

## ИНФОРМАЦИЯ

## НАУЧНАЯ КОНФЕРЕНЦИЯ “СОВРЕМЕННЫЕ РЕЧЕВЫЕ ТЕХНОЛОГИИ”.

Москва, 26–28 января 1999 г.

Конференция была организована Российским акустическим обществом (РАО) в память о Михаиле Андреевиче Сапожке – крупном ученом, заслуженном деятеле науки и техники России.

На конференции было заслушано около 50 докладов и выступлений ученых и специалистов, представлявших высшие учебные заведения, научные и производственные предприятия России и ряда стран СНГ. В докладах о жизни и научно-педагогической деятельности профессора, полковника-инженера М.А. Сапожкова, прочитанных Р.К. Потаповой (Московский государственный лингвистический университет – МГЛУ), В.Г. Михайловым (Московский государственный университет им. М.В. Ломоносова – МГУ), А.В. Рыжковым (16 Центральный научно-исследовательский испытательный институт связи министерства обороны – 16 ЦНИИИС МО), показана выдающаяся роль М.А. Сапожкова в становлении в России приоритетных направлений прикладных речевых исследований, включая автоматическое распознавание и синтез речи, вокодерную связь, тестирование каналов и средств телефонной связи.

В докладах на заседаниях четырех секций конференции были представлены достижения российских ученых и специалистов и намечены пути дальнейшего развития ряда направлений речевых технологий. Представляют интерес доклады по вопросам автоматического распознавания речи (В.Н. Сорокин, Институт проблем передачи информации Российской Академии Наук – ИППИ РАН, В.Я. Чучупал, Вычислительный центр РАН – ВЦ РАН), синтеза звучащей речи (О.Ф. Кривнова, МГУ, М.К. Румянцев, МГУ), тестирования каналов вокодерной связи (В.Е. Муравьев, НИИ радио минсвязи – НИИР), идентификации личности (Л.В. Златоустова, МГУ, И.Н. Тимофеев, Экспертно-криминалистический центр МВД России – ЭКЦ МВД, А.Ш. Каганов, Российский федеральный центр судебной экспертизы министерства юстиции РФ – РФЦСЭ) и ряд других.

В.Н. Сорокин (ИППИ РАН) изложил подход к автоматическому распознаванию речи, основанный на вычислении по акустическому речевому сигналу артикуляторных параметров (площади поперечного сечения речевого тракта и его формы). Для этого используются детекторы артику-

ляторных событий, регистрирующие динамические изменения речевых параметров. Автору удалось построить достаточно точные и устойчивые алгоритмы решения обратной задачи для речевого тракта, давшие положительные результаты по распознаванию гласных. Отмечаются преимущества предложенного подхода по сравнению с методами, основанными на скрытых марковских процессах и искусственных нейронных сетях.

Вместе с тем, следует отметить дальнейшее развитие традиционных подходов к решению проблемы взаимодействия “человек–ЭВМ”. По данным В.Я. Чучупала программный комплекс ВЦ РАН для среды Windows-95 обеспечивает распознавание слитно произносимых фраз произвольного диктора. Система распознавания последовательности слов состоит из акустического процессора для оценки параметров речевого сигнала, набора дискретных марковских моделей фонов, сетевой модели для представления пространства состояний и специальных процедур, учитывающих синтаксические ограничения в построении фразы.

Замечательные результаты получены по синтезу звучащей речи. Система Text-to-Speech (О.Ф. Кривнова, филологический факультет МГУ) обеспечивает высококачественный синтез звучащей русской речи по произвольному тексту, в основу которой положены методы канкатенации и синтеза по правилам. Метод канкатенации при соответствующем наборе исходных элементов синтеза обеспечивает точную передачу спектральных характеристик речевого сигнала с оформлением просодических параметров высказываний по правилам. Продемонстрированная запись синтезированной речи получила высокую оценку экспертов.

Очень значимыми представляются успехи в разработке высококачественных вокодеров, которые сегодня успешно внедряются в телефонную сеть. Достигнутое качество речи при скорости передачи 6.4 кбит/с оказалось выше, чем при АДИКМ-32 (Адаптивная дифференциальная импульсно-кодовая модуляция – Adaptive Pulse Code Modulation – ADPCM). На выставке Связь-98 была продемонстрирована отечественная аппаратура передачи речи с защитой информации “Омега”. Аппарат имеет два режима связи: режим ши-

фрования речи и режим шифрования данных (факсимильные аппараты и ПЭВМ подключаются по стандартному интерфейсу V.24/V.28 RS232C). В режиме шифрования речи используется метод линейного предсказания с высоким качеством речевого синтеза. В комплекте с модемом аппарат "Омега" обеспечивает работу по проводным телефонным, коротковолновым и ультракоротковолновым каналам связи. Аппарат "Омега" совместим с носимой станцией спутниковой связи и имеет следующие технические параметры: скорость передачи – 1200, 2400 бит/с, потребляемая мощность не более 11 ВА, масса 3.5 кг.

В докладе В.Е. Муравьева (НИИ радио) были перечислены стандарты для вокодеров, принятых Европейским институтом телекоммуникационных стандартов для сотовых систем подвижной связи стандарта GSM (Global System Mobile) – вокодер RPELTR (Regular Pulse Excitation – Long Term Prediction) и в дальнейшем – CELP (Code Excitation Linear Prediction) со скоростью передачи 13 кбит/с. Международным союзом электросвязи рекомендован для каналов видеотелефонии вокодер CELP со скоростью передачи 8 кбит/с – Стандарт G.729; для каналов Интернет – вокодер ACELP (Algebraic Code Excited Linear Prediction) со скоростью передачи 6.3 и 5.6 кбит/с – Стандарт G.723. Однако проблема объективной оценки качества речевой связи пока не нашла окончательного решения. Рекомендовано проводить оценку качества речи при реальном разговоре абонентов по дуплексному вокодерному каналу связи на основе применения метода мнений. Были предложены новые идеи и методы исследований, реализованные в компьютерных программах. Л.В. Златоустова (МГУ) проводит глубокие исследования по остаточным диалектным и иноязычным фонетическим явлениям русской речи, направленных на создание компьютерной базы данных. В описании остаточных явлений включены образцы наиболее частотных сегментных и суперсегментных характеристик севернорусских и южнорусских наречий, а также среднерусских говоров. Иноязычные остаточные явления изучены на алтайских языках, включая группу тюркских, корейский и японский. Работа представляет значительный теоретический лингвистический интерес, а также отвечает потребностям судебной фонографической экспертизы.

Экспертная система идентификации личности по фонограммам русской речи ЭКЦ МВД (И.Н. Тимофеев) является результатом почти 30-ти летней работы экспертов-криминалистов, защищена патентом России № 9611162551/09 от 1996 г. и в настоящее время с успехом используется в системе криминалистических лабораторий МВД. К несомненным достижениям этой систе-

мы следует отнести результаты выполнения почти 4000 судебных экспертиз.

Пакет программ речевых исследований ИКАР (Инструментальный комплекс анализа речи) Центра речевых технологий (ЦРТ, Санкт-Петербург, С.Л. Коваль) предназначен для ввода звучащей речи в компьютер, ее прослушивания, просмотра и редактирования осциллограмм (спектрограмм) и идентификационных криминалистических исследований фонограмм. Она находит применение в лабораториях фонографической судебной экспертизы Минюста и МВД России и в научно-исследовательских лабораториях ряда ВУЗов. Система устанавливается на компьютерах Pentium-100 и выше; состоит из пакета программ SIS (Speech Interactive System) и 16-разрядной платы Устройств Ввода/Вывода производства Центра речевых технологий. Система ИКАР обладает гибкими функциональными возможностями и эффективными программами измерения речевых параметров и их статистической обработки, содержит средства анализа, очистки звукозаписей от шумов и помех, компенсации неравномерности амплитудно-частотных характеристик (АЧХ) тракта звукозаписи, а также ускоренного документирования звукозаписей и их распечатки.

Автоматизированная система идентификации (В.Д. Сердюков, РФЦСЭ) оценивает близость голосов по статистическим оценкам дисперсии и корреляции амплитудно-частотных характеристик опорных сегментов фонограмм исследуемого голоса и группы голосов, отобранных из почти 100 эталонных образцов, хранящихся в справочном фонде системы. В результате расчетов система предлагает протокол вероятностной идентификации исследуемого голоса, его образца и одного из наиболее близкого эталона. К достоинствам системы следует отнести ее высокую производительность. В данное время система находится в стадии экспериментальной отработки.

Следует отметить представленные на конференции доклады с описанием модели речеобразования (Л.Г. Скалозуб, Киевский государственный университет, г. Киев), методов выделения основного тона (А.Н. Собакин, МГЛУ и В.Б. Гитлин, Ижевский государственный технический университет, г. Ижевск), а также компьютерной системы обучения (Т.А. Бровченко, Одесский государственный университет, г. Одесса).

На заседании за круглым столом участники конференции вспоминали видных ученых, внесших заметный вклад в исследования русской речи, среди них прозвучали имена: Л.Л. Мясникова, Л.Р. Зиндера, Н.Б. Покровского, Л.А. Варшавского, И.М. Литвака (г. Санкт-Петербург), Н.Н. Акинфиева, В.А. Артемова, Ю.С. Быкова,

Н.И. Жинкина, Л.З. Копелева, Г.И. Цемеля (г. Москва) и других.

В заключение конференции президент РАО и директор Акустического института имени академика Н.Н. Андреева председатель оргкомитета Н.А. Дубровский дал оптимистическую оценку перспективам развития речевых технологий в России и выразил благодарность за оказанную помощь в организации конференции и издании трудов Российскому фонду фундаментальных исследований, компании "Одитек" (г. Санкт-Петербург),

секции Акустические проблемы прикладной лингвистики РАО, а также МГЛУ и членам оргкомитета.

Для справки: Труды конференции, а также информационные материалы Российского акустического общества (РАО) и секции Акустических проблем прикладной лингвистики (АППЛ) приведены в Internet: <http://www.akin.ru> (РАО).

*В.Г. Михайлов*

Сдано в набор 19.03.99 г.

Подписано к печати 28.05.99 г.

Формат бумаги 60 × 88<sup>1</sup>/<sub>8</sub>

Офсетная печать

Усл. печ. л. 18.0

Усл. кр.-отт. 4.4 тыс.

Уч.-изд. л. 18.4

Бум. л. 9.0

Тираж 240 экз.

Зак. 2677