超过 80% 的交通事故责任与驾驶员相关,其中注意力不集中是引发事故的主要原因。此外,预警系统的有效性受到驾驶员反应时间的影响,驾驶员注意力的差异会导致对危险的反应时间不同,精确的数值判断有时反而会给出错误的预警。近年来,研究人员利用内外视觉 (looking-in and looking-out, LiLo) 框架<sup>[16]</sup>实现车辆安全系统性能的提升,在检测车辆前方障碍物的基础上,增加了对驾驶员状态的监测。该框架利用传感器同时获取车辆周围环境、车辆运动状态以及驾驶员行为,系统同时理解车辆外部和内部信息并据此对潜在危险做出判断。在这一框架基础下的部分工作<sup>[17-18]</sup>特别研究了驾驶员对于道路行人的认知。

驾驶员状态的监控对于提高行车安全有着重要的促进作用,驾驶员的注意力与视线方向及头部姿势有着内在的联系,因此眼睛或头部的状态检测技术广泛用于注意分散监测。在实际驾驶条件下,使用非接触式方式进行可靠的视线方向估计非常具有挑战性。就危险预警而言,头部偏移情况下视线的分散程度相对更加严重,因此头部姿态也可以作为驾驶员注意状态判断的可靠指标<sup>[7]</sup>。在实际应用中,头部姿态的获取普遍采用非接触式的基于视觉的方法,通常需要先建立面部形状模型以定位面部特征点,根据特征点间的几何关系还原空间姿态。面部特征点定位的经典方法是 Cootes 等<sup>[8]</sup>提出的主动形状模型 (active shape model, ASM), 近年来广泛地使用有基于回

归森林的方法<sup>[9]</sup>和基于深度学习的方法<sup>[10-11]</sup>。

车辆的行驶状态很大程度上取决于驾驶员是否注意到行人,通过监控驾驶员的注意状态,可以更加准确地衡量碰撞危险。结合对经验丰富的驾驶员驾车行为的研究,本文提出一种基于车内外视觉信息的行人碰撞预警方法,将行人状态分析和驾驶员状态分析同时加入模糊决策过程。本文的方法使用两个单目相机分别获取车外行人方位信息和车内驾驶员头部姿态信息,其中外部信息处理使用通道特征检测器定位行人,根据单目视觉测距原理估计距离与角度,内部信息处理使用级联回归方法定位面部特征点,根据 EPnP 算法估计驾驶员头部姿态,利用模糊推理系统融合内外信息对碰撞风险等级进行评估,可以满足系统决策的准确性和时效性。

## 1 行人位置估计

利用通道特征检测器,可以从车辆前方的道路监控视频中判断每一帧是否存在行人,并提取出行人在图像中的位置,进一步恢复出其在真实场景中的相对距离与角度,作为碰撞风险判断的依据。

聚合通道特征 (aggregated channel features, ACF) 检测器最早由 Dollar 等<sup>[12]</sup>提出,其结构如图 1 所示。检测器的特征构造利用色彩和梯度信息,选取 LUV 色彩空间 3 个通道、6 个方向上的梯度直方图通道和 1 个梯度幅值通道共 10 个通道图像结合而成。对每个通道图像上互不重叠的  $2 \times 2$  的块内像素求和,通道的输出为输入通道每个图像块内像素聚合构成的特征池。分类器由 4 096 棵深度为 4 的决策树通过软级联 (soft cascade) 方式<sup>[13]</sup>构造而成,分类器的训练采用 RealBoost 算法在 Caltech 行人数据集上进行。

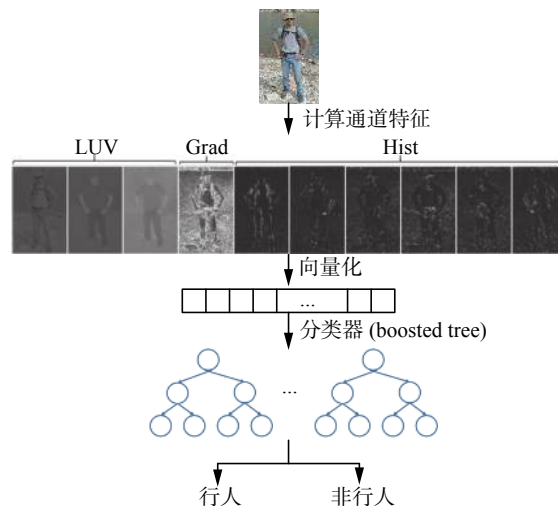


图 1 通道特征检测器

Fig. 1 Channel features detector

为了提高多尺度检测的速度,利用相邻尺度特征估计的方法构造快速特征金字塔。定义  $R(I, s)$  为原始图像  $I$  在尺度  $s$  处的特征采样,  $\Omega$  为通道变换,文献[10]通过对 INRIA 行人数据集图片在不同尺度下提取的通道特征进行统计分析,发现尺度  $s$  的通道特征  $C_s$  与原始图像通道特征  $C = \Omega(I)$  之间满足以下关系:

$$C_s \approx R(C, s) \cdot s^{-\lambda_\Omega} \quad (1)$$

式中  $\lambda_\Omega$  为通道估计系数。那么只要间隔一定尺度由原始图像计算通道特征  $C_s = \Omega(R(I, s))$ ,  $s \in \{1, \frac{1}{2}, \frac{1}{4}, \dots\}$ , 中间尺度的特征  $C_s$  就可以根据其最近尺度的已知特征进行较为精确的估算:

$$C_{s'} = R(C_s, s/s')(s/s')^{-\lambda_\Omega}, s' \in \{1, \frac{1}{2}, \frac{1}{4}, \dots\} \quad (2)$$

这样可以大幅减少常规方法在每个尺度都计算图像特征所需的计算开销。

对于输入的 640 像素  $\times$  480 像素大小的待检测道路图像,以间隔 8 个尺度的通道图像精确计算,其余尺度的通道图像近似估算的方式构造特征金字塔,在不同尺度的特征图上用  $64 \times 32$  的滑动窗口遍历,用训练好的分类器对每个候选区域进行判断,输出包含行人的窗口边框坐标及相应的分类得分,最后用非极大值抑制 (non-maximum suppression, NMS) 消除重叠的窗口。

单目视觉距离估计模型如图2所示,  $X_w, Y_w, Z_w$  为世界坐标系,相机光心在  $O$  点,焦距为  $f$ 。相机光轴  $c$  与成像平面交于点  $O'$ , 对应的像素坐标系原点坐标为  $(u_0, v_0)$ 。世界坐标系中的点  $Q(O, Y)$  在图像坐标系中的投影点为  $Q'(O, y)$ , 对应的像素坐标为  $(u', v)$ ; 点  $P(X, Y)$  在图像坐标系中的投影点为  $P'(x, y)$ , 对应的像素坐标为  $(u, v)$ 。相机距地面的高度为  $h$ , 光轴与水平方向的夹角为  $\alpha$ , 即相机的俯仰角。点  $Q$  与光心的连线与水平方向的夹角为  $\beta$ , 与光轴的夹角为  $\gamma$ 。

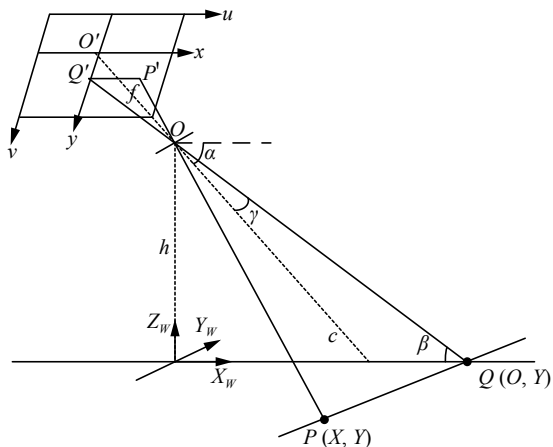


图2 单目视觉距离估计模型

Fig. 2 Monocular vision distance estimation model

由图2可以得到  $\beta = \alpha - \gamma$ ,  $\gamma = \arctan \frac{y}{f}$ 。从而纵向距离可以表示为

$$Y = \frac{h}{\tan \beta} = \frac{h}{\tan(\alpha - \arctan \frac{y}{f})} \quad (3)$$

根据比例关系  $\frac{OQ}{PQ} = \frac{OQ'}{P'Q'}$ , 由  $OQ' = \sqrt{OO'^2 + O'Q'^2} = \sqrt{f^2 + y^2}$ ,  $OQ$  为纵向距离,  $P'Q' = x$ , 可以得到横向距离:

$$X = \frac{P'Q'}{OQ'} Y = \frac{x}{\sqrt{f^2 + y^2}} \frac{h}{\sin(\alpha - \arctan \frac{y}{f})} = \frac{(u - u_0)h}{\sqrt{f^2 + (v - v_0)^2} \sin(\alpha - \arctan \frac{y}{f})} \quad (4)$$

摄像机距地面高度  $h$  和俯仰角  $\alpha$  通过直接测量得到, 本文的实验环境下  $h = 1\,063\text{ mm}$ ,  $\alpha = 9^\circ$ 。相机焦距  $f$  和像素坐标系原点坐标  $(u_0, v_0)$  为相机内部参数, 通过相机标定可以得到  $f = 624.858\,3$ ,  $(u_0, v_0) = (333.091\,9, 222.110\,7)$ 。根据单目视觉定位原理<sup>[14]</sup>, 选取检测到的行人边框下边界的中点为参考点, 可以从单帧静态图像中估计出行人与车辆间的纵向距离 ( $Y\_distance$ ) 与横向距离 ( $X\_distance$ )。道路图像中的行人检测与距离估计结果如图3所示。通道特征检测器选取对行人具有良好表征效果的梯度与色彩特征, 结合邻近特征估计的方法, 兼顾了检测速度与精度。上述行人位置估计方法在 CPU 下就可以达到 14 f/s 的处理速度, 可以满足有效预警中对于目标实时定位的要求。

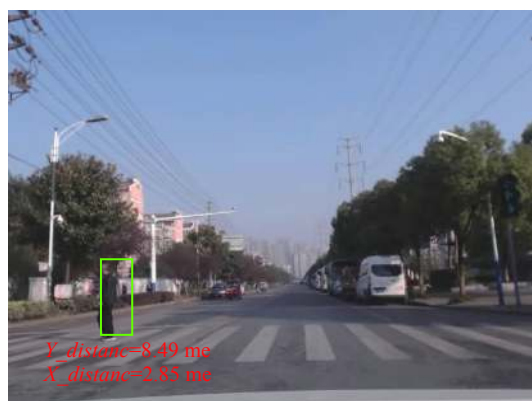


图3 行人检测与距离估计结果

Fig. 3 Pedestrian detection and distance estimation

## 2 驾驶员头部姿态估计

为了计算头部姿态,首先需要建立面部形状模型,得到精确的面部特征点坐标。再根据面部几何模型,利用特征点的几何位置和几何约束关系可以快速估算出头部姿态的方向参数。



多任务级联卷积网络 (multi-task cascaded convolutional networks, MTCNN)<sup>[9]</sup> 是目前速度与精度最好的人脸检测器之一。MTCNN 由 3 个卷积网络级联构成, 分类器级联可以大幅提升检测速度, 利用卷积网络提取的更加高级的特征可以提高分类的准确性, 此外由于卷积网络更好的区分能力, 级联框架的层数也会相应地减少, 从而降低整体的计算开销。同时利用卷积神经网络的多任务学习能力, 实现人脸及特征点的联合检测, 网络输出脸部边框坐标和左眼中心、右眼中心、鼻尖和嘴巴两端 5 个特征点的坐标。

在计算机视觉中, 物体姿态是指其相对相机的方向与位置。由于安置在车内的相机镜头正对驾驶员面部, 因而根据头部姿态可以判断驾驶员的注意方向。头部姿态估计可以视为多点透视 (perspective-n-point, PnP) 问题, 根据 2D-3D 特征点间的投影关系求解头部姿态参数。

针孔相机模型如图 4 所示。O-XYZ 为相机坐标系, F<sub>0</sub>-UVW 为世界坐标系坐标, F<sub>i</sub> 为头部三维模型上的某个特征点, 其对应于图像平面上的特征点为 p<sub>i</sub>(u<sub>i</sub>, v<sub>i</sub>)。f<sub>x</sub>、f<sub>y</sub> 为相机焦距, (c<sub>x</sub>, c<sub>y</sub>) 为相机光心。相机坐标系与世界坐标系之间的对应关系如下所示:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [R|T] \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix}, R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (5)$$

式中:  $R$  为旋转矩阵, 可以表示头部姿态;  $T$  为平移向量。

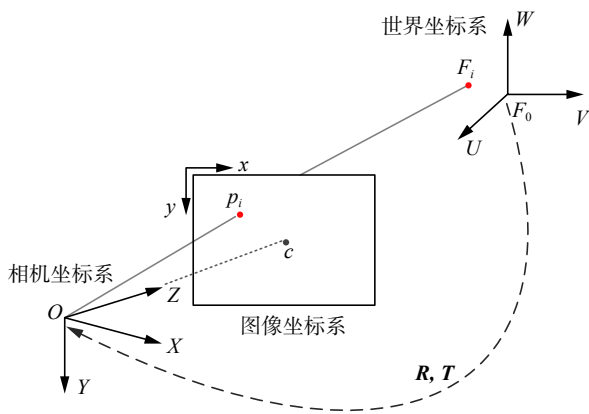


图 4 头部三维模型到图像平面特征点的投影

Fig. 4 Projection of 3D model of the head to the feature points of image plane

本文使用 EPnP 算法<sup>[15]</sup> 求解姿态参数。EPnP 算法利用 4 个虚拟控制点的线性加权表示实际参考点在相机坐标系下的坐标, 4 个虚拟控制点为  $c_j = 1, 2, 3, 4$ , 则世界坐标系中的参考点  $F_i$  可以表示为

$$F_i = \sum_{j=1}^4 \alpha_{ij} c_j^F, \sum_{j=1}^4 \alpha_{ij} = 1 \quad (6)$$

同理, 相机坐标系下对应的参考点可以表示为

$$p_i = \sum_{j=1}^4 \alpha_{ij} c_j^P \quad (7)$$

根据相机成像投影关系, 每个控制点  $c_j^P$  的三维坐标  $(x_j^P, y_j^P, z_j^P)$  与其在图像平面上的投影点  $(u_i, v_i)$  满足

$$\omega_n \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \sum_{j=1}^4 \alpha_{ij} \begin{bmatrix} x_j^P \\ y_j^P \\ z_j^P \end{bmatrix} \quad (8)$$

式中  $\omega_n$  为比例投影系数。进一步可以得到

$$\begin{cases} \sum_{j=1}^4 \alpha_{ij} f_x x_j^P + \alpha_{ij} (c_x - u_i) z_j^P = 0 \\ \sum_{j=1}^4 \alpha_{ij} f_y y_j^P + \alpha_{ij} (c_y - v_i) z_j^P = 0 \end{cases} \quad (9)$$

考虑  $n$  个参考点, 式 (9) 可以表示为线性方程组  $Mx = 0$ , 其中 4 个虚拟控制点共 12 个未知变量,  $M$  为  $2n \times 12$  矩阵,  $x = [c_1^P, c_2^P, c_3^P, c_4^P]^T$  为 12 维未知向量。式 (9) 的解为矩阵  $M$  的核空间, 表示为  $x = \sum_{i=1}^n \beta_i v_i$ ,  $n = 1, 2, 3, 4$ , 式中  $v_i$  为矩阵  $M$  的右奇异向量, 可由  $12 \times 12 M^T M$  矩阵的零特征向量求得,  $\beta_i$  的取值由使得逆投影误差最小的  $N$  值确定。得到虚拟控制点坐标后便可恢复出参考点在相机坐标系中的坐标, 利用 SVD 分解可以得到旋转矩阵。

将旋转矩阵转换为姿态角 pitch、yaw 和 roll, 可以直观地表示头部朝向, 其变换关系如下:

$$\begin{cases} \text{pitch} = \arctan2(r_{32}, r_{33}) \\ \text{yaw} = \arctan2(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2}) \\ \text{roll} = \arctan2(r_{21}, r_{11}) \end{cases} \quad (10)$$

本文使用 MTCNN 输出的左右眼中心、鼻尖、左右嘴角 5 个特征点作为参考点, 估计驾驶员头部姿态角, 结果如图 5 所示。

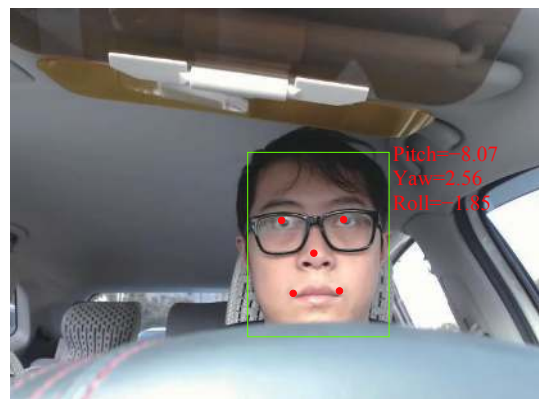


图 5 驾驶员头部姿态估计

Fig. 5 Head pose estimation of the driver

### 3 碰撞风险评估

碰撞保护系统通常根据传感器获取的道路障碍物距离和速度信息计算碰撞时间 (time to collision, TTC), 当碰撞时间小于阈值时做出警告。然而受到行车环境、驾驶员精神及身体状态的影响, 驾驶员对道路危险的反应时间存在差异, 精确的数值分析有时反而会引起错误的预警。实际道路中, 经验丰富的驾驶员并不会精确计算距离或碰撞时间, 而是使用语义信息表达路况, 例如用远、中、近、很近表示前方物体距离, 用很快、快、适中、慢表示当前车速等。可以近似认为, 驾驶员对于车辆前方潜在碰撞风险的估计都是基于一系列 if-then 规则做出的决策, 并且判断结果在大部分情况下已足够准确。

本文构建一种基于道路和驾驶员信息的模糊推理系统, 对输入信息进行决策级融合输出碰撞风险估计, 包括低风险 (low-risk)、中等风险 (mid-risk)、高风险 (high-risk) 和极高风险 (very-high-risk) 4 个等级。系统的输入为道路状态信息和驾驶员状态信息, 式中道路状态信息表现在行人与

车辆间的纵向距离 ( $Y\_distance$ ) 和横向距离 ( $X\_distance$ ), 驾驶员注意方向信息表现在头部垂直方向转角 (pitch) 和水平方向转角 (yaw)。模糊推理使用 Mamdani 规则, 如式 (11) 所示:

$$R^i: \text{if } x_1 \text{ is } A_1^i, x_2 \text{ is } A_2^i, \dots, x_k \text{ is } A_k^i \text{ then } y^i = B^i \quad (11)$$

式中:  $x_k$  为输入变量;  $y^i$  为第  $i$  条规则对应的输出;  $A^i$  和  $B^i$  为模糊集。

为了降低论证的复杂性, 取行人的纵向距离  $Y\_distance$  与横向距离  $X\_distance$ , 驾驶员头部水平转角 yaw 3 个输入参数构建模糊推理系统。输入参数  $Y\_distance$  与  $X\_distance$  模糊化使用三角隶属度函数, 其中  $Y\_distance$  的模糊集论域为  $[0, 20]$ ,  $X\_distance$  的模糊集论域为  $[0, 5]$ , 在论域上的取值均为 {close, mid, far}, 如图 6(a) 和图 6(b) 所示。输入参数 yaw 模糊化使用高斯隶属度函数, 模糊集论域为  $[-30, 30]$ , 在论域上的取值为 {left, center, right}, 如图 6(c) 所示。系统输出参数 Risk 设置 {low, mid, high, veryhigh} 4 个等级, 模糊集论域设为  $[0, 1]$ , 使用高斯隶属度函数, 如图 6(d) 所示。

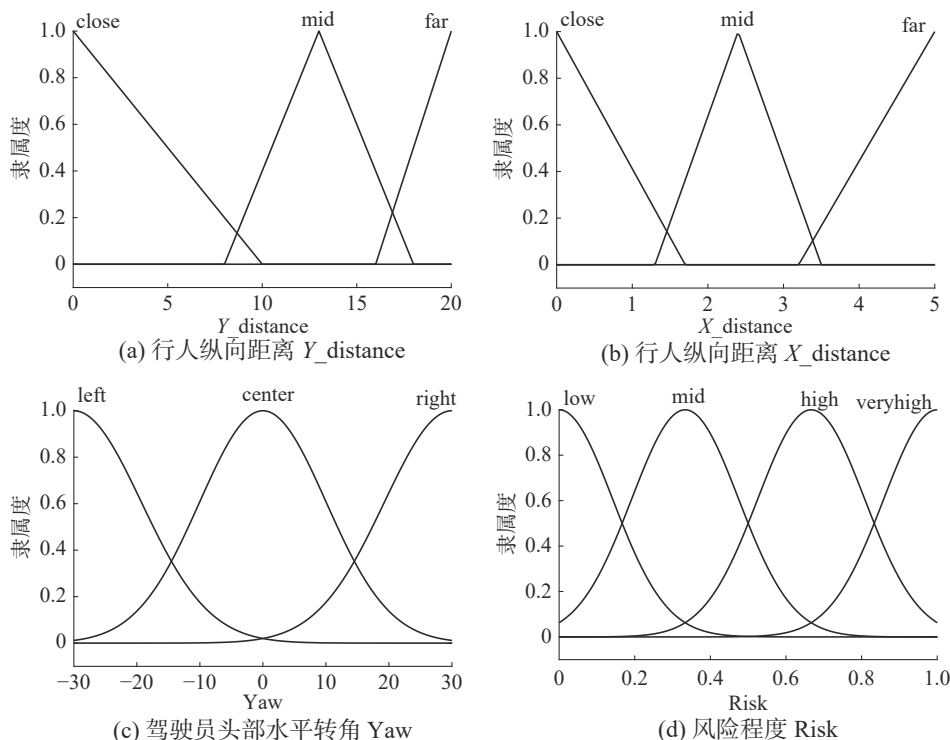


图6 模糊系统输入输出隶属度函数

Fig. 6 Input/output membership functions of fuzzy system

根据 Mamdani 规则构建模糊推理规则, 共设置 27 条模糊规则, 涵盖了输入输出参数在论域上的所有取值, 具体如表 1 所示。

利用重心法进行解模糊化, 根据系统输出的

数值可以判断出对应的碰撞风险等级。由此构建的模糊推理系统通过对车内外环境下的多种输入参数的模糊融合, 可以为车碰撞事故的预警提供一定的指导。

表 1 模糊推理规则

Table 1 Fuzzy inference rules

纵向距离	横向距离	头部转角	风险等级	纵向距离	横向距离	头部转角	风险等级	纵向距离	横向距离	头部转角	风险等级
近	近	左	非常高	中	近	左	高	远	近	左	低
近	近	中	非常高	中	近	中	中	远	近	中	低
近	近	右	非常高	中	近	右	高	远	近	右	低
近	中	左	高	中	中	左	中	远	中	左	低
近	中	中	中	中	中	中	低	远	中	中	低
近	中	右	高	中	中	右	中	远	中	右	低
近	远	左	低	中	远	左	低	远	远	左	低
近	远	中	低	中	远	中	低	远	远	中	低
近	远	右	低	中	远	右	低	远	远	右	低

4 实验分析

图像采集使用安置在车内的两个单目相机, 其中一个相机安放在前挡风玻璃后, 镜头对准前方道路, 另一个相机安放在仪表盘上, 镜头对准

驾驶员。为了提高系统的处理速度, 视频分辨率设为  $640\times480$ 。在实际道路采集的视频中, 选取车辆在路口转弯时的 30 s 片段进行分析, 其中部分时间点采集的车内外图像如图 7 所示。

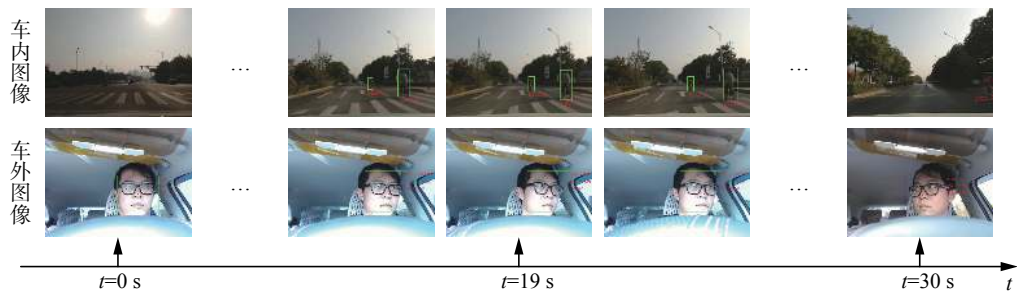


图 7 视频不同时间点采集的车内外图像

Fig. 7 Interior and exterior images collected at different time

根据第 1~3 节所述的方法对车内外视频进行逐帧处理。在同步拍摄得到的图像中, 提取行人相对本车的纵向距离 ( $Y\_distance$ ) 和横向距离

( $X\_distance$ ) 作为道路状态数据, 提取驾驶员的头部方向角 (pitch, yaw, roll) 作为驾驶员状态数据, 车辆内外信息如图 8 所示。

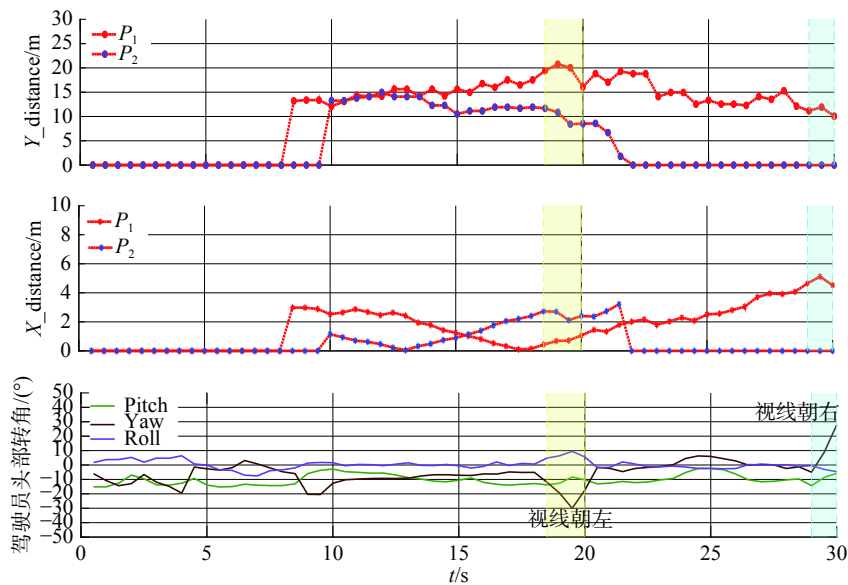


图 8 行人与驾驶员状态数据

Fig. 8 Data of pedestrian and driver status

根据图8所示数据对提出的碰撞预警方法进行分析: 1) 视频开始处车辆驶入路口准备左转, 此时未检测到前方行人, 驾驶员头部水平转角(yaw)在 $-12^{\circ} \sim -10^{\circ}$ 之间, 此时不存在风险。2) 在18~20 s时间段, 驾驶员出现注意力不集中情况, 头部水平转角(yaw)最大达到 $-30^{\circ}$ , 此时驾驶员头部朝左侧。在同一时间段的道路图像序列中检测到2个行人, 分别记为 $P_1$ 与 $P_2$ 。行人 $P_1$ 与车辆的纵向距离 $Y_1\_distance \approx 20$  m, 横向距离

$X_1\_distance \approx 0.7$  m; 行人 $P_2$ 与车辆的纵向距离 $Y_2\_distance \approx 8$  m, 横向距离 $X_2\_distance \approx 2$  m。3) 在28~30 s时间段, 驾驶员头部水平转角(yaw)最大达到 $28^{\circ}$ , 驾驶员注意力集中于道路右侧。与此同时, 道路图像中检测到行人 $P_1$ , 其纵向距离 $Y_1\_distance \approx 2$  m, 横向距离 $X_1\_distance \approx 4.5$  m。

取行人的纵向距离 $Y\_distance$ 与横向距离 $X\_distance$ , 驾驶员头部水平转角yaw 3个输入参数构建模糊推理系统, 如图9所示。

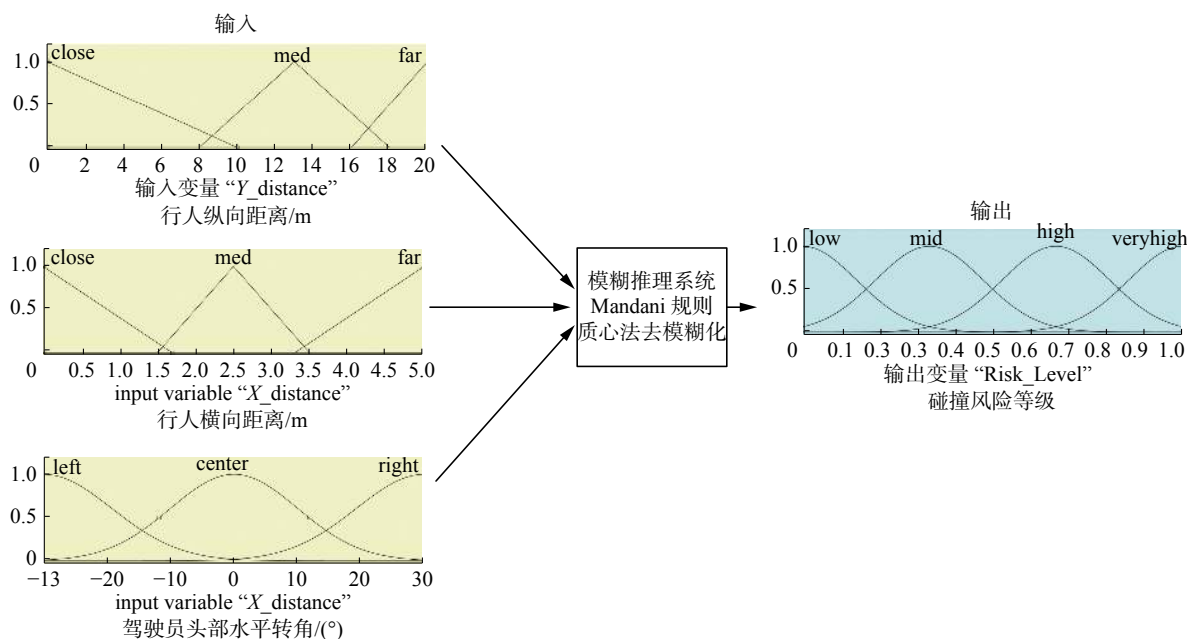


图9 3个输入( $Y\_distance$ ,  $X\_distance$ , Yaw)的模糊推理系统

Fig. 9 Fuzzy inference system of three inputs ( $Y\_distance$ ,  $X\_distance$ , Yaw)

通过模糊推理得到的碰撞风险等级解模糊平面如图10所示。图10(a)为 $t=19$  s时的解模糊平面, 模糊推理系统对行人 $P_1$ 的输出为0.126 6, 判断为低风险(low-risk)等级, 对行人 $P_2$ 的输出为0.647 1, 判断为高风险(high-risk)等级。图10(b)

为 $t=30$  s时的解模糊平面, 模糊推理系统的输出为0.119 5, 由此得到的碰撞风险等级为低风险(low-risk)。在Core i5 CPU, 8 GB RAM, GTX1060 GPU运行环境下, 该碰撞预警系统可以达到4~5 f/s的运行速度, 基本可以满足实时性的要求。

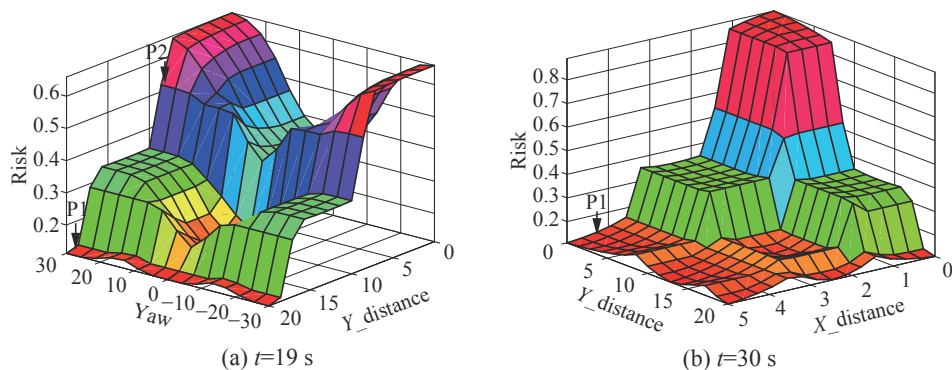


图10 碰撞风险等级的解模糊平面

Fig. 10 Defuzzified surface of the collision risk level

结合图7中道路与驾驶员图像, 分析系统对于风险判断的有效性: 1) 在19 s处, 车辆内部信

息表明驾驶员视线落在车辆左边区域, 无法有效掌握前方路况信息, 因而处于危险状态; 车辆外



部信息中,由于我国城市道路每条车道宽度为3.5 m,行人 $P_1$ 的横向距离表明其在当前车道内,但其纵向距离其余本车相距较远,因而暂时处于低风险区域。行人 $P_2$ 的纵向距离已处在危险距离内,横向距离表明其距本车右侧车道线较近,因而存在碰撞风险,需要提醒驾驶员注意。2)在29 s处,此时根据驾驶员头部姿态可以判断其未注意道路前方,视为危险状态,然而行人 $P_1$ 的横向距离表明其在本车右侧车道且距离较远,虽然纵向距离较近,实际路况为安全状态。上述根据车内外图像分析的结果与模糊推理系统的输出一致,说明本文提出的基于内外视觉信息的行人碰撞预警方法较为可靠,可以在行车过程中为驾驶员提供辅助决策。

## 5 结束语

本文提出一种基于车内外视觉信息的行人碰撞预警方法,利用两个单目相机同时提取道路行人和驾驶员状态数据,通过模糊推理系统结合车辆外部行人信息与车辆内部驾驶员信息进行评估,给出当前场景下的风险等级,对可能发生的人车碰撞做出预判。在实际道路上的实验表明,该方法对于不同场景下的危险判断具有很好的效果,可以在驾驶员注意力不集中时提供有效的指导。然而由于视觉传感器自身的限制,行人的定位非常依赖采集图像的质量,结合多种传感器信息提升车辆外部信息获取的可靠性还有许多需要进一步开展的工作。

## 参考文献:

- [1] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C]//Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, USA, 2001: 1-511-1-518.
- [2] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA, 2005: 886-893.
- [3] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. *IEEE Transactions on pattern analysis and machine intelligence*, 2010, 32(9): 1627-1645.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, USA, 2015: 1440-1448.
- [6] STANTON N A, SALMON P M. Human error taxonomies applied to driving: a generic driver error taxonomy and its implications for intelligent transport systems[J]. *Safety science*, 2009, 47(2): 227-237.
- [7] TAWARI A, MARTIN S, TRIVEDI M M. Continuous head movement estimator for driver assistance: issues, algorithms, and on-road evaluations[J]. *IEEE transactions on intelligent transportation systems*, 2014, 15(2): 818-830.
- [8] COOTES T F, EDWARDS G J, TAYLOR C J. Active appearance models[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2001, 23(6): 681-685.
- [9] REN Shaoqing, CAO Xudong, WEI Yichen, et al. Face alignment at 3000 fps via regressing local binary features[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA, 2014: 1685-1692.
- [10] SUN Yi, WANG Xiaogang, TANG Xiaoou. Deep convolutional network cascade for facial point detection[C]//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA, 2013: 3476-3483.
- [11] ZHANG Kaipeng, ZHANG Zhanpeng, LI Zhifeng, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. *IEEE Signal processing letters*, 2016, 23(10): 1499-1503.
- [12] DOLLÁR P, APPEL R, BELONGIE S, et al. Fast feature pyramids for object detection[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2014, 36(8): 1532-1545.
- [13] BOURDEV L, BRANDT J. Robust object detection via soft cascade[C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA, 2005: 236-243.
- [14] 王牛, 李祖枢, 武德臣, 等. 机器人单目视觉定位模型及其参数辨识[J]. *华中科技大学学报(自然科学版)*, 2008, 36(S1): 57-60.

WANG Niu, LI Zushu, WU Dechen, et al. Robot monocular vision position determination model and its parametric identification[J]. *Journal of Huazhong University of Science and Technology (Nature Science Edition)*,



2008, 36(S1): 57–60.

- [15] LEPETIT V, MORENO-NOGUER F, FUA P. EPnP: an accurate  $O(n)$  solution to the PnP problem[J]. [International journal of computer vision](#), 2009, 81(2): 155–166.
- [16] TAWARI A, SIVARAMAN S, TRIVEDI M M, et al. Looking-in and looking-out vision for Urban Intelligent Assistance: estimation of driver attentive state and dynamic surround for safe merging and braking[C]//Proceedings of 2014 IEEE Intelligent Vehicles Symposium. Dearborn, USA, 2014: 115–120.
- [17] TATEIWA K, YAMADA K. Estimating driver awareness of pedestrians in crosswalk in the path of right or left turns at an intersection from vehicle behavior[C]//Proceedings of 2015 IEEE Intelligent Vehicles Symposium. Seoul, South Korea, 2015: 952–957.
- [18] ROTH M, FLOHR F, GAVRILA D M. Driver and pedestrian awareness-based collision risk analysis[C]//Proceedings of 2016 IEEE Intelligent Vehicles Symposium. Gothenburg, Sweden, 2016: 454–459.

### 作者简介:



杨会成, 男, 1970 年生, 教授, 主要研究方向为图像信息处理、疲劳驾驶检测。主持和参与安徽省自然科学基金项目、安徽省高校自然科学基金项目 6 项。发表学术论文 15 篇。



朱文博, 男, 1992 年生, 硕士研究生, 主要研究方向为图像处理与模式识别。



童英, 女, 1993 年生, 硕士研究生, 主要研究方向为图像处理与模式识别。