姿势度量：角色动作编辑的新范式

Léon Victor里昂大学、里昂国立应用科学学院、LIRIS、法国国家科学研究中心联合研究单位 5205，法国

Alexandre Meyeralexandre.meyer@univ-lyon1.fr 里昂大学、克洛德·贝尔纳里昂第一大学、LIRIS、法国国家科学研究中心联合研究单位 5205，法国

Saïda Bouakaz里昂大学、克洛德·贝尔纳里昂第一大学、LIRIS、法国国家科学研究中心联合研究单位 5205，法国

arXiv:2301.06514v1 [cs.GR] 16 Jan 2023

摘要

在动画中，风格可以被视为动作内容之上的一个独特层次，使角色能够以不同方式完成同一姿势。在保留内容不变的同时，通过编辑现有动画来改变其风格，是一项颇具意义的任务，有助于重复利用动画数据并缩短制作时间。现有的动画编辑方法要么直接处理动作数据，提供精确但繁琐的工具；要么操纵语义风格类别，使用户失去控制权。作为折中方案，我们提出一种全新的角色动作编辑范式，能够进行更高层次的操控，同时不牺牲可控性。我们介绍姿势度量的概念，即可用于编辑动画的目标值函数，并将风格的诠释留给用户。随后，我们提出一种利用姿势度量编辑动画数据的编辑流程。

CCS 概念：• **计算方法** → 动作处理；机器学习。

关键词：角色动画、动画编辑、动画风格、神经网络

1 引言

虚拟角色遍布计算机辅助媒体的各个领域；它们充实虚拟世界、化身为玩家，并讲述故事。为它们制作高质量的动作，对于打造引人入胜的动画、激发观众兴趣至关重要。角色动画软件提供了帮助艺术家实现这些目标的工具和界面。多年来，人们开发了许多此类工具，不断寻求精确度、易用性和可控性之间的平衡。 [6, 25]

人眼通过动作感知大量信息：动作是一种复杂的行为，角色借此传达意图、情绪或个性。同一种姿势可以有多种多样的表现方式，即使由同一个角色来完成也是如此。因此，编辑现有动画、改变其表现方式而不改变姿势本身，是一个值得关注的研究课题。人们已提出许多技术来解决动画风格编辑问题，但其中大多数都从最高层级入手，缺乏用户所寻求的精细控制。

本文提出一种用于角色动作编辑的新框架，以姿势度量作为原始骨骼操控与高层级“风格”标签之间的折中方案。我们介绍这些中间层级描述符，并提出一种基于神经网络的编辑流程，使动画师能够通过操控度量指标来编辑角色的动作。

2 相关工作

传统上，虚拟角色由网格表示，网格上应用了各种纹理和材质。角色的动作由底层骨架驱动，网格则绑定（“蒙皮”）在骨架上。动画师通过编辑特定时间点（称为关键帧）上的骨架姿势来设计角色的动作。除了通过这种方式从头开始制作新的动画外，也可以借助专用的动作捕捉（MoCap）设备捕捉现实演员的动作。动作捕捉数据是获取逼真动作的良好来源，但生成的动画通常仍需清理和编辑，以符合动画师的构想。

以往关于角色动画的研究探讨工具、算法和界面，旨在简化动作创作与编辑过程。动作捕捉系统所获取的自然动作日益普及、质量不断提高，推动了人们对数据导向方法的兴趣。研究者可以从真实标注数据中提取有意义的信息，并将其用于开发更高效的工具，帮助动画师工作。

2.1 动画编辑

无论动画数据最初是由动画师通过关键帧制作，还是在动作捕捉环境中记录的，都可能因各种原因而需要修改：既可能是为了调整细节以符合艺术家的构想，也可能是为了将其重新用于不同的角色。本文建立在大量相关研究的基础之上，这些研究的共同目标是为动画师提供工具，简化编辑流程。常见的动画编辑流程可分为两个阶段：姿势编辑和插值控制。第一阶段中，动画师操控骨架，在指定的关键帧上定义关键姿势。第二阶段中，他们调整用于填补关键帧之间空隙的插值过程。

姿势编辑。 在大多数姿势编辑技术中，骨架被视为一组运动学链 [26]。通过正向运动学，

（FK），修改关节的位置或旋转会沿链向下传递，进而改变链上的其他关节。与之相反的技术——逆向运动学（IK） [3, 34], 会根据目标末端执行器，为运动学链找到合适的参数化方式，让用户能够像操纵木偶一样编辑角色的姿势。这两种方法是大多数现代动画软件的核心，能产生精确的结果，但从艺术家的角度来看，使用起来可能很繁琐。为缓解这一问题，人们还研究了其他技术和界面。

有些方法提出了另一种框架：角色的姿势由用户绘制的一条线生成，类似于画笔笔触 [9, 17]。这种方法也被扩展到虚拟现实中，通过追踪用户的手在三维空间中的运动，并将其动作解读为笔触 [16]。在数据驱动方面，Wu 等人 [37] 运用学习式聚类技术，从运动数据库中提取与用户输入最为匹配的姿势。Victor 等人 [32] 训练深度神经网络来解决逆向运动学问题，从现有数据中学习骨架约束，从而减轻用户遵守这些约束的负担。最后，还有一些研究致力于设计物理设备，以简化姿势编辑过程 [22, 39]。

插值。设定关键姿势后，可以通过不同方法控制软件如何在它们之间进行插值，以生成动作。动画师通常通过调整控制点来编辑动画曲线，从而指定关键帧之间经过的时间内各关节位置和旋转值所使用的插值因子 [7]。同样，也有人提出基于优化的方法，以改进曲线编辑流程 [10, 24]。在数据驱动的解决方案中，Chai 等人 [8] 提出学习低维动作空间，并在其中求解同类优化问题。最近，Harvey等人 [18] 训练深度神经网络，使其学会根据关键帧生成中间帧动作。

创作高质量、引人入胜的动作是一项复杂的任务。要按照自己的设想为角色制作动画，同时呈现预期动作并传达恰当的情感，需要深入了解解剖学和物理规律。前面提到的所有工作都能帮助动画师编辑动画，但它们都侧重于在骨骼层级控制角色的身体。虽然制作过程因此变得更容易，但仍然很复杂。为了让动画师和广大公众能够使用快速且富有表现力的动画工具，我们需要更高层次的概念与工具。

2.2 动作风格操控与迁移

出于复用现有动画方法的需求，另一类研究开始关注通过操控动作风格来编辑动画。相比“内容”这一概念，人们更常使用风格 这一概念。在给定一个动画序列时，这些术语用于区分角色 在做什么 （内容）与如何 做（风格）。例如，多个动画序列可以表现角色以各种方式行走：快或慢、步幅较短或较大、摆动手臂……在这种情况下，它们共享相同的内容——行走，但不同的风格 使它们彼此有所区别。

因此，能够生成风格化动画并修改现有动画的风格，是一个值得追求的目标。这样的控制方式可以让用户轻松尝试不同风格，也能够基于现有序列生成丰富多样的变体。

早期针对风格编辑问题的解决方案，会手动标注风格与内容，以此对现有动作进行分类，并在不同类别之间插值来生成新动作。Rose等人 [30] 采用了这种方法，并结合一种使用动词与副词概念来设定生成动作参数的界面。Brand等人 [5] 介绍了一种学习型状态机，其中的动作序列按共享的风格特征进行组织。近期，Rodriguez等人 [29] 则不再沿用通常的风格范式，转而提出通过高层轨迹参数化来编辑动画。

还有一些方法借鉴了相关领域的研究成果，例如舞蹈理论，以对动作风格进行分类。Durupinar 等人 [12] 基于拉班动作分析[33] ，尝试建立个性与身体动作之间的联系。Aristidou 等人 [4] 将拉班动作特征映射到一套情绪坐标系统中，用户可以通过插值来编辑现有的动画风格。

受图像风格迁移方法所取得成功的启发，Holden 等人 [20] 训练了一个深度卷积神经网络，从动画中提取“风格”特征，并将其应用于另一段动画。同样，研究者也使用了其他类型的神经网络，例如变分自编码器 [11], 、对抗学习 [35] 或生成流 [36]。

与上文讨论的姿势编辑方法处于光谱另一端的这些风格编辑方法，有一个共同的缺点：层次过于 高。由于动画风格的概念仍未得到明确界定（下一节将进一步讨论这一不足），面向广大用户时，将风格用作控制参数的工具可能会让用户感到难以理解。

3 问题表述

先前将风格 定义为与动画的内容 相分离的用法，在计算机动画文献中已普遍被接受。然而，该术语本身所涵盖的范围尚未明确界定。

有些人将动画的风格与角色体验到的情绪联系起来，例如“喜悦”“愤怒”或“悲伤”[15]。有些人则将其与现有的动作分析系统联系起来，例如拉班的 [12, 33], ，或与心理学概念联系起来，例如

唤醒度/效价量表 [13, 28]。还有些人采用任意分类，例如“猩猩”“幼稚”“僵尸”或“中性”[38]。

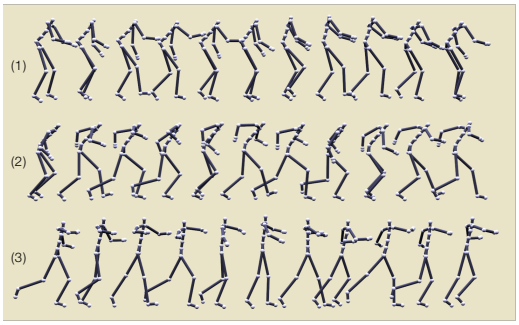

图1. 现有动作数据库 [21, 38] 中风格夸张的行走动画示例。(1)“僵尸”；(2)“老人”；(3)“孩子气”。在没有上下文的情况下，很难区分这些风格。

我们认为，这种分类过于主观，不适合作为提供给用户的表达参数。多种因素都可能使风格类别产生偏差，进而导致基于这些类别的编辑工具无法满足用户的期望。

偏差源于用于校准风格类别的数据是如何记录的。这些数据通常并非在自然情境中采集，而是由演员 [38] 表演的。因此，源序列往往是对演员和导演心目中的类别进行过度演绎、夸张呈现（见图 1）。即使采用最贴近真实情境的采集流程，也会通过预先设定的情境使受试者产生特定的情绪状态。然而，偏差并未因此完全消除；它可能源于实验设计 [19, 27], ，甚至源于不同文化对同一种情绪的不同理解 [14]。

动画编辑研究的主要目标是为动画师提供工具，使他们能够将脑海中对角色动作的构想转化为切实的动画数据。低层级方法虽然精细，却操作繁琐；高层级方法则可能带有偏见，而且会过多剥夺用户的控制权。因此，如何找到折中方案仍是一个重要的开放性问题。既能让新手轻松操作，又能让专业人士充分掌控的高效工具，目前仍有待开发。

参数化编辑范式，如 [29], 所示，为规避偏差问题提供了有趣的思路，同时仍能实现更高层次的控制。沿着这一思路，本文提出了一种替代方案：引入一种能够通过更高层次、但仍具客观性的控制来操纵风格的方法，即姿势度量。

4 方法

我们提出了一种新的角色动画编辑范式，动画师可操控称为姿势度量的中间层参数，并借助神经网络学习并自动处理复杂的形态约束。我们首先介绍这一新概念，随后说明实验中所使用的神经网络流程。最后，我们介绍如何根据用户输入编辑完整动画序列的方法。

在后续章节中，术语骨架 用于定义角色的虚拟骨架，它由一组关节和骨骼构成层级结构 [26]。我们将姿势 定义为一个数据向量，用于描述骨架中每个关节在单个时间帧上的位置和／或朝向。动画序列 则是一系列关键姿势，用来表现角色的动作。

在公式中，符号约定如下：我们用 A 表示一个包含 N 帧的动画序列，用 $p = A_n$ 表示第 n帧的姿势。对于由 J 个关节组成的骨架，可以对姿势进行索引，以访问特定关节的数据： p_{head}。

4.1 姿势度量

能够通过操控更高层级的参数来编辑动画序列是一种理想的能力，但相关文献在低层级控制（使用正向和逆运动学，针对骨骼的位置或朝向等粗略姿势数据进行操作）与高层级控制之间存在两极分化。后者使用风格 标签来描述情绪或行为（即“愤怒”“幼稚”等） [2, 11, 35]。这类风格控制器缺乏控制力，因为它们会根据一种概括性的风格定义一次性生成风格化动画，而用户未必认同这种定义。

为弥合这两者之间的差距，我们提出姿势度量：更高层级、客观且风格无关的数值，可以直接从单个姿势计算得出，并且不依赖任何假定的普适人类感知行为。

这些度量的设计愿景是尽可能灵活，从而不对用户构成限制。每个角色都各不相同，因此这一概念应尽可能少地假设角色的动作方式。这样，用户便可以自行设计度量，在脑海中将其与自己对特定风格的定义联系起来，并通过这些度量编辑动画。

举例来说，设想一位艺术家正在编辑一段角色行走并表现出冷怒的动画序列。在他看来，角色应该迈着短促而快速的步伐，双臂紧绷，伸向地面。使用以往的风格编辑方法，他或许可以按下“愤怒”按钮，也可能用滑块来决定最终效果应有多愤怒。遗憾的是，系统对“愤怒”的理解可能基于另一种定义：角色现在走得更快，双臂弯曲并摆动，

拳头靠近躯干。使用度量指标可以让用户绕过这个问题。创建一些度量指标，例如肘部形成的角度、双脚之间的距离，或双手与地面之间的距离，再通过这些指标编辑动画，就能避免依赖对风格特征的预设。

新度量指标的唯一要求是能够根据单个姿势计算得出，并提供相应的计算函数。只要给出一个用于计算此类用户自定义度量指标的函数，所提出的关联框架就具有足够的可扩展性，能够将其纳入其中，并据此编辑动画序列。

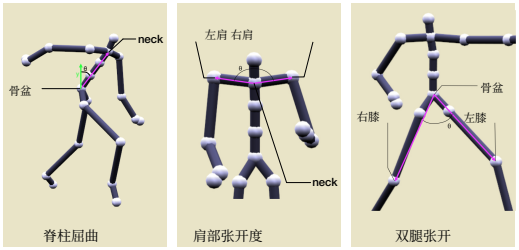

图 2. 本文示例中使用的三种度量指标的示意图。

实验姿态度量。 如前所述，姿势度量由用户提供。为演示本系统的应用，以下实验采用三种适用于人体骨架的示例度量指标：脊柱_屈曲、肩部_张开度和腿部_张开度。

这三者都表示由 $\vec{u}$ 和 $\vec{v}$ 这两个从姿势中获得的向量相交所形成的角度。形式上，交点处的角度按照公式1计算。

$$\theta(\vec{u},\vec{v}) = \arccos\frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\cdot\|\vec{v}\|} \tag{1}$$

脊柱_屈曲度量表示 $\vec{y}$ 世界“向上”向量与脊柱轴之间的夹角。其计算方式如公式2所示。角色匆忙时，可用它让角色向前弯腰；放松时，则可让角色向后弯腰。

$$SF(p) = \theta(p_{neck} - p_{pelvis}, \vec{y}) \tag{2}$$

接下来，肩部_开合度量（公式3）用于描述角色躯干／肩部区域的张开程度。例如，可以抬起角色的肩部来表现焦虑或羞怯，也可以使其舒展开来，表现更加放松或自豪的状态。

$$SO(p) = \theta(p_{spine1} - p_{rshoulder}, p_{lshoulder} - p_{spine1}) \tag{3}$$

最后，腿部_张开程度 （式4）表示角色膝盖彼此之间的距离。可能的用途包括限制行走步幅，或调整角色看起来是否保持平衡。

$$LS(p) = \theta(p_{pelvis} - p_{rknee}, p_{lknee} - p_{pelvis}) \tag{4}$$

在本文的公式中，姿势 p 的度量 M 的值 m 表示为函数 $m = M(p)$。

4.2 学习式度量管线

本节介绍一种使用姿势度量编辑动画序列的流程。该流程在设计时考虑了模块化，因为我们希望用户能够随时决定使用哪种度量指标，并在需要时通过自定义度量指标对其进行扩展。我们首先介绍该流程的整体布局，然后分别介绍其中的各个部分。我们还将说明用户如何扩展该流程，并对其进行配置以满足任意度量指标组合。最后，我们提出一种编辑完整动画片段的方法，利用我们的架构将对选定帧所做的更改传播至序列的其余部分。

流程布局。 我们的流程（图 3）依赖于若干小型神经网络。我们选择这些网络，是因为它们内存占用低、运行时间短。使用较小的独立网络，还能让我们轻松配置流程，添加、移除或修改模块。

流程的外层是一个编码器-解码器模型对，其目的是学习一个潜在姿势空间，在其中姿势表示更加紧凑，也更易于操作。由于解码器只能将潜在编码映射到真实姿势空间，因此在对样本进行均值计算和插值时，不会有意外生成错误姿势的风险。编辑本身由度量网络完成，它仅在潜在空间中运行。度量网络的一个实例经过训练后，可根据一个单一 度量值修改潜在姿势。网络规模较小，而且我们重复使用已训练好的编码器-解码器，因此训练时间相当短。这也进一步方便用户设计自己的度量指标、训练网络，并将其整合到流程中。

编码器-解码器。 编码器 E 和解码器 D 网络经过联合训练，将一个姿势映射 p 到潜在表示 z，如公式5所示。实际上，它们经过训练以最小化 p 与其重构版本 $\hat{p}$ 之间的平方重构误差。这两个网络都是简单的单层感知机，在隐藏层中使用 512 个单元，并采用修正线性单元（ReLU）激活函数。在我们的实验中，潜在空间维度设为 64。

$$\begin{aligned} z &= E(p) \\ \hat{p} &= D(z) \\ \mathcal{L}_{\text{ED}} &= \frac{1}{J}\sum_{j=0}^{J}\left(p_j - \hat{p}_j\right)^2 \end{aligned} \tag{5}$$

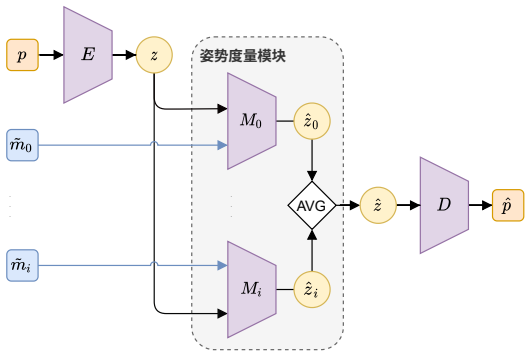

图 3. 所提出的流程及其姿势度量模块概览。该流程根据任意数量的用户提供的目标度量值 $\tilde{m}_{0\ldots i}$修改输入姿势 p。编码器-解码器对（E 和 D）将姿势映射到潜在姿势空间，并从中映射回来，其他模块则在该空间上运行。每个度量指标均由专用网络$M_{0..i}$处理。得到的潜在样本$\hat{z}_{0..i}$取均值，然后映射回姿势空间。

度量网络。。度量网络 N_m 学习生成一个经过修改的潜在姿势$\hat{z}_m$，以一个起始姿势作为输入 z，并与目标度量值拼接 $\tilde{m}$（见公式6）。

$$\hat{z}_m = N_m(z \oplus \tilde{m}) \tag{6}$$

为在训练期间体现这一目标，我们从同一个原始动画序列中采样两个姿势：输入$p = A_t$ 和目标 $\tilde{p} = A_{t+n}$。从相邻帧采样可避免目标姿势与起始姿势相距过远，并确保网络学会对姿势进行编辑，而不是生成一个无关的姿势。在我们的实验中，n 在每次迭代中随机采样，取自 $n \in [-10, 10]$。由于我们的目标是交互式编辑场景，度量值发生较小变化是合理的。输入的目标度量值根据 $\tilde{p}$ 计算，并与 $z = E(p)$一起馈入网络。得到的修改后潜在姿势$\hat{z}$通过解码器映射回姿势空间，即$\hat{p} = D(\hat{z})$。最后，按照损失函数（见公式 7）所示，训练网络以最小化目标姿势与修改后姿势之间的均方误差。

$$\mathcal{L} = \frac{1}{J}\sum_{j=0}^{J}\left(\tilde{p}_j - \hat{p}_j\right)^2 \tag{7}$$

秉持网络越小越好的理念，我们选择了一个具有 126 个隐藏单元和 ReLU 激活的单层感知机。输入层的单元数等于自编码器潜在维度的大小加一（用于计入度量值）。输出的大小与潜在维度相同。根据经验，我们发现这种架构满足我们的需求，更大或更深的网络并未带来可感知的价值。如果用户认为另一种架构更适合他的某个度量指标，那么唯一的要求是输入层和输出层的大小。

数据。。我们使用一个包含约 150 万个姿势的姿势数据集，相关描述见 Victor 等人 [32]。该数据集由多个现有动画数据库合并而成[15, 20, 31]。每个动画片段都被重定向到一个通用骨架上[21]。姿势数据由各关节相对于骨盆关节在地面上的投影的位置组成。用于训练网络时，数据集会经过归一化处理：每个分量值减去均值，再除以标准差。

训练参数。。编码器-解码器和度量网络分别训练，但使用相同的超参数：使用 Adam 优化器 [23]，学习率为 0。0001，以 1024 对姿势为一批进行训练，直至收敛；在 NVIDIA Quadro T1000 上，这一过程约需 15 分钟。

流水线模块化。在需要同时编辑多个度量指标的情况下，可以组合一组度量网络的结果。由于网络的输出是姿势的潜在表示，因此可以将它们组合起来。在 [32],中，作者依次运行多个网络。我们则建议对多个潜在姿势取平均，这样就不必为这些模块设定顺序。

如公式 8所示，源姿势 p 首先映射到潜在空间，然后每个专用度量网络 $M0..i$都会生成经过修改的潜在结果。所有这些中间结果都会取平均，最后由解码器映射回姿势空间。

$$\hat{p} = D\left(\frac{1}{n}\sum_{i=0}^{n} M_i(E(p))\right) \tag{8}$$

这种针对单个关键帧的方法可以与其他姿势处理方案结合使用，也可以与插值方法集成，从而尽可能让用户获得充分的控制。使用共享的潜在空间还使我们能够将文献中的其他模块整合到流程中。例如，[32] 中的逆运动学求解器运行于类似的潜在空间；如果使用该流程的潜在空间进行训练，其输出便可与度量网络的输出取平均。

4.3 动画编辑

上述姿势流程一次只能处理单个姿势。然而，在现实场景中，动画师通常会编辑某个特定的关键帧，并期望所做的更改逐渐应用到相邻帧。为满足这一需求，我们提出一种方法，只需一组目标度量指标即可编辑整个动画序列。

我们的方法（图 4）类似于传统的动画曲线范式，使用用户提供的权重曲线W来控制每个关键姿势受到的影响程度

修改。我们再次利用潜在空间良好的插值特性。

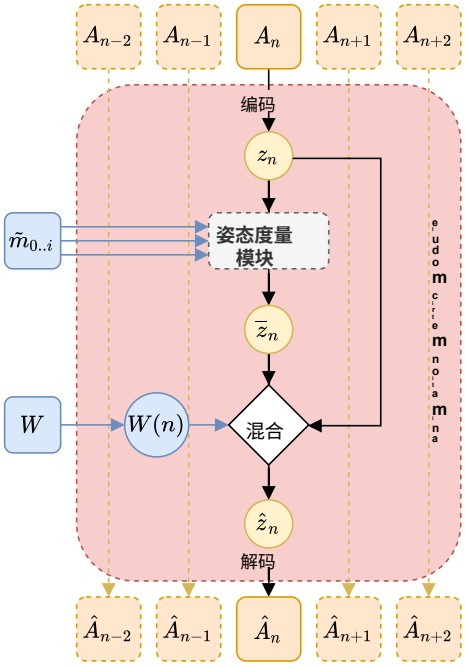

图 4. 所提方法的可视化概览，用于呼应对动画片段中某个关键姿势 A_n 所做的修改。混合操作发生在潜在姿势之间（源姿势z_n 以及经度量模 z_n修改后的姿势），并使用从用户提供的曲线中采样的因子$W(n)$）。

要编辑动画 A，用户选择一个关键帧 A_n并提供任意数量的目标度量值$\tilde{m}_{0..i}$以及一条权重曲线W。对于序列中的每一帧 A_t，都会对潜在表示进行 z_t 编码。随后，姿势度量模块会对每个潜在表示进行处理，得到一个中间潜在姿势 z_t。最终修改后的潜在姿势$\hat{z}_t$通过对 z_t 和 z_t 进行混合得到，混合比例为$W(t)$（公式 9）。

$$\hat{z}_t = \big(1 - W(t)\big) \cdot z_t + W(t) \cdot \bar{z}_t \tag{9}$$

在我们的实验中，W 是一个帽函数，在选定帧处达到峰值 1，并在片段两端降至 0。不过，这个权重函数是一个外部参数：使用正弦函数进行实验可以得到相近的结果，而动画师也可以编辑这条曲线，以符合自己的预期，这与现有的动画曲线相似。我们还注意到，片段中的每个姿势都是独立处理的。因此，编辑可以并行进行，同时保持实时交互。

5 结果

本节展示了该流程在不同场景下的视觉动画编辑结果。更多示例请参见随附视频。

5.1 姿势编辑

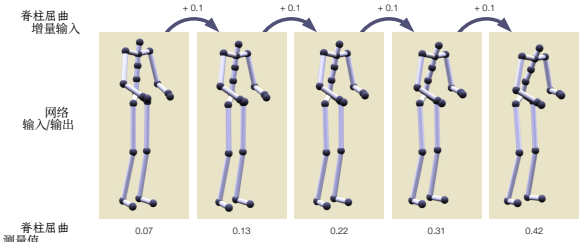

图 5. 通过提供目标 *spine_flexion* 输入逐步修改姿势。每一步中，用户输入都比计算出的度量值增加 0.1。

在图5中，用户通过逐渐增大的目标脊柱_屈曲度量值来修改单个姿势。结果符合预期的变化，如每个姿势下方测得的度量值所示。虽然有人可能会认为，改变骨盆骨骼的朝向角也能产生类似结果，但仔细观察生成的姿势，便能看出潜在姿势空间的有趣特性。每次迭代时，骨架的手臂都会略微抬起，膝盖也会弯曲，以补偿脊柱屈曲导致的平衡失调。

5.2 模块化流程

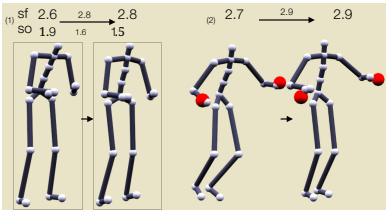

图6。使用多个模块的流程示例结果。在（1）中，同时使用肩部_张开度（so）和脊柱_屈曲。在（2）中，我们将来自 [32] 的潜在逆运动学求解器整合到流程中，并将其用于脊柱_屈曲。

图 6 展示了对多个模块的潜在输出取平均后得到的一些结果。第一个示例展示了同时修改脊柱_屈曲 和肩部_张开度 度量指标的结果。骨架上可以看到这两处修改：脊柱变直，肩部下沉。姿势的其余部分也略作调整，以适应这些变化：手的位置降低，膝盖弯曲得更深。

第二个示例展示了将脊柱_屈曲度量模块与 [32]中的神经逆向运动学方法结合使用的结果。对潜在输出取平均值后，我们的方法会在各项约束之间找到折中方案。这里，骨架的背部变直，并且手部的逆运动学目标也得到了满足。

5.3 动画编辑

图 7 展示了使用前文所述流程编辑动画 A 的结果。用户选择一帧A_n，指定目标腿部_张开度度量值，并设置一个插值因子曲线，即一个在A_n处达到峰值 1、并在两侧的 A_{n-3} 和 A_{n+3}处降至 0 的帽函数。其余帧的因子均为 0。

网络会分别处理每一帧的姿势，并使用该帧对应的因子对输入姿势和输出姿势的潜在表示进行插值，从而得到结果。

因此会输出动画，其中影响在所选帧上最显著，并向两侧逐渐减弱。在第 n 帧，输出姿势的变化最大，腿部夹角的减小十分明显。在第 $n-2$ 帧和第 $n+2$ 帧，这种影响几乎不可见，而更远的帧则完全不受影响，正如预期。

在原始姿势和修改后的姿势之间进行插值，可以让整个序列保持时间连贯性，即使流程每次只处理一帧，且不了解过去和未来帧的上下文。我们还利用了潜在空间的特性：任何潜在空间插值结果都会映射回真实姿势空间，从而避免骨架相互穿插等伪影。

6 结论

为避免动画风格分类所带来的问题，我们引入了姿势度量这一概念：它是在单个姿势上计算得出的数值，用户可以自行定义，以符合自己对风格的自身理解。我们提出了一套通过操控这些度量指标来编辑

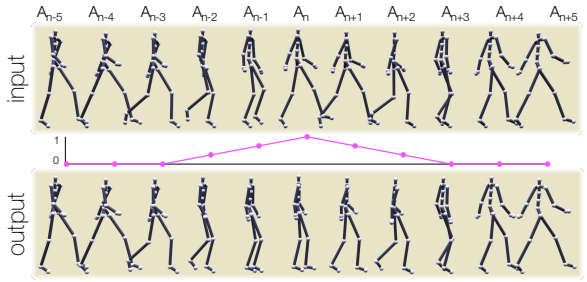

图 7. 根据用户输入编辑动画片段。流程根据选定的帧 A_n、双腿_张开 度量指标的目标值以及权重曲线生成输出。

动画的流程，该流程采用模块化的小型神经网络，可实时运行。用户可以设计自定义度量指标，并在短短 15 分钟内训练相应的度量网络，从而扩展这一流程。

这种方法的主要缺点是，它仅适用于单一的标准化人形骨架，网络也是在该骨架上训练的。这是数据驱动方法的常见问题：若要适配新的骨架，就需要重新训练网络。如果新骨架的现有动画数据难以获取，这可能会很困难；不过，动画重定向领域的最新进展 [1]或许能通过提供更易获取的数据增强来助力。度量指标的模块化也有帮助，因为不同的骨架需要一套新的度量指标。

未来的工作将聚焦于完善所提出的流程：通过自动检测重叠的度量指标来确保流程的可靠性，并开展用户研究，以评估该方法的可用性。

参考文献

[1] Kfir Aberman、Peizhuo Li、Dani Lischinski、Olga Sorkine-Hornung、Daniel Cohen-Or 和 Baoquan Chen。2020年。用于深度动作重定向的骨架感知网络。《ACM 图形学汇刊》39，4（7月 2020 62 62 1 62 62 14 10 1145 3386569 3392462），: : – : :。[2] https://doi.org/ . / . Kfir Aberman、Yijia Weng、Dani Lischinski、Daniel Cohen-Or 和 Baoquan Chen。2020年。从视频到动画的非配对动作风格迁移。《ACM 图形学汇刊》39，4（2020年7月），64:64:1– 64 64 12 10 1145 3386569 3392469 : :。https://doi.org/ . / . [3] Andreas Aristidou 和 Joan Lasenby。2011. FABRIK：一种快速迭代的逆运动学问题求解器。《图形模型》73，5（2011年9月），243–260。https://doi.org/10.1016/j.gmod.2011.05.003[4] Andreas Aristidou、Qiong Zeng、Efstathios Stavrakis、KangKang Yin、Daniel Cohen-Or、Yiorgos Chrysanthou 和 Baoquan Chen。2017年。非结构化舞蹈动作的情感控制。载于 ACM SIGGRAPH / Eurographics 计算机动画研讨会论文集（SCA '17）。美国计算机协会，纽约，NY，美国， 1 10 10 1145 3099564 3099566 –。https://doi.org/ . / . [5] Matthew Brand 和 Aaron Hertzmann。2000. 风格机器。载于第27届年度计算机图形学与交互技术会议论文集（SIGGRAPH '00）。ACM Press/Addison-Wesley Publishing Co.，纽约，NY，美国，183–192。https://doi.org/10.1145/ 344779.344865[6] Edwin Catmull。1978年。计算机辅助动画中的问题。载于第5届年度计算机图形学与交互技术会议论文集（SIGGRAPH '78）。ACM，纽约，NY，美国， 348 353 10 1145 800248 807414 –。https://doi.org/ . / . Ed [7] win Catmull 和 Raphael Rom。1974. 一类局部插值样条。载于《计算机辅助几何设计》，ROBERT E. Barnhill 和 RICHARD F. Riesenfeld（编）。Academic Press，317–326。https: //doi.org/ 10 1016 978 0 12 079050 0 50020 5 . /B - - - - . - Jinxiang Chai 和 [8] Jessica K. Hodgins。2007. 使用统计动力学模型进行基于约束的动作优化。载于 ACM SIGGRAPH 2007 论文集（SIGGRAPH '07）。美国计算机协会，纽约，NY，美国， 8 10 1145 1275808 1276387 – es。https://doi.org/ . / . [9] Byungkuk Choi、Roger Blanco i Ribera、J. P. Lewis、Yeongho Seol、Seokpyo Hong、Haegwang Eom、Sunjin Jung 和 Junyong Noh。2016年。SketchiMo：面向关节角色的基于草图的动作编辑。《ACM 图形学汇刊》35，4（2016年7月），146:1–146:12。https://doi.org/10. 1145/2897824.2925970

[10] Loïc Ciccone、Cengiz Öztireli 和 Robert W. Sumner。2019年。交互式动画控制的切空间优化。《ACM 图形学汇刊》38，4（2019年7月），1–10。https://doi.org/10.1145/ 3306346.3322938[11] Han Du、Erik Herrmann、Janis Sprenger、Klaus Fischer 和 Philipp Slusallek。2019年。使用变分生成模型进行风格化运动建模与合成。载于《运动、交互与游戏》（MIG '19）。ACM，纽约，NY，美国，32:1–32:10。https://doi.org/10.1145/ 3359566.3360083[12] Funda Durupinar、Mubbasir Kapadia、Susan Deutsch、Michael Neff 和 Norman I. Badler。2016年。PERFORM：运用拉班动作分析将 OCEAN 人格融入人体运动的感知方法。《ACM 图形学汇刊》36，1（2016年10月）。https://doi.org/10.1145/ 2983620[13] P. Ekman 和 W. V. Friesen。1967年。情绪判断中的头部与身体线索：一种重新阐释。《知觉与运动技能》24，3（6月 1967 711 724 10 2466 1967 24 3 711），–。[14] https://doi.org/ . /pms. . . . Hillary Anger Elfenbein 和 Nalini Ambady。2002. 情绪识别的普遍性与文化特异性：荟萃分析。《心理学通报》128，2（2002年），203–235。https://doi.org/10.1037/ 0033-2909.128.2.203[15] Nesrine Fourati 和 Catherine Pelachaud。2014年。Emilya：日常动作数据库中的情绪化身体表达。载于第九届国际语言资源与评测会议论文集（LREC-2014）。欧洲语言资源协会（ELRA），雷克雅未克，冰岛，3486–3493。[16] Maxime Garcia、Remi Ronfard 和 Marie-Paule Cani。2019年。空间运动涂鸦：在虚拟现实中利用手势和拉班动作分析绘制动画草图。载于《运动、交互与游戏》（MIG '19）。美国计算机协会，纽约，NY，美国，1–10。https://doi.org/ 10 1145 3359566 3360061 . / . [17] Martin Guay、Marie-Paule Cani 和 Rémi Ronfard。2013年。动作线：用于表现性角色姿势设计的直观界面。《ACM 图形学汇刊》32，6（2013年11月），1–8。https://doi.org/10. 1145/2508363.2508397[18] Félix G. Harvey、Mike Yurick、Derek Nowrouzezahrai 和 Christo-pher Pal。2020年。鲁棒运动补间。《ACM 图形学汇刊》39，4（2020年7月），60:60:1–60:60:12。https://doi.org/10.1145/ 3386569.3392480[19] Joseph Henrich、Steven J. Heine 和 Ara Norenzayan。2010年。世界上最古怪的人？《行为与脑科学》33，2-3（2010年6月），61–83。https://doi.org/10.1017/S0140525X0999152X [20] Daniel Holden、Jun Saito 和 Taku Komura。2016年。用于角色运动合成与编辑的深度学习框架。《ACM 图形学汇刊》35，4（2016年7月），1–11。https://doi.org/10.1145/ 2897824.2925975[21] Daniel Holden、Jun Saito、Taku Komura 和 Thomas Joyce。2015年。使用卷积自编码器学习运动流形。载于 SIGGRAPH ASIA 2015 技术简报（SA '15）。ACM Press，神户，日本，1 4 10 1145 2820903 2820918 –。https://doi.org/ . / . [22] Alec Jacobson 、 Daniele Panozzo 、 Oliver Glauser 、 Cédric Pradalier、Ot-mar Hilliges 和 Olga Sorkine-Hornung。2014年。用于角色关节控制的实体模块化输入设备。《ACM 图形学汇刊》33, 4 2014 82 1 82 12 10 1145 2601097 2601112 （7月），：[23] – :。https://doi.org/ . / . Diederik P. Kingma 和 Jimmy Ba。2015. Adam：一种随机优化方法。载于 arXiv:1412.*6980 [*Cs*]*。美国加利福尼亚州圣迭戈。https://doi.org/ 10 48550 1412 6980 1412 6980 [ ] . / [24] ARXIV. . arXiv: . cs Yuki Koyama 和 Masataka Goto。2018. OptiMo：面向关键帧角色动画的优化引导运动编辑。载于2018年计算机系统中的人为因素国际会议论文集（CHI '18）。ACM Press，加拿大蒙特利尔，1–12。https://doi.org/10.1145/ 3173574.3173735[25] John Lasseter。1987年。应用于三维计算机动画的传统动画原理。《ACM SIGGRAPH 计算机图形学》21，4（ 1987 35 44 10 1145 37402 37407 ），– 。[26] https://doi.org/ 。/。Nadia Magnenat-Thalmann、Richard Laperrire 和 Daniel Thalmann。1988年。用于手部动画和物体抓取的关节相关局部变形。载于《Graphics Interface '88 会议论文集》（GI '88）。加拿大人机通信协会，埃德蒙顿，艾伯塔省，加拿大，26 33 10 20380 1988 04 –。https://doi.org/。/gi。[27] Jim McCambridge、John Witton 和 Diana R. Elbourne。2014年。霍桑效应的系统综述：研究研究参与效应需要新概念。《临床流行病学杂志》67，3（2014年3月），267–277。https://doi.org/10.1016/j.jclinepi.2013.08.015[28] Tanmay Randhavane、Uttaran Bhattacharya、Kyra Kapsaskis、Kurt Gray、Aniket Bera 和 Dinesh Manocha。2019年。利用情感和深度特征识别行走中的情绪。https://doi.org/10.48550/arXiv.1906.11884[29] Alba M. Rios Rodriguez、Steven Poulakos、Maurizio Nitti、Mattia Ryffel 和 Robert Sumner。2019年。参数化动画活动。载于《运动、交互与游戏》（MIG '19）。美国计算机协会，纽约，美国，11 1 11 9 10 1145 3359566 3360062 : – : 。[30] https://doi.org/。/。Charles Rose、Michael F. Cohen 和 Bobby Bodenheimer。1998年。动词与副词：多维运动插值。《IEEE计算机图形学与应用》18，5（1998年9月），32–40。https://doi.org/10.1109/38.708559[31] 卡内基梅隆大学。2003年。卡内基梅隆大学图形实验室动作捕捉数据库。[32] Léon Victor、Alexandre Meyer 和 Saïda Bouakaz。2021年。基于学习的姿势编辑，实现高效交互式设计。《计算机动画与虚拟世界》32，3-4（2021年），e2013。https://doi.org/10.1002/cav.2013[33] Rudolf von Laban 和 Lisa Ulmann。1971年。《动作的掌握》。[34] Li-Chun Tommy Wang 和 Chih Cheng Chen。1991年8月。求解机械操作臂逆运动学问题的组合优化方法。《IEEE机器人与自动化汇刊》7, 4 1991 489 499 10 1109 70 86079 （8月/），–。[35] https://doi.org/。/。Qi Wang、Thierry Artières、Mickael Chen 和 Ludovic Denoyer。2020年。用于人体运动建模的对抗学习。《The Visual Computer》36，1（2020年1月），141–160。https://doi.org/10.1007/s00371-018-1594-7[36] Yu-Hui Wen、Zhipeng Yang、Hongbo Fu、Lin Gao、Yanan Sun 和 Yong-Jin Liu。2021年。结合生成流的自回归风格化运动合成。载于《IEEE/CVF计算机视觉与模式识别会议论文集》。13612–13621。https://doi.org/10.1109/CVPR46437.2021.01340[37] Xiaomao Wu、Maxime Tournier 和 Lionel Reveret。2011年。基于大型运动数据库的自然角色摆姿。《IEEE计算机图形学与应用》31，3（2011年5月），69–77。https://doi.org/10.1109/MCG.2009.111[38] Shihong Xia、Congyi Wang、Jinxiang Chai 和 Jessica Hodgins。2015年。面向无标签异构人体运动的实时风格迁移。《ACM 图形学汇刊》34，4（2015年7月），119:1–119:10。https://doi.org/ 10 1145 2766999 。/ Wataru Yoshizaki、Yuta Sugiura、Albert C. Chiou、Sunao Hashimoto、Masahiko Inami、Takeo Igarashi、Yoshiaki Akazawa、Katsuaki Kawachi、Satoshi Kagami 和 Masaaki Mochimaru。2011年。一种用作控制数字人体模型输入设备的驱动式实体木偶。载于《SIGCHI计算系统中的人为因素会议论文集》（CHI '11）。美国计算机协会，纽约，美国，637–646。https://doi.org/10.1145/1978942.1979034