

8-7-2020

## Pose Estimation Using Limbs Adjustment

Cai Peng

*Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China;*

Dehui Kong

*Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China;*

Baocai Yin

*Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China;*

Huo Yi

*Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China;*

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

---

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

---

## Pose Estimation Using Limbs Adjustment

### Abstract

**Abstract:** Human body is decomposed into several local component templates connected with tree structure by articulated pose estimation with mixture-of-parts, and the corresponding full score is used to estimate body pose. Such a method fails in precisely estimating human pose due to the defects of tree models when human has the complicated pose of body. *To address this problem, pose estimation using limbs adjustment with mixtures-of-parts models was proposed. And pixel discrete direction, local characteristic functions blending full score and pixel discrete direction and other attributes, and pose estimation using limbs adjustment based on local characteristic functions of articulated position were proposed, which is better than the method based on articulated pose estimation with mixtures-of-parts.*

### Keywords

mixture-of-parts, pose estimation, limb adjustment, pixel discrete direction, full score

### Recommended Citation

Cai Peng, Kong Dehui, Yin Baocai, Huo Yi. Pose Estimation Using Limbs Adjustment[J]. Journal of System Simulation, 2015, 27(10): 2344-2351.

## 基于肢体调整的人体姿态估计

蔡鹏, 孔德慧, 尹宝才, 霍奕

(北京工业大学城市交通学院多媒体与智能软件技术北京市重点实验室, 北京 100124)

**摘要:** 基于混合部位模型的人体姿态估计方法将人体拆分成若干由树形结构相连的模板, 利用相应的全分函数进行人体姿态的估计, 由于树形结构的固有缺陷而导致复杂人体姿态估计错误。为了解决这个问题, 提出基于肢体调整的人体姿态估计方法。该方法提出像素离散方向的概念, 融合全分函数和像素离散方向等属性的局部特征函数, 以及基于关节的局部特征进行人体姿态调整的方法。实验表明本文方法的姿态估计效果比以前的基于混合部位的关节模型的方法更好。

**关键词:** 混合部位; 姿态估计; 肢体调整; 像素离散方向; 全分函数

中图分类号: TP391

文献标识码: A

文章编号: 1004-731X (2015) 10-2344-08

### Pose Estimation Using Limbs Adjustment

Cai Peng, Kong Dehui, Yin Baocai, Huo Yi

(Beijing Key Laboratory of Multimedia and Intelligent Software Technology, College of Metropolitan Transportation, Beijing University of Technology, Beijing 100124, China)

**Abstract:** Human body is decomposed into several local component templates connected with tree structure by articulated pose estimation with mixture-of-parts, and the corresponding full score is used to estimate body pose. Such a method fails in precisely estimating human pose due to the defects of tree models when human has the complicated pose of body. To address this problem, pose estimation using limbs adjustment with mixtures-of-parts models was proposed. And pixel discrete direction, local characteristic functions blending full score and pixel discrete direction and other attributes, and pose estimation using limbs adjustment based on local characteristic functions of articulated position were proposed, which is better than the method based on articulated pose estimation with mixtures-of-parts.

**Keywords:** mixture-of-parts; pose estimation; limb adjustment; pixel discrete direction; full score

## 引言

人体姿态估计是计算机视觉中非常重要的研究内容。部位表示<sup>[1]</sup>和图示结构<sup>[2-3]</sup>是姿态估计的两个重要的工具。第一种方法通过部位集合对人体进行建模, 而第二种方法将人体表示为相互连接的

部位模板。与基于部位表示的方法相比, 基于图示结构的方法通过在相邻的部位模板之间增加连接关系, 以提高姿态估计的正确性。然而, 由于多自由度的人体, 衣服样式和姿态多变都可能造成姿态估计困难。本文提出基于肢体调整的人体姿态估计方法, 通过人体的肢体姿态的局部调整, 达到有效估计人体姿态的目的。本文的创新点包括:

提出像素离散方向的概念, 以及融合肢体的全分函数和像素离散方向等属性的局部特征函数。

提出采用人体关节的全分函数和局部特征的融合和选择策略。在此基础上, 提出初始人体姿



收稿日期: 2015-06-07 修回日期: 2015-07-22;  
基金项目: 国家自然科学基金(61227004; 61370120);  
作者简介: 蔡鹏(1977-), 男, 江苏句容, 博士, 助理研究员, 博士后, 研究方向为计算机图形学、计算机视觉; 孔德慧, 女, 博士, 教授, 博导, 研究方向为虚拟现实与图形学; 尹宝才, 男, 博士, 教授, 博导, 研究方向为多媒体技术、计算机图形学、视频编码。

<http://www.china-simulation.com>

• 2344 •

态选择、大腿与上臂姿态调整、小腿与前臂姿态调整以及腿或臂的最优姿态组合选择的四个阶段进行人体姿态估计。

## 1 以前的工作

许多学者在视频和静态图像的人体姿态估计方面提出了许多有效的算法, 我们列出其中的一些重要的方法。

O'Rourke 和 Badler<sup>[4]</sup>, Hogg<sup>[5]</sup>, Rohr<sup>[6]</sup>提出经典的模型方法来解决视频中人体姿态估计问题。静态图像的人体姿态估计正成为起人们关注的焦点, 树形结构在姿态估计中是有效的<sup>[2]</sup>, 但是肢体姿态存在一些困难(如肢体的重复估计)。循环模型需要逼近推断策略, 诸如重要采样<sup>[2,7]</sup>, 循环信令传递<sup>[8]</sup>, 或迭代逼近<sup>[9]</sup>等策略。最近, 基于树的较低级范围的分支定界算法<sup>[10-11]</sup>被用来全局地解决这些问题。Lan 和 Huttenlocher<sup>[12]</sup>考虑通过鲁棒的姿态优先级来跟踪肢体重复估计的情况。Tran 和 Forsyth<sup>[13]</sup>指出上面的解决方案对特定数据库统计的过适现象可能比较敏感。

Yang 和 Ramanan<sup>[3]</sup>提出新的部位模型以估计静态图像中的人体姿态, 这种模型是一种通用的、灵活的混合模型, 它能捕获人体部位之间的语义连贯关系, 并且, 它是一种能对空间关系进行编码的弹性模型。然而, 由于基于混合部位的关节模型采用的树形结构所存在的固有缺陷造成复杂人体姿态估计的无效(如肢体区域的重复估计), 文献[3]的方法无法解决这个问题。

本文提出基于混合部位的关节模型的肢体调整方法。通过关节位置的全分函数和相应局部属性的融合与选择策略, 以及两种肢体姿态的最优组合选择策略, 本文取得比文献[3]中更好的人体姿态估计效果。

## 2 基于肢体调整的人体姿态估计

同文献[3], 本文定义  $I$  为图像, 设  $l_i = (x, y)$  是人体部位  $i$  的像素区域,  $t_i$  是部位  $i$  的混合类型, 其中,  $i \in \{1, 2, \dots, K\}$ ,  $l_i \in \{1, 2, \dots, L\}$ ,

$t_i \in \{1, 2, \dots, T\}$  (本文设  $K=26$  和  $T=6$ )。本文定义  $z_i = (l_i, t_i)$  为部位  $i$  的像素区域和混合类型。

### 2.1 基于肢体调整的人体姿态估计的框架

本文给出基于肢体调整的人体姿态估计方法的框架如下:

- 1) 通过人体的全分函数和全局特征的融合和选择策略, 得到初始的人体姿态。
- 2) 在上臂和大腿的关节点, 通过相应的全分函数和局部特征的融合和选择策略, 以及两种姿态的最优组合选择的策略, 得到最优的上臂和大腿姿态。
- 3) 在前臂和小腿的关节点, 通过相应的全分函数和局部特征的融合和选择策略, 得到最优的前臂或小腿姿态。
- 4) 通过臂(即上臂和前臂)或腿(大腿和小腿)的两种姿态的 4 种组合, 选择最优的臂或腿的姿态。

### 2.2 人体和关节处的全分函数定义

本文利用基于混合部位的关节模型计算人体和肢体的全分函数。在图 1 中, 本文人体的 26 个部位与树形结构的 26 个节点相对应, 上臂、前臂、大腿和小腿分别由相应的节点(或部位)组成, 其中, 圆圈代表人体某个部位, 在圆圈之间连线代表人体部位之间的连接关系。

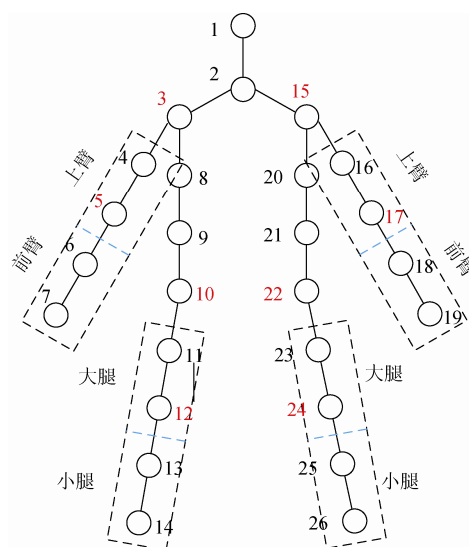


图 1 基于混合部位的关节模型

本文通过设定人体的全分函数的阈值  $T$ ，得到相应  $M$  个人体姿态。

$$score_{i,j}(z_{i,j}) = \phi_{i,j}(I, z_{i,j}) + \sum_{k \in kids(i,j)} m_{i,k}(z_{i,k}) \quad (1)$$

$$m_{i,j}(z_{i,k}) = \max_{z_{i,j}} [score_{i,j}(z_{i,j}) + \psi_{j,k}(z_{i,j}, z_{i,k})] \quad (2)$$

在人体姿态  $i$  的部位  $j$  位置，本文通过式(1)计算相应的全分函数 ( $1 \leq i \leq M, 1 \leq j \leq 26$ )。其中， $z_{i,j}$  是人体姿态  $i$  的部位  $j$  对应的图像位置和混合类型， $\phi_{i,j}(I, z_{i,j})$  是图像  $I$  中人体姿态  $i$  的部位  $j$  的 HoG (Histograms of Oriented Gradients) 向量<sup>[14]</sup>与相应的 HoG 模板的卷积， $\psi_{j,k}(z_{i,j}, z_{i,k})$  与人体姿态  $i$  的部位  $j$  和  $k$  的相对位置相关。 $score_{i,1}(z_{i,1})$  是人体姿态的根部位 1 的全分函数。

本文利用式(2)指定

$$score'_{i,j}(z_{i,j}) = m_{i,j+1}(z_{i,j}) \quad (3)$$

为人体姿态  $i$  的相应肢体在关节点  $j$  的全分函数 ( $j \in \{3, 10, 22, 15, 5, 12, 24, 17\}$ ) (上臂:  $j \in \{3, 15\}$ ; 大腿:  $j \in \{10, 22\}$ ; 前臂:  $j \in \{5, 17\}$ ; 小腿:  $j \in \{12, 24\}$ )。

### 2.3 像素离散方向的定义

本文先给出图像位置某离散方向的局部相对偏差概念，然后计算相应像素的所有离散方向的局部相对偏差，最后给出像素离散方向的概念。本文先将彩色图像转变为相应的灰度图像后，再进行以下与像素离散方向的相关计算。

如图 2(a)所示，本文定义图像  $4 \times 4$  的局部区域，其中心位置按顺时针方向定义 8 个间隔相等的离散方向(1-8)，其中，图 2(b)、(c)和(d)分别是以离散方向 1、3 和 7 为边界的两个相对的局部区域。本文定义以离散方向  $i(1 \leq i \leq 8)$  为边界的局部像素区域的像素和分别为  $S_i$  和  $S'_i$ ，并定义图像位置的离散方向  $i$  的局部相对偏差  $\alpha_i$  如下：

$$\alpha_i = \frac{|S_i - S'_i|}{S_i + S'} \quad (4)$$

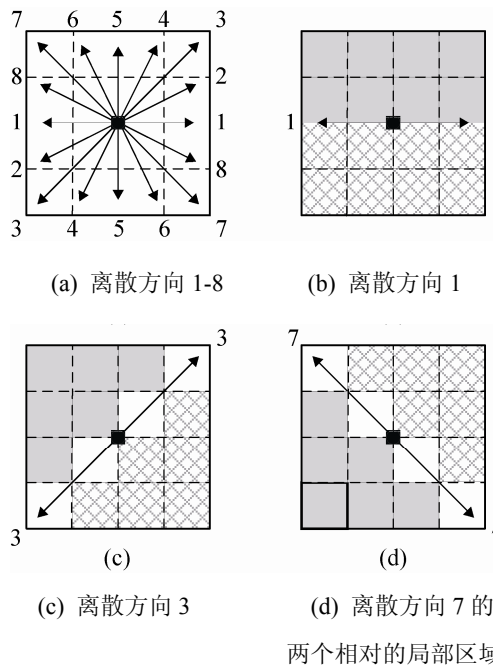


图 2 图像位置的离散方向与相应局部区域

本文定义像素的离散方向  $i(1 \leq i \leq 8)$  的局部相对偏差  $\beta_x^i$ ， $\beta_x^i$  是像素  $p_x$  周围的 4 个图像位置的离散方向  $i$  的局部相对偏差  $\alpha_i$  的平均，定义像素  $p_x$  的离散方向  $n_x$  如下：

$$n_x = \left( \max_{1 \leq i \leq 8} (\beta_x^i) \geq \gamma \right) \cdot i \quad (5)$$

式(5)中逻辑表达式为真时，其值为 1；否则为 0(本文中所有的逻辑表达式的计算与之相同)。 $\gamma$  是局部相对偏差的阈值，当像素  $p_x$  的所有离散方向的  $\beta_x^i$  均小于  $\gamma$  时，像素  $p_x$  的离散向量  $n_x$  为 0(即不指向任何方向)；否则， $n_x$  等于像素  $p_x$  的最大局部相对偏差所对应的离散方向。

### 2.4 人体和肢体的特征函数定义

为了选择最适合的人体初始姿态和肢体调整姿态，本文定义人体和肢体的相应特征函数。

本文定义  $N(A, n)$  为图像的局部区域  $A$  中的像素与离散方向  $n$  接近的像素数目， $N_x(A, n)$  为局部区域  $A$  中的像素与离散方向  $n$  接近的像素在图像  $x$  轴上投影点的数目， $N_y(A, n)$  为局部区域  $A$  中的像素与离散方向  $n$  接近的像素在图像  $y$  轴上投影点的数目 ( $1 \leq n \leq 8$ )。

$$N(A, n) = \sum_{p_x \in A} (n_x = n \text{ or } n_x = n' \text{ or } n_x = n'') \quad (6)$$

$$N_x(A, n) = \sum_{p_x \in A} (\exists p_x \in \overline{p_x} (n_x = n \text{ or } n_x = n' \text{ or } n_x = n'')) \quad (7)$$

$$N_y(A, n) = \sum_{p_y \in A} (\exists p_x \in \overline{p_y} (n_x = n \text{ or } n_x = n' \text{ or } n_x = n'')) \quad (8)$$

其中,  $n' = \begin{cases} n-1 & \text{if } n > 1 \\ 8 & \text{otherwise} \end{cases}$ ,  $n'' = \begin{cases} n+1 & \text{if } n < 8 \\ 1 & \text{otherwise} \end{cases}$

$p_x$  是图像区域  $A$  中像素,  $\overline{p_x}$  是图像区域  $A$  中与图像  $x$  轴投影点对应的像素集合,  $\overline{p_y}$  是图像区域  $A$  中与图像  $y$  轴投影点对应的像素集合。

本文定义

$$A_{i,j}^a = \overline{A_{i,j}^a} + k \cdot A \quad (9)$$

是臂或腿的末部位(即部位( $j+4$ ))和人体相应部位的非重叠区域的面积(即相应的像素个数, 下同),  $i$  是上臂或大腿的第  $i$  个姿态,  $j$  是臂或腿的关节点 ( $j \in \{3, 10, 22, 15\}$ ),  $\overline{A_{i,j}^a}$  是臂或腿的末部位和除部位 1、2 和( $j+1$ )到( $j+4$ )外 ( $j \in \{3, 15\}$ )或除部位 1、2、( $j+1$ )到( $j+4$ )、4 到 7 和 16 到 19 外 ( $j \in \{10, 22\}$ )的人体部位非重叠区域的面积,  $k$  是与臂或腿的相关系数,  $A$  是一个部位的面积。当  $\overline{A_{i,j}^a}$  越大, 即臂或腿的末部位与人体相应部位的重叠面积越小, 这能有效减少由树形结构固有缺陷引起的姿态估计错误,  $k \cdot A$  能对  $\overline{A_{i,j}^a}$  的相对大小起调节作用。

图像中人体部位之间的单位向量是由部位中心的位置决定, 本文定义

$$P_{i,j} = P_{i,j}^{j-1,j+2} \cdot P_{i,j}^{j,j+1} \quad (10)$$

臂或腿的两对部位之间单位向量的点积,  $i$  是前臂或小腿的姿态的序号,  $j$  是前臂或小腿的关节位置 ( $j \in \{5, 12, 24, 17\}$ ),  $P_{i,j}^{j-1,j+2}$  是部位( $j-1$ )与( $j+2$ )中心之间的单位向量,  $P_{i,j}^{j,j+1}$  部位  $j$  与( $j+1$ )中心之间的单位向量。真实姿态的肢体的部位中心之间单位向量的点积  $P_{i,j}$  满足  $0.5 \leq P_{i,j} \leq 1$ , 可作为下面的局

部特征函数中  $P_{i,j}$  的约束条件。

本文定义  $H_i^k$  为人体姿态  $i$  的部位  $k$  的 HoG 向量与相应 HoG 模板的卷积( $1 \leq i \leq N$ ), 并定义

$$H_i = \sum_{k=1}^{26} H_i^k \quad (11)$$

为人体姿态  $i$  的 26 个部位的 HoG 向量卷积值的和。

### 2.5 肢体姿态的调整过程

本文给出人体姿态估计方法的框架如图 3 所示, 包括初始人体姿态选择、大腿与上臂的姿态调整、小腿与前臂姿态调整以及腿或臂的最优姿态组合选择的 4 个过程。

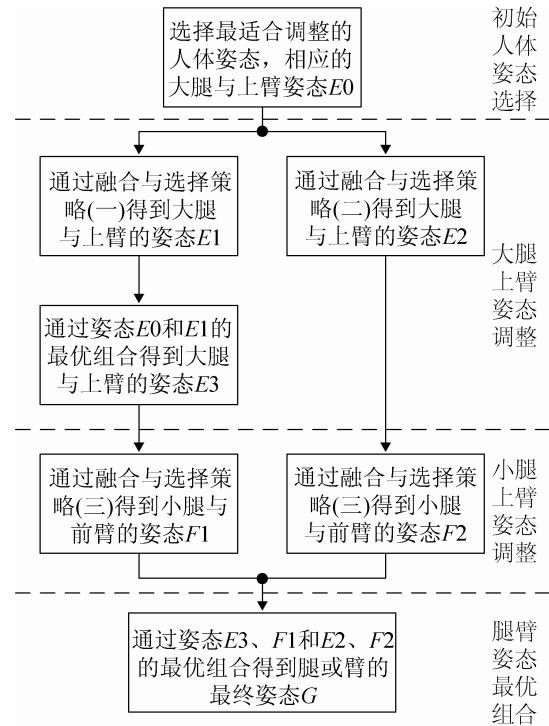


图 3 人体姿态的肢体调整过程

本文通过设置阈值  $T$  得到  $M$  个人体姿态, 并定义

$$S_i = score_{i,1}(z_{i,1}) - T \quad (12)$$

为人体姿态  $i$  的全分函数( $1 \leq i \leq M$ ), 其中,  $score_{i,1}(z_{i,1}) \geq T$  (即  $S_i \geq 0$ ), 其中,  $score_{i,1}(z_{i,1})$  是人体姿态  $i$  在根部位 1 的全分函数(式(1))。

对于上面的  $M$  个人体姿态中的第  $i$  个姿态,

本文定义

$$S'_i = S_i \cdot H_i \cdot A_i^b \quad (13)$$

为相应的融合函数，其中， $S_i$  是人体姿态  $i$  的全分函数(式(12))， $H_i$  是人体姿态  $i$  的所有部位的 HoG 向量卷积值的和(式(11))， $A_i^b$  是人体姿态  $i$  的所有部位的包围盒面积。

在初始人体姿态的选择阶段，本文通过设置阈值  $T$ ，得到  $M$  个人体姿态，通过式(13)得到与  $S'_i$  的最大值对应的人体姿态，即相应得到下一阶段初始的大腿(或上臂)姿态  $E_0$ 。

本文定义

$$S'_{i,j} = S_{i,j} \cdot N_{i,j}^b \cdot A_{i,j}^a \quad (14)$$

$$S'_{i,j} = S_{i,j} \cdot A_{i,j}^a \quad (15)$$

分别为人体姿态  $i$  的关节点  $j(1 \leq i \leq N, j \in \{3, 10, 22, 15\})$  的融合与选择策略(一)和(二)的局部特征函数，其中，

$$N_{i,j}^b = \sqrt{N_x^2(A, n) + N_y^2(A, n)}, \quad (16)$$

$S_{i,j} = score'_{i,j}(z_{i,j})$  是人体姿态  $i$  在关节点  $j$  的全分函数(式(3))， $A$  为部位  $(j+1)$  和以部位  $(j+1)$  和  $(j+2)$  各顶点平均为顶点的区域的覆盖区域， $n$  为部位  $(j+1)$  和  $(j+2)$  中心之间的离散方向  $(1 \leq n \leq 8, \text{下同})$ ， $N_x(A, n)$  和  $N_y(A, n)$  分别为图像区域  $A$  中与离散方向  $n$  接近的像素在图像  $x$  和  $y$  轴上投影点的数目(式(7)和(8))， $A_{i,j}^a$  为腿(或臂)在关节部位  $j$  的末部位和人体相应部位的非重叠面积(式(9))。

在大腿与上臂姿态调整阶段，本文通过融合与选择策略(一)的局部特征函数(式(14))，在初始大腿(或上臂)姿态  $E_0$  的基础上，产生与降序的  $S_{i,j}$  所对应的  $N$  个大腿(或上臂)姿态，并得到与  $S'_{i,j}$  的最大值所对应的大腿(或上臂)姿态  $E_1$ 。

左、右大腿(或上臂)的两种姿态  $E_0$  和  $E_1$  构成 4 种不同的姿态组合，计算每种姿态组合的值

$$N(A_{ab}, n_a, n_b), \quad (17)$$

其中， $A_{ab}$  是左、右大腿(或上臂)所有部位的覆盖区域， $n_a$  和  $n_b$  分别是左、右大腿(或上臂)的两个部位中心之间的离散方向， $N(A_{ab}, n_a, n_b)$  是图像区域  $A_{ab}$  中与所在部位的离散方向  $n_a$  或  $n_b$  接近的像素数目(与式(6)类似)。本文以式(17)的最大值所对应的左、右大腿(或上臂)姿态组合为其姿态  $E_3$ 。

本文通过融合与选择策略(二)的局部特征函数(式(15))，在初始大腿(或上臂)姿态  $E_0$  的基础上，产生与降序的  $S_{i,j}$  所对应的  $N$  个大腿(或上臂)姿态，并得到与  $S'_{i,j}$  的最大值所对应的大腿(或上臂)姿态  $E_2$ 。

本文定义

$$S'_{i,j} = S_{i,j} \cdot N(A_{ab}, n_a, n_b) \cdot (P_{i,j} > 0) \cdot (N(A_2, n_{2,3}) + N(A_3, n_{2,3}) > 0.4 \overline{N_{2,3}}) \quad (18)$$

$$S''_{i,j} = S_{i,j}^{1.15} \cdot$$

$$\max \{N_x(A_{ab}, n_a, n_b), N_y(A_{ab}, n_a, n_b)\} \cdot A_{i,j}^c \quad (19)$$

为人体姿态  $i$  的关节点  $j(j \in \{5, 12, 24, 17\})$  的融合与选择策略(三)的两个局部特征函数，其中， $S_{i,j} = score'_{i,j}(z_{i,j})$  是腿(或臂)在关节点  $j$  的全分函数(式(3))； $A_{ab}$  是关节点  $j$  所在的腿(或臂)的四个部位的覆盖区域， $n_a$  是大腿(或上臂)的两部位中心之间的离散方向， $n_b$  是小腿(或前臂)的两部位中心之间的离散方向， $N(A_{ab}, n_a, n_b)$  是图像区域  $A_{ab}$  中与所在部位的离散方向  $n_a$  或  $n_b$  接近的像素数目(与式(6)类似)， $N_x(A_{ab}, n_a, n_b)$  是图像区域  $A_{ab}$  中与所在部位的离散方向  $n_a$  或  $n_b$  接近的像素在图像  $x$  轴上投影点的数目(与式(7)类似)， $N_y(A_{ab}, n_a, n_b)$  是图像区域  $A_{ab}$  中与所在部位的离散方向  $n_a$  或  $n_b$  接近的像素在图像  $y$  轴上投影点的数目(与式(8)类似)； $A_2$  和  $A_3$  是关节点  $j$  所在的腿(或臂)的第 2、3 个部位的区域， $n_{2,3}$  是腿(或臂)的第 2、3 个部位中心之间的离散方向， $N(A_2, n_{2,3})$  和  $N(A_3, n_{2,3})$  分别是腿(或臂)的第 2、3 个部位中与离散方向  $n_{2,3}$  接近的像素数目(式(6))， $\overline{N_{2,3}}$  是所有产生的  $N$  个姿态的  $(N(A_2, n_{2,3}) + N(A_3, n_{2,3})) / N$ 。

$n_{2,3}$ )的平均值;  $P_{ij}$  是关节  $j$  所在的腿(或臂)的两对部位中心之间单位向量的点积(式(10));  $A_{i,j}^c$  是在关节位置  $j$  的腿(或臂)的第 1、3 部位的覆盖面积。

在小腿与前臂姿态调整阶段, 以大腿(或上臂)两种姿态( $E_3$  和  $E_2$ )为初始姿态(见图 3), 本文通过融合与选择策略(三)的第 1 个局部特征函数(式(18)), 产生与降序的  $S_{i,j}$  所对应的  $N$  个小腿(或前臂)姿态, 并得到与  $S_{i,j}'$  的最大  $N'$  个值所对应的小腿(或前臂)姿态( $N' = 0.1N$ ); 通过融合与选择策略(三)的第 2 个局部特征函数(式(19)), 在上面的  $N'$  个小腿(或前臂)姿态中, 选择与  $S_{i,j}''$  的最大值所对应的小腿(或前臂)两种姿态( $F_1$  和  $F_2$ )。

经过肢体的姿态调整后, 腿(或臂)有两种姿态: 一种是大腿(或上臂)的姿态  $E_3$  和小腿(或前臂)的姿态  $F_1$ , 另一种是大腿(或上臂)的姿态  $E_2$  和小腿(或前臂)的姿态  $F_2$ (见图 3)。

在腿或臂的最优姿态组合选择阶段, 对于左、右腿(或臂)有 4 种不同姿态组合, 计算每种姿态组合的值, 其中,  $A_a$  和  $A_b$  分别是左、右腿(或臂)所在部位的覆盖区域,  $A_{ab}$  是左、右腿(或臂)所有部位的覆盖区域,  $n_{a1}$  和  $n_{b1}$  分别是左、右大腿(或上臂)的两个部位中心之间的离散方向,  $n_{a2}$  和  $n_{b2}$  分别是左、右小腿(或前臂)的两个部位中心之间的离散方向,  $N_x(A_a, n_{a1}, n_{a2})$ 、 $N_y(A_a, n_{a1}, n_{a2})$ 、 $N_x(A_b, n_{b1}, n_{b2})$  和  $N_y(A_b, n_{b1}, n_{b2})$  分别为  $A_a$  和  $A_b$  中与所在部位的离散方向接近的像素在图像  $x$  和  $y$  轴上投影点的数目,  $N_x(A_{ab}, n_{a1}, n_{a2}, n_{b1}, n_{b2})$  和  $N_y(A_{ab}, n_{a1}, n_{a2}, n_{b1}, n_{b2})$  分别为  $A_{ab}$  中与所在部位的离散方向接近的像素在图像  $x$  和  $y$  轴上投影点的数目(与式(7)和(8)类似)。

$$(N_x(A_{ab}, n_{a1}, n_{a2}, n_{b1}, n_{b2}) + N_y(A_{ab}, n_{a1}, n_{a2}, n_{b1}, n_{b2})) \cdot (N_x(A_a, n_{a1}, n_{a2}) + N_y(A_a, n_{a1}, n_{a2})) + N_x(A_b, n_{b1}, n_{b2}) + N_y(A_b, n_{b1}, n_{b2}) \quad (20)$$

本文通过式(20)的最大值选择腿(或臂)的最优姿态组合, 即得到经过姿态调整的 4 个阶段后最终的肢体姿态  $G$ 。

### 3 实验结果

本文实验在单个 Intel® 3.20 GHz CPU, 4 G RAM 的硬件平台执行, 比较文献[3]方法与本文方法的实验结果。

本文的混合部位的关节模型的人体部位数目为 26, 人体部位的 HoG 单元的尺寸设置为  $5 \times 5$ , 每个部位有 6 个混合类型。对于文献[3]的方法, 人体姿态的全分函数的阈值:  $T$ 。对于本文的方法, 设置的参数包括:

$T$ : 人体姿态的全分函数的阈值

$N$ : 大腿、上臂、小腿和前臂的姿态数目

$k$ : 与大腿和上臂相关的系数

$\gamma$ : 像素离散方向的局部相对偏差的阈值

对于文献[3],  $T$  为 -1; 对于本文方法,  $T$  为 -1,  $N$  为 100,  $k$  为 3(上臂),  $k$  为 1(大腿),  $\gamma$  为 0.05。

在图 4 中, 本文给出 6 组人体姿态估计的比较效果。每组左图是文献[3]的方法的姿态估计效果, 右图是本文方法姿态估计效果。实验结果表明本文的方法比文献[3]的方法好。因为本文的肢体全分函数与局部属性的融合与选择策略, 能从肢体的指数级数目的姿态中, 选择最适合的肢体姿态进行人体姿态估计。

在表 1 中, 文献[3]的姿态估计时间大致在 3~8s, 这比本文的 20~32 s 的时间少。本文的时间可以通过减少肢体的姿态数目的阈值  $N$ , 达到大幅度地减少计算时间的目的。



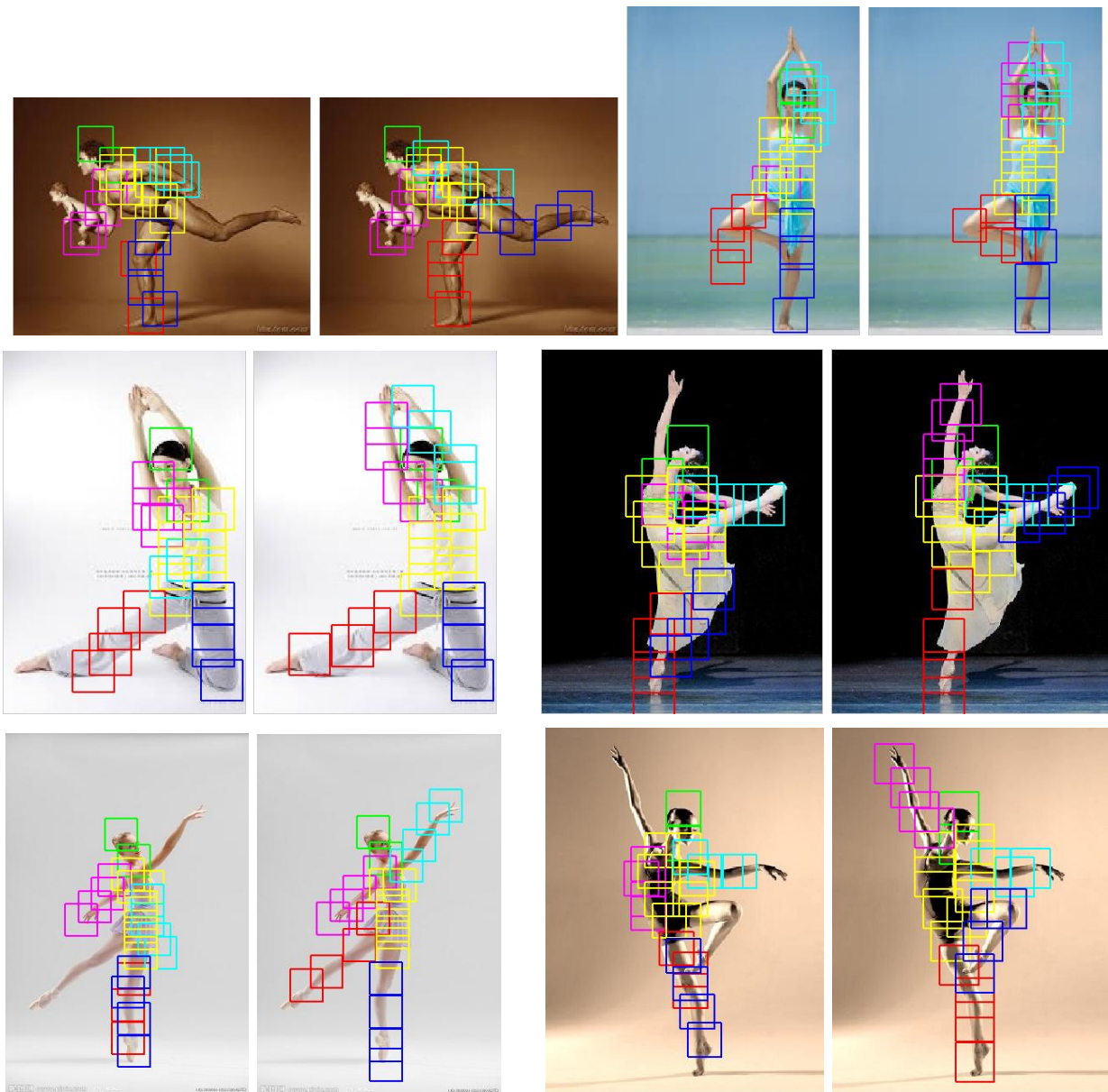


图 4 人体姿态估计效果的比较(左: 文献[3]的方法; 右: 本文的方法)

表 1 图 4 中两种人体姿态估计方法的时间 /s

文献[3]的方法		本文的方法	
3.82	3.75	22.98	21.36
3.71	5.72	20.93	27.14
7.87	7.47	32.50	30.80

#### 4 结论

本文提出基于混合部位模型的肢体调整的人体姿态估计方法。本文的方法基于肢体的全分函数和局部特征属性的融合与调整的策略, 可以充分利

用肢体的全分函数和局部特征属性的指数级数目的姿态组合, 进行肢体调整以达到提高人体姿态估计的准确性的目的。实验表明本文方法的人体姿态估计效果比文献[3]的方法好。我们将提出更有效的基于肢体调整的人体姿态估计方法, 以更加鲁棒和精确地估计人体姿态。

#### 参考文献:

- [1] D Marr, H Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes [C]// Proceedings of the Royal Society of London. Series B,

- Biological Sciences, 200(1140), 1978. London, UK: Royal Society of Great Britain, 1978: 269-294.
- [2] P Felzenszwalb, D Huttenlocher. Pictorial structures for object recognition [J]. IJCV (S0920-5691), 2005, 61(1): 55-79.
- [3] Y Yang, D Ramanan. Articulated human detection with flexible mixtures of parts [J]. TPAMI (S0162-8828), 2013, 35(12): 2878-2890.
- [4] J O'Rourke, N Badler. Model-based image analysis of human motion using constraint propagation [J]. TPAMI (S0162-8828), 1980, 2(6): 522-536.
- [5] D Hogg. Model-based vision: a program to see a walking person [J]. Image and Vision Computing (S0262-8856), 1983, 1(1): 5-20.
- [6] K Rohr. Towards model-based recognition of human movements in image sequences [J]. CVGIP-Image Understanding (S1077-3142), 1994, 59(1): 94-115.
- [7] M. Lee, I. Cohen. Proposal maps driven mcmc for estimating human body pose in static images [C]// CVPR. Washington, DC, USA: IEEE, 2004: 334-341.
- [8] L Sigal, M Black. Measure locally, reason globally: Occlusion-sensitive articulated pose estimation [C]// CVPR. New York, NY, USA: IEEE, 2006: 2041-2048.
- [9] Y Wang, G Mori. Multiple tree models for occlusion and spatial constraints in human pose estimation [C]// ECCV. Marseille, France: Springer, 2008: 710-724.
- [10] T Tian, S Sclaroff. Fast globally optimal 2D human detection with loopy graph models [C]// CVPR. San Francisco, CA, USA: IEEE, 2010: 81-88.
- [11] V Singh, R Nevatia, C Huang. Efficient inference with multiple heterogeneous part detectors for human pose estimation [C]// ECCV. Crete, Greece: Springer, 2010: 305-318.
- [12] X Lan, D Huttenlocher. Beyond trees: common-factor models for 2D human pose recovery [C]// ICCV. Beijing, China: IEEE, 2005: 470-477.
- [13] D Tran, D Forsyth. Improved human parsing with a full relational model [C]// ECCV. Crete, Greece: Springer, 2010: 227-240.
- [14] N Dalal, B Triggs. Histograms of oriented gradients for human detection [C]// CVPR. San Diego, CA, USA: IEEE, 2005: 886-893.



(上接第 2343 页)

- [7] Liyu T, Chongcheng C, Jie Z, *et al.* OntoPlant: An Integrated Virtual Plant Software Package for Different Scale Applications [C]// Spatial Data Mining and Geographical Knowledge Services (ICSDM), 2011 IEEE International Conference on; June 29 2011-July 1, 2011. USA: IEEE, 2011: 308-314.
- [8] 林定, 陈崇成, 唐丽玉, 等. 基于颜色编码的虚拟树木交互式修剪技术及其实现 [J]. 计算机辅助设计与图形学学报, 2011, 23(11): 1799-1807.
- [9] 邹杰, 林郁欣, 陈崇成, 等. 基于虚拟植物冠层的直射光合有效辐射三维分布模拟研究 [J]. 福州大学学报 (自然科学版), 2011, 39(6): 838-844.
- [10] 彭群生, 鲍虎军, 金小刚. 计算机真实感图形算法基础 [M]. 北京: 科学出版社, 1999.
- [11] Nusselt W. Graphische Bestimmung des Winkel-verhältnisses bei der Wärmestrahlung [J]. Zeitschrift des Vereines Deutscher Ingenieure, 1928, 72(20): 673.
- [12] Cohen M F, Greenberg D P. The Hemi-cube: a Radiosity Solution for Complex Environments [J]. SIGGRAPH Comput Graph (S1049-8907), 1985, 19(3): 31-40.
- [13] 周近宇, 黄天暄, 蒙林, 等. 视角因子方法在间接驱动 ICF 中的应用 [J]. 强激光与粒子束, 2010, 22(5): 1053-1058.
- [14] 俞洲. 基于 CUDA 的光线追踪实时成像模拟技术研究 [D]. 北京: 中国科学院空间研究所中国科学院空间科学与应用研究中心, 2011.
- [15] D'Azevedo E, Hu Z, Su S-Q, *et al.* Solving a Large Scale Radiosity Problem on GPU-based Parallel Computers [J]. Journal of Computational and Applied Mathematics (S0377-0427), 2014, 270: 109-120.
- [16] M T, Trumbore B. Fast, Minimum Storage Ray-triangle Intersection [J]. J Graph Tools, 1997, 2(1): 21-28.