

三 维 动 画 和 交 互 设 计 课 程 读 书 报 告



作者姓名 梁家坤

作者学号 21851125

年级专业 2018级软件工程

所在学院 软件学院

完成日期 2018.11.30

Research on Computer Animation and Interaction Design

A Report Submitted to

Zhejiang University

Major Subject: Software Engineering

Advisor: Li Qilei

By

Liang Jiakun

College of Software Technology, Zhejiang University

2018

**目录**

[一、 论文简介 4](#_Toc532369555)

[二、 该文研究背景与问题描述 4](#_Toc532369556)

[2.1 研究背景 4](#_Toc532369557)

[2.2 问题描述 4](#_Toc532369558)

[2.2 论文贡献 5](#_Toc532369559)

[三、 模型架构实现 6](#_Toc532369560)

[3.1 模型概述 6](#_Toc532369561)

[3.2 概要设计 6](#_Toc532369562)

[3.3 详细设计 7](#_Toc532369563)

[四、 实验结果与分析 10](#_Toc532369564)

[五、 思考与扩展建议 12](#_Toc532369565)

[参考文献 13](#_Toc532369566)

# 论文简介

论文题目：Deep learning of biomimetic sensorimotor control for biomechanical human animation，在人体生物力学动画中基于深度学习的仿生感觉运动控制研究

作者：5人，UCLA（加州大学洛杉矶分校）的研究团队

发表于：ACM Transactions on Graphics，2018年7月

# 该文研究背景与问题描述

## 2.1 研究背景

基于人体解剖学的图形字符建模在计算机动画领域中引起了广泛关注。原则上，越来越精确的人体生物力学建模应该可以产生更逼真的动画。然而，对于复杂的人体肌肉骨骼模型（通常具有超过200个骨骼和大约1000个收缩肌肉致动器）来说，如何进行控制仍然是一项艰巨的挑战。

## 2.2 问题描述

这篇文章针对以下具体研究问题进行了探索：

1）仿生运动控制：生物运动控制机制与以往的在线控制方法不同，在线控制方法依赖于反向运动学、最小努力优化等方法来在线计算产生运动的肌肉激活，而大脑能够从历史经验中来学习肌肉控制，这天然就很适合用基于如DNN等深度学习方法训练的方法来模拟控制肌肉。

2）仿生感觉控制：为了更加真实地模拟人体感觉控制，这篇文章基于人类视觉的基本原理，采用仿生学方法构建了一个三维眼睛模型，来为人体模型提供计算机视觉。该仿生眼球具有眼球运动的能力，它的视网膜由光感受器填充，然后通过DNN训练从光感受器中的光感反应来推断出基本的视觉信息。

结合上面两个方面，这篇文章介绍了一种用于仿生感觉运动控制的人体模型。在模型仿生眼球的视网膜上分布不均匀的光感受器，通过计算辐照度来对周围环境进行视觉采样。利用生物力学模拟了人体肌肉骨骼模型，通过激活信号来驱动肌肉。通过实验表明，该仿生人体模型能自动学习眼睛、头部和四肢的高效、在线、主动的视觉运动控制，以执行对目标物体的视觉追踪、视觉引导等，目前已经具备拦截移动目标的行动，以及绘图和书写等能力。

如下所示，图1为该仿生人体模型在通过虚拟眼睛中的光感受器对周围环境进行视觉采样；图2中展示了该模型可以驱动头颈部来移动虚拟眼睛的视线、跟踪小球的运动轨迹，并用胳膊和腿进行简单的拦截动作；图3中展现了该模型具备一定的写字和绘画能力。

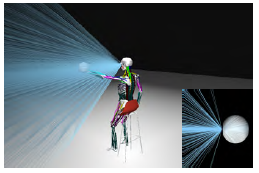


图1 通过仿生眼睛进行视觉采样

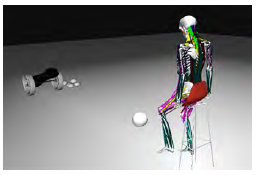


图2用胳膊和腿拦截运动中的小球

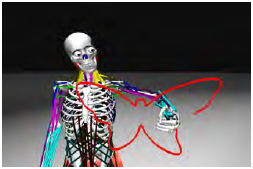


图3画蝴蝶过程

## 2.2 论文贡献

这篇文章的贡献是在于建立了人体感觉运动控制的新型仿生模型，包括使用复杂的生物力学人体设备来模拟可收缩的肌肉执行器。该模型不仅可以完成静态动态的视觉目标追踪，还具备包括如进行自然协调的头颈部、肢体运动，以及手工书写和绘制等复杂动作的能力。就像在使用现代机器学习方法来实现对现实肌肉骨骼系统的在线运动控制以及对活跃的感知进行在线视觉处理一样，这篇文章构建了一套完整的模块化DNN模型来实现这一目标，所有DNN都通过生物力学模型本身合成的训练数据进行离线自动训练。

# 模型架构实现

## 3.1 模型概述

该仿生人体模型的感觉运动控制系统包含20个用于训练的深度神经网络（DNN）。10个DNN用来构成神经肌肉运动子系统，其中有5个神经肌肉控制器，用来控制颈部肌肉驱动肌肉骨骼复合体以产生自然头部运动，以及控制手臂肌、腿肌等肢体。另外10个DNN用于构成视觉感知子系统，其中2个DNN驱动眼睛和头部运动，8个DNN提取指导手臂和腿部动作所需的视觉信息，如前所述，信息直接来自光感受器响应。

## 3.2 概要设计

整个感觉运动控制框架的肌肉骨骼模型中有823个肌肉执行器，其中有352个肌肉执行器需要通过指定一个随时间变化的传出神经激活信号 a(t) 来控制，所有这些激活信号都由训练得到的DNN生成。图4为整个感觉运动控制系统的神经网络结构图，包括感觉和运动子系统，可以看到总共包含20个深度神经网络（DNN），编号为1-20。

视觉感知子系统：包含10个视觉DNN。图4(a) 每个视网膜感光器将光线投射到虚拟世界中以计算感光器捕获的辐照度。图4(b)R和图4(b)L为左右眼视网膜上3,600个感光器的排列情况。每个视网膜输出10,800维光学神经矢量，并作为10个视觉DNN的输入。其中用图4(c)R和图4(c)L输出左右眼的角度差异，最后驱动眼睛图4(e)运动到视觉目标位置。

运动子系统：包含10个运动DNN。神经肌肉运动控制器图4(f) 输入图4 (c)R和图4(c)L 的DNN输出结果平均值，并将激活输出到颈部肌肉组。图4(d)R和图4(d)L输出的肢体到目标的视觉差异估计平均值为输入，输出到肢体肌肉骨骼复合体的肢体神经肌肉运动控制器图4(g)、图4(h)，即各自的手臂和腿部肌肉群，以此完成驱动功能。

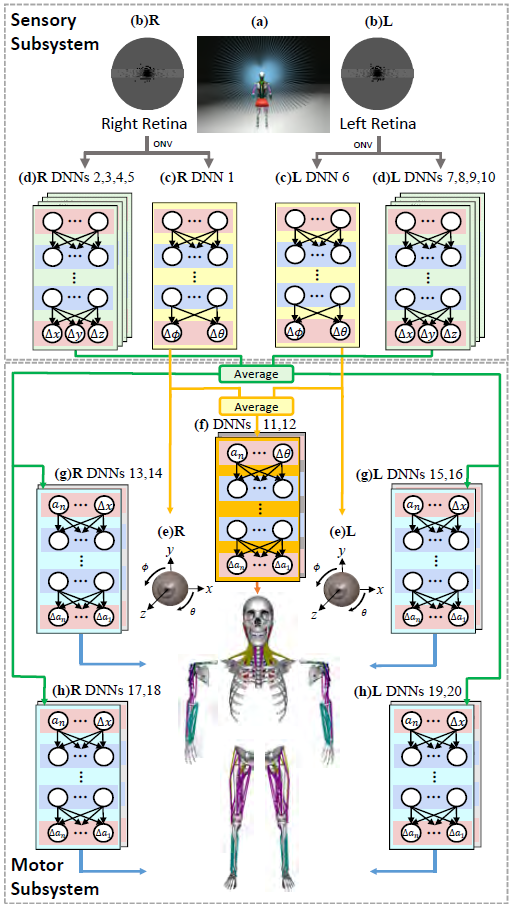


图4 系统完整的感觉运动控制结构

## 3.3 详细设计

对于图4中运动子系统的11,13,15,17,19号DNN，设计如下。头颈运动DNN(如图5所示)的功能是在实现头部真实运动的同时，要向颈部肌肉发出激活信号，以保持头部在颈部上方的平衡状态。四肢运动DNN(如图6所示)的功能是向四肢肌肉产生传出的激活信号，以进行控制胳膊和腿的运动，如伸展四肢达到目标。除了输入层和输出层的尺寸不同之外，其余三个运动DNN的结构都是相同的。

对于图4中运动子系统的12,14,16,18,20号DNN，设计如下。该仿生模型的反射运动控制器是通过离线合成的训练数据学习到所需控制功能的运动神经网络，该过程类似于头颈运动控制器的训练过程。

|  |  |  |
| --- | --- | --- |
|  |  |  |
| 图5 头颈运动DNN | 图6 四肢运动DNN | 图7 反射运动DNN |

对于图4中的视觉感知子系统，如前所述，用10个训练好的DNN模型来进行视觉采样。对于虚拟眼睛和视网膜模型，这篇文章根据人体生理数据对眼睛进行建模，将其建模为12毫米半径、可以绕中心垂直和水平旋转任意角的球体，并用传统的射线追踪方法计算眼球后部视网膜球面上任意一点的辐照度。通过针孔将视网膜表面光感受器位置的样本射线投射到三维空间中，与虚拟物体的可见表面递归相交，并根据局部光照模型将阴影射线与光源结合，这些射线返回的RGB值即决定了照射到视网膜光感受器上的辐照度，图8为视网膜成像过程。其中光感受器在视网膜上的位置可见图9。

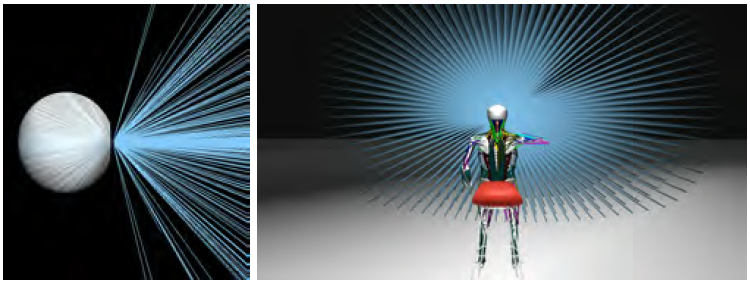


图8 视网膜成像过程

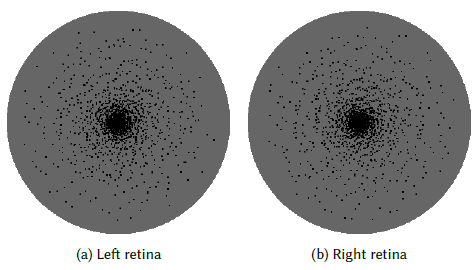


图9光感受器在视网膜上的位置（黑点）

对于图4中编号为1,6的视觉DNN，设计如下。与眼睛一起，左右神经系统构成眼动子系统，它们的第一个作用是通过驱动眼球运动来跟踪目标物体，从而在注视方向上产生自动变化。如图9所示，一个白色球体正在运动，从右下方进入视野，刺激视网膜左上方的一些外围光感受器。眼球运动的最大速度是900度/秒，眼球几乎立即旋转以跟踪视觉目标。为了帮助神经系统进行视觉跟踪，眼球运动可以引起头部的辅助性运动。第二个作用是通过驱动头颈神经肌肉自主运动DNN来控制头部运动。如图10所示，输入层有10800维的ONV，输出层有2维来分别代表眼睛转动调整的度数，隐藏层有6层。

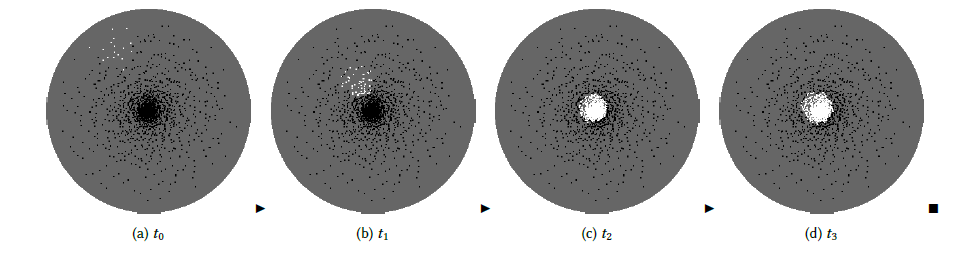


图9 光感受器在目标物体（白色球体）靠近时的变化

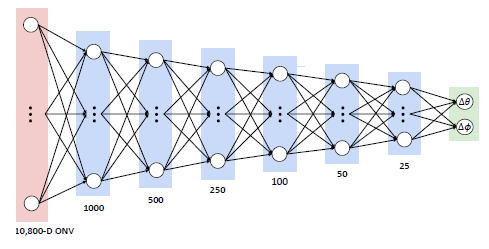


图10 前向视觉DNN

对于图4中编号为2,3,4,5,7,8,9,10的肢体视觉DNN，设计如下。左右肢（手臂和腿）视觉DNN是为了估算出在三维空间中肢体离视觉目标的位置，以此来驱动相关的肢体神经肌肉控制器接触目标。肢体视觉DNN与上面编号为1和6的DNN之间只有输出层的大小不同。如图11所示，一个固定的红色球体从右下方进入眼睛的视野，刺激视网膜左上方的外围光感受器，然后一个绿色的肢体（胳膊或腿）向球体伸展。同样，眼睛视网膜中的光感受器的响应是通过光线追踪3D场景来计算的，它们作为各自ONV的RGB分量输出。对于给定的ONV输入，期望训练网络的输出是现实空间中和目标位置不同维度之间的三个差异值Δx、Δy和Δz。重复将球体置于视野中任意位置，并在空间中随机驱动关节肢体，得到输入输出对的训练数据集。经过388次迭代后，反向传播DNN训练过程收敛为一个小误差，从而触发早期停止条件，避免过拟合。

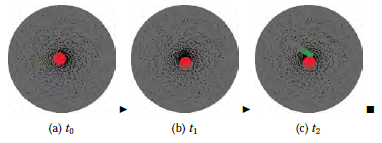


图11在一个手臂伸向一个移动的目标球时光感受器的变化

考虑到图4的组成部分，图12中的图说明了具体的感觉运动仿真迭代过程，具体如下：眼睛输出的ONV由视觉神经网络处理，来驱动到运动目标位置的眼球增量旋转。视觉神经神经网络的输出同时输入给神经肌肉运动控制器，这些控制器为头颈和四肢肌肉骨骼复合体产生肌肉激活信号。生物力学模拟器通过时间推进肌肉骨骼系统的状态，从而在外力(如重力)作用下递增地驱动五个复合体。三个主要的仿真组件可以异步运行。

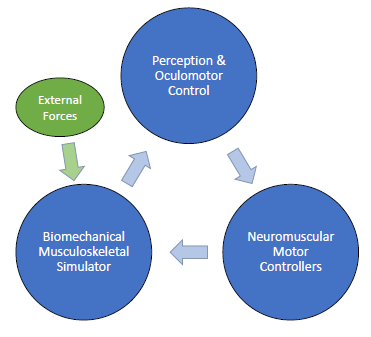
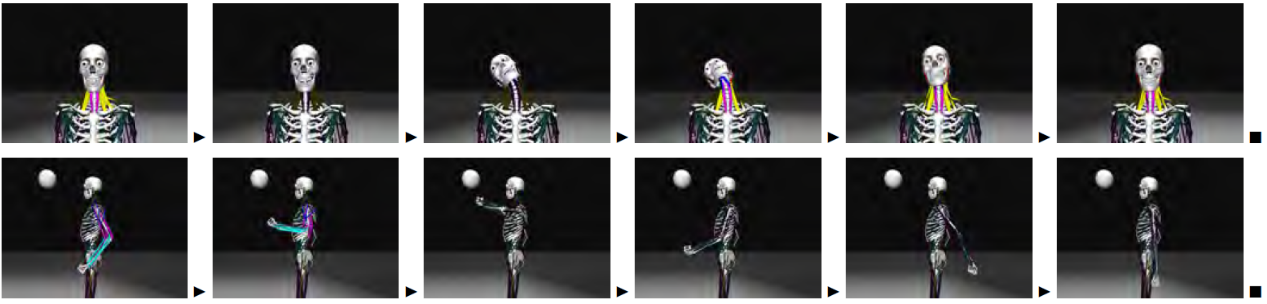


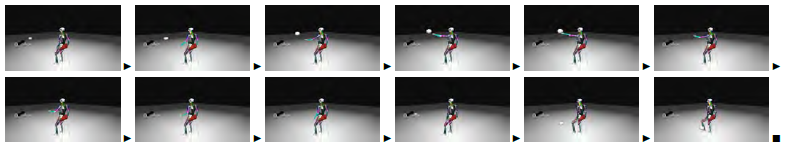
图12 感知运动仿真迭代过程

# 实验结果与分析

这篇文章的中所有的DNN都是使用Theano库[3]进行实现和训练的。实验环境如下：CPU为Intel Core i7、GPU为NVIDIA Titan X、OS为Ubuntu16.04。

图13(第1行)显示的是在神经肌肉运动控制器下模拟头颈肌骨骼复合体的一系列帧的序列。当控制器失活时，头部会由于重力而自然向后倾斜。控制器重新激活后，颈部肌肉将头部抬起、平衡在颈椎上，并控制头部运动。类似地，图13(第2行)显示了在经肌肉运动控制器下左臂肌肉骨骼复合体的模拟图，图中展示了手臂的完全动态控制。在图13(第3-4行)中有一个发射器向仿生人体模型发射小球，仿生人体模型会主动感知视网膜中央的小球。通过视觉DNNs处理来自眼睛的ONVs，使入球形成凹窝和视觉追踪，再加上配合头部运动控制器驱动的头部动作，以追求注视方向。同时，通过视觉神经神经网络的输入，肢体神经肌肉运动控制器能有效地控制手臂和腿部，截击来球。图13(第5-6行)显示了该仿生人体模型能主动执行到达动作拦截一个在多个倾斜平面上弹跳的球的动画中的一系列帧。眼睛和头部相互配合来追逐目标小球，而右臂也自然地跟踪球，直到它到达。头部、眼睛和手臂的运动由视觉目标的瞬时位置自动诱导，该位置由眼睛ONV输出的视觉神经网络视觉估计。眼睛做出快速的运动，快速固定移动的视觉目标并对其进行追踪，而头部则以一种自然的、合作的方式主动定向，尽管由于它的体积较大，动作较为迟缓。虽然球最初是遥不可及的，但我们的虚拟人会自动产生逼真的预备臂动作，以接近球。





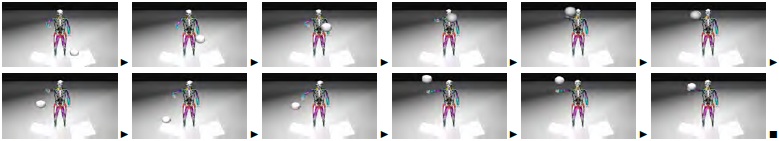


图 13 头颈部运动、用胳膊和腿拦截球

图14显示了坐在凳子向目标球伸展手臂上的一系列帧。图中蓝线表示从视网膜光感受器位置发射的光线，用于对3D场景进行采样，计算光感受器处的辐照度。首先，球被眼睛感知，然后由视觉神经中枢神经系统DNN处理，通过眼球运动与由头颈神经肌肉运动控制器控制的肌肉驱动的头部运动相结合，形成跟踪。手臂神经肌肉运动控制器对肌肉驱动的伸展运动进行主动控制。

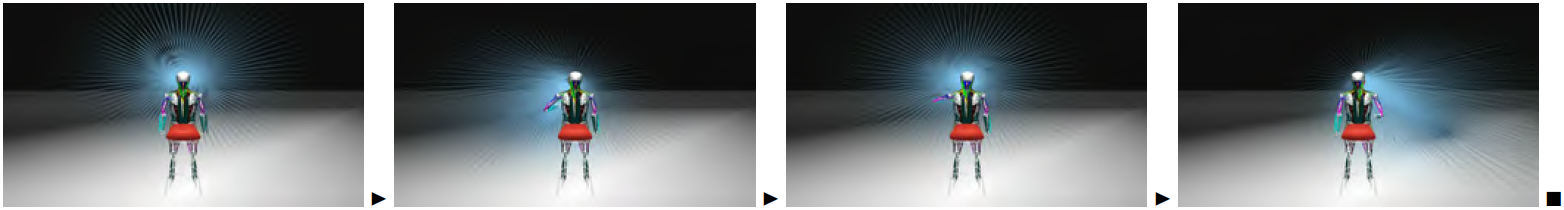


图14 通过仿生眼睛进行视觉采样

为了研究生物力学人体肌肉骨骼模型的感知运动系统的稳健性，这篇文章还考虑了神经肌肉运动控制器是否能够应用到还未经过专门训练的情况。首先，通过应用一个连续的正弦位移函数，对凳子的三维位置进行了垂直和水平的运动干扰。对于振幅为10cm，频率为3.75 Hz的扰动，发现感知运动系统仍然能够稳健地工作，而对于振幅为70cm，频率为6.0 Hz的干扰，感知运动系统往往无法用胳膊和腿拦截运动的球。第二，当该仿生人体在仰卧位时，神经肌肉足够强大的控制器在不经过训练的情况下仍能够计算结合各个方位的重力。图15显示了该模型坐在办公椅上，不经过任何训练，连续上下倾斜，书写“SIGGRAPH 2018”的动画画面。



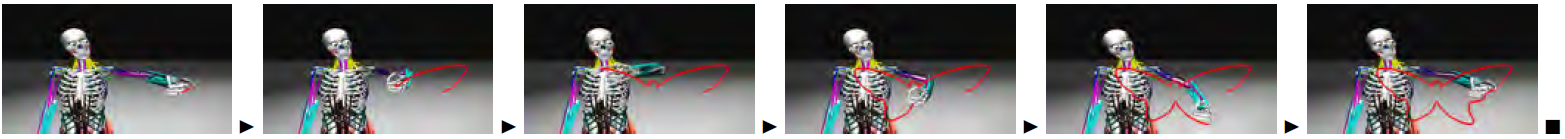


图15 写字和绘画能力

# 思考与扩展建议

这篇文章的原型感觉运动系统中的DNN是进行离线训练的，而后会直接应用于在线感觉运动控制的任务中，这无疑是不太合适的。为了解决该问题，还需要在模型中加入一个在线的、可持续的、更深层的强化学习方案。

这篇文章目前的眼睛模型基于理想的针孔相机和均为的运动旋转球体，为了进一步创造一个更逼真的眼睛模型，我认为还需要考虑到以下情形：1）针对不同的入射光强度，需要微调针孔相机的孔径，来模拟瞳孔的扩张和收缩，还需要通过角膜和镜片等组件来折射光线。2）根据生物力学眼模型，正常人体眼球质量约为7.5克，由眼外肌驱动，而不仅仅是这篇文章提到的4个直肌和2个斜肌。而且根据调研，通过直肌计算、通过斜肌控制凝视方向周围的扭转运动这两部分的工作还有待改进[1, 2]。

虽然这篇文章工作已经证明，在感觉运动控制中最为关键的一步是感觉运动任务，这篇文章也足以直接从感光器响应映射到控制输出，但为了执行更多复杂的任务，仍然需要引入更多的可视化处理阶段。为此，我认为可以适当增加视网膜中光感受器的数量，并尝试不同的非均匀感光器的分布。

最后，我认为这篇文章可以继续完善，来建立一个完整的生物力学肌肉骨骼模型，添加额外的神经肌肉运动控制器，来完成包括以下一些文中未提及但我认为可能有关的一些工作：1）控制其整个脊柱和躯干，使其能够直立、自动保持重力平衡；2）能够独立执行各种主动视觉任务，包括识别物体、提取特征、区分物种等。

# 参考文献

[1]. Lesmana M, Pai D K. A biologically inspired controller for fast eye movements[C]//Robotics and Automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 3670-3675.

[2]. Lesmana M, Landgren A, Forssén P E, et al. Active gaze stabilization[C]//Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing. ACM, 2014: 81.

[3]. J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley, and Y. Bengio. 2010. Theano: A CPU and GPU math compiler in Python. In Proc. 9th Python in Science Conference. Austin, TX, 1–7.