

点状特征柔性物体三维运动捕获方法^①

廖芳, 史金龙, 龚肖

(江苏科技大学 计算机学院, 镇江 212003)

通讯作者: 史金龙, E-mail: jlshifudan@gmail.com

摘要: 针对具有点状特征的柔性物体, 提出了一种三维运动捕获方法. 首先, 该方法利用两个标定的高速摄像机拍摄柔性物体的运动视频, 并对图像进行立体校正; 然后, 采用 DOG (Difference Of Gaussian) 算法获取点状特征的位置, 并提取特征点极值; 其次, 在一定范围的窗口上搜索匹配对, 匹配左右图像的特征点; 再次, 通过三角测量法进行三维重建; 最后, 利用搜索策略进行时间序列上的匹配, 实现动态柔性物体的三维运动捕获, 并计算空间坐标、速度、加速度参数. 实验结果表明, 相比于采用 sift 算法匹配特征点捕获柔性运动物体的方法, 本方法精度更高.

关键词: 柔性物体; 匹配; 三维重建; 运动捕获

引用格式: 廖芳, 史金龙, 龚肖. 点状特征柔性物体三维运动捕获方法. 计算机系统应用, 2018, 27(7): 230-235. <http://www.c-s-a.org.cn/1003-3254/6447.html>

3D Motion Capture Method of Point Features Flexible Object

LIAO Fang, SHI Jin-Long, GONG Xiao

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: In this study, a 3D motion capture method is proposed for a flexible object with point features. Firstly, the method utilizes two calibrated high-speed cameras to capture motion video of a flexible object, and makes a stereo rectification of the image. Secondly, through the Difference Of Gaussian (DOG) algorithm, the positions of the feature points are obtained, and the extreme of feature points is extracted. Thirdly, matching pairs in a certain range of windows are searched, the feature points of the left and right images are matched. Thirdly, 3D reconstruction is realized by triangulation. Finally, using the search strategy to match the time series, the 3D motion capture of a dynamic flexible object is realized, and the spatial coordinates, velocity, and acceleration parameters are calculated. The experimental results show that the method is more accurate than the method of using SIFT algorithm to match the feature points to capture the flexible moving object.

Key words: flexible objects; matching; 3D reconstruction; motion capture

近年来三维运动捕获技术日趋成熟, 市场上生成了大量三维运动捕获系统, 该技术被应用于动画制作、影视特效、体育训练、医学仿真等领域之中.

文献[1-4]提出运动捕获从工作原理来划分, 主要分为: 机械式运动捕获、声学式运动捕获、电磁式运动捕获、光学式运动捕获和基于视频的运动捕获. 文

献[5,6]提出了基于视频的运动捕获, 将人体模型从视频序列中分离出来后, 再通过传统的匹配算法匹配和重建. 该方法对于人体节点简易模型进行匹配效果较好, 但是对无模型的多点柔性物体效果并不理想. 文献[7]提出了一种基于稠密的非刚性重建方法, 对薄的物体进行重建, 效果较为理想, 但是对于稀疏的点状柔性物体并

① 基金项目: 江苏省六大人才高峰专项 (1612991602); 中国博士后基金 (2014M560417)

Foundation item: Specific Foundation for Six Top Talents in Jiangsu Province (1612991602); Postdoctoral Foundation in China (2014M560417)

收稿时间: 2017-11-19; 修改时间: 2017-12-15; 采用时间: 2017-12-27; csa 在线出版时间: 2018-06-27

不适用. 文献[8]提出了一种基于图形硬件的点的曲面重构和可视化技术, 由于点不能很好地适应多边形网格图形硬件, 所以效果一般且设备较昂贵. 文献[9]提出了一种无标记的服装变化运动捕获方法, 该方法基于多目立体视觉原理和多视图几何的关系来恢复服装的三维数据^[10]. 该方法重建效果较好, 但会产生庞大的数据量, 因此处理起来极其复杂.

本文提出了一种基于双目视觉的点状特征柔性物体运动捕获方法, 该方法针对传统的 SIFT 算法^[11,12]对较大的圆点匹配较困难的问题, 提出了一种改进的匹配方法. 该方法首先对柔性物体上的点状特征进行匹配, 再进行三维重建和时间序列上的匹配, 最终实现动态柔性物体的三维运动捕获.

1 特征提取

第一步: 利用 DOG 方法进行角点检测^[13], 如公式(1)所示. 首先利用不同的高斯核对图像进行卷积; 然后再进行归一化处理; 最后获得高斯差分后的图像. 经过高斯差分后的图像, 点状特征的边缘是暗的, 越靠近圆心越亮, 此结果表明圆心的灰度值最大.

$$f(x; \mu, \sigma_1, \sigma_2) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma_1^2}\right) - \frac{1}{\sigma_2 \sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma_2^2}\right) \quad (1)$$

第二步: 检测极值获取点状特征的圆心初始值. 以 3×3 的滤波器为例, 对图像中的每一个点进行遍历, 把这个点与其二维领域其他 8 个点进行比较, 如果其灰度值最大则为极大值, 然而有些极值并不是圆心, 而是圆心邻域的点, 将坐标值仅相差一个或两个像素范围的定位为同一圆心邻域的点, 再对同一邻域的点取均值从而获得精确的圆心.

2 特征匹配

本文对传统的 sift 算法进行改进, 采用搜索策略来描述关键点. 首先, 经过立体标定和立体校正后^[14,15], 按上节所述特征提取方法获取左右两幅图像的特征点坐标; 再分别比较单幅图像上前后坐标值是否重复, 通过取均值剔除掉重复点; 其次, 将左图像特征点坐标在一定像素范围的窗口上搜索右图像上相应匹配点的坐标, 该窗口可以如下公式(2)的形式表示:

$$\begin{cases} |x_i^m - x_j^m| < \Delta\sigma_1^m \\ |y_i^m - y_j^m| < \Delta\sigma_2^m \end{cases} \quad (i, j = 1, 2, 3, \dots) \quad (2)$$

其中, 在第 m 帧的时刻, 左图像上像素点用 (x_i, y_i) 表示, 右图像上像素点用 (x_j, y_j) 表示, 若左图像上的像素坐标与右图像上的像素坐标能在给定的阈值 $\sigma_1 \times \sigma_2$ 窗口内, 则分别给左右图像的匹配点坐标一个索引序号, 即每个匹配对将有唯一的标识; 最后, 通过绘制关键点方法匹配所有点状特征的圆心, 本文根据实验结果得出结论, 相比于传统的 SIFT 算法^[12]和 SURF 算法^[16]对较大的圆点特征提取率极低, 匹配效果不理想, 本文方法大大提高了匹配精度.

3 重建

3.1 空间上的重建

世界坐标系转换成图像坐标系经过了刚体变换和透视投影变换, 而三维重建是该过程的一个逆向操作, 因此可以通过图像坐标和投影矩阵反向求解出匹配点的三维空间坐标. 由于本文是建立在基于双目立体视觉的基础上, 所以只需要考虑两台摄像机的相对位置关系, 如果把左相机坐标系当作世界坐标系, 则投影矩阵可由公式(3)表示:

$$\begin{aligned} P_1 &= K_1[I|0] \\ P_2 &= K_2[R|T] \end{aligned} \quad (3)$$

其中 K 为内参数矩阵, R 为旋转矩阵, T 为平移矩阵. 左投影矩阵 P_1 可以进一步分解为 P_{11}, P_{12}, P_{13} , 如公式(4)所示的形式:

$$S \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} \\ P_{12} \\ P_{13} \end{bmatrix} \tilde{X}_i \Rightarrow \begin{bmatrix} P_{13}u_i - P_{11} \\ P_{13}v_i - P_{12} \end{bmatrix} \tilde{X}_i = 0 \quad (4)$$

而左图像的任意匹配点 i 齐次坐标为 $(u_i, v_i, 1)$, \tilde{X}_i 为该点的三维空间齐次坐标, S 为常量因子. 同理, 如公式(5)右投影矩阵 P_2 如上分解, 两式联立.

$$\begin{bmatrix} P_{13}u_i - P_{11} \\ P_{13}v_i - P_{12} \\ P_{23}u'_i - P_{21} \\ P_{23}v'_i - P_{22} \end{bmatrix} \tilde{X}_i = 0 \quad (5)$$

由公式(5)可知, 四个方程求解三个未知数, 通过最小二乘法原理求解出 \tilde{X}_i 的值, 该公式可以简化为 $A\tilde{X}_i = 0$, 如公式(6)的形式:

$$A = USV^T \quad (6)$$

再对 A 进行奇异值分解^[17], 则 \tilde{X}_i 等于矩阵 V 的最后一列.

3.2 时间上的连续

对点状特征柔性物体的运动轨迹进行时间上的重建, 步骤如下:

第一步, 利用已获取的当前帧图像的二维坐标, 在一定像素范围的小窗口内搜寻下一帧相应的坐标, 如公式 (7) 所示:

$$\begin{cases} x_i^m - \Delta\sigma < x_i^{m+1} < x_i^m + \Delta\sigma \\ y_i^m - \Delta\sigma < y_i^{m+1} < y_i^m + \Delta\sigma \end{cases} \quad (i = 1, 2, 3, \dots) \quad (7)$$

第二步, 同理, 在右图像上找到目标点连续帧内的坐标.

第三步, 将左右图像时间上连续的坐标进行匹配重建, 就可以获得该点在时间上连续的空间坐标.

按 PAL 标准 1 s 分为 25 帧, 所以两帧之间的时间就是 Δt , 两帧之间运动的距离为 ΔS , 如公式 (8) 所示:

$$\begin{cases} \Delta s = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \\ \Delta t = \frac{1}{24} s \end{cases} \quad (8)$$

因此可以得到目标点在两帧之间运动的速率 Δv ,

由公式 (9) 所示. 同理也可以通过速率求得加速度 Δa 如公式 (10).

$$\Delta v = \frac{\Delta s}{\Delta t} \quad (9)$$

$$\Delta a = \frac{\Delta v_{i+1}^2 - \Delta v_i^2}{2\Delta s} \quad (10)$$

4 实验结果分析及应用

本文进行了三组实验分别针对: 布满点状特征的白纸、布满点状特征的服装以及画有点状特征的人脸, 采用两台型号为 SONY FDR-AX30 的高速摄像机拍摄运动视频, 再经过立体标定, 立体校正, 特征点匹配, 三维重建等过程, 最终实现动态柔性物体的三维运动捕获, 并用 Meshlab 重建三维空间点效果图.

4.1 实验结果分析

本实验的实验对象分别如图 1(a) 点状特征的白纸、图 1(d) 点状特征的服装、图 1(g) 点状特征的人脸所示, 图 1(b)、图 1(e) 和图 1(h) 是重建结果的正视图, 图 1(c)、图 1(f) 和图 1(i) 是相应的侧视图. 经分析图片可以清晰地看到柔性物体表面的点状特征在某一时刻的三维空间位置被精准地重建出来, 因此说明本文对于点状柔性物体的三维运动捕获方法效果比较理想.

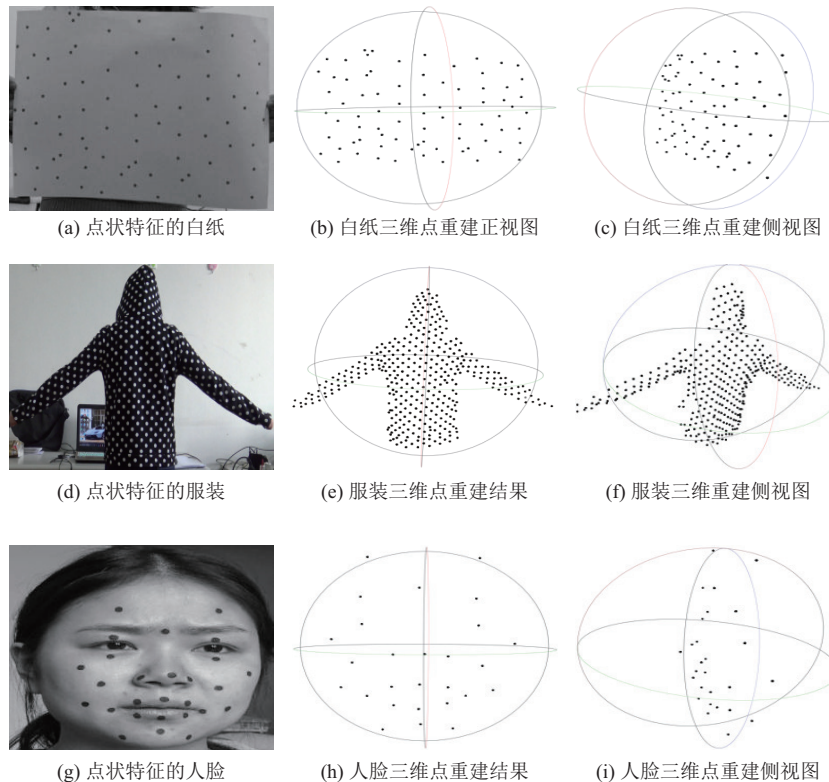


图 1 三维重建效果图

本实验将 DOG 方法与最小二乘法拟合椭圆的方法^[18]作比较. 图 2 为 DOG 角点检测图, 图 3 为椭圆拟合的结果, 圆中亮点为特征点圆心位置; 图 4 为图 3 中右下角某点状特征放大图. 首先, 因为 DOG 获得的是与特征点相同的圆, 而椭圆拟合获取的是与特征点类似的椭圆, 所以 DOG 方法获得的圆心理论上更接近精准圆心位置. 其次, 图 4 中灰色的点为 DOG 方法获得的圆心位置, 黑色的点为椭圆拟合方法获得的圆心位置, 图中黑色的点偏离了灰色点, 并且不在圆内的灰色点则为椭圆拟合没有检测到的圆心位置, 此结果表明 DOG 获取的特征点位置更加精准.

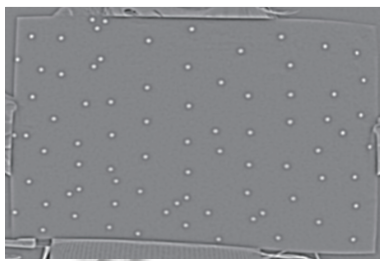


图 2 DOG 角点检测

本文通过 DOG 方法检测特征点, 再对检测到的特征点运用搜索策略进行特征匹配. 如图 5(a) 是在经过 DOG 检测特征点后, 利用搜索策略进行特征匹配的效果图. 本文不仅给出了搜索策略的匹配效果图, 还对 sift 算法和 surf 算法进行了实验, 图 5(b) 和图 5(c) 分别是两方法的匹配效果图, 经过对实验结果的比对, 发现本文的匹配方法更加理想和精准, 说明了本文方法

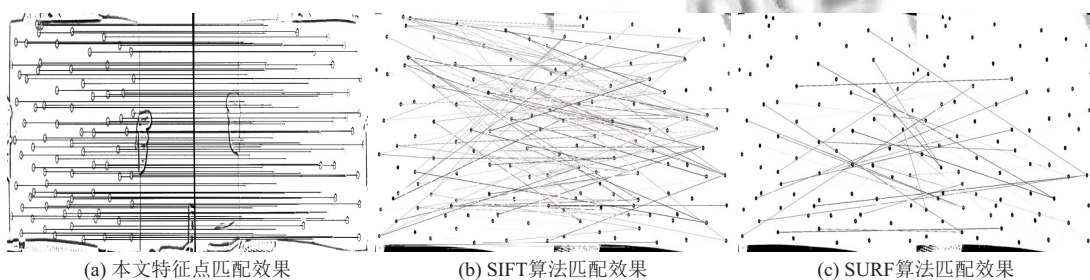


图 5 特征点检测与匹配效果图

假设某点的三维空间坐标为 (X, Y, Z) , 如表 1 所示是图 6(a) 中 1 号点的前 5 帧空间坐标. 表 2 是对应的速度和加速度数据. 从表格中的数据可以看出该点前五帧的运动趋势是先加速再减速, 即可以预测该点处于在前后抖动的运动趋势.

4.2 人脸表情应用

本实验对画有点状特征的人脸进行三维运动捕获,

对于较大的点状特征提取率高且匹配更加精准.

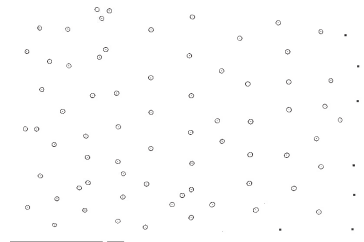


图 3 椭圆拟合获取特征点



图 4 圆心坐标比对效果图

如图 6 所示, 是白纸上 9 个点在连续 15 帧内的运动轨迹. 图 6(a) 是在柔性物体上选取 9 个目标点并进行标注, 图 6(b) 是这 9 个点在 15 帧内的运动轨迹正视图, 图 6(c) 是相应的俯视图. 从图 6(b) 和图 6(c) 可看出, 图 7 的目标点是在进行前后运动的, 个别点运动的幅度大, 个别运动的幅度小. 总体上是在进行前后运动的趋势, 实现了目标点在时间上运动轨迹的重建.

重建结果如图 1(h), 1(i) 所示. 本实验通过计算点的速度、加速度分析点的运动趋势.

如图 7 所示, 本文用黑色箭头表示第一帧到最后一帧的运动趋势. 图 7(a) 是人脸做出微笑表情时点的运动趋势, 如图所示微笑的时候嘴唇和脸颊周围的点呈上升趋势, 而鼻子额头上的点呈下降趋势. 图 7(b) 是人脸做出大笑表情时点的运动趋势, 如图所示上嘴

唇、脸颊、鼻子、额头上的点都呈上升趋势,而只有下嘴唇和下巴上的点呈下降趋势.图7(c)是人脸做出惊讶的表情,如图所示唇部与大笑表情几乎一样的运动趋势,但是有一个很大的区别就是两嘴角却是向内

收敛的运动趋势.图7(d)是人脸做出愤怒的表情,如图所示与其他表情不一样,该图中脸部的点都在做呈向上运动的趋势,上下嘴唇中间点,先向下运动再向上运动回到原点.

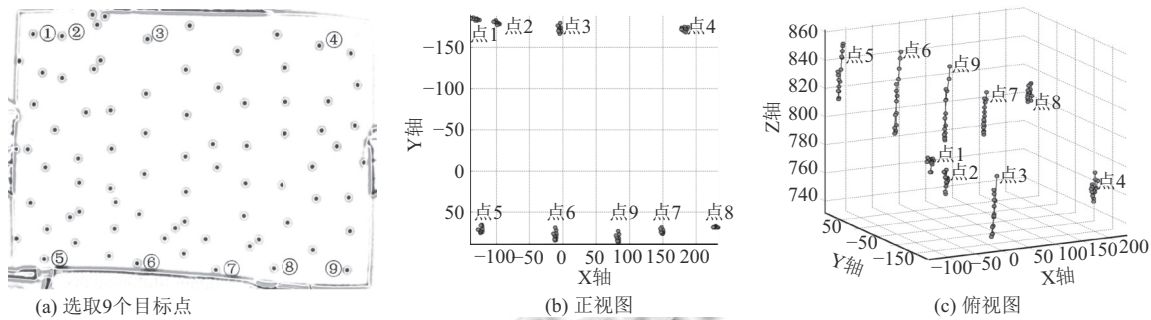


图6 目标点连续15帧空间运动轨迹

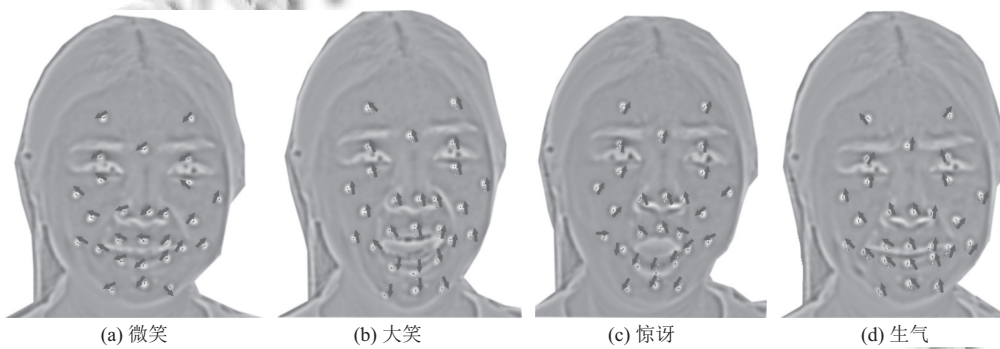


图7 人脸表情上点运动趋势图

表1 1号点的前五帧坐标(单位:mm)

No.	X	Y	Z
1	-66.1750	-205.8712	786.3466
2	-64.0884	-201.5432	773.2901
3	-62.1495	-199.4280	770.3859
4	-61.0566	-198.5511	768.9124
5	-61.7321	-197.9965	768.1859

表2 1号点前五帧的速度和加速度

No.	Velocity (mm/s)	Acceleratio (mm/s ²)
1	333.898 179	8013.556 305
2	97.983 4688	-5661.953 054
3	-516.542 894	-1180.419 97
4	27.276 682	-516.542 894

在图7人脸数据中选取上嘴唇中间点和下嘴唇中间点,分别用up和down表示.表3记录了这两个点在图7所示四种表情的首、v_end、a_start、a_end分别代表特征点首帧速度,尾帧速度,首帧加速度,尾帧加速度.如表3所示,图7(b)的速度值比图7(a)大,加速值比其小,说明人在大笑时唇部肌肉运动的速度比微笑时快,但速度的变化比微笑时慢,从而表明了人大笑时面部表情比微笑时保持得更久;图7(c)的速度值比图7(b)大,加速度值比其大,说明人在惊讶时唇部肌肉运动速度比大笑时快,而且速度变化得快,从而表明了这个表情在一瞬间产生;图7(d)的速度值均偏小,说明人在生气时,唇部肌肉运动的速度慢,a_up的值较大表明上嘴唇上方的肌肉运动速度变化较快,a_down的值为负数说明下嘴唇下方的肌肉从开始时运动很快又立即变慢.

表3 点 up 和 down 速度加速度

Fig_point	v_start (mm)	v_end (mm)	a_start (mm/s ²)	a_end (mm/s ²)
Fig. 7(a)_up	98.954 535 014 823 85	87.361 318 671 366 22	201.041 469 723 985	-278.237 192 242 9830
Fig. 7(b)_up	144.499 134 946	273.642 102 0238	-11.979 238 6921	92.262 824 3649
Fig. 7(c)_up	139.942 845 476 287 23	182.778 554 540 7338	1287.975 155 039 879	2309.887774310354
Fig. 7(d)_up	89.442 719 099 992 73	129.243 955 371 2281	1821.471 932 256 987	164.819 689 080 030 12
Fig. 7(a)_down	86.533 230 611 135 74	120.598 507 453 4506	1281.830 756 763 63	1454.364 178 882 816
Fig. 7(b)_down	288	72	368.548 028 813 32	11.520 000 000 00
Fig. 7(c)_down	682.630 207 359 7388	83.522 452 071 282 66	4153.595 069 334 515	1397.381 538 958 4688
Fig. 7(d)_down	87.361 318 671 366 22	108.295 890 965 4473	643.987 577 519 9395	-2124.449 823 284 0006

5 结论

本文提出了一种实用的基于点状特征的柔性物体运动捕获方法. 实验结果显示, DOG 方法比椭圆拟合法获取特征点的位置更加精准; 采用搜索策略实现匹配, 相比传统的 SIFT 算法, SURF 算法大大提高了匹配精度. 此外, 在三维重建的基础上进行扩充, 增加时间序列上的匹配, 从而实现动态柔性物体的三维运动捕获, 其次通过分析速度、加速度等数据来应用到人脸表情分析, 为后期用动态数据驱动静态图像打下基础.

参考文献

- 1 向泽锐, 支锦亦, 徐伯初, 等. 运动捕捉技术及其应用研究综述. 计算机应用研究, 2013, 30(8): 2241-2245.
- 2 李晓丹, 肖明, 曾莉. 人体动作捕捉技术综述以及一种新的动作捕捉方案陈述. 中国西部科技, 2011, 10(15): 35-37. [doi: 10.3969/j.issn.1671-6396.2011.15.018]
- 3 梁丰, 张志利, 李向阳, 等. 基于光学运动捕捉数据的虚拟人下肢运动控制方法. 系统仿真学报, 2015, 27(2): 327-335.
- 4 Luo Q. Research on motion capture instruments in sports. Applied Mechanics and Materials, 2013, 346: 151-155. [doi: 10.4028/www.scientific.net/AMM.346]
- 5 罗忠祥, 庄越挺, 潘云鹤, 等. 基于视频的运动捕获. 中国图像图形学报, 2002, 7(8): 752-758.
- 6 庄越挺, 刘小明, 潘云鹤, 等. 运动图像序列的人体三维运动骨架重建. 计算机辅助设计与图形学学报, 2000, 12(4): 245-250.
- 7 Ummenhofer B, Brox T. Point-based 3D reconstruction of thin objects. Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 969-976.
- 8 Boubekour T, Reuter P, Schlick C. Visualization of point-based surfaces with locally reconstructed subdivision surfaces. Proceedings of the 2005 International Conference on Shape Modeling and Applications. Cambridge, MA, USA. 2005. 23-32.
- 9 Bradley D, Popa T, Sheffer A, et al. Markerless garment capture. ACM Transactions on Graphics (TOG), 2008, 27(3): 99.
- 10 Seitz SM, Curless B, Diebel J, et al. A comparison and evaluation of multi-view stereo reconstruction algorithms. Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA. 2006. 519-528.
- 11 Peng KJ, Chen X, Zhou DX, et al. 3D reconstruction based on SIFT and Harris feature points. Proceedings of the 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO). Guilin, China. 2009. 960-964.
- 12 Lindeberg T. Scale invariant feature transform. Scholarpedia, 2012, 7(5): 10491. [doi: 10.4249/scholarpedia.10491]
- 13 王青松, 赵西安, 吕京国, 等. 基于高斯差分的改进 Harris 特征点提取算法. 测绘科学, 2014, 39(4): 119-122, 134.
- 14 Zhang Z. A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334. [doi: 10.1109/34.888718]
- 15 陈蓉, 罗晓曙. 一种基于双目立体视觉的图像校正算法设计. 测控技术, 2016, 35(5): 16-19, 24.
- 16 Wang XL, Yang L, Wang LR, et al. Characteristic point match algorithm based on the SURF in binocular stereo vision. Proceedings of the 5th International Conference on Intelligent Networks and Intelligent Systems (ICINIS). Tianjin, China. 2012. 302-305.
- 17 Cui YM, Wang H, Zhang J. 3D restructuring and splicing technology of shape space frame. Applied Mechanics and Materials, 2014, 644-650: 1438-1442. [doi: 10.4028/www.scientific.net/AMM.644-650]
- 18 马向南, 李航, 刘丽丽, 等. 最小二乘改进算法及其在椭圆拟合中的应用. 河南科技大学学报: 自然科学版, 2014, 35(3): 18-21.