

文章编号:1671-8879(2016)05-0079-09

# 非约束条件下的驾驶人嘴唇检测方法

程文冬<sup>1,2</sup>, 付 锐<sup>1</sup>, 马 勇<sup>1</sup>, 张名芳<sup>1</sup>, 刘 通<sup>1</sup>

(1. 长安大学 汽车运输保障技术交通行业重点实验室, 陕西 西安 710064;

2. 西安工业大学 机电工程学院, 陕西 西安 710032)

**摘 要:**基于机器视觉的嘴唇识别是驾驶人行为监测的关键技术之一,理想光照、正面姿态与闭嘴状态下的嘴唇识别率较高,但实际驾驶环境中存在光照变化、头部转动与嘴部运动等非约束条件。提出了一种实车环境下的驾驶人面部识别与嘴唇分割方法。考虑到肤色在像素空间邻域与灰度值域的平滑相似性,首先给出一种基于双边滤波器与单尺度 Retinex 的光照均衡方法,提升非均匀光照下肤色聚类的稳定性,进而建立 Adaboost-YCb'Cr' 动态肤色模型识别人脸,该方法可以有效适应复杂环境下的光照变化与头部姿态。提出一种基于 YCbCr 与 HSI 色彩空间的复合色度累加策略,在有效增强唇色与肤色阈值可分性的基础上,采用降维的 2D-OTSU 方法准确地分割嘴唇区域。研究表明:提出的算法能有效检测实际行车环境下的驾驶人嘴唇区域,对于非均匀光照、头部转动与嘴部动作产生的色度畸变、嘴部阴影与边缘模糊等干扰具有较好的鲁棒性;该算法可应用于检测实际行车中驾驶人的说话、打哈欠等行为。

**关键词:**交通工程;驾驶行为监控;机器视觉;光照均衡;肤色模型;2D-OTSU

**中图分类号:**U491      **文献标志码:**A

## Driver's lip detection method under non-restraint conditions

CHENG Wen-dong<sup>1,2</sup>, FU Rui<sup>1</sup>, MA Yong<sup>1</sup>, ZHANG Ming-fang<sup>1</sup>, LIU Tong<sup>1</sup>

(1. Key Laboratory of Automotive Transportation Safety Technology of Ministry of Transport, Chang'an University, Xi'an 710064, Shaanxi, China; 2. School of Mechatronic Engineering, Xi'an Technological University, Xi'an 710032, Shaanxi, China)

**Abstract:** Lip recognition based on machine vision is one of the key techniques of driver behavior surveillance. Lip region can be detected with a desirable recognition rate under ideal conditions, including good illumination, stationary state of head and lip. However, non-restraint conditions, including unequalized illumination, head rotation and lip motion, seriously restrict the lip recognition efficiency. To adapt these unconstrained conditions, face detection and lip segmentation methods under real driving situation were put forward. Firstly, an illumination equalization method based on the algorithms of bilateral filter and single Retinex (BF-SSR) was proposed to enhance the clustering stability of skin color under unequalized illumination. Dynamic skin model based on Adaboost and YCb'Cr' algorithms was then set up for face recognition. Finally, a multi-color accumulate strategy was introduced to enhance the threshold separability

收稿日期:2016-05-16

基金项目:国家自然科学基金项目(61374196,61473046);陕西省自然科学基金项目(2016JQ5096);

陕西省教育厅专项科研项目(16JK1375);西安工业大学校长基金项目(XAGDXJJ15006)

作者简介:程文冬(1981-),男,河北保定人,西安工业大学讲师,长安大学工学博士研究生,E-mail:chengwendong@foxmail.com。

between lip and face, and 2D-OTSU criterion with dimension-reduction was proposed for lip segmentation. The results show that the method can detect driver's lip region efficiently. The proposed algorithm is robust to lip shadow and edge blur caused by unequalized illumination and head pose. The algorithm is available for speaking and yawn detection in real driving environment. 1 tab, 14 figs, 20 refs.

**Key words:** traffic engineering; driver behavior surveillance; machine vision; illumination equalization; skin model; 2D-OTSU

## 0 引言

基于机器视觉的驾驶人嘴唇识别(driver lip recognition, DLR)能够估计驾驶人的精神分散状态(如说话、打哈欠等),是检测驾驶人行为状态的一项重要内容。DLR对于完善辅助驾驶系统(advanced driving assistant system, ADAS)、提升车辆主动安全性等具有重要意义,其关键技术包括:鲁棒的面部识别、有效的唇色增强以及嘴唇轮廓分割方法等。基于样本结构规则的 Adaboost 等算法对正面姿态及静态的人脸、嘴唇具有良好的识别效果,对动态光照也有较强的耐受性<sup>[1]</sup>。基于色度模型的嘴唇识别利用唇色的聚类特性提取相关性像素,再利用边缘、角点等像素信息来匹配嘴唇的细节特征<sup>[2-3]</sup>。国内外学者主要采取色度转换与增强、光照均衡化等方法来提高算法的光照耐受性与自适应收敛性。Cheung 等假设嘴唇区域的亮度呈线性变化,提出一种纵横双向的灰度均衡算法来提高嘴唇的光照平滑性<sup>[4]</sup>;文献<sup>[5]</sup>提出一种基于 YUV 空间的唇色滤波器来提高嘴唇色度的聚类稳定性,计算简单,但在实际驾驶环境中该方法易受光照干扰;文献<sup>[6]</sup>采用 Fisher 变换放大嘴唇与皮肤的类间距,同时缩小类内距离,此方法需要训练大量样本以获取色度最佳投影向量,同时需要确保训练样本与测试图像中嘴唇色度的相似性;文献<sup>[7]</sup>提出一种基于梯度信息和先验知识的 Level Set 嘴唇边缘精确拟合方法,由于其拟合精度依赖于静态嘴唇的曲线演化与迭代收敛,因此对于头部姿态与嘴部动作的适应性不高。

面部光照、头部姿态与嘴部是制约 DLR 鲁棒性与稳定性的主要非约束条件。目前 DLR 研究中尚存在以下 2 个方面的关键问题有待进一步探索。第一,适应复杂光照变化与头部姿态的驾驶人面部识别是实现 DLR 的首要条件。基于样本纹理结构规则的人脸识别对于光照条件比较鲁棒,但对头部姿态和面部动作的容忍性较差。基于皮肤色度分割的人脸识别方法虽然与头部姿态、面部动作和部分遮

挡不相关,但在实际行车过程中色度聚类范围易受到光照干扰而发生漂移,导致色度判别规则欠缺环境适应性。对此需要深入开展融合策略的人脸识别,以提高算法对光照、头部姿态与面部动作的自适应性。第二,目前基于单一色彩空间<sup>[5,8]</sup>、基于色度样本训练<sup>[6]</sup>、梯度边缘<sup>[7]</sup>等方法的嘴唇识别对均衡光照与静态面部的依赖严重,或者须满足离线样本与测试图像的特征一致性,因此难以满足 DLR 的实际应用。

由上述分析可见,实车环境下的 DLR 仍是一项具有挑战的研究课题。鉴于此,本文首先提出一种基于光照均衡的 Adaboost-YCb'Cr' 动态肤色模型来识别驾驶人脸;进而提出一种基于双色彩空间的复合色度累加策略,增强变化光照下嘴唇与皮肤的色度可分性;再采用降维的 2D-OTSU 方法实现嘴唇的准确分割。实车试验表明,所提出的融合算法有效提高了 DLR 的光照鲁棒性与动作适应性。

## 1 面部光照预处理

驾驶人面部光照干扰主要有 2 类,由漫反射产生的面部亮度不均匀与“自遮挡”产生的局部阴影。这些因素会导致肤色分布出现发散或漂移,不易建立稳定的肤色模型。本文提出一种融合双边滤波器的单尺度 Retinex 算法来进行面部光照均衡处理,提升肤色的聚类稳定性。

### 1.1 单尺度 Retinex 算法

单尺度 Retinex(single scale retinex, SSR)算法认为,同一目标在动态光照下遵循颜色恒定原则,其光滑性近似为常数,输入图像  $I(x, y)$  可分解为反射层分量  $R(x, y)$  和光照层分量  $L(x, y)$ <sup>[9]</sup>,因此可对层域空间上的光照强度进行估计并对像素灰度赋予新值,还原目标的真实特征。SSR 算法为

$$\lg[R_i(x, y, f_s)] = \lg[I_i(x, y)] - \lg[f_s(x, y)], \quad i \in R, G, B \quad (1)$$

式中:  $f_s(x, y)$  为环绕平滑函数,通常取 Gauss 函数与  $I(x, y)$  作卷积运算;  $R, G, B$  分别为红、绿、蓝

色彩。

SSR 算法的关键问题在于如何对光照层分量  $L(x, y)$  进行准确预测,即环绕平滑函数  $f_s(x, y)$  的选型问题。对于光照平滑过渡的图像, Gauss 环绕函数能够保持目标图像的颜色恒定与边缘细节,但由于驾驶人面部并不严格遵循光照平滑变化的假设,故这种情况下 SSR 易放大噪声,丢失细节边缘以及产生光晕。本文引入一种二维空间滤波器来估计输入图像的亮度变化特性,提升 SSR 算法对光照的平滑性能,兼顾实现噪声抑制与边缘保留。

## 1.2 双边滤波器

双边滤波器 (bilateral filter, BF) 是一种基于 Gaussian 滤波的非线性二维滤波方法,由值域滤波和空间域滤波组成<sup>[10]</sup>。设像素亮度函数为  $f(x)$ , 值域滤波传递函数  $h_r(x)$  和空间域滤波传递函数  $h_d(x)$  为

$$h_r(x) = \frac{1}{k_r(x)} \int_{-\infty}^{+\infty} f(x) S[f(x), f(\xi)] d\xi \quad (2)$$

$$h_d(x) = \frac{1}{k_d(x)} \int_{-\infty}^{+\infty} f(x) D(x, \xi) d\xi \quad (3)$$

式中:  $k_r(x)$ 、 $k_d(x)$  为归一化常数;  $S[f(x), f(\xi)]$ 、 $D(x, \xi)$  分别为中心点  $x$  与邻域点  $\xi$  的光照相似度与距离相似度。

则双边滤波传递函数  $h(x)$  可对式(1)与式(2)进行融合,有

$$h(x) = \frac{1}{k(x)} \iint_{-\infty}^{+\infty} f(x) S[f(x), f(\xi)] \cdot D(\xi, x) d\xi dx \quad (4)$$

式中:  $k(x)$  为融合后的归一化参数,可表示为

$$k(x) = \iint_{-\infty}^{+\infty} S[f(x), f(\xi)] D(\xi, x) d\xi dx \quad (5)$$

$S[f(x), f(\xi)]$  与  $D(\xi, x)$  可表达为欧氏距离参数的高斯函数,如式(6)所示

$$\begin{cases} S[f(x), f(\xi)] = e^{-0.5 \left( \frac{\|f(x) - f(\xi)\|}{\sigma_r} \right)^2} \\ D(x, \xi) = e^{-0.5 \left( \frac{\|f(\xi) - f(x)\|}{\sigma_d} \right)^2} \end{cases} \quad (6)$$

式中:  $\sigma_r$ 、 $\sigma_d$  分别为高斯函数的灰度标准差和距离标准差。

BF 在处理邻域像素灰度时不仅考虑了二维空间的临近关系,也兼顾了光照平滑相似性。由式(6)可知,光照相似度  $S[f(x), f(\xi)]$  随中心点与邻域点欧氏距离的增大而减小;距离相似度  $D(x, \xi)$  也随中心点与邻域点光照差值的增大而减小。对于灰度变化平缓区域, BF 具有高斯低通滤波器的特性。而对

于灰度变化剧烈的区域(特征边缘), BF 将邻域亮度均值赋予中心像素点,因此能够平滑图像并保持边缘。图 1 是来自网络“Lena”图像(增加高斯噪声)的滤波比较<sup>[11-12]</sup>。高斯滤波在平滑噪声的同时抑制了边缘细节,中值滤波与之相反,噪声与边缘细节同时得到保留。而 BF 对噪声抑制与边缘细节的兼顾性更好。



图 1 “Lena”图像(高斯噪声)的滤波比较

Fig. 1 Filtering comparison to image “Lena” (with Gauss noise)

## 1.3 BF-SSR 光照均衡方法

基于 BF-SSR 算法的面部图像光照均衡流程如图 2 所示,图中驾驶人图像来自于实车试验中建立的图像数据库。首先对输入图像  $I(x, y)$  进行对数转换;然后针对  $R, G, B$  分量采用  $\text{Gray} = 0.2R + 1.2G + 0.2B$  的方法进行简单的对比度增强,与此同时对  $I(x, y)$  进行 BF 光照估计并为像素赋予新值;最后将求得的对数反函数  $\exp$  输出光照均衡化图像  $R(x, y)$ 。下页图 3 是 BF-SSR 光照均衡示例。第 1 行是驾驶人面部图像( $G$  通道),第 2 行是 BF-SSR ( $\sigma_r = 2.6$ ,  $\sigma_d = 5$ ) 的光照均衡结果。通过对比 3 组灰度直方图可发现,原图中离散严重的低灰度与高亮度区域均得到显著的均衡化处理,肤色的聚类性得到增强,为建立稳健的驾驶人肤色模型提供了重要前提。

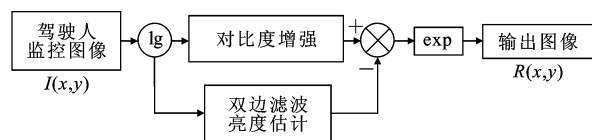


图 2 面部光照均衡流程

Fig. 2 Flow of face illumination equalization

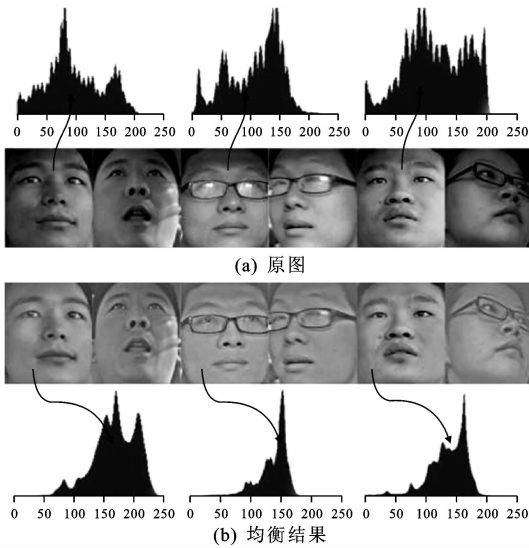


图3 光照均衡示例

Fig. 3 Examples of BF-SSR illumination equalization

## 2 面部定位方法

为适应驾驶人头部姿态,克服肤色发散与漂移问题,本文给出一种结合 Adaboost 与  $YCb'Cr'$  算法的驾驶人面部定位方法。 $YCb'Cr'$  是  $YCbCr$  空间的非线性变换,可削弱色度分量  $Cb$ 、 $Cr$  与亮度分量  $Y$  的非线性依赖关系,进而获取更为稳定的肤色聚类范围<sup>[13]</sup>。由于  $Cb'$ 、 $Cr'$  分量在 2D 平面上近似遵循椭圆形状的分布规律,因此可采用椭圆方程作为肤色判别的边界条件。面部区域的检测流程如图 4 所示,具体步骤如下所述。

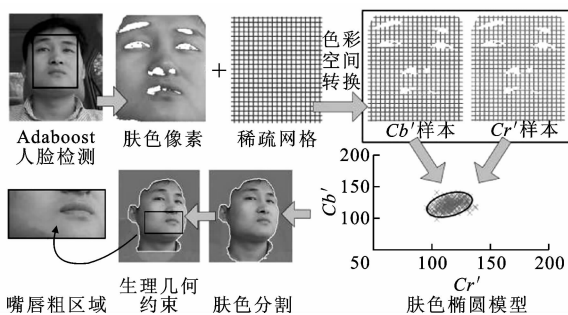


图4 肤色检测流程

Fig. 4 Flow of skin detection

Step1: 当驾驶人头部接近于正面姿态时,采用 Adaboost 方法检测人脸,并以 3 : 4 矩形提取人脸中心区域,排除掉面部周围的头发、背景等无用信息<sup>[1]</sup>。

Step2: 采用简单阈值方法,去除掉矩形区域内的眼睛、鼻孔、眉毛等低灰度区域,获取干净的肤色像素并进行 BF-SSR 光照均衡<sup>[2]</sup>。

Step3: 建立稀疏网格掩膜图像,单元大小为 20 pixels × 20 pixels,线宽为 4 pixels,对肤色像素进行间隔采样。该方法在考虑全局肤色信息的同时大大减少了冗余像素的计算量,提升了算法的实时性。

Step4: 将肤色像素样本转换到  $YCb'Cr'$  色彩空间,获取  $Cb'$ 、 $Cr'$  分量值。建立  $Cb'-Cr'$  坐标系,采用最小二乘法拟合肤色椭圆方程<sup>[14]</sup>。

Step5: 当 Adaboost 算法失效时,采用肤色椭圆模型提取面部肤色区域。若在后续图像中检测的肤色区域面积  $A_{skin}$  超出  $A_{min} \sim A_{max}$  阈值范围,说明肤色分布范围受到很大的光照影响已发生漂移,出现较大的漏检或误检。此时返回 Step1,通过 Adaboost 重新采样建立  $Cb'-Cr'$  肤色模型。

Step6: 根据人脸与五官的先验性生理几何结构,采用长宽比 2 : 1 的矩形粗略定位嘴唇<sup>[6]</sup>。

对实车试验中的驾驶人面部检测与嘴唇粗定位示例如图 5 所示。驾驶人肤色与光照条件各不相同,左右面部也存在光照不均衡现象。图 5 中白色封闭曲线是由  $Cb'-Cr'$  椭圆模型提取的面部肤色区域,黑色矩形框是基于五官几何约束的嘴唇粗区域。本文算法在这些场景下均能够稳定分割出面部肤色,对于多样性的头部姿态和嘴部动作同样具有鲁棒性。



图5 面部检测与嘴唇粗定位

Fig. 5 Face detection and lip coarse position

## 3 基于 2D-OTSU 准则的嘴唇分割

肤色与唇色的聚类特性都很显著,由两者组成的色度直方图具有典型的双峰分布特征,可采用 OTSU 算法实现最佳阈值分割<sup>[15]</sup>。但是在头部与嘴唇运动时,嘴唇区域附近会产生较多的阴影与边缘,这些随机性干扰对嘴唇的分割精度影响较大。而作为 OTSU 的扩展方法,2D-OTSU 更适用于存在噪声、边缘和灰度分布交叉的嘴唇图像,具有更优的分割准确性与抗噪性<sup>[16-17]</sup>。

### 3.1 2D-OTSU 准则

2D-OTSU 是以原始图像和邻域平滑图像的 2D 灰度直方图为依据,以目标和背景的最大类间方差

为准则的阈值自动搜索方法。设定大小为  $M \times N$  的图像  $f(x, y)$ , 其邻域均值图像为  $g(x, y)$ , 两者灰度范围均为  $0 \sim (L-1)$ 。设  $f(x, y)$  像素灰度级为  $i$  且  $g(x, y)$  灰度级为  $j$  的像素数量为  $N(i, j)$ , 则二元组  $(i, j)$  在  $f(x, y)$  与  $g(x, y)$  中出现的 2D 联合概率  $p(i, j) = N(i, j) / (M \times N)$ ,  $0 \leq p(i, j) \leq 1$ 。则  $L \times L$  级的 2D 灰度直方图如图 6 所示。图 6 中, 阈值  $(s, t)$  将图像分为 4 个区域, A、B 分别对应目标和背景, C、D 分别对应边缘和噪声。A、B 的联合概率分别为  $P_A(s, t)$  和  $P_B(s, t)$ , 相应的类内均值向量分别为  $\mu_A$  和  $\mu_B$ , 则 2D 直方图的总均值向量为  $\mu_T$

$$\begin{cases} P_A(s, t) = \sum_{i=0}^{s-1} \sum_{j=0}^{t-1} p(i, j) \\ P_B(s, t) = \sum_{i=s}^{L-1} \sum_{j=t}^{L-1} p(i, j) \end{cases} \quad (7)$$

$$\begin{cases} \mu_A = [\mu_{Ai}, \mu_{Aj}]^T = \left[ \frac{\sum_{i=0}^{s-1} \sum_{j=0}^{t-1} ip(i, j)}{P_A(s, t)}, \frac{\sum_{i=0}^{s-1} \sum_{j=0}^{t-1} jp(i, j)}{P_A(s, t)} \right]^T \\ \mu_B = [\mu_{Bi}, \mu_{Bj}]^T = \left[ \frac{\sum_{i=s}^{L-1} \sum_{j=t}^{L-1} ip(i, j)}{P_B(s, t)}, \frac{\sum_{i=s}^{L-1} \sum_{j=t}^{L-1} jp(i, j)}{P_B(s, t)} \right]^T \end{cases} \quad (8)$$

$$\mu_T = [\mu_{Ti}, \mu_{Tj}]^T = \left[ \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} ip(i, j), \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} jp(i, j) \right]^T \quad (9)$$

定义目标 A 和背景 B 的类间离散矩阵  $S_B$  为

$$S_B = P_A \begin{bmatrix} (\mu_{Ai} - \mu_{Ti})^2 & (\mu_{Ai} - \mu_{Ti})(\mu_{Aj} - \mu_{Tj}) \\ (\mu_{Ai} - \mu_{Ti})(\mu_{Aj} - \mu_{Tj}) & (\mu_{Aj} - \mu_{Tj})^2 \end{bmatrix} + P_B \begin{bmatrix} (\mu_{Bi} - \mu_{Ti})^2 & (\mu_{Bi} - \mu_{Ti})(\mu_{Bj} - \mu_{Tj}) \\ (\mu_{Bi} - \mu_{Ti})(\mu_{Bj} - \mu_{Tj}) & (\mu_{Bj} - \mu_{Tj})^2 \end{bmatrix} \quad (10)$$

式中:点坐标  $(s, t)$  如图 6 所示;  $\mu_{Ai}, \mu_{Aj}, \mu_{Bi}, \mu_{Bj}, \mu_{Ti}, \mu_{Tj}$  分别为相应的类内均值向量的元素。

以矩阵  $S_B$  的迹  $\text{tr}S_B$  来度量类间离散程度

$$\text{tr}S_B = P_A [(\mu_{Ai} - \mu_{Ti})^2 + (\mu_{Aj} - \mu_{Tj})^2] + P_B [(\mu_{Bi} - \mu_{Ti})^2 + (\mu_{Bj} - \mu_{Tj})^2] \quad (11)$$

最优分割阈值  $(s^*, t^*)$  可定义为

$$(s^*, t^*) = \text{Arg} \max_{0 < s < L-1} \max_{0 < t < L-1} [\text{tr}S_B(s, t)] \quad (12)$$

2D-OTSU 兼顾到像素点与其邻域的灰度信息与空间信息, 像素灰度与其邻域均值灰度的相似度越高, 越可能为目标或者背景。统计发现, 嘴唇区域 2D 灰度直方图中概率  $p(i, j)$  不为 0 的点分布在 diagonal 附近, 而被平滑的边缘和噪声点远离主 diagonal。

由此采用一种基于区域约束的阈值搜索策略。如图 7 所示, 图中  $a, b$  为直线方程截距。2D 灰度直方图中主 diagonal 方程为  $g = f$ , 则散点分布在直线  $g = f + N/2$  和  $g = f - N/2$  之间。由这 2 条直线组成的窄带便是最佳阈值的约束边界,  $N$  是窄带宽度, 经大量试验测试, 取  $N = 30$ 。设分割阈值为  $(s, t)$ , 则将直线  $g = s + t - f$  作为分割准则, 直线左下侧为目标, 右上侧为背景。在计算类间离散矩阵的轨迹时将  $s + t$  整体作为阈值, 而不需要分别对  $s$  和  $t$  进行循环搜索, 最优阈值搜索空间由 2 维转化成了 1 维, 计算复杂度大大降低。

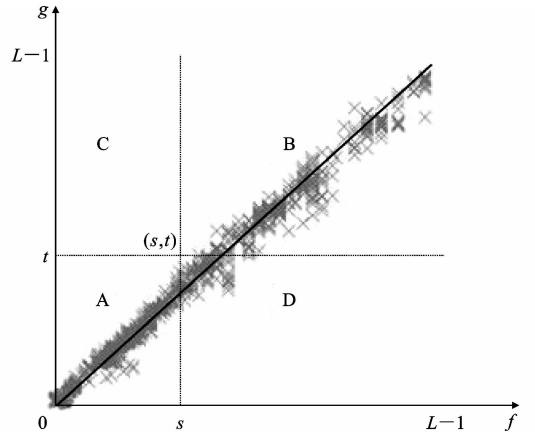


图 6 2D 灰度直方图

Fig. 6 2D gray histogram

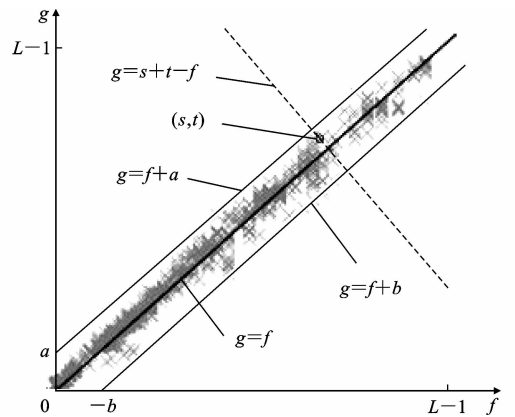


图 7 2D 分区策略

Fig. 7 2D partition strategy

### 3.2 嘴唇分割方法

强化唇色聚类性是 2D-OTSU 分割嘴唇的关键前提。本文利用唇色在 YCbCr 与 HSI 这 2 种色彩空间中的不同聚类特性, 提出一种不依赖样本训练的复合色度累加策略, 增强嘴唇与皮肤的可分性, 再运用 2D-OTSU 实现唇色分割, 步骤如下所述<sup>[18]</sup>。

Step1: 将嘴唇粗区域转换至 YCbCr 空间, 提取 Cb 分量。同时将嘴唇粗区域转换至 HSI 空间, 提

取  $H$  分量。见图 8(a)~(c)。

Step2: 分别对  $Cb$  分量与  $H$  分量进行灰度均衡拉伸, 结果为  $Cb_{\text{equ}}$  与  $H_{\text{equ}}$ , 如图 8(d)、(e) 所示<sup>[19-20]</sup>。  $Cb_{\text{equ}}$  分量凸显了嘴唇的高灰度特性, 而  $H_{\text{equ}}$  分量使得嘴唇的低灰度聚类特性更加显著, 周围肤色均体现出与嘴唇相反的灰度属性。

Step3: 对  $Cb_{\text{equ}}$  与  $H_{\text{equ}}$  分量进行色度累加, 算法见式(13)

$$G_{\text{ray}} = \lambda(I_{Cb_{\text{equ}}} - H_{\text{equ}}) + \xi \quad (13)$$

式中:  $G_{\text{ray}}$  为色度累加结果;  $I_{Cb_{\text{equ}}}$  为  $Cb_{\text{equ}}$  的计算参数;  $\lambda$  为线性增益因子;  $\xi$  为灰度补偿因子; 经试验统计  $\lambda$  取 1.3,  $\xi$  取 60。

该方法进一步增强嘴唇区域的高灰度聚类特性, 提高了唇色对面部运动与光照变化的鲁棒性, 结果如图 8(f) 所示。

Step4: 运用 2D-OTSU 对色度累加结果进行自动阈值分割, 再根据生理几何分布剔除掉上方鼻孔区域, 得到精确的嘴唇区域, 见图 8(g)~(i)。

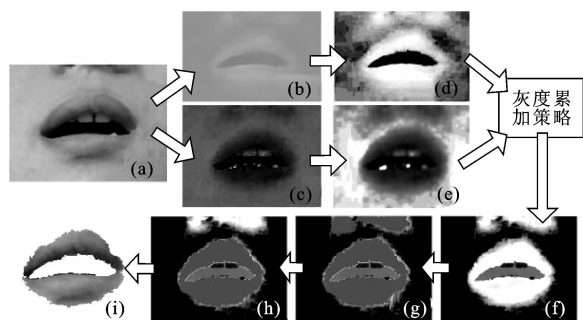


图 8 嘴唇分割流程

Fig. 8 Flow of lip segmentation

## 4 试验结果及分析

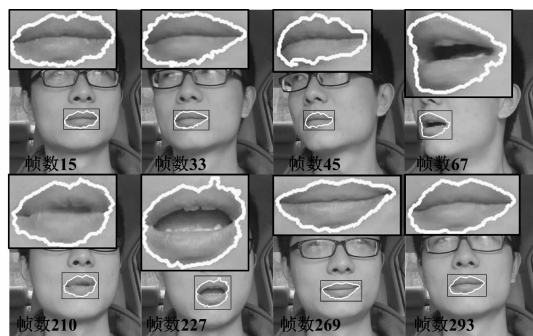
### 4.1 驾驶人图像数据库

在实际行驶环境下采集驾驶人面部图像。为确保交通安全, 试验在西安市南三环路辅道上进行, 车速不高于 30 km/h。CCD 摄像头安装在仪表盘内, 分辨率为  $640 \times 480$ , 帧率为 20 帧/s。算法运行环境为 Intel Core i7 3770K CPU, 3.5 GHz, 4 GB 内存。为确保图像光照条件的多样性, 监控视频在白天不同时间段分类采集, 分为 4 类: 强烈(13:00~14:00)、较强烈(15:00~16:00)、适中(10:00~11:00 与 16:00~17:00)、昏暗(傍晚)(18:00~19:00)。测试驾驶人共 10 名(7 男 3 女, 6 人戴眼镜), 头部姿态随机, 嘴唇运动包括闭嘴、说话、打哈欠。

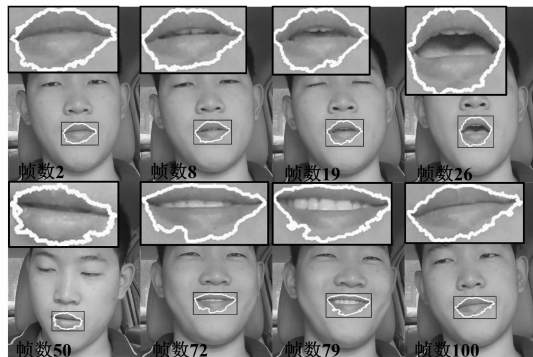
### 4.2 嘴唇识别试验

试验针对 10 名驾驶人(共计 32 536 帧图片)进

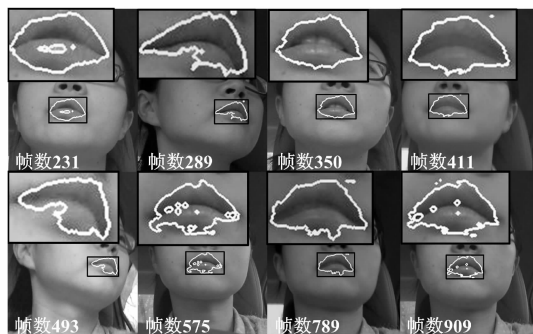
行嘴唇识别测试。图 9 为其中 3 名驾驶人的嘴唇检测结果。第 1、2 组视频的光照强度适中且比较稳定, 驾驶人存在复杂的头部运动和嘴部动作。在此条件下算法能够非常准确地分割嘴唇区域; 第 3 组视频中光照强度较弱且波动性较大, 而头部转动与嘴部运动较少。与前 2 组试验相比, 嘴唇轮廓边缘的平滑性与准确性较差, 但可以通过后续的形态学处理来提高嘴唇分割的准确性。



(a) 第1组



(b) 第2组



(c) 第3组

图 9 驾驶人嘴唇连续检测

Fig. 9 Continuous lip detection

试验在 4 类光照条件下分别选取 5 名驾驶人(各 1 000 幅图像)进行目标识别率测试, 结果如下页图 10 所示。Adaboost 算法的人脸平均识别率为 84.3%。由于本文提出的 Adaboost 结合  $Cb'-Cr'$  椭圆模型算法能够自适应地拟合肤色, 因此在各种头部姿态下的人脸平均识别率更高, 达到 90.9%。在

正确检测人脸的基础上,嘴唇平均识别率达到 91.7%,这是由于本文嘴唇分割算法有效增强了嘴唇与肤色的灰度类间距,对各种光照条件具有较高的自适应水平。

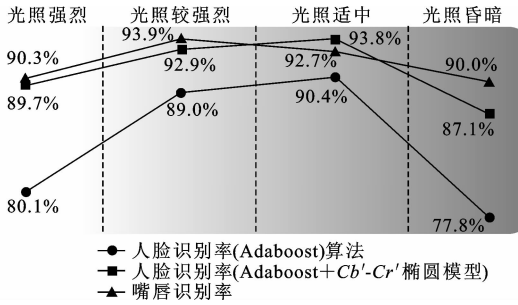


图 10 不同光照条件下的目标识别率

Fig. 10 Target recognition rate under different illuminations

嘴唇轮廓外接矩形(见图 9)的高宽比  $\delta$  在不同的动作下具有一定的变化规律,能够用于描述驾驶人闭嘴、说话、打哈欠等动作。如图 11 所示,某驾驶人随机的说话和打哈欠,黑色曲线为  $\delta$  变化规律,淡灰色虚线为连续累积 20 帧  $\delta$  的样本方差  $S^2$ 。当驾驶人打哈欠时, $\delta$  值明显高于“闭嘴”与“说话”状态下的  $\delta$  值,而且具有一定的延时特性,因此当  $\delta$  值持续大于特定阈值(图 11 中  $\delta$  阈值为 0.90)时,可认为驾驶人在打哈欠。在进一步区分驾驶人“闭嘴”与“说话”状态时,可通过  $\delta$  连续累积值的离散性进行判别。经过试验可知,“闭嘴”过程的  $S^2$  总是小于 0.01,而“说话”行为的  $S^2$  在 0.02~0.05 之间波动,因此可通过特定时间窗口内的  $S^2$  阈值来判别“闭嘴”和“说话”行为。

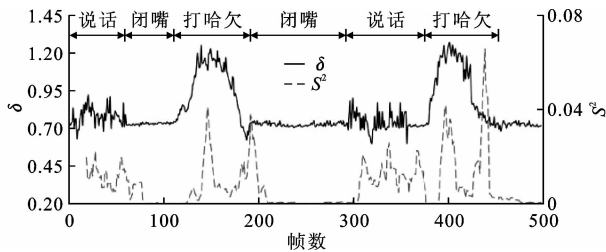


图 11 嘴唇动作检测

Fig. 11 Lip movement detection

### 4.3 嘴唇区域分割精度试验

试验以嘴角目标来检验算法的准确性,检测过程中要求驾驶人进行转头与说话动作。图 12 是实车试验数据库中 6 名驾驶人的嘴唇区域色度累加与 OTSU、2D-OTSU 的分割结果。由于  $Cb_{equ}$  与  $H_{equ}$  取自 2 类不同的色彩空间并均剔除了光照分量,因此该方法能够稳定地增强肤色与唇色的类间距。图 12 中的白色与黑色曲线分别为 2D-OTSU 与 OT-

SU 分割结果。对于驾驶人正面姿态且无嘴部运动的图像,2D-OTSU 与 OTSU 分割结果都比较准确。而存在头部和嘴部运动时,嘴角区域易产生阴影、模糊的边缘与扩张性的伪边界,此时 OTSU 更容易将这些干扰分割进来,扩大了嘴唇区域面积,而 2D-OTSU 对嘴唇轮廓的分割效果更为准确。

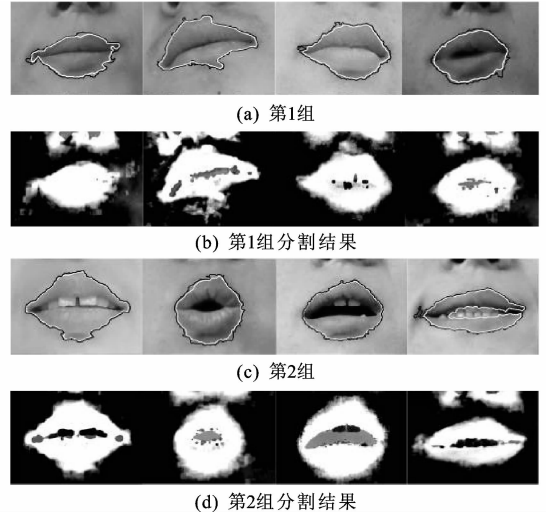


图 12 嘴唇分割示例

Fig. 12 Examples of lip segmentation

下页图 13 为光照稳定条件下的某驾驶人左嘴角坐标检测结果,粗实线曲线为人工手动标定获取的左嘴角实际坐标值,细实线曲线与虚线分别为 2D-OTSU 与 OTSU 检测的坐标值。在前 600 帧内驾驶人分别向左、向右转头,后 600 帧为说话行为。整体而言,OTSU 检测结果具有较大的波动性,在说话时的波动性更为显著,而 2D-OTSU 具有更高的检测精度。

试验选取“闭嘴”、“说话”和“向左转头”3 种动作的图像各 500 帧,分别运用 2D-OTSU 与 OTSU 算法分割嘴唇,统计出左嘴角的平均像素误差  $E_{ave}$ 、最大像素误差  $E_{max}$  和误差标准偏差  $\sigma$ ,结果如下页图 14 所示。由统计结果可知,2D-OTSU 的检测误差明显低于 OTSU 算法,对头部转动与嘴部动作的适应性更强。

### 4.4 算法实时性试验

试验针对各种光照条件的 10 名驾驶人(各 1 000 帧图像,共计 10 000 帧)进行算法耗时测试,统计结果见下页表 1。在所有子算法中, $Cb'-Cr'$  椭圆肤色模型需要进行平面维度的数据统计与边缘拟合,因此计算量较大。由于本文利用了稀疏网格进行肤色的间断采样,所以处理数据量降低了约 68%。另外  $Cb'-Cr'$  椭圆模型并不需要针对每 1 帧实时更新,仅需要在肤色像素出现较大漏检或误检

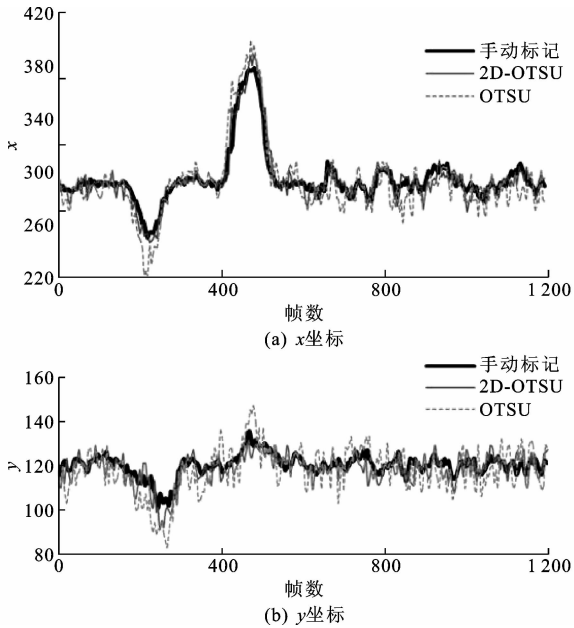


图13 左嘴角坐标检测

Fig. 13 Coordinate value detection of the left lip corner

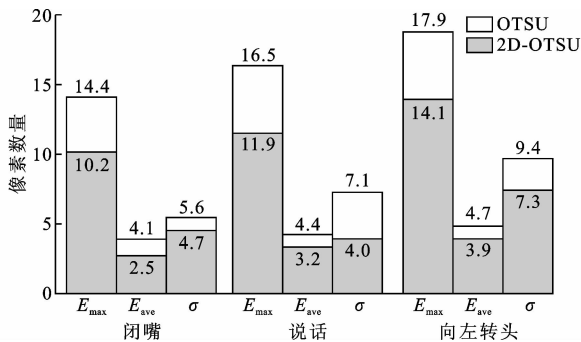


图14 左嘴角坐标误差统计

Fig. 14 Error statistics of the left lip corner coordinate

时运用 Adaboost 检测人脸并重新建立肤色模型,因此对系统时效性的影响很小。2D-OTSU 的数据计算量也很大,由于本文将最优阈值的搜索空间进行了降维,同时需要处理的嘴唇粗区域大小仅为  $55 \times 30$  像素,因此 2D-OTSU 表现出较好的运算效率。本文提出的算法能够实现高于 4 帧/s 的处理速度。若植入具有较高硬件配置的车载专用系统,算法的实时性可进一步提高,能够满足 DLR 的实时监测。

表1 算法耗时统计

Tab. 1 Operation time statistics of algorithms ms

子算法	①	②	③*	④	⑤	总耗时
最大耗时	52.2	53.7	75.8	26.5	84.0	227.8
最小耗时	41.8	43.4	65.9	21.2	70.1	185.0
平均耗时	46.6	51.2	72.0	24.7	76.4	201.3

注:①BF-SSR 光照均衡;②Adaboost 人脸检测;③基于  $Cb'-Cr'$  椭圆模型的肤色分割;④ $Cb_{equ}$  与  $H_{equ}$  灰度累加;⑤2D-OTSU; \* 表示当子算法②失效时启动子算法③。

## 5 结 语

(1)提出一种用于驾驶人嘴部行为监测的嘴唇检测方法,主要创新点有:①在 BF-SSR 光照均衡化的基础上,建立了 Adaboost- $YCb'Cr'$  自适应肤色模型来识别驾驶人脸,能有效适应复杂的光照变化与头部姿态,同时确保了算法的实时性;②提出基于  $YCbCr$  与 HSI 色彩空间的复合色度累加策略,有效增强唇色与肤色的灰度类间距离,再运用降维的 2D-OTSU 准则分割嘴唇。该方法对具有动作和阴影的嘴唇区域分割更为准确。

(2)经试验验证,驾驶人脸的识别率为 90.9%,嘴唇识别率为 91.7%,平均耗时 201.3 ms。算法对于白天环境下的非均匀光照、头部姿态与嘴部动作等非约束条件具有较好的鲁棒性,可应用于驾驶人嘴部行为监测与精神状态估计。

(3)本文方法适用于白天不同光照强度的驾驶环境,而并不适用于夜间的红外图像。下一步将深入开展针对红外单通道图像的 DLR 研究,以提升 DLR 的全天候适用性。另外,驾驶人嘴部动作与精神状态的映射模型、与车辆运行状态耦合的嘴部行为预警策略是提升 ADAS 主动安全水平的重要内容,值得进一步深入研究。

## 参考文献:

### References:

- [1] FASEL I, FORTENBERRY B, MOVELLAN J. A generative frame-work for real time object detection and classification[J]. Computer Vision and Image Understanding, 2005, 98(1):182-210.
- [2] KAKUMANU P, MAKROGIANNIS S, BOURBAKIS N. A survey of skin-color modeling and detection methods[J]. Pattern Recognition, 2007, 40(3):1106-1122.
- [3] 宋怀波,齐关锋,钱程.基于 YUV 颜色空间的脸部区域特征点定位方法[J].吉林大学学报:工学版, 2013, 43(增1):39-42.  
SONG Huai-bo, QI Guan-feng, QIAN Cheng. Human face feature points detection method based on YUV color space[J]. Journal of Jilin University: Engineering and Technology Edition, 2013, 43(S1):39-42. (in Chinese)
- [4] CHEUNG Y M, LIU X, YOU X G. A local region based approach to lip tracking[J]. Pattern Recognition, 2012, 45(9):3336-3347.
- [5] 姚鸿勋,吕雅娟,高文.基于色度分析的唇动特征

- 提取与识别[J]. 电子学报, 2002, 30(2):168-172.
- YAO Hong-xun, LU Ya-juan, GAO Wen. Lip-movement features extraction and recognition based on chroma analysis[J]. Acta Electronica Sinica, 2002, 30(2):168-172. (in Chinese)
- [6] 王晓平, 郝玉峰, 付德刚, 等. 一种自动的唇部定位及唇轮廓提取、跟踪方法[J]. 模式识别与人工智能, 2007, 20(4):485-491.
- WANG Xiao-ping, HAO Yu-feng, FU De-gang, et al. An automatic approach to lip localization, contour extraction and tracking[J]. Pattern Recognition and Artificial Intelligence, 2007, 20(4):485-491. (in Chinese)
- [7] 奉小慧, 贺前华, 王伟凝, 等. 基于 PS-Level Set 的嘴唇几何形状定位模型[J]. 华南理工大学学报: 自然科学版, 2010, 38(2):121-125.
- FENG Xiao-hui, HE Qian-hua, WANG Wei-ning, et al. Detection model of geometric shape of lip based on PS-Level Set[J]. Journal of South China University of Technology: Natural Science Edition, 2010, 38(2):121-125. (in Chinese)
- [8] 盛敬. 基于机器视觉的驾驶员嘴唇特征点定位算法[J]. 交通运输工程学报, 2009, 9(4):118-121.
- SHENG Jing. Locating algorithm of driver lip feature points based on machine vision[J]. Journal of Traffic and Transportation Engineering, 2009, 9(4):118-121. (in Chinese)
- [9] KIMMEL R, ELAD M, SHAKED D, et al. A variational frame-work for Retinex[J]. International Journal of Computer Vision, 2003, 52(1):7-23.
- [10] TOMASI C, MANDUCHI R. Bilateral filtering for gray and color images[C]//IEEE. Proceedings of the Sixth International Conference on Computer Vision. Bombay: IEEE, 1998:839-846.
- [11] KRYSTEK M. A fast Gauss filtering algorithm for roughness measurements[J]. Precision Engineering, 1996, 19(2/3):198-200.
- [12] 孙宏琦, 施维颖, 巨永锋. 利用中值滤波进行图像处理[J]. 长安大学学报: 自然科学版, 2003, 23(2):104-106.
- SUN Hong-qi, SHI Wei-ying, JU Yong-feng. Image processing with medium value filter[J]. Journal of Chang'an University: Natural Science Edition, 2003, 23(2):104-106. (in Chinese)
- [13] HSU R L, ABDEL-MOTTALEB M, JAIN A K. Face detection in color images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5):696-706.
- [14] LI L, FENG Z, HE K. A randomized algorithm for detecting multiple ellipses based on least square approach[J]. Opto-Electronics Review, 2005, 13(1):61-67.
- [15] OTSU N. A threshold selection method from gray-level histogram[J]. IEEE Transaction on Systems, Man and Cybernetics, 1979, 9(1):62-66.
- [16] GONG J, LI L Y, CHEN W N. Fast recursive algorithm for two-dimensional thresholding[J]. Pattern Recognition, 1998, 31(3):295-300.
- [17] 郝颖明, 朱枫. 2维 Otsu 自适应阈值的快速算法[J]. 中国图象图形学报: A 辑, 2005, 10(4):484-488.
- HAO Ying-ming, ZHU Feng. Fast algorithm for two-dimensional Otsu adaptive threshold algorithm[J]. Journal of Image and Graphics, 2005, 10(4):484-488. (in Chinese)
- [18] CHAVES-GONZÁLEZ J M, VEGA-RODRÍGUEZ M A, GÓMEZ-PULIDO J A, et al. Detecting skin in face recognition systems: a colour spaces study[J]. Digital Signal Processing, 2010, 20(3):806-823.
- [19] 丁爱玲, 周秦武, 郑春红. 基于直方图均衡与小波变换的超声图像增强[J]. 长安大学学报: 自然科学版, 2004, 24(6):84-87.
- DING Ai-ling, ZHOU Qin-wu, ZHENG Chun-hong. Ultrasound image contrast enhancement based on histogram equalization and wavelet transformation[J]. Journal of Chang'an University: Natural Science Edition, 2004, 24(6):84-87. (in Chinese)
- [20] 孙文荣, 周先春, 嵇亚婷. 基于直方图均衡化、PCA 和 SVM 算法的人脸识别[J]. 软件, 2014, 35(8):11-15.
- SUN Wen-rong, ZHOU Xian-chun, JI Ya-ting. Face recognition based on histogram equalization, principal components analysis and support vector machine algorithms[J]. Software, 2014, 35(8):11-15. (in Chinese)