

Формантная разборчивость как универсальный показатель информативности речевой подписи и биоподписи документа

А. М. Алюшин

Национальный исследовательский ядерный университет «МИФИ»;

ФГБОУ ВО «РЭУ им. Г.В. Плеханова»

Alyshin@list.ru

Аннотация. Речевая подпись (РП) и биоподпись (БП) выделены в качестве перспективных цифровых технологий для защиты важных финансовых, юридических и распорядительных документов от подделки и фальсификации. Обоснована актуальность использования универсальной метрики для оценки качества передаваемой в составе РП и БП информации о биопараметрах автора документа. В работе обоснована целесообразность использования в качестве такой метрики формантной разборчивости (ФР) речевой информации, передаваемой в составе РП и БП. Показана применимость ФР для оценки качества содержащихся в РП и БП биопараметров автора документа, ответственных за передачу контекстного содержания важных фрагментов документа, идентификацию его личности, а также оценку его функционального и психоэмоционального состояния. Расчетный метод определения ФР выбран в качестве основного на этапах разработки и предварительных испытаний образцов новой техники. Определены области возможного использования различных графических форматов представления изображения РП и БП в зависимости от уровня износа бумажного, либо пластикового носителя документа.

Ключевые слова: защита документов; биоподпись; речевая подпись; надежность передачи информации

I. ВВЕДЕНИЕ

В настоящее время активно разрабатываются биометрические технологии защиты информации, а также правовые аспекты их применения на практике [1]. Традиционные методы защиты документов не позволяют обеспечить в полной мере требуемый сегодня уровень защиты [2]. Одними из перспективных методов защиты важных финансовых, юридических и распорядительных документов от подделки и фальсификации являются технологии их маркирования на основе речевой подписи (РП) [3, 4], либо биоподписи (БП) [5].

Такие маркеры представляют из себя фрагменты графических изображений, например, спектрограмм [1–3], позволяющих в сжатом виде передавать информацию о биопараметрах автора документа, а также данные о контекстном содержании его важных разделов. РП и БП, встроенные в документ, позволяют использовать стандартное полиграфическое оборудование для копирования, сканирования и размножения документов.

РП и БП дают возможность осуществить идентификацию автора документа по голосу [6]. При этом БП позволяет дополнительно оценить

функциональное и психоэмоциональное состояние [7–10] (ФПЭС) автора документа, что дает возможность выявлять случаи получения подписи под принуждением, либо при неадекватном состоянии автора.

Износ и старение носителя документа могут приводить к потере информации, содержащейся в таких маркерах. Существующие графические форматы кодирования изображения РП, либо БП обеспечивают различный уровень надежности сохранения информации в маркерах. К сожалению, в настоящее время отсутствуют единые показатели, позволяющие одновременно охарактеризовать качество передачи контекстного содержания РП и БП, идентификационных признаков автора документа, а также биопараметров, определяющих его ФПЭС. По этой причине разработка и обоснование единой метрики, позволяющей оценить качество передаваемой в составе РП и БП комплексной биометрической и контекстной информации, являются актуальными.

Целью исследования является разработка и обоснование единой метрики для оценки качества, передаваемой в составе РП и БП, комплексной информации о документе и его авторе.

II. РАЗРАБОТКА И ОБОСНОВАНИЕ МЕТРИКИ

В работе анализировались следующие аспекты информативности РП и БП, обуславливающие:

- разборчивость речевого сигнала (РС), как показатель качества и достоверности передачи контекстного содержания важных разделов документа;
- качество передачи индивидуальных особенностей голоса автора документа;
- качество передачи признаков голоса, используемых для оценки ФПЭС автора документа.

A. Разборчивость РС

Среди общепризнанных подходов [11] к определению разборчивости речи были выделены субъективные и объективные методы, предполагающие соответственно использование специально подобранных экспертов (аудиторов) в соответствии с ГОСТ Р 50840-95 [12] и расчетных методик [13]. Основным недостатком субъективных методов является их трудоемкость, что делает их малоприменимыми на практике, особенно на этапе разработки и проведения многочисленных предварительных испытаний. Как правило, их

целесообразно использовать на заключительных этапах разработки при проведении итоговых приемо-сдаточных испытаний.

Объективные методы являются существенно менее затратными и позволяют определить ФР (артикуляционную) разборчивость РС путем анализа качества передачи и уровня восприятия различных частот человеком. В основе определения ФР лежит анализ формантной структуры наиболее информативных вокализованных участков речи [13–15]. По этой причине эти методы были выбраны в качестве основных.

В. Качество передачи индивидуальных особенностей голоса автора документа

Индивидуальные особенности голоса во многом обусловлены такими его характеристиками, как тембр, высота, окраска [14, 15]. Эти характеристики определяются числом гармоник вокализованных участков речи и их параметрами (амплитуда, ширина). Поэтому качество передачи индивидуальных особенностей голоса может быть охарактеризовано с помощью метрики, определяющей качество передачи его формантной структуры, то есть ФР.

С. Качество передачи признаков голоса для оценки ФПЭС автора документа

При оценке изменения ФПЭС человека по голосу обычно анализируется изменение таких характеристик, как громкость, прерывистость, эмоциональность [7, 8, 10]. Изменения указанных характеристик, прежде всего, проявляются в изменении параметров фонем вокализованных участков РС, например, в изменении энергии отдельных фонем, а также уровня вибраций (тремора) основного тона (первой фонемы). Эти изменения обусловлены, в том числе, степенью напряжения голосовых связок, которое отчетливо проявляется в состоянии стресса, либо усталости.

По этой причине качество передачи информации о фонемной структуре речи может быть использовано в качестве метрики для оценки признаков, используемых для определения ФПЭС автора документа.

Таким образом, ФР может быть использована в качестве единой универсальной метрики для комплексной оценки информативности РП и БП при передаче РС.

III. ТРЕБОВАНИЯ К НЕОБХОДИМОМУ УРОВНЮ ФР

ФР РС (обозначена на рис. 1 как D) позволяет определить с помощью эмпирически полученных зависимостей слоговую $S(D)$ и словесную $W(S)$ разборчивость [13]. На рис. 1 и 2 представлены эти зависимости.

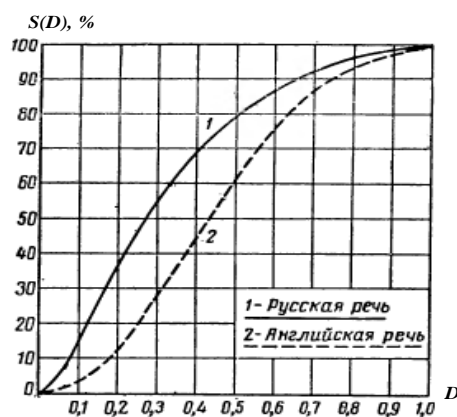


Рис. 1. Зависимость слоговой разборчивости от формантной [13]

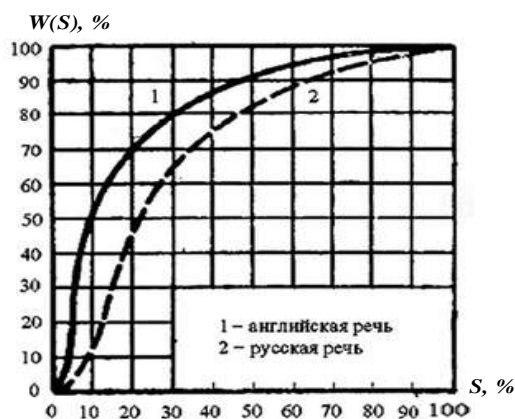


Рис. 2. Зависимость словесной разборчивости от слоговой [13]

Задавая требуемый уровень, например словесной разборчивости, можно оценить необходимый уровень ФР, который должен обеспечить тот, либо иной графический формат представления изображения РП и БП.

В табл. 1 представлены необходимые уровни ФР для обеспечения требуемого уровня словесной разборчивости при заданном качестве передачи контекстного содержания документа [13].

ТАБЛИЦА I ФОРМАНТНАЯ И СЛОВЕСНАЯ РАЗБОРЧИВОСТЬ [13]

№ пп	Качество передачи контекстного содержания документа с помощью РС	Разборчивость, %	
		Словесная (W)	Формантная (D)
1	Предельно-допустимое	75-87	60-67
2	Удовлетворительное	87-93	67-73
3	Хорошее	93-98	73-85
4	Отличное	>98	>85

IV. ОЦЕНКА УРОВНЯ ИЗНОСА НОСИТЕЛЯ ДОКУМЕНТА

В работе рассматривалось старение и износ бумажного, либо пластикового носителя документа в виде выцветания темных участков изображения и появления серых компонентов у белых участков изображения РП и БП, что суммарно приводит к уменьшению контрастности изображения и искажению яркости, в первую очередь, информативных вокализованных участков спектрограммы РС, передаваемой в составе этих маркеров. Для учета влияния указанных факторов в работе предложено

использовать показатель контрастности графического изображения ($0 \leq \text{КГИ} \leq 1$) РП, либо БП:

$$\text{КГИ} = 0,5 \cdot (I_{\text{WMAX}} - I_{\text{BMAX}}) / I_{\text{MEAN}},$$

где I_{WMAX} – максимальные значения яркости белых пикселей изображения; I_{BMAX} – минимальные значения яркости черных пикселей изображения; I_{MEAN} – среднее значение яркости пикселей изображения.

На рис. 3 показаны зависимости диапазона изменения яркости пикселей I_n и формантной разборчивости D от параметра КГИ для изображения спектрограммы в черно-белом формате – $D_{\text{B\&W}}$ и в оттенках серого цвета – D_{GS} .

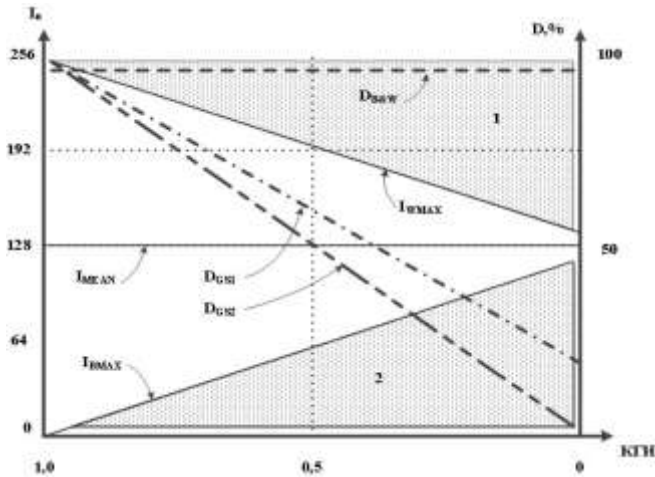


Рис. 3. Зависимости диапазона изменения яркости пикселей I_n и формантной разборчивости D от параметра КГИ

Область 1 соответствует шумам, обусловленным уменьшением яркости белых тонов при старении бумаги. Область 2 соответствует выцветанию черного цвета. Уменьшение параметра КГИ до нуля соответствует полной деградации качества изображения спектрограммы и преобладанию серого оттенка средней интенсивности (128). Для изображения спектрограммы в черно-белом формате полнота передачи информации о величине формант остается постоянной во всем диапазоне изменения параметра $1 \geq \text{КГИ} > 0$. При этом ФР остается также постоянной $D_{\text{B\&W}} = 1$.

Иначе обстоит дело с изображением спектрограммы в оттенках серого (Gray Scale). На качество передачи информации о величине формант в этом случае влияют два фактора, обусловленных деградацией светлых и темных тонов изображения:

- увеличение уровня шумов, приводящее к уменьшению разборчивости формант;
- уменьшение передаваемой амплитуды формант.

Первый эффект отражен в зависимости D_{GS1} , которая соответствует уменьшению формантной разборчивости до $D_{\text{GS1}} = 60\%$ при отношении сигнал/шум, равном единице (наблюдается при $\text{КГИ} = 0,5$) [15]. Второй эффект обусловлен уменьшением на 10–12% эффективного уровня ощущения формант при уменьшении их амплитуды в два раза (для точки $\text{КГИ} = 0,5$). Результирующая зависимость, учитывающая оба эффекта, – D_{GS2} .

Для оценки влияния амплитуды формант на их восприятие был проанализирован метод определения ФР (D) Покровского Н.Б., предполагающий использование выражения [13]:

$$D = \sum_{k=1}^K D_k = \sum_{k=1}^K p_k \times P(E'_k),$$

где D_k – разборчивость для k -ой частотной полосы; p_k – вероятность нахождения формант в k -ой частотной полосе; $P(E'_k)$ – коэффициент восприятия речи; E'_k – эффективный уровень ощущения формант в k -ой частотной полосе (в дБ).

Коэффициент восприятия речи $P(E'_k)$ зависит от громкости речи, а, значит, и от точности передачи оттенков серого в изображении спектрограммы, которые и определяют уровень звучания фонем. Для упрощения вычислений коэффициента $P(E'_k)$ была осуществлена линейная аппроксимация его зависимости от громкости речи. Так, при уменьшении интенсивности звучания в 10 раз (в диапазоне от +10 дБ до -10 дБ) коэффициент восприятия речи уменьшается на 60% (с 80% до 20%) [12-14]. Соответственно, при уменьшении энергии в два раза – коэффициент восприятия речи уменьшается соответственно на 10-12%.

V. РЕЗУЛЬТАТЫ СРАВНЕНИЯ РАЗЛИЧНЫХ ГРАФИЧЕСКИХ ФОРМАТОВ ПРЕДСТАВЛЕНИЯ РП И БП

Для удобства сравнения различных графических форматов представления РП и БП, для зависимости ФР от параметра КГИ – D_{GS2} (Рис. 1) была использована линейная аппроксимация в виде:

$$D_{\text{GS2}} = 100\% \cdot \text{КГИ}.$$

Выигрыш Δ в качестве получаемой ФР при использовании черно-белого формата изображения по сравнению с форматом Grayscale можно определить следующим образом ($1 \geq \text{КГИ} > 0$):

$$\Delta = 100\% \cdot (1 - \text{КГИ}).$$

Говоря об износе бумажного, либо пластикового носителя документа, целесообразно выделить несколько возможных градаций износа:

- невысокий уровень износа при $1 \geq \text{КГИ} > 0,85$;
- средний уровень износа носителя при $0,85 \geq \text{КГИ} > 0,7$;
- высокий уровень износа носителя при $0,7 \geq \text{КГИ} > 0,5$;
- предельный уровень износа при $0,5 \geq \text{КГИ} > 0,3$.

Для качественной передачи в составе РП, либо БП речевой информации, обеспечивающей отличную словесную разборчивость $W > 98\%$ (табл. 1) после восстановления сообщения из графического изображения спектрограммы, необходимо обеспечить ФР на уровне $D > 85\%$. Такое качество передачи спектральной информации может быть обеспечено использованием форматов изображения спектрограмм Black&White и Gray Scale. При этом использование формата Black&White позволяет получить в среднем 7,5% выигрыша по качеству передачи голосовой информации.

В случае более сильного износа носителя только формат Black&White может обеспечить необходимую формантную $D > 85\%$ и словесную $W > 98\%$ разборчивость для воспроизведения речевого сообщения отличного качества. При этом для среднего, высокого и предельного уровней износа носителя документа формат Black&White позволяет получить средний выигрыш соответственно в 22 %, 40 % и 60 %.

Аналогичным образом обстоит дело с качеством передачи в составе РП, либо БП информации о биопараметрах автора защищаемого документа, необходимых для идентификации его личности, а также определения ФПЭС с целью выявления случаев получения подписи под принуждением, либо при неадекватном состоянии автора.

VI. ЗАКЛЮЧЕНИЕ

Таким образом, ФР может быть использована в качестве универсального показателя информативности передаваемой в составе РП, либо БП речевого сообщения автора документа, интегрирующего в своем составе информацию о контекстном содержании его важных разделов, а также биометрические данные его автора, необходимые для идентификации личности и оценки ФПЭС.

Графический формат Black&White, используемый для представления изображения РП, либо БП, позволяет обеспечить необходимую ФР речи автора документа при различных уровнях износа его носителя. При этом выигрыш в ФР по сравнению с использованием графического формата Grayscale достигает 7,5–60 % в зависимости от уровня износа носителя.

СПИСОК ЛИТЕРАТУРЫ

[1] Кочеткова О.В. Особенности правового регулирования использования биометрических документов в Европейском Союзе и Российской Федерации // *Фундаментальные исследования*. 2015. № 2 (часть 5). С.1118–1122.

[2] Коншин А.А. Защита полиграфической продукции от фальсификации. М.: ООО «Синус», 1999. 160 с.

[3] Дворянкин С.В. Речевая подпись. М.: РИО МТУСИ, 2003. 184 с.

[4] Дворянкин С.В. Цифровая обработка изображений динамических спектрограмм аудиосигналов в задачах безопасности речевой связи // *Специальная техника*. 2000. № 3. С. 37–44.

[5] Alyushin A.M. Document protection technology in the digital economics using cognitive biometric methods. *Procedia Computer Science*, 2020, vol. 169, pp. 887–891. Available at: <https://www.sciencedirect.com/science/article/pii/S1877050920302702?via%3Dihub>

[6] Bimbot F., Bonastre J.-F., Fredouille C. A tutorial on text-independent speaker verification. *EURASIP Journal on Advances in Signal Processing*, 2004, no. 4, pp. 430–451. Available at: <https://aspreurasiipjournals.springeropen.com/track/pdf/10.1155/S1110865704310024.pdf>

[7] Alyushin M.V., Kolobashkina L.V., Rozhanskaya Y.N. Measurement of static and dynamic bio-parameters of a person in remote systems for current psycho-emotional and functional state monitoring. *Proc. of the IEEE 2018 Third International Conference on Human Factors in Complex Technical Systems and Environments (ERGO)*, Russia, St. Petersburg, 2018, pp. 161–165.

[8] Demenko, G., Jastrzębska M. Analysis of natural speech under stress. *Acta Physica Polonica A*, 2012, vol. 121, no. 1-A, pp. A92–A95.

[9] Alyushin A.M. Assessment of a person's psychoemotional state by his signature on the document as the basis for his disavowal. *Voprosy Psikhologii*, 2018, no. 2, pp. 133–140.

[10] Giddens C.L., Barron K.W., Byrd-Craven J., Clark K.F., Winter A.S. Vocal indices of stress: a review. *Journal of Voice*, 2013, vol. 27, no. 3, pp. 1–9.

[11] Сравнительный анализ некоторых методов оценки разборчивости речи / Гавриленко А.В., Дидковский В.С., Продеус А.Н. // *Акустичний симпозиум КОНСОНАНС-2007: тезиси докладов / Київ. ІГМ нан України, 25–27 вересня 2007. С. 54–65.*
ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. Дата введения 01.01.1997. М.: ИПК Издательство стандартов, 1996.

[12] Покровский Н.Б. Расчет и измерение разборчивости речи. М.: Гос. изд-во лит. по вопросам связи и радио, 1962. 391 с.

[13] Пирогов А.А. Основы фонетической теории речи // *Русская мысль*. 2010. № 1/12. С. 99–109.

[14] Фланеган Дж.Л. Анализ, синтез и восприятие речи. Пер. с англ. / Под ред. А.А. Пирогова. М.: Связь, 1968. 396 с.