

На правах рукописи

Инь Шуай

**ДИАЛОГОВАЯ СИСТЕМА УПРАВЛЕНИЯ КОЛЛАБОРАТИВНЫМ  
МЕДИЦИНСКИМ РОБОТОМ**

Специальность 05.02.05 –  
Роботы, мехатроника и робототехнические системы  
(в технических системах)

Автореферат

Диссертация на соискание ученой степени  
кандидата технических наук

*Инь Шуай*

Москва — 2021

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Московский государственный технический университет имени Н. Э. Баумана (национальный исследовательский университет)» (МГТУ им. Н.Э. Баумана).

Научный руководитель: **Ющенко Аркадий Семенович**  
доктор технических наук, профессор

Официальные оппоненты: **Жданов Александр Аркадьевич**  
Доктор физико-математических наук, профессор, главный научный сотрудник, «Институт точной механики и вычислительной техники имени С.А. Лебедева РАН» (ИТМиВТ)  
**Климов Даниил Дмитриевич**  
Кандидат технических наук, старший научный сотрудник, «Московский Государственный Технологический Университет СТАНКИН»

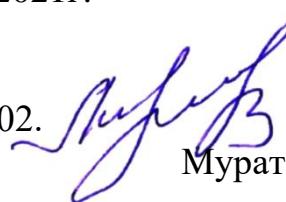
Ведущая организация: Федеральное государственное бюджетное учреждение науки «Институт проблем управления им. В. А. Трапезникова РАН»  
(ИПУ им. В. А. Трапезникова РАН)

Защита состоится «15» июня 2021 г. в 16 часов 30 мин. на заседании диссертационного совета Д212.141.02 при Московском государственном техническом университете им. Н.Э. Баумана по адресу: 105005, г. Москва, Госпитальный пер., д.10, ауд. 613м.

С диссертацией можно ознакомиться в библиотеке МГТУ им. Н. Э. Баумана и на сайте [www.bmstu.ru](http://www.bmstu.ru).

Отзыв на автореферат в двух экземплярах, заверенные печатью учреждения, просьба направлять по адресу: 105005, г. Москва, ул. 2-я Бауманская, д.5. стр. 1, МГТУ им. Н.Э. Баумана на имя ученого секретаря диссертационного совета Д212.141.02.

Автореферат разослан «\_\_\_» \_\_\_\_ 2021г.

Ученый секретарь  
Диссертационного совета Д212.141.02.  
Кандидат технических наук, доцент   
Муратов Игорь Валентинович

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность работы.** Диссертационная работа посвящена новому направлению в робототехнике, получившему название «коллаборативная робототехника». Несмотря на современную тенденцию к полной автоматизации производственной и иных сфер деятельности человека, существуют области, в которых человека невозможно заменить полностью. В силу сложности и не предсказуемости необходимых действий, а также ответственности, которую несет исполнитель. Одной из таких областей является медицина. Задача коллаборативной робототехнической системы (КоРТС) состоит в создании роботов, способных выполнять необходимые операции совместно с человеком, в том числе, работая в общей с ним рабочей зоне, и безопасно для него.

Актуальность роботизации вспомогательных действий во время проведения хирургических операций может быть обусловлена недостатком квалифицированного вспомогательного медицинского персонала. Особенно, во время пандемии, природных бедствий, во время боевых действий. Такая помощь хирургу может оказаться актуальной и при дистанционном проведении операций с помощью роботов-хирургов, примером которых может служить РТС «Да Винчи».

В настоящее время уже разработаны основные элементы этой структуры. В том числе, способы обработки распознавания речевой информации и способы формирования речевых сообщений. Речевые диалоговые системы широко используются для получения экспертных оценок, медицинской и технической диагностики, в различных формах информационного обслуживания населения. Принципиальная особенность диалоговых систем в робототехнике состоит в том, что они ориентированы не на информационную, а на физическую деятельность робота, контролируемую человеком, во внешнем мире. Отсюда следует, что робот, используя собственные информационно-сенсорные системы, в первую очередь, компьютерное зрение, должен создать в своем «сознании» модель мира и внешних объектов в нем, аналогичную модели, существующей в сознании человека, которую мы будем предполагать адекватной действительности. Второе требование состоит в том, что робот должен адекватно воспринимать действия, указанные оператором и быть способен их выполнить. Наконец, он должен быть в состоянии спланировать эти действия таким образом, чтобы выполнить поставленную перед ним задачу, не причиняя вред человеку, находящемуся в том же рабочем пространстве. Таким образом, речь идёт о достаточно сложной деятельности робота, управляемой с помощью речевого диалога. Системы, обладающие всеми перечисленными признаками, можно было бы назвать системами диалогового управления опосредованной (роботом) деятельности человека.

Обобщая полученные к настоящему времени результаты, можно сделать вывод, что с учетом сложности и неопределенности условий выполнения задач, возникающих перед КоРТС, традиционные методы теории автоматического управления в данном случае неприемлемы. Решение следует искать в области обучаемых и самообучаемых на базе нейронных сетей КоРТС, которые решали

бы следующие задачи: поддерживали бы речевой диалог на профессионально-ориентированном языке, обладали собственной СТЗ и были бы способны различать объекты внешнего мира в соответствии с их речевыми маркерами, а также выполнять манипуляционные операции в соответствии с их речевыми маркерами в условиях неполной определенности. При этом они должны обеспечить полную безопасность человека.

Разработки диалоговой системы управления колаборативным роботом при хирургической операции ведутся в различных организациях Китая, России, США, Европы и Японии. Ключевыми организациями в данном научном направлении являются Шанхайский университет Цзяо Тонг (Китай), Центр робототехники в институте RIKEN (Япония), Технологический институт Джорджии (Georgia Institute of Technology, США). В РФ такие работы ведутся в Институте проблем Механики РАН, в МГТУ им. Н.Э. Баумана, в Университете МосСтанкин и в ряде других организаций.

**Целью** диссертационной работы является разработка профессионально-ориентированной колаборативной системы управления вспомогательным роботом медицинского назначения – роботом-ассистентом хирурга, выполняющим функции хирургической (операционной) сестры.

Для достижения этой цели в диссертационной работе рассматриваются **следующие задачи:**

1. Определение задач КоРТС, её структуры и технических требований к её подсистемам и элементам на основе анализа структуры вспомогательных действий хирурга и хирургической сестры при проведении операций, включая речевой диалог

2. Выбор методов и разработка алгоритмов распознавания речи с учетом особенностей решаемых задач, а также организация диалога между медицинским персоналом и роботом.

3. Разработка метода и алгоритма поиска и распознавания хирургических инструментов роботом по его речевому маркеру с помощью системы компьютерного зрения и конволюционной нейронной сети.

4. Разработка методики создания КоРТС с речевым диалоговым управлением, включая организацию совместной работы человека и робота в КоРТС.

5. Организация совместной работы человека и робота при выполнении вспомогательных операций.

6. Проведение экспериментальных исследований отдельных подсистем разработанной КоРТС, а также системы в целом.

#### **Научная новизна:**

1. Предложена методика проектирования профессионально-ориентированной диалоговой колаборативной системы управления роботом с использованием нейросетевых алгоритмов.

2. Разработана система организации диалога «человек – робот» в процессе действий робота. Эта система позволяет человеку поддерживать двустороннюю голосовую связь с роботом при выполнении задач таким образом, чтобы своевременно корректировать его работу.

3. Предложена структура параллельной сверточной (конволюционной) сети, позволяющей решить, как задачу поиска и распознавания объектов, так и определить их положение, что необходимо для захвата объектов манипулятором.

4. Разработана методика определения пространственного положения хирурга, позволяющая работу безопасно передать инструмент непосредственно в его руку.

**Практическая ценность.** Предложенная методика проектирования профессионально-ориентированной диалоговой колаборативной системы управления роботом с использованием нейросетевых алгоритмов и разработанное программно-алгоритмическое обеспечение могут быть применены для создания колаборативных роботов с речевым диалоговым управлением не только при решении рассмотренных в диссертационной работе задач, но и для сервисной робототехники широкого назначения, в том числе, для реабилитации инвалидов, для помощи людям с дефектами зрения и органов движения.

**Положения, выносимые на защиту:**

1. Методика проектирования профессионально-ориентированной диалоговой колаборативной системы управления роботом с использованием нейросетевых алгоритмов.

2. Организация параллельной сверточной сети, позволяющей одновременно решить, как задачу поиска медицинских инструментов по голосовому маркеру, так и определить положение инструментов на столе хирурга, что необходимо для их захвата манипулятором.

3. Алгоритм организации диалога и действий робота при участии пользователя, обеспечивающий возможность своевременной коррекции работы КоРТС.

4. Результаты экспериментальных исследований, подтверждающие основные теоретические и методические положения диссертационной работы.

**Апробация работы.** Основные результаты работы докладывались и получили положительную оценку на следующих научных формах:

1. Международная конференция «International Symposium on Automation, Information and Computing» (Пекин, 02-04 декабря 2020 г.);

2. Международная конференция «экстремальная робототехника» (Санкт-Петербург, 13-15 июня 2019 г.);

3. Международная конференция «International Conference on Interactive Collaborative Robotics» (Стамбул, 20-25 августа 2019 г.);

4. Научная конференция для студентов, молодых ученых и аспирантов МГТУ им. Н. Э. Баумана, посвященная 105-летию академика РАН Е. П. Попова (Москва, 27-28 марта 2019 г.);

5. Всероссийская конференция «Мультиконференция по проблемам управления» (23-28 сентября 2019 г.);

6. Всероссийская студенческая конференция «Студенческая научная весна» (Москва, 2018 г.).

**Публикации.** Основные результаты по теме диссертации изложены 6 печатных изданиях, в том числе, одна опубликована в журнале,

рекомендованного ВАК РФ, две в журналах с индексом цитирования Scopus и две публикации в тезисах докладов конференций.

**Структура и объем работы.** Диссертационная работа состоит из введения, пяти глав, заключения и приложения. Полный объем диссертации составляет 161 страницу, включая 62 рисунка и 17 таблиц. Список литературы содержит 89 наименований.

## СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность темы исследования, сформулированы цель и задачи, раскрыта научная новизна и практическая значимость работы, представлена структура диссертации, приведены защищаемые положения.

В **первой главе** представлены основные разработки в области КоРТС медицинского назначения. Описаны основные подходы к построению медицинских роботов и к применяемым в медицине методам управления КоРТС. Приведен анализ существующих методов диалогового управления КоРТС.

Предметом диссертации является диалоговый способ управления колаборативными роботами. Для реализации диалогового управления КоРТС необходимо решить задачи обработки речевого сообщения, организации диалога и действий робота, включая поиск и распознавание объектов, а также их передачу хирургу. Решение этих задач включает организацию диалога и действий робота, разработку метода распознавания речи, разработку методов распознавания и поиска объектов и методов определения пространственного положения объекта. С помощью этих методов, используемых совместно, организуется КоРТС, схема которой показана на Рис. 1.

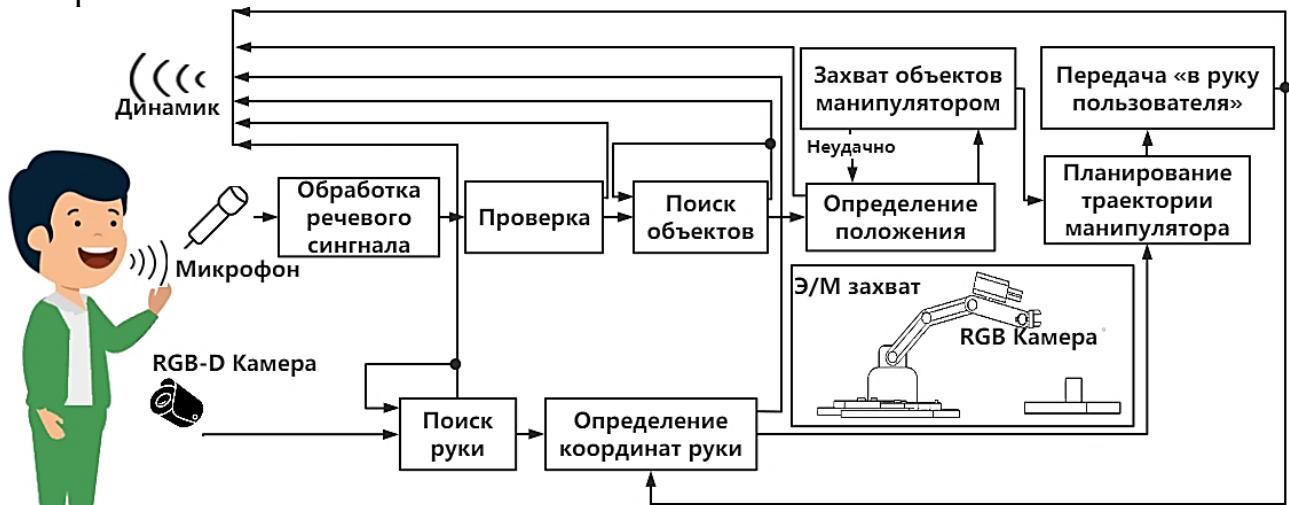


Рис. 1. Функциональная схема КоРТС

Сначала с использованием метода распознавания речи обрабатывается речевой запрос пользователя. Затем проверяется корректность и полнота полученной информации в виде текста. При необходимости пользователю будет предложено дополнить информацию о задаче. После формулировки задачи (например, «передать определенный объект пользователю») в блоке поиска производится поиск объектов с помощью манипулятора и системы

компьютерного зрения. Далее по положению объекта в полученном изображении вычисляется положение объекта в системе координат робота. Выполняется задача захвата объекта манипулятором. В завершение этого цикла выполняется задачи поиска руки хирурга и затем безопасной передачи объекта в его руку.

Во второй главе приводится сравнительный анализ современных методов распознавания речи и разрабатывается система распознавания профессионально-ориентированной речи с использованием глубокой нейронной сети, а также языковой модели n-gram. Алгоритм распознавания речи с использованием разработанной глубокой нейронной сети способен повысить точность распознавания при различных акцентах и темпах произношения. К тому же этот метод может быть реализован на недорогих устройствах и не предполагает подключения к сети Интернет.

Система распознавания непрерывной речи состоит из четырех частей (Рис. 2). Модуль обработки сигнала извлечения признаков выполняет обработку речевого сигнала для устранения шума и улучшения полезной информации. Модуль акустической модели решет проблемы, связанные с переменной длиной вектора признаков и с вариативностью голосового сигнала, который зависит от пола и состояния говорящего. Модуль языковой модели определяет связи последовательности слов. Модуль оптимального решения – определение текстовых последовательностей  $W^*$  по формуле:  $W^* = \arg \max P(X|W)P(W)/P(X)$ , где  $P(X)$  – априорная вероятность входного голосового сигнала, которая не зависит от акустической и языковой модели и может быть принята равной 1;  $P(W)$  – вероятность текстовой последовательности, которая вычисляется языковой моделью;  $P(X|W)$  – вероятность входного голосового сигнала, при условии текстовой последовательности  $W$ , вычисляемой акустической моделью.

В диссертации для обработки звукового сигнала использовался метод MFCC (Mel-Frequency Cepstral Coefficient). Для реализации акустической модели использовалась разработанная глубокая нейронная сеть на основе скрытой Марковской модели. Для реализации языковой модели использовалась модель n-gram.



Рис. 2. Схема системы распознавания речи

Схема скрытой Марковской модели, использованной в работе, приведена на Рис. 3. Она может быть описана вектором состояний  $S = \{S_1, S_2, \dots, S_N\}$ ,

матрицей вероятностей переходов между состояниями  $A = [a_{ij}], i, j = 1, 2, 3, \dots, N$ , вектором наблюдаемых вероятностей  $B = b_j(o_t) = P(o_t | S_j), j = 1, 2, \dots, N$ . На схеме обозначено  $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$  – начальные вероятности состояний и  $V = o_t = \{v_1, v_2, \dots, v_n\}$  – наблюдаемые векторы.

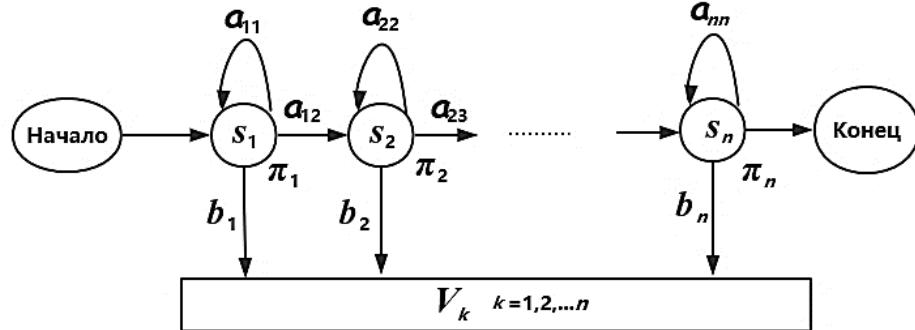


Рис. 3. Схема скрытой Марковской модели

Существует метод распознавания непрерывной речи с использованием глубокой нейросети и скрытой Марковской модели, т.е. метода DNN-HMM (Deep Neural Network - Hidden Markov Model) – поиск сравнения между звуковыми сигналами и словами. Для распознавания непрерывной речи в работе предлагается применить скрытую Марковскую модель совместно с языковой моделью n-gram. Этот алгоритм реализован с помощью глубокой нейронной сети. Входом глубокой нейронной сети являются голосовые признаки, преобразуемые методом MFCC. Схема этой модели, разработанная в диссертации, показана на Рис. 4.

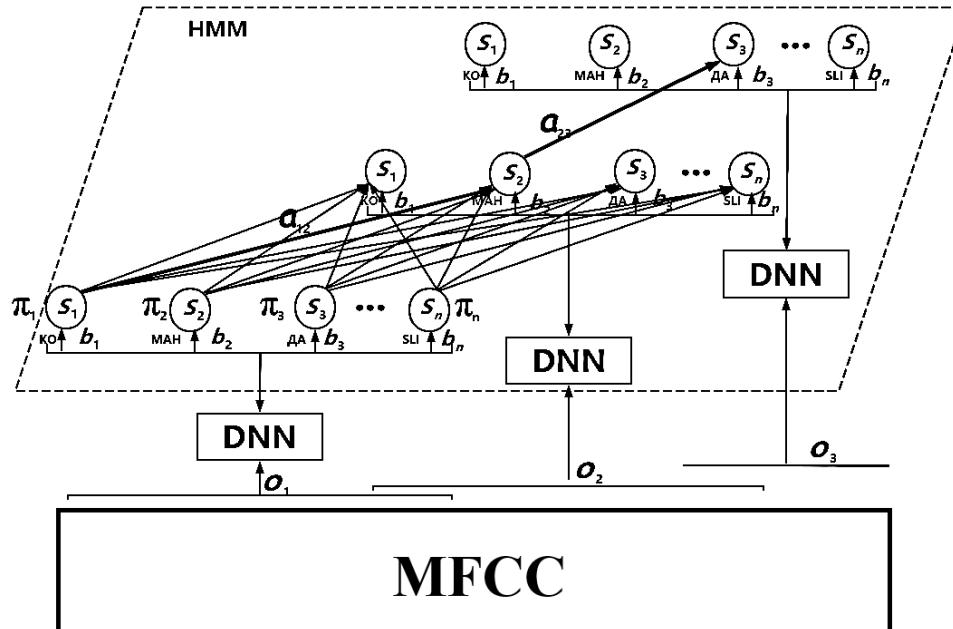


Рис. 4. Схема распознавания речи с использованием глубокой нейронной сети

В модели DNN-HMM глубокая нейронная сеть используется для вычисления вектора наблюдаемых вероятностей  $B$  при известных параметрах  $A$  – вероятность между состояниями,  $\pi$  – начальные вероятности состояний и  $V$

– наблюдаемые векторы. Глубокая нейронная сеть обучается с помощью обучающей выборки, в которой входами являются наблюдаемые векторы  $\mathbf{o}^m$ , а выходами – наблюдаемые вероятности  $\mathbf{B}$ . Чтобы определить наблюдаемые вероятности  $\mathbf{B}$  при известных наблюдаемых векторах  $\mathbf{o}^m$ , необходимо обучить модели GMM-HMM (Gaussian Mixture Model-Hidden Markov Model). Задача обучения модели Гауссовых смесей состоит в том, чтобы определить эти параметры модели Гауссовых смесей  $\Theta = \{c_m, \mu_m, \Sigma_m\}$  с использованием алгоритма максимального правдоподобия. После обучения модели GMM-HMM обученная модель GMM-HMM используется для вычисления наблюдаемых вероятностей  $\mathbf{B}$  при известных наблюдаемых векторах  $\mathbf{o}^m$ . Глубокая нейронная сеть обучается с использованием полученных обучающих выборок ( $\mathbf{o}^m, \mathbf{B}$ ).

Алгоритм обучения модели DNN-HMM включает следующие основные этапы:

0. Предобработку голосовых сигналов с помощью метода MFCC.

1. Обучение модели GMM-HMM (Gaussian Mixture Model-Hidden Markov Model) с использованием алгоритма максимального правдоподобия.

2. Инициализация скрытой Марковской модели с использованием модели GMM-HMM, полученной на первом шаге.

3. Предварительное обучение глубокой нейронной сети с помощью глубокой сети доверия (Deep Belief Network), которая представляет собой вероятностную генеративную модель, состоящую из нескольких ограниченных машин Больцмана (RBM) для определения начальных параметров глубокой нейронной сети.

4. Попарное совпадение наблюдаемых векторах  $\mathbf{o}^m$  с наблюдаемыми вероятностями  $\mathbf{B}$ , полученными на шаге 2, обеспечивает получение обучающих выборок для обучения DNN.

5. Обучение DNN с использованием алгоритма обратного распространения ошибки (Backward Error Propagation, BP) при известных обучающей выборки ( $\mathbf{o}^m, \mathbf{B}$ ), полученных на шаге 4.

6. Переоценка точности (процент для тестовых данных) распознавания речи с совместным применением глубокой нейронной сети и Скрытой Марковской модели.

7. Если точность на шаге 6 больше не повышается, следует выходить из алгоритма, в противном случае процесс возвращается к шагу 5.

8. После обучения модели DNN-HMM глубокая нейронная сеть вычисляет наблюдаемые вероятности  $\mathbf{B}$  при известных наблюдаемых векторах  $\mathbf{o}^m$ .

В диссертации приведено экспериментальное исследование распознавания речи с использованием разработанной глубокой нейронной сети. Из полученных результатов следует, что метод распознавания речи с использованием глубокой нейронной сети способен преобразовать голосовые сигналы в текстовую информацию в реальном времени близкому к темпу естественной речи – задержка распознавания речи составляет примерно 0.12 с. Точность распознавания является достаточно высокой и составляет 97%. Это соответствует требованиям к проектированию диалоговой системы.

В третьей главе рассматривается организация системы компьютерного

зрения КоРТС. Она включает две подсистемы. RGB- камера устанавливается на последнем звене манипулятора и обеспечивает поиск и распознавание объектов на рабочем столе по голосовому запросу хирурга. Вторая RGB-D камера предназначена для внешнего наблюдения, и для передачи инструмента хирургу с помощью манипулятора.

Для решения первой задачи разработана параллельная сверточная (конволюционная) сеть, позволяющая одновременно найти сам объект и определить его положение на рабочем столе, что необходимо для захвата объекта манипулятором. С учетом требования минимального времени захвата объекта для решения этой задачи предлагается разработанная автором параллельная сверточная сеть (P-CNN), структура которой приведена на Рис. 5 Сеть состоит из двух параллельных каналов. Верхняя сверточная сеть разделяет области распознавания на три группы: головная часть объекта, хвостовая часть объекта и оставшаяся часть объекта. Нижняя сверточная сеть анализирует полное изображение для определения названия объекта. По дополнительной формуле №1 вычисляются координаты центра тяжести и углы поворота объекта. Дополнительная формула №2 определяет серийный номер объекта.

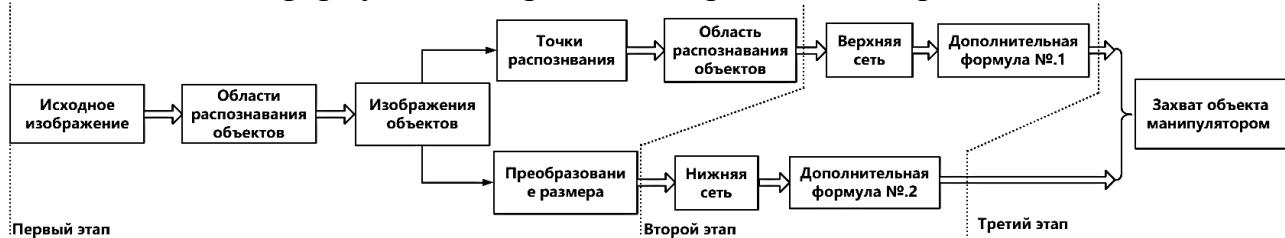


Рис. 5. Универсальная схема захвата объектов с использованием разработанной параллельной сверточной сети

Угол поворота объекта вычисляется по формуле:

$$\delta_{\text{угол поворота}} = \tan^{-1}\left(\frac{v_g - v_x}{u_g - u_x}\right) \quad (1)$$

где  $(u_g, v_g)$  – координаты центра головной части;  $(u_x, v_x)$  – координаты центра хвостовой части (Рис 6).

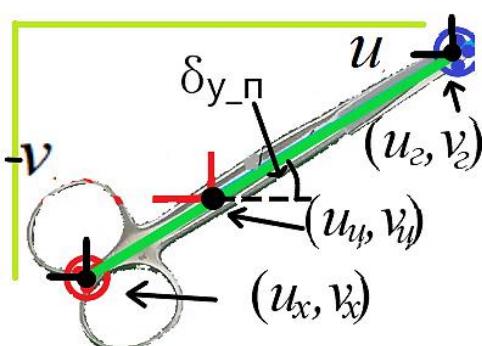


Рис. 6. К вычислению центра тяжести хирургических ножниц

В данной работе положение центра тяжести объекта необходимо для определения положения электромагнитного захвата манипулятора. Центр тяжести объекта вычисляется по формуле, справедливой для плоских симметричных объектов:

$$\begin{cases} u_{\text{ц}} = k \cdot u_g + (1 - k) \cdot u_x \\ v_{\text{ц}} = k \cdot v_g + (1 - k) \cdot v_x \end{cases} \quad (2)$$

где  $(u_{\text{ц}}, v_{\text{ц}})$  – координаты центра тяжести;  $(u_g, v_g)$  – координаты центра головной части;  $(u_x, v_x)$  – координаты центра хвостовой части;  $k$  – коэффициент центра тяжести, определяющий его реальное положение.

Координаты центра тяжести объекта относительно робота вычисляются с

помощью однородного преобразования по следующей формуле.

$$s \begin{bmatrix} u_u \\ v_u \\ 1 \end{bmatrix} = A_{in} \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3)$$

где  $A_{in} = \begin{bmatrix} s_1 & 0 & u_0 \\ 0 & s_2 & v_0 \\ 0 & 0 & 1 \end{bmatrix}$  – матрица внутренних параметров камеры;  $s$  –

произвольный масштабный коэффициент; параметры  $r_{11}, r_{12}, \dots, r_{33}$  – определяют 3x3 матрицу вращения между роботом и камерой;  $[T_x \ T_y \ T_z]^T$  – 3x1 вектор перемещения;  $s_1$  и  $s_2$  – масштабные коэффициенты по осям  $u$  и  $v$  изображения.

Для проверки работоспособности разработанного метода были проведены эксперименты, в ходе которых верхняя сеть и нижняя сеть на Рис.5 были обучены по отдельности методом обратного распространения ошибки. Результаты поиска объектов с использованием разработанной параллельной сверточной сети показаны на Рис. 7.

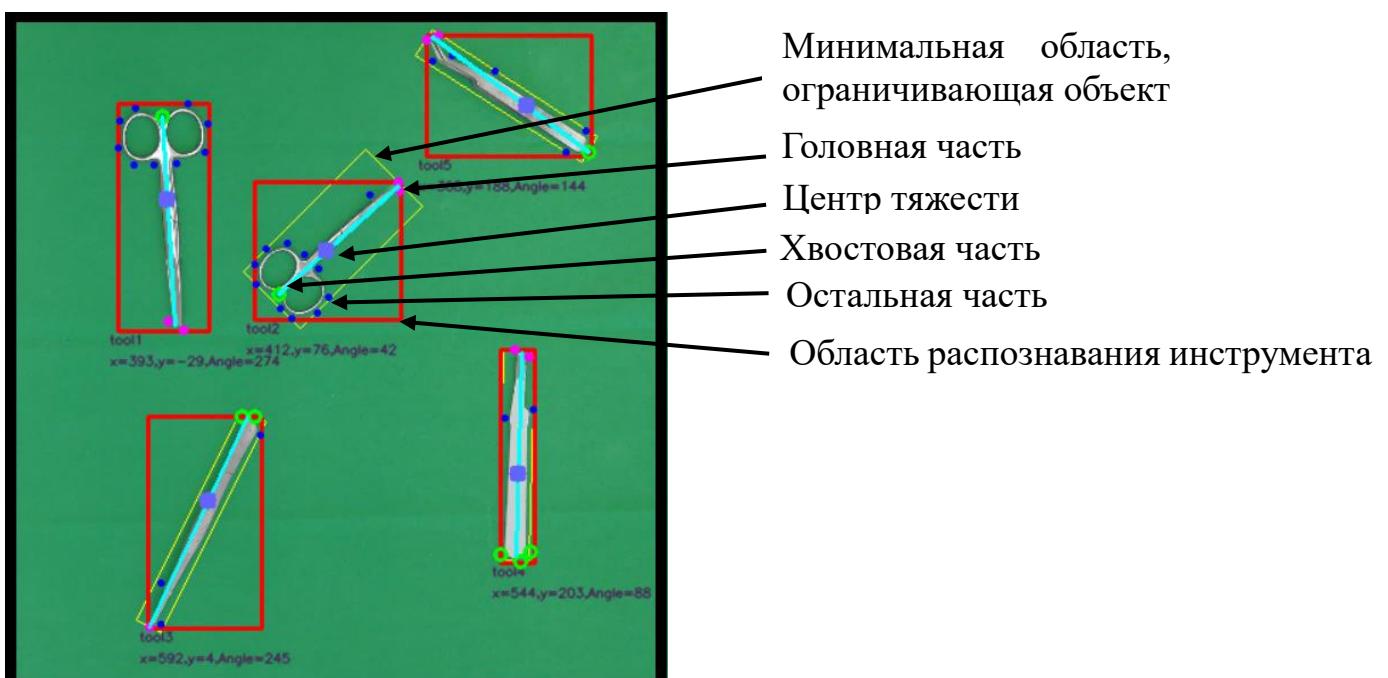


Рис. 7. Результаты работы параллельной сверточной сети

Из полученных результатов экспериментального исследования следует, что точность распознавания обучающей выборки достигает достаточно высокого уровня (96%). Время поиска объекта при этом составляет примерно 0.5 с., что соответствует требованиям к выполнению операции. Это значение обеспечивает выполнение операции поиска и захвата инструмента за время, близкое к времени при обычной организации операции.

Задача определения пространственного положения руки (ладони) с помощью RGB-D камеры состоит из поиска руки пользователя и определения пространственной координаты руки. Поиск руки пользователя, в свою очередь,

разделяется на три этапа: а) поиск корпуса человека в рабочем пространстве с помощью RGB-D камеры; б) поиск фрагментов корпуса, включая руки, с помощью алгоритма построения рандомизированных деревьев (randomized trees); в) распознавание отдельных суставов, в том числе, суставов руки, с использованием нейросетевых алгоритмов. В конечном счете определяется изображение кисти (ладони) руки хирурга (Рис. 8).

Для определения пространственных координат руки человека необходимо обеспечить совпадение цветного изображения с изображением глубины. Эта процедура имеет два этапа:

1. Вычисление положения RGB-камеры относительно D-камеры по формуле:

$$[\mathbf{R}_{ir2c}, \mathbf{T}_{ir2c}] = \frac{\mathbf{A}_c[\mathbf{R}_c^i, \mathbf{T}_c^i]}{\mathbf{A}_{ir}[\mathbf{R}_{ir}^i, \mathbf{T}_{ir}^i]}, \quad (4)$$

где обозначено:  $[\mathbf{R}_c^i, \mathbf{T}_c^i]$  – матрица поворота и вектор перемещения между мировой системой координат и RGB-камерой соответственно;  $[\mathbf{R}_{ir}^i, \mathbf{T}_{ir}^i]$  – матрица поворота и вектор перемещения между мировой системой координат и D-камерой.

2. Вычисление коэффициента масштаба  $Res_{ir2c}$  между пиксели D-изображения и пиксели RGB-изображения

$$\frac{Res_{ir2c}}{Res_{ir}} = \frac{f_c}{f_{ir}} \quad (5)$$

где  $Res_{ir}$  – базовое значение (равно 1);  $f_{ir}$  – фокусное расстояние D-камеры.  $f_c$  – фокусное расстояние RGB камеры.

Восстановление трехмерной точки, соответствующей пикселию  $(u, v)$  с глубиной  $z$ , выполняется с помощью формулы:

$$[x, y, z]^T = \left[ \frac{(u - c_x)z}{f_x}, \frac{(v - c_y)z}{f_y}, z \right]^T \quad (6)$$

где  $c_x, c_y, f_x, f_y$  – внутренние параметры камеры;  $(u, v)$  – пиксель в изображении;  $z$  – глубина объекта до камеры.

В эксперименте по определению пространственных координат руки с помощью RGB-D камеры Kinect 360 была проведены отдельная калибровка геометрии D и RGB камеры с помощью пакета OpenCV и было обеспечено совпадение цветного изображения с изображением глубины из RGB-D камеры с использованием методики, рассмотренной выше. При проведении экспериментов требовалось чтобы пользователь был расположен лицом к RGB-D камере на расстоянии 1-3 метра напротив камеры. На Рис. 8 показан результат поиска руки и определения координат центра в RGB изображении.

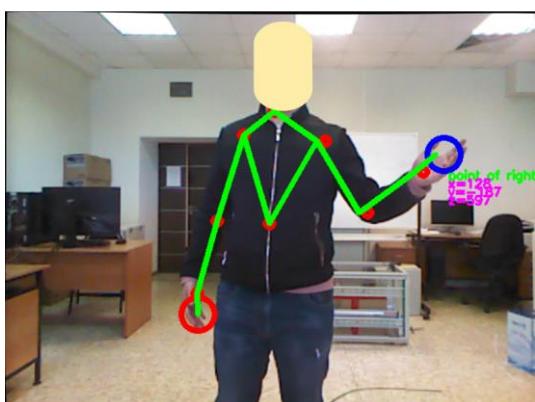


Рис. 8. Результаты распознавания руки и определения координат кисти

Эксперимент показал, что средняя погрешность измерения пространственных координат руки (ладони) составляет не более 10 мм, что достаточно для размещения инструмента в руку пользователя. Время распознавания и определения координат руки пренебрежимо мало ( $\approx 0.03$  секунд на одно изображение).

В четвертой главе предложен процесс организации диалога «человек-робот» и действий робота при участии пользователя, позволяющий роботу поддерживать двустороннюю связь с человеком при выполнении задач. Процесс организации диалога и действий робота включает пять этапов: 1. Распознавание голосовых сигналов с использованием алгоритма DNN-HMM с их предварительной обработкой с помощью метода MFCC. 2. Проверка результатов распознавания голосовых запросов пользователя с помощью базы данных. 3. Определение задачи для подсистемы компьютерного зрения и подсистемы управления роботом по голосовым запросам пользователя. 4. Этап захвата медицинского инструмента манипулятором. 5. Этап проверки результатов решения поставленной пользователем задачи с помощью СТЗ.

В диссертации для организации действий робота на основе диалога «человек-робот» была разработана архитектура системы диалогового управления, представленная на Рис. 9.

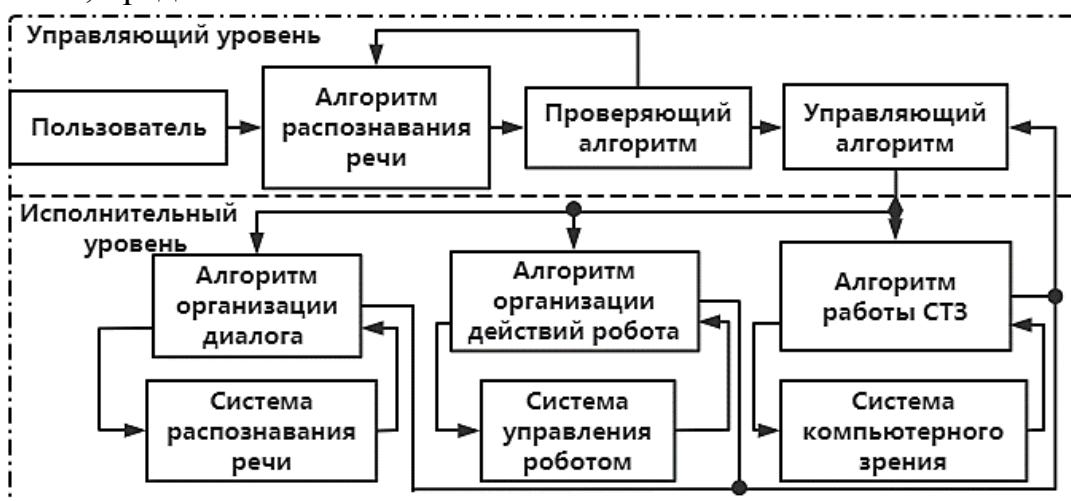


Рис. 9. Архитектура системы организации диалога и действий робота при участии пользователя

Эта система включает управляющий уровень всей системы и исполнительный уровень. К последнему можно отнести алгоритм организации диалога, алгоритм организации действий робота и алгоритм работы СТЗ. Управляющий уровень назначает задачи исполнительному уровню в момент времени  $t$  на основе информации обратной связи от исполнительного уровня в предыдущий момент времени  $t-1$ .

Для профессионально-ориентированной диалоговой системы управления роботом в диссертации разработаны четыре типа сценария диалога: (1) сценарий диалога для постановки задачи, (2) для уточнения требований к задаче, (3) для получения помощи пользователя и (4) для пояснения запроса. Эти четыре типа сценария диалога разработаны в соответствии с ситуациями, которые могут

возникнуть во время профессионально-ориентированного диалога.

В пятой главе описан разработанный автором экспериментальный комплекс, на базе которого проведено экспериментальное исследование разработанной диалоговой системы управления. Комплекс включает колаборативный манипулятор KUKA LBR 14, а также всю необходимую измерительную и вычислительную аппаратуру (Рис. 10). При выполнении экспериментов исследовалось время выполнения одного запроса, процент успешного выполнения запросов, степень влияния различных факторов на успешность выполнения диалогового управления роботом. В том числе, влияние освещенности с учетом настольных материалов (светоотражающие материалы, не отражающие материалы), характер источников света (солнечный свет или искусственное освещение). Исследовалась безопасность работы при помещении манипулятором инструмента в руку хирурга и при опасности столкновения манипулятора и руки человека.

Экспериментально установлено, что допустимый уровень шума в операционной лежит в пределах от 30 дБ до 40 дБ. Оптимальная интенсивность освещения составляет 200-350 Lux. На деревянной столешнице при равномерном освещении система распознавания работает лучше. Возможная сила столкновения не превышает 50 Н, что безопасно для человека.

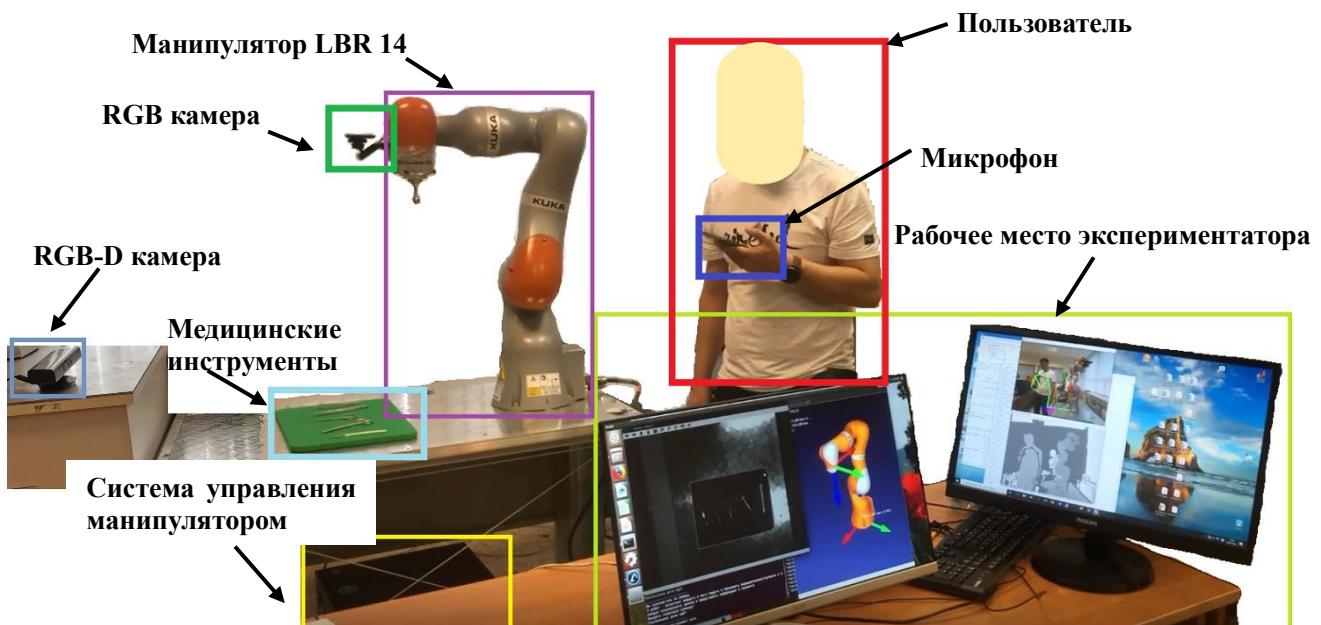


Рис. 10. Общий вид экспериментального комплекса

Схема экспериментального комплекса КоРТС показана на Рис. 11. Управляющий алгоритм, алгоритм распознавания речи, алгоритм поиска объекта реализованы на одном компьютере (XiaoMi Pro), а алгоритм поиска руки - на другом компьютере, к которому подключится RGB-D камера Kinect 360. Эти компьютеры обмениваются данными по протоколу Socket. Компьютер, на котором реализован алгоритм поиска руки, непрерывно отправляет информацию от пользователя на компьютер, на котором реализован управляющий алгоритм.

Компьютер, на котором реализован управляющий алгоритм, сохраняет полученную информацию в текстовом формате. Управляющий алгоритм постоянно обращается к этому файлу в рабочем процессе для извлечения необходимой информации. Для реализации коммуникации между управляющим алгоритмом и контроллером управления манипулятором Sunrise Cabinet использовано программное обеспечение IIWA\_Ros.java и Ros\_IIWA.py. Контроллер Sunrise Cabinet непрерывно возвращает информацию от робота в управляющий алгоритм для отображения состояния робота оператору. Алгоритмы поиска объекта и распознавания речи начинают работать только после получения команды управляющего алгоритма, и затем передают результаты решения задачи в управляющий алгоритм. Через WiFi RGB камера и микрофон подключаются к компьютеру, на котором реализован управляющий алгоритм.

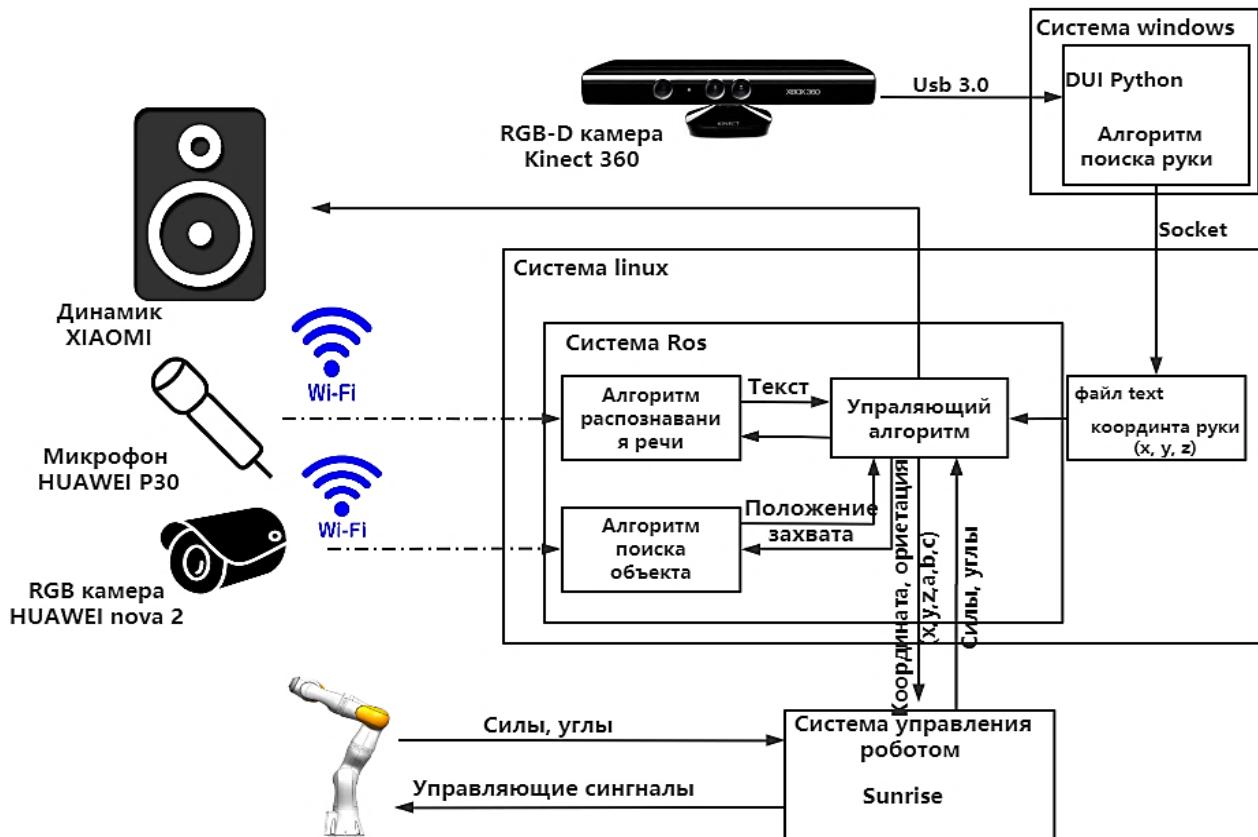


Рис. 11. Функциональная схема экспериментального комплекса КоРТС

Вначале работы диалоговая система загружает нужные подсистемы и дополнительные модули. Затем она проверяет подключение робота, положение медицинских инструментов и наличие руки пользователя в рабочей зоне манипулятора. Если рука не находится в рабочей зоне или имеет неправильную ориентацию, то диалоговая система просит пользователя переместить руку в рабочую зону и придать возможное для захвата положение кисти руки. После проверки диалоговая система уточняет задачу с помощью речевого диалога и затем выполняет задачу. В процессе выполнения задачи диалоговая система сообщает пользователю рабочий статус робота и при необходимости запрашивает пользователя для необходимой помощи.

Далее приводятся основные результаты экспериментального исследования КоРТС в целом. Экспериментатор произносил речевой запрос «вложить ножницы в руку», производились обработка речевого высказывания, поиск нужного инструмента и обработка его изображения, захват его манипулятором и передача инструмента в руку экспериментатора. На Рис. 12 показаны после получения голосового запроса пользователя отдельные этапы операции «захват инструмента и передача инструмента пользователю манипулятором по его голосовому запросу», включая начальное (Рис. 12, а) и конечное положение манипулятора (Рис. 12, в), а также промежуточное значение в момент захвата (Рис. 12, б).

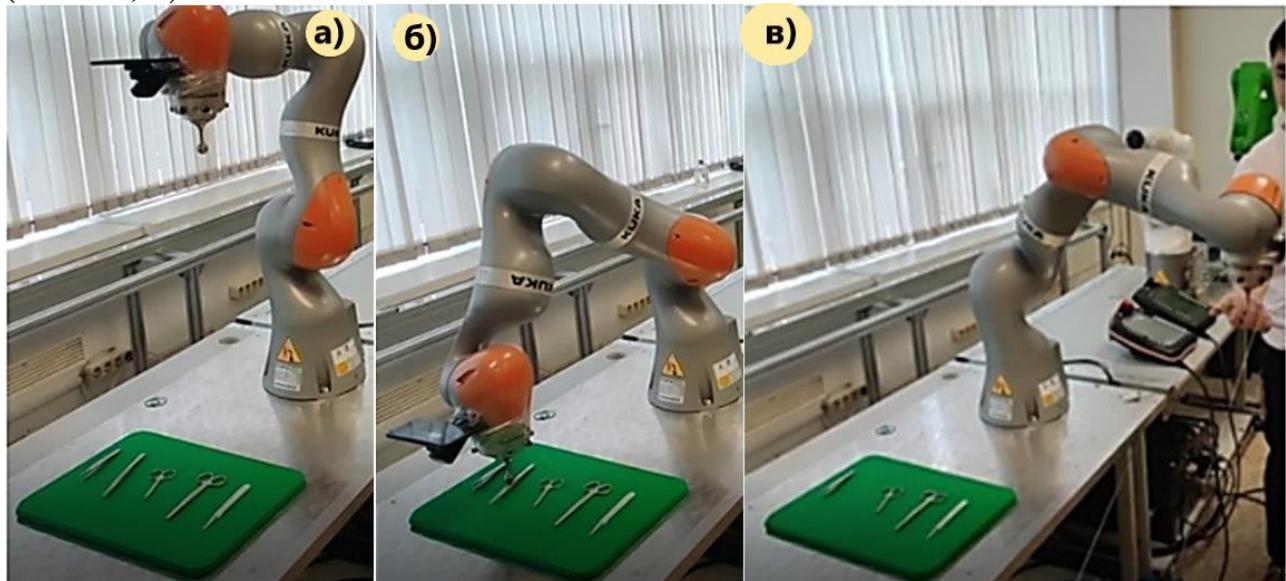


Рис. 12. Захват инструмента и передача инструмента пользователю манипулятором по его голосовому запросу

Обобщенные результаты эксперимента для выполнения одного голосового запроса пользователя (при отсутствии диалога) показаны в Таблице 2.

Таблица 2

Операция	Распознавание речи	Обработка изображения	Действие манипулятора	Суммарное время
Время	0.12с	0.5с	6с	6.62с
Точность	97%	96%	± 0.1мм	-

Результаты экспериментального исследования комплексной диалоговой системы управления коллаборативным манипулятором, включая отдельные этапы её работы, показаны в Таблице 3, где обозначено: Команда №1 – вложить большие ножницы в руку; Команда №2 – вложить маленькие ножницы в руку; Команда №3 – вложить пинцет в руку; Команда №4 – вложить скальпель в руку; Команда №5 – вложить пинцет с правильной ориентацией в руку. В процессе выполнения экспериментов человек отправлял каждую команду 20 раз в диалоговую систему. Таким образом, всего было проведено 100 реализаций голосового запроса.

Таблица 3

Голосовой запрос	Количество экспериментов	Число успешных распознаваний команд	Число успешных захватов инструмента	Число успешных передач инструмента
Команда №1	20	20	19	19
Команда №2	20	20	20	20
Команда №3	20	19	17	20
Команда №4	20	20	19	20
Команда №5	20	18	17	19
Процент успеха	—	97%	92%	98%

Из представленных результатов можно сделать вывод о том, что после получения голосового запроса диалоговая система выполняет задачу примерно за 6.62 с. В том числе 0.12 с. составляет задержка распознавания речи, менее 0.5 с. процедура обработки изображения и выполнение движения манипулятора. Таким образом, полное время выполнения операции передачи инструмента хирургу по голосовому запросу не превышает времени выполнения этой операции хирургической сестрой. Точность работы манипулятора достаточна для захвата плоского инструмента с помощью электромагнитного захвата и безопасного вложения инструмента в руку хирурга. Разработанная система диалогового управления работает достаточно надежно – в проведенных экспериментах процесс успешно выполненных операций составил не менее 92 %.

Таким образом, экспериментальные исследования подтвердили работоспособность и эффективность разработанной коллaborативной системы «Робот-ассистент хирурга»

В Заключении приведены основные результаты диссертационной работы.

В Приложении к диссертации приведены листинги программ диалоговой системы управления манипулятора, а также программы обработки изображений на языке Python и обработки голосовых сигналов.

## ОСНОВНЫЕ РЕЗУЛЬТАТЫ ДИССЕРТАЦИИ

1. Разработана методика проектирования профессионально-ориентированной диалоговой коллаборативной системы управления роботом с использованием нейросетевых алгоритмов.
2. Разработаны алгоритмы распознавания речи с учетом особенностей решаемых задач, а также организация диалога между медицинским персоналом и роботом, позволяющие вести диалог в темпе, близком к темпу естественной речи.
3. Разработаны способы и алгоритмы нейросетевой реализации задачи распознавания хирургических инструментов, которые могут лежать на хирургическом столе в любом месте и с произвольной ориентацией, по голосовому запросу хирурга.

4. Разработаны способы и программное обеспечение колаборативной робототехнической системы, обеспечивающие автоматический поиск роботом с помощью СТЗ нужного инструмента на рабочем столе, или на стеллаже, и его захват манипулятором.
5. Решена задача надежной и безопасной передачи роботом инструмента хирургу, включая поиск кисти руки хирурга в рабочем пространстве робота с помощью RGB-D камеры, и автоматическое наведение захвата манипулятора с инструментом.
6. Разработана структура организации совместной работы человека и робота при выполнении вспомогательных операций.
7. Проведено экспериментальное исследование отдельных подсистем разработанной КоРТС, а также системы в целом, подтвердившее правильность основных теоретических положений диссертации.
8. Результаты проведенных экспериментальных исследований позволяют утверждать, что на основе разработанной методики и с использованием разработанного программно-алгоритмического обеспечения возможно создание эффективных диалоговых робототехнических систем для широкого круга задач сервисной робототехники

## **СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ**

1. Шуай Инь, А.С. Ющенко. Диалоговая система управления роботом на базе теории конечных автоматов // Мехатроника, автоматизация и управление. 2019. №11. С.686-695. (0.625п.л./0.325п.л.)
2. Shuai Yin, A.S. Yuschenko. Coollaborative robot – surgeon assistant // Extreme robotics, Санкт-Петербург, 2019. С.568-575. (0.5п.л./0.315п.л.)
3. Shuai Yin, A.S. Yuschenko. The application of the convolutional neural network to organize the work of a collaborative robot - surgeon assistant // In: Ronzhin A., Rigoll G., Meshcheryakov R. (eds) Interactive Collaborative Robotics. ICR 2019. Lecture Notes in Computer Science, vol 11659. P.287-297. (0.688п.л./0.425п.л.)
4. Shuai Yin, A.S. Yuschenko. Object Recognition of the Robotic System with Using a Parallel Convolutional Neural Network // Industry 4.0 Issues & New Intelligent Control Paradigms. Studies in Systems, Decision and Control, Springer. 2020, Vol. 272. P.3-11. (0.5625п.л./0.31п.л.)
5. Шуай Инь, А.С. Ющенко. Интеллектуальная система управления мобильным манипуляционным роботом // Всероссийская студенческая конференция «Студенческая научная весна», г. Москва, 2018, С.27-28. (0.75п.л./0.45п.л.)
6. А.С. Ющенко, Шуай Инь, А.Швандт Диалоговое управление коллаборативными роботами с использованием средств нечеткой логики, нейронных сетей и дополненной реальности // Материалы XII мультиконференции по проблемам управления (МКПУ-2019). г Геленджик, 2019, С.180-182. (0.188п.л./0.05п.л.)