

ГОЛОСОВАЯ ИДЕНТИФИКАЦИЯ В УСЛОВИЯХ ВНЕШНЕГО ШУМА

П.В. Малинин, В.В. Поляков
Алтайский государственный университет
г. Барнаул

В работе рассмотрен подход к голосовой идентификации диктора, основанный на использовании проекционных методов многомерного анализа данных. В качестве информативных признаков для математической обработки голосовых записей использованы мел-частотные кепстральные коэффициенты. Рассмотрено влияние внешнего акустического шума на результат идентификации. Показана эффективность предложенного подхода в условиях высокого уровня внешнего шума.

Ключевые слова: голосовая идентификация, акустический шум, мел-частотные кепстральные коэффициенты, метод главных компонент.

Одним из факторов, призванных повысить информационную безопасность современных автоматизированных систем, является разработка новых подходов к предотвращению угрозы доступа к информационным ресурсам посторонних лиц [1]. Это достигается за счет создания новых систем и способов идентификации личности [2]. Одним из таких способов является идентификация по голосовым данным диктора, находящая все большее распространение. Суть голосовой идентификации заключается в выявлении индивидуальных акустических характеристик и последующем сравнении голосовых записей, предъявляемых при доступе к информационному ресурсу, со специализированной базой данных [3, 4].

Особое место голосовая идентификация личности занимает при расследовании преступлений в сфере компьютерной информации, прежде всего при формировании доказательной базы таких расследований. Так, результаты идентификации используются при проведении фоноскопических экспертиз, при поиске подозреваемых по голосовым записям, при проведении экспертного криминалистического исследования на основе теории криминалистической идентификации и т.д. [5, 6].

Идентификация личности по голосовым записям, полученным в реальных условиях, встречается с серьезными трудностями. Прежде всего это связано с наложением на запись акустических сигналов внешних механических шумов, которые неизбежно подавляют и искажают индивидуальные голосовые

характеристики. В силу этого системы идентификации при анализе реальной речевой записи с накладывающимися внешними шумами показывают значительно более низкие надежность и эффективность, чем в идеализированных условиях.

В настоящей работе рассматривается зависимость результатов голосовой идентификации личности от величины внешнего шума.

Для успешного проведения идентификации необходим адекватный выбор исходных информативных признаков, характеризующих голосовой сигнал. В качестве признаков, описывавших индивидуальные характеристики голоса диктора, в работе использовались мел-частотные кепстральные коэффициенты. Их достоинством являлась независимость получаемых информативных характеристик от длины исходного голосового сигнала, выделения особенностей этого сигнала, связанных с дальними гармониками, снижение уменьшения вклада от случайных искажений, а учет нелинейных свойств человеческого слуха [7-9].

Для кластеризации и последующей классификации признаков применялся аппарат проекционных методов многомерного анализа данных, прежде всего метод главных компонент [10 - 13]. Именно, по вычисленным, согласно [14], коэффициентам мел-кепстрального разложения производился расчет главных компонент ГК1, ГК2, ГК3 и т.д. с выбором наиболее информативных пар компонент. Наглядная интерпретация получаемых результатов проводилась с помощью

графиков проекций многомерных голосовых данных на плоскости главных компонент. Отдельная точка на таком графике соответствовала конкретной голосовой записи, близкое расположение точек свидетельствовало о сходстве акустических характеристик соответствующих записей. Вся многомерная совокупность записей, относящихся к одному диктору, на плоскости имела вид эллипса, строившегося по максимальному разбросу точек. Принадлежность тому или иному диктору неизвестных голосовых данных определялась по попаданию точки, описывавшей эти данные, в соответствующий эллипс.

В качестве акустического шума, накладывавшегося на исходные голосовые данные, использовались специально сделанные записи шума улицы. Этот шум записывался в реальных условиях оживленных автомагистралей. Уровень шума задавался с помощью коэффициента

$$R_{SN} = 20 \log_{10} \left(\frac{A_S}{A_N} \right),$$

где A_S – среднеквадратичное значение амплитуды сигнала, A_N – среднеквадратичное значение амплитуды шума.

На рисунке 1 в качестве примера изображено представление голосовых записей в главных компонентах в условиях различного уровня шума улицы, величина R_{SN} задавалась в пределах от 15 дБ до -15 дБ. Эллипсом отмечены голосовые записи конкретного диктора (точки «д»), залитые точки изображают записи с наложенным шумом. Как следует из рис. 1, расположение точек 7-9 за пределами эллипса, характеризующего исходные голосовые сигналы, свидетельствует о невозможности проведения идентификации при таком уровне шума, при меньшем уровне (точки 1-6) идентификация возможна.

Важной задачей является расширение возможностей голосовой идентификации для наиболее неблагоприятных случаев, отличающихся значительными значениями R_{SN} . Для выяснения надежности и достоверности предложенного подхода и определения границ его применимости при наложении шума были проведены расчеты ошибок идентификации 1-го и 2-го рода. При этом под ошибками 1-го рода понимались ложно отвергнутые голосовые записи, а под ошибками 2-го рода – ложно принятые сигналы. Отметим, что для практических целей наиболее опасными являются ошибки 2-го рода, то есть допуск к защищаемой информации лиц, не имеющих на это право. Итоговое значение ошибки

идентификации находилось как среднее значение ошибок 1-го и 2-го рода.

На рисунке 2 в виде кривой 1 приведены зависимости числа правильно идентифицированных дикторов от величины отношения «полезный сигнал/шум». Для сопоставления на рисунке 2 представлены также данные, полученные с помощью иного подхода – по методу гауссовых смесей из работ [15] (кривая 2), [16] (кривая 3), [17] (кривая 4). Видно, что все кривые ведут себя качественно одинаковым образом, именно, доля правильно идентифицируемых дикторов быстро падает с увеличением уровня шума. В то же время из рисунка 2 следует, что предложенный в настоящей работе метод обеспечивает более медленное ухудшение результатов идентификации при росте шума.

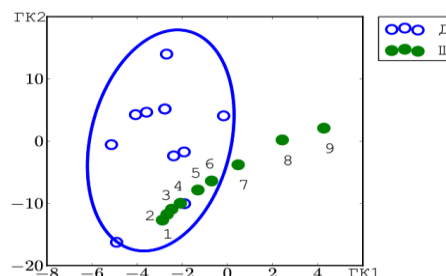


Рисунок 1 – Голосовые записи в главных компонентах при наложении шума: д– данные без шума, ш– данные с наложенным шумом R_{SN} : 1 – 15 дБ, 2 – 10 дБ, 3 – 7 дБ, 4 – 5 дБ, 5 – 0 дБ, 6 – -5 дБ, 7 – -7 дБ, 8 – -10 дБ, 9 – -15 дБ

Полученные результаты могут быть использованы при разработке и тестировании технических устройств защиты информации от утечек по акустическому каналу, а также при создании новых систем голосовой идентификации, устойчивых к влиянию внешнего акустического сигнала.

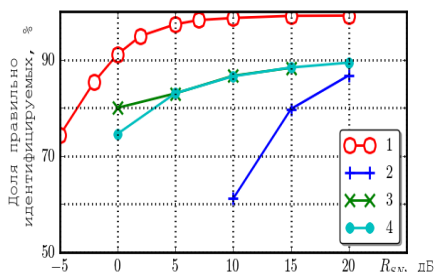


Рисунок 2 – Доля правильно идентифицируемых дикторов при наложении шума: 1 – настоящая работа, 2 – [15], 3 – [16], 4 – [17]

Публикация подготовлена в рамках поддержанного РГНФ научного проекта №16-33-01160.

СПИСОК ЛИТЕРАТУРЫ

1. Мазуров В.А., Головин А.В., Поляков В.В. Информационная безопасность: основы правовой и технической защиты информации // Барнаул: Изд-во Алт. ун-та, 2005. 196 с.
2. Дшхунян В.Л., Шаньгин В.Ф. Электронная идентификация. Бесконтактные электронные идентификаторы и смарт-карты // М.: ООО «Издательство АСТ», Издательство «НТ Пресс», 2004. – 695 с.
3. Дворянкин С.В. О необходимости новых подходов к оценке эффективности технических средств акустозащиты // Информация и безопасность. – 2002. – №2. – С. 244–245.
4. Ролдугин С.В., Голубинский А.Н., Вольская Т.А. Модели речевых сигналов для идентификации личности по голосу // Радиотехника. – 2002. – № 11. – С. 79–81.
5. Андреев Б.В., Пак П.Н., Хорст В.Н. Расследование преступлений в сфере компьютерной информации // М.: Юрлитинформ, 2001. – 152 с.
6. Каганов А.Ш. Криминалистическая идентификация личности по голосу и звучащей речи // М.: Юрлитинформ, 2009. – 291 с.
7. Ахмад Х.М. Сравнительное исследование эффективности различных методов кепстрального описания речевых сигналов в задачах распознавания // Вестник ТГТУ. – 2007. – Том 13, № 4. – С. 887–891.
8. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу // М.: Радио и связь, 1981. – 224 с.
9. Григорян Р.Л., Коршунов С.С., С.А. Репалов С.С., Григорян М.Ю., Хрящев Р.Л. Сравнение различных способов оценки схожести распределений частоты основного тона в задаче идентификации диктора по его речи // Речевые технологии. – 2010. – № 3. – С. 35–38.
10. Малинин П.В., Поляков В.В. Иерархический

- подход в задаче идентификации личности по голосу с помощью проекционных методов классификации многомерных данных // Доклады ТУСУР. – 2010. – № 1. Ч.1. – С. 128–130.
11. Малинин П.В., Поляков В.В. Применение методов анализа многомерных данных к задаче идентификации личности по голосу // Известия АлтГУ. – 2010. – № 1. – С. 140–142.
 12. Егоров А.В., Лепендин А.А., Поляков В.В. Методы обработки экспериментальных данных при акустической и электромагнитной диагностике: монография. Барнаул: Изд-во Алт. ун-та, 2012. – 66 с.
 13. Polyakov V.V., Egorov A.V., Pirogov A.A., and Kolubaev E.A. Application of Projection Methods of Multivariate Data Analysis in Eddy Current Testing of Materials. AIP Conference Proceedings. AIP Conference Proceedings. 2015. Vol. 1683. Pp. 0201851-0201854. Doi:10.1063/1.4932875.
 14. Заковряшин А.С. Малинин П.В., Лепендин А.А. Применение распределений мел-частотных кепстральных коэффициентов для голосовой идентификации личности // Известия АлтГУ. – 2014. – №1/1. – С. 156–160.
 15. Reynolds D.A., Glass J.R., Hazen T.J., Ming J. Robust Speaker Recognition in Noisy Conditions // Audio, Speech, and Language Processing. July – 2007. – vol. 15 (5). – pp. 1711–1723.
 16. Crookes D., Ming J., McLaughlin N. Speaker recognition in noisy conditions with limited training data // 19th European Signal Processing Conference (EUSIPCO 2011): Book of abstracts. – Barcelona, 2011. – pp. 1294–1298.
 17. Graciarena M., Kajarekar S., Stolcke A., Shriberg E. Noise Robust Speaker Identification for Spontaneous Speech // ICASSP 2007. IEEE International Conference – 2007. – pp. 245–248.

Малинин Петр Владимирович – к.т.н., доцент, тел.: (3852) 364809, e-mail: phys.asu.ru;
Поляков Виктор Владимирович – д.ф.-м.н., профессор, заведующий кафедрой прикладной физики, электроники и информационной безопасности, тел.: (3852) 367059, e-mail: pvv@asu.ru.