

基于 YCbCr 颜色空间背景建模与手势阴影消除

史东承*,倪康

(长春工业大学 计算机科学与工程学院,吉林 长春 130012)

摘要:为了提高动态手势检测的精确度,本文将基于 YCbCr 颜色空间的混合高斯背景建模应用于动态手势识别中,并且提出手势阴影消除的有效算法。首先,对待检测视频帧通过抠图抠出手势图像,在 YCb'Cr'颜色空间进行椭圆拟合,统计建立椭圆肤色模型,继而在 YCbCr 颜色空间进行混合高斯背景建模检测出动态手势,点乘原图像得到含有阴影的 RGB 手势图像,对检测出的含有阴影的手势图像利用已建立的椭圆肤色模型进行阴影消除,最后将手势图像连成视频序列。实验结果表明,该算法在复杂背景下进行动态手势的检测率可达 91.4%,高出传统方法 10% 左右,能够满足动态手势检测基本要求,且具有较高的实用价值。

关键词:手势检测;混合高斯背景建模;椭圆拟合;阴影去除

中图分类号:TP391.9 文献标识码:A doi:10.3788/CO.20150804.0589

Background modeling based on YCbCr color space and gesture shadow elimination

SHI Dong-cheng*, NI Kang

(College of Computer Science and Engineering, Changchun University of Technology, Changchun 130012, China)

* Corresponding author, E-mail: dcshi@foxmail.com

Abstract: To improve the accuracy of the dynamic gesture detection, Gaussian mixture background modeling based on YCbCr color space is applied to the dynamic gesture recognition, and the effective gesture shadow elimination algorithm is proposed. First of all, the gesture image is cut out from video frame to be detected, and space ellipse fitting is developed in YCb'Cr' color. Oval color model is established statistically, and then dynamic gesture in the YCbCr color space through Gaussian mixture background modeling is detected. Original image is dotted product to get the gesture RGB image containing shadows. The shadows contained in the detected gestures image were eliminated by using ellipse color model, and finally we take gesture images together into a video sequence. Experimental results show that in the algorithm of dynamic gesture detection rate is 91.4% under a complex background, about 10% higher than that by the traditional methods. So it can meet the basic requirements of dynamic gesture detection, and has a high practical value.

Key words: gesture detection; Gaussian mixture background modeling; ellipse fitting; shadow removal

1 引言

在信息迅速膨胀的今天,各种智能化产品登上人们生活的舞台,推动了人工交互的迅速发展。手势识别技术是人机交互的重要组成部分,近年来各种手势识别的应用产品已投入到大众需求中,例如:基于手势识别的远程电视控制系统和视频游戏的用户界面、哑语识别及交警手势识别等。手势识别是集人工智能、神经识别、图像处理及智能分析等于一体,涉及计算机语言、计算机视觉、图像处理等学科。手势识别技术的迅速发展带动了各个领域的共同发展,在处于信息交互频繁的今天,手势识别已经成为人机交互发展的主流。

手势识别过程中首先要进行手势检测,检测结果的好坏直接影响到后面识别的精确度。而在运动目标检测中,运动物体受到风吹、光照等外界因素的干扰容易产生阴影及其他噪声,对运动目标检测的精确性造成较大影响^[1-3]。目前运动目标检测运用最广泛的方法就是背景建模法,国内外对于手势检测的研究也取得了突破性进展,手势分割技术也从最初的借助数据手套等工具向着由纹理、颜色等特征的方向发展^[4-6]。如文献[7]和文献[8]中提出的背景建模在简单背景下能够较好地进行运动目标检测,但是不能有效地抑制阴影,容易造成运动目标的误检测;文献[1]和文献[4]提出的复杂环境下的手势检测,能够在较为复杂的背景中准确提取手势,但仅局限于静态手势;文献[9]和文献[10]提出基于GMM的彩色运动目标检测和联合两种颜色空间的运动目标阴影消除,这两种方法实现了对运动目标的提取及阴影的部分消除,但是该方法对光照及外界环境的要求高,具有一定的局限性,实用性不高。文献[11]和文献[12]提出的方法能够在复杂背景下进行运动目标检测,但也不能有效地抑制阴影,

具有一定的局限性。

本文提出的基于YCbCr颜色空间的混合高斯背景建模和基于YCb'Cr'颜色空间椭圆肤色模型相结合的动态手势检测不仅能够复杂背景下准确进行手势检测,还能够有效消除手势阴影,提高了手势检测的精确度,具有一定的实用价值。

2 YCb'Cr'颜色空间椭圆模型的建立

2.1 手势图像抠图及YCb'Cr'颜色空间的转化

将待检测手势图像拆分成70帧,利用Photoshop软件手动将前20帧含背景信息的手势图像中抠出手势图像,并将抠出的手势图像非线性转换到YCb'Cr'颜色空间,抠出的手势图像如图1所示。

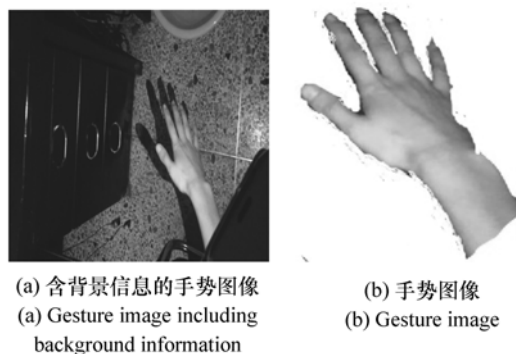


图1 手势原图及手势抠图

Fig. 1 Hand gesture and cutout style of hand gestures

YCb'Cr'颜色空间是YCbCr非线性转换形成的。研究发现在YCbCr颜色空间,肤色亮度和色度在高亮度和低亮度区域并非严格线性无关,这种依赖关系很大程度上影响了肤色的检测效果^[13]。YCb'Cr'颜色空间不仅解除了这种依赖关系,肤色在该颜色空间有着很好的聚类性,并且这种肤色聚类区域不随着亮度的变化而变化。YCb'Cr'颜色空间转换公式如下:

$$C'_i = \begin{cases} (C_i(Y) - \overline{C_i(Y)}) \frac{W_{C_i}}{W_{C_i}(Y)} + \overline{C_i(Y)}, Y \notin [K_l, K_h] \\ C_i(Y), Y \in [K_l, K_h] \end{cases}, \quad (1)$$

式中: $W_{C_i}(Y)$ 、 $\overline{C_i(Y)}$ 的具体运算公式在文献[1]中有具体的求解。 $W_{C_i}(Y)$ 描述的是在Y-C_i坐标

中肤色区域的宽度, C_i 表示 C_b 和 C_r , $\overline{C_i(Y)}$ 、 $C_b(Y)$ 用4个边界来限定聚类区域。这样更加有

利于模型更好地适应亮度过暗或过亮的区域,以此提高系统的鲁棒性,提高对环境的适应能力。 K_l 、 K_h 表示非线性分段色彩变换的分段阈值。经过大量统计实验数据得出^[1]: $K_l = 125$, $K_h = 148$, $W_{C_b} = 46.97$, $W_{C_r} = 38.76$, $WL_{C_b} = 23$, $WL_{C_r} = 20$ 。

2.2 最小二乘法椭圆拟合及椭圆模型的建立

经研究发现,肤色在 YCb'Cr'颜色空间聚类形状类似椭圆^[1],且具有较强的聚类性。把视频序列的第7帧进行非线性转换,并且将肤色点在Cb'-Cr'坐标系下的投影,投影结果如图2所示,图中横坐标为Cb'像素值,纵坐标为Cr'像素值。

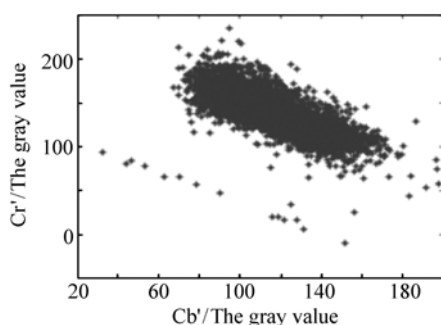


图2 肤色在Cb'-Cr'坐标系下的投影

Fig.2 Projection of oval color in color Cb'-Cr' coordinate system

从投影图可以看出,用椭圆模型来表述肤色在此颜色空间的分布比较精确,能够找到一个能包含所有肤色点的最小椭圆是手势提取的关键。本文选用最小二乘法进行椭圆拟合,最小二乘法通过最小化误差的平方和来寻找最佳函数匹配,距离拟合和几何曲线拟合是最小二乘法的两种拟合方式^[1]。本文采用距离拟合椭圆的两种表达式如下:

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1, \quad (2)$$

$$F(x) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F. \quad (3)$$

式(3)拟合过程运算较简便,且式(3)中参数能与式(2)中的椭圆五参数进行转换^[14]。本文椭圆拟合过程如下:首先,为避免椭圆方程有零解,并将解的整数倍误认为同一方程,需将式(3)加限制条件 $A + C = 1$ ^[15],然后根据聚类图形手动选取聚类图形的6个具有代表性的边缘像素点

$p(x_i, y_i)$,其中包括4个椭圆长轴短轴与椭圆交点,再另选2个椭圆边缘曲率大的对称像素点。继而构造误差平方和函数,如式(4):

$$f = \sum_{i=1}^6 (F(x_i, y_i) - P(x_i, y_i))^2. \quad (4)$$

由极值原理可知,可得以下线性方程:

$$\frac{\partial f}{\partial A} = \frac{\partial f}{\partial B} = \frac{\partial f}{\partial C} = \frac{\partial f}{\partial D} = \frac{\partial f}{\partial E} = \frac{\partial f}{\partial F} = 0, \quad (5)$$

继而求解该线性方程式可求出参数A、B、C、D、E、F。

本文对视频序列中前20帧抠出的手势图像进行最小二乘法椭圆拟合,然后求出参数并取其均值,最后进行公式转化^[14]后可得椭圆的五参数具体数值为: $\theta = 2.80$, $ex_y = 2.35$, $ec_x = 2.05$, $a = 27.50$, $b = 14.30$ 。

此时拟合的椭圆方程已确定,继而建立椭圆肤色模型,图像中的像素点可用下式表示:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos 2.80 & \sin 2.80 \\ -\sin 2.80 & \cos 2.80 \end{bmatrix} \begin{bmatrix} C'_b - 109.38 \\ C'_r - 152.02 \end{bmatrix}, \quad (6)$$

此时令

$$M = \frac{(x - 2.05)^2}{27.50^2} + \frac{(y - 2.35)^2}{14.30^2}. \quad (7)$$

若 $M > 1$,该像素位于椭圆之外认定为背景像素,标记为0,若 $M < 1$,该像素位于椭圆内部认定为前景像素,即为肤色,标记为1。经过非线性转换容易形成伪肤色,即亮度小于80的非肤色点视为肤色,故本文将 $Y < 80$ 的像素点认定为背景像素,即标记为0。

3 基于 YCbCr 颜色空间的混合高斯背景建模

3.1 模型简介以及参数初始化

混合高斯模型中的背景像素的每一个像素可用 N 个不同的高斯分布表示,本文取 $N = 3$ ^[11]。每一个高斯分布都有其对应的权值 W_k 。高斯模型还有2个重要的参数:均值 μ 与方差 Ω 。对于手势视频帧中的像素点 (x, y) 在 t 时刻的观察值表示为 $X_t = (I_t^r, I_t^b, I_t^g)^T$ 。方差矩阵和均值矢量

表示为:

$$\Omega_{k,t} = \begin{bmatrix} (\sigma_{k,t}^y)^2 & 0 & 0 \\ 0 & (\sigma_{k,t}^{cb})^2 & 0 \\ 0 & 0 & (\sigma_{k,t}^{cr})^2 \end{bmatrix}, \quad (8)$$

$$\mu_{k,t} = [\mu_{k,t}^y, \mu_{k,t}^{cb}, \mu_{k,t}^{cr}]^T. \quad (9)$$

参数初始化过程中将权值设置尽可能小,方差设置尽可能大,均值大小取第一帧图像中的每一个像素点的像素值。本文取 $W_{k,t} = 0.05$; $\Omega_{k,t} = 15$ 。

3.2 模型参数实时更新

读取手势视频中第一帧, 现有的高斯模型相匹配与帧中像素点进行匹配, 若匹配, 继而判断该像素与混合高斯模型中的某个高斯分布均值的距离是否小于标准差与偏差阈值的乘积(本文偏差阈值 $D = 2.0$)。若满足, 则该像素点与高斯模型中高斯分布相匹配, 将该像素点归入高斯模型中, 并按照文献[6]进行参数更新。在更新过程中, α 反映的是背景的更新速度, 其取值大小对模型的更新有较大的影响。经测试, 本文 α 取值为 0.03 时效果最佳。若该像素与混合高斯模型中高斯分布不匹配, 则按照下式减小权重 W_k , 即: $W_{k,t+1} = (1 - \alpha) W_{k,t}$ 。若像素点与已有的高斯模型不匹配, 新的高斯模型将取代权重最小的高斯模型。经过参数不断更新和训练, 得到混合高斯分布模型。

3.3 混合高斯背景模型的建立及手势分割

选取权重大于某一阈值 T 的所有高斯分布建立混合高斯背景模型, 经测试, 本算法 $T = 0.40$ 时效果最佳。此时判断混合高斯背景模型与当前帧的某一像素点 X_t 是否匹配, 即: $|I_{k,t}^i - \mu_{k,t}^j| < D\sigma_{k,t}^j$, 其中, D 为偏差阈值, $\sigma_{k,t}$ 为标准差。若上式成立, 该像素点标记为 0, 即为背景像素点; 若不成立, 该像素点即为 1, 即为前景像素点[6]。继而进行一下形态学操作, 形态开操作消除部分噪声, 连通域操作删除面积较小的噪声, 运用合适结构元填充图像空洞区域。最终得到运动手势图像。经过以上处理, 混合高斯背景建模第 17 帧图像动态手势检测结果如图 3 所示。

经过混合高斯建模得到的手势图像含有明显的阴影, 从图 3 可看出, 手势阴影已严重影响手势



图3 视频序列第17帧

Fig.3 Video sequence frame 17

检测的精确度, 本文对这些手势图像做了进一步处理。

首先归一化得到手势图像, 然后用相应的原图像点乘所得到的阴影手势图像, 继而得到仅含有手势阴影的 RGB 手势图像, 该图像仅含有手势阴影和带有肤色信息的 RGB 手势图像, 为后期手势阴影的去除做铺垫。处理得到的手势图像如图 4 所示。



图4 RGB 手势图像

Fig.4 RGB gesture image

4 椭圆肤色模型与阴影消除

4.1 椭圆肤色模型与阴影消除原理

手势阴影是由于光源照射到背景的光线收到手势的阻挡而形成的, 故阴影会随着手势的运动而运动[8]。由于混合高斯背景模型是基于像素的操作, 手势阴影也被误检测为运动目标, 为后续的研究带来影响。受到光照等影响, 经过混合高斯模型检测出的动态手势中经常会含有手势阴影, 而在 $YCb'Cr'$ 颜色空间进行椭圆拟合建立的椭圆肤色模型可以有效地分割出肤色点与非肤色

点,手势阴影不属于肤色点,这样就可以有效地消除手势阴影,而保留手势图像,从而精确地检测出动态手势。

本文将上述得到的 RGB 手势图像中每一个像素点与已得到的椭圆肤色模型相匹配,以此来区分图像中的像素点与非像素点,继而消除手势阴影而保留手势图像,最终将得到的手势图像连成视频序列,该算法完成了动态手势的提取及其阴影的消除。本文将基于 YCbCr 颜色空间的混合高斯背景建模和基于 YCb'Cr'颜色空间椭圆肤色模型相结合,总体算法的系统框图如图 5 所示。

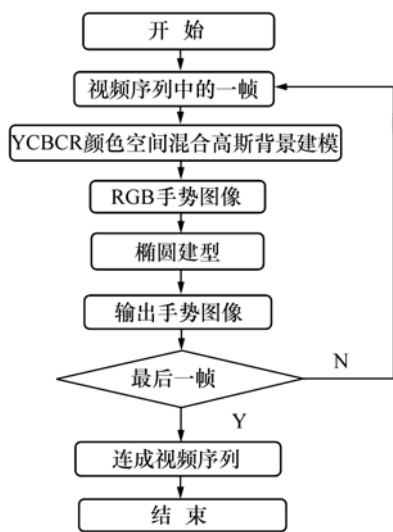


图5 算法系统框图

Fig. 5 System block diagram of algorithm

4.2 实验结果分析

本算法选择光照条件下的静态复杂背景,拍摄的运动手势视频大约 3.5 s,本文还对该视频运用传统的混合高斯模型进行运动手势检测,以作对比。实验环境为 Pentium (R) dual-core CPU 3.0GHZ,安装内存为 2.0 GB,操作系统为 WINDOW7,使用 MATLAB 2009b 进行仿真实验。经 MATLAB 仿真实验后,仿真结果如图 6 所示。

从仿真结果可以看出,经过基于 YCbCr 颜色空间的混合高斯背景建模检测出的结果中含有较大的噪声信息,并且阴影的存在已经严重影响到手势的精确性。经本算法,手势阴影被有效去除,并且能够较精确地保留手势图像的有效信息。

将实验使用的手势视频经过传统混合高斯模

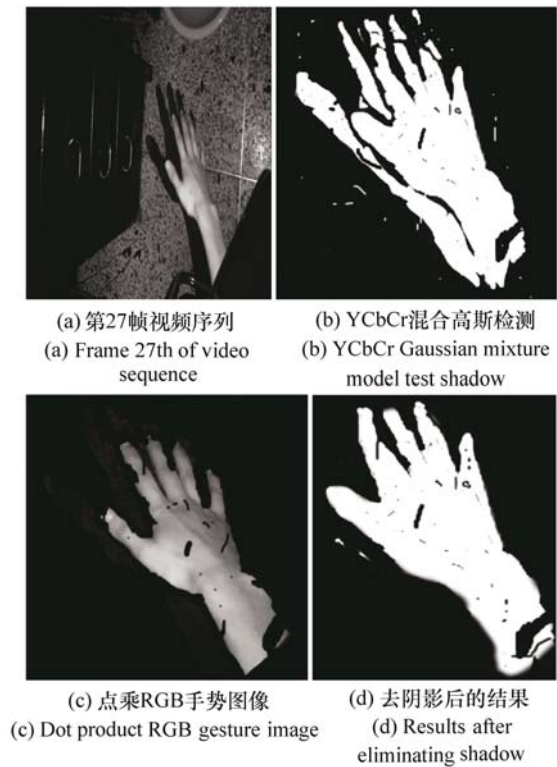


图6 仿真结果

Fig. 6 Simulation results

型进行仿真实验后,取视频中的第 27 帧进行对比,对比结果如图 7 所示。

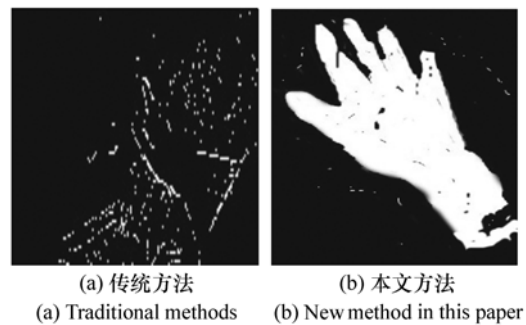


图7 对比实验结果

Fig. 7 Comparative results

经过对比实验,本算法与传统的混合高斯背景建模运动手势检测相比较,本算法可以较精确地检测出运动手势信息,而在该实验的复杂背景和阴影的干扰下,传统的检测方法基本不能检测出完整的手势信息,存在大量的误检区域,并且噪声特别严重。本实验运动手势视频时长约3.5 s,并拆分成 70 帧,然后使用不同的方法进行手势检

测。检测率 DR 定义如下:

$$DR = \frac{TP}{TP + FP}, \quad (7)$$

式中: TP 为正检帧数, FP 为错检帧数。表 1 为检测结果统计。

表 1 检测结果统计
Tab. 1 Statistics of test results

| | TP | FP | DR |
|----------|----|----|-------|
| 传统高斯模型检测 | 57 | 13 | 81.4% |
| 本文算法 | 64 | 6 | 91.4% |

经过对以上实验结果和数据进行分析, 本实验能够对检测出手势图像进行阴影的消除, 且能够精确地检测出运动手势, 有利于后续目标跟踪和识别方面的研究。

参考文献:

- [1] WEN J T, GAO Y D, HAN S, *et al.*. Gesture segmentation based on YCb'Cr' color space ellipse fitting skin color modeling[C]. Control and Decision Conference (CCDC), Taiyuan, China, 2012:1905-1908.
- [2] YU X C, YANG CH, YU T C. Hand skin color segmentation using online training artificial neural network[C]. Natural Computation (ICNC), Shenyang, China, 2013:242-246.
- [3] WEI W, JING P. Hand segmentation using skin color and background information[C]. Machine Learning and Cybernetics (CMLC), Xi'an, China, 2012:1484-1492.
- [4] AVINASH B D, GHOSH D K, ARI S. Color hand gesture segmentation for images with complex background[C]. International Conference on Circuits, Power and Computing Technologies (ICCPCT), Nagercoil, India, 2013:1127-1131.
- [5] YOU L ZH, CHENG H, JIA J CH. Face detection method based on multi-feature fusion in ycbcr color space[C]. Image and Signal Processing (CISP), Chongqing, China, 2012:1249-1252.
- [6] JING L, GENG W. A shadow detection method based improved gaussian mixture model[C]. Electronics Information and Emergency Communication (CEIEC), Beijing, China, 2013:62-65.
- [7] 杨秀芳, 戚银城, 李婷, 等. 基于 YCb'Cr' 肤色特征和 Ostu 算法的人眼定位优化方法[J]. 南京邮电大学学报(自然科学版), 2014, 34(1):99-102.
YANG X F, QI Y CH, LI T, *et al.*. Eye location optimization method based on YCb'Cr' color features and Ostu algorithm [J]. *Nanjing University of Posts and Telecommunications (Natural Science)*, 2014, 34(1):99-102. (in Chinese)
- [8] 邱鹏瑞. 基于 YCbCr 肤色模型改进算法及区域特征的人脸检测研究[J]. 计算机与现代化, 2012(4):179-182.
QIU P R. Face Detection YCbCr color model based on human and regional characteristics of the improved algorithm[J]. *Computer and Modernization*, 2012(4):179-182. (in Chinese)
- [9] AI Y Y, JING J L, AI X W. Color moving object segmentation based on mixture gaussian models[C]. Natural Computation (ICNC), Yantai, Shandong, China, 2010:1208-1211.
- [10] 黄建强, 曹腾飞, 郭文静, 等. 联合两种颜色空间的阴影检测算法[J]. 计算机技术与发展, 2014, 24(7):95-98.
HUANG J Q, CAO T F, GUO W J, *et al.*. shadow detection algorithm combined with two color spaces[J]. *Computer Technology and Development*, 2014, 24(7):95-98. (in Chinese)
- [11] 张红颖, 胡正. CenSure 特征和时空信息相结合的运动目标检测[J]. 光学精密工程, 2013, 21(9):2452-2462.
ZHANG H Y, HU ZH. Moving target detecte combinates CenSure characteristics and spatial temporal information[J].

5 结 论

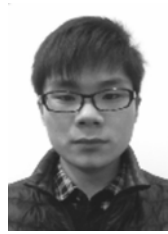
本文把基于 YCbCr 颜色空间的混合高斯背景建模运用到动态手势检测中, 并且结合 YCb'Cr' 颜色空间的椭圆肤色模型, 实现了复杂背景下动态手势的精确检测和手势运动检测中产生的阴影消除问题。本算法在复杂背景下的手势检测率可达 91.4%, 相比于传统混合高斯模型, 检测率提高了 10%, 有效提高了动态手势检测的精确度, 具有较高的实用价值。但由于混合高斯模型是基于像素的操作, 加上本文的阴影消除方法也是建立在像素的操作上, 运行速度较慢, 实时性较差, 所以进一步的研究将放在算法优化及实时性提高方面。

- Opt. Precision Eng.*, 2013, 21(9):2452-2462. (in Chinese)
- [12] 王梅,屠大维,周许超. SIFT 特征匹配和差分相乘融合的运动目标检测[J]. 光学精密工程, 2011, 19(4):892-899.
WANG M, TU D W, ZHOU X CH. Moving Object Detection combines SIFT feature matching and differential multiplied integration[J]. *Opt. Precision Eng.*, 2011, 19(4):892-899. (in Chinese)
- [13] 韩廷祥,张志胜,郝飞,等. 灰度序列图像中基于纹理特征的移动阴影检测[J]. 光学精密工程, 2013, 21(11):2931-2942.
HAN Y X, ZHANG ZH SH, HAO F, et al. . Grayscale image sequence moving shadow detection based on texture features [J]. *Opt. Precision Eng.*, 2013, 21(11):2931-2942. (in Chinese)
- [14] 卢官明,郎苏娟. 基于 YCbCr 颜色空间的背景建模及运动目标检测[J]. 南京邮电大学学报(自然科学版), 2009, 29(6):17-22.
LU G M, LANG S J. background modeling and moving target detection based on YCbCr color space[J]. *Nanjing University of Posts and Telecommunications (Natural Science)*, 2009, 29(6):17-22. (in Chinese)
- [15] 闫蓓,王斌,李媛. 基于最小二乘法的椭圆拟合改进算法[J]. 北京航空航天大学学报, 2008, 34(3):295-298.
YAN B, WANG B, LI Y. Least squares fitting based on the elliptical improved algorithm[J]. *J. Beijing University of Aeronautics and Astronautics*, 2008, 34(3):295-298. (in Chinese)

作者简介:



史东承(1959—),男,吉林长春人,教授,1982年于北京邮电学院获得学士学位,1988年于东北电力学院获得硕士学位,主要从事图像处理与机器视觉方面的研究。E-mail:dcshi@foxmail.com



倪康(1991—),男,山东枣庄人,硕士研究生,2013年于青岛滨海学院获得学士学位,主要从事图像处理与机器视觉方面的研究。E-mail:9500467@qq.com