采用深度学习的先进机器人抓取系统

摘要—由于机器人手和物体建模的不确定性、未知的接触类型和物体刚度等特性，机器人手抓取物体具有一定的挑战性。为了克服这些挑战，实现机器人手的数学模型，对物体和物体与手的接触进行建模是必不可少的。在已知物体特性的前提下，建立了耦合系统的智能手-物体接触模型。在Matlab Simulink/ simmechanical、神经网络工具箱和计算机视觉系统工具箱中对控制进行仿真。

1. 介绍

机器人技术正朝着允许机器人进入我们日常生活的技术研发方向发展。最理想的机器人助手应该与人共享一个环境，能够以一种非常友好的方式应对人的存在和互动。要创建这样的应用程序，需要解决许多问题，包括调换日常任务中使用的动作，以及找出如何解释人类互动，以及如何使用所有这些知识来创建能够成功充当助手的机器人。

拥有智能机器人的需求意味着编程的复杂性必须大大降低，机器人的自主性必须变得更加自然。这一挑战与新一代机器人有关，它们必须与人互动，并在人类环境中运行。

在现实世界中，有许多类型的物体可以被人类的手操纵。为了确保机器人有足够的技能与我们的环境互动，我们需要实现人手的技能。自主机器人的手需要适应不同情况下的各种抓取任务。为了解决这样复杂的问题，需要人工的认知技能，使机器人平台能够为执行每一个特定的任务做出决策，并适应人类环境。

由于未知的物体和姿势，机器人抓取极其困难。Martinez和Collet制作了框架mope (multi Object Pose Estimation and Detection)，解决了目标识别和姿态估计问题。

以人手作为末端执行器的机械手在目前和未来的应用中都将非常有效。为了减少传统抓取器的限制，提高机械手执行抓取和操作任务的效率，首次引入了多指手。索尔兹伯里首先成功地设计了一种类似人的机械手，这是一种先进的末端执行器，用于抓取研究。从那时起，许多机器人手被设计用于工业、危险和偏远的地方，如发电厂和太空应用。犹他麻省理工学院手，DLR手和巴雷特手是一些例子，广泛应用于工业和作为抓取和操纵研究平台。为了提高机械手执行抓取任务的效率，对机械手的建模和接触控制进行了研究。

抓取研究关注通过确保在接触过程中物体与手之间不发生运动，提高机械手灵巧地抓取多种形状和刚度物体的能力。在全指操作中，手指的各个环节都与物体发生接触，给手-物体动力学带来了更大的复杂性。

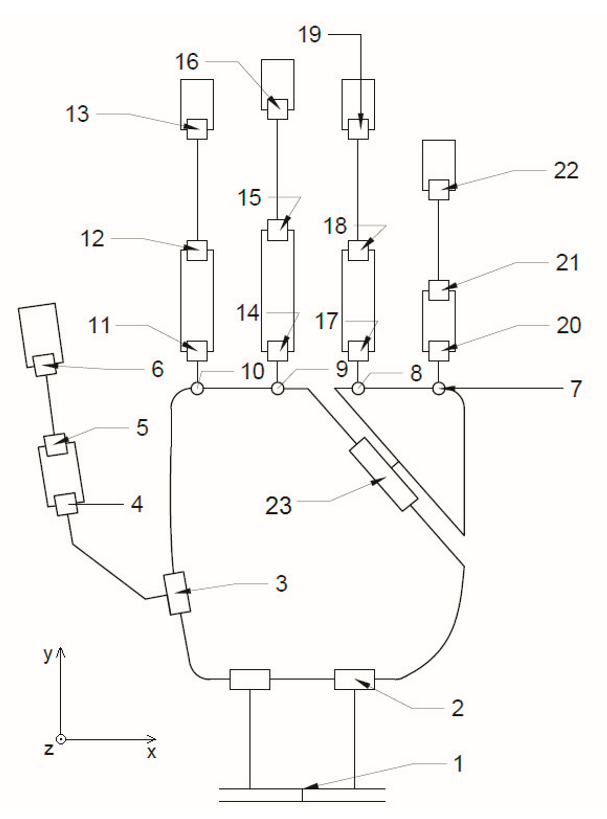
为了提高多指手抓取的效率和灵巧性，必须解决手的建模、手-物体接触建模和控制等问题。手的抓握性能取决于物体属性的准确信息，手的造型，手与物体之间的接触造型，根据物体位置和物体与手的接触点位置把握规划，根据手-物体动力学原理对手的接触力进行估计和控制。考虑到上述问题，多指手抓取物体的过程分为以下四个步骤: **1. 对象的位置。**目标是通过视觉传感器来定位目标的位置。目标识别、视觉跟踪和三维姿态估计是环境中目标识别的前沿技术。**2. 抓握计划。**根据安装在指尖和电机上的传感器数据计划抓取。在这个阶段，需要定义接触点的个数来抓取物体。抓握是根据这些接触点来计划的。多个接触点导致复杂的抓取规划。**3.接触模型。**接触模型对于了解手与物体之间的接触位置、估算接触力以及避免摩擦引起的任何滑移都是必不可少的。**4. 抓握控制。**当手对象模型先验未知时，抓取控制是复杂的。接触类型和接触次数是控制抓握的因素。

控制任务分两步进行:首先，手到达想要的位置，然后与物体接触;其次，根据物体的刚度，需要在物体上施加一个理想的力，使其不发生滑移。本文选择三指手作为研究机械手抓取物体的对象。

1. 一般问题公式化

抓握的主要任务是物体和手之间的相互作用。让我们考虑一个具有已知坐标和机械手手指的固定对象。抓握的任务可以分为两个截然不同的问题。第一个任务是命令手到达目标位置。接触力是手接触物体时产生的，需要处理。接触力需要有界，以防止任何滑移或损坏的对象。

**抓取的手运动学**

人手运动模型如图1所示。采用23自由度的结构进行运动学分析。

2.1．正向运动学模型

首先，根据标准的Denavit-Hartenberg约定，为手的每个手指k的每个连杆分配一个坐标系。接下来，为每个坐标系定义一组Denavit-Hartenberg参数。利用Matlab，可以计算出这些项。

* 1. 逆运动学模型

角θk2和角θk3可以通过考虑到手指的两个指骨包含在一个平面kII中来计算。

因此，将余弦定理应用于BCD三角形，可以得到如下公式来计算角度θk3:

最后，由图3中角α和角β的差得到角θk2:

图片包含 文字, 地图, 天空

描述已自动生成θ=α-β=

* 1. 逆动力学模型

对于每个手指k，其逆动态模型可以表示为:

图片包含 物体, 时钟, 手表

描述已自动生成

1. 简化仿人机械手CAD模型

通过运动学分析，建立了简化(三指)人形手的三维模型。为了完成这项任务，使用了Autodesk Inventor软件。在CAD模型中，我们没有假设执行器。CAD模型(图3)仅供我们更好地想象和观察可能的运动和手工操作。

可以将Autodesk Inventor中的模型导入Matlab Simmechanics进行进一步的分析。我们使用了smlink\_linkinv工具。导入成功后，我们还获得了Simulink模型的手部模型。

图片包含 玩具

描述已自动生成

图 3 仿人机器人手的CAD模型

在Matlab Simulink模型中，我们可以看到从Autodesk Inventor导入的仿真机械手的部分(图3)。

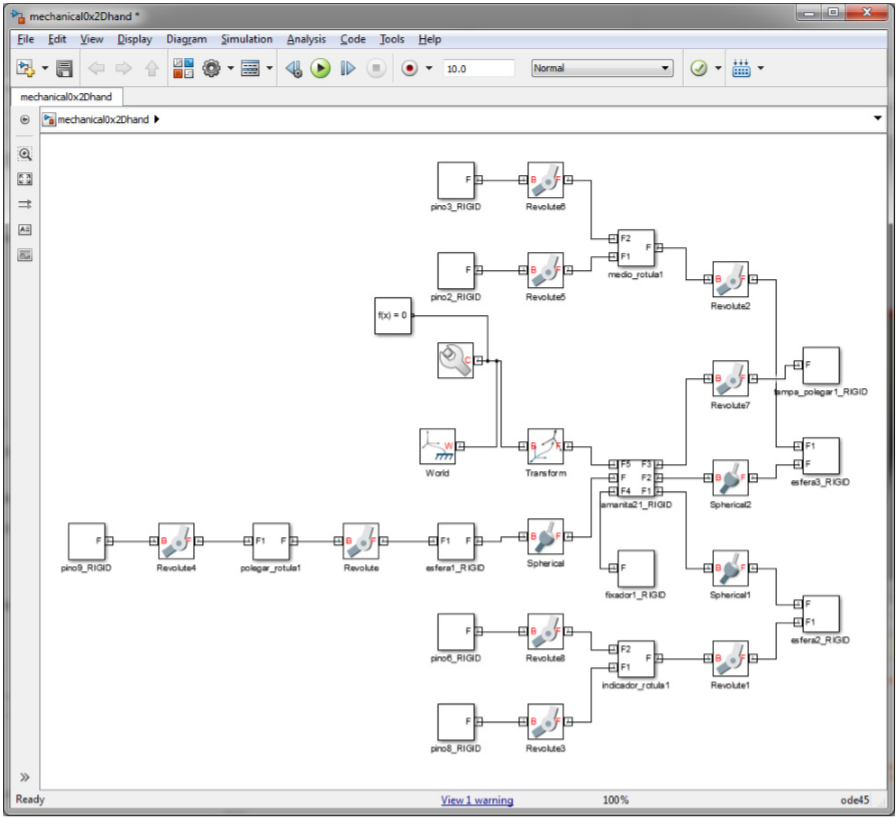


图 4 导入Matlab Simulink仿人机械手模型

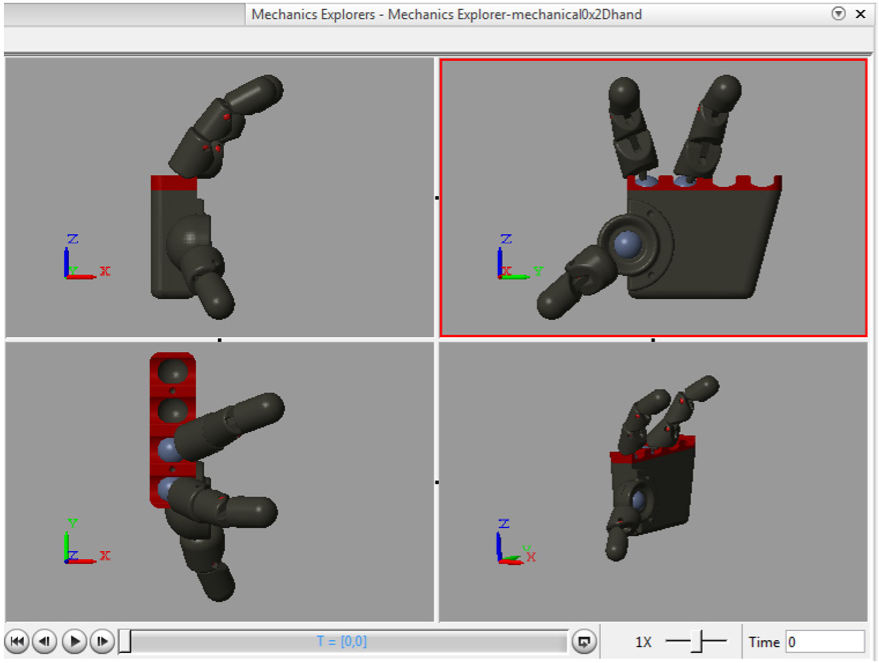


图 5 Matlab力学资源管理器中仿人机械手的三维模型

1. 手臂控制，目标检测，识别和姿态估计

在执行操作任务期间，可能会发生对象与环境的意外或预先计划的交互。控制的主要目的是确保机器人系统不会失去目标，并且交换的力仍然有限。关键的一点是控制物体和手指之间的接触力。保持接触力在一定范围内是很重要的，有几个原因。一方面，这些力必须足够高，以保证满足摩擦锥的约束; 另一方面，接触力不能过高，以避免电机饱和和能量浪费，以及保护材料。在这种情况下，指尖力传感器的存在对手臂和手的控制起着重要的作用。

视觉目标检测是机器人抓取过程中最重要的步骤。目前已报道了许多方法。由于几乎所有的目标识别方法都严重依赖于前景目标检测的准确性，因此需要高效可靠的方法。

4.1. 人工神经网络简介

神经网络是受生物神经网络功能启发而建立的一种计算模型。人工神经网络由一组人工神经元组成，这些神经元通过相互连接来处理信息。尽管有许多不同版本的神经网络，但它们都具有相似的特性。尽管有许多不同版本的神经网络，但它们都具有相似的特性。它们有许多神经元，这些神经元与其他神经元相连。一个典型神经元的方案如图6所示。

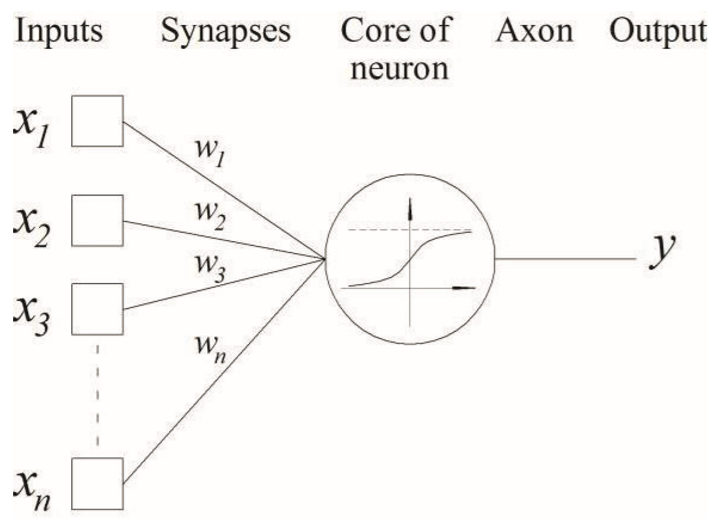


图 6 神经元的结构

人工神经元有连接神经元输入和核心的突触。其次是处理输入信号的核心。它还有轴突连接到下一层的神经元。每个突触都有一个权重，它定义了神经元的输入如何影响其状态。

神经元的状态定义为:

其中：

n—神经元的输入数;

—第i个神经元的值；

—第i个突触的权重。

轴突的值计算公式如下: Y = f(S)

其中f是一个激活函数。我们使用高斯激活函数，如下图所示：

在机器学习和相关领域，人工神经网络(ANNs)是一种计算模型，灵感来自动物的中枢神经系统(尤其是大脑)，它能够进行机器学习和模式识别。人工神经网络(图7)通常表示为相互连接的“神经元”系统，这些“神经元”可以从输入计算值。

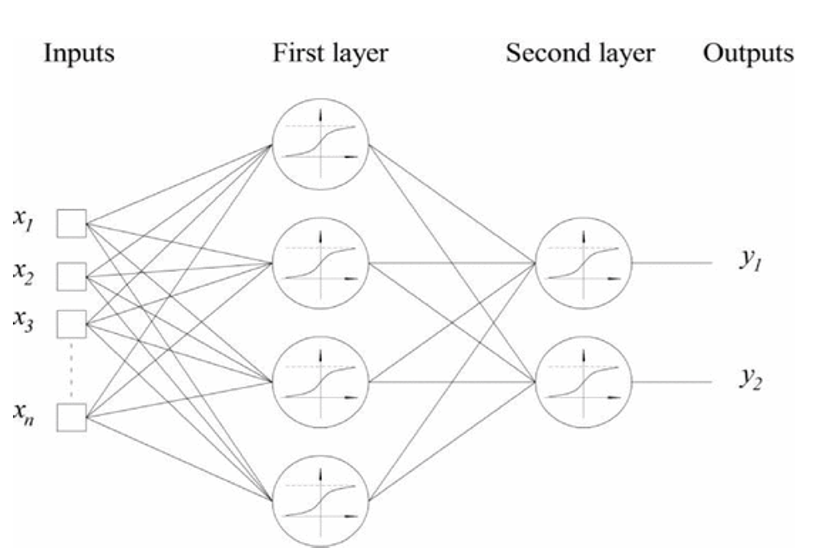


图 7神经网络的结构

* 1. 深度神经网络与深度学习简介

在机器学习中，深度信念网络(deep belief network, DBN)是一种生成图形模型，或者说是一种深度神经网络，由多层潜在变量(“隐藏单元”)组成，各层之间有连接，但各层之间没有连接。深度学习是机器学习中的一组算法，它试图通过使用由多个非线性转换组成的模型体系结构来对数据中的高级抽象进行建模。各种深度学习体系结构，如深度神经网络、卷积深度神经网络、深度信念网络等，已被应用于计算机视觉、自动语音识别、自然语言处理、音乐/音频信号识别等领域，并在这些领域中对各种任务产生了最先进的结果。

* 1. 卷积神经网络(CNN)

CNN由一个或多个卷积层组成，这些卷积层的顶部是完全连接的层(与典型的人工神经网络中的层匹配)。它还使用绑定权重和池化层。这种架构允许CNNs利用输入数据的2D结构。与其他深层结构相比，卷积神经网络在图像和语音应用方面都开始显示出较好的效果。它们也可以用标准的反向传播进行训练。与其他常规的深度前馈神经网络相比，CNNs更容易训练，而且需要估计的参数更少，这使得CNNs成为一个非常有吸引力的体系结构。

与传统的基于深度学习模型的目标识别方法相比，本文重点研究了包括目标识别在内的目标姿态估计问题。深度学习方法通过学习小组模式的稀疏特征，具有识别或预测大组模式的能力。利用这一优势，我们可以使用一组小的姿态来训练深度学习模型，然后用该模型预测一组大的姿态。

对于一般目标识别和图像分类任务，卷积神经网络(CNNs)的变种已经成为机器人的监督特征学习和分类工具，尤其是与max-pooling (MPCNN)结合使用时(图8)。

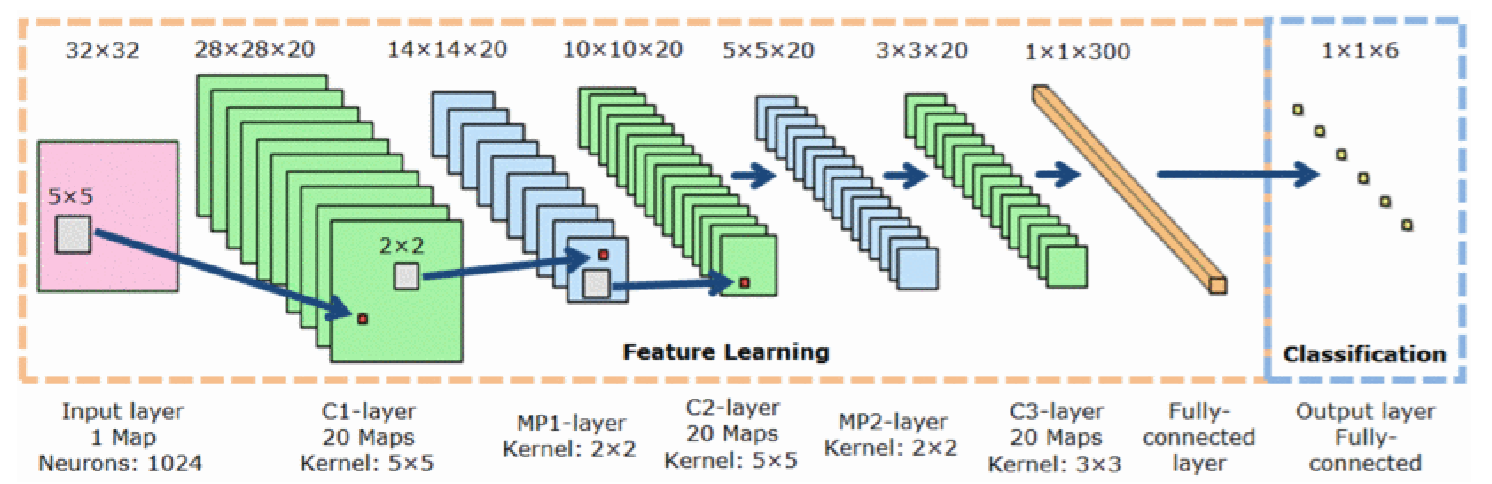


图 8 使用交替卷积和最大池层的MPCNN架构

MPCNNs包括卷积层和子采样层。MPCNNs根据卷积层和子采样层的训练和实现的不同而有所不同。

卷积层

卷积层的参数为:映射数、映射大小和内核大小。每个层(L)包括映射(M)。在输入图像的有效区域上移动一个大小为K的内核。Ln层中的每个映射都连接到Ln-1层中的所有映射。给定映射的神经元共享它们的权重，但具有不同的输入字段

最大池化层

最大池化层的输出由非重叠矩形区域上的最大激活量决定。最大池提高了泛化性能

分类层

为了完成MPCNN，我们使用了一个浅的多层感知器(MLP)。在分类任务中，输出层每个类有一个神经元。

1. 实验

5.1. 目标检测

该系统由Matlab仿真机器人手场景模型和仿真摄像机组成。虚拟对象在视觉系统的范围内，通过Matlab计算机视觉工具箱进行识别，并实现了MPCNN模型。

输入图像来自RGBD相机数据，相对于简单的二维图像数据，已经证明可以显著提高抓取检测结果。

视觉目标检测是机器人抓取的第一个关键动作。需要使用或开发可靠的方法来有效地识别存在于视觉系统输入端的前台对象，并从后台检测这些对象。

一种可能的方法是使用稀疏编码或k -均值聚类。第一步是构建对象字典。聚类使我们能够将带有背景和对象的图像的颜色组件分组。每个组件在特征空间中都有一个位置，重要的是要找到分区，以便每个集群中的组件彼此尽可能地靠近，并且尽可能地远离其他集群中的组件。

* 1. 目标识别和姿态估计

通过仿真，我们利用深度学习的方法实现了目标位姿估计和位姿估计，该方法通过学习小组模式的稀疏特征，具有识别或预测大组模式的能力。利用这一优势，我们可以使用一组小的姿态来训练深度学习模型，然后用该模型预测一组大的姿态。

* 1. 机器人抓取

这一阶段介绍了一种通过改变夹持器的位置和方向来抓取物体的系统。

图片包含 室内, 墙壁

描述已自动生成

图表 1仿人机器人手部抓取物体的三维模型

1. 结论

前面部分所述的开发模型已经在选定的Matlab工具箱(Computer Vision System Toolbox、Deep Learning Toolbox和simmechanical)中实现。

本文提出了一种多指机械手抓取任务的仿真模型。该模型包括运动学、动力学、物体表示和接触模型。接触模型决定了机械手对物体施加的力。目标检测、目标识别和机械手姿态估计都是基于最常用的深度学习模型之一的最大池卷积神经网络。使用深度学习可以避免手工工程特性，而是学习它们。