

基于多模板配准的人体上半身三维运动跟踪

陈 姝^{1,2}, 梁文章³, 伍 靓^{1,2}

(1. 湘潭大学信息工程学院 湖南 湘潭 411105; 2. 智能计算与信息处理教育部重点实验室 湖南 湘潭 411105;
3. 广西卡斯动漫有限公司 南宁 530003)

【摘要】单目视频下的人体三维运动跟踪过程中,二维肢体模板由于遮挡而无法配准,从而造成跟踪丢失。针对这一问题,建立人体肢体的三维纹理模型用于人体运动跟踪。首先在初始帧中得到纹理模型在直立姿态下的二维模板及各像素点的局部坐标,在其它帧中利用像素灰度一致性约束优化进行跟踪,同时由跟踪结果生成新的二维模板,形成一个模板库。随着跟踪帧数增加,模板库动态更新,最终跟踪结果由模板库中各模板配准结果中的最优值决定。实验结果表明,与单一二维模板配准算法相比,多模板匹配能够克服由于肢体自旋转而造成的遮挡,取得了更好的跟踪结果。

关键词 约束优化; 多模板配准; 运动跟踪; 姿态提取

中图分类号 TP391

文献标志码 A

doi:10.3969/j.issn.1001-0548.2014.06.019

3D Human Upper Body Pose Tracking Based on Multi-Template Register

CHEN Shu^{1,2}, LIANG Wen-zhang³, WU Liang^{1,2}

(1. College of Information Engineering, Xiangtan University Xiangtan Hunan 411105;

2. Key Laboratory of Intelligent Computing & Information Processing, Ministry of Education Xiangtan Hunan 411105;

3. Guangxi Cast Animation Limited Company Nanning 530003)

Abstract In three-dimensional (3D) human motion tracking from monocular video sequences, two-dimensional (2D) body template cannot register accurately due to occlusion, which lead to track fail. To overcome this problem, a 3D texture model is used to track human motion. First, the template of the texture model and the local coordinates of each pixel under straight pose are initialized in the first frame. Then, the human motion is optimized by the pixel intensity consistency constraint, and a new 2D template is obtained by the tracked result. To contain the latest body information, the oldest template in the collection is dumped while a new template arrives, and the finally tracking result is determined by the output of template with the highest similarity. The experimental results demonstrate that the proposed algorithm is more effective in solving occlusion caused by segment self-rotation than the single 2D template matching method.

Key words constrained optimization; multi-template register; motion tracking; pose extraction

人体运动跟踪可以广泛应用于智能监控、互动娱乐、人机交互等领域。基于单目视频的人体三维运动跟踪使用成本低廉、使用范围广,相比多目视频运动跟踪具有明显优势。

随着基于视频的人体三维运动跟踪的广泛应用,越来越多的研究者投入该领域,并取得了许多研究成果。文献[1-2]对基于视频的人体运动跟踪及分析作了详细介绍并指出了当前的研究难点及热点。模型匹配作为该领域研究方法的一种,具有跟踪准确的优点。其基本思想是将预先定义的人体模型投影到图像上进行模型匹配,根据匹配过程采用

不同策略可以分为局部优化^[3-5]和全局优化^[6-7]两种。粒子滤波作为全局优化的代表,具有跟踪鲁棒性高的特点,可以适用于非线性非高斯应用场合,但是需要大量粒子来模拟概率密度函数,因而计算复杂度高,跟踪速度慢。文献[8-9]分别在运动模型、观测模型及采样效率方面作了改进。局部优化为单模式优化方法,跟踪速度快,但其跟踪效果依赖于初值选择。文献[10]提出一种联合线性混合蒙皮和Snake的局部优化算法进行人体外形和运动跟踪,但是该算法要求预先得到被跟踪人体高精度三维模型,且输入数据源为多目视频,不便于实际应用。

收稿日期: 2013-09-16; 修回日期: 2014-08-29

基金项目: 国家自然科学基金(61100139, 61040009); 科技部科技型中小企业技术创新基金(国科发计[2011]242号); 广西千亿元产业重点攻关工程项目(桂科攻11107006-13号)

作者简介: 陈姝(1976-), 男, 博士, 副教授, 主要从事计算机视觉方面的研究。

文献[3]提出一种基于比例正交模型的局部优化算法进行人体运动跟踪, 但是该算法不能处理强透视投影效果下的人体三维运动跟踪。在单目视频下采用单一模板匹配进行人体三维运动跟踪, 由于肢体自旋转或级连旋转使得用作匹配的模板转到面向摄像机肢体背面, 从而模板不可见造成模板匹配失效, 最终导致跟踪失败。

为了解决以上单目视频下基于模板匹配的人体三维运动跟踪算法中的两个问题, 提出了一种基于多模板匹配的方法。该算法的创新在于: 1) 基于透视投影成像模型下的约束优化, 可以从具有强透视投影效果的单目视频中恢复人体运动的全局姿态及关节局部姿态; 2) 采用多模板匹配方法既可以防止单一模板被遮挡情况, 又可以使模板反映肢体最新信息, 抵抗光照、形变对模板的影响。

1 采用的模型

1.1 人体模型

采用文献[11]中的树型结构建立人体模型如图1a所示, 人体肢体以关节相连, 其中胸部关节为根关节。左右胸及头部肢体以二维平面纸板模型表示, 由于左右臂肢体具有自旋转, 因此以三维圆柱建模。人体模板由左右胸、头部二维平面模板以及上下臂三维圆柱模板共7个模板组成。

1.2 坐标系及其变换模型

根据多刚体运动学知识可知, 整个人体运动表示成一个运动链, 某肢体的运动表示成相对于运动链中父肢体的相对运动。每个肢体固结一个局部坐标系, 坐标系的方向如图1b所示, 坐标系的原点位于对应的关节上。因而整个运动可以看作一个变换树, 树根是基坐标系, 叶子分别是头及前臂局部坐标系, 图中不同的虚线段代表不同的运动链。基坐标系的原点位于根关节, 方向为初始化时的方向, 在整个运动中保持不变。

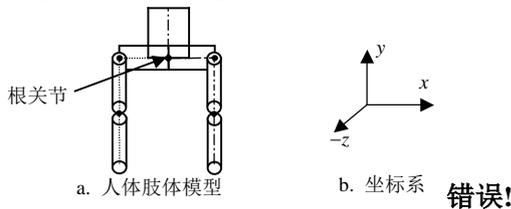


图1 人体模型

人体运动姿态以21维向量 P 表示: $P = \{p, p_1, \dots, p_5\}$, 其中 p 表示全局位姿定义为 $(x, y, s, \alpha, \beta, \gamma)$, (x, y, s) 代表全局平移(s 说明见式(5)), (α, β, γ) 代表全局旋转, p_1, p_2, \dots, p_5 分别为其

他5个关节的局部姿态, $p_1 = (\alpha_1, \beta_1, \gamma_1)$ 代表右肩关节的旋转欧拉角, 其他参数类似定义。

2 系统初始化

用作初始化的视频帧要求图像中的人体与成像面平行, 通过初始化, 可以得到以下参数: 各肢体长度及各关节在局部坐标系中的坐标, 各模板中像素灰度值及其在局部坐标系下的坐标, 初始姿态下对应的可变因子 s , 这些参数的具体计算方法参见文献[11]。手臂各模板中的像素值及局部坐标计算解释如下, 由于手臂肢体以圆柱建模, 故初始模板中的像素点在局部坐标系下的 z 坐标值不为零, 以平行于 xoy 平面的中心截面将圆柱体分成两部分, 故圆柱纹理模型以两个半圆柱体表示。如图2所示, 以纵切面分成两部分(前部及背部)。由标定信息可以得到手臂某肢体模板像素点在关节局部坐标系 x, y 方向下的局部坐标 (x_i, y_i) [11], 根据图2可知, z 方向的局部坐标 z_i 满足 $x_i^2 + z_i^2 = r^2$, 其中 r 为该肢体圆柱横切圆半径(标定可得到)。由于背部被遮挡, 故只能得到肢体前部的像素灰度值, 但可以得到背部中像素点局部坐标。

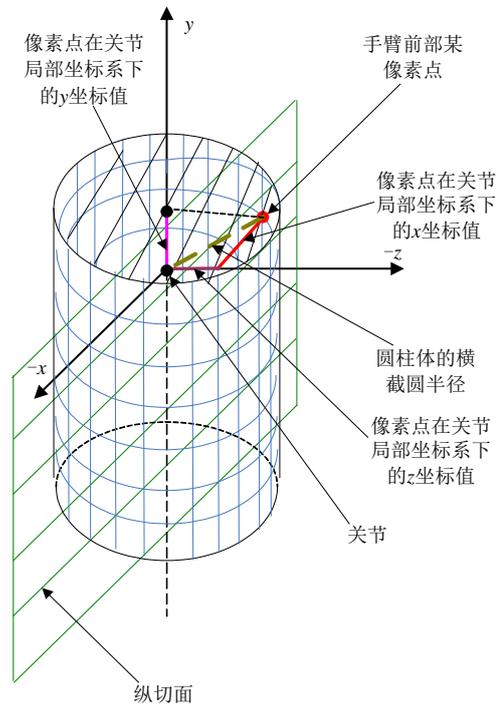


图2 手臂模板及模板中像素局部坐标计算示意图

3 基于模板配准的全局姿态估算

人体运动姿态采用基于区域的运动估计方法恢复, 其中全局姿态估计所用模板为左右胸肢体。

3.1 形变模型

根据运动学知识可知,在一定旋转下模板中某像素点旋转前的三维坐标与其在旋转后的三维坐标满足以下方程:

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} \quad (1)$$

式中, $[x_l, y_l, z_l]^T$ 为该像素点旋转前的三维坐标,在这里旋转为人体在当前帧的姿态相对于人体在初始帧姿态的旋转,因此 $[x_l, y_l, z_l]^T$ 为初始帧时获得的局部坐标; $[x_t, y_t, z_t]^T$ 为当前帧人体姿态下像素点在局部坐标系中的三维坐标; \mathbf{R} 为旋转矩阵定义如下:

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2)$$

式中,有:

$$r_{11} = \cos \alpha \cos \beta, r_{12} = -\sin \alpha \cos \gamma + \cos \alpha \sin \beta \sin \gamma$$

$$r_{13} = \sin \alpha \sin \gamma + \cos \alpha \sin \beta \cos \gamma$$

$$r_{21} = \sin \alpha \cos \beta$$

$$r_{22} = \cos \alpha \cos \gamma + \sin \alpha \sin \beta \sin \gamma$$

$$r_{23} = -\cos \alpha \sin \gamma + \sin \alpha \sin \beta \cos \gamma$$

$$r_{31} = -\sin \beta, r_{32} = \cos \beta \sin \gamma, r_{33} = \cos \beta \cos \gamma$$

α, β, γ 分别为绕 z 轴、 y 轴及 x 轴的旋转欧拉角。因此该像素点在成像空间中的三维坐标为:

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} + \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} \quad (3)$$

式中, $[x_r, y_r, z_r]^T$ 为根关节在三维空间中的坐标。根据透视投影模型可得,该像素旋转后其在图像上的坐标为:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{s} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad (4)$$

式中, $[u, v]^T$ 为该像素投影到成像面上的坐标; s 为可变因子,采用文献[11]所用方法可得:

$$s = z_c \cdot ds \quad (5)$$

式中, ds 为经过简单标定后得到的参数,表示相对于一个单位 z 变化量对应的 s 变化量。将式(3)代入式(4)得:

$$u = \frac{A}{B}, v = \frac{A}{C} \quad (6)$$

式中,有:

$$A = \cos \alpha \cos \beta \cdot x_l + (-\sin \alpha \cos \gamma + \cos \alpha \sin \beta \sin \gamma) \cdot$$

$$y_l + (\sin \alpha \sin \gamma + \cos \alpha \sin \beta \cos \gamma) \cdot z_l + s_r \cdot x$$

$$B = \sin \alpha \cos \beta \cdot x_l + (\cos \alpha \cos \gamma + \sin \alpha \sin \beta \sin \gamma) \cdot$$

$$y_l + (-\cos \alpha \sin \gamma + \sin \alpha \sin \beta \cos \gamma) \cdot z_l + s_r \cdot y$$

$$C = (-\sin \beta \cdot x_l + \cos \beta \sin \gamma \cdot y_l) \cdot ds + s_r$$

$(x, y, s_r, \alpha, \beta, \gamma)$ 为待估算的全局参数(x, y 为根关节在图像上的坐标, s_r 为根关节对应的可变因子);

(x_l, y_l, z_l) 为模板中像素点在根关节局部坐标系中的局部坐标。定义形变模型如下:

$$\mathbf{W}(x; \mathbf{p}) = \begin{bmatrix} W_x(x; \mathbf{p}) \\ W_y(x; \mathbf{p}) \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \quad (7)$$

式中, (u, v) 定义见式(6); $\mathbf{p} = [x, y, s_r, \alpha, \beta, \gamma]^T$; \mathbf{x} 为模板中的像素。

3.2 姿态估计

根据像素灰度不变性约束定义以下目标函数:

$$\sum_x [I(\mathbf{W}(x; \mathbf{p})) - T(x)]^2 \quad (8)$$

式中, $T(x)$ 为模板中的像素 \mathbf{x} 灰度值; \mathbf{p} 为当前帧下待恢复的全局姿态; $\mathbf{W}(x; \mathbf{p})$ 为形变模型, $I(\mathbf{W}(x; \mathbf{p}))$ 为模板像素 \mathbf{x} 根据形变模型 $\mathbf{W}(x; \mathbf{p})$ 映射到当前帧图像上的灰度值。最优姿态 \mathbf{p} 应该为初始姿态下模板像素与在姿态 \mathbf{p} 下模板投影到图像上的像素灰度差之和最小,即满足下式:

$$\hat{\mathbf{p}} = \min_{\mathbf{p}} \sum_x [I(\mathbf{W}(x; \mathbf{p})) - T(x)]^2 \quad (9)$$

式(9)为非线性优化问题,采用 Gauss-Newton 迭代优化方法来得到最优参数。假设第 $k+1$ 次迭代时 \mathbf{p}_k 估计值已知,则求取 \mathbf{p} 的增量 $\Delta \mathbf{p}$ 使得下式最小化,有:

$$\sum_x [I(\mathbf{W}(x; \mathbf{p} + \Delta \mathbf{p})) - T(x)]^2 \quad (10)$$

然后更新参数 $\mathbf{p}_{k+1} \leftarrow \mathbf{p}_k + \Delta \mathbf{p}$, 重复以上两步直至收敛。将 $I(\mathbf{W}(x; \mathbf{p}))$ 按照一阶泰勒展开得:

$$I(\mathbf{W}(x; \mathbf{p} + \Delta \mathbf{p})) \approx I(\mathbf{W}(x; \mathbf{p})) + \nabla I \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \Delta \mathbf{p} \quad (11)$$

式中, $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$ 为模板中的像素点根据形变模

型 $\mathbf{W}(x; \mathbf{p})$ 映射到当前帧上灰度梯度; $\frac{\partial \mathbf{W}}{\partial \mathbf{p}}$ 为形变

模型的雅克比矩阵定义如下:

$$\frac{\partial \mathbf{W}}{\partial \mathbf{p}} = \begin{pmatrix} \frac{\partial W_x}{\partial \alpha} & \frac{\partial W_x}{\partial \beta} & \frac{\partial W_x}{\partial \gamma} & \frac{\partial W_x}{\partial x} & \frac{\partial W_x}{\partial y} & \frac{\partial W_x}{\partial s_r} \\ \frac{\partial W_y}{\partial \alpha} & \frac{\partial W_y}{\partial \beta} & \frac{\partial W_y}{\partial \gamma} & \frac{\partial W_y}{\partial x} & \frac{\partial W_y}{\partial y} & \frac{\partial W_y}{\partial s_r} \end{pmatrix} \quad (12)$$

将式(11)代入式(10),并对 $\Delta \mathbf{p}$ 求导得:

$$2 \sum_x \left[\nabla I \frac{\partial W}{\partial \mathbf{p}} \right]^T \left[I(W(\mathbf{x}; \mathbf{p})) + \nabla I \frac{\partial W}{\partial \mathbf{p}} \Delta \mathbf{p} - T(\mathbf{x}) \right] \quad (13)$$

令式(13)等于零得:

$$\Delta \mathbf{p} = \left(\sum_x \left[\nabla I \frac{\partial W}{\partial \mathbf{p}} \right]^T \left[\nabla I \frac{\partial W}{\partial \mathbf{p}} \right] \right)^{-1} \times \left(\sum_x \left[\nabla I \frac{\partial W}{\partial \mathbf{p}} \right]^T [T(\mathbf{x}) - I(W(\mathbf{x}; \mathbf{p}))] \right) \quad (14)$$

全局姿态模板配准算法以伪代码描述如下:

Algorithm 1: Template Register

REPEAT

Warp I with $W(\mathbf{x}; \mathbf{p})$ to compute $I(W(\mathbf{x}; \mathbf{p}))$

Compute the intensity difference $T(\mathbf{x}) - I(W(\mathbf{x}; \mathbf{p}))$

Warp the gradient ∇I with $W(\mathbf{x}; \mathbf{p})$

Evaluate the Jacobian $\frac{\partial W}{\partial \mathbf{p}}$ at $(\mathbf{x}; \mathbf{p})$ using equ. (12)

Compute $\Delta \mathbf{p}$ using equ. (14)

Update the parameters $\mathbf{p} \leftarrow \mathbf{p} + \Delta \mathbf{p}$

UNTIL $\|\Delta \mathbf{p}\| < \varepsilon$

4 基于模板配准的关节局部姿态估算

其他关节的局部姿态估算方法与全局姿态估算方法类似, 所用模板为该关节旋转带动的肢体(如计算右肘关节姿态则模板为右前臂)。由上节可知, 模板中某一像素点旋转后其在三维空间中的坐标可以由式(3)计算得到, 但在计算关节局部姿态时, $[x_r, y_r, z_r]^T$ 为当前帧下该关节在运动链中父节点旋转后其在三维空间中的坐标, 根据前向运动学知识可知由式(3)计算得到, 本文以右肘关节姿态恢复为例解释其计算过程。

在估算右肘关节的局部姿态时右肩关节的姿态已估算完, 可得右肩关节相对于基坐标系的旋转矩阵为 \mathbf{R}_{rs} 以及右肩关节在成像三维空间中的坐标为 $[x_{rs}, y_{rs}, z_{rs}]^T$, 则根据式(3)可计算右肘关节在三维空间中的坐标 $[x_{re}, y_{re}, z_{re}]^T$ 如下:

$$\begin{bmatrix} x_{re} \\ y_{re} \\ z_{re} \end{bmatrix} = \mathbf{R}_{rs} \cdot \begin{bmatrix} 0 \\ -L \\ 0 \end{bmatrix} + \begin{bmatrix} x_{rs} \\ y_{rs} \\ z_{rs} \end{bmatrix} \quad (15)$$

式中, $[0, -L, 0]^T$ 为初始状态下右肘关节在右肩局部坐标系中的坐标(初始化时得到), L 为右上臂长度。

采用3.1节同样方法可得:

$$u = \frac{A}{C}, v = \frac{B}{C} \quad (16)$$

$$\begin{aligned} A &= \cos \alpha \cos \beta \cdot x_l + (-\sin \alpha \cos \gamma + \cos \alpha \sin \beta \sin \gamma) \cdot \\ &\quad y_l + (\sin \alpha \sin \gamma + \cos \alpha \sin \beta \cos \gamma) \cdot z_l + x_{re} \\ B &= \sin \alpha \cos \beta \cdot x_l + (\cos \alpha \cos \gamma + \sin \alpha \sin \beta \sin \gamma) \cdot \\ &\quad y_l + (-\cos \alpha \sin \gamma + \sin \alpha \sin \beta \cos \gamma) \cdot z_l + y_{re} \\ C &= (-\sin \beta \cdot x_l + \cos \beta \sin \gamma \cdot y_l) \cdot ds + z_{re} \cdot ds \end{aligned}$$

式中, $[x_{re}, y_{re}, z_{re}]^T$ 为根据式(15)计算得到的右肘关节在三维空间中的坐标, (α, β, γ) 为待恢复的右肘关节局部坐标系相对于基坐标系的旋转欧拉角。

采用3.2节类似方法可以计算得到右肘关节局部坐标系相对于基坐标系的旋转矩阵 \mathbf{R}_{re} 。实际应用中需要计算相对于父关节坐标系的旋转欧拉角 $(\alpha_r, \beta_r, \gamma_r)$ 。以右肘关节为例计算相对于右肩关节坐标系的旋转欧拉角, 先计算右肘局部坐标系相对于右肩局部坐标系的旋转矩阵 \mathbf{R}_{es} 如下:

$$\mathbf{R}_{es} = \mathbf{R}_{re} \cdot \mathbf{R}_{rs}^{-1} \quad (17)$$

设 $\mathbf{R}_{es} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$, 则 $(\alpha_r, \beta_r, \gamma_r)$ 计算得:

$$\begin{cases} \beta_r = \text{atan2}(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2}) \\ \alpha_r = \text{atan2}\left(\frac{r_{21}}{\cos \beta}, \frac{r_{11}}{\cos \beta}\right) \\ \gamma_r = \text{atan2}\left(\frac{r_{32}}{\cos \beta}, \frac{r_{33}}{\cos \beta}\right) \end{cases} \quad (18)$$

式中, $\text{atan2}(x, y)$ 为计算参数为 y/x 的反正切函数。

5 多模板配准

当人体只有左右及前后运动时(即没有肢体自旋转, $\beta = 0$), 以上方法可以准确地恢复人体上半身三维运动姿态。但是当全局姿态或关节局部姿态中的欧拉角 $\beta \neq 0$ 时, 则初始化时提取的模板会旋转到肢体相对于摄像机的背面, 发生部分或全部遮挡。在这种情况下, 由初始化获得的单一模板跟踪效果急剧下降, 严重时会发生跟踪丢失现象。因此本文采用多模板配准来解决该问题, 原理以伪代码描述如下:

Algorithm 2: Multi-template Register

FOR $i=1:N_t$

Draw template T_i from collection C

Use the methods proposed in sections 4 and 5 to reconstruct the pose of human upper body \mathbf{P}_i , and calculate the mean intensity difference, E_i , according to equ. (19)

IF $E_i < E$

```

    E = Ei
    P = Pi
  END IF
END FOR
Output the optimal pose P
Obtain the template Tc according to pose P
IF E < T
  IF Nr < Nmax
    Add Tc into template collection C
  ELSE
    Replace the oldest template with Tc
  END IF
END IF
END IF

```

其中, N_i 为模板库的模板数; N_{\max} 为模板库中允许的最大模板数。每次跟踪后, 计算每个模板中像素平均灰度差值 E_i 如下:

$$E_i = \frac{\sum_x [I(W(x; p)) - T(x)]^2}{n} \quad (19)$$

式中, n 为模板中的像素点数。

$$E = \min\{E_i | i = 1: N_i\} \quad (20)$$

如果最小差值 E 小于阈值 T , 则表示当前跟踪结果可靠, 可以生成新的模板。

6 实验及结果分析

为了验证算法的有效性, 本文采用 vs2010+opencv2.2 实现了一个原型系统。实验所用视频数据为三星 S600 数码相机拍摄, 分辨率为 640*480, 实验视频未做任何预处理, 人体跟踪均在原始图像上进行。实验中使用模板数为 10 个。实验在主频为 1.58-GHZ 的计算机上进行, 视频跟踪的平均处理速率为 1 分钟/帧, 原因在于基于金字塔的多模板迭代收敛过程比较耗时。目前的系统还不能达到实时跟踪的目标, 但是将代码移植到并行机上运行, 并进行代码优化, 应该可以达到实时跟踪的效果。

图3为一段视频总长为 689 帧小孩蹦跳运动跟踪部分结果, 在整个运动过程中, 小孩不断改变自身姿态并且摄像机不停移动, 造成初始化时提取的肢体前部模板发生部分遮挡, 如第 5 个视频帧中, 人体右臂已旋转到侧面, 其初始模板大部分在图像中不可见, 本文采用了多模板配准方法可以处理该情形。从实验结果可以看出二维跟踪结果与肢体所在图像位置基本一致。图4分别为前 70 帧提取的左肩及左肘

关节旋转欧拉角, 由图可知, 旋转角轨迹平滑, 没有明显的跳跃点。因此, 本文算法跟踪比较准确, 鲁棒性强。

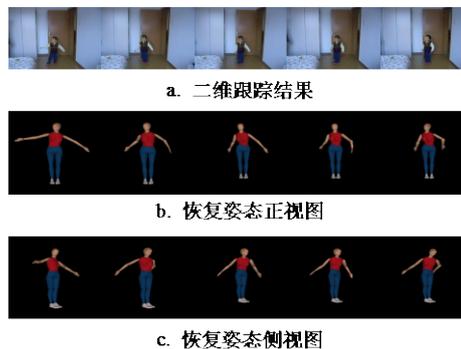


图3 部分遮挡视频跟踪结果

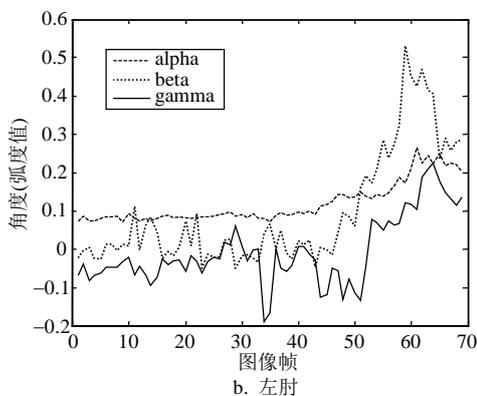
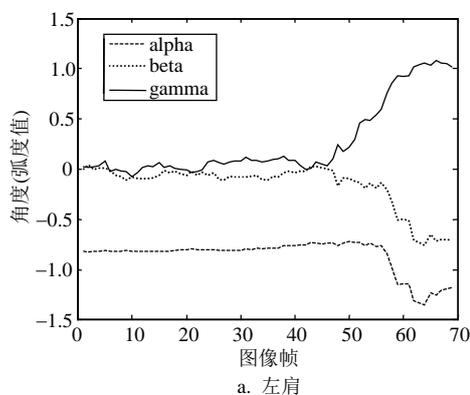


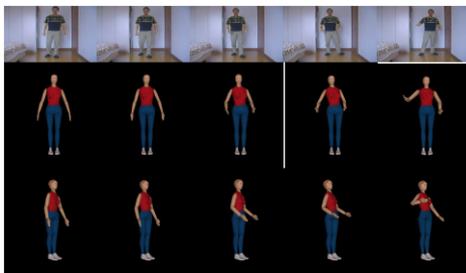
图4 左臂关节欧拉角轨迹图

图5为本文算法与文献[3]跟踪一段时长 365 帧的人体运动结果比较。由于视频中的人体姿态具有强透视投影效果, 而文献[3]采用比例正交投影模型不能对真实投影进行准确建模, 故本文算法的跟踪效果要优于文献[3], 从图5的跟踪结果比较可以验证。本文采用手工标定方法从各视频帧中得到图像上关节二维坐标真实值, 并将关节二维跟踪结果与真实值比较。表1显示了本文算法与文献[3]跟踪算法各关节平均跟踪误差, 从比较结果可以看出, 本文算法的平均误差要小于文献[3]的平均误差。

表1 二维关节跟踪误差统计表

关节名	平均跟踪误差	
	本文算法	文献[3]所提算法
右肘关节	6.35	8.26
右腕关节	7.73	10.25
左肘关节	6.14	8.06
左腕关节	7.86	11.32

图6为本文算法与文献[11]跟踪一段时长568帧的人体运动结果比较。在该运动过程中由于全局姿态变化或手臂局部姿态变化造成初始模板部分不可见。在这种情况下,文献[11]由于采用单一模板,不能进行模板更新,而本文算法新增模板能够反映最新肢体信息,故在实际跟踪中取得了更好的效果。



a. 本文算法跟踪结果



b. 文献[3]算法跟踪结果

图5 与比例正交投影模型匹配算法视频跟踪结果比较



a. 本文跟踪结果



b. 文献[13]跟踪结果

图6 与单模板匹配算法视频跟踪结果比较

7 结论

本文针对单一模板跟踪不能应用于模板被遮挡下的跟踪问题,提出了一种多模板匹配跟踪算法。该算法通过动态增加新模板来不断更新模板信息,由于使用了多模板,很难发生所有模板都不可见现象,以提高算法鲁棒性。同时由于算法采用了透视投影模型,可以处理强透视投影下的人体三维运动

跟踪,使得算法应用更加广泛。实验表明,算法在鲁棒性及适用性方面都较单模板算法有所提高。

参 考 文 献

- [1] THOMAS B M, ADRIAN H, VOLKER K. A survey of advances in vision-based human motion capture and analysis[J]. Computer Vision and Image Understanding, 2006(104): 90-126.
- [2] RONALD P. Vision-based human motion analysis: An overview[J]. Computer Vision and Image Understanding, 2007(108): 4-18.
- [3] CHRISTOPH B, JITENDRA M. Tracking people with twists and exponential maps[C]//IEEE Conference on Computer Vision and Pattern Recognition. Santa Barbara, USA: IEEE, 1998: 8-15.
- [4] GALL J, ROSENHAHN B, BROX T, et al. Optimization and filtering for human motion capture-a multi-layer framework[J]. International Journal of Computer Vision, 2010, 87(1): 75-92.
- [5] VONDRAK M, SIGAL L, JENKINS C. Physical simulation for probabilistic motion tracking[C]//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, USA: IEEE, 2008: 1-8.
- [6] NING H, TAN T, WANG L. People tracking based on motion model and motion constraints with automatic initialization[J]. Pattern Recognition, 2004, 37(7): 1423-1440.
- [7] BREGLER C, MALIK J, PULLEN K. Twist based acquisition and tracking of animal and human kinematics[J]. International Journal of Computer Vision, 2004(56): 179-194.
- [8] DAUBNEY B, XIE X. Tracking 3D human pose with large root node uncertainty[C]//IEEE Conference on Computer Vision and Pattern Recognition. Colorado, USA: IEEE, 2011: 1321-1328.
- [9] CHANG I, LIN S. 3D human motion tracking based on a progressive particle filter[J]. Pattern Recognition, 2010(43): 3621-3635.
- [10] 陈加, 吴晓军. 联合LBS和Snake的3D人体外形和运动跟踪方法[J]. 计算机辅助设计与图形学学报, 2012, 24(3): 357-371.
CHEN Jia, WU Xiao-jun. 3D human body shape and motion tracking by LBS and Snake[J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(3): 357-371.
- [11] 邹北骥, 陈姝, 彭小宁, 等. 适用于单目视频的无标记三维人体运动跟踪[J]. 计算机辅助设计与图形学学报, 2008, 20(8): 1047-1055.
ZOU Bei-ji, CHEN Shu, PENG Xiao-ning, et al. Markerless 3D human motion tracking for monocular video sequences[J]. Journal of Computer-Aided Design & Computer Graphics, 2008, 20(8): 1047-1055.

编辑 蒋 晓