

4-12-2023

## Construction Technology of Hand Posture Dataset Based on Virtual Simulation Method

Jiaxin Chen

*Harbin Normal University, Harbin 150080, China;* 844528295@qq.com

Guohui Zhou

*Harbin Normal University, Harbin 150080, China;* zhouguohui@hrbnu.edu.cn

Jianbai Yang

*Harbin Normal University, Harbin 150080, 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.

---

# Construction Technology of Hand Posture Dataset Based on Virtual Simulation Method

## Abstract

**Abstract:** Hand posture is an important carrier of human-computer interaction, and the acquisition and recognition of posture information largely depends on the hand posture dataset. Existing datasets can be divided into two categories, real datasets and synthetic datasets. As real data is limited by equipment, environment, and other factors, the classification of hand posture is insufficient and the annotation is mixed with a lot of manual errors. The existing synthetic data can solve the data scale problem of real data, but the synthetic hand posture volume is limited and with some unreasonable kinematic postures of which the data form are only RGB images. *By studying the hand anatomical structure, a virtual 3D kinematic model of hand is created based on virtual simulation, and a hand posture simulation generator is built.* The hand posture dataset constructed by the generator can not only effectively solve the problems of the insufficient data volume of real data, low accuracy of annotation, and large workload of manual annotation, but also solve the problems of insufficient hand posture category and single data mode of synthetic data.

## Keywords

virtual simulation, hand posture dataset, kinematics modeling, hand posture estimation, human-computer interaction

## Recommended Citation

Jiaxin Chen, Guohui Zhou, Jianbai Yang. Construction Technology of Hand Posture Dataset Based on Virtual Simulation Method[J]. Journal of System Simulation, 2023, 35(4): 862-870.

# 基于虚拟仿真方法的手部姿态数据集构建技术

陈佳昕, 周国辉\*, 杨建柏

(哈尔滨师范大学, 黑龙江 哈尔滨 150080)

**摘要:** 手部姿态是人机交互的重要载体, 姿态信息的获取和识别很大程度上依赖于手部姿态数据集。现有的数据集分成 2 类: 真实数据集和合成数据集。真实数据受限于设备与环境等因素, 手部姿态类别不足且标注掺入大量人工误差。而目前存在的合成数据虽然可以解决真实数据的数据量问题, 但生成的手部姿态有限且存在一些不合理的运动学姿态, 获取的数据往往也只有 RGB 图像。通过研究手部的解剖学结构, 基于虚拟仿真方法, 创建了手部的虚拟三维运动学模型和一个手部姿态仿真生成器。利用该生成器构建的手部姿态数据集既可以有效解决真实数据数据量不足、标注准确度低、人工标注工作量大等问题, 也可以解决当前合成数据手部姿态类别不足且数据模式单一的问题。

**关键词:** 虚拟仿真; 手部姿态数据集; 运动学建模; 手部姿态估计; 人机交互

中图分类号: TP391.9 文献标志码: A 文章编号: 1004-731X(2023)04-0862-09

DOI: 10.16182/j.issn1004731x.joss.21-1265

**引用格式:** 陈佳昕, 周国辉, 杨建柏. 基于虚拟仿真方法的手部姿态数据集构建技术[J]. 系统仿真学报, 2023, 35(4): 862-870.

**Reference format:** Chen Jiaxin, Zhou Guohui, Yang Jianbai. Construction Technology of Hand Posture Dataset Based on Virtual Simulation Method[J]. Journal of System Simulation, 2023, 35(4): 862-870.

## Construction Technology of Hand Posture Dataset Based on Virtual Simulation Method

Chen Jiaxin, Zhou Guohui\*, Yang Jianbai

(Harbin Normal University, Harbin 150080, China)

**Abstract:** Hand posture is an important carrier of human-computer interaction, and the acquisition and recognition of posture information largely depends on the hand posture dataset. Existing datasets can be divided into two categories, real datasets and synthetic datasets. As real data is limited by equipment, environment, and other factors, the classification of hand posture is insufficient and the annotation is mixed with a lot of manual errors. The existing synthetic data can solve the data scale problem of real data, but the synthetic hand posture volume is limited and with some unreasonable kinematic postures of which the data form are only RGB images. By studying the hand anatomical structure, a virtual 3D kinematic model of hand is created based on virtual simulation, and a hand posture simulation generator is built. The hand posture dataset constructed by the generator can not only effectively solve the problems of the insufficient data volume of real data, low accuracy of annotation, and large workload of manual annotation, but also solve the problems of insufficient hand posture category and single data mode of synthetic data.

**Keywords:** virtual simulation; hand posture dataset; kinematics modeling; hand posture estimation; human-computer interaction

收稿日期: 2021-12-10 修回日期: 2022-02-14

基金项目: 研究生学术创新项目(HSDSSCX2021-118)

第一作者: 陈佳昕(1994-), 男, 硕士, 研究方向为数字孪生技术。Email: 844528295@qq.com

通讯作者: 周国辉(1973-), 男, 教授, 博士, 研究方向为机器视觉与模式识别。Email: zhouguohui@hrbnu.edu.cn

## 0 引言

手部姿态估计是指从图像或者视频中获取人体手部关节的位置, 从而进一步识别人体手部动作的行为隐喻。手部姿态估计已成为研究热点, 并提出了各种各样不同的方法且取得了不错的进展。基于手部姿态估计的应用也不断涌现, 并在人机交互<sup>[1]</sup>、虚拟现实和增强现实<sup>[2]</sup>、计算机辅助设计<sup>[3]</sup>等方面都有重要应用。在元宇宙系统中<sup>[4]</sup>, 手是其与虚拟世界交互的重中之重, 更是对手部姿态提出了更高的要求。一个高效而准确的手部姿态估计深度学习模型是基于手部姿态应用广泛使用的基本前提。而影响手部姿态深度学习模型性能的一个重要因素就是数据集。不同类别的手部姿态训练数据的缺乏与标注准确度等问题限制了基于深度学习方法的泛化性能。James Steven 等<sup>[5]</sup>指出大多数手部姿态估计系统在训练集外的泛化能力都不足, 需要有更好和更多的数据用以训练。Yuan 等<sup>[6]</sup>通过对3D手部姿态估计现状的研究发现对极端视点的手部姿态估计远远不够。Yuan 等<sup>[7]</sup>的研究证实使用最标准的CNN网络训练百万级别数据可以达到state-of-art, 但训练结果受到多视角、遮挡等固有因素影响, 结果表现不一致。因此, 扩大不同类别的多视角手部姿态数据规模、提升数据标注准确性是优化手部姿态估计模型学习性能的关键。

为解决人工标注准确性问题与劳动密集性问题, 国内外许多学者采用半自动或自动标注的方式对数据集进行标注。Zimmermann 等<sup>[8]</sup>使用半自动的方式对RGB图像数据进行标注。首先使用经过训练的模型对图像关节标注位置进行预测, 然后手工对标注进行调整。另外一种解决此类问题的方法是使用合成数据, 合成数据自动对模型数据进行精确标注。Kulon 等<sup>[9]</sup>为确保合成数据的标注质量, 在OpenPose获取的2D标注基础上, 回归MANO(hand model with articulated and non-rigid deformations)手部模型参数, 根据置信度进行阈值

分割, 去除置信度低的阈值来确保标注质量。为解决多视角手部姿态数据量不足的问题, Chen 等<sup>[10]</sup>使用三维模型创建了MVHM(multi-view hand mesh benchmark)合成数据集, 并提出了一种多视图融合的估计方法, 提高了手部姿态估计器的性能, 但只考虑水平面上的手部多视角, 并未考虑到纵向上的极端视角。

目前, 国内对手部姿态估计数据集方面的研究严重缺乏。本文使用虚拟仿真方法, 在ROS(robot operating system)机器人操作系统平台上建立一个动态仿真环境。基于解剖学关于手部尺寸与关节运动角度限制等基本事实, 构建一个手部姿态生成器模型, 用于生成全类别的手部姿态完备数据集。该手部姿态生成器生成的手部姿态, 可以有效解决真实环境中手部姿态类别数据量不足的问题, 同时也能避免生成目前合成数据集中存在的反关节、模型变形等违背手部运动学和手部形态特征的姿态。同时, 姿态生成器通过从多个视角(包含极端视角)采集手部姿态生成RGB、Depth 2种模态数据。

## 1 常用数据集

目前, 国内外很多研究人员根据各自不同的需求发布了多种手部姿态数据集。手部姿态数据的采集方式分为2种: ①使用传感器采集的真实数据; ②在虚拟环境使用虚拟传感器采集的合成数据。真实数据受限于实验场景的空间有限, 习惯性姿态动作和简单姿态动作在数据集中大量冗余, 同时视角变化范围有限, 导致手部姿态数据类别不足。数据采集也受到物理设备精度与环境的影响。构建大规模真实数据的难点在于数据标注的工程量与精度, 采用手工标注的方法是一件费时费力的任务, 且难免会产生不准确的标注数据, 限制了手部姿态真实数据集的数据规模。而合成数据集由三维软件建模生成, 获取的训练数据不会受到物理设备的影响, 数据标注自动生成, 不需要人工进行标注, 避免了人工标注误差。但

是，合成数据严重依赖于手部建模的准确性。目前存在的手部姿态合成数据基于建模软件进行数字化建模，或者使用生成式对抗网络等三维重建方法生成，并未考虑到手部解剖学上的关节运动限制与尺寸比例等限制，且目前这些合成数据只有 RGB 图像，极端视角下的手部姿态数据匮乏。

### 1.1 真实数据

(1) ICVL 数据集<sup>[11]</sup>：使用 Intel 交互式动作摄像机捕捉手部姿态的深度数据。先采用 3D 骨骼追踪的方式标注数据，再手工细化。数据集虽包含 18 万张带标注的图像，但只包含了深度图像数据，如图 1(a)所示。

(2) NYU 数据集<sup>[12]</sup>：使用 3 个 Kinect 传感器捕获手部 RGBD 数据。数据集包含 8 万张深度图像，同时包含合成的深度数据，但深度合成本身具有误差，同时 RGB 图像数据分辨率较低，如图 1(b)所示。

(3) BigHand2.2M<sup>[7]</sup>：使用 Kinematic6D 摄像机捕获手部姿态数据，共包含 220 万幅自动标记的手部深度图像。数据集规模大但却只有深度数据，如图 1(c)所示。

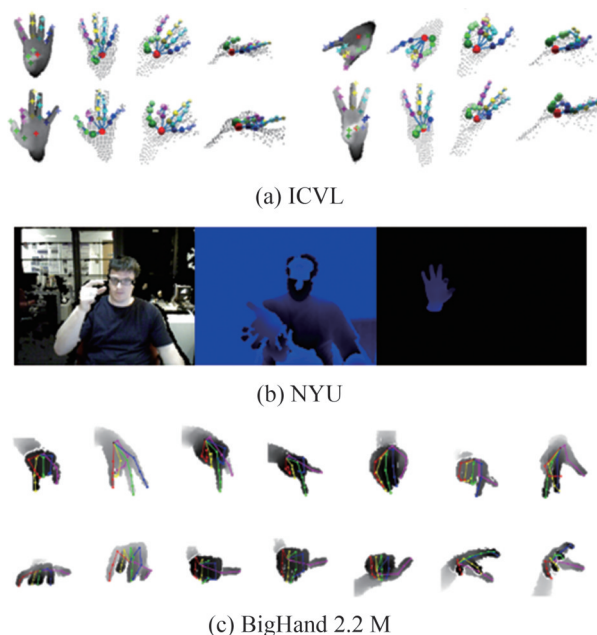


图 1 真实数据集  
Fig. 1 Real datasets

### 1.2 合成数据

(1) RHP 数据集<sup>[13]</sup>：使用三维手部模型和其他软件生成相应的手部动画并渲染。使用了 20 个不同的人物角色执行 39 个动作，但只有 RGB 图像数据，如图 2(a)所示。

(2) SynHand5M 数据集<sup>[14]</sup>：共提供 5M 数据，创建了一个手部三维模型并使用虚拟相机获取 RGB 图像，数据集规模大，但只有 RGB 图像数据，如图 2(b)所示。

(3) MVHM 数据集<sup>[10]</sup>：使用高仿真三维手部模型生成手部姿态并渲染。数据集中每张图像都包含 21 个关节点和 2 651 个网格点的完整注释，但缺少深度数据且数据规模小，如图 2(c)所示。

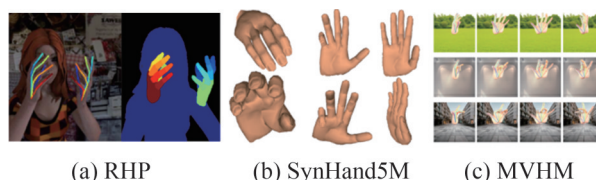


图 2 合成数据集  
Fig. 2 Synthetic datasets

## 2 手部建模

### 2.1 手的解剖学结构

根据解剖学原理，手骨共有 27 块，包括腕骨、掌骨、指骨 3 部分。腕骨由 8 块短骨排成 2 列，每列由 4 块小骨组成；掌骨共 5 块，由桡侧向尺侧分别称为第 1~5 掌骨；指骨共 14 块，除拇指为 2 节外，其余各 3 节，由近侧至远侧依次为近节指骨、中节指骨、远节指骨，如图 3 所示。每根骨段可以视作刚体，并且由指间关节串联形成一个整体。关节类型分为 2 类，髁突关节和转动关节。指骨由这 2 类关节连接在一起并围绕关节在一定的范围内运动<sup>[15]</sup>。

表 1~2 展示了手部关节的运动范围。CMC 表示腕掌关节，MCP 表示掌指关节，IP 表示大拇指的指间关节，PIP 和 DIP 分别表示除大拇指外其他

四指的近端指间关节和远端指尖关节。其中, E/F表示手指的纵向的伸展, Ab/Ad分别表示横向的外展和内收(手指握拳方向为纵向, 手指张开方向为横向)。

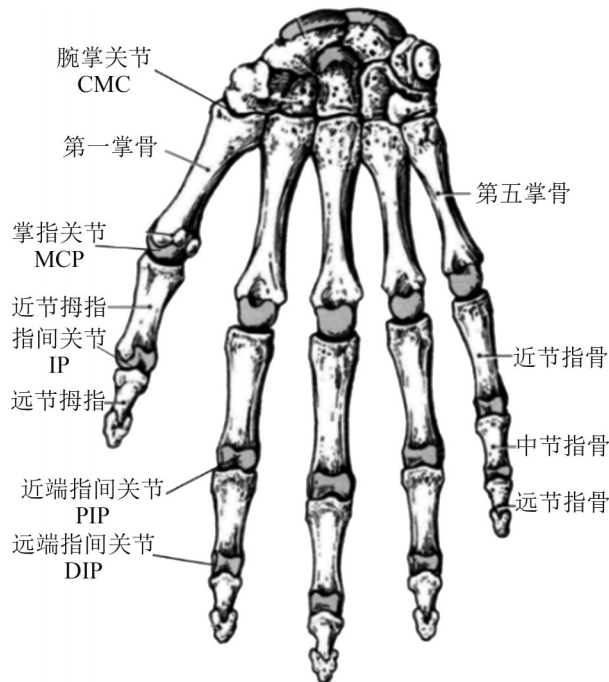


图3 手部骨骼解剖

Fig. 3 Skeletal anatomy of hand

表1 大拇指关节运动范围

Table 1 Range of motion of thumb joint

大拇指关节	旋转类型	角度限制/(°)
CMC	Ad/Ab	0/60
	E/F	25/35
MCP	E/F	-10/55
	Ad/Ab	0/60
IP	E/F	-15/80

表2 其余手指关节运动范围

Table 2 Range of motion of other finger joints (°)

手指	MCP(E/F)	MCP(Ad/Ab)	PIP(E/F)	DIP(E/F)
II	0/80	13/42	0/100	-10/90
III	0/80	8/35	0/100	-10/90
IV	0/80	14/20	0/100	-20/90
V	0/80	19/33	0/100	-30/90

## 2.2 运动学建模

在解剖学原理基础上, 许多研究者根据各自

需求, 对手部运动学模型进行了一定程度上的简化。Esteban等提出了一种25个自由度的手部骨骼模型, 在小指和环指的腕掌关节处分别添加了2个自由度, 可以形成掌弓<sup>[16]</sup>。Immaculada等提出的25个自由度的手部骨骼模型则是在食指、中指、环指、小指的腕掌关节处分别有1个自由度, 并且提出了手部拟人化程度的评价指标<sup>[17]</sup>。由于手部在自然运动情况下, 环指和小指在横向方向上具有微小运动, 所以本文在Esteban等研究的基础上, 建立一个具有25个自由度的运动学模型。其中5个在大拇指上, 中指和食指各4个, 环指和小指各5个, 具体关节自由度分布位置如图4(a)所示。

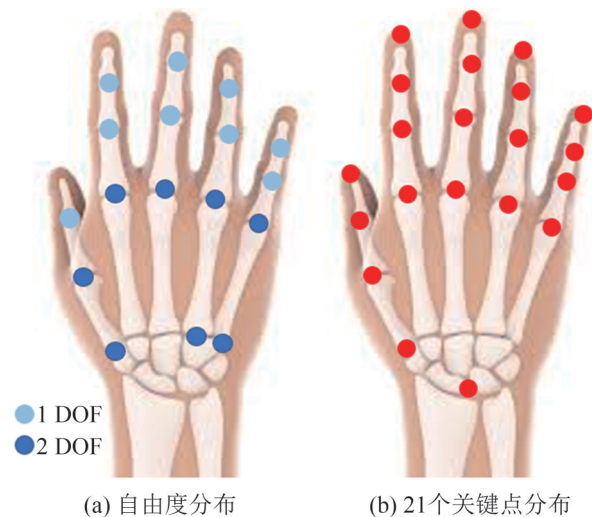


图4 手部关节自由度与关键点分布

Fig. 4 Degree of freedom and keypoints distribution of hand joints

由于指骨不容易发生形变, 我们将其视为刚体, 并使用DH参数法建立运动学模型。DH参数法<sup>[18]</sup>是由Denavit和Hartenberg于1995年提出的一种为关节链中每一杆件建立坐标系的矩阵方法, 主要用在机器人运动学上, 这种方法在每个连杆上建立一个坐标系, 通过齐次坐标变换来实现2个连杆上坐标的变换, 在多连杆串联的系统中, 多次使用齐次坐标变换, 就可以建立连杆各个端点的空间坐标变换关系。对于从腕部到4个指尖

的运动链，以腕部关节中心为世界坐标系原点，以每节指段近端关节中心点为该指段的本地坐标系原点。 ${}^i T_{i+1}$ 表示在以近端关节中心点为坐标系原点下远端关节中心点 $i+1$ 的齐次变换矩阵。由此，可以得到该手部各手指关节的正向运动学方程：

$${}^0 T_i = {}^0 T_1 \cdots {}^{i-1} T_i \quad (1)$$

由此，可以根据此手指关节的正向运动学方程计算出每个关节在以腕部为空间坐标原点的空间坐标：

$${}^0 P_{i+1} = {}^0 T_i {}^i P_{i+1} \quad (2)$$

式中： ${}^i P_{i+1}$ 为每根手指上远端关节中心点在以近端关节中心点 $i+1$ 为本地坐标系原点的空间坐标。

### 3 自碰撞检测

在现实中手指的协同操作过程，受限于手指与手指之间、手指指段与指段之间几何空间和关节角度，手部运动不会发生碰撞。在没有人体意识的参与下，虚拟空间中手部运动的碰撞检测显得尤为重要。在运动过程中，手指的姿态是随时间变化的，同一手指由于关节的限制不会发生碰撞，而不同的手指，各个指段之间都有可能发生碰撞，使得手部指节之间的自碰撞运动尤为复杂。

目前，有部分学者对自碰撞检测问题有较为深入的研究，并将其利用在双臂机器人的自碰撞避免方面与虚拟空间中。Shen等<sup>[19]</sup>通过将机械手的碰撞问题简化成4种简单几何体的碰撞问题，提出了一种基于基本几何元素矢量关系的机械手快速碰撞检测算法。Hou等<sup>[20]</sup>针对双臂机器人的协调运动规划方面的自碰撞问题，提出了一种基于快速探索随机树算法的避障算法。吴长征等<sup>[21]</sup>则在双臂机器人的自碰撞问题上，提出了一种基于空间向量几何的机械臂自碰撞检测方法，建立了一个处理9类自碰撞的碰撞检测模型。Wang等<sup>[22]</sup>对可变形模型提出了一种精确的检测算法，使用了一种利用曲面法向锥和双法向锥的新方法。

手部的运动方式和双臂机器人的运动方式相似，但更复杂。本文在笛卡尔空间中分析了空间向量的几何关系，计算出空间向量的距离，建立了手部指段的自碰撞检测模型。将手部的空间碰撞分为两大类情况：共面和异面。而在这两大类情况下，又分别具有3种不同的情况。

共面情况下：

(1) 最近的端点 $A$ 、 $C$ 的垂点 $M$ 、 $N$ 分别在 $\overline{CD}$ 、 $\overline{AB}$ 的延长线上，最短距离为 $\overline{AC}$ ，如图5(a)所示。

(2) 最近的端点 $A$ 、 $C$ 的垂点，其中一个在其中一线段的延长线上，另外一个的垂点则在另一线段上。 $C$ 点的垂点 $N$ 在 $\overline{AB}$ 上， $A$ 点垂点 $M$ 在的 $\overline{CD}$ 延长线上，最短距离为 $\overline{CN}$ ，如图5(b)所示。

(3) 最近的端点 $A$ 、 $C$ 的垂点 $M$ 、 $N$ 都在 $\overline{CD}$ 、 $\overline{AB}$ 上，最短距离为0，如图5(c)所示。

异面情况下：

(1)  $\overline{AB}$ 、 $\overline{CD}$ 的公垂线 $\overline{MN}$ 与 $\overline{AB}$ 、 $\overline{CD}$ 的交点 $M$ 、 $N$ 在 $\overline{AB}$ 、 $\overline{CD}$ 的延长线上，最短距离为 $\overline{AC}$ ，如图5(d)所示。

(2)  $\overline{AB}$ 、 $\overline{CD}$ 的公垂线 $\overline{MN}$ 与 $\overline{AB}$ 、 $\overline{CD}$ 的交点 $M$ 、 $N$ ，其中一个交点在另一线段的延长线上，另一个交点则在线段上， $M$ 在 $\overline{AB}$ 上， $N$ 在 $\overline{CD}$ 的延长线上，最短距离为 $\overline{CM}$ ，如图5(e)所示。

(3)  $\overline{AB}$ 、 $\overline{CD}$ 的公垂线 $\overline{MN}$ 与 $\overline{AB}$ 、 $\overline{CD}$ 的交点 $M$ 、 $N$ 分别在 $\overline{AB}$ 、 $\overline{CD}$ 上，最短距离为 $\overline{MN}$ ，如图5(f)所示。

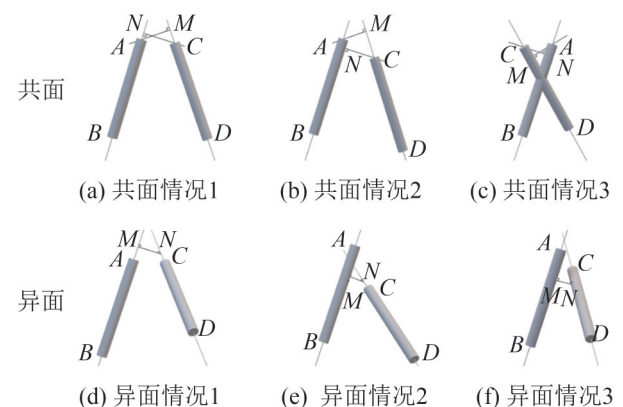


图5 手指段自碰撞检测模型

Fig. 5 Self-collision detection model of finger segmental

在手部运动的过程中, 手部的每个关节点都随着时间不断地发生变化, 每节指段的空间位置也随时间发生变化。现实中手部的指节粗细具有差别, 而在虚拟环境中我们需要实时进行自碰撞检测。为了提高检测的运算速度, 将手指视为半径为  $r$  的圆柱体。以食指和中指其中一节指段为例, 设这2个指段分别为  $A_i$  和  $B_j$ , 以腕部为空间坐标系原点, 两指段的端点坐标分别为  $P_i^s$ 、 $P_i^c$  和  $P_j^s$ 、 $P_j^c$ , 则指段关节中心点连线所在的方向向量分别为

$$d_i = P_i^c - P_i^s \quad (3)$$

$$d_j = P_j^c - P_j^s \quad (4)$$

计算叉乘可以得出异面直线公垂线向量

$$C_{ij} = d_i \times d_j \quad (5)$$

根据2个指段端点空间坐标, 计算其向量在异面直线公垂线上的投影长度, 即2条异面空间直线之间的距离  $D_{ij}$ 。

$$E = P_j^c - P_i^c \quad (6)$$

$$D_{ij} = \|C_{ij} \cdot E\| / \|C_{ij}\| \quad (7)$$

由此, 可以得到一个与异面直线公垂线向量共线、其大小等于异面空间直线距离的向量:

$$S = D_{ij} C_{ij} / \|C_{ij}\| \quad (8)$$

在此向量上移动其中一条线段即可使得2条直线共面。从而将异面和共面两大类情况统一到共面上, 从而简化异面空间直线关系的复杂性。假设空间中两指段4个端点之间距离最短的2个端点分别为  $M_i$  和  $M_j$ , 其中端点  $M_i$  在指段  $A$  上, 端点  $M_j$  在指段  $B$  上, 分别在指段  $A$  和指段  $B$  上取样  $n$  个点, 分别命名为  $P_{a1}, P_{a2}, \dots, P_{an}$  和  $P_{b1}, P_{b2}, \dots, P_{bn}$ 。则由  $M_i$  到  $P_{b1}, P_{b2}, \dots, P_{bn}$  的距离序列为  $SR_i$ , 由  $M_j$  到  $P_{a1}, P_{a2}, \dots, P_{an}$  的距离序列为  $SR_j$ 。对  $SR_i$  和  $SR_j$  进行单调性检测。

单调性检测情况分类:

(1)  $SR_i$ 、 $SR_j$  同时具有单调性。说明  $M_i$  的垂点在  $B$  的延长线或端点上。  $M_j$  的垂点在指段  $A$  的延长线或端点上。对应图 5(a)(d)。

(2)  $SR_i$  不具单调性,  $SR_j$  有单调性; 或者  $SR_i$

具有单调性,  $SR_j$  不具单调性。说明  $M_i$  的垂点在指段  $B$  上,  $M_j$  的垂点在指段  $A$  的延长线上; 或者相反。

(3)  $SR_i$ 、 $SR_j$  不具单调性。说明  $M_i$ 、 $M_j$  的垂点在指段  $A$ 、指段  $B$  上。对应图 5(c)。

最后, 由距离序列的单调性和最小值可计算出共面下, 指段  $A$ 、 $B$  之间的最短距离  $d_c$ 。如果是异面, 则根据勾股定理即可得出异面下指段  $A$ 、 $B$  之间的最短距离。

$$d_b = \sqrt{\|S\|^2 + d_c^2} \quad (9)$$

当最短距离大于等于指段直径时, 即  $d_b \geq r$  时, 指段  $A$ 、 $B$  将不会发生碰撞; 当最短距离小于指段直径时, 即  $d_b < r$  时, 指段  $A$ 、 $B$  将发生碰撞。

## 4 手部姿态生成

### 4.1 架构设计

本文基于 ROS 机器人操作系统构建手部姿态仿生成器。如图 6 所示, 基于手部解剖学原理, 对手部进行运动学建模, 并结合自碰撞检测算法随机生成完备的手部姿态, 并监听手部 21 个关键点的空间三维坐标。当手部姿态数目达到所需数据集规模时, 生成 BVH 动作捕捉文件。利用 Gazebo 仿真软件获取 BVH 文件关节数据, 在具有皮肤与骨架的手部仿真模型上上演, 并通过 Gazebo 中图像传感器采集手部姿态的 RGB 和 Depth 图像。

### 4.2 姿态生成方法

当前, 合成数据集只提供手部姿态的 RGB 图像数据, 且生成的数据量较小, 生成的手部姿态类别少, 尤其是极端视角下的手部姿态数据更少。使用这类数据集训练的深度学习模型性能往往不能达到很好的效果, 在训练集外的泛化能力也往往不足。基于 Yuan 等<sup>[6]</sup>的观点, 单张图像的 3D 手部姿态估计在 [70, 120] 的范围内误差最小, 因为此角度上的手部自遮挡的部位最少。受此启发, 极端视角下的手部姿态数据采集尤为重要。



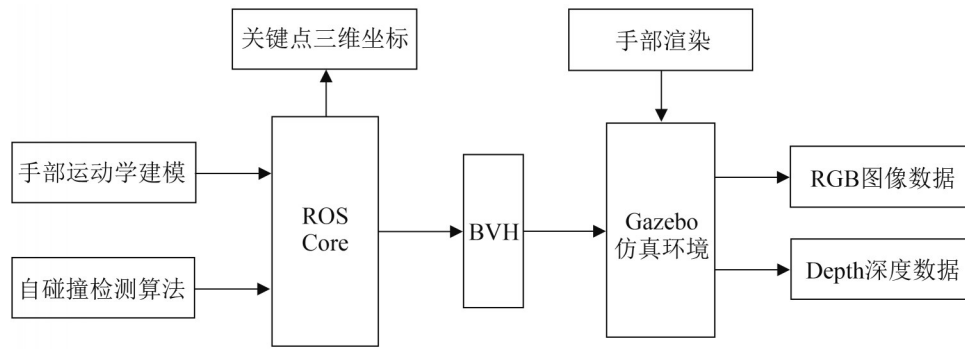


图6 手部姿态仿真生成器架构设计  
Fig. 6 Architecture design of hand posture simulation generator

手部运动学模型共有25个自由度，分别有25个转动控制器独立控制不包含腕部关节在内的16个关节的25个自由度的运动，并以固定的角度步长控制25个转动控制器的转动，模拟16个关节的伸展(EF)、外展和内收(Ab/Ad)运动，并通过自碰撞检测算法随机生成不发生自碰撞的手部姿态。

本文从12个方位采集虚拟空间手部模型的手部姿态。如图7(a)所示，展示了12个摄像机所处方位。其中，底部8个摄像机指向手掌中心，上部4个摄像机则与手部水平方向成45°，如图7(b)所示。其中，12个方位采集的数据包括RGB，Depth 2种模态数据。并通过ROS系统监听21个关键点的三维空间坐标，关键点分布如图4(b)所示。

本文手部模型25个自由度控制器控制17个关节的转动，并以固定的角度步长随机生成手部姿态。设 $R_{\min}^i$ 和 $R_{\max}^i$ 为第 $i$ 个转动控制器旋转角度限制的最小值和最大值， $S$ 为角度步长，其中， $n=0, 1, \dots$ 。则每个控制器所控制的角度集合为

$$G^i = [R_{\min}^i + nS, \dots, R_{\max}^i] \quad (10)$$

其中项数为

$$N^i = \lceil (R_{\max}^i - R_{\min}^i) / S \rceil + 1 \quad (11)$$

由此，则手部姿态的笛卡尔积为

$$H = N^1 \times N^2 \times \dots \times N^i \times \dots \times N^{25} \quad (12)$$

在上述的所有生成的手部姿态中，包含有三大类：

(1) 发生碰撞的姿态。利用第3节的自碰撞检

测在姿态生成前进行预测剔除。

(2) 非自然手部姿态。手指间受压产生或与物体交互受压产生的运动能量消耗较大的手部姿态。

(3) 自然手部姿态。日常经常使用的手部姿态。动能量消耗最小。显示中常使用的手势语言属于此类。

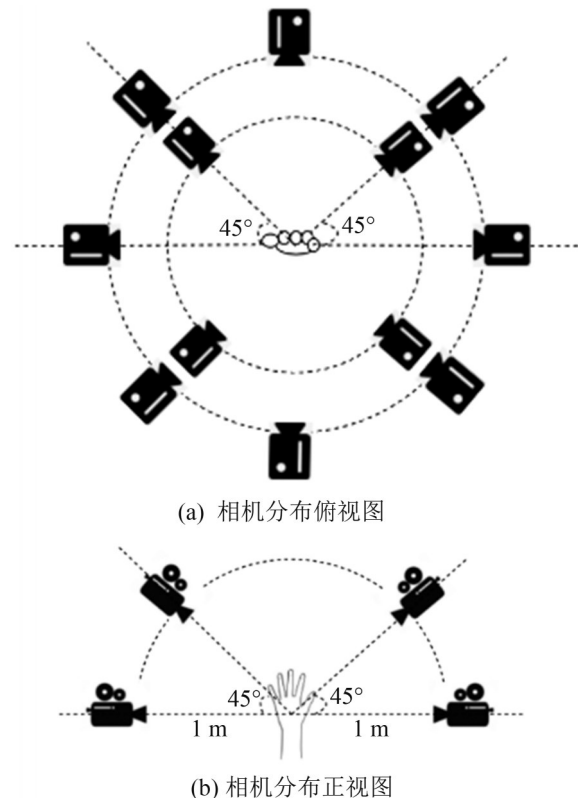
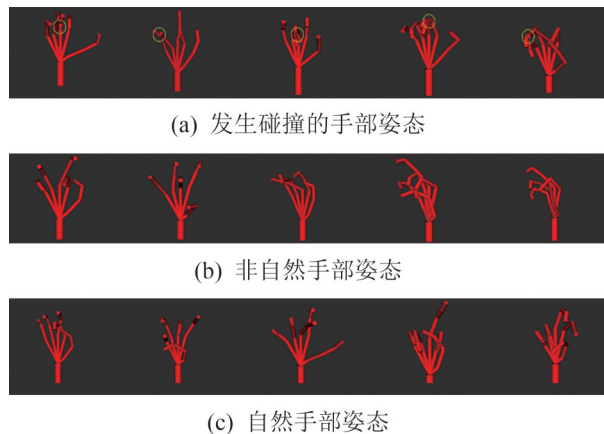


图7 手部姿态采集时相机分布  
Fig. 7 Camera distribution during hand posture acquisition

图8展示了使用本文手部姿态生成器生成的手部姿态示例。图9展示了最终采集的RGB和Depth图像。



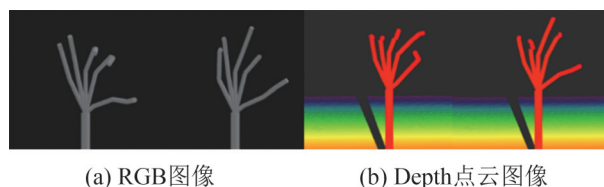
(a) 发生碰撞的手部姿态

(b) 非自然手部姿态

(c) 自然手部姿态

图8 三类手部姿态

Fig. 8 Three types of hand postures



(a) RGB图像

(b) Depth点云图像

图9 手部姿态RGB图像与Depth点云图像

Fig. 9 RGB images and Depth point-cloud images of hand postures

## 5 结论

本文基于虚拟仿真方法, 结合手部解剖学结构, 在ROS平台上构建了一个手部姿态仿真生成器, 并结合自碰撞检测算法生成不发生碰撞的手部姿态。在虚拟空间中从12个方位(包括极端视角)采集了RGB、Depth 2种模态的手部姿态数据, 基于运动学原理对手部姿态数据中手部21个关键点自动标注。有效解决了目前真实数据集数据量不足、标注准确度低、人工标注工作量大等问题, 也解决了当前合成数据手部姿态类别不足且数据模态单一的问题。

由于手部自由度多而导致的运动复杂性, 目前, 本文提出的手部运动暂时无法体现出拟人化运动的细腻性。虽然完善了目前手部姿态数据集

的现有缺点, 但这些改善对手部姿态估计任务的性能提升有多大帮助还未可知。未来的工作将着重于姿态仿真生成器运动的拟人化与数据集性能提升量化上。数据集获取请访问 <https://github.com/chenjiaxinEd/Hand>, 后续将陆续更新。

## 参考文献:

- [1] Gao Q, Chen Y, Ju Z, et al. Dynamic Hand Gesture Recognition Based on 3D Hand Pose Estimation for Human-Robot Interaction[J]. IEEE Sensors Journal (S1530-437X), 2021, 22: 17421-17430.
- [2] Wu M Y, Ting P W, Tang Y H, et al. Hand Pose Estimation in Object-Interaction Based on Deep Learning for Virtual Reality Applications[J]. Journal of Visual Communication and Image Representation (S1047-3203), 2020, 70: 102802.
- [3] Brereton M, Khan S, Rajapakse H, et al. GesCAD: An Intuitive Interface for Conceptual Architectural[C]// 29th Australian Conference on Computer-Human Interaction. 2017: 402-406.
- [4] Lee L H, Braud T, Zhou P. All One Needs to Know about Metaverse: A Complete Survey on Technological Singularity, Virtual Ecosystem, and Research Agenda [EB/OL]. (2021-11-03) [2022-03-30]. <https://arxiv.org/abs/2110.05352.pdf>.
- [5] Supančić J S, Rogez G, Yang Y, et al. Depth-Based Hand Pose Estimation: Data, Methods, and Challenges[C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1868-1876.
- [6] Yuan S, Garcia-Hernando G, Stenger B. Depth-Based 3D Hand Pose Estimation: From Current Achievements to Future Goals[C]// IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Salt Lake City, UT, USA: IEEE, 2018: 2636-2645.
- [7] Yuan S, Ye Q, Stenger B, et al. Bighand2.2M Benchmark: Hand Pose Dataset and State of the Art Analysis[C]//IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 4866-4874.
- [8] Zimmermann C, Ceylan D, Yang J, et al. FreiHAND: A Dataset for Markerless Capture of Hand Pose and Shape from Single RGB Images[C]// 2019 IEEE/CVF International Conference on Computer Vision(ICCV). Seoul, Korea (South): IEEE, 2019: 813-822.
- [9] Kulon D, Güler R A, Kokkinos I, et al. Weakly-Supervised Mesh-Convolutional Hand Reconstruction in the Wild[C]//2020 IEEE/CVF Conference on Computer

- Vision and Pattern Recognition(CVPR). Seattle, WA, USA: IEEE, 2020: 4989-4999.
- [10] Chen L, Lin S Y, Xie Y, et al. MVHM: A Large-Scale Multi-View Hand Mesh Benchmark for Accurate[C]// IEEE/CVF Winter Conference on Applications of Computer Vision(WACV). Waikoloa, HI, USA: IEEE, 2021: 836-845.
- [11] Tang D, Chang H J, Tejani A, et al. Latent Regression Forest: Structured Estimation of 3D Articulated Hand Posture[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus, OH, USA: IEEE, 2014: 3786-3793.
- [12] Tompson J, Stein M, Lecun Y, et al. Real-Time Continuous Pose Recovery of Human Hands Using Convolutional Networks[J]. ACM Transactions on Graphics(TOG)(S0730-0301), 2014, 33(5): 1-10.
- [13] Zimmermann C, Brox T. Learning to Estimate 3D Hand Pose from Single RGB Images[C]// 2017 IEEE International Conference on Computer Vision(ICCV). Venice, Italy: IEEE, 2017: 4913-4921.
- [14] Malik J, Elhayek A, Nunnari F, et al. DeepHPS: End-to-End Estimation of 3D Hand Pose and Shape by Learning from Synthetic Depth[C]// 2018 International Conference on 3D Vision(3DV). Verona, Italy: IEEE, 2018: 110-119.
- [15] Tubiana R, Thomine J M, Mackin E. Examination of the Hand and Wrist[M]. London, UK: Collaborative Research Centres Press, 1998.
- [16] Peña-Pitarch E, Falguera N T, Yang J. Virtual Human Hand: Model and Kinematics[J]. Computer Methods in Biomechanics and Biomedical Engineering(S1025-5842), 2014, 17(5): 568-579.
- [17] Llop-Harillo I, Pérez-González A, Gracia-Ibáñez V. Anthropomorphism Index of Mobility for Artificial Hands[J]. Applied Bionics and Biomechanics (S1176-2322), 2019: 1-11.
- [18] Lynch K M, Park F C. Modern Robotics[M]. Cambridge UK: Cambridge University Press, 2017.
- [19] Shen Y, Jia Q, Chen G, et al. Study of Rapid Collision Detection Algorithm for Manipulator[C]//2015 IEEE 10th Conference on Industrial Electronics and Applications(ICIEA). Auckland, New Zealand: IEEE, 2015: 934-938.
- [20] Hou Z, Ma S, Zeng Q, et al. Kinematics Analysis and Self-collision Detection of Truss Type Multi-robot Cooperative Welding Platform[J]. Procedia CIRP (S2212-8271), 2019, 81: 488-493.
- [21] 吴长征, 岳义, 韦宝琛, 等. 双臂机器人自碰撞检测及其运动规划[J]. 上海交通大学学报, 2018, 52(1): 45-53.  
Wu Changzheng, Yue Yi, Wei Baochen, et al. Self-Collision Detection and Motion Planning for Dual-Arm Robot[J]. Journal of Shanghai Jiaotong University, 2018, 52(1): 45-53.
- [22] Wang T, Tang M, Wang Z, et al. Accurate Self-collision Detection Using Enhanced Dual-cone Method[J]. Computers & Graphics(S0097-8493), 2018, 73: 70-79.