

А. С. Ющенко, д-р техн. наук, проф., arkadyus@mail.ru, **Шуай Инь**, аспирант, shuai.yin@yandex.ru,
Московский государственный технический университет им. Н. Э. Баумана
(национальный исследовательский университет)

Диалоговое управление коллаборативными роботами с помощью искусственных нейронных сетей

Рассмотрен коллаборативный робот как мобильное робототехническое устройство, способное по речевой команде человека выполнять те или иные действия, причем не обязательно в том же рабочем пространстве, в котором находится сам оператор. При этом предполагается двусторонний диалог человека и робота в целях уточнения задачи, ситуации, состояния робота и, возможно, самого оператора.

Комплексная задача управления, а скорее, задача взаимодействия человека с таким искусственным партнером потребовала применения новых средств управления, распознавания ситуации, ведения речевого диалога. В качестве средств, позволяющих не только решить весь комплекс задач, но и провести предварительное обучение робота, в этой работе выбрана комбинация нейронных сетей различного типа — конволюционных для распознавания образов, глубоких сетей для распознавания речи, нейросетей LSTM для автономного управления движением робота по наблюдаемой ситуации.

Работы в области управления мобильными и манипуляционными роботами, в том числе с участием человека, в течении ряда лет проводились на кафедре "Робототехнические системы и мехатроника" МГТУ им. Н. Э. Баумана, отмечающей свое семидесятилетие. Читатель может познакомиться с этими работами по библиографическому списку. По существу, результатом всех этих работ является законченная модель сервисного манипуляционного робота, которая может найти самое широкое применение.

Ключевые слова: мобильный коллаборативный робот, нейронная сеть, диалоговое управление, конволюционная сеть LSTM, планирование траектории, управление движением, распознавание образов, SLAM

Введение

Коллаборативная робототехника все активнее входит в нашу жизнь. Основными ее особенностями является возможность безопасного взаимодействия робота с человеком в общем рабочем пространстве и возможность такого взаимодействия без специальной подготовки в области робототехники. Примерами таких систем могут быть коллаборативные робототехнические системы (КРТС), используемые в медицине для помощи хирургу при выполнении хирургических операций, для реабилитации людей с ограниченными возможностями, для помощи инвалидам, для оказания поддержки в работе космонавтов, в военном деле и т. п. В настоящее время получили практическое применение коллаборативные манипуляторы, которые работают в общем рабочем пространстве с оператором и безопасны для него. Но коллаборативными могут быть и мобильные роботы (КМР), которые перемещаются в сложном пространстве, в котором могут находиться препятствия и другие участники движения, в том числе люди. В более общем случае КМР включает в себя и подвижную платформу и манипуляторы, что позволяет реализовать принципиально более широкий

круг задач, например, человек может попросить робота принести нужный предмет или выполнить несложную операцию. Будем предполагать также, что КМР обладает диалоговым интерфейсом, обеспечивающим человеку возможность двустороннего речевого диалога в рамках ограниченного профессионально-ориентированного языка, близкого по своему синтаксису к естественному. Теперь КМР становится полноправным партнером человека.

Решение такой задачи, определяющей новый шаг в разработке КРТС, требует применения нового аппарата, который в равной мере был бы применим для решения всего комплекса возникающих задач, включая распознавание речи и организацию диалога, распознавание образов по речевому маркеру, выполнение манипуляционных операций, наконец, перемещение в ограниченном пространстве с не определенными заранее препятствиями, в том числе движущимися. Универсальным средством решения всего комплекса задач сегодня являются обучаемые нейронные сети различного типа и их комбинированное применение при решении поставленных перед КРТС задач.

Задача управления КРТС манипуляционного типа рассматривалась в ряде работ, в том числе выполненных в МГТУ им. Н. Э. Баума-

на, и опубликованных ранее. Ниже мы кратко остановимся на полученных результатах и уделим основное внимание задаче диалогового управления мобильными роботами с использованием нейронных сетей, которая представляется актуальной при организации диалогового управления КМР широкого назначения.

Диалоговое управление коллаборативным мобильным роботом

Мобильные роботы, дистанционно управляемые человеком, находят широкое применение в задачах инспекции территорий и помещений, при спасении людей после аварий и катастроф, при реабилитации людей с ограниченными возможностями к передвижению и т. д. Во многих случаях управление такими роботами происходит дистанционно, в непрерывном режиме контроля и управления с помощью средств визуализации и компьютерной поддержки действий оператора. Эволюция таких систем заключается в постепенной автоматизации процесса управления, в результате которой создаются системы управления, в которых человек только ставит задачу роботу. При этом робот должен самостоятельно решать такие задачи, как определение собственного положения в пространстве, маневрирование при появлении новых, в том числе движущихся, объектов. Если перемещение робота происходит в не полностью определенной обстановке, то всегда могут возникнуть проблемы, решение которых не было заранее предусмотрено. Поэтому в таких случаях целесообразен диалог с человеком-оператором, который поставил задачу и контролирует ее выполнение роботом.

Учитывая, что речь идет о коллаборативных системах, мы не предполагаем никаких специальных знаний пользователя в области робототехники. Диалог должен вестись в рамках профессионально-ориентированного языка, доступного пользователю и близкого к естественному. Заметим, что такой подход, названный ранее диалоговым управлением [1, 2], подразумевает визуальное распознавание роботом окружающей обстановки, объектов, препятствий и сообщение об обстановке пользователю на том же языке. Таким образом, робот должен быть оснащен системой распознавания не только речи, но и визуальных образов, а также системой планирования маршрута движения.

Уточним, что мы понимаем под термином "коллаборативный мобильный робот". Это робот, в общем случае действующий в условиях частично определенной среды, модель которой может быть предварительно построена тем же роботом с использованием технологии SLAM (Self Localization and Mapping) при участии оператора форме диалога. Робот выполняет команды оператора, определяющие цель движения в данной среде, самостоятельно решая задачи планирования маршрута, обхода неожиданных препятствий, в том числе динамических, обращаясь к оператору с вопросами в тех случаях, когда решение не определено или неоднозначно. Робот снабжен манипулятором, позволяющим выполнить задание оператора во внешнем мире.

Отличие КМР от КРТС в прежнем понимании заключается в том, что робот и оператор функционируют в разных пространствах. Однако при этом вовсе не исключается взаимодействие с другим человеком-партнером, которому, например, надо что-то передать. Остается задача распознавания объектов внешней среды, в которой робот должен "узнавать" и своего "хозяина" и "партнера", а также относить к посторонним объектам других людей, которые могут встретиться в процессе движения и которых нужно обходить.

Вопросы организации диалогового управления коллаборативными манипуляционными роботами рассматривались нами при разработке робота-ассистента хирурга [3–5]. Ряд проблем, связанных с реализацией диалогового управления мобильными роботами, могут быть решены в данном случае идентично. Это относится к системе распознавания речи и организации диалога, к подсистеме поиска и распознавания объектов с помощью системы компьютерного зрения, наконец, к системе распознавания положения самого оператора, поскольку во многих случаях речь идет об обслуживании пользователя. Эти блоки функциональной схемы системы управления КМР показаны на рис. 1. В случае мобильных роботов задача существенно усложняется тем, что робот должен также решать задачу SLAM, т. е. самостоятельно строить "карту" местности, определять собственную локализацию на этой карте и выполнять поставленную оператором задачу в построенной модели среды.

При традиционном подходе к управлению автономными мобильными роботами решение

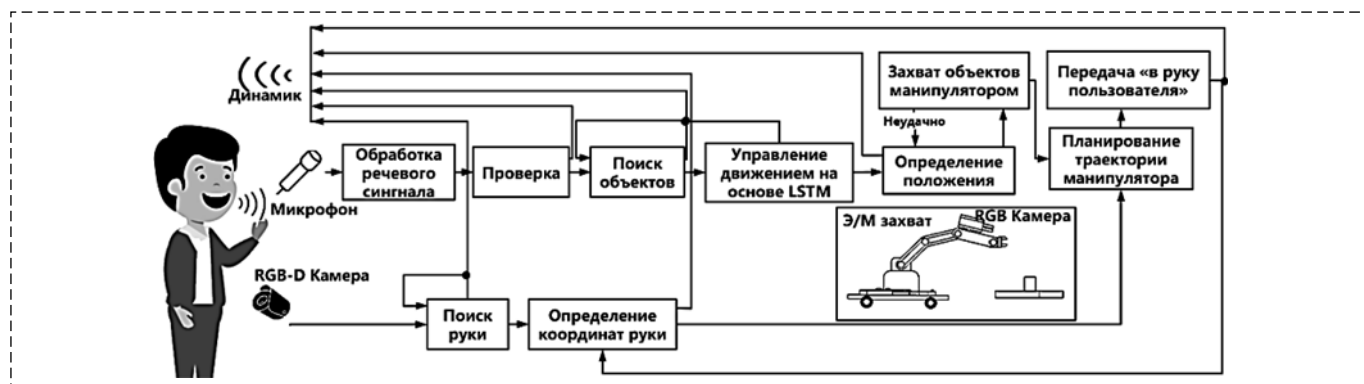


Рис. 1. Функциональная схема коллаборативного мобильного робота

Fig. 1. Functional diagram of a collaborative mobile robot

этой задачи разбивается на три этапа: построение карты местности при автономной навигации, планирование движения и реализация движения с помощью системы управления робота [1]. Такой подход требует много времени, что в ряде случаев неприемлемо. В связи с этим в данной работе предлагается применить convolutional LSTM (Long Short-Term Memory), позволяющую решать задачу управления непосредственно по текущим наблюдениям и по заданной цели движения, подобно тому, как это делает человек. При этом управление роботом может сопровождаться диалогом с пользователем с целью уточнить наблюдаемые объекты и цели движения, выбрать наилучший из возможных маршрутов движения при появлении препятствий. После выбора маршрута система должна самостоятельно генерировать сигналы управления угловой и линейной скоростями робота в соответствии с задачей, имеющимися ограничениями и текущими наблюдениями. Комплексирование задач, которые приходится решать последовательно при традиционном походе к управлению, позволяет существенно снизить требуемые для этого вычисления и обеспечить решение задачи диалогового управления в реальном масштабе времени.

Реализация такого подхода предъявляет ряд требований к структуре системы управления КМР. Поскольку задача ставится пользователем путем речевого сообщения и возможного диалога, то система должна обладать возможностью надежно распознавать голосовые сообщения и правильно их интерпретировать в соответствии с построенной или заданной заранее картой. При построении маршрута система должна генерировать возможные варианты маршрута и снабжать их комментарием, облегчающим выбор пользователю. Для создания самой 2D- или

3D-карты и определения собственного положения робота он должен быть, в общем случае, оснащен RGB-D-камерой или иным аналогичным устройством, обладающим возможностью кругового обзора окружающей среды. Сеть LSTM должна непосредственно управлять движением робота по наблюдаемому образцу ситуации. Иными словами, входной сигнал сети LSTM представляет собой RGB- или RGB-D-образ ситуации с указанием текущего положения робота и цели, а выходные сигналы — это сигналы управления линейной и угловой скоростями платформы мобильного робота.

Организация распознавания речи и диалога "пользователь—робот" с использованием нейронной сети

В соответствии со схемой на рис. 1 первая задача, которая должна быть решена при диалоговом управлении, это собственно реализация речевого диалога между человеком и роботом. Структура системы диалогового управления включает четыре блока: распознавание естественной речи, "понимание" речи, преобразование текстовой информации в инструкции мобильному роботу, речевая обратная связь с информацией о ситуации и получение разрешений на выбранные действия (рис. 2).

Схема блока распознавания естественной речи, основанная на применении нейронных сетей и удовлетворяющая сформулированным требованиям, показана на рис. 3. Она включает словарь используемых терминов профессионально-ориентированного языка, акустическую модель и речевую модель [4].

Акустическая модель предназначена для распознавания не только отдельных слов, но и от-

дельных фраз, предложений. Для ее реализации целесообразно применять глубокие нейронные сети (Deep Neural Network, DNN), которые содержат не менее пяти скрытых слоев. Однако нейронные сети не способны непосредственно моделировать голосовые сигналы. Поэтому, чтобы использовать способность нейронных сетей к классификации, целесообразно рассмотреть комбинированную акустическую модель, которая состоит из глубоких нейронных сетей DNN и скрытой марковской модели HMM (Hidden Markov Network). Последняя используется для описания динамики голосовых сигналов. Обучение нейронной сети проводится с помощью алгоритма обратного распространения ошибки. Акустическая модель с использованием модели DNN-HMM способна распознавать естественную речь в реальном масштабе времени, что особенно важно при разработке робототехнической системы коллаборативного типа.

Речевая модель в структуре, показанной на рис. 3, служит средством распознавания речи. Для решения этой задачи целесообразно выбрать речевую модель n -gram, которая позволяет вычислять вероятность следующего слова по словам, которые появились раньше. Иными словами, речевая модель n -gram способна про-

гнозировать состояние по предыдущим n состояниям.

Полученная из акустического блока текстовая информация поступает в блок **"понимания" естественной речи** (см. рис. 2), задача которого состоит в том, чтобы определить поставленную роботу задачу по ключевым словам. Так, если обнаружены глаголы "захватить", "взять", то определяется задача захвата. При этом заранее определяется условие, что эта задача может быть выполнена, только если объект находится в рабочей зоне манипулятора. В противном случае определяется задача движения к объекту. Если обнаружены глаголы "двигаться", "обходить" и т. п., то определяется задача движения. В общем случае задача описывается фреймом, который можно представить в следующем виде: <объект управления> <имя операции> <сопутствующие объекты> <условия выполнимости операции>. Объектом управления по умолчанию является сам мобильный робот, возможности и ограничения которого имеются в базе данных. Эти возможности определяют условия выполнимости операции с учетом текущей ситуации и свойств рабочей среды. В случае отсутствия необходимой информации робот запрашивает об этом оператора. Заметим, что при этом могут

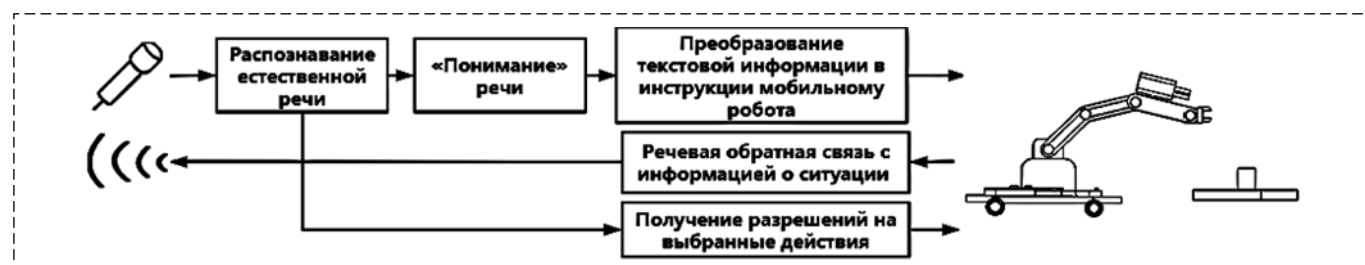


Рис.2. Структура системы диалогового управления КМР

Fig. 2. The structure of the dialogue management system of the CMR

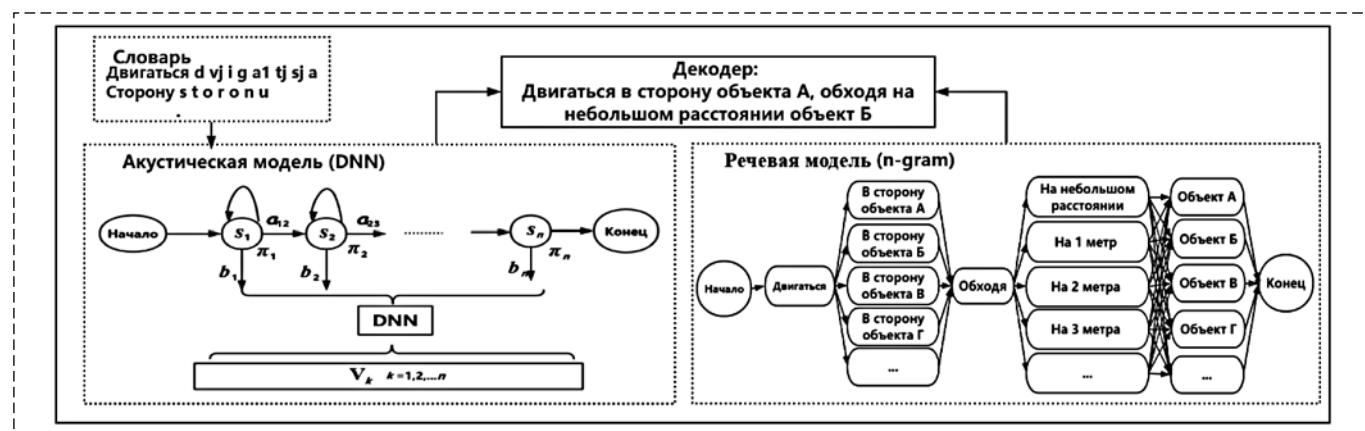


Рис. 3. Структура системы распознавания естественной речи с использованием глубоких нейронных сетей

Fig. 3. The structure of a natural speech recognition system using deep neural networks

быть использованы лингвистические переменные [5]. Условия выполнимости операции могут включать в себя и проверку постусловий, которые должны быть выполнены после завершения операции. Таким образом, в результате работы системы "понимания" речи заполняются слоты фрейма операции, которые, в свою очередь, преобразуются в команды управления роботом. Если оператор не определил те или иные слоты, формируется речевой запрос оператору для уточнения задачи.

Интерфейс оператора должен содержать ряд сценариев диалога, которые используются в зависимости от задачи робота и текущей ситуации. Это может быть запрос о состоянии внешней (по отношению к роботу) среды, включая новые препятствия, или о состоянии робота, полученных повреждениях, запасе энергопитания и т. п. Отметим, что для реализации управления диалогом может быть успешно применен подход, основанный на использовании модифицированной сети Петри [6].

Решение задач локализации, картографирования и генерации траекторий мобильного робота

Задача, условно называемая картографированием (mapping), представляет собой преобразование показаний датчиков робота в трехмерный образ окружающей среды. В большинстве случаев задачи картографирования и позиционирования, т. е. определения текущего положения робота, решаются одновременно.

Связь между наблюдаемыми координатами точек на двумерном изображении и соответствующими трехмерными координатами в абсолютной системе координат выражается формулой

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \quad (1)$$

Здесь u, v — координаты точки (пикселя) на двумерном изображении; f_x, f_y, c_x, c_y — внутренние параметры камеры; $r_{11}, \dots, r_{33}$ — параметры однородной матрицы, определяющие поворот камеры относительно сцены; t_x, t_y, t_z — параметры, определяющие поступательное перемещение; наконец, x, y, z — это координаты соответствующей точки в инерциальной системе координат.

Заметим, что последовательность параметров, определяющих вращение и поступательное перемещение датчиков, установленных на роботе, в процессе его движения определяет процедуру так называемой визуальной одометрии [7]. С помощью визуальной одометрии удастся решить и ряд других задач управления КМР, например, задачу возвращения робота к своему "хозяину" при потере связи [8]. Существуют и другие методы одометрии, в том числе методы, комбинирующие информацию от двигателей, гироскопов и других сенсоров [9], методы визуальной одометрии, основанные на выделении характерных точек на изображении внешних объектов [10]. Поскольку задача построения карты обстановки достаточно хорошо известна, мы не будем на ней останавливаться более подробно, тем более что во многих случаях, в том числе при работе в известном помещении, эта карта известна априори.

На построенной карте может быть отмечено собственное положение робота, которое определяется с помощью собственной сенсорной системы, и цель движения, заданная оператором и формализованная системой распознавания речи. Теперь задача состоит в построении маршрута движения с минимальной функцией "стоимости", в качестве которой обычно используется длина траектории. При традиционном планировании обычно используются сплайны третьего порядка, позволяющие управлять скоростью робота. Для улучшения качества движения, включая плавный разгон и плавное торможение, применяются сплайны четвертого порядка. Заметим, что эта задача аналогична планированию траектории манипулятора с учетом его динамики [11]. По полученным сплайнам формируются и управляющие сигналы двигателей робота с учетом возможностей приводов и оптимизации движения [12, 13]. Задача управления может быть решена и при управлении группой роботов в ограниченном пространстве [14].

В данной работе для решения задачи автономной навигации робота и для выбора нужного движения применяется другой метод, основанный на применении конволюционной сети LSTM, который описывается ниже. После того как робот получает указание оператора, например, "принеси инструмент А, который лежит на лабораторном столе В", робот осуществляет поисковые движения RGB-D-камерой, используя известные алгоритмы обнаружения и распозна-

вания объектов [15]. В том случае, если объект не может быть обнаружен при текущем положении робота, он занимает новое положение с учетом того, что ему известно собственное положение на карте и положение лабораторного стола. Полученное изображение, содержащее искомый объект, передается в генератор образов. Туда же передается информация о текущем положении робота. Генератор образов формирует комплексный RGB-образ, включающий положение робота и цели, и передает его в качестве входного образа на конволюционную сеть LSTM, как показано на схеме системы управления КМР (рис. 4). Сеть непосредственно формирует управляющие сигналы линейной и угловой скоростями робота в соответствии с образом ситуации и поставленной задачей. На следующем такте работы системы робот определяет свое новое положение с помощью RGB-D-камеры и фотоэлектрического декодера и вновь передает информацию в генератор образов, который, в свою очередь, посылает образ новой ситуации в нейросеть. Сеть корректирует управляющие команды роботу и т. д.

После того как робот подъехал к столу, он должен найти и взять нужный инструмент. Задача обнаружения нужного инструмента на рабочем

столе и захвата его манипулятором с использованием конволюционной сети была решена нами ранее применительно к роботу-ассистенту хирурга, который искал требуемый хирургом инструмент [16]. Особенностью предложенного решения была параллельная конволюционная сеть, которая обнаруживала предмет по головному маркеру и одновременно определяла его положение на столе в целях формирования команд манипулятору для захвата объекта.

После того как робот захватил нужный объект, он должен передать его пользователю. С этой целью, используя внешнюю или собственную RGB-D-камеру, робот ищет оператора, определяет положение корпуса и рук оператора в пространстве помещения и перемещается к нему с помощью сети LSTM. При необходимости инструмент может быть безопасно передан непосредственно в руку оператора [17].

Конволюционная сеть LSTM

Структура сети. Остановимся более подробно на структуре конволюционной сети LSTM, которая и обеспечивает движение робота непосредственно по наблюдаемым обра-

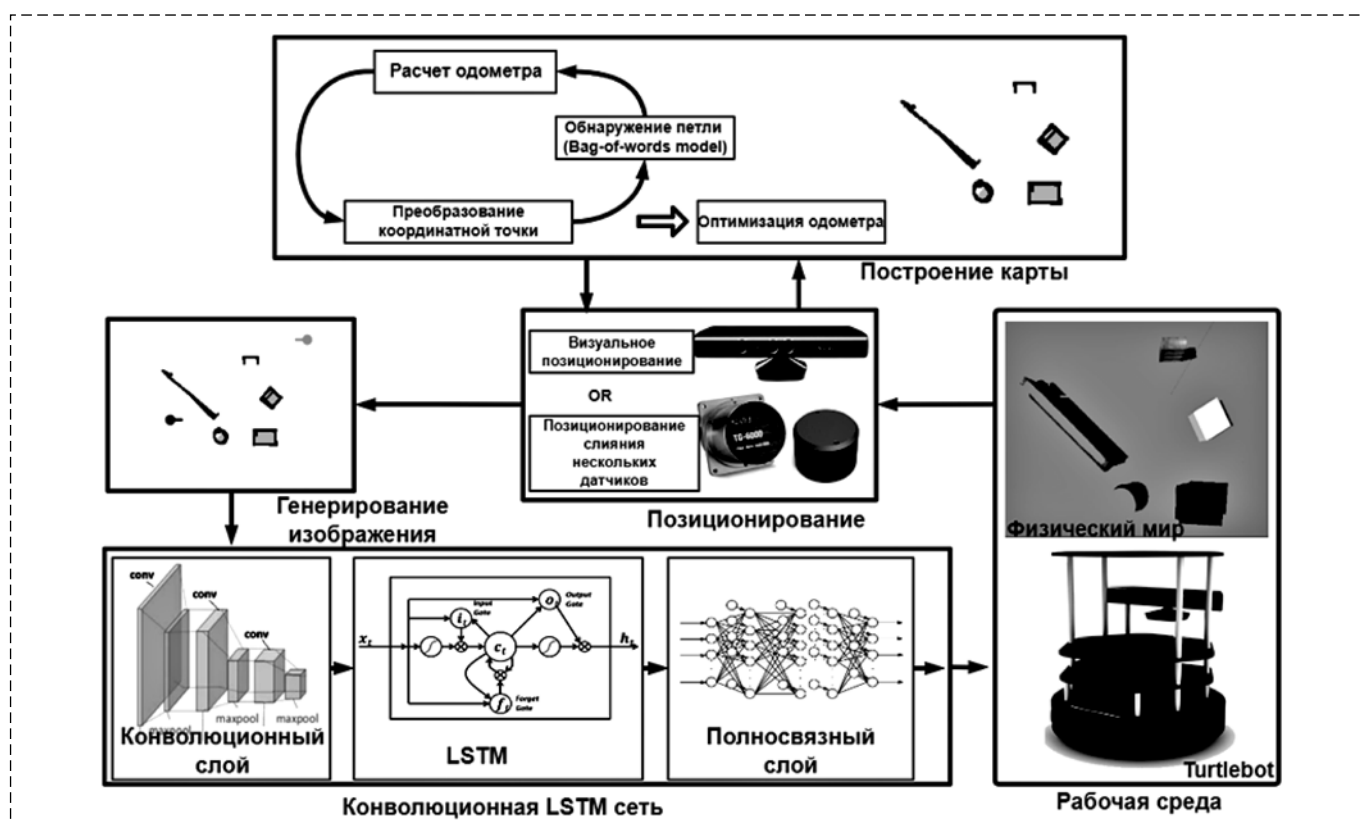


Рис. 4. Структура системы управления мобильным роботом Turtlebot, основанная на применении нейросети LSTM

Fig. 4. The structure of the control system of the mobile robot Turtlebot, based on the use of the LSTM neural network

зам. Сеть LSTM (Long short-term memory) была предложена в 1997 г. [18]. Ее название можно перевести как "нейронная сеть с долгой краткосрочной памятью". Описание работы сети можно найти, например, в работе [19]. Как и другие рекуррентные нейронные сети, сеть LSTM имеет форму цепочки повторяющихся модулей. Основное назначение сети — запоминание информации в течение длительных периодов времени. Благодаря своей структуре рекуррентная нейронная сеть анализирует полученную ранее информацию для решения последующих задач, например, используя уже полученные видеофрагменты для анализа последующих

Сеть состоит из трех основных частей, показанных на рис. 4: конволюционного слоя, блока LSTM и суммирующего полносвязного слоя.

Конволюционный слой используется для уменьшения нерелевантной информации в изображении и улучшения существенной информации. Он, в свою очередь, состоит из двух слоев: свертки и субдискретизации. Первый слой содержит ряд фильтров, предназначенных для удаления несущественной информации в изображении и выделения важной информации. Его можно описать формулой

$$x_{i,j} = f \left(\sum_{m=1}^F \sum_{n=1}^F w_{m,n} a_{i+m,j+n} + w_b \right), \quad (2)$$

где $x_{i,j}$ — характерный элемент изображения; $w_{m,n}$ — коэффициенты фильтра; $a_{i+m,j+n}$ — входные значения; w_b — порог срабатывания.

Слой субдискретизации обрезает изображение для того, чтобы уменьшить объем вычислений следующего за ним LSTM-слоя в соответствии с соотношением

$$\delta_{i,j}^l = \max(x_{i,j}), \quad (3)$$

где $\delta_{i,j}^l$ — параметр образа, полученного после субдискретизации.

Конволюционный слой содержит также слой, обеспечивающий преобразование трехмерной матрицы размерности $[W, H, N]$, где W — ширина полученного изображения, H — его длина, N — глубина, в вектор размерности $[1, WHN]$, который рассматривается в качестве входа для сети LSTM.

В качестве *функции ошибки* при настройке сети может использоваться среднее арифметическое значение или среднеквадратическая ошибка. В первом случае ошибка вычисляется как

$$E(y_{pre}, y_{true}) = \frac{k}{m} \sum_{i=1}^m |y_{pre}^i - y_{true}^i|, \quad (4)$$

где k — коэффициент усиления; m — число выборок; y_{pre}^i — желаемый вектор выхода; y_{true}^i — реальный вектор. Во втором случае

$$E(y_{pre}, y_{true}) = \frac{k}{m} \sqrt{\sum_{i=1}^m (y_{pre}^i - y_{true}^i)^2}. \quad (5)$$

В качестве вектора выхода сети, как уже говорилось, рассматриваются векторы линейной и угловой скоростей робота. Для ускорения сходимости процесса оптимизации рекомендуется в случае больших значений ошибки использовать вторую формулу, а при малых ошибках более эффективной оказывается первая.

Универсальным методом настройки (оптимизации) сети является метод градиентного спуска. В этом случае веса нейросети w находятся по формулам

$$w = w + \Delta w; \quad (6)$$

$$\Delta w = -\eta \nabla_w E(w), \quad (7)$$

где $\nabla_w E(w)$ — градиент ошибки, а параметр η определяет скорость сходимости. Однако его сложно применять к глубоким нейросетям, имеющим значительное число слоев. Для оптимизации таких нейросетей были предложены специальные методы, учитывающие предысторию изменения градиента в целях некоторой адаптации процесса сходимости. Одним из таких методов является метод RMSProp (Root Mean Square Propagation). В этом случае вычисляется усредненный по "истории" процесса сходимости квадрат модуля градиента:

$$r = \gamma r + (1 - \gamma) |g|^2, \quad (8)$$

в которой градиент $g = \nabla_w E(w)$, а параметр γ определяет "вес" накопленного среднего значения. Коррекция весов выполняется по формуле (6), в которой надо положить

$$\Delta w = -\frac{\eta}{\sqrt{\delta + r}} \nabla_w E(w).$$

Здесь параметр r определяется через усредненный квадрат модуля градиента по формуле (8), параметр η определяет скорость сходимости, а параметр δ назначается для того, чтобы в процессе оптимизации знаменатель не стал равен нулю. В нашей работе использовался именно этот метод.

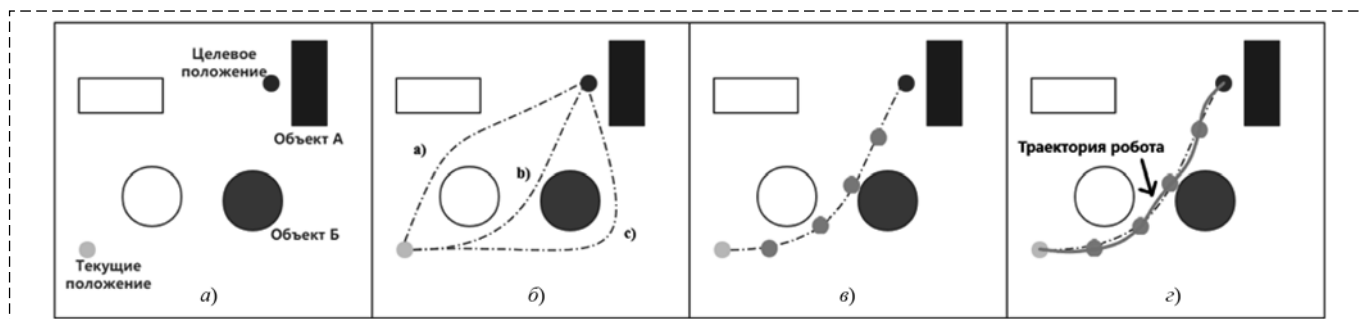


Рис. 5. Процесс диалогового управления мобильным коллаборативным роботом

Fig. 5. The process of interactive control of a mobile collaborative robot

Обучающая выборка. Для того чтобы обучить нейронную сеть, нужна обучающая выборка, которая представляет собой множество последовательных во времени изображений ситуации, включающей робота, заданную цель и внешнюю обстановку. Предполагается, что эти изображения могут быть получены с помощью RGB-D-камеры, установленной на роботе. Однако при работе в помещении задача может быть упрощена, поскольку для прокладки маршрута может быть достаточно двумерного плана помещения. В этом случае обучающая выборка представляет собой последовательность снимков ситуации в равноудаленные моменты времени. Пример показан на рис. 5. Оператор дал указание роботу: "Переместиться из текущего положения в целевое положение, обходя объект Б".

На рис. 5 показаны этапы процесса планирования движения: рис. 5, а — это карта, содержащая текущее положение робота, "объект А" и "объект Б"; рис. 5, б — на карте получены три маршрута, построенных традиционным методом (A^*). Здесь возможен диалог с оператором, который подтвердил маршрут *b*; рис. 5, в — на этом маршруте в опорных точках выбираются значения скорости и ускорения с помощью метода сплайнов третьего порядка с учетом заданных ограничений скорости МКР; рис. 5, г — на последнем этапе может проводиться моделирование движения робота с учетом его динамики с демонстрацией полученной траектории оператору.

Моделирование обученной системы управления МКР

В эксперименте в качестве входа конволюционной сети LSTM была выбрана матрица размерности $224 \times 224 \times 3$, а выходами сети при движении на плоскости являлись линейная и угловая скорости. Базовая обучающая выборка была получена традиционным методом путем

записи скорости мобильного робота и его соответствующих пространственных положений. Для обработки данных использовался метод скользящего окна. В каждой точке траектории t_i , $i = 1, 2, \dots, 20$, выбираются еще четыре точки в пределах длины скользящего окна, смещенные относительно текущего положения на $90k^\circ$, $k = 1, 2, 3, 4$. Таким образом, у нас получается 100 входных изображений, для можно получить 200 значений линейных и угловых скоростей.

Задачу формирования обучающей выборки можно обобщить. Пусть положение цели известно, а исходное положение робота может быть любым в данном помещении. Тогда на карту помещения можно наложить сетку, например, сетку 10×10 . Тогда полученное выше число выборок нужно увеличить в 100 раз. Это, конечно, увеличит время настройки сети, но зато придаст ей большую универсальность, поскольку задача, поставленная оператором, будет выполнена, каково бы ни было исходное положение робота. Процедура обучения занимает больше времени, но зато проводится только один раз.

Процедура моделирования работы обученного мобильного робота проводилась в программном обеспечении GAZEBO системы ROS. Были построены динамическая и кинематическая модель двухколесного робота Turtlebot. Предполагается, что для построения 2D-карты и определения положения робота в эксперименте используется RGB-D-камера, а положение и направление движения робота на 2D-карте передаются в генератор изображений. Программное обеспечение GAZEBO используется для моделирования движения МКР. Для управления движением МКР применяется предварительно обученная конволюционная сеть LSTM. Поскольку физическая среда компьютерного моделирования занимает часть центрального процессора, был создан новый узел ROS для вычисления конволюционной сети LSTM.

Результаты моделирования подтвердили возможность предлагаемого способа управления мобильным роботом. Подчеркнем, что конволюционная сеть, настроенная по сетке возможных положений, позволяет определять управление роботом при произвольном задании стартовой и целевой точек в рабочем пространстве. Ясно, что точность работы системы зависит от размерности сетки возможных положений. Однако, как уже отмечалось выше, задача существенно упрощается, если заранее задано исходное положение робота или настройка проводится на ряд заранее известных объектов в пределах одного и того же помещения.

Заключение

Проведенные исследования позволяют сделать вывод о том, что нейронные сети различного типа могут быть использованы на всех этапах создания КРТС, включая речевой диалог, распознавание объектов внешнего мира, автономное перемещение КМР в рабочем пространстве. В последнем случае предложенный способ управления КМР с использованием конволюционной сети LSTM позволяет вычислять команды управления роботом в реальном времени непосредственно по результатам наблюдения, что обеспечивает диалоговое управление сервисными КМР в реальном масштабе времени. Многочисленные эксперименты подтверждают, что этот метод обладает и хорошей отказоустойчивостью.

Основная проблема, связанная с управлением мобильным коллаборативным роботом с помощью нейронных сетей, заключается в сложности обучения, которое должно предусматривать, в общем случае, произвольное начальное положение робота и произвольное же задание цели в рабочем пространстве. Возможно, эту проблему в дальнейшем можно будет решить с использованием обучения с подкреплением. Однако в тех случаях, когда возможны ограничения как на исходное положение робота, так и на выбор маршрутов движения, предложенный метод управления может быть достаточно эффективен.

Нужно принять во внимание и эффект участия в управлении коллаборативным роботом человека-оператора, что предъявляет ряд требований к его психофизиологическим возможностям, включая пространственное мышление

и способность принимать решение в реальном масштабе времени [20].

Список литературы

1. Михайлов Б. Б., Назарова А. В., Ющенко А. С. Автономные мобильные роботы — навигация и управление // Известия ЮФУ. Технические науки, 2016. № 2(175). С. 48—67.
2. Ющенко А. С. Диалоговое управление роботами на основе нечеткой логики // Труды конференции "Экстремальная робототехника". С-Пб.: Политехника-сервис, 2015. С. 143—146.
3. Инь Шуай, Ющенко А. С. Коллаборативный робот — ассистент хирурга. С-Пб.: Политехника-сервис, 2019. С. 568—575.
4. Шуай Инь, Ющенко А. С. Диалоговая система управления роботом на базе теории конечных автоматов // Мехатроника, автоматизация и управление. 2019. Т. 20, № 11. С. 686—695.
5. Ющенко А. С. Интеллектуальное планирование в деятельности роботов // Мехатроника, автоматизация, управление. 2005. № 3. С.5—18.
6. Жонин А. А. Алгоритм обучения менеджера диалога речевой диалоговой системы управления роботом // Интегрированные модели и мягкие вычисления в искусственном интеллекте. Сб. научных трудов международной конференции. М.: ФИЗМАТЛИТ, 2011. С.395—406
7. Nist'ér D., Naroditsky O., Bergen J. Visual odometry for ground vehicle applications. // Journal of Field Robotics. Vol. 23(1). 2006. P. 3—20
8. Девятериков Е. А. Алгоритм описания траектории мобильного робота по данным визуального одометра для автоматического возвращения к оператору // Наука и Образование. 2014. № 12. С.705—715.
9. Wang X., Lian Y., Li L. Localization of Autonomous Cars Using Multi-Sensor Data // Fusion Chinese Automation Congress. 2018. P. 4152—4155
10. Девятериков Е. А., Михайлов Б. Б. Визуальный одометр // Вестник МГТУ им. Н. Э. Сер Приборостроение, Спецвыпуск "Робототехнические системы". 2012. № 6. С. 68—82.
11. Ющенко А. С., Зенкевич С. Л. Основы теории управления манипуляционными роботами. М.: Изд. МГТУ им. Н. Э. Баумана, 2004. 480 с.
12. Mellinger D., Kumar V. Minimum snap trajectory generation and control for quadrotors // IEEE International Conference on Robotics and Automation. 2011. P. 2520—2525.
13. Carius J., Ranftl R., Koltun V., Hutter M. Trajectory Optimization with Implicit Hard Contacts // IEEE Robotics and Automation Letters. 2018. Vol 3. P. 3326—3323.
14. Zenkevich S. L., Nazarova A. V., Zhu H. Logical Control a Group of Mobile Robots // Smart Electromechanical Systems Studies in Systems. 2019. Vol. 174. P. 32—43.
15. Zhao Z. Q., Zheng P., Xu S., Wu X. Object Detection with Deep Learning: A Review // IEEE Transactions on Neural Networks and Learning Systems 2019. Vol. 30. P. 3212—3232.
16. Yin S., Yuschenko A. Object Recognition of the Robotic System with Using a Parallel Convolutional Neural Network // Industry 40 Issues & New Intelligent Control Paradigms. 2019. Vol 272. P. 3—11.
17. Yin S., Yuschenko A. The application of the convolutional neural network to organize the work of a collaborative robot — surgeon assistant // Interactive Collaborative Robotics. 2019. P. 287—297.
18. Hochreiter S., Schmidhuber J. Long short-term memory // Neural Computation. 1997. N. 9 (8). P. 1735—1780.
19. Борисов Е. С. О рекуррентных нейронных сетях. URL: <http://mechanoid.kiev.ua/neural-net-rnn.html>
20. Ющенко А. С. Эргономические проблемы коллаборативной робототехники // Робототехника и техническая кибернетика. 2019. № 7 (2). С. 85—93.

Dialogue Control of Collaborative Robots Based on Artificial Neural Networks

A. S. Yuschenko, arkadyus@mail.ru, **Shuai Yin**, shuai.yin@yandex.ru,
Bauman Moscow State Technical University (BMSTU), Moscow, 105005, Russian Federation

Corresponding author: **Yuschenko Arkady S.** Dr. Science Tech., Professor,
Bauman Moscow State Technical University, Moscow, 105005, Russian Federation,
e-mail: arkadyus@mail.ru

Accepted on July 27, 2021

Abstract

Collaborative robotics progress is based on the possibility to apply robots to the wide range activity of peoples. Now the user can control the robot without any special knowledge in robotics and safe. The price of such possibilities is complication of control system of robot which now has to acquire an opportunity of autonomous behavior under human's control, using the necessary sensors and elements of artificial intelligence. In our research we suppose the collaborative robot as mobile robotic device possible to fulfil some work under the human's speech demands not only in the same space with the human. We also suppose the necessity of bilateral dialogue human-robot to make it clear the task, the current situation, the state as robot as human. The complex task of control, or may be the collaboration of human with his artificial partner need new means of control, situation recognition, speech dialogue management. As a mean to solve the whole complex of problems we propose the combination of different artificial neural networks. Such as convolution networks for image recognition, deep networks for speech recognition, LSTM networks for autonomous movement of robot control in current situation. Investigations in the field of mobile and manipulation robots including the human-robot control have been proceeded for some years in the department "Robotic systems and mechatronics" BMSTU celebrating now it 70th years Jubilee. The reader may find some of the works in the bibliography. In result of all these investigations we obtain the service robot model which may find a wide application

Keywords: mobile collaborative robot, neural network, convolutional LSTM network, dialogue control, trajectory planning, movement control, images recognition, SLAM

For citation:

Yuschenko A. S., Shuai Yin Dialogue control of collaborative robots based on artificial neural networks, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2021, vol. 22, no. 11, pp. 567–576.

DOI: 10.17587/mau.22.567-576

References

1. **Mikhailov B. B., Nazarova A. V., Yuschenko A. S.** Autonomous mobile robots — navigation and control, *Izvestiya SFedU, Engineering Sciences*, 2016, vol. 2(175), pp. 48–67.
2. **Yuschenko A. S.** Dialogue Mode of Robot Control on the Base of Fuzzy Logic, *Transaction of the Conference Extreme Robotics*, S-Pb., 2015, pp. 143–146.
3. **Shuai Yin, Yuschenko A. S.** Collaborative robot — surgeon assistant, *Transaction of the Conference Extreme Robotics*, S-Pb., 2019, pp. 568–575.
4. **Shuai Yin, Yuschenko A. S.** Dialogue system of controlling robot based on the theory of finite automata, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2019, vol. 20, no. 11, pp. 686–695.
5. **Yuschenko A. S.** Intelligent Planning of Work of Robots, *Mekhatronika, Avtomatizatsiya, Upravlenie*, 2005, no. 3, pp. 5–18.
6. **Zhonin A. A.** Training Algorithm for dialogue manager of a speech dialogue system, *Integrated Models and Soft Calculations in Artificial Intelligence. Proceedings of International Conference*, 2011, pp. 395–406.
7. **Nist'er D., Naroditsky O., Bergen J.** Visual odometry for ground vehicle applications, *Journal of Field Robotics*, 2006, vol. 23(1), pp. 3–20.
8. **Devtjarikov E. A.** Algorithm of the trajectory planning for a mobile robot return to operator using the visual odometer datum, *Nauka i Obrazovanie*, 2014, no. 12, pp. 705–715.
9. **Wang X., Lian Y., Li L.** Localization of Autonomous Cars Using Multi-Sensor Data, *Fusion Chinese Automation Congress*, 2018, pp. 4152–4155.
10. **Devtjarikov E. A., Mikhailov B. B.** Visual odometer, *Vestnik BMSTU Priboroostroenie*, 2012, Special issue no. 6 "Robotic Systems", pp. 68–82.
11. **Yuschenko A. S., Zenkevich S. L.** Basic of the manipulation robots control, Moscow, Publishing house of BMSTU, 2004, 480 p.
12. **Mellinger D., Kumar V.** Minimum snap trajectory generation and control for quadrotors, *IEEE International Conference on Robotics and Automation*, 2011, pp. 2520–2525.
13. **Carius J., Ranftl R., Koltun V., Hutter M.** Trajectory Optimization with Implicit Hard Contacts, *IEEE Robotics and Automation Letters*, 2018, vol. 3, pp. 3326–3323.
14. **Zenkevich S. L., Nazarova A. V., Zhu H.** Logical Control a Group of Mobile Robots, *Smart Electromechanical Systems Studies in Systems*, 2019, vol. 174, pp. 32–43.
15. **Zhao Z. Q., Zheng P., Xu S., Wu X.** Object Detection with Deep Learning: A Review, *IEEE Transactions on Neural Networks and Learning Systems*, 2019, vol. 30, pp. 3212–3232.
16. **Yin S., Yuschenko A.** Object Recognition of the Robotic System with Using a Parallel Convolutional Neural Network, *Industry 40 Issues & New Intelligent Control Paradigms*, 2019, vol. 272, pp. 3–11.
17. **Yin S., Yuschenko A.** The application of the convolutional neural network to organize the work of a collaborative robot — surgeon assistant, *Interactive Collaborative Robotics*, 2019, pp. 287–297.
18. **Hochreiter S., Schmidhuber J.** Long short-term memory, *Neural Computation*, 1997, vol. 9, no. 8, pp. 1735–1780.
19. **Borisov E. S.** About recurrent neural networks, available at: <http://mechanooid.kiev.ua/neural-net-rnn.html>.
20. **Yuschenko A. S.** Ergonomic problems of collaborative robotics, *Robototekhnika i Technicheskaya Kybernetika*, 2019, no. 7 (2), pp. 85–93.