

# 基于马尔科夫随机场的多特征人脸跟踪算法

蔡荣太<sup>1,2</sup>, 朱 鹏<sup>1,2</sup><sup>1</sup> 福建师范大学光电与信息工程学院, 福建 福州 350007;<sup>2</sup> 福建师范大学医学光电科学与技术教育部重点实验室, 福建 福州 350007

**摘要** 为实现稳健和精确的人脸跟踪,充分挖掘了人脸中的颜色信息、梯度方向信息和空间结构信息。在人脸中提取眼睛、鼻子和嘴巴等显著特征子块作为跟踪子块,在每个子块中选择最显著的特征作为跟踪的依据,并用马尔科夫随机场建立各个子块之间的空间约束关系,实现稳健的人脸跟踪。和若干典型跟踪算法的比较,实验结果表明,所提出的跟踪算法具有较好的稳健性和精确性。

**关键词** 图像处理; 人脸跟踪; 马尔科夫随机场; 多特征跟踪; 分块跟踪; 粒子滤波

**中图分类号** TP391.4      **文献标识码** A

**doi:** 10.3788/LOP54.021002

## Face Tracking with Multi-Feature Based on Markov Random Field

Cai Rongtai<sup>1,2</sup>, Zhu Peng<sup>1,2</sup><sup>1</sup> College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou, Fujian 350007, China;<sup>2</sup> Key Laboratory of Optoelectronic Science and Technology for Medicine of Ministry of Education, Fujian Normal University, Fuzhou, Fujian 350007, China

**Abstract** To achieve a robust and precise face tracking, the color information, gradient direction information and spatial structure information of face are fully exploited. The eyes, nose and mouth patches are employed as tracking regions from human face. The dominant features in these patches are extracted as the basis for tracking. Markov random fields are used to build the spatial constraints between these patches, and a robust tracking algorithm is realized. Experimental results show that, compared with several typical tracking algorithms, the proposed algorithm has well performance in robustness and precision.

**Key words** image processing; face tracking; Markov random field; multi-feature tracking; part-based tracking; particle filter

**OCIS codes** 100.4999; 150.1135; 100.2000

## 1 引 言

脸部是人类最显著的特征之一,是身份鉴别的重要依据。在表情识别、智能人机交互和远程视频会议等领域具有广泛的应用。而人脸总是会随着人体的运动而移动。因此,在实现上述应用之前,必须对人脸进行有效的跟踪定位。

视觉跟踪涉及到两个重要的技术:1)目标的描述,即在目标中抽取某种特征作为目标的描述,如纹理<sup>[1]</sup>、梯度<sup>[2]</sup>和颜色<sup>[3]</sup>等;2)目标搜索策略,即以一定的方法在图像中寻找目标的位置和姿态,如粒子滤波(PF)<sup>[4]</sup>、核跟踪方法<sup>[5]</sup>等。其中,目标的描述需要克服目标形变、光照变化、物体干扰等问题。为实现稳健的目标跟踪,通常采用多特征进行目标描述。比如文献[1]基于纹理描述实现目标初步定位,再由多维特征时空上下文模型实现目标的精确定位,达到稳健跟踪。郭鹏宇等<sup>[6]</sup>基于机器学习的思想和外观信息,提出一种在线选择纹理和形状特征的混合随机朴素贝叶斯视觉跟踪器,但是多特征表达没有解决目标受到局部遮

收稿日期: 2016-08-25; 收到修改稿日期: 2016-10-25

基金项目: 国家自然科学基金(61179011)、福建省自然科学基金(2014J01224)

作者简介: 蔡荣太(1979—),男,博士,副教授,主要从事图像处理和识别方面的研究。E-mail: gjrtcai@163.com

挡的问题。因此,又提出了利用多个子块描述目标或者分情况描述目标的方法,如 Adam 等<sup>[7]</sup>提出了 Fragtrack 方法,即在水平和垂直方向上,将目标分割成若干大小相等的矩形块,再计算每个子块的颜色直方图,然后与当前候选区域相应子块的颜色直方图进行匹配,最终通过加权投票的方式确定当前帧目标的中心位置。仇春春等<sup>[8]</sup>对目标外观采用混合表示方法,若目标未被遮挡则使用增量主成分分析与高斯观测噪声进行表示,反之则使用连续均匀概率分布进行表示,对混合模型进行能量最小化求解来实现对目标的跟踪。这些算法可以在一定程度上解决目标受到局部遮挡的问题,但是如果使用单一特征,当该特征失效时,也不能维持目标的持续跟踪。

为克服上述方法的局限性,同时考虑到人脸的特性,本文设计了一种分块多特征的人脸跟踪方法。以人脸中的眼睛、鼻子和嘴巴作为跟踪子块。和传统的多特征跟踪方法不同,本文在每个子块中仅抽取最显著的一种特征作为目标子块的描述,避免无效特征降低目标跟踪的精度。然后以这些图像子块的空间结构约束关系为依据,估计目标的真实状态。不仅充分利用了特征块构成的空间结构信息,而且避免了某图像块被遮挡引起的跟踪偏移。

## 2 人脸子块及其显著特征

多子块目标跟踪可以避免在单一图像块跟踪时,目标受到局部遮挡,整个图像块的计算都受到影响的缺点,有利于提高目标跟踪的稳健性。眼睛、鼻子、嘴巴和耳朵等器官是人脸中的显著特征块。但是,耳朵容易被头发遮挡,不适合作为人脸跟踪的子块。因此,本文采用眼睛、鼻子和嘴巴作为人脸跟踪的子块。

为提高目标描述的稳健性,通常采用多特征描述方法。即同时提取和采用多个特征,利用加权等方法融合多种特征,提高跟踪的稳健性。但是这种方法,无效特征和有效特征同时存在,无效特征找到的目标位置可能严重偏离真实目标位置,而有效特征找到的目标位置可能是准确的。两者加权融合后,最终的目标位置会严重偏离准确目标位置,降低目标跟踪的精度。与这种常规多特征融合方法不同,本文在每个图像块中只提取一种特征。由于眼珠和眉毛的存在,人眼区域的颜色和人脸其他区域的肤色有显著的区别,而边缘轮廓的统计分布和嘴巴等区域的分布则比较相似。因此,称前者为显著特征,后者为非显著特征,并采用颜色直方图(HoC)描述眼睛区域图像块。而人的鼻子和嘴巴通常在颜色上和脸部其他区域的颜色没有显著的区别,但是边缘轮廓是显著不同的。因此,采用梯度方向直方图(HoG)描述鼻子和嘴巴区域的图像子块。

## 3 基于子块的人脸马尔科夫随机场表示

显然,各个子块并不是人脸的完整描述,采用马尔科夫随机场(MRF)<sup>[8]</sup>对人脸进行建模。具体如下:将各个子块用 MRF 的节点表示,子块间的空间约束关系用 MRF 节点  $i$  和节点  $j$  之间的边  $e_{ij}$  表示。每个节点包括状态和观测,状态用  $s_i = \{x_i, y_i, \Delta x_i, \Delta y_i\}$  表示。其中  $(x_i, y_i)$  是第  $i$  个子块的中心位置,  $(\Delta x_i, \Delta y_i)$  是第  $i$  个子块的宽度和高度。用每个子块的显著特征直方图表示(眼睛用 HoC,其他子块用 HoG),整个人脸的状态由这些子块状态的联合  $s = (s_1, s_2, \dots, s_i)$  表示。节点状态  $s_i$  和观测  $o_i$  的联合概率分布为<sup>[9]</sup>

$$p(s_i, o_i) = \frac{1}{c} \prod_{i,j \in Z} \psi(s_i, s_j) \prod_{i \in V} \varphi(s_i, o_i), \quad (1)$$

式中  $c$  是归一化常数,  $\psi(s_i, s_j)$  是节点  $i$  和节点  $j$  的相关势函数,描述两个节点之间的空间约束关系,即 MRF 的结构模型;  $\varphi(s_i, o_i)$  是节点  $i$  的似然函数,描述节点  $i$  的状态和观测之间的关系,即 MRF 的观测模型。

势函数定义为

$$\psi(s_i, s_j) = \exp\{-K [d(s_i, s_j) - \xi_{ij}]^2\}, \quad (2)$$

式中  $\xi_{ij}$  为常数,是节点  $i$  和节点  $j$  的初始位置关系;  $K$  为常量,用来调节图模型中边的弹性;  $d(s_i, s_j)$  是结合了空间位置信息和子块尺度大小的度量函数,用欧式距离度量。势函数用于计算节点空间约束下,某个节点的可信度。

似然函数定义为

$$\varphi(s_i, o_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{d^2(t_i, q_i)}{2\sigma^2}\right], \quad (3)$$

式中  $d(t_i, q_i)$  表示模板中第  $i$  个子块的特征  $t_i$  与待匹配图像中节点  $i$  的特征  $q_i$  之间的相似性, 用欧式距离表示;  $\sigma$  表示方差。似然函数用于计算获得观测条件下, 某个节点的可信度。

采用反向传播(BP)算法<sup>[9]</sup>来计算节点的信任度。BP算法通过迭代的方式, 利用节点之间相互传递信息来更新整个 MRF 的标记状态, 直到所有的节点信任度不再发生变化。在 BP 算法的第  $n$  次迭代中, 节点  $j$  发送一个消息到节点  $i$  为

$$m_{ji}(n) = \sum_{s_j} \psi(s_j, s_i) \varphi_j(s_j, o_j) \prod_{\Gamma(j) \setminus i} m_{ji}(n-1), \quad (4)$$

式中  $\Gamma(j) \setminus i$  是  $j$  的邻域节点集, 但是扣除节点  $i$ 。

综合节点  $i$  所得到的邻域消息后, 可得到节点  $i$  的信任度为

$$b(s_i) = \alpha \varphi(s_i, o_i) \prod_{\Gamma(i) \setminus i} m_{ji}(n), \quad (5)$$

式中  $\alpha$  为归一化系数。

各节点的信任度取决于多方面因素, 比如受遮挡的程度、姿态变化的大小、光照变化的程度等。不同子块在确认人脸最终状态中的权重应该由它们的信任度决定, 第  $i$  个子块的权重更新为

$$\omega(s_i)(n) = \exp[b(s_i)] \omega(s_i)(n-1). \quad (6)$$

## 4 基于多子块多特征马尔科夫随机场的人脸跟踪

获得各个节点的权重之后, 可对整个人脸的状态进行估计。在此之前, 本文采用粒子滤波器(PF)跟踪各个图像子块, 获得各个图像子块的状态, 并以它作为 MRF 计算的依据。整个人脸跟踪算法如下。

初始化: 在人脸中提取眼睛、鼻子、嘴巴等图像子块。

初始化每个子块的状态和整个人脸的状态。

初始化各个粒子及其权重, 初始化所有粒子的权重为 1。

初始化 MRF 的各个参数, 初始化所有节点的信息为 1, 权重为 1。

for  $k =$  第 2 帧图像: 最后一帧图像;

for  $i =$  图像块 1: 最后一个图像块;

1) 用 PF 算法估计图像块  $i$  的状态;

2) 根据(2)~(6)式, 利用 MRF 计算图像块  $i$  的权重。

End;

根据各个子块的权重更新人脸状态。

End;

## 5 实验结果和分析

首先采用 FaceOcc1 人脸图像序列的跟踪来验证本文算法的有效性。该序列存在镜头移动、局部遮挡等干扰, 序列长度为 500 帧, 图像大小为 352 pixel  $\times$  288 pixel, 目标大小为 96 pixel  $\times$  150 pixel。参数设定如下: 目标位置方差为 5, 尺度方差为 0.1, 每个子块用 15 个粒子跟踪。

部分结果如图 1 所示。由图 1 可见, 第 29 帧, 同时存在遮挡与镜头移动两种干扰的情况下, 本文算法不

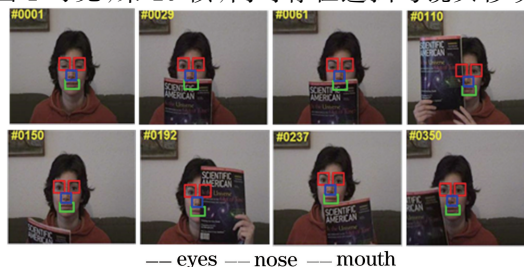


图 1 FaceOcc1 序列跟踪结果

Fig. 1 Tracking results of FaceOcc1 sequence

仅可以定位眼睛、鼻子等未被遮挡部位的位置信息,还可以根据未遮挡部位确定嘴巴的位置;第 110 帧和 192 帧,干扰物从左侧或右侧出现,遮住人脸的三分之一时,本文算法还可以继续维持跟踪。

图 2 给出了单特征跟踪算法、多特征融合跟踪算法与本文算法的跟踪误差曲线。由图 2 可见,和单特征跟踪算法以及传统多特征融合跟踪算法相比,本文算法具有较好的跟踪精度。

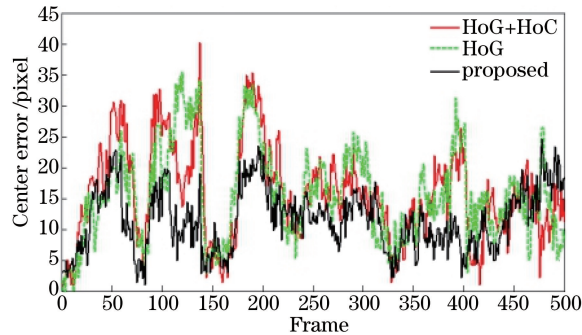


图 2 FaceOcc1 序列跟踪误差曲线

Fig. 2 Tracking errors of FaceOcc1 sequence

此外,将本文算法和块跟踪算法(Frag)<sup>[7]</sup>、分布场跟踪算法(DFT)<sup>[10]</sup>、稳健的超分辨率跟踪算法(RS)<sup>[11]</sup>等三种具有代表性的跟踪算法进行比较,验证本文算法的有效性。采用的图像序列是 David1 序列,该序列具有光照变化和姿态变化等干扰情况,序列大小为 320 pixel×240 pixel,序列长度 471 帧。算法参数设置同前。部分结果如图 3 所示。

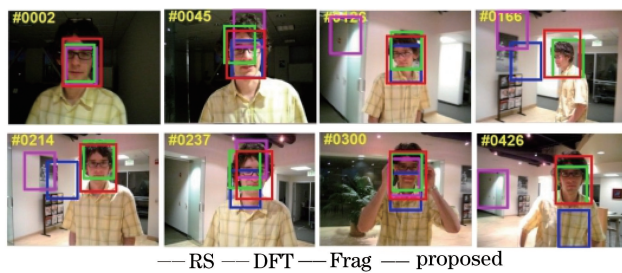


图 3 David1 序列跟踪结果

Fig. 3 Tracking results of David1 sequence

从图 3 可以看出,在跟踪过程中,Frag 算法由于光照变化最早失去跟踪目标,而其他算法对于光照变化具有良好的稳健性;在第 166 帧中,人物做了一个侧身动作导致 DFT 算法跟踪失败,而 RS 算法会有一点点的中心偏移,本文算法跟踪稳定;在后续几帧中,目标人物做了摘掉眼镜和重新戴上眼镜的动作,本文算法可保持良好的跟踪。

图 4 所示为跟踪误差曲线。由图 4 可见,和其他三种代表性跟踪算法比较,本文算法较为稳健和精确。

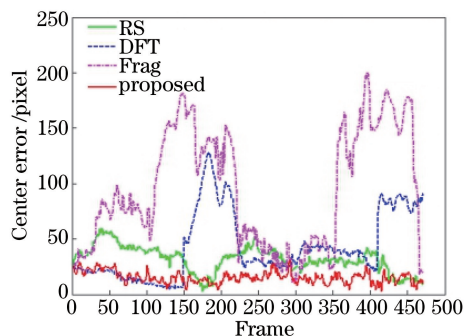


图 4 David1 序列跟踪误差曲线

Fig. 4 Tracking error curves of David1 sequence

## 6 结 论

设计了一种基于多子块多特征的人脸描述方法,利用 MRF 建立各个子块的空间关系,利用粒子滤波器跟踪每个图像子块,根据粒子滤波器的跟踪结果和 MRF 的空间约束关系,估计目标的真实状态。有效地克服单一特征、单一跟踪块的局限性,充分利用了人脸中多个子块的空间结构信息、颜色信息和梯度方向信息,实现比较稳健和精确的人脸跟踪。

## 参 考 文 献

- [1] Liu Wei, Zhao Wenjie, Li Cheng. Long-term visual tracking based on spatio-temporal context[J]. Acta Optica Sinica, 2016, 36(1): 0115001.  
刘 威, 赵文杰, 李 成. 时空上下文学习长时目标跟踪[J]. 光学学报, 2016, 36(1): 0115001.
- [2] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE, 2005, 1: 886-893.
- [3] Wang J, Yagi Y. Integrating color and shape-texture features for adaptive real-time object tracking [J]. IEEE Transactions on Image Processing, 2008, 17(2): 235-240.
- [4] Gordon N J, Salmond D J, Smith A F M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation [J]. IEEE Proceedings F-Radar and Signal Processing, 1993, 140(2): 107-113.
- [5] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(5): 564-575.
- [6] Guo Pengyu, Su Ang, Zhang Hongliang, *et al.* Online mixture of random naive Bayes tracker combined texture with shape feature[J]. Acta Optica Sinica, 2015, 35(3): 0315002.  
郭鹏宇, 苏 昂, 张红良, 等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器[J]. 光学学报, 2015, 35(3): 0315002.
- [7] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, 1: 798-805.
- [8] Qiu Chunchun, Li Qingwu, Wang Tian, *et al.* An improved IVT algorithm for object tracking [J]. Laser & Optoelectronics Progress, 2016, 53(1): 011002.  
仇春春, 李庆武, 王 恬, 等. 一种改进的 IVT 目标跟踪算法[J]. 激光与光电子学进展, 2016, 53(1): 011002.
- [9] Koller D, Friedman N. Probabilistic graphical models: principles and techniques[M]. Massachusetts: The MIT Press, 2009.
- [10] Sevilla-Lara L, Learned-Miller E. Distribution fields for tracking [C]. IEEE Conference on Computer Vision and Pattern Recognition, Providence, 2012: 1910-1917.
- [11] Yang F, Lu H, Yang M, *et al.* Robust superpixel tracking[J]. IEEE Transactions on Image Processing, 2014, 23(4): 1639-1651.