

融合颜色信息与深度信息的运动目标检测方法

胡良梅* 段琳琳 张旭东 杨静
(合肥工业大学计算机与信息学院 合肥 230009)

摘要: 基于颜色信息的运动目标检测易受光照、阴影等影响, 基于深度信息的运动目标检测存在目标边缘噪声大, 无法检测距离背景较近的目标等问题。针对上述问题, 该文利用 CCD 相机获取的颜色信息及 TOF 相机获取的深度信息分别为每个像素建立颜色与深度信息的分类器, 根据像素点的深度特征及前一帧的检测结果, 自适应地为每个分类器的输出分配不同的权值, 实现运动目标的检测。该文采集多组视频序列进行实验, 实验结果表明该方法能有效解决单独利用颜色或深度信息进行运动目标检测时出现的问题。

关键词: 运动目标检测; 融合; TOF 相机; 深度信息; 颜色信息

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2014)09-2047-06

DOI: 10.3724/SP.J.1146.2013.01763

Moving Object Detection Based on the Fusion of Color and Depth Information

Hu Liang-mei Duan Lin-lin Zhang Xu-dong Yang Jing

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: Color-based moving object detection performs poorly when illumination changes or shadow exists. Depth-based moving object detection is affected by the high level of depth-data noise at object boundaries, and it fails when foreground objects move close to the background. For these reasons, a novel approach that establishes color and depth classifier for each pixel is presented by making full use of color information obtained by CCD camera and depth information obtained by TOF camera. In order to realize the effective detection, different weights are assigned adaptively for each output of the classifier by considering foreground detections in the previous frames and the depth feature. Multi video sequences are captured to verify the proposed method, and the experimental results show that the proposed approach can effectively solve the limitations of color-based or depth-based detection and realize the effective detection.

Key words: Moving object detection; Fusion; TOF camera; Depth information; Color information

1 引言

从视频序列中分离出运动目标是计算机视觉中的重要研究内容, 在视频监控、人机交互、视频会议等领域有着广泛的应用。目前最常用的运动目标检测是基于颜色信息的检测方法, 但其主要存在以下问题: (1)易受光照变化影响, (2)难以克服阴影干扰, (3)运动背景颜色相似时检测困难。针对上述问题, 国内外学者提出了许多有效的检测算法^[1-5], 但这些方法受环境因素的限制较大, 收敛速度较慢, 且均难以处理(3)的情况。

利用深度信息进行运动目标检测, 可以有效解决上述问题。近年来深度信息的获取发展迅速, 基

于飞行时间(Time-Of-Flight, TOF)原理的3维测距相机^[6,7]凭借其无需扫描、高帧率、不依赖与外界光照等特点, 在运动目标检测跟踪等方面有广泛应用^[8-10]。但其获取的深度信息存在噪声大、分辨率低^[11]的问题, 且当运动目标与背景距离相近时出现漏检。

近年来, 部分学者提出了融合颜色和深度信息检测运动目标的方法。文献[12]用灰度信息与深度信息分别检测目标, 然后用逻辑或运算将检测结果融合, 解决了目标与背景颜色及距离相近的问题, 但引入了光照及阴影的影响; 文献[13]采用区域生长的方法在距离图像中提取前景目标, 利用颜色信息细化前景目标, 但难以处理复杂场景; 文献[14]利用彩色图像与立体视觉获取的深度图像分别检测目标, 通过一系列的逻辑判断将两者的检测结果进行融合, 但仍难以克服边缘噪声及漏检问题。

2013-11-08 收到, 2014-01-23 改回

国家自然科学基金(61273237)和安徽省自然科学基金(11040606 M149)资助课题

*通信作者: 胡良梅 liangmeihu@gmail.com

本文提出一种融合颜色信息与深度信息的运动目标检测算法,首先为每个像素建立基于颜色、深度信息的贝叶斯分类器,依据像素的边缘特征及上一帧的检测结果,自适应地为分类器的输出赋予不同的权重进行融合,有效地实现运动目标的检测。

2 距离图像与彩色图像的配准融合

本文自行搭建 2D/3D 系统,由 TOF 相机和 CCD 相机构成,实时获取场景的视频序列。首先对 TOF 相机获取的距离图像和 CCD 相机获取的彩色图像进行配准融合。采用文献[15]的标定法分别对 TOF 相机和 CCD 相机进行标定,获取两个相机的内外部参数,然后利用文献[16]中的配准融合方法实现图像的配准融合,最终每一帧距离图像都可以得到与彩色图像同分辨率的距离图像。

3 基于颜色信息与深度信息融合的运动目标检测

本文算法的整体框架如图 1 所示,首先利用配准融合后的距离图像和彩色图像建立基于像素的贝叶斯分类器 CL_D, CL_C , 然后对输入的像素 s , 将其深度信息 s_D 、颜色信息 s_C 分别送入分类器 CL_D, CL_C , 两个分类器的输出结果组成矩阵 Q 。

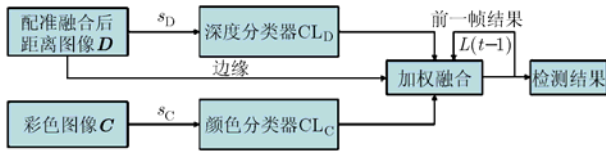


图1 运动目标检测算法的整体框架

$$Q(s) = \begin{bmatrix} P(\omega_{bg} | s_D) & P(\omega_{fg} | s_D) \\ P(\omega_{bg} | s_C) & P(\omega_{fg} | s_C) \end{bmatrix} \quad (1)$$

其中 $P(\omega_{bg} | s_D)$, $P(\omega_{fg} | s_D)$ 代表分类器 CL_D 输出的像素 s 属于背景 ω_{bg} 和目标 ω_{fg} 的概率, $P(\omega_{bg} | s_C)$, $P(\omega_{fg} | s_C)$ 代表分类器 CL_C 输出的像素 s 属于背景 ω_{bg} 和目标 ω_{fg} 的概率。

对图像的不同区域,深度信息和颜色信息对检测结果的贡献程度不同,因此我们为每个分类器的输出分配权重 W_i ($i \in \{C, D\}$), 则像素 s 属于 ω_i 的概率 $P(\omega_i | s)$ 为

$$P(\omega_i | s) = W_D \cdot P(\omega_i | s_D) + W_C \cdot P(\omega_i | s_C), \quad i \in \{bg, fg\} \quad (2)$$

其中 $W_D + W_C = 1$ 。最后将像素 s 归为较大的 $P(\omega_i | s)$ 对应的 ω_i 类。

3.1 权重选择策略

配准融合后的距离图像存在两个问题:(1)TOF

相机与 CCD 相机拍摄的场景范围不同,配准融合后的距离图像存在无法获取颜色信息的像素;(2)TOF 相机边缘噪声较大,配准融合后的距离图像在边缘区域可信度低。针对上述问题,充分考虑深度信息和颜色信息在不同区域对检测结果的贡献程度,本文提出一种权重选择策略,具体步骤如下:

(1)判断 t 帧时像素 s 是否获取深度信息,若获取深度信息,执行步骤(2),若未获取深度信息,只利用颜色信息进行目标检测,即

$$\begin{cases} W_D(s, t) = 0 \\ W_C(s, t) = 1 \end{cases}, \quad s \text{ 未获取深度信息} \quad (3)$$

(2)利用 Sobel 算子对像素 s 进行边缘检测,像素 s 的梯度 $G(s)$ 大于阈值 T , 归为边缘区域,梯度小于等于阈值 T , 归为平坦区域,若像素 s 属于边缘,执行步骤(3),否则执行步骤(4);

(3)为了减少深度信息边缘噪声大的影响,分类器 CL_C 起主要作用,按式(4)分配权重:

$$\begin{cases} W_D(s, t) = W_{\min} \\ W_C(s, t) = W_{\max} \end{cases}, \quad G(s) > T \quad (4)$$

(4)为了解决运动目标逐渐靠近背景时分类器 CL_D 出现的漏检情况,本文引入前一帧的检测结果 $L(s, t-1)$ 对权重进行分配。

(a)若 $L(s, t-1)$ 为背景,按照式(5)分配权重:

$$\begin{cases} W_D(s, t) = W_{\max} \\ W_C(s, t) = W_{\min} \end{cases}, \quad G(s) < T, \quad L(s, t-1) \in \omega_{bg} \quad (5)$$

(b)若 $L(s, t-1)$ 为运动目标,计算 $t-1$ 帧时像素 s 的深度信息与分类器 CL_D 的背景模型(见 3.2.1 节)的距离关系:

$$\lambda(s, t-1) = |s_{D,t-1} - \mu_{s,t-1}| / \sigma_{s,t-1} \quad (6)$$

$\lambda(s, t-1)$ 越大表示像素 s 对应的目标离背景越远,相反, CL_D 表示像素 s 对应的目标离背景越近。为了避免目标与背景距离相近时采用深度信息会出现漏检, W_D 应随 λ 的减小而减小。本文采用广义逻辑函数分配 W_D , 权重分配如式(7)所示。

$$\begin{cases} W_D(s, t) = W_{\min} \\ \quad + \frac{W_{\max} - W_{\min}}{\left(1 + Q \exp(-B(\lambda(X_s, t-1) - M))\right)^{1/v}}, \\ W_C(s, t) = 1 - W_D(X_s, t) \end{cases}, \quad G(s) < T, \quad L(s, t-1) \in \omega_{fg} \quad (7)$$

3.2 贝叶斯分类器

本文采用贝叶斯分类器,每个贝叶斯分类器输出像素 s 属于目标和背景的后验概率 $P(\omega_i | s)$, 假设

目标与背景先验概率相同，则根据贝叶斯公式，式(2)可以写成：

$$P(\omega_i | s) = W_D \cdot \frac{p(s_D | \omega_j)}{p(s_D | \omega_{fg}) + p(s_D | \omega_{bg})} + W_C \cdot \frac{p(s_C | \omega_j)}{p(s_C | \omega_{fg}) + p(s_C | \omega_{bg})}, j \in \{bg, fg\} \quad (8)$$

3.2.1 背景与目标模型 采用单高斯模型对每个像素进行背景建模，背景模型如式(9)：

$$p(s_t | \omega_{bg}) = \eta(s_t, \mu_{s,t}, \Sigma_{s,t}) \quad (9)$$

其中 s_t 表示 t 帧时像素 s 的深度或颜色信息， $\mu_{s,t}$, $\Sigma_{s,t}$ 分别代表像素 s 的均值和协方差矩阵。

对每个像素建立目标模型，由于场景中的运动目标特征未知，因此可以认为该前景目标像素点可能出现任意的特征值，且每个特征值出现的概率是相同的，本文采用文献[17]的方法，利用均匀概率分布模型对运动目标进行建模。

$$p(s | \omega_{fg}) = \frac{1}{256^N} \quad (10)$$

其中，256 是每个通道的离散值个数， N 是像素 s 值的通道数。

3.2.2 模型初始化及更新 利用获取的 n 帧彩色图像和融合后的距离图像，求得每个像素的均值和方差作为分类器 CL_C , CL_D 背景模型的初始值。背景初始化后，按照本文算法对像素进行分类。若像素 s 被判为背景，对两个分类器背景模型的均值进行更新，更新步骤如式(11)：

$$\mu_{s,t+1} = (1 - \alpha)\mu_{s,t} + \alpha s_{t+1} \quad (11)$$

其中 α 为更新速率。

为解决背景物体突然发生运动引起的“鬼影”问题，本文提出一种基于像素深度信息的有效解决方法。背景物体突然运动，原位置的深度信息变大，利用这一特征可以有效解决“鬼影”问题。若像素 s 判为目标，将像素 s 的深度信息与分类器 CL_D 背景模型的均值比较，在误差允许范围内，像素 s 的深度信息满足式(12)，将像素 s 改判为背景。

$$s_{D,t} - \mu_{s,t} > K \cdot \sigma_s \quad (12)$$

同时，重新初始化像素点 s 的两个背景模型的均值， $\mu_{s,t+1} = s_t$ 。

4 实验结果与分析

实验采用的 TOF 相机为 PMD 公司生产的 PMD[vision]® CamCube 2.0 相机，分辨率为 204×204 ，测量的距离范围为 $0 \sim 7.5$ m；CCD 相机为大恒公司的 DH-SV1421FC，分辨率为 1392×1040 。运行环境为 Matlab.R2008a，运行在主频为 2.4 GHz，内存为 2 G 的 PC 机上。为验证本文方法的

有效性，本文采集了目标与背景颜色相近、目标逐渐靠近背景、背景物体移动、光照突变这 4 种情况下的室内视频序列，并与文献[12]的检测结果进行比较。

4.1 参数设置

本文算法包含有多个参数，这些参数的设置共同影响着算法的有效性，为了使算法整体取得好的性能，在大量实验的基础上本文采用如表 1 所示的参数值。

表 1 本文算法选取的参数值

参数	参数值	参数	参数值	参数	参数值
W_{\min}	0.10	Q	2.50	v	0.50
W_{\max}	0.90	B	1.50	a	0.05
T	0.29	M	3.00	K	4.00

4.2 定性分析

4.2.1 目标与背景颜色相近 序列 1 的背景如图 2(a) 所示，如图 2(b) 所示绿色盒子经过与之颜色相同的背景。为了更好地分析结果，本文截取了图像中的盒子部分，实验结果如图 3 所示。由于目标与背景颜色相近，如图 3(a) 所示分类器 CL_C 无法完整检测出目标；如图 3(b) 所示，分类器 CL_D 利用深度信息可以保证运动物体完整检测出来，然而由于边缘噪声大，造成边缘检测模糊；如图 3(c) 所示，文献[12]方法将两个检测结果直接逻辑运算，仍存在边缘模糊的问题；如图 3(d) 所示本文算法可以综合两者优势，有效检测出运动目标，与图 3(e) 手工检测结果最为接近。

4.2.2 目标与背景距离相近 序列 2 人手持纸板使其逐渐靠近背景，本文分别选取第 690, 766, 840 帧图像，690 帧时，纸板离背景较远，766 帧时纸板逐渐靠近背景，并且在场景中产生阴影，840 帧时纸板几乎贴近背景，3 帧的检测结果如图 4, 图 5, 图 6 所示。由检测结果可以看出本文算法可以在运动目标逐渐靠近背景时完整检测出运动目标，并且可以克服阴影的影响。

4.2.3 背景物体移动 序列 3 的背景如图 7(a) 所示，从 145 帧开始背景中的物体开始运动，图 7(c) 为 187 帧对应的融合距离图像。如图 8 所示，分类器 CL_C 由于只对背景像素更新，在原位置产生“鬼影”；分类器 CL_D 按照本文提出的更新背景模型方法，快速地判断背景中移动的物体，且可以克服阴影影响，但边缘出现模糊；文献[12]方法无法消除“鬼影”及阴影的影响，本文方法可以迅速判断背景物体的移动，并消除阴影影响。

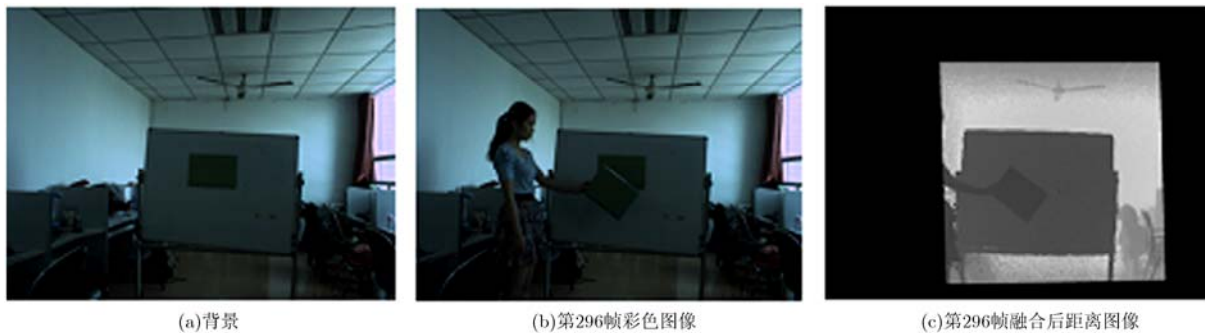


图2 序列1 第296帧图像

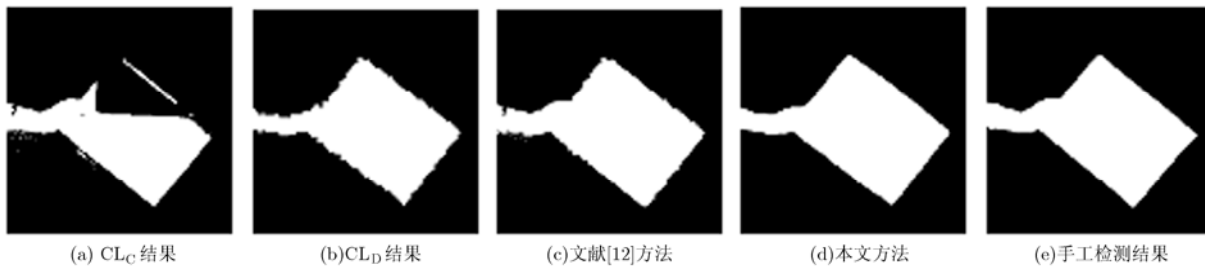


图3 序列1 第296帧检测结果



图4 序列2 第690帧结果



图5 序列2 第766帧结果

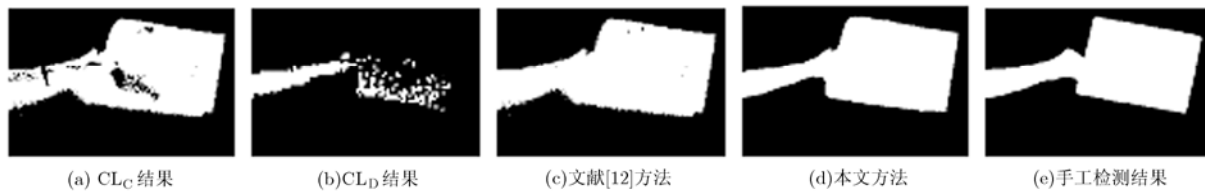


图6 序列2 第840帧结果



图7 序列3 背景及第187帧图像



图 8 序列 3 第 187 帧检测结果

4.2.4 光照突变 序列 4 为光照变化的视频序列，第 1411 帧时光线突变。本文仍截取图像的目标部分，为了使截取的部分具有一般性，本文截取了部分未获取距离信息的像素。如图 9 所示，分类器 CL_C 及文献[12]方法均无法适应光照的突然变化，而本文算法虽然对无法获取距离信息的部分仍然受光照影响，但整体上可以有效克服阴影及光照的变化。

4.3 定量分析

常见的目标检测性能评价标准有：精确度 P ，查全率 R 及 F 数，其计算公式如下：

$$P = \frac{TP}{TP + FP} \tag{13}$$

$$R = \frac{TP}{TP + FN} \tag{14}$$

$$F = 2RP / (R + P) \tag{15}$$

其中 TP 为被正确检测为运动目标的像素数； TN 为被正确检测为背景的像素数； FN 为被误判为背景的像素数； FP 为被误判为运动目标的像素数。

F 数综合了精确度 P 与查全率 R 两个评价指

标，能够更全面地评价算法的性能，本文给出了上述 4 个序列中部分帧数的 F 数的曲线，如图 10~图 13 所示。综合图 10~图 13 可知，单独基于颜色或距离的目标检测及文献[12]的方法均会在不同的情况下出现算法性能下降甚至失效，而本文算法在处理目标与背景颜色或距离相近、背景物体移动及光照阴影影响等多种复杂情况下均能保持较高的检测性能。

5 结束语

论文针对单独利用颜色信息或深度信息检测目标存在的问题，提出一种融合颜色信息与深度信息的运动目标检测方法。对每个像素建立基于颜色和深度的贝叶斯分类器，为综合颜色与深度信息在不同区域的置信度，分配不同的权重将两者有效融合，并且针对背景物体运动产生的“鬼影”问题，利用其深度信息提出了一种快速简单的解决方法。多组实验表明，本文方法可以有效解决单独利用颜色或深度信息检测难以克服的问题。

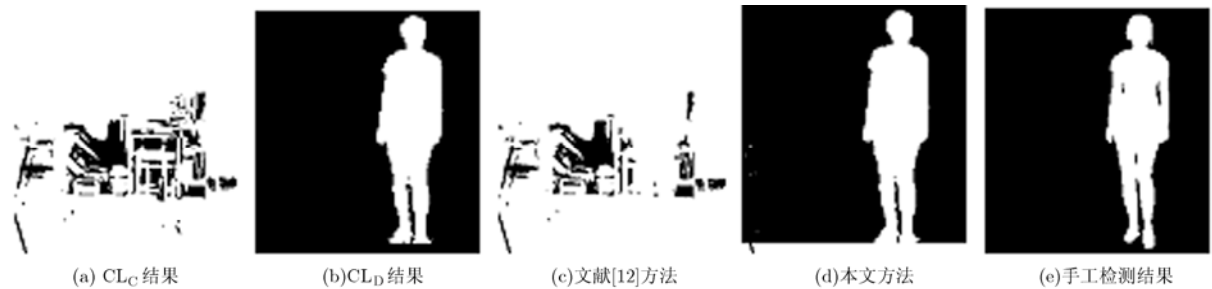


图 9 序列 4 第 1411 帧检测结果

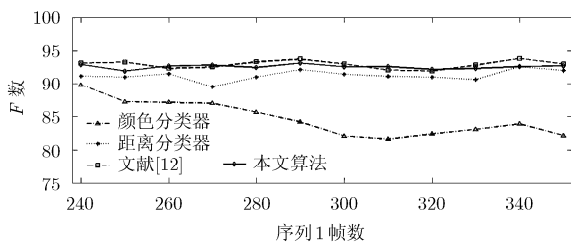


图 10 序列 1 的 F 数

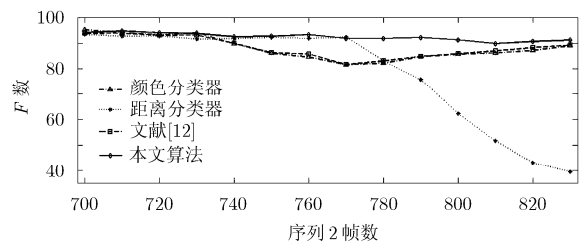


图 11 序列 2 的 F 数

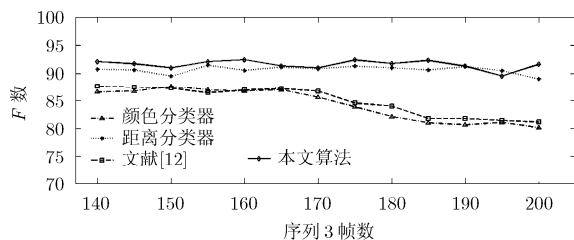


图12 序列3的F数

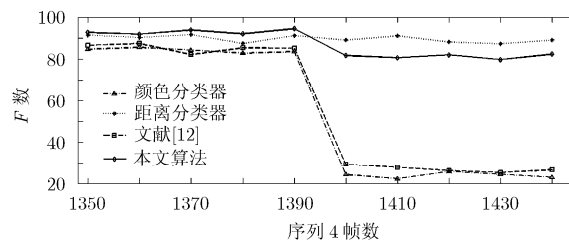


图13 序列4的F数

参考文献

- [1] Stauffer C and Grimson W E L. Adaptive background mixture models for real-time tracking[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, 1999: 246-252.
- [2] Cucchiara R, Grana C, Piccardi M, *et al.* Improving shadow suppression in moving object detection with HSV color information[C]. Proceedings of the IEEE Intelligent Transportation Systems, Oakland, CA, 2001: 334-339.
- [3] 李拥军, 曾标, 徐克付, 等. 复杂背景下基于贝叶斯-全概率联合估计的前景检测[J]. 电子与信息学报, 2012, 34(2): 388-392.
Li Yong-jun, Zeng Biao, Xu Ke-fu, *et al.* Foreground object detection in complex background based on Bayes-total probability joint estimation[J]. *Journal of Electronics & Information Technology*, 2012, 34(2): 388-392.
- [4] Feng W, Liu R, Jia B, *et al.* An efficient pixel-wise method for moving object detection in complex scenes[C]. IEEE International Conference on Advanced Video and Signal Based Surveillance, Krakow, Poland, 2013: 389-394.
- [5] 周建英, 吴小培, 张超, 等. 基于滑动窗的混合高斯模型运动目标检测方法[J]. 电子与信息学报, 2013, 35(7): 1650-1656.
Zhou Jian-ying, Wu Xiao-pei, Zhang Chao, *et al.* A moving object detection method based on sliding window Gaussian mixture model[J]. *Journal of Electronics & Information Technology*, 2013, 35(7): 1650-1656.
- [6] Zhang D and Lu G. The Principle of 3D Camera Imaging[M]. New York: 3D Biometrics, Springer, 2013: 233-256.
- [7] Foix S, Alenya G, and Torras C. Lock-in Time-of-Flight (ToF) cameras: a survey[J]. *IEEE Sensors Journal*, 2011, 11(9): 1917-1926.
- [8] Lee C H, Su Y C, and Chen L G. An intelligent depth-based obstacle detection for mobile applications[C]. IEEE International Conference on Consumer Electronics, Berlin, 2012: 223-225.
- [9] Li Y, Wang G, Lin X, *et al.* Real-time depth-based segmentation and tracking of multiple objects[C]. Proceedings of IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems, Bangkok, Thailand, 2012: 429-433.
- [10] Tran A and Harada K. Depth-aided tracking multiple objects under occlusion[J]. *Journal of Signal and Information Processing*, 2013, 4(3): 299-307.
- [11] Hu Liang-mei, Zhang Ya-qiong, Zhang Xu-dong, *et al.* The 3D reconstruction of PMD camera and high resolution color camera[C]. Proceedings of 3rd International Conference on Signal Processing Systems, Yantai, China, 2011, (2): 470-472.
- [12] Leens J, Piérard S, Barnich O, *et al.* Combining color, depth, and motion for video segmentation[J]. *LNCS*, 2009, 5815: 104-113.
- [13] Mirante E, Georgiev M, and Gotchev A. A fast image segmentation algorithm using color and depth map[C]. IEEE 3DTV-Conference on The True Vision-Capture, Transmission and Display of 3D Video, Antalya, Turkey, 2011: 1-4.
- [14] Ottonelli S, Spagnolo P, Mazzeo P L, *et al.* Improved video segmentation with color and depth using a stereo camera[C]. IEEE International Conference on Industrial Technology, Western Cape, South Africa, 2013: 1134-1139.
- [15] Zhang Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(11): 1330-1334.
- [16] 贾保柱. 融合2D与3D图像的三维重建系统实现[D]. [硕士论文], 合肥工业大学, 2011.
Jia Bao-zhu. 3D reconstruction from 2D and 3D images[D]. [Master dissertation], Hefei University of Technology, 2011.
- [17] Landabaso J L. A unified framework for consistent 2D/3D foreground object detection[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2008, 18(8): 1040-1051.

胡良梅: 女, 1974年生, 副教授, 研究方向为信息融合、图像理解、模式识别、智能信息处理。

段琳琳: 女, 1990年生, 硕士, 研究方向为智能信息处理。

张旭东: 男, 1966年生, 教授, 研究方向为机器视觉、传感器技术、智能信息处理及相关应用系统的开发。