

СИСТЕМА АВТОМАТИЗИРОВАННОГО ПОИСКА ГОЛОСОВ

О.В. Рыбальский¹, В.И. Соловьев², А.Н. Шабля³, В.В. Журавель⁴

¹Национальная академия внутренних дел,
пл. Соломянская, 1, Киев, 03056, Украина; e-mail: rybalsky_ol@mail.ru

²Национальный университет им. В. Даля,
ул. Советская, 51, Северодонецк, 93400, Украина; e-mail: mettilyd@mail.ru

³Одесский научно-исследовательский институт судебных экспертиз МЮ Украины,
ул. Ланжероновская, 21, Одесса, 65044, Украина; e-mail: alik.shablya@gmail.cjm

⁴Государственный научно-исследовательский экспертно-криминалистический центр МВД Украины,
ул. Б. Окружная, 4а, Киев, 03056, Украина; e-mail: fonoscopia@ukr.net

Рассмотрен отдельный вид экспертных материалов цифровой звукозаписи - фонограммы, содержащие сообщения, поступающие в дежурные части ОВД Украины, среди которых могут быть ложные. Для указанных экспертных материалов решаются задачи идентификации по параметрам сигналов речи личности, подозреваемой в предоставлении ложного сообщения, в случае ее задержания или наличия образцов ее голоса, полученных оперативным путем, а также выявления в базе данных голосов фигурантов с близкими параметрами сигналов устной речи для установления потенциальных подозреваемых в предоставлении ложной информации, для чего разработана автоматизированная система быстрого поиска подозреваемых.

Ключевые слова: база данных, голос, идентификация, фигурант, цифровая фонограмма

Вступление

Материалы звукозаписи составляют значительную часть доказательств, используемых при расследовании и рассмотрении в суде дел, связанных с коррупцией, взяточничеством, вымогательством, рэкетом, похищениями людей и тому подобное [1-3].

Но существуют особенные цифровые фонограммы, содержащие сообщения, поступающие в дежурные части ОВД Украины. Они образуют отдельный вид экспертных материалов цифровой звукозаписи. Это объясняется тем, что при их экспертизе всегда используют цифровую копию фонограммы и не проводят экспертизу оригинальности и подлинности фонограмм. Еще одна особенность таких фонограмм – малая длительность записываемых сообщений. Принимая во внимание, что в настоящее время значительную часть таких фонограмм составляют ложные сообщения относительно ожидаемых терактов, перед экспертными подразделениями МВД Украины предстала проблема помощи в розыске и следствии выявлением подозреваемых в таких правонарушениях экспертным путем.

Основными задачами такой экспертизы являются:

- идентификация по параметрам сигналов речи личности, подозреваемой в предоставлении ложного сообщения, в случае ее задержания или наличия образцов ее голоса, полученных оперативным путем;
- выявление в базе данных голосов фигурантов с близкими параметрами сигналов устной речи для установления потенциальных подозреваемых в предоставлении ложной информации.

Успешное решение этих задач может быть обеспечено соответствующими экспертными подразделениями Государственного научно-исследовательского экспертно-криминалистического центра (ГНИЭКЦ) МВД Украины. Для облегчения их выполнения требовалась разработка автоматизированной системы быстрого поиска подозреваемых в имеющейся базе голосов для последующей идентификации личности по физическим параметрам сигналов ее устной речи.

В настоящее время такая система разработана. Она была подвергнута испытаниям и предоставлена на апробацию в ГНИЭКЦ и Одесский НИИ судебных экспертиз.

Основная часть

В системе предусматривается автоматическое разделение голосов дежурного работника ОВД и заявителя. Такая функция во многих случаях исключает необходимость предыдущей подготовки экспертом записей, поступающих или уже зафиксированных в базе данных. Однако ее применение не обеспечивает гарантированного разделения голосов в фонограммах, записанных на одноканальной аппаратуре. Поэтому введение этой опции в программу не исключает необходимости технически и методически предусмотреть возможность проведения предыдущей подготовки записей, зафиксированных в базе данных. Такой подготовкой является разделение голосов экспертом, и такая опция также предусмотрена в программе.

Также следует отметить, что система является поисковой, поэтому результаты поиска отнюдь не тождественны идентификации личности по голосу, поскольку являются лишь результатом ранжирования по степени близости отдельных параметров голосовых сигналов. Также следует отметить, что сравнения физических параметров голосов всегда носят вероятностный характер, который определяется величиной вероятностей ошибок I и II рода. В системе использован метод ранжирования по величине ошибки I рода. Проведенные эксперименты с использованием первой версии программы AVATAR показали, что при поиске разных образцов голоса одного человека, записанного на разных фонограммах, величина вероятности ошибки I рода не превышает значения 0,08. Результаты поиска предоставляются в виде ранговой таблицы. В процессе ввода фонограмм в базу данных производится расчет параметров сигналов зафиксированных на них голосов.

Первоначально расчет параметров голосов и ранжирование проводилось только по частоте основного тона. Но при расчете ранговой таблицы весьма часто случалось так, что хотя на ее первой позиции размещалась фонограмма с записью голоса диктора, голос которого записан и на образцовой фонограмме, однако на последующих позициях находились фонограммы с голосами других дикторов. И это несмотря на то, что в базе было несколько фонограмм с записью голоса искомого диктора. Но при этом во всех экспериментах на первую позицию, как правило, приходился голос диктора, который записан и на образцовой фонограмме. Это квалифицировалось авторами, как существенный недостаток программы.

Была создана программа, использующая ранжирование не только по частоте основного тона, но по дополнительным признакам, имеющим индивидуальный характер. В результате отработки предлагаемого метода были отобраны три метода ранжирования по признакам, содержащимся в голосовых сигналах, пригодные для построения окончательного варианта системы. К ним относятся:

- вычисление близости кривых функций двумерной плотности вероятности для частоты основного тона и расположению в спектре семи формант, выделяемых из речи, зафиксированной на фонограмме;
- вычисление близости кривых функций плотности вероятности для каждого из этих признаков отдельно;

▪ вычисление меры близости абсолютных максимумов спектров формант, выделяемых из речи, зафиксированной на фонограмме.

Особенностью разработанного метода является спектральный анализ фрагментов речи на малых временных интервалах (20-30 мс). Для этого используется вейвлет-преобразование, в частности, используется вейвлет Морле. Фрагменты речи сканируются вейвлетом Морле с минимально возможным смещением окна сканирования. Величина этого смещения обратна частоте дискретизации. В каждом временном окне вычисляется преобразование на основе вейвлета Морле. Определяется частота основного тона – F_b как корреляционная функция максимумов вейвлет преобразования в каждом окне. Важно отметить, что частота основного тона F_b определяется для каждого временного окна.

Таким образом, для фрагмента речи длительностью, например, 1 сек при частоте дискретизации 11025 Гц и размере окна анализа 20 мс, массив частот основного тона составляет величину порядка десять тысяч отсчетов, что весьма существенно при статистических оценках.

Аналогично, в каждом временном окне рассчитывается и спектр. Особенностью применяемого метода расчета спектра является то, что расчет спектральных характеристик проводится с разрешающей способностью 1 Гц (неортогональные преобразования в малом временном окне, для чего использована теория каркасов [4]). В спектре каждого временного окна выделяются максимумы первых семи формант.

Для вычисления близости кривых функций двумерной плотности вероятности рассчитывается функция F_f – функция характеристик голоса. Эта функция является нелинейной эмпирической функцией зависимостей амплитуды и частоты максимумов формант.

Диапазон частот основного тона F_b разбивается на интервалы в диапазоне от 80 до 600 Гц. Аналогично на частотные интервалы в диапазоне от 80 до 2500 Гц разбиваются значения F_f .

Для всех фрагментов речи определяется двумерная плотность вероятностей по двум координатам – частоте основного тона F_b и функции характеристик голоса F_f (см. рис. 1). Таким образом, в основу системы заложен комбинированный метод оценки на основе частоты основного тона и спектральных характеристик. Степень близости двух характеристик голоса определяется абсолютными разностями между двумерными плотностями вероятностей.

Разумеется, что проекции двумерных плотностей вероятностей на каждую из одномерных осей координат дает распределение частоты основного тона F_b и распределение функции характеристик голоса F_f , которые также можно представить в виде отдельных зависимостей. Это позволяет произвести вычисление близости кривых функций плотности вероятности для каждого из этих признаков отдельно.

Точность оценки частоты основного тона в вариантах фрагментов речи с явно выраженными участками частоты основного тона определяется произведением $D_t \cdot F_t$, где D_t – период дискретизации речевого файла.

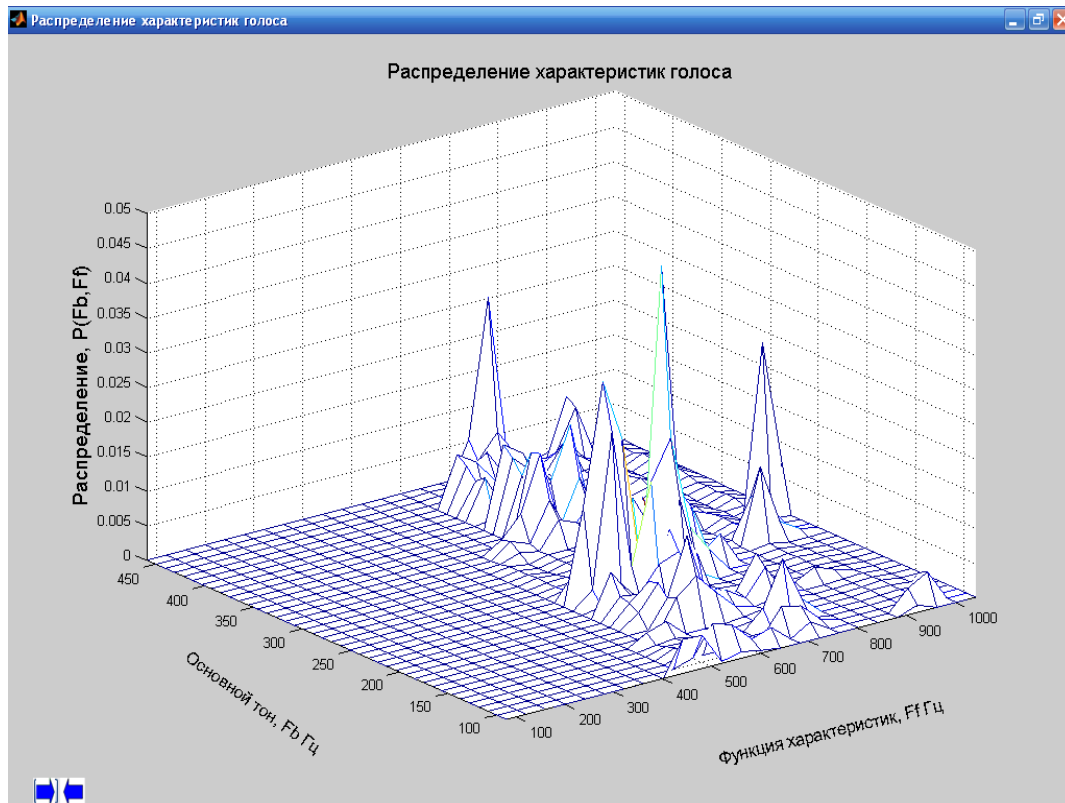


Рис. 1. Двумерная функция плотности вероятности характеристик голоса

Для ранжирования по плотностям распределения основного тона F_f и функциям характеристик голоса F_f в качестве меры близости используется величина, эквивалентная критерию согласия Колмогорова:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n IX_i \leq x, \quad (1)$$

где $IX_i \leq x$ показывает, попала ли величина X_i в область $(-\infty, x]$ при условии, что

$$IX_i \leq x = \begin{cases} 1, & X_i \leq x; \\ 0, & X_i > x; \end{cases} \quad (2)$$

Для вычисления меры близости абсолютных максимумов спектров формант, выделяемых из речи, зафиксированной на фонограмме, используется критерий

$$S_i = \text{abs}\{\max [P_1(x_1, x_2) - \max P_2(x_1, x_2)]\}, \quad (3)$$

где $P_1(x_1, x_2), P_2(x_1, x_2)$ – двумерные плотности вероятностей для двух разных записей с аргументами, как частоты основного тона, так и характеристик голоса, сравниваемых раздельно.

Результаты поиска голосов с близкими характеристиками представляются в виде ранговых таблиц. Пример таблицы для ранжирования по плотностям вероятностей основного тона F_f и функциям характеристик голоса F_f показан на рис. 2.

Следует отметить, что все вычисления характеристик голоса осуществляются для фрагментов записей, содержащих речевые сигналы. Разделение на паузы и фрагменты речи в системе осуществляется автоматически на основе двух критериев – уровня нормируемого аудиосигнала и наличия фрагмента записи, содержащего звуки речи. Система выделения звуков, является вспомогательной системой, встроенной в систему.

Она осуществляет общее разделение фрагментов, в которых присутствуют звуки речи, на основе специального метода, близкого к методологии скрытых марковских цепей и не является предметом рассмотрения данной статьи.

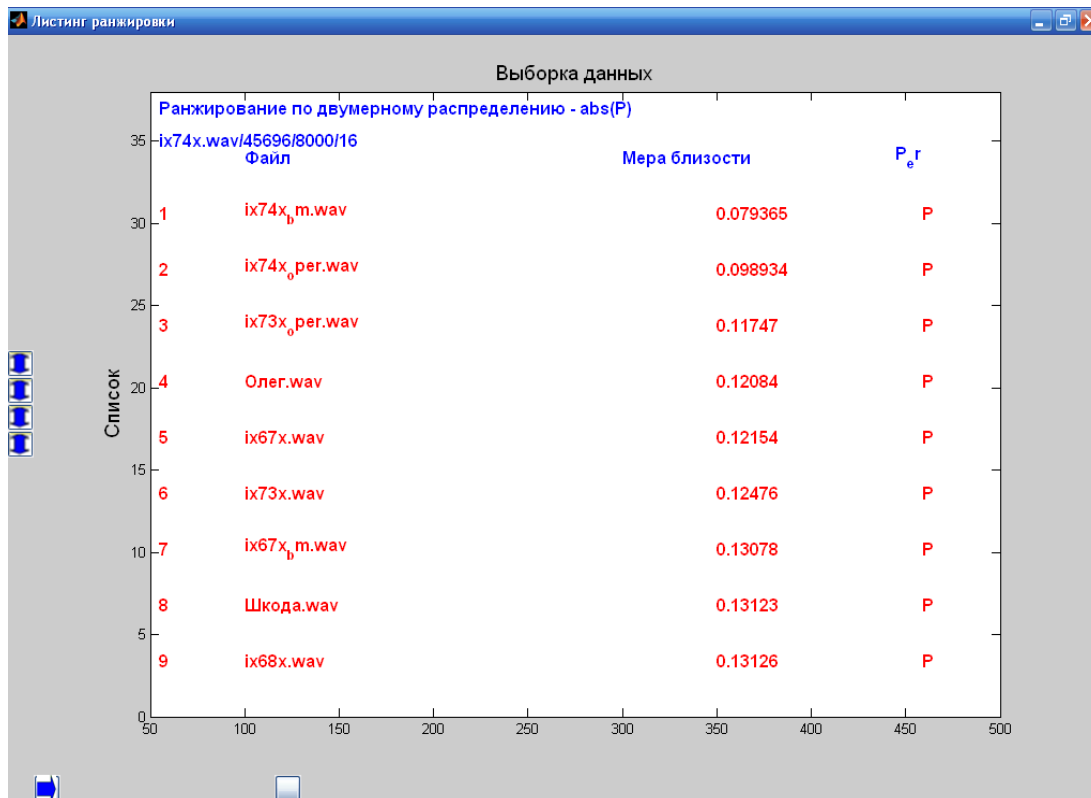


Рис. 2. Пример ранжирования характеристик голоса по двумерному распределению

Выводы

Создана система, обеспечивающая автоматический поиск по базе данных личностей с голосами, наиболее близкими к голосу фигуранта, записанного на образцовой фонограмме, что призвало повысить эффективность работы экспертных подразделений МВД Украины.

Список литературы

1. Кобозева, А.А. Матричный анализ - основа общего подхода к обнаружению фальсификации цифрового сигнала / А.А. Кобозева, О.В. Рыбальский, Е.А. Трифонова // Вісник Східноукраїнського національного університету ім. Володимира Даля. – 2008. – №8(126), ч.1. – С. 62-72.
2. Рыбальский, О.В. Комплексный подход к экспертизе подлинности материалов цифровой звукозаписи / О.В. Рыбальский, А.А. Кобозева, И.А. Струк, Е.А. Трифонова // Вісник Східноукраїнського національного університету ім. Володимира Даля. – 2009. – №6(136), ч.1. – С. 75-79.
3. Кобозева, А.А. Разработка общей теории выявления следов цифровой обработки сигналов и ее реализация аппаратно-программным комплексом «Теорема-М» / А.А. Кобозева, О.В. Рыбальский, В.И. Соловьев // Сучасна спеціальна техніка. – 2010. – №1(20). – С.5-14.
4. Малла, С. Вэйвлеты в обработке сигналов: Пер. с англ. / С. Мала. – М.: Мир, 2005. – 671 с.

СИСТЕМА АВТОМАТИЗОВАНОГО ПОШУКУ ГОЛОСІВО.В. Рибальський¹, В.І. Соловйов², А.Н. Шабля³, В.В. Журавель⁴¹Національна академія внутрішніх справ,
пл. Солом'янська, 1, Київ, 03056, Україна; e-mail: rybalsky_ol@mail.ru²Національний університет ім. В. Даля,
вул. Радянська, 51, Сєверодонецьк, 93400, Україна; e-mail: mettilyd@mail.ru³Одеський науково-дослідний інститут судових експертиз МЮ України,

вул. Ланжеронівська, 21, Одеса, 65044, Україна; e-mail: alik.shablya@gmail.cjm

⁴Державний науково-дослідний експертно-криміналістичний центр МВС України,
вул. Б. Окружна, 4а, Київ, 03056, Україна; e-mail: fonoscopia@ukr.net

Розглянуто окремих вид експертних матеріалів цифрового звукозапису - фонограми, що містять повідомлення, які надходять в чергові частини ОВС України, серед яких можуть бути хибні. Для зазначених експертних матеріалів вирішуються задачі ідентифікації за параметрами сигналів мови особистості, підозрюваної у наданні неправдивого повідомлення, у разі її затримання або наявності зразків її голосу, отриманих оперативним шляхом, а також виявлення в базі даних голосів фігурантів з близькими параметрами сигналів усного мовлення для встановлення потенційних підозрюваних у наданні неправдивої інформації, для чого розроблена автоматизована система швидкого пошуку підозрюваних.

Ключові слова: база даних, голос, ідентифікація, фігурант, цифрова фонограма

AUTOMATED SYSTEM VOICE SEARCHO.V. Rybalskiy¹, V.I. Solovyov², A.N. Chablis³, V.V. Zhuravel⁴¹National Academy of Internal Affairs,
1, Solomenskaya Sq., Kiev, 03056, Ukraine; e-mail: rybalsky_ol@mail.ru²National University. Dal,
51, Soviet st., Severodonetsk, 93400, Ukraine; e-mail: mettilyd@mail.ru³Odessa scientific research institute of forensic examinations Justice of Ukraine,
21, Lanzheronovskaya st., Odessa, 65044, Ukraine; e-mail: alik.shablya@gmail.cjm⁴State Research Forensic Center Interior Ministry of Ukraine,
4a, B. District st., Kiev, 03056, Ukraine; e-mail: fonoscopia@ukr.net

Is considered a separate type of digital audio expert materials - soundtrack containing messages received in the duty of MIA of Ukraine, among which may be false. For these expert materials solved the problem of identification of parameters of speech signals of the individual suspected of providing false messages in the event of detention or the availability of samples of its votes obtained by surgery and to identify the database votes defendants with similar parameters of signals of speech to determine potential suspected of providing false information, which developed the automated system of quick search of suspects.

Keywords: database, voice authentication, suspected, digital phonogram