

透视投影下基于轨迹基的非刚体三维结构恢复研究

王 晶, 包晓敏, 汪亚明

(浙江理工大学信息学院, 杭州 310018)

摘 要: 基于预定义轨迹基重建的非刚体运动重建算法通常采用因式分解方法,该方法在仿射摄像机模型下恢复物体的结构信息和运动信息,未考虑深度信息,从而使算法在深度变化明显时具有较大重建误差甚至呈现病态。在基于预定义轨迹基的非刚体运动重建算法中引入一种线性迭代过程,使图像深度信息得以恢复,从而使算法在透视投影模型下可以适用。人体运动重建结果表明,提出的方法是可行和有效的。

关键词: 非刚体; 轨迹基; 深度信息; 透视投影; 仿射模型

中图分类号: TP391.4 **文献标志码:** A

0 引 言

在计算机视觉领域中,从图像序列中恢复物体的结构信息和运动信息是一个重难点,尤其是对于非刚体物体的恢复。在过去的20年中,大量的方法被不断提出,其中,由于因式分解法具有较好的稳定性和精确性,已经被广泛应用于各个领域。

Tomasi等^[1]首次提出了因式分解法,该方法运用奇异值分解(SVD)方法,在秩约束的条件下,将由图像二维特征点组成的测量矩阵直接分解为旋转矩阵和结构矩阵。他们提出的方法比较适合应用在正交投影模型下,Poelman等^[2]随后将该法分别推广到透视和弱透视投影模型下。Breger等^[3]将非刚体看作是刚性形状基(shape base)的线性加权组合,基于该思想,他们首次将因式分解法应用于非刚体的三维重建。Xiao等^[4]引进了基约束来解决结果的歧义性。基于三维空间中特征点的时间平滑性,Akhter等^[5]之后又提出了一种基于轨迹基的非刚体结构和运动恢复的算法,即将特征点的运动轨迹看作是一系列轨迹基(trajjectory base)的线性加权组合。大多数的因式分解方法都是基于在仿射摄像机模型下恢复物体的结构信息和运动信息的。这

种模型是对真实透视投影模型的零阶近似(弱透视)模型或一阶近似(透视)模型,该假设只有在深度变化和物体的尺寸相对于物体到摄像机的距离比较小的时候才可成立。这就意味着在恢复过程中会丢失大量的深度信息。

针对这一问题,本文在轨迹空间中,将因式分解法和线性迭代方法相结合,根据弱透视投影与透视投影之间的近似关系,通过对矩阵进行加权,运用迭代算法计算出深度信息,最终用弱透视投影模型下的结果来近似得到在真实透视投影模型下的重建结果,从而实现在透视投影模型下非刚体的结构和运动恢复。

1 仿射模型下的非刚体结构

1.1 非刚体的因式分解

在形态空间中,运动结构恢复(structure from motion)的问题可描述为:给定 F 帧图像序列的 P 个特征点 $\{x_{ij} | i=1, \dots, F, j=1, \dots, P\}$ 组成的矩阵 $W_{2F \times P}$,恢复出物体三维结构矩阵 S 和旋转矩阵 R 。

研究通常假设非刚体模型是 K 个刚体形状基的近似加权线性组合,即

收稿日期: 2013-09-05

基金项目: 国家自然科学基金(61272311, 61070063)

作者简介: 王 晶(1988-),女,河南辉县人,研究方向为图像处理。

通信作者: 包晓敏, E-mail: zjzhbxm@gmail.com

$$\tilde{S}_i = \sum_{l=1}^K \omega_{il} B_l \quad (1)$$

式(1)中, ω_{il} 为加权系数, B_l 为形态基, K 为形态基个数。当 $K=1, \omega_{il}=1$ 时, 则对应刚体结构的情况。

在仿射模型中, 即非刚体的深度变化远小于它到摄像机的距离, 对第 i 帧图像, 有如下关系

$$\begin{bmatrix} u_{i1} & \cdots & u_{iP} \\ v_{i1} & \cdots & v_{iP} \end{bmatrix} = \bar{\mathbf{R}}_i \left(\sum_{l=1}^K \omega_{il} B_l \right) + \bar{\mathbf{T}}_i \mathbf{e}_n^T \quad (2)$$

式(2)中, $\bar{\mathbf{R}}_i$ 是旋转矩阵 $\mathbf{R}_i (3 \times 3)$ 的前两行, $\bar{\mathbf{T}}_i$ 是平移向量 $\mathbf{T}_i (3 \times 3)$ 的前两个元素, $\mathbf{e}_n^T = [1, \dots, 1] (1 \times n)$ 是单位行向量。

对每幅图像特征点进行坐标变换, 可以使图像特征点坐标的原点位于该幅图像点集的质心, 这样可以消除掉式(2)中的平移向量 $\bar{\mathbf{T}}_i$, 则上式可以简化为

$$\begin{bmatrix} u_{i1} & \cdots & u_{iP} \\ v_{i1} & \cdots & v_{iP} \end{bmatrix} = \bar{\mathbf{R}}_i \left(\sum_{l=1}^K \omega_{il} B_l \right) \quad (3)$$

则非刚体的测量矩阵 $\tilde{\mathbf{W}}_{2F \times P}$ 可以表示为

$$\tilde{\mathbf{W}}_{2F \times P} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1P} \\ v_{11} & v_{12} & \cdots & v_{1P} \\ \vdots & \vdots & \ddots & \vdots \\ u_{F1} & u_{F2} & \cdots & u_{FP} \\ v_{F1} & v_{F2} & \cdots & v_{FP} \end{bmatrix} = \begin{bmatrix} \omega_{11} \bar{\mathbf{R}}_1 & \cdots & \omega_{1K} \bar{\mathbf{R}}_1 \\ \vdots & \ddots & \vdots \\ \omega_{F1} \bar{\mathbf{R}}_F & \cdots & \omega_{FK} \bar{\mathbf{R}}_F \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_K \end{bmatrix} = \mathbf{M}_{2F \times 3K} \mathbf{B}_{3K \times P} \quad (4)$$

式(4)中, $\mathbf{M}_{2F \times 3K}$ 表示非刚体的运动矩阵, $\mathbf{B}_{3K \times P}$ 表示形状矩阵。

1.2 轨迹空间中的非刚体因式分解

根据轨迹基和形态基的对偶性定理^[6], 即轨迹基系数就是形态基, 而轨迹基则是形态基的系数, 本文采用预定义的普通离散小波变换(discrete wavelet transform, DWT)基。所以结构矩阵可以表示为一个包含轨迹基和轨迹基系数的乘积的反向投影矩阵, 即

$$\mathbf{S}_{3F \times P} = \Theta_{3F \times 3K} \mathbf{A}_{3K \times P} \quad (5)$$

则式(4)可被分解为

$$\tilde{\mathbf{W}} = \mathbf{R} \Theta \mathbf{A} = \mathbf{A} \mathbf{A} \quad (6)$$

式(6)中 $\mathbf{A} = \mathbf{R} \Theta$ 是一个 $2F \times 3K$ 的矩阵。当 K 个基分别在 X, Y, Z 三个方向上重建轨迹基时, 测量矩阵 $\tilde{\mathbf{W}}$ 的秩最大为 $3K$ 。这是秩约束的对偶定理^[7]。

对 $\tilde{\mathbf{W}}$ 进行 SVD 分解, 可得 $\tilde{\mathbf{W}} = \hat{\mathbf{\Lambda}} \hat{\mathbf{A}}$, 其中 $\hat{\mathbf{A}}$ 是

一个 $2F \times 3K$ 的矩阵, $\hat{\mathbf{A}}$ 是一个 $3K \times P$ 的矩阵。但由于 SVD 分解的不惟一性, 所以 $\hat{\mathbf{A}}$ 并不等于 \mathbf{A} , $\hat{\mathbf{A}}$ 也不等于 \mathbf{A} 。为了恢复出结构矩阵, 需要估计一个校正矩阵 \mathbf{Q} (非奇异的正交矩阵^[8]) 来满足下面的约束条件:

$$\begin{aligned} \mathbf{A} &= \hat{\mathbf{A}} \mathbf{Q} \\ \mathbf{A} &= \mathbf{Q}^{-1} \hat{\mathbf{A}} \end{aligned} \quad (7)$$

为了计算校正矩阵 \mathbf{Q} , 考虑矩阵 \mathbf{A} :

$$\mathbf{A} = \begin{bmatrix} \theta_{11} R_1 & \cdots & \theta_{1K} R_1 \\ \vdots & & \vdots \\ \theta_{F1} R_F & \cdots & \theta_{FK} R_F \end{bmatrix} \quad (8)$$

由式(8)可知, 无需得到 \mathbf{Q} 的整个矩阵, 只需求出前三列向量即可, 不妨定义 $\mathbf{Q}_{||}$ 为 \mathbf{Q} 的前三列向量。结合式(7)和(8), 可得

$$\hat{\mathbf{A}} \mathbf{Q}_{||} = \begin{bmatrix} \theta_{11} R_1 \\ \vdots \\ \theta_{F1} R_F \end{bmatrix} \quad (9)$$

式(9)表明: 当 $\mathbf{Q}_{||}$ 已知时, 摄像机的投影矩阵可以计算得出, 则摄像机的旋转矩阵 \mathbf{R} 也就可以计算出来, \mathbf{R} 乘以预定义的轨迹基 $\Theta_{3F \times 3K}$ 便可以恢复出矩阵 \mathbf{A} , 即

$$\mathbf{A}_{2F \times 3K} \mathbf{A}_{3K \times P} = \mathbf{W}_{2F \times P} \quad (10)$$

为了计算 $\mathbf{Q}_{||}$, 使用摄像机投影矩阵 \mathbf{R}_i 的正交约束条件, 类似于文献[1]中的方法。对于每张图像 i , 都满足正交约束条件^[9]。

$$\hat{\mathbf{A}}_{2i-1:2i} \mathbf{Q}_{||} \mathbf{Q}_{||}^T \hat{\mathbf{A}}_{2i-1:2i}^T = \theta_{i1}^2 \mathbf{I}_{2 \times 2} \quad (11)$$

式(11)中, $\hat{\mathbf{A}}_{2i-1:2i}$ 表示矩阵 $\hat{\mathbf{A}}$ 的第 $2i-1$ 行到第 $2i$ 行的行向量, $\mathbf{I}_{2 \times 2}$ 是一个单位矩阵。对 F 帧图像, 将有 $3F$ 个约束条件, $9K$ 个未知量 ($\mathbf{Q}_{||}$), 则至少需要 $3K$ 张非退化图像来计算 $\mathbf{Q}_{||}$ 。最后根据校正矩阵计算出结构矩阵和旋转矩阵。

2 透视投影下的非刚体结构恢复

采用因式分解法得到的非刚体三维结构信息, 大都是基于仿射摄像机模型的。且在实验过程中, 采用简化摄像机模型, 设所有特征点的质心均处于图像的几何中心, 因此已忽略了非刚体结构的深度信息, 这样会造成较大的误差。本文根据弱透视投影和透视投影之间的关系, 采用线性迭代的方法, 对弱透视投影模型下的恢复结果进行线性迭代, 得出深度信息, 来逼近真实透视投影模型下的结果。

2.1 迭代法简介

迭代是数值分析中, 从一个初始估计出发寻找

一系列近似解来解决问题的过程,为实现这一过程所使用的方法统称为迭代法^[10]。其基本思想如下:

对于给定的线性方程组 $Ax=b \Leftrightarrow x=Mx+g$, 取初始向量 $x^{(0)} \in R^n$, 构造如下单步定常线性迭代公式:

$$x^{(k+1)} = Mx^{(k)} + g \quad (k=0,1,2,\dots) \quad (12)$$

以此来产生近似向量序列 $x^{(1)}, x^{(2)}, \dots$ 。当 k 充分大时, $x^{(k)} \approx x^*$ (x^* 为线性方程组的解)。

2.2 深度信息的获得

假设摄像机已标定,则在透视投影模型下,空间点 X_i 的投影点是 x_i ,他们的关系如下:

$$m_i x_i = \mathbf{R}X_i + t \quad (13)$$

式(13)中, m_i 是一非零系数, $x_i = [u_i \ v_i \ 1]^T$ 是图像特征点的齐次坐标, $\mathbf{R} = [r_1^T \ r_2^T \ r_3^T]^T$ 为 3×3 旋转矩阵, $\mathbf{T} = [t_x \ t_y \ t_z]^T$ 为 3×3 平移向量。

根据式(13),求得:

$$\begin{cases} u_i = \frac{\frac{r_1 X_i + t_x}{t_z}}{1 + \frac{r_3 X_i}{t_z}} = \frac{\frac{r_1 X_i + t_x}{t_z}}{1 + \epsilon_i} \\ v_i = \frac{\frac{r_2 X_i + t_y}{t_z}}{1 + \frac{r_3 X_i}{t_z}} = \frac{\frac{r_2 X_i + t_y}{t_z}}{1 + \epsilon_i} \end{cases} \quad (14)$$

假设物体的深度和尺寸变化相对于其到摄像机的距离很小,则弱透视投影就是透视投影的零阶近似,即 $\epsilon_i = 0$ 。令 $\lambda_i = 1 + \epsilon_i$, 便可得到弱透视与透视投影的关系如式(15):

$$x_i = \lambda_i x_i \quad (15)$$

图1为透视投影模型及其近似模型的比较。

根据该关系,可对式(4)进行加权,即:

$$\hat{\mathbf{W}}_{2F \times P} = \begin{bmatrix} \hat{x}_{11} & \cdots & \hat{x}_{1P} \\ \vdots & \ddots & \vdots \\ \hat{x}_{F1} & \cdots & \hat{x}_{FP} \end{bmatrix} = \begin{bmatrix} \lambda_{11} x_{11} & \cdots & \lambda_{1P} x_{1P} \\ \vdots & \ddots & \vdots \\ \lambda_{F1} x_{F1} & \cdots & \lambda_{FP} x_{FP} \end{bmatrix} \quad (16)$$

首先初始化 $\lambda_{ij} = 1 (i=1,2,\dots,F; j=1,2,\dots,P)$, 根据式(16)和式(12)更新测量矩阵 \mathbf{W} , 对测量矩阵进行线性迭代,可以求出迭代系数 λ_{ij} , 然后重复该法直至 λ_{ij} 收敛(即本次求得 λ_{ij} 的值与上次求得的 λ_{ij} 值达到误差允许的范围), 最终近似恢复出透视投影模型下的结构和运动信息。

图1中, x_0, x_w, x_p 和 x 分别表示特征点 P 在正交投影、弱透视投影、平行透视投影和透视投影模型下的投影点, G 表示被恢复物体的质心。

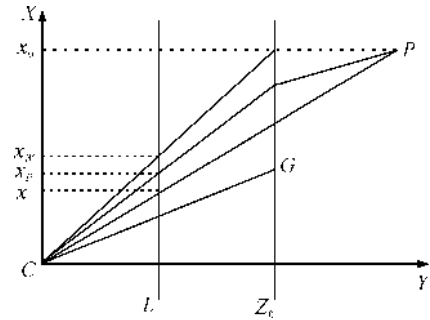


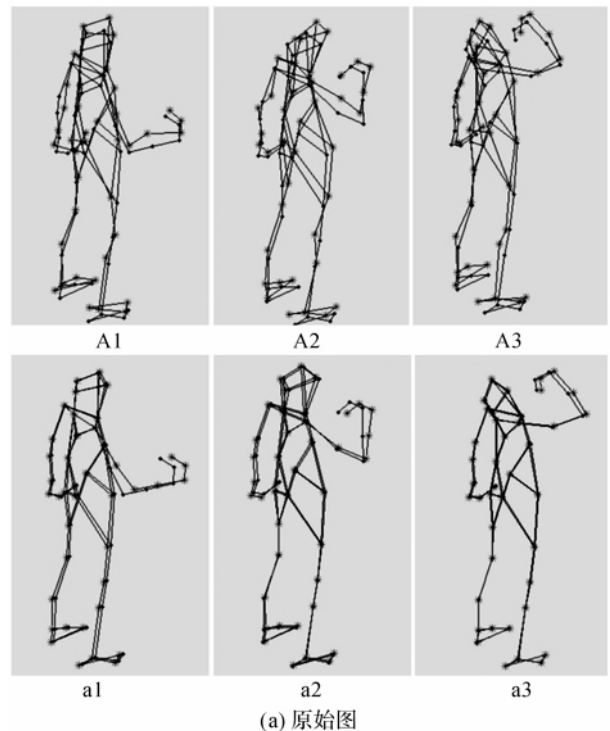
图1 透视投影模型及其近似模型比较

3 实验结果及分析

实验采用了1102帧drink的真实图像特征点序列来对以上算法进行试验,每帧图像提取41个特征点。本实验采用简化摄像机模型,假设所有特征点的质心均处于图像的几何中心,并且摄像机的各个内置参数在整个摄影过程中均保持不变。实验中,轨迹基数目 K 都取为12。本文使用的数据来源于 <http://cvlab.lums.edu.pk/nrsfm>。

以下是任选三幅图像的实验结果。“•”是真实结构特征点,“*”是通过算法恢复出来的点。

由图2可直观地看出本文的算法较之前的算法有更好的效果,误差有了明显的改善。其中在手臂部位还有较明显的误差,存在不足,需要进一步针对这一问题展开讨论。



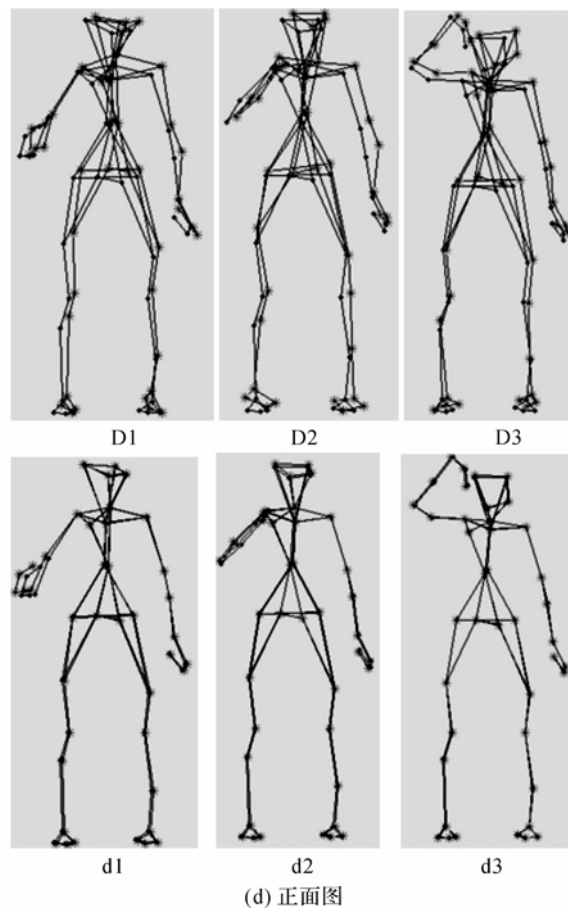
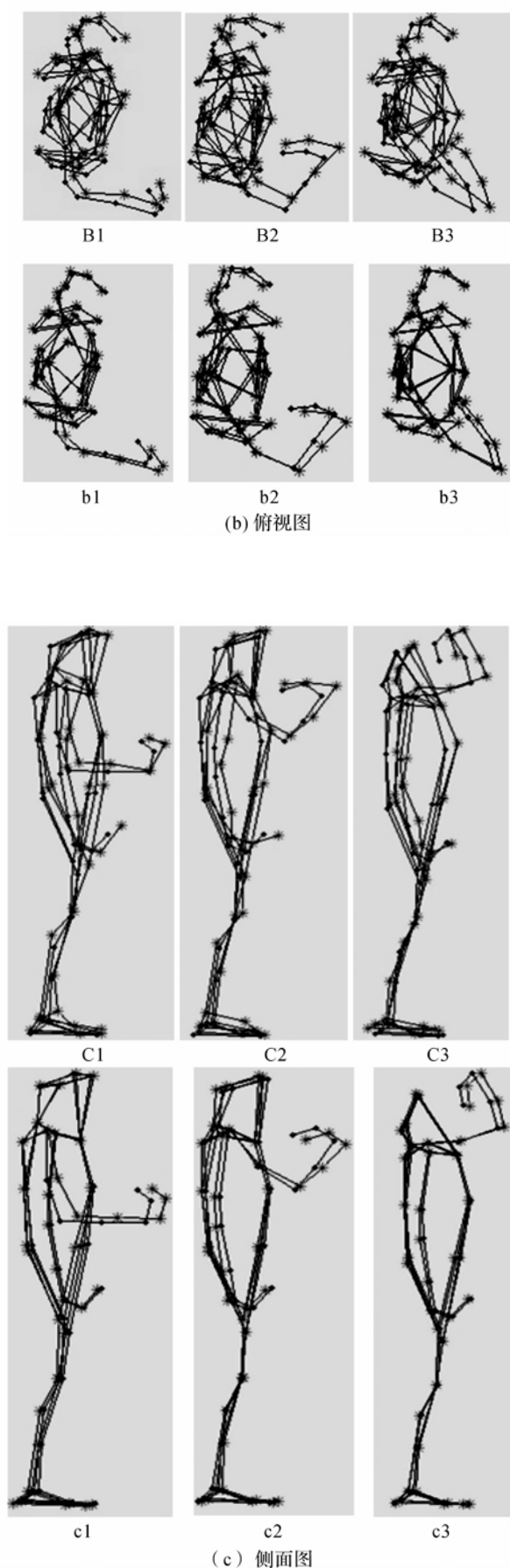


图 2 3 帧图像的实验结果

注: A1、A2、A3, B1、B2、B3, C1、C2、C3, D1、D2、D3 分别是放射模型下的实验结果, a1、a2、a3, b1、b2、b3, c1、c2、c3, d1、d2、d3 是透视投影模型下相对应的实验结果。

为了定量地分析和比较本算法的性能,根据反投影误差^[10],表 1 给出了先前方法和与本文方法的重建结果的误差对比结果。

反投影误差定义如下:

$$\sigma = \frac{\|\mathbf{W} - \mathbf{W}_r\|_F}{\|\mathbf{W}\|_F} \times 100\%$$

其中, \mathbf{W} 为原始测量矩阵, \mathbf{W}_r 为恢复结果的反投影测量矩阵。

表 1 结构恢复误差分析比较

试验方法	仿射模型方法	改进后的透视投影模型方法
反投影误差/%	0.16	0.10

4 结 语

本文在基于轨迹基的因式分解的基础上,介绍了一种线性迭代方法。根据弱透视投影与透视投影的关系,对分解矩阵进行加权。通过迭代法计算出投影深度,从而减小误差。对真实的图像序列特征点进行实验,具有较好的结果。本文只是通过迭代

法来近似透视投影模型下的结果,为确保重建算法的精确性和适用性,需进一步研究直接在透视投影模型下三维结构的恢复算法。

参考文献:

- [1] Tomasi C, Kanade T. Shape and motion from image streams; a factorization method[J]. International Journal of Computer Vision, 1992, 9(2): 137-154.
- [2] Poelman C, Kanade T. A paraperspective factorization method for shape and motion recovery[J]. IEEE Trans Pattern Anal Mach Intell, 1997, 19(3): 206-218.
- [3] Bregler C, Hertzmann A, Biermann H. Recovering non-rigid 3D shape from image stream[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2000, 2: 690-696.
- [4] Xiao J, Chai J, Kanade T. A closed-form solution to non-rigid shape and motion recovery[J]. International Journal of Computer Vision, 2006, 67: 573-587.
- [5] Akhter I, Sheikh Y, Khan S, et al. Nonrigid structure from motion in trajectory space[C]//Neural Information Processing Systems, 2008: 1534-1541.
- [6] Akhter I, Sheikh Y, Khan S. Trajectory space: a dual representation for nonrigid structure from motion[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(7): 1442-1456.
- [7] 孙国强. 基于图像序列的非刚性物体三维结构及运动恢复[D]. 长春: 吉林大学, 2010.
- [8] Torresani L, Hertzmann A, Bregler C. Nonrigid structure from motion: estimating shape and motion with hierarchical priors[C]//IEEE Trans. Pattern Analysis and Machine Intelligence, 2008, 30(5): 878-892.
- [9] Akhter I, Sheikh Y, Khan S. In defense of orthonormality constraints for nonrigid structure from motion[C]//IEEE Computer Vision and Pattern Recognition (CVPR), Miami FL, June 2009: 1534-1541.
- [10] 李庆扬, 王能超, 易大义. 数值分析[M]. 武汉: 华中科技大学出版社, 1986: 145-150.
- [11] 倪晓庆. 织物的三维动态仿真研究[D]. 杭州: 浙江理工大学, 2012.

Research on Recovery of Three-dimensional Structure of Non-rigid Body Based on Trajectory Basis under Perspective Projection

WANG Jing, BAO Xiao-min, WANG Ya-ming

(The School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: To improve the stability of motion restructuring algorithm of non-rigid body, people put forward a restructuring method based on predefined trajectory basis. However, this method is still based on the affine model which does not consider depth information, thus making the algorithm have big restructuring error and even show ill condition when depth changes greatly. In allusion to this problem, this paper introduces a linear iteration process for motion restructuring algorithm of non-rigid body based on predefined trajectory basis, thus recovering image depth information and making the algorithm applicable under perspective projection model. The analysis on human motion restructuring result shows that the method put forward in this paper is feasible and effective.

Key words: non-rigid body; trajectory basis; depth information; perspective projection; affine model

(责任编辑: 陈和榜)