



АНАЛИТИЧЕСКИЙ ОБЗОР СИСТЕМ ДОПОЛНЕННОЙ И СМЕШАННОЙ РЕАЛЬНОСТИ В КОНТЕКСТЕ ИНДУСТРИИ 4.0

А. А. Голубкин¹, Р. Я. Пирмагомедов^{2*}

¹Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

²Университет Тампере, г. Тампере, 33720, Финляндия

*Адрес для переписки: rustam.pirmagomedov@tuni.fi

Аннотация—Системы дополненной реальности приобретают популярность для решения задач в области оптимизации производства и обучения сотрудников. В статье рассматриваются системы дополненной реальности в контексте Индустрии 4.0. В работе был проведен анализ, обобщение и синтез научных публикаций по указанной теме. Данная статья представляет актуальную информацию о научном и технологическом состоянии дел в области использования систем дополненной реальности на промышленных предприятиях, которая может быть полезна как научным, так и инженерным работникам.

Ключевые слова—дополненная реальность, смешанная реальность, индустрия 4.0.

Информация о статье

УДК 338.24.01

Язык статьи – русский.

Поступила в редакцию 25.10.2021, принята к печати 20.12.2021.

Ссылка для цитирования: Голубкин А. А., Пирмагомедов Р. Я. Аналитический обзор систем дополненной и смешанной реальности в контексте индустрии 4.0 // Информационные технологии и телекоммуникации. 2021. Том 9. № 4. С. 1–27. DOI 10.31854/2307-1303-2021-9-4-1-27.



ANALYTICAL REVIEW OF AUGMENTED AND MIXED REALITY SYSTEMS IN THE CONTEXT OF INDUSTRY 4.0

A. Golubkin¹, R. Pirmagomedov^{2*}

¹The Bonch-Bruevich Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

²Tampere University, PO Box 774 FI-00074 CGI Finland

*Corresponding author: rustam.pirmagomedov@tuni.fi

Abstract—Augmented reality systems are gaining popularity for solving problems in the field of production optimization and employee training. The article discusses augmented reality systems in the context of Industry 4.0. The analysis, generalization and synthesis of scientific publications on this topic was carried out. This article presents up-to-date information about the scientific and technological state of affairs in the field of the use of augmented reality systems in industrial enterprises, which can be useful to both scientific and engineering workers.

Keywords—augmented reality, mixed reality, industry 4.0.

Article info

Article in Russian.

Received 25.10.2021, accepted 20.12.2021.

For citation: Golubkin A., Pirmagomedov R.: Analytical Review of Augmented and Mixed Reality Systems in the Context of Industry 4.0 // Telecom IT. 2021. Vol. 9. Iss. 4. pp. 1–27 (in Russian). DOI 10.31854/2307-1303-2021-9-4-1-27.



Введение

С каждым годом промышленные предприятия расширяются и становятся более технологичными. Появляется все больше сложного и дорогостоящего оборудования. Считается, что системы дополненной реальности могут помочь в оптимизации производства и обучении работников, что может исправить дефицит высококвалифицированных кадров. Из-за этого, все больше компаний прибегают к дополненной реальности для решения внутренних задач. Также, дополненная реальность все чаще применяется в относительно новой области здравоохранения – «Internet of Medical Things», для помощи при диагностике заболеваний [1].

Однако, текущие системы имеют множество ограничений и нерешенных проблем, такие как задержка, отсутствие мобильности, нагрузка на вычислительное оборудование. Трафик, создаваемый современными мобильными устройствами, демонстрирует высокую временную и пространственную динамику. В результате, традиционные подходы к предоставлению услуг могут оказаться недостаточными в системах за пределами пятого поколения [2]. Поэтому до сих пор ведутся исследования по выбору общей архитектуры и технологий для реализации. Одним из вариантов решений было предложено создать эффективные сеточные топологии на основе mmWave, в которых используются передовые механизмы множественного подключения для улучшения уровней подключения [3]. Также предлагается модель трафика дополненной реальности, построенная в зависимости от изменения местоположения пользователя [4].

Можно выделить основные 7 критерии классификации систем дополненной реальности:

- 1) Тип представления информации.
- 2) Способ распознавания объекта.
- 3) Способ взаимодействия с пользователем.
- 4) Тип мобильности.
- 5) Целевое назначение;
- 6) Тип безопасности.
- 7) Степень воздействия на окружающую среду [5].

В данной работе был произведен обзор научных статей, связанных с системами дополненной реальности в контексте «Индустрин 4.0». Рассматриваются основные типы систем, примеры реализации, а также анализируются причины неудач в данной области. В первой части рассматривается общая информация по системам дополненной реальности. Во второй части затрагиваются маркерные системы. Третья часть описывает безмаркерные системы. В заключительной части затрагивается интерфейс мозг-компьютер.

Дополненная реальность в контексте индустрии 4.0

В контексте Индустрии 4.0 проектирование и оценка эксплуатационных показателей промышленного оборудования представляют собой решающий фактор, который может быть поддержан новыми технологиями, такими как дополненная реальность (AR) и смешанная реальность (MR), с помощью которых можно радикально обновить и улучшить производственные процессы. [6]



Дополненная реальность (AR) – это прямое или косвенное представление о физической среде реального мира, элементы которого дополняются сенсорным вводом, генерируемым компьютером, таким как звук, видео, графика или данные GPS [7]. Дополненная реальность характеризуется внедрением виртуальных элементов в физический мир, за счет вспомогательного устройства вывода информации. [8]

За последние несколько лет было реализовано множество прототипов приложений дополненной реальности, охватывающих такие сферы как обучение, маркетинг, сборочные производства и многое другие. Несмотря на огромный объем рынка, большинство прототипов не смогли превратиться в коммерческие продукты.

Можно выделить важную причину такого положения: текущему AR оборудованию не хватает вычислительной мощности для необходимых задач. Например, Google Glass, представленный в 2012 году, страдает от перегрева и короткого времени автономной работы при обработке требовательных алгоритмов AR. Но даже более современные устройства, такие как Microsoft HoloLens имеют ограничения, например, в количестве отображаемых полигонов. Также это относится к смартфонам и планшетам. Во-вторых, поддержка AR приложений в них в актуальном состоянии на разных мобильных устройствах (то есть на другом оборудовании, ОС, платформе и т. д.) очень затратно. Поэтому традиционные (не AR) приложения часто основываются на архитектуре клиент-сервер, где необходимо настроить и поддерживать только одно центральное устройство (сервер).

В отличие от виртуальной реальности (VR), дополненная реальность (AR) не ставит целью заменить реальный мир симуляцией и часто классифицируется как система смешанной реальности (MR), где реальность смешивается с виртуализацией. Поэтому с каждым годом интерес к дополненной реальности увеличивается.

После анализа публикаций [9] наиболее популярными ключевыми словами в области дополненной реальности, не включая самого словосочетания "Augmented reality", были: 3D, tracking, interaction, assembly и mobile (рис. 1).

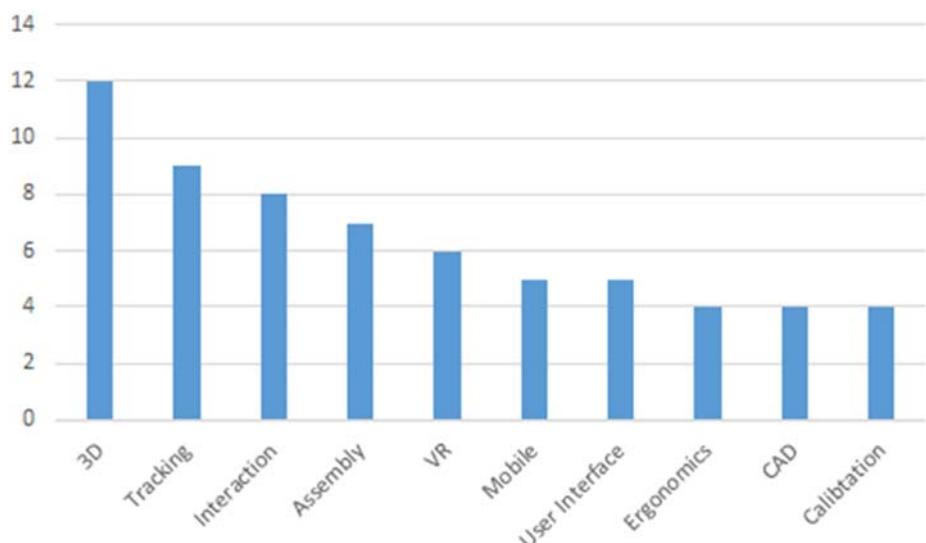


Рис. 1. Частота использования ключевых слов



Assembly встречается часто из-за того, что процесс сборки (*assembly process*) является одной из наиболее дорогих частей производства, однако стоимость может быть значительно сокращена, если изделие собирается в соответствии с хорошо спланированной последовательностью сборки (*assembly sequence*). Результаты исследования научных работ подтверждают, что сборка и обслуживание являются наиболее популярными областями применения AR.

Проанализировав результаты 2015–2020 годов можно сказать что интерес к внедрению дополненной реальности в области сборки и обслуживания со временем вырос, и большинство статей затрагивают именно эти темы.

В настоящее время совокупный годовой темп роста промышленного AR-рынка в период с 2018 по 2025 год прогнозируется на уровне 74 %. Совокупный рынок промышленного AR к 2025 году ожидается на уровне 76 миллиардов долларов. Этот рост, вероятно, будет поддерживаться или ускоряться за счет растущей зрелости технологии дополненной реальности и широкого спектра приложений в отрасли.

Компоненты AR систем

Общая информация

Типичная система дополненной реальности включает в себя такие важные компоненты, как (1) устройство визуализации / захвата, (2) устройство взаимодействия и (3) система слежения.

Технологии захвата – это решения, принятые для захвата сцены, которая затем может быть просмотрена пользователем с наложенной информацией. Следовательно, эти технологии несут ответственность за сбор информации об окружающей среде. Устройства визуализации используются для отображения результатов обработки изображений, обогащенного дополнительной полезной информацией.

Устройства взаимодействия используются для команд со стороны пользователя, влияющих на обработку и отображения дополнительной информации.

Наконец, технологии отслеживания обеспечивают распознавание ключевых компонентов в захваченной сцене и, таким образом, предоставляют исходную информацию для семантического анализа текущей ситуации, включая определения положения пользователя в промышленной среде. Далее эта информация используется для интеграции в изображение визуальных объектов дополненной реальности (например, подсказки пользователю).

Варианты систем захвата и визуализации

На основе изученных публикаций можно сделать вывод, что камера, подключенная к монитору, является наиболее часто используемым решением для захвата сцены внешней среды. В этом случае монитор отображает сцену для пользователей с соответствующей дополнительной информацией.

Еще одно технологическое решение, часто используемое для визуализации – это HMD (*Head-mounted display*). По сравнению с первым решением, HMD легче транспортировать и не требуется установка камер и мониторов в производ-



ственной зоне. Использование HMD-решений со временем несколько увеличилось. Однако в некоторых публикациях использование HMD было подвергнуто критике [10], особенно при их использовании в системах удаленной помощи. Фактически, использование головных дисплеев ограничивает движение головы, чтобы стабильно использовать удаленного помощника.

В некоторых исследованиях также сообщалось, что пользователи могут страдать от снижения остроты зрения, глядя на физическую цель через HMD [11]. Поэтому для систем дополненной реальности также используются мобильные устройства (телефоны, планшеты).

Основные блоки мобильной системы

В [12] выделяется 6 функциональных блоков мобильной системы дополненной реальности:

1. Уменьшение качества кадра.
2. Обнаружение целей в поле зрения камеры мобильного устройства и идентификация области интереса (*regions of interest*) для целей.
3. Извлечение характерных точек из каждой области интереса и распознавание объектов, чтобы определить исходное изображение, хранящееся в базе данных.
4. Сравнение целевого объекта с распознанным исходным изображением.
5. Определение положения и ориентации.
6. Рендеринг информации.

Методы отслеживания объектов в системах дополненной реальности

Общее положение

Технология отслеживания обычно относится к решениям, используемых для того, чтобы система AR могла распознавать ключевые компоненты в захваченной сцене и, таким образом, предоставлять правильную информацию при расширении самой сцены. Технологии отслеживания также важны для определения положения пользователя в промышленной среде.

Системы отслеживания для AR можно разделить на две основные категории, а именно: отслеживание на основе маркеров (QR, Штрих-коды, RFID, визуальные маркеры) и отслеживание без маркеров (гибридный метод слежения, объединяющий слежение по площади и по признакам; использование естественных маркеров).

Роль машинного зрения

Чтобы добиться правильного рендеринга, необходимо отслеживать интересующий объект, который необходимо дополнить [13]. Это можно сделать с помощью методов машинного зрения. Целевой объект, представляющий интерес, может быть обнаружен и отслежен в видеопотоке в реальном времени. Целью может быть простой планарный маркер [14] или любой трехмерный (3D) объект [15]. Алгоритмы обнаружения используют подход, основанный на признаках, когда точки интереса сопоставляются с эталонным видом [16, 17].



Методы глубокого обучения успешно используются в сложных задачах распознавания, таких как распознавание письменных цифр и классификация объектов. В частности, сверточные нейронные сети (CNN) успешно используются в задачах обнаружения [18]. Однако, надежное обнаружение объектов для AR по-прежнему остается сложной задачей. Поскольку приложения AR все больше перемещаются на мобильные устройства, ведется поиск эффективных и надежных алгоритмов обнаружения.

AR-отслеживание на основе модели (или целевого изображения) включает два основных этапа. Сначала, цель обнаруживается во входящем видеопотоке с использованием алгоритма обнаружения, после происходит обработка информации за счет алгоритма отслеживания.

Во время этапа обнаружения, определяется положение камеры по отношению к известной цели. Эта начальная поза, в свою очередь, используется алгоритмом отслеживания на втором этапе. Обычно отслеживание является более надежным и требует меньше времени и ресурсов, чем обнаружение. Поэтому, насколько это возможно, используется алгоритм отслеживания, а алгоритм обнаружения используется только при необходимости.

Процедура обнаружения хорошо отработана, и текущая работа сосредоточена на повышении устойчивости алгоритмов сопоставления. Разрабатываются новые дескрипторы функций, которые работают быстрее (с большим вниманием к мобильным устройствам) [19] и более надежны в отношении освещения и изменений позы.

Типы дополненной реальности

Рассмотрим основные варианты дополненной реальности. В [20] выделяются следующие типы:

Проекция: использует виртуальные образы, чтобы дополнить то, что вы видите вживую. Некоторые мобильные устройства могут отслеживать движения с помощью камеры. Виртуальные клавиатуры, которые можно проецировать практически на любую плоскую поверхность, являются примерами устройств дополненной реальности, которые используют интерактивную проекцию. **Распознавание:** это тип дополненной реальности, который использует распознавание форм, лиц или других предметов реального мира для предоставления дополнительной информации пользователю в режиме реального времени. Карманное устройство, такое как смартфон с соответствующим программным обеспечением, может использовать распознавание для считывания штрих-кодов продуктов и предоставления соответствующей информации, такой как обзоры и цены, или для чтения лиц, а затем предоставлять ссылки на профили людей в социальных сетях.

Местоположение: можно использовать смартфон с GPS для определения своего местоположения, а затем наложить экранные стрелки на живое изображение того, что находится перед пользователем, и указать ему направление, куда ему необходимо двигаться. Эту технологию также можно использовать для определения ближайших остановок общественного транспорта.



Схема: это тип дополненной реальности, которая объединяет очертания человеческого тела или части тела с виртуальными материалами, позволяя пользователю подбирать и иным образом манипулировать объектами, которые не существуют в реальности. Один из примеров этого можно найти в некоторых музеях и научных центрах в форме виртуального волейбола. Хотя игрок может стоять и двигаться по реальной площадке, мяч проецируется на стену позади него, и он может управлять им с помощью своего очертания, которое также проецируется на стену.

Аппаратные проблемы и особенности

Рассмотрим основные поставленные задачи, варианты решений, а также проблемы, связанные с аппаратным обеспечением приложений дополненной реальности в контексте Индустрии 4.0.

Системы дополненной реальности очень сильно зависят от основных аппаратных компонентов (системы слежения, графической системы, дисплея и т. д.). Например, системы слежения могут сильно различаться по точности, частоте обновления и задержке в зависимости от производителя и технологии [21].

Была проделана большая работа по сравнению нескольких систем VR и AR, но текущие результаты трудно обобщить, поскольку они зависят от общих систем, используемых в этих экспериментах. Pausch и др. [22] изучали визуальный поиск между системой виртуальной реальности и настольной системой. Они обнаружили, что пользователям, в системах VR где не отслеживалась голова, требовалось значительно больше времени, чтобы определить, существует ли виртуальная цель в комнате, когда она отсутствует.

В аналогичной задаче с AR Wither и др. [23] оценили различные устройства отображения для задач выбора и аннотации в AR. Авторы сравнили HMD, ручной дисплей с линзами и дисплей планшета, который держится на уровне талии.

Было предположено, что визуальный поиск виртуальных букв должен был быть быстрее с HMD из-за его более высокого уровня погружения, но результаты были неубедительными, хотя пользователи, в целом, согласились с этой гипотезой. Эти разные выводы по похожим задачам подчеркивают сложность обобщения результатов на основе реальных систем AR / VR. Аппаратное обеспечение и уровень погружения систем были слишком разными. Возможно, другие различия в отображении могли повлиять на результаты, и главный реальный вывод состоит в том, что эти результаты зависят от соответствующих сценариев.

Задержка

Одной из самых важных аспектов проблем при работе с проектами в системах дополненной реальности является задержка. Виртуальная информация, связанная с объектами в сцене и отображаемая в окне просмотра, должна регулярно обновляться, чтобы поддерживать правильное выравнивание с объектами в реальном мире [24]. Следовательно, изменения точки обзора необходимо отслеживать и передавать обратно в систему отображения, чтобы повторно отображать виртуальную информацию в правильном месте и своевременно. Padmos и Mild-



ers [25] указывают, что для иммерсивной реальности время обновления (задержка) должно быть ниже 40 мс. Для дополненной реальности ограничения будут намного строже. Они предполагают, что смещение объектов между двумя кадрами не должно превышать 15 угловых минут ($0,25^\circ$), что потребует максимальной задержки в 5 мс, даже когда наблюдатель вращает головой с умеренной скоростью $50^\circ/\text{с}$. К аналогичным значениям максимальной задержки пришли и другие авторы [26, 27]. Во время типичных движений головы могут возникать скорости до $370^\circ/\text{с}$, но маловероятно, что наблюдатели, вращающие голову так быстро, заметят небольшие смещения объектов. Многие авторы предполагают, что 10 мс будет приемлемым для AR [26, 28].

Маркерные системы

Общие принципы

Рассмотрим работу AR систем с отслеживанием на основе маркеров.

Различные типы маркеров дополненной реальности – это изображения, которые могут быть обнаружены камерой и использованы с программным обеспечением в качестве местоположения виртуальных объектов, размещенных в сцене. Большинство из них черно-белые, хотя можно использовать прочие цвета, если камера может правильно распознать контраст между ними. Простые маркеры дополненной реальности могут состоять из одной или нескольких основных фигур, состоящих из черных квадратов на белом фоне. Более сложные маркеры могут быть созданы с использованием простых изображений, которые все еще правильночитываются камерой [20]. Также на рис. 2 представлена архитектура маркерной дополненной системы, представленная в данном источнике.

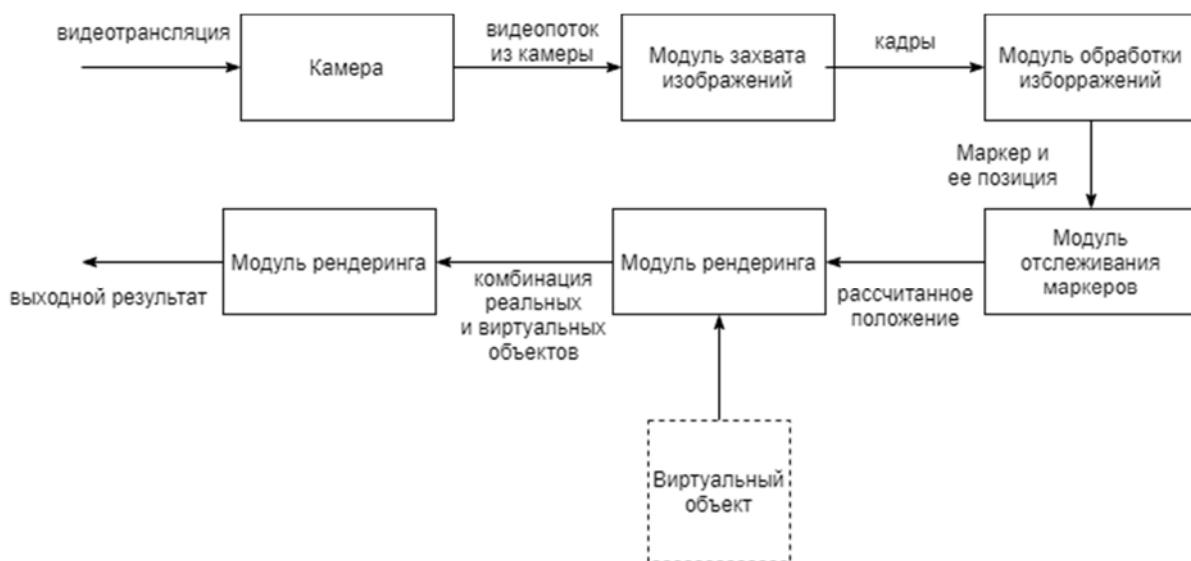


Рис. 2. Архитектура маркерных систем

Ограничения на типы маркеров дополненной реальности, которые можно использовать, зависят от программного обеспечения, которое их распознает.

В [29] расширяют эту идею и уточняют, что для ускорения реального проецирования контента и использования вычислительной мощности устройства,



приложение в значительной степени полагается на отдельные потоки. Кроме того, чтобы убедиться, что модуль, отображающий предварительный просмотр камеры, не замедляется из-за другой обработки, была использована модель программирования – конвейер. Каждый модуль считывает данные из входного буфера и после обработки записывает данные в другой буфер, чтобы другой модуль мог их использовать.

Типы маркерных систем

Есть много практических систем, которые используют двумерные шаблоны для передачи информации. Область их применения варьируется от промышленных систем, где маркеры предназначены для маркировки деталей или передачи определенной информации, до систем, в которых маркеры используются для локализации.

Примерами первого случая являются Maxicode, используемой почтовой службой США, DataMatrix и QR (*Quick Response*), в промышленных условиях для маркировки деталей. Примерами для второго случая являются ARToolKit, ARTag и ARStudio, три системы для используемые в дополненной реальности.

Чтобы снизить чувствительность к настройкам камеры, в планарных системах маркеров обычно используются битональные маркеры. Таким образом, нет необходимости определять оттенки серого, и решение, принимаемое для каждого пикселя, сводится к пороговому решению. Кроме того, многие системы маркеров используют некоторые биты данных маркера для передачи избыточной информации, что позволяет обнаруживать и исправлять ошибки. Дизайн маркеров во многом зависит от области применения.

Требования к маркерным системам

Для приложений дополненной реальности очень важно, чтобы маркеры можно было найти в большом поле обзора. Это означает, что они должны быть распознаны даже если кажутся искаженными на изображении. Кроме того, информация, хранящаяся внутри маркера, не должна быть слишком плотной, чтобы увеличить расстояние, на котором данные могут быть восстановлены с маркера. К счастью, этого легко добиться, поскольку в приложениях дополненной реальности маркер несет меньше информации, обычно это только идентификатор, чтобы отличить его от других маркеров.

Большинство систем дополненной реальности также работают, если виден только один маркер. Следовательно, маркер должен иметь несколько отдельных точек, по крайней мере, четыре, чтобы можно было вычислить позу маркера на камере. Обычно такие маркеры имеют четырехугольный контур, а четыре угловые точки используются для расчета трехмерной позы.

Пример реализации маркерной системы

В примере, описанном в [30] затрагивается реализация AR-интерфейса, который позволяет нескольким участникам динамически размещать и упорядочивать виртуальные объекты на рабочем пространстве смешанной реальности.



В этой системе пользователь носит головной дисплей с небольшой прикрепленной камерой. Данный проект, под названием Tiles был разработан для быстрого прототипирования и оценки приборных панелей самолетов. Это совместная исследовательская инициатива, осуществляемая при поддержке DASA / EADS Airbus и DaimlerChrysler AG.

Трехмерное положение и ориентация виртуальных объектов определяется с помощью методов отслеживания компьютерного зрения, а также с помощью квадратных маркеров, которые могут быть прикреплены к любому физическому объекту. Виртуальные объекты визуализируются относительно этих маркеров, и, манипулируя помеченными физическими объектами, пользователь может манипулировать их виртуальными значениями, без необходимости использовать любые дополнительные устройства ввода. Рабочее пространство и интерфейс Tiles состоят из карточек (15 на 15 сантиметров каждая) с прикрепленными к ним шаблонами отслеживания. У каждой из этих карточек есть магнит на обратной стороне, поэтому ее можно прикрепить на доску и снять с нее.

Белая доска действует как общее рабочее пространство для совместной работы, где пользователи могут быстро нарисовать приблизительную компоновку виртуальных инструментов, а затем визуализировать ее, разместив плитки с виртуальными инструментами на доске. Вместо того, чтобы взаимодействовать с цифровыми данными, манипулируя значками с помощью мыши, пользователь взаимодействует с цифровыми данными, физически манипулируя соответствующими элементами. Все плитки можно перемещать в пространстве и размещать на доске: пользователь просто берет любую из них, исследует ее содержимое и помещает на доску. Операции между плитками вызываются путем размещения двух из них рядом друг с другом (на расстоянии менее 15 % от размера плитки). Например, чтобы скопировать инструмент в плитку данных, пользователь сначала находит нужный виртуальный инструмент, а затем помещает любую пустую плитку данных рядом с инструментом.

Систему протестировали около семидесяти пользователей. Самая распространенная жалоба была на внешний вид плитки, а также пользователи имели тенденцию перекрывать границу отслеживания, что приводило к сбою.

Маркерная система как способ рассчитать расстояние

Целью исследования в [31] являлось найти алгоритм измерения размеров и формы помещения с помощью маркерных технологий AR. Этот процесс включает в себя проектирование маркеров, их отслеживание и использование полученной от них информации для расчета и создания формы и размера комнаты. В данном случае был выбран тип маркеров ArUco, поскольку он соответствует требованиям с точки зрения надежности, прочности и универсальности. Кроме того, их распознавание происходит быстро, с низким уровнем неоднозначности положения в помещениях. Обычно маркер ArUco представляет собой черно-белый квадратный узор из блока 7×7 . Черная рамка имеет толщину 1 блок, поэтому внутренний узор представляет собой подматрицу размером 6×6 блоков.

Выделяется 3 этапа отслеживания маркера: обнаружение кандидатов в маркеры: на этом этапе информация с камеры будет проанализирована, чтобы про-



верить с использованием техники адаптивного порога, есть ли квадратный маркер или нет, а затем отфильтровать другую информацию. Анализ маркеров: на этом этапе использовалась пороговая техника Оцу (*Otsu threshold technique*) для разделения черного и белого битов. После расчета маркер может определить, есть ли его идентификатор в базе данных или нет. Техника оценки позы: техника, используемая для оценки положения маркера путем объединения координат создателя в реальном и виртуальном мире. Процесс даст положение и поворот маркера.

Поскольку необходимо найти расстояние между двумя позициями: то есть между двумя стенами комнаты или двумя сторонами одной стены, то маркеры этих двух позиций должны быть собраны или зафиксированы на одном и том же кадре.

При каждом захвате, данные, полученные от маркера, такие как идентификатор, положение и угол поворота, будут храниться в базе данных для дальнейшего анализа. В случае небольшой комнаты, где два маркера с противоположной стены могут быть сняты одновременно, данные с обоих могут храниться в базе данных напрямую без каких-либо дополнительных изменений. Если комната слишком большая, необходимо поставить между ними еще один маркер. В этом случае придется снимать два и более раза. После завершения сбора всех данных с маркеров система поставит их в базу данных приложения в виде набора данных, например:

Комната = {{Левый, Правый}, {Левый, Задний}, {Левый, Передний}, {Левый, Пол}, {Пол, Потолок}}. Каждое подмножество означает один набор данных из одного захвата. Набор данных из общего маркера будет использоваться в качестве связующего звена с другим набором данных.

Чтобы соединить два набора снимков, нужно сопоставить общий маркер в том же положении и с одинаковым углом поворота. Поскольку это один и тот же маркер, то набор данных изменится на: Комната = {Левый, Правый, Задний, Передний, Пол, Потолок}.

Подводя итог эксперимента можно сказать, что расчет результатов по расстояниям маркеров обеспечил высокую точность: размеры всех трех помещений, полученные из системы, отличаются от ручного измерения менее чем на 3 %. Этот результат подтвердил, что измерение AR на основе маркеров можно использовать для эффективного управления пространством.

Пример взаимодействия AR системы и робота

Описание

Система, рассмотренная в [32] состоит из гуманоидной роботизированной платформы с двумя руками с шестью степенями свободы (DOF), стола с набором блоков разных цветов, компьютерной станции для записи видео и потоковой передачи, а также планшетного устройства, которое удерживается пользователем и обеспечивает интерфейс смешанной реальности для взаимодействия с роботом.

Робот запрограммирован на получение по Wi-Fi команд, которые содержат желаемые позы, чтобы инструмент робота мог подбирать и размещать объекты



в своем рабочем пространстве. Робот окружен симметричным рабочим пространством, состоящим из двух прямоугольных секций, по одной с каждой стороны и одной полукруглой секции перед роботом. К поверхностям рабочего пространства прикреплены двенадцать визуальных маркеров, которые расположены известным образом в заранее определенных местах.

Работа с интерфейсом

Чтобы исследовать смешанную реальность как инструмент HRI (*Human-robot interaction*) и изучить способы, которыми мобильные технологии могут улучшить HRI в общих пространствах, разработаны три мобильных интерфейса, в которых используются различные стратегии проектирования.

Мобильное приложение должно выполнять несколько задач: каждый видеокадр обрабатывается с использованием метода адаптивной пороговой обработки для определения местоположения изображения центров маркеров, прикрепленных к рабочему пространству и интересующим объектам. Затем относительные позы преобразуются так, чтобы позы всех интересующих объектов представлялись относительно фиксированной системы координат, установленной маркерами рабочего пространства.

В первом интерфейсе пользователи взаимодействуют с роботом, получая визуальную обратную связь с точки зрения одной из камер робота. Ограниченнное поле зрения этой камеры способно захватить только один из трех интересующих участков рабочего пространства в любой момент времени. Таким образом, пользователи вынуждены панорамировать и наклонять камеру между этими тремя секциями, если объекты должны перемещаться из одной секции рабочего пространства в другую. Интерфейс использует акселерометр и гироскоп устройства для оценки положения устройства, чтобы предоставить пользователям наиболее естественное управление панорамированием и наклоном системы камеры робота, сохраняя при этом жесты сенсорного экрана для взаимодействия с виртуальными элементами в среде смешанной реальности,

Для того, чтобы избавиться от ограничений, с которыми сталкивается первый интерфейс при наблюдении за рабочим пространством, второй интерфейс предоставляет пользователям расширенный и более естественный вид с камеры, установленной на потолке над рабочим пространством.

В третьем интерфейсе пользователи держат планшет так, чтобы его задняя камера была направлена на робота и его рабочее пространство с произвольной точки зрения. Камера планшета захватывает видео, используемое для визуализации среды смешанной реальности на экране при взаимодействии с роботом.

Тестирование

Чтобы оценить пользовательский опыт и производительность при взаимодействии с роботом-гуманоидом с использованием разработанных интерфейсов, было проведено экспериментальное исследование с участниками, каждого из которых попросили дать команду роботу выполнить задачу манипулирования объ-



ектами с помощью одного из интерфейсов. Задача заключалась в том, чтобы подобрать каждый из трех блоков в рабочей области и поместить его в контейнер соответствующего цвета за как можно меньшее время.

Сначала каждый участник выполнил предварительную оценку системы, которая использовалась для проверки его знакомства с мобильными устройствами, дополненной реальностью и роботами. Затем участникам была объяснена задача, которую необходимо выполнить, и интерфейс, который будет использоваться для ее выполнения.

Когда участники выполняли задание, они стояли лицом к роботу за полу-круглой частью рабочего пространства (рис. 3). В третьем интерфейсе предлагалось перемещаться по мере необходимости, чтобы увидеть блоки, которые могут быть спрятаны за роботом. Чтобы оценить выполнение поставленной задачи с использованием каждого из интерфейсов, фиксировалось время выполнения каждого испытания.

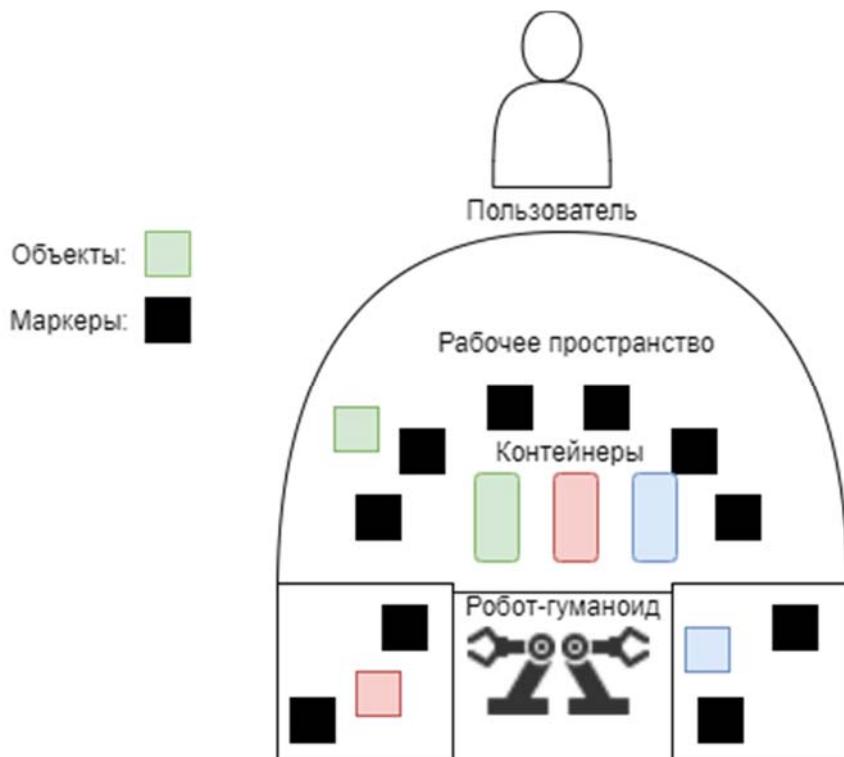


Рис. 3. Работа с роботом

После выполнения задания, участникам было предложено оставить оценку взаимодействия, состоящую из двух частей. В первой части оценивания, участников попросили указать уровень их согласия с девятью положительными и отрицательными утверждениями по 5-балльной шкале. Во второй части, для оценки рабочей нагрузки, связанной с использованием каждого интерфейса, использовался индекс нагрузки NASA (NASA-TLX) [33]. Этот индекс, который рассчитывается по шкале от 0 до 100, характеризует сложность выполнения задачи для участника и состоит из шести категорий: умственная нагрузка, физическая нагрузка, времененная нагрузка, производительность, усилия и разочарование. Результаты опроса приведены на рис. 4. На графике видно, что наиболее удоб-



ным интерфейсом по мнению пользователей является третий (мобильный). Критика мобильного интерфейса не была связана с перспективой, предоставляемой камерой, или какой-либо задержкой видео, а была сосредоточена на вопросах удобства использования, связанных с визуальными и интерактивными аспектами интерфейса. Предложения включали альтернативные способы отображения расширенной сетки и блоков в окружающей среде (например, изменение внешнего вида и снижение прозрачности визуальных элементов, предоставление контейнеров дополненной реальности и т. д.).



Рис. 4. Процент типа комментариев, оставленных для каждого интерфейса

Недостатки маркерных систем

Однако, в традиционной маркерной системе есть определенные проблемы. В [34] говорится о том, что поскольку зарегистрированная информация является независимой для разных систем AR, маркеры, используемые в одной системе, не могут применяться в другой, если не происходит дополнительная процедура регистрации. Это говорит о том, что такая система применима только для частного использования, а не в приложениях общего пользования. Поэтому использование QR-кода в качестве маркера, будет более удобна.

Также можно выделить проблему распознавания маркеров: например, если маркер отражает свет или переход между цветами маркера слишком плавный, то это может привести к некорректной работе.

Безмаркерные системы

Пример реализации безмаркерной системы

Перейдем к безмаркерным системам дополненной реальности.

В [35] технология AR используется для отображения информации о задачах или для выделения определенных частей промышленного манипулятора. Например: инструкции по выполнению задач, отображаются на интерфейсе рабочего стола, для поддержки людей-операторов во время совместной сборки человека и робота. Также, проекционный интерфейс AR используется для отображения интерактивных инструкций по сборке. Список доступных функций проецируется на сенсорный стол. Пользователи могут добавлять данные для выполнения требований задачи. Кроме того, проектируемые интерфейсы AR позволяют визуализировать объекты, которыми будут управлять промышленные роботы.

В ходе работы определили 4 типа проблем: проблемы датчиков, проблемы в системе управления, проблемы механической конструкции, проблемы с перегрузкой.

Также выделяется общая схема рабочего процесса:

- 1) представление неисправности (FR) (в виде 2D значков);
- 2) Размещение значка неисправности (FIP);
- 3) экспериментальные испытания (ET);
- 4) Анализ результатов (RA).

Система, описанная в [35] состоит из трех основных элементов: промышленного манипулятора, AR клиента и удаленного сервера. Клиентом AR является шлем Microsoft HoloLens. Это система позволяет пользователям держать руки свободными от AR-клиента при диагностике неисправностей производственного оборудования. Так как алгоритм сегментации был разработан с использованием библиотек обработки изображений, для которых может потребоваться высокая вычислительная мощность, была использована архитектура клиент-сервер, чтобы ограничить потребление заряда батареи и снизить вычислительные затраты HoloLens. Сервер получает от робота: место ошибки, тип ошибки, конфигурацию суставов: значения суставов робота. Общая архитектура проекта представлена на рис. 5.

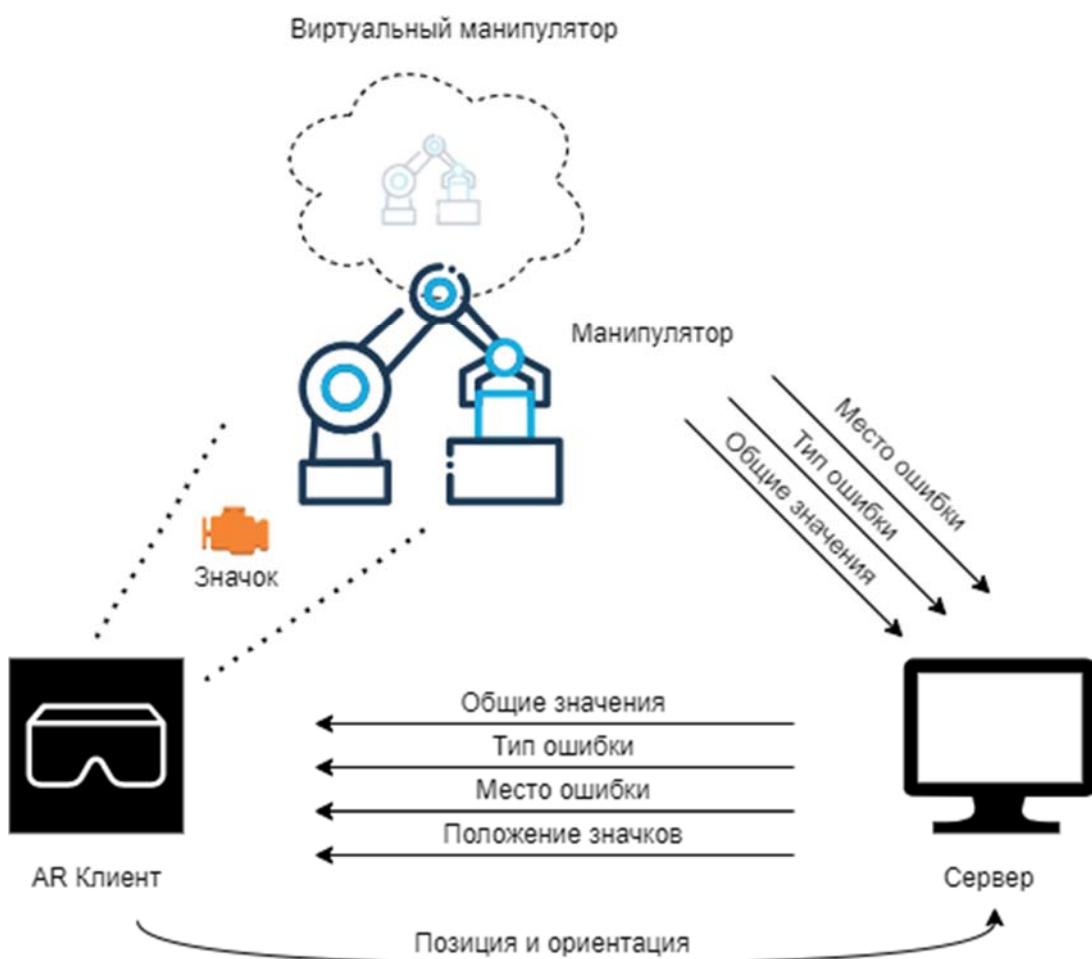


Рис. 5. Общая архитектура проекта



Сам сервер работает в двух режимах: адаптивный и неадаптивный. При работе в адаптивном режиме сервер получает от клиента: позицию пользователя: 3 плавающих значения; ориентацию пользователя: 3 плавающих значения. Эти значения отправляются каждые 20 мс и используются для анимации виртуальной камеры, которая представляет фактическое положение и ориентацию пользователя.

В конце были сделаны выводы, связанные со временем, ошибками и количеством действий для выявления неисправностей. Поскольку виртуальные значки были расположены близко к стыку, затронутому повреждением, пользователи могли одновременно распознавать как типологию неисправности, так и номер соединения, без необходимости часто менять их местоположение и точку обзора. Кроме того, механизм автоматического масштабирования и ориентирования позволяет удерживать значок в узком поле зрения устройства HoloLens, сохраняя его четко узнаваемым. Правильное распознавание суставов манипулятора предполагает, что принятый способ отслеживания был признан подходящим.

Реализация безмаркерной системы для мобильных устройств

В [11] рассматривается работа с дополненной реальностью через мобильные устройства. Чтобы избежать выполнения ресурсоемких задач (например, извлечение и распознавание объектов) на мобильных устройствах, системы дополненной реальности могут выгружать их в облако. Существует два распространенных сценария разгрузки в зависимости от того, где выполняется извлечение функций.

1) Системы, такие как Overlay [36], разгружают задачи, начиная с извлечения признаков, в облако, загружая кадры камеры;

2) Мобильные устройства также могут выполнять извлечение признаков локально и отправлять сжатые точки в облако для распознавания объектов (например, *VisualPrint*).

Проблемы

Одной из фундаментальных проблем реализации мобильной системы дополненной реальности является гибкое и точное отслеживание объектов, которое особенно необходимо при непрерывном распознавании, когда пользователи могут свободно перемещать свои мобильные устройства во время распознавания. С появлением точных датчиков и высококачественных камер на мобильных устройствах AR переходит на безмаркерную систему и позволяет пользователям размещать аннотации в физическом мире без распознавания объектов. Безмаркерная AR (также известная как *Dead Reckoning*) обычно решает проблему SLAM (одновременной локализации и сопоставления) во время выполнения.

Проект представленный в данном исследовании состоит из двух частей: клиент и сервер. Клиент Jaguar использует метод отслеживания ARCore и ARKit и извлекает выгоду из возможности крупномасштабного распознавания объектов сервера Jaguar. Цель сервера Jaguar – не только разгрузить ресурсоемкие вычислительные задачи в облаке, но и использовать аппаратное ускорение с помощью



графических процессоров, чтобы преодолеть предел сквозной задержки для мобильной AR. Еще одна причина, по которой используется облачная архитектура, заключается в том, что она необходима для крупномасштабного поиска изображений в наборе данных с тысячами изображений, что невозможно на мобильных устройствах из-за их ограниченного хранилища. Данный проект развернут под системы Android, однако нет проблем с переносом для ARKit на систему IOS.

Традиционные решения AR, основанные на поиске изображений, хорошо справляются с распознаванием объектов, но страдают от низкого качества или даже полной потери отслеживания из-за затенения, размытия движения или изменения изображения. AR SDK, такие как ARCore и ARKit, решают эту проблему, используя как инерционные измерения от IMU, так и характерные точки из захваченных изображений камеры. Однако ключевой функцией, необходимой для систем AR, является понимание контекста, которое обеспечивается распознаванием объектов. Эта важная функция отсутствует в начальных выпусках ARCore и ARKit. Jaguar является надстройкой для ARCore и улучшает функцию отслеживания движения. ARCore берет на себя управление оборудованием камеры и использует OpenGL для визуализации вида камеры. Jaguar извлекает текущий кадр камеры из обзора ARCore, затем отправляет запрос распознавания, содержащий этот кадр, на пограничные облачные серверы через Network Manager.

Рендеринг изображений

Для рендеринга содержимого Jaguar преобразует 2D-границу объекта в 3D-позу с шестью степенями свободы (6DoF). В OpenGL преобразование трехмерной координаты в двумерную экранную координату осуществляется посредством матричных преобразований:

$$\nu 2D = P \cdot V + M \cdot \nu 3D,$$

где M – матрица модели, описывающая размещение объекта в 3D, V – матрица вида, описывающая размещение камеры в 3D, а P – матрица проекции.

В Jaguar P и V предоставляются ARCore для каждого кадра, $\nu 2D$ (местоположения вершин объекта в 2D-обзоре камеры) и $\nu 3D$ (полученные из размера объекта в пикселях) возвращаются с сервера. Однако, нет возможности вывести M из-за отношения «много к одному» относительно левой и правой части уравнения. То есть для каждого $\nu 2D$ существует несколько возможных M , соответствующих ему, потому что информация о физическом размере объекта все еще отсутствует.

Чтобы определить физический размер объекта во время выполнения, Jaguar использует как информацию о сопоставлении от клиента, так и результат распознавания от сервера. Первый моделирует среду в физическом масштабе, а второй определяет границу и местоположение объекта в среде. ARCore создает разреженное трехмерное облако точек среды как часть своей функциональности SLAM, которая описывает расположение каждой точки в масштабе физического мира. В OpenGL плоскость изображения камеры помещается между камерой и объектами, а трехмерные точки внутри усеченной пирамиды перед камерой проецируются на плоскость изображения. Когда происходит отбрасывание луча из камеры



в направлении 2D-вершины объекта в поле зрения камеры, первая плоскость в облаке точек, на которую попадает луч, является плоскостью объекта, а точка пересечения - это в точности трехмерное местоположение соответствующая вершине физического объекта. На основе этих вершин Jaguar определяет физический размер и матрицу модели M объекта, так что может точно выровнять содержимое с объектом. В настоящее время Jaguar рассчитывает физический размер только плоских объектов.

Уменьшение задержки за счет разгрузки

В идеале задержка должна быть ниже, чем интервалы кадров камеры, которые обычно составляют около 33 мс (т. е. 30 кадров в секунду), чтобы обеспечить наилучшее качество взаимодействия с пользователем. Предыдущая работа [37] продемонстрировала, что разгрузка облака может значительно уменьшить задержку обработки AR с нескольких секунд до нескольких сотен миллисекунд [36, 38]. Еще одно преимущество разгрузки в пограничное облако (*edge cloud*) – это сокращение задержки передачи данных по сети, которая может составлять всего несколько миллисекунд. Естественно, можно разделить обработку на стороне сервера на две части: автономная подготовка и онлайн-распознавание.

Автономная подготовка. Автономная обработка Jaguar включает 4 этапа: после извлечения характерных точек из эталонных изображений в базе данных подлежащих распознаванию объектов, он сначала уменьшает их размер и обучает статистическую модель, чтобы сжать их в компактное представление. Затем Jaguar создает хэш-таблицы сжатых точек функций для более быстрого распознавания объектов в Интернете. Для каждого изображения Jaguar выполняет обнаружение и извлечение признаков, чтобы получить дескрипторы признаков. Существует огромное количество алгоритмов извлечения признаков, таких как SIFT, SURF, ORB, FAST + FREAK и BRISK. Среди них SIFT превосходит другие с точки зрения точности, особенно в случаях масштабирования, поворота и размытия движения [39, 40]. Уменьшение размеров: каждая характерная точка, извлеченная с помощью SIFT, имеет 128 измерений, и некоторые из них могут быть коррелированы. Jaguar использует анализ главных компонентов (PCA), чтобы уменьшить размер дескрипторов функций до 82, что также увеличивает точность распознавания объектов. Кодирование функций: поскольку количество характерных точек колеблется от 500 до 1000 для разных изображений, цель кодирования признаков - создать более компактное представление с одинаковым размером для каждого изображения. Jaguar создает статистическую модель (например, *Gaussian Mixture Model*, GMM [41]), используя характерные точки, преобразованные с помощью PCA, а затем использует вектор Фишера (FV) [42] для дальнейшего сжатия их в векторы фиксированной длины. LSH Хеширование: для дальнейшего ускорения онлайн-поиска Jaguar использует локальное чувствительное хеширование (LSH) и вставляет все FV вектора эталонных изображений в хеш-таблицы LSH. Без хеш-таблиц нам нужно поочередно вычислять расстояние между FV целевого объекта и каждого контрольного изображения, что не масштабируется для большого набора данных.



Распределенные вычисления для AR

Туманные вычисления

Для разработки приложений, связанных с описанными ранее вариантами использования, системы AR могут использовать технологию туманных вычислений. Туманные вычисления – это парадигма распределенных вычислений, которая действует как промежуточный уровень между облачными центрами обработки данных и устройствами. Предлагаются вычислительные и сетевые мощности, а также хранилища, чтобы облачные сервисы могли быть расширены ближе к устройствам и датчикам интернета вещей. [43] Туманные вычисления расширяют возможности облачных вычислений, перемещая часть вычислительных и коммуникационных возможностей облака ближе к узлам датчиков [44]. Такой шаг дает системе AR следующие преимущества:

- Локальные туманные узлы позволяют минимизировать задержку ответа. Это требуется в динамических системах AR реального времени, которые обычно страдают от высоких задержек при доступе к данным, хранящимся в облаке или на удаленных серверах.
- «Туман» может распределять вычислительную мощность и хранилище между локальными шлюзами. Доступ к таким ресурсам осуществляется через службы, которые особенно полезны, когда операторы носят устройства AR, которые обычно ограничены в ресурсах с целью сохранения легкости и продления срока службы батареи. Следовательно, такие устройства могут делегировать «туману» определенные ресурсоемкие задачи или хранение больших файлов.

Архитектура

Архитектура туманных вычислений AR Navantia состоит из трех уровней: уровень узла: он включает все устройства AR, которые взаимодействуют со службами, предоставляемыми слоем тумана. Слой «тумана» также обменивается данными с сетями датчиков, которые соответствуют экосистеме IIoT верфи, и считывателями RFID, которые используют систему позиционирования труб [45]. Слой тумана: он состоит из одного или нескольких одноплатных компьютеров (SBC), которые установлены в фиксированных положениях в цехах верфи и на корабле. Каждый SBC действует как шлюз и предоставляет услуги «тумана». В случае службы AR она обеспечивает устройства локализованными данными и отвечает быстрее, чем облако, таким образом выступая в качестве прокси-кэширующего сервера для данных. Облако: здесь хранятся данные, полученные из нескольких источников верфи. Это также место, где запускаются свои собственные ресурсоемкие сервисы и сервисы, предлагаемые через стороннее программное обеспечение. Связь между каждым устройством и слоем «тумана» осуществляется через соединения Wi-Fi. Данная архитектура представлена на рис. 6.

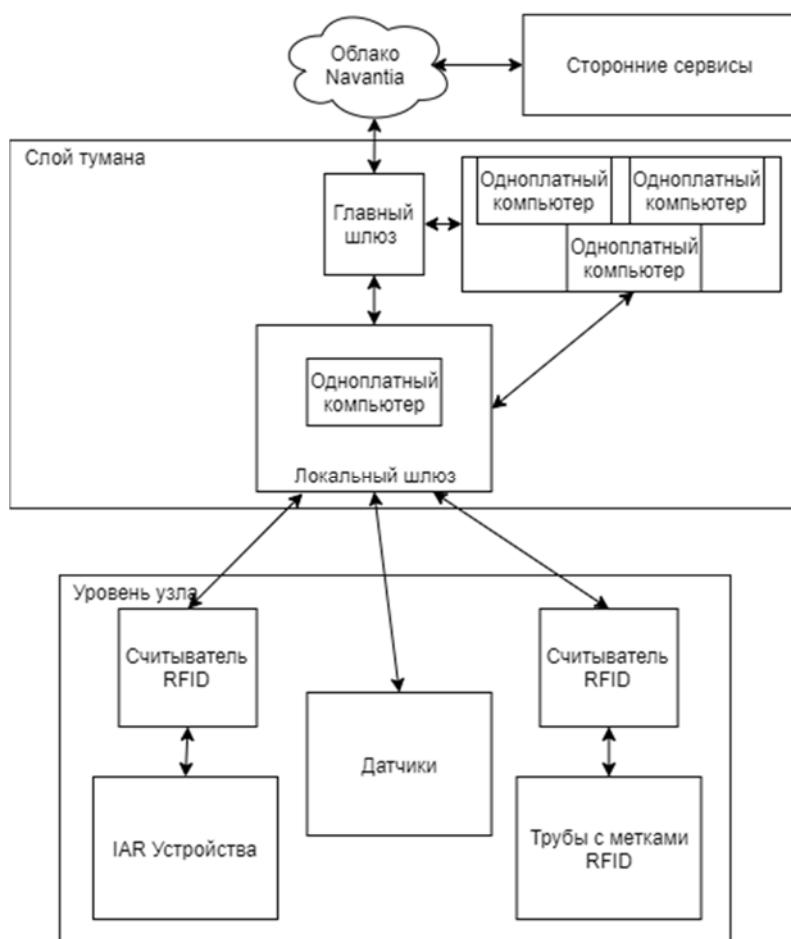


Рис. 6. Архитектура туманных вычислений AR Navantia

Интерфейс мозг-компьютер

Предпосылки создания

Системами дополненной реальности можно управлять не только за счет манипуляций с физическими или виртуальными объектами на прямую. В [46] говорится о том, что интерфейс мозг-компьютер (BCI) может стать мощным инструментом, для улучшения связи между человеком и машиной. Исторически BCI всегда был связан с медицинскими приложениями, гарантируя альтернативные способы общения людям с серьезными двигательными нарушениями, но при этом считается, что комбинация интерфейсов мозг-компьютер с AR получает широкое признание в сегодняшнем мире. Хоть такая комбинация обещает большое удобство при взаимодействии с машинами, технология все еще недостаточно развита для практического применения вне лаборатории

Среди широкого разнообразия технологий BCI в [47] работе основное внимание уделяется одноканальным не инвазивным интерфейсам мозг-компьютер на основе SSVEP за счет ЭЭГ (Электроэнцефалографии), которые являются недорогими и простыми в использовании, требуют минимального обучения и имеют высокое отношение сигнал / шум. Эта технология сочетается с очками дополненной реальности на базе Android, которые легко настраиваются и переносятся.



Проблемы и реализация

Проблема использования дополненной реальности с BCI на основе SSVEP заключается в качестве стимульного сигнала, поскольку постоянство окружающей среды может ослабить интенсивность эффекта мерцания и, следовательно, вызвать сигнал SSVEP.

В статье рассматриваемая система разделена на две основные части: генератор стимулов и блок сбора и обработки. В первом подходе эти два блока независимы друг от друга, но конечная цель – интегрировать всю систему на единой (переносной) платформе.

Система BCI, основанная на SSVEP, требует правильных стимулов, чтобы вызвать электрический потенциал в мозге. Эти стимулы обычно представляют собой мерцающие огни, которые могут быть светодиодами или значками на экране. Для сбора и обработки данных использовалось устройство ЭЭГ.

Эксперимент

Для экспериментальной фазы, испытуемого 25-летнего мужчину попросили надеть ранее описанную систему. В первой экспериментальной кампании было необходимо изучить эффекты стимулов, пытаясь минимизировать другие эффекты. Движения или другие действия могут генерировать нежелательные гармоники в спектре сигналов мозга, тем самым внося шум в измерения.

Сначала проводились эксперименты, вызывающие зрительные потенциалы с помощью светодиодов. Для запуска использовался один светодиод, а затем использовались оба светодиода одновременно, хотя пользователь BCI должен был сосредоточить свое внимание только на одном из них за одно испытание. Целью измерения потенциалов, вызванных этими стимулами, было получение эталонных результатов для следующей фазы.

Во второй части эксперимента в качестве генератора стимулов использовались AR-очки. Сначала на экране мигал единственный значок, и мозг регистрировал сигналы. Затем одновременно отображались оба мерцающих значка, но пользователь BCI сосредоточил внимание только на одном значке за испытание.

Экспериментальные результаты показывают, что можно использовать очки AR для генерации стимулов в BCI на основе SSVEP. Следует отметить, что, если пользователь BCI перемещается во время использования системы, шум увеличивается. Это важный вопрос для системы AR-BCI, так как такая система будет использоваться в промышленном контексте, позволяя пользователю участвовать в нескольких задачах одновременно. Однако амплитуды сигналов, предполагают, что частотные компоненты все же можно идентифицировать.

Заключение

В заключение можно выделить следующие основные выводы:

- Наиболее подходящей архитектурой для систем дополненной реальности на данном этапе их развития является «Клиент – Сервер».
- Одной из важных проблем является задержка. Ее можно уменьшить за счет «конвейера» или edge cloud.



- Определяющее значение в эффективности работы системам дополненной реальности играет машинное зрение.

- Выделяют два основных вида AR систем: маркерные и безмаркерные.

Например, для получения дополнительной информации об объекте или человеке используется маркерная система, а при задачах, где необходимы положение и ориентация – безмаркерная. Также маркерная система больше подходит для задач на массовую аудиторию с разнообразными устройствами в качестве AR-клиента, за счет меньшей нагрузки и большей стабильности. Нередко наблюдается комбинированное использование данных технологий.

- На текущий момент мобильные системы являются одними из наиболее удобных технологий при работе с AR.

- С каждым годом спрос на системы дополненной реальности увеличивается.

Литература

1. Маколкина М. А., Шарлаев М. В. Применение порогового преобразования для сегментации изображений // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2020). IX Международная научно-техническая и научно-методическая конференция : сб. науч. ст. в 4-х т. СПб. : СПбГУТ, 2020. Т. 1. С. 697–702.
2. Pirmagomedov, R., A. Ometov, D. Moltchanov, Xi Lu, R. Kovalchukov, E. Olshannikova, S. Andreev, Y. Koucheryavy and M. Dohler, 2020. Applying Blockchain Technology for User Incentivization in mmWave-Based Mesh Networks. IEEE Access (Volume: 8), pp: 50983–50994.
3. Pirmagomedov, R., D. Moltchanov, A. Ometov, M. Khan, S. Andreev and Y. Koucheryavy, 2019. Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence. IEEE Access (Volume: 7), pp: 180700–180712.
4. Маколкина М. А. Исследование модели движения пользователя и моделей трафика для приложений дополненной реальности // Электросвязь. 2019. № 8. С. 13–22.
5. Маколкина М. А., Кучеряный А. Е. Классификация приложений дополненной реальности // Информационные технологии и телекоммуникации. 2020. Том 8. № 1. С. 11–21. DOI 10.31854/2307-1303-2020-8-1-11-21.
6. Bruno, F., L. Barbieri and M. Muzzupappa, 2020. A Mixed Reality system for the ergonomic assessment of industrial workstations. International Journal on Interactive Design and Manufacturing (IJIDeM), 14: 805–812.
7. Milgram, P. and F. Kishino, 1994. A Taxonomy of Mixed Reality Visual Displays. IEICE Trans. Information Systems, 12: 1321–1329.
8. Swaminathan, A., Y. Mao and M. Wu, 2006. "Robust and secure image hashing. IEEE Trans. Inf. Forensics Security, 2: 215–230.
9. Bottani, E. and G. Vignali, 2019. Augmented reality technology in the manufacturing industry: A review of the last decade. IISE Transactions, 51: 284–310.
10. Gurevich, P., J. Lanir and B. Cohen, 2015. Design and Implementation of TeleAdvisor: a Projection-Based Augmented Reality System for Remote Collaboration. Computer Supported Cooperative Work (CSCW), 24: 527–562.
11. Livingston M. A., C. Zanbaka, J. E., Smallman H. S. (2005). Objective measures for the effectiveness of augmented reality. Proceedings of the Virtual Reality conference (VR 2005), Bonn (Germany), 12-16 March 2005, pp.287–288.
12. Zhang, W., B. Han and P. Hui, 2018. Jaguar: Low Latency Mobile Augmented Reality with Flexible Tracking. Proceedings of the 26th ACM international conference on Multimedia, MM, pp: 355–363.
13. Akgul, O., H.I. Penekli and Y. Genc, 2016. Applying Deep Learning in Augmented Reality Tracking. International Conference on Signal-Image Technology & Internet-Based Systems (issue 12), IEEE, pp: 47-54.



14. Genc, Y., S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, 2002. Marker-less tracking for AR: a learning-based approach. Proceedings. International Symposium on Mixed and Augmented Reality, IEEE.
15. Klein, G., 2006. Visual Tracking for Augmented Reality, PhD thesis, University of Cambridge, Cambridge.
16. Harris, C. and M. Stephens, 1988. A combined corner and edge detector, The Plessey Company plc.
17. Rosten, E. and T. Drummond, 2006. Machine Learning for High-Speed Corner Detection. European Conference on Computer Vision, pp: 430–443.
18. Farabet, C., C. Couprie, L. Najman and Y. LeCun, 2013. Learning hierarchical features for scene labeling. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35: 1915–1929.
19. Taylor, S. and T. Drummond, 2011. Binary histogrammed intensity patches for efficient and robust matching. International Journal of Computer Vision, 94: 241–265.
20. Katiyar, A., K. Kalra and C. Garg, 2015. Marker Based Augmented Reality. Advances in Computer Science and Information Technology, 5: 441–445.
21. Lee, C., S. Bonebrake, T. Hollerer and D.A. Bowman, 2010. The role of latency in the validity of AR simulation. 2010 IEEE Virtual Reality Conference (VR), IEEE, pp: 11–18.
22. Pausch, R., D. Proffitt and G. Williams, 1997. Quantifying immersion in virtual reality. SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co., pp: 13–18.
23. Wither, J., S. DiVerdi and T. Hollerer, 2007. Evaluating display types for ar selection and annotation. In ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, IEEE, pp: 1–4.
24. Pasman, W., A. van der Schaaf, R.L. Lagendijk and F.W. Jansen, 1999. Low latency rendering for mobile augmented reality. ASCI Research School Conference, Ubicom-Publication, pp: 1–5.
25. Padmos, P. and M.V. Milders, 1992. Quality criteria for simulator images: A literature review. Human Factors: The Journal of the Human Factors and Ergonomics Society, 34: 727–748.
26. Azuma, R. and G. Bishop, 1994. Improving static and dynamic registration in an optical see-through HMD. SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques, IEEE, pp: 197–204.
27. Olano, M., J. Cohen, M. Mine and G. Bishop, 1995. Combatting rendering latency. Proceedings of the 1995 symposium on interactive 3D graphics, Association for Computing Machinery, pp: 19–24.
28. Poot, H.J., 1995. Monocular perception of motion in depth, PhD thesis, University of Utrecht, Utrecht.
29. Gherghina, A., A.C. Olteanu and N. Tapus, 2013. A marker-based augmented reality system for mobile devices. 2013 11th RoEduNet International Conference.
30. Poupyrev, I., D. Tan, M. Billinghurst, H. Kato and H. Regenbrecht, 2001. Tiles: A Mixed Reality Authoring Interface. Proceedings of Human-Computer Interaction INTERACT'01, IOS Press IFIP.
31. Boonbrahm, S., P. Boonbrahm and C. Kaewrat, 2020. The Use of Marker-Based Augmented Reality in Space Measurement. International Conference on Industry 4.0 and Smart Manufacturing (ISM 2019) (issue 42), Francesco Longo, Feng Qiao, Antonio Padovano, pp: 337–343.
32. Frank, J. A., M. Moorhead and V. Kapila, 2017. Mobile Mixed-Reality Interfaces That Enhance Human–Robot Interaction in Shared Spaces, Frontiers.
33. Hart, S. G., 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 9: 904–908.
34. Kan, T. W., C. H. Teng and M. Y. Chen, 2011. Handbook of Augmented Reality. Springer, pp: 339–354.
35. Avalle, G., F. D. Pace, C. Fornaro, F. Manuri and A. Sanna, 2019. An Augmented Reality System to Support Fault Visualization in Industrial Robotic Tasks. IEEE Access, 7.
36. Puneet, J., J. Manweiler and R.R. Romit, 2015. OverLay: Practical Mobile Augmented Reality. MobiSys '15, Association for Computing Machinery, pp: 331–344.
37. Zhang, W., B. Han and P. Hui, 17. On the Networking Challenges of Mobile Augmented Reality. VR/AR Network '17, COMM.



38. Chen, T. Y., L. Ravindranath, S. Deng, P. Bahl and H. Balakrishnan, 2015. Glimpse: Continuous, Real-Time Object Recognition on Mobile Devices. *SenSys '15*, Association for Computing Machinery, pp: 155–168.
39. Bekele, D., M. Teutsch and T. Schuchert, 2014. Evaluation of binary keypoint descriptors. *2013 IEEE International Conference on Image Processing*, IEEE.
40. Juan, L. and O. Gwun, 2009. A Comparison of SIFT, PCA-SIFT and SURF. *International Journal of Image Processing (IJIP)*, 4: 143–152.
41. Reynolds, D. A., T. F. Quatieri and R. B. Dunn, 2000. Speaker Verification Using Adapted Gaussian Mixture Models. *Digital Signal Processing*, 10: 19–41.
42. Perronnin, F., Y. Liu, J. Sánchez and H. Poirier, 2010. Large-scale image retrieval with compressed Fisher vectors. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE.
43. Dastjerdi, A. V., H. Gupta, R. N. Calheiros, S. K. Ghosh and R. Buyya, 2016. Chapter 4 – Fog Computing: principles, architectures, and applications. *Internet of Things Principles and Paradigms*, Morgan Kaufmann.
44. Blanco-Novoa, O., T. M. Fernández-Caramés, P. Fraga-Lamas and M.A. Vilar-Montesinos, 2018. A Practical Evaluation of Commercial Industrial Augmented Reality Systems in an Industry 4.0 Shipyard. *IEEE Access*, 6: 8201–8218.
45. Bonomi, F., R. Milito, J. Zhu and S. Addepalli, 2012. Fog Computing and its Role in the Internet of Things. *Proceedings of the first edition of the MCC workshop on Mobile cloud computing (issue MCC '12: Proceedings of the first edition of the MCC workshop on Mobile cloud computing)*, Association for Computing Machinery, pp: 13–16.
46. Fraga-Lamas, P., D. Noceda-Davila, T.M. Fernández-Caramés, M.A. Díaz-Bouza and M. Vilar-Montesinos, 2016. Smart Pipe System for a Shipyard 4.0., *Sensors*, 12: 2186.
47. Angrisani, L., P. Arpaia, N. Moccaldi and A. Esposito, 2018. Wearable Augmented Reality and Brain Computer Interface to Improve Human-Robot Interactions in Smart Industry: A Feasibility Study for SSVEP Signals. *IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI)*, IEEE.

References

1. Makolkina M., Sharlaeva M. Use of the Otsu Method to Calculate the Adaptive Segmentation Threshold // 9th International conference on advanced infotelecommunications, ICAIT, 2020. V. 1. PP. 697–702.
2. Pirmagomedov, R., A. Ometov, D. Moltchanov, Xi Lu, R. Kovalchukov, E. Olshannikova, S. Andreev, Y. Koucheryavy and M. Dohler, 2020. Applying Blockchain Technology for User Incentivization in mmWave-Based Mesh Networks. *IEEE Access* (Volume: 8), pp: 50983–50994.
3. Pirmagomedov, R., D. Moltchanov, A. Ometov, M. Khan, S. Andreev and Y. Koucheryavy, 2019. Facilitating mmWave Mesh Reliability in PPDR Scenarios Utilizing Artificial Intelligence. *IEEE Access* (Volume: 7), pp: 180700–180712.
4. Makolkina M. A. Research of the user movement and traffic models for applications of augmented reality // *Electrosvyaz*. 2019. No. 8. pp. 13–22.
5. Makolkina M., Koucheryavy A.: Augmented Reality Applications Classification // *Telecom IT*. 2020. Vol. 8. Iss. 1. pp. 11-21 (in Russian). DOI DOI 10.31854/2307-1303-2020-8-1-11-21.
6. Bruno, F., L. Barbieri and M. Muzzupappa, 2020. A Mixed Reality system for the ergonomic assessment of industrial workstations. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 14: 805–812.
7. Milgram, P. and F. Kishino, 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems*, 12: 1321–1329.
8. Swaminathan, A., Y. Mao and M. Wu, 2006. "Robust and secure image hashing. *IEEE Trans. Inf. Forensics Security*, 2: 215–230.
9. Bottani, E. and G. Vignali, 2019. Augmented reality technology in the manufacturing industry: A review of the last decade. *IIE Transactions*, 51: 284–310.



10. Gurevich, P., J. Lanir and B. Cohen, 2015. Design and Implementation of TeleAdvisor: a Projection-Based Augmented Reality System for Remote Collaboration. *Computer Supported Cooperative Work (CSCW)*, 24: 527–562.
11. Livingston M. A., C. Zanbaka, J. E., Smallman H. S. (2005). Objective measures for the effectiveness of augmented reality. *Proceedings of the Virtual Reality conference (VR 2005)*, Bonn (Germany), 12-16 March 2005, pp.287–288.
12. Zhang, W., B. Han and P. Hui, 2018. Jaguar: Low Latency Mobile Augmented Reality with Flexible Tracking. *Proceedings of the 26th ACM international conference on Multimedia, MM*, pp: 355–363.
13. Akgul, O., H.I. Penekli and Y. Genc, 2016. Applying Deep Learning in Augmented Reality Tracking. *International Conference on Signal-Image Technology & Internet-Based Systems (issue 12)*, IEEE, pp: 47-54.
14. Genc, Y., S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, 2002. Marker-less tracking for AR: a learning-based approach. *Proceedings. International Symposium on Mixed and Augmented Reality, IEEE*.
15. Klein, G., 2006. Visual Tracking for Augmented Reality, PhD thesis, University of Cambridge, Cambridge.
16. Harris, C. and M. Stephens, 1988. A combined corner and edge detector, *The Plessey Company plc*.
17. Rosten, E. and T. Drummond, 2006. Machine Learning for High-Speed Corner Detection. *European Conference on Computer Vision*, pp: 430–443.
18. Farabet, C., C. Couprie, L. Najman and Y. LeCun, 2013. Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35: 1915–1929.
19. Taylor, S. and T. Drummond, 2011. Binary histogrammed intensity patches for efficient and robust matching. *International Journal of Computer Vision*, 94: 241–265.
20. Katiyar, A., K. Kalra and C. Garg, 2015. Marker Based Augmented Reality. *Advances in Computer Science and Information Technology*, 5: 441–445.
21. Lee, C., S. Bonebrake, T. Hollerer and D.A. Bowman, 2010. The role of latency in the validity of AR simulation. *2010 IEEE Virtual Reality Conference (VR)*, IEEE, pp: 11–18.
22. Pausch, R., D. Proffitt and G. Williams, 1997. Quantifying immersion in virtual reality. *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., pp: 13–18.
23. Wither, J., S. DiVerdi and T. Hollerer, 2007. Evaluating display types for ar selection and annotation. In *ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE, pp: 1–4.
24. Pasman, W., A. van der Schaaf, R.L. Lagendijk and F.W. Jansen, 1999. Low latency rendering for mobile augmented reality. *ASCI Research School Conference, Ubicom-Publication*, pp: 1–5.
25. Padmos, P. and M.V. Milders, 1992. Quality criteria for simulator images: A literature review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34: 727–748.
26. Azuma, R. and G. Bishop, 1994. Improving static and dynamic registration in an optical see-through HMD. *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, IEEE, pp: 197–204.
27. Olano, M., J. Cohen, M. Mine and G. Bishop, 1995. Combatting rendering latency. *Proceedings of the 1995 symposium on interactive 3D graphics*, Association for Computing Machinery, pp: 19–24.
28. Poot, H.J., 1995. Monocular perception of motion in depth, PhD thesis, University of Utrecht, Utrecht.
29. Gherghina, A., A.C. Olteanu and N. Tapus, 2013. A marker-based augmented reality system for mobile devices. *2013 11th RoEduNet International Conference*.
30. Poupyrev, I., D. Tan, M. Billinghurst, H. Kato and H. Regenbrecht, 2001. Tiles: A Mixed Reality Authoring Interface. *Proceedings of Human-Computer Interaction INTERACT'01*, IOS Press IFIP.
31. Boonbrahm, S., P. Boonbrahm and C. Kaewrat, 2020. The Use of Marker-Based Augmented Reality in Space Measurement. *International Conference on Industry 4.0 and Smart Manufacturing (ISM 2019) (issue 42)*, Francesco Longo, Feng Qiao, Antonio Padovano, pp: 337–343.
32. Frank, J. A., M. Moorhead and V. Kapila, 2017. Mobile Mixed-Reality Interfaces That Enhance Human–Robot Interaction in Shared Spaces, *Frontiers*.



33. Hart, S. G., 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 9: 904–908.
34. Kan, T. W., C. H. Teng and M. Y. Chen, 2011. Handbook of Augmented Reality. Springer, pp: 339–354.
35. Avalle, G., F. D. Pace, C. Fornaro, F. Manuri and A. Sanna, 2019. An Augmented Reality System to Support Fault Visualization in Industrial Robotic Tasks. IEEE Access, 7.
36. Puneet, J., J. Manweiler and R.R. Romit, 2015. OverLay: Practical Mobile Augmented Reality. MobiSys '15, Association for Computing Machinery, pp: 331–344.
37. Zhang, W., B. Han and P. Hui, 17. On the Networking Challenges of Mobile Augmented Reality. VR/AR Network '17, COMM.
38. Chen, T. Y., L. Ravindranath, S. Deng, P. Bahl and H. Balakrishnan, 2015. Glimpse: Continuous, Real-Time Object Recognition on Mobile Devices. SenSys '15, Association for Computing Machinery, pp: 155–168.
39. Bekele, D., M. Teutsch and T. Schuchert, 2014. Evaluation of binary keypoint descriptors. 2013 IEEE International Conference on Image Processing, IEEE.
40. Juan, L. and O. Gwun, 2009. A Comparison of SIFT, PCA-SIFT and SURF. International Journal of Image Processing (IJIP), 4: 143–152.
41. Reynolds, D. A., T. F. Quatieri and R. B. Dunn, 2000. Speaker Verification Using Adapted Gaussian Mixture Models. Digital Signal Processing, 10: 19–41.
42. Perronnin, F., Y. Liu, J. Sánchez and H. Poirier, 2010. Large-scale image retrieval with compressed Fisher vectors. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE.
43. Dastjerdi, A. V., H. Gupta, R. N. Calheiros, S. K. Ghosh and R. Buyya, 2016. Chapter 4 – Fog Computing: principles, architectures, and applications. Internet of Things Principles and Paradigms, Morgan Kaufmann.
44. Blanco-Novoa, O., T. M. Fernández-Caramés, P. Fraga-Lamas and M.A. Vilar-Montesinos, 2018. A Practical Evaluation of Commercial Industrial Augmented Reality Systems in an Industry 4.0 Shipyard. IEEE Access, 6: 8201–8218.
45. Bonomi, F., R. Milito, J. Zhu and S. Addepalli, 2012. Fog Computing and its Role in the Internet of Things. Proceedings of the first edition of the MCC workshop on Mobile cloud computing (issue MCC '12: Proceedings of the first edition of the MCC workshop on Mobile cloud computing), Association for Computing Machinery, pp: 13–16.
46. Fraga-Lamas, P., D. Noceda-Davila, T.M. Fernández-Caramés, M.A. Díaz-Bouza and M. Vilar-Montesinos, 2016. Smart Pipe System for a Shipyard 4.0,. Sensors, 12: 2186.
47. Angrisani, L., P. Arpaia, N. Moccaldi and A. Esposito, 2018. Wearable Augmented Reality and Brain Computer Interface to Improve Human-Robot Interactions in Smart Industry: A Feasibility Study for SSVEP Signals. IEEE 4th International Forum on Research and Technology for Society and Industry (RTSI), IEEE.

Голубкин Александр А.

магистрант Санкт-Петербургского государственного
университета телекоммуникаций им. проф.
М. А. Бонч-Бруевича

Golubkin Alexandr A.

Undergraduate, The Bonch-Bruevich Saint-Petersburg
State University of Telecommunications

Пирмагомедов Рустам Ярахмедович

кандидат технических наук, Университет Тампере,
rustam.pirmagomedov@tuni.fi

Pirmagomedov Rustam Y.

Candidate of Engineering Sciences, Tampere University,
rustam.pirmagomedov@tuni.fi