

БАЗОВЫЕ ЭЛЕМЕНТЫ КОГНИТИВНОЙ МОДЕЛИ МЕХАНИЗМА ВОСПРИЯТИЯ РЕЧИ НА ОСНОВЕ МУЛЬТИАГЕНТНОГО РЕКУРСИВНОГО ИНТЕЛЛЕКТА

З.В. НАГОЕВ¹, И.А. ГУРТУЕВА²

¹ФГБНУ «Федеральный научный центр
«Кабардино-Балкарский научный центр Российской академии наук»
360002, КБР, г. Нальчик, ул. Балкарова, 2
E-mail: cgrkbnrcan@bk.ru

²Институт информатики и проблем регионального управления –
филиал ФГБНУ «Федеральный научный центр
«Кабардино-Балкарский научный центр Российской академии наук»
360000, КБР, г. Нальчик, ул. И. Арманд, 37-а
E-mail: iipru@rambler.ru

В данной работе проанализирована обобщенная архитектура, лежащая в основе практически всех современных систем автоматического распознавания речи. Кратко изложена необходимость разработки принципиально нового подхода к решению проблем распознавания речи. Предлагается формальное описание структуры акта речевосприятия для применения в качестве общей теоретической основы при разработке универсальных систем автоматического распознавания речи, высокоэффективных в условиях высокой зашумленности и ситуациях «cocktail party». Разработана общая структурная динамика процесса распознавания речи, позволяющая учесть лингвистические и экстралингвистические аспекты речевого сообщения. Доказана необходимость использования понятия артикуляционного события в качестве минимального базового паттерна распознавания звукового образа.

Процесс распознавания структурирован на основе функциональной детерминанты «ситуация». Необходимость анализа многочисленных источников информации, сопровождающих звуковое сообщение, отказ от поиска инварианта носят здесь принципиальный характер.

Формальными средствами для реализации выбраны мультиагентные системы. Мультиагентный подход позволяет дифференцировать и анализировать звуки разной природы. Это делает предложенную модель уникальной и дает ей преимущества в ситуации так называемой «cocktail party», а также в задачах, где уровень шумов крайне высок.

Ключевые слова: мультиагентные системы, искусственный интеллект, искусственные нейронные сети, распознавание речи.

ЛИТЕРАТУРА

1. Abdel-Hamid O., Mohamed A., Jiang H., Penn G. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition, Proc. IEEE Int. Conf. Acoust., Speech, Signal Process, 2012. Pp. 4277-4280.
2. Chomsky N.A. A Review of Skinner's Verbal Behavior, [Readings in the Psychology of Language], Prentice-Hall, Upper Saddle River, New Jersey, 1967. Pp. 636.
3. Coates A., Ng A.Y. Learning feature representations with k -means, *Neural Networks: Tricks of the Trade*, 2012. Pp. 561-580.
4. De Mulder W., Bethard S., Moens M.-F. A Survey on the Application of Recurrent Neural Networks to Statistical Language Modeling, *Computer Speech and Language*, 2015. № 30(1). Pp. 61-98.
5. Де Соссюр Ф. Курс общей лингвистики. Екатеринбург: Изд-во Уральского университета, 1999. Pp.256.

6. *Deng L., Li X.* Machine Learning Paradigms for Speech Recognition: An Overview, *IEEE Transactions on Audio, Speech, and Language Processing*, 2013. № 21(5). Pp. 1060-1089.
7. *Gazzaniga M.S.* Conversations in the Cognitive Neuroscience. The MIT Press, Cambridge, 1996. Pp. 752.
8. *Ghai W., Singh N.* Literature Review on Automatic Speech Recognition // *International Journal of Computer Applications*. 2012. № 41 (8). Pp. 42-50.
9. *Ghitza O.* Auditory nerve representation as a front-end for speech recognition in a noisy environment // *Computer Speech and Language*, 1986. Vol. 1. Pp. 109-130.
10. *Gupta V.* A Survey of Natural Language Processing Techniques // *International Journal of Computer Science & Engineering Technology*. 2014. № 5 (1). Pp. 14-16.
11. *Haikonen P.* The Cognitive Approach to Conscious Machines, imprint Academic, Exeter, UK, 2003. Pp. 300.
12. *Hinton G., Deng L., Yu D., et al.* Deep neural networks for acoustic modeling in speech recognition // *IEEE Signal Process. Mag.* 2012. № 29(6). Pp. 82-97.
13. *Juan B.H.* Speech Recognition in Adverse Environments. *Computer Speech and Language*, 1991. Vol. 5. Pp. 275-294.
14. *Jurafsky D., Martin J.* Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition. Prentice Hall, Boston, 2008. Pp. 1032.
15. *Kotseruba I., Tsotsos J.K.* A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications, 2016 available at: arxiv.org/abs/1610.08602.
16. *Мазуренко И.Л.* Компьютерные системы распознавания речи // *Интеллектуальные системы*. 1998. Т. 3. Вып. 1-2. С. 117-134.
17. *Minsky M.* The Society of Mind. Simon and Shuster. New York, 1988. Pp. 336.
18. *Mohamed A., Dahl G., Hinton G.* Acoustic modeling using deep belief networks // *IEEE Audio, Speech, Lang. Process.* 2012. № 20(1). Pp. 14-22.
19. *Морозов В.П., Вартамян И.А., Галунов В.И.* Восприятие речи: вопросы функциональной асимметрии мозга. Ленинград: Наука, 1988, 135 с.
20. *Нагоев З.В.* Интеллектика, или Мышление в живых и искусственных системах. Нальчик: Изд-во КБНЦ РАН, 2013. С. 232.
21. *Нагоев З.В., Нагоева О.В.* Извлечение знаний из многомодальных потоков неструктурированных данных на основе самоорганизации мультиагентной когнитивной архитектуры мобильного робота // *Известия КБНЦ РАН*. 2015. № 6 (68). С. 73-85.
22. *Нагоев З.В., Нагоева О.В.* Зрительный анализатор интеллектуального робота для обработки неструктурированных данных на основе мультиагентной нейрокогнитивной архитектуры // *Перспективные системы и задачи управления: материалы XII Всероссийской научно-практической конференции (г. Ростов-на-Дону, 2017 г.)*. ЮФУ, 2017. С. 457-467.
23. *Нагоев З.В., Денисенко В.А., Лютикова Л.А.* Система обучения автономного сельскохозяйственного робота распознаванию статических изображений на основе мультиагентных когнитивных архитектур // *Устойчивое развитие горных территорий*. 2018. № 2. С. 289-297.
24. *Nagoev Z., Lyutikova L., Gurtueva I.* Model for Automatic Speech Recognition Using Multi-Agent Recursive Cognitive Architecture, Annual International Conference on Biologically Inspired Cognitive Architectures BICA, 2018, Prague, Czech Republic <http://doi.org/10.1016/j.procs.2018.11.089>
25. *Newell A.* Unified Theories of Cognition. Harvard University Press, Cambridge, Massachusetts, 1990. Pp. 576.
26. *Rabiner L.R., Schafer R.W.* Digital Processing in Speech Signal. Moscow: Radio and communications, 1981. Pp. 496.

27. *Reddy R.* Speech Recognition by Machine: A Review, Proceedings of the IEEE, 1976. № 64 (4). Pp. 501-531.
28. *Ронжин А.Л., Карпов А.А., Кагиров И.А.* Особенности дистанционной записи и обработки речи в автоматах самообслуживания // Информация и системы управления. 2009. № 5. Pp. 32-38.
29. *Schunk D.H.* Learning Theories: An Educational Perspective, Pearson Merrill Prentice Hall, Boston, 2011. Pp. 576.
30. *Van Veen B.D., Buckley K.M.* Beamforming: A Versatile Approach to Spatial Filtering // IEEE ASSP Magazine. 1988. № 5(2). Pp. 4-24.
31. *Waibel A., Lee K.-F.* Readings in Speech Recognition, Morgan Kaufman, Berlington, 1990. Pp. 680.
32. *Wooldridge M.* An Introduction to Multi-Agent Systems, Wiley, Hoboken, 2009. Pp. 366.
33. *Зиндер Л.Р.* Общая фонетика, М.: Высш. школа, 1979. С. 312.
34. *Zion Golumbic E.M., Ding N., Bickel S. et al.* Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party", 2013. Neuron, 77 (5). Pp. 980-991.

Нагоев Залимхан Вячеславович, к.т.н., врио председателя Кабардино-Балкарского научного центра Российской академии наук.

360000, КБР, г. Нальчик, ул. И. Арманд, 37-а.

Тел.: (8662) 42-65-62, факс: (8662) 42-65-62.

E-mail: zaliman@mail.ru

Гуртуева Ирина Асланбековна, н.с. отдела «Компьютерная лингвистика» Института информатики и проблем регионального управления – филиала Кабардино-Балкарского научного центра РАН.

360000, КБР, г. Нальчик, ул. И. Арманд, 37-а.

Тел.: (8662) 42-65-62, факс: (8662) 42-65-62.

E-mail: gurtueva-i@yandex.ru