

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021年7月22日 (22.07.2021)



(10) 国际公布号
WO 2021/143261 A1

- (51) 国际专利分类号:
G06T 13/40 (2011.01)
- (21) 国际申请号: PCT/CN2020/123677
- (22) 国际申请日: 2020年10月26日 (26.10.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202010059845.3 2020年1月19日 (19.01.2020) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (72) 发明人: 林群芬 (LIN, Qunfen); 中国广东省深圳市南山区高新区科技中一路腾讯大厦35层, Guangdong 518057 (CN)。
- (74) 代理人: 广州华进联合专利商标代理有限公司 (ADVANCE CHINA IP LAW OFFICE); 中国广东省广州市天河区珠江东路6号4501房(部位: 自编01-03和08-12单元)(仅限办公用途), Guangdong 510623 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: ANIMATION IMPLEMENTATION METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 一种动画实现方法、装置、电子设备和存储介质

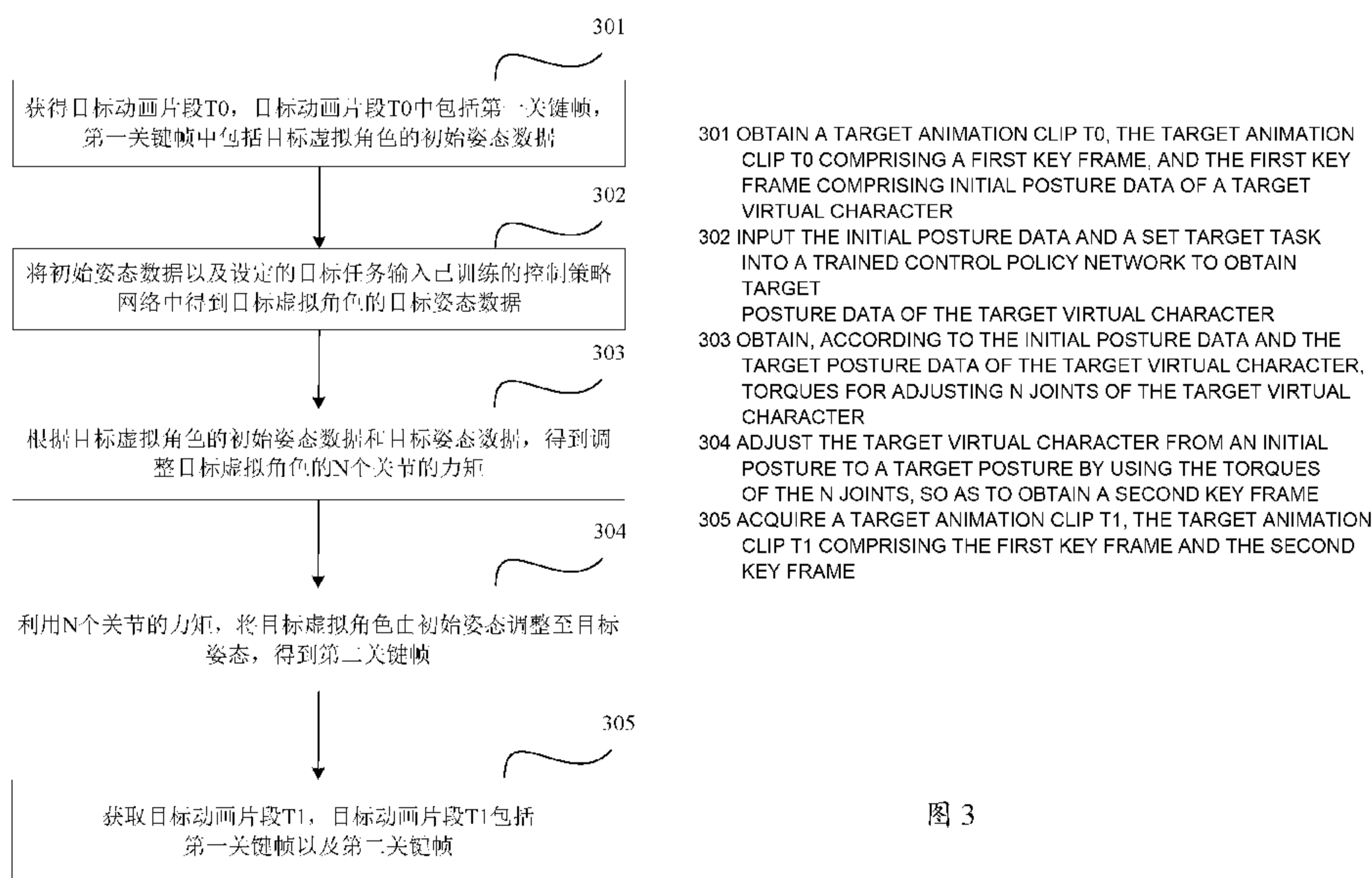


图3

(57) Abstract: An animation implementation method, comprising: obtaining a first target animation clip, the first target animation clip comprising a first key frame, and the first key frame comprising initial posture data of a target virtual character; inputting the initial posture data and a set target task into a trained control policy network to obtain target posture data of the target virtual character; obtaining, according to the initial posture data and the target posture data of the target virtual character, torques for adjusting a first preset number of joints of the target virtual character; adjusting the target virtual character from an initial posture to a target posture by using the torques of the first preset number of joints, so as to obtain a second key frame; and acquiring a second target animation clip, the second target animation clip consisting of at least the first key frame and the second key frame.

WO 2021/143261 A1

PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种动画实现方法, 包括: 获得第一目标动画片段, 第一目标动画片段中包括第一关键帧, 第一关键帧中包括目标虚拟角色的初始姿态数据; 将初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据; 根据目标虚拟角色的初始姿态数据和目标姿态数据, 调整目标虚拟角色的第一预设数量的关节的力矩; 利用第一预设数量的关节的力矩, 将目标虚拟角色由初始姿态调整至目标姿态, 得到第二关键帧; 获取第二目标动画片段, 第二目标动画片段由至少第一关键帧以及第二关键帧组成。

一种动画实现方法、装置、电子设备和存储介质

本申请要求于 2020 年 01 月 19 日提交中国专利局，申请号为 2020100598453，申请名称为“一种动画实现方法、装置、电子设备和存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及计算机技术领域、更涉及图像处理技术领域，尤其涉及一种动画实现方法、装置、电子设备和存储介质。

背景技术

随着科学技术的飞速发展，动画制作在各行各业中都起到较为广泛的作用。比如，对于游戏产业来说，不论是单机游戏、网游还是手游的开发，通常都会涉及到动画制作。

目前，动画制作主要是由人工绘制关键帧用于动画角色中，比如，游戏中动画人物的动作制作方法，主要是人工绘制关键帧用于游戏中的动画角色中。然而，人工绘制关键帧的方法需要逐帧设计，十分耗时耗力。

发明内容

根据本申请提供的各种实施例，提供了一种动画实现方法、装置、电子设备和存储介质。

根据本申请实施例的第一方面，提供了一种动画实现方法，由电子设备执行，包括：

获得第一目标动画片段，第一目标动画片段中包括第一关键帧，第一关键帧中包括目标虚拟角色的初始姿态数据；

将初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据；

根据目标虚拟角色的初始姿态数据和目标姿态数据，调整目标虚拟角色

的第一预设数量的关节的力矩；

利用第一预设数量的关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧；及

获取第二目标动画片段，第二目标动画片段包括第一关键帧以及第二关键帧。

第二方面，本申请实施例提供一种动画实现装置，设置于电子设备中，装置包括：

动画处理单元，用于获得目标虚拟角色的第一目标动画片段，第一目标动画片段中包括第一关键帧，第一关键帧中包括目标虚拟角色的初始姿态数据；

姿态获取单元，用于将初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据；

力矩获取单元，用于根据目标虚拟角色的初始姿态数据和目标姿态数据，得到调整目标虚拟角色的第一预设数量的关节的力矩；

姿态调整单元，用于利用第一预设数量的关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧；

动画生成单元，用于获取第二目标动画片段，第二目标动画片段包括第一关键帧以及第二关键帧。

第三方面，本申请实施例还提供一个或多个计算机可读存储介质，计算机可读存储介质内存储有计算机可读指令，计算机可读指令被一个或多个处理器执行时，实现本申请各实施例所述的动画实现方法。

第四方面，本申请实施例还提供一种电子设备，包括存储器和一个或多个处理器，存储器上存储有可在一个或多个处理器上运行的计算机可读指令，当计算机可读指令被一个或多个处理器执行时，使得一个或多个处理器实现本申请各实施例所述的动画实现方法。

本申请的一个或多个实施例的细节在下面的附图和描述中提出。基于本

申请的说明书、附图以及权利要求书，本申请的其它特征、目的和优点将变得更加明显。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例描述中所需要使用的附图作简要介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域的普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1 为本申请实施例提供的一种动画实现方法的系统架构示意图；

图 2 为本申请实施例提供的一种动画实现方法的应用场景示意图；

图 3 为本申请实施例提供的一种动画实现方法的流程图；

图 4 为本申请实施例中一种人型虚拟角色的关节位置示意图；

图 5 为本申请实施例提供的完成一个走路目标任务的示意图；

图 6 为本申请实施例提供的微分比例控制器的一个处理流程示意图；

图 7 为本申请实施例提供的一种控制策略网络的训练过程的流程图；

图 8 为本申请实施例提供的一种 actor 网络的网络结构示意图；

图 9 为本申请实施例提供的另一种 actor 网络的网络结构示意图；

图 10 为本申请实施例提供的一种价值评估网络的训练过程的流程图；

图 11 为本申请实施例提供的一种训练控制策略网络的收敛曲线的示意图；

图 12 示出了本申请实施例中一种已训练的控制策略网络输出效果展示图；

图 13 为本申请实施例提供的一种动画实现装置的结构示意图；及

图 14 为本申请实施例提供的一种电子设备的结构示意图。

具体实施方式

为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请作进一步地详细描述，显然，所描述的实施例仅仅是本申请一部分实施

例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其它实施例，都属于本申请保护的范围。

下文中所用的词语“示例性”的意思为“用作例子、实施例或说明性”。作为“示例性”所说明的任何实施例不必解释为优于或好于其它实施例。

文中的术语“第一”、“第二”仅用于描述目的，而不能理解为明示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征，在本申请实施例的描述中，除非另有说明，“多个”的含义是两个或两个以上。

以下对本申请实施例中的部分用语进行解释说明，以便于本领域技术人员理解。

骨骼动画：骨骼动画是模型动画中的一种，模型具有互相连接的“骨骼”组成的骨架结构，通过改变骨骼的朝向和位置来为模型生成动画。

虚拟角色：应用中的可操控对象，通过可操控对象的行为来推动应用的进程，若应用为 MOBA 游戏或者 RTS 游戏，虚拟对象为游戏中的可控制游戏角色，若应用为虚拟现实应用，虚拟对象为虚拟现实应用中的虚拟现实角色。

目标任务：用于指示虚拟角色完成指定动作的任务，例如，“前进”、“射击”、“翻跟斗”、“旋风踢球”等。每个虚拟角色可以完成多种类型的目标任务，不同类型的虚拟角色可以完成不同类型的目标任务。可以预先设置不同的控制指令，指示虚拟角色完成不同的目标任务，例如，玩家可以通过控制按键触发对应的控制指令，为虚拟角色设定当前的目标任务。

物理动画：在物理引擎中播放的动画称为物理动画，动画当前时刻的状态由前一时刻的状态和当前的受力情况决定，由于物理动画是对真实世界的模拟，因此可以得到较为逼真的动画效果。

物理引擎：指通过计算机可读指令模拟物理规律的引擎，主要用在计算

物理学、电子游戏以及计算机动画中，可以使用质量、速度、摩擦力及阻力等变量，预测虚拟角色在不同情况下的动作效果。

力矩 (torque): 也可以称为扭转的力，在物理学中，力矩指作用力促使物体绕着转动轴或支点转动的趋向；在本申请实施例中，力矩指作用力促使骨骼绕关节转动的趋向。

强化学习: 机器学习中的一个领域，强调如何基于环境而行动，以取得最大化的预期利益，被广泛应用于运动控制问题。

PPO: (Proximal Policy Optimization, 近端策略优化算法) 强化学习中策略梯度算法群中的一种，属于 off-policy (离策略) 算法，具有 actor (控制策略网络) 和 critic (价值评估网络) 两个子网络，该方法的优势是可以应用于连续型动作，收敛性快且可以实现分布式训练，是目前强化学习领域的主流算法。

Episode: 强化学习中的一个概念，强化学习中的虚拟角色与环境交互的一个连续片段称为 Episode，本申请实施例中为样本动画片段。

PD 控制器: (比例-微分控制器) 是一个在工业控制应用中常见的反馈回路部件，比例单元 P 控制当前位置和目标位置的误差，微分单元 D 控制当前速度和目标速度的误差。控制器通过输出与误差相关的值反馈到系统来实现系统稳定达到目标状态。PD 控制器输出的结果根据具体场景而定，如对于旋转问题，PD 控制器输出力矩，对于平移问题，PD 控制器输出力。

StablePD 控制器: PD 控制器的一种改进算法，传统的 PD 算法对控制参数 spring 和 damper 敏感，当 spring 参数设置过大时，容易出现高比例增益的稳定性问题。StablePD 同时考虑了下一个时间周期的位置和加速度实现更快更稳定的控制，试验证明，StablePD 可以使用更长的控制间隔来实现比传统 PD 更稳定的控制效果。

Retarget: 重定向技术，细分为角色重定向和环境重定向。角色重定向是指将动画数据从一个角色拷贝到另一个角色的过程，两个角色间可能骨架略

有不同，可能物理参数不同。环境重定向指将动画数据从一个环境拷贝到另一个环境的过程，如地形差异。

下面结合附图及具体实施例对本申请作进一步详细的说明。

为了解决相关技术中的技术问题，本申请实施例提供了一种动画实现方法、装置、电子设备和存储介质。本申请实施例涉及人工智能（Artificial Intelligence, AI）和机器学习（Machine Learning, ML）技术，基于人工智能中的计算机视觉（Computer Vision, CV）技术和机器学习而设计。

人工智能是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。换句话说，人工智能是计算机科学的一个综合技术，它企图了解智能的实质，并生产出一种新的能以人类智能相似的方式做出反应的智能机器。人工智能也就是研究各种智能机器的设计原理与实现方法，使机器具有感知、推理与决策的功能。人工智能技术主要包括计算机视觉技术、语音处理技术、以及机器学习/深度学习等几大方向。

随着人工智能技术研究和进步，人工智能在多个领域展开研究和应用，例如常见的智能家居、图像检索、视频监控、智能音箱、智能营销、无人驾驶、自动驾驶、无人机、机器人、智能医疗等，相信随着技术的发展，人工智能将在更多的领域得到应用，并发挥越来越重要的价值。

计算机视觉技术(Computer Vision, CV)是一门研究如何使机器“看”的科学，更进一步的说，就是指用摄影机和电脑代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图形处理，使电脑处理成为更适合人眼观察或传送给仪器检测的图像。作为一个科学学科，计算机视觉研究相关的理论和技术，试图建立能够从图像或者多维数据中获取信息的人工智能系统。计算机视觉技术通常包括图像处理、图像识别、图像语义理解、图像检索、OCR、视频处理、视频语义理解、视频内容/行为识别、三维物体重建、3D技术、虚拟现实、增强现实、同步定位与地图构建等技术，还包括常见的人脸识别、指

纹识别等生物特征识别技术。

机器学习是一门多领域交叉学科，涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科。专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。机器学习是人工智能的核心，是使计算机具有智能的根本途径，其应用遍及人工智能的各个领域。机器学习和深度学习通常包括人工神经网络、置信网络、强化学习、迁移学习、归纳学习等技术。本申请实施例采用了基于深度强化学习的控制策略网络对包含参考虚拟角色的姿态序列的参考动画片段进行学习，并在生成目标动画片段的过程中，利用学习得到的控制策略网络针对不同的虚拟角色生成动画片段。

本申请实施例提供的动画实现方法可以应用于 3D (3 Dimensions, 三维) 立体游戏、3D 动画电影以及 VR (Virtual Reality, 虚拟现实) 等场景中。例如，在 3D 立体游戏中，一般包含大量的虚拟角色。这里的虚拟角色也可以称为物理角色，在物理引擎中，这些虚拟角色可以拥有质量，受到重力的作用等等。在一些实施例中，虚拟角色可以由骨骼构成，骨骼是指由关节搭建出的且可以活动的骨架，是活动的虚拟主体，驱动整个虚拟角色运动。在另一些实施例中，虚拟角色可以由骨骼和蒙皮构成，蒙皮是指包裹在骨骼周围的三角形网格，网格的每个顶点被一个或多个骨骼控制，当骨骼周围包裹有蒙皮时，骨骼不会在游戏画面中被渲染出来。

在游戏中，虚拟角色可以由玩家进行控制，也可以按照游戏进度自动实现角色的控制。虚拟角色的类型也多种多样，如“战士”、“法师”、“射手”、“运动员”等。不同类型的虚拟角色有一部分动作类型是相同的，比如，跑、走、跳以及蹲等，还有一部分动作类型是不同的，比如攻击方式和防御方式等。并且，不同类型的虚拟角色可以完成相同类型的目标任务，也可以完成不同类型的目标任务。本申请实施例提供的动画实现方法，可以根据针对虚拟角色设定目标任务，生成动画片段。

本申请实施例提供的动画实现方法的一种应用场景可以参见图 1 所示，该应用场景中包括终端设备 101 和游戏服务器 102。终端设备 101 与游戏服务器 102 之间可以通过通信网络进行通信。该通信网络可以是有线网络或无线网络。

其中，终端设备 101 为可以安装各类应用程序，并且能够将已安装的应用程序的运行界面进行显示的电子设备，该电子设备可以是移动的，也可以是固定的。例如，手机、平板电脑、各类可穿戴设备、车载设备或其它能够实现上述功能的电子设备等。各终端设备 101 通过通信网络与游戏服务器 102 连接，游戏服务器 102 可以是游戏平台的服务器，可以是一台服务器或由若干台服务器组成的服务器集群或云计算中心，或者是一个虚拟化平台。

请参考图 2，其示出了本申请一个实施例提供的动画实现方法的应用场景的示意图。终端设备 101 中安装有某 AR (Augmented Reality, 增强现实) 游戏的客户端。如图 2 所示，在该 AR 游戏的游戏界面中，采集现实环境的视频信息，即办公桌 220 的桌面，终端设备中的物理引擎结合办公桌 220 的桌面情况，生成虚拟角色 210 的动画，从而显示虚拟角色 210 在办公桌 220 上的 AR 动画。

在该 AR 游戏中，虚拟角色可以跟用户、跟当前场景交互。在在一个实施例中，客户端接收到用户通过控制按键输入的指示虚拟角色 210 执行某一项目标任务（如蹲下）的操作，获取虚拟角色 210 当前第一关键帧的状态信息以及该目标任务并发送至游戏服务器 102。游戏服务器 102 上针对不同的目标任务存储有已经训练好的控制策略网络，控制策略网络是根据包含参考虚拟角色的参考姿态序列的参考动画片段训练得到的。游戏服务器 102 将虚拟角色 210 在第一关键帧的状态信息以及目标任务输入控制策略网络，控制策略网络输出第二关键帧的目标姿态数据，并将目标姿态数据发送回客户端。客户端根据虚拟角色 210 的初始姿态数据和目标姿态数据，计算得到虚拟角色 210 的各个关节的力矩。客户端基于物理引擎，根据各个关节的力矩调整

第一关键帧中虚拟角色 210 的各个关节，得到第二关键帧中的虚拟角色的目标姿态，进而生成第一关键帧与第二关键帧之间的固定帧以及第二关键帧。客户端依次显示第一关键帧、第一关键帧与第二关键帧之间的固定帧、第二关键帧，获得虚拟角色执行设定的目标任务的动画片段。

在另一种实施例中，上述过程可以由终端设备 101 中安装的客户端独立完成。客户端接收到用户通过控制按键输入的指示虚拟角色 210 执行某一目标任务的操作，获取虚拟角色 210 当前第一关键帧的状态信息以及该目标任务，终端设备 101 中针对不同的目标任务存储有已经训练好的控制策略网络，控制策略网络是根据包含参考虚拟角色的参考姿态序列的参考动画片段训练得到的。终端设备 101 中将虚拟角色 210 在第一关键帧的状态信息以及目标任务输入控制策略网络，控制策略网络输出第二关键帧的目标姿态数据。客户端根据虚拟角色 210 的初始姿态数据和目标姿态数据，计算得到虚拟角色 210 的各个关节的力矩。客户端基于物理引擎，根据各个关节的力矩调整第一关键帧中虚拟角色 210 的各个关节，得到第二关键帧中的虚拟角色的目标姿态，进而生成第一关键帧与第二关键帧之间的固定帧以及第二关键帧。客户端依次显示第一关键帧、第一关键帧与第二关键帧之间的固定帧、第二关键帧，获得虚拟角色执行设定的目标任务的动画片段。

需要说明的是，本申请提供的动画实现方法可以应用于游戏服务器 102，也可以应用于终端设备的客户端中，由终端设备 101 实施本申请提供的动画实现方法，还可以由游戏服务器 102 与终端设备 101 中的客户端配合完成。

图 3 示出了本申请一个实施例提供的动画实现方法的流程图。如图 3 所示，该方法包括如下步骤：

步骤 S301，获得目标动画片段 T0，目标动画片段 T0 中包括第一关键帧，第一关键帧中包括目标虚拟角色的初始姿态数据。

其中，目标虚拟角色可以是人物形态，可以是动物、卡通或者其它形态，本申请实施例对此不作限定。目标虚拟角色可以通过三维形式展示，也可以

通过二维形式展示。目标虚拟角色具有骨骼，相邻的骨骼之间具有关节，通过改变每个关节的位置和旋转角度，可以改变目标虚拟角色的姿态，将目标虚拟角色的一系列姿态连贯后，即可形成连贯的动作。

本申请实施例中利用骨骼动画进行计算和呈现，可以让动画效果做得更丰富。并且，做到了虚拟对象和动画分离，即只需要记录虚拟对象对于骨骼的蒙皮权重，就可以单独的去制作骨骼的动画，在保证蒙皮信息和骨骼信息一致的情况下，还可以多个虚拟对象之间共享骨骼动画。此外，也大大的节省资源容量。

目标动画片段 T0 可以是终端设备的物理引擎中预先保存的动画片段，或者是游戏中已生成的动画片段，也可以是用户输入的动画片段。目标动画片段 T0 中包含第一关键帧。帧就是动画中最小单位的单幅影像画面，相当于电影胶片上的每一格镜头，在动画软件的时间轴上帧表现为一格或一个标记。关键帧相当于二维动画中的原画，指角色或者物体运动或变化中的关键动作所处的那一帧，由于一个视频很多帧其实与所做的动作无关，将与动作有关的帧作为关键帧。本申请实施例中，关键帧与关键帧之间的动画为固定帧，其中，相邻两关键帧之间相差的帧数为预先设置，例如，可以为每 20 帧设置一个关键帧，即动画片段的第 0 帧、第 20 帧、第 40 帧……设置为关键帧，则第 1 帧至第 19 帧、第 21 帧至第 39 帧……为固定帧。

为了便于介绍，请参阅图 4，图 4 为本申请实施例中一种人型虚拟角色的关节位置示意图，如图所示，假设人型角色有 15 个关节，其中，A15 所指示的关节为根（root）关节，在人型角色的盆骨位置，也是最顶层的父节点。其余的 14 个关节分别为 A1 所指示的胸膛、A2 所指示的脖子、A3 所指示的右腿、A4 所指示的左腿、A5 所指示的右膝、A6 所指示的左膝、A7 所指示的右踝、A8 所指示的左踝、A9 所指示的右大臂、A10 所指示的左大臂、A11 所指示的右肘、A12 所指示的左肘、A13 所指示的右手腕以及 A14 所指示的左手腕。其中肘关节和膝关节是转动关节，腕关节是固定关节，其他

关节均为球关节，因此，虚拟角色可以配置共 43 个自由度。可以理解的是，人型虚拟角色还可以包括其他数量的关节，此处仅为一个示意，不应理解为本申请的限定。

在一个实施例中，虚拟角色所有关节的位置数据和旋转数据构成了该虚拟角色的姿态数据。角色的关节位置可以用三维向量表示，角色的关节旋转则可以使用四元数来表示，因此，一个角色关节的姿态可以用七维向量表示。

上述目标虚拟角色的初始姿态数据可以由终端设备基于物理引擎获得，即将目标动画片段 T0 输入物理引擎中，物理引擎可直接输出该动画片段 T0 中第一关键帧中的初始姿态数据。

步骤 S302、将初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据。

在游戏过程中，用户可以通过控制按键控制虚拟角色执行不同的动作。在本申请实施例中，每个控制按键对应一项目标任务，用户可以通过控制按键为虚拟角色设定目标任务，虚拟角色将执行该目标任务对应的动作。例如，用户通过显示界面上的控制按键输入“跳跃”的控制指令，使目标虚拟角色从地面跳起，则针对目标虚拟角色设定的目标任务即为“跳跃”。目标任务还可以是其它任务，如使目标虚拟角色沿给定的方向前进，或者让目标虚拟角色使用旋风踢的动作，踢到指定的位置，本申请实施例对此不作限定。

在一个实施例中，目标姿态数据可以为终端设备直接将至少初始姿态数据以及设定的目标任务输入控制策略网络中得到，此时，终端设备中存储有已训练的控制策略网络。

在一个实施例中，目标姿态数据也可以为终端设备通过与服务器交互得到，此时，服务器中存储有已训练的控制策略网络，终端设备将至少初始姿态数据以及设定的目标任务向服务器发送，服务器将至少初始姿态数据以及设定的目标任务输入控制策略网络中得到目标姿态数据，再将计算得到的目标姿态数据反馈回终端设备。

步骤 S303、根据目标虚拟角色的初始姿态数据和目标姿态数据，得到调整目标虚拟角色的 N 个关节的力矩。其中，N 为大于或者等于 1 的正整数。

力矩指作用力促使骨骼绕关节转动的趋向。在一个实施例中，可以利用 StablePD 控制器计算力矩值，即将目标虚拟角色的初始姿态数据以及目标姿态数据输入 StablePD 控制器中，得到作用于目标虚拟角色的各个关节的力矩，并将力矩反馈到物理引擎中从而实现将目标虚拟角色稳定的控制至目标姿态。

步骤 S304、利用 N 个关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧。

具体实施过程中，可以由物理引擎将各个关节的力矩作用于目标虚拟角色的各个关节，对上一帧动画画面中目标虚拟角色的姿态进行调整，得到下一帧动画画面中目标虚拟角色的姿态。通过物理引擎将力矩直接作用于目标虚拟角色的各个关节，从而可以对各个关节的姿态进行准确地调整，使得目标虚拟角色呈现出相应的动作姿态，动画效果更佳丰富自然，实现了传统动画技术达不到的动作效果，使目标虚拟角色的动作姿态更加逼真。

步骤 S305、获取目标动画片段 T1。

其中，目标动画片段 T1 包括第一关键帧以及第二关键帧。第二关键帧为第一关键帧的下一关键帧。

进一步的，可以将第二关键帧中目标虚拟角色的姿态数据作为初始姿态数据，至少将第二关键帧中的初始姿态数据以及目标任务输入策略控制网络得到第三关键帧中目标虚拟角色的目标姿态数据，第三关键帧为第二关键帧的下一关键帧。利用第二关键帧中的初始姿态数据和第三关键帧中目标虚拟角色的目标姿态数据得到调整第二关键帧中目标虚拟角色各个关节的力矩，并利用力矩，将目标虚拟角色由第二关键帧中的初始姿态调整至第三关键帧中的目标姿态，得到第二关键帧与第三关键帧之间的固定帧以及第三关键帧。以此类推，可以依次生成后续的动画帧。

将目标虚拟角色的一系列姿态连贯后，即可形成连贯的动作，得到目标

虚拟角色通过一系列动作执行目标任务的动画片段。

可以理解，在其他实施例中，也可以将目标动画片段 T0 和目标动画片段 T1 分别用“第一目标动画片段”和“第二目标动画片段”来做区分性表述。

本申请实施例利用控制策略网络得到第二关键帧中目标虚拟角色的目标姿态数据，并通过目标虚拟角色的初始姿态数据和目标姿态数据，得到调整目标虚拟角色的力矩，利用力矩调整目标虚拟角色的初始姿态，得到第二关键帧，即可以根据第一关键帧和目标任务生成目标虚拟角色的姿态序列，进而得到目标虚拟角色的动画片段，从而缩短了工作人员的工作时间，提升了工作效率。此外，由于目标虚拟角色是利用力矩调整目标虚拟角色的姿态，因此赋予了目标虚拟角色真实的物理属性，并基于此来计算目标虚拟角色的运动得到目标虚拟角色的姿态，因此可以得到更符合实际场景的目标姿态，实现更逼真的动作效果。

本申请实施例中，目标虚拟角色的目标姿态数据基于强化学习模型计算输出。相关技术中，有利用深度学习模型进行虚拟角色运动轨迹的控制。深度学习主要通过大量的有标记样本来进行训练，基于深度模型强大的标识学习的能力，在拥有大量有标记数据的场景下往往能够得到显著优越的效果。

然而，在很多场景下有标记的样本很难获得，代价很大。例如本申请实施例中的游戏为例，若利用深度学习技术，就必须用大量的人类玩家在游戏内的操作数据来训练模型。这就意味着这种方法只在那些已经上线且记录了用户的大量操作数据的游戏中有实现的可能性，适用范围有限。角色动画等场景也存在着相似的问题，本身就不存在足够的数据来进行训练。

与深度学习不同的是，深度强化学习是让虚拟角色在环境中进行探索来学习策略，不需要大量经过标记的样本，无需进行数据预处理，效率较高，更为适应本申请实施例中的动画游戏场景。

针对不同的目标任务和应用场景，向强化学习模型中输入的数据类型不同。

一种可能的实施例中，上述第二关键帧中目标虚拟角色的目标姿态数据通过以下方式获取：

获取目标虚拟角色在第一关键帧中的状态信息；

将目标虚拟角色在第一关键帧中的状态信息和目标任务输入控制策略网络，获得控制策略网络输出的目标虚拟角色的目标姿态数据。

其中，控制策略网络是根据参考动画片段训练得到的，参考动画片段包含参考虚拟角色完成目标任务的参考姿态序列。

参考虚拟对象是指已经制作好动画的虚拟对象，参考虚拟对象的动作称为参考动作，可以用运动学来表示参考动作。可选地，目标虚拟对象和参考虚拟对象是两个骨架相同或相近的虚拟对象。

骨架相同，可以表示为目标虚拟对象和参考虚拟对象的骨架中的所有骨骼的尺寸、形状，以及骨骼之间的连接关系完全相同；骨架相近，包括但不限于以下至少一种：目标虚拟对象和参考虚拟对象的骨架中的骨骼的尺寸相近、形状相近、骨骼之间的连接关系相近。

参考虚拟动画支持单个虚拟角色的任意动画片段，可以通过动作捕捉设备获取的数据，也可以是简单的视频片段。参考虚拟动画中，参考虚拟角色可以完成走路，奔跑，旋转踢，后空翻等多个动作。

目标虚拟角色的状态信息用于表征目标虚拟对象的物理状态，可以包括目标虚拟角色的当前相位数据、当前的初始姿态数据、当前速度数据和历史姿态序列。

其中，当前相位数据用于表征目标虚拟角色在第一关键帧中所处的阶段。当前相位数据的取值范围为 0 至 1，用于定义目标虚拟角色的当前状态在动作片段中所处的阶段。对于可循环动画而言，整个学习的动画序列存在周期性，而不同周期的同一阶段的动作具备高度相似性，本申请实施例中利用相位信息来使得模型在相同相位输出类似的动作信息，增强模型的拟合能力。

例如，完成一个走路目标任务需要如图 5 所示的 6 帧动画画面，第一关

键帧动画画面为图 5 中的第 5 帧动画画面，则相位数据为 $5/30=1/6$ 。相位数据也可以通过时间进行标识。例如，完成某一个目标任务需要 30 帧动画画面，播放这 30 帧动画画面所需的总时间长度为 T ，将第一帧动画画面对应的播放时刻记为开始时刻，设上一帧动画画面对应的播放时刻 t ，则上一帧动画画面中目标虚拟角色的相位数据 Ph 可以表示为 $Ph=t/T$ 。

当前的初始姿态数据用于表征目标虚拟角色当前的姿态，即目标虚拟角色在第一关键帧中的姿态。具体实施过程中，虚拟角色的姿态数据包括虚拟角色所有关节的位置数据和旋转数据。其中，虚拟角色的关节位置可以用三维向量表示，虚拟角色的关节旋转则可以使用四元数来表示，因此，在一个实施例中，一个虚拟角色的姿态可以用七维向量表示。可以理解的是，姿态数据中的三维向量表示关节的空间坐标，而旋转数据中的单位四元数表示三维空间里的旋转情况。当然，姿态数据还可以采用其他形式来表征，采用多种形式的的数据表示目标虚拟角色的姿态，可以更准确地确定目标虚拟角色的姿态。

其中，四元数是一种超复数。复数是由实数加上虚数单位 i 组成，相似地，四元数都是由实数加上三个虚数单位 i 、 j 、 k 组成，并且，三个虚数之间具有如下关系： $i^2=j^2=k^2=-1$ ， $i^0=j^0=k^0=1$ 。每个四元数都是 1 、 i 、 j 和 k 的线性组合，四元数一般可表示为： $a+bi+cj+dk$ ，其中 a 、 b 、 c 、 d 是实数。 i 、 j 、 k 本身的几何意义可以理解为一种旋转，其中 i 旋转代表 X 轴与 Y 轴相交平面中自 X 轴正向向 Y 轴正向的旋转， j 旋转代表 Z 轴与 X 轴相交平面中自 Z 轴正向向 X 轴正向的旋转， k 旋转代表 Y 轴与 Z 轴相交平面中 Y 轴正向向 Z 轴正向的旋转， $-i$ 旋转表示 i 旋转的反向旋转， $-j$ 旋转表示 j 旋转的反向旋转， $-k$ 旋转表示 k 旋转的反向旋转。

当前速度数据用于表征目标虚拟角色当前的速度状态，即目标虚拟角色在第一关键帧中的速度。具体实施过程中，目标虚拟角色的当前速度数据包括虚拟角色所有关节的线速度和角速度。由于每个关节的线速度可以用三维

向量表示，分别为 X 轴上的速度、Y 轴上的速度以及 Z 轴上的速度；且角速度也可以用三维向量表示，分别为 X 轴上的速度、Y 轴上的速度以及 Z 轴上的速度。因此，一个角色关节的速度可以用六维向量表示。若目标虚拟角色包括 N 个关节，则目标虚拟角色的速度维数可以是 $N \times 6$ 。采用线速度和角速度的组合来表示目标虚拟角色的速度数据，有利于更准确地确定目标虚拟角色的速度。

历史姿态序列用于表征目标虚拟角色在历史时间段内的姿态。具体实施过程中，可以将目标虚拟角色的根关节的历史时间段内的姿态信息作为目标虚拟角色的历史姿态序列，例如在 1s 的历史时间窗口内采 10 帧的根关节，由于每个根关节的姿态用七维向量表示，目标虚拟角色的历史姿态序列则使用 70 维的向量来表示。目标虚拟角色的历史姿态序列可以更全面地描述目标虚拟角色的当前状态。

本申请实施例中目标虚拟角色的状态信息还可以包括其他形式的表征数据，将多种维度多种形式的输入数据输入强化学习模型，用于描述目标虚拟角色的当前物理状态，从而能够得到更为准确的计算结果。

将目标虚拟角色在第一关键帧中的状态信息和目标任务输入控制策略网络，获得目标虚拟角色的目标姿态数据。目标任务输入控制策略网络时，目标任务可以使用向量表示。例如，假设目标任务是使目标虚拟角色沿给定的方向前进，给定的方向可以通过在水平平面上的一个二维向量表示。可以将目标虚拟角色在上一帧动画画面中的状态信息和表示目标任务的向量拼接在一起输入控制策略网络，控制策略网络将输出用于调目标虚拟角色的各个关节的力矩。

本申请实施例中将目标虚拟角色的状态信息和目标任务输入控制策略网络，获得目标虚拟角色的目标姿态数据。在一个实施例中，采用基于强化学习的玩家评判 (Actor-Critic, AC) 算法框架进行训练，AC 算法框架包括 Actor 网络以及 Critic 网络，本申请实施例中的控制策略网络即为 Actor 网络，价

值评估网络即为 Critic 网络。其中，控制策略网络训练的是当前策略，输出姿态数据，而价值评估网络用于指导控制策略网络的学习。训练的时候需要对控制策略网络和价值评估网络均进行训练，实际应用的时候仅使用控制策略网络即可。控制策略网络和价值评估网络的具体训练过程将在下文中详细介绍。

在一个实施例中，还需要考虑目标虚拟角色所处的场景环境对目标虚拟角色的影响。即当实际场景中的环境与参考动画中的环境不一致时，需要结合目标虚拟角色所处的场景环境。则此时，第二关键帧中目标虚拟角色的目标姿态数据通过以下方式获取：

获取目标虚拟角色在第一关键帧中的状态信息和目标虚拟角色所处的场景环境的环境信息；

将目标虚拟角色在第一关键帧中的状态信息、目标任务和目标虚拟角色所处的场景环境的环境信息输入控制策略网络，获得控制策略网络输出的目标虚拟角色的目标姿态数据。

其中，控制策略网络是根据参考动画片段训练得到的，参考动画片段包含参考虚拟角色完成目标任务的参考姿态序列。

环境信息用于表征目标虚拟对象所处的虚拟环境。虚拟环境可以由物理引擎模拟的物理环境，在模拟的物理环境中，虚拟对象遵守动力学规律，使得虚拟对象的运动接近现实情况。

其中，上述场景环境可以是物理引擎在终端设备中运行时显示(或提供)的场景，该场景环境是指营造出的供目标虚拟对象进行活动(如游戏竞技)的场景。该场景环境可以是对真实世界的仿真环境，也可以是半仿真半虚构的环境，还可以是纯虚构的环境。场景环境可以是二维虚拟环境，也可以是2.5 维虚拟环境，或者是三维虚拟环境，本申请实施例对此不作限定。

例如，如果物理引擎中的环境与参考动画中的环境不同是地形不同时，则环境信息为当前角色周围地形的高度图。

本申请实施例中，通过上述控制策略网络，输出目标虚拟角色的目标姿态数据。目标姿态数据用于表征目标虚拟角色下一个时间片段需要达到的目标姿态，而这个目标姿态数据具体用于力矩值的计算，因此，这里的目标姿态数据为目标虚拟角色所有关节的旋转数据，因此，一个关节的目标姿态数据可以用四元数表示。

通过上述基于强化学习模型的方法，可以实现角色的重定向和环境的重定向，即只需生成一个参考虚拟角色基于某一场景的参考动画序列，则针对与该参考虚拟角色骨架相同或骨架相近的目标虚拟角色的类似场景动画，可以直接利用参考动画序列，通过模型生成目标虚拟角色类似场景的动画，大大缩短了工作时间，提升了工作效率。

上述通过控制策略网络获取目标虚拟角色的目标姿态数据的过程可以在终端设备中实现，也可以在服务器中实现并由服务器将目标姿态数据发送至终端设备。

进一步地，第一关键帧与第二关键帧之间包括M个固定帧，则上述步骤利用N个关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧，包括：利用N个关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第一关键帧与第二关键帧之间的固定帧以及第二关键帧。

这种情况下，目标动画片段T1由至少第一关键帧、M个固定帧以及第二关键帧组成。

具体来说，终端设备获取目标姿态数据后，则可根据初始姿态数据，通过计算得到调整各个关节的力矩，并进一步得到第二关键帧，以及第一关键帧与第二关键帧之间的固定帧。其中，固定帧为两关键帧之间的动画帧，可以利用插值的方法确定固定帧，也可以为力矩调整目标虚拟角色的姿态而得到固定帧。

可以理解，在其他实施例中，可以将本申请各实施例中的“N”和“M”分别用“第一预设数量”和“第二预设数量”进行区分性表述。

在一个实施例中，调整目标虚拟角色的 N 个关节的力矩，以及固定帧、第二关键帧通过如下方式得到：

根据初始姿态数据和目标姿态数据，得到调整第一关键帧中目标虚拟角色的 N 个关节的力矩；

利用力矩调整目标虚拟角色的初始姿态，得到第一固定帧，第一固定帧是第一关键帧的下一帧动画；

针对除第一固定帧和第 M 固定帧之外的每一固定帧，获取上一固定帧中目标虚拟角色的姿态数据，利用上一帧中目标虚拟角色的姿态数据和目标姿态数据，得到调整固定帧中目标虚拟角色的 N 个关节的力矩；利用力矩调整固定帧中目标虚拟角色的姿态，得到下一固定帧；

获取第 M 固定帧中目标虚拟角色的姿态数据，利用第 M 固定帧中目标虚拟角色的姿态数据和目标姿态数据，得到调整第 M 固定帧中目标虚拟角色的 N 个关节的力矩；

利用力矩调整第 M 固定帧中目标虚拟角色的姿态，得到第二关键帧，第二关键帧是第 M 固定帧的下一帧动画。

具体实施过程中，采用了力矩来控制目标虚拟角色进行运动。这样可以实现较为真实的物理效果，且可以感知碰撞等相互作用。利用力矩对目标虚拟角色进行控制有多种方式，例如基于速度控制的方法，在该方法中，控制策略网络输出的信息为角速度，然后采用 PD 控制算法动态控制角速度信息，基于角速度信息计算出作用于目标虚拟角色各个关节的力矩，然后由物理引擎根据力矩控制目标虚拟角色进行运动。

本申请实施例中利用基于位置控制的方法。其对目标虚拟角色的控制效果较好，从而实现较为自然的动作效果。

为了便于理解，可以参考图 6，图 6 为本申请实施例中微分比例控制器的一个处理流程示意图，如图 6 所示，整个闭环控制系统相当于分层控制，将目标虚拟角色在第一关键帧的状态信息 S_t 输入至控制策略网络，由该控制

策略网络输出第二关键帧的目标姿态数据。这里计算力矩利用的目标姿态数据为各个关节的旋转数据，利用 PD 控制器动态地将目标虚拟角色的旋转数据控制到第一关键帧的姿态 A_t ，基于第一关键帧的姿态 A_t 计算得到力矩，在物理引擎中采用力矩对目标虚拟角色进行控制，由此得到第二关键帧的状态信息 S_{t+1} ，于是将作为控制策略网络的输入，从而得到第二关键帧的姿态 A_{t+1} ，依次类推，可以实现对目标角色的连续控制。

具体地，可以根据以下公式计算每个关节的力矩：

$$\tau^n = -k_p(q^n + \Delta t \dot{q}^n - \hat{q}^{n+1}) - k_d(\ddot{q}^n + \Delta t \dot{\ddot{q}}^n - \hat{\dot{q}}^{n+1}) \quad \dots\dots \text{公式 1}$$

其中， τ^n 表示第 n 帧的力矩； k_p 表示比例增益； k_d 表示微分增益； q^n 表示当前时刻的旋转数据， \dot{q}^n 表示当前时刻的角速度， q^n 和 \dot{q}^n 可以直接从物理引擎中获得； \ddot{q}^n 表示当前时刻的角加速度，通过逆动力学方式计算获得； \hat{q}^{n+1} 为下一时刻的旋转数据，即控制策略网络的输出； $\hat{\dot{q}}^{n+1}$ 表示下一时刻的角速度，这里总是置为 0。 Δt 表示周期时间，即当前时刻与下一时刻之间的时长。

通过上述方式，控制策略网络输出各个关节的旋转数据，再通过 PD 控制动态的将目标角色控制到对应的位置，相比力矩控制姿态更稳定，且控制策略网络输出每个关节的旋转数据，其分布方差较小，且采样样本小，因此控制策略网络收敛速度快。

以下详细介绍本申请实施例所采用的控制策略网络和价值评估网络的训练过程。

在本申请实施例中，针对同一类型的目标任务，可以训练一个控制策略网络，针对不同类型的目标任务，可以分别训练相应的控制策略网络。例如，“向前射击”、“向左射击”和“向右射击”均属于射击，仅是射击的方向不同，因此属于同一类型的目标任务，可以训练一个控制策略网络。而“射击”和“跑步”属于不同类型的目标任务，可以分别训练相应的控制策略网络。

本申请在一个实施例中，采用基于强化学习的 AC 算法框架进行训练，AC 算法框架包括 Actor 网络以及 Critic 网络，本申请实施例中的控制策略

网络即为 Actor 网络，价值评估网络即为 Critic 网络。其中，控制策略网络训练的是当前策略，输出姿态数据，而价值评估网络用于指导控制策略网络的学习。训练的时候需要对控制策略网络和价值评估网络均进行训练。其中，价值评估网络与控制策略网络的网络结构可以是一致的，也可以是不同的。本申请实施例中价值评估网络与控制策略网络采用相同的输入结构和输入数据，价值评估网络与控制策略网络的输出不同，控制策略网络的输出为目标姿态数据，而价值评估网络的输出为一维数据，表示的当前时刻的状态价值。

具体地，控制策略网络的训练过程可以如图 7 所示，包括如下步骤：

步骤 S701、将样本动画片段中训练虚拟对象在当前时刻的状态信息和设定的训练任务输入控制策略网络，得到控制策略网络输出的下一时刻的训练虚拟对象的姿态数据。

其中，每一时刻对应一关键帧动画。可以理解，本申请实施例中的控制策略网络，为待训练的（包括未训练和训练中）的控制策略网络。

当前时刻为当前关键帧的样本动画画面对应的播放时刻，下一时刻为下一关键帧样本动画画面对应的播放时刻。训练虚拟对象的状态信息包括当前相位数据、当前的初始姿态数据、当前速度数据和历史姿态序列，与上文中目标虚拟角色的状态信息相同，这里不做赘述。

训练虚拟对象是在训练过程中控制策略网络输出的控制策略的作用对象。训练虚拟对象和参考虚拟对象是两个骨架相同或相近的虚拟对象。骨架相同，可以表示为训练虚拟对象和参考虚拟对象的骨架中的所有骨骼的尺寸、形状，以及骨骼之间的连接关系完全相同；骨架相近，包括但不限于以下至少一种：训练虚拟对象和参考虚拟对象的骨架中的骨骼的尺寸相近、形状相近、骨骼之间的连接关系相近。在训练控制策略网络之前，需要先获取包含参考虚拟角色的姿态序列的样本动画片段。

训练任务对应于使用过程中的目标任务进行设定。示例性地，训练任务可以是使训练虚拟对象沿给定的方向前进，或者是让训练虚拟对象使用旋风

踢的动作，踢到指定的位置。

控制策略网络（actor 网络）可以是带目标的网络，包括输入层、隐藏层和输出层。其中，隐藏层中可以包括一层神经网络，也可以包括多层神经网络，可以根据实际情况进行设定，本申请实施例对此不作限定。隐藏层中的神经网络层可以是全连接层。例如，隐藏层中可以包括两层全连接层，其中，第一全连接层可以包括 1024 个神经元，第二全连接层可以包括 512 个神经元。当隐藏层包括两层或两层以上的神经网络层时，神经网络层之间的激活函数为 ReLU（Rectified Linear Unit, 线性整流函数）函数。

在一种实施例中，actor 网络的网络结构如图 8 所示，可以将训练虚拟对象在当前时刻的状态信息和设定的训练任务输入控制策略网络，得到控制策略网络输出的下一时刻的控制策略，该控制策略为下一时刻训练虚拟对象的姿态。其中，状态信息和训练任务都是一维信息，将状态信息和训练任务信息做拼接，共同输入到两层的全连接隐藏层。模型的输出为线性输出，假设角色姿态的分布满足高斯分布，模型输出高斯分布的均值，而高斯分布的方差则作为网络的超参数。

在另一种实施例中，如图 9 所示，可以获取训练虚拟对象所处的场景环境的环境信息，场景环境的环境信息可以是场景环境的地形图。将场景环境的环境信息、训练对象在当前时刻的状态信息和训练任务输入控制策略网络，得到控制策略网络输出的下一时刻的训练虚拟对象的姿态。其中状态信息和训练任务都是一维信息，地形图是二维数据，地形图使用三层卷积网络完成平面信息提取，将输出结果展平成一维后与状态信息和训练任务做拼接，共同输入到两层的全连接隐藏层。模型的输出为线性输出，假设角色姿态的分布满足高斯分布，模型输出高斯分布的均值，而高斯分布的方差则作为网络的超参数。

例如，对于一些视觉任务，需要自适应地形环境，控制策略网络还可以包括特征提取网络，由多层卷积网络和全连接层组成，用于从输入的场景环

境的环境信息中提取地形特征，将提取的地形特征与输入的训练对象的状态信息和训练任务进行合并，确定下一时刻训练虚拟对象的姿态数据。

步骤 S702、将训练虚拟对象在当前时刻的状态信息和设定的训练任务输入价值评估网络，得到价值评估网络输出的当前时刻的状态价值。

其中，价值评估网络是根据参考动画片段训练得到的。具体实施过程中，由价值评估网络输出的状态价值来衡量控制策略网络的输出姿态的好坏，即训练虚拟对象执行完当前的动作之后，状态信息变化为 s ，状态价值 $V(s)$ 评估当前状态的好坏，间接衡量动作的好坏，可以理解的是，状态价值 $V(s)$ 越大表示状态越好。Critic 网络输出的状态价值 $V(s)$ 的学习标准是由环境反馈的一系列奖励信息计算而来的，即得到多个时刻的奖励信息之后，可以通过时间差分学习 (temporal-difference learning) 估计出当前状态价值 $V(s)$ 。

时间差分学习是强化学习中的一个中心思想，类似蒙特卡洛方法，时间差分学习能够直接从经验中学习而不需要对于环境的完整知识。类似动态规划方法，时间差分学习能够在现有的估计结果上进行提升而不需要等待整个事件结束。

步骤 S703、根据状态价值调整控制策略网络的参数并对调整参数后的控制策略网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的控制策略网络。

在一个实施例中，电子设备可以根据状态价值、以及控制策略网络输出的下一时刻的训练虚拟对象的姿态数据，调整控制策略网络的参数并对调整参数后的控制策略网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的控制策略网络。

本申请实施例中，根据 Actor 网络 (即控制策略网络) 输出的姿态数据，以及 Critic 网络 (即价值评估网络) 输出的状态价值，对该 Actor 网络进行训练。从而 Actor 网络根据 Critic 网络得到学习标准，因此可以根据损失函数计算误差和梯度，并对 Actor 网络进行训练，最终将训练得到的 Actor 网

络作为已训练的控制策略网络。

本申请实施例中的价值评估网络 (Critic 网络) 与控制策略网络的网络结构可以是一致的, 也可以是不同的。价值评估网络用于对控制策略网络输出的控制策略进行评价, 确定训练虚拟对象模仿参考虚拟角色及完成训练任务的奖励值。价值评估网络也是经过训练得到的。价值评估网络的训练过程如图 10 所示, 包括如下步骤:

步骤 S1001、根据训练虚拟对象和参考动画片段中参考虚拟角色在下一时刻的状态信息及设定的训练任务, 确定训练虚拟对象当前时刻的即时奖励值。

其中, 当前时刻的即时奖励值包括任务目标奖励和模仿目标奖励两部分。模仿目标奖励用于激励训练虚拟对象的姿态与参考虚拟角色的姿态保持一致, 本质是在相同相位时, 比较参考虚拟角色的姿态和训练对象的姿态的相似程度, 两者越接近, 则模仿目标奖励越高; 反之, 模仿目标奖励越低。任务目标奖励根据训练虚拟对象完成训练任务的情况确定。

根据上述描述, 在一个实施例中, 根据训练虚拟对象在下一时刻的状态信息及设定的训练任务, 确定当前时刻的任务目标奖励, 用于评价目标任务的完成程度;

根据训练虚拟对象在下一时刻的状态信息与参考虚拟角色在下一时刻的状态信息, 确定当前时刻的模仿目标奖励, 用于评价训练虚拟对象与参考虚拟对象的相似程度, 进而判断训练虚拟对象的动作是否自然;

根据任务目标奖励和模仿目标奖励, 确定当前时刻的即时奖励值。

具体的, 将任务目标奖励和模仿目标奖励进行加权, 加权后的值作为即时奖励值 r_t , 具体如下公式所示:

$$r_t = w^I r_t^I + w^G r_t^G \quad \dots\dots\text{公式 2}$$

其中, r_t^I 表示时刻 t 对应的模仿目标奖励, w^I 为模仿目标奖励的权重, r_t^G 表示时刻 t 对应的任务目标奖励, w^G 为任务目标奖励的权重。 w^I 和 w^G 与价值

评估网络的网络参数相关。

在一个实施例中，可以将模仿目标奖励 r_t^I 细分为四个方面：姿态相似度、速度相似度、末端关节相似度、质心姿态相似度，四个部分加权后的值作为最后的即时奖励，具体的，模仿目标奖励 r_t^I 可以表示为：

$$r_t^I = w^p r_t^p + w^v r_t^v + w^e r_t^e + w^c r_t^c \quad \dots\dots\text{公式 3}$$

其中， w^p 为姿态相似度 r_t^p 的权重， w^v 为速度相似度 r_t^v 的权重， w^e 为末端关节相似度 r_t^e 的权重， w^c 为质心姿态相似度 r_t^c 的权重。

其中，姿态相似度 r_t^p 用于描述训练对象与参考虚拟角色的姿态相似度，各个关节的位置和旋转度之间的相似度，可以表示为：

$$r_t^p = \exp[-2(\sum_j ||q_t^{*j} - q_t^j||^2)] \quad \dots\dots\text{公式 4}$$

q_t^{*j} 表示在时刻 t ，参考虚拟角色的第 j 个关节的姿态数据，可以用四元数表示。 q_t^j 表示在时刻 t ，训练对象的第 j 个关节的姿态数据。

速度相似度 r_t^v 用于描述训练对象与参考虚拟角色的速度相似度，包括各个关节的角速度与线速度与目标姿态之间的相似度，可以表示为：

$$r_t^v = \exp[-0.1(\sum_j ||v_t^{*j} - v_t^j||^2)] \quad \dots\dots\text{公式 5}$$

v_t^{*j} 表示在时刻 t ，参考虚拟角色的第 j 个关节的速度数据。 v_t^j 表示在时刻 t ，训练对象的第 j 个关节的速度数据。

末端关节相似度 r_t^e 用于描述训练对象与参考虚拟角色末端关节姿态的相似度，包括四肢关节，可以表示为：

$$r_t^e = \exp[-40(\sum_e ||p_t^{*e} - p_t^e||^2)] \quad \dots\dots\text{公式 6}$$

p_t^{*e} 表示在时刻 t ，参考虚拟角色的第 e 个末端关节的姿态数据， p_t^e 表示在时刻 t ，训练对象的第 e 个末端关节的姿态数据。

质心姿态相似度 r_t^c 用于描述训练对象与参考虚拟角色的重心位置的相似度，可以表示为：

$$r_t^c = \exp[-10(\|p_t^{*c} - p_t^c\|^2)] \quad \dots\dots\text{公式 7}$$

p_t^{*c} 表示在时刻 t ，参考虚拟角色的重心在世界坐标系中的位置， p_t^c 表示在时刻 t ，训练对象的重心在世界坐标系中的位置。

当然，模仿目标奖励 r_t^l 还可以包括其他相似度，例如根关节相似度等。一般来说，相似度个数越多，计算结果越准确，但计算量会越多。

根据训练任务的类型不同，任务目标奖励的设置也可以不同。例如，如果训练任务是使训练对象模仿参考虚拟角色走路姿态，并在走路过程中完成转向的任务，设定一个针对走路方向的任务目标奖励，以鼓励训练对象朝指定方向以给定的速度前进。如果训练任务是让训练对象使用旋风踢的动作，踢到指定的位置。例如，在训练对象周围指定一个随机的目标球体，训练对象用旋风踢来踢到指定的目标球体。该训练任务的任务向量由两部分组成，一个是给定的目标球体的位置 p_t^{tar} ，可以由空间中一个三维向量表示，另一个是一个二值的标志 h ，表明目标在前一个时间周期是否被击中。

在踢中目标球体的训练任务中，任务目标奖励可以表示为：

$$r_t^G = \exp[-4\|p_t^{\text{tar}} - p_t^e\|^2] \quad \dots\dots\text{公式 8}$$

其中， p_t^{tar} 表示目标球体的位置， p_t^e 表示时刻 t 训练虚拟对象的末端关节的位置。该训练任务的目标是，不仅能够保证旋风踢姿态的同时，还能够准确踢到指定的目标，完成任务。

步骤 S1002、根据样本动画片段中每个时刻训练虚拟对象的即时奖励值和状态价值，确定训练虚拟对象的期望奖励值。

收集到一个样本动画片段后，使用 GAE (map autoencoder, 图自编码器)

算法实现训练虚拟对象的期望奖励值，具体的，训练虚拟对象的期望奖励值 G_t^λ 可以根据以下公式计算：

$$G_t^\lambda = (1 - \lambda) \sum_{n=1}^{\infty} \lambda^{n-1} G_t^{(n)} \dots \dots \text{公式 9}$$

其中， $G_t^{(n)} = R_{t+1} + \gamma R_{t+2} + \dots + \gamma^{n-1} R_{t+n} + \gamma^{n-1} V(S_{t+n}) \dots \dots \text{公式}$

10

上述公式 9 中， $G_t^{(n)}$ 为 t 时刻计算得出的收获值， λ 为 0 到 1 的参数； R_t 为时刻 t 的即时奖励值， γ 为衰减因子， $V(S_t)$ 为 t 时刻价值评估网络输出的状态价值，n 为样本动画片段中关键帧的数量。

进一步地，根据以下至少一项条件确定样本动画片段结束：

样本动画片段时长到达时长阈值；训练虚拟对象的姿态数据满足摔倒阈值；训练虚拟对象的姿态数据与参考虚拟对象的姿态数据之差大于差异阈值；训练虚拟对象的速度数据大于速度阈值。

具体来说，样本动画片段时长超过设定的时长阈值，则认为样本动画片段结束。或者，训练虚拟对象摔倒，则认为样本动画片段结束，其中，摔倒的定义为指定的关节接触到了地面。或者，样本动画片段与参考动画片段的差异过大，这里主要指根关节的旋转角度差异超过角度阈值，一般设置为 90 度。或者，训练虚拟对象的速度数据大于速度阈值。

根据上述方式确定样本动画片段结束，从而保证训练虚拟对象与参考虚拟对象的相似程度以及训练虚拟对象的真实程度，进而提高训练的准确性。

步骤 S1003、根据期望奖励值调整价值评估网络的参数并对调整参数后的价值评估网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的价值评估网络。

其中，训练结束条件可以是训练次数达到设定次数、连续 N 次训练得到的期望奖励值的变化幅度在设定幅度之内或期望奖励值达到设定阈值。

示例性地，可以采用 PPO (Proximal Policy Optimization, 近端策略优化) 算法、SAC (Soft Actor-Critic, 柔性致动/评价) 算法或 DDPG (Deep

Deterministic Policy Gradient, 深度确定性策略梯度) 算法等用于处理连续控制问题的深度强化学习算法对上述模型进行训练。

图 11 示出了上述方法的训练效果。以复现旋转踢动作为例, 在运行 15000 次迭代, 大概 24h 后, 模型收敛。图 11 中横坐标表示训练的迭代次数, 即训练次数, 纵坐标表示反馈的奖励值, 其中, 曲线 1001 为训练数据的每一条数据的平均即时奖励 (即, 图 11 中的 Train_Avg_Return_0), 可达到 0.78; 曲线 1002 为测试数据的每一条数据的平均即时奖励 (即, 图 11 中的 Test_Avg_Return_0), 可达到 0.82。

图 12 示出了本申请实施例中已训练的控制策略网络输出效果展示图, 图 12 中在某一时刻的动画画面中, 参考虚拟角色 1201 的姿态与目标虚拟角色 1202 的姿态基本一致, 目标虚拟角色很好的复现了参考动画片段中的参考虚拟角色。

本申请实施例利用控制策略网络得到第二关键帧中目标虚拟角色的目标姿态数据, 并通过目标虚拟角色的初始姿态数据和目标姿态数据, 得到调整目标虚拟角色的力矩, 利用力矩调整目标虚拟角色的初始姿态, 得到第一关键帧与第二关键帧之间的固定帧以及第二关键帧, 即可以根据第一关键帧和目标任务生成目标虚拟角色的姿态序列, 进而得到目标虚拟角色的动画片段, 从而缩短了工作人员的工作时间, 提升了工作效率。此外, 由于目标虚拟角色是利用力矩调整目标虚拟角色的姿态, 因此赋予了目标虚拟角色真实的物理属性, 并基于此来计算目标虚拟角色的运动得到目标虚拟角色的姿态, 因此可以得到更符合实际场景的目标姿态, 实现更逼真的动作效果。

以下通过具体实例说明本申请实施例提供的动画实现方法的实现过程。

终端设备中安装有游戏客户端, 客户端与服务器进行交互实现游戏角色实现旋风踢的动作, 并踢中目标球体。假设在游戏中用户通过控制按键输入旋风踢指令, 指示由其控制的游戏角色执行向目标球体踢的动作。

游戏客户端基于已有的动画片段 T0，获取动画片段 T0 包含的 A0 关键帧中的目标虚拟角色的状态信息，A0 关键帧可以是显示界面中正在显示的当前动画帧。将 A0 关键帧作为上一帧动画，将目标虚拟角色在 A0 关键帧中的状态信息和旋风踢球的目标任务向服务器发送。其中，旋风踢球的目标任务也可以是一个任务向量，该向量中包括目标球体的位置坐标。

服务器中存储有已训练的控制策略网络。服务器将 A0 关键帧中的状态信息和旋风踢球的目标任务输入已训练的控制策略网络，得到 A1 关键帧中目标虚拟角色的目标姿态数据，A1 关键帧为 A0 关键帧的下一关键帧。

服务器将 A1 关键帧中的目标虚拟角色的目标姿态数据发送回游戏客户端。

游戏客户端获取 A0 关键帧中目标虚拟角色的初始姿态数据，以及 A1 关键帧中目标虚拟角色的目标姿态数据。根据初始姿态数据和目标姿态数据，利用公式 1 计算作用于 A0 关键帧中目标虚拟角色的 N 个关节的力矩。

游戏客户端基于物理引擎将得到的力矩作用于 A0 关键帧中目标虚拟角色的 N 个关节，调整目标虚拟角色的姿态，得到 B0 固定帧。固定帧为 A0 关键帧与 A1 关键帧之间的动画帧，这里 A0 关键帧与 A1 关键帧之间设置有 20 个固定帧。

游戏客户端获取 B0 固定帧中目标虚拟角色的 B0 姿态数据。根据 B0 姿态数据和目标姿态数据，利用公式 1 计算作用于 B0 固定帧中目标虚拟角色的 N 个关节的力矩。

游戏客户端基于物理引擎将得到的力矩作用于 B0 固定帧中目标虚拟角色的 N 个关节，调整目标虚拟角色的姿态，得到 B1 固定帧。重复上述步骤，直至获取 B19 固定帧中目标虚拟角色的 B19 姿态数据。根据 B19 姿态数据和目标姿态数据，利用公式 1 计算作用于 B19 固定帧中目标虚拟角色的 N 个关节的力矩得到 A1 关键帧。

重复上述步骤，游戏客户端继续将 A1 关键帧作为上一帧动画，将目标虚

拟角色在 A1 关键帧中的状态信息和目标任务向服务器发送，并接收服务器发送的 A2 关键帧的目标姿态数据。游戏客户端基于物理引擎，得到 A1 关键帧与 A2 关键帧之间的固定帧，以及 A2 关键帧。

以此类推，可以生成多个动画帧，获得目标虚拟角色完成旋风踢球的目标任务的动画片段 T1。动画片段 T1 中包括上述的 A0 关键帧、B0 至 B19 固定帧、A1 关键帧、B20 至 B29 固定帧、A2 关键帧以及后续生成的多个动画帧。在该实施例中，控制策略网络是基于参考虚拟角色执行旋风踢球任务的样本动画片段进行训练得到的，因此，控制策略网络可以确定动画片段 T1 中所包含的动画帧的数量。

下述为本申请装置实施例，对于装置实施例中未详尽描述的细节，可以参考上述一一对应的方法实施例。

请参考图 13，其示出了本申请一个实施例提供的动画实现装置的结构方框图。该动画实现装置设置于电子设备中，该装置可以通过硬件或者软硬件的结合实现成为图 1 中终端设备 101 的全部或者一部分。该装置包括：动画处理单元 1301、姿态获取单元 1302、力矩获取单元 1303、姿态调整单元 1304、和动画生成单元 1305。

其中，动画处理单元 1301，获得目标虚拟角色的目标动画片段 T0，目标动画片段 T0 中包括第一关键帧，第一关键帧中包括目标虚拟角色的初始姿态数据。

姿态获取单元 1302，用于将初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据。

力矩获取单元 1303，用于根据目标虚拟角色的初始姿态数据和目标姿态数据，得到调整目标虚拟角色的 N 个关节的力矩；N 为大于或者等于 1 的正整数。

姿态调整单元 1304，用于利用 N 个关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧。

动画生成单元 1305，用于获取目标动画片段 T1，目标动画片段 T1 包括第一关键帧以及第二关键帧。

在在一个实施例中，第一关键帧与第二关键帧之间包括 M 个固定帧；

姿态调整单元 1304，用于利用 N 个关节的力矩，将目标虚拟角色由初始姿态调整至目标姿态，得到第一关键帧与第二关键帧之间的固定帧以及第二关键帧；目标动画片段 T1 由至少第一关键帧、M 个固定帧以及第二关键帧组成。

在在一个实施例中，姿态获取单元 1302，具体用于：

获取目标虚拟角色在第一关键帧中的状态信息；

将目标虚拟角色在第一关键帧中的状态信息和目标任务输入控制策略网络，获得控制策略网络输出的目标虚拟角色的目标姿态数据；控制策略网络是根据参考动画片段训练得到的，参考动画片段包含参考虚拟角色完成目标任务的参考姿态序列。

在在一个实施例中，姿态获取单元 1302，具体用于：

获取目标虚拟角色在第一关键帧中的状态信息和目标虚拟角色所处的场景环境的环境信息；

将目标虚拟角色在第一关键帧中的状态信息、目标任务和目标虚拟角色所处的场景环境的环境信息输入控制策略网络，获得控制策略网络输出的目标虚拟角色的目标姿态数据；控制策略网络是根据参考动画片段训练得到的，参考动画片段包含参考虚拟角色完成目标任务的参考姿态序列。

在在一个实施例中，状态信息包括目标虚拟角色的当前相位数据、当前的初始姿态数据、当前速度数据和历史姿态序列；其中，当前相位数据用于表征目标虚拟角色在第一关键帧中所处的阶段，当前的初始姿态数据用于表征目标虚拟角色当前的姿态，当前速度数据用于表征目标虚拟角色当前的速度状态，历史姿态序列用于表征目标虚拟角色在历史时间段内的姿态。

在在一个实施例中，装置还包括网络训练单元 1306，用于：

将样本动画片段中训练虚拟对象在当前时刻的状态信息和设定的训练任务输入控制策略网络，得到控制策略网络输出的下一时刻的训练虚拟对象的姿态数据；每一时刻对应一关键帧动画；

将训练虚拟对象在当前时刻的状态信息和设定的训练任务输入价值评估网络，得到价值评估网络输出的当前时刻的状态价值；价值评估网络是根据参考动画片段训练得到的；

根据状态价值调整控制策略网络的参数并对调整参数后的控制策略网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的控制策略网络。

在在一个实施例中，网络训练单元 1306，用于：

根据训练虚拟对象和参考动画片段中参考虚拟角色在下一时刻的状态信息及设定的训练任务，确定训练虚拟对象当前时刻的即时奖励值；

根据样本动画片段中每个时刻训练虚拟对象的即时奖励值和状态价值，确定训练虚拟对象的期望奖励值；

根据期望奖励值调整价值评估网络的参数并对调整参数后的价值评估网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的价值评估网络。

在在一个实施例中，网络训练单元 1306，用于：

根据训练虚拟对象在下一时刻的状态信息及设定的训练任务，确定当前时刻的任务目标奖励；

根据训练虚拟对象在下一时刻的状态信息与参考虚拟角色在下一时刻的状态信息，确定当前时刻的模仿目标奖励；

根据任务目标奖励和模仿目标奖励，确定当前时刻的即时奖励值。

在在一个实施例中，模仿目标奖励包括以下至少一项：姿态相似度、速度相似度、末端关节相似度、质心姿态相似度；

姿态相似度用于表征训练虚拟对象与参考虚拟角色的姿态数据的相似度；

速度相似度用于表征训练虚拟对象与参考虚拟角色的速度数据的相似度；末端关节相似度用于表征训练虚拟对象与参考虚拟角色的末端关节的姿态数据的相似度；质心姿态相似度用于表征训练虚拟对象与参考虚拟角色的重心位置的相似度。

在在一个实施例中，网络训练单元 1306，用于根据以下至少一项确定样本动画片段结束：

样本动画片段时长到达时长阈值；训练虚拟对象姿态数据满足摔倒阈值；训练虚拟对象姿态数据与参考虚拟对象的姿态数据之差大于差异阈值；训练虚拟对象的速度数据大于速度阈值。

在在一个实施例中，网络训练单元 1306，用于：

获取训练虚拟对象所处的场景环境的环境信息；

将环境信息、训练虚拟对象在当前时刻的状态信息和训练任务输入控制策略网络，得到控制策略网络输出的下一时刻的训练虚拟对象的姿态。

与上述方法实施例相对应地，本申请实施例还提供了一种电子设备。该电子设备可以是终端设备，如图 1 中所示的终端设备 101，也可以是智能手机、平板电脑，手提电脑或计算机等电子设备，该电子设备至少包括用于存储数据的存储器和用于数据处理的一个或多个处理器。其中，对于用于数据处理的一个或多个处理器而言，在执行处理时，可以采用微一个或多个处理器、CPU、GPU（Graphics Processing Unit，图形处理单元）、DSP 或 FPGA 实现。对于存储器来说，存储器中存储有操作指令，该操作指令可以为计算机可执行代码，通过该操作指令来实现上述本申请实施例的视频筛选方法的流程中的各个步骤。

图 14 为本申请实施例提供的一种电子设备的结构示意图；如图 14 所示，本申请实施例中该电子设备 140 包括：一个或多个处理器 141、显示器 142、

存储器 143、输入设备 146、总线 145 和通讯设备 144；该一个或多个处理器 141、存储器 143、输入设备 146、显示器 142 和通讯设备 144 均通过总线 145 连接，该总线 145 用于该一个或多个处理器 141、存储器 143、显示器 142、通讯设备 144 和输入设备 146 之间传输数据。

其中，存储器 143 可用于存储软件程序以及模块，如本申请实施例中的动画实现方法对应的程序指令/模块，处理器 141 通过运行存储在存储器 143 中的软件程序以及模块，从而执行电子设备 140 的各种功能应用以及数据处理，如本申请实施例提供的动画实现方法。存储器 143 可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个应用的应用程序等；存储数据区可存储根据电子设备 140 的使用所创建的数据（比如动画片段、控制策略网络）等。此外，存储器 143 可以包括高速随机存取存储器，还可以包括非易失性存储器，例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。

处理器 141 是电子设备 140 的控制中心，利用总线 145 以及各种接口和线路连接整个电子设备 140 的各个部分，通过运行或执行存储在存储器 143 内的软件程序和/或模块，以及调用存储在存储器 143 内的数据，执行电子设备 140 的各种功能和处理数据。在一个实施例中，处理器 141 可包括一个或多个处理单元，如 CPU、GPU（Graphics Processing Unit，图形处理单元）、数字处理单元等。

本申请实施例中，处理器 141 将生成的动画片段通过显示器 142 展示给用户。

处理器 141 还可以通过通讯设备 144 连接网络，如果电子设备是终端设备，则处理器 141 可以通过通讯设备 144 与游戏服务器之间传输数据。如果电子设备是游戏服务器，则处理器 141 可以通过通讯设备 144 与终端设备之间传输数据。

该输入设备 146 主要用于获得用户的输入操作，当该电子设备不同时，该输入设备 146 也可能不同。例如，当该电子设备为计算机时，该输入设备 146 可以为鼠标、键盘等输入设备；当该电子设备为智能手机、平板电脑等便携设备时，该输入设备 146 可以为触控屏。

本申请实施例还提供了一种计算机存储介质，该计算机存储介质中存储有计算机可执行指令，该计算机可执行指令用于实现本申请任一实施例的动画实现方法。

在一些可能的实施方式中，本申请提供的动画实现方法的各个方面还可以实现为一种程序产品的形式，其包括程序代码，当程序产品在计算机设备上运行时，程序代码用于使计算机设备执行本说明书上述描述的根据本申请各种示例性实施方式的动画实现方法的步骤，例如，计算机设备可以执行如图 3 所示的步骤 S301 至 S306 中的动画生成流程。

程序产品可以采用一个或多个可读介质的任意组合。可读介质可以是可读信号介质或者可读存储介质。可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件，或者任意以上的组合。可读存储介质的更具体的例子（非穷举的列表）包括：具有一个或多个导线的电连接、便携式盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、光纤、便携式紧凑盘只读存储器（CD-ROM）、光存储器件、磁存储器件、或者上述的任意合适的组合。

可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号，其中承载了可读程序代码。这种传播的数据信号可以采用多种形式，包括——但不限于——电磁信号、光信号或上述的任意合适的组合。可读信号介质还可以是可读存储介质以外的任何可读介质，该可读介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备和方法，

可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的，例如，单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，如：多个单元或组件可以结合，或可以集成到另一个系统，或一些特征可以忽略，或不执行。另外，所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口，设备或单元的间接耦合或通信连接，可以是电性的、机械的或其它形式的。

上述作为分离部件说明的单元可以是、或也可以不是物理上分开的，作为单元显示的部件可以是、或也可以不是物理单元，即可以位于一个地方，也可以分布到多个网络单元上；可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

另外，在本申请各实施例中的各功能单元可以全部集成在一个处理单元中，也可以是各单元分别单独作为一个单元，也可以两个或两个以上单元集成在一个单元中；上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

以上，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。

权利要求书

1、一种动画实现方法，由电子设备执行，其特征在于，包括：

获得第一目标动画片段，所述第一目标动画片段中包括第一关键帧，所述第一关键帧中包括目标虚拟角色的初始姿态数据；

将所述初始姿态数据以及设定的目标任务输入已训练的控制策略网络中，得到目标虚拟角色的目标姿态数据；

根据所述目标虚拟角色的初始姿态数据和目标姿态数据，调整所述目标虚拟角色的第一预设数量的关节的力矩；

利用所述第一预设数量的关节的力矩，将所述目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧；

获取第二目标动画片段，所述第二目标动画片段包括第一关键帧以及所述第二关键帧。

2、根据权利要求1所述的方法，其特征在于，所述第一关键帧与所述第二关键帧之间包括第二预设数量的固定帧；所述第二目标动画片段，包括第一关键帧、所述第二预设数量的固定帧以及所述第二关键帧；

所述利用所述第一预设数量的关节的力矩，将所述目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧，包括：

利用所述第一预设数量的关节的力矩，将所述目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧、以及所述第一关键帧与所述第二关键帧之间的固定帧。

3、根据权利要求1所述的方法，其特征在于，所述将所述初始姿态数据以及设定的目标任务输入已训练的控制策略网络中，得到目标虚拟角色的目标姿态数据，包括：

获取所述目标虚拟角色在所述第一关键帧中的状态信息；所述状态信息中包括所述初始姿态数据；

将所述目标虚拟角色在所述第一关键帧中的状态信息和所述目标任务输入控制策略网络，获得所述控制策略网络输出的所述目标虚拟角色的目标姿态数据；所述控制策略网络是根据参考动画片段训练得到的，所述参考动画片段包含参考虚拟角色完成所述目标任务的参考姿态序列。

4、根据权利要求1所述的方法，其特征在于，所述将所述初始姿态数据以及设定的目标任务输入已训练的控制策略网络中，得到目标虚拟角色的目标姿态数据，包括：

获取所述目标虚拟角色在所述第一关键帧中的状态信息和所述目标虚拟角色所处的场景环境的环境信息；所述状态信息中包括所述初始姿态数据；

将所述目标虚拟角色在所述第一关键帧中的状态信息、所述目标任务和所述目标虚拟角色所处的场景环境的环境信息输入控制策略网络，获得所述控制策略网络输出的所述目标虚拟角色的目标姿态数据。

5、根据权利要求3或4所述的方法，其特征在于，所述状态信息包括目标虚拟角色的当前相位数据、当前的所述初始姿态数据、当前速度数据和历史姿态序列；其中，所述当前相位数据用于表征所述目标虚拟角色在所述第一关键帧中所处的阶段，所述当前的所述初始姿态数据用于表征所述目标虚拟角色当前的姿态，所述当前速度数据用于表征所述目标虚拟角色当前的速度状态，所述历史姿态序列用于表征所述目标虚拟角色在历史时间段内的姿态。

6、根据权利要求3或4所述的方法，其特征在于，所述控制策略网络的训练步骤，包括：

将样本动画片段中训练虚拟对象在当前时刻的状态信息和设定的训练任务输入待训练的控制策略网络，得到所述控制策略网络输出的下一时刻的所述训练虚拟对象的姿态数据；每一时刻对应一关键帧动画；

将训练虚拟对象在当前时刻的状态信息和设定的训练任务输入价值评估网络，得到价值评估网络输出的当前时刻的状态价值；所述价值评估网络是

根据所述参考动画片段训练得到的；

根据所述状态价值、以及所述下一时刻的所述训练虚拟对象的姿态数据，调整所述控制策略网络的参数并对调整参数后的控制策略网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的控制策略网络。

7、根据权利要求6所述的方法，其特征在于，所述价值评估网络的训练步骤，包括：

根据所述训练虚拟对象和参考动画片段中参考虚拟角色在下一时刻的状态信息及设定的训练任务，确定所述训练虚拟对象当前时刻的即时奖励值；

根据样本动画片段中每个时刻所述训练虚拟对象的即时奖励值和状态价值，确定所述训练虚拟对象的期望奖励值；

根据所述期望奖励值调整所述价值评估网络的参数并对调整参数后的价值评估网络继续进行训练，直至达到设定的训练结束条件为止，得到已训练的价值评估网络。

8、根据权利要求7所述的方法，其特征在于，所述根据所述训练虚拟对象和参考动画片段中参考虚拟角色在下一时刻的状态信息及设定的训练任务，确定所述训练虚拟对象当前时刻的即时奖励值，包括：

根据所述训练虚拟对象在下一时刻的状态信息及设定的训练任务，确定当前时刻的任务目标奖励；

根据所述训练虚拟对象在下一时刻的状态信息与所述参考虚拟角色在下一时刻的状态信息，确定当前时刻的模仿目标奖励；

根据所述任务目标奖励和所述模仿目标奖励，确定当前时刻的即时奖励值。

9、根据权利要求8所述的方法，其特征在于，所述模仿目标奖励包括以下至少一项：姿态相似度、速度相似度、末端关节相似度、以及质心姿态相似度；

其中，所述姿态相似度用于表征所述训练虚拟对象与所述参考虚拟角色

的姿态数据的相似度；所述速度相似度用于表征所述训练虚拟对象与所述参考虚拟角色的速度数据的相似度；所述末端关节相似度用于表征所述训练虚拟对象与所述参考虚拟角色的末端关节的姿态数据的相似度；所述质心姿态相似度用于表征所述训练虚拟对象与所述参考虚拟角色的重心位置的相似度。

10、根据权利要求 7 所述的方法，其特征在于，根据以下至少一项确定所述样本动画片段结束：

所述样本动画片段时长到达时长阈值；

所述训练虚拟对象姿态数据满足摔倒阈值；

所述训练虚拟对象姿态数据与所述参考虚拟对象姿态数据之差大于差异阈值；

所述训练虚拟对象的速度数据大于速度阈值。

11、根据权利要求 7 所述的方法，其特征在于，所述将样本动画片段中训练虚拟对象在当前时刻的状态信息和设定的训练任务输入待训练的控制策略网络，得到所述控制策略网络输出的下一时刻的所述训练虚拟对象姿态数据，包括：

获取所述训练虚拟对象所处的场景环境的环境信息；

将所述环境信息、所述训练虚拟对象在当前时刻的状态信息和所述训练任务输入待训练的所述控制策略网络，得到控制策略网络输出的下一时刻的所述训练虚拟对象姿态数据。

12、一种动画实现装置，设置于电子设备中，其特征在于，所述装置包括：

动画处理单元，用于获得目标虚拟角色的第一目标动画片段，所述第一目标动画片段中包括第一关键帧，所述第一关键帧中包括目标虚拟角色的初始姿态数据；

姿态获取单元，用于将所述初始姿态数据以及设定的目标任务输入已训练的控制策略网络中得到目标虚拟角色的目标姿态数据；

力矩获取单元，用于根据所述目标虚拟角色的初始姿态数据和目标姿态数据，得到调整所述目标虚拟角色的第一预设数量的关节的力矩；

姿态调整单元，用于利用所述第一预设数量的关节的力矩，将所述目标虚拟角色由初始姿态调整至目标姿态，得到第二关键帧；

动画生成单元，用于获取第二目标动画片段，所述第二目标动画片段包括至少所述第一关键帧以及所述第二关键帧。

13、一个或多个计算机可读存储介质，所述计算机可读存储介质内存储有计算机可读指令，其特征在于：所述计算机可读指令被一个或多个处理器执行时，实现权利要求 1~11 任一项所述的方法。

14、一种电子设备，其特征在于，包括存储器和一个或多个处理器，所述存储器上存储有可在所述一个或多个处理器上运行的计算机可读指令，当所述计算机可读指令被所述一个或多个处理器执行时，使得所述一个或多个处理器实现权利要求 1~11 任一项所述的方法。

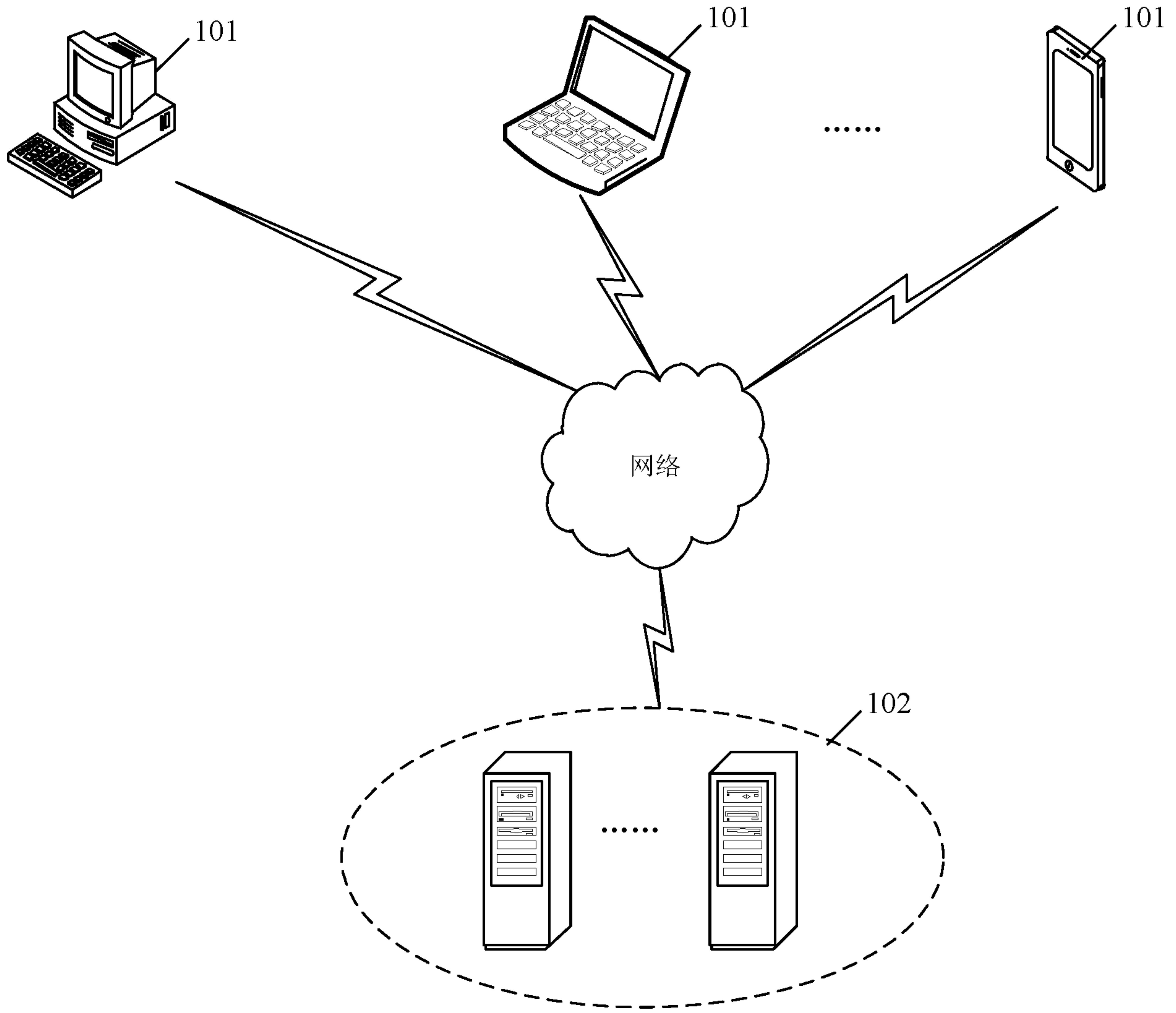


图 1

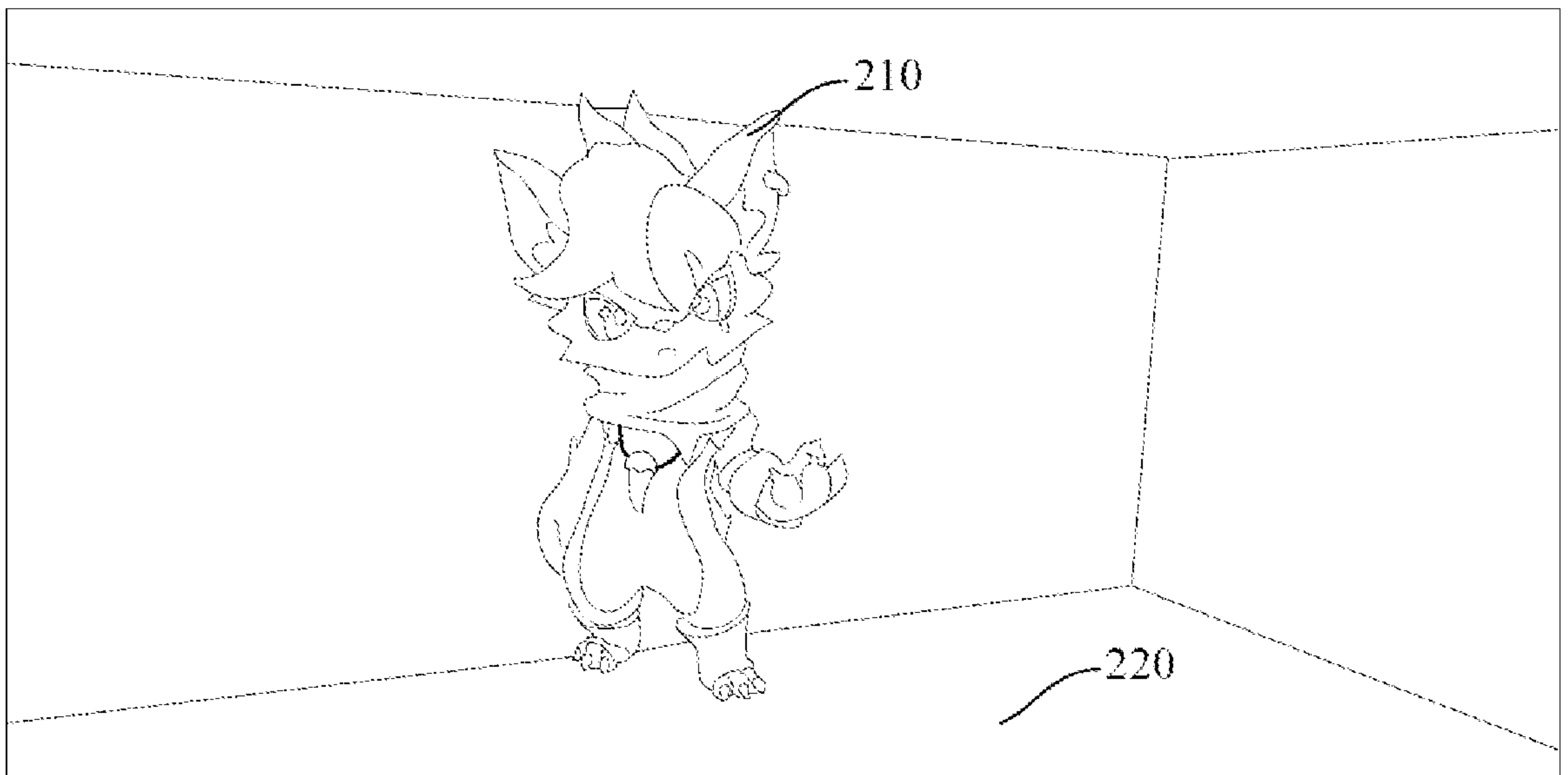


图 2

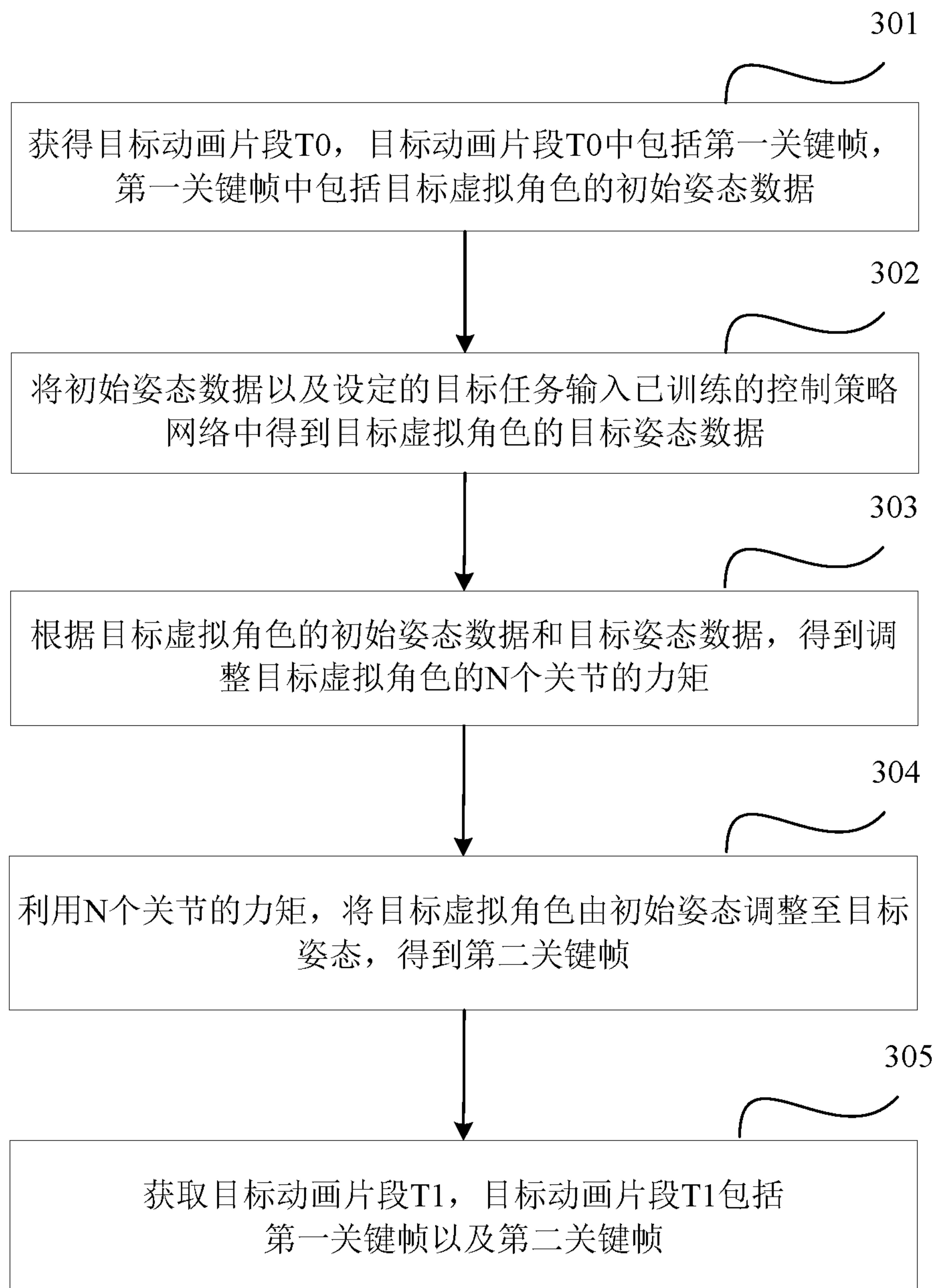


图 3

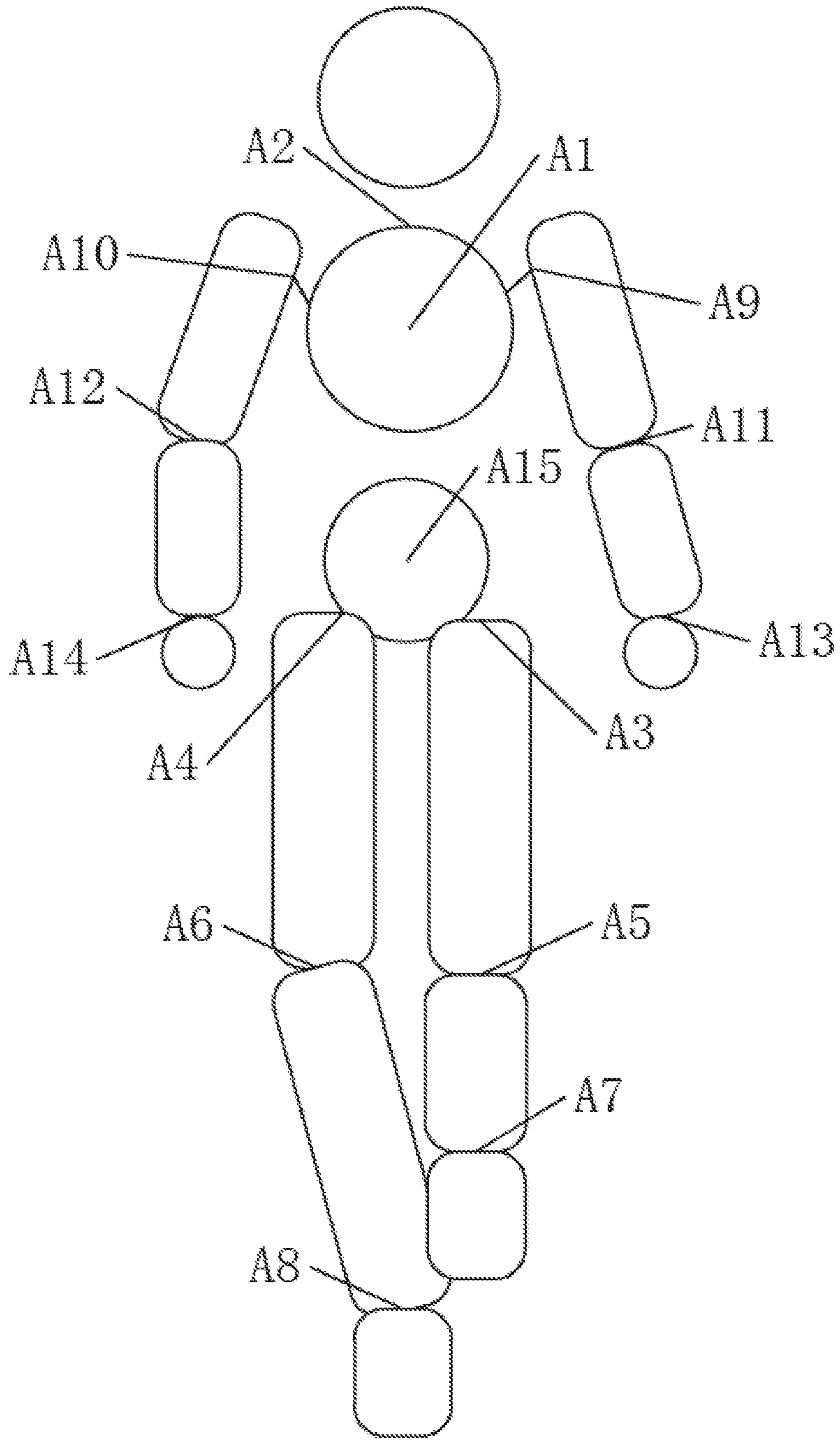


图 4

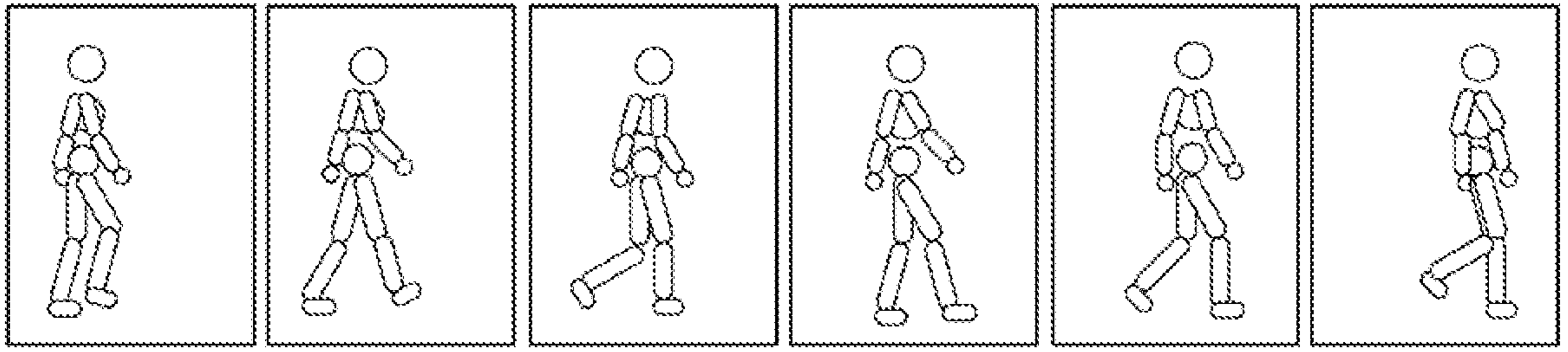


图 5

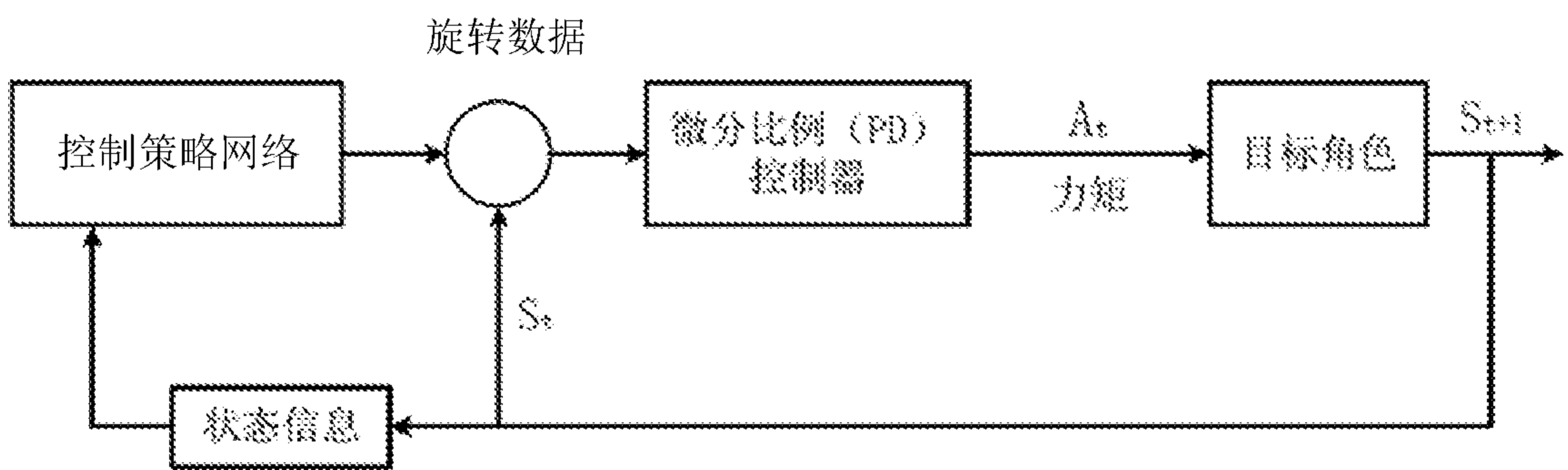


图 6

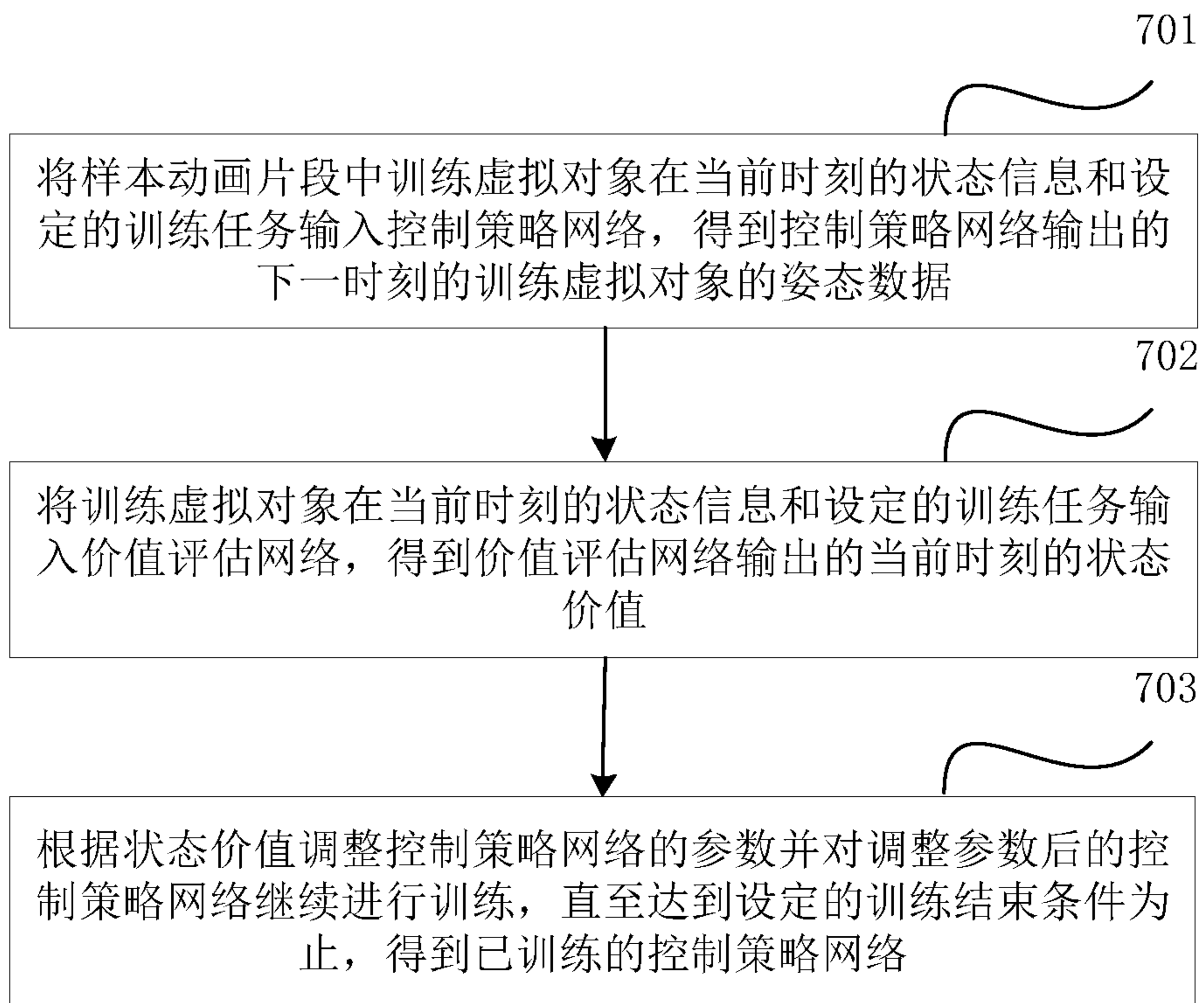


图 7

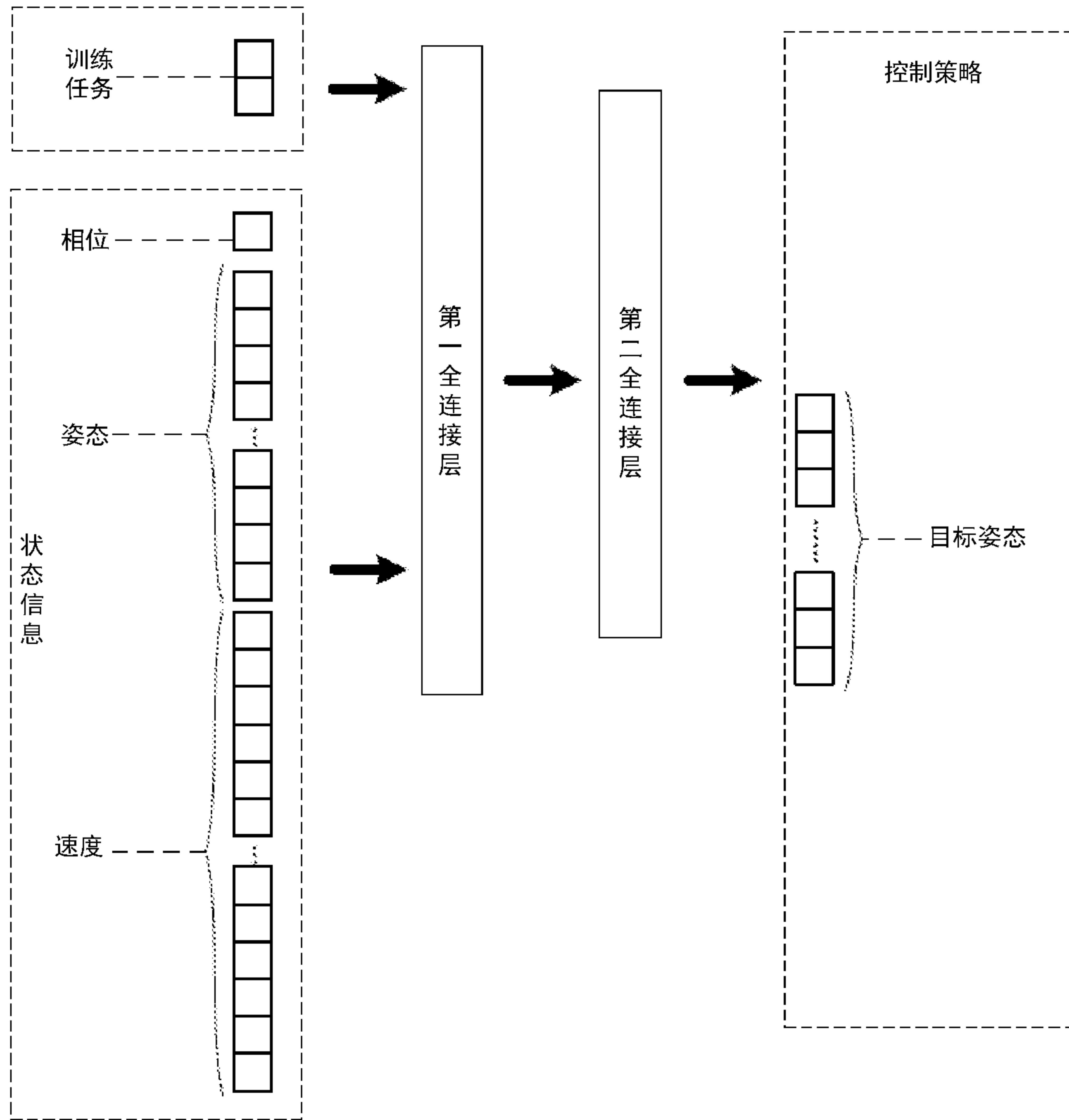


图 8

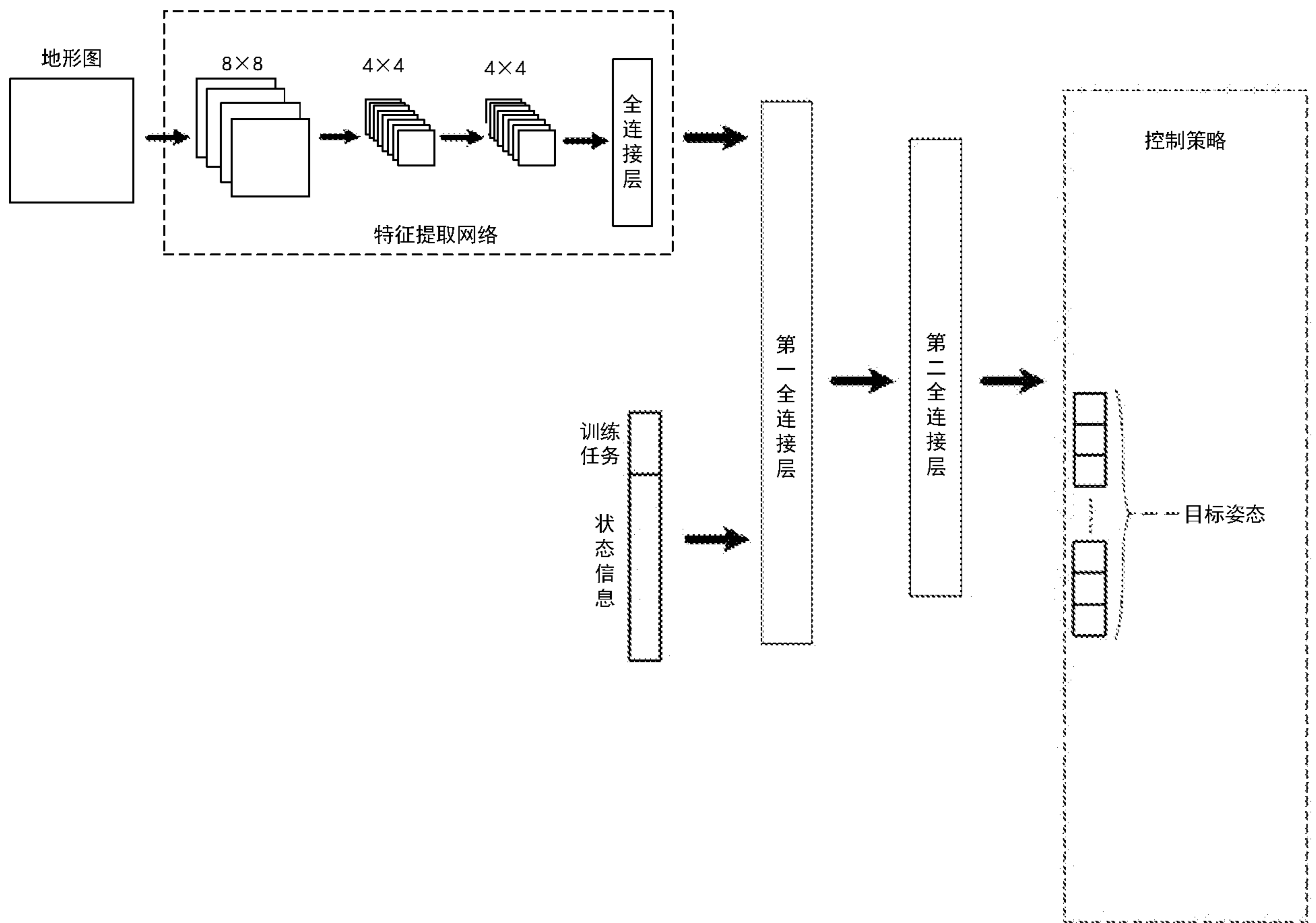


图 9

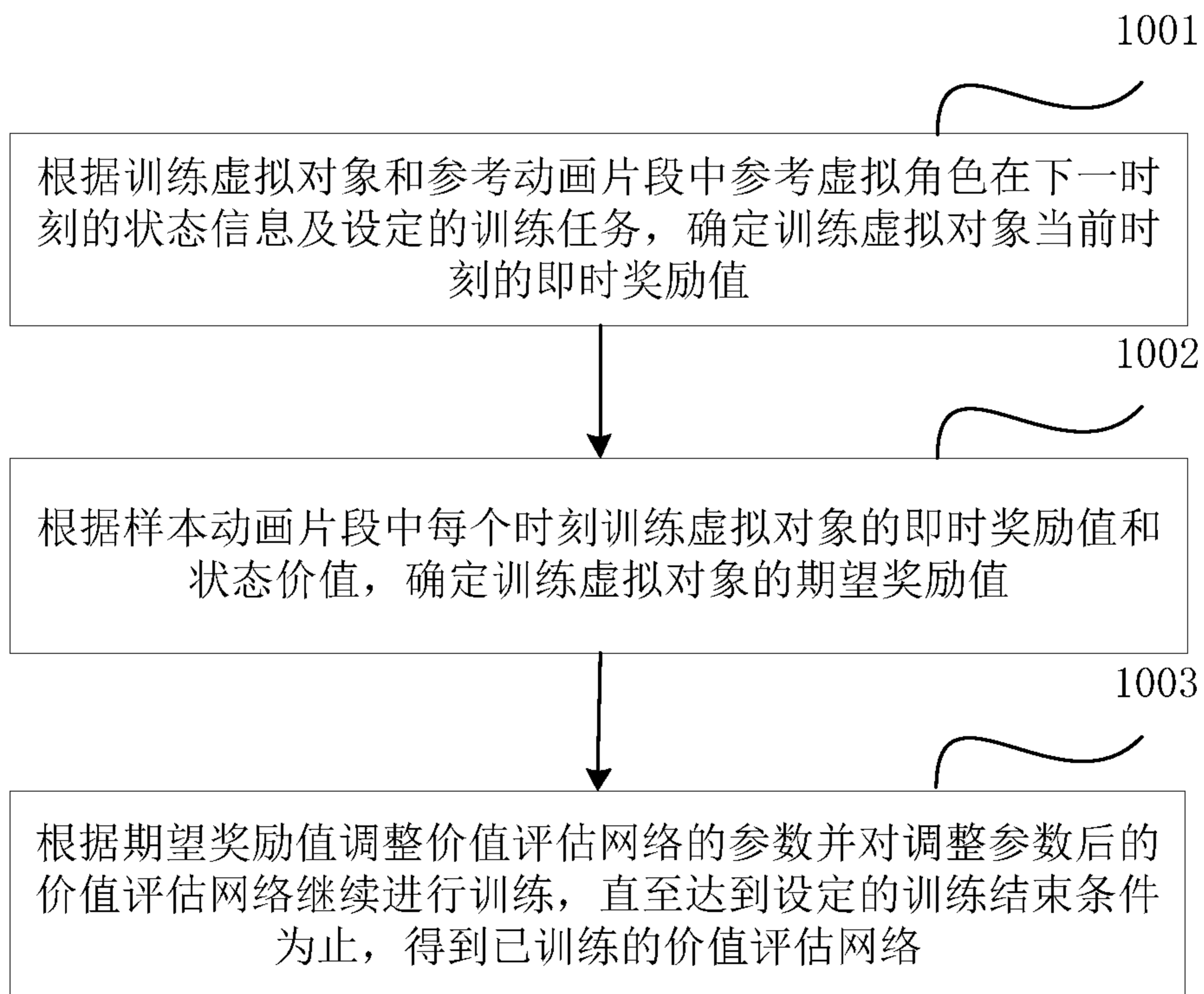


图 10

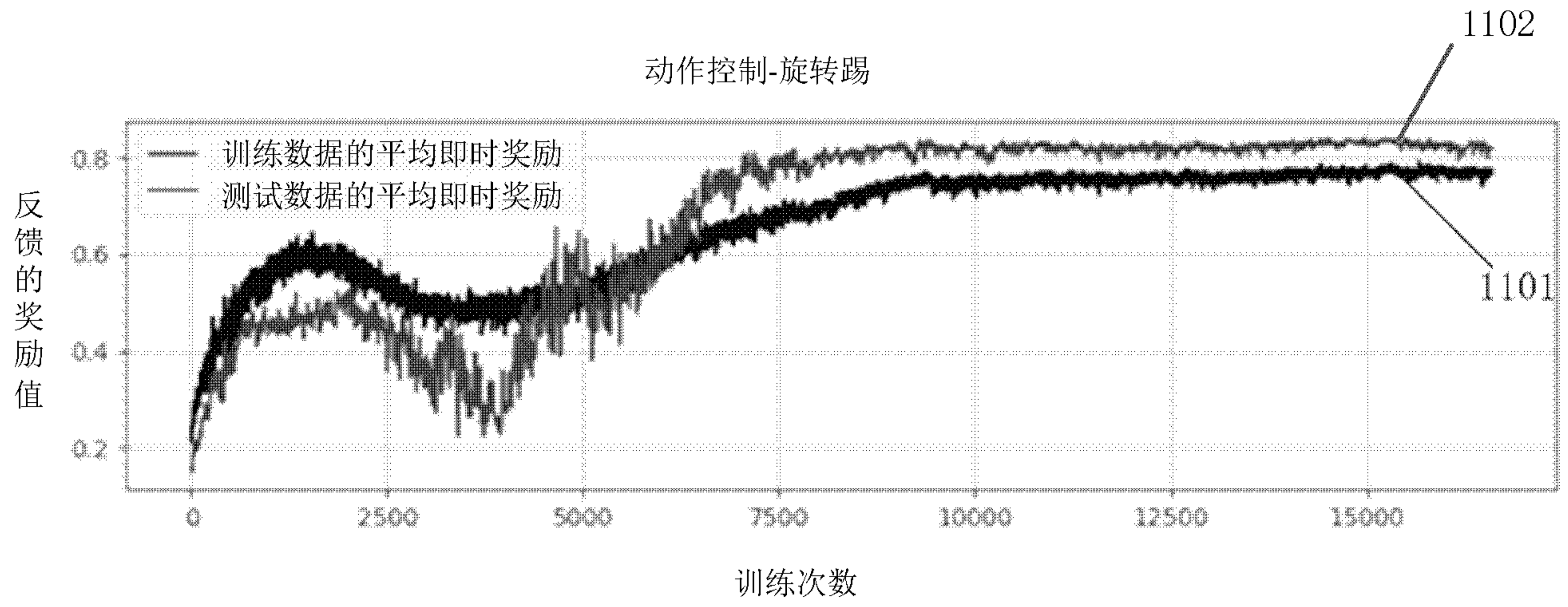


图 11

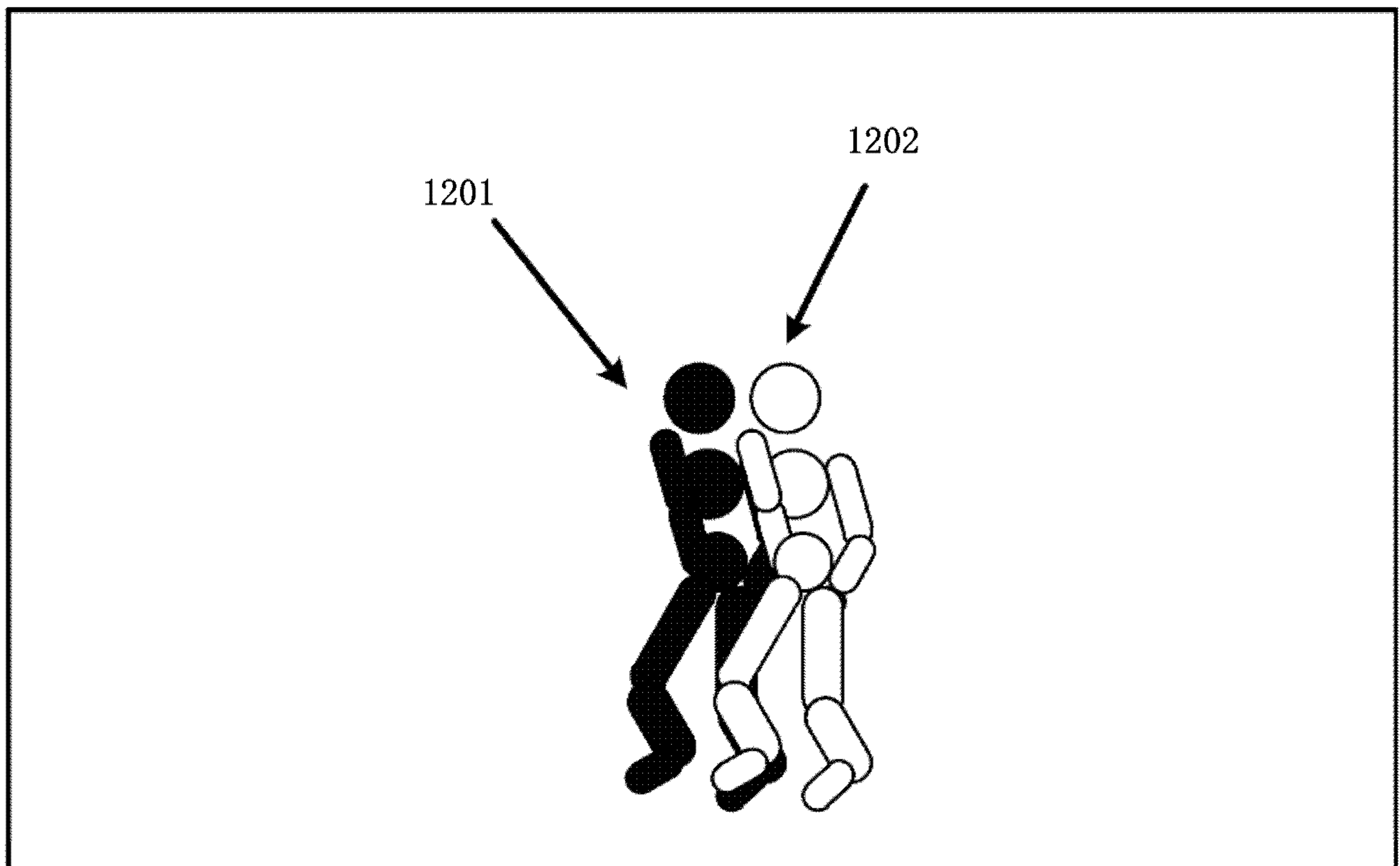


图 12

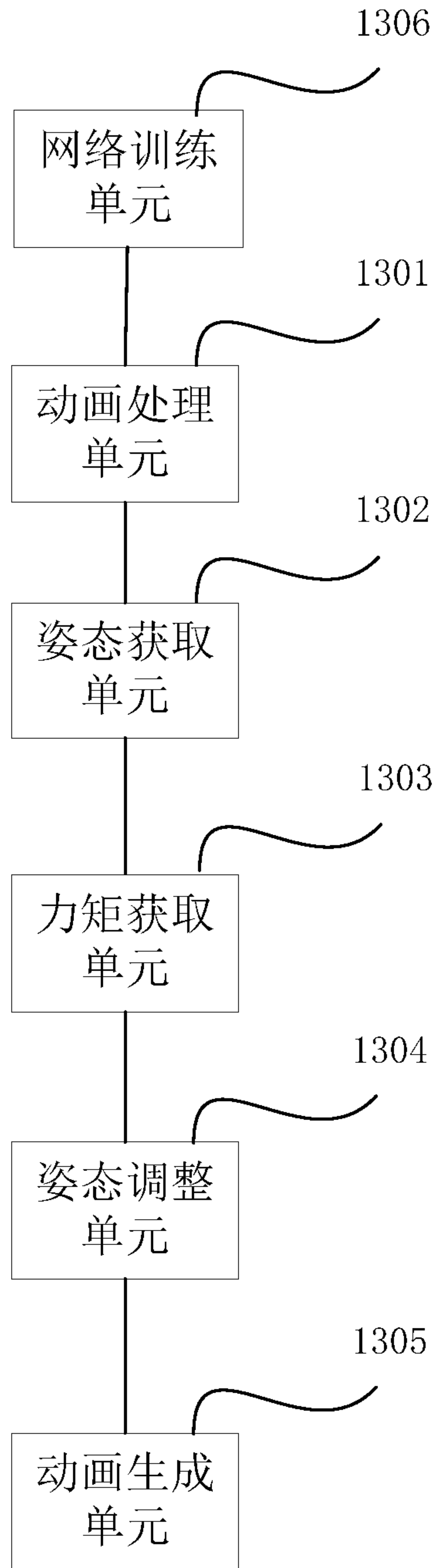


图 13

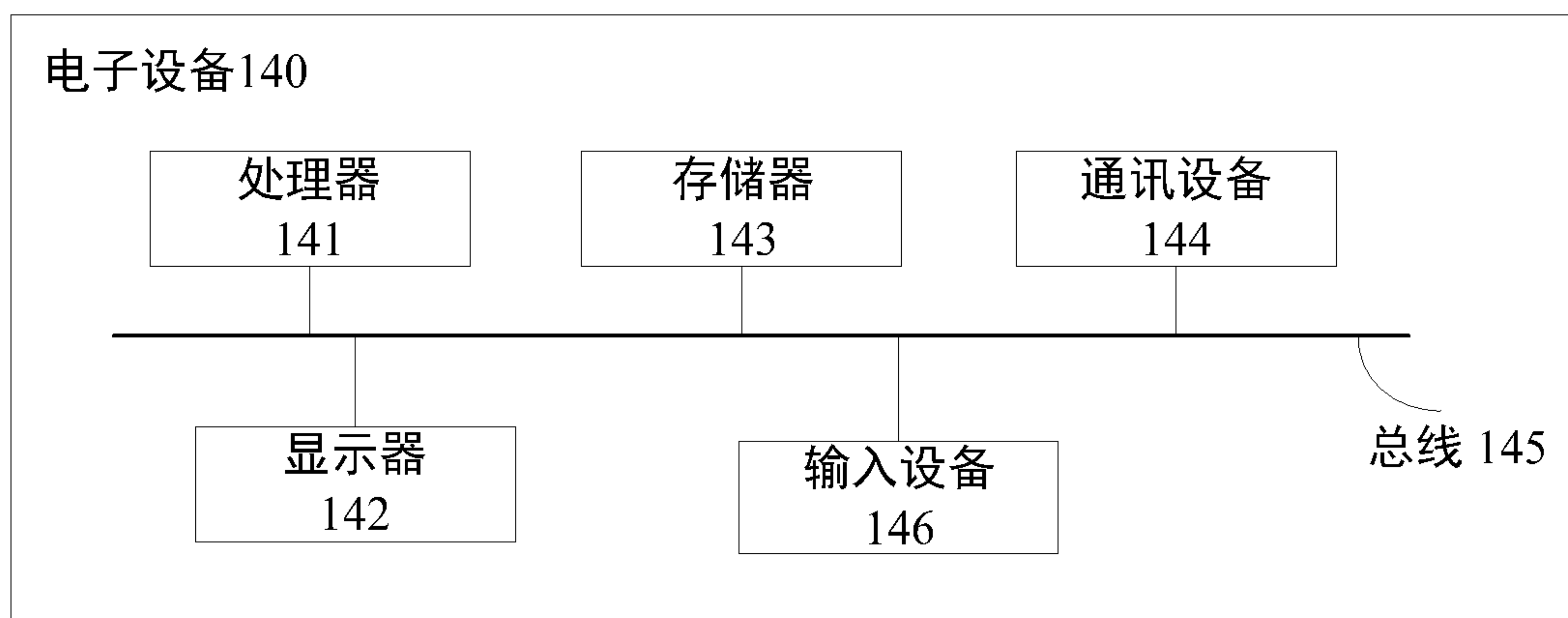


图 14

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/123677

A. CLASSIFICATION OF SUBJECT MATTER G06T 13/40(2011.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06T; G06F Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNPAT, CNKI, WPI, EPODOC, IEEE: 当前, 初始, 动画, 目标, 下一, 帧, 生成, 实现, 虚拟, 力矩, 关节, 网络, current, initial, animation, target, next, frame, generat+, realize+, virtual, moment, joint, network		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111260762 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 09 June 2020 (2020-06-09) claims 1-14	1-14
Y	CN 108182719 A (BEIJING DREAMWID TECH CO., LTD.) 19 June 2018 (2018-06-19) description, paragraphs [0042]-[0150]	1-14
Y	CN 106600668 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 26 April 2017 (2017-04-26) description, paragraphs [0061]-[0107]	1-14
Y	CN 110516389 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 29 November 2019 (2019-11-29) description, paragraphs [0041]-[0142]	3, 6-11
A	CN 109471712 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 15 March 2019 (2019-03-15) entire document	1-14
A	CN 102855470 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCES) 02 January 2013 (2013-01-02) entire document	1-14
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 24 December 2020		Date of mailing of the international search report 27 January 2021
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088 China		Authorized officer
Facsimile No. (86-10)62019451		Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/123677**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 101883287 A (GRADUATE SCHOOL AT SHENZHEN, TSINGHUA UNIVERSITY) 10 November 2010 (2010-11-10) entire document	1-14
A	CN 108022286 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 11 May 2018 (2018-05-11) entire document	1-14

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/CN2020/123677

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111260762	A	09 June 2020	None	
CN	108182719	A	19 June 2018	None	
CN	106600668	A	26 April 2017	None	
CN	110516389	A	29 November 2019	None	
CN	109471712	A	15 March 2019	WO 2020103723 A1	28 May 2020
CN	102855470	A	02 January 2013	None	
CN	101883287	A	10 November 2010	None	
CN	108022286	A	11 May 2018	WO 2019105396 A1	06 June 2019

国际检索报告

国际申请号

PCT/CN2020/123677

<p>A. 主题的分类</p> <p>G06T 13/40(2011.01) i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																													
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06T; G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNPAT, CNKI, WPI, EPODOC, IEEE: 当前, 初始, 动画, 目标, 下一, 帧, 生成, 实现, 虚拟, 力矩, 关节, 网络, current, initial, animation, target, next, frame, generat+, realize+, virtual, moment, joint, network</p>																													
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>PX</td> <td>CN 111260762 A (腾讯科技深圳有限公司) 2020年 6月 9日 (2020 - 06 - 09) 权利要求1-14</td> <td>1-14</td> </tr> <tr> <td>Y</td> <td>CN 108182719 A (北京聚力维度科技有限公司) 2018年 6月 19日 (2018 - 06 - 19) 说明书第[0042]-[0150]段</td> <td>1-14</td> </tr> <tr> <td>Y</td> <td>CN 106600668 A (中国科学院自动化研究所) 2017年 4月 26日 (2017 - 04 - 26) 说明书第[0061]-[0107]段</td> <td>1-14</td> </tr> <tr> <td>Y</td> <td>CN 110516389 A (腾讯科技深圳有限公司) 2019年 11月 29日 (2019 - 11 - 29) 说明书第[0041]-[0142]段</td> <td>3、6-11</td> </tr> <tr> <td>A</td> <td>CN 109471712 A (腾讯科技深圳有限公司) 2019年 3月 15日 (2019 - 03 - 15) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 102855470 A (中国科学院自动化研究所) 2013年 1月 2日 (2013 - 01 - 02) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 101883287 A (清华大学深圳研究生院) 2010年 11月 10日 (2010 - 11 - 10) 全文</td> <td>1-14</td> </tr> <tr> <td>A</td> <td>CN 108022286 A (腾讯科技深圳有限公司) 2018年 5月 11日 (2018 - 05 - 11) 全文</td> <td>1-14</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	PX	CN 111260762 A (腾讯科技深圳有限公司) 2020年 6月 9日 (2020 - 06 - 09) 权利要求1-14	1-14	Y	CN 108182719 A (北京聚力维度科技有限公司) 2018年 6月 19日 (2018 - 06 - 19) 说明书第[0042]-[0150]段	1-14	Y	CN 106600668 A (中国科学院自动化研究所) 2017年 4月 26日 (2017 - 04 - 26) 说明书第[0061]-[0107]段	1-14	Y	CN 110516389 A (腾讯科技深圳有限公司) 2019年 11月 29日 (2019 - 11 - 29) 说明书第[0041]-[0142]段	3、6-11	A	CN 109471712 A (腾讯科技深圳有限公司) 2019年 3月 15日 (2019 - 03 - 15) 全文	1-14	A	CN 102855470 A (中国科学院自动化研究所) 2013年 1月 2日 (2013 - 01 - 02) 全文	1-14	A	CN 101883287 A (清华大学深圳研究生院) 2010年 11月 10日 (2010 - 11 - 10) 全文	1-14	A	CN 108022286 A (腾讯科技深圳有限公司) 2018年 5月 11日 (2018 - 05 - 11) 全文	1-14
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																											
PX	CN 111260762 A (腾讯科技深圳有限公司) 2020年 6月 9日 (2020 - 06 - 09) 权利要求1-14	1-14																											
Y	CN 108182719 A (北京聚力维度科技有限公司) 2018年 6月 19日 (2018 - 06 - 19) 说明书第[0042]-[0150]段	1-14																											
Y	CN 106600668 A (中国科学院自动化研究所) 2017年 4月 26日 (2017 - 04 - 26) 说明书第[0061]-[0107]段	1-14																											
Y	CN 110516389 A (腾讯科技深圳有限公司) 2019年 11月 29日 (2019 - 11 - 29) 说明书第[0041]-[0142]段	3、6-11																											
A	CN 109471712 A (腾讯科技深圳有限公司) 2019年 3月 15日 (2019 - 03 - 15) 全文	1-14																											
A	CN 102855470 A (中国科学院自动化研究所) 2013年 1月 2日 (2013 - 01 - 02) 全文	1-14																											
A	CN 101883287 A (清华大学深圳研究生院) 2010年 11月 10日 (2010 - 11 - 10) 全文	1-14																											
A	CN 108022286 A (腾讯科技深圳有限公司) 2018年 5月 11日 (2018 - 05 - 11) 全文	1-14																											
<p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>																													
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</p> <p>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了解发明之理论或原理的在后文件</p> <p>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</p> <p>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</p> <p>“&” 同族专利的文件</p>																													
<p>国际检索实际完成的日期</p> <p>2020年 12月 24日</p>		<p>国际检索报告邮寄日期</p> <p>2021年 1月 27日</p>																											
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>张千</p> <p>电话号码 86-(10)-53961316</p>																											

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2020/123677

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	111260762	A	2020年 6月 9日	无			
CN	108182719	A	2018年 6月 19日	无			
CN	106600668	A	2017年 4月 26日	无			
CN	110516389	A	2019年 11月 29日	无			
CN	109471712	A	2019年 3月 15日	WO	2020103723	A1	2020年 5月 28日
CN	102855470	A	2013年 1月 2日	无			
CN	101883287	A	2010年 11月 10日	无			
CN	108022286	A	2018年 5月 11日	WO	2019105396	A1	2019年 6月 6日