

значительно меньше времени работы алгоритма свертки в классической плекс-грамматике. К примеру, для плекса на рис. 1.6 продолжительность свертки оценивается в 24 шага ($MEN = 3$; $MNA = 2$; $HT = 3$), в то время как для алгоритма из [9] она бы оценивалась в 125 шагов. На самом деле данный плекс сворачивается за 12 шагов.

Выводы. Разработанные математические основы позволяют эффективно использовать типизированные плекс-грамматики в задачах анализа и преобразования семантически нагруженных графовых структур, представляющих собой диаграммы движения данных.

Список литературы: 1. Formal syntax and semantics of programming languages: a laboratory based approach / *Kenneth Slonneger, Barry L. Kurtz*. – Addison-Wesley PC, 1995. – 637 p. 2. *Колдовский В.* Знакомство с Windows Workflow Foundation – Компьютерное Обозрение, 2006, <http://itc.ua/article.phtml?ID=23217> 3. *Карнов В.* Теория и технология программирования. Основы построения трансляторов. – СПб.: БХВ-Петербург, 2005. – 272 с. 4. *Rekers, J., Schürr A.* Defining and Parsing Visual Languages with Layered Graph Grammars // Journal of Visual Languages and Computing. – 1997. – Vol.10. – P. 27 – 55. 5. *Keukelaar J. H. D.* Topics in Soft Computing: Doctoral Dissertation – Royal Institute of Technology, Department of Numerical Analysis and Computer Science. – Stockholm, 2002. – 176 p. 6. *Kam-Fai Chan, Dit-Yan Yeung* Elastic structural matching for recognising on-line handwritten alphanumeric characters: Technical report HKUST-CS98-07 / Department of Computer Science, The Hong Kong University of Science & Technology, 1998. – 29 p. 7. *Алексеев В.Б., Носов В.А.* NP-полные задачи и их полиномиальные варианты. Обзор // Обозрение промышленной и прикладной математики. – 1997. – Т.4. – №2. – С. 165 – 193. 8. *Льюис Ф., Розенкранц Д., Стирнз Р.* Теоретические основы проектирования компиляторов – М.: Мир, 1990. – 654 с. 9. *Da Qian Zhang, Kang Zhang, Jiannong Cao* A Context-sensitive Graph Grammar Formalism For the Specification of Visual Languages // The computer journal. – 2001. – Vol. 44. – № 3. – 15 p.

Поступила в редакцию 30.09.2006

УДК 621.391

Е.Г. ЖИЛЯКОВ, д-р техн. наук, БелГУ (г. Белгород, Россия),
С.П. БЕЛОВ, канд. техн. наук, БелГУ (г. Белгород, Россия),
Е.И. ПРОХОРЕНКО

ОБ ОДНОМ СПОСОБЕ ОБНАРУЖЕНИЯ ПАУЗ В РЕЧЕВЫХ ДАННЫХ

У статті пропонується один з шляхів рішення задачі зменшення об'єму бітового представлення мовних даних при їх передачі і зберіганні в інформаційно-телекомунікаційних системах. Це новий спосіб виявлення і кодування пауз, заснований на обліку відмінностей в розподілі енергетичних частотних складових звуків мови і сигналу у паузі.

In article is offered one of the ways of decision a problem of reduction of volume of bit presentation voice data. This is the new method founded on highlighting and coding in input an voice a pause based on the account of differences in distributing power frequency constituents of sounds of speech and signal in a pause.

Постановка задачи. В настоящее время наблюдается непрерывное увеличение объемов речевых данных в общем потоке информации, циркулирующем в информационно-телекоммуникационных системах [1, 2]. В связи с этим возникает необходимость решения задачи обеспечения высокой эффективности их хранения и передачи. Одним из путей решения данной задачи может стать разработка новых методов сжатия (кодирования) речевых данных, применение которых позволит значительно сократить объем их битового представления, и, как следствие, уменьшить скорость их передачи по каналам связи, а также ресурсы памяти при их записи на физические носители. Данная работа посвящена рассмотрению именно этих вопросов.

Анализ литературы. При решении задачи кодирования речевых данных целесообразно учитывать ряд особенностей речевого обмена. Речь, как известно [3, 4], обладает смысловой и сигнальной избыточностью и допускает определенный уровень потерь и искажений при осуществлении преобразований. В зависимости от избыточности цифровых данных и длительности обрабатываемых сегментов речи, допустимый уровень потерь может составлять от 1 до 50% времени активности диктора [5]. Различные методы устранения избыточности в процессе преобразования речевых данных представляют широкий диапазон возможных уменьшений объема их битового представления.

Особенностью речевых сигналов является высокая доля пауз, составляющая в среднем, например, при телефонных переговорах, 56 % от длительности диалога [6]. Речь содержит множество кратких перерывов длительностью от 5 до 200 мс, существующих как внутри слов, так и между

словами в слитной речи. Суммарная продолжительность перерывов в среднем занимает около 15 % от продолжительности слитной речи.

Таким образом, сокращение объема речевых данных может быть достигнуто за счет кодирования пауз, объем которых, исходя из вышесказанного, в речевых данных значителен. Кодирование пауз заключается в определении интервала, на котором отсутствуют звуки речи, фиксации начала этого интервала и его длительности. Для того, чтобы осуществить обнаружение пауз в речи и произвести последующее их кодирование с целью уменьшения объема речевых данных при их хранении и передаче, необходимо проанализировать наличие отличительных признаков сигнала в паузе от звуков речи.

Пауза в речевом сигнале – отрезок, содержащий более или менее нерегулярные случайные изменения, т.е. процесс который можно отнести к классу случайных. Если принять, что вероятностная структура паузы не изменяется со временем, то тогда сигнал можно считать случайным стационарным процессом.

Сигнал, соответствующий звукам, не является стационарным, так как формируется при активном воздействии речевого аппарата. Процедура обнаружения и кодирования пауз, как было рассмотрено в [7], может быть основана на принципе обнаружения отличий характеристик сигналов на данном интервале по сравнению с характеристиками сигнала в паузе.

Другим методом сокращения объема речевых данных за счет обнаружения и кодирования пауз может служить метод, основанный на учете отличий в распределении энергетических частотных составляющих звуков речи и сигнала в паузе.

Цель статьи. В данной работе рассматривается задача уменьшения объема речевых данных при их передаче и хранении в информационно-телекоммуникационных системах. При этом основное внимание уделяется новому способу обнаружения и кодирования пауз, основанному на учете отличий в распределении энергетических частотных составляющих звуков речи и сигнала в паузе.

Основы способа обнаружения пауз. Проведенные исследования тонкой структуры частотных спектров звуков речи и сигнала в паузе показывают, что энергия фрагмента сигнала, соответствующего паузе распределена относительно равномерно по всему диапазону частот, тогда как энергия фрагмента сигнала, соответствующего звуку, сосредоточена в узких полосах частот. Таким образом, процедура обнаружения пауз может быть основана на принципе обнаружения отличий распределения энергии сигнала по частотному диапазону, соответствующего звуку, по сравнению с распределением энергии сигнала в паузе. Предлагается разбить частотный диапазон на R одинаковых частотных интервалов, для каждого из которых

определять долю энергии сигнала соответствующего паузе, приходящуюся на данный интервал.

Сформулируем следующую гипотезу:

H_0 : энергия исходного отрезка $\vec{x}_i$, $i = 1, \dots, N$ в r -м частотном интервале $(v_{2r} - v_{1r})$ обусловлена внешними шумами.

Для точного вычисления доли энергии отрезка сигнала $P_{r,M,i}$ в определенном частотном диапазоне будем использовать новый инструмент частотного анализа, приведенный в [8], а именно представление вида:

$$P_r = \vec{x}^T A_r \vec{x}, \quad (1)$$

где $A_r = \{a_{m,n}^r\}$, $m, n = 1, \dots, N$,

$$a_{m,n}^r = \begin{cases} \frac{\sin v_{2r}(m-n) - \sin v_{1r}(m-n)}{\pi(m-n)}, & m \neq n, \\ \frac{v_{2r} - v_{1r}}{\pi}, & m = n. \end{cases} \quad (2)$$

Непосредственное применение (1) затруднено, в виду необходимости вычисления квадратичных форм, особенно когда длительности анализируемых отрезков велики. Поэтому в дальнейшем все вычисления целесообразно проводить с использованием выражений вида:

$$P_r(\vec{x}) = \sum_{k=1}^M (y_{rk})^2; \quad (3)$$

$$y_{rk} = \sqrt{\lambda_{rk}} (\vec{q}_{rk}, \vec{x}), \quad (4)$$

где $r = 1, \dots, R$ – число частотных интервалов; $k = 1, \dots, M$ – число

собственных чисел $\lambda \approx 1$, $\vec{q}_{rk}$ – собственный вектор,

тогда для отрезка сигнала соответствующего паузе можно определить

$$P_r^\Pi = E[\hat{P}_r^\Pi]; \quad (5)$$

$$\hat{P}_r^\Pi = \sum_{k=1}^M (y_{rk}^\Pi)^2, \quad (6)$$

т.е. P_r^Π является математическим ожиданием доли энергии паузы в r -м частотном интервале. Здесь y_{rk}^Π означает компоненты вектора, полученные

на участке паузы. При этом считаем, что $R = \frac{\pi}{v_{2r} - v_{1r}} = const$ (для всех r одинаковое), причем целое и имеет место $RM = N$; N – четное, M – четное, так что $M = 2 \left[N \frac{v_{2r} - v_{1r}}{2\pi} \right]$ – целая часть, $N \frac{v_{2r} - v_{1r}}{2\pi}$ – целое число.

Математическое ожидание доли энергии паузы в r -ом частотном интервале P_r^{Π} (5) получаем путем усреднения значений, определенных для некоторого количества $N_{отр}$ отрезков сигнала, соответствующих паузе. Чем больше таких отрезков вовлечено в анализ, тем меньше погрешность вычисления

$$P_r^{\Pi} = \frac{1}{N_{отр}} \sum_{i=1}^{N_{отр}} P_r^i; \quad \delta^{\Pi} \approx \frac{c}{\sqrt{N_{отр}}}, \quad (7)$$

где c – некоторая константа.

Введем решающую функцию

$$S_r^i = \frac{\hat{P}_r^i}{P_r^{\Pi}}, r = 1, 2, \dots, R. \quad (8)$$

В случае, когда i -й анализируемый отрезок сигнала $\vec{x}_i$ представляет собой отрезок паузы, правая часть (8) представляет собой случайную величину, распределенную по закону

$$\chi_{M_1}^2: \frac{1}{2^{\frac{M_1}{2}} \Gamma\left(\frac{M_1}{2}\right)} (S_r^i)^{\frac{M_1}{2}-1} e^{-\frac{S_r^i}{2}}, \quad (9)$$

т.е. можно установить порог h_{α} , который удовлетворяет неравенству

$$\frac{1}{2^{\frac{M_1}{2}} \Gamma\left(\frac{M_1}{2}\right)} \int_{h_{\alpha}}^{\infty} y^{\frac{M_1}{2}-1} e^{-\frac{y}{2}} dy \leq \alpha, \quad (10)$$

где α – так называемая, вероятность ложной тревоги; M_1 – число степеней свободы, которое можно приблизительно вычислить как $M_1 = \sum_{i=1}^M \lambda_i$.

Для увеличения вероятности правильного обнаружения границы пауза/звук с порогом сравнивается максимальное значение решающей функции, так как энергия сигнала, соответствующего звуку может быть сосредоточена в сравнительно узком диапазоне частот, что приходится учитывать при разбиении всего частотного диапазона на небольшое количество интервалов.

Таким образом, если имеет место

$$\max(S_r^i) \geq h_{\alpha}, \quad (11)$$

то отвергается нулевая гипотеза H_0 .

Если же выполняется неравенство

$$\max(S_r^i) < h_{\alpha}, \quad (12)$$

то считаем справедливой H_0 и относим данный отрезок речевого сигнала, как соответствующий паузе.

Выполнение неравенств вида (12) для всех частотных интервалов соответствует случаю, когда весь анализируемый отрезок $\vec{x}_i$ является отрезком паузы.

Вычислительные эксперименты. Для проверки работоспособности метода были проведены вычислительные эксперименты, которые состояли из нескольких последовательных этапов:

1. Запись речевых данных: x_1 – отрезок сигнала являющийся фрагментом лекции, прочитанной в аудитории. Длительность сигнала 132 с, частота дискретизации 7350 Гц.

2. Определение длительности окна анализа речевого сигнала N и количества частотных интервалов R .

Длительность окна анализа должна быть достаточно малой, чтобы последовательность окон более точно отражала динамику изменения речевого сигнала и, в частности, появления сравнительно коротких пауз. С другой стороны, длительность окна должна быть достаточно большой, чтобы исследуемые характеристики сигнала для каждого окна отражали характеристики квазистационарного участка речевого сигнала. Таким образом, длительность окна принималась 8 мс, т.е. $N = 60$ отсчетов, т.к. только окна с длиной не менее 8 мс отражают периодический характер сигнала, связанный с проявлением основного тона. Экспериментально было установлено, что для определения распределения энергии сигнала достаточно $R = 10$ частотных диапазонов. При этом $M = N/R$, т.е. выполняется неравенство $M \geq 4$ и собственные числа в (4) обладают свойством $\lambda_1 \approx \lambda_2 \approx \lambda_3 \approx \lambda_4 \approx 1$.

3. Вычисление матрицы A с элементами вида (2), при $N = 60$, $R = 10$.

4. Вычисление матрицы AA с элементами вида

$$AA = \begin{pmatrix} \sqrt{L_1^1} (Q_1^1)^T \\ \sqrt{L_1^2} (Q_1^2)^T \\ \dots \\ \sqrt{L_1^R} (Q_1^R)^T \end{pmatrix},$$

где $Q_1 = (\bar{q}_1, \dots, \bar{q}_J)$, $L_1 = \text{diag}(\lambda_1, \dots, \lambda_J)$ получены на основе матрицы A .

5. Выделение отрезков x_2 , соответствующих паузе речевого сигнала и вычисление в каждом из частотных диапазонов подвекторов вида

$$\begin{pmatrix} \bar{y}_1^1 \\ \bar{y}_1^2 \\ \dots \\ \bar{y}_1^R \end{pmatrix} = AA\bar{x},$$

и доли энергии сигнала паузы (6)

$$\hat{P}_r^\Pi = \sum_{k=1}^M (y_{rk}^\Pi)^2.$$

Доля энергии в каждом из частотных диапазонов (6) определялась для 500 отрезков анализа и усреднялась для каждого частотного диапазона

$$P_r^\Pi = \frac{1}{N_{\text{отр}}} \sum_{i=1}^{N_{\text{отр}}} \hat{P}_r^i.$$

Полученные численные значения приведены в табл. 1.

Таблица 1

Границы частотных диапазонов	$\nu_1 = 0;$ $\nu_2 = \pi/10$	$\nu_1 = \pi/10;$ $\nu_2 = 2\pi/10$	$\nu_1 = 2\pi/10;$ $\nu_2 = 3\pi/10$	$\nu_1 = 3\pi/10;$ $\nu_2 = 4\pi/10$	$\nu_1 = 4\pi/10;$ $\nu_2 = 5\pi/10$
Среднее значение $P_r^\Pi \times 10^{-5}$	1,4049	17,4150	3,7250	5,8881	4,3204
Границы частотных диапазонов	$\nu_1 = 5\pi/10;$ $\nu_2 = 6\pi/10$	$\nu_1 = 6\pi/10;$ $\nu_2 = 7\pi/10$	$\nu_1 = 7\pi/10;$ $\nu_2 = 8\pi/10$	$\nu_1 = 8\pi/10;$ $\nu_2 = 9\pi/10$	$\nu_1 = 9\pi/10;$ $\nu_2 = 10\pi/10$
Среднее значение $P_r^\Pi \times 10^{-5}$	5,3262	5,4420	10,8220	3,8715	0,1181

6. Установление величины порогового значения решающей функции.

Величину порога h_α необходимо определять из тех соображений, что наибольшую опасность, с точки зрения разборчивости речи после кодирования пауз и последующего воспроизведения, представляет

«соединение» части отрезка сигнала, соответствующего звуку речи при определении границы пауза/звук (особенно если звук является невокализованным). С другой стороны, если установить небольшую величину порога, то увеличится вероятность ложной тревоги и часть отрезков сигнала, соответствующих паузе, будет определена, как звук, что отрицательно скажется на эффективности сокращения объема данных.

Для установления приемлемой с обеих точек зрения величины порога были проведены эксперименты, в ходе которых было установлено, что при достаточно малой вероятности ложной тревоги и соответствующей величине порога $h_\alpha = 2M_1$, выбранной по таблицам вероятностей распределения χ^2 [9] при $N = 60$, $R = 10$, число переходов решающей функции $\max(S_r^i)$ через установленный порог на отрезках сигнала, соответствующих паузе, не превышает заданную величину. Наиболее типичные результаты приведены в табл. 2. Здесь вероятность ложной тревоги определялась как

$$\alpha = \frac{N_{\text{пер}}}{N_\Pi},$$

где N_Π – длина паузы, $N_{\text{пер}}$ – количество переходов решающей функции через порог.

Таблица 2

Длина паузы, N_Π	30000	25000	14000	6000	14000	20000
Количество переходов решающей функции через порог, $N_{\text{пер}}$	1500	1440	0	120	360	600
Вероятность ложной тревоги, α	0,05	0,0576	0	0,02	0,025714	0,03

С другой стороны, на участках соответствующих звукам речи решающая функция превышала установленный порог с вероятностью 90 – 95%. В табл. 3 приведены результаты исследования поведения решающей функции на отрезках сигнала, соответствующих слитной речи.

Таблица 3

Длина фрагмента слитной речи, N_p	4516	17672	8451
Количество переходов решающей функции через порог, $N_{\text{пер}}$	4156	16832	7671
Вероятность правильного обнаружения, $P_{\text{по}}$	0,92	0,95	0,9077

Причем отрезки сигнала, на которых величина решающей функции не превышала установленный порог, как правило, являлись паузами малой

длительности между фонемами, или слитно произнесенными словами, например, как показано на рис. 2. Здесь приведена фраза "ну, для начала я хочу", произнесенная слитно, и отрезок, определенный как пауза, соответствует паузе между фрагментами фразы "ну" и "для начала я хочу".

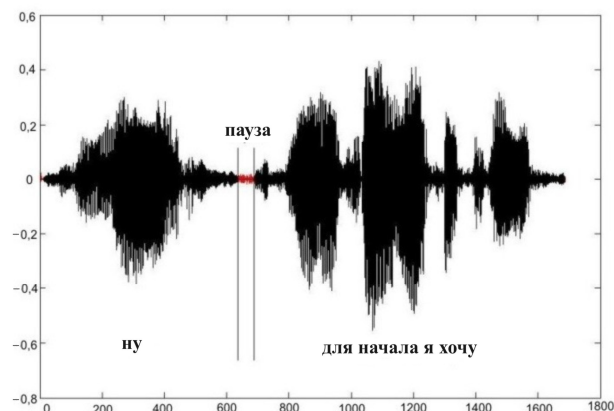


Рис. 2. Обнаружение паузы между словами фразы

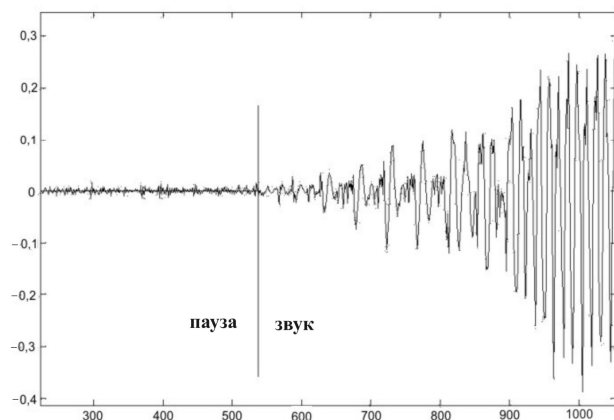


Рис. 3. Граница пауза/звук, определенная в начале фразы

В другом случае (рис. 4) анализировалась фраза, содержащая слово "фактически". Здесь отрезки, определенные как паузы соответствуют паузам между звуко сочетаниями "фак" - "ти" - "чески".

В состав этого слова входят звуки "к", "т" и "ч", которые принадлежат к невокализованным звукам малой длительности, вероятность пропуска которых наиболее велика, из-за их малого (по сравнению с вокализованными звуками) уровня, и в данном случае отрезки определенные как паузы приходятся на окончание звука "к" и начало звуков "т" и "ч". Срезание начала

звуков в этих случаях особенно нежелательно, так как это может снизить разборчивость речи. Визуально сложно определить, на какой части этих звуков значение решающей функции не превышало установленного порога, но экспертная оценка при воспроизведении сигнала с удаленными паузами показала, что звуки "к", "т" и "ч" в слове "фактически" четко различимы.

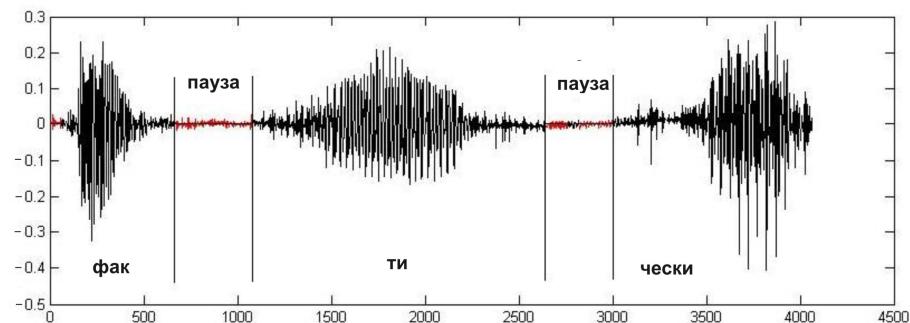


Рис. 4. Обнаружение пауз между звуко сочетаниями

Определенную роль при обнаружении звуков играет уровень шума, присутствующего в исходном речевом сигнале. Для определения чувствительности метода к отношению сигнал/шум были проведены следующие вычислительные эксперименты.

В речевом сигнале, соответствующему звуку, изменялось отношение сигнал/шум и определялась вероятность правильного обнаружения $P_{по}$ звука как

$$P_{по} = \frac{N_{пер}}{N},$$

где N – длина звука, $N_{пер}$ – количество переходов решающей функции через порог.

Наиболее типичные результаты приведены в табл. 4 и на рис. 5.

Из данных табл. 4 и рис. 5 можно сделать вывод, что для достижения величины вероятности правильного обнаружения 95% вокализованных звуков, величина отношения сигнал/шум может быть значительно меньшей, чем для достижения той же вероятности правильного обнаружения невокализованных звуков. Но такие звуки, как правило, произносятся слитно с другими, что увеличивает вероятность их обнаружения. Для большей уверенности, что часть сигнала, соответствующего звукам, не будет потеряна, можно ввести ограничение длины паузы, т.е. если длина паузы не превышает определенной величины, то данный отрезок определяется, как соответствующий звуку. Такое же ограничение можно использовать и на отрезках сигнала, соответствующих паузе, для случаев превышения решающей функцией установленного порога на коротких фрагментах паузы, что может быть обусловлено случайными внешними помехами.

7. Вычисление решающей функции.

На всей реализации сигнала x_1 рассчитана решающая функция (8). Для сравнения с выбранным порогом $h_\alpha = 12$ принималось максимальное из значений решающей функции S_r^i , определенных для каждого анализируемого частотного интервала. На рис. 6 показаны фрагмент речевого сигнала x_1 включающего как паузу, так и звук и значения S_r^i на всем интервале его анализа. Значения РФ на участках, соответствующих звукам речи значительно превышают пороговое значение.

Таблица 4

№	Звук "а" 213,6 мс		Звук "с" 204 мс		Звук "р" 149,7 мс		Звук "к" 8,2 мс		Звук "ч" 108,8 мс	
	$\frac{\text{сигнал}}{\text{шум}}$	$P_{\text{по}}$	$\frac{\text{сигнал}}{\text{шум}}$	$P_{\text{по}}$	$\frac{\text{сигнал}}{\text{шум}}$	$P_{\text{по}}$	$\frac{\text{сигнал}}{\text{шум}}$	$P_{\text{по}}$	$\frac{\text{сигнал}}{\text{шум}}$	$P_{\text{по}}$
1	0,04	0,54	0,01	0,71	0,02	0,61	0,07	0,75	0,02	0,60
2	0,22	0,58	0,15	0,71	0,14	0,61	0,60	0,75	0,32	0,73
3	0,48	0,62	0,37	0,79	0,33	0,67	1,40	0,88	0,79	0,77
4	0,66	0,65	0,51	0,82	0,46	0,72	1,93	0,88	1,10	0,83
5	0,83	0,65	0,65	0,87	0,59	0,72	2,46	0,88	1,41	0,87
6	1,01	0,69	0,80	0,89	0,72	0,72	3,00	0,88	1,72	0,90
7	1,18	0,85	0,94	0,89	0,84	0,83	3,53	0,88	2,03	0,90
8	1,36	0,92	1,08	0,89	0,97	0,83	4,06	0,94	2,34	0,90
9	1,53	0,92	1,23	0,92	1,10	0,89	4,60	0,94	2,65	0,93
10	1,71	0,92	1,37	0,92	1,23	0,89	5,13	0,94	2,96	0,93
11	1,88	0,96	1,51	0,92	1,35	0,89	5,66	0,94	3,27	0,93
12	2,06	0,96	1,66	0,92	1,48	0,94	6,19	0,94	3,57	0,93
13	2,23	1	1,80	0,92	1,61	0,94	6,73	0,94	3,88	0,93
14	2,41	1	1,95	0,92	1,74	0,94	7,26	0,94	4,19	0,93
15	2,58	1	2,09	0,95	1,86	0,94	7,79	0,94	4,50	0,93
16	2,76	1	2,23	0,95	1,99	0,94	8,33	0,94	4,81	0,93
17	2,93	1	2,38	0,97	2,12	0,94	8,86	0,94	5,12	0,97
18	3,11	1	2,52	0,97	2,25	0,94	9,39	0,94	5,43	0,97
19	3,28	1	2,66	0,97	2,37	0,94	9,92	1	5,74	0,97
20	3,46	1	2,81	0,97	2,50	1	10,46	1	6,05	0,97
21	3,63	1	2,95	0,97	2,63	1	10,99	1	6,36	0,97
22	3,81	1	3,09	0,97	2,75	1	11,52	1	6,67	0,97
23	3,98	1	3,24	1	2,88	1	12,06	1	6,98	0,97
24	4,16	1	3,38	1	3,01	1	12,59	1	7,29	0,97
25	4,33	1	3,53	1	3,14	1	13,12	1	7,60	0,97

Аналогичные эксперименты были проведены для других сигналов, записанных различными дикторами (табл. 5). Степень сжатия определялась как отношение количества отсчетов исходного сигнала к количеству отсчетов сигнала с удаленными отрезками сигнала – паузами.

8. Оценка качества воспроизведения речи.

Для подтверждения правильности определения границ пауза/звук было оценено качество воспроизведения речи, при удалении и последующем восстановлении обнаруженных пауз. Если значения РФ на i -м интервале не превышали пороговых значений, то отсчеты речевого сигнала x_1 на данном интервале удалялись и кодировались (значения номеров начальных отсчетов пауз и длительность паузы), таким образом формировалась последовательность, представляющая собой "сжатый" речевой сигнал.

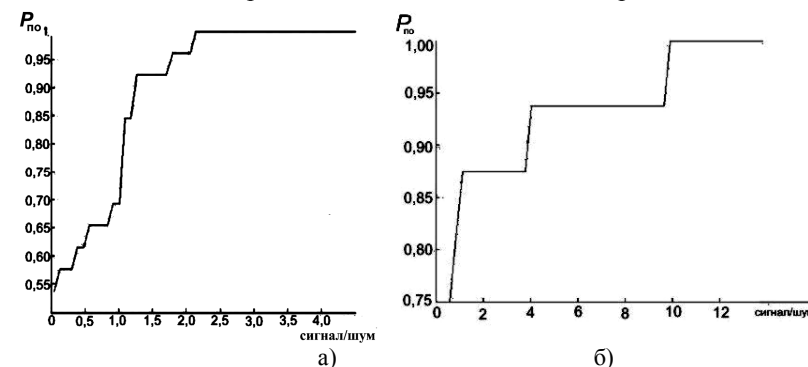


Рис. 5. Зависимость вероятности правильного обнаружения звукового сигнала от величины отношения сигнал/шум для звуков "а" (а) и "к" (б)

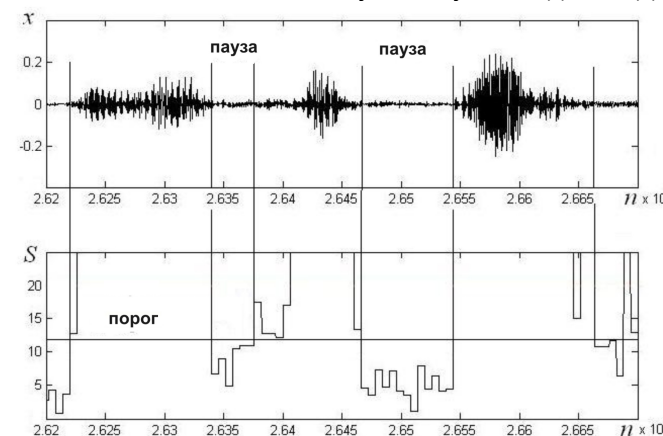


Рис. 6. Отрезок речевого сигнала x_1 и значения решающей функции на всем интервале его анализа

Таблица 5

Результаты сжатия и оценки качества воспроизведения речи	
<i>Фрагмент лекции (132,1680 с)</i>	
Степень сжатия, Ксж	2,7147
Оценка качества воспроизведения речи	4,4
<i>Слово «аппроксимация» (1,4967 с)</i>	
Степень сжатия, Ксж	1,6034
Оценка качества воспроизведения речи	4,5
<i>Слитная речь (161,8750 с)</i>	
Степень сжатия, Ксж	1,7152
Оценка качества воспроизведения речи	4,7

Выводы. Результаты экспериментов свидетельствуют о том, что предлагаемый способ позволяет обнаруживать границу пауза/звук на достаточно коротком интервале анализа, что позволяет повысить степень точности обнаружения границы пауза/звук. Таким образом, предложенный новый метод сжатия речевых сигналов за счет обнаружения и кодирования пауз, основанный на учете отличий в распределении энергетических составляющих звуков речи и сигнала в паузе в частотной области, является эффективным и может быть использован в информационно-телекоммуникационных системах для хранения или передачи речевых сигналов по цифровым каналам связи.

Список литературы: 1. Быков С.Ф., Журавлев В.И., Шалимов И.А. Цифровая телефония. – М.: Радио и связь, 2003. – 144 с. 2. Росляков А.В., Самсонов М.Ю., Шибаева И.В. IP-телефония. – М.: Эко-Тредз, 2001. – 250 с. 3. Калинин Ю.К. Разборчивость речи в цифровых вокодерах. – М.: Радио и связь, 1991. – 220 с. 4. Сжатие данных в системах сбора и передачи информации / В.И.Орищенко, В.Г. Саннико, В.А. Свириденко. Под ред. В.А. Свириденко. – М.: Радио и связь, 1985. – 184 с. 5. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи / Под ред. О.И. Шелухина. – М.: Радио и связь, 2000. – 456 с. 6. Михайлов В.Г., Златоустова Л.В. Измерение параметров речи / Под ред. М.А. Сапожкова. – М.: Радио и связь, 1987. – 168 с. 7. Жиликов Е.Г., Белов С.П., Прохоренко Е.И. О сжатии речевых сигналов // Вестник Национального технического университета "ХПИ". – Харьков: Изд-во НТУ "ХПИ". – 2005. – Вып. 56. – С. 32 – 41. 8. Жиликов Е.Г., Белов С.П., Прохоренко Е.И. Вариационные методы частотного анализа звуковых сигналов // Труды учебных заведений связи / СПб.: СПбГУТ, 2006. – № 174. – С. 163 – 170. 9. Таблицы математической статистики / Л.Н. Большев, Н.В. Смирнов. – М.: Наука. Гл. ред. ф-м. лит., 1983. – 416 с. 10. Гонтмахер Ф.Р. Теория матриц. – М.: Физматлит, 2004. – 560 с.

Поступила в редакцию 20.10.2006

УДК 621.317

И.П. ЗАХАРОВ, д-р техн. наук, ХНУВД (г. Харьков)

АНАЛИЗ ЧИСЛЕННЫХ МЕТОДОВ ОЦЕНИВАНИЯ НЕОПРЕДЕЛЕННОСТИ В ИЗМЕРЕНИЯХ

Проведено аналіз чисельних методів, які застосовуються для оцінювання невизначеності у вимірюваннях. Даються рекомендації щодо їхньої практичної реалізації.

The analysis of the numerical methods used for an evaluation of uncertainty in measurements is carried out. Recommendations on their practical realization are given.

Постановка проблемы. Для оценивания неопределенности измерений широко применяют аналитические методы [1]. В основе их реализации лежит закон распространения неопределенности, заключающийся в приближении исходного модельного уравнения линейными членами ряда Тейлора, который в сочетании с методом суммирования дисперсий и ковариаций позволяет получить выражение для вычисления суммарной стандартной неопределенности (неопределенности результата измерения) $u_c(y)$ в виде

$$u_c(y) = \sqrt{\sum_{i=1}^m c_i^2 u^2(x_i) + 2 \sum_{i=1}^{m-1} \sum_{j=i+1}^m r_{ij} c_i c_j u(x_i) u(x_j)}, \quad (1)$$

где $u(x_i)$ – неопределенность i -й входной величины x_i ; c_i – коэффициент чувствительности; r_{ij} – коэффициент попарной корреляции между i -й и j -й входными величинами. Применение такого подхода при существенно нелинейном модельном уравнении дает смещенную оценку результата измерений и недостоверную оценку суммарной стандартной неопределенности $u_c(y)$. Расширенная неопределенность в [1] оценивается на основе определения эффективного числа степеней свободы. Такой подход имеет ряд существенных ограничений даже при отсутствии попарной корреляции между входными величинами и совершенно неприемлем при ее наличии [2].

Анализ литературы. Устранить все перечисленные недостатки базового метода оценивания неопределенности измерений может применение численных методов. Среди численных методов наиболее известными являются метод дискретного уравнения свертки [3], метод частных приращений [4] и метод Монте-Карло [5].

Цель статьи. Произвести сравнительный анализ существующих численных методов оценивания неопределенности в измерениях для определения диапазона их применимости и путей их дальнейшего развития.

Метод дискретного уравнения свертки. Метод уравнения свертки