

МУЛЬТИМОДАЛЬНЫЙ ПОДХОД К ОБРАБОТКЕ ДАННЫХ: ТЕКСТ, АУДИО, ВИЗУАЛЬНЫЕ И СЕНСОРНЫЕ СИГНАЛЫ

Дяченко А.С., преподаватель,

Комендаков Д.И., студент,

Тимофеев Б.И., студент,

ФГБОУ ВО «МелГУ», г. Мелитополь, Россия

Аннотация. Статья посвящена мультимодальному подходу к обработке данных, включающему текст, аудио, визуальные и сенсорные сигналы, и их интеграции в образовательные системы. Рассматриваются возможности повышения эффективности обучения через персонализированные и интерактивные интерфейсы, а также ключевые вызовы реализации мультимодальных платформ.

Ключевые слова: мультимодальные данные, интеграция, образование, визуальные сигналы, сенсорика, интерфейсы.

Современные информационные системы всё чаще сталкиваются с необходимостью обработки разнообразных источников данных, выходящих за рамки традиционного текстового контента. Концепция мультимодальных данных предполагает интеграцию информации, поступающей из различных каналов – текстов, аудиозаписей, визуальных материалов и сенсорных сигналов – с целью более полного понимания и моделирования сложных процессов. Такой подход позволяет преодолевать ограничения однотипных источников информации, обеспечивая более богатое представление о предметной области и способствуя формированию интерактивных, адаптивных и контекстно-зависимых систем. В образовательной сфере мультимодальные данные становятся особенно ценными, так как они способны учитывать разнообразные формы восприятия учащихся и улучшать эффективность взаимодействия с учебным контентом [2, 4].

Современные образовательные технологии демонстрируют значительные достижения в персонализации обучения, однако они по-прежнему ограничены в способности интегрировать различные типы информации в единую когнитивную модель пользователя. Большинство платформ ориентировано на текстовые материалы или статические визуальные компоненты, что не позволяет полноценно учитывать разнообразие каналов восприятия учащихся. Более того, традиционные подходы редко используют аудиовизуальные и сенсорные данные для анализа эмоционального и когнитивного состояния обучающихся, что снижает потенциал адаптивной обратной связи и ограничивает возможности создания действительно интерактивной образовательной среды [1, 3]. Таким образом, внедрение мультимодального подхода становится необходимым условием для преодоления существующих ограничений и повышения качества образовательных систем.

Для наглядного сравнения особенностей различных типов данных целесообразно представить их ключевые характеристики в виде таблицы 1, которая позволяет сопоставить цели, способы восприятия, преимущества и ограничения каждого канала информации.

Таблица 1–Сравнение типов данных в мультимодальных образовательных системах

Тип данных	Цель	Способ восприятия	Преимущества	Ограничения
Текст	Передача знаний	Чтение	Структурированность, анализ	Ограниченная вовлечённость
Аудио	Усиление понимания	Слух	Эмоциональная окраска, память	Не подходит для всех условий
Визуальные	Визуализация концепций	Зрение	Наглядность, мгновенное понимание	Требует внимания и концентрации
Сенсорные	Мониторинг состояния	Осязание, физиология	Адаптация интерфейса, интерактивность	Сложность сбора, конфиденциальность

Данная таблица демонстрирует, как каждый тип данных играет уникальную роль в образовательном процессе, дополняя другие каналы восприятия и обеспечивая более полное и интерактивное освоение материала.

Особое внимание в контексте образовательных интерфейсов уделяется текстовым данным, которые являются одним из ключевых элементов передачи знаний. Текст обеспечивает структуру информации, позволяет формализовать учебный материал и служит основой для автоматического анализа, извлечения знаний и генерации рекомендаций. Однако эффективность текстовых данных возрастает многократно при их интеграции с аудио- и визуальными источниками, а также с сенсорными сигналами, которые отражают физическое и эмоциональное состояние пользователя. Например, сочетание анализа речи, визуальных выражений лица и сенсорных данных о внимании позволяет создавать адаптивные интерфейсы, способные подстраиваться под индивидуальные особенности обучающегося, прогнозировать затруднения и предлагать оптимальные стратегии обучения.

Аудиоинформация играет ключевую роль в формировании комплексного образовательного опыта, позволяя усилить восприятие и понимание материала за счет многоканального воздействия на когнитивные процессы учащихся. Звуковые сигналы, включая речь, музыкальные элементы и звуковые эффекты, способны передавать эмоциональные оттенки, структурировать информацию и облегчать запоминание сложных концепций. В образовательных интерфейсах аудио становится инструментом не только передачи знаний, но и создания динамичной обратной связи, реагирующей на поведение обучающегося. Интеграция аудиоданных с текстовой информацией способствует формированию более глубоких когнитивных связей, стимулирует внимание и способствует активному вовлечению в учебный процесс, что особенно важно для адаптивных и персонализированных образовательных систем.

Не менее значимую роль в мультимодальных образовательных платформах играют визуальные данные, которые обеспечивают мгновенное представление информации и упрощают понимание сложных явлений.

Графики, схемы, анимации и видеоматериалы позволяют визуализировать абстрактные концепции, создавать наглядные связи между элементами учебного материала и формировать целостное восприятие знаний. В сочетании с текстом и аудио визуальные данные повышают эффективность обучения, предоставляя учащемуся возможность усвоить информацию через разнообразные каналы восприятия. Использование визуализации также открывает новые возможности для адаптивного обучения: анализ взглядов, отслеживание внимания и реакции на визуальные стимулы позволяет создавать интерактивные среды, которые подстраиваются под индивидуальные особенности каждого пользователя.

Современные мультимодальные системы приобретают еще больший потенциал при интеграции сенсорных данных, фиксирующих физические и поведенческие параметры обучающихся. Сенсорная информация, включая данные о движении, прикосновениях, биометрические показатели и физиологические реакции, позволяет создавать интерактивные образовательные среды, реагирующие на состояние учащегося в реальном времени [1, 2]. Например, использование сенсоров для отслеживания жестов или мимики может позволить платформе корректировать сложность заданий, предлагать дополнительные подсказки или модифицировать подачу материала в зависимости от уровня внимания и эмоциональной вовлеченности. Такой подход обеспечивает глубокое понимание взаимодействия пользователя с учебным контентом и открывает перспективы для персонализированного, адаптивного обучения, способного максимально эффективно сочетать текст, аудио и визуальные элементы с сенсорными сигналами [5].



Рисунок 1 –Схема взаимодействия с пользователем

Таким образом, объединение аудио, визуальных и сенсорных данных в мультимодальном подходе создает комплексные образовательные системы, которые способны обеспечивать более естественное, интерактивное и персонализированное восприятие знаний. Эти интегрированные технологии формируют основу для новых методик обучения, в которых каждая форма данных взаимно дополняет другие, повышая вовлеченность, понимание и эффективность образовательного процесса в целом.

Эффективная интеграция мультимодальных данных требует разработки комплексных подходов и технологий, способных объединять текст, аудио, визуальные и сенсорные сигналы в единую когнитивную модель. Современные методы включают использование алгоритмов машинного обучения, нейронных сетей и гибридных моделей, которые позволяют синхронизировать данные разных типов, выявлять скрытые взаимосвязи и извлекать значимую информацию для персонализированных образовательных решений. Центральное значение имеет обработка и согласование временных и семантических характеристик каждого канала, поскольку успешная интеграция подразумевает не просто совмещение информации, но и обеспечение когерентного, контекстно-зависимого восприятия данных пользователем.

При этом реализация мультимодальных интерфейсов сталкивается с рядом вызовов и проблем, которые ограничивают их широкое внедрение. Среди ключевых сложностей можно выделить вопросы стандартизации данных, обеспечение их совместимости и синхронизации, а также высокие требования к вычислительным ресурсам и алгоритмической точности. Дополнительные трудности возникают при учёте вариативности восприятия и когнитивных особенностей пользователей, что требует создания адаптивных моделей с высокой степенью персонализации. Этические и конфиденциальные аспекты также играют значительную роль, поскольку сенсорные данные и биометрические показатели включают чувствительную информацию, требующую защиты и корректного использования [3].

Для структурирования проблем и потенциальных решений целесообразно представить посредством рисунка 2 основные вызовы мультимодальных систем и соответствующие направления их преодоления.



Рисунок 2 – Вызовы и перспективы развития мультимодального подхода

Инфографика позволяет наглядно оценить ключевые ограничения текущих технологий и перспективные направления их развития, помогая понять, какие шаги необходимы для создания более эффективных мультимодальных образовательных платформ.

Перспективы развития мультимодальных подходов обусловлены как технологическими инновациями, так и растущими требованиями к персонализированному и интерактивному обучению. Будущие исследования могут быть направлены на совершенствование методов слияния разнородных данных, разработку более эффективных алгоритмов анализа, а также на

создание универсальных стандартов представления мультимодальной информации. Кроме того, актуальным направлением является изучение влияния мультимодальных систем на когнитивные и эмоциональные процессы пользователей, что позволит разрабатывать интерфейсы с более высоким уровнем адаптивности и вовлечённости.

Таким образом, развитие мультимодального подхода открывает новые горизонты для образовательных технологий, стимулирует инновации и создаёт основу для формирования интеллектуальных, интерактивных и персонализированных образовательных платформ будущего.

Литература

1. Anadkat K., Solanki A., Patel D., Thakkar V. Enhancing emotion recognition with multimodel approach using deep neural networks // RT&A. – 2025. – № 1 (82). – URL: <https://cyberleninka.ru/article/n/enhancing-emotion-recognition-with-multimodel-approach-using-deep-neural-networks> (датаобращения: 22.10.2025).

2. Tsankov N. Multimodal competence of future teachers – a context of understanding and opportunities for development in university education // IJCRSEE. – 2025. – № 2. – URL: <https://cyberleninka.ru/article/n/multimodal-competence-of-future-teachers-a-context-of-understanding-and-opportunities-for-development-in-university-education> (датаобращения: 22.10.2025).

3. Дяченко А. С., Мухина Н. Ю. Цифровая этика в эпоху данных: баланс приватности, безопасности и свободы в обществе и образовании / А. С. Дяченко, Н. Ю. Мухина // Философские, социологические и психолого-педагогические проблемы современного образования. – 2025. – № 7. – С. 42-47. – EDN NАНХТА.

4. Кац Н. Г., Рубцова А. В., Аитов В. Ф. Иноязычный мультимодальный текст как основа разработки учебных материалов для цифровой образовательной среды вуза // Вестн. Том. гос. ун-та. – 2024. – № 504. – URL: <https://cyberleninka.ru/article/n/inoyazychnyy-multimodalnyy-tekst-kak-osnova->

razrabotki-uchebnyh-materialov-dlya-tsifrovoy-obrazovatelnoy-sredy-vuza (дата обращения: 22.10.2025).

5. Сечкарь Ю. А. Методологическое проектирование системы профессиональной подготовки педагогов с применением парадигмы продуктового подхода / Ю. А. Сечкарь // Научный потенциал. – 2025. – № 3-2(50). – С. 51-53. – EDN ICZGZM.