
СИСТЕМЫ РАСПОЗНАВАНИЯ ЛИЧНОСТЕЙ ПО ГОЛОСУ

УДК 343.98

Е. В. БУЛГАКОВА, Е. В. КРАСНОВА

ЭКСПЕРТНЫЕ СИСТЕМЫ И МЕТОДЫ ИДЕНТИФИКАЦИИ ДИКТОРА

Проанализированы методы проведения фонографической экспертизы. Рассмотрено используемое в целях идентификации диктора программное обеспечение. Описывается звуковой редактор SIS II.

Ключевые слова: криминалистическая фонографическая экспертиза, идентификация дикторов, программное обеспечение.

Введение. В мировой криминалистической практике в настоящее время фонографическая экспертиза, несомненно, играет важную роль. Однако в России и за рубежом наблюдается различие в подходах к проведению данного вида криминалистической экспертизы, а также в технических средствах для идентификации диктора. Традиция использования различных технических средств и подходы к разработке программного обеспечения также различаются. В отличие от отечественной традиции, за рубежом при проведении фонографической экспертизы решающую роль играют личность эксперта и степень доверия к его профессиональной квалификации. В связи с этим за рубежом эксперты, как правило, применяют программное обеспечение, не предназначенное непосредственно для задачи идентификации диктора. Обычно в качестве таких программ выступают звуковые редакторы с относительно стандартным функционалом. Также некоторыми зарубежными экспертами предлагается использовать при проведении фонографической экспертизы программное обеспечение, разработанное для специалистов в других научных областях (например, в области медицины). В России специально для криминалистических целей разработан ряд программных комплексов, широко используемых правоохранительными органами. Однако функционал этих программ также варьирует, предоставляя экспертам различный набор исследовательских возможностей при проведении фонографической экспертизы [1].

В настоящей работе анализируются методы идентификации диктора, принятые в криминалистической практике, рассматриваются преимущества использования звукового редактора SIS II, разработанного в ООО „ЦРТ“.

Под автоматическим подходом как в отечественной, так и в зарубежной традиции понимается процесс сравнения фонограмм и принятия решения о тождестве или различии звучащих на них голосов без участия эксперта [4].

Аудитивный метод идентификации диктора в отечественной традиции подразумевает, в первую очередь, работу эксперта-фонетиста. За рубежом под аудитивным методом идентификации понимается любая идентификация, не привлекающая дополнительных инструментов. Таким образом, в понятие „аудитивный метод“ в зарубежной традиции входит как прослуши-

вание опытным фонетистом предоставленных спорных фонограмм и их сравнение с фонограммами речи подозреваемых, так и прослушивание фонограмм „наивным“ слушателем.

Акустический подход в отечественной традиции и акустико-фонетический в зарубежной также могут быть соотнесены. Однако в отечественной традиции акустико-фонетический анализ сегментного и суперсегментного уровней может использоваться в рамках лингвистического подхода, в то время как акустический подход в зарубежной традиции предполагает только инструментальный анализ физических параметров звукового сигнала без непосредственной соотнесенности с лингвистической информацией.

Лингвистический подход не получил за рубежом отражения в том виде, в котором он широко применяется в отечественной экспертной практике (с использованием признакового пространства). И, наконец, зарубежный спектрографический метод в отечественной традиции развит не был.

Анализ речевого сигнала с использованием различных технических средств как в России, так и за рубежом, проводится в рамках различных подходов.

В зарубежной криминалистической практике программы, используемые для решения задач идентификации диктора, — это, как правило, звуковые редакторы [5], функционал которых включает возможность визуализации речи (например, Praat, Multi-speech и Wave Surfer). Программа Praat, разработанная в Амстердамском университете, создавалась в первую очередь для научно-исследовательских целей, она предназначена для лингвистов, исследующих звучащую речь. Программа предоставляет возможности для визуализации речевого сигнала, сегментации речевого потока, анализа и синтеза речи. Программа Multi-speech, разработанная компанией KayPentax, — это компьютерная речевая лаборатория, в которой имеются возможности анализа речи, инструменты для визуализации речевого сигнала различными способами. Она предназначена для акустического анализа речевого сигнала, а также анализа устной речи в медицинских целях. Звуковой редактор Wave Surfer, разработанный в Королевском технологическом институте (Стокгольм), активно используется для решения задач, связанных с акустической фонетикой, благодаря широким возможностям визуализации речи, а также возможности считывать и записывать различные форматы файлов транскрипции.

Некоторые зарубежные специалисты в области идентификации диктора предлагают использовать для решения идентификационных задач программы, изначально предназначенные для иных целей. Например, программа Glottex, первоначально разработанная для использования в медицине, позволяет получать информацию о свойствах источника голоса, включая физическую структуру голосовых складок. По мнению авторов статьи [6], эти данные могут быть использованы для судебного сравнения голосов.

Однако ошибочно было бы утверждать, что за рубежом не существует программных комплексов, разработанных специально для целей фонографической экспертизы. В качестве примера таких разработок можно привести программу SAUSI (Semi-Automatic Speaker Identification System), созданную специалистом в области идентификации диктора Г. Холлиеном. В программе анализируются составляемые на основе фонограмм векторы значений, включающие четыре модификации: спектр длительных фрагментов, частота основного тона, временное распределение энергии, характерные особенности формант гласных. На первом этапе четыре приведенных вектора подвергаются самостоятельному анализу, а затем — в сочетаниях друг с другом, на основе чего формируются результаты сравнительного анализа [7].

Одной из систем разработанных непосредственно для идентификации диктора при проведении фонографической экспертизы, является система OTExpert [8], разработанная специалистами ООО „ОТ-Контакт“ (Москва) и используемая в некоторых криминалистических лабораториях Министерства юстиции РФ. Система предназначена для инструментального криминалистического исследования фонограмм. Функционал программы включает возможности визуализации (осциллограмма, спектрограмма, кепстрограмма, КЛП-спектрограмма с прорисовкой

локальных максимумов), сегментации и редактирования звукового сигнала, автоматического статистического анализа параметров основного тона и формант, а также используемого в работе эксперта сравнительного анализа речевых сигналов на разных уровнях.

Разработанный специалистами ООО „Целевые технологии“ [9] комплекс Justiphone представляет собой систему криминалистического исследования фонограмм речи, обеспечивающую шумоочистку фонограмм, установление дословного содержания, выявление идентификационных признаков говорящего, техническое исследование фонограмм на предмет наличия или отсутствия признаков монтажа. Программный комплекс имеет гибкую конфигурацию. Наиболее интересной с точки зрения идентификации диктора является комплектация Justiphone-voice analyzer, которая обеспечивает исследование акустических параметров голоса и звучащей речи. Justiphone-voice analyzer рассчитывает статистические характеристики основного тона и оценивает средневзвешенное относительное отклонение полученных параметров спорной фонограммы от параметров фонограммы-образца.

Автоматизированная система „Диалект“ основана на методике идентификации диктора, разработанной ведущими отечественными специалистами в области фонографической экспертизы [10, 11], включающей акустический и лингвистический методы исследования фонограмм (аудитивный анализ в данном случае является составляющей лингвистического метода). Функционал программы позволяет эксперту обрабатывать речевые сигналы, вычислять акустические и выявлять лингвистические признаки, сравнивать параметры и принимать решения по проводимому исследованию. Эксперт имеет возможность проводить углубленный анализ микроструктуры речевого сигнала и акустических шумов фонограммы на этапах определения пригодности фонограмм для идентификационных исследований и идентификации устной речи неизвестного и подозреваемого.

Итак, как показал анализ, при проведении фонографической экспертизы используется широкий спектр различных программных средств. Однако в описанных выше программах не предусмотрена возможность автоматического принятия идентификационного решения. Эта функция реализована в специализированном звуковом редакторе SIS II v2.0 (далее SIS II), разработанном ООО „ЦРТ“. Наличие модуля обобщенного решения отличает этот редактор от других программ, предназначенных для проведения идентификации диктора.

Редактор SIS II предназначен для анализа речевых сигналов, их шумоочистки и автоматизации выполнения криминалистических экспертиз фонограмм на всех этапах. SIS II обладает уникальными инструментами для идентификационного исследования. Биометрические алгоритмы и экспертные модули позволяют автоматизировать и формализовать многие процессы фонографического исследования, например поиск одинаковых слов и звуков, отбор сравниваемых звуковых и мелодических фрагментов, сравнение дикторов по формантам и основному тону, аудитивный и лингвистический типы анализа. Результаты по каждому методу исследования представляются в виде численных показателей общего идентификационного решения. Рассмотрим модули редактора.

1. Модуль автоматической идентификации (рис. 1) позволяет выполнять сравнение в режиме „один-к-одному“ с использованием следующих алгоритмов: спектрально-формантного, статистики основного тона, смеси гауссовых распределений, обобщенного метода.

Значения вероятности совпадения и различия дикторов рассчитываются не только для каждого из методов, но и для их совокупности. Все результаты сравнения речевых сигналов, получаемые в модуле автоматической идентификации, основаны на выделении в них идентификационно значимых признаков и вычислении меры близости между полученными наборами признаков.

2. Модуль сравнения формант. Процесс идентификации с использованием модуля (рис. 2) может быть разделен на два этапа. Сначала эксперт осуществляет поиск и отбор опорных звуковых фрагментов:

- по диаграмме рассеяния с треугольником Фанта путем выделения области поиска;
- указанием частотных диапазонов поиска формант;
- по установленным горизонтальным маркам с заданием допуска в герцах и процентах;
- путем поиска всех звуков.

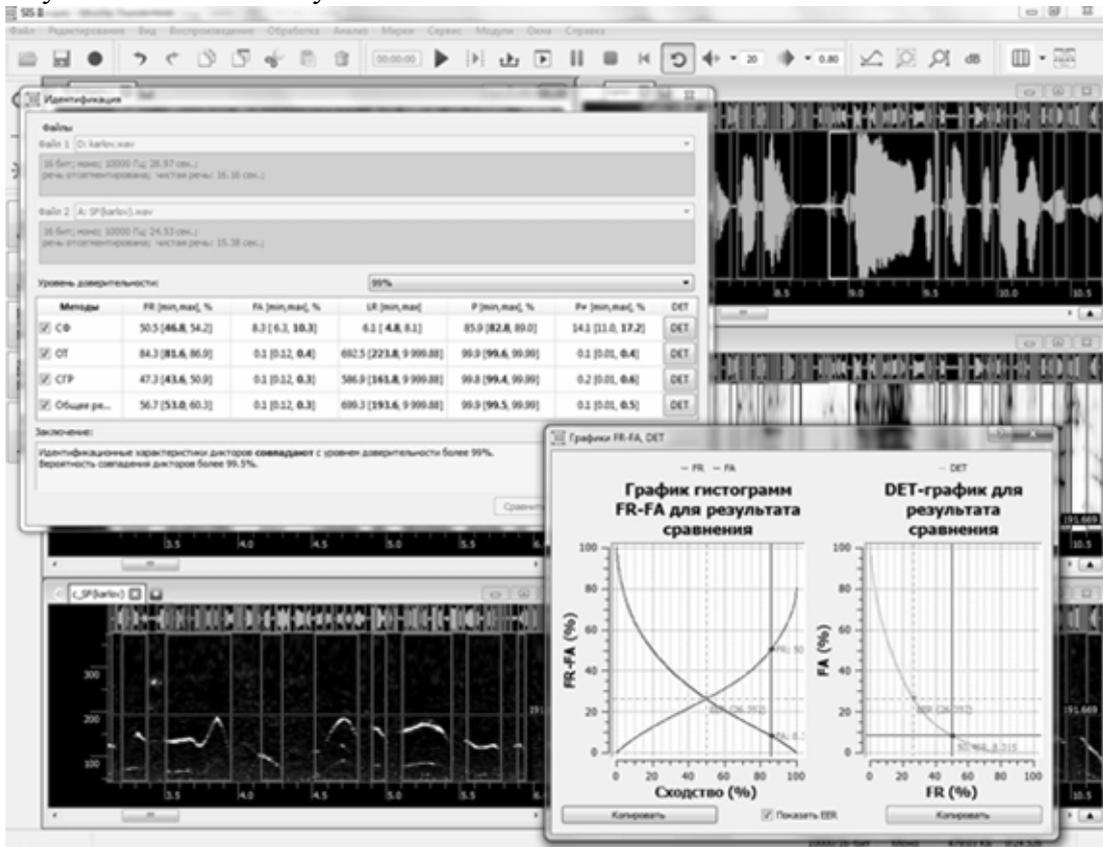


Рис. 1

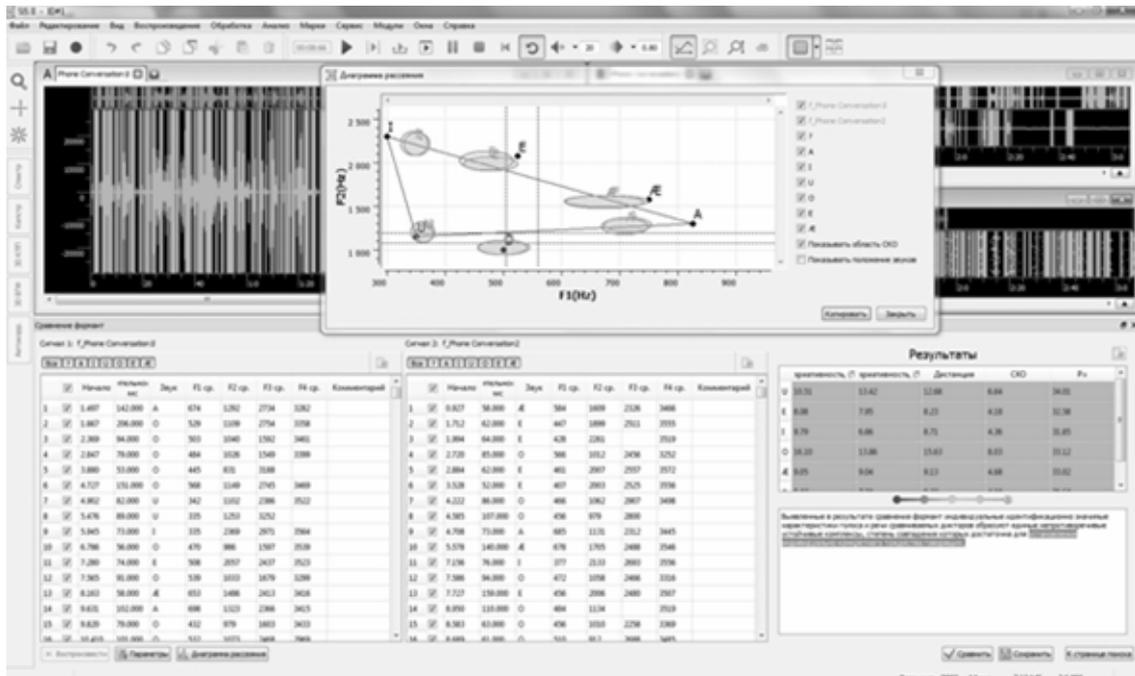


Рис. 2

После того как опорные фрагменты для известного и неизвестного дикторов набраны, эксперт может начать сравнение. Модуль автоматически рассчитывает внутридикторскую и

междикторскую вариативность формантных траекторий для выбранных звуков и принимает решение о положительной/отрицательной идентификации или неопределенном результате.

3. Модуль сравнения основного тона позволяет автоматизировать процесс идентификации дикторов с помощью метода анализа мелодического контура. Метод предназначен для сравнения речевых образцов на основе параметров реализации однотипных элементов структуры мелодического контура. Возможен анализ 18 типов фрагментов контура и 15 параметров их описания, включая значения минимума, максимума, скорости изменения тона, эксцесса, скоса и др. Модуль возвращает результаты сравнения в виде процентного совпадения для каждого из параметров и принимает решение о положительной или отрицательной идентификации или неопределенном результате. Все данные могут экспортироваться в текстовый отчет.

4. Модуль лингвистического анализа позволяет выполнять лингвистическую часть идентификационного исследования и проводить сравнение основного тона на материале русской речи. Он реализован в виде реестра признаков (290 позиций) со звуковыми примерами. Эксперт может отметить три степени выраженности и частотности каждого признака, а также отметить пример его реализации признака на фонограмме. После заполнения таблицы признаков для двух дикторов в модуле лингвистического анализа будет представлена общая статистика по несовпавшим, полностью совпавшим и частично совпавшим признакам.

5. Модуль аудитивного анализа позволяет проводить идентификацию по аудитивным признакам голоса и речи дикторов. Модуль реализован в виде списка из 126 признаков с примерами их звучания. Для каждого признака эксперт может указать три степени его выраженности и частотности, а также отметить пример реализации признака на фонограмме. После заполнения таблицы признаков для двух дикторов в модуле аудитивного анализа можно увидеть общую статистику по несовпавшим, полностью совпавшим и частично совпавшим признакам.

6. Модуль обобщенного решения. Основываясь на результатах работы каждого из модулей, программа позволяет автоматически принять единое идентификационное решение [12]. При этом пользователь имеет возможность изменять границы применимости каждого из отдельных методов исследования, изменяя таким образом его весовой коэффициент в общем решении.

Важными преимуществами специализированного звукового редактора SIS II являются автоматическая оценка качества фонограмм, высокое качество вычисления спектрограмм и автоматического выделения формант. Благодаря уникальному набору функций, обеспечивающих высокую эффективность работы эксперта и расширяющих его возможности при проведении идентификации диктора, SIS II широко используется в отечественных и зарубежных экспертных лабораториях.

Заключение. Анализ имеющихся решений показал, что только в специализированном звуковом редакторе SIS II имеется модуль обобщенного решения, который позволяет автоматизированно с использованием данных, вводимых экспертом, ответить на ключевой вопрос фонографической экспертизы — тождественны ли голоса на сравниваемых фонограммах.

Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

СПИСОК ЛИТЕРАТУРЫ

1. Хитров М. В. и др. Фоноскопическое исследование фонограмм речи: Исследование достоверности фонограмм. Кн. I. СПб: Юридический центр-Пресс, 2011.
2. Добрякова М. В. Зависимость между вербальной спецификой эксперта и надежностью идентификации иноязычного говорящего: Автореф. дис. ... канд. филол. наук. М., 2003.

3. Kersta L. G. Voiceprint Identification // Nature. 1962. Vol. 196, N 4861. P. 1253—1257.
4. Матвеев Ю. Н. Технологии биометрической идентификации личности по голосу и другим модальностям // Вестник МГТУ им. Н. Э. Баумана. Сер. „Приборостроение“. 2012. № 3(3). С. 46—61.
5. Harrison P. Variability of Formant Measurements. Submitted in partial fulfilment of the degree of MA. University of York, UK, 2004.
6. Enzinger E., Zhang C., Morrison G. S. Voice source features for forensic voice comparison — an evaluation of the Glottex software package // Proc. of Odyssey. Singapore, 2012. P. 78—85.
7. Hollien H. Forensic voice identification. NY, 2002.
8. Программный комплекс „ОТExpert“ [Электронный ресурс]: <http://фоноскопическая.рф/forensic_audio-video/forensic_tools/otexpert/>.
9. Каганов А. Ш. Криминалистическая экспертиза звукозаписей. М.: Юрлитинформ, 2005.
10. Идентификация лиц по фонограммам русской речи на автоматизированной системе „Диалект“: Пособие для экспертов / Под ред. А. В. Фесенко. М., 1996.
11. Тимофеев И. Н. Применение автоматизированной системы „Диалект“ на базе компьютерной речевой лаборатории CLS (США) при решении задач идентификации дикторов: Метод. рекомендации. М.: ЭКЦ МВД России, 2000.
12. Матвеев Ю. Н. Оценка доверительного интервала общего решения ансамбля классификаторов // Изв. вузов. Приборостроение. 2013. Т. 56, № 2. С. 74—79.

Сведения об авторах

Елена Владимировна Булгакова

— аспирант; Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кафедра речевых информационных систем;
E-mail: bulgakova@speechpro.com

Екатерина Викторовна Краснова

— ООО „ЦРТ“, Санкт-Петербург; научный сотрудник;
E-mail: krasnova@speechpro.com

Рекомендована кафедрой
речевых информационных систем

Поступила в редакцию
22.10.13 г.

УДК 004.93+57.087.1

Д. В. ДЫРМОВСКИЙ, С. Л. КОВАЛЬ, М. В. ХИТРОВ

КОНЦЕПЦИЯ СИСТЕМЫ НАЦИОНАЛЬНОГО ФОНОУЧЕТА И ГОЛОСОВОГО БИОМЕТРИЧЕСКОГО ПОИСКА

Представлена концепция системы национального фоноучета и голосового биометрического поиска.

Ключевые слова: фоноучет, голосовой биометрический поиск, инфраструктура системы.

Введение. Многие государства в настоящее время столкнулись с целым рядом задач, связанных с отражением угроз нелегальной эмиграции и контрабанды, предотвращением террористических атак и применением новых подходов при создании систем национальной и международной безопасности. Особую важность в связи с этим приобретает совершенствование систем учета, поиска, выявления и идентификации лиц, нарушивших закон. Решение подобных задач в глобальном масштабе требует применения передовых технологических решений [1].