

УДК 681.3.07

П.В. Малинин, В.В. Поляков

Математическое моделирование искажений спектра акустического сигнала при голосовой идентификации личности

P.V. Malinin, V.V. Polyakov

Simulation of Distortion of the Speech Signal for Voice Identification

Предложена математическая модель искажений голосового сигнала, имитирующая влияние психофизиологических особенностей и факторов, вносимых техническими средствами записи и передачи акустических сигналов. С помощью проекционного метода главных компонент проведены количественные оценки влияния искажений на голосовую идентификацию личности.

Ключевые слова: идентификация личности, искажения сигнала, математическая модель, проекционные методы.

Введение. Одним из методов биометрической идентификации личности является метод, основанный на использовании индивидуальных голосовых данных. Достоверность получаемых при этом результатов в определенной степени зависит от такого фактора, как психофизиологическое состояние человека, вносящее искажения в индивидуальные амплитудно-частотные характеристики голоса. Кроме того, искажения исходных акустических сигналов могут быть вызваны использованием различных технических средств при записи или передаче речи, а также в случае применения технических средств обработки голосовой информации [1]. Указанные обстоятельства требуют оценки надежности методов, применяемых при голосовой идентификации личности. В настоящей работе предлагается подход, позволяющий проводить такую оценку путем математического моделирования искажений в спектральном разложении голосового акустического сигнала.

1. Моделирование искажений спектра голосового сигнала. Для математического моделирования искажения голосовых сигналов был применен алгоритм передискретизации, основанный на использовании дискретного преобразования Фурье и позволяющий повышать частоту дискретизации сигнала в задаваемое целое или дробное число раз [2].

A mathematical model of the distortion of the voice signal. The analysis of the spectrum of the distorted voice signal. Quantitative assessment of the impact of distortion on the speaker identification.

Key words: cryptographic protocol, authentication, model checking, linear temporal logic, automatic verification.

Моделирование конкретного искажения осуществлялось следующим образом. Пусть исходный сигнал характеризовался конечным числом отсчетов x_n . На первом шаге алгоритма проводилось вычисление коэффициентов X_k прямого преобразования Фурье:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi \frac{k}{N}n}, \quad k = 0, 1, 2, \dots, N-1.$$

На втором шаге в область возле отсчета с номером $N/2$ спектра вставлялись нулевые компоненты, количество которых задавалось значениями начального числа отсчетов N и числа отсчетов в передискретизованном сигнале M . Коэффициенты Y_i передискретизованного спектра в случае четных чисел N определялись формулами

$$\begin{cases} Y_i = X_i, & 1 \leq i \leq \frac{N+1}{2} \\ Y_i = 0, & \frac{N+1}{2} + 1 \leq i \leq \frac{N+1}{2} + M - N \\ Y_i = X_{i-M+N}, & \frac{N+1}{2} + M - N + 1 \leq i \leq M \end{cases}$$

в случае нечетных N – формулами

$$\begin{cases} Y_i = X_i, & 1 \leq i \leq \frac{N}{2} \\ Y_i = \frac{X_{\frac{N}{2}+1}}{2}, & i = \frac{N}{2} + 1 \\ Y_i = 0, & \frac{N}{2} + 2 \leq i \leq \frac{N}{2} + M - N \\ Y_i = \frac{X_{\frac{N}{2}+1}}{2}, & i = \frac{N}{2} + M - N + 1 \\ Y_i = X_{i-M+N}, & \frac{N}{2} + M - N + 2 \leq i \leq M \end{cases}$$

На заключительном шаге алгоритма вычислялись отсчеты y_m обратного дискретного преобразования Фурье с нормировкой:

$$y_k = \frac{1}{M} \sum_{k=0}^{M-1} Y_k \cdot e^{-i2\pi \frac{k}{M} m}, m = 0, 1, 2, \dots, M-1,$$

на этом формирование искаженного сигнала заканчивалось.

В качестве величины, количественно характеризующей искажения, был использован коэффициент нелинейных искажений K , который вводился как отношение среднеквадратичной суммы спектральных компонентов выходного сигнала, отсутствующих в спектре входного сигнала, к среднеквадратичной сумме спектральных компонентов входного сигнала:

$$K = \frac{\sqrt{\frac{1}{m} \sum_{j=1}^m Y_j^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}}.$$

Искажения генерировались таким образом, что частота дискретизации F ,

$$F = M / t$$

(здесь t – длительность исходного сигнала) увеличивалась за счет изменения количества отчетов от N до M , при этом для каждого искаженного сигнала рассчитывался коэффициент нелинейных искажений K .

2. Влияние искажений голосового сигнала на идентификацию личности. Исходный и искаженные акустические сигналы подвергались процедуре идентификации, описанной в работах [3, 4] и основанной на применении проекционных методов анализа многомерных данных. Полученные при расчетах данные представлялись в виде графиков счетов в пространстве первых главных компонент [5], позволявших наглядно интерпретировать результат идентификации.

Для апробации подхода в качестве исходных сигналов использовались записи голоса диктора, представлявшие из себя фразу «один два три четыре пять шесть семь восемь девять», с частотой дискретизации 8 кГц и разрядностью 16 бит. Искажения в исходный сигнал вносились по описанному в предыдущем разделе алгоритму и заключались в добавлении узкого участка спектра с отсчетами, имевшими равную нулю амплитуду. В качестве графической иллюстрации на рисунке 1 приведены участки передискретизованных спектров, подвергавшихся искажениям, соответствовавшим коэффициентам $K = 0,15; 0,20; 0,25; 0,30$.

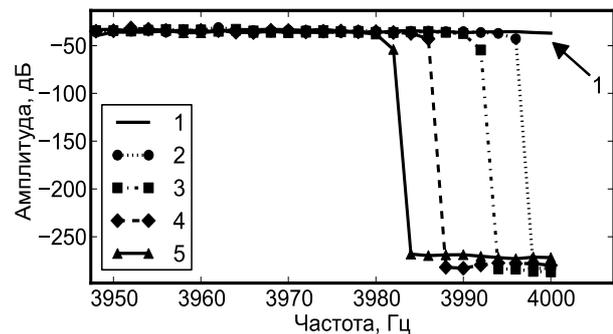


Рис. 1. Участок акустического спектра голосовых сигналов с искажениями: 1 – исходный сигнал, искаженные сигналы с коэффициентом искажения; 2 – $K = 0,15$; 3 – $K = 0,20$; 4 – $K = 0,25$; 5 – $K = 0,30$

Рассчитанный для исходного и искаженных сигналов график счетов представлен на рисунке 2. Здесь отдельная точка соответствует одной голосовой записи, исходные сигналы (использовавшиеся в качестве калибровочного набора) изображены незалитыми точками I, искаженные сигналы с разными значениями коэффициента искажения K – залитыми точками II (точка 1 – $K = 0,15$; 2 – $K = 0,20$; 3 – $K = 0,25$; 4 – $K = 0,30$).



Рис. 2. Влияние искажений голосового сигнала на идентификацию. I – исходные сигналы, II – искаженные сигналы. Коэффициент искажения: 1 – $K = 0,15$; 2 – $K = 0,20$; 3 – $K = 0,25$; 4 – $K = 0,30$

Все исходные (неискаженные) сигналы представлены компактной областью, выделенной эллипсом, который строился по максимальному разбросу соответствующих точек согласно [3]. Попадание точек, описывавших искаженные сигналы, в выделенную область означало правильную идентификацию диктора. Случаи, отвечавшие точкам вне выделенной области, означали, что диктор воспринимался как «чужой», т.е. искажения сигнала оказывались настолько значительными, что идентификация не достигалась. Из рисунка 2 видно, что при коэффициентах искажения $K = 0,15$ и $K = 0,20$ (точки **1** и **2**) диктор, несмотря на искажения сигнала, идентифицировался правильно.

При коэффициентах $K = 0,25$ и $K = 0,30$ (точки **3** и **4**, расположенные вне эллипса) идентификация уже не достигалась.

Заключение. Предложен подход к оценке надежности голосовой идентификации личности, основанный на математическом моделировании искажений амплитудно-частотных характеристик акустического спектра голосового сигнала. Результаты расчетов показали возможность количественной оценки величины искажений, при которых возможна правильная идентификация личности. Разработанный подход может быть применен для анализа надежности методов голосовой идентификации.

Библиографический список

1. Галяшина Е.И. Прикладные основы фоноскопической экспертизы // Теория и практика судебной экспертизы. – СПб., 2003.
2. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов / под ред. М.В. Назарова, Ю.Н. Прохорова. – М., 1981.
3. Малинин П.В., Поляков В.В. Иерархический подход в задаче идентификации личности по голосу с помощью проекционных методов классификации многомерных данных // Доклады ТУСУР. – 2010. – Ч. 1.
4. Малинин П.В., Поляков В.В. Исследование влияния шума на идентификацию личности по голосу // Проблемы информационной безопасности государства, общества и личности, безопасность нанотехнологий. – Томск, 2012.
5. Эсбенсен К. Анализ многомерных данных. Избранные главы / пер. с англ. С.В. Кучерявского; под ред. О.Е. Родионовой. – Черноголовка, 2005.