

ние), могут помогать вычислительной модели, обученной на первичной модальности. Этот вопрос актуален, когда основной способ имеет ограниченные ресурсы – отсутствие аннотированных данных, шумный ввод и ненадежные метки. Перенос вызова обусловлен передачей информации из вторичной модальности, порождает новое поведение, ранее невидимое в первичной модальности. Выделяются три типа подходов к переносу: кросс-модальный перенос, мультимодальное взаимосвязанное обучение и индукция модели [6, 8].

Количественная оценка – техническая задача, цель которой – обеспечить более глубокое эмпирическое и теоретическое исследование, чтобы лучше понять неоднородность между модальностями, межмодальные взаимодействия и мультимодальный процесс обучения. Измерение неоднородности – количественная оценка измерений неоднородности в мультимодальных данных и то, как они впоследствии влияют на моделирование и обучение; взаимодействия взаимосвязей модальности – количественная оценка наличия и типа связей и взаимодействий в мультимодальных наборах данных и обученных моделях; процесс обучения – характеристика задач обучения и оптимизации, возникающих при обучении на основе разнородных данных. Всестороннее понимание этих основных задач в контексте мультимодальных проблем позволяет получить представление и повысить их надежность, интерпретируемость и надежность в реальных мультимодальных приложениях [8, 4].

#### Список литературы

- 1 Barsalou, Lawrence W. Grounded cognition / L. W. Barsalou // Annual review of psychology. – 2008. – Vol. 59. – Pt. 1. – P. 617–645.
- 2 Kagnet: Knowledge-aware graph networks for commonsense reasoning / Bill Yuchen Lin [et al.] // In Proceedings of the 2019 Conference on Empirical Methods in Natural Language and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). – 2019. – P. 2829–2839.
- 3 Compression, transduction, and creation: A unified framework for evaluating natural language generation / M. Deng [et al.] // In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. – 2021. – P. 7580–7605.
- 4 Multibench: Multiscale benchmarks for multimodal representation learning / P. P. Liang [et al.] // arXiv preprint arXiv:2107.07502. – 2021.
- 5 Louwerse, Max M. Symbol interdependency in symbolic and embodied cognition / Max M. Louwerse // Topics in Cognitive Science. – 2011. – Vol. 3. – Pt. 2. – P. 273–302.
- 6 Humor knowledge enriched transformer for understanding multimodal humor / Md Kamrul, Hasan [et al.] // In Proceedings of the AAAI Conference on Artificial Intelligence. – 2021. – Vol. 35. – P. 12972–12980.
- 7 Reasoning with neural tensor networks for knowledge base completion / R. Socher [et al.] // In Advances in neural information processing systems. – 2013. – P. 926–934.
- 8 Vision guided generative pre-trained language models for multimodal abstractive summarization / Tiezheng Yu [et al.] // In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. – 2021. – P. 3995–4007.
- 9 Characterizing and overcoming the greedy nature of learning in multi-modal deep neural networks / N. Wu [et al.] // International Conference on Machine Learning. – 2022. – P. 24043–24055. – PMLR.
- 10 Heterogeneous graph learning for visual commonsense reasoning / Yu Weijiang [et al.] // Advances in Neural Information Processing Systems. – 2019. – Vol. 32.

УДК 656.078.1

## ОСОБЕННОСТИ ПОСТРОЕНИЯ МУЛЬТИМОДАЛЬНЫХ СИСТЕМ

*О. В. КОРНЕЕВ*

*Белорусский государственный университет транспорта, г. Гомель*

Необходимость применения мультимодальных структур характерна для многих отраслей жизнедеятельности человека, так как возросла потребность в гибком, мультикритериальном подходе к исследованию, прогнозированию и управлению процессами, учитывая неоднородность и зашумленность представляемого объема данных [7, 4]. Одна из основных задач при проектировании мультимодальных систем (далее – МС) заключается в том, чтобы создать условия для самостоятельного научения системы комбинировать различные режимы в стратегически интегрированную целостную систему [5].

С точки зрения проектирования и организации системы эффект Ломбарда является рефлексивным [1], это означает, что он не может быть краткосрочно устранен с помощью инструкции, тре-

нинга или подавлен выборочно при появлении шума. Отсюда определяются два принципа строения мультимодальности систем: неоднородность и взаимосвязанность. Модальности неоднородны, потому информация, приведенная в разных модальностях, демонстрирует разные качества, структуры и представления. Модальности не являются независимыми объектами, а скорее взаимосвязаны, потому что содержат в себе дополненную информацию, которая имеет способность взаимодействовать, в случае интеграции модальностей. Информация создается одной физической единицей посредством разных способов и поступает в другую физическую единицу посредством тех же способов. Трудности взаимодействия и интерпретации могут возникать, когда встречаются разные доминирующие модальности или разные по структуре модальности.

Принцип неоднородности отражает наблюдения о том, что информация, представленная в разных модальностях в основном, демонстрирует разные качества. Гетерогенность рассматривается как спектр, а не просто как бинарный выбор между однородным и гетерогенным [2]. Размер неоднородности говорит, каким образом модальности могут отличаться по представлению элементов – единицы анализа для каждой модальности, присущие им основные элементы и то, как можно их представить; распределению по частоте и вероятности присутствия элементов; структуре; информации; шуму в естественных данных; релевантности к задачам и контекстам (определенные модальности могут быть более полезны для определенных задач).

Принцип взаимосвязанности прослеживается в процедуре обработки данных, хотя информация неоднородна, она не обрабатывается независимо и часто связана дополнительной информацией которой обменивается. Эти взаимосвязи имеют два аспекта: модальные связи (описывают как модальности, часто связаны и имеют общие черты, например, соответствие между одним и тем же понятием в языке и изображении или пространстве) и модальные взаимодействия (как элементы взаимодействуют, чтобы дать начало новой информации, когда объединяются вместе для вывода) [2].

Существует различие между модальными связями и взаимодействием: связи существуют внутри самих модальных единиц (модальных данных), в то время как взаимодействие возникает только тогда, когда модальности интегрированы и обрабатываются вместе. Другими словами, модальные взаимодействия возникают, когда некоторый вывод выполняется по двум и более модальностям, приводя к новому ответу.

Модальность ввода показывает количество элементов модальности ввода, которые взаимодействуют друг с другом. Бимодальные взаимодействия включают элементы из двух модальностей, мультимодальные – из большого числа модальностей, что влечет за собой сбор параллельных данных и высокий объем вычислений (рисунок 1 [2]).

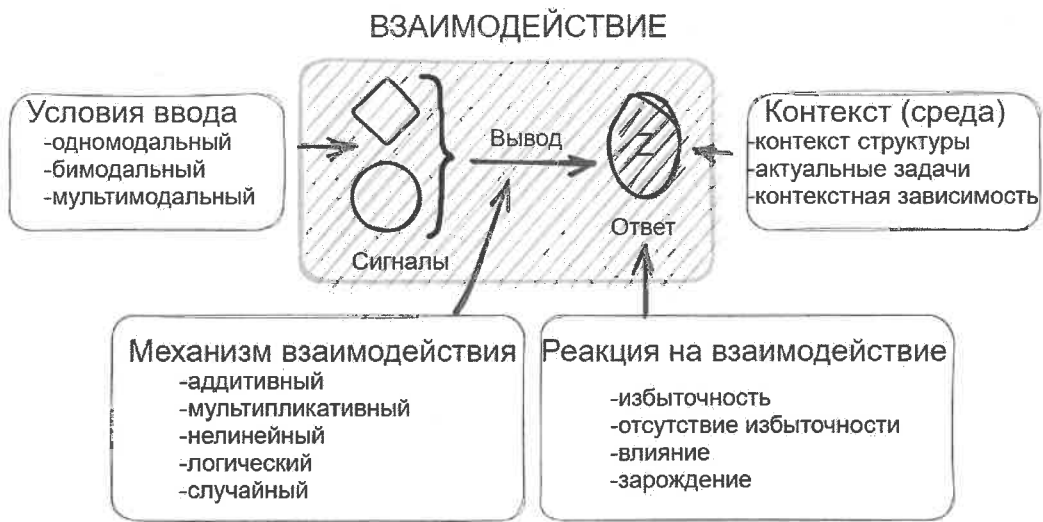


Рисунок 1 – Элементы модального взаимодействия

Механика взаимодействия показывает элементы, задействованные при интеграции элементов модальности для вывода. Например, взаимодействия могут существовать в виде аддитивной функции  $f(x, y) = g_1(x) + g_2(y)$  унимодальных подфункций  $g_1, g_2$  [1, 6, 8, 9]. Аналогично мультипликативные взаимодействия включают функцию (порядка  $k$ )  $f(x, y) = g_1(x) \otimes W \otimes g_2(y)$  тензорное произ-

ведение тензора  $W$  порядка  $k$ . Помимо мультипликативных отношений механика взаимодействия может также включать нелинейные, логические или причинно-следственные операторы.

Реакция на взаимодействие показывает изменение предполагаемого ответа в присутствии элементов из нескольких модальностей [2]. Тип модального взаимодействия зависит от задачи прогнозирования или контекста. Для одного и того же набора элементов в разных модальностях, различные задачи ввода могут влиять на тип взаимодействий, задействованных между модальностями ввода, их механику и реакцию на взаимодействие. Контекст показывает влияние задачи прогнозирования на тип взаимодействия.

#### Список литературы

- 1 **Friedman, Jerome H.** Predictive learning via rule ensembles / H. Friedman, Jerome, E. Popescu Bogdan // The annals of applied statistics. – 2008. – Vol. 2. – P. 916–954.
- 2 **Liang, P. P.** Foundations and recent trends in multimodal machine learning: Principles, challenges, and open questions / P. P. Liang, A. Zadeh, L.-P. Morency // arXiv preprint arXiv:2209.03430. – 2022.
- 3 **Lombard, E.** Le signe de l'élévation de la voix / E. Lombard // Annales des Maladies de L'Oreille du Larynx du Nez et du Pharynx. – 1911. – Vol. 37. – P. 101–119.
- 4 Multimodality in VR: A survey / D. Martin [et al.] // ACM Computing Surveys (CSUR). – 2022. – Vol. 54, no. 10. – P. 1–36.
- 5 **Oviatt, Sharon.** Multimodal system processing in mobile environments / Sharon Oviatt // In Proceedings of the 13th annual ACM symposium on User interface software and technology. – 2000. – P. 21–30.
- 6 Detecting statistical interactions with additive groves of trees / D. Sorokina [et al.] // Proceedings of the 25th international conference on Machine learning. – 2008. – P. 1000–1007.
- 7 **Tobisková, N.** Multimodal Augmented Reality and Subtle Quidance for Industrial Assembly—A Survey and Ideation Method / N. Tobisková, L. Malmköld, T. Pederson // In International Conference on Human-Computer Interaction. – Springer. – Cham. – 2022. – P. 329–349.
- 8 **Tsang, M.** Detecting statistical interactions from neural network weights / M. Tsang, D. Cheng, Y. Liu // arXiv preprint arXiv:1705.04977. – 2017.
- 9 Feature interaction interpretability: A case for explaining ad-recommendation systems via neural interaction detection / M. Tsang [et al.] // arXiv preprint arXiv:2006.10966. – 2020.

УДК 004.8, 624.9

### ПРИМЕНЕНИЕ АЛГОРИТМА ХЕББА ДЛЯ ОПРЕДЕЛЕНИЯ УРОВНЯ ТЕХНИЧЕСКОГО СОСТОЯНИЯ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ

*О. И. КОС, В. Ю. СМЕРНОВ*

*Московский авиационный институт (НИИ), Российская Федерация*

В настоящее время реализуется схема, которая называется «управление техническим состоянием СТС по нормативу». Для реализации более прогрессивной стратегии необходимо в любой заданный момент времени иметь возможность оценить техническое состояние всех элементов сложных технических систем (СТС) и СТС в целом. Предлагается использовать интеллектуальный алгоритм для классификации текущего технического состояния элементов СТС.

В нашей стране эксплуатируется огромное количество СТС. Для того, чтобы снизить затраты на эксплуатацию и одновременно обеспечить заданный уровень надежности, необходимо применение вероятностных методов и алгоритмов при принятии управленческих решений.

Рассмотрим следующие уровни технического состояния СТС.

- 1 Дефекты отсутствуют.
- 2 Присутствуют мелкие дефекты, не оказывающие влияния на влияющие на безотказность работы.
- 3 Дефекты, развитие которых создаст угрозу безотказности работы.
- 4 Серьезные дефекты, оказывающие влияние на безотказность работы, требуют особых условий эксплуатации вплоть до введения ограничений.

Классификация элементов СТС по перечисленным уровням предлагается производить с помощью адаптированного алгоритма Хебба [3]. На основании входных значений описывающих повре-