

УДК 004.93

**ОБЗОР МЕТОДОВ АНАЛИЗА ДЛЯ ПОСТРОЕНИЯ АДАПТИВНОГО
ИНТЕРФЕЙСА СЦЕНЫ ДОПОЛНЕННОЙ РЕАЛЬНОСТИ НА ОСНОВЕ
ТЕКСТОВОГО ЗАПРОСА****Самодурова Валерия Романовна,**

студент, магистрант 2 курс, факультет "Отдел магистратуры",
кафедра «Институт сквозных технологий»
Донской государственной технической университет
Россия, г. Ростов-на-Дону
Mironova22092001@mail.ru

Ралко Ольга Максимовна,

студент, магистрант 2 курс, факультет "Отдел магистратуры",
кафедра «Институт сквозных технологий»
Донской государственной технической университет
Россия, г. Ростов-на-Дону
olga32-13-33@yandex.ru

Аннотация

Современные технологии искусственного интеллекта (ИИ) и нейронных сетей трансформируют подходы к созданию адаптивных пользовательских интерфейсов. Одним из ключевых понятий такого интерфейса является предварительный анализ, который необходим ещё до начала построения сцены дополненной реальности. В данной работе будут рассмотрены различные методы анализа, которые позволяют сделать дополненную реальность более персонализированной и точной, обеспечить качественную интеграцию в реальный мир.

Ключевые слова: гиперсцена, дополненная реальность, тональность текста, анализ текста, генерация гиперсцены

**A REVIEW OF ANALYSIS METHODS FOR BUILDING AN ADAPTIVE
AUGMENTED REALITY SCENE INTERFACE BASED ON TEXTUAL
QUERIES****Samodurova Valeria Romanovna,**

Student,
Master's student 2st year,
Department of "Institute of End-to-End Technologies",
Department of Media Technologies Don State Technical University,
Rostov-on-Don, Russia.
Mironova22092001@mail.ru

Ralko Olga Maximovna,

Student, Master's student 2st year,

Department of "Institute of End-to-End Technologies",

Department of Media Technologies Don State Technical University,

Rostov-on-Don, Russia

olga32-13-33@yandex.ru

ABSTRACT

Modern artificial intelligence (AI) and neural network technologies are transforming approaches to creating adaptive user interfaces. One of the key concepts of such an interface is preliminary analysis, which is necessary even before the construction of an augmented reality scene begins. This paper will examine various analysis methods that enable augmented reality to become more personalized and accurate, ensuring high-quality integration into the real world.

Keywords: hyperscene, augmented reality, text sentiment, text analysis, hyperscene generation

Понятие сцены дополненной реальности

Сцена дополненной реальности (гиперсцена) – это интерактивное пространство в дополненной реальности (AR), состоящее из различных элементов и объектов, объединенных в одну сцену. Гиперсцена позволяет пользователям взаимодействовать с визуальными элементами с помощью жестов, голосовых команд или других средств управления, перемещаться по сцене, получать информацию, запускать анимации и медиаконтент [1].

Предлагаемый метод генерации гиперсцены

Исходя из потребности в создании интерфейса, которая реагирует на эмоциональное состояние пользователя [2], были выделены следующие этапы построения гиперсцены;

Анализ промпта, определение тональности текста и пользовательского окружения для правильного отображения контента;

Генерация сцены [3] на основе текста пользователя и определённой тональности;

Референцирование отдельных элементов гиперсцены к релевантным источникам в сети;

Взаимосвязь этапов генерации гиперсцены приведена на рисунке 1.

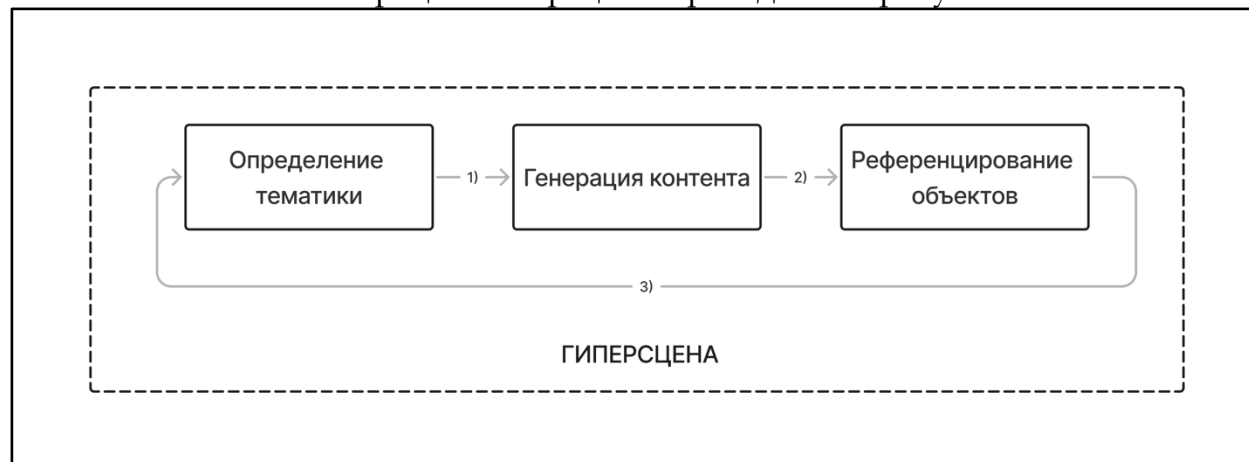


Рисунок 1 - Взаимосвязь компонентов генерации гиперсцены

Особое внимание стоит уделить первому этапу, который включает в себя анализ перед генерацией контента. Данный этап можно разбить на три ключевых момента: анализ промпта, анализ тональности текста и анализ окружения [3, 4,5]. Такой подход не только обеспечивает согласованность компонентов гиперсцены, но и делает пользовательский опыт индивидуальным. Поговорим подробнее о каждом из этих шагов анализа перед генерацией контента.

Анализ текста

В процессе формирования контента на основе анализа промптов [6] активно используются различные методы искусственного интеллекта, включая трансформеры, модели обучения с подкреплением, методы ранжирования и поиска, сетевые модели для креативного контента, а также модели эмбедингов. Каждый из этих подходов имеет свои особенности.

Трансформеры (например, BERT, T5) эффективно анализируют текст, понимая контекст и взаимосвязи, что позволяет получать ключевую информацию и анализировать синтаксис и семантику. Однако они требуют значительных вычислительных ресурсов.

Модели обучения с подкреплением адаптируются через взаимодействие с пользователем, повышая точность анализа текста, но их обучение может быть длительным и структурным.

Сравнительный анализ методов для анализа текста представлен в таблице 1.

Таблица 1 – Сравнительный анализ методов анализа текста

Метод ИИ	Применение	Преимущества	Ограничение	Ресурсы и сложности
Трансформеры (BERT, T5)	Понимание контекста, извлечение информации, анализ синтаксиса и семантики	Высокая точность, хорошее понимание контекста и взаимосвязей.	Требуются большие вычислительные ресурсы, контекст ограниченной длины	Высокие вычислительные ресурсы, сложность обучения
Модели обучения с подкреплением	Улучшение анализа через взаимодействие с пользователем, адаптация	Адаптивность, возможность улучшения моделей при меньшем использовании	Длительное время обучения, сложность настройки	Требуют времени на обучение и настройку
Методы ранжирования (BM25, ColBERT)	Поиск, фильтрация и сортировка информации	Эффективность фильтрации и извлечения важных данных	Ограничения при генерации нового контента, не рисовать текст	Относительно низкие требования к вычислениям
Сетевые модели для анализа текста	Анализ тональности, эмоций, смысловых связей	Хорошо анализируют тональность и эмоции, выделение	Ограничения в обработке сложных контекстов,	Средние требования к вычислениям

		смысловых элементов.	многозначных слов	
Модели эмбедингов (Word2Vec, GloVe)	Семантический анализ текста, выделение аналогичных и противоположных слов	Простота применения, эффективность для анализа детских слов	Не всегда можно обрабатывать сложные контексты или многозначность.	Низкие требования к вычислениям

Методы ранжирования (например, BM25) эффективно фильтруют и сортируют информацию, извлекая наиболее релевантные данные, но ограничены в создании нового контента.

Сетевые модели позволяют анализировать тональность, эмоции и смысловые связи, однако их способность работать с более сложным контекстом ограничена.

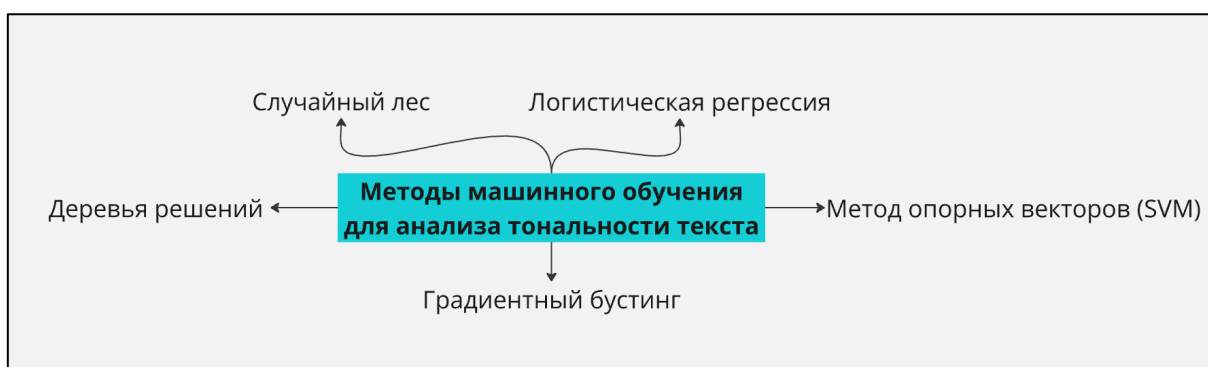
Таким образом, выбор зависит от задач, которые необходимо решить: для высокого анализа текста на сложном слое контекста и формирования лучшего результата подходят трансформеры, для быстрой фильтрации информации – методы ранжирования, для анализа эмоций и тональности – сетевые модели.

Анализ тональности текста

Анализ контента по его настроению может применяться во многих сферах жизни: экономической, информационной, спортивной, СМИ и другие. Текстовый анализ тональности имеет ряд сложностей, связанных с субъективной оценкой, сарказмом, многозначностью. Сейчас выделяют несколько основных подходов для решения данной проблемы: подход, основанный на машинном обучении и подход, использующий словари тональной лексики.

Существует множество методов традиционного машинного обучения. Они представлены на рисунке 2.

Рисунок 2 - Методы машинного обучения для анализа тональности



Произведём сравнительный анализ методов по следующим критериям: принцип работы, простота реализации, точность, необходимость в больших данных, склонность к переобучению, эффективность на больших данных, работа с несбалансированными данными, чувствительность к качеству данных, объяснимость результатов. Результат сравнения представлен в таблице 2.

Таблица 2 – Сравнительный анализ методов машинного обучения для анализа тональности текста

Критерии	Логистическая регрессия	Метод опорных векторов (SVM)	Наивный байесовский классификатор	Метод максимума энтропии	Деревья решений	Случайный лес	Градиентный бустинг
Принцип работы	Линейный классификатор, оценивающий вероятность принадлежности объекта к классу. Использует регуляризацию.	Находит оптимальную гиперплоскость, разделяющую классы, максимизируя отступ до ближайших точек.	Использует теорему Байеса с предположением независимости признаков.	Вероятностный классификатор, максимизирующий неопределенность при заданной информации.	Древовидная структура, где ветви представляют условия, а листья – вероятности классов.	Ансамбль решающих деревьев, объединяющий их результаты для повышения точности.	Итеративно улучшает модель, строя деревья с учетом ошибок предыдущих этапов.
Простота реализации	+	-	+	+	+	-	-
Высокая точность	+	+	-	+	-	+	+
Необходимость больших данных	-	-	-	-	+	-	+
Склонность к переобучению	-	-	-	-	-	-	+
Эффективность на больших данных	+	+	-	+	-	+	+
Работа с несбалансированными данными	+	+	-	+	-	-	+
Чувствительность к качеству данных	+	+	-	+	+	+	-
Объяснимость результата	+	-	+	-	-	-	-

В зависимости от итоговой потребности при использовании модели оптимальными могут являться разные методы [6]. При необходимости быстрой оценки, где не используется глубокая настройка модели, подходят метод логистической регрессии и наивный байесовский классификатор.

Таким образом, каждый из методов имеет свои слабые и сильные стороны и выбирается исходя из требований к реализации.

Анализ окружения

Для плавной интеграции виртуального мира и реального мира требует провести глубокий анализ окружения пользователя. Он включает в себя следующие этапы: сканирование и картографирование, распознавание поверхностей и объектов [5], оценка освещения и динамики [4].

Для того, чтобы выполнить анализ окружения приходится использовать множество технологий искусственного интеллекта. Основным аспектом является распознавание объектов и их контекста окружения, что требует применения методов компьютерного зрения [3]. Эти методы позволяют идентифицировать различные объекты в реальном времени, а также сегментировать сцены для более точной интеграции виртуальных объектов в реальный мир.

Для обеспечения точного позиционирования и навигации устройства в пространстве используется технология одновременной локализации и картографирования (SLAM). SLAM-системы анализируют пространство, создавая 3D-карты и отслеживая положение устройства, что обеспечивает стабильное размещение виртуальных объектов в изменяющейся среде. Кроме того, использование сенсоров и алгоритмов для анализа глубины позволяет точно воспроизводить трехмерные структуры реального мира.

Особое внимание уделяется анализу, позволяющему сделать сцену более реалистичной. Методы ИИ, использующие нейронные сети, способны оценивать освещенность и корректировать визуальные элементы с учетом окружающего света, создавая иллюзию естественного взаимодействия виртуальных объектов с реальной средой. Дополнительно, для обработки динамических изменений в сценах, таких как движение объектов или изменение положения камеры, применяются алгоритмы предсказания траекторий и анализа временных последовательностей.

Таблица 3 - Применение ИИ для анализа окружения

Область анализа	Используемые ИИ-модели и алгоритмы	Применение
Компьютерное зрение	- CNNs (ResNet, MobileNet, EfficientNet)	Распознавание объектов, текстур, поверхностей.
	- Mask R-CNN, U-Net	Сегментация объектов и пространств.
	- SIFT, ORB, SURF	Определение ключевых точек и особенностей объектов.
	- Оптический поток	Анализ движения объектов в реальном времени.
SLAM (Simultaneous Localization and Mapping)	- Deep Learning SLAM (PoseNet, DROID-SLAM)	Построение 3D-карт и определение положения устройства.
	- ORB-SLAM	Точное позиционирование с минимальными вычислительными ресурсами.
	- RGB-D SLAM	Работа с RGB и глубинными сенсорами.
Данные о глубине и пространстве	- PointNet, PointNet++	Анализ облаков точек для реконструкции сцены.
	- MonoDepth, DeepLab	Оценка глубины сцены на основе 2D-изображений.
Распознавание объектов и сцен	- YOLO (You Only Look Once)	Быстрое и точное обнаружение объектов.
	- OpenPose	Анализ поз и движений людей.

	- Places365	Распознавание типа сцены (улица, помещение и т. д.).
Обработка освещения и теней	- Neural Radiance Fields (NeRF)	Реалистичная обработка освещения в 3D-сценах.
	- Lighting Estimation Models	Анализ и интеграция интенсивности и направления света.
Генеративные модели	- GANs (Pix2Pix, StyleGAN)	Генерация текстур, дополнение недостающих частей сцены.
	- Autoencoders	Восстановление повреждённых данных и текстур.
Анализ движений и траекторий	- Recurrent Neural Networks (RNNs)	Анализ временных последовательностей данных.
	- Long Short-Term Memory (LSTM)	Предсказание траекторий движущихся объектов.
Интеграция с AR-платформами	- Google ARCore, Apple ARKit, Microsoft Azure Spatial Anchors	Интеграция алгоритмов анализа освещения, поверхности и контекста сцены.
Аппаратное ускорение	- NVIDIA Jetson, Apple Neural Engine (ANE)	Ускорение работы ИИ для анализа окружения в реальном времени.

Для создания интерактивных сцен дополненной реальности необходимо обеспечить реализацию сложных вычислительных задач в реальном времени, что обеспечивается с помощью интеграции с современными AR-платформами и использовании специализированных аппаратных решений.

В таблице 3 приведены области анализа, их применение и используемые для этого ИИ-модели.

Вывод

В статье рассмотрены этапы генерации гиперсцены по текстовому запросу, методы для анализа текста, нахождения тональности текста и анализа окружения. Представлены сравнительные таблицы.

Список литературы:

1. Иванова, З. И. (2021). Учебные материалы с дополненной реальностью в высшем профессиональном образовании. БГЖ, №1 (34). URL: <https://cyberleninka.ru/article/n/uchebnye-materialy-sdopolnennoy-realnostyu-v-vysshem-professionalnom-obrazovanii> (дата обращения: 25.12.2024).
2. Александрова, И. П., & Соловьев, А. Н. (2018). Модели и алгоритмы для построения сцен дополненной реальности в системах обучения. Известия высших учебных заведений. Прикладная нелинейная динамика, 26(4), 42-50.
3. Кузнецов, И. А., & Котляров, И. Ю. (2020). Применение методов компьютерного зрения для построения сцен дополненной реальности. Информационные технологии и вычислительные системы, 5(3), 95-102.
4. Азума, Р. (1997). Обзор дополненной реальности. Presence: Teleoperators & Virtual Environments, 6(4), 355-385.
5. Джеральд, Дж. (2015). The VR Book: Human-Centered Design for Virtual Reality. ACM Books.

6. Панг, Б., & Ли, Л. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

References:

1. Ivanova, Z. I. (2021). Educational Materials with Augmented Reality in Higher Professional Education. BGJ, No. 1 (34). URL: <https://cyberleninka.ru/article/n/uchebnye-materialy-sdopolnennoy-realnostyu-v-vysshem-professionalnom-obrazovanii> (accessed: December 25, 2024).
2. Aleksandrova, I. P., & Solovyev, A. N. (2018). Models and Algorithms for Building Augmented Reality Scenes in Educational Systems. Izvestiya Vysshikh Uchebnykh Zavedeniy. Applied Nonlinear Dynamics, 26(4), 42-50.
3. Kuznetsov, I. A., & Kotlyarov, I. Yu. (2020). Application of Computer Vision Methods for Building Augmented Reality Scenes. Information Technologies and Computing Systems, 5(3), 95-102.
4. Azuma, R. (1997). A Survey of Augmented Reality. Presence: Teleoperators & Virtual Environments, 6(4), 355-385.
5. Jerald, J. (2015). The VR Book: Human-Centered Design for Virtual Reality. ACM Books.
6. Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.