

ПРИМЕНЕНИЕ КОМПЬЮТЕРНОГО ЗРЕНИЯ ДЛЯ УПРАВЛЕНИЯ РЕКОНФИГУРИРУЕМОЙ ИНТЕЛЛЕКТУАЛЬНОЙ ПОВЕРХНОСТЬЮ С ЦЕЛЬЮ РАСШИРЕНИЯ ЗОНЫ ПОКРЫТИЯ В СИСТЕМАХ МОБИЛЬНОЙ СВЯЗИ

Павлов Виталий Александрович,
Санкт-Петербургский политехнический университет Петра Великого
(СПбПУ), г. Санкт-Петербург, Россия, pavlov_va@spbstu.ru

Белов Андрей Александрович,
Санкт-Петербургский политехнический университет Петра Великого
(СПбПУ), г. Санкт-Петербург, Россия, belov@spbstu.ru

Завьялов Сергей Викторович,
Санкт-Петербургский политехнический университет Петра Великого
(СПбПУ), г. Санкт-Петербург, Россия, zavyalov_sv@spbstu.ru

Саянский Андрей Дмитриевич,
Университет ИТМО, г. Санкт-Петербург, Россия,
a.sayanskiy@metalab.ifmo.ru

Вабищевич Даниил Андреевич,
Университет ИТМО, г. Санкт-Петербург, Россия,
d.vabishchevich@metalab.ifmo.ru

DOI: 10.36724/2072-8735-2024-18-9-28-35

Manuscript received 10 August 2024;
Accepted 07 September 2024

Ключевые слова: реконфигурируемые
интеллектуальные поверхности, мобильная связь,
компьютерное зрение, сопровождение объектов,
обнаружение объектов

В статье рассматривается вариант решения задачи увеличения зоны покрытия в современных высокочастотных системах мобильной связи с помощью применения компьютерного зрения для управления диаграммами рассеяния отражательных реконфигурируемых интеллектуальных поверхностей (РИП). РИП фактически является отражательной антенной решеткой, состоящей из набора антенных элементов с индивидуальным управлением фазой отраженного каждым элементом сигнала, обеспечивающей формирование диаграмм рассеяния различной формы. Испытания подсистемы компьютерного зрения на тестовом наборе видеопоследовательностей, включающем различные количества людей, перемещающихся внутри помещения, показало высокое качество сопровождения абонентов по метрикам МОТА (80,85%) и МОТР (86,99%). Результаты работы разработанного прототипа комплекса для оптического управления диаграммой рассеяния РИП, использующего портативную энергоэффективную вычислительную платформу NVIDIA Jetson Nano показали высокую скорость обработки видео, достигающую до 15 кадров в секунду. Показано, что применение РИП, работающей под управлением системы компьютерного зрения, для создания канала связи в теневой зоне передатчика при работе в помещении сложной формы (Г-образный коридор), позволяет существенно увеличить мощность принимаемого сигнала (на величину, достигающую 15-20 дБ), делая ее достаточной для надежной работы абонентского оборудования.

Информация об авторах:

Павлов Виталий Александрович, к.т.н., доцент, Институт электроники и телекоммуникаций, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), г. Санкт-Петербург, Россия. Orcid: 0000-0003-0726-6613

Белов Андрей Александрович, ведущий инженер, Институт электроники и телекоммуникаций, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), г. Санкт-Петербург, Россия. Orcid: 0000-0003-0617-4514

Завьялов Сергей Викторович, к.т.н., доцент, Институт электроники и телекоммуникаций, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ), г. Санкт-Петербург, Orcid: 0000-0003-3398-3616

Саянский Андрей Дмитриевич, к.ф.-м.н., научный сотрудник, Физико-технический мегафакультет, Университет ИТМО, г. Санкт-Петербург. Orcid: 0000-0001-7689-9516

Вабищевич Даниил Андреевич, инженер, Физико-технический мегафакультет, Университет ИТМО, Санкт-Петербург. Orcid: 0009-0005-1400-3485

Для цитирования:

Павлов В.А., Белов А.А., Завьялов С.В., Саянский А.Д., Вабищевич Д.А. Применение компьютерного зрения для управления реконфигурируемой интеллектуальной поверхностью с целью расширения зоны покрытия в системах мобильной связи // Т-Comm: Телекоммуникации и транспорт. 2024. Том 18. №9. С. 28-35.

For citation:

Pavlov V.A., Belov A.A., Zavjalov S.V., Sayanskiy A.D., Vabishchevich D.A. (2024). Application of computer vision to control reconfigurable intelligent surface to extend coverage area in mobile communication systems. *T-Comm*, vol. 18, no.9 pp. 28-35. (in Russian)

Введение

С ростом количества абонентов и устройств, требующих высокую скорость передачи данных, повышается требование к увеличению полосы пропускания и качеству обслуживания. Одним из путей решения является переход на миллиметровый диапазон частот (например, стандарт 5G FR2) 1, который позволяет обеспечить большую пропускную способность и высокую скорость передачи данных. Однако, данный диапазон обладает существенным недостатком, поскольку во время передачи данных по радиоканалу внутри и снаружи помещений возникают значительные затенения сигнала строительными конструкциями, приводящие к появлению теневых зон с недостаточным уровнем сигнала. Это касается и других стандартов связи, например, WiFi, который переходит на использование повышенных рабочих частоты (5-7 ГГц), что приводит к экранированию сигнала точек доступа стенами при размещении внутри зданий. Одним из путей решения данной проблемы является использование активных и пассивных ретрансляторов.

В работе 2 для усиления сигнала в миллиметровом диапазоне частот был использован активный ретранслятор, устанавливаемый внутри помещений. Это позволило повысить среднюю пропускную способность и снизить коэффициент битовых ошибок. Другим возможным решением проблемы является использование реконфигурируемой интеллектуальной поверхности (РИП). РИП содержит множество пассивных элементов, которые позволяют изменять направление отражения радиоволн с помощью фазовых сдвигов. Пассивная РИП является недорогим и энергоэффективным решением для увеличения зоны обслуживания и позволяет увеличивать мощность сигнала у абонента, расширять зоны покрытия и создавать в помещениях виртуальные линии прямой видимости между приемопередатчиками 3.

Перспективным направлением для оптимального управления реконфигурируемыми антенными решетками является использование алгоритмов компьютерного зрения на основе глубокого обучения 4–9. Такой подход позволяет оптимально перенаправить сигнал базовой станции на мобильное устройство пользователя путем формирования диаграмм направленности ретранслятора с использованием видеоинформации. Такие ретрансляторы могут работать независимо от базовой станции и располагаться на зданиях и беспилотных летательных аппаратах. Для эффективного использования система должна иметь низкое энергопотребление и в то же время высокую точность локализации и отслеживания пользователей.

В работах 4, 5 показан пример сопряжения системы компьютерного зрения на базе алгоритма FairMOT 14 и РИП, работающей в диапазоне частот миллиметровых волн, вне помещения. Для определения наиболее подходящей диаграммы направленности авторы в 4 применяют генетический алгоритм. В 5 показано, что при использовании таких подходов и высокой загрузке сети можно снизить средний уровень взаимных помех на 40%. Таким образом, можно сделать вывод о перспективности использования комбинированного подхода и с точки зрения снижения взаимных помех.

В 6 авторы сформулировали парадигму, лежащую на стыке компьютерного зрения и беспроводной связи, позволяющую создавать критически важные приложения, выходящие за рамки 5G (автономные/дистанционно управляемые транспортные средства, визуально-тактильная виртуальная реальность и другие киберфизические приложения). Авторы рассмотрели

различные сценарии совместной работы алгоритмов компьютерного зрения и беспроводной связи и продемонстрировали, как компьютерное зрение позволяет прогнозировать сценарий блокировки канала миллиметрового диапазона до того, как блокировка действительно произойдет.

В 7 приведен пример моделирования применения компьютерного зрения для формирования диаграммы направленности mmWave MIMO на синтетическом наборе данных 11. В работе предложена структура сети, состоящая из ResNet 12, 3D ResNext 13, модуля объединения признаков (FFM) 14 и прогнозирующей сети LSTM 15. Результаты моделирования показали, что визуальные данные могут значительно повысить точность прогнозирования луча.

В 8 в качестве алгоритма компьютерного зрения для управления РИП используется YOLOv3 17. Моделирование проводилось на синтетическом наборе данных 11. Авторы показали, что предлагаемый выбор луча РИП с помощью визуального контроля может обеспечить почти оптимальную достижимую скорость со значительно меньшими затратами на обучение луча. Это подчеркивает потенциал использования визуальной информации для улучшения работы систем связи с помощью РИП.

В 9 в качестве оптического сенсора используется стереокамера. Для обнаружения объектов применяется YOLOX 10. Размер РИП 20×20, частота 5,4 ГГц. Результаты экспериментов показывают, что РИП может быстро корректировать коэффициенты отражения для динамического отслеживания абонента с помощью визуальной информации. Авторы делают предположение о перспективности применения лидаров и инфракрасных датчиков в ночное время. К недостатку данного подхода следует отнести увеличенную сложность и падение точности результатов стереообработки с увеличением расстояния до объектов обнаружения.

Из обзора и анализа литературы 4–9 можно сделать вывод, что задача сопряжения алгоритмов компьютерного зрения и РИП является перспективной и относительно мало проработанной, в то же время, на данный момент не описано экспериментально проверенного прототипа подобной системы, который сочетал бы в себе высокую точность и низкое энергопотребление. Отдельно стоит отметить актуальность исследования сопряжения алгоритмов компьютерного зрения и РИП применительно к работе внутри помещения. Данная проблема не исследована в должной мере и представляет, как научный, так и практический интерес.

Цель данной работы заключается в разработке и исследовании РИП, управляемой с использованием алгоритмов компьютерного зрения.

Описание установки

Для отработки элементов концепции был создан прототип пассивного ретранслятора, работающего на частоте 5 ГГц. Структура используемой РИП, управляемой с использованием алгоритмов компьютерного зрения показана на рисунке 1. Система состоит из источника сигнала, камеры, РИП и встраиваемого вычислительного модуля для обработки видеопотока и управления РИП.

Однобитная РИП работает на центральной частоте 5 ГГц и состоит из 10×10 патч-антенн, фаза которых индивидуально управляется варакторными диодами. Решетка является частью большей антенны размером 20×20, подробно описанной в 18.

Для управления системой и обработки видеопотока используется одноплатный вычислительный модуль NVIDIA Jetson Nano 19. Этот модуль содержит 128-ядерный графический сопроцессор, который позволяет ускорить нейросетевую обработку видеопотока для оценки координат пользователей. Координаты пользователей используются для расчета фаз элементов РИП и направления луча на пользователей

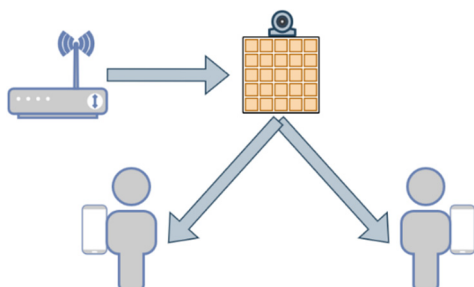


Рис. 1. Структура предложенной РИП, управляемая с помощью системы компьютерного зрения

Обнаружение и сопровождение

Анализ литературы 20 и собственный опыт авторов показали, что архитектура нейронной сети TinyYOLOv7 20 весьма перспективна для обработки видео с использованием портативной энергоэффективной вычислительной платформы NVIDIA Jetson Nano. Этот нейросетевой детектор имеет низкую вычислительную сложность — 5,8 GFLOPS (при входном разрешении $416 \times 416 \times 3$ пикселей), его средняя точность (AP) в наборе данных MS COCO 21 составляет 35,2%. При реализации на платформе NVIDIA Jetson Nano скорость обработки составляет около 15 кадров в секунду при разрешении входного видео $416 \times 416 \times 3$ пикселей.

Для сопровождения абонента, на основе анализа литературы, был выбран алгоритм сопровождения ByteTrack 22. Алгоритм совместной работы алгоритма обнаружения и сопровождения состоит из 4 этапов (обнаружение, двухстадийная ассоциация данных и управление жизненным циклом траекторий):

- Обнаружение: алгоритм обнаружения идентифицирует объекты в каждом кадре и присваивает каждому коэффициент достоверности.

- Ассоциация с высокой степенью достоверности: ByteTrack сначала принимает обнаруженные данные с высокой степенью достоверности и связывает их между кадрами, используя модель движения на основе фильтра Калмана и алгоритма назначения. Этот шаг связывает обнаружения одного и того же объекта в разных кадрах и создаёт траекторию для каждого объекта.

- Ассоциация с низкой достоверностью. После связывания обнаружений с высокой достоверностью ByteTrack включает обнаружения с низкой достоверностью, которые обычно игнорируются. Он связывает эти обнаружения с существующими траекториями, если они соответствуют ожидаемому местоположению и внешнему виду отслеживаемого объекта. Этот шаг имеет решающее значение для сохранения идентичности объектов, когда они закрыты или плохо видны.

- Управление жизненным циклом траекторий: ByteTrack управляет созданием и удалением траекторий, инициализируя новые траектории для обнаруженных объектов, которые не соответствуют существующим

Результаты сопровождения

Для количественной оценки системы компьютерного зрения был использован набор данных из 23, 24. Предварительно были отобраны и размечены 10 видеопоследовательностей с различными ракурсами, сценариями поведения объектов и расположением камеры под потолком. Описание видеопоследовательностей представлено в таблице 1.

Таблица 1

Описание видеопоследовательностей

Номер	Описание
1	Всего объектов: 2. Максимальное количество объектов в кадре: 2. Есть длительное исчезновение объекта из поля зрения камеры.
2	Всего объектов: 2. Максимальное количество объектов в кадре: 2. Есть длительное исчезновение объекта из поля зрения камеры.
3	Всего объектов: 2. Максимальное количество объектов в кадре: 2. Есть длительное исчезновение объекта из поля зрения камеры.
4	Всего объектов: 2. Максимальное количество объектов в кадре: 2. Есть длительное исчезновение объекта из поля зрения камеры и пересечение траекторий
5	Всего объектов: 2. Максимальное количество объектов в кадре: 2. Есть длительное исчезновение объекта из поля зрения камеры.
6	Всего объектов: 14. Максимальное количество объектов в кадре: 3.
7	Всего объектов: 4. Максимальное количество абонентов в кадре: 4. Есть пересечение траекторий
8	Всего объектов: 6. Максимальное количество абонентов в кадре: 6. Есть длительное исчезновение объекта из поля зрения камеры и пересечение траекторий.
9	Всего объектов: 14. Максимальное количество абонентов в кадре: 6. Есть пересечение траекторий.
10	Всего объектов: 14. Максимальное количество объектов в кадре: 5. Есть пересечение траекторий и кратковременные окклюзии.

В таблице 2 представлены количественные оценки системы компьютерного зрения с помощью специализированных метрик. Подробное описание метрик представлено в 25.

Таблица 2

Описание видеопоследовательностей

Параметр (усреднённое значение по 10 видеопоследовательностям)	TinyYOLOv7 + ByteTrack
IDF1	80,96 %
IDP	87,40 %
IDR	75,60 %
Recall	83,90 %
Precision	96,80 %
MOTA	80,85 %
MOTP	86,99 %

Примеры работы TinyYOLOv7 + ByteTrack приведены на рисунках 2 а-г.

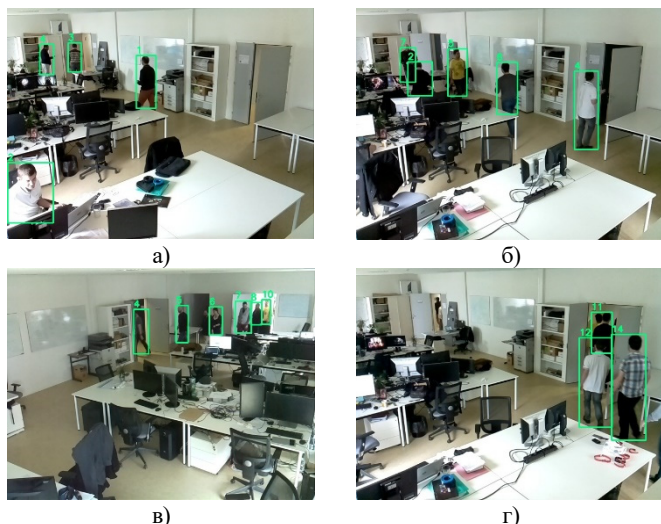


Рис. 2. Примеры работы подсистемы компьютерного зрения

Система координат

По умолчанию в изображении пространственные координаты задаются с помощью внутренней системы координат, в которой началом является точка в верхнем левом углу изображения. Эта система соответствует пиксельным индексам изображения. Для преобразования координат объекта в прямоугольную систему координат, а затем в сферическую необходимо выполнить шаги, которые будут описаны ниже.

На первом этапе выполняется нормировка координат по высоте и ширине изображения: $x_{0norm} = x_0/w$ и $y_{0norm} = y_0/h$. Затем необходимо преобразовать диапазон значений и сместить центр системы координат в центр изображения: $r \in [0, 1]$ – старый диапазон и $t \in [-1, 1]$ – новый диапазон.

$$\begin{aligned} x_{new} &= \frac{x_{0norm} - r_{min}}{r_{max} - r_{min}} \times (t_{max} - t_{min}) + t_{min} \\ y_{new} &= \frac{y_{0norm} - r_{min}}{r_{max} - r_{min}} \times (t_{max} - t_{min}) + t_{min} \end{aligned} \quad (1)$$

Для расчёта фаз необходимо найти углы θ и φ по следующим формулам:

$$\theta = \left(\sqrt{x^2 + y^2} \right) \cdot 50^\circ \quad (2)$$

где 50° это угол обзора камеры (определяется по результатам тестирования камеры).

$$\varphi = \arctg\left(\frac{y}{x}\right) + \frac{180^\circ}{\pi} \quad (3)$$

Рабочий цикл и режимы работы

Рабочий цикл можно разделить на следующие этапы:

1. Считывание параметров для инициализации.
2. Инициализация:
 - камера: высота, ширина, кадровая частота;
 - нейросетевой детектор: загрузка конфигурации и обученных весовых коэффициентов;
 - координаты передатчика в сферической системе координат: θ , φ , ρ

3. Захват кадра.
4. Передача кадра в нейросетевой детектор и формирование выходных данных:

- координаты ограничивающих объект прямоугольников (bounding boxes),
- номера классов,
- вероятности принадлежности объектов к заданным классам.

5. Формирование координат объекта. Данный этап зависит от режима работы:

• Режим ближнего. Осуществляется определение координат ближнего объекта.

• Режим дальнего. Осуществляется определение координат дальнего объекта.

• Режим сопровождения с ручным выбором. Из множества сопровождаемых системой одновременно объектов, пользователь вручную выбирает объект, координаты которого будут отслеживаться в дальнейшем.

• Режим сопровождения одного объекта с автоматическим выбором (инициализация с помощью «ловушки»). Из множества сопровождаемых одновременно объектов, выбирается тот объект, который попал в область инициализации («ловушка») заданную пользователем. Координаты этого объекта будут отслеживаться в дальнейшем до его выхода из зоны работы системы.

6. Коррекция координат объекта с учётом искажений, вносимых объективом камеры.

7. Преобразование скорректированных координат объекта в сферическую систему координат.

8. Передача углов θ , φ , ρ в алгоритм формирования луча.

9. Расчёт фазовых состояний элементов решетки

10. Передача данных на решетку по ИК каналу связи.

11. Возврат к пункту 3.

Проведение испытаний полной системы

На рисунке 3 представлена фотография демонстрационного стенда, включающего в себя разработанный программно-аппаратный комплекс, и монитор для визуального отображения поля зрения видеокамеры и сопровождаемых абонентов.



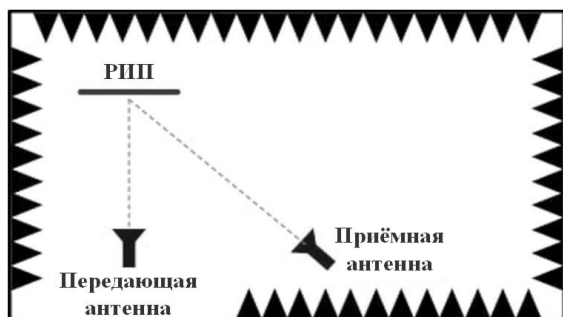
Рис. 3. Фотография демонстрационного стенда

Тестирование работы комплекса осуществлялось в несколько этапов:

- 1) Тестирование в условиях безэховой камеры (БЭК)
- 2) Тестирование в реалистичных условиях.

На первом этапе производилось тестирование в условиях БЭК для минимизации паразитных переотражений от

окружающих предметов. Фотография экспериментального стенда приведена на рисунке 4. Передающая антенна располагалась на расстоянии 4 метров от РИП. Приемная антенна располагалась на расстоянии 3 метра от РИП. Конфигурация расположения антенн была выбрана таким образом, чтобы минимизировать прямой сигнал.



а)



б)

Рис. 4. а) Схема эксперимента;
б) Фотография экспериментального стенда

На рисунке 5 показана зависимость уровня сигнала, принятого стационарным приемником, установленном под углом 32 градусов к РИП при разных угловых положениях абонента.

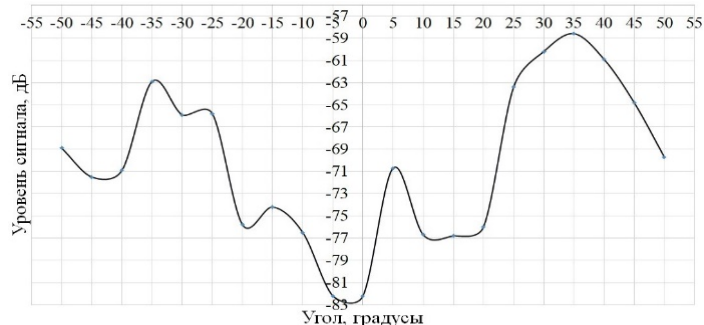
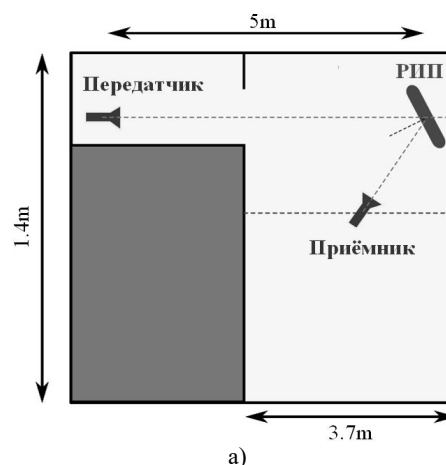


Рис. 5. Зависимость уровня сигнала на стационарном приемнике, располагающемся под углом 32 градусов к «стенке» при разных угловых положениях абонента в БЭК

Фактически, полученные результаты демонстрируют диаграмму направленности отражательной антенной решетки, а также наглядно демонстрируют корректность работы комплекса. Видно резкое увеличение уровня принимаемого сигнала при угловом положении абонента, соответствующему положению приемной антенны, а также высокий лепесток квантования 1-битной решетки, расположенный зеркально относительно нормали.

На следующем этапе проверялась корректность работы разработанного комплекса в условиях реального помещения, показанного на рисунках 6 и 7. Приемник и передатчик частотно-модулированного сигнала были построены на основе программно-определяемой радиосистемы HackRF и одноплатных компьютеров Raspberry Pi4. Целью эксперимента была демонстрация непрерывности передачи при перемещении абонента с приемником. Во всех случаях расположение передающей антенны выбиралось из требований уменьшения прямого сигнала до уровня, не обеспечивающего надежный прием сигнала.

Результаты моделирования зоны покрытия демонстрационного стенда №1 при одном из положений луча приведены на рисунке 7. При моделировании использовались следующие параметры оборудования — частота 5 ГГц, мощность передатчика 100 мВт, коэффициент усиления передающей антенны 6 дБ, коэффициент усиления приемной антенны 6 дБ.



а)



б)

Рис. 6. а) Схема эксперимента;
б) Фотография демонстрационного стенда №1

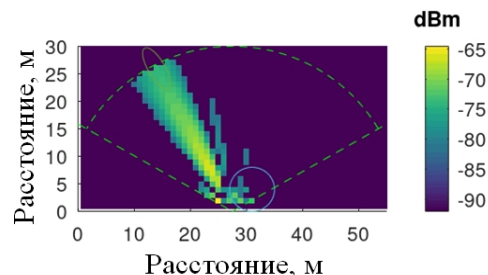


Рис. 7. Результаты моделирования зоны покрытия демонстрационного стенда №1 при одном из положений луча

На рисунке 8 показана зависимость уровня сигнала, принятого стационарным приемником, установленным под углом 34 градуса к РИП при разных угловых положениях абонента. Источник находится на расстоянии от решётки – 5м, приёмник – 3м.

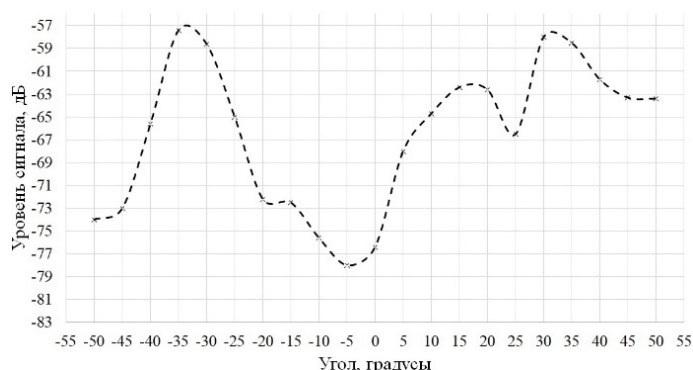


Рис. 8. Зависимость уровня сигнала на стационарном приемнике, располагающемся под углом 32 градусов к «стенке» при разных угловых положениях абонента в помещении

Заключение

В данной статье рассмотрено применение реконфигурируемой интеллектуальной поверхности, управляемой системой компьютерного зрения для создания канала связи в теневой зоне передатчика, в которой не обеспечивается мощность сигнала, достаточная для надежной работы абонентского оборудования. Связь обеспечивается за счет переотражения сигнала передатчика в направлении абонента. Для определения пространственного положения абонента и слежения за его изменениями используются методы компьютерного зрения. Измерения, проведенные на разработанном авторами экспериментальном стенде, продемонстрировали работоспособность данного подхода в условиях помещений сложной формы. Увеличение мощности принятого сигнала в теневой зоне передатчика составило от 15 до 20 дБ, обеспечивая его надежный прием.

Авторами работы был разработан и впервые экспериментально исследован прототип системы компьютерного зрения для управления реконфигурируемой интеллектуальной поверхностью с целью расширения зоны покрытия в системах мобильной связи, который сочетает в себе высокую точность и низкое энергопотребление. Технологии, использованные в данном прототипе, могут быть использованы и для наружного применения в системах связи 5 и 6 поколения.

Испытания подсистемы компьютерного зрения на тестовом наборе данных показало высокое качество сопровождения абонентов по метрикам МОТА 80,85% и МОТР 86,99%. Результаты тестирования разработанного прототипа на портативной энергоэффективной вычислительной платформе Jetson Nano показали высокую скорость обработки видео, достигающую до 15 кадров в секунду при энергопотреблении, не превышающем 10 Вт.

Литература

1. Busari S.A., Huq K.M.S., Mumtaz S., Dai L., Rodriguez J. Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey // IEEE Communications Surveys & Tutorials, vol. 20, no. 2, pp. 836-869, Secondquarter 2018, doi: 10.1109/COMST.2017.2787460.

2. Randy V-P, Alonso J.I. Performance Analysis of Two-Hop mmWave Relay Nodes over the 5G NR Uplink Signal // Applied Sciences, vol 11, no. 13, 5828, 2021. doi:10.3390/app11135828.

3. Pan C. et al. Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions // IEEE Communications Magazine, vol. 59, no. 6, pp. 14-20, June 2021, doi: 10.1109/MCOM.001.2001076.

4. Loscri V., Rizza C., Benslimane A., Vegni A.M., Innocenti E., Giuliano R. BEST-RIM: A mmWave Beam Steering Approach Based on Computer Vision-Enhanced Reconfigurable Intelligent Metasurfaces // IEEE Transactions on Vehicular Technology, vol. 72, no. 6, pp. 7613-7626, June 2023, doi: 10.1109/TVT.2023.3243358.

5. Loscri V., Vegni A.M., Innocenti E., Giuliano R., Mazzenga F. A joint Computer Vision and Reconfigurable Intelligent Meta-surface Approach for Interference Reduction in Beyond 5G Networks // 2021 IEEE 22nd International Conference on High Performance Switching and Routing (HPSR), Paris, France, 2021, pp. 1-6, doi: 10.1109/HPSR52026.2021.9481794.

6. Nishio T., Koda Y., Park J., Bennis M., Doppler K. When Wireless Communications Meet Computer Vision in Beyond 5G // IEEE Communications Standards Magazine, vol. 5, no. 2, pp. 76-83, June 2021, doi: 10.1109/MCOMSTD.001.2000047.

7. Tian Y., Pan G., Alouini M.-S. Applying Deep-Learning-Based Computer Vision to Wireless Communications: Methodologies, Opportunities, and Challenges // IEEE Open Journal of the Communications Society, vol. 2, pp. 132-143, 2021, doi: 10.1109/OJCOMS.2020.3042630.

8. Jiang S., Hindy A., Alkhateeb A. Camera Aided Reconfigurable Intelligent Surfaces: Computer Vision Based Fast Beam Selection // ICC 2023 - IEEE International Conference on Communications, Rome, Italy, 2023, pp. 2921-2926, doi: 10.1109/ICC45041.2023.10279112.

9. Ouyang M., Gao F., Wang Y., Zhang S., Li P., Ren J. Computer Vision-Aided Reconfigurable Intelligent Surface-Based Beam Tracking: Prototyping and Experimental Results // IEEE Transactions on Wireless Communications, vol. 22, no. 12, pp. 8681-8693, Dec. 2023, doi: 10.1109/TWC.2023.3264752.

10. Ge Z., Liu S., Wang F., Li Z., Sun J. YOLOX: Exceeding YOLO series in 2021 // Preprint arXiv. 2021. Source: <https://arxiv.org/abs/2107.08430>.

11. Alrabeiah M., Booth J., Hredzak A., Alkhateeb A. ViWi vision-aided mmWave beam tracking: Dataset task and baseline solutions // Preprint arXiv. 2020. Source: <http://arXiv:2002.02445>.

12. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

13. Hara K., Kataoka H., Satoh Y. Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet? // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 2018 pp. 6546-6555. doi: 10.1109/CVPR.2018.00685.

14. Wang B., et al. Controllable Video Captioning With POS Sequence Guidance Based on Gated Fusion Network // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019 pp. 2641-2650. doi: 10.1109/ICCV.2019.00273.

15. Hochreiter S., Schmidhuber J. Long Short-Term Memory // Neural Computation, 1997, vol. 9, no. 8, pp. 1735-1780. doi: 10.1162/neco.1997.9.8.1735.

16. Zhang Y., Wang C., Wang X. et al. FairMOT: On the fairness of detection and re-identification in multiple object tracking // International Journal of Computer Vision, 2021, vol. 129, no. 11, pp. 3069-3087. doi: 10.1007/s11263-021-01513-4

17. Redmon J., Farhadi A. YOLOv3: An incremental improvement // Preprint arXiv. 2018. Source: <https://arxiv.org/abs/1804.02767>.

18. Sayanskiy A. et al. A 2D-Programmable and Scalable Reconfigurable Intelligent Surface Remotely Controlled via Digital Infrared Code // IEEE Transactions on Antennas and Propagation, vol. 71, no. 1, pp. 570-580, Jan. 2023, doi: 10.1109/TAP.2022.3217327.

19. [online] Available: <https://www.nvidia.com>.

20. Wang C.-Y., Bochkovskiy A., Liao H.-Y.M. YOLOv7:

Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors // 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 7464-7475, doi: 10.1109/CVPR52729.2023.00721.

21. Microsoft Common Objects in Context dataset cite: <https://cocodataset.org/>

22. Zhang Y. et al. ByteTrack: Multi-object Tracking by Associating Every Detection Box // Avidan, S., Brostow, G., Cissé, M., Fari-nella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13682. Springer, Cham. https://doi.org/10.1007/978-3-031-20047-2_1.

23. Marroquin R., Dubois J., Nicolle C. WiseNET: An indoor multi-camera multi-space dataset with contextual information and annotations for people detection and tracking // Data in Brief, 2019, vol. 27, pp. 1-14. doi: 10.1016/j.dib.2019.104654.

24. WiseNET: Multi-camera dataset cite: https://data.4tu.nl/articles/_/12714416/1

25. Dendorfer P., Ošep A., Milan A. et al., "MOTChallenge: A Benchmark for Single-Camera Multiple Target Tracking // International Journal of Computer Vision; 2021; vol. 129, pp. 845-881. doi: 10.1007/s11263-020-01393-0

APPLICATION OF COMPUTER VISION TO CONTROL RECONFIGURABLE INTELLIGENT SURFACE TO EXTEND COVERAGE AREA IN MOBILE COMMUNICATION SYSTEMS

Vitalii A. Pavlov, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia, pavlov_va@spbstu.ru

Andrey A. Belov, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia, belov@spbstu.ru

Sergey V. Zavjalov, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia, zavjalov_sv@spbstu.ru

Andrey D. Sayanskiy, ITMO University, St. Petersburg, Russia, a.sayanskiy@metalab.ifmo.ru

Daniil A. Vabishchevich, ITMO University, St. Petersburg, Russia, d.vabishchevich@metalab.ifmo.ru

Abstract

The paper considers a variant of solving the problem of increasing the coverage area in modern high-frequency mobile communication systems by using computer vision to control the scattering patterns of reflective reconfigurable intelligent surfaces (RIS). The RIS is actually a reflective antenna array consisting of a set of antenna elements with individual control of the phase of the signal reflected by each element, providing the formation of scattering patterns of different shapes. Tests of the computer vision subsystem on a test set of video sequences, including different numbers of people moving inside the room, showed high quality of subscriber tracking by MOTA (80.85%) and MOTP (86.99%) metrics. The results of the developed prototype of the complex for optical control of the scattering pattern of RIS, using a portable energy-efficient computing platform NVIDIA Jetson Nano showed a high speed of video processing, up to 15 frames per second. It has been shown that the use of a computer vision system to create a communication channel in the shadow zone of the transmitter when working in a complex-shaped room (?-shaped corridor), allows to significantly increase the received signal power (by a value reaching 15-20 dB), making it sufficient for reliable operation of subscriber equipment.

Keywords: reconfigurable intelligent surface (RIS), mobile communications, computer vision, object tracking, object detection.

References

- [1] S. A. Busari, K. M. S. Huq, S. Mumtaz, L. Dai and J. Rodriguez, "Millimeter-Wave Massive MIMO Communication for Future Wireless Systems: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 836-869, Secondquarter 2018, doi: 10.1109/COMST.2017.2787460.
- [2] V-P Randy, and J. I. Alonso. "Performance Analysis of Two-Hop mmWave Relay Nodes over the 5G NR Uplink Signal," *Applied Sciences*, vol 11, no. 13, 5828, 2021. doi:10.3390/app11135828.
- [3] C. Pan et al., "Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14-20, June 2021, doi: 10.1109/MCOM.001.2001076.

- [4] V. Loscri, C. Rizza, A. Benslimane, A. M. Vegni, E. Innocenti and R. Giuliano, "BEST-RIM: A mmWave Beam Steering Approach Based on Computer Vision-Enhanced Reconfigurable Intelligent Metasurfaces," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 6, pp. 7613-7626, June 2023, doi: 10.1109/TVT.2023.3243358.
- [5] V. Loscri, A. M. Vegni, E. Innocenti, R. Giuliano and F. Mazzenga, "A joint Computer Vision and Reconfigurable Intelligent Meta-surface Approach for Interference Reduction in Beyond 5G Networks," *2021 IEEE 22nd International Conference on High Performance Switching and Routing (HPSR)*, Paris, France, 2021, pp. 1-6, doi: 10.1109/HPSR52026.2021.9481794.
- [6] T. Nishio, Y. Koda, J. Park, M. Bennis and K. Doppler, "When Wireless Communications Meet Computer Vision in Beyond 5G," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 76-83, June 2021, doi: 10.1109/MCOMSTD.001.2000047.
- [7] Y. Tian, G. Pan and M. -S. Alouini, "Applying Deep-Learning-Based Computer Vision to Wireless Communications: Methodologies, Opportunities, and Challenges," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 132-143, 2021, doi: 10.1109/OJCOMS.2020.3042630.
- [8] S. Jiang, A. Hindy and A. Alkhateeb, "Camera Aided Reconfigurable Intelligent Surfaces: Computer Vision Based Fast Beam Selection," *ICC 2023 - IEEE International Conference on Communications*, Rome, Italy, 2023, pp. 2921-2926, doi: 10.1109/ICC45041.2023.10279112.
- [9] M. Ouyang, F. Gao, Y. Wang, S. Zhang, P. Li and J. Ren, "Computer Vision-Aided Reconfigurable Intelligent Surface-Based Beam Tracking: Prototyping and Experimental Results," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 8681-8693, Dec. 2023, doi: 10.1109/TWC.2023.3264752.
- [10] Z. Ge, S. Liu, F. Wang, Z. Li and J. Sun, "YOLOX: Exceeding YOLO series in 2021," *Preprint arXiv*. 2021. Source: <https://arxiv.org/abs/2107.08430>.
- [11] M. Alrabeiah, J. Booth, A. Hredzak, A. Alkhateeb, "ViWi vision-aided mmWave beam tracking: Dataset task and baseline solutions," *Preprint arXiv*. 2020. Source: <http://arXiv:2002.02445>.
- [12] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [13] K. Hara, H. Kataoka and Y. Satoh, "Can Spatiotemporal 3D CNNs Retrace the History of 2D CNNs and ImageNet?," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, UT, USA, 2018 pp. 6546-6555. doi: 10.1109/CVPR.2018.00685.
- [14] B. Wang, et al., "Controllable Video Captioning With POS Sequence Guidance Based on Gated Fusion Network," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 2019 pp. 2641-2650. doi: 10.1109/ICCV.2019.00273.
- [15] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, 1997, vol. 9, no. 8, pp. 1735-1780. doi: 10.1162/neco.1997.9.8.1735.
- [16] Y. Zhang, C. Wang, X. Wang and et al., "FairMOT: On the fairness of detection and re-identification in multiple object tracking," *International Journal of Computer Vision*, 2021, vol. 129, no. 11, pp. 3069-3087. doi: 10.1007/s11263-021-01513-4
- [17] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," *Preprint arXiv*. 2018. Source: <https://arxiv.org/abs/1804.02767>.
- [18] A. Sayanskiy et al., "A 2D-Programmable and Scalable Reconfigurable Intelligent Surface Remotely Controlled via Digital Infrared Code," *IEEE Transactions on Antennas and Propagation*, vol. 71, no. 1, pp. 570-580, Jan. 2023, doi: 10.1109/TAP.2022.3217327.
- [19] [online] Available: <https://www.nvidia.com>.
- [20] C.-Y. Wang, A. Bochkovskiy and H. -Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 2023, pp. 7464-7475, doi: 10.1109/CVPR52729.2023.00721.
- [21] Microsoft Common Objects in Context dataset cite: <https://cocodataset.org/>
- [22] Y. Zhang and et al., "ByteTrack: Multi-object Tracking by Associating Every Detection Box," In: Avidan, S., Brostow, G., Ciss, M., Farinella, G.M., Hassner, T. (eds) *Computer Vision - ECCV 2022*. ECCV 2022. Lecture Notes in Computer Science, vol 13682. Springer, Cham. https://doi.org/10.1007/978-3-031-20047-2_1.
- [23] R. Marroquin, J. Dubois and C. Nicolle, "WiseNET: An indoor multi-camera multi-space dataset with contextual information and annotations for people detection and tracking," *Data in Brief*, 2019, vol. 27, pp. 1-14. doi: 10.1016/j.dib.2019.104654.
- [24] WiseNET: Multi-camera dataset cite: https://data.4tu.nl/articles/_/12714416/
- [25] P. Dendorfer, A. Os'ep, A. Milan and et al., "MOTChallenge: A Benchmark for Single-Camera Multiple Target Tracking," *International Journal of Computer Vision*, 2021, vol. 129, pp. 845-881. doi: 10.1007/s11263-020-01393-0

Information about authors:

Vitalii A. Pavlov, Associate professor, PhD, Institute of Electronics and Telecommunications, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia. Orcid: 0000-0003-0726-6613

Andrey A. Belov, Lead Engineer, Institute of Electronics and Telecommunications, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia. Orcid: 0000-0003-0617-4514

Sergey V. Zavjalov, Associate professor, PhD, Institute of Electronics and Telecommunications, Peter the Great St.Petersburg Polytechnic University, Saint-Petersburg, Russia. Orcid: 0000-0003-3398-3616

Andrey D. Sayanskiy, Researcher, PhD, School of Physics and Engineering, ITMO University, St. Petersburg, Russia. Orcid: 0000-0001-7689-9516

Daniil A. Vabishchevich, Engineer, School of Physics and Engineering, ITMO University, St. Petersburg, Russia. Orcid: 0009-0005-1400-3485