

文章编号: 1006-4710(2014)02-0200-04

一种基于颜色直方图的人体部位外观模型

韩贵金^{1,2}, 朱虹¹

(1. 西安理工大学 自动化与信息工程学院, 陕西 西安 710048;

2. 西安邮电大学 自动化学院, 陕西 西安 710121)

摘要: 为提高单幅静态图像中人体姿态估计的准确度, 提出了一种仅利用单幅静态图像建立的基于颜色直方图的人体部位外观模型。模型的建立主要包括三个步骤, 即: ① 计算人体部位减小后状态空间中的每个状态所对应图像区域的 HOG 特征与基于 HOG 特征的部位外观模型的似然度; ② 利用似然度较高的状态学习定位概率; ③ 根据学习得到的定位概率建立基于颜色直方图的人体部位外观模型。仿真实验表明当将所提人体部位外观模型用于单幅静态图像的人体姿态估计时可以得到更好的估计效果。

关键词: 人体部位外观模型; 颜色直方图; 人体姿态估计

中图分类号: TP242.6 **文献标志码:** A

A kind of human part appearance model based on color histogram

HAN Guijin^{1,2}, ZHU Hong¹

(1. Faculty of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China;

2. School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

Abstract: For improving the accuracy of human pose estimation from a single static image, an human part appearance model based on color histogram is proposed just using a single static image. There are three steps in the process of establishing the part appearance model: ① the local appearance likelihood of HOG feature corresponding with each state of the reduced part state space with the HOG-based part appearance model is computed; ② the location probability of human part is learned from states with higher local appearance likelihood; ③ the part appearance model based on color histogram is established on the basis of location probability. When the proposed human part appearance model is used to human pose estimation from a single static image, the simulation experiments show it can get better estimation effect.

Key words: human part appearance model; color histogram; human pose estimation

人体姿态估计是通过对静态图像或视频帧图像进行搜索以确定人体各部位最佳姿态的过程。根据人体部位姿态分布空间维数的不同, 人体姿态估计可分为二维和三维两种类型。二维人体姿态估计一般使用线段或者矩形等描述人体部位在图像二维平面的投影, 三维人体姿态一般使用关节树模型来描述人体部位在三维空间中的位置和角度信息, 而二维人体姿态估计又往往是三维人体姿态估计的第

一步^[1]。

在对静态图像进行人体姿态估计的过程中面临着诸多困难, 主要有:

- 1) 不同图像中光照条件的不同和人体形体的不同会导致人体部位外观特征的变化;
- 2) 人体部位的自由转动可能会导致部位之间出现相互遮挡;
- 3) 描述人体部位姿态的状态所在空间的维数

收稿日期: 2013-10-10

基金项目: 国际科技合作项目子项目基金资助(2011DRF10480)。

作者简介: 韩贵金, 男, 讲师, 博士生, 研究方向为智能视频监控。E-mail: hgjin123@126.com。

朱虹, 女, 博士, 教授, 研究方向为数字图像处理。E-mail: zhuhong@xaut.edu.cn。

较高且状态空间很大。

因此,人体姿态估计成为了计算机视觉领域中一个非常具有挑战性的课题^[2],是近年来一个非常热门的研究方向,为此,已有多种人体姿态估计算法^[2-8]被提出。

图结构模型^[3-9]是适用于部位之间通过关节相连的各种运动目标的姿态估计的一种概率模型,该模型包含部位似然度和相连部位之间的结构先验两项,其中部位似然度概率项即部位姿态为某一状态时所对应图像区域的外观特征与部位的真实外观特征之间的相似程度,但实际上人体部位的真实外观特征是未知的,往往由人体部位外观模型来代替。因此,似然度概率项的准确度取决于所建立的人体部位外观模型是否能准确地描述人体部位的真实外观特征。

部位之间的结构先验概率项即相连部位的关节符合运动目标关节几何约束的程度,图结构模型利用推导算法来确定人体各部位的最佳姿态。

鉴于人体部位外观模型的重要性,近年来对于人体姿态估计问题的研究中有相当大的部分集中在人体部位观测模型的改进上,并已经提出了很多人体部位外观模型^[3,10-15]。已有的人体部位观测模型中大部分都是通过通过对多幅训练图像^[11-13]或视频序列图像^[3,14]的训练建立的。为适应不同图像光照环境以及人体姿态的不同,部分人体部位观测模型^[10-15]仅利用待处理图像建立更适用于本幅图像的人体部位外观模型。文献[10]首先利用训练图像集学习人体部位相对于人体上半身矩形框的定位概率,然后对待处理图像通过上半身检测和前景增强得到人体前景区域,最后根据该图像的定位概率建立基于边缘特征的人体部位外观模型。文献[15]采用与文献[10]相同的方法学习人体部位的定位概率,然后对待处理图像根据定位概率建立基于颜色直方图的人体部位外观模型。

文献[10]和[15]在对不同的待处理图像建立人体部位外观模型时都采用相同的定位概率,但实际上由于人体姿态的变化,不同人体的同一部位在不同图像中经常处于不同的位置,若待处理图像中人体部位恰好位于定位概率很低的区域,则该图像中人体部位定位区域对应的颜色直方图特征与根据定位概率求解得到的基于颜色直方图的人体部位外观模型之间的似然度将会较低,进而导致最终的人体姿态估计准确度也较差。为了解决这个问题,本研究提出了一种新的仅利用单幅静态图像建立人体部位外观模型的方法,并将其用于人体上半身姿态的

估计。

1 树形图结构模型

人体上半身的树形图结构模型见图 1。如图 1(a)所示,人体上半身由 6 个刚体部位组成,相连部位之间通过关节相连,人体部位的姿态用定位矩形框 $l = \{x, y, r, l, w\}$ 来表示,其中, (x, y) 为矩形框中心在图像中的坐标, r 为矩形框相对于垂直方向所偏移的角度, l 为矩形框长度, w 为矩形框宽度。人体上半身结构可以用一个无向图 $G = (V, E)$ 来表示,建立如图 1(b)所示的人体树形图结构模型,其中 $V = \{v_1, \dots, v_6\}$ 为无向图中所有顶点的集合,一个顶点对应一个人体部位, E 为无向图中所有相连的顶点对的集合。

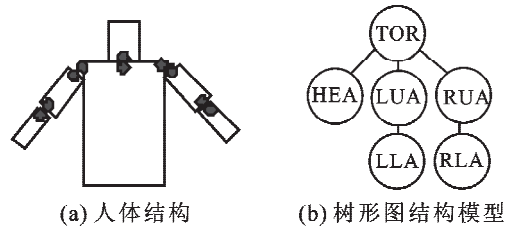


图 1 人体树形图结构模型
Fig. 1 Human model of tree structure

根据如图 1(a)所示的人体的结构特征,人体姿态估计结果的准确度主要由人体部位的姿态与部位真实特征的匹配程度以及相连部位之间的连接符合人体几何约束的程度,即相连部位各自的关节点相匹配的程度两个因素决定。

人体姿态估计问题可以转化为:

$$L^* = \arg \min_L \left(\sum_{i=1}^6 m_i(l_i) + \sum_{(v_i, v_j) \in E} d_{ij}(l_i, l_j) \right) \quad (1)$$

其中, $m_i(l_i)$ 表示第 i 个部位状态为 l_i 时对应的图像特征与该部位的外观模型不匹配的程度, $d_{ij}(l_i, l_j)$ 表示相连的第 i 个部位和第 j 个部位对应状态分别为 l_i 和 l_j 时各自对应的关节点不匹配的程度。

与文献[3]相同,上述优化问题可以转化为概率估计问题,并可由贝叶斯估计解决,即:

$$p(L | I) \propto p(I | L) p(L) \quad (2)$$

其中, I 为待处理静态图像, $L = \{l_1, \dots, l_6\}$ 为人体各部位的姿态集合, $P(L)$ 为先验概率,表示人体各部位姿态为 L 时所有相连部位之间的连接符合人体几何约束的程度, $p(I | L)$ 表示人体各部位姿态为 L 时对应的图像特征与对应的部位外观模型的匹配程度, $P(L | I)$ 为后验概率,表示静态图像 I 中人体各部位姿态为 L 的概率。

假设人体每个部位是相互独立的,即:

$$p(I | L) = \prod_{i=1}^6 p(I | l_i) \quad (3)$$

式(2)中 $p(L)$ 求解为:

$$p(L) = p(L | E, c) = \prod_{(v_i, v_j) \in E} p(l_i, l_j | c_{ij}) \quad (4)$$

其中, $p(l_i, l_j | c_{ij})$ 表示第 i 个部位和第 j 个部位的状态分别为 l_i 和 l_j 时各自对应的关节点相匹配的程度,可以用转换后状态空间中的高斯函数求解^[3]为:

$$p(l_i, l_j | c_{ij}) = N(T_{ij}(l_i), C_{ij}) \quad (5)$$

其中, C_{ij} 为相连的第 i 个部位和第 j 个部位之间的协方差,函数 $T()$ 可表示为:

$$T_{ij}(l_i) = \begin{bmatrix} x_i + d_x^i \cos \theta_i - d_y^i \sin \theta_i \\ y_i + d_x^i \sin \theta_i + d_y^i \cos \theta_i \\ \theta_i + \bar{\theta}_{ij} \end{bmatrix} \quad (6)$$

其中, $d^i = (d_x^i, d_y^i)^T$ 表示相连的第 i 个部位和第 j 个部位的平均相对位置, $\bar{\theta}_{ij}$ 为两部位之间的相对角度。参数 d^i 和 C_{ij} 可利用最大似然估计方法进行学习。

本研究采用文献[3]中的维泰尔比递推算法进行人体姿态的推导。

2 基于颜色直方图的人体部位外观模型

对于不同的静态图像,不论光照条件如何变化和人体体形如何不同,人体都有基本相同的边缘特征,所以不同人体 HOG 特征变化不大,同样不同人体的相同部位具有大致相同的 HOG 特征。颜色特征在人体姿态估计领域中是应用非常广泛的一种图像特征,主要有颜色直方图^[10-16]和颜色的对称性^[17-18]两种应用方式。人在不同的场合会穿着不同颜色的衣服,而且即使同样颜色的衣服在不同的光照条件下也会有很大的变化,所以与 HOG 特征不同的是不同人体相同部位的颜色特征可能相差很大。

由于不同人体同一部位的 HOG 特征变化不大,所以无论人体部位位于什么位置,其定位区域对应的 HOG 特征与该部位基于 HOG 特征的外观模型之间一定具有较高的似然度,则其一定会位于定位概率较高的区域,基于此建立的部位外观模型用于人体姿态估计对于任何待处理图像都会得到较好的估计效果。

本研究人体部位外观模型的建立可以分为三步,即:① 确定人体部位的分布区域;② 求解人体部位定位区域中各像素点的定位概率;③ 依据定位

概率求解人体部位的颜色直方图,即为人体部位外观模型。

2.1 确定分布区域

对于待处理的静态图像,本研究采用文献[19]提出的方法减小人体部位状态空间,然后根据减小后的状态空间确定人体部位的分布区域。

图 2 给出了某待处理图像左上臂分布区域的确定过程,图 2(a) 为待处理静态图像,左上臂最初的状态空间为整幅图像。图 2(b) 为利用人体上半身检测器检测确定的人体上半身矩形框。图 2(c) 为经过减小后的状态空间,其中对于似然度阈值本研究采用所有大于零的似然度的均值。图 2(d) 为根据减小后的状态空间所确定的左上臂的分布区域。

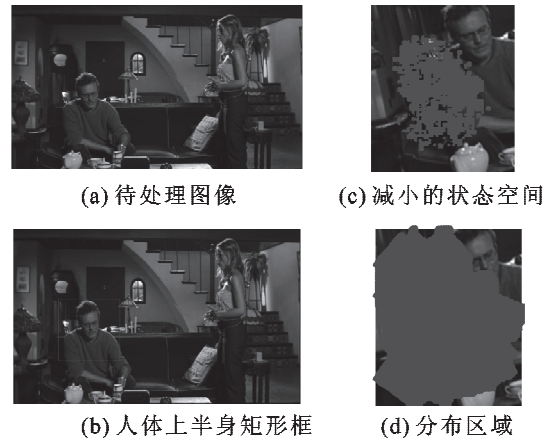


图 2 定位区域确定示意图

Fig. 2 Example of reduction of location region

2.2 求解定位概率

人体部位减小后的状态空间中每个状态对应的图像区域的 HOG 特征与基于 HOG 特征的人体部位外观模型^[19]具有不同大小的似然度,很显然似然度越高的状态对应的图像区域中像素点属于人体真实部位的可能性也越大。定位区域中的每一个像素点都会被多个状态对应的图像区域所包含,但考虑到似然度不同的状态属于真实部位的概率不同,定位区域中的每一个像素点的定位概率定义为:

$$L_i P_i(x, y) = \frac{\sum_{l_j \in L_i} p(I | l_j)}{num} \quad (7)$$

其中, L_i 为所有包含该像素点的状态集合, num 为人体部位减小后的状态空间的总状态数目。

2.3 建立部位外观模型

对于待处理图像中人体部位分布区域,笔者把每一个像素点的定位概率作为权值来求解分布区域的颜色直方图,采用线性插值方法对颜色直方图进行均衡,最终得到的颜色直方图即为待处理图像对应的人体部位的外观模型。

公式(3)中的似然度 $p(I | l_i)$ 可以用状态为 l_i 时对应的颜色直方图与外观模型归一化的欧式距离来计算,即:

$$D_i(x_i, z_i) = \frac{1}{n} \sum_{k=1}^n \left(1 - \frac{|x_{ik} - z_{ik}|}{\max(x_{ik}, z_{ik})} \right) \quad (8)$$

其中, x_i 是人体部位状态 l_i 时对应的颜色直方图, n 是 x_i 的维数, z_i 是该部位对应的基于颜色直方图的外观模型。

3 实验结果与分析

为便于与文献[10]和[15]的处理效果比较,本研究选择与其相同的训练图像集和待处理图像集,并选择相同的评价标准。

图3给出了几个在不同条件和姿态下的几幅图像的姿态估计结果。表1给出了将本研究所提人体部位外观模型用于人体姿态估计的估计结果与文献[10]和[15]的人体姿态估计结果的比较。文献[10]和[15]只给出了人体所有部位的平均姿态估计结果,而本算法给出了躯干、头部、上臂和小臂各自的估计结果。从表1可以看出,与文献[10]和[15]相比,本算法的人体姿态估计准确度得到了较大幅度的提高。

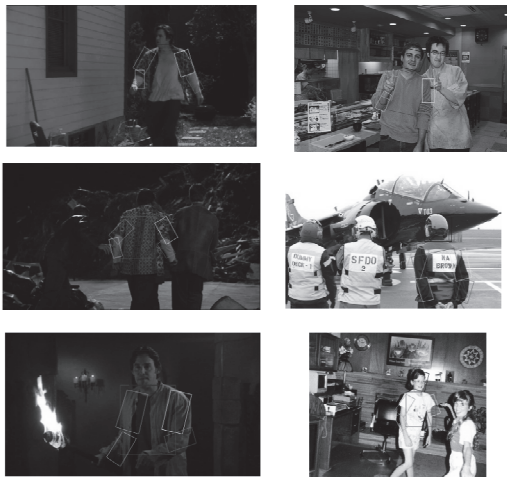


图3 人体姿态估计示例

Fig. 3 Examples of human pose estimation

表1 人体姿态估计结果比较

Tab. 1 Comparison of human pose estimation

图像库	方法	躯干	头部	上臂	小臂	总体
Buffy	文献[10]					61.4
	文献[15]					72.2
	本研究	100	100	89.4	60.2	83.2
PASCAL	文献[10]					52.2
	文献[15]					59.1
	本研究	100	100	81.3	48.1	76.5

4 结 论

1) 提出了一种新的基于颜色直方图特征的人体部位外观模型,并将其用于静态图像中人体上半身的姿态估计。

2) 首先减小人体部位的状态空间,然后利用与基于 HOG 特征的人体部位外观模型具有较高似然度的状态来学习人体部位的定位概率,最后将定位概率作为权值求解颜色直方图以构成人体部位的外观模型。

实验结果表明所提人体部位外观模型效果更佳。

参考文献:

- [1] Agarwal A, Triggs B. 3d human pose from silhouettes by relevance vector regression [C]//Proceedings of CVPR, Washington, 2004.
- [2] Jiang H. Finding human poses in videos using concurrent matching and segmentation [C]//Proceedings of ACCV, Queenstown, New Zealand, 2010.
- [3] Felzenszwalb P F, Huttenlocher D P. Pictorial structures for object recognition[J]. International Journal of Computer Vision, 2005, 61(1):55-79.
- [4] Sigal L, Black J. Measure locally, reason globally: occlusion-sensitive articulated pose estimation [C]//Proceedings of CVPR, New York, 2006.
- [5] Freifeld O, Weiss A, Zuffi S, et al. Contour people: a parameterized model of 2D articulated human shape [C]//Proceedings of CVPR, San Francisco, USA, 2010.
- [6] Wang Y, Tran D, Liao Z C. Learning hierarchical poselets for human parsing[C]//Proceedings of CVPR, Colorado Springs, USA, 2011.
- [7] Ukita N. Articulated pose estimation with parts connectivity using discriminative local oriented contours[C]//Proceedings of CVPR, Providence, USA, 2012.
- [8] Andriluka M, Roth S, Schiele B. Pictorial structures revisited; people detection and articulated pose estimation [C]//Proceedings of CVPR, Miami, USA, 2009.
- [9] Fischler M, Elschlager R. The representation and matching of pictorial structures[J]. IEEE Transactions on Computer, 1973, 22(1):67-92.
- [10] Ferrari V, Marin-Jimenez M, Zisserman A. Progressive search space reduction for human pose estimation [C]//Proceedings of CVPR, Anchorage, USA, 2008.

(下转第 219 页)