

DOI: 10.7652/xjtuxb201409003

# 利用密度描述符对应的视觉跟踪算法

查宇飞<sup>1</sup>, 杨源<sup>2</sup>, 王锦江<sup>3</sup>, 张立朝<sup>1</sup>

(1. 空军工程大学航空航天工程学院, 710038, 西安; 2. 空军工程大学空管领航学院, 710038, 西安;  
3. 空军工程大学综合电子信息系统与电子对抗技术研究中心, 710038, 西安)

**摘要:** 针对稀疏描述符由于关键点检测不稳定及其误匹配造成跟踪失败的问题,提出了一种利用密度描述符对应实现目标跟踪的算法。该算法通过计算目标在相邻两帧之间的密度描述符流,考虑目标的空间分布特性和描述符的权重,得到目标的运动矢量,获得目标在当前帧中的估计;根据密度描述符的运动矢量与目标运动矢量的关系及其匹配程度,更新密度描述符的权重。在大量测试数据上对所提算法进行实验,并与相关跟踪算法进行比较,定性和定量分析表明:当发生光照、遮挡、姿态变化时,所提算法能够稳定地跟踪目标,跟踪成功率均在 90% 以上,跟踪误差要小于其他算法。

**关键词:** 目标跟踪;密度描述符;稀疏描述符

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 0253-987X(2014)09-0013-06

## A Visual Object Tracking Algorithm Using Dense Descriptors Correspondences

ZHA Yufei<sup>1</sup>, YANG Yuan<sup>2</sup>, WANG Jinjiang<sup>3</sup>, ZHANG Lichao<sup>1</sup>

(1. School of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an 710038, China; 2. School of Air Control and Navigation, Air Force Engineering University, Xi'an 710051, China; 3. Research Center for Integrated Electronic & Information System and Electronic Countermeasure Technology, Air Force Engineering University, Xi'an 710051, China)

**Abstract:** An object tracking algorithm is proposed to overcome the problem that the sparse local invariance feature descriptor depends on the feature detection and always leads to failure. The proposed algorithm uses the dense descriptors correspondences. The dense descriptor flows of objects between consecutive frames are calculated, and object estimations are obtained by considering the spatial distribution and weights of dense flows. Then the weights of descriptors are updated according to the relation and the matching degree between the descriptor motion and the object motion. Qualitative and quantitative analyses on challenging benchmark image sequences show that the average tracking successes rate of the proposed algorithm is over 90% and the performance of the algorithm is better than the performances of several state-of-art methods that assume the constant brightness and template.

**Keywords:** object tracking; dense descriptors; sparse descriptors

鲁棒视觉跟踪系统的关键在于寻找目标的不变特征,光流<sup>[1]</sup>假定目标在相邻图像之间的亮度保持恒定。然而在实际情况下,当目标发生一些微小变化时,如光照变化、目标形变或视频中噪音较大等,

收稿日期: 2014-04-01。 作者简介: 查宇飞(1979—),男,博士,副教授。 基金项目: 国家自然科学基金资助项目(61202339,61203628);陕西省自然科学基金资助项目(2012JQ8034)。

网络出版时间: 2014-07-03 网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20140703.1130.002.html>

容易丢失目标,很难获得满意的跟踪效果。颜色直方图特征<sup>[2]</sup>描述了颜色分布的恒常性,文献[3-4]分别在粒子滤波和均值漂移框架下,利用目标直方图不变特征来实现鲁棒的目标跟踪。然而,直方图本质上是一种统计特征,未考虑目标的结构信息。子空间跟踪算法<sup>[5]</sup>利用子空间的不变性,但当目标发生遮挡或姿态变化时,主成分比重降低,易丢失目标。尺度不变特征转换(SIFT)描述符<sup>[6]</sup>、BRIEF描述符<sup>[7]</sup>和DBrief描述符<sup>[8]</sup>等不变特征描述符,利用局部相关像素对关键点进行描述,使关键点具有尺度和旋转不变性。文献[9]利用SIFT算法检测到的稀疏关键点及其描述符构建相关图来表征目标,然而在实际中相关图的构建比较困难,从而导致跟踪失败,可将SIFT描述符作为均值漂移的度量实现跟踪<sup>[10]</sup>。

以上基于局部不变特征描述符的跟踪算法均属于稀疏描述符,关键点检测结果对跟踪结果影响很大。与稀疏描述符不同,密度描述符不需要检测关键点,可直接对每一个像素建立描述符,DAISY描述符<sup>[11]</sup>是2010年提出的一种密度描述符,由像素点附近的多个采样点所生成的卷积方向图向量组成,该方法对仿射变换和光照差异都有较好的鲁棒性,在稠密立体匹配中取得了较好的效果。Liu等人利用密度SIFT<sup>[12]</sup>寻求图像对应关系<sup>[13]</sup>,但密度SIFT没有考虑尺度的影响。为克服稀疏描述符在关键点检测中的不稳定性,本文提出了一种基于密度描述符的跟踪算法,避免了关键点检测不稳定及其误匹配带来的影响。所提算法通过计算目标在相邻两帧之间的密度描述符流,考虑目标的空间分布特性和描述符的权重,得到目标的运动矢量,获得目标在当前帧中的估计;根据密度描述符的运动矢量与目标运动矢量的关系及其匹配程度,更新密度描述符的权重。如此迭代,从而实现对感兴趣目标的鲁棒跟踪。

## 1 密度描述符及其运动矢量的计算

尺度不变特征转换是计算机视觉具有里程碑意义的工作,包含关键点的检测、描述和匹配3个步骤,但本文所提算法不检测关键点,只是对图像中像素点规则采样,直接构建图像的密度描述。SIFT描述符在检测到的关键点周围取一个 $16 \times 16$ 的邻域像素区域,并将邻域划分为16个子区域,每个子区域的直方图分布为

$$H_i(b) = \sum_{p=1}^N Q(I_i(p), b) \\ i = 1, 2, \dots, 16; b = 1, 2, \dots, 8 \\ Q(I(p), b) = \begin{cases} 1, & I(p) \in b \\ 0, & \text{其他} \end{cases} \quad (1)$$

式中: $N$ 表示该区域中的像素个数; $H_i$ 表示第 $i$ 个子区域的直方图,将图像分成16个子区域; $b$ 表示图像的灰度级,一般设为8个灰度级; $p$ 表示像素在图像中的位置; $Q(I(p), b)$ 为指示函数,若像素属于灰度级 $b$ ,则输出为1,否则为0。因此,每个子区域对应一个含有8个灰度级的直方图来描述该像素的局部描述符,然后将每个子区域直方图合并,得到一个128维的向量,即SIFT描述符

$$F_{\text{sift}} = [H_1, \dots, H_{16}] \in R^{128} \quad (2)$$

SIFT描述符在匹配时,要预先知道每个描述符的尺度和方向参数。SURF描述符<sup>[14]</sup>利用积分图像,降低SIFT描述符计算,获得接近SIFT描述符的效果。对于目标的描述,需首先检测目标中的关键点,然后对每个关键点进行描述,构成目标模型

$$M_s = \{(P_i, \sigma_i, \theta_i, F_i) \mid i = 1, 2, \dots, N\} \quad (3)$$

式中: $P_i$ 表示关键点的位置; $\sigma_i$ 表示关键点的尺度; $\theta_i$ 表示关键点的主方向; $F_i$ 表示描述符。然而,在实际中并不能总是在同一个位置、尺度和方向上得到相同的描述符,这样会导致错误的匹配,当检测到的特征很少时,少量的匹配错误将会使目标模型无法有效地表示目标。为了解决这一问题,本文提出用密度特征描述符表征目标模型,这种描述符不依赖特征检测,直接对目标中的规则像素点进行描述。

设原始图像 $I$ 的尺寸是 $M \times N$ ,构建的密度特征描述符记作 $F$ ,其维度为 $M \times N \times 128$ ,最后一维表示SIFT特征描述符的维数。由于密度特征描述符不是建立在特征检测的基础上,故没有尺度和方向信息。因此,目标模型表示为

$$M_d = \{F_{i,j,b} \mid i = 1, \dots, N, \\ j = 1, \dots, M, b = 1, \dots, B\} \quad (4)$$

在时刻 $t$ ,图像中位于 $(x, y)$ 处的像素值表示为 $I(x, y, t)$ ,而光流假设相邻两帧像素灰度值保持不变,即 $I(x, y, t) = I(x + u, y + v, t + 1)$ ,从而获得像素的运动矢量 $(u, v)$ 。假设相邻帧的描述符保持不变,即 $F(x, y, t) = F(x + u, y + v, t + 1)$ ,同样获得各个像素的运动矢量 $(u, v)$ 。特征描述符含有该像素的局部信息,而不仅仅考虑单个像素的信息。

取 $p = (x, y)$ 作为与图像等价的网格,取 $w(p) = (u(p), v(p))$ 作为 $p$ 点的运动向量,得约束密度

SIFT 描述符为

$$F(p, t) - F(p + w(p), t + \Delta t) = 0 \quad (5)$$

式中:  $F(p, t)$  为  $t$  时刻时  $p$  点的密度特征描述符。

$$E(w) = \underbrace{\sum_p (F(p + w(p), t + 1) - F(p, t))}_{\text{数据项}} + \underbrace{\sum_p \beta (|u(p)| + |v(p)|)}_{\text{位移项}} + \underbrace{\sum_{(p, q) \in \epsilon} \min(\alpha |u(p) - u(q)|, d) + \min(\alpha |v(p) - v(q)|, d)}_{\text{平滑项}} \quad (6)$$

式(6)中的能量函数包含数据项、位移项和平滑项 3 项。其中:数据项主要保证目标在相邻帧之间具有相同的描述符;位移项用来惩罚目标在相邻帧之间的运动量;平滑项用来约束目标内部的相邻像素应具有相同的运动矢量。在目标函数的位移和平滑项中引入阈值  $d$ , 通过简化  $L1$  范数, 消除外点匹配和流不连续性的影响, 其中  $\beta$  和  $\alpha$  分别为数据项和平滑项的系数, 代表对应项在整个能量函数中的比重。式(6)可以采用 SIFT 流<sup>[14]</sup>、PatchMatch<sup>[15]</sup>、可变空域金字塔匹配<sup>[16]</sup>等方法进行优化。

## 2 目标跟踪算法

通过上节的分析, 可以获得密度描述符的运动矢量, 而目标在视频中的位置不断发生变化, 同时自身的状态(如尺度、方向等)也不断发生变化, 因此要实现对目标的鲁棒跟踪, 需要解决 2 个问题: 如何由单个特征描述符的运动矢量去估计目标的运动矢量; 如何有效地更新目标模型, 以适应目标的变化。本节将在上节获得的密度描述符运动矢量的基础上, 详细描述跟踪的过程。

(1) 初始化。在初始化图像中, 手动标定目标在图像中的位置, 用矩形  $R = \{x, y, w, h\}$  来表示, 其中  $(x, y)$  表示矩形左上角的位置,  $(w, h)$  表示矩形的宽和高。目标图像为  $x = I(R)$ , 即矩形框所对应的图像。构建目标的密度描述符, 同时初始化每个描述符, 使其权重相同, 即  $w_{ij} = 1/(M \times N)$ 。

(2) 跟踪过程。当获得下一帧图像后, 在上一帧估计的目标位置上获取候选目标图像  $x_t = I_t(R_{t-1})$ , 同时获得对应的密度特征描述符  $F_t = \text{Dense}(x_t)$ 。通过计算相邻两帧目标的密度特征描述符对应关系, 最小化能量函数获得运动矢量

$$[u, v] = \arg \min E \quad (7)$$

式中:  $u, v$  为  $M \times N$  矩阵, 表示目标中每个像素点的运动矢量。整个目标的运动矢量还需要考虑目标本身的空间分布特性和特征描述符的重要性。本文采用文献<sup>[16]</sup>所提出的可变空域金字塔算法, 可快速

考虑到跟踪问题中目标运动的平滑性, 利用一阶马尔科夫随机场构造能量函数

实现相邻两帧目标的密度特征描述符对应关系。在跟踪过程中, 一般采用矩形框标识目标, 因此离矩形中心越近的像素, 属于目标的可能性越大, 同时不同位置的描述符对表征目标信息的贡献也不一样, 因此可采用下式来估计目标的运动矢量

$$\begin{aligned} u^* &= \sum_{i=1}^M \sum_{j=1}^N \exp\left(\frac{||p_{ij} - p_c||^2}{2\delta^2}\right) w_{ij} u_{ij} \\ v^* &= \sum_{i=1}^M \sum_{j=1}^N \exp\left(\frac{||p_{ij} - p_c||^2}{2\delta^2}\right) w_{ij} v_{ij} \end{aligned} \quad (8)$$

式中:  $p_c$  表示矩形框的中心位置;  $w$  表示描述符的重要度量;  $\delta$  为空域分布参数。可以看出, 目标的运动矢量由密度描述符流、空间分布和描述符的重要程度共同决定。

(3) 更新过程。在跟踪过程中需要适应不断变化的目标, 而每个描述符表征目标的重要程度不一样, 需要不断更新权重。权重更新如下

$$w_{ij} = \exp(\lambda (\|u^* - u_{ij}\|^2 + \|v^* - v_{ij}\|^2)) + (1 - \lambda) \|F_{ij}^* - F_{ij}\|^2 \quad (9)$$

式中:  $F_{ij}^*$  和  $F_{ij}$  分别表示估计后的特征描述符和上一帧的特征描述符;  $\lambda$  为 0 到 1 之间的数, 表示流和特征匹配的比重。最终的权重  $w$  是一个归一化常数, 表示密度描述符中每个特征描述符对表示目标的贡献。单个特征的运动矢量与目标的运动矢量越相近, 权重越大。

## 3 实验分析

所有实验都是在 Intel Core Duo 3.0 GHz、2 GB 内存的笔记本上进行的。算法通过 MATLAB 实现, 核心算法采用 MEX 文件, 其中计算 SIFT 流的算法采用文献<sup>[16]</sup>的公开代码, 能量函数构建中的参数取值为:  $\alpha = 2, \beta = 0.05, d = 40$ , 空域分布参数  $\delta$  取目标矩形框宽高均值的 1/4, 权重更新参数  $\lambda = 0.7$ 。跟踪过程中目标状态只考虑目标在二维空间的状态, 本文算法效率与目标框的大小相关, 平均为 20 帧/s。相对于稀疏描述符, 密度描述符不需要检测关键点, 得益于文献<sup>[16]</sup>提出的快速优化算法, 因

此大大加快了匹配速度,保证了本文算法的实时性。

实验数据为 MILT 数据库<sup>[17]</sup>和 PROST 数据库<sup>[18]</sup>中的 8 段视频,目标真实状态通过相应的数据库获得,目标初始状态由数据库给出。比较算法包括基于子空间不变的跟踪算法 (IVT)、基于正反向光流的跟踪算法 (Median Flow)<sup>[19]</sup>和基于稀疏不变特征的稀疏描述符<sup>[20]</sup>。IVT 跟踪和 Median Flow 跟踪的代码分别来自公开的代码,稀疏描述符的代码是由 VLFEAT 和 RANSAC 工具箱<sup>[21]</sup>编写的。

### 3.1 定性分析

图 1 是部分实验结果图,IVT 跟踪算法、Median Flow 跟踪算法、稀疏描述符跟踪算法和本文所提算法 (DDC 算法) 分别用虚线、点线、点划线和实线的矩形框表示。下面将从光照变化、姿态变化和部分遮挡 3 个方面对实验结果进行定性分析。

(1) 光照变化。在视频“Sylvester”和“David”中,都有不同程度的光照变化,特别是“David”中,目标由暗处走向亮处,光照变化较大,光照变化使目标灰度值突然变化,导致目标外观改变,增加了跟踪的难度。基于稀疏描述符的跟踪依赖特征点的检测从而导致跟踪的不稳定。IVT 跟踪受光照的影响,导致正样本更新产生误差,而误差积累会导致跟踪结果偏离真实目标。Median Flow 跟踪是利用光流在相邻两帧之间预测目标运动矢量,在光照变化不大的情况下,能够较好地跟踪目标。密度特征描述符利用相邻像素相对分布关系,而目标整体灰度的变化不会改变相对分布,从而避免了光照变化带来的影响,故能稳定地跟踪目标。

(2) 姿态变化。视频“Girl”、“Jumping”、“Lemming”和“Box”中,目标发生了大的姿态变化。虽然目标的变化很大,但是它们都是渐变的,也就是目标在时序上具有连续性,这将考验跟踪算法的在线学习能力。IVT 跟踪是每隔 5 帧更新,对特征基向量进行更新,不能适应目标的快速变化;Median Flow 跟踪只考虑相邻两帧之间的情况,误差积累导致跟踪结果偏离真实的目标;基于稀疏描述符的跟踪能适应目标大的变化,但误匹配和特征检测的不稳定会导致跟踪失败;本文的算法由于采用密度特征描述符,能够适应目标的变化,同时利用动态特征描述符,从而消除了误差积累的影响。

(3) 部分遮挡。视频“Faceocc”和“Faceocc2”主要测试遮挡对算法的影响,其中“Faceocc2”中的目标还发生了旋转变化的。受遮挡影响最大的是基于稀疏描述符的跟踪算法,由于目标遮挡后无法检测到



(a)“Sylvester”序列的跟踪结果(从左到右第 304、400、610 帧)



(b)“David”序列的跟踪结果(从左到右第 60、150、200 帧)



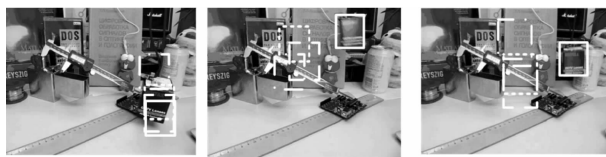
(c)“Girl”序列的跟踪结果(从左到右第 2、352、462 帧)



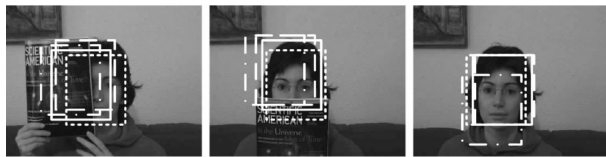
(d)“Jumping”序列的跟踪结果(从左到右第 20、180、300 帧)



(e)“Lemming”序列的跟踪结果(从左到右第 60、350、410 帧)



(f)“Box”序列的跟踪结果(从左到右第 151、351、396 帧)



(g)“Faceocc”序列的跟踪结果(从左到右第 100、300、600 帧)



(h)“Faceocc2”序列的跟踪结果(从左到右第 150、425、600 帧)

— — — IVT    ..... Median flow    - · - Sparse    —— DDC

图 1 跟踪结果

特征,从而导致误匹配;IVT 跟踪是将目标整体当作向量看待,遮挡使得特征基模糊,偏离了真实目标;Median Flow 跟踪由于只用匹配较好的点,则目标部分遮挡对算法的影响不是很大;本文算法基于密度特征,使得部分遮挡并不能导致匹配失败。因此,在目标发生各种变化和背景干扰很大的情况下,本文算法能够很好地跟踪目标,与 Median Flow 算法、IVT 算法和稀疏描述符算法相比,能够更好地处理光照、姿态变化和遮挡的影响。

3.2 定量分析

本文用重叠度、均方根误差 2 个标准来衡量本文算法与其他算法的优劣。重叠度是检测目标与真实目标重叠部分与两者的并集之比,即

$$o = \frac{\hat{S} \cap S}{\hat{S} \cup S} \tag{10}$$

式中: $\hat{S}$  和  $S$  分别表示估计值和真实值。当检测目标与真实目标完全重叠时,重叠度为 1;当检测到错误目标时,重叠度会降低;当完全检测不到目标时,重叠度为 0。若重叠度大于 0.5,可认为跟踪成功,否则认为目标丢失。

均方根误差的计算如下

$$E_{\text{RMS}} = ((x - \hat{x})^2 + (y - \hat{y})^2)^{1/2} \tag{11}$$

式中: $(x, y)$  为算法得出的跟踪框的中心位置坐标; $(\hat{x}, \hat{y})$  为标定的跟踪框的中心位置坐标。均方根误差表示了目标中心与真实目标中心的误差,值越小表明跟踪精度越高。

表 1 和表 2 分别为 4 种算法在各个视频上的跟踪成功率和均方根误差。加粗和加线的表示在该视频中的成功率排名第 1 和第 2。与基于子空间不变的 IVT 和基于灰度不变的 Median Flow 相比,本文算法获得了更好的结果,部分原因在于 SIFT 描述符能更好地描述变化目标的不变属性。与稀疏描述方法相比,本文算法由于不需要检测关键点和计算仿射矩阵,同样获得了更好的结果。

表 1 不同算法的跟踪成功率

| 视频名称     | 跟踪成功率       |             |             |             |
|----------|-------------|-------------|-------------|-------------|
|          | ST          | IVT         | MF          | DDC         |
| Syly     | 0.47        | 0.81        | <u>0.97</u> | <b>0.98</b> |
| David    | 0.48        | <b>0.98</b> | 0.30        | <u>0.96</u> |
| Girl     | 0.82        | 0.21        | 0.94        | <u>0.92</u> |
| Jumping  | 0.23        | 0.51        | <b>1</b>    | <u>0.99</u> |
| Lemming  | 0.78        | 0.79        | <u>0.90</u> | <b>0.92</b> |
| Box      | 0.82        | 0.70        | <u>0.83</u> | <b>0.91</b> |
| Faceocc  | 0.95        | <b>0.98</b> | 0.78        | <u>0.95</u> |
| Faceocc2 | <u>0.97</u> | <b>0.98</b> | 0.91        | 0.96        |

表 2 不同算法的均方根误差比较

| 视频名称     | $E_{\text{RMS}}$ |       |              |              |
|----------|------------------|-------|--------------|--------------|
|          | ST               | IVT   | MF           | DDC          |
| Syly     | <u>14.65</u>     | 60.85 | 20.67        | <b>13.50</b> |
| David    | 38.76            | 21.45 | <b>8.65</b>  | <u>9.02</u>  |
| Girl     | <b>25.34</b>     | 48.09 | 123.56       | <u>27.03</u> |
| Jumping  | <u>10.35</u>     | 45.36 | 18.67        | <b>8.54</b>  |
| Lemming  | <u>10.45</u>     | 16.34 | 15.67        | <b>9.76</b>  |
| Box      | 60.34            | 35.12 | <u>20.61</u> | <b>9.78</b>  |
| Faceocc  | 18.76            | 14.78 | <b>8.89</b>  | <u>10.45</u> |
| Faceocc2 | 13.98            | 18.49 | <u>8.76</u>  | <b>7.75</b>  |

4 结 论

鲁棒视觉跟踪的一个关键问题在于如何快速有效地描述视频中的目标,本文算法通过计算目标在相邻两帧之间的密度描述符流,估计得到目标的运动矢量,同时更新密度描述符的权重。在实验中,针对光照变化、姿态变化和遮挡等情况,对本文算法和其他相关算法进行了对比实验,结果表明本文算法能够很好地适应复杂的变化,能取得比其他算法更好的结果。本文算法的运行效率需要进一步提升,主要原因在于密度描述符对应的计算量较大。

参考文献:

[1] SENST T, EISELEIN V, SIKORA T. Robust local optical flow for feature tracking [J]. IEEE Transactions on Circuits and Systems, 2012, 22(9): 1377-1387.

[2] LEE K Y, PARK R H, LEE S W. Color matching for soft proofing using a camera [J]. IET Image Process, 2012, 6(3): 292-300.

[3] PEREZ P, HUE C, VERMAAK J, et al. Color-based probabilistic tracking [C]//Proceedings of 7th European Conference on Computer Vision. Berlin, Germany: Springer, 2002: 661-675.

[4] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.

[5] ROSS D, LIM J, LIN Ruei-Sung, et al. Incremental learning for robust visual tracking [J]. International Journal of Computer Vision, 2007, 77(1): 125-141.

[6] DAVID G L. Scale & affine invariant interest point detectors [J]. International Journal of Computer Vision, 2004, 60(1): 63-86.

[7] CALONDER M, LEPETIT V, FUA P. Binary robust

- independent elementary features [C]//Proceedings of 11th European Conference on Computer Vision. Berlin, Germany: Springer, 2010: 778-792.
- [8] HEINLY J, DUMN E, FRAHM J M. Comparative evaluation of binary features [C]//Proceedings of 12th European Conference on Computer Vision. Berlin, Germany: Springer, 2012: 369-382.
- [9] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [10] ZHOU Huiyu, YUAN Yuan, SHI Chunmei. Object tracking using SIFT features and mean shift [J]. Computer Vision and Image Understand, 2009, 113(3): 345-352.
- [11] ENGIN T, VINCENT L, PASCAL F. An efficient dense descriptor applied to wide baseline stereo [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(5): 815-830
- [12] VEDALDI A, FULKERSON B. An open and portable library of computer vision algorithms [C]//Proceedings of the International Conference on Multimedia. New York, USA: ACM, 2010: 1469-1472
- [13] LIU Ce, JENNY Y, ANTONIO T, et al. Dense correspondence across different scenes [C]//Proceedings of 9th European Conference on Computer Vision. Berlin, Germany: Springer, 2008: 28-42.
- [14] BAY H, TUYTELAARS T, GOOL L V. Speeded up robust features [C] // Proceedings of 7th European Conference on Computer Vision. Berlin, Germany: Springer, 2006: 404-417
- [15] BARNES C, SHECHTMAN E, GOLDMAN D, et al. The generalized PatchMatch correspondence algorithm [C]//Proceedings of 11th European Conference on Computer Vision. Berlin, Germany: Springer, 2010: 29-43.
- [16] KIM J, LIU Ce, SHA Fei, et al. Deformable spatial pyramid matching for fast dense correspondences [C]//26th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2013: 2307-2314.
- [17] BABENKO B, YANG Ming-Hsuan, BELONGIE S. Visual tracking with online multiple instance learning [C]//24th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2009: 983-990.
- [18] SANTNER J, LEISTNER C, SAFFARI A, et al. Parallel robust online simple tracking [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2010: 723-730.
- [19] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking learning detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 6(1): 1-14
- [20] 蔺海峰, 马宇峰, 宋涛. 基于 SIFT 特征目标跟踪算法研究 [J]. 自动化学报, 2010, 36(8): 1204-1208.
- LIN Haifeng, MA Yufeng, SONG Tao. Research on object tracking algorithm based on SIFT [J]. Acta Automatica Sinica, 2010, 36(8): 1204-1208.
- [21] ZULIANI M, KENNEY C S, MANJUNATH B S. The multiRANSAC algorithm and its application to detect planar homographies [C]//Proceedings IEEE International Conference on Image Processing. Piscataway, NJ, USA: IEEE, 2005: 153-156.

(编辑 赵炜)

(上接第5页)

- [8] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2001: 511-518.
- [9] MEYNET J. Fast face detection using AdaBoost pattern recognition[EB/OL]. [2003-07-16]. [http://ftp.utcluj.ro/pub/users/nedeveschi/AV/References/B3\\_Face%20detection/Meynet 2003\\_923.pdf](http://ftp.utcluj.ro/pub/users/nedeveschi/AV/References/B3_Face%20detection/Meynet%2003_923.pdf).
- [10] 赵锋, 薛惠锋, 王伟, 等. 采用多级分类器的手写数字识别技术研究 [J]. 微电子学与计算机, 2006, 30(12): 30-32.
- ZHAO Feng, XUE Huifeng, WANG Wei, et al. Handwritten digit recognition technology based on multistage classifying device [J]. Microelectronics & Computer, 2006, 30(12): 30-32.

(编辑 赵炜)