

面向 AR 沙盘异地协同标绘的动作重构技术

李维浩¹, 姚世明¹, 李蔚清², 苏智勇¹

(1. 南京理工大学 自动化学院, 江苏, 南京 210094; 2. 南京理工大学 计算机科学与工程学院, 江苏, 南京 210094)

摘要: 为真实再现增强现实环境下全息沙盘异地协同标绘的交互过程, 提出了一种基于视频的异地协同标绘动作重构技术. 使用单目 RGB 相机采集作业人员的实时视频信息, 基于 YOLO 实现多人检测, 并确定作业人员在三维空间中的相对位置关系; 基于改进的 ResNet 网络对检测到的每个作业人员进行关节的三维位置预测; 然后将估算出的本地作业人员的关节位置实时传送到异地, 经过逆向运动学及低通滤波器的优化驱动虚拟化身的运动. 实验结果表明, 能够准确地再现异地作业人员的标绘动作, 满足实际应用需求.

关键词: 协同标绘; 多人检测; ResNet; 三维姿态估计; 动作重构

中图分类号: TP391 **文献标志码:** A **文章编号:** 1001-0645(2019)12-1298-07

DOI: 10.15918/j.tbit.1001-0645.2018.431

A Motion Reconstruction Technology for Distributed Collaborative Plotting of AR Sand Table

LI Wei-hao¹, YAO Shi-ming¹, LI Wei-qing², SU Zhi-yong¹

(1. School of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China;

2. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China)

Abstract: In order to realistically reproduce the interactive process of collaborative plotting of holographic electronic sand table under the augmented reality environment, a video-based distributed collaborative plotting reconstruction method was proposed. Firstly, the operators were tracked based on YOLO using a single RGB camera simultaneously to estimate the relative position of those people in 3D space. And an improved ResNet network was used to predict 3D joint positions for every operator. Then the estimated joint positions of the local operators were sent to other place to drive the avatar's movement in real time through low-pass filter and the method of inverse kinematics. The experiment results show that the proposed method can reproduce the operator's motion and meet the application requirements.

Key words: collaborative plotting; multi-person detection; ResNet; 3D pose estimation; motion reconstruction

随着增强现实(augmented reality, AR)技术的发展,使得基于 AR 电子沙盘的异地研讨成为可能. 相较于传统沙盘,基于 AR 电子沙盘的异地研讨能够克服指挥作业人员不在同一地域的问题,同时能够获得比传统沙盘研讨更形象直观的战场态势展示

效果. 为了使得在基于 AR 电子沙盘的研讨中,能够让研讨过程更加的真实、直观,一个重要问题就是解决多人异地的协同标绘.

为了实现 AR 环境下的异地协同标绘,需要解决多人标绘动作的实时动作重构问题. 动作重构的基

收稿日期: 2018-11-02

基金项目: “十三五”装备预研资助项目(61409230104, 315100104, 41401010203); 中央高校基本科研业务费专项资金资助项目(30918012203); 上海航天科技创新基金资助项目(SAST2019009)

作者简介: 李维浩(1995—),男,硕士生,E-mail: 734234272@qq.com.

通信作者: 苏智勇(1981—),男,副教授,E-mail: su@njust.edu.cn.

础是姿态估计,现在主流的姿态估计方法有基于深度相机和基于深度学习的方法. 基于深度相机例如 Kinect 的方法利用了其能够直接获得深度信息的能力,直接预测出人体的姿态,但是其可视范围较小且精度稍差;在深度学习三维姿态估计领域主要有 Bogo 等^[1]提出的 SMPL 方法,使用 SMPL 模型创建 3D 网格形状,从而通过 2D 姿态估计理解 3D 形态;Zhou 等^[2]提出的一种基于 2D 标注和人类骨架先验知识的弱监督学习的方法;Tekin 等^[3]提出的使用过完备的自动编码的深度神经网络来学习人体的高维姿态表示的方法,以及 Mehta 等^[4]提出的 VNect 方法,通过附加位置图(location-map)解决关节点 2D 到 3D 的映射问题,从而恢复出检测对象的三维姿态. 以上基于神经网络的三维姿态恢复方法很好地解决了捕捉设备昂贵、人物运动受限等问题,但是都只实现了单人的姿态估计而没有解决对多人姿态的实时重构问题. 为了解决异地协同标绘过程中多人动作的实时重构问题,本文基于文献[4]的 VNect 方法提出了面向 AR 沙盘异地协同标绘的动作重构技术.

1 系统框架

系统分为 3 个部分,如图 1 所示:第 1 部分为多人检测跟踪,对采集的视频数据进行预处理,包括多人检测跟踪以及对初始帧图像的深度图估计,基于此估算出视频中多人的相对位置关系;第 2 部分为三维姿态估计,对第 1 部分检测到的每个个体通过卷积神经网络做姿态回归,得到二维热图(heatmap)以及 x, y, z 方向的位置图,并根据这些输出合成预测的三维关节点位置;第 3 部分驱动虚拟化身,将多人之间的相对位置关系以及每个人的骨骼点位

置信息传输到异地,通过低通滤波器的优化以及基于运动学的骨架拟合驱动虚拟化身的运动.

2 多人检测跟踪

2.1 基于 YOLO 的多人检测

多人检测跟踪作为多目标检测跟踪(multiple object tracking, MOT)在行人领域内的一部分,近年来随着机器学习的发展,基于神经网络的 MOT 方法越来越多,这些方法在精度以及效率上都要优于传统的方法. YOLO^[5]就是其中一种方法,其核心思想是将目标检测问题转换成一个回归问题,将整张图像作为输入,直接在输出层回归出被检测物体的位置和它所属的类别. 将输入图像划分为 $S \times S$ 的栅格,每个栅格都预测出 B 个矩形框(bounding boxes),这里的栅格划分是逻辑上的划分,其预测向量的长度 L 为

$$L = SS(5B + C), \quad (1)$$

式中: S 为划分的格子数目,一般 S 取 7; B 为每个格子预测的边框数,一般 B 取 2; C 为分类的类别数. 不难得出这个预测向量包括:① $S \times S \times C$ 个类别信息,表示每一个格子可能属于哪一类;② $S \times S \times B$ 个置信度,表示每一个格子的 B 个框的置信度(confidence),一般只保留置信度为 0.5 以上的边框;③ $S \times S \times B \times 4$ 个位置信息,4 个位置信息分别是 x, y, z, h ,其中 x, y 分别为预测边框中心与栅格边界的相对值, w, h 分别为预测边框的宽度和高度相对于整张图像的比例. 该方法的效率较高,且能很好地避免背景错位,但在精度上相对于其他优秀的检测算法要差一些,且对小物体的检测效果不太理想.

在本应用当中,由于被检测人员均立于摄像机可视范围之内,且周围环境并不复杂,人物突出,所以在精度方面的误差可以忽略,且能保证较高的实时性.

2.2 基于深度图的相对位置估计

多人检测给出了不同人在二维图像中的位置关系,但却无法获得他们在三维空间中的相对位置关系. 本文通过结合多人检测结果和估计深度图的方法推算视频中多人的空间位置关系,生成不同作业人员在电子沙盘研讨中所处的方位信息. 目前,本文只考虑了在人员位置与沙盘位置关系事先约定好的情况下,着重解决人员之间的空间位置关系问题.

应用中使用相同的虚拟化身模型,所以在估算位置的时候不用考虑 y 轴上的差值,只用考虑 $x-z$ 平面上人物的位置站姿. 且对单人的姿态估计的关节点坐标是相对于根节点而言的,故对多人

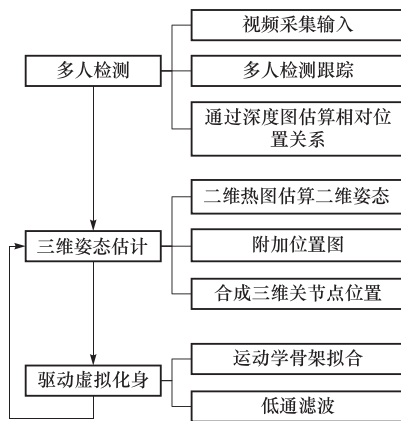


图 1 系统框架图

Fig. 1 Flow chart of system framework

之间的相对三维位置估计问题即转化为对多人根节点在 $x-z$ 平面上的二维位置估计问题. 估计方法如图 2 所示, 首先通过边框的中心点确定指挥

作业人员在像素空间内 x 轴的位置, 以图像最左侧为 x 轴 0 坐标, 这样可以得到检测对象在 x 轴上的位置关系.

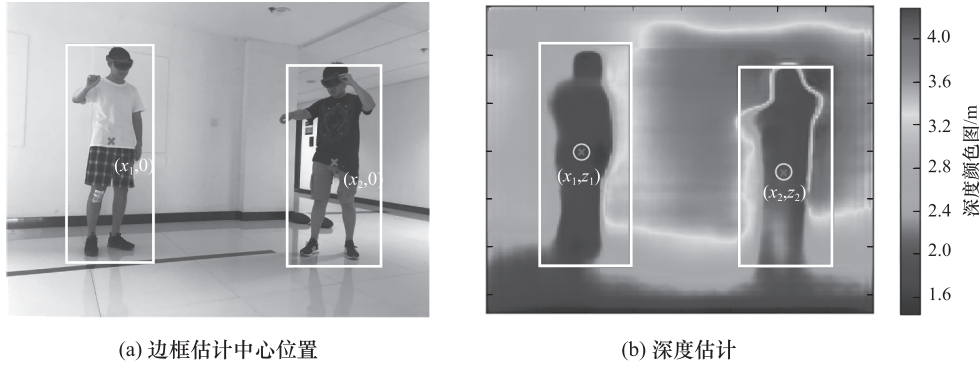


图 2 结合边框以及深度图估计相对位置
Fig. 2 Estimate relative position with bounding box and depth map

然后对图像中的人物进行深度估算^[6]. 考虑到只需对根节点的位置进行估算, 故取边框中心位置作为估算点. 与 x 轴位置估算不同的是, 深度估算本身存在一定的误差, 同一目标的不同位置的深度信息也不一样, 为尽可能低地消除这种误差, 选择中心点附近的 n 个点, 选择的范围半径为边框长度的 $1/5$. 对其深度值求平均得到目标的深度坐标为

$$z_i = \sum_{j=1}^n d_{ij} / n, \quad (2)$$

式中: z_i 为第 i 个用户的深度坐标; d_{ij} 为第 i 个用户选取第 j 个点的深度坐标. 本文中 $n=9$. 这样就获得了多指挥作业人员之间的相对空间关系, 结合先验的与沙盘的空间位置关系将人员动态的加载到应用当中.

3 基于 Vnect 的三维姿态估计

3.1 网络结构

使用已经在 MPII^[7]、MPI-INF-3DHP^[8] 和 Human3.6 m^[9] 数据集上预训练好的模型 Vnect 进行图像的特征值提取. 但是该模型在 Human3.6 m 的数据集上测试的精度并不理想, 故针对本文的样本数据重新训练模型.

该模型基于 ResNet50 (deep residual learning network) 改进而来^[10]. 深度残差网络 ResNet 的出现为训练更深的网络提供了方法, 其主要思想是跨层连接. 让这些层不是通过每一个堆叠层直接适合一个所需的底层映射, 而是适合残差映射. 改进后的网络结构如图 3 所示. 加入了中继监督, 通

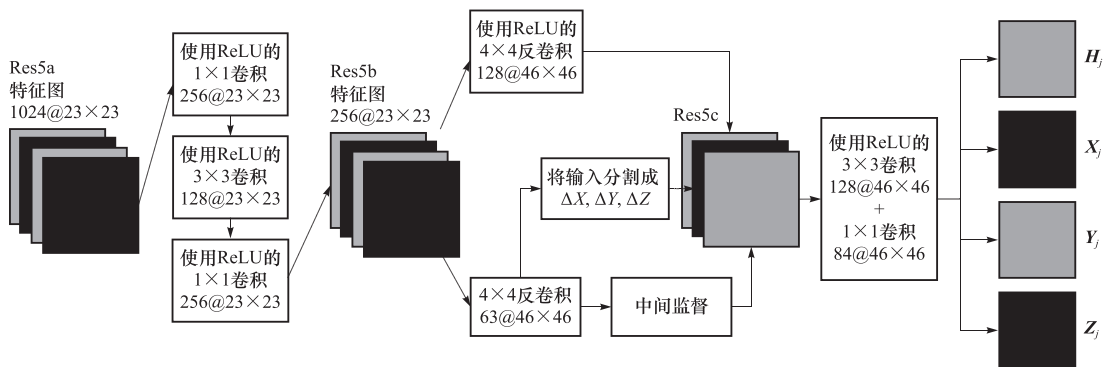


图 3 网络结构(从 res5a 开始)
Fig. 3 Network structure (start from res5a)

过计算骨骼的长度来指导最终的预测特征.

如式(3)所示, 中继监督环节通过两个关节之间的位置差值计算骨骼的长度.

$$L_j = \sqrt{\Delta X_j \odot \Delta X_j + \Delta Y_j \odot \Delta Y_j + \Delta Z_j \odot \Delta Z_j}, \quad (3)$$

式中: L_j 为关节 j 距父节点的骨骼长度图; ΔX_j ,

$\Delta Y_j, \Delta Z_j$ 为关节点 j 相对于其父节点在 x, y, z 方向上的距离位置图; \odot 为哈达玛积.

输入为 368×368 的图片, 经过卷积运算到 res5a 时是 1 024 张 23×23 的特征图, 其中 1 024 对

应卷积核的个数. 最终的输出是每个预测的关节点 j 对应的二维热图 H_j 以及附加的位置图 (location-maps) X_j, Y_j, Z_j . H_j 以及 X_j 的输出如图 4 所示, Y_j, Z_j 同理, 其中 j 为 21.

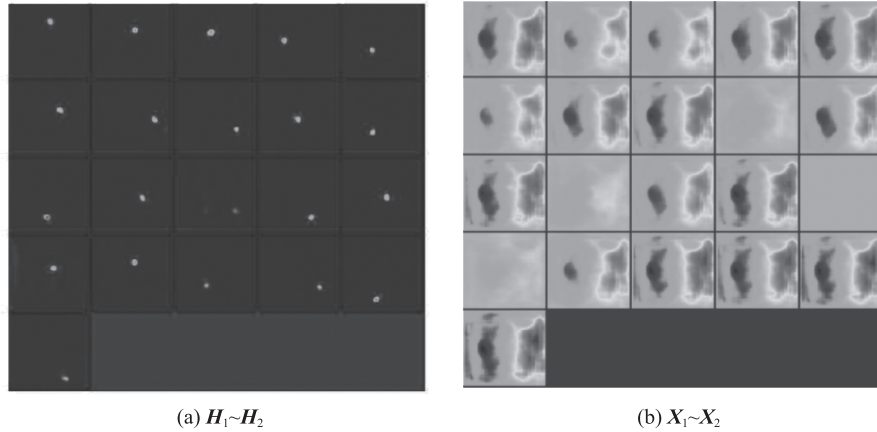


图 4 网络部分输出结果

Fig. 4 Part outputs of network

3.2 关节点位置预测

由图 4 可知, 输出是 21 个关节点的特征图, 其中包括每个关节点的二维热图以及 3 个不同方向的位置图. 位置图是被检测对象在某个方向上的位置分布. 通过二维热图可以快速恢复出二维姿态, 但对于三维空间而言, 由于多了纵向的维度, 如果单用二维的热图难以恢复出深度信息. 故通过附加的 3 个不同方向的位置图, 结合热图一起预测出关节点的三维位置.

先从输出的热图中获得关节点的图像坐标, 接着在该关节点的其他位置图中同样的位置标记出这个点, 得到这个点的坐标数值, 结合 3 张位置图, 得到此关节点相对于根节点的坐标 x_j, y_j, z_j . 选取其中一个关节点的预测过程如图 5 所示.

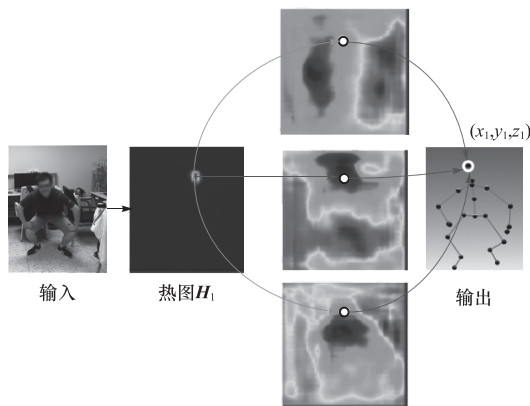


图 5 头部关节点的预测过程

Fig. 5 Prediction process of head joint point

对每个方向的位置来说, 损失函数如下 (GT 表示 ground truth)

$$\text{Loss}(x) = \| \mathbf{H}_j^{\text{GT}} \odot (\mathbf{X}_j - \mathbf{X}_j^{\text{GT}}) \|_2, \quad (4)$$

$$\text{Loss}(y) = \| \mathbf{H}_j^{\text{GT}} \odot (\mathbf{Y}_j - \mathbf{Y}_j^{\text{GT}}) \|_2, \quad (5)$$

$$\text{Loss}(z) = \| \mathbf{H}_j^{\text{GT}} \odot (\mathbf{Z}_j - \mathbf{Z}_j^{\text{GT}}) \|_2, \quad (6)$$

式中: \mathbf{H}_j^{GT} 为对于关节点 j 的热度图的标准; $\mathbf{X}_j^{\text{GT}}, \mathbf{Y}_j^{\text{GT}}, \mathbf{Z}_j^{\text{GT}}$ 分别为位置图 $\mathbf{X}_j, \mathbf{Y}_j, \mathbf{Z}_j$ 的标准值. 这些损失函数强化了二维热图与位置图之间的联系.

4 虚拟化身驱动

得到每帧图像中人物的关节点位置之后驱动虚拟化身的运动. 由于人的运动是高度非线性的, 直接将坐标赋值给虚拟化身会导致两个问题: 一是人物运动过程僵硬的问题; 二是人物运动在时域上当运动过快导致的动作滞后以及运动过慢导致的抖动问题. 故需要对虚拟化身的运动进行优化.

使用 FABRIK (forwards and backwards reaching inverse kinematics)^[11] 求解逆向运动学问题并在每一次的姿态预测之后, 使用自适应的一阶低通滤波器 (the 1 € filter)^[12] 进行扰动消除. 滤波器根据信号速度的变化自适应调整截止频率. 在被跟踪对象低速运动时, 低截止频率会通过降低抖动使信号稳定, 随着目标运动速度的增加, 截止频率会逐渐提高以减少滞后. 一阶低通滤波器的公式为

$$\hat{X}_i = \alpha X_i + (1 - \alpha) \hat{X}_{i-1}, \quad (7)$$

式中: X_i 为在时间 i 时的采样值; \hat{X}_i 为在时间 i 时的滤波输出值; α 为 $[0, 1]$ 上的平滑因子. 为了自适应输入信号的波动, 将 α 写成关于采样周期 T_c 和时间常数 τ 的函数, 滤波器的表示如下

$$\alpha = \frac{1}{1 + \tau/T_c}, \quad (8)$$

$$\tau = \frac{1}{2\pi f_c}, \quad (9)$$

$$\hat{X}_i = \frac{1}{1 + \tau/T_c} \left(X_i + \frac{\tau}{T_c} \hat{X}_{i-1} \right), \quad (10)$$

式中 f_c 为截止频率. 由于采样周期 T_c 可由时间戳自动算得, 故变量只有 f_c , 将 f_c 写成关于信号运动速度 \dot{X}_i 的函数, 为

$$f_c = f_{\min} + \beta |\dot{X}_i|. \quad (11)$$

由式(11)可知该自适应低通滤波器只与 2 个变量 f_{\min} 和 β 有关, 本文中取 $f_{\min} = 1.0$, $\beta = 0.008$.

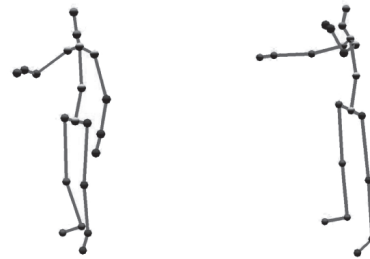
5 仿真实验

本文对全息电子沙盘协同标绘过程中常用的几个动作进行了仿真分析. 在 Ubuntu 系统下, 基于 Caffe 搭建了网络结构, 使用 Unity3D 创建了一个简易的可供多人协同操作的电子沙盘. 首先 2 个指挥作业人员头戴 HoloLens 站在室外操作电子沙盘, 使用一个普通 RGB 摄像机获取作业人员的视频信息并送往服务器处理, 将输出的三维姿态以及相对位置通过 Socket 传送给后方异地的指挥作业人员, 在后方人员 HoloLens 中显示身处异地的 2 个人的姿态.

在 6 核 Xeon 处理器和单 GTX1080Ti 的条件下, 目标检测分割时间约为 5~7 ms, 卷积神经网络对单人的处理时间约为 20 ms, 数据传输和滤波处理时间约为 6 ms, 应用中的运动学优化处理时间约为 7~9 ms. 仿真结果如图 6 和图 7 所示.



(a) 二维关节点预测效果



(b) 三维关节点预测效果



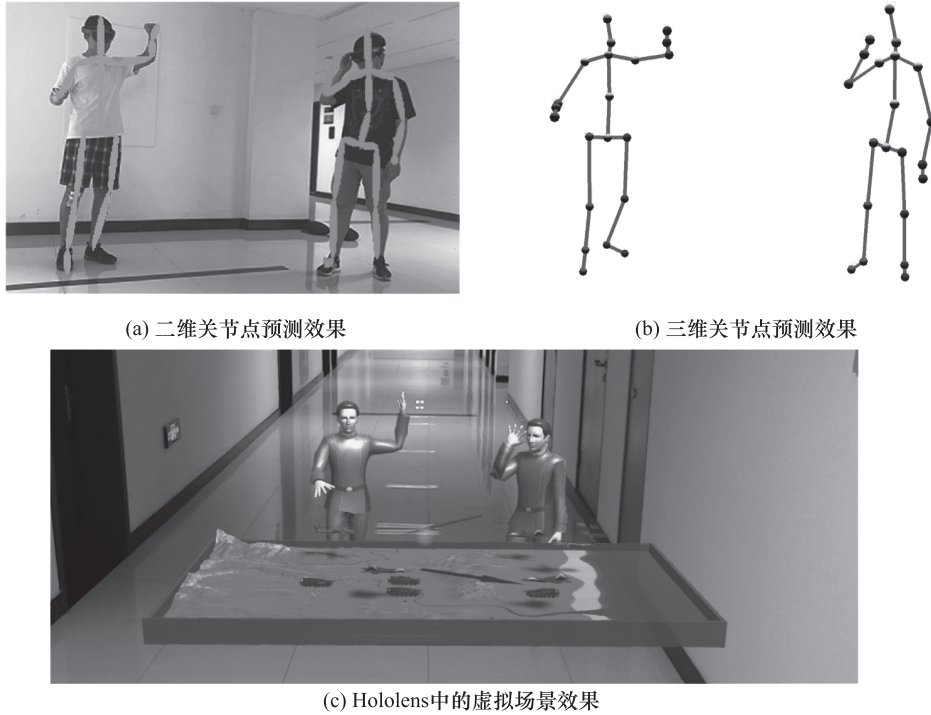
(c) HoloLens中的虚拟场景效果

图 6 协同标绘中拖拽及选取操作

Fig. 6 Drag and select operations in collaborative plotting

图 6 显示了协同标绘过程中的拖拽以及选取的动作, 图 7 显示了协同标绘过程中的放大缩小的操作. 在 AR 仿真过程中, 由于没有对人手进行检测以及对人脸的识别, 故人物的手掌无法变化, 且头部无

法转向. 图 8 给出了在实验室条件下对人体部分关节点在 3D 空间中的检测正确率, 此准确率的标准是当检测到的关节点的位置距离真实位置的误差不超过 150 mm, 即认为检测准确.



(a) 二维关节点预测效果

(b) 三维关节点预测效果

(c) HoloLens中的虚拟场景效果

图 7 协同标绘中放大及缩小操作

Fig. 7 Zoom in and out operations in collaborative plotting

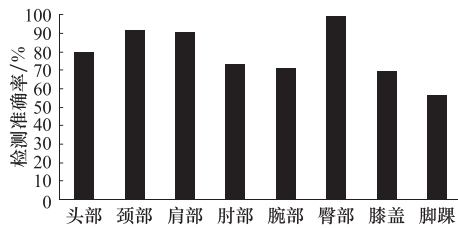


图 8 各个关节点的检测准确率

Fig. 8 Detection accuracy of joints

由图 8 分析可知,臀部关节,即检测中的根节点的准确率最高,这也是方法的基础,因为其他节点的位置都是相对于根节点而言的. 而膝盖关节以及脚踝关节的检测准确率最低,偏差最大. 整体而言,该方法的平均准确率达到 75.4%,满足应用背景下的需求.

6 结 论

本文采用结合二维热度图和附加位置图的方法实现了对单人的三维姿态恢复,并提出了一种通过多人检测跟踪并基于深度图定位的方法,实现了面向 AR 沙盘协同标绘的多人动作实时重构. 实验结果表明本文提出的方法在协同标绘应用中能够较为准确地对用户的动作进行重构恢复,并且在重构效率上基本可以满足当前协同标绘应用的需求.

参考文献:

[1] Bogo F, Kanazawa A, Lassner C, et al. Keep it SMPL: automatic estimation of 3D human pose and shape from a single image[C]// Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 561 - 578.

[2] Zhou X, Huang Q, Sun X, et al. Towards 3d human pose estimation in the wild: a weakly-supervised approach[C]// Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017: 398 - 407.

[3] Tekin B, Rozantsev A, Lepetit V, et al. Direct prediction of 3D body poses from motion compensated sequences [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 991 - 1000.

[4] Mehta D, Sridhar S, Sotnychenko O, et al. Vnect: real-time 3d human pose estimation with a single RGB camera[J]. ACM Transactions on Graphics, 2017, 36 (4): 44 - 50.

[5] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016: 779 - 788.

(下转第 1320 页)

- 2011,40(6):830-837.
- [2] 黄磊,吴朝彦. B2B品牌导向对品牌绩效的影响机制研究:供应商资源投入的关键作用[J]. 管理评论,2017,29(9):181-192.
Huang Lei,Wu Chaoyan. The impact of B2B brand orientation on brand performance:the role of suppliers' resource commitment [J]. Management Review, 2017, 29(9):181-192. (in Chinese)
- [3] Adamson B,Dixon M,Toman N. The end of solution sales [J]. Harvard Business Review,2012,90(7-8):60-68.
- [4] Han S L,Sung H S. Industrial brand value and relationship performance in business markets-a general structural equation model [J]. Industrial Marketing Management,2008,37(7):807-818.
- [5] Herbst U,Merz M A. The industrial brand personality-scale:building strong business-to-business brands[J]. Industrial Marketing Management, 2011, 40(7): 1072-1081.
- [6] Hills S B,Sarin S. From market driven to market driving:an alternative paradigm for marketing in high technology industries [J]. Journal of Marketing Theory and Practice,2003,11(3):13-24.
- [7] Wise R,Zednickova J. The rise and rise of the B2B brand [J]. Journal of Business Strategy,2009,30(1):4-13.
- [8] Aaker D A,Jacobson R. The value relevance of brand attitude in high-technology markets[J]. Journal of Marketing Research,2001,38:485-493.
- [9] Lynch J,de Chernatony L. The power of emotion;brand communication in business-to-business markets[J]. Journal of Brand Management,2004,11(5):403-419.
- [10] 郑春东,邹孟. 网购消费者信息搜索模式研究[J]. 科技管理研究,2014,34(15):177-181.
Zheng Chundong,Zou Meng. Research on information searching mode of online shopping consumers[J]. Science and Technology Management Research, 2014, 34(15):177-181. (in Chinese)
- [11] Bradley S,Kim C,Kim J, et al. Toward an evolution strategy for the digital goods business[J]. Management Decision,2012,50(2):234-252.
- [12] Halligan B,Shah D. Inbound marketing[M]. [S. l.]:Wiley,Hoboken,2010.

(责任编辑:刘雨)

(上接第1303页)

- [6] Godard C,AodhaM O,Brostow G J. Unsupervised monocular depth estimation with left-right consistency[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2017,2(6):7-14.
- [7] Andriluka M,Pishchulin L,Gehler P, et al. 2D human pose estimation; new benchmark and state of the art analysis[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2014;3686-3693.
- [8] Mehta D,Rhodin H,Casas D, et al. Monocular 3D human pose estimation in the wild using improved CNN supervision[C]// Proceedings of International Conference on 3D Vision. New York:IEEE,2016;506-516.
- [9] Ionescu C,Papava D,Olaru V, et al. Human3.6m: large scale datasets and predictive methods for 3D human sensing in natural environments[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(7):1325-1339.
- [10] He K,Zhang X,Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2016;770-778.
- [11] Aristidou A,Lasenby J. FABRIK;a fast,iterative solver for the inverse kinematics problem[M]. New York:Academic Press,2011.
- [12] Casiez G,Roussel N,Vogel D. 1 filter;a simple speed-based low-pass filter for noisy input in interactive systems[C]// Proceedings of Sigchi Conference on Human Factors in Computing Systems. New York:ACM, 2012;2527-2530.

(责任编辑:李兵)

- 2011,40(6):830-837.
- [2] 黄磊,吴朝彦. B2B品牌导向对品牌绩效的影响机制研究:供应商资源投入的关键作用[J]. 管理评论,2017,29(9):181-192.
Huang Lei,Wu Chaoyan. The impact of B2B brand orientation on brand performance:the role of suppliers' resource commitment [J]. Management Review, 2017, 29(9):181-192. (in Chinese)
- [3] Adamson B,Dixon M,Toman N. The end of solution sales [J]. Harvard Business Review,2012,90(7-8):60-68.
- [4] Han S L,Sung H S. Industrial brand value and relationship performance in business markets-a general structural equation model [J]. Industrial Marketing Management,2008,37(7):807-818.
- [5] Herbst U,Merz M A. The industrial brand personality-scale:building strong business-to-business brands[J]. Industrial Marketing Management, 2011, 40(7): 1072-1081.
- [6] Hills S B,Sarin S. From market driven to market driving:an alternative paradigm for marketing in high technology industries [J]. Journal of Marketing Theory and Practice,2003,11(3):13-24.
- [7] Wise R,Zednickova J. The rise and rise of the B2B brand [J]. Journal of Business Strategy,2009,30(1):4-13.
- [8] Aaker D A,Jacobson R. The value relevance of brand attitude in high-technology markets[J]. Journal of Marketing Research,2001,38:485-493.
- [9] Lynch J,de Chernatony L. The power of emotion;brand communication in business-to-business markets[J]. Journal of Brand Management,2004,11(5):403-419.
- [10] 郑春东,邹孟. 网购消费者信息搜索模式研究[J]. 科技管理研究,2014,34(15):177-181.
Zheng Chundong,Zou Meng. Research on information searching mode of online shopping consumers[J]. Science and Technology Management Research, 2014, 34(15):177-181. (in Chinese)
- [11] Bradley S,Kim C,Kim J, et al. Toward an evolution strategy for the digital goods business[J]. Management Decision,2012,50(2):234-252.
- [12] Halligan B,Shah D. Inbound marketing[M]. [S. l.]:Wiley,Hoboken,2010.

(责任编辑:刘雨)

(上接第1303页)

- [6] Godard C,AodhaM O,Brostow G J. Unsupervised monocular depth estimation with left-right consistency[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2017,2(6):7-14.
- [7] Andriluka M,Pishchulin L,Gehler P, et al. 2D human pose estimation; new benchmark and state of the art analysis[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2014;3686-3693.
- [8] Mehta D,Rhodin H,Casas D, et al. Monocular 3D human pose estimation in the wild using improved CNN supervision[C]// Proceedings of International Conference on 3D Vision. New York:IEEE,2016;506-516.
- [9] Ionescu C,Papava D,Olaru V, et al. Human3.6m: large scale datasets and predictive methods for 3D human sensing in natural environments[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(7):1325-1339.
- [10] He K,Zhang X,Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2016;770-778.
- [11] Aristidou A,Lasenby J. FABRIK;a fast,iterative solver for the inverse kinematics problem[M]. New York:Academic Press,2011.
- [12] Casiez G,Roussel N,Vogel D. 1 filter;a simple speed-based low-pass filter for noisy input in interactive systems[C]// Proceedings of Sigchi Conference on Human Factors in Computing Systems. New York:ACM, 2012;2527-2530.

(责任编辑:李兵)