

Результаты работы первого семинара «Обеспечение расследования, раскрытия и профилактики преступлений с использованием фоноскопических экспертиз»



А.Н. Слепич,

И.В. Рыжкова

В соответствии с планом мероприятий МВД России по подготовке и проведению Международного салона средств обеспечения безопасности «Комплексная безопасность-2008», утверждённом 24.11.2007, ГУ НПО «СТИС» МВД России совместно с ЭКЦ МВД России, провели семинар «Обеспечение расследования, раскрытия и профилактики преступлений с использованием фоноскопических экспертиз».

Проведение семинара направлено на поиск перспективных научно-технических решений для проведения фоноскопических исследований и мониторинг разработок в области речевых технологий.

В семинаре приняли участие ведущие организации, проводящие разработки в области речевых технологий, такие как Институт проблем передачи информации им. А.А. Харкевича РАН, Московский государственный лингвистический университет, Московский технический университет связи и информатики, Московский технический университет им. Н.Э. Баумана, ООО «Центр речевых технологий», ЗАО «НПП «ИСТА-Системс», Рязанский государственный радиотехнический радиоуниверситет, ООО «ЮНИКОР микросистемы», ООО «Скинер». Среди участников — 6 докторов и 12 кандидатов наук.

На семинаре были представлены доклады по трём направлениям развития речевых технологий, результаты которых используются в МВД России, это:

- технические средства для расследования преступлений с использованием фоноскопической экспертизы;
- средства контроля доступа на основе использования речевых технологий;
- преобразователи речи.

Расследование, раскрытие и профилактика преступлений основываются на возможности получения доказательной базы, предоставляемой фоноскопической экспертизы, эффективность которой определяется уровнем развития соответствующих технических средств.



На семинаре отмечено, что наиболее перспективным научно-техническим решением для проведения фоноскопической экспертизы на сегодняшний день является подход, использованный сотрудниками ЭКЦ МВД России, заключающийся в создании единого «типового технологического процесса» проведения экспертизы.

Этот подход получил развитие в разработке АПК «Сапфир». «Типовой технологический процесс» представлен в виде последовательности, автоматизированных основных этапов (видов) работ, выполняемых экспертом в процессе производства фоноскопической экспертизы.

Основными этапами работ являются:

- установление дословного содержания исследуемой фонограммы (путём многократного прослушивания фонограммы, дифференциации и атрибуции реплик участников разговора);
- сегментация речевых сигналов интересующих эксперта лиц;
- определение технических параметров и лингвистических характеристик речи исследуемых лиц с целью решения вопросов пригодности речевых сигналов для проведения акустического и лингвистического анализа;
- лингвистический анализ речи исследуемых лиц (определение индивидуальных и групповых признаков речи на уровне речевого потока, уровне фразы и слова и фонетическом уровне). Особую роль при проведении отдельного и сравнительно лингвистического анализа играют признаки фонетического уровня, которые могут даже при минимально допустимой длительности речевых фрагментов составить основную часть индивидуальной совокупности признаков речи исследуемого диктора. Анализ признаков фонетического уровня включает в себя сегментацию сопоставимых речевых фрагментов (как правило, триад звуков), которые используются для сравнения экспертом внутридикторских вариантов произнесения той или иной фонемы и для сопоставления вариантов произнесения данных фонем разными дикторами методами аудитивного и лингвистического анализа. Сопоставимые триады, подвергшиеся лингвистическому анализу, в дальнейшем используются для проведения акустического анализа;
- проведение акустического анализа сигнала: интегрального (макроанализ) и сегментного (микроанализ). На практике при проведении акустического анализа возникают ситуации, где решающую роль для принятия решения о сходстве или различии исследуемых дикторов играют признаки микроанализа звуков — наиболее трудоёмкого вида анализа, сопровождающегося поиском экспертом в спорных фонограммах идентичных речевых фрагментов (триад звуков) и измерении для них спектральных и временных характеристик.

Аппаратурная реализация «типового технологического процесса» проведения фоноскопической экспертизы в АПК «САПФИР» обеспечила следующие функциональные возможности:

1. Производить ввод/вывод сигнала из файла и со звуковой платы или других устройств.
2. Одновременно открывать несколько звуковых файлов (до 60).
3. Получать визуальное отображение сигнала, содержащегося в открытом звуковом файле или введённого с внешнего устройства (сигналы визуализируются в так называемых «Окнах осциллограммы», таких окон может быть открыто

несколько, в одном «Окне осциллограммы» может отображаться несколько звуковых файлов, имеющих разную частоту дискретизации).

4. Прослушивать звуковой сигнал в различных режимах (в режиме псевдостерео, в ускоренном или замедленном темпе без изменения тембра голоса, прослушивать сигнал, преобразованный различными способами).
5. Проводить различные виды спектрального и временного анализа сигнала (вычислять и визуализировать мгновенный спектр, средний спектр, спектрограмму, кепстрограмму, траектории формант, контур основного тона) — для визуализации результатов анализа предусмотрен удобный интерфейс, ориентированный на типичные действия, выполняемые экспертом-фоноскопистом.
6. Проводить исследование фонограммы на предмет наличия или отсутствия следов монтажа (путём исследования непрерывности фазовой составляющей выбранной гармонике спектра, непрерывности спектральных характеристик фоновых шумов, анализ спектрального состава сигнала фонограммы).
7. Создавать описание дословного содержания фонограммы: предусмотрен интерфейс для проведения сегментации фонограммы (на временные интервалы), составления дословного описания содержащихся в сегментах реплик и классификации реплик по дикторам. Введённое через АПК «САПФИР» дословное описание фонограммы сохраняется в файле, имеющем текстовый формат, и может быть прочитано, например, текстовым редактором Microsoft Word с отображением результатов сегментации и атрибуции сегментов по дикторам.
8. Просматривать дословное содержание фонограммы, помещённое в таблицу текстовых интервалов, и/или отображаемое под осциллограммой сигнала синхронно во времени.
9. Автоматически прослушивать и отображать в окне осциллограммы временной интервал, соответствующий выбранной экспертом реплике фонограммы.
10. Проводить автоматический поиск слов или их частей (триад) в дословном содержании фонограммы.
11. Создавать в автоматическом режиме на основе исходной фонограммы и составленного её дословного содержания новый звуковой файл, содержащий реплики выбранного экспертом диктора («дикторский» звуковой файл).
12. Копировать в автоматическом режиме результаты редактирования дословного содержания «дикторского» звукового файла в текстовый файл, содержащий дословное содержание исходной фонограммы.
13. Проводить автоматизированный поиск идентичных фрагментов речи в двух спорных фонограммах (алгоритм поиска ориентирован на то, что исследуемые фрагменты речи будут соответствовать триадам звуков, например, согласный — гласный — согласный).

В АПК «САПФИР» предусматривается система подсказок, помогающая эксперту выбрать оптимальный список параметров описания конкретного речевого фрагмента. АПК рассчитана на экспертов, обладающих самой разной квалификацией. Также следует отметить, что в связи со значительным увеличением в последнее время количества материалов, содержащих речь не на русском языке, предоставляемых для проведения фоноскопической экспертизы, необходимость в проведении автоматизированного поиска идентичных речевых фрагментов для решения задач идентификации иноязычных дикторов может стать особенно острой.



До недавнего времени использование речи в системах средства контроля доступа носило только вспомогательный характер. На прошедшем семинаре доложены результаты использования новых методов анализа, обеспечивающих верификацию диктора с достоверностью выше 99%. Реализация такого верификатора рассчитана на использование в виде «пароля» числительных русского языка. Количество числительных в «пароле» — десять. Для предотвращения несанкционированного доступа «пароль» меняется случайным образом при каждом обращении к системе.

В отдельных применениях для сокращения времени произнесения «пароля» число слов в нём может быть от трёх до семи. Для повышения достоверности число слов может быть увеличено до двенадцати.

Верификатор устойчив к стационарным шумам при SNR выше 12 дБ и не чувствителен к случайным помехам высокой интенсивности, в виде посторонних разговоров и музыки. Предусмотрена защита этого средства контроля доступа от попытки «вторжения» с использованием записи и последующего воспроизведения чужого голоса.

Разработка проведена в ИППИ им. А.А. Харкевича РАН. Испытание речевого средства контроля доступа на базе данных, включающей 429 дикторов (в объёме 125 часов речи), показали, что суммарная ошибка (несанкционированный пропуск и ошибочный отказ в доступе) для 90% дикторов составила менее 0,01%.

Вопросы преобразования речевой информации в текст были представлены на семинаре сотрудниками ООО «СКИНЕР», разработавшими в 2007 году систему автоматизированного преобразования русскоязычной речи в письменный вид.

При разработке систем автоматического распознавания речи и преобразование её в текстовый формат возникает проблема вариативного произношения слов (как одним человеком в различных эмоциональных состояниях, так и разными людьми). Кроме того, на входящий сигнал влияют многочисленные факторы, такие как окружающий шум, отражение, эхо и помехи в канале с заранее неизвестными параметрами.

Система распознавания состоит из следующих взаимосвязанных и взаимодействующих модулей:

- модуля ввода произнесённых слов;
- модуля настройки на особенности речи пользователя, формирования цифровых моделей вводимых слов и ведения словаря;
- модуля распознавания произнесённых слов;
- интеллектуального интерфейса.

Разработанная система распознавания речи основана на фоновно-ориентированном методе.

Разработанная система распознавания речи имеет следующие возможности:

1. Не требуется предварительная настройка системы на пользователя, что существенно упрощает эксплуатацию.
2. Реальная точность распознавания слов при однократном произнесении составляет не менее 75% при объёме словаря до 7000 слов.
3. Не требуется написание строгой грамматики для распознавания речи. Позволяет распознавать не только отдельные слова, но и предложения ограниченного размера.
4. Затрачивает сравнительно небольшое время на обработку звукового сигнала.

Такая система может применяться в call-центрах, для упрощения обслуживания клиентов, а также эту систему можно использовать в службах безопасности организаций.

Мониторинг существующих технических средств для проведения фоноскопической экспертизы показал, что

- положительный опыт использования технических средств, разработанных специально для использования в ЭКП ОВД России;
- появление систем, автоматического поиска записи речи известного лица (диктора) в сколь угодно больших массивах фонетического материала.

Отличительными особенностями технических средств, разработанных для использования в ЭКП ОВД России, является высокий уровень их автоматизации при выполнении идентификационных исследований по голосу и речи; автоматизация при проведении технического исследования фонограмм (идентификация средств звукозаписи, установление аутентичности фонограмм речи); возможность использования различных видов визуализации при проведении анализа звуковых сигналов.

Положительно отмечены АПК «Учёт-Ф», АПК «Фонексия» и АПК «Виртуоз», нашедшие своё применение в ЭКП МВД России.

АПК «Учёт-Ф» позволяет формировать и хранить базы данных, содержащие кодовые образы особенностей устной речи, установочные данные и фонограммы речи объектов, а также позволяет осуществлять поиск в учётном массиве объектов, характеристики речи которых (по установленным критериям) идентичны характеристикам объекта.

АПК «Учёт-Ф» позволяет:

- производить централизованный учёт для нужд МВД, производящих расследование по уголовным делам;
- проводить фоноскопические исследования в интересах оперативно-розыскных подразделений.

АПК «Фонексия» предназначен для автоматизированной диагностики акцентов и диалектов русской устной речи. Этот комплекс используется для проведения фоноскопических экспертиз, направленных на диагностику места рождения или места длительного проживания неизвестного субъекта (диктора), фонограмма речи которого исследуется.

АПК «Виртуоз» предназначен для поиска на речевых фонограммах признаков нарушения их аутентичности, в частности, следов монтажа фонограмм, а также следов цифровой обработки фонограмм и следов выборочной звукозаписи.

Системы автоматического поиска записи речи известного лица (диктора) в сколь угодно больших массивах фонетического материала могут использоваться для поиска лиц, представляющих оперативный интерес и для систематизации постоянно накапливающегося фонетического материала.

К таким системам относятся: «Трал Х», имеющая высокую, более чем 100 фонограмм в минуту, скорость поиска, а также «Голос», работающая в реальном масштабы времени.

Дополнительно отмечено, что для аппаратной поддержки АПК проведения фоноскопической экспертизы хорошо зарекомендовали себя устройства для измерения характеристик и формирования электрических сигналов в звуковом диапазоне частот STC-H246 «Камертон». Устройство «Камертон» сертифицировано как средство измерения.



Выводы по результатам работы семинара

1. По известным техническим средствам проведения фоноскопических экспертиз положительную оценку заслуживают: АПК «Виртуоз», АПК «Учёт-Ф», АПК «Фонексия»; дополнительно могут быть рекомендованы к внедрению в службы ЭКП МВД России программное средство «Ривьера» и АПК «Сапфир».
2. Расширение возможности, снижение времени проведения фоноскопической экспертизы связаны с разработкой технических средств, основанных на «технологии проведения фоноскопической экспертизы», объединяющей в единое целое все этапы экспертизы». Первой разработкой, намеченной к внедрению в ЭКП ОВД в 2008 году, является АПК «Сапфир».
3. Безусловным достижением речевых технологий является впервые обеспеченная возможность использования голоса человека для верификации в системах ограничения доступа. Достигнутый уровень достоверности превышает 0,99.
4. Отмечена перспективность начатых разработок в области создания преобразователей русской речи в текст, хотя используемые в настоящее время преобразователи имеют недостаточные функциональные возможности, ограничивающие их практическое применение.
5. Перспективными направлениями речевых технологий, с точки зрения использования подразделениями ОВД России, являются:
 - создание автоматических средств проведения фоноскопической экспертизы, направленной для расследования преступлений;
 - создание средств и технологий проведения лингвистической экспертизы;
 - разработка преобразователей речи на основе нейроинформационных технологий.

В целях постоянного отслеживания достижений в области речевых технологий, оценки возможности использования их результатов в МВД России, а также апробации результатов ведомственной науки принято решение о придании настоящему семинару статуса постоянно действующего.

Слепич Андрей Николаевич —

*начальник отдела научно-исследовательского института специальной техники ГУ НПО «СТИС» МВД России, образование высшее.
Контактный тел. (495) 673-91-49. e-mail: an@slepich.ru*

Рыжкова Ирина Владимировна —

*старший научный сотрудник Научно-исследовательского института специальной техники ГУ НПО «СТИС» МВД России, Образование высшее.
Контактный тел. (495) 673-91-06. e-mail: irina-ryzhkova@inbox.ru*