

基于深度信息的弱光条件下人脸检测

廖广军^{1,2a}, 李致富^{2b,2c}, 刘 屹^{2a,2b}, 胡跃明^{2a,2b}

(1. 广东警官学院 刑事技术系, 广州 510320; 2. 华南理工大学 a. 自动化科学与工程学院,
b. 精密电子制造装备教育部工程研究中心, c. 机械与汽车工程学院, 广州 510640)

摘要: 为了解决弱光环境下人脸检测问题, 研究基于深度信息的人脸快速定位方法. 首先, 使用2D Chamfer Match方法实现人脸的快速预定位, 并对定位人脸进行区域增长和增强运算; 然后, 使用AdaBoost算法结合扩展的Harr特征训练出弱光条件下人脸检测的强分类器, 以实现准确的人脸定位. 实验表明, 在弱光条件下, 基于深度信息实现人脸定位和过滤, 可以减小搜索范围, 加快检测速度, 具有较强的鲁棒性和时效性.

关键词: 弱光条件; 人脸检测; 2D Chamfer Match; 扩展Harr特征; AdaBoost算法

中图分类号: TP391

文献标志码: A

Human face detection under weak light based on depth information

LIAO Guang-jun^{1,2a}, LI Zhi-fu^{2a,2c}, LIU Yu^{2a,2b}, HU Yue-ming^{2a,2b}

(1. Faculty of Forensic Science and Technology, Guangdong Police College, Guangzhou 510320, China; 2a. School of Automation Science and Engineering, 2b. Engineering Research Center for Precision Electronic Manufacturing Equipments of the Ministry of Education, 2c. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China. Correspondent: LIAO Guang-jun, E-mail: gdliaogi@gmail.com)

Abstract: In order to solve the problem of human face detection under weak light, a quick locating method based on the depth information by Kinect device is proposed. Firstly, 2D Chamfer Match method is used to rapidly preliminary locate the human face, and region-growing algorithm is applied on the preliminary locating face. Then, combined with extended Harr features, the AdaBoost algorithm is used to train the strong classifier under weak light, which realizes the precise face detection. The experiments show that rationally utilizing the depth information to detect the human face under the normal light can reduce the scope of the search window, which accelerates the detection speed, and the proposed method has strong robustness and timeliness.

Key words: weak light; face detection; 2D Chamfer Match; extended Harr feature; AdaBoost algorithm

0 引言

人脸检测是人脸识别、姿态估计、表情识别等问题研究的基础, 已受到许多科研工作者的广泛关注. Viola等^[1]提出了基于AdaBoost算法的实时人脸检测方法, 即将弱分类器组合成为强分类器, 并对训练出来的强分类器级联组合, 较好地解决了彩色图像空间中的人脸检测问题. 后续的研究有Real AdaBoost^[2]和LogitBoost^[3]等改进算法的研究. 文献[4]利用场景中的先验知识, 诸如运动跟踪、肤色检测等方法实现了人脸检测的窗口搜索和预估, 以节省全局搜索时间; Hu等^[5]提出使用基于肤色和表情的融合特征实

现人脸检测, 在光照变化的情况下具有一定稳定性; Zhang等^[6]使用小波分析处理人脸检测的光照影响; Yun等^[7]通过算法实现光照补偿, 降低光照对人脸检测的影响; Katsimerou等^[8]提出将颜色约束应用于创建的稳定肤色模型, 以实现人脸检测. 上述研究大多针对彩色图像中有一定的光照条件或光照变化并不复杂的环境, 对于特别复杂的光照环境, 特别是极端弱光条件下, 相关研究相对较少. 文献[9]提出利用马尔科夫随机场的原理来训练人脸区域的决策模型, 文献[10]通过对人脸区域进行关键点的提取来描述人脸特征, 能在一定程度上应付复杂场景下的人脸检测,

收稿日期: 2013-06-19; 修回日期: 2013-09-19.

基金项目: 国家自然科学基金项目(61203060); 公安部应用创新课题项目(2012YYCXGDST081); 广东省战略新兴产业项目(2010A081002007); 中央高校基本科研业务费专项资金项目(2013ZM0098, 2013ZG010); 广东省省部产学研合作专项资金项目.

作者简介: 廖广军(1981—), 男, 博士生, 从事模式识别的研究; 胡跃明(1960—), 男, 教授, 博士生导师, 从事非线性控制、智能控制等研究.

但仍存在改进空间。

针对上述问题,同时考虑到信息采集设备成本及民用化需求,本文采用 Kinect 设备,提出不依赖人脸区域肤色信息,而是通过对人脸深度信息的特征分析实现人脸定位.在算法上,本文首先使用 2D Chamfer Match^[11]在人脸深度图中全局搜索出人脸区域的候选位置;然后通过对该区域深度信息的增强处理,并结合扩展 Haar 特征训练出适合弱光下的人脸检测强分类器,进而完成人脸检测^[12].

1 2D Chamfer Match

弱光条件下,彩色图像信息常常包含大量的噪声,不能真实反映出场景信息,而深度信息在弱光条件下的表现相对稳定.本文主要考虑利用场景的深度信息实现弱光条件下的人脸检测.本文借鉴文献[13]在人体检测方面的处理方法,首先使用经典的 Canny 算法实现深度信息的边缘提取,并用滤波算法过滤边缘较小的区域,以剔除人脸区域的干扰信息;然后使用 2D Chamfer Match 算法定位人脸候选区域。

鉴于 2D Chamfer Match 在处理模版匹配中具有很好的效果,本文首先使用经典的 Canny 算法实现深度信息的边缘提取,并过滤边缘较小的区域以剔除人脸区域的干扰信息;然后使用 Chamfer Match 算法定位人脸的候选区域.图 1 为 Chamfer Match 人脸定位过程,采集设备为 Kinect.

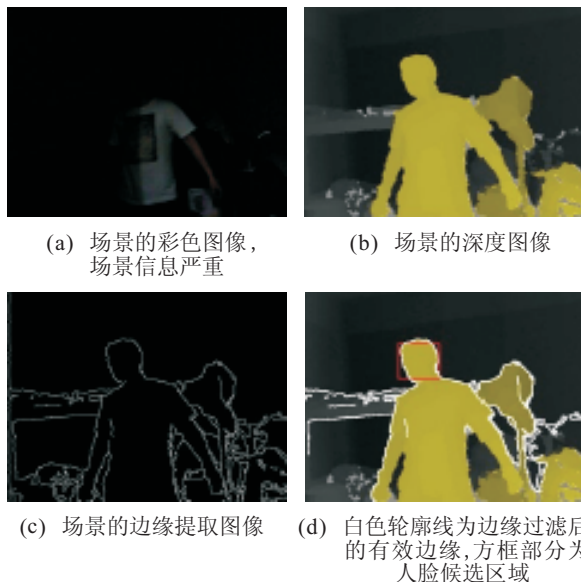







图 1 Chamfer Match 人脸定位过程

受透视规律的影响,不同深度条件下人脸区域的半径会有所变化.本文通过人脸深度信息的统计分析来确定人脸区域的提取半径.由文献[14]可知, Kinect 感应的有效范围约为 1.2~3.6 m.在此范围内,通过对人脸区域深度信息的统计分析确定人脸区域的提取半径.表 1 为部分人脸区域深度信息与人脸区域半径的实验数据.

表 1 部分人脸区域深度和人脸区域半径实验数据

实验图像数据	人脸区域半径/像素	人脸区域平均深度/m
	34	3.616 4
	36	3.238 4
	38	2.940 5
	51	2.206 4
	75	1.441 8

利用最小二乘法对人脸区域半径和人脸区域平均深度进行拟合,得到如下关系式:

$$y = 8.2830x^2 - 61.1502x + 146.9780. \quad (1)$$

其中: y 表示人脸区域半径的大小,以像素为单位; x 表示人脸区域所处的深度,以 m 为单位,其拟合结果如图 2 所示.

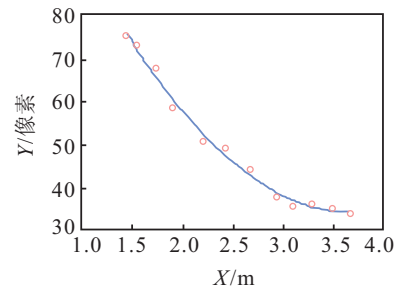


图 2 人脸区域深度与人脸半径的拟合数据

2 深度图增强

与传统的利用 ToF (time of flight) 技术获取深度信息的设备相比, Kinect 设备具有分辨率高、成本低、速度快等优点.但是,在弱光条件下,由于遮挡和红外光被吸收的原因, Kinect 获取的深度信息会出现大量的空白区域,这些区域对后续基于深度信息的人



图 3 Kinect 深度信息和彩色信息及校正

脸检测将造成巨大的困扰. 图3为人脸深度信息与彩色信息的对应情况, 在图3(c)中人脸鼻子部分的深度信息有明显的缺失, 在图像上形成空洞.

为了降低缺失深度信息对人脸检测的影响, 本文采用一种改进的区域增长算法来实现深度信息图像的增强. 实验表明, 场景中彩色图像中同一颜色的物体在深度上也是相近的. 为此, 本文结合场景中的彩色信息实现深度信息中空洞部分的估计和平滑. 算法描述如下.

记 $D_{i,j}$ 为当前深度图中空洞的像素点 $D(i,j)$, 其深度信息为

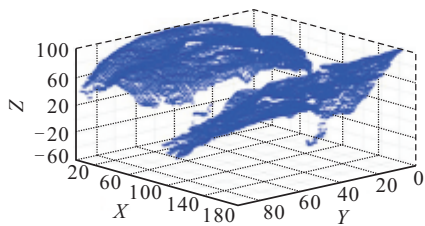
$$D_{i,j} = \sum_{p=i-2}^{i+2} \sum_{q=j-2}^{j+2} w_{p,q} D_{p,q}. \quad (2)$$

其中: $D_{p,q}$ 为深度图中点 $D(p,q)$ 的深度; $w_{p,q}$ 为当前深度图中点 $D(p,q)$ 对点 $D(i,j)$ 深度的贡献值, 越大表示贡献越大, 反之越小; $w_{p,q}$ 由下式决定:

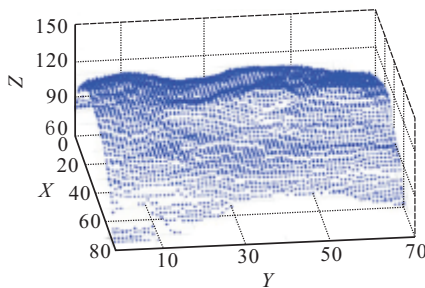
$$w_{p,q} = 1 - \frac{|C_{i,j} - C_{p,q}|}{\sum_{p=i-2}^{i+2} \sum_{q=j-2}^{j+2} |C_{i,j} - C_{m,n}|}, \quad (3)$$

$C_{i,j}$ 为深度图中点 $D(i,j)$ 在彩色图中的像素值, $C_{p,q}$ 为深度图中点 $D(p,q)$ 在彩色图中的像素值.

使用上述算法对图中的深度信息缺失部分进行恢复, 并对深度图使用区域增长算法, 以排除其他部分对人脸区域造成的干扰, 处理结果如图4所示.



(a) 人脸数据



(b) 处理后的人脸数据

图4 人脸三维信息恢复图

由图4可以看出, 填补了图中的空洞部分, 能够很好地恢复人脸的轮廓信息, 为深度信息中的人脸检测强分类器提供了准确的样本. 为了简化计算流程, 提高系统在实时性方面的表现, 本文使用HSV空间中的色度空间 H 来完成上述运算. 需要说明的是, 在极端的弱光下, 人脸区域的彩色信息几乎完全丢失,

深度信息的增强及平滑, 可以认为周围灰度都是一样的, 此时深度信息增强便退化成简单的均值滤波.

3 人脸深度过滤

通过对深度区域信息的分析, 能够判断当前检测的人脸是否是真实的人脸而非照片人脸. 具体实现方法如下: 记 $d(x,y)$ 为图像中像素点 $p(x,y)$ 的深度信息, 则人脸区域的平均深度 Avg_d 可以表示为

$$\text{Avg}_d = \frac{1}{n} \sum_1^n \{d(x,y) | d(x,y) \neq 0, p(x,y) \in \text{Area}_{\text{face}}\}; \quad (4)$$

深度信息的方差 Vari_d 表示为

$$\text{Vari}_d = \frac{1}{n} \sum_1^n \{(d(x,y) - \text{Avg}_d)^2 | d(x,y) \neq 0, p(x,y) \in \text{Area}_{\text{face}}\}. \quad (5)$$

最后, 通过对 Vari_d 值的判断来确认是否已通过深度信息过滤.

4 扩展 Haar 特征

传统的 Haar 特征表征的对象是灰度图像中的颜色变化, 在弱光条件下并不适合. 考虑到人脸具有丰富的深度信息变化, 比如人脸的鼻子部分比眼睛部分高, 嘴唇的位置比周围高等, 这些人脸的生理信息特征相对于传统的 Haar 特征所使用的彩色空间颜色深浅变化更为稳定, 也更能准确地描述人脸区域的特征. 因此, 为了适应弱光环境下的人脸检测, 本文采用一种扩展到3D的 Haar 特征, 以准确地表示出深度信息上的人脸变化特征^[15]. 扩展 Haar 特征充分利用了人脸区域在深度信息上的变化, 在弱光环境中具有稳定性, 能为人脸检测分类器的训练提供稳定特征.

5 AdaBoost 算法

AdaBoost 是一种通过迭代的方式实现分类器训练的算法, 其核心思想是通过训练集训练出的弱分类器进行组合, 最终形成强分类器. 本文使用扩展 Haar 特征作为 AdaBoost 训练中的弱分类器, 通过如下方式完成最终强分类器的训练.

给出训练集 $(x_1, y_1), \dots, (x_m, y_m)$. 其中: $x_i \in X$, $y_i \in Y = \{-1, +1\}$, 这里 X 代表特征空间, Y 代表目标空间. 针对文中的人脸检测问题, 用 -1 代表非人脸区域, 1 代表人脸区域. 对于 $i = 1, 2, \dots, m$ 个弱分类器, 均匀化初始权值分布 $D_1(i) = 1/m$, $D_1(i)$ 表示第1次迭代中第 i 个特征的权值, 进行 $t = 1, 2, \dots, T$ 次迭代.

记 H 为所有弱分类器的集合, 根据其权值 D_t 找出本次迭代过程中最大化阈值的弱分类器为

$$h_t = \arg \max_{h_t \in H} |0.5 - \phi|, \quad (6)$$

其中 $\phi = \sum_{i=1}^m D_t I(y_i \neq h_t(x_i))$.

如果 $|0.5 - \phi| \leq \beta$, β 为预先设置的阈值, 则停止迭代, 输出强分类器. 选择一个实数 $\alpha_t \in R$, 通过下式计算确定:

$$\alpha_t = \frac{1}{2} \ln \frac{1 - \phi}{\phi}. \quad (7)$$

根据下式更新所有权值分布, 继续迭代直至停止:

$$D_{t+1}(i) = \frac{D_t \exp(-\alpha_t y_i h_t(x_i))}{\sum_i D_t \exp(-\alpha_t y_i h_t(x_i))}. \quad (8)$$

最后输出强分类器组合, 即

$$H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right). \quad (9)$$

本文共使用了 500 张弱光下人脸深度信息分布图作为正样本, 1000 张非人脸区域的深度信息分布图作为负样本, 以实现弱光下人脸检测分类器的训练.

6 实验

本文使用 Kinect 采集场景的深度信息来完成人脸区域的定位, 很好地解决了弱光条件下人脸检测所遇到的困难. 研究中, 设计了 4 组对比实验, 以验证算法的稳定性和有效性. 第 1 组实验为光照充分且较为稳定的情况, 常用人脸检测算法与本文检测算法进行对比; 第 2 组实验为较弱光环境下, 本文算法与其他人脸检测算法的对比研究; 第 3 组实验为在极弱光条件下进行人脸检测, 以验证本文算法对极弱光环境的适应能力; 第 4 组实验使用红外图像作为常用的人脸检测算法的输入, 与所研究的方法进行性能比较.

为了使数据能更客观地体现出各种检测算法的性能, 在实验中统计了 3 种检测数据: 1) 正确检测数, 即出现人脸图像帧, 且检测到的次数; 2) 误检数, 即未出现人脸图像帧, 却检测出的次数; 3) 漏检数, 即出现

人脸图像帧, 但未被检测出的次数. 实验所采用的比较方法分别为文献 [9] 和文献 [10] 所描述的人脸检测算法. 图 5 为 Kinect 在不同光照条件下采集的视频序列数据.

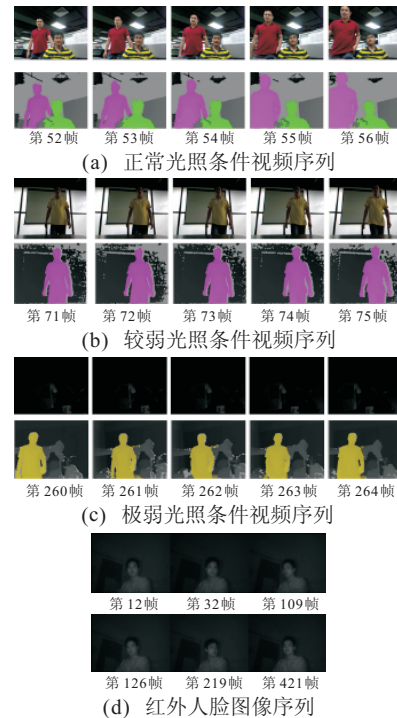


图 5 各种光照条件下 Kinect 采集的视频序列数据

第 1 组实验视频序列如图 5(a) 所示, 实验统计数据如表 2 所示. 从表 2 可以看出, 与传统的人脸检测算法相比, 本文所提出的算法在正确数上并不具备优势, 但在误检数方面优势明显. 这是因为通过深度过滤, 滤除了部分深度信息不明显的人脸 (如照片人脸), 从而降低了错误总数.

表 2 正常光照下人脸检测对比实验结果

人脸检测方法	正确数	误检数 (非人脸报告为人脸)	漏检数 (人脸报告为非人脸)	错误总数
Skin + Markov Random Fields	268/300	45/200	32/300	77/500
Interest Point Descriptors	278/300	30/200	22/300	52/500
Ours	272/300	12/200	28/300	40/500

表 3 较弱光照下人脸检测对比实验结果

人脸检测方法	正确数	误检数 (非人脸报告为人脸)	漏检数 (人脸报告为非人脸)	错误总数
Skin + Markov Random Fields	189/312	45/198	123/312	168/500
Interest Point Descriptors	240/312	70/198	72/312	142/500
Ours	287/312	17/198	25/312	42/500

表 4 极弱光照下人脸检测对比实验结果

人脸检测方法	正确数	误检数 (非人脸报告为人脸)	漏检数 (人脸报告为非人脸)	错误总数
Skin + Markov Random Fields	19/200	38/300	181/200	219/500
Interest Point Descriptors	123/200	54/300	77/200	131/500
Ours	145/200	21/300	55/200	76/500

表 5 红外图像光照下人脸检测对比实验结果

人脸检测方法	正确数	误检数 (非人脸报告为人脸)	漏检数 (人脸报告为非人脸)	错误总数
Skin + Markov Random Fields	163/237	38/263	74/237	112/500
Interest Point Descriptors	178/237	54/263	59/237	113/500
Ours	211/237	21/300	26/237	47/500

第 2 组实验视频序列如图 5(b) 所示, 实验统计数据如表 3 所示. 第 2 组实验在较弱光照条件下进行, 这种条件下, 环境光照度偏弱, 场景中的彩色图像信息不够完整, 但依然保留了部分彩色信息. 从表 3 可以看出, 在较弱光照条件下, 传统的人脸检测方法已经很难胜任检测工作, 但本文所提出的方法在正确数和错误总数方面均维持在原有的水准, 能比较好地满足检测需求.

第 3 组实验视频序列如图 5(c) 所示, 实验统计数据如表 4 所示. 第 3 组实验在极弱光条件下进行, 这种条件下, 彩色图像基本上难以获得有效信息. 从表 4 可以看到, 在极弱光条件下, 本文提出的人脸检测算法在错误总数方面表现显著, 基本能够适应极弱光条件下的人脸检测.

第 4 组实验使用红外图像作为常用的人脸检测算法的输入, 与文中提到的基于深度信息的人脸检测算法进行性能比较, 实验使用红外摄像头采集到的一段时长 5 min 的视频, 从中随机抽取 500 帧红外图像, 其中包含人脸图像 237 帧, 不包含人脸的图像 263 帧. 实验统计数据如表 5 所示.

常见的人脸检测算法由于红外图像中缺乏图像的彩色信息, 前两种算法受到了较大的影响, 而本文所提出的算法在人脸正确数和错误总数方面均维持良好的效果, 算法具有较好的适应性.

7 结 论

基于深度信息的人脸检测方法在光照变化剧烈的情况下, 依然能够实现快速而稳定的人脸定位, 相比于利用颜色信息的人脸定位, 本文方法具有更强的鲁棒性. 本文的主要贡献在于以下几个方面: 1) 提出了一种基于深度信息的人脸快速定位算法; 2) 提出了一种融合彩色信息和深度信息的人脸区域增强算法; 3) 实现了在极弱光环境下人脸的快速定位.

本文提出的方法在面对人脸姿态变化, 特别是对于大角度侧脸以及物体遮挡时, 并不能很好地定位出场景中的人脸区域. 在未来的研究中将主要针对环境中人脸姿态、角度变化以及物体部分遮挡情况提出人脸检测的解决方案.

参考文献(References)

- [1] Viola P, Jones M J. Robust real-time face detection[J]. *Int J of Computer Vision*, 2004, 57(2): 137-54.
- [2] 余益民, 黄廷辉, 桑涛. 基于 Real AdaBoost 算法的肤色分割方法[J]. *计算机应用*, 2011, 31(12): 3370-3373.
(Yu Y M, Huang Y H, Sang T. Skin segmentation based on Real AdaBoost algorithm[J]. *J of Computer Applications*, 2011, 31(12): 3370-3373.)
- [3] Li P. Robust logitboost and adaptive base class(ABC) logitboost[C]. *Proc of the Conf on Uncertainty in Artificial Intelligence*. California: AUAI Press, 2010: 302-311.
- [4] 朱凯, 杨倩倩. 基于肤色模型和 CAMShift 算法的人脸检测与跟踪研究[J]. *现代计算机*, 2009(6): 60-63.
(Zhu K, Yang Q Q. Research on face detect and tracking based on skin-color model and CAMShift algorithm[J]. *Modern Computer*, 2009(6): 60-63.)
- [5] Hu W, Yang C, Huang D, et al. Feature-based face detection against skin-color like backgrounds with varying illumination[J]. *J of Information Hiding and Multimedia Signal Processing*, 2011, 2(2): 123-132.
- [6] Zhang T, Fang B, Yuan Y, et al. Multiscale facial structure representation for face recognition under varying illumination[J]. *Pattern Recognition*, 2009, 42(2): 251-258.
- [7] Yun J, Lee H, Paul A K, et al. Face detection for video summary using illumination-compensation and morphological processing[J]. *Pattern Recognition Letters*, 2009, 30(9): 856-860.
- [8] Katsimerou C, Redi J A, Heynderickx I. Face detection in intelligent ambiances with colored illumination[C]. *Int Symposium on Ambient Intelligence*. Salamanca: Springer, 2013: 195-204.
- [9] Maglogiannis I, Vouyioukas D, Aggelopoulos C. Face detection and recognition of natural human emotion using Markov random fields[J]. *Personal and Ubiquitous Computing*, 2009, 13(1): 95-101.
- [10] Stein S, Fink G A. A new method for combined face detection and identification using interest point descriptors[C]. *Proc of IEEE on Automatic Face & Gesture Recognition and Workshops*. Santa Barbara: IEEE Press, 2011: 519-524.
- [11] Thayananthan A, Stenger B, Torr P H, et al. Shape context and chamfer matching in cluttered scenes, in editor[C]. *Proc of IEEE on Computer Vision and Pattern Recognition*. Monona Terrace: IEEE Press, 2003, I: 127-133.
- [12] Jung F, Kirschner M, Wesarg S. Bildverarbeitung für die medizin[M]. Berlin: Springer, 2013: 320-325.
- [13] Xia L, Chen C, Aggarwal J. Human detection using depth information by kinect[C]. *Proc of IEEE on Computer Vision and Pattern Recognition Workshops*. Colorado Springs: IEEE Press, 2011: 15-22.
- [14] Leyvand T, Meekhof C, Wei Y, et al. Kinect identity: Technology and experience[J]. *Computer*, 2011, 44(4): 94-96.
- [15] Cui X, Liu Y, Shan S, et al. 3d haar-like features for pedestrian detection[C]. *Proc of IEEE on Multimedia and Expo*. Beijing: IEEE Press, 2007: 1263-1266.

(责任编辑: 曹洪武)