

О развитии речевых технологий в Белоруссии

Б.М. Лобанов,
доктор технических наук



Описывается 40-летняя история создания и развития в Белоруссии компьютерных систем распознавания и синтеза речи, включающая: начальный этап с 1965 по 1974 гг. в рамках группы исследования речи в МРТИ, развитие исследований с 1974 по 1988 гг. в составе лаборатории обработки речевых сигналов МОНИИС и новейшую историю разработок компьютерных систем распознавания и синтеза речи в лаборатории ИТК (ныне ОИПИ) НАН Беларуси. Описываются последние научные и практические разработки лаборатории и, в частности, компьютерное клонирование персонального голоса и дикции человека.

Введение

Из всего живого только человека Создатель наградил даром речи, благодаря чему ему удалось столь значительно развить свои интеллектуальные способности и, по мнению многих философов, стать человеку человеком. Нечто подобное происходит на наших глазах и с компьютером, интенсивно овладевающим широким спектром речевых технологий: от работы с телефонными сигналами до синтеза, распознавания и понимания речи ([см. http://www.speech.su.oz.au/comp.speech/](http://www.speech.su.oz.au/comp.speech/)). Вполне реально надеяться, что дальнейшее развитие человеко-машинных систем речевого общения неизбежно приведёт к постоянному совершенствованию языковой «грамотности» компьютеров и, в конечном итоге, к непрерывному приближению их интеллектуальных способностей к человеческим. Приведём уместную в данном контексте цитату из книги академика РАН Вячеслава Всеволодовича Иванова «Лингвистика третьего тысячелетия»:

«Мы впервые в истории вида начинаем широко пользоваться техническими говорящими орудиями — иначе говоря, не только изготавливать орудия (чем человек отличается от животных), но и обучать их нашему языку (чем мы начинаем отличаться от всех ранее живших людей)».

Однако на пути создания практически востребованных систем, реализующих речевые технологии, всё ещё предстоит преодолеть немало трудностей. Характерен при-

мер высказывания вице-президента компании **Microsoft** Гурдипа Синха: «... Билл Гейтс расстроен тем, насколько долго технологиям распознавания речи приходится «пробивать» себе дорогу. Ведь корпорация вкладывает деньги в этот проект, по крайней мере, с 1991 года. ... Говоря о новых средствах управления компьютерами, Гейтс подчёркивает: «естественные интерфейсы, возможно, одно из самых недооценённых явлений нынешней цифровой революции»...» (по материалам CNET News).

1. Начальный этап развития речевых исследований

Начало современной истории речевых исследований в СССР датируется серединой 60-х годов прошлого века, когда впервые начала работать Всесоюзная школа-семинар по автоматическому распознаванию слуховых образов (АРСО), собиравшая в лучшие годы до 300 участников. К этому же времени относится и начальный этап развития речевых исследований в Белоруссии. В 1965 году в научной лаборатории кафедры радиоприёмных устройств Минского радиотехнического института под руководством автора этой статьи была организована группа исследования речевых сигналов. В то время в неё входили Н.П. Дегтярёв, Б.В. Панченко, М. Фатеев и др., которые ещё долгое время, а некоторые из них и по сей день, работают в этом направлении.

Первые исследования группы были связаны с разработкой общих принципов анализа речевых сигналов и выделения информативных признаков, которые позволили бы представить непрерывный речевой сигнал последовательностью фонетических сегментов. Результаты этих исследований были обобщены в диссертации автора «Некоторые вопросы анализа речевых сигналов», защищённой в 1968 г. в Московском ГосНИИ Радио. Наиболее важные результаты этой работы позднее были опубликованы в авторитетных международных журналах [1–2]. На базе этих исследований впервые в СССР было разработано относительно простое устройство распознавания речевых команд «СЕЗАМ-2», получившее в 1968 г. серебряную медаль ВДНХ СССР. Устройство состояло из двух блоков: анализатора признаков речевого сигнала, таких как «Голосовой», «Шумный», «Гласный» и др., и счётчика количества признаков в речевой команде. Достигнута достаточно высокая надёжность распознавания 20 команд (включая названия цифр) независимо от голоса диктора, громкости и темпа произношения. В тот же период времени были разработаны специализированные приборы для экспериментально-фонетических исследований речи: анализатор динамических спектров и интонограф, с помощью которых в последующие годы проведены многочисленные исследования в фонетических лабораториях Института языкознания АНБ и Минского института иностранных языков.

Исследования динамических спектров речи дали толчок к развитию нелинейных методов сопоставления распознаваемых слов устной речи с их эталонами. Спектральные изображения речи, в отличие от обычных визуальных изображений объектов, могут подвергаться неконтролируемым нелинейным искажениям временной оси. На *рис. 1* приведены два спектральных изображения слова «автомашина», произнесённых одним и тем же диктором с различным темпом речи. Из рисунка видно, что при ускоренном темпе (нижняя спектрограмма) при общем сокращении длительности слова на 30% длительность

звуков ударного слога практически не изменилась, а некоторые звуки в безударных слогах сократились более чем в 2 раза. Из приведённого примера видно, что простого масштабирования спектральных изображений недостаточно для их надёжного распознавания. Ситуация, образно говоря, схожа с той, которая могла бы возникнуть в условиях «кривого зеркала» при распознавании зрительных образов.

Решение фундаментальной проблемы распознавания речи, связанной с нелинейными искажениями временной оси, было предложено независимо и практически одновременно Г.С. Слуцкером (Московский ГосНИИ Радио) и Т.К. Винцюком (Институт кибернетики АН УССР) во второй половине 60-х годов. Суть предложенного решения заключалась в нахождении методами динамического программирования (ДП-метод) оптимального пути на матрице локальных расстояний между временными отсчётами векторов распознаваемого и эталонного спектров. В 1969 году автором совместно с сотрудниками Московского ГосНИИ Радио была опубликована статья [3], в которой дано дальнейшее развитие ДП-метода для исключительно важного практического случая, когда границы распознаваемого слова неизвестны, т.е. для решения задачи обнаружения и распознавания

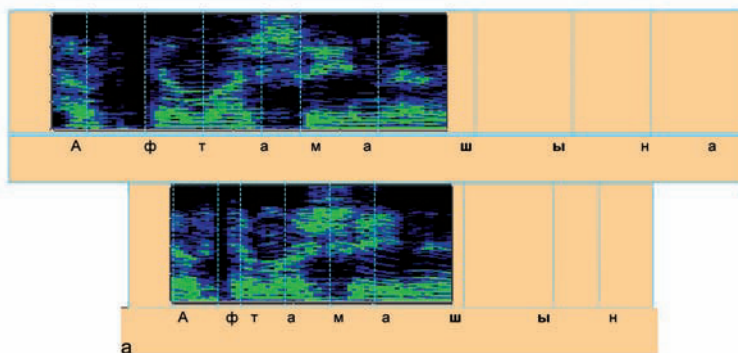


Рис.1. Спектральные изображения речи

и распознавания звуко сочетаний в непрерывном речевом сигнале. ДП-метод получил широкое признание зарубежных исследователей и, наряду с методом скрытых марковских моделей (СММ), до сих пор используется в современных системах распознавания речи.

К концу 60-х годов относится также начало работ по созданию синтезаторов речи. Стимулом послужило осознание того, что разработка и исследование



На фотографии Б.М. Лобанов, Т.К. Винцюк и В.Н. Трунин-Донской на семинаре у В.И. Галунова в мае 1982 г. Фото А.А. Харламова

моделей синтеза речи — это прямой путь к получению более детальных знаний о природе образования и свойствах речевого сигнала, опираясь на которые в дальнейшем можно будет построить более совершенные алгоритмы анализа и распознавания речи. Немаловажную роль в освоении мирового технологического уровня синтеза речи того времени сыграла научная стажировка автора этой статьи в 1970-м году в Лаборатории профессора Лоренца (Эдинбургский университет), где была разработана одна из первых

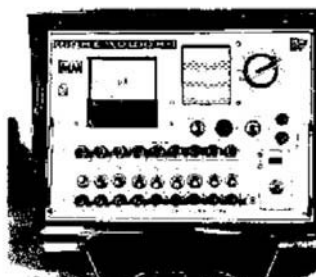


Рис. 2. Синтезатор «Фонемафон-1»

формантных моделей синтеза речевых сигналов. С помощью синтезатора этой лаборатории были впервые получены высококачественные образцы синтезированной русской речи.

Первая, пока ещё не вполне совершенная модель синтезатора русской речи по тексту «ФОНЕМАФОН-1» (рис. 2), «заговорила» в начале 70-х гг. и успех в её создании связан, прежде всего, с разработкой новых методов аппаратной реализации формантного синтеза речевых сигналов, которые основаны на моделировании свойств источников возбуждения (голосового и шумового) и резонансных (формантных) характеристик речевого аппарата человека. В результате экспериментальных исследований был создан полный набор формантных «портретов» фонем, позволивший впервые осуществить синтез русской речи по произвольному тексту. Позднее появилась улучшенная версия синтезатора — «Фонемафон-2», с дополнительным блоком преобразования «фонема — аллофон».

2. История «средних» лет (70-е — 80-е годы)

Решающую роль в ориентировке на решение прикладных задач использования речевых технологий сыграло создание в 1974 г. Лаборатории обработки речевых сигналов в составе Минского отдела центрального НИИ связи (ЦНИИС). В 1976 г. на Всесоюзном семинаре АРСО-9, проведённом в Минске, был впервые продемонстрирован прототип автоматической телефонной справочной службы с синтезированным речевым ответом, а уже с начала 80-х годов в Минске длительное время работала система автоматической обзвонки должников за междугородние переговоры (в их число иногда попадали и авторы этой разработки). К середине 80-х годов эта система была внедрена во многих городах СССР — от Бреста до Петропавловска-Камчатского. Успешному внедрению синтезаторов речи предшествовала длительная работа по совершенствованию как качественных показателей синтезированной речи, так и технологии их реализации в качестве нового класса внешних устройств ЭВМ. Основным недостатком первых моделей синтезаторов «Фонемафон-1, 2» были невысокие разборчивость и качество синтезированной речи, обусловленные прежде всего использованием предельно упрощённых моделей взаимодействия звуков в процессе речеобразования (эффектов коартикуляции и редукции) и недостаточно проработанной модели интонирования речи по тексту. В следующей модели — «Фонемафон-3» были введены дополнительные блоки моделирования процессов артикуляции и интонирования речи [4], что существенно повысило качественные показатели синтезированной речи.



Рис. 3. «Фонемафон-3» на выставке в Женеве

В 1979 г. «ФОНЕМАФОН-3» демонстрировался на Всемирной выставке «Телеком-79» в Женеве (Рис. 3), где известный фантаст Артур Кларк, посетив павильон СССР и ознакомившись с синтезатором речи, записал в книгу отзывов: «Вы предвосхитили мои фантазии из фильма «Космическая Одиссея — 2001», а швейцарская газета «Обозреватель» опубликовала статью: «Теперь русские изучают иностранные языки с помощью компьютера, который говорит».

Важную роль в создании промышленных синтезаторов речи сыграла разработка полностью цифровой модели синтезатора речи «Фонемафон-4», системы

распознавания речи «Сезам» и речевого терминала — «Марс» (рис. 4). Их серийный выпуск впервые в СССР (1983 г.) был налажен на ПО «Кварц» г. Калининграда благодаря энтузиазму сотрудников конструкторского отдела, возглавляемого Валерием Афонасьевым. Ключевую роль в их создании сыграли разработки сотрудников лаборатории МОЦНИИС Н. Дегтярёва и В. Шатерника. В речевом терминале «МАРС» впервые были интегрированы функции распознавания и синтеза речи. В основу алгоритмов распознавания речи положен упомянутый ранее ДП-метод принятия словесных решений на базе набора формантных признаков речевого сигнала. Опытные образцы систем «Сезам» и «Марс» были выполнены на микропроцессорной основе и по параметрам назначения не уступали лучшим зарубежным аналогам того времени. Оригинальность технических решений, использованных при создании систем «Фонемафон», «Сезам» и «Марс», защищена многочисленными авторскими свидетельствами СССР на изобретения.



Рис. 4. Речевой терминал «МАРС»

К началу 1984 г. относится окончательная формулировка, теоретическая и экспериментальная разработка единого лингво-акустического подхода к решению проблемы синтеза речи по тексту, его реализация в виде технических систем и практическое внедрение в составе автоматизированных систем управления и связи. Результаты этих исследований были обобщены в докторской диссертации автора «Методы автоматического синтеза русской речи по тексту», защищённой в 1984 г. в Институте электроники и вычислительной техники АН Латвийской ССР. Позднее полученные результаты были адаптированы для систем синтеза речи на других европейских языках. В частности, к 1987 г. благодаря сотрудничеству с профессором Минского института иностранных языков Е. Карневской была разработана англоязычная версия синтезатора [5], демонстрировавшаяся на Всемирном конгрессе фонетических наук и получившая высокую оценку англоязычных специалистов. Вот факсимиле отзыва об этой демонстрации одного из виднейших в мире исследователей речи Гуннара Фанта (см. рис. 5).

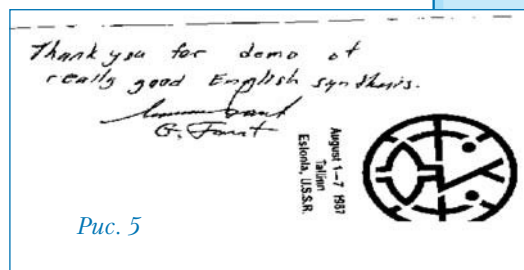


Рис. 5

3. Новейшая история

В 1988 г. на базе лаборатории МОНИИС в ИТК АНБ была создана лаборатория распознавания и синтеза речи, на должность заведующего которой дирекцией Института был приглашён автор этой статьи. Как все ещё хорошо помнят, конец 1980-х годов ознаменовался появлением первых персональных компьютеров, так что естественным образом в планах работ лаборатории появилась тематика, связанная с оснащением ПК системой речевого ввода — вывода информации. Формантный метод, который долгое время играл ключевую роль в системах синтеза речи по тексту, не подходил для этой цели из-за необходимости большого объёма вычислений в реальном времени.

В конце 80-х годов был предложен новый микроволновой (МВ) метод синтеза речевых сигналов [6], в котором вместо вычислений формантных колебаний (звуковых волн) использовался подготовленный заранее набор микроволн естественного речевого сигнала, который со-

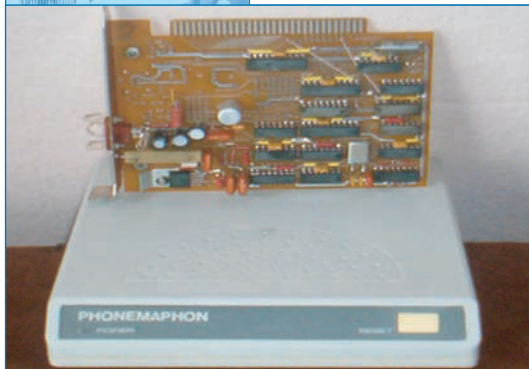


Рис. 6. Микроволновой синтезатор речи

стоял из отрезков сигнала, равных длительности периода, а их количество, необходимое для генерации любого звука речи, достигало нескольких сотен. МВ-метод воплощён Александром Ивановым в синтезаторе «ФОНЕМАФОН-5» в виде специализированного ПО синтезатора, ориентированного на работу с внутренней звуковой платой либо с автономным устройством, подключаемым к порту RS-232 (см. рис. 6). Удивительная для многих, компактность его ПО (всего 64К байт) позволила оснастить синтезом речи уже первые IBM PC-XT и даже отечественные ПК ЕС1840. Синтезатор речи был востребован во многих практических приложениях, но особенно широко, и до сих пор ещё он используется незрячими пользователями ПК (более

сотни комплектов специализированных аппаратно-программных продуктов для незрячих были созданы и распространены Георгием Лосиком в России, Украине и Белоруссии в первой половине 90-х годов). До сих пор ещё его вполне разборчивое звучание можно услышать, приобретая на рынке CD ROM «Говорящая мышь». В дальнейшем на основе МВ-метода разработаны версии для чешского и польского языков (примерно за 3 месяца пребывания в стране по приглашению заказчиков), а также автономный одноплатный модуль синтеза речи, украинско-язычная версия которого некоторое время работала на линии киевского метро.

Сложная экономическая ситуация, сложившаяся в стране в середине 90-х годов, заставляла искать источники финансирования исследований за рубежом, в первую очередь, в форме совместных международных проектов. Первым из них стал международный проект: «Двухязычный синтез речи — немецкий / русский» (1995–1996), выполнявшийся совместно с Дрезденским техническим университетом и финансировавшийся Германским научным фондом FTU Karlsruhe. Основные результаты этого проекта изложены в [7].

Следующим был проект «Анализ естественного языка и речи» (1996–1997), выполнявшийся совместно с Саарбрюкенским университетом (Германия), Манчестерским университетом (Великобритания) и Институтом проблем передачи информации (Россия) и финансировавшийся европейским фондом INTAS. Участие в этом проекте было связано с дальнейшим развитием моделей синтеза речи [8] путём их интеграции в системы обработки естественного языка методами компьютерной лингвистики.

Важную роль в интеграции белорусских исследователей в области лингвистики и речи в европейское сообщество стало участие в международном проекте «Развитие Европейской компьютерной сети по лингвистике и речи в восточном направлении» (1997–1998), финансировавшимся европейским фондом COPERNICUS. С 1998 г. лаборатория распознавания и синтеза речи ОИПИ является координационным центром этой сети в Белоруссии.

Кроме европейских научных организаций интерес к сотрудничеству с лабораторией проявили в эти годы и некоторые коммерческие организации. В 1996 г. французская фирма «Секстант Авионик» предложила реализовать научный проект «Распознавание речевых команд в условиях шумов в кабине самолёта». Проект финансировался фондом Министерства обороны Франции. Несмотря на исключительную сложность поставленной задачи, проект был успешно выпол-

нен и в 1997 г. принят заказчиком. Основные научные результаты этого проекта изложены в [9]. Другой коммерческой разработкой стал проект создания интеллектуального телефонного автоответчика, выполнявшийся с 1997 г. по договору с фирмой NovCom NV (США). Суть его заключалась в решении задачи распознавания произносимых по телефону имён абонентов и другой служебной информации с тем, чтобы система смогла выполнять функции телефонного автосекретаря. Проект завершён в 1999 г., а его основные научные результаты опубликованы в [10]. В выполнении этих международных проектов ключевую роль сыграли сотрудники лаборатории Татьяна Левковская, Александр Иванов и Андрей Кубашин. Работы по этим проектам явились стимулом к реанимации и дальнейшему развитию алгоритмов распознавания речи, предложенных ещё в 1969 г. [3].

4. Современные исследования и разработки

4.1. Компьютерное клонирование персонального голоса и дикции

Многолетние исследования, проведённые в XX веке, позволили создать синтезаторы, обеспечивающие качество и разборчивость речи, вполне пригодные для широкого спектра практических приложений. Однако, несмотря на все усилия, синтезированная речь оставалась ещё далёкой по качеству от натуральной и обладала узнаваемым машинным акцентом. Причиной этому были не столько уровень наших знаний о процессах речеобразования и о фонетике, сколько нехватка вычислительных ресурсов компьютеров того времени. Сейчас мы можем не ограничивать себя ни объёмом оперативной и дисковой памяти, ни требуемым объёмом вычислений, и это явилось предпосылкой к постановке задачи синтеза речи по тексту с максимально возможным приближением по звучанию к голосу и манере чтения конкретного человека.

Такая постановка задачи, хотя и отдалённо, напоминает широкоизвестную биологическую проблему клонирования, когда на основе носителя генетической информации делается попытка воспроизвести копию живого существа. В нашем случае, в отличие от классической задачи клонирования, делается попытка создания близкой копии, но не биологической, а компьютерной, и не всего существа в целом (в данном случае человека), а только одной из его интеллектуальных функций: чтение произвольного орфографического текста. При этом ставится задача максимально полного сохранения персональных акустических особенностей голоса, фонетических особенностей произношения и акцента, а также просодической индивидуальности речи (мелодика, ритмика, динамика). В принципе в генетике рассматривается и такая возможность, как создание своеобразных «химер» из разнородного генетического материала. Применительно к клонированию голоса и речи — это тот случай, когда в основу синтеза закладываются, например, акустика голоса одного диктора, фонетические особенности произношения — другого, а просодическая индивидуальность речи — третьего.

Общая структурная схема аллофонно-волнового синтезатора речи, на базе которого осуществляется клонирование, показана на [рис.7](#).

Синтезатор состоит из четырёх процессоров: лингвистического, просодического, фонетического и акустического. Каждый из процессоров использует для осуществляемых им преобразований специализированные базы данных (БД). В этих БД заложены как общие языковые правила (лингвистические, просодические, фонетические, акустические), так и правила, связанные с индивидуальными особенностями голоса и речи диктора.

Клонирование акустических характеристик голоса. Персональные акустические характеристики голоса человека обусловлены множеством факторов, таких как анатомические осо-

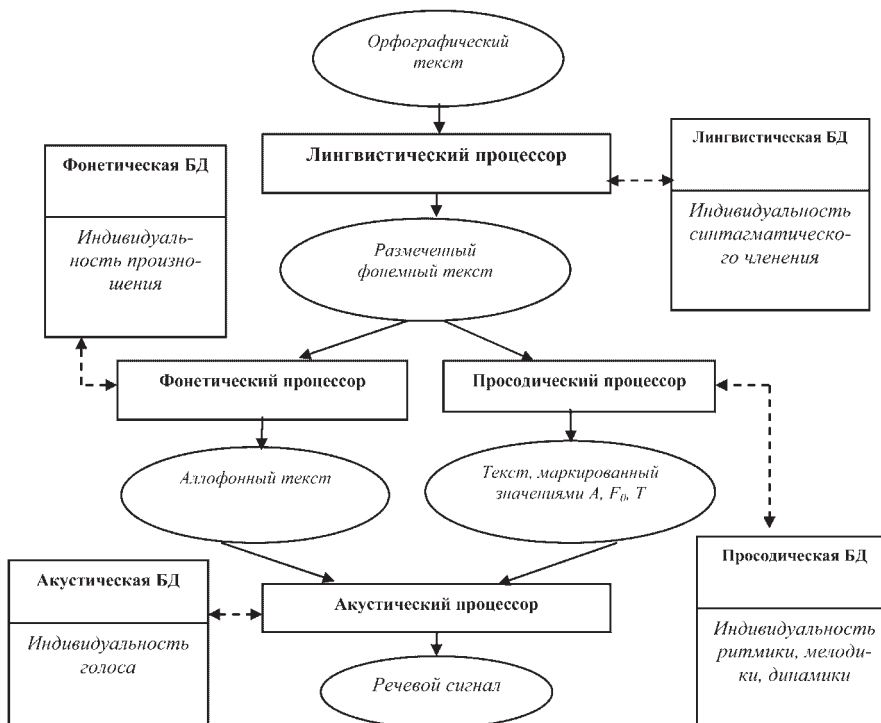


Рис. 7. Общая структурная схема аллофонно-волнового синтезатора речи

бенности строения и функционирования элементов речевого аппарата (гортань, голосовые связки, глотка, полость рта и др.), динамические особенности взаимодействия колебаний голосовых связок и резонаторов речевого аппарата, а также многое другое. Как известно, попытки имитации персональных характеристик голоса в системах «текст — речь» на основе моделирования физиологических и акустических процессов речеобразования из-за их чрезвычайной сложности до сих пор не привели к ощутимым результатам. В связи с этим наиболее разумным представляется использование отрезков натуральной речевой волны в качестве минимального «генетического материала» для клонирования голоса. В качестве такого отрезка целесообразно выбрать аллофон как наиболее изученную фонетическую субстанцию, причём ограниченный набор аллофонов способен обеспечить порождение устной речи произвольного содержания. При этом звуковая волна содержит в себе все существенные персональные особенности голосообразования, проявляющиеся в данном конкретном аллофоне.

Клонирование персональных фонетических особенностей произношения. В отличие от персональных акустических характеристик голоса, обусловленных в основном статическими параметрами речевого аппарата, фонетические особенности произношения обусловлены главным образом динамикой артикуляторных движений, осуществляемых в процессе речи. Присущие данному индивиду скорость артикуляторных движений, индивидуальные особенности артикуляции того или иного звука (например /P/), региональный или иностранный акцент обуславливают возникновение своеобразных позиционных и комбинаторных оттенков фонем и создают уникальную систему аллофонов. Таким образом, успешное клонирование персональных фонетических осо-

бенностей произношения может быть достигнуто путём имитации особенностей фонемно-аллофонного преобразования, присущего данному человеку в процессе речи.

Клонирование персональных просодических характеристик речи. Комплекс просодических (интонационных) характеристик речи, включающий мелодику, ритмику и энергетику, задаётся закономерными изменениями во времени частоты основного тона, длительности звуков и амплитуды звуковых сигналов. Характер этих изменений определяется не только конкретным текстом, но и персональной манерой его чтения. Решение задачи клонирования просодических характеристик речи заключается в создании достаточно полного набора персональных интонационных «портретов» речи.

Технология клонирования и её приложения. Для успешного клонирования персональных характеристик голоса и дикции необходимо создать достаточно полные наборы звуковых волн аллофонов и интонационных «портретов» речи. В случае, если клонируемый диктор физически доступен, для этой цели используется специально разработанный компактный звуковой массив слов и отрывков текста, начитываемый им в студии или в обычных условиях. Если же клонируемый диктор недоступен, то используются уже имеющиеся записи его голоса на радио, телевидении и др. Первые результаты по клонированию (на примере персонального голоса и дикции автора этой статьи) были получены в 2000 году и опубликованы в феврале 2001 года [11]. В 2001-м году получен клон женского голоса, а к концу 2005-го набор клонов состоял уже из 3-х мужских и 2-х женских голосов [12]. В настоящее время создана компьютерная технология клонирования фонетико-акустических [13] и просодических [14] характеристик, речи, позволяющая в значительной степени автоматизировать и ускорить процесс создания речевых клонов диктора.

Отметим некоторые возможные коммерческие и практические аспекты компьютерного клонирования. Наверняка найдётся большое количество пользователей компьютера, желающих, чтобы их PC заговорил его собственным голосом или, например, голосом близкого ему человека, или любимого актёра. Интересным может быть также проект оживления давно ушедших от нас голосов великих людей по оставшимся от них граммафонным или студийным записям. Разработка технологии создания голосовых клонов может оказаться кардинальным средством борьбы с так называемым телефонным терроризмом, обеспечив идентификацию личности по голосу путём автоматического сравнения оперативной записи голоса с содержимым БД голосовых клонов потенциальных правонарушителей [15].

4.2. Компьютерная модель речевого виртуального собеседника

Компьютерная модель устно-речевого виртуального собеседника (система «РЕВИРС») — новая разработка лаборатории распознавания и синтеза речи, в которой интегрированы оригинальные научно-технические решения, полученные сотрудниками лаборатории в течение последних лет. Система РЕВИРС позволяет создавать сценарии диалогов для разнообразных приложений и осуществлять их посредством устно-речевого человеко-машинного общения. Уникальность системы РЕВИРС заключается в том, что в ней реализуется:

- надёжное распознавание *ключевых слов* запроса в *непрерывном потоке речи*;
- *многодикторное* распознавание *ключевых слов* в *условиях акустических помех и искажений*;
- *многоголосый* синтез речи по произвольному тексту;
- возможность «*клонирования*» *голоса личности* в процессе синтеза речи;
- *дуплексный режим* в *реальном времени* (возможность прерывания голосового ответа).

Структурная схема системы приведена на [рис.8](#).

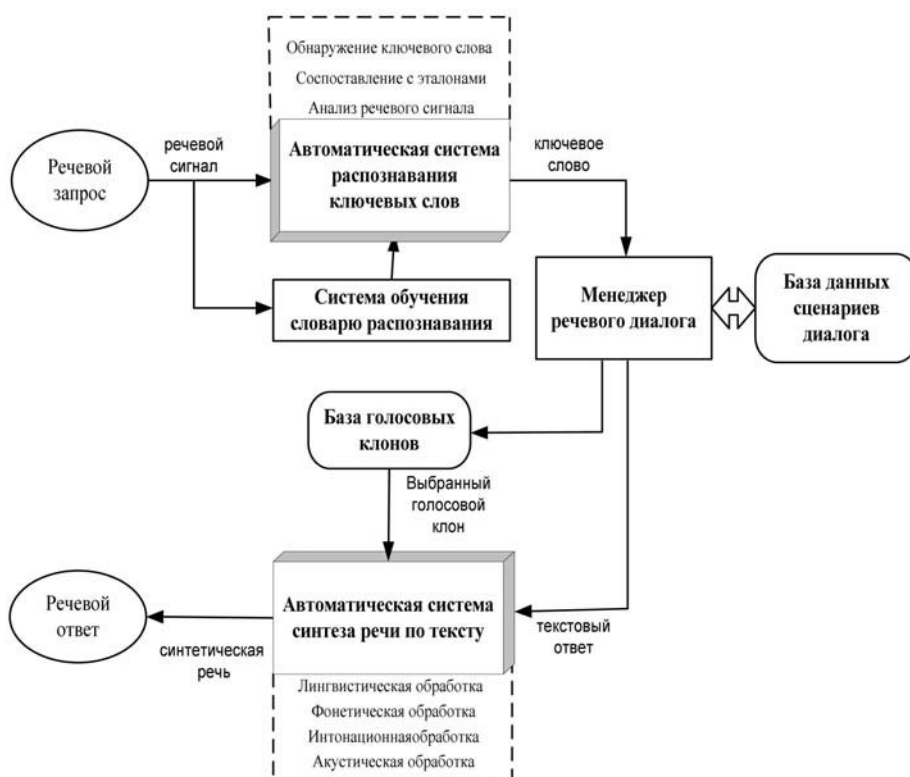


Рис. 8
Структурная
схема системы
РЕБИРС

Речевой сигнал поступает на вход системы распознавания речи, которая осуществляет анализ информативных признаков сигнала, их сопоставление с эталонами ключевых слов, обнаружение и принятие решения о произнесённом слове. Если распознающий модуль обнаружил ключевое слово, оно перенаправляется менеджеру речевого диалога, который формирует текстовый ответ. Менеджер речевого диалога выбирает также голосовой клон для синтеза речевого ответа. Выбранный клон и текст ответа поступают на вход системы синтеза речи, которая осуществляет лингвистическую, интонационную, фонетическую и акустическую обработку, в результате которой текст преобразуется в звучащую речь заданного голосового клона.

Заклучение

Изложенный в настоящей статье краткий очерк истории создания и развития речевых технологий в Белоруссии не претендует, конечно, на полноту освещения всех научно-технических результатов 40-летних исследований. Не было возможности даже кратко упомянуть работы всех авторов, внёсших определённый вклад в развитие этой отрасли знаний в Белоруссии (в общей сложности ими опубликовано не менее 300 научных работ, в т.ч. 5 монографий).

Несмотря на то что проделан большой путь и получены ощутимые научные и практические результаты, полагаю, что история речевых исследований в Белоруссии на этом не закончится. Как и 40 лет назад, разгадка природы человеческой речи остаётся увлекательной для молодых исследователей, а возможные приложения речевых технологий как никогда становятся актуальными.

Литература:

1. Lobanov B.M. More About Speech Signal and the Main Principles of its Analysis // IEEE Transactions on Audio and Electroacoustics. 1970. № 3. P. 316–318.
2. Lobanov B.M. Classification of Russian Vowels Spoken by Different Speakers // The Journal of the Acoustical Society of America. 1971. № 2 (2). P. 521–524.
3. Лобанов Б.М., Слуцкер Г.С., Тизик А.П. Автоматическое распознавание звукоочетаний в текущем речевом потоке // Труды НИИР. М., 1969. С. 67–75.
4. Lobanov B.M. Articulatory-Formant Speech Synthesis from Printed Text // Proceedings of Franco-Sovietique Symposium on Speech. Paris, 1981. P. 221–251.
5. Lobanov B.M. The Phonemophon Text-to-Speech System // Proceedings of the XI-th ICPHS, Tallin, 1987. P. 100–104.
6. Lobanov B.M., Karnevskaia E.B. MW — Speech Synthesis from Text // Proceedings of the XII International Congress of Phonetic Sciences.— Aix-en-Provence, France 1991.—P. 387–391
7. Lobanov B.M. et al. A Bilingual German/Russian Text-to-Speech System // Proceedings of the 3rd International Workshop «Speech and Computer» — SPECOM'98. St.-Petersburg, 1998. P. 327–330.
8. Lobanov B.M. et al. Generation of Intonation and Accentuation of Synthetic Speech on the Base of Morpho-Syntactic Knowledge // Proceedings of the International Workshop «Integration of Language and Speech». Moscow, 1996. P. 11–28.
9. Lobanov B.M., Levkovskaya T.V. Continuous Speech Recognizer for Aircraft Application // Proceedings of the 2nd International Workshop «Speech and Computer» — SPECOM'97. Cluj-Napoca, 1997. P. 97–102.
10. Lobanov B.M. et al. An Intelligent Answering System Using Speech Recognition // Proceedings of the 5th European Conference on Speech Communication and Technology — EUROSPEECH'97. V. 4. Rhodes-Greece, 1997. P. 1803–1806.
11. Лобанов Б.М. и др. Синтезатор персонализированной речи по тексту «ЛобаноФон-2000» // Труды Международной конференции, посвящённой 100-летию российской экспериментальной фонетики. СПб., 1–4 февраля 2001 г. С. 101–104.
12. Lobanov B.M., Tsiulnik L.I. Phonetic-Acoustical Problems of Personal Voice Cloning by TTS // Proceedings of the International Conference «Speech and Computer» — SPECOM'2004. St.-Petersburg, 2004.
13. Лобанов Б.М., Цирульник Л.И. Технология компьютерного клонирования и синтеза персональных характеристик речи диктора // Труды международной конференции Диалог'2007. М.: Издательский центр РГГУ, 2007. С. 559–565.
14. Lobanov, B., Tsiulnik L., Sizonov O. AUP's Modeling of Speaker Specific Intonation Contour Peculiarities // Speech and Computer: proceedings of the 12-th International conference SPECOM'2007, Moscow, Russia, 15–18 October, 2007 / Moscow State Linguistic University. — Moscow, 2007. V. 1. P. 312–317.
15. Лобанов Б.М. и др. Система экспресс-идентификации голоса личности методом клонирования акустических характеристик // Сб. докл. Международной конференции «Теория и практика речевой коммуникации». М., 2004.

Лобанов Борис Мефодьевич

Доктор технических наук.

Почётный радист СССР (1981), обладатель серебряной и бронзовой медалей ВДНХ СССР (1983), главного приза международного конкурса фирмы HEWLETT-PACKARD за работу «Распознавание голоса» (1992). С 1987 — член Международного акустического общества, с 1994 — координатор Белорусского отделения Европейской сети по компьютерной лингвистике и речи, с 1995 — член Европейской ассоциации речевых исследований, с 2001 — эксперт Европейской сети языковых технологий. Член докторских советов по защите диссертаций (ОИПИ НАН Беларуси, БГУИР, БГУ, МГЛУ), С 1998 — профессор БГУИР, с 2003 — профессор Университета в Белостоке.