

# 三维深度图像在自动客流计数系统中的应用

尹章芹 顾国华 陈 钱 钱惟贤

(南京理工大学电光学院, 江苏 南京 210094)

**摘要** 在客流计数中,针对目前基于普通彩色相机的图像处理受光照、人体姿态、遮挡等因素影响使系统整体性能降低等问题,提出了一种基于人体头肩三维信息检测客流的新方法,即利用人体头部与肩部空间位置关系特征以及头部形状、面积等特征检测客流。以 Xtion 传感器为平台,通过综合分析传感器获取的包含人体表面三维特征的深度图像,识别并跟踪人体头部,统计进出的人数。实验结果表明,这种方法不仅提高了检测的精度和效率,减少了运算数据量,实现了视频流实时处理,而且解决了基于普通彩色图像中目标检测需要建立训练库等问题。系统精度达到 96%,具有实际的应用价值。

**关键词** 图像处理;客流;Xtion 传感器;三维特征;深度图

中图分类号 TN29 文献标识码 A doi: 10.3788/CJL201441.0609003

## Application of Three-Dimensional Depth Image in Automatic Statistic Passenger Flow System

Yin Zhangqin Gu Guohua Chen Qian Qian Weixian

(College of Electric & Optic, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China)

**Abstract** A new method of detecting passenger flow based on human head and shoulder three-dimensional information is proposed, detecting passenger flow according to spatial relationship features between head and shoulders as well as head shape, size and other characteristics, which can effectively solve system overall performance reduce problems owing to ordinary color-camera image processing being affected by light, posture, shelter and so on. With Xtion sensor as a platform, identifying and tracking human head, counting the passenger flow according to comprehensive analysis on depth image of body surface three-dimensional features. The experimental results show that this method not only improves the precision and efficiency of detection, reducing the amount of computation of data, real-time video stream can be processed, but also solves the need to establish training detection in color images. The system accuracy can be up to 96%, and has practical value.

**Key words** image processing; passenger flow; Xtion sensor; three-dimensional features; depth image

**OCIS codes** 100.2000; 100.6890; 110.2970; 140.6910

## 1 引 言

自动客流计数系统是用来检测和分析客流量的自动化系统,系统适用于旅游业、交通业和安防领域,可应用于商场、超市、公园景区、体育场馆、机场车站等公共场所的人流检测,系统为用户提供人流

数量、滞留人数、人流规律、人数趋势等信息,为管理者提供科学的决策依据,尤其在交通系统中,要求计数系统能精准、快速统计每个时间段、每路车、每个车站等的客流量,以便更好发挥智能调度系统,提高资源利用率。

收稿日期: 2013-12-02; 收到修改稿日期: 2014-01-10

基金项目: 江苏省“六大人才高峰”计划支持(2010-DZXX-022)、国家自然科学基金(61101199)、总装预先研究基金(40405050303)、教育部新世纪优秀人才(NCET-12-0630)

作者简介: 尹章芹(1988—),女,硕士研究生,主要从事图像处理、模式识别和三维构建等方面的研究。

E-mail: yinzhangqin\_lkfl@163.com

导师简介: 顾国华(1966—),男,教授,博士生导师,主要从事光电信息探测与实时图像处理等方面的研究。

E-mail: gghnjust@163.com

国内传统的客流计数系统主要有红外遮挡系统和压力传感系统,它们虽然易实现、成本低,但计数不准确,尤其在客流高峰期拥挤状况严重时,且应用场所受限。相对而言,图像信息的容量更大、更丰富,因此,图像处理技术的兴起和不断发展为传统客流计数技术中存在的一系列问题的解决提出了许多新方法。目前国内外主流的自动客流计数系统是基于图像处理的客流计数系统。文献[1-2]中基于图像处理的自动客流计数过程概括起来一般包括运动目标提取,特征识别,特征匹配和跟踪计数四个过程。基于彩色图像的自动客流计数主要是利用图像中的特征识别和模式匹配识别头部,算法数据量很大,只适用于背景简单的情况,且无法解决人体相互遮挡等问题,算法应用受到很多限制,应用不具备普遍性。基于双目立体视觉的自动客流计数系统<sup>[3]</sup>是通过双摄像头对输出的图像匹配计算运动目标的景深信息并进行客流头部识别,但系统算法涉及图像数据的自动立体匹配计算与头部识别跟踪等,算法量大,并且双摄像头需要进行复杂的标定,硬件成本

太高,不是目前市场上主流的客流计数系统。

针对目前基于图像处理的自动客流计数中存在的问题,设计了基于 Xtion 传感器三维深度图的自动客流计数系统。

## 2 三维深度图像

### 2.1 Xtion 传感器介绍

华硕公司推出的 Xtion 传感器是一种便携式、集成化的三维传感器,实物及原理如图 1 所示。COMS 红外摄像机接收红外发射器发出的近红外激光散斑编码,利用 Xtion 传感器内部集成的处理芯片 PS1080 处理接收到的红外光编码,生成准确的场景深度图像。由于 Xtion 传感器内部集成算法处理芯片,可以直接获取包含三维信息的数据,不需要考虑双目成像中的相机同步、立体匹配计算深度信息等问题,减少了算法量。Xtion 传感器通过 USB 端口将深度图输出,进一步简化了系统视频流的采集。

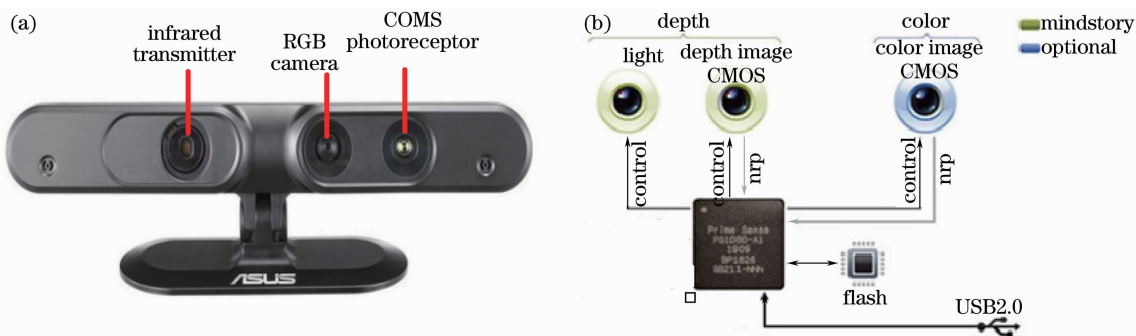


图 1 Xtion 体传感器和原理图。(a) Xtion 体传感器; (b) 原理图

Fig. 1 Xtion sensor and its principle diagram. (a) Xtion sensor; (b) schematic diagram

深度图像中的像素灰度值表示传感器到物体表面的距离。Xtion 传感器对距离的标定以 mm 为单位,(1)式表示传感器到物体表面的距离与灰度级的关系。

$$H = \alpha \times G, \quad (1)$$

式中  $H(300 < H < 1000)$  表示传感器到物体表面的距离, $G$  表示深度图灰度值,如果 COMS 红外摄像机没有接收到红外光编码,则对应的像素灰度值为 0,如果空间距离超出有效值范围,则对应的像素灰度值为 255, $\alpha$  是深度距离与灰度值转换系数。

### 2.2 深度图像的三维特征

采用文献[3-6]中提到的头顶正上方俯视的角度,通过 USB 端口 Xtion 传感器输出的深度图如图 2(a)所示。利用(2)式对深度图灰度值取反,结果如

图 2(b)。利用伪彩色算法显示取反后深度图的三维曲面,结果如图 2(c)。由于不是基于立体匹配计算深度信息,因此获取的三维信息精度高,数据完整。分析图 2(c)得出:头部灰度值是局部区域极大值,且头部彼此分离,头部是类椭圆形,连通域面积较大,头部和肩部数据是高内聚低耦合关系等。在客流统计中,利用完整的头部信息特征识别头部。同时,深度图像与空间颜色无关,只表示空间距离信息,降低了基于彩色图像自动客流统计中阴影、光照变化、环境变化等因素的干扰,提高头部识别精度,且适用于拥挤人流的检测。

$$\begin{bmatrix} H - H_1 \\ H - H_2 \\ H - H_3 \end{bmatrix} = \alpha \times \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix}, \quad (2)$$

系统自动识别统计身高 1 m 以上的人。 $H$  表示传感器到地面的距离, 距离地面  $H_1, H_2$  对应的灰度

值为  $G_1$  和  $G_2$ ,  $50 < H_1 < H_2 < 100$ ,  $230 < H_3 < H$ ,  $G_3$  表示距离地面  $H_3$  对应的灰度值。

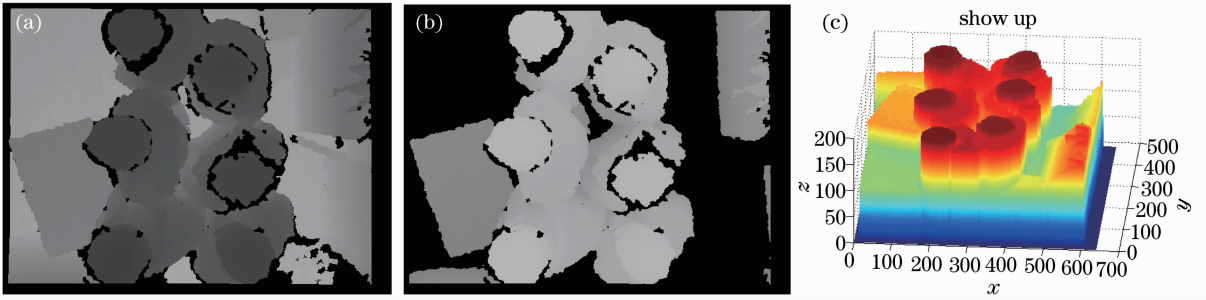


图 2 深度图和三维直方图。(a) 深度图; (b) 取反后的深度图; (c) 图 2(b) 的三维直方图

Fig. 2 Depth maps and the three dimensional histogram. (a) Depth map; (b) negated depth map; (c) three-dimensional histogram of fig. 2 (b)

### 3 客流统计系统

系统算法的设计思路概括为目标分割、目标识

别、目标跟踪、目标运动轨迹建立, 算法流程图如图 3 所示。

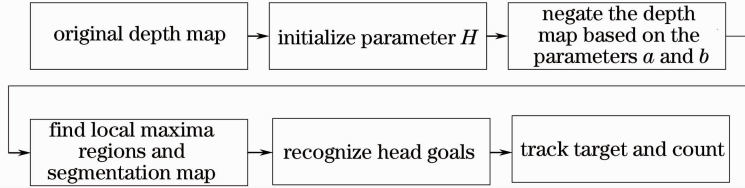


图 3 算法总体流程图

Fig. 3 Algorithm flow chart of the overall

#### 3.1 头部识别

头部识别包括目标分割与识别, 基于深度图中头部特征识别头部。利用九宫格算法寻找灰度值极大值区域, 公式如下:

$$\bar{G} > \bar{G}_k \quad k = \{1, 2, 3, \dots, 8\}. \quad (3)$$

深度图的宽高分别为  $W$  和  $H$ , 由于图像边缘可能出现极大值区域, 利用(4)式将深度图四周分别扩充  $A$  ( $20 < A < 70$ ), 得到新的宽高  $W'$  和  $H'$ 。基于九宫格算法原理, 图像分割算法流程如下:

1) 以  $A \text{ pixel} \times A \text{ pixel}$  为窗口, 将取反得到的深度图分割成  $\left(\frac{H'}{A}\right) \times \left(\frac{W'}{A}\right)$  个区域, 结果如图 4, 其中  $\frac{H'}{A}$ 、 $\frac{W'}{A}$  均为正整数; 遍历每个区域, 统计区域灰度值满足  $255 - G_2 < G_i(x, y) < 255 - G_3$  对应的像素个数, 计算每一区域灰度均值并确定区域最大灰度值  $G_{\max_i}$ , 其中  $G_i(x, y) \left[ 0 \leq i < \left(\frac{H'}{A}\right) \times \left(\frac{W'}{A}\right) \right]$  表示第  $i$  区域坐标  $(x, y)$  的像素灰度值;

2) 以  $3 \text{ pixel} \times 3 \text{ pixel}$  为窗口, 遍历区域, 利用(3)式找出满足中心区域灰度均值大于或等于其他 8 个区域的灰度均值对应的区域, 标记为  $N_i$  ( $0 \leq$

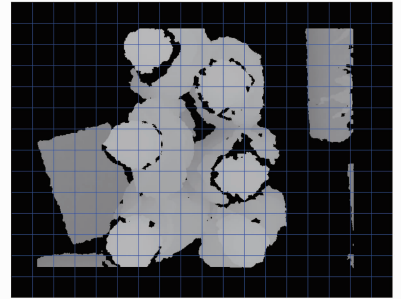


图 4 深度图分割结果

Fig. 4 Result of the depth map segmentation  $N_i (\leq 255)$ ,  $N_i$  表示第  $i$  区域的标识符;

3) 区域分割, 以  $N_i$  区域为中心, 分割区域增扩至  $3A \text{ pixel} \times 3A \text{ pixel}$ , 找出增扩后的区域中灰度值满足  $G_{\max_i} - G_{N_i}(x, y) < \lambda$  ( $4 < \lambda < 15$ ) 条件的所有像素坐标, 并标记为  $K_i$ , 共生成  $T$  个标记符, 如图 5;

4) 遍历  $T$  个灰度图, 将标识符区域生成  $T$  个二值图序列和对应的灰度图。遍历二值图序列和对应的灰度图, 基于以下特征识别人体头部<sup>[7]</sup>: 1) 类椭圆形; 2) 面积大于 1000; 3) 基于(5)式,  $h'$  与  $w'$  的取值小于 25 且大于 10。记录所有满足条件的区域起始坐标、宽高以及区域均值  $[(x, y), h, w, \bar{G}]$ , 即

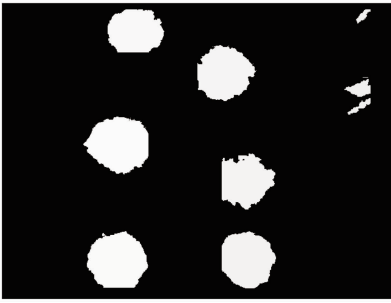


图 5 极大值区域结果

Fig. 5 Result of local maxima regions

头部信息。

$$\begin{bmatrix} H + 2 \times A \\ W + 2 \times A \end{bmatrix} = \begin{bmatrix} H' \\ W' \end{bmatrix}, \quad (4)$$

$$\beta \times (255.0 - \bar{G}) \times \begin{bmatrix} h \\ w \end{bmatrix} = \begin{bmatrix} h' \\ w' \end{bmatrix}, \quad (5)$$

式中  $h'$  和  $w'$  表示头部尺寸,  $h$  和  $w$  表示二值图中求出的区域宽和高,  $\beta$  是比例系数,  $\bar{G}$  表示取反后的区域灰度均值。

### 3.2 头部跟踪和计数

确定头部信息后,基于自适应波门<sup>[8]</sup>的多航迹管理,建立头部运动航迹,跟踪目标,统计进出的客流量。以头部中心坐标为点迹,波门形状为矩形,波门大小为头部的宽高值,波门的最小尺寸为 50,由点迹  $(x, y)$  与波门  $(\Delta x, \Delta y)$  确定以点迹预测值  $(\hat{x}, \hat{y})$  为中心的空间搜索区域,其关系如(6)式,其中点迹  $(x, y)$  与航迹相关。假定航迹,第  $K$  帧初次记录目标,基于第  $K + 1$  帧得到的坐标对目标的速度进行估计,如果估计的速度在取值范围内,则生成一条暂时航迹。其次,基于第  $K + 2$  帧目标的位置进行预测,并以预测位置为中心,确定一个关联区域,任何落在关联区域内的点迹扩展一条暂时的航迹,继续估计速度值,基于速度值对下一帧的位置进行预测并建立相关区域,落在相关区域内的点迹将生成一条新的航迹。对所有生成的航迹用二次曲线进行拟合,如果航迹上的点和拟合曲线的误差在一定范围内,则确定该航迹,如果不满足,则删除该航迹。确定目标运动航迹后,基于航迹确定目标的进出方向,在经过第三步分割处理后的灰度图的第  $L_1 = \frac{H \times 4}{7}$  行和第  $L_2 = \frac{H \times 5}{7}$  行画两条判断线,如果当前目标点迹  $x < L_1$  且航迹存在行坐标大于  $L_2$  的点迹,则进的人数加 1。反之如果当前目标点迹  $x > L_2$  且航迹存在行坐标小于  $L_1$  的点迹,则出的人数加 1。同时,删除该航迹。结果如图 6 所示,同时以文本记

录进出客流。

$$\begin{cases} |x - \hat{x}| \leq \Delta x \\ |y - \hat{y}| \leq \Delta y \end{cases} \quad (6)$$

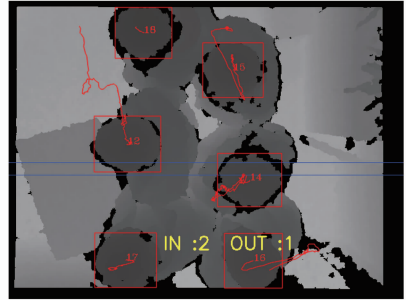


图 6 人头检测结果

Fig. 6 Result of heads' detection

## 4 实验分析

基于深度图的三维特征设计统计客流进出的算法,为了验证本系统算法的可行性、精准性和适用性,不仅搭建模拟平台,实时仿真,验证算法的有效性,而且还在长途汽车客运站的进出口大厅实时检测人流的进出情况,全天候统计客流量,并将客流的进出人数和时间以“time.txt”文件的形式记录下来,并与实际的客流信息进行比较。

### 4.1 不同环境的测试

为了测试系统程序稳定性、通用性与实用性,在客运站大厅、实验室模拟平台等不同环境下对系统进行实验,并针对可能的目标干扰信息的因素进行测试,测试环境如下:

#### 1) 客运站客流进出口处

客运站客流进出口处,干扰信息比较少,主要的干扰因素是小孩,因为系统自动识别统计身高 1 m 以上的行人。在客运站中,传感器距离地面的高度  $H$  为 3.7 m,设定距离地面距离  $H_1$  等于 0.5 m 为取反的距离,系统自动识别距离地面  $H_2$  等于 0.9 m 到  $H_3$  等于 2.3 m 之间的头部目标。系统参数取决于  $H$ ,  $H$  由地面灰度值直方图的最大值决定。因此,本系统是全自动化式的客流计数系统。图 7 是客运站一组客流统计的例子。第 353 帧是在视场内第二次识别到目标,并预测了目标航迹,用绿色框表示;第 354 帧是在视场内第三次识别到目标,确定运动目标航迹,确定 ID;第 370 帧是目标运动航迹接近进出判断线的边际,第 373 帧是目标航迹撤销后重新生成航迹,目标的 ID 加 1,出的人数加 1。视场内的儿童由于身高小于 0.9 m,因此,儿童不在目标识别的范围内。

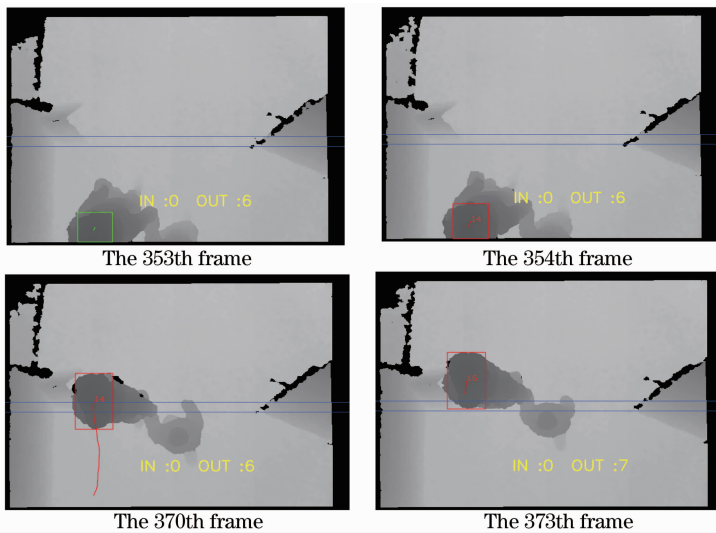


图7 长途汽车站的一组客流检测

Fig. 7 Figures of passenger flow detection at the coach station

### 2) 实验室的模拟进出口平台

实验室天花板距离地面距离  $H$  为 2.7 m, 其他参数由  $H$  对应的灰度值  $G$  可以计算得出。实验室的模拟平台主要是为了验证本系统适用于拥挤情况

下的客流统计, 图 8 是在实验室的模拟进出口平台下的客流检测。说明第 244 帧, 第 332 帧, 第 345 帧由于 Xtion 传感器的视场范围有限, 在俯视距离 2.7 m 的情况下, 视场范围人流拥挤。

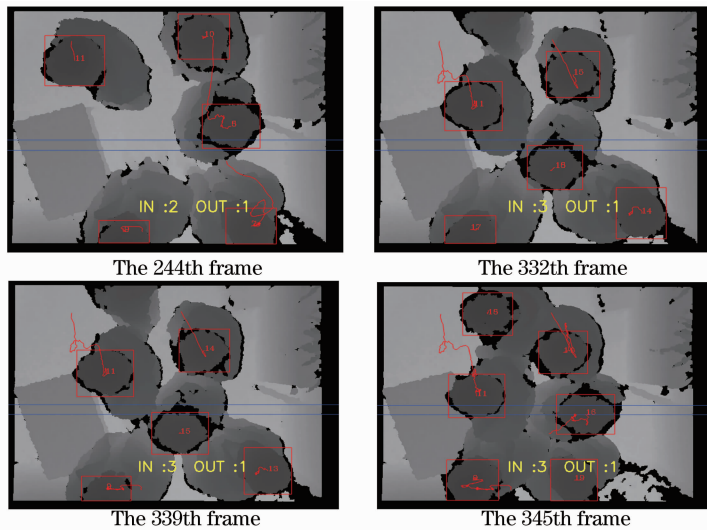


图8 在实验室的模拟进出口平台下的客流检测

Fig. 8 Figures of passenger flow detection at laboratory simulation platform

### 3) 实验室模拟台阶的平台

在实验室的门口设两级阶梯, 图 9 表示有台阶的平台情况下的客流检测。第 28 帧是人站在第一阶梯的情况, 头部在视场内第二次被识别到, 用绿色的框表示; 第 30 帧是站人在第二阶梯的情况, 确定目标, 航迹稳定; 第 34 帧头部目标航迹满足航迹撤销条件; 第 35 帧是航迹被撤销后重起航迹, 进方向的人数加 1。

### 4.2 精度与实时性测试

(7) 式为精度计算公式, 分析在不同平台下客流进出的精度<sup>[7]</sup>。其中  $C_{in}$ ,  $R_{in}$ ,  $C_{out}$  和  $R_{out}$  分别表示车站不同时刻统计与实际的进出方向的人数, 客运站安检进出口统计结果以及精度结果如表 1。同时统计每秒已处理的帧数, 即帧频, 以每 5 s 一个时间段, 利用(8)式计算帧频,  $N_T$  与  $N_{T+5}$  分别表示某一时刻以及其 5 s 后已处理的帧数, 帧频计算结果如表 2。

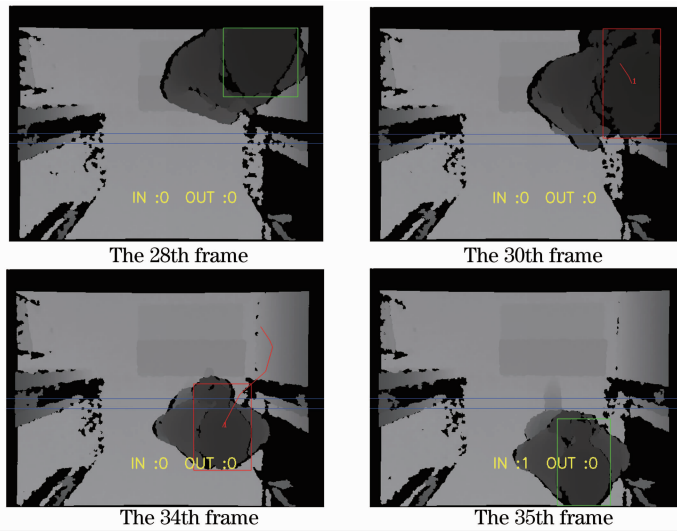


图 9 有台阶的平台情况下的客流检测

Fig. 9 Figures of passenger flow detection in the case of a stepped platform

$$\epsilon = \frac{1}{2} \times \left[ \frac{C_{in}}{R_{in}} \text{ or } \frac{R_{in}}{C_{in}} + \frac{C_{out}}{R_{out}} \text{ or } \frac{R_{out}}{C_{out}} \right] \times 100\%, \quad (7)$$

$$f_{ps} = \frac{N_{T+5} - N_T}{5}. \quad (8)$$

表 1 实际和统计的进出人数

Table 1 Practical and statistical number of in and out

Time	$C_{in}/C_{out}$	$R_{in}/R_{out}$	$\epsilon$
11:00—11:15	114/67	112/64	0.9688
11:20—11:35	94/48	90/46	0.9579
14:05—14:15	117/64	119/69	0.9553
15:00—15:15	80/44	84/42	0.9640
18:00—18:15	8/15	8/14	0.9666
Average			0.9625

表 2 实时处理的帧频

Table 2 Frame frequency of real-time processing

$N_T$	$N_{T+5}$	$f_{ps}$
1039	1188	29.8
2976	3125	29.8
5510	5659	29.8
6403	6555	30.4
7603	7753	30.0

表 1 是长途汽车站一天不同时刻的客流情况，从表中可以看出，中午与下午时间段的人流量比较大，晚上的人流量比较少，且由于传感器是主动发射空间光编码，系统不需要辅助光就可以获取夜间的人流深度信息。每 15 min 为一时间段，统计实际的进出人数，并且通过“time.txt”计算出每个时间段统计的人流量，求出系统精度在 96% 左右，且精度偏差小，说明系统客流统计稳定，并适应于夜间环境，不受光线影响。

表 2 验证系统的实时性，硬件系统标配为 Intel Core TM i3 CPU(3.3 GHz)、DDR-3.0 8GB 内存，USB2.0 输出的视频流帧频为 30 fps。系统处理视频流的同时，记录处理的帧号，以 5 s 为一时间段，计算这段时间处理的帧数，系统的处理帧频为 29.8 fps，近似于视频流采集的帧频，说明系统算法可以实现视频流实时处理。

本系统的硬件成本低，且解决了目前立体视觉中深度信息获取的问题，系统算法虽然简单，但是精度高。分析目前国内外主流的几种基于图像处理的客流计数系统<sup>[9-10]</sup>，对比其性能，结果如表 3。

表 3 基于图像处理的客流计数系

Table 3 Passenger counting system based on image processing

	Passenger counting system based on color video stream	Passenger counting system based on stereo vision	The system
System component	DSP and colorful CCD	DSP and double CCD	PC or ARM and Xtion sensor
Key technology	Training library, identification, tracking	Camera calibration, stereo matching, identification	Feature recognition, tracking
Cost /RMB	3000	4000	2000
System accuracy /%	93	92	96

	Passenger counting system based on color video stream	Passenger counting system based on stereo vision	The system
System merits	High precision	High precision, solving the environmental interference; suitable for crowded environments	High precision; low cost; portable; Integrated; Suitable for dark environment
System defects	Algorithm complexity is high; accuracy affected by environmental factors is large; high cost.	It is difficult to get double CCD data synchronously; Calibration and matching pairs of CCD are complex	Effective height is limited.

## 5 结 论

基于 Xtion 传感器输出的包含人体三维信息深度图,设计了一种新的客流计数方法。该算法利用人体三维空间特征统计人流量,解决了基于彩色图和基于立体视觉中的许多问题,实验证实了该系统的稳定性和高精度,具有实际的应用价值。并且 Xtion 传感器有 Android 平台的驱动,可以在 Android 平台终端实现系统算法,实现系统的硬件化。并且系统可以同时驱动多 Xtion 传感器,实现多摄像机多视角同步检测统计客流。

### 参 考 文 献

- 1 S Mukherjee, B Saha, I Jamal, *et al.*. A novel framework for automatic passenger counting[C]. IEEE International Conference on Image Processing, 2011. 2969—2972.
- 2 T Yabiaoui, L Khoudour, C Meurie. Real-time passenger counting in buses using dense stereovision [J]. Journal of Electronic Imaging, 2010, 19(3): 031202.
- 3 G J Brostow, R Cipolla. Unsupervised bayesian detection of independent motion in crowds [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, 1: 594—601.
- 4 Haibin Yu, Zhiwei He, Jilin Li. A vision-based method to estimator passenger flow in bus [C]. Proceedings of 2007

International Symposium on Intelligent Signal Processing and Communication Systems, Xiamen, 2007. 654—657.

- 5 F Bartolini, V Cappellini, A Mecocci. Counting people getting in and out of a bus by real-time image-sequence processing [J]. Image and Vision Computing, 1994, 12(1): 36—41.
- 6 Chaoho Chen, Yinchan Chang, Tsongyi Chen. People counting system for getting in/out of a bus based on video processing[C]. Proceedings of Eighth International Conference on Intelligent Systems Design and Applications, 2008, 3: 565—569.
- 7 Yin Zhangqin, Gu Guohua, Bai Xiaofeng. Passenger flow statistics across the field of view based on the depth map of the double xtion sensors[C]. SPIE, 2013: 89080X.
- 8 Liang Min, Liu Guixi. Multi-object tracking algorithm based on adaptive mixed filtering[J]. Acta Optica Sinica, 2010, 30(9): 2554—2561.  
梁 敏, 刘贵喜. 基于自适应混多目标跟踪算法[J]. 光学学报, 2010, 30(9): 2554—2561.
- 9 Jin Jing. Image Processing Based Pedestrians Detection Algorithms Research and Framework Design Using Single Camera [D]. Shanghai: Shanghai Jiao Tong University, 2009.  
靳 晶. 基于图像处理技术的单目客流检测系统研究[D]. 上海: 上海交通大学, 2009.
- 10 Zhu Qiuyu, Tang Li, Yu Ming, *et al.*. A passenger counting method based on stereovision[J]. Journal of Image and Graphics, 2009, 14(11): 2391—2395.  
朱秋煜, 唐 利, 郁 铭, 等. 一种基于立体视觉的公交车客流计数方法[J]. 中国图像图形学, 2009, 14(11): 2391—2395.

栏目编辑:张浩佳