

ISSN 2500—0381

Электронная версия

ISSN 0021—3454

Печатная версия

ИЗВЕСТИЯ ВЫСШИХ УЧЕБНЫХ ЗАВЕДЕНИЙ

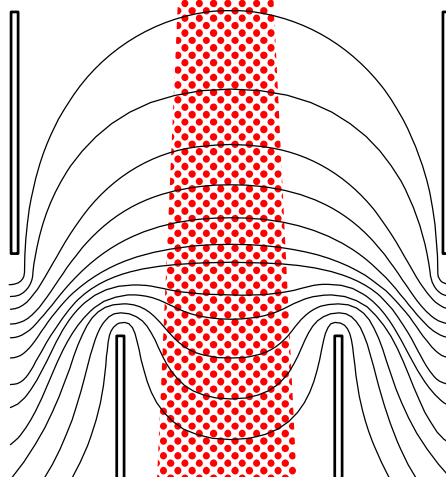
ПРИБОРОСТРОЕНИЕ

НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ



№ **3**
Т. 65
2022

Journal of **I**nstrument **E**ngineering



ГЛАВНЫЙ РЕДАКТОР

проф. А. А. БОБЦОВ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

А. П. АЛЕШКИН, проф., ВКА им. А. Ф. Можайского, Санкт-Петербург, Россия
В. В. ГРИГОРЬЕВ, проф., Университет ИТМО, Санкт-Петербург, Россия
И. П. ГУРОВ, проф., Университет ИТМО, Санкт-Петербург, Россия
В. В. ЕФИМОВ, проф., ВКА им. А. Ф. Можайского, Санкт-Петербург, Россия
Т. А. ИСМАИЛОВ, проф., ДГТУ, Махачкала, Россия
В. М. МУСАЛИМОВ, проф., Университет ИТМО, Санкт-Петербург, Россия
В. О. НИКИФОРОВ, проф., Университет ИТМО, Санкт-Петербург, Россия
А. А. ПЫРКИН (зам. главного редактора), проф., Университет ИТМО, Санкт-Петербург, Россия
В. Я. РАСПОПОВ, проф., ТГУ, Тула, Россия
Л. А. СЕВЕРОВ, проф., ГУАП, Санкт-Петербург, Россия
Б. В. СОКОЛОВ, проф., СПбФИЦ РАН, Санкт-Петербург, Россия
В. С. ТИТОВ, проф., ЮЗГУ, Курск, Россия
В. С. ТОМАСОВ, доц., Университет ИТМО, Санкт-Петербург, Россия
И. Б. ФЕДОРОВ, акад. РАН, МГТУ им. Н. Э. Баумана, Москва, Россия
Н. Б. ФИЛИМОНОВ (зам. главного редактора), проф., МГУ им. М. В. Ломоносова, Москва, Россия
В. Б. ШЛИШЕВСКИЙ, проф., СГГА, Новосибирск, Россия
А. С. ЮЩЕНКО, проф., МГТУ им. Н. Э. Баумана, Москва, Россия

ИНОСТРАННЫЕ ЧЛЕНЫ РЕДКОЛЛЕГИИ

С. В. АБЛАМЕЙКО, акад. НАН Беларуси, БГУ, Минск, Беларусь
Р. М. АЛГУЛИЕВ, акад. НАН Азербайджана, Институт информационных технологий НАНА, Баку, Азербайджан
Т. ЗЕЕГЕР, проф., Университет Зигена, Зиген, Германия
Т. ОРЛОВСКА-КОВАЛЬСКА, проф., Университет Технологий, Вроцлав, Польша
Э. РОМЕРО-КАДАВАЛ, проф., Университет Экстремадура, Бадайоз, Испания
Р. СТЖЕЛЕЦКИ, проф., Институт Электротехники в Варшаве, Гдыня, Польша

Ответственный секретарь *М. В. Лебецкая*

Редактор *И. М. Игошина*

Набор, верстка: *М. В. Герасимова*

Перевод: *Ю. И. Копилевич*

Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49, лит. А, Университет ИТМО

Телефон: (812) 480-02-73, E-mail: pribor@itmo.ru

[http:// pribor.ifmo.ru/](http://pribor.ifmo.ru/)

Подписано в печать 04.04.2022 г.

Отпечатано в учреждении „Университетские телекоммуникации“

Адрес: 197101, Санкт-Петербург, Кронверкский пр., 49, лит. А

Телефон: (812) 915-14-54, E-mail: Prudentov@tibir.ru

Редколлегия просит авторов и рецензентов отнестись с пониманием к ужесточению требований к рукописям и изменению формы подачи материала в журнале.

Аннотация объемом 150—250 слов должна отражать содержание статьи: предмет исследования, цель работы, метод проведения работы, краткое описание экспериментальных или теоретических исследований, полученные результаты и рекомендации по их применению.

Ключевые слова, служащие идентификаторами при предметном поиске, должны однозначно отражать содержание статьи. Основное ключевое слово/словосочетание указывается первым, общее количество ключевых слов — не менее 5.

Список литературы: для обзорной статьи рекомендуется приводить не менее 50 источников, для полнотекстовой статьи — не менее 15, для краткого сообщения — не менее 8 (при составлении списка преимущество следует отдавать изданиям, включенным в международные базы цитирования Scopus и Web of Science).

Новые требования к оформлению рукописи размещены на сайте журнала <http://pribor.itmo.ru/>



Адрес: 197101, Санкт-Петербург,
Кронверкский пр., 49, лит. А, Университет ИТМО
Телефон: (812) 480-02-73
[http:// pribor.ifmo.ru/](http://pribor.ifmo.ru/)
E-mail: pribor@itmo.ru

Журнал публикует научные обзоры, полнотекстовые статьи и краткие сообщения, отражающие результаты практических и теоретических исследований в области приборостроения.

Журнал содержит следующие разделы: информационные технологии и системы, приборы и системы автоматического управления, вычислительная техника, приборы навигации, элементы и устройства вычислительной техники и систем управления, оптические и оптико-электронные приборы и системы, акустические приборы и системы, приборы и методы контроля природной среды, веществ, материалов и изделий, приборы, системы и изделия медицинского назначения, технология приборостроения, научные и практические разработки. По договоренности с Редакцией может быть опубликован специальный тематический выпуск, посвященный достаточно узкой проблеме приборостроения.

Рукопись подается по электронной почте: pribor@itmo.ru. Редакция принимает рукопись к рецензированию при условии получения полного комплекта документов, включающего:

- экспертное заключение о возможности опубликования в открытой печати (формат PDF);
- файлы рукописи статьи, оформленные в соответствии с требованиями журнала;
- сведения об авторах, заполненные по шаблону (на русском и английском языке);
- договор о передаче авторских прав.

Рукопись должна содержать:

- название статьи (на русском и английском языке), индекс УДК;
- аннотацию (150—250 слов, на русском и английском языке), ключевые слова (не менее 5 и не более 25, на русском и английском языке), основной текст, список литературы;
- рисунки: каждый на отдельном листе формата А4 (не заверстывать в текст), таблицы.

Объем текста (оформленного в соответствии с требованиями журнала — шрифт Times New Roman, размер — 14 пт, междустрочный интервал 1,5, поля — сверху и слева 25 мм, снизу 20 мм, справа 10 мм) должен составлять 10—20 страниц, краткого сообщения — 3—5. Объем обзора предварительно согласовывается с Редакцией.

Список литературы (рекомендуется): для обзорной статьи — не менее 50, для полнотекстовой статьи — не менее 15, для краткого сообщения — не менее 8 литературных источников (не менее половины из них должны представлять собой ссылки на издания, включенные в международные базы цитирования Scopus и Web of Science). Список литературы должен формироваться в порядке появления ссылок в тексте.

Не допускается выравнивание с помощью пробелов, табуляций и символов конца абзаца. Эти символы при необходимости могут вставляться не более одного раза подряд (**не переносить с помощью дефисов**).

Формулы и отдельные символы набираются с использованием редактора формул MathType (Microsoft Equation). **Не вставлять** формулы из пакетов MathCad и MatLab. Размер шрифта в формулах — по умолчанию. Все переменные набираются курсивом, греческие буквы — прямым шрифтом, полужирные символы в формуле — стилем „матрица вектор“, русские символы — стилем „текст“, векторы следует набирать полужирным шрифтом без стрелок.

Рисунки выполняются в градациях серого в одном из форматов приложений Microsoft (Word, Excel, PowerPoint), предпочтителен векторный формат. Кривые на графиках следует помечать цифрами (1, 2, 3 и т.д.) или/и задавать линиями разного стиля (пунктир, штрихпунктир, жирная); толщина осей должна быть не менее 0,5 пт, иначе качество печати не гарантируется.

Датой поступления статьи считается последняя после доработки.

Плата за публикацию не взимается.

СОДЕРЖАНИЕ

Под редакцией доктора технических наук, профессора С. А. Колюбина

ПРОЕКТИРОВАНИЕ АДАПТИВНЫХ И ЭНЕРГОЭФФЕКТИВНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ

- Хомутов Е. Э., Иволга Д. В., Борисов И. И., Молчанов Н. А., Максимов И. А.
Структурно-параметрический синтез плоских неполноприводных механизмов
для антропоморфных роботизированных кистей 155
- Иволга Д. В., Хомутов Е. Э., Борисов И. И., Молчанов Н. А., Максимов И. А.,
Колюбин С. А. Программно-аппаратная архитектура системы управления ан-
тропоморфной кистью робота..... 164
- Насонов К. В., Борисов И. И. Проектирование неполноприводного прыгающего
робота с гибкими сочленениями 174

КОМПЬЮТЕРНОЕ ЗРЕНИЕ И ПЛАНИРОВАНИЕ ДВИЖЕНИЯ РОБОТОВ В ЗАДАЧАХ МАНИПУЛИРОВАНИЯ

- Довгополик И. С., Артемов К., Борисов О. И., Забихифар С., Семочкин А. Н.
Алгоритм модифицированного интеллектуального двунаправленного случайно-
го дерева для планирования движения антропоморфных манипуляторов..... 185
- Ненахов И. Д., Артемов К., Забихифар С., Семочкин А. Н., Колюбин С. А. Сег-
ментация объектов с функцией дообучения 194

МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СИСТЕМАХ НАВИГАЦИИ МОБИЛЬНЫХ РОБОТОВ

- Беркаев А. Р., Мохрат М., Бурков А. М., Колюбин С. А. Метрико-семантическое
картирование на основе глубоких нейронных сетей для систем автономной на-
вигации в помещениях 204
- Махмуд Ж., Ха Тхе Лонг В., Бурков А. М., Колюбин С. А. Оптимизационные ал-
горитмы повышения точности и робастности визуальной одометрии наземных
мобильных роботов 218

EDITOR-IN-CHIEF

prof. ALEXEY BOBTSOV

EDITORIAL BOARD

ALESHKIN ANDREY, prof., A. F. Mozhaisky Military Space Academy, St. Petersburg, Russia

GRIGORIEV VALERY, prof., ITMO University, St. Petersburg, Russia

GUROV IGOR, prof., ITMO University, St. Petersburg, Russia

EFIMOV VLADIMIR, prof., A. F. Mozhaisky Military Space Academy, St. Petersburg, Russia

ISMAILOV TAGIR, prof., DSTU, Makhachkala, Russia

MUSALIMOV VIKTOR, prof., ITMO University, St. Petersburg, Russia

NIKIFOROV VLADIMIR, prof., ITMO University, St. Petersburg, Russia

PYRKIN ANTON, prof., Deputy Chief Editor, ITMO University, St. Petersburg, Russia

RASPOPOV VLADIMIR, prof., TSU, Tula, Russia

SEVEROV LEONID, prof., SUAI, St. Petersburg, Russia

SOKOLOV BORIS, prof., SPbFRC RAS, St. Petersburg, Russia

TITOV VITALY, prof., SWSU, Kursk, Russia

TOMASOV VALENTIN, prof., ITMO University, St. Petersburg, Russia

FEDOROV IGOR, acad., N. E. Bauman MSTU, Moscow, Russia

FILIMONOV NIKOLAY, prof., Deputy Chief Editor, M. V. Lomonosov MSU, Moscow, Russia

SHLISHEVSKY VIKTOR, prof., SSGA, Novosibirsk, Russia

YUSHCHENKO ARKADY, prof., N. E. Bauman MSTU, Moscow, Russia

FOREIGN MEMBERS

ABLAMEYKO SERGEY, acad., Belarusian State University, Minsk, Belarus

ALGULIYEV RASIM, acad., Institute of Information Technology of ANAS, Baku, Azerbaijan

ORLOWSKA-KOWALSKA TERESA, prof., University of Technology, Wroclaw, Poland

ROMERO-CADAVAL ENRIQUE, prof., University of Extremadura, Badajoz, Spain

SEEGER THOMAS, prof., University of Siegen, Siegen, Germany

STRZELECKI RYSZARD, prof., Electrotechnical Institute in Warsaw, Gdynia, Poland

Executive secretary Marina V. Lebetskaya

Editor Irina M. Igoshina

Page-proof, makeup Maria V. Gerasimova

Translation of Yuriy I. Kopilevich

Address: ITMO University, Kronverksky pr. 49, bldg. A, St. Petersburg, Russia, 197101

Tel: (812) 480-02-73

Site: pribor.ifmo.ru

E-mail: pribor@itmo.ru

JOURNAL OF INSTRUMENT ENGINEERING

ITMO UNIVERSITY

The Journal is published since January 1958

Vol. 65

MARCH 2022

No 3

CONTENTS

By Edition of S. A. Kolyubin, Dr. Sci., Professor

DESIGN OF ADAPTIVE AND ENERGY-EFFICIENT ROBOTIC SYSTEMS

- Khomutov E. E., Ivolga D. V., Borisov I. I., Molchanov N. A., Maksimov I. A.** Structural-Parametric Synthesis of Planar Underactuated Linkages for Anthropomorphic Robotic Hands 155
- Ivolga D. V., Khomutov E. E., Borisov I. I., Molchanov N. A., Maksimov I. A., Kolyubin S. A.** Hardware and Software Architecture for Anthropomorphic Robot Hand Control System 164
- Nasonov K. V., Borisov I. I.** Design of an Underactuated Jumping Robot with Flexible Joints..... 174

COMPUTER VISION AND ROBOT MOVEMENT PLANNING IN MANIPULATION TASKS

- Dovgopolik I. S., Artemov K., Borisov O. I., Zabihiyar S., Semochkin A. N.** Modified Intelligent Bidirectional Random Tree Algorithm for Planning the Movement of Anthropomorphic Manipulators 185
- Nenakhov I. D., Artemov K., Zabihiyar S., Semochkin A. N., Kolyubin S. A.** Objects Segmentation with Retraining Function 194

ARTIFICIAL INTELLIGENCE METHODS IN MOBILE ROBOT NAVIGATION SYSTEMS

- Berkaev A. R., Mohrat M., Burkov A. M., Kolyubin S. A.** Metric-Semantic Mapping Based on Deep Neural Networks for Systems of Indoor Autonomous Navigation 204
- Mahmoud J., Ha The Long V., Burkov A. M., Kolyubin S. A.** Optimization Algorithms for Improving the Accuracy and Robustness of Visual Odometry of Ground-Based Mobile Robots 218

To the Authors!

The Journal of Instrument Engineering publishes scientific reviews, full-text articles and short messages, reflecting the unpublished results of theoretical and practical research in the field of fundamental instrument making.

The Journal publishes articles on the following themes: information technologies and systems, devices and automatic control systems, computer engineering, gyroscopic and navigation systems, instruments of precision mechanics, electronic and electromagnetic devices, optical and optoelectronic devices and systems, thermal regimes and reliability of devices and systems, technology and instrumentation, scientific and practical development. In agreement with the Editorial Board can be published a special themed issue devoted to a fairly narrow issue of instrumentation.

The material of the article accepted in electronic form and on any media, executed in a text editor Word without formatting along with a hard copy on A4. Electronic version of the paper may be sent by E-mail: **pribor@mail.ifmo.ru**

The size of the article (in printed form) must not exceed eight pages, a short presentation of two. Size scientific review previously agreed with the Editorial Board.

Submission includes two copies of:

— the manuscript text; list of references (recommended): review article — not less than 50, for the full text of the article — no less than 15, for short messages — at least 8 of the literary source (at least half of them should be links to the editions listed in international database of citation Scopus and Web of Science); number citations consecutively as they appear in the text;

— extended abstract (150—250 words);

— keywords (not less than 5 and not more than 25);

— illustrations: each on a separate sheet of A4 (not overestimate in the text) — big, with a maximum fill;

— the recommendation of the Department (laboratory);

— information about the authors on a template (surname, name, patronymic, name of the organization in accordance with the latest edition of the Charter, the address of the organization, academic degree and title, position, e-mail address).

The article must also enclose agreement on the transfer of copyright.

The main text. The following requirements are common to good practice in the design documents.

Font — Times New Roman, size 14, line spacing 1.5, margins — top and left — 25 mm, bottom — 20 mm, right — 10 mm.

Not allowed alignment with spaces, tabs and end of a paragraph (in the layout will need to be cleared). These symbols must be inserted only when necessary, and not more than once in a row (not to carry with dashes).

Don't insert formulas from MathCad or MatLab!

Figures and tables are to be inserted after the list of references. Graphics are accepted in one of Microsoft formats (Word, Excel, and PowerPoint); vector format is preferable. Don't use color graphics; it is better to apply dotted, dash-and dot lines, etc., or mark lines with numerical symbols.

Editorial Staff

Уважаемые читатели!

Серьезное развитие технологий искусственного интеллекта позволяет находить решения разнообразных задач, связанных с обработкой информации и принятием решений, которые „классическими методами“ зачастую реализуются хуже. С другой стороны, для дальнейшего развития, с фокусом на сильный/общий искусственный интеллект, востребованными являются исследования по физическому воплощению самообучающихся алгоритмов в агентах (роботах), решающих задачи сенсорно-моторной координации в активном взаимодействии с динамичной средой и человеком. Симбиоз между искусственным и естественным интеллектом способен дать замечательные результаты, но перенос алгоритмов машинного обучения из симуляторов на реальных роботов сопряжен с рядом „узких мест“ искусственного интеллекта: слабое понимание нейросетями контекста и адаптация к нему; сложность интерпретации работы и верификации; „жадность“ к объемам передаваемой и хранимой информации и вычислительной мощности, что также приводит к значительным энергетическим затратам и увеличению массогабаритных показателей бортовых электронных модулей; недостаточная робастность к параметрическим, структурным и сигнальным возмущениям. На реализацию упомянутых возможностей и преодоление отмеченных барьеров и нацелены совместные исследования и разработки лаборатории биомехатроники и энергоэффективной робототехники Университета ИТМО и лаборатории робототехники ПАО „Сбербанк“, некоторые из результатов которых представлены в настоящем выпуске журнала.

Выпуск содержит три раздела: проектирование адаптивных и энергоэффективных робототехнических систем; компьютерное зрение и планирование движения роботов в задачах манипулирования; методы искусственного интеллекта в системах навигации мобильных роботов. Первый раздел посвящен использованию принципов морфологических вычислений при синтезе „умной“ механики локомоционных роботов и роботизированных захватных устройств, способных адаптироваться к неструктурированной среде и устойчиво с ней взаимодействовать. Во втором разделе представлены статьи, посвященные задачам управления манипуляционными роботами на тактическом уровне, а именно автоматическому определению и классификации объектов манипулирования и планированию опорных траекторий для их перемещения. Цель этих работ состоит в ускорении процесса поиска оптимального решения и минимизации необходимых для этого вычислительных ресурсов. Третий раздел содержит результаты по повышению автономности наземных мобильных роботов. Здесь речь идет об использовании нейросетевых подходов и методов численной оптимизации для повышения качества и информативности генерируемых роботом трехмерных карт окружающего пространства, а также для улучшения робастности и точности визуально-инерциальной одометрии, используемой для локализации роботов в априори неизвестной динамичной среде.

*Доктор технических наук,
профессор факультета систем управления
и робототехники Университета ИТМО
С. А. КОЛЮБИН*

*Руководитель
лаборатории робототехники ПАО „Сбербанк“
кандидат технических наук
А. С. ГОННОЧЕНКО*

To the Readers!

The serious development of artificial intelligence technologies makes it possible to find solutions to various tasks related to information processing and decision-making, which are often implemented worse by „classical methods“. On the other hand, for further development, with a focus on strong/general artificial intelligence, research on the physical embodiment of self-learning algorithms in agents (robots) solving problems of sensory-motor coordination in active interaction with a dynamic environment and a person is in demand. The symbiosis between artificial and natural intelligence is able to give remarkable results, but the transfer of machine learning algorithms from simulators to real robots is associated with a number of artificial intelligence „bottlenecks“: weak understanding of the context by neural networks and adaptation to it; difficulty in interpreting work and verification; „greed“ for the volumes of transmitted and stored information and computing power, which also translates into significant energy costs and an increase in the weight and size indicators of on-board electronic modules; insufficient robustness to parametric, structural and signal disturbances. Joint research and development of the Laboratory of Biomechatronics and Energy-Efficient Robotics of ITMO University and the Robotics Laboratory of Sberbank PJSC are aimed at realizing the mentioned opportunities and overcoming the noted barriers, some of the results of which are presented in this issue of the journal.

The issue contains three sections: design of adaptive and energy-efficient robotic systems; computer vision and robot movement planning in manipulation tasks; artificial intelligence methods in mobile robot navigation systems. The first section is devoted to the use of the principles of morphological calculations in the synthesis of „smart“ mechanics of locomotion robots and robotic gripping devices capable of adapting to an unstructured environment and interacting with it stably. The second section presents articles devoted to the tasks of controlling manipulative robots at the tactical level, namely, automatic detection and classification of objects of manipulation and planning reference trajectories for their movement. The purpose of these works is to accelerate the process of finding the optimal solution and minimize the computing resources necessary for this. The third section contains the results on increasing the autonomy of ground-based mobile robots. Here we are talking about the use of neural network approaches and numerical optimization methods to improve the quality and informativeness of three-dimensional maps of the surrounding space generated by the robot, as well as to improve the robustness and accuracy of visual inertial odometry used to localize robots in an a priori unknown dynamic environment.

S. A. KOLYUBIN

Dr. Sci., Professor

ITMO University, Faculty of Control Systems and Robotics

A. S. GONNOCHENKO

*PhD, Head of the Robotics Laboratory
of Sberbank PJSC*

ПРОЕКТИРОВАНИЕ АДАПТИВНЫХ И ЭНЕРГОЭФФЕКТИВНЫХ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ

DESIGN OF ADAPTIVE AND ENERGY-EFFICIENT ROBOTIC SYSTEMS

УДК 621.837.5
DOI: 10.17586/0021-3454-2022-65-3-155-163

СТРУКТУРНО-ПАРАМЕТРИЧЕСКИЙ СИНТЕЗ ПЛОСКИХ НЕПОЛНОПРИВОДНЫХ МЕХАНИЗМОВ ДЛЯ АНТРОПОМОРФНЫХ РОБОТИЗИРОВАННЫХ КИСТЕЙ

Е. Э. Хомутов¹, Д. В. Иволга¹, И. И. Борисов^{1*}, Н. А. Молчанов², И. А. Максимов²

¹ Университет ИТМО, Санкт-Петербург, Россия

* borisovii@itmo.ru

² Сбербанк, Москва, Россия

Аннотация. Представлены результаты структурно-параметрического синтеза плоских неполноприводных механизмов со звеньями переменной длины. Синтез основан на принципе морфологического расчета, согласно которому желаемые свойства робота „программируются“ на уровне механической конструкции, а активное управление используется лишь для мягкой корректировки движений, обусловленных естественным поведением системы. Разработанный алгоритм был апробирован при создании адаптивной компактной кисти для антропоморфного робота iCub, способной выполнять все фундаментальные захваты и обладающей при сопоставимых габаритах более высокой грузоподъемностью по сравнению с оригинальным аналогом на тросовой тяге. Приведен пример синтеза механизма указательного пальца и представлены результаты тестирования опытного образца устройства. Разработанный прототип обладает 14 степенями свободы, но при этом управляется с помощью четырех двигателей. Минимизация количества приводов позволила упростить систему управления, уменьшить количество требуемых сенсоров и массогабаритные параметры и снизить затраты на комплектующие.

Ключевые слова: антропоморфные хватные устройства, морфологический расчет, манипулирование объектами, звенья переменной длины, моделирование контактного взаимодействия

Ссылка для цитирования: Хомутов Е. Э., Иволга Д. В., Борисов И. И., Молчанов Н. А., Максимов И. А. Структурно-параметрический синтез плоских неполноприводных механизмов для антропоморфных роботизированных кистей // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 155—163. DOI: 10.17586/0021-3454-2022-65-3-155-163.

STRUCTURAL-PARAMETRIC SYNTHESIS OF PLANAR UNDERACTUATED LINKAGES FOR ANTHROPOMORPHIC ROBOTIC HANDS

E. E. Khomutov¹, D. V. Ivolga¹, I. I. Borisov^{1*}, N. A. Molchanov², I. A. Maksimov²

¹ ITMO University, St. Petersburg, Russia

* borisovii@itmo.ru

² Sberbank, Moscow, Russia

Abstract. The results of structural-parametric synthesis of planar underactuated linkages with variable length links are presented. The proposed synthesis is based on the principle of morphological computation, which means that robot's desired properties are „programmed“ at mechanical level, while active control is used only for soft motion correction caused by the natural behavior of the system. The proposed synthesis algorithm was tested to create an adaptive compact hand for the anthropomorphic robot iCub, capable of performing all fundamental grips and having a higher load capacity with dimensions similar to the original cable-driven hand design. An example of the synthesis for an index finger mechanism is given and the results of testing a prototype device are presented. The developed prototype has 14 de-

degrees of freedom, but it is controlled by just four motors. Minimizing the number of motors made it possible to simplify the hardware and software of its control system, reduce the number of required sensors and weight and size parameters and reduce the cost of components.

Keywords: grippers and other end-effectors, compliant joint/mechanism, variable length links, flexible robots, mechanism design

For citation: Khomutov E. E., Ivolska D. V., Borisov I. I., Molchanov N. A., Maksimov I. A. Structural-parametric synthesis of planar underactuated linkages for anthropomorphic robotic hands. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 155—163 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-155-163.

Введение. Во многих робототехнических устройствах, например таких, как пальцы антропоморфных захватных устройств [1—3], роботы-манипуляторы [4], ноги шагающих роботов [5], используется простая, в понятиях механики, полноприводная конструкция открытой кинематики. Такой подход универсален, так как позволяет алгоритмически реализовать любую траекторию в рамках конфигурационного пространства, однако требует нетривиальных алгоритмов управления скоординированным движением всех сочленений робота одновременно. Кроме того, для полноприводной системы требуется количество приводов, равное количеству степеней свободы, что влияет на габариты, массу и стоимость устройства. Примерами антропоморфных захватов на основе полноприводной системы с открытой кинематической цепью являются DLR Hand II [1], Shadow EDC Hand [2] и бионическая рука [3]. Помимо сложности системы управления, из-за большого количества двигателей подобные роботизированные руки снабжаются массивными предплечьями, в которых располагаются двигатели, что ограничивает их практическое использование.

Для робототехнических устройств, движение которых не требует высокой степени реконфигурации, возможно синтезировать механизмы с минимумом управляющих двигателей, в том числе рычажные механизмы с одной степенью свободы. В работе [6] представлен метод синтеза многозвенных рычажных механизмов аниматронных персонажей, исполнительные элементы которых движутся по сложным траекториям, при этом используется минимум двигателей для обеспечения движения: так как конечности персонажей движутся по одной траектории, нет необходимости реконфигурировать траекторию их перемещения.

Некоторые протезы кисти оснащаются механизмами пальцев с одной степенью свободы, благодаря чему они надежны в эксплуатации и предсказуемы по поведению. Однако такие механизмы не способны адаптироваться к форме захватываемых объектов и поэтому количество жестов, которые может выполнять рука, сильно ограничено [7]. В робототехнике многие устройства тоже должны выполнять однотипные движения, но если устройство контактирует с внешней средой, оно должно быть способно адаптироваться, т.е. реконфигурироваться под действием силы контакта.

В настоящей статье представлен алгоритм синтеза неполноприводных рычажных механизмов замкнутой кинематики с эластичными элементами, необходимыми для требуемой реконфигурации при взаимодействии механизма с окружающей средой. Хотя рассматриваемый принцип применим для различных механизмов, например механизм ноги в работе [8] был синтезирован аналогичным образом, — в рамках данной статьи разработанный алгоритм описывается на примере синтеза механизмов пальцев антропоморфного захватного устройства. Предлагаемый алгоритм позволяет упростить управление, но при этом сохранить большое количество степеней свободы разрабатываемых механизмов с помощью интеграции элементов переменной геометрии и эластичности: пружин сжатия/растяжения и работающих на изгиб гибких элементов.

В рамках исследования была поставлена задача создать захватное устройство, способное обеспечить силовой адаптивный и точный щипковый типы захвата, а также повысить грузоподъемность, минимизировать массу кисти и количество двигателей, приводимых в

действие простыми алгоритмами управления, по сравнению с оригинальной конструкцией кисти робота iCub. Существующая оригинальная конструкция кисти робота iCub, построенная на основе неполноприводной схемы с открытой кинематической цепью с системой тросов и шкивов, обладает компактными размерами, высокой адаптивностью и богатой жестикующей, однако для нее свойственны ненадежный щипковый захват, низкая повторяемость движений и малая грузоподъемность (до 350 г) [9, 10].

Структурно-параметрический синтез механизмов пальцев. Цель синтеза механизмов пальцев заключается в определении *структуры* и *геометрических параметров* проектируемого механизма, при которых движение пальцев будет иметь требуемый характер.

В [11] предложен метод проектирования универсальных адаптивных захватных устройств, способных обеспечить оба фундаментальных вида захвата: точный щипковый и силовой обхватывающий, благодаря использованию звеньев переменной длины и механической декомпозиции каналов управления. Базирующийся на этом методе алгоритм структурно-параметрического синтеза механизма пальца может быть описан следующим образом.

Шаг 1. Задать требуемый характер движения (траекторию, рабочее пространство и пр.), количество фаланг и сформировать полноприводный механизм пальца открытой кинематики.

Шаг 2. Замкнуть кинематическую схему механизма путем дополнения голономных связей в целях перемещения приводов в корпус ладони без изменения их количества.

Шаг 3. Рассчитать возможные структуры присоединяемых групп звеньев для редуцирования количества степеней свободы механизма и выбрать из них оптимальную в смысле наиболее простой конструкторской реализации.

Шаг 4. Решить задачу оптимизации длин и расположения точек крепления звеньев выбранной присоединяемой группы для получения полноприводного механизма с редуцированным числом степеней свободы.

Шаг 5. Определить, какое из звеньев присоединяемой группы следует заменить на звено переменной длины для получения неполноприводных механически адаптивных пальцев захвата, соответствующих предъявляемым требованиям.

Поскольку указательный, средний, безымянный пальцы и мизинец обладают единой структурой, то пример синтеза дан только для механизма указательного пальца. Синтез механизма большого пальца приведен в [12, 13].

Открытая кинематическая схема и характер движения. При синтезе механизма указательного пальца в качестве базовой была взята полноприводная открытая кинематическая схема с двумя степенями свободы. На рис. 1, а показан требуемый характер сгибания пальца при отсутствии контактного взаимодействия. Палец вращается относительно корпуса ладони при сохранении расположения фаланг друг относительно друга.

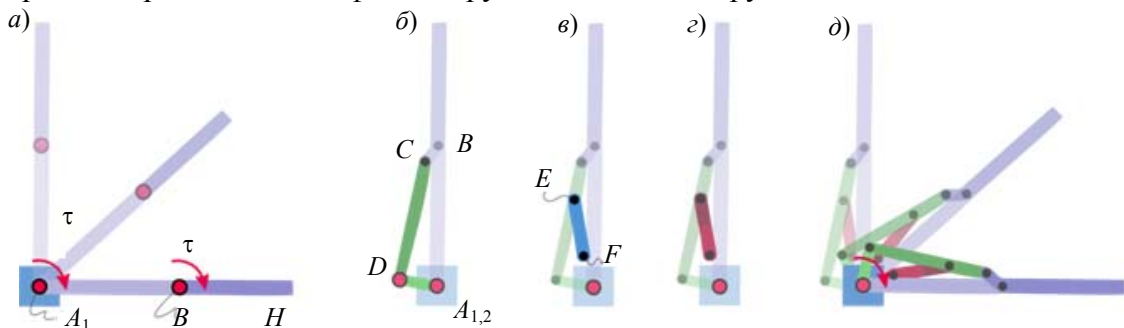


Рис. 1

Перемещение двигателей в корпус ладони. При проектировании антропоморфного захватного устройства важно сохранить пропорции человеческой кисти, что накладывает ограничения на габариты комплектующих элементов. Для обеспечения высокой грузоподъемности требуются габаритные двигатели, которые невозможно уместить в суставах пальца.

Двигатели можно разместить в корпусе ладони, а движение на суставы передать с помощью тросов, как в оригинальной конструкции iCub, либо с помощью рычажных механизмов, в частности механизмов замкнутой кинематики (см. далее).

У базового механизма пальца две степени свободы (см. рис. 1, а), что может быть рассчитано по формуле Чебышева для плоских механизмов

$$W_0 = 3n - 2P_5 = 3 \cdot 2 - 2 \cdot 2 = 2,$$

где n — число подвижных звеньев, P_5 — число кинематических пар с одной степенью свободы.

Для замыкания кинематической цепи базового механизма к нему следует присоединить группу звеньев, которая не дополняет и не редуцирует степень свободы, т.е. надо присоединить группу звеньев Ассур с количеством степеней свободы $W_a = 0$. С помощью формулы Чебышева можно рассчитать сочетания звеньев и кинематических пар групп звеньев, которые могут быть присоединены для замыкания механизма:

$$W_a = 3n_a - 2P_a = 0,$$

где возможны группы $n_a = 2$ и $P_a = 3$ или $n_a = 4$ и $P_a = 6$, также может быть рассчитана группа с большим количеством звеньев, но увеличение числа звеньев не рационально. Самой простой группой звеньев в рассматриваемом примере является двухповодковая с $n_a = 2$ и $P_a = 3$.

Для замыкания кинематической цепи один крайний поводок присоединяемой группы должен крепиться к крайнему звену базового механизма, т.е. к дистальной фаланге, а второй крайний поводок должен крепиться к стойке, т.е. к корпусу. При этом необходимо обеспечить следование пальца по требуемой траектории, для чего должна быть проведена параметрическая оптимизация координат точек крепления звеньев и их длин.

На рис. 1, б изображен механизм пальца с присоединенной группой звеньев с контуром ADC. Группа свободных параметров, определяемых в процессе оптимизации, состоит из длин l_1 и l_2 присоединяемых звеньев AD и DC соответственно и координат точек их присоединения к дистальной фаланге и стойке — координат точек C и A₂ в локальных системах координат звеньев. Группа заданных проектировщиком фиксированных параметров, помимо длин фаланг пальца, определяется требуемым положением линии равновесия (подробности приведены ниже). Численная оптимизация параметров была проведена в программной среде MatLab с использованием пакета Global Optimization Toolbox и метода patternsearch. Целью оптимизации является минимизация длин присоединяемых звеньев и расстояния между требуемой и фактической траекториями двух крайних точек дистальной фаланги:

$$F_1 = k_1 l_1 + k_2 l_2 + k_3 \sqrt{\frac{1}{q} \sum_i \|p_{1,i}^* - p_{1,i}\|^2 + \|p_{2,i}^* - p_{2,i}\|^2},$$

где $p_{1,i}^* = [x_{1,i}^*, y_{1,i}^*]^T$ и $p_{2,i}^* = [x_{2,i}^*, y_{2,i}^*]^T$ — координаты требуемых траекторий движения крайних точек дистальной фаланги в момент времени i ; $p_{1,i}$ и $p_{2,i}$ — координаты фактических траекторий движения крайних точек дистальной фаланги в момент времени i ; q — количество измерений, k_i — весовые коэффициенты.

Минимизация длин присоединяемых звеньев гарантирует, что размеры присоединяемой группы будут наиболее компактными, потребуется меньше материала, а значит, механизм будет наиболее легким. Минимизация расстояния между желаемой и фактической траекториями гарантирует, что синтезированный механизм будет следовать по требуемой траектории. В процессе решения задачи оптимизации выполняется поиск координат точек крепления p_1 , p_2 и длин звеньев l_1 , l_2 , присоединяемых к базовому механизму, при которых обеспечивается минимум целевой функции F_1 .

Редуцирование количества степеней свободы механизма. Полученный на предыдущем этапе механизм замкнутой кинематики имеет две степени свободы (см. рис. 1, б). Реду-

цирование количества степеней свободы до одной проведем путем наложения дополнительных голономных связей, присоединив группу звеньев с W_f степенями свободы:

$$W_f = -(W_n - A),$$

где W_n — степень свободы механизма замкнутой кинематики, A — требуемое число приводов пальца, в рассматриваемом примере $A=1$.

Аналогично предыдущему этапу, с помощью формулы Чебышева можно рассчитать сочетания звеньев и кинематических пар группы звеньев, которые могут быть присоединены для редуцирования степени свободы. В рассматриваемом примере $A=1$, а $W_f=-1$, т.е.

$$W_f = 3n_f - 2P_f = -1,$$

где возможны группы $n_f=1$ и $P_f=2$ или $n_f=3$ и $P_f=5$; также может быть рассчитана группа с большим количеством звеньев, но увеличение числа звеньев не рационально. Самым простым элементом в рассматриваемом примере является присоединение одного звена с $n_f=1$ и $P_f=2$.

При интегрировании этого звена в механизм замкнутой кинематики необходимо обеспечить следование по требуемой траектории, для чего проводится параметрическая оптимизация положения точек крепления присоединяемого звена. Группа свободных параметров, определяемых в процессе оптимизации, состоит из координат точек крепления присоединяемого звена, выраженных в локальных системах координат звеньев, к которым звено присоединяется. Для двух звеньев определяется пара точек p_a и p_b , абсолютное расстояние между которыми изменяется минимально на протяжении всего цикла движения. Среднеквадратическое расстояние между двумя точками рассчитывается как

$$l_{ab} = \frac{1}{q} \sum_i \|p_{a,i} - p_{b,i}\|^2,$$

где $p_{a,i}$ и $p_{b,i}$ — координаты точек p_a и p_b в мировой системе координат в момент времени i .

Отклонение среднеквадратического расстояния между точками крепления звена

$$F_2 = \frac{1}{q} \sum_i \left(\|p_{a,i} - p_{b,i}\|^2 - l_{ab} \right)^2$$

используется в качестве целевой функции [9]; минимизация функции F_2 гарантирует, что величина l_{ab} постоянна при движении механизма по требуемой траектории.

Интеграция звеньев переменной длины. На рис. 1, в изображен синтезированный механизм указательного пальца, в котором посредством звена EF вводятся дополнительные голономные связи. Полученная полноприводная закрытая кинематическая цепь обладает одной степенью свободы и приводится в движение одним приводом. Однако данный механизм не способен адаптироваться к форме объекта.

Для придания свойств адаптивности пальцу захватного устройства необходимо выбрать звено, которое будет заменено звеном переменной длины для обеспечения реконфигурации между щипковым и адаптивным захватами. Под звеном переменной длины будем понимать систему из двух твердых тел, образующих поступательную кинематическую пару, и упругого элемента в виде пружины сжатия, контролирующего расстояние в поступательной кинематической паре.

На рис. 1, г показано абсолютно твердое звено EF , замененное звеном переменной длины, которое способно растягиваться при приложении нагрузки к рабочим поверхностям фаланг. При отсутствии контактного взаимодействия палец следует требуемой траектории (рис. 1, д).

Взаимодействие указательного пальца с объектами манипулирования продемонстрировано на рис. 2. Для синтезированного неполноприводного механизма пальца замкнутой кинематики с двумя степенями свободы свойственна *линия равновесия* [14], разделяющая рабочую поверхность дистальной фаланги на зону щипкового захвата и зону адаптивного захвата. При захвате объекта кончиками пальцев будет выполняться щипковый захват (рис. 2, а), а если точки касания пальцев с объектами находятся ближе к сочленению, то будет выполнен адаптивный захват (рис. 2, б). Особенности параметрического синтеза механизмов пальца с учетом требуемого положения линии равновесия приведены в [14]. На рис. 2, в изображена адаптация пальца при захвате сферического объекта.

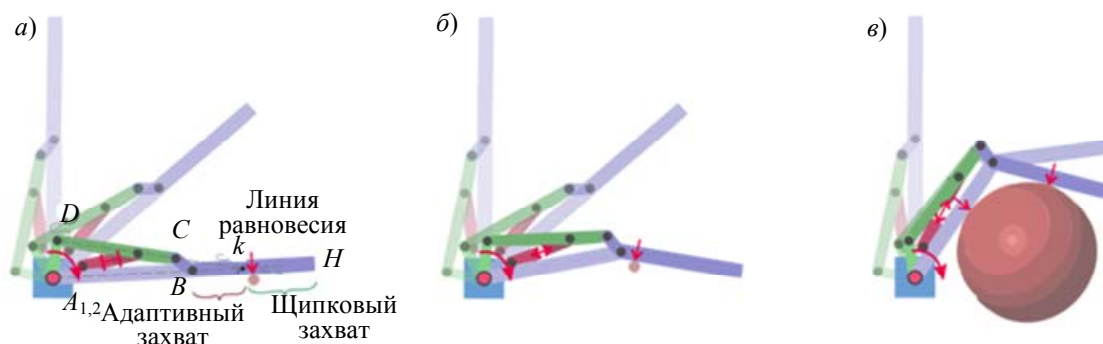


Рис. 2

Тестирование захватного устройства. На основе предложенного алгоритма синтеза механизмов пальцев была спроектирована конструкция антропоморфного реконфигурируемого неполноприводного захватного устройства для гуманоидного робота iCub. Была поставлена задача спроектировать конструкцию роботизированной кисти в габаритах оригинальной кисти iCub, массой до 0,5 кг, грузоподъемностью не менее 1 кг, которая способна выполнять большое количество жестов, захватывать и надежно удерживать объекты разных размеров, произвольной массы, формы и жесткости.

На рис. 3 представлен опытный образец захватного устройства, который производит захват произвольных объектов, выполняя как точные щипковые, так и силовые адаптивные захваты. Разработанный захват обладает 14 степенями свободы, но требует только четыре управляющих двигателя. Рабочие зоны захвата с предельными размерами объектов манипулирования изображены на рис. 4. Основные сравнительные характеристики оригинального захвата робота iCub и разработанной конструкции захватного устройства приведены в таблице.

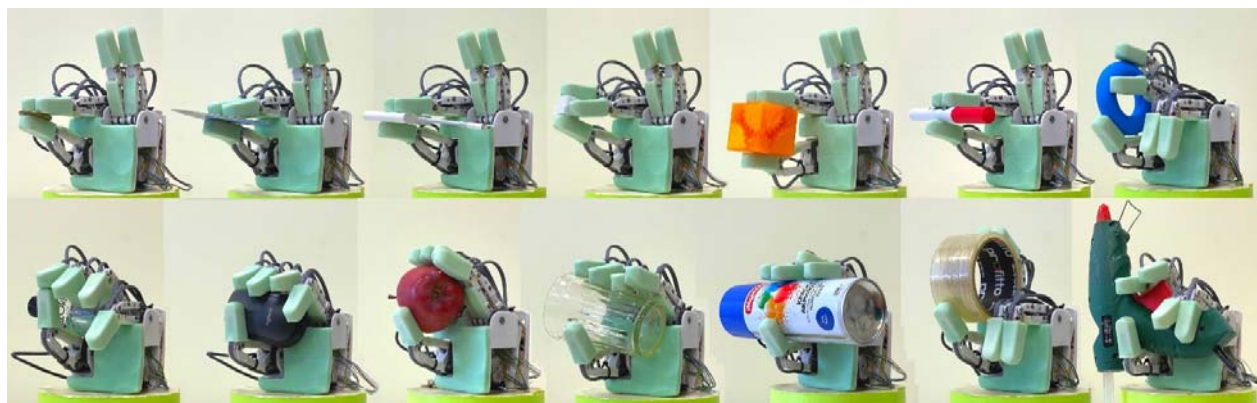


Рис. 3

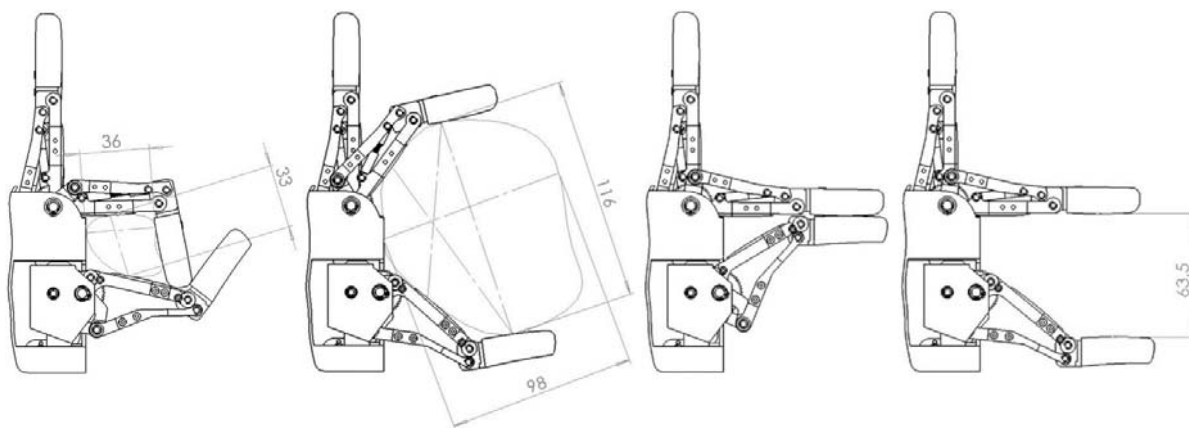


Рис. 4

Характеристика	iCub [12]	Захватное устройство
Усредненная сила, создаваемая одним пальцем, Н	До 2	До 9,6
Размеры, мм	165×138×40	186×170×40
Масса, г	250	495
Грузоподъемность, г	350	Более 1000
Количество актуаторов, шт	9	4
Габаритный диаметр сечения объекта при адаптивном захвате, мм	14—100	36—116
Габаритный диаметр сечения объекта при щипковом захвате, мм	25	0—63,5

Целями тестирования захвата антропоморфной кисти были следующие:

- проверка кинематики синтезированных пальцев для верификации их поведения при выполнении жестов и захватов;
- проверка работы эластичных элементов и адаптации пальцев к форме объекта при контакте;
- проверка функционирования системы управления с реализацией работы датчиков касания и датчиков углового положения.

Результаты тестирования показали, что кинематика синтезированных пальцев работает согласно ожиданиям и соотносится с результатами моделирования. При контакте только кончиков дистальных фаланг с объектом осуществляется щипковый захват. Когда задействованы проксимальные фаланги, звено переменной длины увеличивается и захват начинает адаптироваться к форме объекта. Несмотря на то, что указательный и средний пальцы приводятся в движение одним приводом, они могут адаптироваться к форме объекта благодаря подпружиненности входных звеньев пальцев относительно общего приводного вала. Если в контакте участвует только указательный палец, то средний палец продолжает двигаться по заданной траектории, и наоборот.

Заключение. Предложенный алгоритм структурно-параметрического синтеза адаптивных неполноприводных механизмов на основе принципа морфологического проектирования позволяет создавать робототехнические устройства, которые часть задачи управления выполняют благодаря своей механической конструкции, упрощая сложность вычислений алгоритмов системы управления и требования к техническому оснащению устройства. В результате в качестве апробации алгоритма синтеза спроектирована конструкция захватного устройства с 14 степенями свободы, но требующего только четыре двигателя для приведения в движение всего устройства: два для большого пальца, один для указательного и среднего пальцев и один для безымянного пальца и мизинца, тогда как для управления оригинальной кистью робота iCub требуется 9 приводов. Адаптация пальцев захвата к форме объектов манипулирования

происходит автоматически за счет механической структуры и не требует сложных алгоритмов управления для скоординированного движения большого количества приводов.

СПИСОК ЛИТЕРАТУРЫ

1. Liu H. et al. Multisensory five-finger dexterous hand: The DLR/HIT Hand II // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. IEEE, 2008. P. 3692—3697.
2. Rothling F. et al. Platform portable anthropomorphic grasping with the bielefeld 20-dof shadow and 9-dof tum hand // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. IEEE, 2007. P. 2951—2956.
3. Xu Z., Todorov E. Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration // IEEE Intern. Conf. on Robotics and Automation (ICRA). IEEE, 2016. P. 3485—3492.
4. Raiola G. et al. Development of a safety-and energy-aware impedance controller for collaborative robots // IEEE Robotics and Automation Letters. 2018. Vol. 3, N 2. P. 1237—1244.
5. Katz B., Di Carlo J., Kim S. Mini cheetah: A platform for pushing the limits of dynamic quadruped control // Intern. Conf. on Robotics and Automation (ICRA). IEEE, 2019. P. 6295—6301.
6. Thomaszewski B. et al. Computational design of linkage-based characters // ACM Trans. on Graphics (TOG). 2014. Vol. 33, N 4. P. 1—9.
7. Abd M. A. et al. Surface Feature Recognition and Grasped Object Slip Prevention With a Liquid Metal Tactile Sensor for a Prosthetic Hand // Proc. of the 8th IEEE RAS/EMBS Intern. Conf. for Biomedical Robotics and Biomechatronics (BioRob). IEEE, 2020. P. 1174—1179.
8. Borisov I. I. et al. Study on elastic elements allocation for energy-efficient robotic cheetah leg // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2019. P. 1696—1701.
9. Metta G. et al. The iCub humanoid robot: An open-systems platform for research in cognitive development // Neural Networks. 2010. Vol. 23, N 8—9. P. 1125—1134.
10. Schmitz A. et al. Design, realization and sensorization of the dexterous iCub hand // Proc. of the 10th IEEE-RAS Intern. Conf. on Humanoid Robots. IEEE, 2010. P. 186—191.
11. Borisov I. I., Khomutov E. E., Kolyubin S. A., Stramigioli S. Computational design of reconfigurable underactuated linkages for adaptive grippers // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2021.
12. Borisov I. I. et al. Novel optimization approach to development of digit mechanism for bio-inspired prosthetic hand // Proc. of the 7th IEEE Intern. Conf. on Biomedical Robotics and Biomechatronics (Biorob). IEEE, 2018. P. 726—731.
13. Борисов И. И., Защитин Р. А., Борисова О. В., Колюбин С. А. Алгоритм структурно-параметрического синтеза механизмов адаптивных захватных устройств со звеньями переменной длины // Изв. вузов. Приборостроение. 2020. Т. 63, № 5. С. 467—475.
14. Birglen L., Laliberté T., Gosselin C. M. Underactuated Robotic Hands. Springer, 2007. Vol. 40.

Сведения об авторах

- | | |
|-----------------------------------|---|
| Евгений Эдуардович Хомутов | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: e.khomutov@itmo.ru |
| Дмитрий Викторович Иволга | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: ivolga.dv@itmo.ru |
| Иван Игоревич Борисов | — канд. техн. наук; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; научный сотрудник; E-mail: borisovii@itmo.ru |
| Никита Алексеевич Молчанов | — ПАО „Сбербанк“, лаборатория робототехники; ст. инженер-разработчик; E-mail: molchanov.n.al@sberbank.ru |
| Иван Алексеевич Максимов | — ПАО „Сбербанк“, лаборатория робототехники; ст. инженер-разработчик; E-mail: maksimov.i.a@sberbank.ru |

Поступила в редакцию 28.12.21; одобрена после рецензирования 10.01.22; принята к публикации 18.01.22.

REFERENCES

1. Liu H. et al. 2008 *IEEE/RSJ International Conference On Intelligent Robots and Systems*, IEEE, 2008, pp. 3692–3697.
2. Rothling F. et al. 2007 *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2007, pp. 2951–2956.
3. Xu Z., Todorov E. 2016 *IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 3485–3492.
4. Raiola G. et al. *IEEE Robotics and Automation Letters*, 2018, no. 2(3), pp. 1237–1244.
5. Katz B., Di Carlo J., Kim S. 2019 *International Conference on Robotics and Automation (ICRA)*, 2019, pp. 6295–6301.
6. Thomaszewski B. et al. *ACM Transactions on Graphics (TOG)*, 2014, no. 4(33), pp. 1–9.
7. Abd M.A. et al. 2020 *8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 2020, pp. 1174–1179.
8. Borisov I.I. et al. 2019 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 1696–1701.
9. Metta G. et al. *Neural Networks*, 2010, no. 8-9(23), pp. 1125–1134.
10. Schmitz A. et al. 2010 *10th IEEE-RAS International Conference on Humanoid Robots*, 2010, pp. 186–191.
11. Borisov I.I., Khomutov E.E., Kolyubin S.A., Stramigioli S. 2021 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 6117–6123.
12. Borisov I.I. et al. 2018 *7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, 2018, pp. 726–731.
13. Borisov I.I., Zashchitin R.A., Borisova O.V., Kolyubin S.A. *Journal of Instrument Engineering*, 2020, no. 5(63), pp. 467–475, DOI: 10.17586/0021-3454-2020-63-5-467-475. (in Russ.)
14. Birglen L., Laliberté T., Gosselin C.M. *Underactuated Robotic Hands*, Springer, 2007, vol. 40.

Data on authors

Evgeny E. Khomutov	—	Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; E-mail: e.khomutov@itmo.ru
Dmitry V. Ivolga	—	Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; E-mail: ivolga.dv@itmo.ru
Ivan I. Borisov	—	PhD; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Researcher; E-mail: borisovii@itmo.ru
Nikita A. Molchanov	—	Sberbank, Robotics Laboratory; Senior Engineer-Designer; E-mail: molchanov.n.al@sberbank.ru
Ivan A. Maksimov	—	Sberbank, Robotics Laboratory; Senior Engineer-Designer; E-mail: maksimov.i.a@sberbank.ru

Received 28.12.21; approved after reviewing 10.01.22; accepted for publication 18.01.22.

ПРОГРАММНО-АППАРАТНАЯ АРХИТЕКТУРА СИСТЕМЫ УПРАВЛЕНИЯ АНТРОПОМОРФНОЙ КИСТЬЮ РОБОТА

Д. В. Иволга¹, Е. Э. Хомутов¹, И. И. Борисов^{1*},
Н. А. Молчанов², И. А. Максимов², С. А. С. Колубин¹

¹Университет ИТМО, Санкт-Петербург, Россия

*borisovii@itmo.ru

²Сбербанк, Москва, Россия

Аннотация. Представлены результаты разработки и реализации программно-аппаратной архитектуры системы управления для адаптивной антропоморфной роботизированной кисти. Рассматривается возможность интеграции разработанного захватного устройства в контур управления робота iCub с сохранением функциональных возможностей и гибкости реализации алгоритмов управления. Это достигается посредством прототипирования системы управления как самостоятельного модуля, подключаемого к роботу iCub через сетевой интерфейс Ethernet. Обмен данными между захватным устройством и роботом iCub имеет высокую стабильность и производительность с частотой управления 2 кГц при задержке не более 310 мкс и джиттере ниже 50 мкс. Тестирование программно-аппаратной архитектуры системы управления продемонстрировало высокую точность управления положением ($\pm 1^\circ$) и тактильным усилием ($\pm 0,15$ Н) проксимальных фаланг пальцев.

Ключевые слова: антропоморфные захватные устройства, система управления, аппаратное обеспечение, программный интерфейс

Ссылка для цитирования: Иволга Д. В., Хомутов Е. Э., Борисов И. И., Молчанов Н. А., Максимов И. А., Колубин С. А. Программно-аппаратная архитектура системы управления антропоморфной кистью робота // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 164—173. DOI: 10.17586/0021-3454-2022-65-3-164-173.

HARDWARE AND SOFTWARE ARCHITECTURE FOR ANTHROPOMORPHIC ROBOT HAND CONTROL SYSTEM

D. V. Ivolga¹, E. E. Khomutov¹, I. I. Borisov^{1*},
N. A. Molchanov², I. A. Maksimov², S. A. Kolyubin¹

¹ITMO University, St. Petersburg, Russia

*borisovii@itmo.ru

²Sberbank, Moscow, Russia

Abstract. The results of the development and implementation of hardware and software architecture for the control system of an adaptive anthropomorphic robotic hand are presented. The research focuses on possibility of integrating the proposed robotic hand into the iCub robot's control system while preserving the functionality and flexibility of implementing control algorithms. It has been achieved by prototyping the control system as an independent module connected to the iCub robot via network interface Ethernet. The data exchange between the gripper and the iCub robot has high stability and performance with a control frequency of 2 kHz with a delay less than 310 us and a jitter below 50 us. Testing of the control system's software and hardware architecture demonstrated high accuracy in position control ($\pm 1^\circ$) and force control ($\pm 0,15$ N) for fingers' proximal phalanges.

Keywords: anthropomorphic grippers, control system, hardware, choice of element base, sensors, software interface

For citation: Ivolga D. V., Khomutov E. E., Borisov I. I., Molchanov N. A., Maksimov I. A., Kolyubin S. A. Hardware and software architecture for anthropomorphic robot hand control system. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 164—173 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-164-173.

Введение. В настоящее время актуальны исследования по воссозданию сенсорных и моторных функций человеческой руки, способной манипулировать объектами произвольной

формы. Кисть имеет не только функциональное назначение, но и социальную значимость при выражении эмоций, рукопожатии, приветствии и т.д. Более того, развитию области способствует идея о том, что восприятие — это активный процесс, в котором роботу необходимо взаимодействовать с окружающими объектами для получения более точной информации [1—3]. В подобном контексте разработка сложных и оснащенных различными датчиками роботизированных рук имеет фундаментальное значение.

Одной из наиболее современных и функциональных робототехнических платформ для научных исследований в области манипуляции и локомоции, человеко-машинного взаимодействия и т. п. является гуманоидный робот iCub, разрабатываемый консорциумом The RobotCub Project [4]. Кисть робота iCub (рисунок 1, а, б) имеет пятипальную антропоморфную конструкцию, способную совершать множество жестов, однако для этого требуются 9 электроприводов и система управления для скоординированного движения, что влечет за собой увеличение габаритов, массы и стоимости устройства. Фаланги робота приводятся в движение системой блоков и тросов. Подобное решение позволяет разместить привода в кисти и предплечье для сохранения габаритов руки, но в то же время не позволяет достичь высокой точности позиционирования и грузоподъемности из-за растяжения и риска обрыва тросов. Независимое управление проксимальными и дистальными фалангами пальцев кисти iCub обуславливает возможность ее адаптации к объектам произвольной формы, но не позволяет надежно захватывать объекты малого размера точным щипковым схватом.

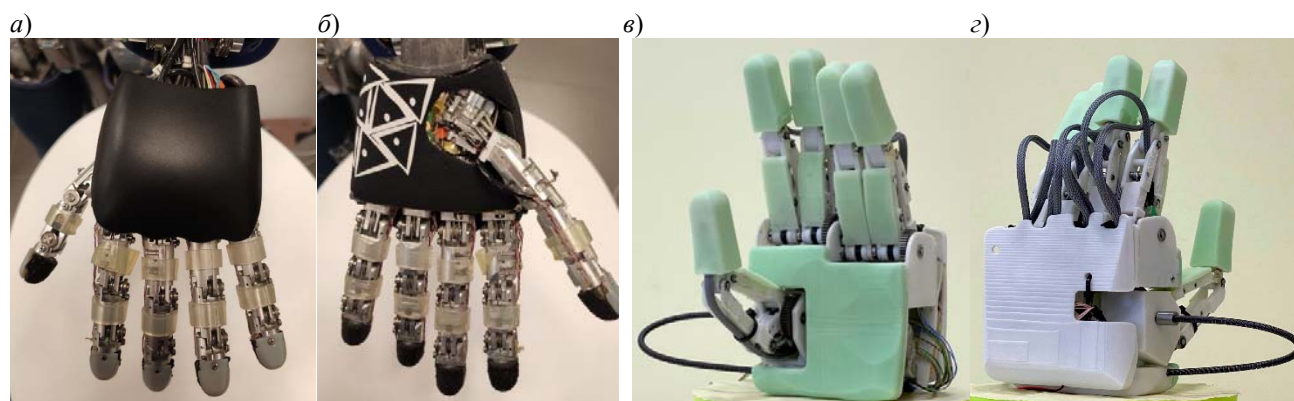


Рис. 1

Для решения указанных проблем была разработана новая конструкция пятипального адаптивного антропоморфного захватного устройства (рис. 1, в, г). Отличительная особенность кисти — синтезированные неполноприводные механизмы пальцев со звеньями *переменной длины*, благодаря которым устройство способно совершать адаптивные силовые и точные щипковые захваты объектов разных размеров и произвольной формы и массы [5—7].

Захватное устройство оснащено пятью активными пальцами, но требует только четыре двигателя для управления всеми 14 степенями свободы. Первый мотор-редуктор используется для приведения в действие указательного и среднего пальцев; второй — для приведения в действие безымянного пальца и мизинца. Механизм большого пальца оснащен двумя приводами: первый для сгибания, а второй для переориентации относительно ладони, что позволяет добиться множества жестов. Механизмы указательного, среднего, безымянного пальцев и мизинца прикреплены к общим валам. Все пальцы соединены с входным валом через вращательный гибкий элемент, что предотвращает от заклинивания при контакте с объектом. Рабочие поверхности фаланг и ладони покрыты силиконом с высоким коэффициентом трения для предотвращения выскальзывания.

К аппаратной и программной частям системы управления новым захватным устройством сформулированы следующие требования:

— возможность определения углового положения звеньев в диапазоне до 110° ;

— датчик тактильного усилия должен обладать миниатюрными размерами (не более $10 \times 15 \times 10$ мм) и измерять тактильное усилие в диапазоне от 0 до 10 Н с точностью не менее 0,2 Н;

— скорость выполнения захвата должна составлять 48,7 °/с;

— необходимый постоянный крутящий момент на входном валу пальцев — не менее 1,7 мН·м;

— диаметр мотор-редуктора не должен превышать 13 мм, а суммарная длина — 35 мм;

— возможность интеграции в программно-аппаратную архитектуру робота iCub.

Основные функциональные части захватного устройства. Новый адаптивный антропоморфный захват разработан специально для установки на iCub, поэтому изменения в архитектуре робота должны быть минимальны, чтобы обеспечить простоту интеграции. Система управления захватным устройством должна обеспечивать возможность коммуникации с бортовым компьютером робота через имеющиеся сетевые интерфейсы. На основе работы [4] и конструкторской документации робототехнической платформы [8, 9] были определены элементы системы управления новым адаптивным схватом, структурная схема которой представлена на рис. 2.

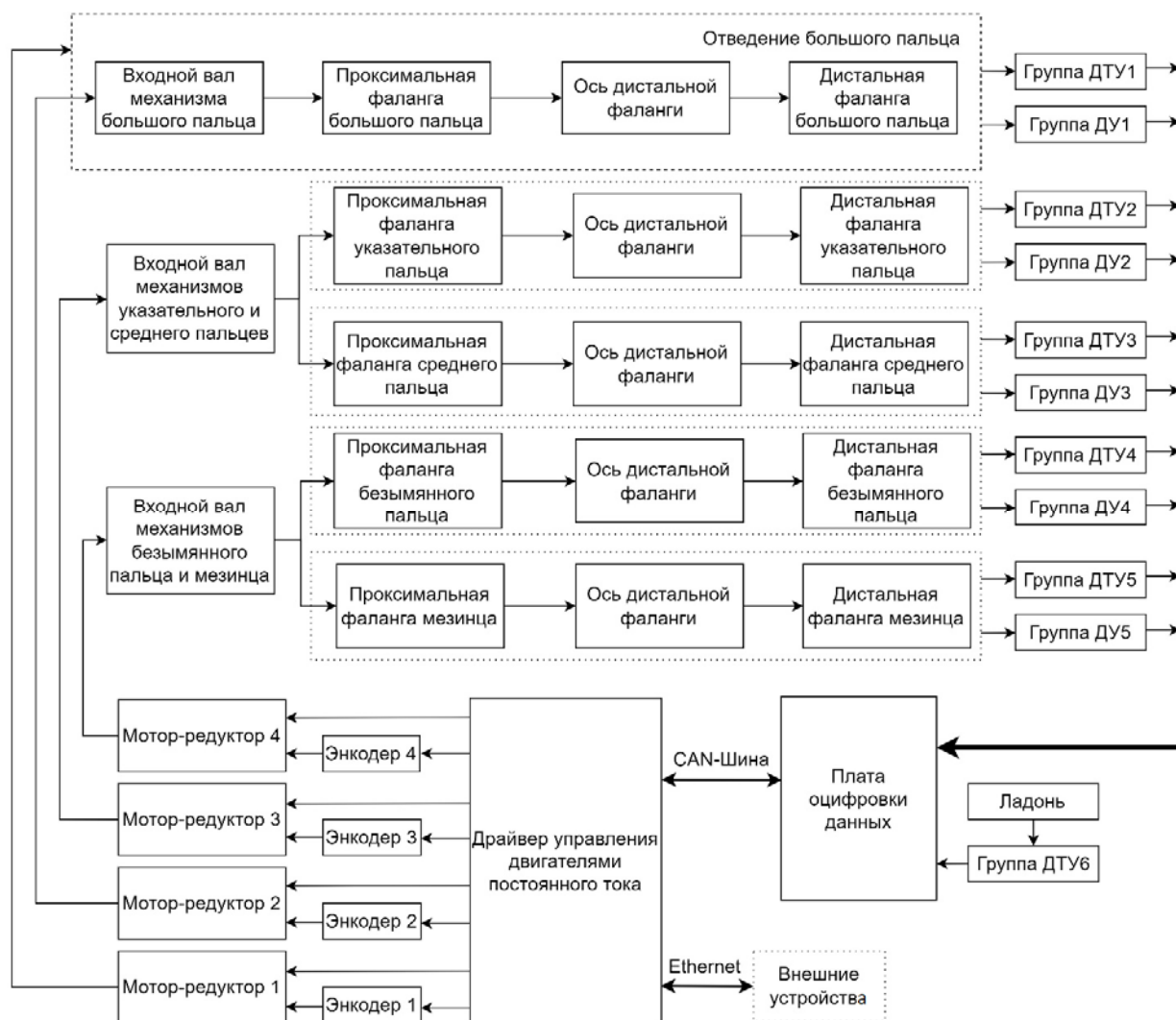


Рис. 2

Основным элементом системы является плата управления мотор-редукторами пальцев захвата. Информация с датчиков кисти передается на плату управления мотор-редукторами через CAN-шину. Драйвер мотор-редукторов соединен с бортовым компьютером робота че-

рез Ethernet. Вращение двигателя постоянного тока передается через зубчатые механизмы на входные звенья рычажных механизмов пальцев. Мотор-редуктор 1 приводит в движение механизм отведения и сведения большого пальца, мотор-редуктор 2 — проксимальную и дистальную фаланги большого пальца, мотор-редуктор 3 — указательный и средний пальцы, мотор-редуктор 4 — безымянный палец и мизинец. Мотор-редукторы 2, 3 и 4 приводят в движение входные звенья пальцев, далее механизмы замкнутой кинематики со звеньями переменной длины преобразовывают движение входных звеньев, жестко соединенных с валами двигателя, в выходное движение проксимальных и дистальных фаланг. Для определения положения фаланг используются датчики углового положения (ДУ), которые установлены на проксимальных и дистальных фалангах. Под рабочими поверхностями дистальных и проксимальных фаланг, а также на ладони установлены датчики тактильного усилия (ДТУ) для определения силы взаимодействия с объектом.

Система управления может работать в двух режимах: а) управление по положению на основе данных о состоянии проксимальных фаланг; б) управление по силе взаимодействия на основе данных о состоянии проксимальных и дистальных фаланг. При контакте с поверхностью объекта манипулирования скорость движения пальцев уменьшается пропорционально развиваемой силе взаимодействия и производится фиксация предмета с необходимым усилием. Если при управлении по силе взаимодействия пальцы не соприкасаются с объектом, то механизмы останавливаются в конечном положении [9].

Датчики углового положения фаланг. Всего в кисти необходимо расположить 11 датчиков углового положения, измеряющих углы между проксимальной фалангой и ладонью, дистальной и проксимальной фалангами, а также угол отведения большого пальца относительно ладони. При этом положения фаланг должны измеряться при максимальном угле раскрытия в 110° .

Для данных целей было предложено использовать миниатюрные датчики Холла SS495A, преимущество которых является бесконтактное измерение углового положения; кроме того, такие датчики не усложняют конструкцию, не вносят дополнительные силы сопротивления при движении фаланг и не приводят к существенному увеличению веса.

Датчик оснащен аналоговым выходом по напряжению, значение которого пропорционально магнитной индукции, наводимой внешним магнитом. Значение магнитной индукции зависит от геометрических параметров и расположения магнита. Для определения этих параметров было проведено моделирование в программном пакете EMWorks для SolidWorks [10]. Для новой конструкции адаптивного схвата выбран кольцевой диаметрально поляризованный магнит. Материал магнита N4212, внешний диаметр — 8 мм, внутренний диаметр — 3 мм, высота — 2 мм.

На основе модели магнита и силы магнитной индукции была выбрана область, которая обеспечивает необходимую для датчика SS495A магнитную индукцию в пределах от $-0,067$ до $+0,067$ Тл. Как наиболее подходящая и конструктивно реализуемая, выбрана область прямого прилегания датчика к краю магнита на расстоянии 4,2 мм. Схема расположения датчика и магнита в фаланге представлена на рис. 3, а.

Датчики тактильного усилия. При адаптивном манипулировании объектами произвольной формы местоположение и сила контакта не могут быть заранее полностью предсказаны, поэтому для обеспечения надежного захвата и удержания объекта захватное устройство оснащено датчиками тактильного взаимодействия. Количество и наилучшее расположение точек измерений было определено по результатам имитационного моделирования адаптивного схвата в программной среде MatLab Simscape Multibody. На рабочей поверхности захвата расположено 30 точек измерения тактильного усилия: по 2 на каждой проксимальной фаланге, 3 на дистальной фаланге (2 на подушечке пальца, 1 на кончике) и 5 точек контакта на ладони. Соответственно, датчик должен иметь малые габариты для встраивания в фаланги

пальцев и ладонь. Схват спроектирован для взаимодействия с предметами, окружающими человека в быту, масса подавляющего большинства которых составляет до 1 кг, т.е. для их захвата и удержания необходимо развивать силу не менее 10 Н.

Согласно данным требованиям был выбран тензорезистивный датчик силы HSFPAR003A, имеющий широкое распространение, точность измерения до 0,01 Н, большой цикл повторения до 10^6 нажатий.

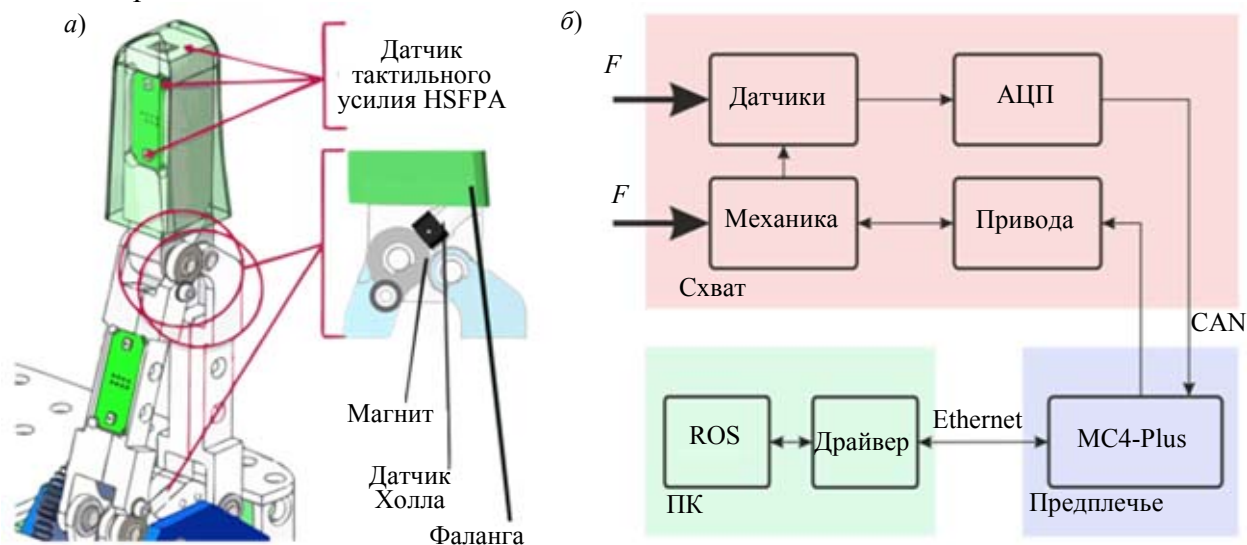


Рис. 3

Датчик тактильного усилия (см. рис. 3, а) устанавливается на печатной плате, жестко закрепленной на фаланге. Поверх датчика устанавливается железная пластинка толщиной 0,2 мм, через которую передается усилие на чувствительный элемент датчика. Для расширения диапазона измерений был применен метод, основанный на покрытии тензорезистивного моста слоем упругого материала, например силикона [11]. Необходимая толщина слоя для конкретного материала определяется с учетом его упругих свойств либо экспериментальным путем в ходе калибровочной процедуры. Для конструкции был выбран слой покрытия датчика в 3 мм, так как при этом сохраняются миниатюрные габариты сенсора, а диапазон измерений увеличивается до 12 Н.

Плата управления приводами. В оригинальной конструкции iCub для управления приводами применяется драйвер MC4-Plus на базе микроконтроллера STM32F407VGT6. Данная плата поддерживает подключение до четырех коллекторных двигателей с напряжением питания 12 В и потребляемым током 1 А (пиковая токовая нагрузка 2 А). Для управления приводами используются абсолютные и инкрементальные энкодеры, а также на плате расположен токовый резистивный сенсор. Интерфейс взаимодействия драйвера с бортовым компьютером робота — Ethernet и CAN-шина. Плата имеет сравнительно малые габариты 80×30×11 мм и массу 35 г. Данное решение — наиболее подходящее для системы управления адаптивным захватом, так как плата является оригинальным компонентом от iCub, а также имеется возможность подключения четырех моторов и изменения прошивки микроконтроллера [12].

Приводы. Выбор мотор-редуктора для данной конструкции схвата имеет ряд ограничений по мощностным и габаритным характеристикам, которые обусловлены конструкцией адаптивного схвата [5]. Захват объекта осуществляется с усилием 10 Н, при этом скорость движения должна быть сопоставима с нормальной скоростью движения пальцев человека, а именно 48,7 °/с [8]. Исходя из требований и параметров оригинального драйвера MC4-Plus, был выбран коллекторный мотор-редуктор FAULHABER 1224 в совокупности с редуктором FAULHABER 12/4 64:1. Данная связка мотор-редуктора дополняется редуктором и зубчатой передачей для более компактного размещения внутри кисти и увеличения крутящего момента до 1,7 мН·м.

Плата оцифровки показаний датчиков. Так как в конструкции новой антропоморфной кисти используются 11 аналоговых датчиков углового положения и 30 аналоговых датчиков тактильного усилия с дифференциальным выходом, необходимое количество каналов для оцифровки данных составляет 71. Для обработки сигналов сенсоров руки и передачи показаний в плату управления мотор-редукторами была спроектирована оригинальная плата с малыми размерами (45×50 мм) для размещения внутри ладони схвата [13].

Датчик тактильного усилия обладает малой чувствительностью (3,7 мВ/В/Н), следовательно, для обработки сигнала необходимо его усиление и оцифровка с помощью аналого-цифрового преобразователя (АЦП). Для упрощения разводки печатной платы и сохранения малых габаритов выбран многоканальный быстродействующий АЦП ADS1261 с программируемым коэффициентом усиления. АЦП имеет цифровой интерфейс взаимодействия SPI с разрешением в 24 бита. Выбранная микросхема позволит обрабатывать группу сигналов сенсоров с частотой дискретизации 3 кГц.

Датчик углового положения обладает внутренним усилителем и пассивным фильтром, следовательно, для обработки сигнала достаточно любого современного микроконтроллера со встроенным АЦП.

Микроконтроллер STM32F415RGT6, выбранный для управления АЦП и коммуникации с платой управления двигателями, имеет высокую частоту работы, малые размеры и набор необходимых аппаратных средств. Подключение к модулю оцифровки сигнала с датчиков осуществляется по CAN-шине.

Архитектура программной части системы управления. Общий вид системы управления антропоморфным адаптивным схватом представлен на рис. 3, б. Главным компонентом системы является плата управления двигателями MC4-Plus. Соединение MC4-Plus с ПК или бортовым компьютером робота производится посредством Ethernet. Коммуникация с платой оцифровки измерений сенсоров осуществляется через CAN. Драйвер управления мотор-редукторами является клиентом, а ПК или бортовой компьютер с установленным программным обеспечением — сервером [4, 14].

Взаимодействие iCub со схватом осуществляется через систему ROS (Robot Operating System), которая организует канал приема-передачи информации и команд управления на драйвер схвата. Пользователь посредством программного интерфейса имеет возможность выбора режима работы схвата, задания целевого положения и силы удержания, независимого управления вращением мотор-редукторов, настройки параметров управления рукой, оцифровки и фильтрации сенсоров, а также возможность получения контрольно-измерительной информации о конфигурации руки и силовом контакте.

В режиме позиционирования проксимальных фаланг контур управления состоит из двухкаскадного регулятора: по положению и по току. На рис. 4 представлена схема контура управления для мотор-редуктора, который приводит в движение указательный и средний пальцы. Применение такого регулятора позволяет получить привод с постоянным крутящим моментом и переменным положением [15]:

$$\left. \begin{aligned} e_1 &= p^* - p; \\ u_1 &= k_{11}e_1 + k_{12} \frac{de_1}{dt}; \\ e_2 &= u_1 - i; \\ u_2 &= k_{21}e_2 + k_{22} \frac{de_2}{dt} + k_{23} \int e_2 dt, \end{aligned} \right\} \quad (1)$$

где e_1 — ошибка по положению, e_2 — ошибка по току, u_1 — сигнал управления по току, u_2 — сигнал управления по напряжению, k_i — коэффициенты регуляторов, p^* и p — требуемое и реальное сочленения соответственно, i — реальное значение силы тока.

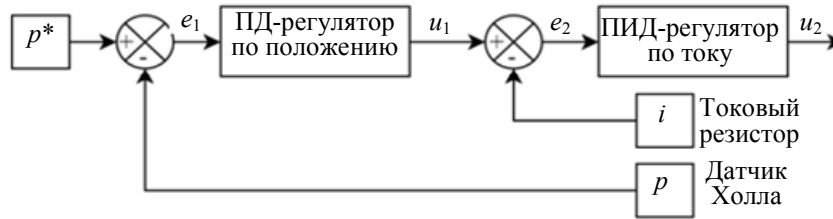


Рис. 4

Работа контура управления положением для проксимальной фаланги указательного пальца продемонстрирована на рис. 5. Целевое положение (q) составляет 40° . График показывает точность работы системы регулирования и среднюю скорость движения, равную $49^\circ/\text{с}$.

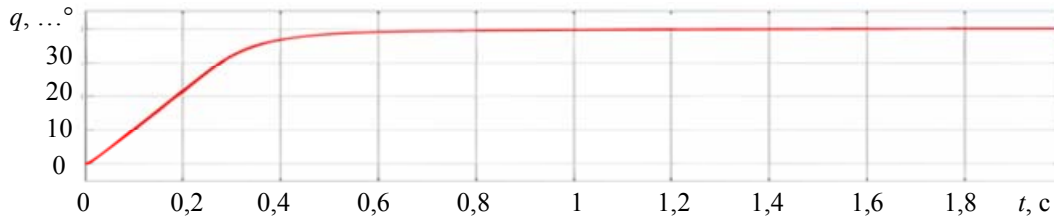


Рис. 5

В режиме силового захвата контур управления тактильным усилием на дистальной фаланге (рис. 6) состоит из двух частей: регулятора скорости и двухкаскадного регулятора по усилию и току:

$$\left. \begin{aligned} e_3 &= w^* + w; \\ u_3 &= k_{31}e_3 + k_{32} \frac{de_3}{dt}, \text{ если } F = 0; \end{aligned} \right\} \quad (2)$$

$$\left. \begin{aligned} e_4 &= F^* + F; \\ u_4 &= k_{41}e_4 + k_{42} \frac{de_4}{dt} + k_{43} \int e_4 dt; \\ e_5 &= u_4 + i; \\ u_5 &= k_{51}e_5 + k_{52} \frac{de_5}{dt} + k_{53} \int e_5 dt, \text{ если } F > 0, \end{aligned} \right\} \quad (3)$$

где w^* и w — требуемая и реальная скорость вращения мотор-редуктора соответственно, e_3 — ошибка по скорости, e_4 — ошибка по силе, e_5 — ошибка по току, u_3 и u_5 — сигнал управления напряжением на мотор-редукторе, u_4 — сигнал управления по току, F^* — требуемая сила контактного взаимодействия дистальной фаланги, F — текущая сила взаимодействия тактильного датчика силы и дистальной фаланги.

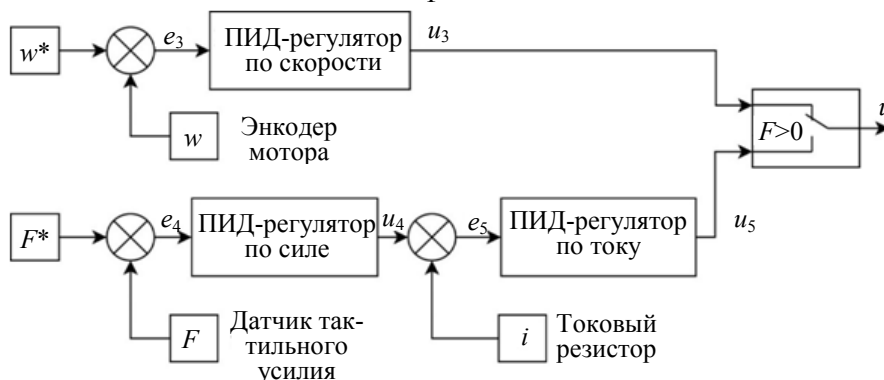


Рис. 6

Тактильное усилие для пальцев, приводимых в движение одним мотор-редуктором, рассчитывается как сумма усилий, приложенных к дистальным фалангам:

$$F = F' + F'',$$

где F' и F'' — усилие приложенное к указательному и среднему пальцам соответственно.

Первый регулятор (2) задействован во время сгибания пальца до соприкосновения с объектом. После соприкосновения силовой захват выполняется с помощью двухкаскадного регулятора (3). Если в процессе движения захват не соприкасается с объектом, то происходит его остановка в крайних положениях [16]. Работа контура управления при захвате статического объекта продемонстрирована на рис. 7. Целевое усилие для дистальных фаланг пальцев соответствует следующим значениям: большой палец — 3,1 Н, указательный и средний пальцы — 1,5 Н, безымянный палец и мизинец — 2 Н.

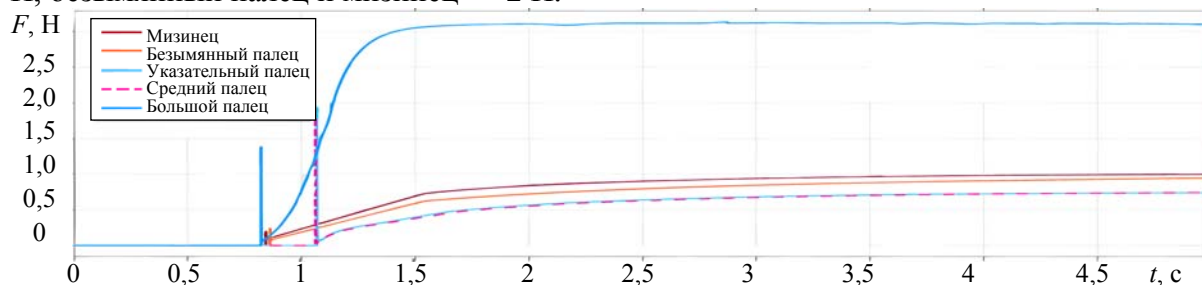


Рис. 7

Заключение. Представлена разработанная программно-аппаратная архитектура системы управления новым адаптивным антропоморфным схватом для iCub. Приведено описание аппаратной части системы управления схватом, в состав которой входят: плата управления двигателями MC4-Plus, мотор-редукторы FAULHABER, датчик углового положения SS495A в сочетании с кольцевым магнитом, датчик тактильного усилия HSFAR003A с расширенным диапазоном измерения до 1,2 кг и чувствительностью 3,7 мВ/В/Н, разработанная плата оцифровки сигналов с сенсоров кисти.

Испытание системы проводилось на прототипе адаптивного схвата (рис. 8). Результаты лабораторных испытаний показали работоспособность применяемых решений и соблюдение сформулированных технических требований. Тестирование системы управления продемонстрировало высокую точность в $\pm 1^\circ$ в режиме позиционирования проксимальных фаланг пальцев и удержание предметов с тактильным усилием $\pm 0,15$ Н. Интерфейс имеет стабильность и высокую производительность с частотой управления 2 кГц при задержке менее 310 мкс и джиттере ниже 50 мкс.

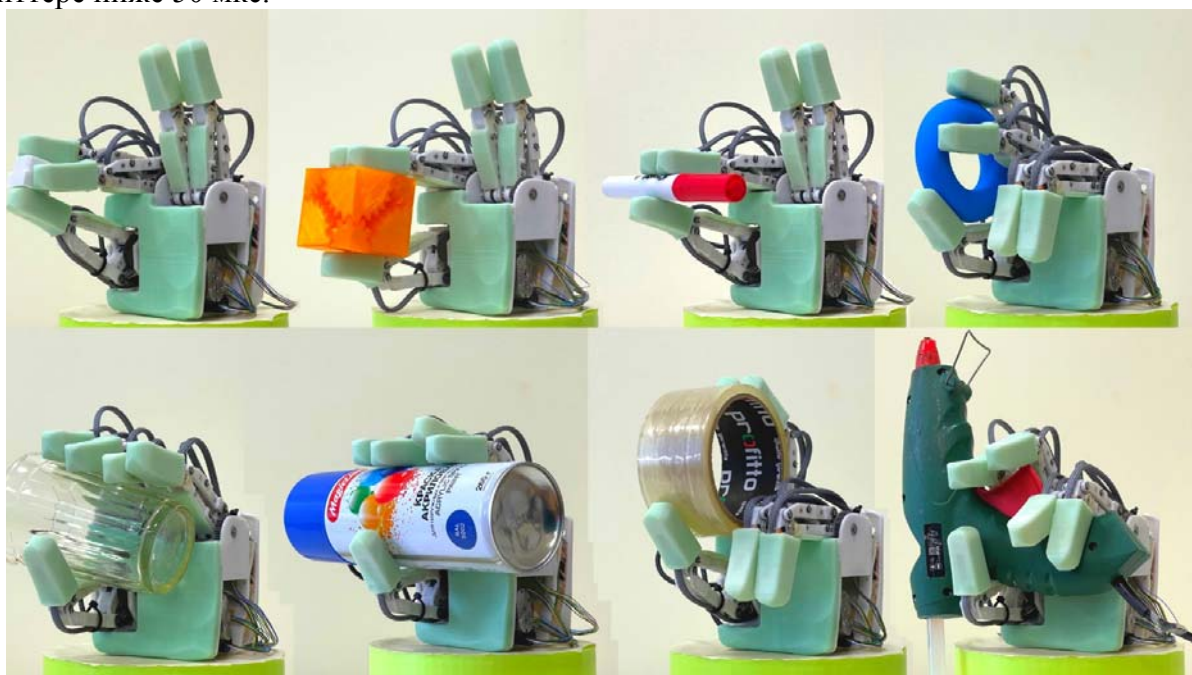


Рис. 8

Дальнейшие исследования будут сосредоточены на интеграции схвата в робототехническую платформу iCub и проведении натурных испытаний по захвату объектов произвольной формы.

СПИСОК ЛИТЕРАТУРЫ

1. Piazza C. et al. A Century of Robotic Hands // Annual Review of Control, Robotics, and Autonomous Systems. 2019. Vol. 2. P. 1—32. DOI: 10.1146/annurev-control-060117-105003.
2. Simpkins C. Mechanisms and Machine Science: Grasping in Robotic System // Robotics & Automation Magazine, IEEE. 2013. Vol. 20. P. 94—94. DOI: 10.1109/MRA.2012.2236252.
3. Metta G., Fitzpatrick P. Better vision through manipulation // Adaptive Behavior. 2003. Vol. 11, N 2. P. 109—128. DOI: 10.1177/10597123030112004.
4. Parmiggiani A., Maggiali M., Natale L. et al. The Design of the iCub Humanoid Robot // Intern. Journal of Humanoid Robotics. 2012. Vol. 9, N 4. DOI: 10.1142/S0219843612500272.
5. Borisov I. I., Khomutov E. E., Kolyubin S. A., Stramigioli S. Computational design of reconfigurable underactuated linkages for adaptive grippers // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2021.
6. Müller V. C., Hoffmann M. What is morphological computation? On how the body contributes to cognition and control // Artificial Life. 2017. Vol. 23, N 1. P. 1—24.
7. Борисов И. И., Защитин Р. А., Борисова О. В., Колюбин С. А. Алгоритм структурно-параметрического синтеза механизмов адаптивных захватных устройств со звеньями переменной длины // Изв. вузов. Приборостроение. 2020. Т. 63, № 5. С. 467—475. DOI: 10.17586/0021-3454-2020-63-5-467-475.
8. Jamali N. et al. A new design of a fingertip for the iCub hand // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2015. DOI: 10.1109/IROS.2015.7353747.
9. Schmitz A. et al. Design, realization and sensorization of the dexterous iCub hand // Proc. of the 10th IEEE-RAS Intern. Conf. on Humanoid Robots. IEEE, 2011. P. 186—191. DOI: 10.1109/ICHR.2010.5686825.
10. Sönmez Ö. Numerical Analysis of an Electromagnetic Plunger // European Journal of Science and Technology. 2020. P. 170—175. DOI: 10.31590/ejosat.803129.
11. Tenzer Y., Jentoft L., Howe R. The feel of mems barometers: Inexpensive and easily customized tactile array sensors // Robotics Automation Magazine, IEEE. 2014. Vol. 21. P. 89—95. DOI: 10.1109/MRA.2014.2310152.
12. Casalino G., Giorgi F. et al. Embedded FPGA-based control of a multifingered robotic hand // IEEE Intern. Conf. on Robotics and Automation. 2003. N 3. DOI: 10.1109/robot.2003.1242014.
13. Zeng J., Yi P. et al. MESO-ADC: The ADC design using MESO device // Microelectronics Journal. 2021. P. 116. DOI: 10.1016/j.mejo.2021.105235.
14. Lee Y. K. Design of exoskeleton robotic hand/arm system for upper limbs rehabilitation considering mobility and portability // Proc. of the 11th Intern. Conf. on Ubiquitous Robots and Ambient Intelligence (URAI). 2014. DOI: 10.1109/URAI.2014.7057385.
15. Liu H., Wu K. et al. A dexterous humanoid five-fingered robotic hand // Proc. of the 17th IEEE Intern. Symp. on Robot and Human Interactive Communication. 2018. DOI: 10.1109/roman.2008.4600694.
16. Sartori E., Fiorini P., Muradore R. Cutaneous feedback in teleoperated robotic hands // 42nd Annual Conf. of the IEEE Industrial Electronics Society. IECON 2016. DOI: 10.1109/iecon.2016.7792990.

Сведения об авторах

- | | |
|-----------------------------------|---|
| Дмитрий Викторович Иволга | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: ivolga.dv@itmo.ru |
| Евгений Эдуардович Хомутов | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: e.khomutov@itmo.ru |
| Иван Игоревич Борисов | — канд. техн. наук; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; научный сотрудник; E-mail: borisovii@itmo.ru |

- Никита Алексеевич Молчанов** — ПАО „Сбербанк“, лаборатория робототехники; ст. инженер-разработчик; E-mail: molchanov.n.al@sberbank.ru
- Иван Алексеевич Максимов** — ПАО „Сбербанк“, лаборатория робототехники; ст. инженер-разработчик; E-mail: maksimov.i.a@sberbank.ru
- Сергей Алексеевич Колюбин** — д-р техн. наук, доцент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; вед. научный сотрудник; E-mail: s.kolyubin@itmo.ru

Поступила в редакцию 28.12.21; одобрена после рецензирования 12.01.22; принята к публикации 18.01.22.

REFERENCES

1. Piazza C. et al. *A Century of Robotic Hands. Annual Review of Control, Robotics, and Autonomous Systems*, 2019, vol. 2, pp. 1–32, DOI: 2. 1–32. 10.1146/annurev-control-060117-105003.
2. Simpkins C. *Robotics & Automation Magazine*, 2013, vol. 20, pp. 94–94, DOI: 10.1109/MRA.2012.2236252.
3. Metta G. and Fitzpatrick P. *Adaptive Behavior*, 2003, no. 2(11), pp. 109–128, DOI: 10.1177/10597123030112004.
4. Parmiggiani A. et al. *International Journal of Humanoid Robotics*, 2012, no. 4(9), pp. 9, DOI: 10.1142/S0219843612500272.
5. Borisov I.I., Khomutov E.E., Kolyubin S.A., and Stramigioli S. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021, pp. 6117–6123.
6. Müller V.C., Hoffmann M. *Artificial Life*, 2017, no. 1(23), pp. 1–24.
7. Borisov I.I., Zashchitin R.A., Borisova O.V., Kolyubin S.A. *Journal of Instrument Engineering*, 2020, no. 5(63), pp. 467–475, DOI: 10.17586/0021-3454-2020-63-5-467-475. (in Russ.)
8. Jamali N. et al. *A new design of a fingertip for the iCub hand*, 2015, DOI:10.1109/IROS.2015.7353747.
9. Schmitz A. et al. *10th IEEE-RAS International Conference of Humanoid Robots*, 2011, pp. 186–191, DOI: 10.1109/ICHR.2010.5686825.
10. Sönmez Ö. *European Journal of Science and Technology*, 2020, pp. 170–175, DOI: 10.31590/ejosat.803129.
11. Tenzer Y., Jentoft L., and Howe R. *Robotics Automation Magazine*, IEEE, 2014, vol. 21, pp. 89–95, DOI: 10.1109/MRA.2014.2310152.
12. Casalino G., Giorgi F. et al. *IEEE International Conference on Robotics and Automation*, 2003, no. 03, DOI:10.1109/robot.2003.1242014.
13. Zeng J., Yi P. et al. *Microelectronics Journal*, 2021, pp. 116, DOI: 10.1016/j.mejo.2021.105235.
14. Lee Y. et al. *11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. 2014. DOI: 10.1109/URAI.2014.7057385.
15. Liu H., Wu K. et al. *17th IEEE International Symposium on Robot and Human Interactive Communication*, 2018, DOI:10.1109/roman.2008.4600694.
16. Sartori E., Fiorini P., Muradore R. *42nd Annual Conference of the IEEE Industrial Electronics Society, IECON*, 2016, DOI:10.1109/iecon.2016.7792990.

Data on authors

- Dmitry V. Ivolga** — Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; E-mail: ivolga.dv@itmo.ru
- Evgeny E. Khomutov** — Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; E-mail: e.khomutov@itmo.ru
- Ivan I. Borisov** — PhD; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Researcher; E-mail: borisovii@itmo.ru
- Nikita A. Molchanov** — Sberbank, Robotics Laboratory; Senior Engineer-Designer; E-mail: molchanov.n.al@sberbank.ru
- Ivan A. Maksimov** — Sberbank, Robotics Laboratory; Senior Engineer-Designer; E-mail: maksimov.i.a@sberbank.ru
- Sergey A. Kolyubin** — Dr. Sci., Associate Professor; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Leading Researcher; E-mail: s.kolyubin@itmo.ru

Received 28.12.21; approved after reviewing 12.01.22; accepted for publication 18.01.22.

**ПРОЕКТИРОВАНИЕ НЕПОЛНОПРИВОДНОГО ПРЫГАЮЩЕГО РОБОТА
С ГИБКИМИ СОЧЛЕНЕНИЯМИ**

К. В. НАСОНОВ, И. И. БОРИСОВ*

Университет ИТМО, Санкт-Петербург, Россия

*borisovii@itmo.ru

Аннотация. Представлены результаты исследования по проектированию и изготовлению прототипа энергоэффективного прыгающего робота с гибкими сочленениями с использованием принципов морфологического расчета. Гибкие элементы позволяют роботам адаптивно подстраиваться к окружающей среде при контактном взаимодействии, перенаправляя энергию взаимодействия с пластической деформации твердых тел в упругую деформацию эластичных тел, что способствует рекуперации энергии в системе. В отличие от традиционных низших и высших кинематических пар, гибкие сочленения обеспечивают перемещения звеньев только в ограниченном диапазоне в пределах зоны упругой деформации. Решена задача проектирования эластичных полимерных перекрестных сочленений на примере плоского механизма ноги неполноприводного прыгающего робота замкнутой кинематики, приводимого в движение от единственного серводвигателя с присоединенными последовательно эластичными элементами. При синтезе такого робота необходимо оптимизировать не только кинематические параметры рычажного механизма, но и топологию и эластостатические параметры самих эластичных сочленений.

Ключевые слова: биомиметика, гибкие сочленения, податливые роботы, численная оптимизация, морфологический расчет

Ссылка для цитирования: Насонов К. В., Борисов И. И. Проектирование неполноприводного прыгающего робота с гибкими сочленениями // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 174—184. DOI: 10.17586/0021-3454-2022-65-3-174-184.

DESIGN OF AN UNDERACTUATED JUMPING ROBOT WITH FLEXIBLE JOINTS

K. V. Nasonov, I. I. Borisov *

ITMO University, St. Petersburg, Russia

*borisovii@itmo.ru

Abstract. Results of a study on the design and manufacture of a prototype of an energy-efficient jumping robot with flexible joints using the principles of morphological calculation are presented. Flexible elements allow robots adaptation to the environment during contact interaction, redirecting the interaction energy from the plastic deformation of solids to the elastic deformation of elastic bodies, which contributes to energy recovery in the system. Unlike traditional lower and higher kinematic pairs, flexible joints provide movement of links only in a limited range within the elastic deformation zone. The problem of designing elastic polymer cross joints is solved by the example of a flat leg mechanism of an incomplete jumping robot of closed kinematics, driven by a single servo motor with elastic elements connected in series. When synthesizing such a robot, it is necessary to optimize not only the kinematic parameters of the lever mechanism, but also the topology and elastic-static parameters of the elastic joints themselves.

Keywords: biomimetics, flexible joints, soft robots, numerical optimization, morphological computation

For citation: Nasonov K. V., Borisov I. I. Design of an underactuated jumping robot with flexible joints. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 174—184 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-174-184.

Введение. Шагающие и галопирующие роботы сформировались и приобрели следующие характерные черты: двигатели с высокой плотностью момента [1], импедансные регуляторы [2] с планировщиками траекторий движения [3], топология конструкции в виде абсолютно жесткого тела и ног открытой кинематики для обеспечения динамической локомоции [4—7]. При

этом актуальным остается вопрос обеспечения энергоэффективности движения роботов, так как она напрямую влияет на запас хода и длительность работы [8]. Несмотря на универсальность и эффективность упомянутых роботов [4—7] они тратят большое количество энергии на движение. Помимо высоких требований к энергетическому запасу, для функционирования таких роботов необходимы высокоточные датчики, специально спроектированные электродвигатели с высокой плотностью момента, высокопроизводительные контроллеры.

Для повышения энергоэффективности применяются методы бионики, когда принципы движения живых систем переносятся в искусственные системы. В частности, такой вид движения, как энергоэффективный вращательный галоп, присущий псовым и кошачьим, является источником вдохновения для создания робототехнических систем [9]. Эластичность сухожилий и мышц животных позволяет поглощать удары от контакта с поверхностью и рекуперировать часть энергии при ходьбе, беге, прыжках и прочих циклических движениях [10, 11]. Использование приводов и механизмов переменной жесткости в конструкциях ног нашло отражение при проектировании роботов [12, 13].

Одним из параметров для оценки энергоэффективности является *стоимость перемещения* (Cost of Transport), которая рассчитывается одним из двух способов [10]:

$$\text{COT} = \frac{E}{mgd} = \frac{P}{mgv}, \quad (1)$$

где E — энергия, необходимая для перемещения груза массой m на расстояние d , g — ускорение свободного падения; P — мощность, необходимая для перемещения груза массой m со скоростью v ; стоимость перемещения является безразмерной величиной.

Механизмы с эластичными сочленениями используются, как правило, в целях повышения точности позиционирования в широком спектре областей: прецизионные датчики, микрохирургия, аэрокосмические модули, разворачиваемые механизмы, оптическое прецизионное оборудование, потребительские товары. Прецизионность достигается за счет отсутствия люфта, низкого упругого гистерезиса, отсутствия механического износа, неподверженности влиянию засоренности. Однако аналогичный подход может быть использован и в целях повышения адаптивности, безопасности и энергоэффективности робототехнических устройств.

Настоящая статья посвящена проектированию прыгающего на месте робота с *гибкими* сочленениями, разработанного с использованием принципов морфологического проектирования.

Преодоленная дистанция и скорость перемещения при прыжке на месте равны нулю, поэтому стоимость перемещения (1) как метрика энергоэффективности неприменима. Проектирование осуществлялось в предположении, что минимизация количества приводов, управляющего момента, массы робота и трения способствует снижению энергопотребления. В качестве численного показателя энергопотребления использован числитель выражения (1), т. е. мощность P и энергия E . В результате был спроектирован, изготовлен и протестирован прототип с гибкими полимерными сочленениями, который динамически взаимодействует с твердой поверхностью.

Эта работа является частью проекта по разработке галопирующего робота с гибкой спиной, в котором динамика и особенности механической конструкции используются для максимизации метрик производительности, таких как энергоэффективность, генерируемый момент, податливость, соотношение грузоподъемности к собственному весу и пр., при этом максимально бережно расходуются бортовые запасы энергии. Например, в предыдущей работе авторов [8] были решены задачи по распределению массы между передней и задней частями тела галопирующего робота и по оптимизации параметров модели пространственной пассивной пружины гибкой спины робота для повторения, в частности, характера динамики гепарда, использующего мышцы и сухожилия спины при вращательном галопе.

Синтез и оптимизация рычажного механизма ноги робота. Для создания и анализа шагающих, прыгающих и галопирующих роботов используется модель перевернутого маятника

(Spring Loaded Inverted Pendulum — SLIP), согласно которой основная часть массы должна быть сосредоточена в теле робота, а инерция ноги должна быть минимальна [15]. В ходе исследований был апробирован метод синтеза конструкции, в которой все вращательные кинематические пары заменены на гибкие сочленения, что позволило существенно снизить массу ноги и „поднять“ центр масс робота к корпусу. Цель синтеза — определение структуры и геометрических параметров рычажного механизма совместно с топологией и эластостатическими параметрами сочленений.

Для прыжка робота на месте точка его контакта F с землей, расположенная строго над центром масс робота, должна перемещаться вдоль вертикальной прямой. Данную траекторию можно обеспечить рычажными механизмами разной структуры, например открытой кинематической цепью, однако в этом случае необходимо использовать столько приводов, сколько степеней свободы имеет механизм. Увеличение числа двигателей приводит к увеличению массы конструкции, усложнению системы управления и повышению энергопотребления. Поэтому была решена задача синтеза рычажного механизма замкнутой кинематики с одной степенью свободы, который следует требуемой траектории.

На рис. 1 представлены основные этапы синтеза рычажного механизма замкнутой кинематики ноги прыгающего робота. На первом этапе был задан двухзвенный механизм открытой кинематики с контуром ABF , который может совершать требуемое движение вдоль вертикальной прямой (рис. 1, а). Для приведения в движение механизма открытой кинематики с двумя звеньями необходимы два двигателя, расположенные в сочленениях A и B . Для обеспечения такого простого движения, как прыжок на месте, использование двух двигателей не рационально. Необходимо редуцировать количество двигателей до одного. Для получения механизма только с одним приводом была проведена оптимизация структуры в соответствии с методом, описанным в [16].

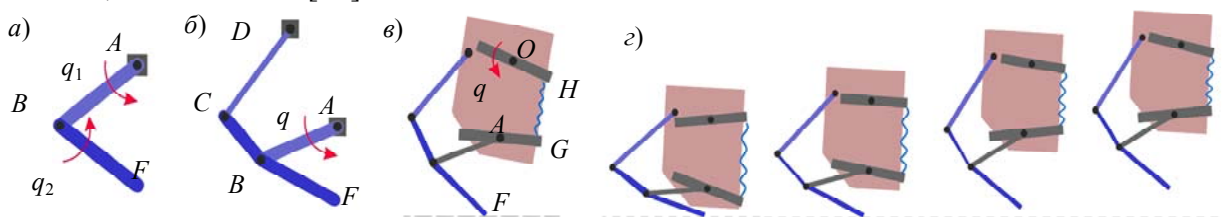


Рис. 1

Суть метода основана на наблюдении, что если расстояние между двумя компонентами кинематической цепи не меняется в течение всего цикла движения, то эти компоненты могут быть соединены через новое абсолютно твердое звено. Введение звена накладывает дополнительные голономные связи и редуцирует степень свободы механизма. Изложенный принцип обобщается в следующее правило: поддержание движения по заданной траектории при удалении актуаторов достигается путем присоединения звеньев.

Кинематическая цепь механизма представляет собой набор звеньев l_i , взаимосвязанных голономными связями. При плоском движении каждое звено l_i обладает тремя степенями свободы. Координаты $p_i^w \in \mathbb{R}^2$ физической точки p_i , принадлежащей звену l_i , в мировой системе координат Ψ_w определяются согласно следующему выражению:

$$p_i^w = R_l^w p_i^l + o_l^w,$$

где $R_l^w \in \text{SO}(2)$ — матрица ориентации системы координат Ψ_l относительно Ψ_w ; $p_i^l \in \mathbb{R}^2$ — координаты физической точки p_i в системе координат Ψ_l ; $o_l^w \in \mathbb{R}^2$ — точка на многообразии $\mathcal{M}$, описывающем положение системы координат Ψ_l относительно Ψ_w .

В общем случае наложение ограничения в виде фиксации относительного расстояния между точками p_i и p_j , принадлежащими звеньям l_i и l_j , приводит к отклонению движения от изначальной траектории механизма. В рассматриваемом случае целью оптимизации является поиск точек крепления присоединяемого звена l_{CD} к точке p_C звена l_{CBF} и точке p_D звена l_D (рис. 1, б), при которых наложение ограничения приводит к минимальному изменению изначальной траектории механизма на всем цикле движения, в частности, к минимальному изменению расстояния между точками $|p_C^w(t_i) - p_D^w(t_i)|$, выраженному в мировой системе координат Ψ_w . Среднеквадратическое изменение квадрата расстояния d_{CD}^2 между точками p_C и p_D вдоль всей траектории определяется в соответствии с выражением

$$d_{CD}^2 = \frac{1}{n_s} \sum_i^{n_s} (p_C^w(t_i) - p_D^w(t_i))^2,$$

где n_s — количество элементов в массиве данных, описывающих движение.

Соответственно, оптимизация направлена на минимизацию дисперсии квадрата расстояния между точками p_C и p_D :

$$f = \frac{1}{n_s} \sum_i^{n_s} \left[(p_C^w(t_i) - p_D^w(t_i))^2 - d_{CD}^2 \right]^2.$$

Оптимизация была проведена в программной среде MatLab численным методом Хука — Дживса (patternsearch), который входит в состав пакета Global Optimization Toolbox. Были введены ограничения на значения координат точек, обусловленные желаемыми габаритными размерами робота. В процессе оптимизации вычисляются значения координат точек p_C и p_D , при которых целевая функция принимает минимальное значение, а замкнутая кинематическая цепь с контуром $ABCD$ способна обеспечивать движение точки F вдоль требуемой траектории. Результат выполнения оптимизации — механизм, приведенный на рис. 1, б.

Приведение в движение. Одним из способов передачи движения от двигателя к исполнительному звену является использование редукторов с большим передаточным отношением. Такой способ позволяет значительно увеличить приводной момент, однако накладывает также ряд ограничений, связанных с наличием трения и люфтов между зубчатыми колесами. Например, передаточное отношение N редуктора означает возрастание инерции в N^2 раз, что обуславливает более высокие ударные нагрузки, передающиеся на зубья шестерен [17]. Кроме того, у приводов с высоким передаточным числом редуктора, как правило, отсутствует обратный ход (non-back drivability). Это может привести, например, к поломке редуктора при внезапном воздействии внешней нагрузки на выходное звено.

Другое решение — использование привода с последовательной эластичностью (Series Elastic Actuation — SEA). Отличительной особенностью данного типа актуатора, принцип работы которого подробно описан в [17], является наличие пружин между приводным валом и звеном, приводимым в движение. Упрощенное представление интеграции SEA в модель робота изображено на рис. 1, в. Пружина устанавливается между точками H и G . Помимо обеспечения возможности обратного хода (back drivability), основными преимуществами привода SEA являются устойчивость к ударам, более точное регулирование приводного момента, а также снижение пикового усилия, передающегося обратно на вал сервопривода.

В пакете MatLab Simscape было проведено имитационное моделирование синтезированной конструкции. Визуализация фазы полета робота при прыжке представлена на рис. 1, г. При построении модели были учтены масса конструкции, жесткость пружин привода и поверхности, с которой контактирует робот. Модель контактного взаимодействия описывается

уравнениями силы реакции вдоль вертикальной $\hat{y}$ и горизонтальной $\hat{x}$ осей. Вертикальная составляющая соответствует вязкоупругой модели Кельвина — Войта [18]:

$$F_{\hat{y}} = K\delta + b\dot{\delta},$$

где δ — глубина проникновения контактной поверхности ноги в поверхность земли, м; K — коэффициент жесткости контактного взаимодействия, Н/м; b — коэффициент демпфирования контактного взаимодействия, Н·с/м.

Для горизонтальной составляющей используется модель кулоновского трения с учетом проскальзывания [19]:

$$F_{\hat{x}} = -\mu F_y \operatorname{sgn} v_x,$$

где μ — коэффициент статического трения; v_x — горизонтальная составляющая скорости, м/с.

Моделирование позволило оценить высоту (h) прыжка (рис. 2, а), диапазоны углов вращения сочленений механизма и действующие в них нагрузки. Эти данные необходимы для синтеза гибких сочленений. На рис. 2, б представлен график изменения моментов, генерируемых в сочленениях робота (M_A, M_B, M_C, M_D), в процессе его работы на протяжении 5 с из состояния покоя: зеленая зона соответствует этапу нагрузки пружин, т. е. накопления энергии; розовая зона — этапу разгрузки пружин, т. е. высвобождения энергии; белая зона — этапу свободного полета робота. На графике отчетливо видны всплески значений моментов, возникающие при переходе между этапами. Амплитуды всплесков зависят напрямую от коэффициентов жесткости K и демпфирования b модели контактного взаимодействия и являются исключительно индикаторами катастрофических событий — приземления робота и смены характера работы пружин (рис. 2, б). На увеличенном графике (рис. 2, в) видны изменения моментов между катастрофическими событиями. Во время полета на сочленения не действуют нагружающие моменты. По полученным значениям моментов с учетом коэффициента запаса были синтезированы гибкие сочленения.

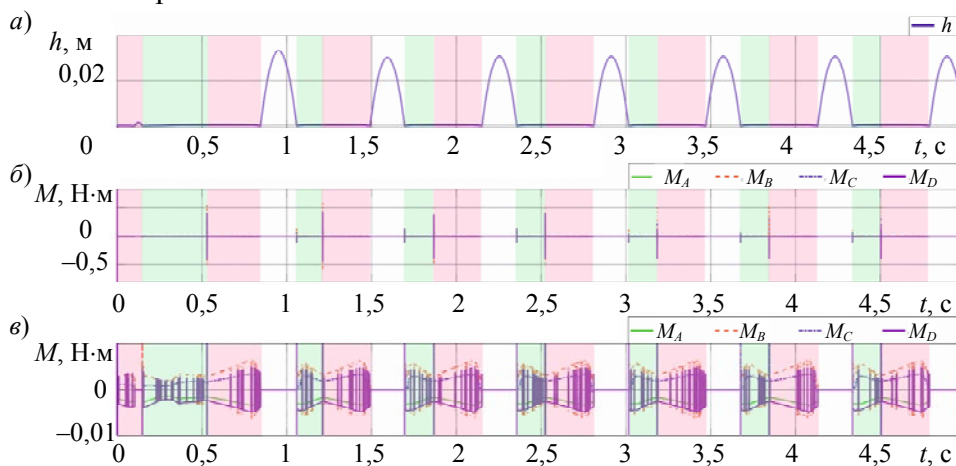


Рис. 2

На рис. 3 представлены графики для оценки энергопотребления прыгающего робота: момент серводвигателя M , угловая скорость ω , мощность P и итоговая потребленная энергия E (данные получены с использованием имитационной модели). Для управления роботом использован реактивный регулятор, принцип работы которого, описанный в [20], состоит во „впрыскивании“ энергии в систему в определенные моменты движения: когда сочленения начинают нагружаться, двигатель вращается, чтобы помочь ноге робота максимально сжаться; во время разгибания ноги двигатель вращается в обратную сторону для создания дополнительного усилия при отталкивании.

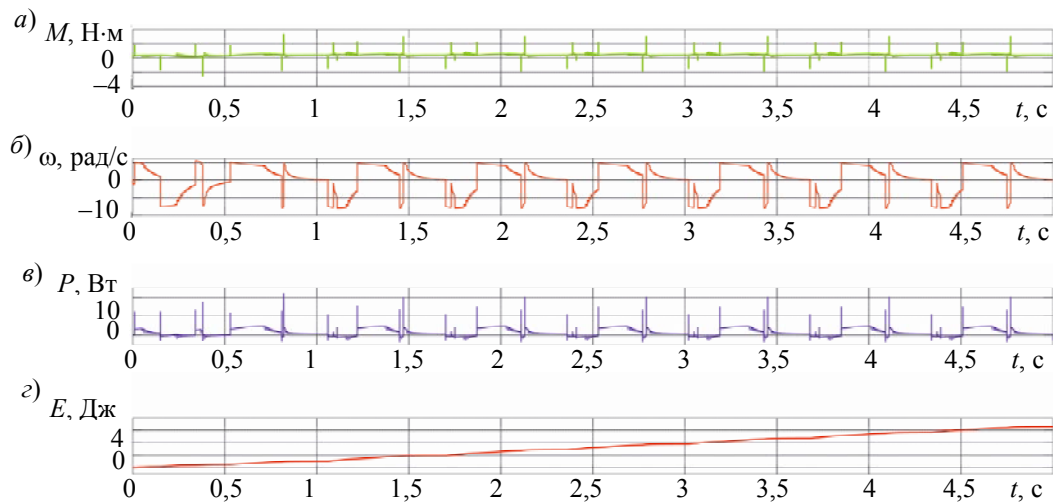


Рис. 3

Синтез гибких сочленений. Гибкие сочленения — разновидность податливых механизмов и зачастую используются в качестве замены классическим вращательным шарнирам [21]. Важными преимуществами гибких сочленений являются отсутствие люфтов, низкий эластичный гистерезис, а также более низкий вес и меньшие габариты по сравнению с традиционной компоновкой „подшипник-ось“. Гибкие сочленения характеризуется малой площадью в центре петли эластичного гистерезиса, что обуславливает малое количество энергии, рассеиваемой из-за внутреннего трения материала [22].

Существует множество топологий гибких сочленений и методов их проектирования. Например, в работах [23, 24] приведены структуры сочленений с одной вращательной степенью свободы; их существенные недостатки — малый диапазон углов вращения (до 10°) и низкая податливость в направлении вращения. Другим решением является метод проектирования гибких сочленений из листовых пружин [25, 26]. Протез пальца с двумя гибкими сочленениями, которые обеспечивают соответствующие степени свободы, представлен в работе [27]: синтезированный протез способен обеспечивать 20 Н контактного усилия в диапазоне движения 80° .

Оптимизированная структура, позволяющая роботу совершать прыжок на месте, содержит четыре сочленения. Для апробации рассматриваемого метода проектирования все вращательные шарниры предложено заменить на гибкие сочленения. Для прыгающего робота была выбрана описанная в [25] топология сочленения, обладающая рядом преимуществ: большой диапазон углов вращения (до 90°), сохранение податливости по оси вращения на всем интервале движения, возможность анализа жесткости в направлениях, перпендикулярных оси вращения.

Для синтеза гибких сочленений использовался инструмент SPACAR Toolbox [28], который позволяет оптимизировать топологию сочленения в зависимости от действующих нагрузок и желаемых углов вращения. Оптимизация направлена на увеличение значения первой паразитной собственной частоты сочленения f , что способствует увеличению точности воспроизводимого движения в ограниченном диапазоне [29]. Целевая функция определяется как

$$F(x) = \min f(x, \theta)^{-1}, \quad (2)$$

где x — вектор параметров сочленения, θ — угол отклонения.

Вектор параметров выбранного сочленения для прыгающего робота имеет вид $x = [L, W, H, T]^T$, где L , W , H — длина, ширина и глубина сочленения соответственно; T — толщина листовых пружин.

Для того чтобы предотвратить неограниченный рост параметров во время оптимизации и гарантировать, что алгоритм выполнит построение геометрии сочленения, применяется

функция ограничений. Ограничения на максимальный момент $M(x, \theta) < M_{\max}$, при котором достигается максимальный угол $\theta_{\max}$, и на максимальное возникающее напряжение $\sigma(x, \theta) < \sigma_{\max}$ определяют нелинейную функцию ограничения:

$$C(x) = \max \left\{ \begin{array}{l} \max \sigma(x, \theta) - \sigma_{\max} \\ M(x, \theta) - M_{\max} \end{array} \right\} \forall \theta,$$

где допустимый вектор параметров должен удовлетворять $C(x) < 0$.

Помимо наложенных ограничений, при оптимизации учитываются основные свойства материала (модуль Юнга E , плотность ρ , модуль сдвига G), из которого планируется изготавливать сочленение.

Результатом оптимизации гибкого сочленения является трехмерная модель, изображенная на рис. 4, а.

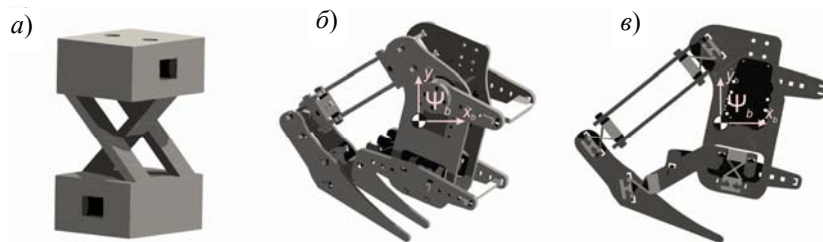


Рис. 4

Изготовление прототипа. В соответствии с полученными результатами в САПР SolidWorks была построена твердотельная модель разрабатываемого робота. На рис. 4, б представлена визуализация модели вместе с отображением центра масс, который находится над точкой контакта в области корпуса робота, согласно аналитической модели перевернутого маятника.

Робот был спроектирован таким образом, что каждое гибкое сочленение может быть заменено в случае необходимости. Основные корпусные детали нарезаны на лазерном гравюре из материала полиацеталь. Гибкие сочленения изготовлены посредством трехмерной печати. В качестве материала использовался полиуретан марки BF FLEX. Для крепления деталей между собой не использовались метизы — детали крепятся посредством соединения типа „шип — паз“, что способствует уменьшению общей массы конструкции. В качестве приводного двигателя выбран маломощный сервомотор Dynamixel AX-12A [30].

На рис. 5 представлен прототип, корпус которого присоединен к направляющей алюминиевой трубке, обеспечивающей перемещения вдоль вертикальной прямой $\hat{y}$.

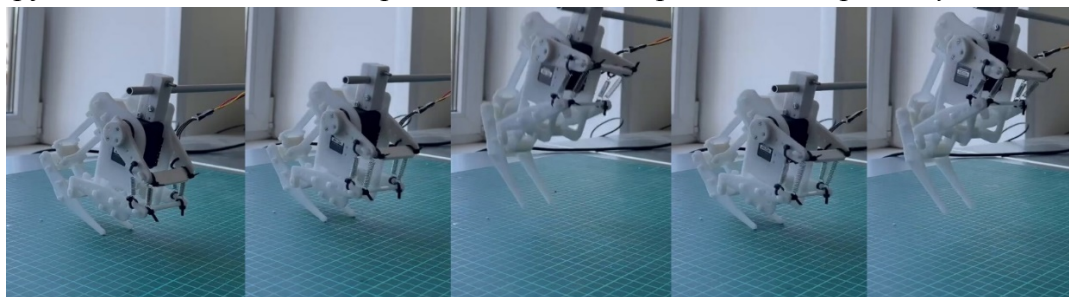


Рис. 5

Получение данных, регистрируемых роботом, осуществляется с помощью тензорезистивных датчиков. Принцип их действия заключается в изменении значения сопротивления в зависимости от степени деформации. Это обусловлено свойствами металлической резистивной фольги, из которой датчики изготавливаются [31]. Так как гибкие сочленения эластично деформируются в процессе движения, тензорезисторы прикрепляются к гибким элементам сочленения с обеих сторон. В результате образуется электрическая цепь, позволяющая выяв-

лять мельчайшие изменения сопротивления — мост Уитстона. Такое решение позволяет преобразовать электрический сигнал, соответствующий деформации, в угол поворота.

На рис. 6 представлены графики, отражающие показания тензорезистора на одном из сочленений: q_{ref} отображает сигнал, полученный с энкодера серводвигателя AX12-A, соответствующий синусоиде с частотой 3 Гц и амплитудой 20° ; q_h — сигнал, полученный с тензорезистора. Представленные данные получены в ходе эксперимента, для которого была собрана стендовая установка для тестирования прототипа (рис. 7). Двигатель воспроизводит движение, аналогичное нагружению сочленения робота при прыжке.

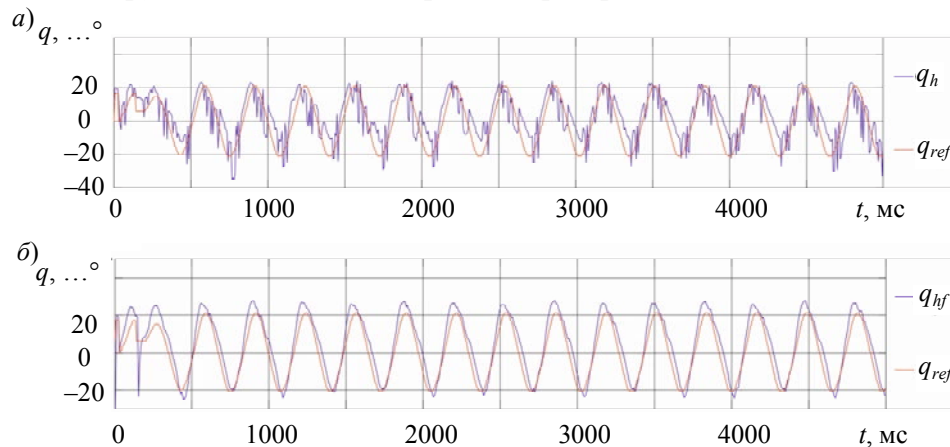


Рис. 6

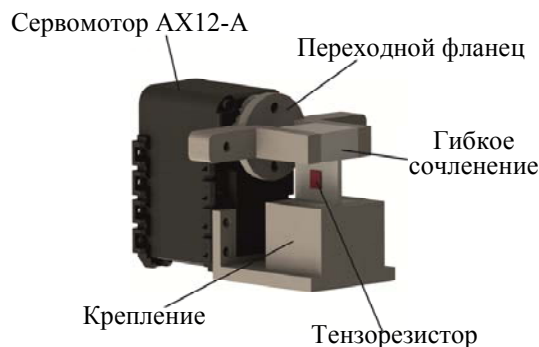


Рис. 7

На рис. 6, а отчетливо видны шумы и выбросы, которые необходимо фильтровать для корректной обработки сигнала. Для данной системы был выбран фильтр бегущего среднего, который задается следующим выражением:

$$\zeta = \zeta + (\eta - \zeta)k,$$

где ζ — фильтрованное значение; η — значение, которое требуется фильтровать; k — коэффициент фильтрации, значение которого настраивается в пределах от 0,0 до 1,0 — чем меньше значение, тем плавнее фильтрация.

Данный алгоритм был выбран вследствие его простой реализации и скорости вычислений.

На рис. 6, б приведен график после фильтрации: q_{hf} соответствует значению угла поворота после обработки фильтра. Среднеквадратическое отклонение за период работы составляет $5,58^\circ$, максимальное отклонение от желаемого значения составляет $11,23^\circ$. Сигнал тензорезистора совпадает по форме и частоте с референсным сигналом, полученным с энкодера серводвигателя. Это говорит о том, что гибкое сочленение, параметры которого определены в процессе оптимизации, способно точно воспроизводить желаемую траекторию.

Заключение. Представлены результаты исследования по проектированию и изготовлению прототипа энергоэффективного прыгающего робота с гибкими сочленениями с использованием принципов морфологического расчета. Синтезирована и оптимизирована кинематическая схема робота, позволяющая совершать прыжок на месте. Она представляет собой замкнутую, четырехзвенную структуру с одной степенью свободы. Параметрическая оптимизация звеньев позволила обеспечить следование точки контакта вдоль желаемой траектории. В качестве актуатора был выбран привод с последовательной эластичностью, что позволило уменьшить требуемую мощность электродвигателя.

На основе данных, полученных с использованием имитационной модели, синтезированы гибкие сочленения, заменяющие кинематические пары. Данное решение принято для уменьшения массы. Из-за свойств материала для сочленений условно принято, что вся накопленная энергия рассеивается вследствие высокого демпфирования. Поэтому гибкие сочленения не рассматриваются как элементы рекуперации энергии.

Интеграция гибких сочленений для уменьшения массы конструкции и избавления от трений в узлах, а также использование маломощного двигателя являются приемами по увеличению энергоэффективности прототипа.

В САПР SolidWorks спроектирована твердотельная модель робота. Посредством трехмерной печати и резки на лазерном гравере были изготовлены основные детали. Результатом выполнения работы стал прототип прыгающего робота с гибкими сочленениями, который демонстрирует работоспособность предложенного способа проектирования.

СПИСОК ЛИТЕРАТУРЫ

1. *Seok S. et al.* Actuator design for high force proprioceptive control in fast legged locomotion // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. IEEE, 2012. P. 1970—1975.
2. *Hauser S., Dujany M., Arreguit J., Ijspeert A., Iida F.* Coupling-dependent convergence behavior of phase oscillators with tegotae-control // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). 2021.
3. *Ye K., Karydis K.* Modeling and Trajectory Optimization for Standing Long Jumping of a Quadruped With a Preloaded Elastic Prismatic Spine // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). 2021.
4. *Fankhauser P., Hutter M.* ANYmal: a unique quadruped robot conquering harsh environments // Research Features. 2018. N 126. P. 54—57.
5. *Bledt G. et al.* MIT Cheetah 3: Design and control of a robust, dynamic quadruped robot // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). IEEE, 2018. P. 2245—2252.
6. *Park H. W., Wensing P. M., Kim S.* High-speed bounding with the MIT Cheetah 2: Control design and experiments // Intern. Journal of Robotics Research. 2017. Vol. 36, N 2. P. 167—192.
7. *Zhong Y. et al.* Analysis and research of quadruped robot's legs: a comprehensive review // Intern. Journal of Advanced Robotic Systems. 2019. Vol. 16, N 3. P. 1729881419844148.
8. *Borisova O., Borisov I., Kolyubin S., Stramigioli S.* Design of Galloping Robots with Elastic Spine: Tracking Relations between Dynamic Model Parameters Based on Motion Analysis of a Real Cheetah // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). 2021.
9. *Folkertsma G. A.* Energy-Based and Biomimetic Robotics. Enschede: Univ. of Twente, 2017.
10. *Seok S. et al.* Design principles for energy-efficient legged locomotion and implementation on the MIT cheetah robot // IEEE/ASME Trans. on Mechatronics. 2014. Vol. 20, N 3. P. 1117—1129.
11. *Folkertsma G. A., Kim S., Stramigioli S.* Parallel stiffness in a bounding quadruped with flexible spine // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. IEEE, 2012. P. 2210—2215.
12. *Hurst J. W., Chestnutt J. E., Rizzi A. A.* An actuator with physically variable stiffness for highly dynamic legged locomotion // Proc. IEEE Intern. Conf. on Robotics and Automation, ICRA'04. IEEE, 2004. Vol. 5. P. 4662—4667.

13. *Vu H. Q.* et al. A variable stiffness mechanism for improving energy efficiency of a planar single-legged hopping robot // Proc. of the 16th Intern. Conf. on Advanced Robotics (ICAR). IEEE, 2013. P. 1—7.
14. *Radhakrishnan V.* Locomotion: dealing with friction // Proc. of the National Academy of Sciences. 1998. Vol. 95, N 10. P. 5448—5455.
15. *Haberland M.* et al. The effect of swing leg retraction on running energy efficiency // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. IEEE, 2011. P. 3957—3962.
16. *Thomaszewski B.* et al. Computational design of linkage-based characters // ACM Trans. on Graphics (TOG). 2014. Vol. 33, N 4. P. 1—9.
17. *Pratt G. A., Williamson M. M.* Series elastic actuators // Proc. IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots. IEEE, 1995. Vol. 1. P. 399—406.
18. *Pappalardo A.* et al. Hunt–crossley model based force control for minimally invasive robotic surgery // Biomedical Signal Processing and Control. 2016. Vol. 29. P. 31—43.
19. *Townsend W., Salisbury J.* The effect of coulomb friction and stiction on force control // Proc. IEEE Intern. Conf. on Robotics and Automation. IEEE, 1987. Vol. 4. P. 883—889.
20. *Lakatos D., Albu-Schäffer A.* Switching based limit cycle control for compliantly actuated second-order systems // IFAC Proc. Volumes. 2014. Vol. 47, N 3. P. 6392—6399.
21. *Howell L. L.* Compliant mechanisms // 21st Century Kinematics; Ed. *I. M. McCarthy*. Springer, 2013. P. 189—216.
22. *Ashby M. F., Jones D. R. H.* Engineering Materials 1: An Introduction to Properties, Applications and Design. Elsevier, 2012. Vol. 1.
23. *Chen G. M., Jia J. Y., Li Z. W.* Right-circular corner-fillet flexure hinges // IEEE Intern. Conf. on Automation Science and Engineering. IEEE, 2005. P. 249—253.
24. *Lobontiu N.* Compliant Mechanisms: Design of Flexure Hinges. CRC Press, 2002.
25. *Naves M.* et al. Flexure-based 60 degrees stroke actuator suspension for a high torque iron core motor // Precision Engineering. 2020. Vol. 63. P. 105—114.
26. *Naves M., Aarts R., Brouwer D. M.* Large-stroke flexure hinges: Building-block-based spatial topology synthesis method for maximising flexure performance over their entire range of motion // Mikroniek. 2017. Vol. 57, N 3. P. 5—9.
27. *Boers A. S. B.* et al. Inverted curved flexure hinge with torsional reinforcements in a printed prosthetic finger // 33rd ASPE Annual Meeting. 2018.
28. SPACAR Wiki. 2021 [Электронный ресурс]: <<http://www.spacar.nl/wiki/doku.php?id=start>>, 10.04.2021.
29. *Fix M. E., Brouwer D. M., Aarts R. G. K. M.* Building Block Based Topology Synthesis Algorithm to Optimize the Natural Frequency in Large Stroke Flexure Mechanisms // Intern. Design Engineering Technical Conf. and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, 2020. Vol. 83914. P. V002T02A007.
30. ROBOTIS e-Manual AX-12A [Электронный ресурс]: <<https://emanual.robotis.com/docs/en/dxl/ax/ax-12a/>>, 25.06.2021.
31. Тензорезисторы KYOWA. 2018 [Электронный ресурс]: <<http://kyowa.ru/produksiya/tenzorezistori.html>>, 20.06.2021.

Сведения об авторах

Кирилл Вячеславович Насонов

— магистрант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; инженер; E-mail: kvnasonov@itmo.ru

Иван Игоревич Борисов

— канд. техн. наук; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; научный сотрудник; E-mail: borisovii@itmo.ru

Поступила в редакцию 28.12.21; одобрена после рецензирования 11.01.22; принята к публикации 18.01.22.

REFERENCES

1. Seok S. et al. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 1970–1975.
2. Hauser S., Dujany M., Arreguit J., Ijspeert A., & Iida F. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS 2021, 2021.
3. Ye K. and Karydis K. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS 2021, 2021.
4. Fankhauser P., Hutter M. *Research Features*, 2018, no. 126, pp. 54–57.
5. Bledt G. et al. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2018, pp. 2245–2252.
6. Park H.W., Wensing P.M., Kim S. *International Journal of Robotics Research*, 2017, no. 2(36), pp. 167–192.
7. Zhong Y. et al. *International Journal of Advanced Robotic Systems*, 2019, no. 3(16), pp. 1729881419844148.
8. Borisova O., Borisov I., Kolyubin S., Stramigioli S. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, IROS 2021, 2021.
9. Folkertsma G.A. *Energy-based and biomimetic robotics*, 2017.
10. Seok S. et al. *IEEE/ASME transactions on mechatronics*, 2014, no. 3(20), pp. 1117–1129.
11. Folkertsma G.A., Kim S., Stramigioli S. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2012, pp. 2210–2215.
12. Hurst J.W., Chestnutt J.E., Rizzi A.A. *IEEE International Conference on Robotics and Automation*, Proceedings. ICRA'04, 2004, vol. 5, pp. 4662–4667.
13. Vu H.Q. et al. *2013 16th International Conference on Advanced Robotics (ICAR)*, IEEE, 2013, pp. 1–7.
14. Radhakrishnan V. *Proceedings of the National Academy of Sciences*, 1998, no. 10(95), pp. 5448–5455.
15. Haberland M. et al. *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2011, pp. 3957–3962.
16. Thomaszewski B. et al. *ACM Transactions on Graphics (TOG)*, 2014, no. 4(33), pp. 1–9.
17. Pratt G.A., Williamson M.M. *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Human Robot Interaction and Cooperative Robots, IEEE, 1995, vol. 1, pp. 399–406.
18. Pappalardo A. et al. *Biomedical Signal Processing and Control*, 2016, vol. 29, pp. 31–43.
19. Townsend W., Salisbury J. *Proceedings 1987 IEEE International Conference on Robotics and Automation*, IEEE, 1987, vol. 4, pp. 883–889.
20. Lakatos D., Albu-Schäffer A. *IFAC Proceedings Volumes*, 2014, no. 3(47), pp. 6392–6399.
21. Howell L.L. *21st century kinematics*, Springer, London, 2013, pp. 189–216.
22. Ashby M.F., Jones D.R.H. *Engineering materials 1: an introduction to properties, applications and design*, Elsevier, 2012, vol. 1.
23. Chen G.M., Jia J.Y., Li Z.W. *IEEE International Conference on Automation Science and Engineering*, 2005, pp. 249–253.
24. Lobontiu N. *Compliant mechanisms: design of flexure hinges*, CRC press, 2002.
25. Naves M. et al. *Precision engineering*, 2020, vol. 63, pp. 105–114.
26. Naves M., Aarts R., Brouwer D.M. *Mikroniek*, 2017, no. 3(57), pp. 5–9.
27. Boers A.S.B. et al. *33rd ASPE Annual Meeting 2018*, 2018.
28. SPACAR Wiki, 2021, <http://www.spacar.nl/wiki/doku.php?id=start>.
29. Fix M.E., Brouwer D.M., Aarts R.G.K.M. *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, 2020, vol. 83914, pp. V002T02A007.
30. *ROBOTIS e-Manual AX-12A*, <https://manual.robotis.com/docs/en/dxl/ax/ax-12a/>.
31. <http://kyowa.ru/produktsiya/tenzorezistori.html>.

Data on authors

- | | |
|--------------------------|---|
| Kirill V. Nasonov | — Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Engineer; E-mail: kvnasonov@itmo.ru |
| Ivan I. Borisov | — PhD; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Researcher; E-mail: borisovii@itmo.ru |

Received 28.12.21; approved after reviewing 11.01.22; accepted for publication 18.01.22.

КОМПЬЮТЕРНОЕ ЗРЕНИЕ И ПЛАНИРОВАНИЕ ДВИЖЕНИЯ РОБОТОВ В ЗАДАЧАХ МАНИПУЛИРОВАНИЯ

COMPUTER VISION AND ROBOT MOVEMENT PLANNING IN MANIPULATION TASKS

УДК 004.021
DOI: 10.17586/0021-3454-2022-65-3-185-193

АЛГОРИТМ МОДИФИЦИРОВАННОГО ИНТЕЛЛЕКТУАЛЬНОГО ДВУНАПРАВЛЕННОГО СЛУЧАЙНОГО ДЕРЕВА ДЛЯ ПЛАНИРОВАНИЯ ДВИЖЕНИЯ АНТРОПОМОРФНЫХ МАНИПУЛЯТОРОВ

И. С. ДОВГОПОЛИК^{1*}, К. АРТЕМОВ¹, О. И. БОРИСОВ¹,
С. ЗАБИХИФАР², А. Н. СЕМОЧКИН²

¹ Университет ИТМО, Санкт-Петербург, Россия

*isdovgopolik@itmo.ru

² Сбербанк, Москва, Россия

Аннотация. Рассматривается алгоритм планирования движения многозвенной робототехнической системы в пространстве с препятствиями. Основные требования, предъявляемые к данной задаче, — высокое быстродействие и эффективное использование памяти в процессе работы. Представлен алгоритм планирования пути на основе метода двунаправленного быстроисследующего случайного дерева с использованием подхода, исключающего добавление новых вершин к дереву, если по их расположению в пространстве можно однозначно определить нецелесообразность их использования для построения пути. Эта модификация позволяет ускорить планирование движения и уменьшить объем памяти, необходимой для хранения данных об анализе пространства.

Ключевые слова: планирование пути, быстроисследующие случайные деревья, исключение неиспользуемых вершин, iCub

Ссылка для цитирования: Довгополик И. С., Артемов К., Борисов О. И., Забихифар С., Семочкин А. Н. Алгоритм модифицированного интеллектуального двунаправленного случайного дерева для планирования движения антропоморфных манипуляторов // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 185—193. DOI: 10.17586/0021-3454-2022-65-3-185-193.

MODIFIED INTELLIGENT BIDIRECTIONAL RANDOM TREE ALGORITHM FOR PLANNING THE MOVEMENT OF ANTHROPOMORPHIC MANIPULATORS

I. S. Dovgopolik^{1*}, K. Artemov¹, O. I. Borisov¹, S. Zabihifar², A. N. Semochkin²

¹ ITMO University, St. Petersburg, Russia

*isdovgopolik@itmo.ru

² Sberbank, Moscow, Russia

Abstract. An algorithm for planning the movement of a multi-link robotic system in an environment with obstacles is considered. The main requirements for this task are high performance and efficient memory usage during operation. An algorithm for path planning based on the method of a bidirectional fast-investigating random tree is presented. An approach is used which excludes addition of new vertices to the tree if their location in space can unambiguously conclude that it is inappropriate to use them in the path construction. This modification makes it possible to speed up movement planning and reduce the amount of memory needed to store the environment analysis data.

Keywords: path planning, rapidly-exploring random trees, exclusion of unused vertices, iCub

For citation: Dovgopolik I. S., Artemov K., Borisov O. I., Zabihifar S., Semochkin A. N. Modified intelligent bidirectional random tree algorithm for planning the movement of anthropomorphic manipulators. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 185—193 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-185-193.

Введение. Проблема планирования движения составляет большинство задач, выполняемых роботами. Эта проблема актуальна как для мобильных роботов, так и для многосвязных, которые выполняют поиск пути для взаимодействия с объектами окружающей среды, для захвата объектов и манипулирования ими. Задача планирования движения в первую очередь заключается в определении траектории перемещения, к которой предъявляются три основных требования [1]. Во-первых, траектория должна проходить от заданной начальной точки к заданной конечной точке; во-вторых, она должна проходить только через свободную часть рабочего пространства — какие-либо пересечения с зонами, занятыми препятствиями, недопустимы; в-третьих, полученная траектория должна быть оптимальной по параметру, который определяется спецификой задачи.

Для многих роботов, в частности антропоморфных, важным требованием является автономность, что накладывает значительные ограничения на вычислительную мощность и объем доступной для работы внутренней памяти. Задачу планирования движения можно назвать рутинной для робототехнической системы, поэтому она должна выполняться быстро и не требовать применения значительных ресурсов. В настоящей статье предлагается решение данной задачи, для чего необходимо:

- исследовать линейку алгоритмов быстроисследующих случайных деревьев (Rapidly Exploring Random Trees — RRT);
- провести моделирование работы линейки алгоритмов и оценить эффективность использования памяти;
- усовершенствовать лучший по результатам моделирования алгоритм и использовать его для сокращения объема памяти;
- оценить эффективность работы модифицированного алгоритма.

Алгоритм RRT и его модификации. Алгоритм, предложенный в [2], работает следующим образом. Исходными данными являются начальная и целевая конфигурации робота и данные о пространстве. Первая вершина дерева — это начальная конфигурация. В пространство добавляется случайная точка на расстоянии, равном шагу алгоритма от текущей точки. Если проверка пройдена, вершина и линия добавляются в дерево. Если случайная и целевая точки совпадают, построение траектории выполнено. В противном случае действия повторяются. Структурная схема алгоритма представлена на рис. 1. В этом методе не используются алгоритмы оптимизации при определении траектории.

В [3] представлен оптимальный вариант RRT, при котором траектория строится таким образом, чтобы сумма стоимостей перехода между вершинами дерева от начальной точки к конечной была минимальной. Когда количество случайных точек стремится к бесконечности, этот подход обеспечивает асимптотически оптимальное решение, понимаемое как решение, наиболее близкое к оптимальному по сравнению с другими решениями [4].

Для ускорения процесса поиска траектории было предложено двунаправленное быстроисследующее случайное дерево [5]. В этом алгоритме инициализируются два дерева, одно — из начальной точки траектории, второе — из конечной точки. Если какие-либо вершины первого и второго деревьев находятся на расстоянии, равном или меньшем, чем шаг алгоритма, и между этими вершинами нет областей ограничения, деревья соединяются. Структурная схема этого алгоритма представлена на рис. 2. Такой подход позволяет увеличить скорость по-

строения траектории и уменьшить объем требуемой памяти, поскольку уменьшается необходимое количество итераций для успешного определения пути.

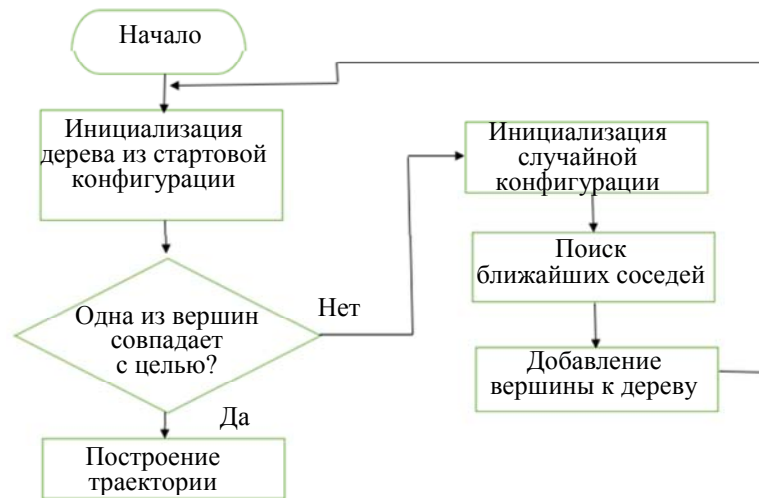


Рис. 1

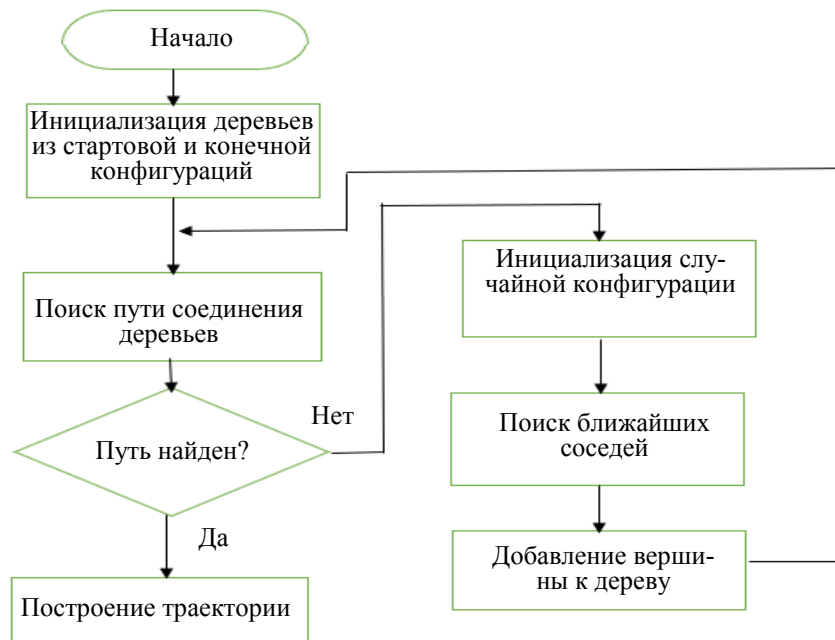


Рис. 2

В работах [6, 7] предложен модифицированный подход, в котором реализован способ выбора вершин для присоединения. Модификация получила название „интеллектуальное двунаправленное быстроисследующее случайное дерево“. Алгоритм определяет, к какому из деревьев целесообразно добавить случайную точку, и оценивает расстояние между нею и ближайшей вершиной дерева. Вершина присоединяется к ближайшему дереву. Структурная схема алгоритма представлена на рис. 3.

Подход, предусматривающий ограничение области поиска для увеличения скорости планирования и уменьшения объема используемой памяти, представлен в [8]. Подпространство для поиска пути предложено определять как фигуру эллипсоидальной формы, размеры которой определяются линейным расстоянием между начальной и конечной точками траектории и длиной первого пути, который найден в пространстве посредством оригинального алгоритма. Такое подпространство должно уменьшаться, постепенно „складываясь“ в прямую, соединяющую начало и конец траектории, что необходимо для поиска оптимального решения. Во избежание необходимости поиска решения стандартным алгоритмом RRT, в [9] был предложен подход к поиску траектории путем инициализации случайных точек

в пространстве группами. В данном случае построение подпространства для поиска опирается только на линейное расстояние между начальной и конечной точками траектории. Случайные точки инициализируются группами, каждая из которых содержит n точек. Если после инициализации всех точек не был найден ни один путь, подпространство поиска расширяется и вновь инициализируются n точек. После того как первый путь был найден, подпространство начинает сужаться после инициализации каждой группы точек. Такой подход имеет выраженные недостатки — необходимость расширения области поиска и вероятность отсутствия пути в подпространстве из-за ограниченного количества инициализированных точек.

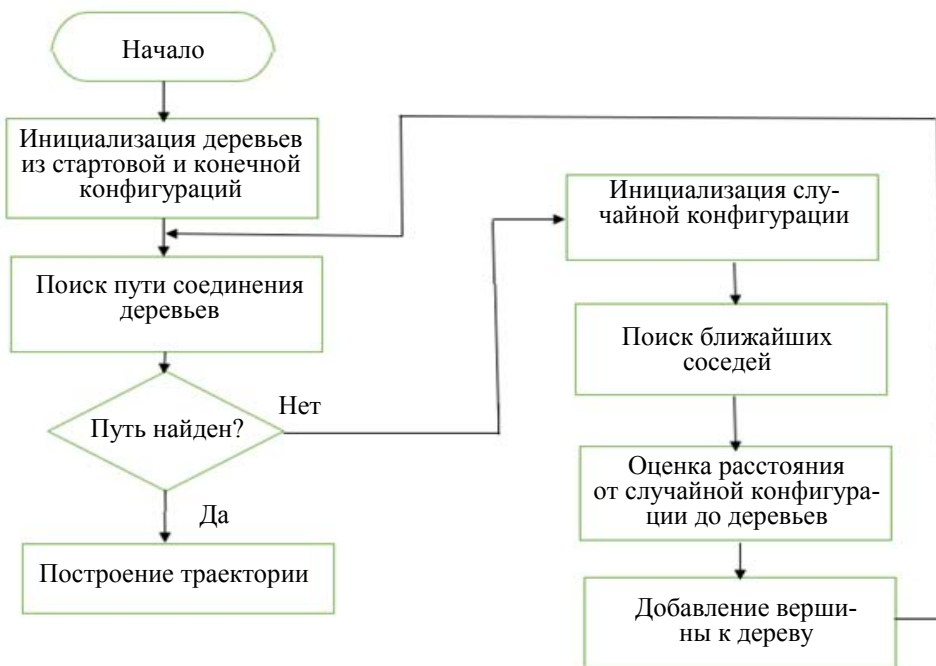


Рис. 3

Методы на основе RRT приобретают широкое распространение [1, 10, 11]. Для определения траектории не нужно анализировать все пространство, достаточно лишь часть, соответственно не требуется большой объем памяти. С увеличением размерности пространства сложность вычислений почти не возрастает, что особенно важно для многосвязных робототехнических систем.

Моделирование работы алгоритмов RRT. Эксперимент проводился для двух вариантов начальной и конечной точек траектории робота в конфигурационном пространстве. Показателем используемого объема памяти является количество вершин. Траектории, построенные с использованием алгоритмов двунаправленного RRT и интеллектуального двунаправленного RRT, представлены на рис. 4, а, б соответственно (здесь q_1, q_2, q_3 — углы поворота сочленений робота).

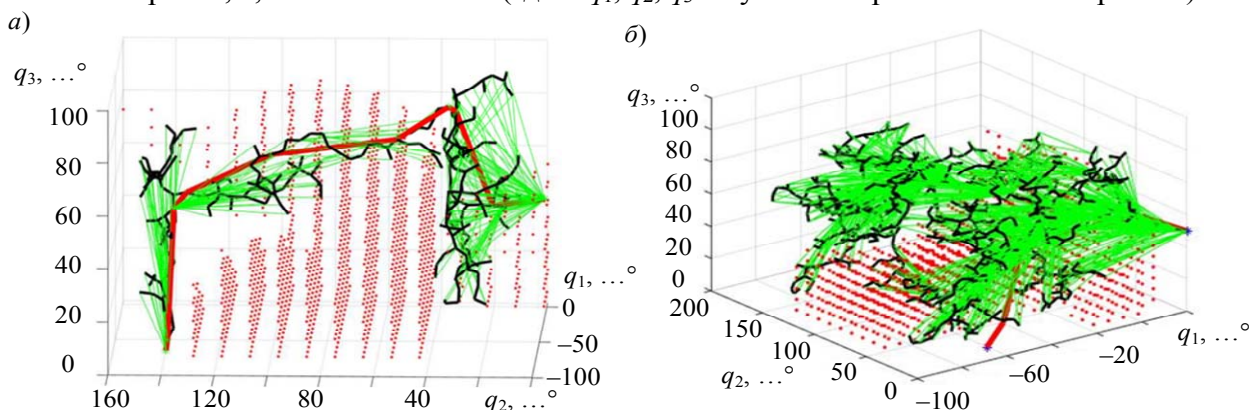


Рис. 4

Анализ рисунков показывает, что каждый алгоритм создает большое количество ветвей дерева, которые можно назвать заведомо неиспользуемыми. Подразумевается, что в процессе поиска пути в графе они никогда не будут востребованы в качестве составляющих частей траектории, поскольку построены не в пространстве между начальной и конечной точками траектории. Таким образом, часть памяти, используемая для хранения дерева, может быть освобождена путем исключения этих ветвей.

Модифицированное интеллектуальное двунаправленное RRT. Для сокращения используемого объема памяти необходимо организовать добавление к дереву вершин, которые располагаются в пространстве между начальной и конечной точками траектории. Это геометрическая задача, рассмотрим ее для двухмерного пространства. Предположим, что новая случайная конфигурация расположена в пространстве между начальной и конечной точками пути, для четкого определения ее положения введем критерий оценки. Рассмотрим рис. 5, а, где q_{start1} и q_{start2} — деревья, инициализированные из начальной и конечной точек траектории соответственно, q_{rand} — случайная конфигурация, q_{near1} и q_{near2} — ближайшие к случайной конфигурации вершины инициализированных деревьев. Соединив q_{near1} , q_{near2} и q_{rand} в треугольник, получим углы α_1 и α_2 , которые являются острыми.

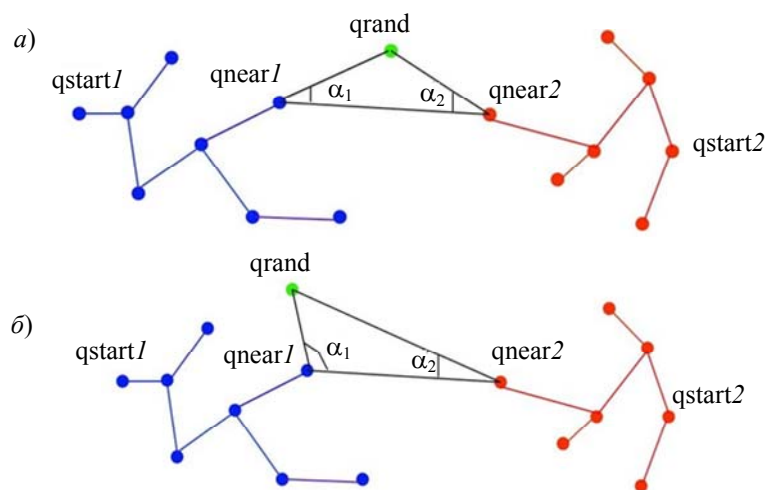


Рис. 5

Рассмотрим другой вариант (рис. 5, б) — новая случайная конфигурация расположена вне пространства между инициализированными деревьями. Если повторить аналогичные предыдущему варианту операции, получим углы α_1 и α_2 , при этом α_1 — тупой угол, α_2 — острый. Также из рисунка следует, что в случае присоединения случайной точки к дереву q_{start2} новая ветвь будет построена в пространстве между инициализированными деревьями, но если присоединить точку к дереву q_{start1} , то новая ветвь будет построена в сторону от пространства между начальной и конечной точками траектории, т.е. такая ветвь будет заведомо ложной. Таким образом, если в качестве критерия оценки расположения случайной точки выбирать углы α_1 и α_2 , то можно однозначно оценивать целесообразность ее присоединения к дереву и исключать заведомо ненужные вершины.

Структурная схема предлагаемого модифицированного алгоритма представлена на рис. 6. Осуществляется инициализация деревьев из начальной и конечной точек траектории. Затем определяется угол расширения дерева, по умолчанию равный 90° , что позволяет обходить вертикальные препятствия перед деревом. Однако необходимо проверить, не окружено ли дерево препятствиями так, что обойти их, учитывая ограничение на значение этого угла, невозможно.

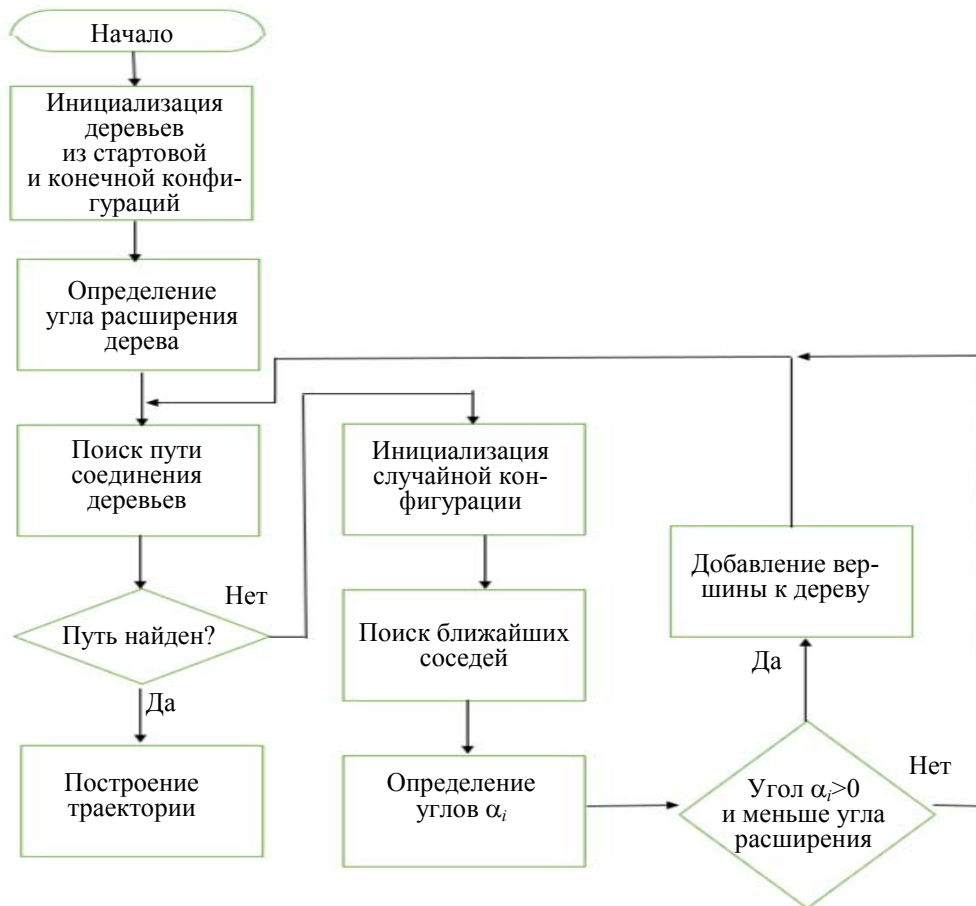


Рис. 6

Для проверки предлагается следующий алгоритм (рис. 7). От стартовой точки дерева q_{start} опускаются две перпендикулярные к линии $q_{start}-q_{goal}$ и друг к другу прямые. Выполняется поиск пересечения этих прямых с препятствиями. Если есть пересечение, то выполняется поиск крайних точек сфер и сравнение их отдаленности от q_{start} . На расстоянии, равном шагу алгоритма от самой удаленной точки, вводится другая точка и осуществляется построение треугольника abc , что позволяет определить угол расширения α_{lim} , который вычисляется с использованием теоремы косинусов по формуле

$$\alpha_{lim} = \arccos\left(\frac{b^2 + c^2 - a^2}{2bc}\right) + 90.$$

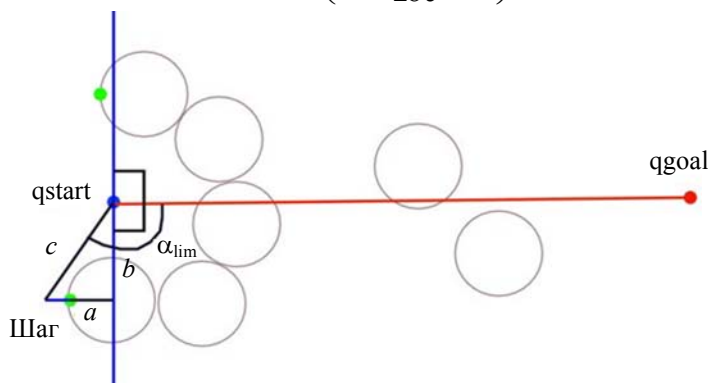


Рис. 7

Следующий этап модифицированного алгоритма — проверка возможности соединения деревьев. Если соединение возможно, то происходит построение траектории, если же невозможно, то инициализируется случайная точка, выполняется поиск ближайших соседей и оп-

ределяются углы α_i . Если угол α_i не превышает значения угла расширения, то точка добавляется как вершина дерева и запускается следующая итерация.

Представленный модифицированный алгоритм исключает построение лишних ветвей дерева, что позволяет ускорить процесс планирования движения и снизить объем используемой памяти. Для определения степени улучшения этих показателей проводятся моделирование и сравнительный анализ алгоритмов.

Сравнение алгоритмов. Сцена для проведения эксперимента представляет собой антропоморфного робота iCub, стену с окном, стол с предметами (рис. 8).



Рис. 8

Задача робота — переместить руку с захваченным предметом через препятствие (горизонтальное или вертикальное) без столкновений с объектами окружающей среды. Для решения задействуется три сочленения робота: два в плече и одно в локте. Это необходимо для визуализации конфигурационного пространства, в котором происходит планирование, поскольку пространство размерностью выше трех визуализировать невозможно, а управления этими тремя сочленениями достаточно для перемещения руки по всему рабочему пространству. Каждый алгоритм запускается 10 раз для каждой задачи. На рис. 9, *а–в* представлены несколько траекторий, полученных по результатам работы алгоритмов: *а* — двунаправленного RRT; *б* — интеллектуального двунаправленного RRT; *в* — модифицированного интеллектуального двунаправленного RRT.

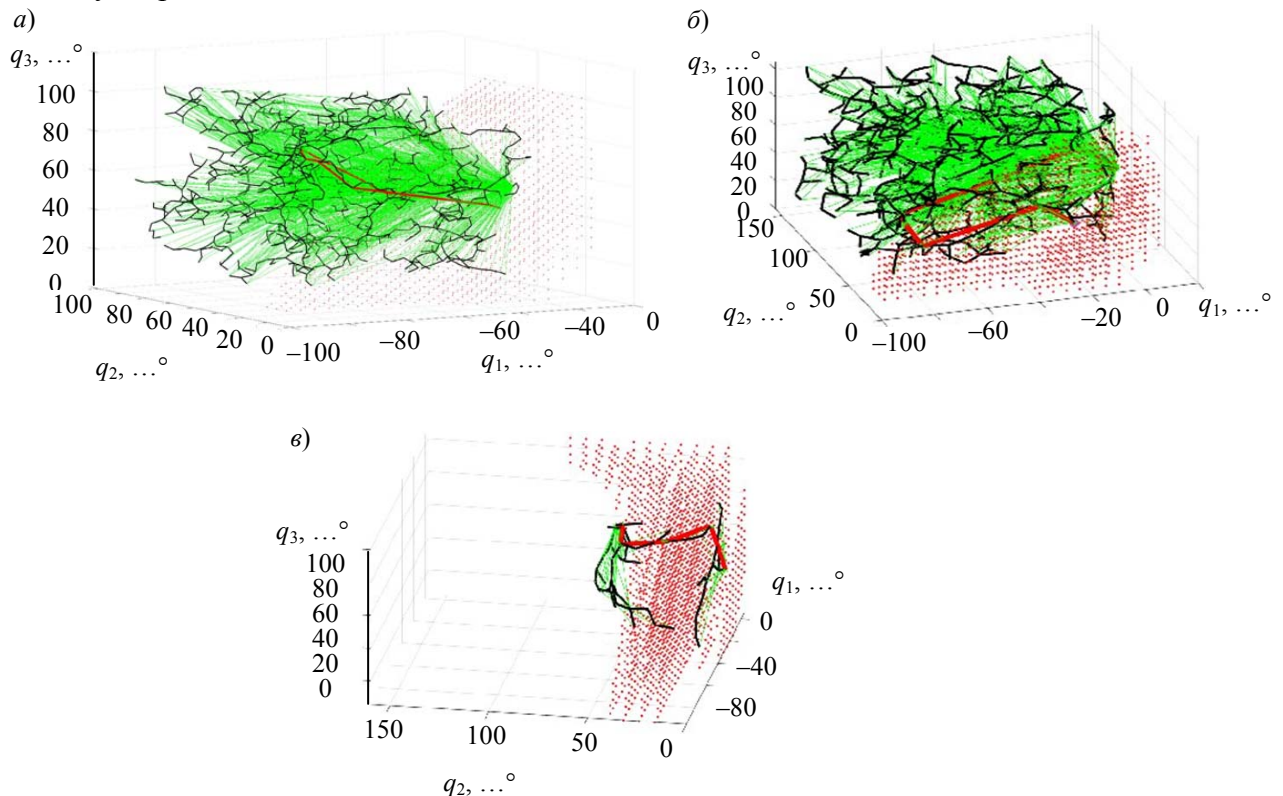


Рис. 9

Проанализировав представленные рисунки, можно сделать вывод, что алгоритм модифицированного интеллектуального двунаправленного быстроисследующего случайного дерева выполняет построение траектории без лишних ветвей в отличие от других алгоритмов, которые заполняют ими пространство. Показатели работы алгоритмов для задачи с горизонтальным (horisont obstacle — HO) и вертикальным (vertical obstacle — VO) препятствиями представлены в таблице.

Алгоритм	Среднее время планирования, с		Среднее число вершин		Количество успешных реализаций, %		Длина пути в конфигурационном пространстве, о.е.	
	HO	VO	HO	VO	HO	VO	HO	VO
RRT	938	703	3724	2343	50	50	640	691
Двунаправленное RRT	41	35	580	457	90	90	570	541
Интеллектуальное двунаправленное RRT	55	71	604	633	90	90	530	623
Модифицированное интеллектуальное двунаправленное RRT	13	5	147	53	100	100	500	411

Анализ таблицы показывает, что модифицированный алгоритм имеет лучшие показатели, чем любой другой из рассмотренных. Этот алгоритм позволяет быстрее находить траекторию и использует меньший объем памяти. Показатель стоимости (длины) пути означает, что для заданного числа случайных точек алгоритм находит более короткое решение.

Заключение. Представлен новый алгоритм планирования пути для робототехнических систем — модифицированное интеллектуальное двунаправленное быстроисследующее случайное дерево; используется подход, исключающий добавление лишних вершин к деревьям на основе оценивания их положения в пространстве планирования. Данный алгоритм превосходит предыдущие модификации по скорости обнаружения пути и эффективности использования памяти.

СПИСОК ЛИТЕРАТУРЫ

1. Лю В. Методы планирования пути в среде с препятствиями (обзор) // Математика и математическое моделирование. 2018. № 1. С. 15—58. DOI: 10.24108/mathm.0118.0000098.
2. LaValle S. M. Rapidly-exploring random trees : A new tool for path planning // The Annual Research Report. Iowa State University, 1998. October.
3. Karaman S., Frazzoli E. Sampling-based algorithms for optimal motion planning // Intern. Journal of Robotics Research. 2011. Vol. 30. P. 846—894.
4. Левин Б. П. Теоретические основы статической радиотехники. М.: Радио и связь, 1989. 656 с.
5. Jordan M., Perez A. Optimal bidirectional rapidly-exploring random trees // Tech. Rep. MITCSAIL-TR-2013-021. Massachusetts Institute of Technology, Cambridge, MA. 2013. August.
6. Qureshi A. H., Ayaz Y. Intelligent bidirectional rapidly-exploring random trees for optimal motion planning in complex cluttered environments // ArXiv. 2015. Vol. abs/1703.08944.
7. Tahir Z., Qureshi A. H., Ayaz Y., Nawaz R. Potentially guided bidirectionalized RRT* for fast optimal path planning in cluttered environments // Robotics Auton. Syst. 2018. Vol. 108. P. 13—27.
8. Gammell J., Srinivasa S., Barfoot T. Informed RRT*: Optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems. 2014. P. 2997—3004.
9. Burget F., Bennewitz M., Burgard W. Bi2RRT*: An efficient sampling-based path planning framework for task-constrained mobile manipulation // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). 2016. P. 3714—3721.

10. Kingston Z., Moll M., Kavraki L. Exploring implicit spaces for constrained sampling-based planning // Intern. Journal of Robotics Research. 2019. Vol. 38. P. 1151—1178.
11. Hauser K. Motion and Path Planning. Berlin, Heidelberg: Springer, 2020. P. 1—11.

Сведения об авторах

- | | |
|--------------------------------------|---|
| Илья Сергеевич Довгополик | — магистр; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; инженер; E-mail: isdovgopolik@itmo.ru |
| Кирилл Артемов | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; инженер-исследователь; E-mail: kaartemov@itmo.ru |
| Олег Игоревич Борисов | — канд. техн. наук; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; доцент; E-mail: borisov@itmo.ru |
| СейедХассан Забихифар | — канд. техн. наук; ПАО „Сбербанк“, лаборатория робототехники; инженер-разработчик; E-mail: zabikhifar.s@sberbank.ru |
| Александр Николаевич Семочкин | — канд. физ.-мат. наук, доцент; ПАО „Сбербанк“, лаборатория робототехники; гл. инженер-разработчик; E-mail: Semochkin.A.N@sberbank.ru |

Поступила в редакцию 28.12.21; одобрена после рецензирования 10.01.22; принята к публикации 18.01.22.

REFERENCES

1. Liu W. *Mathematics and Mathematical Modelling*, 2018, no. 01, pp. 15—58, DOI: 10.24108/mathm.0118.0000098.
2. LaValle S. *TR 98-11*, Computer Science Dept., Iowa State University, October 1998.
3. Karaman S. and Frazzoli E. *The International Journal of Robotics Research*, 2011, vol. 30, pp. 846—894.
4. Levin B.R. *Teoreticheskiye osnovy staticheskoy radiotekhniki* (Theoretical Foundations of Static Radio Engineering), Moscow, 1989, 656 p. (in Russ.)
5. Jordan M. and Perez A. *Tech. Rep. MITCSAIL-TR-2013-021*, Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, August 2013.
6. Qureshi A.H. and Ayaz Y. *ArXiv*, 2015, vol. abs/1703.08944.
7. Tahir Z., Qureshi A.H., Ayaz Y., and Nawaz R. *Robotics Auton. Syst.*, 2018, vol. 108, pp. 13—27.
8. Gammell J., Srinivasa S., and Barfoot T. *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 2997—3004.
9. Burget F., Bennewitz M., and Burgard W. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 3714—3721.
10. Kingston Z., Moll M., and Kavraki L. *The International Journal of Robotics Research*, 2019, vol. 38, pp. 1151—1178.
11. Hauser K. *Motion and Path Planning*, Berlin, Heidelberg, Springer, 2020, pp. 1—11.

Data on authors

- | | |
|-------------------------------|--|
| Ilya S. Dovgopolik | — MSc; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Engineer; E-mail: isdovgopolik@itmo.ru |
| Kirill Artemov | — Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Engineer-Researcher; E-mail: kaartemov@itmo.ru |
| Oleg I. Borisov | — PhD; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Associate Professor; E-mail: borisov@itmo.ru |
| Seyedhassan Zabihifar | — PhD; Sberbank, Robotics Laboratory; Engineer-Designer; E-mail: zabikhifar.s@sberbank.ru |
| Aleksandr N. Semochkin | — PhD, Associate Professor; Sberbank, Robotics Laboratory; Chief Engineer-Designer; E-mail: Semochkin.A.N@sberbank.ru |

Received 28.12.21; approved after reviewing 10.01.22; accepted for publication 18.01.22.

СЕГМЕНТАЦИЯ ОБЪЕКТОВ С ФУНКЦИЕЙ ДООБУЧЕНИЯ

И. Д. НЕНАХОВ^{1*}, К. АРТЕМОВ¹, С. ЗАБИХИФАР²,
А. Н. СЕМОЧКИН², С. А. КОЛЮБИН¹¹Университет ИТМО, Санкт-Петербург, Россия

*ivdmne@gmail.com

²Сбербанк, Москва, Россия

Аннотация. Рассматриваются способы расширения набора распознаваемых классов объектов для задачи их сегментирования, где необходимо построить маску объекта, а также узнать его класс. Для первой задачи использованы методы, не зависящие от классов предметов и являющиеся наиболее устойчивыми к изменениям формы; для второй задачи проанализированы методы, основанные на итеративном обучении (iterative learning), и методы метрического обучения (metric learning). Вторым подходом выбран в качестве основного, и для него протестированы различные архитектуры нейронных сетей. Проведена классификация объектов с использованием алгоритма k ближайших соседей. В качестве набора данных для обучения нейронной сети использован набор COIL-100, а после обученная модель протестирована на собственном наборе данных. Проведенные эксперименты показывают, что используемый метод позволяет обрабатывать 7-8 изображений в секунду на видеокарте GTX 1050 ti с 4 Гбайт видеопамати с точностью классификации в 99 %.

Ключевые слова: метрическое обучение, итеративное обучение, сегментация, классификация, сверточные нейронные сети, робототехника, распознавание изображений

Ссылка для цитирования: Ненахов И. Д., Артемов К., Забихифар С., Семочкин А. Н., Колюбин С. А. Сегментация объектов с функцией дообучения // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 194—203. DOI: 10.17586/0021-3454-2022-65-3-194-203.

OBJECTS SEGMENTATION WITH RETRAINING FUNCTION

I. D. Nenakhov^{1*}, K. Artemov¹, S. Zabihifar²,
A. N. Semochkin², S. A. Kolyubin¹¹ ITMO University, St. Petersburg, Russia Россия

*ivdmne@gmail.com

² Sberbank, Moscow, Russia

Abstract. Ways to expand the set of recognized object classes for the task of segmenting them, where it is necessary to build an object mask, as well as to find out its class, are considered. For the first task, methods that do not depend on the classes of subjects and are the most resistant to shape changes were used; for the second task, methods based on iterative learning and methods of metric learning are analyzed. The second approach is chosen as the main one, and various neural network architectures are tested for it. The classification of objects using the k nearest neighbors algorithm is carried out. The COIL-100 set is used as a data set for training a neural network, and after that the trained model was tested on its own data set. The experiments show that the method used allows processing 7-8 images per second on a GTX 1050 ti graphics card with 4 GB of video memory with a classification accuracy of 99%.

Keywords: metric learning, iterative learning, segmentation, classification, convolutional neural networks, robotics, image recognition

For citation: Nenakhov I. D., Artemov K., Zabihifar S., Semochkin A. N., Kolyubin S. A. Objects segmentation with retraining function. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 194—203 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-194-203.

Введение. Распознавание образов является одним из ключевых навыков человека для восприятия окружающего мира. Аналогично, роботу для имитации человека необходимо понимать, что его окружает. Как для человека основным органом чувств является зрение, так и

для робота основная часть информации о мире поступает через видеокамеры в виде двухмерных изображений. Однако недостаточно только получать данные с камер, их необходимо надлежащим образом обрабатывать.

Сегментация изображений — одна из основных задач в компьютерном зрении. Данная задача, как и большинство задач компьютерного зрения, сегодня решается с помощью алгоритмов, основанных на сверточных нейронных сетях (CNN), активное развитие которых отмечается в последнее десятилетие с достижением высоких показателей производительности компьютеров и, в частности видеокарт.

Компьютерное зрение в гуманоидной робототехнике имеет несколько отличительных характеристик и задач, отделяющих его от других областей компьютерного зрения, основанных на глубоком обучении. Автономный робот должен быть способен эффективно взаимодействовать со средой, предназначенной для человека, в частности, различать предметы в изменяющемся окружении и быть способным итеративно обучаться распознаванию объектов.

В настоящей статье рассматриваются различные методы итеративного обучения и исследуется методика метрического обучения, а также предлагается алгоритм, минимизирующий участие человека в процессе адаптации к новым классам объектов.

Распознавание образов. *Основные архитектуры сверточных нейронных сетей.* Операция свертки заключается в применении скользящего фильтра (как правило, малого размера) по изображению. Сама по себе свертка применима для нахождения примитивов: линий, углов, точек и т.д. Однако если использовать свертку не на исходном изображении, а на результате свертки этого изображения с каким-либо фильтром, то можно находить более сложные образы.

Первой удачной архитектурой сверточной нейронной сети является Alexnet [1], в 2013 г. превзошедшая классические алгоритмы в соревновании по классификации изображений на наборе данных ImageNet. Наряду со сверточными слоями в Alexnet применяются слой пулинга и полносвязный слой. Первый уменьшает размер карты признаков, деля тензор на сетку и выбирая среднее или максимальное его значение в элементах каждой ячейки. Полносвязный слой служит для классификации полученных высокоуровневых признаков.

Впоследствии нейронные сети стали все более глубокими. В очень глубоких нейронных сетях наблюдается проблема угасания градиента по мере его прохождения во время обновления весов. Нововведением архитектуры ResNet (2015) [2] является добавление обходных соединений. На определенных слоях создается соединение, пропускающее сигнал через несколько следующих слоев с последующей конкатенацией.

С увеличением глубины сети возрастает как вычислительная сложность, так и объем занимаемой памяти. Архитектура MobileNet [3] создавалась как компактная версия глубокой нейронной сети. В такой архитектуре для уменьшения пространственных размеров вектора используются блоки со свертками с шагом 2, а не пулинг, как применялось прежде. Также отличительной особенностью MobileNet является малый объем занимаемой памяти (16 Мбайт).

Сегментация изображений. Описанные выше архитектуры сверточных сетей применяются не только для классификации, но и служат остовом (backbone) моделей для остальных задач, таких как детектирование и сегментация. Рассмотрим подробнее задачу сегментации изображений.

При семантической сегментации ставится задача каждому пикселу входного изображения присвоить класс объекта, к которому он относится. При этом если два объекта одного класса соприкасаются, то на выходе модели нельзя определить их границу. Задача сегментации объектов (instance segmentation) решает эту проблему. Наряду с классом объекта необходимо определить, к какому объекту относится пиксел.

Наиболее успешное и популярное решение в задаче сегментации объектов — Mask R-CNN [4] является модификацией двухфазового (two-stage) детектора объектов Faster R-CNN [5]. На первом этапе изображение обрабатывается backbone-сетью. Далее берутся результаты с различных слоев этой сети и подаются в сеть для предсказания регионов (Region Proposal Network — RPN), которая выделяет области, где может содержаться объект. На этапе RoIAlign для каждого региона выделяется соответствующая ему карта признаков из карты признаков исходного изображения. Каждый найденный регион затем классифицируется полносвязными слоями и при наличии объекта уточняются границы региона и осуществляется поиск его маски. Эти три этапа производятся параллельно на признаках, полученных backbone-моделью. Структура модели Mask R-CNN приведена на рис. 1.

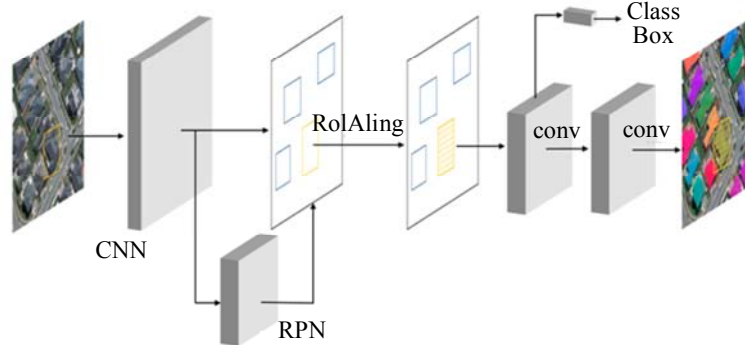


Рис. 1

Адаптация классификации объектов к расширению набора классов. Поставленную задачу можно сформулировать следующим образом. Пусть имеется исходное конечное множество объектов $N_x = \{n_{x1}, n_{x2}, \dots, n_{xn}\}$ и множество соответствующих объектам классов $N_y = \{n_{y1}, n_{y2}, \dots, n_{yk}\}$. Также имеется счетное множество $M_x = \{m_{x1}, m_{x2}, \dots\}$ дополнительных заранее не определенных объектов с соответствующими метками классов $M_y = \{m_{y1}, m_{y2}, \dots\}$, мощность которого также заранее не известна, но может быть оценена сверху как m^* . Множества N_y и M_y не пересекаются, $N \cap M = \emptyset$. Изначально алгоритм должен быть способен классифицировать объекты из множества классов $C_0 \subset N_y$. При этом в процессе работы алгоритма в явном виде в качестве новых обучающих данных могут выступать объекты, относящиеся к произвольному классу из множества M_y , требующие запуска процесса адаптации алгоритма. Следовательно, на каждой i -й итерации обучения алгоритм должен адаптироваться и расширять множество классифицируемых классов $C_i = C_{i-1} \cup \{m_{yi}\}$. В конечном счете требуется решить задачу классификации объектов из множества $N \cup M$ максимальной мощностью $(n+m^*)$ с точностью не ниже заданной:

$$A = \frac{T_p + T_n}{k + m^*},$$

где T_p — количество истинно положительных прогнозов, T_n — количество истинно отрицательных прогнозов, а $(k + m^*)$ — количество объектов в выборке.

В настоящее время существует две области машинного обучения, способные решить данную задачу: итеративное обучение (iterative learning) и метрическое обучение (metric learning).

Итеративное обучение. Процесс обучения состоит в определении оптимальных параметров модели. Для каждого обучающего примера оптимальные параметры различаются, и цель обучения — найти такие параметры, которые позволят получить удовлетворительное решение на всем множестве обучающих данных. Для этого при обучении данные выбираются в случайном порядке. Но когда данные поступают последовательно, оптимальные параметры, найденные в текущей итерации, при последующих итерациях сойдутся к другим значениям. Таким образом, модель потеряет способность распознавать объекты предшествующих итераций. Этот эффект называется катастрофическим забыванием.

В сверточных нейронных сетях разработано множество методов итеративного обучения, направленных на преодоление катастрофического забывания. Их можно разделить на три вида. Первые — регуляризационные подходы [6, 7], ограничивающие возможность изменения наиболее „важных“ весов; вторые — архитектурные подходы [8, 9], которые основываются на изменении архитектуры после каждого обученного класса; третьи [10] — подходы, никак не влияющие на параметры сети, но использующие часть данных от прежних классов для „напоминания“ в последующих итерациях обучения. Также существуют работы [11], в которых рассматривается комбинация этих методик. Однако применение методик итеративного обучения сопряжено с достаточно длительным временем обучения или необходимостью изменять модель при каждой итерации.

Метрическое обучение. Задача метрического обучения — построить такой экстрактор признаков, чтобы векторы, соответствующие данным одного класса, находились близко друг к другу в пространстве признаков, а векторы, соответствующие данным разных классов, — далеко друг от друга. При этом в экстракторе не содержится информации о том, какому именно классу принадлежат входные данные.

Экстрактор признаков можно представить в виде функции, которая отображает входное изображение на пространство признаков фиксированной размерности — $F: R^{whc} \rightarrow R^z$, где w, h, c — размерность изображений, z — размерность итогового пространства признаков. Таким образом, каждому изображению ставится в соответствие вектор $f \in R^z$.

Для обучения такого экстрактора разработана функция потерь триплета (triplet loss) [12]

$$\text{Loss} = \sum_{i=1}^{k+m} \left[\|f_i^a - f_i^p\|_2^2 - \|f_i^a - f_i^n\|_2^2 + \alpha \right],$$

принимая на вход тройку векторов: опорный f_i^a , положительный f_i^p и отрицательный f_i^n . Расстояние от опорного вектора до положительного должно быть меньше, чем от опорного до отрицательного как минимум на величину α . При этом за положительный принимается вектор, класс которого совпадает с классом опорного, а за отрицательный — вектор, класс которого не совпадает с классом опорного. Далее необходимо сохранить векторы, представляющие изображения набора данных, и впоследствии производить классификацию новых изображений с помощью алгоритма k ближайших соседей (k -Nearest Neighbors — kNN).

Метод k ближайших соседей — один из наиболее тривиальных алгоритмов машинного обучения для классификации и регрессии. Изначально сохраняются точки обучающей выборки в пространстве признаков. Затем задается параметр k , указывающий на количество соседей, среди которых будет проводиться классификация. Для каждого тестового объекта находят k ближайших соседей в пространстве признаков и присваивается класс, наиболее часто встречающийся среди найденных соседей.

Реализация процесса сегментации с возможностью дообучения. *Описание алгоритма.* Так как алгоритмы метрического обучения позволяют быстро добавлять новые объекты, то было решено использовать обученный по данной методике экстрактор признаков и алгоритм kNN для классификации объектов, а для сегментации — Mask R-CNN. При этом экстрактор признаков, сеть для прогноза регионов и голова сети для определения маски предмета (mask head) в модели Mask R-CNN при обучении не имеют информации о классах предметов, поэтому возможно использовать эти компоненты для сегментации изображения без классификации. Найденные объекты после удаления фона подаются в экстрактор признаков и с помощью алгоритма kNN производится их классификация.

Тестирование алгоритма. Для решения задачи классификации были протестированы несколько архитектур экстракторов признаков: ResNet50, MobileNetV3_small, MobileNetV3_large.

Архитектура ResNet была выбрана, так как является одной из современных (state-of-the-art) моделей из набора данных ImageNet. Архитектуры MobileNet — легковесные модели, в некоторых задачах склонные к лучшей обобщающей способности, чем громоздкие аналоги.

Также имеет значение размерность выходного вектора — чем он больше, тем больший объем информации можно закодировать им и тем точнее будет классификация. Однако вместе с ростом размерности увеличивается размер файла, в котором хранятся ранее полученные признаки, а также увеличивается объем оперативной памяти во время работы программы.

Каждый из используемых экстракторов признаков формирует вектор заданной размерности. Для регулирования размерности после экстрактора признаков был добавлен один полносвязный слой с количеством нейронов, равным желаемой размерности вектора. В данном эксперименте были протестированы векторы следующих размерностей: 16, 32, 64 и 128 элементов.

Для тестирования был выбран набор данных COIL-100 [13]. Примеры объектов показаны на рис. 2. Набор данных содержит 100 объектов по 72 изображения на каждый. Для обучения были выбраны первые 50 объектов из набора данных. Обучение проводилось на протяжении 10 эпох. Для тестирования использовались оставшиеся 50 объектов. Работа алгоритмов оценивалась с помощью видеокарты Nvidia RTX 3080. Результаты обучения различных архитектур приведены в таблице.



Рис. 2

Архитектура	Размерность вектора	Точность, %	Время обучения, с	Время обработки одного изображения, мс
Resnet50	16	97,9	293	10
MobileNetV3 small	16	96,9	303	8
MobileNetV3 large	16	98,1	321	11
Resnet50	32	98,8	409	10
MobileNetV3 small	32	98,7	307	8
MobileNetV3 large	32	99,1	321	10
Resnet50	64	99,1	413	10
MobileNetV3 small	64	99,2	309	8
MobileNetV3 large	64	99,5	324	10
Resnet50	128	99,4	418	10
MobileNetV3 small	128	99,2	312	8
MobileNetV3 large	128	99,5	327	10
Resnet50	256	99,6	426	10
MobileNetV3 small	256	99,3	319	8
MobileNetV3 large	256	99,6	333	10
Resnet50	512	99,4	435	10
MobileNetV3 small	512	99,4	331	9
MobileNetV3 large	512	99,6	347	10

Как видно из таблицы, более легковесные архитектуры немного превосходят ResNet по точности и значительно по скорости обучения и времени обработки. Архитектура ResNet по сравнению с MobileNet имеет большее число параметров. Следовательно, модель может „запомнить“ больше обучающих примеров вместо выведения закономерностей, тем самым более склонна к переобучению. Этим объясняется несколько меньшая точность ResNet по сравнению с MobileNet.

Размерность пространства признаков оказывает влияние как на точность и скорость классификации, так и на объем памяти для хранения признаков. При увеличении размерности модель имеет больше способов разнести в пространстве точки разных классов, что повышает точность. В то же время расчет расстояния между двумя точками в таком пространстве будет занимать больше времени. Также кратно возрастает объем памяти для хранения векторов признаков.

Проецирование векторов признаков из 128-мерного пространства на плоскость с помощью алгоритма снижения размерности Umap представлено на рис. 3. Визуализация отображает точки, соответствующие векторам тестовой части (50 объектов) набора данных. Точки, соответствующие одному классу, сгруппированы друг с другом, тогда как точки разных классов находятся на значительном расстоянии друг от друга практически для всех классов объектов.

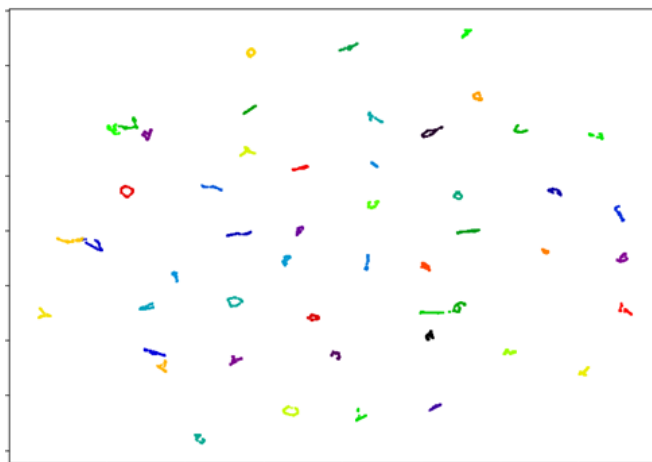


Рис. 3

Тестирование алгоритма на реальных объектах. После тестирования алгоритма на заранее подготовленном наборе данных был проведен эксперимент с использованием реальных объектов: кружка (cup), коробка (box), кубик (cube), компьютерная мышь (mouse) и антисептик (cleaner).

Для каждого объекта было снято видео длительностью 15 с, показывающее объект с разных сторон. Далее изображения были разделены на обучающую и тестовую части в соотношении 70 на 30. Обучающие 70 % изображений были использованы для получения векторов в пространстве признаков и обучения алгоритма k ближайших соседей, а оставшиеся 30 % — для оценки классификации. Точность обученной системы составила 98 %.

В то время как обученные классы объектов модель определяет с удовлетворительной точностью, в базовой конфигурации она часто распознает незнакомые предметы как обученные, т.е. имеют место ложно-положительные срабатывания. В этом случае необходимо фильтровать результаты классификации в соответствии с процентным соотношением количества объектов спрогнозированного класса среди k ближайших соседей, чтобы исключить объекты, находящиеся между кластерами известных классов. При тестировании алгоритма был выбран порог в 80 %. Также могут появиться объекты, все ближайшие соседи которых принадлежат к одному классу, но расстояние до первого из них велико. Тогда считается, что объект не принадлежит ни к одному известному классу. При тестировании был выбран порог для

косинусного расстояния в 0,2. Примеры сегментации для объектов из набора данных представлены на рис. 4.

Результаты 2D визуализации обучающих данных в пространстве признаков представлены на рис. 5. Можно заметить, что за редким исключением репрезентации различных объектов явно разделены. Используя видеокарту Nvidia GTX 1050ti с 4 Гбайт видеопамяти, алгоритм способен обрабатывать 7-8 изображений в секунду.

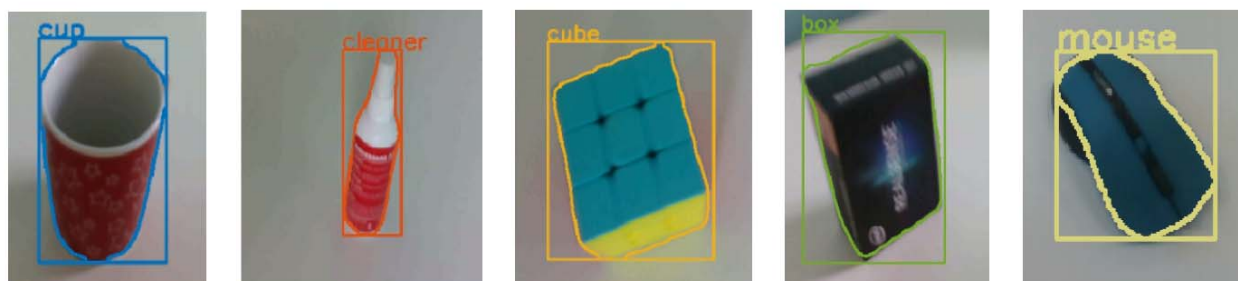


Рис. 4

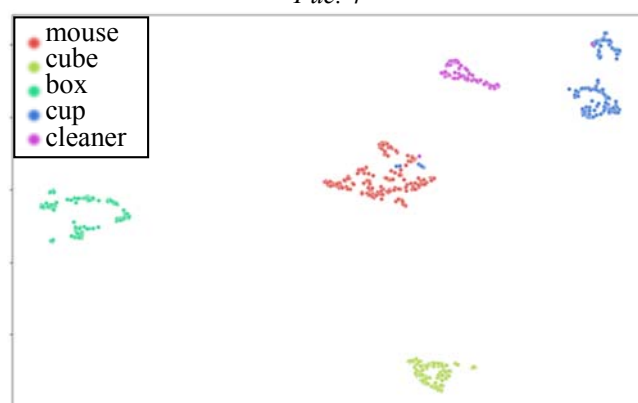


Рис. 5

При реализации данного алгоритма были использованы фреймворки Detectron2 [14] и PyTorch [15] и библиотека Pytorch Metric learning [16].

Алгоритм автоматического дообучения. Ключевой особенностью робототехники является минимизация участия человека в рабочем процессе. При поступлении объектов новых классов робот должен быть способен распознавать эти объекты как не принадлежащие ни к одному известному классу и запускать процесс адаптации к ним. При наличии манипулятора и видеокамеры, робот может осуществить это по алгоритму, структурная схема которого приведена на рис. 6.

Как только появляется объект, вектор которого лежит дальше установленного порога d_p , запускается процесс проверки: определяются координаты центра объекта в 3D-пространстве и выполняются N снимков этого объекта с разных ракурсов. Если среди векторов собранных изображений среднее минимальных расстояний $\bar{d}_{i\min}$ меньше d_p , то объекту присваивается наиболее частый класс и собранные векторы добавляются в обучающую выборку. В противном случае запускается процесс дообучения: производится R снимков объекта, определяются их векторы признаков, которые затем добавляются в базу данных с меткой класса $j+1$, где j — текущее количество обученных классов. Из эвристических соображений следует выбирать параметры N и R таким образом, чтобы $N < R$, так как для уточнения классификации необходимо, как правило, меньше данных, чем на дообучение.

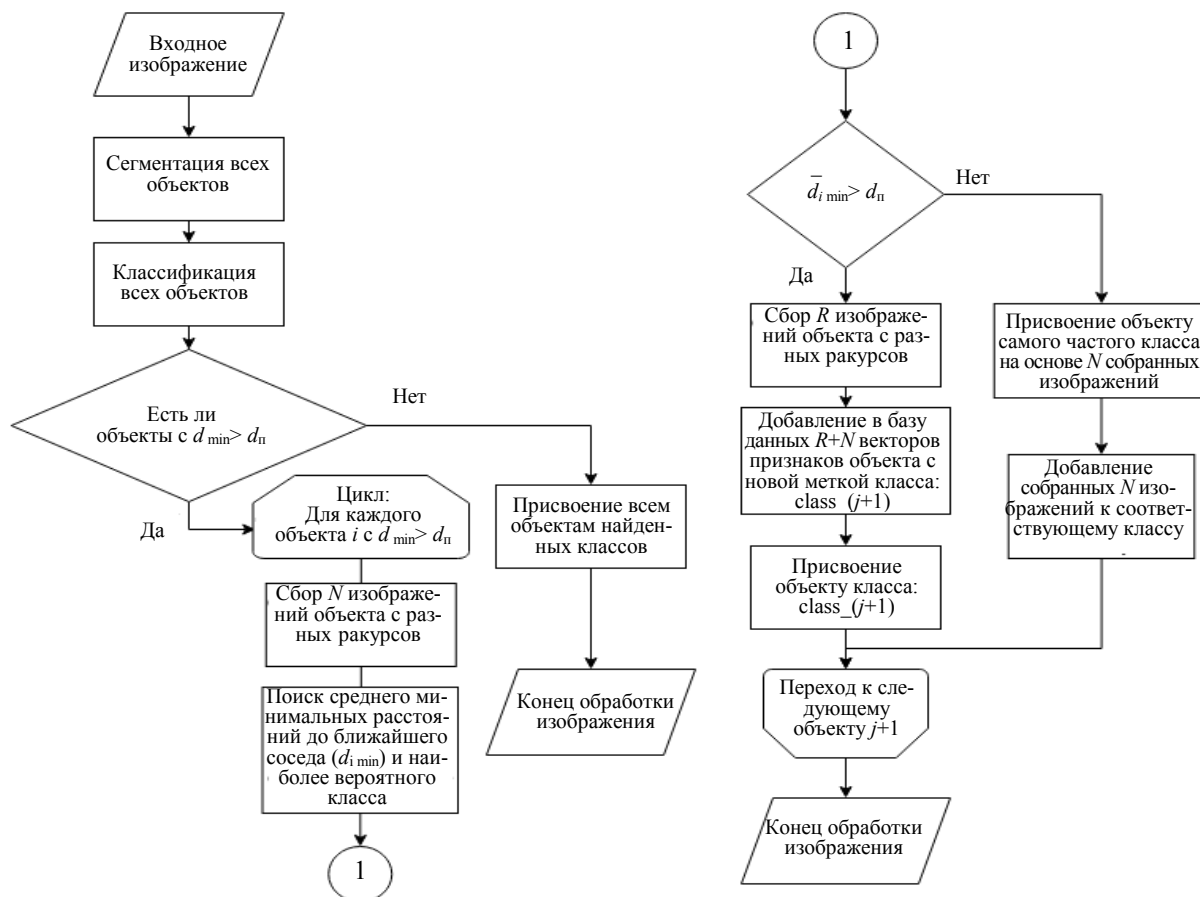


Рис. 6

Заключение. Исследована возможность применения предобученной модели для сегментации Mask R-CNN и обучения метрики для решения задачи расширения распознаваемых классов. В отличие от методики итеративного обучения данный метод не подвержен влиянию катастрофического забывания, а также не требует времени на дообучение.

Перспективные исследования связаны с соединением экстрактора признаков в Mask R-CNN и классификатора для ускорения работы алгоритма, а также с уменьшением количества сохраняемых точек в пространстве признаков и удалением выбросов в данных. Для улучшения качества сегментации планируется обучить Mask R-CNN на наборе данных, приспособленном для задач мобильной робототехники.

СПИСОК ЛИТЕРАТУРЫ

1. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet Classification with Deep Convolutional Neural Networks // Advances in Neural Information Processing Systems / Ed.: F. Pereira, C. J. C. Burges, L. Bottou, K. Q. Weinberger. Curran Associates Inc., 2012. Vol. 25. P. 1097—1105.
2. Deep Residual Learning for Image Recognition / Kaiming He, Xi-Anyu Zhang, Shaoqing Ren, Jian Sun // CoRR. 2015. Vol. ab-s/1512.03385. [Электронный ресурс]: <<http://arxiv.org/abs/1512.03385>>.
3. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications / A. G. Howard, Menglong Zhu, Bo Chen et al. // CoRR. 2017. Vol. abs/1704.04861. [Электронный ресурс]: <<http://arxiv.org/abs/1704.04861>>.
4. Mask R-CNN / Kaiming He, G. Gkioxari, P. Doll'ar, R. B. Girshick // CoRR. 2017. Vol. abs/1703.06870. [Электронный ресурс]: <<http://arxiv.org/abs/1703.06870>>.

5. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks / *Shaoqing Ren, Kaiming He, R. Girshick, Jian Sun* // *Advances in Neural Information Processing Systems* / Ed.: C. Cortes, N. Lawrence, D. Lee et al. Curran Associates Inc., 2015. Vol. 28. P. 91—99. [Электронный ресурс]: <<https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf>>.
6. Kirkpatrick J., Pascanu R., Rabinowitz N. et al. Overcoming catastrophic forgetting in neural networks // *Proc. of the National Academy of Sciences*. 2017. N 114(13). P. 3521—3526.
7. Zenke F., Poole B., Ganguli S. Continual Learning Through Synaptic Intelligence // *Proc. of the 34th Intern. Conf. on Machine Learning*. Sydney, Australia. 2017. Vol. 70. P. 3987—3995.
8. Lomonaco V., Maltoni D. CORE50: A New Dataset and Benchmark for Continuous Object Recognition // *Proc. of the 1st Annual Conf. on Robot Learning*. PMLR. 2017. Vol. 78. P. 17—26.
9. Progressive Neural Networks / A. A. Rusu, N. C. Rabinowitz, G. Desjardins, H. Soyer, J. Kirkpatrick, K. Kavukcuoglu, R. Pascanu, R. Hadsell. // *arXiv preprint arXiv:1606.04671*, 2016.
10. Hayes T. L., Cahill N. D., Kanan C. Memory Efficient Experience Replay for Streaming Learning // *arXiv preprint arXiv:1809.05922*, 2018.
11. Rebuffi S. A., Kolesnikov A., Sperl G., Lampert C. H. iCaRL: Incremental Classifier and Representation Learning // *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii. 2017.
12. Hoffer E., Nir A. Deep metric learning using Triplet network // *Intern. Workshop on Similarity-Based Pattern Recognition*. Springer, Cham, 2015.
13. Columbia Object Image Library (COIL-100) / S. A. Nene, S. K. Nayar, H. Murase. // *Tech. Report CUCS-006-96*. 1996. February.
14. Wu Yuxin. Detectron2, 2019. [Электронный ресурс]: <<https://github.com/facebookresearch/detectron2>>.
15. Pytorch: An Imperative Style, High-Performance Deep Learning Library / A. Paszke, S. Gross, F. Massa et al. // *Advances in Neural Information Processing Systems 32* / Ed.: H. Wallach, H. Larochelle, A. Beygelzimer et al. Curran Associates Inc., 2019. P. 8024—8035.
16. Musgrave K. , Belongie S., Lim S.-N. Pytorch metric learning // *arXiv preprint, arXiv:2008.09164*, 2020.

Сведения об авторах

- | | |
|--------------------------------------|---|
| Иван Дмитриевич Ненахов | — студент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: ivdmne@gmail.com |
| Кирилл Артемов | — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; инженер-исследователь; E-mail: kaartemov@itmo.ru |
| СейедХассан Забихифар | — канд. техн. наук; ПАО „Сбербанк“, лаборатория робототехники; инженер-разработчик; E-mail: zabikhifar.s@sberbank.ru |
| Александр Николаевич Семочкин | — канд. физ.-мат. наук, доцент; ПАО „Сбербанк“, лаборатория робототехники; гл. инженер-разработчик; E-mail: Semochkin.A.N@sberbank.ru |
| Сергей Алексеевич Колюбин | — д-р техн. наук, доцент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; вед. научный сотрудник; E-mail: s.kolyubin@itmo.ru |

Поступила в редакцию 28.12.21; одобрена после рецензирования 11.01.22; принята к публикации 18.01.22.

REFERENCES

1. Krizhevsky A., Sutskever I., Hinton G.E. *Advances in Neural Information Processing Systems*, 2012, vol. 25, pp. 1097—1105, <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.
2. He K., Zhang X., Ren Sh., Sun J. *CoRR*, 2015, vol. abs/1512.03385, <http://arxiv.org/abs/1512.03385>.
3. Howard A.G., Zhu M., Chen B. et al. *CoRR*, 2017, vol. abs/1704.04861, <http://arxiv.org/abs/1704.04861>.
4. He K., Gkioxari G., Dollár P., Girshick R.B. *CoRR*, 2017, vol. abs/1703.06870, <http://arxiv.org/abs/1703.06870>.
5. Ren Sh., He K., Girshick R., Sun J. *Advances in Neural Information Processing Systems*, 2015, vol. 28, pp. 91—99, <https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf>.

6. Kirkpatrick J., Pascanu R., Rabinowitz N., Veness J., Desjardins G., Rusu A.A., Milan K., Quan J., Ramalho T., Grabska-Barwinska A., Hassabis D., Clopath C., Kumaran D., and Hadsell R. *Overcoming catastrophic forgetting in neural networks*, 2017, no. 13(114), pp. 3521–3526.
7. Zenke F., Poole B., and Ganguli S. *Proceedings of the 34th International Conference on Machine Learning*, Sydney, Australia, 2017, vol. 70, pp. 3987–3995.
8. Lomonaco V. and Maltoni D. *Proceedings of the 1st Annual Conference on Robot Learning*, PMLR, 2017, vol. 78, *Proceedings of Machine Learning Research*, pp. 17–26.
9. Rusu A.A., Rabinowitz N.C., Desjardins G., Soyer H., Kirkpatrick J., Kavukcuoglu K., Pascanu R., and Hadsell R. *arXiv preprint*, arXiv:1606.04671, 2016.
10. Hayes T.L., Cahill N.D., and Kanan Ch. *arXiv preprint*, arXiv:1809.05922, 2018.
11. Rebuffi S.-a., Kolesnikov A., Sperl G., and Lampert Ch.H. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, Hawaii, 2017.
12. Hoffer E. and Nir A. *International Workshop on Similarity-Based Pattern Recognition*, Springer, Cham, 2015.
13. Nene S.A., Nayar S.K., and Murase H. *Columbia Object Image Library (COIL-100)*, Technical Report CUCS-006-96, February 1996.
14. Wu Y. Detectron2, <https://github.com/facebookresearch/detectron2>, 2019.
15. Paszke A., Gross S., Massa F. et al. *Advances in Neural Information Processing Systems 32*, Curran Associates, Inc., 2019, pp. 8024–8035.
16. Musgrave K., Belongie S., and Lim S.-N. *arXiv preprint*, arXiv:2008.09164, 2020.

Data on authors

Ivan D. Nenakhov	— Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; E-mail: ivdmne@gmail.com
Kirill Artemov	— Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Engineer-Researcher; E-mail: kaartemov@itmo.ru
Seyedhassan Zabihifar	— PhD; Sberbank, Robotics Laboratory; Engineer-Designer; E-mail: zabikhifar.s@sberbank.ru
Aleksandr N. Semochkin	— PhD, Associate Professor; Sberbank, Robotics Laboratory; Chief Engineer-Designer; E-mail: Semochkin.A.N@sberbank.ru
Sergey A. Kolyubin	— Dr. Sci., Associate Professor; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronics and Energy-Efficient Robotics; Leading Researcher; E-mail: s.kolyubin@itmo.ru

Received 28.12.21; approved after reviewing 11.01.22; accepted for publication 18.01.22.

МЕТОДЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СИСТЕМАХ НАВИГАЦИИ МОБИЛЬНЫХ РОБОТОВ

ARTIFICIAL INTELLIGENCE METHODS IN MOBILE ROBOT NAVIGATION SYSTEMS

УДК 004.896 + 007.52 + 004.93'1
DOI: 10.17586/0021-3454-2022-65-3-204-217

МЕТРИКО-СЕМАНТИЧЕСКОЕ КАРТИРОВАНИЕ НА ОСНОВЕ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СИСТЕМ АВТОНОМНОЙ НАВИГАЦИИ В ПОМЕЩЕНИЯХ

А. Р. БЕРКАЕВ^{1*}, М. МОХРАТ¹, А. М. БУРКОВ², С. А. КОЛЮБИН¹

¹Университет ИТМО, Санкт-Петербург, Россия

*berkaevamiran@mail.ru

²Сбербанк, Москва, Россия

Аннотация. Представлены результаты исследования, направленного на разработку интеллектуальной автономной навигационной системы для складской и офисной логистики с использованием глубоких нейронных сетей. Проанализированы современные и наиболее универсальные средства для получения карт глубин и семантической сегментации данных на изображениях в различных средах. Проведен сравнительный анализ карт глубин, формируемых RGB-D-камерой, а также с помощью нейросетевых алгоритмов и модифицированного алгоритма Хиршмюллера. Результаты тестирования, проведенного на специально подготовленном наборе данных, снятых в офисном пространстве, демонстрируют, что предложенное решение превосходит альтернативные по точности и позволяет сократить затраты вычислительных ресурсов.

Ключевые слова: семантическая сегментация, карты глубин, одновременная локализация и картирование, метрико-семантическая карта, мобильный робот, логистика, глубокие нейронные сети, оценка глубины, интеллектуальные системы

Ссылка для цитирования: Беркаев А. Р., Мохрат М., Бурков А. М., Колюбин С. А. Метрико-семантическое картирование на основе глубоких нейронных сетей для систем автономной навигации в помещениях // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 204—217. DOI: 10.17586/0021-3454-2022-65-3-204-217.

METRIC-SEMANTIC MAPPING BASED ON DEEP NEURAL NETWORKS FOR SYSTEMS OF INDOOR AUTONOMOUS NAVIGATION

A. R. Berkaev^{1*}, M. Mohrat¹, A. M. Burkov², S. A. Kolyubin¹

¹ITMO University, St. Petersburg, Russia

*berkaevamiran@mail.ru

²Sberbank, Moscow, Russia

Abstract. Results of a study aimed at developing an intelligent autonomous navigation system for warehouse and office logistics using deep neural networks, are presented. The modern and most versatile tools for depth maps retrieval and semantic data segmentation on images in different environments are analyzed. A comparison of depth maps retrieved hardware from RGB-D camera, neural network algorithms, and a modified Hirschmuller algorithm is carried out. Results of testing performed with a specially prepared dataset shot in an office space, including many complex objects such as glass, mirrors, and multiple light sources demonstrate that the proposed solution outperforms the alternatives in accuracy and uses fewer computational resources in the process.

Keywords: segmentation, depth maps, simultaneous localization and mapping, metric-semantic map, mobile robot, logistics, deep neural network, depth estimation, intelligent systems

For citation: Berkaev A. R., Mohrat M., Burkov A. M., Kolyubin S. A. Metric-semantic mapping based on deep neural networks for systems of indoor autonomous navigation. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 204—217 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-204-217.

Введение. Совершенствование методов одновременной локализации и картирования имеет важное значение для функционирования мобильных роботов, способных автономно перемещаться в помещениях различного назначения. Многие известные методы разработаны применительно к условиям конкретного пространства, однако все чаще возникает потребность в универсальных подходах. Это связано с появлением новых сценариев для мобильных роботов, например в сфере логистики: роботу-курьеру необходимо забрать посылку со склада, преодолеть определенное расстояние во внешнем пространстве, а затем снова заехать в помещение для доставки адресату. Поэтому настоящая работа сосредоточена на выявлении универсальных и эффективных методов метрико-семантической одновременной локализации и картирования (Simultaneous Localization and Mapping — SLAM) на основе глубоких нейронных сетей (DNN).

Для создания максимально эффективной метрико-семантической SLAM-системы были протестированы различные методы генерации облаков точек. Одним из наиболее частых решений является использование специальных датчиков с активным зрением, таких как лидар, или датчиков со структурированным светом, таких как RGB-D-камеры.

Использование камер структурированного света — популярное решение для работы в помещении или на открытом воздухе. Однако, как и все датчики, они имеют не только преимущества, но и недостатки. Одно из преимуществ — возможность использования вне помещений, даже при сильном влиянии солнечного света. С другой стороны, наличие высокого уровня шумов обуславливает невысокое качество информации о пространственной глубине: на изображениях с объектами, находящимися на значительном расстоянии от камеры, теряется большая часть деталей или наблюдается сильное искажение. Такие камеры используются для расчета глубины проецирования инфракрасных лучей на поверхность объектов. В связи с этим облака точек, полученные для прозрачных объектов, недостаточно информативны. Кроме того, как правило, получаемые таким образом облака точек имеют невысокий коэффициент заполнения, т.е. процент вокселей, имеющих ненулевые значения глубины.

Следовательно, представляется актуальной разработка альтернативных решений для создания высококачественных семантически аннотированных 3D-карт окружающего пространства с использованием возможностей искусственного интеллекта.

Основное значение и особенность настоящей статьи заключаются в системном анализе существующих подходов и разработке интегрированной метрико-семантической SLAM-системы, использующей современные методы на основе DNN для программного восстановления карт глубин и семантической сегментации и аннотирования получаемых трехмерных карт. Предлагаемое решение базируется на алгоритмах, обеспечивающих наилучшие показатели по качеству (точности и плотности заполнения облаков точек) и скорости работы.

Обзор известных подходов. Построение семантически аннотированной карты — это современная задача в области мобильной робототехники, решение которой позволяет получить 2D- или 3D-карты окружающего пространства, представляющие информацию не только о пустых и заполненных пространствах, но также и о характере окружающих предметов. В метрико-семантических SLAM-системах необходимо добиться уверенного детектирования элементов пространства (полов, потолков, стен, базовой мебели, источников света в помещении) и, конечно, динамических объектов (людей, автомобилей, других роботов и т. д.). Также

важно знать точное местоположение всех объектов и их геометрию. Возникающие в данном случае сложности обусловлены не только шумом датчиков, искажающим объекты на изображении, но и проблемами, связанными с динамичностью самой системы, которая к тому же перемещается в пространстве, содержащем и статические, и динамические объекты. Кроме того, добавляются погрешности измерений, которые накапливаются со временем, что приводит к отклонениям оцениваемой траектории робота от реальной, а также искажает точность формируемых 2D- или 3D-карт окружающего пространства.

Восстановление карт глубин по изображению на основе DNN. Извлечение информации о глубине сцены важно для восприятия окружающей среды и оценки ее состояния. Качественная оценка расстояния до окружающих объектов особенно необходима для работы автономных мобильных систем. С ускоренным развитием глубоких нейронных сетей их использование для данной задачи показало обнадеживающие результаты с точки зрения точности при обучении end-to-end (e2e) способом, требующим в большинстве случаев только одного изображения в качестве входных данных [1].

Расчет расстояния до объектов по данным монокулярной камеры затруднен из-за внутренней неоднозначности, вызванной шумами и вибрациями на изображении. Поэтому вначале для решения этой проблемы были исследованы методы, основанные на статистических признаках.

В ходе исследования было протестировано несколько алгоритмов в целях выбора лучших с точки зрения точности и работы в реальном времени с частотой выше 10 кадров/с. Такие алгоритмы, как DiverseDepth [2] и MiDaS [3], показали впечатляющие результаты в плане точности, поскольку обучение моделей производилось на большом и разнообразном количестве наборов данных. DELTAS [4] и DeepVideoMVS [5] представляют алгоритмы оценки глубины сцены, основанные на многоракурсной технике, где используются два или три последовательных кадра, а также геометрическая информация о положении и ориентации камеры между этими кадрами для создания плотного облака точек, связанного с оценкой метрических значений глубины. Алгоритмы FastDepth [6], PydNet [7] и FCNN_Node [8] показали наилучшую производительность с точки зрения выполнения в реальном времени. С использованием FastDepth и FCNN_Node были реализованы эффективные сетевые архитектуры с программными решениями во время выполнения для снижения сложности моделей, тогда как в PydNet вычисления производятся только на центральном процессоре.

Другим способом косвенного получения карт глубин является реализация бинокулярных алгоритмов оценки глубины на основе DNN, принимающих на вход пары стереоизображений RGB для создания карт смещений, процесс преобразования которых в карту глубин принципиально прост. В работе [9] представлено множество решений с различными архитектурами — от моделей, обеспечивающих производительность в режиме реального времени, до моделей, дающих высокую точность оценки смещений. Более того, существуют алгоритм MADNet [10], производительность которого может быть адаптирована во время работы, и AANet [11], имеющий архитектуру без слоя 3D-свертки для быстрого оценивания глубины.

Семантическое аннотирование. Семантическая сегментация изображений — это разделение изображений на сегменты в зависимости от того, какому объекту принадлежит каждый пиксел.

В последние годы исследования по семантической сегментации изображений значительно активизировались благодаря появлению различных архитектур сверточных нейросетей (CNN). Точность прогнозирования постоянно повышается, и в целом можно заметить, что проблема сегментации изображения на два-три класса уже решена [12]. Кроме того, отмечается достаточно высокая скорость работы DNN-алгоритмов, обученных на конкретных наборах данных или сценах, например DeepLabV3+ [13] и U-Net [14]. В ряде работ описываются классы универсальных архитектур, таких как HRNet [15] и SegNet [16], которые имеют множество

модификаций для обеспечения лучшей сегментации данных и более быстрого прогнозирования. Другие решения считаются приблизительно универсальными, например ESANet [17], построенная на архитектуре ResNet [18], которая первоначально предназначалась для внутренних сцен, но хорошо работает и для сцен вне помещений.

Семантическое картирование. Благодаря заметному увеличению доступных вычислительных ресурсов современных встраиваемых систем появилась возможность построения семантически аннотированных 3D-карт в режиме реального времени. Такие карты меняют уровень понимания роботом окружающей обстановки. Классические карты дают информацию об окружающем пространстве в виде препятствий и свободных пространств. В свою очередь, метрико-семантические карты содержат множество дополнительной информации, которая позволяет роботу понять, какие именно объекты находятся вокруг него и как далеко они расположены. Эта информация используется для перехода на более интеллектуальный уровень локализации.

В настоящей работе рассматриваются фреймворки, использующие воксельное представление карт или усеченное поле знаковых расстояний (Truncated Signed Distance Field — TSDF). Преимуществами таких карт являются визуальная составляющая, а также возможность преобразования в евклидовы поля знаковых расстояний (Euclidean Signed Distance Field — ESDF), используемые для планирования траекторий. Данные определения и их преимущества наиболее полно объясняются в работе по Voxblox [19].

В ходе исследования были изучены такие современные программные пакеты, как Kimera [20], DA-RNN [21] и Voxblox++ [22], являющиеся фреймворками семантического картирования на основе DNN. Voxblox++ базируется на экземплярно-семантической сегментации с использованием архитектуры Mask R-CNN, тогда как DA-RNN построен на собственной архитектуре, основанной на полностью сверточной сети (Fully Convolutional Network — FCN). Также интересной представляется современная работа Voxgraph [23], фреймворк которой позволяет строить глобально согласованную карту и корректировать ее во время построения в соответствии с замыканиями цикла, но при этом здесь отсутствуют семантические данные.

Для исследования был выбран программный пакет Kimera, отличающийся от других следующими характеристиками:

- метрико-семантическая SLAM-система основана на визуально-инерциальной одометрии (VIO);
- использование вычислительных ресурсов исключительно центрального процессора, что важно с точки зрения энергопотребления для мобильной робототехники;
- наличие обширного и гибкого функционала: например, имеется возможность исключения из построения классов объектов, в частности динамических объектов, что позволяет избежать появления артефактов на карте в виде треков движущихся объектов.

С учетом способности эффективно работать со сценами различной сложности, а также уровня производительности, достаточного для работы в режиме реального времени, были выбраны алгоритмы MiDaS и MADNet в качестве лучших для получения карт глубин, формируемых монокулярной и бинокулярной камерами соответственно. Кроме того, на основе сопоставительного анализа WildDash2 [24] выбран пакет MSeg [25] в качестве решения, имеющего универсальную таксономию для большого количества наборов данных и удовлетворительно работающего в различных сценариях. MSeg-semantic построен на базе HRNet-W48 и обучен сегментировать до 194 классов.

Система метрико-семантического картирования. Структура системы. Схема предлагаемой метрико-семантической SLAM-системы представлена на рис. 1 (здесь GPU — графический процессор). На вход системы поступают изображения, формируемые левой и правой видеокамерами, а также синхронизированные со снимками по времени данные IMU-сенсора. Входными данными служат расстояния до некоторых характерных точек, разреженно расположенных в поле зрения камер. На выходе системы формируется плотная,

точная, сглаженная, семантически аннотированная 3D-карта окружающего пространства, на которой отмечены заранее задаваемые ключевые классы объектов (пол, стены, дверные и оконные проемы, мебель и т.п.).

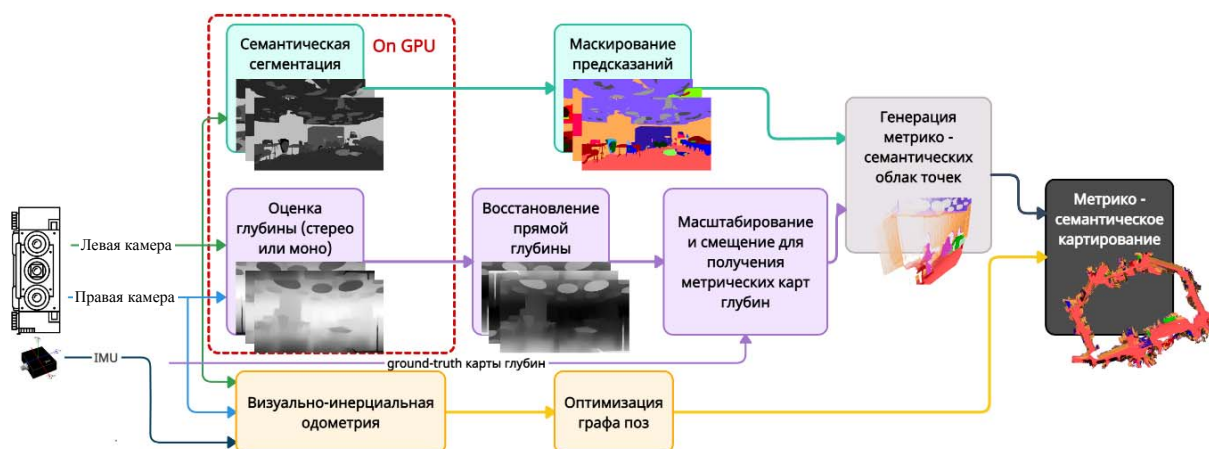


Рис. 1

Выделяются три основных параллельных потока, каждый из которых отвечает за генерацию данных, необходимых для построения метрико-семантической 3D-карты окружающего пространства. *Первый поток* отвечает за генерацию семантически аннотированных изображений и последующее маскирование прогнозируемых сцен в оттенках серого до более наглядных и понятных человеку цветных изображений. *Второй поток* — это расчет плотных RGB-D облаков точек по данным монокулярных или бинокулярных изображений, этот поток отвечает за генерацию 3D-карт: здесь сначала для каждого кадра рассчитывается массив обратных (относительных) глубин, а затем с использованием разреженных прямых измерений расстояния или внутренних параметров камер (в случае бинокулярных изображений) происходит масштабирование полученных обратных глубин и строится карта абсолютных метрических глубин. Далее данные первого и второго потоков объединяются для построения семантического облака точек. *Третий поток* отвечает за работу модуля визуально-инерциальной одометрии (VIO) и модуля оптимизации графа поз (PGO), которые необходимы для получения надежных данных о текущей позиции и перемещениях камеры.

Генерация плотных облаков точек. Для получения облаков точек абсолютной метрической глубины используются обратная глубина, генерируемая монокулярными „оценщиками“ глубины на основе DNN, и глубина, формируемая RGB-D-камерой. Процесс реализации был проведен на наборе данных, который никогда раньше не подавался на вход обученных DNN-алгоритмов MiDaS и MADNet, а схема представлена на рис. 1 вторым потоком.

Монокулярная оценка глубины изображения на основе DNN. Для получения высокого уровня оценки глубины монокулярного изображения были использованы нейросетевая модель MiDaS и 10 наборов данных, имеющих различные характеристики масштаба и сдвига. Исходя из такого большого количества источников, необходимо было предложить функцию потерь, объединяющую стратегии обучения данных.

Процесс генерации обратных карт глубин, являющихся выводом алгоритма MiDaS, происходит параллельно с сохранением полученных облаков точек в виде файлов в формате *rfm*, которые содержат исходные значения, предсказанные моделью MiDaS.

В зависимости от набора данных, на которых была обучена модель MiDaS, обратные карты глубин получаются с неоднозначными масштабом и сдвигом, которые изменяются от кадра к кадру и не соответствуют реальным метрическим значениям. Возможный способ использования модели MiDaS — генерация карт смещений, которые пропорциональны обратным картам глубин.

В соответствии с этим выполняется покадровая постобработка данных. Для получения абсолютных метрических глубин для некоторых пикселей сцены используются значения глубины d^* , формируемой стереокамерой RealSense D435i. Таким образом, необходимо определить наилучшее соответствие между глубиной d^* и алгоритмическим прогнозом на основе критерия наименьших квадратов. Для восстановления предсказанной глубины в реальном метрическом пространстве используется следующая процедура:

— инвертирование глубины d^* для того, чтобы ее значения соответствовали одной и той же области смещения:

$$\bar{d}^* = 1/d^* ;$$

— расчет масштаба и сдвига карты глубин на основе критерия наименьших квадратов:

$$(s, t) = \operatorname{argmin}_{s, t} \sum_{i=1}^P (s\bar{d}_i + t - \bar{d}_i^*)^2 ,$$

где $\bar{d}_i$ — глубина, оцениваемая алгоритмом MiDaS; s и t — коэффициенты масштаба и сдвига, P — количество достоверных пикселей, получаемых из карты глубин, формируемой RGB-D-камерой RealSense;

— выравнивание глубины $\bar{d}_i$ с помощью найденных коэффициентов масштаба и сдвига:

$$\hat{\bar{d}}_i = s\bar{d}_i + t ,$$

где $\hat{\bar{d}}_i$ — выровненная обратная глубина $\bar{d}_i$;

— инвертирование значения $\hat{\bar{d}}_i$ для получения реальных значений метрической глубины:

$$\hat{D}_i = 1/\hat{\bar{d}}_i ,$$

где $\hat{D}_i$ — метрическое предсказание глубины, согласованное с данными стереокамеры RealSense.

Все сгенерированные облака точек глубины будут упакованы в один файл для дальнейшей работы с наборами данных.

Бинокулярная оценка глубины изображения на основе DNN. Адаптация нейронных сетей к окружающей среде в режиме реального времени при выводе изображений введена в [10]. Адаптация здесь означает применение метода обратного распространения ошибки и его модификаций для редактирования весов модели к данным, которые подаются на вход сети впервые. Это позволяет получать точные карты смещений даже для сцен, на которых модель не была обучена. В MADNet существуют три различных варианта вывода, а именно:

— отсутствие режима адаптации — не применяется метод обратного распространения ошибки и сеть работает так, как она обучалась;

— режим полной адаптации — онлайн-адаптация осуществляется полностью;

— режим MAD — модификации осуществляются быстрее, чем в полном режиме.

Режим полной адаптации работает наилучшим образом, но с более высокой задержкой, тогда как режим MAD представляет собой компромисс по скорости и качеству.

Генерация семантических прогнозов. Как определено выше, одно из наиболее точных универсальных решений — MSeg-semantic. Это решение используется в качестве ядра семантической сегментации.

Предварительно обученная модель MSeg (3 млн параметров, разрешение входящих изображений 480p) была выбрана для выполнения семантической сегментации данных на изображениях по нескольким причинам. Во-первых, эта модель требует минимального объема видеопамяти, что является немаловажным фактором, и, во-вторых, модель характеризуется высоким быстродействием.

Тестирование модели производилось в режиме одномасштабного вывода, что означает использование одного конкретного масштаба изображения (например, оригинального $\times 1$). Однако одномасштабный вывод приводит к некачественным прогнозам. Так как при более высоком разрешении входного изображения алгоритм MSeg „разбивает“ изображение на более мелкие фрагменты, то на границах этих фрагментов возникают ошибки в предсказании классов объектов. Во избежание этого был использован многомасштабный вывод, который заключается в выполнении многократного прогнозирования для одних и тех же изображений при различном масштабировании и в последующем усреднении прогнозов для каждого пиксела.

Многомасштабный вывод позволяет исключить отрицательное влияние некачественных семантических прогнозов на SLAM-систему, провести более наглядные эксперименты и получить более точные и плавные предсказания. Однако многомасштабный вывод намного дороже по расходу ресурсов и не работает в режиме реального времени. Поэтому в будущем, при реализации системы в режиме реального времени, планируется использовать одномасштабный вывод.

Изначально на выходе модели MSeg-semantic формируются изображения в оттенках серого, представляющие семантические метки, где, согласно универсальной таксономии [25], 194 метки соответствуют каждому известному классу, а 195-я метка выделяется для неизвестных объектов. Для более корректного представления 2D-семантических изображений, семантического облака точек и результирующей семантической карты была выполнена собственная цветовая маскировка изображений, базирующаяся на Detectron2 [26], и, таким образом, получены RGB семантически-аннотированные изображения.

Экспериментальные результаты. В качестве аппаратного обеспечения для выполнения семантической сегментации 2D-изображений с помощью модели MSeg и для пакета Kimera использовался ноутбук с процессором Intel I5-9300H 2,4 ГГц $\times$ 8 и графическим процессором NVIDIA GeForce GTX 1050; для работы DNN-алгоритмов по оценке глубины использован одноплатный компьютер NVIDIA Jetson AGX Xavier.

Основной набор данных, предоставленный лабораторией робототехники ПАО „Сбербанк“, представляет собой данные, записанные с мобильного наземного робота в офисном помещении, имеющем большие открытые пространства, узкие коридоры и большое количество окон, что создает дополнительные требования к тестируемым решениям и позволяет более точно определить лучшие из них. Кроме того, этот набор содержит все необходимые данные, включая RGB-изображения, инфракрасные стереоизображения, необработанные измерения IMU-сенсора, карты глубин, формируемые RealSense D435i, характеристики стереокамеры. Для оценки качества фреймворка Kimera-VIO использовались данные одометрии, полученные SLAM-системой робота-курьера ПАО „Сбербанк“.

Были обработаны 11 475 RGB-моноизображений с разрешением 1280 $\times$ 720 с помощью моделей MiDaS и MSeg-semantic, а также 11 282 стереоизображения с разрешением 1280 $\times$ 720, сформированные правой и левой камерами RealSense, с использованием алгоритма MADNet в полном режиме. Результат обработки представлен на рис. 2. Для сравнения дополнительно приведены данные, полученные с помощью модифицированного алгоритма Хиршмюллера (Stereo Semi-Global Matching — SGM), используемого по умолчанию в Kimera в случае отсутствия других данных о глубине.

Для фреймворка Kimera объединены плотные данные различных алгоритмов на основе DNN и семантические 2D-изображения с исходным набором данных в целях получения расширенного набора данных. Поскольку генерируемые данные записываются в режиме реального времени, их необходимо адаптировать ко времени, связанному с исходным набором данных. Поэтому временные метки сгенерированных данных переназначаются, чтобы они соответствовали тем, которые существовали в исходном наборе данных.

Для того чтобы все модули Kimera работали в режиме реального времени и во избежание больших затрат ресурсов, используется масштаб 0,5 для левого и правого изображений. Кроме того, произведена модификация Kimera-Semantics для работы с 195 классами и создана карта цветов для 2D-семантических RGB-изображений, которая подается на вход Kimera-Semantics.

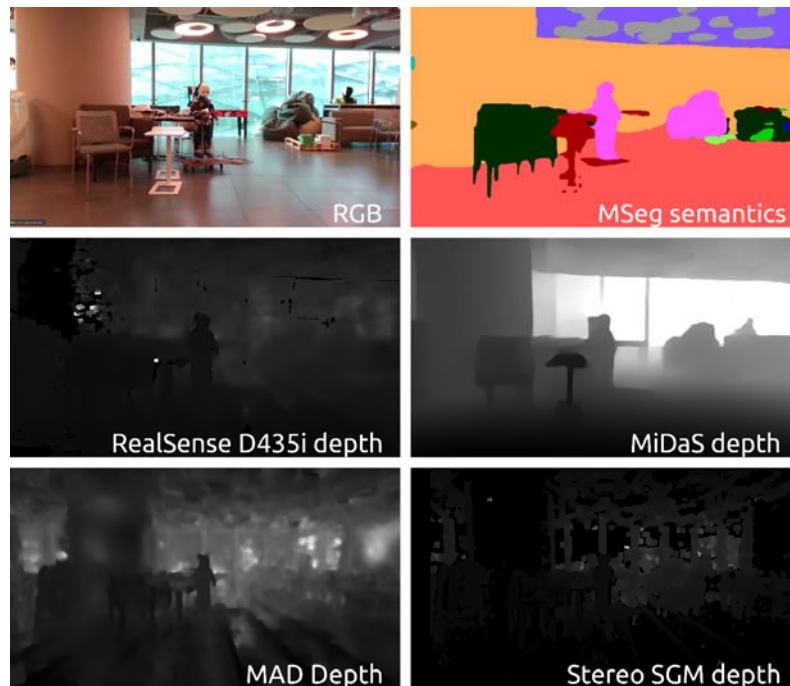


Рис. 2

Объединяя облака точек глубины и семантические метки, получаем семантические облака точек. На рис. 3 представлены семантические облака точек, полученные для каждого решения. Визуальная оценка их качества наглядно демонстрирует превосходство MiDaS над остальными алгоритмами по показателю отсутствия шума. Однако алгоритмом MiDaS не решается проблема детектирования зеркал. Кроме того, использование RealSense обеспечивает наиболее точную глубину для каждого пиксела, тогда как MiDaS не позволяет сделать это напрямую, если отсутствуют дополнительные данные о метрике пространства. Тем не менее плотные облака точек, сформированные RGB-D-камерой RealSense, не обеспечивают получение достоверной информации о прозрачных объектах и отражающих поверхностях, что приводит к большим погрешностям и, следовательно, к повреждению 3D-карт. Более того, облака точек, сгенерированные с использованием MiDaS, имеют более высокое качество с геометрической точки зрения. С другой стороны, облака точек, полученные с помощью MADNet, содержат больше всего артефактов, что объясняется неудовлетворительным предсказанием глубины инфракрасных монохромных изображений, поскольку эта сеть была обучена на стерео RGB-изображениях.

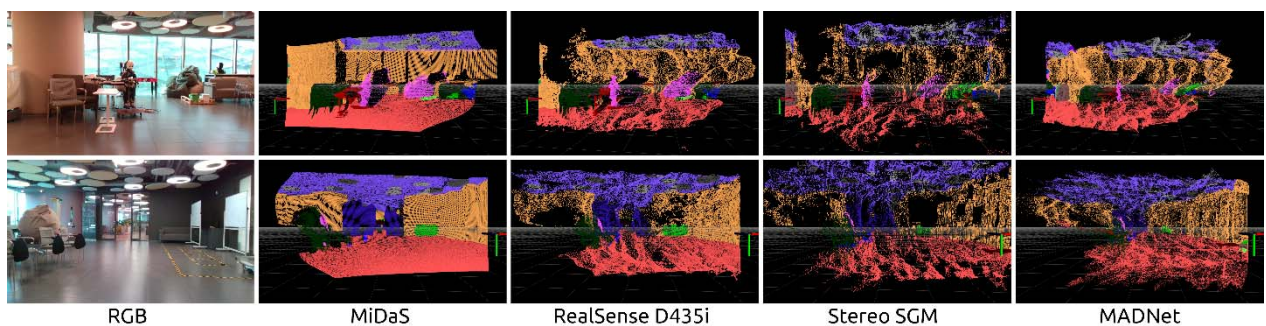


Рис. 3

Для каждого решения были созданы метрико-семантические 3D-карты (рис. 4, вид сверху), а сравнение карт в нескольких сценах представлено на рис. 5 (в скобках указан размер воксела). Визуальный анализ показывает, что качество построенной с помощью MiDaS семантической карты с использованием карт глубин является наилучшим, тогда как решение MADNet наихудшее. Все решения, кроме MiDaS, показали значительные выбросы, особенно в областях сцен, содержащих прозрачные объекты, такие как окна. Визуализация построения метрико-семантической 3D-карты с использованием карт глубин, полученных с помощью алгоритма MiDaS, представлена в электронном ресурсе <https://youtu.be/VtPTjaiF8Jw>.

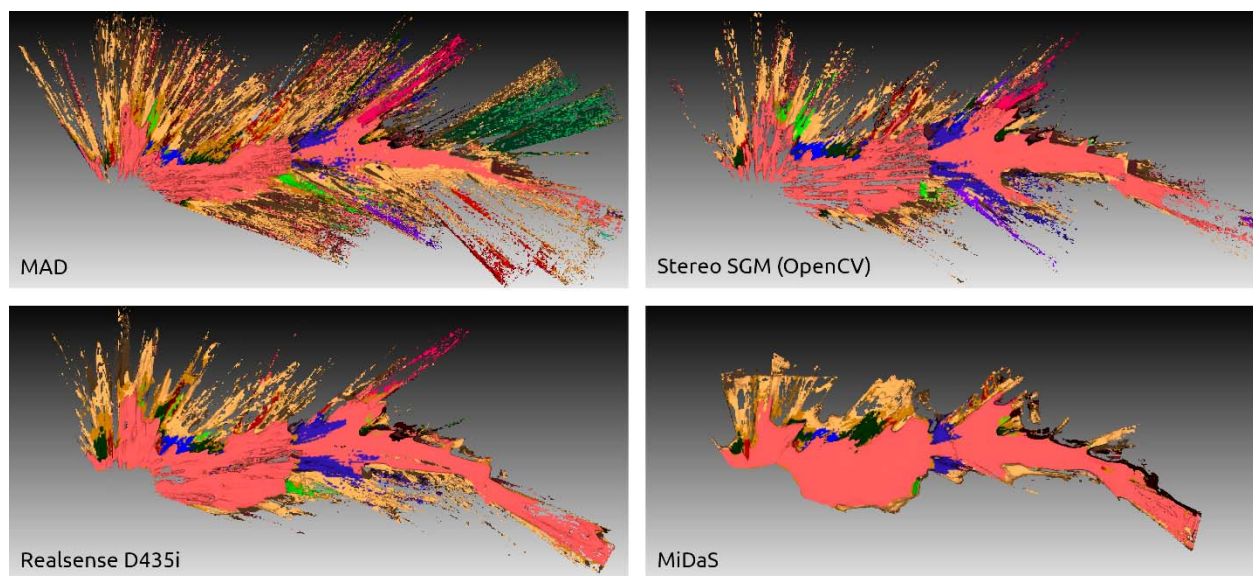


Рис. 4

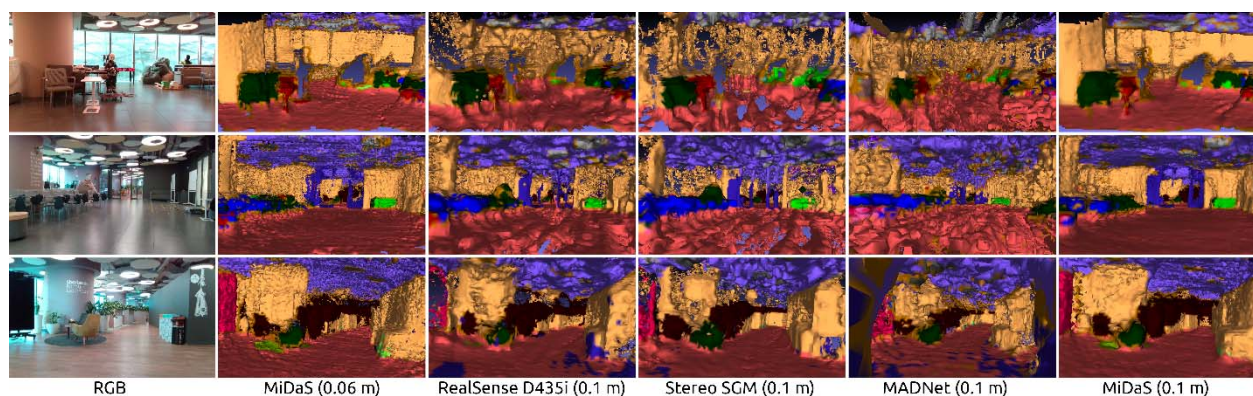


Рис. 5

Для трехмерного случая в качестве основных метрик для оценки качества карт использовались метрики Хаусдорфа. При этом карта, полученная на основе информации о глубине, сформированной камерой RealSense, использовалась как „наземная истина“ (ground-truth). В общем случае одностороннее расстояние Хаусдорфа определяется как максимальное из всех возможных расстояний от каждой точки одного множества до ближайшей к ней точки второго множества. Двустороннее расстояние Хаусдорфа обобщается до максимального из односторонних расстояний Хаусдорфа относительно каждого из множеств, т.е.:

$$D_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} |xy|, \sup_{y \in Y} \inf_{x \in X} |xy| \right\}.$$

Дополнительной метрикой качества является среднее расстояние между полученными 3D-картами, а точнее, между каждым вокселем одной карты и ближайшим к нему вокселем другой карты:

$$\bar{D}(X, Y) = \frac{1}{X} \sum_{x \in X} \inf_{y \in Y} |xy|; \quad \bar{D}(Y, X) = \frac{1}{Y} \sum_{y \in Y} \inf_{x \in X} |xy|,$$

которое целесообразно представить в виде среднего хаусдорфова расстояния:

$$\bar{D}_H(X, Y) = \frac{\bar{D}(X, Y) + \bar{D}(Y, X)}{2}.$$

Результаты оценивания точности восстановления карт глубин с помощью метрик Хаусдорфа представлены в табл. 1. Следует отметить, что показатели метрик имеют достаточно большие значения, особенно двустороннее расстояние Хаусдорфа. Это вызвано множеством причин, в частности несовершенством выбранных в качестве ground-truth данных. Отдаленные выбросы на 3D-картах в наибольшей степени влияют на двустороннее расстояние Хаусдорфа, поэтому в таблице представлены также другие метрики для более справедливого сравнения. Кроме того, в качестве визуальной метрики для выявления явных выбросов и шумов использована тепловая карта ошибок между метрико-семантическими 3D-картами, построенными с применением карт глубин от MiDaS и RealSense (рис. 6).

Таблица 1

Решение (размер воксела)	Количество вершин	Расстояние Хаусдорфа		
		Двустороннее	Среднее	Среднеквадратическое
Stereo SGM (0,1 м)	939667	6,284989	0,315699	0,520958
MADNet (0,1 м)	2342929	8,420206	0,663223	1,148033
MiDaS (0,1 м)	347426	1,084647	0,158118	0,214312
MiDaS (0,06 м)	1105870	1,139520	0,162516	0,223605

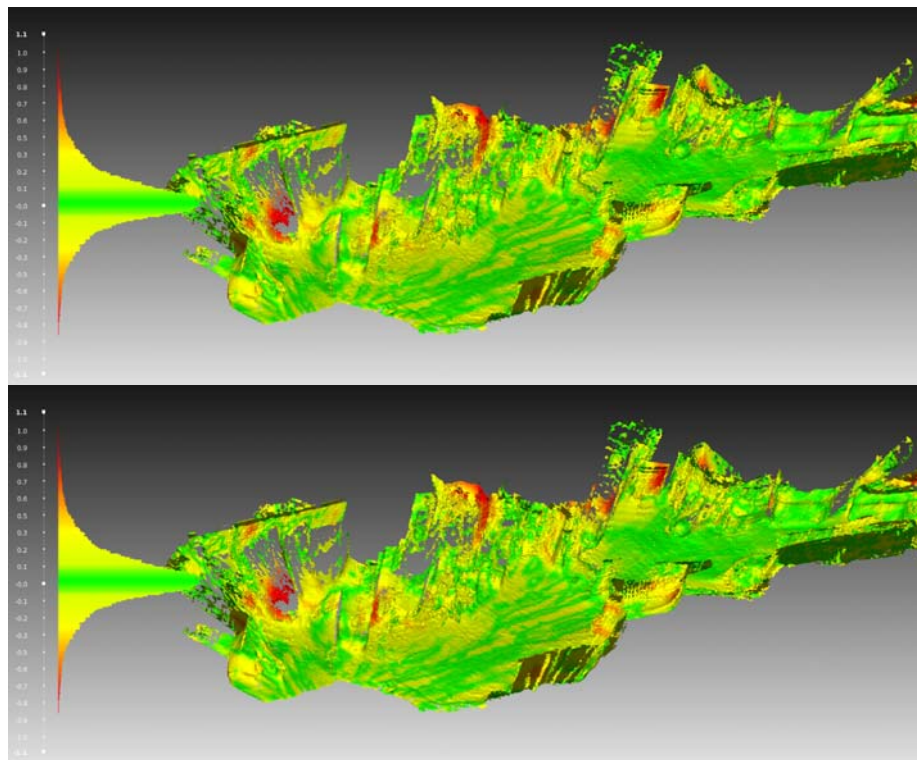


Рис. 6

Еще один важный фактор, который следует отметить при построении карт с использованием глубины, формируемой MiDaS, — меньшее использование оперативной памяти. В табл. 2 приведены данные по использованию оперативной памяти (ОЗУ) в процессе

построения одного и того же фрагмента карты для каждого из решений, а также данные по частоте генерации облаков точек глубины с помощью этих решений. Согласно рис. 5 и табл. 2 заметна зависимость между качеством карты и объемом используемой памяти — чем лучше карта, тем меньше памяти она использует. Это объясняется геометрическим совпадением точек у соседних кадров, в результате чего образуется меньшее количество вокселей.

Таблица 2

Решение (размер вокселя)	ОЗУ, Гбайт	Частота, Гц
RealSense D435i (0,1 м)	2,5	30
Stereo SGM (0,1 м)	2,9	30
MADNet (0,1 м)	4,1	1,17
MiDaS (0,1 м)	1,5	16
MiDaS (0,06 м)	4,6	16

Заключение. Для обеспечения качественной работы метрико-семантической VIO-SLAM-системы были протестированы различные решения по генерации облаков точек глубины изображения. Выбор лучших решений основан на их общей производительности в реальном времени и универсальности для различных сцен с естественными условиями. Аналогичный критерий был задан при выборе пакета MSeg в качестве семантического предиктора.

Для получения правильно масштабированных облаков точек глубины, генерируемых алгоритмом MiDaS, необходимо иметь как минимум несколько точных результатов измерений расстояния. Таким образом, было выполнено слияние алгоритма MiDaS и решения, полученного с помощью RGB-D-камеры RealSense, для генерации более реалистичного метрического облака точек глубины.

Представлена разработанная собственная SLAM-система, основанная на нескольких DNN-алгоритмах, которая может работать благодаря всего одному устройству получения данных — RGB-D-стереокамере. Апробация работы системы проведена на реальных данных, предоставленных лабораторией робототехники ПАО „Сбербанк“, и дополнительно проведены эксперименты, результаты которых демонстрируют высокие требования к алгоритмам. Работа системы показала, что карты, созданные с помощью DNN-алгоритмов, более стабильны и детальны по сравнению с картами, сформированными RGB-D-камерой RealSense, особенно в сценах, содержащих яркие объекты.

В дальнейшем планируется развитие системы с использованием только данных визуально-инерциальной одометрии для обеспечения корректного масштабирования получаемых облаков точек с исключением необходимости в подаче на вход карт глубин из RGB-D-камер и расширением, таким образом, разрабатываемого решения для использования с любыми RGB-камерами.

СПИСОК ЛИТЕРАТУРЫ

1. Zhao C., Sun Q., Zhang C., Tang Y., Qian F. Monocular depth estimation based on deep learning: An overview // Science China Technological Sciences. 2020. Vol. 63, N 9. P. 1612—1627. DOI: 10.1007/s11431-020-1582-8.
2. Yin W., Liu Y., Shen C. Virtual Normal: Enforcing Geometric Constraints for Accurate and Robust Depth Prediction // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2021. DOI: 10.1109/TPAMI.2021.3097396.
3. Ranftl R., Lasinger K., Hafner D., Schindler K., Koltun V. Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-Shot Cross-Dataset Transfer // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2022. Vol. 44, N 03. P. 1623—1637.
4. Sinha A., Murez Z., Bartolozzi J., Badrinarayanan V., Rabinovich A. DELTAS: Depth Estimation by Learning Triangulation and Densification of Sparse Points // Computer Vision – ECCV. 2020. Vol. 12366. P. 104—121. DOI: 10.1007/978-3-030-58589-1_7.

5. Duzceker A., Galliani S., Vogel C., Speciale P., Dusmanu M., Pollefeys M. DeepVideoMVS: Multi-View Stereo on Video with Recurrent Spatio-Temporal Fusion // Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA. 2021. June. P. 15324—15333. DOI: 10.1109/CVPR46437.2021.01507.
6. Wofk D., Ma F., Yang T.-J., Karaman S., Sze V. FastDepth: Fast monocular depth estimation on embedded systems // Proc. IEEE Intern. Conf. on Robotics and Automation. 2019. Vol. 2019, May. P. 6101—6108. DOI: 10.1109/ICRA.2019.8794182.
7. Poggi M., Aleotti F., Tosi F., Mattoccia S. Towards Real-Time Unsupervised Monocular Depth Estimation on CPU // IEEE Intern. Conf. on Intelligent Robots and Systems. 2018. P. 5848—5854. DOI: 10.1109/IROS.2018.8593814.
8. Bokovoy A., Muravyev K., Yakovlev K. Real-time Vision-based Depth Reconstruction with NVidia Jetson // European Conf. on Mobile Robots (ECMR). 2019. P. 1—6. DOI: 10.1109/ECMR.2019.8870936.
9. Smolyanskiy N., Kamenev A., Birchfield S. On the importance of stereo for accurate depth estimation: An efficient semi-supervised deep neural network approach // IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Workshops. 2018. Vol. 2018, June. P. 1120—1128. DOI: 10.1109/CVPRW.2018.00147.
10. Tonioni A., Tosi F., Poggi M., Mattoccia S., L. Stefano D. Real-time self-adaptive deep stereo // Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2019. Vol. 2019. June. P. 195—204. DOI: 10.1109/CVPR.2019.00028.
11. Xu H., Zhang J. AANET: Adaptive aggregation network for efficient stereo matching // Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition. 2020. P. 1956—1965. DOI: 10.1109/CVPR42600.2020.00203.
12. Xia L., Cui J., Shen R., Xu X., Gao Y., Li X. A survey of image semantics-based visual simultaneous localization and mapping: Application-oriented solutions to autonomous navigation of mobile robots // Intern. Journal of Advanced Robotic Systems. 2020. Vol. 17. P. 1—17. DOI: 10.1177/1729881420919185.
13. Chen L.-C., Zhu Y., Papandreou G., Schroff F., Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation // Computer Vision – ECCV. 2018. Vol. 11211. P. 833—851. DOI: 10.1007/978-3-030-01234-2_49.
14. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation // Medical Image Computing and Computer-Assisted Intervention – MICCAI. 2015. Vol. 9351. P. 234—241. DOI: 10.1007/978-3-319-24574-4_28.
15. Wang J., Sun K., Cheng T., Jiang B., Deng C., Zhao Y., Liu D., Mu Y., Tan M., Wang X., Liu W., Xiao B. Deep High-Resolution Representation Learning for Visual Recognition // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2021. Vol. 43, N 10. P. 3349—3364. DOI: 10.1109/TPAMI.2020.2983686.
16. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2017. Vol. 39, N 12. P. 2481—2495. DOI: 10.1109/TPAMI.2016.2644615.
17. Seichter D., Köhler M., Lewandowski B., Wengelfeld T., Gross H. -M. Efficient RGB-D Semantic Segmentation for Indoor Scene Analysis // IEEE Intern. Conf. on Robotics and Automation (ICRA). 2021. P. 13525—13531. DOI: 10.1109/ICRA48506.2021.9561675.
18. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // IEEE Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. P. 770—778. DOI: 10.1109/CVPR.2016.90.
19. Oleynikova H., Taylor Z., Fehr M., Siegwart R., Nieto J. Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). 2017. P. 1366—1373. DOI: 10.1109/IROS.2017.8202315.
20. Rosinol A., Abate M., Chang Y., Carlone L. Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping // IEEE Intern. Conf. on Robotics and Automation (ICRA). 2020. P. 1689—1696. DOI: 10.1109/ICRA40945.2020.9196885.
21. Xiang Y., Fox D. DA-RNN: Semantic Mapping with Data Associated Recurrent Neural Networks // Robotics: Science and Systems. 2017. Vol. 13. DOI: 10.15607/RSS.2017.XIII.013.
22. Grinvald M. et al. Volumetric Instance-Aware Semantic Mapping and 3D Object Discovery // IEEE Robotics and Automation Letters. 2019. Vol. 4, N 3. P. 3037—3044. DOI: 10.1109/LRA.2019.2923960.

23. Reijgwart V., Millane A., Oleynikova H., Siegwart R., Cadena C., Nieto J. Voxgraph: Globally Consistent, Volumetric Mapping Using Signed Distance Function Submaps // *IEEE Robotics and Automation Letters*. 2020. Vol. 5, N 1. P. 227—234. DOI: 10.1109/LRA.2019.2953859.
24. Zendel O., Honauer K., Murschitz M., Steininger D., Dominguez G. F. WildDash — Creating Hazard-Aware Benchmarks // *Proc. of the European Conf. on Computer Vision (ECCV)*. 2018. Sept. P. 402—416.
25. Lambert J., Liu Z., Sener O., Hays J., Koltun V. MSeg: A Composite Dataset for Multi-Domain Semantic Segmentation // *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2020. P. 2876—2885. DOI: 10.1109/CVPR42600.2020.00295.
26. Kirillov A., Wu Y., He K., Girshick R. PointRend: Image Segmentation As Rendering // *IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2020. P. 9796—9805. DOI: 10.1109/CVPR42600.2020.00982.

Сведения об авторах

- | | |
|----------------------------------|---|
| Амиран Рустамович Беркаев | — студент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: berkaevamiran@mail.ru |
| Малик Мохрат | — студент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: malik.mohrat@gmail.com |
| Алексей Михайлович Бурков | — ПАО „Сбербанк“, лаборатория робототехники; ведущий инженер-разработчик; E-mail: amburkoff@gmail.com |
| Сергей Алексеевич Колюбин | — д-р техн. наук, доцент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; вед. научный сотрудник; E-mail: s.kolyubin@itmo.ru |

Поступила в редакцию 28.12.21; одобрена после рецензирования 12.01.22; принята к публикации 18.01.22.

REFERENCES

1. Zhao C., Sun Q., Zhang C., Tang Y., and Qian F. *Science China Technological Sciences*, 2020, no. 9(63), pp. 1612–1627, DOI: 10.1007/s11431-020-1582-8.
2. Yin W., Liu Y. and Shen C. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, DOI: 10.1109/TPAMI.2021.3097396.
3. Ranftl R., Lasinger K., Hafner D., Schindler K., and Koltun V. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, no. 03(44), pp. 1623–1637.
4. Sinha A., Murez Z., Bartolozzi J., Badrinarayanan V., and Rabinovich A. *Computer Vision – ECCV 2020, Springer International Publishing, Cham*, 2020, vol. 12366, pp. 104–121, DOI: 10.1007/978-3-030-58589-1_7.
5. Düzçeker A., Galliani S., Vogel C., Speciale P., Dusmanu M., and Pollefeys M. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, June 2021, pp. 15324–15333, DOI: 10.1109/CVPR46437.2021.01507.
6. Wofk D., Ma F., Yang T.-J., Karaman S., and Sze V. *Proceedings - IEEE International Conference on Robotics and Automation*, 2019, vol. 2019, May, pp. 6101–6108, DOI: 10.1109/ICRA.2019.8794182.
7. Poggi M., Aleotti F., Tosi F., and Mattoccia S. *IEEE International Conference on Intelligent Robots and Systems*, 2018, pp. 5848–5854, DOI: 10.1109/IROS.2018.8593814.
8. Bokovoy A., Muravyev K., and Yakovlev K. *2019 European Conference on Mobile Robots (ECMR)*, 2019, pp. 1–6, DOI: 10.1109/ECMR.2019.8870936.
9. Smolyanskiy N., Kamenev A., and Birchfield S. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2018, vol. 2018-June, pp. 1120–1128, doi: 10.1109/CVPRW.2018.00147.
10. Tonioni A., Tosi F., Poggi M., Mattoccia S., and Stefano L.D. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019, vol. 2019-June, pp. 195–204, DOI: 10.1109/CVPR.2019.00028.
11. Xu H. and Zhang J. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020, pp. 1956–1965, DOI: 10.1109/CVPR42600.2020.00203.
12. Xia L., Cui J., Shen R., Xu X., Gao Y., and Li X. *Int. J. Adv. Robot. Syst.*, 2020, vol. 17, p. 172988142091918, DOI: 10.1177/1729881420919185.
13. Chen L.-C., Zhu Y., Papandreou G., Schroff F., and Adam H. *Computer Vision – ECCV, Springer International Publishing, Cham*, 2018, vol. 11211, pp. 833–851, DOI: 10.1007/978-3-030-01234-2_49.
14. Ronneberger O., Fischer P., and Brox T. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham*, 2015, vol. 9351, pp. 234–241, DOI: 10.1007/978-3-319-24574-4_28.
15. Wang J., Sun K., Cheng T., Jiang B., Deng C., Zhao Y., Liu D., Mu Y., Tan M., Wang X., Liu W. and Xiao B. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, no. 10(43), pp. 3349–3364, DOI: 10.1109/TPAMI.2020.2983686.
16. Badrinarayanan V., Kendall A., and Cipolla R. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,

- 2017, no. 12(39), pp. 2481–2495, DOI: 10.1109/TPAMI.2016.2644615.
17. Seichter D., Köhler M., Lewandowski B., Wengefeld T., and Gross H.-M. *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 13525–13531, DOI: 10.1109/ICRA48506.2021.9561675.
 18. He K., Zhang X., Ren S., and Sun J. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778, DOI: 10.1109/CVPR.2016.90.
 19. Oleynikova H., Taylor Z., Fehr M., Siegwart R., and Nieto J. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 1366–1373, DOI: 10.1109/IROS.2017.8202315.
 20. Rosinol A., Abate M., Chang Y., and Carlone L. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1689–1696, DOI: 10.1109/ICRA40945.2020.9196885.
 21. Xiang Y. and Fox D. *2017 Robotics: Science and Systems*, 2017, july, vol. 13, DOI: 10.15607/RSS.2017.XIII.013.
 22. Grinvald M. et al. *IEEE Robot. Autom. Lett.*, 2019, no. 3(4), pp. 3037–3044, DOI: 10.1109/LRA.2019.2923960.
 23. Reijgwart V., Millane A., Oleynikova H., Siegwart R., Cadena C., and Nieto J. *IEEE Robot. Autom. Lett.*, 2020, no. 1(5), pp. 227–234, DOI: 10.1109/lra.2019.2953859.
 24. Zendel O., Honauer K., Murschitz M., Steininger D., and Dominguez G.F. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 402–416.
 25. Lambert J., Liu Z., Sener O., Hays J., and Koltun V. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2876–2885, DOI: 10.1109/CVPR42600.2020.00295.
 26. Kirillov A., Wu Y., He K. and Girshick R. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 9796–9805, DOI: 10.1109/CVPR42600.2020.00982.

Data on authors

Amiran R. Berkaev	—	Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; E-mail: berkaevamiran@mail.ru
Malik Mohrat	—	Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; E-mail: malik.mohrat@gmail.com
Alexey M. Burkov	—	Sberbank, Robotics Laboratory; Lead Engineer-Designer; E-mail: amburkoff@gmail.com
Sergey A. Kolyubin	—	Dr. Sci., Associate Professor; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; Leading Researcher; E-mail: s.kolyubin@itmo.ru

Received 28.12.21; approved after reviewing 12.01.22; accepted for publication 18.01.22.

**ОПТИМИЗАЦИОННЫЕ АЛГОРИТМЫ
ПОВЫШЕНИЯ ТОЧНОСТИ И РОБАСТНОСТИ ВИЗУАЛЬНОЙ ОДОМЕТРИИ
НАЗЕМНЫХ МОБИЛЬНЫХ РОБОТОВ**

Ж. МАХМУД^{1*}, В. ХА ТХЕ ЛОНГ¹, А. М. БУРКОВ², С. А. КОЛЮБИН¹

¹ Университет ИТМО, Санкт-Петербург, Россия

*Jaafar.a.mahmoud@itmo.ru

² Сбербанк, Москва, Россия

Аннотация. Рассматривается задача повышения точности и робастности методов одновременной локализации и картирования с использованием численной оптимизации с ограничениями. В основе предлагаемого решения лежит модификация алгоритма ORB-SLAM3, учитывающая особенности кинематики наземных роботов и использующая комплексирование данных визуальной и колесной одометрии, блочное уравнивание (bundle adjustment) для настройки параметров, комплексно характеризующих состояние визуального сенсора, а также алгоритм замыкания петли для корректировки карты. Результаты апробации на наборе данных OpenLoris показывают, что для ряда сценариев предлагаемое решение существенно превосходит по точности и робастности известный алгоритм ORB-SLAM3.

Ключевые слова: мобильные роботы, локализация, оптимизация, одометрия

Ссылка для цитирования: Махмуд Ж., Ха Тхе Лонг В., Бурков А. М., Колюбин С. А. Оптимизационные алгоритмы повышения точности и робастности визуальной одометрии наземных мобильных роботов // Изв. вузов. Приборостроение. 2022. Т. 65, № 3. С. 218—226. DOI: 10.17586/0021-3454-2022-65-3-218-226.

**OPTIMIZATION ALGORITHMS
FOR IMPROVING THE ACCURACY AND ROBUSTNESS OF VISUAL ODOMETRY
OF GROUND-BASED MOBILE ROBOTS**

J. Mahmoud^{1*}, V. Ha The Long¹, A. M. Burkov², S. A. Kolyubin¹

¹ ITMO University, St. Petersburg, Russia

*Jaafar.a.mahmoud@itmo.ru

² Sberbank, Moscow, Russia

Abstract. The problem of improving the accuracy and robustness of simultaneous localization and mapping methods using numerical optimization with constraints is considered. The proposed solution is based on a modification of the ORB-SLAM3 algorithm, which takes into account the peculiarities of the kinematics of ground robots and uses the complexing of visual and wheel odometry data, bundle adjustment for setting parameters that comprehensively characterize the state of the visual sensor, as well as the loop closure algorithm to correct the map. Results of the approach approbation with an OpenLoris dataset demonstrate that for several scenarios the proposed solution is significantly superior in accuracy and robustness to the ORB-SLAM3 algorithm.

Keywords: mobile robots, localization, optimization, odometry

For citation: Mahmoud J., Ha The Long V., Burkov A. M., Kolyubin S. A. Optimization algorithms for improving the accuracy and robustness of visual odometry of ground-based mobile robots. *Journal of Instrument Engineering*. 2022. Vol. 65, N 3. P. 218—226 (in Russian). DOI: 10.17586/0021-3454-2022-65-3-218-226.

Введение. Современные навигационные системы мобильных роботов должны эффективно решать задачи одновременной локализации и картирования (Simultaneous Localization and Mapping — SLAM) [1]. Для этих целей часто используются алгоритмы визуальной одометрии. Несмотря на серьезные преимущества, этот подход не лишен ряда недостатков, приводящих как к дрейфу, так и к выбросам при локализации робота. В настоящей статье пред-

лагается комплексное решение, направленное на повышение точности и робастности существующих SLAM-систем.

Положение робота в трехмерном евклидовом пространстве ($SE(3)$) описывается через линейные координаты позиции и угловые координаты ориентации. Для наземных колесных роботов эта параметризация является избыточной, так как для однозначного определения положения на плоскости (в пространстве $SE(2)$) достаточно знать значения двух линейных координат положения и угла рыскания. Однако ограничения $SE(2)$ недостаточно точны при неровности дороги или при тряске мобильного робота. Для решения этой проблемы предлагается простой, но мощный алгоритм, позволяющий напрямую вычислить положение колесного робота по измерениям координат за пределами пространства $SE(2)$ с учетом возможных возмущений. Это оптимизационный алгоритм, который использует нелинейные ограничения типа $SE(2)$ -XYZ, чтобы связать положение робота в $SE(2)$ с расположением 3D-ориентиров визуальной одометрии.

Типовым сочетанием датчиков мобильного робота являются различные видеокамеры и колесные энкодеры. Комплексирование информации, поступающей с этих датчиков, позволяет значительно повысить точность и надежность системы, особенно когда исключительно визуальные данные не могут считаться достоверным источником информации, например в условиях плохого освещения. При этом для комплексирования асинхронно поступающих измерений от системы технического зрения и колесных энкодеров временные последовательности данных от энкодеров были интерполированы с использованием алгоритма предварительной интеграции и привязаны к моментам времени поступления кадров, формируемых камерами.

Еще одним преимуществом предлагаемого подхода является его робастность к потере визуальных маркеров на более высоких скоростях перемещения робота. Благодаря комплексированию с колесной одометрией обеспечивается более высокая точность локализации робота, что позволяет лучше прогнозировать расположение отслеживаемых точек при изменении положения робота в интервале между получением последовательных кадров в условиях, когда использование других алгоритмов визуальной одометрии привело бы к прерыванию выполнения этой операции.

В результате на основе модификации подхода ORB-SLAM3 [2] авторами настоящей статьи была разработана полная визуально-одометрическая система одновременной локализации и картирования для колесного наземного мобильного робота. Данная система была протестирована на открытом наборе данных OpenLoris [3].

Анализ алгоритмов одновременной локализации и картирования. Алгоритмы SLAM можно разделить на два класса: системы на основе фильтров и системы на основе оптимизации. Системы на основе фильтров, такие как [4, 5], работают путем объединения данных различных датчиков, так что неопределенность положения робота может быть значительно уменьшена. Одним из наиболее популярных методов является SLAM расширенного фильтра Калмана (Extended Kalman Filter for SLAM: EKF-SLAM) [1, 6]. Для обычного EKF-SLAM характерна проблема рассогласования фильтров, вызванная невыполнением гипотезы относительного гауссова распределения шума и процессом линеаризации [7]. Более того, вычислительная нагрузка, связанная с растущей сложностью вычисления матриц Якоби с увеличением размера карты, также ограничивает области применения данного алгоритма.

В последнее время алгоритмы, основанные на оптимизации, используются все чаще. Их легче реализовать и адаптировать для добавления большого набора датчиков. В этом случае задача локализации и картирования формулируется как восстановление графа последовательных положений робота и представляет собой задачу нелинейной оптимизации. Примеры SLAM-систем, работающих с использованием оптимизации графа положений, представлены, например, в работах [2, 8, 9]. В настоящей работе используется графовый алгоритм блочного

уравнивания (bundle adjustment), работающий на комплексированных данных визуальной и колесной одометрии.

Теоретические основы оптимизации. Оптимизация графа положений формирует задачу оценки состояния робота в виде графа G , при этом определяются координаты положения робота, которые минимизируют функцию затрат следующего вида:

$$F(x) = \sum_{k \in G} e_k(x)^T \Omega_k e_k(x),$$

где e_k — функция ошибки оптических и одометрических измерений, e_k можно рассматривать как ребро k в графе G ; Ω_k — информационная матрица измерений ребра k .

Состояние мобильного робота (положение, скорость, показания ИНС-датчиков и др.) может быть представлено в виде вершины графа G , а ограничения на изменение состояния представлены в виде ребер графа, при этом информационная матрица Ω_k служит весовым коэффициентом перехода между ребрами графа, как правило, определяемым обратной матрицей ковариации.

Цель состоит в том, чтобы минимизировать функцию затрат $F(x)$, найдя состояние x :

$$x^* = \arg \min_x F(x).$$

Функцию затрат $F(x)$ можно аппроксимировать следующим образом:

$$F(x \oplus \Delta x) \approx \sum_{k \in G} \|e_k + J_k \Delta x\|_{\Omega_k} = \sum_{k \in G} \|e_k\|_{\Omega_k} + 2e_k^T \Omega_k J_k \Delta x + \|\Delta x\| J_k^T \Omega_k J_k,$$

где $\oplus$ — специальный оператор суммы для обновления состояния с возмущением Δx ; $J_k = \frac{\delta(e_k(x \oplus \Delta x))}{\delta(\Delta x)}$ — якобианы, $H_k = J_k^T \Omega_k J_k$ — гессианы аппроксимирующей модели.

Тогда представленная оптимизационная задача решается итеративно, где вычисление значения аргумента на каждом шаге процедуры происходит по правилу $x^* = x \oplus \Delta x$, где приращение аргумента обновляется как

$$\Delta x = H^{-1}b,$$

где

$$H = \sum_{k \in G} H_k, \quad b = \sum_{k \in G} b_k = \sum_{k \in G} (e_k^T \Omega_k J_k)^T.$$

Оптимизация графа с добавлением ребер в пространстве SE(2)-XYZ. Общая архитектура системы унаследована от системы ORB-SLAM3 [2], где ключевые кадры выбираются на основе оптических и временных критериев. После создания нового ключевого кадра он отправляется в поток картирования. Работа алгоритма блочного уравнивания в составе предлагаемого решения проиллюстрирована на рис. 1.

Процесс оптимизации графа выполняется с использованием двух типов ребер. Первое ребро отражает ошибку проекции между обнаруженным объектом в сцене и проецируемой 3D-точкой объекта. Второе ребро представляет собой одометрическое ограничение, которое содержит информацию о предварительном интегрировании между двумя временными метками двух последующих ключевых кадров. Информация о предварительном интегрировании содержит числовые значения, которые описывают движение робота. Два ребра должны быть описаны относительно пространства SE(2)-XYZ.

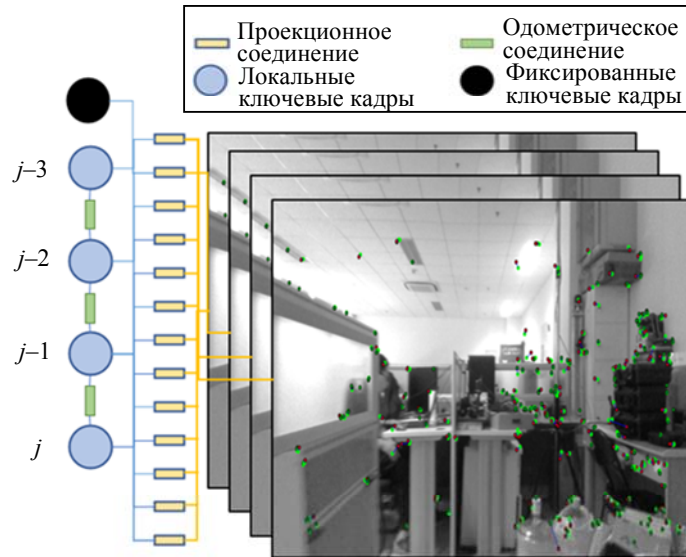


Рис. 1

Рассмотрим точку l в мировой системе координат, точка l наблюдается в ключевом кадре j . Согласно калибровке датчиков робота осуществляется преобразование между системами координат базы робота и камеры $[R_{CB}, p_{CB}]$, где R_{CB} — вращение, p_{CB} — перемещение между камерой C и роботом B . Модель измерения соответствующих координат оптической точки в системе координат камеры представима как

$$u(l, j) = \Gamma(R_{CB}R_j^T(l - p_j) + p_{CB}) + \eta_u,$$

где $\eta_u \sim N(0_{2 \times 1}, \sigma_u^2 \mathbf{I}_{2 \times 2})$ — гауссов шум в измерениях с нулевым средним и дисперсией $\sigma_u^2 \mathbf{I}_{2 \times 2}$; $\Gamma(\cdot)$ — функция проекции; $[R_j, p_j]$ — ориентация и положение робота во время ключевого кадра j .

Если для колесного робота вводятся жесткие ограничения при расчете матрицы поворота R_j и положения p_j робота в пространстве $SE(2)$, так что $R_j = \exp([0 \ 0 \ \theta_z]^T)$, где θ_z — угол поворота робота вокруг оси z , $\exp(\cdot)$ — функция для вычисления матрицы поворота в пространстве $SO(3)$ по углам Эйлера в пространстве $\mathbb{R}^3$, а $p_j = [x_j \ y_j \ 0]^T$, где $r_j = [x_j \ y_j]^T$, то это может привести к снижению точности локализации. Во избежание этого неопределенность out-of- $SE(2)$ инкапсулируется вместе с ограниченным шумом в измерениях по всем возможным координатам перемещения робота в трехмерном пространстве. Иными словами, $R_j \leftarrow \exp([\eta_{\theta_{xy}}^T \ 0]^T)R_j$, где $\theta_{xy} = [\theta_x \ \theta_y]^T$ и $\eta_{\theta_{xy}} = [\eta_{\theta_x} \ \eta_{\theta_y}]^T$, $p_j \leftarrow p_j + [0 \ 0 \ \eta_z]^T$, что ведет к

$$\begin{aligned} u(l, j) &= \Gamma \left[R_{CB}R_j^T \exp \left(\begin{bmatrix} -\eta_{\theta_{xy}} \\ 0 \end{bmatrix} \right) \left(l - p_j - \begin{bmatrix} 0_{2 \times 1} \\ \eta_z \end{bmatrix} \right) + p_{CB} \right] + \eta_u \approx \\ &\approx \Gamma(R_{CB}R_j^T(l - p_j) + p_{CB}) + J_{\eta_{\theta_{xy}}} \eta_{\theta_{xy}} + J_{\eta_z} \eta_z + \eta_u, \end{aligned}$$

где возмущение по линейной вертикальной координате $\eta_z \sim N(0, D_z^2)$ — это гауссов шум с нулевым средним и дисперсией σ_z^2 , а по угловой координате $\eta_{\theta_{xy}} \sim N(0_{2 \times 1}, D_{\theta_{xy}})$ — это гауссов шум с нулевым средним и дисперсией $D_{\theta_{xy}}$.

Затем для процесса оптимизации вычисляются якобианы первого ребра.

Одометрическое ограничение выводится из теоремы предварительного интегрирования, аналогичной [10], но должно быть сформулировано в пространстве SE(2). Предполагается, что энкодеры обеспечивают измерения $\tilde{\mathbf{v}}_k = [\tilde{r}_k \quad \tilde{\theta}_k]^T$ в дискретном времени между двумя последовательными кадрами $[k, k+1]$, где $\tilde{r}_k = [r_k^{k+1} + \eta_{r_k}]$ — движение робота на этом интервале по осям X и Y , а $\tilde{\theta}_k = [\theta_k^{k+1} + \eta_{\theta_k}]$ — угол поворота робота на этом интервале вокруг оси Z (здесь k — индекс, соответствующий измерениям между двумя кадрами, для ключевых кадров используются индексы i, j):

$$\tilde{\mathbf{v}}_k = \mathbf{v}_k^{k+1} + \eta_{\mathbf{v}_k} = \begin{bmatrix} r_k^{k+1} \\ \theta_k^{k+1} \end{bmatrix} + \begin{bmatrix} \eta_{r_k} \\ \eta_{\theta_k} \end{bmatrix},$$

где $\eta_{\mathbf{v}_k} \sim N(0_{3 \times 1}, D_{\mathbf{v}_k})$ — гауссов шум с нулевым средним и дисперсией $D_{\mathbf{v}_k}$.

Позиция робота r_{k+1} зависит от угла рыскания θ_k и рассчитывается на основе выражения

$$\mathbf{v}_{k+1} = \begin{bmatrix} r_k + \Phi(\theta_k)(\tilde{r}_k - \eta_{r_k}) \\ \theta_k + (\tilde{\theta}_k - \eta_{\theta_k}) \end{bmatrix}, \quad \Phi(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

Проводя измерения между двумя ключевыми кадрами i и j вне зависимости от ключевого кадра i , можно записать:

$$\Phi(-\theta_i)r_{k+1} = \Phi(-\theta_i)r_k + \Phi(\theta_k - \theta_i)(\tilde{r}_k - \eta_{r_k}) = \Phi(-\theta_i)r_i + \sum_{n=i}^k \Phi(\theta_n - \theta_i)(\tilde{r}_n - \eta_{r_n}).$$

Далее можно записать выражения для предварительно интегрированных измерений и соответствующих шумов между ключевыми кадрами i и j :

$$\begin{aligned} \theta_i^j &= \sum_{k=i}^{j-1} (\tilde{\theta}_k - \eta_{\theta_k}) = \sum_{k=i}^{j-1} \tilde{\theta}_k - \sum_{k=i}^{j-1} \eta_{\theta_k} := \tilde{\theta}_i^j - \delta\theta_i^j; \\ r_i^j &= \sum_{k=i}^{j-1} \Phi(\tilde{\theta}_k^i) \Phi(\delta\theta_k^i) (\tilde{r}_k - \eta_{r_k}) \approx \sum_{k=i}^{j-1} \Phi(\tilde{\theta}_k^i) (\mathbf{I}_{2 \times 2} - \left[\delta\theta_k^i \right]_{\times}) (\tilde{r}_k - \eta_{r_k}) \approx \\ &\approx \sum_{k=i}^{j-1} \Phi(\tilde{\theta}_k^i) \tilde{r}_k - \sum_{k=i}^{j-1} \Phi(\tilde{\theta}_k^i) (\eta_{r_k} - \left[\delta\theta_k^i \right]_{\times} \tilde{r}_k) := \tilde{r}_i^j - \delta r_i^j, \end{aligned}$$

где $\left[\begin{bmatrix} x & y & z \end{bmatrix}^T \right]_{\times} = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}$ — кососимметричная матрица.

Для расчета величины шума в процессе интегрирования используется итерационная формула:

$$\begin{aligned} \delta\theta_i^{k+1} &= \delta\theta_i^k + \eta_{\theta_k}; \\ \delta r_i^{k+1} &= \delta r_i^k + \Phi(\tilde{\theta}_i^k) (\eta_{r_k} + \left[\delta\theta_k^i \right]_{\times} \tilde{r}_k) = \delta r_i^k + \Phi(\tilde{\theta}_i^k) \eta_{r_k} + \Phi(\tilde{\theta}_i^k) \left[1 \right]_{\times} \tilde{r}_k \delta\theta_k^i \Rightarrow \\ &\Rightarrow \begin{bmatrix} \delta r_i^{k+1} \\ \delta\theta_i^{k+1} \end{bmatrix} := \delta\mathbf{v}_i^{k+1} = A_k \delta\mathbf{v}_i^k + B_k \eta_{\mathbf{v}_k}, \end{aligned}$$

где

$$A_k = \begin{bmatrix} \mathbf{I}_{2 \times 2} & \Phi(\tilde{\theta}_i^k) \lfloor 1 \rfloor_{\times} \tilde{r}_k \\ 0 & 1 \end{bmatrix}, \quad B_k = \begin{bmatrix} \Phi(\tilde{\theta}_i^k) & 0 \\ 0 & 1 \end{bmatrix}.$$

Учитывая ковариацию шума одометрического измерения (зависит от энкодера, используемого в роботе) на каждой итерации, ковариацию δv_i^k можно вычислять постепенно вместе с предварительно интегрированными измерениями $\tilde{\theta}_i^j, \tilde{r}_i^j$:

$$D_{\delta v_i^{k+1}} = A_k D_{\delta v_i^k} A_k^T + B_k D_{v_k} B_k^T.$$

Теперь можно построить модель ошибок для оптимизации на основе предварительно интегрированных одометрических измерений. С учетом двух ключевых кадров, характеризующих последовательные положения v_i и v_j робота в пространстве SE(2), разница между ними определяется как

$$e_i^j = \begin{bmatrix} \Phi(-\theta_i)(r_j - r_i) \\ \theta_j - \theta_i \end{bmatrix} - \begin{bmatrix} \tilde{r}_i^j \\ \tilde{\theta}_i^j \end{bmatrix}.$$

Затем для процесса оптимизации вычисляются якобианы второго ребра.

Интегрированное решение. Архитектура модифицированной системы представлена на рис. 2. В потоке „отслеживание“ процесс оценивания положения робота выполняется на основе предварительной интеграции одометрических измерений с датчиков. Рассчитанное положение выражается в пространстве SE(2), и ключевые кадры отправляются в поток „локальное картирование“, где выполняется блочное уравнивание. В потоке „закрывание петли“ выполняется распознавание прежних оптических сцен и исправляется смещение, вызванное зашумленными оптическими измерениями. Все вычисления в этом потоке выполняются в пространстве SE(3), а затем на заключительном этапе осуществляется глобальное блочное уравнивание оптической одометрии в пространстве SE(2).

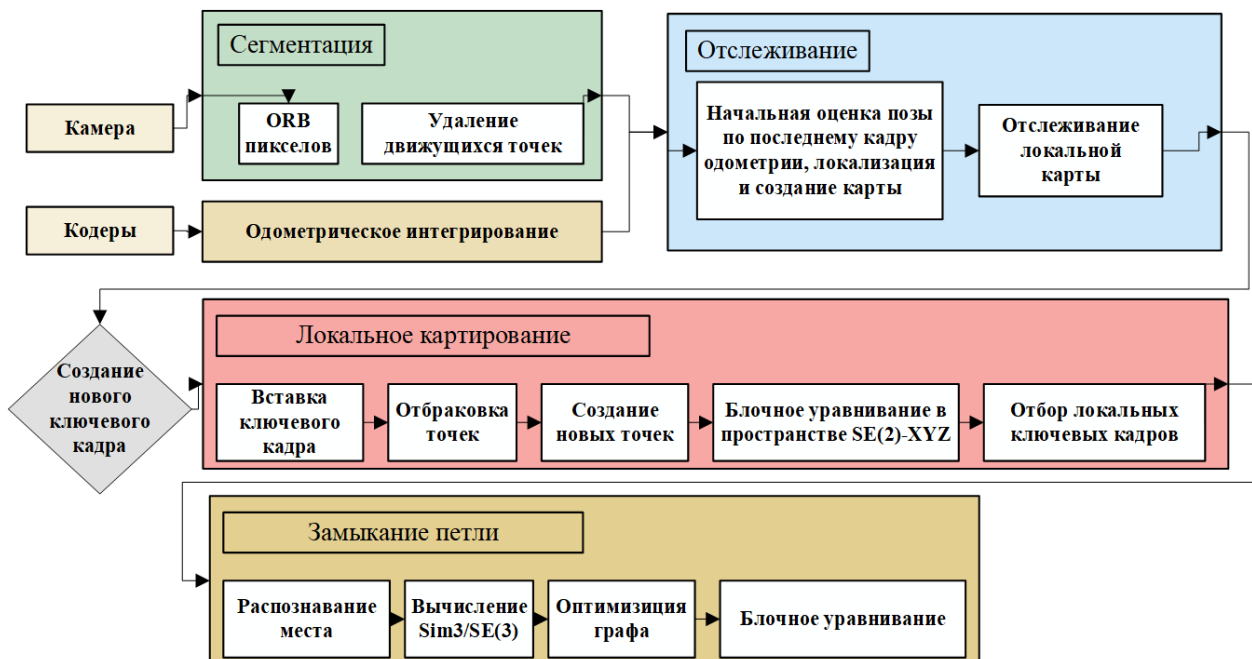


Рис. 2

Апробация модифицированной системы. Разработанная система была протестирована на наборе данных с открытым доступом OpenLoris [3]. Этот набор данных содержит несколько последовательных изображений, записанных наземным колесным мобильным роботом с соответствующими показаниями энкодера. Результаты оценки точности локализации

робота на основе разработанного алгоритма представлены в таблице, где оценка ошибки траектории — среднеквадратическое отклонение (RMSE); здесь же для сопоставления приведены результаты тестирования системы ORB-SLAM3. Анализ приведенных данных показывает, что предложенная модификация значительно улучшает точность работы визуальной SLAM-системы при различных сценариях перемещения роботов в разных помещениях.

Сценарий тестирования (последовательность данных OpenLoris)		Ошибка определения положения робота, м	
		ORB-SLAM3	Предлагаемое решение
Офис	Тест 1	0,063	0,020
	Тест 2	0,076	0,026
	Тест 3	0,028	0,022
	Тест 4	0,080	0,048
	Тест 5	0,227	0,081
	Тест 6	0,067	0,024
	Тест 7	0,060	0,022
RMSE (офис)		0,086	0,035
Квартира	Тест 1	0,424	0,221
	Тест 2	0,371	0,299
	Тест 3	0,351	0,116
	Тест 4	0,299	0,085
	Тест 5	0,257	0,069
RMSE (квартира)		0,340	0,158
Кафе	Тест 1	0,102	0,213
	Тест 2	0,115	0,422
RMSE (кафе)		0,108	0,318
Коридор	Тест 1	5,040	1,831
	Тест 2	6,119	2,147
	Тест 3	11,273	0,159
	Тест 4	2,916	0,235
	Тест 5	4,464	0,420
RMSE (коридор)		5,962	1,139
Рынок	Тест 1	15,771	2,590
	Тест 2	9,126	5,351
	Тест 3	11,431	4,523
RMSE (рынок)		12,109	4,155

На рис. 3 приведены графики оценки траектории движения робота (1 — точные данные, 2 — предлагаемое решение, 3 — ORB-SLAM3), демонстрирующие наряду с повышением точности предлагаемого решения устойчивое движение по оси Z благодаря одометрическим ограничениям.

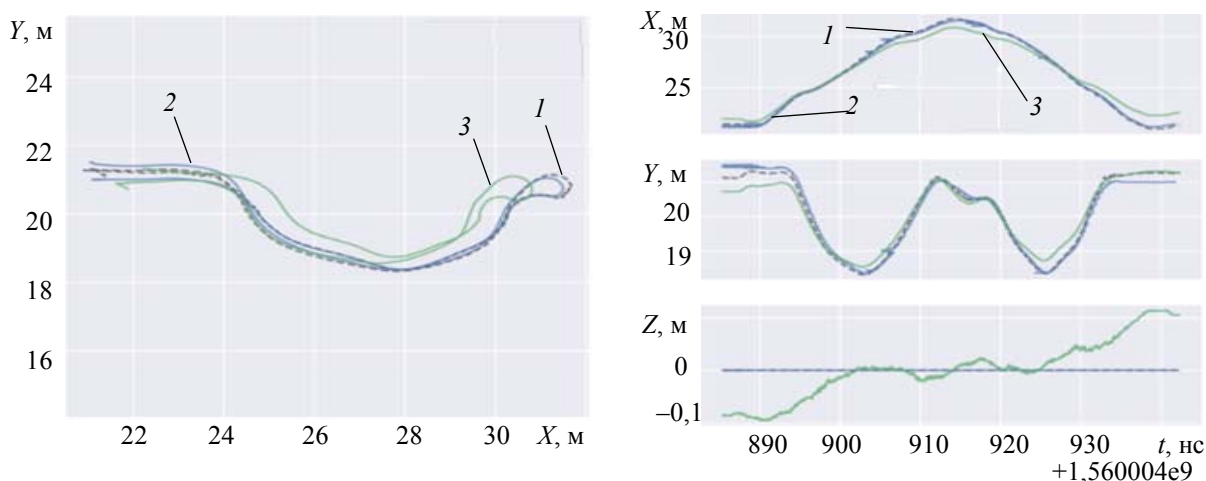


Рис. 3

Надежность рассмотренных систем продемонстрирована на рис. 4: *а* — система ORB-SLAM3 потеряла визуальные маркеры, *б* — предложенная система продолжает отслеживать оптические точки благодаря предварительной интеграции одометрических измерений в потоке „отслеживание“.

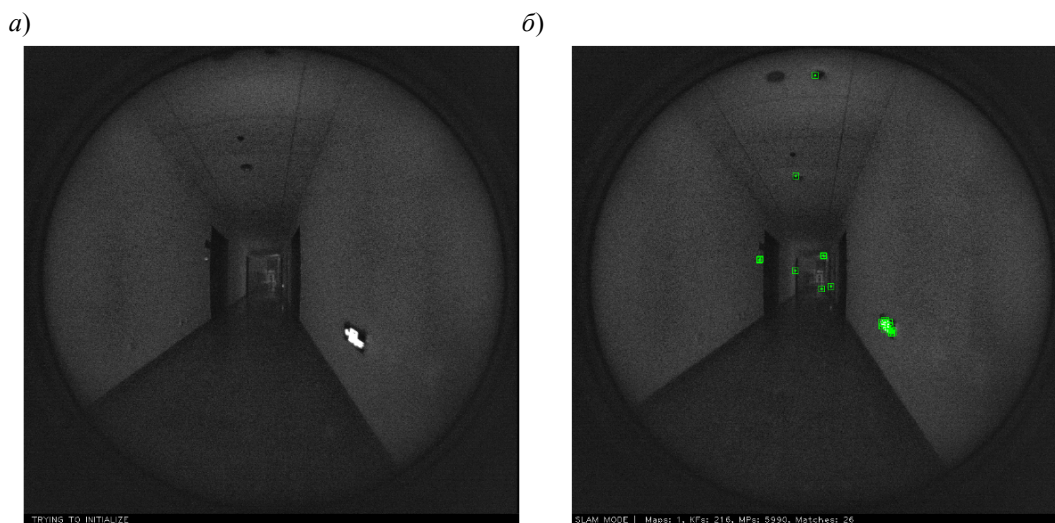


Рис. 4

Заключение. На основе модификации системы ORB-SLAM3 разработана полная визуально-одометрическая система одновременной локализации и картирования для более эффективной и точной работы мобильных наземных роботов. Идея состоит в том, чтобы воспользоваться преимуществами ограниченного плоского движения этих роботов, добавив ограничение в процедуру оптимизации графа положений. Система была протестирована с использованием набора данных OpenLoris. Результаты апробации продемонстрировали, что предложенный подход значительно превосходит известные аналоги по робастности и точности при работе в сложных условиях динамического окружения.

СПИСОК ЛИТЕРАТУРЫ

1. Bailey T., Durrant-Whyte H. Simultaneous localization and mapping (SLAM): Part II. // IEEE Robotics & Automation Magazine. 2006. Vol. 13(3). P. 108—117.
2. Campos C., Elvira R., Rodriguez J., Montiel J. M. Tardos J. D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM // IEEE Trans. on Robotics. 2021. P. 1—17.
3. Shi X., Li D. et al. Are We Ready for Service Robots? The OpenLORIS-Scene Datasets for Lifelong SLAM // Intern. Conf. on Robotics and Automation (ICRA), Paris, France. 2020. P. 3139—3145.
4. Bloesch M., Burri M., Omari S., Hutter M., Siegwart R. Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback // Intern. Journal of Robotics Research. 2017. Vol. 36(10). P. 1053—1072.
5. Geneva P., Eckenhoff K., Lee W., Yang Y., Huang G. 2020. OpenVINS: A Research Platform for Visual-Inertial Estimation // IEEE Intern. Conf. on Robotics and Automation (ICRA). 2020.
6. Kummerle R., Grisetti G., Strasdat H., Konolige K., Burgard W. G2o: A general framework for graph optimization // IEEE Intern. Conf. on Robotics and Automation. 2011.
7. Shoudong Huang, Gamini Dissanayake. Convergence and Consistency Analysis for Extended Kalman Filter Based SLAM // IEEE Trans. on Robotics. 2007. Vol. 23(5). P. 1036—1049.
8. Qin T., Li P., Shen S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator // IEEE Trans. on Robotics. 2018. Vol. 34(4). P. 1004—1020.
9. Usenko V., Demmel N., Schubert D., Stuckler J., Cremers D. Visual-Inertial Mapping With Non-Linear Factor Recovery // IEEE Robotics and Automation Letters. 2020. Vol. 5(2). P. 422—429.

10. Quan M., Piao S., Tan M., Huang S. Tightly-Coupled Monocular Visual-Odometric SLAM Using Wheels and a MEMS Gyroscope // *IEEE Access*. 2019. N 7. P. 97374–97389.

Сведения об авторах

- Жаафар Махмуд** — аспирант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: Jaafar.a.mahmoud@itmo.ru
- Вьонг Ха Тхе Лонг** — магистрант; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; E-mail: longvuonghathe@gmail.com
- Алексей Михайлович Бурков** — ПАО „Сбербанк“, лаборатория робототехники; ведущий инженер-разработчик; E-mail: amburkoff@gmail.com
- Сергей Алексеевич Колюбин** — д-р техн. наук, доцент; Университет ИТМО, факультет систем управления и робототехники, лаборатория биомехатроники и энергоэффективной робототехники; вед. научный сотрудник; E-mail: s.kolyubin@itmo.ru

Поступила в редакцию 28.01.21; одобрена после рецензирования 13.01.22; принята к публикации 18.01.22.

REFERENCES

1. Bailey T. and Durrant-Whyte H. *IEEE Robotics & Automation Magazine*, 2006, no. 3(13), pp. 108–117.
2. Campos C., Elvira R., Rodriguez J., Montiel M. et al. *IEEE Transactions on Robotics*, 2021, pp. 1–17.
3. Shi X., Li D. et al. *International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 3139–3145.
4. Bloesch M., Burri M., Omari S., Hutter M. and Siegwart R. *The International Journal of Robotics Research*, 2017, no. 10(36), pp. 1053–1072.
5. Geneva P., Eckenhoff K., Lee W., Yang Y. and Huang G. *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020.
6. Kummerle R., Grisetti G., Strasdat H., Konolige K. and Burgard W. *2011 IEEE International Conference on Robotics and Automation*, 2011.
7. Huang Sh. and Dissanayake G. *IEEE Transactions on Robotics*, 2007, no. 5(23), pp. 1036–1049.
8. Qin T., Li P. and Shen S. *IEEE Transactions on Robotics*, 2018, no. 4(34), pp. 1004–1020.
9. Usenko V., Demmel N., Schubert D., Stuckler J. and Cremers D. *IEEE Robotics and Automation Letters*, 2020, no. 2(5), pp. 422–429.
10. Quan M., Piao S., Tan M. and Huang S. *IEEE Access*, 2019, vol. 7, pp. 97374–97389.

Data on authors

- Jaafar Mahmoud** — Post-Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; E-mail: Jaafar.a.mahmoud@itmo.ru
- Vuong Ha The Long** — Graduate Student; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; E-mail: longvuonghathe@gmail.com
- Alexey M. Burkov** — Sberbank, Robotics Laboratory; Lead Engineer-Designer; E-mail: amburkoff@gmail.com
- Sergey A. Kolyubin** — Dr. Sci., Associate Professor; ITMO University, Faculty of Control Systems and Robotics, International Laboratory of Biomechatronic and Energy-Efficient Robotics; Leading Researcher; E-mail: s.kolyubin@itmo.ru

Received 28.01.21; approved after reviewing 13.01.22; accepted for publication 18.01.22.