

УДК 621.391

**Я.В. ЯНСОНС, А.Ю. СТРЮК***Полтавский военный институт связи, Украина***МЕТОДИКА ОЦЕНКИ ПОМЕХОУСТОЙЧИВОСТИ ПАРАМЕТРОВ  
СЖАТОГО СИГНАЛА КОДЕКОВ РЕЧИ**

Предложена методика оценки помехоустойчивости параметров сжатого сигнала кодеков речи с использованием объективной оценки качества речевых сигналов. Введено понятие помехоустойчивости кодека речи.

**помехоустойчивость, сжатие речи, кодирование речи, объективная оценка качества речи, коротковолновые, ультракоротковолновые радиосети**

**Введение**

На данном этапе развития техники речевой связи разработано и внедрено значительное количество разнообразных цифровых кодеров и декодеров речи (кодеры источника для кодирования речи называются вокодерами (VOiceCODER)), которые обеспечивают приемлемое качество работы в различных сетях связи. В связи с тем, что на рынке оборудования существует большое количество кодеков речи для коротковолновых (КВ), ультракоротковолновых (УКВ) радиосетей, потребитель имеет возможность выбирать необходимый продукт, исходя лишь из компромисса цены и информационной скорости (производительности) представленного продукта.

Постановка задачи определения, расчета и обоснования оценки помехоустойчивости того или иного кодека речи является актуальной на сегодняшний день, исходя из недостатков систем подвижной радиосвязи:

- высокий уровень ошибок (до 10%);
- резкое затухание сигнала на трассе радиосвязи;
- зависимость качества связи от времени суток года и состояния ионосферы и т. д. [1 – 4].

**Анализ литературы.** Существует несколько основных типов построения кодеков речи: полосовые (канальные), формантные, ортогональные, гомоморфные, с линейным предсказанием речи.

Широкое применение нашли кодеки речи с ли-

нейным предсказанием, которые позволяют при достаточно низкой скорости сохранять приемлемую разборчивость речи [5 – 8]. Известны следующие методы линейного предсказания:

- с возбуждением от импульсов основного тона – LPC (Linear Predictive Coding);
- с многоимпульсным возбуждением – MPELP (Multi Pulse Excited Linear Predictive) или MPLPC (Multi Pulse Excited LPC);
- с возбуждением от остатка предсказания – RELP (Residual Excited LP);
- предсказание типа линейной спектрально пары – LSP (Line Spectrum Pair);
- многоимпульсное квантование с максимальным правдоподобием – MP-MLQ (MultiPulse Maximum Likelihood Quantization);
- линейное предсказание с возбуждением от кода – CELP (Codebook Excited Linear Predictive);
- (генерируемое кодом с низкой задержкой) с возбуждением от кода и малой задержкой – LD-CELP (Low Delay - Code Excited Linear Prediction);
- с возбуждением от набора векторов (которые содержатся в кодовых книгах) – VSELP (Vector Sum Excited Linear Prediction);
- (генерируемое алгебраическим кодом) с линейным предсказанием управляемым алгебраическим кодом – ASELP (Algebraic Code Excited Linear Predictive);

- (генерируемое алгебраическим кодом сопряженной структуры) сопряженная структура с управляемым алгебраическим кодом линейным предсказанием – CS-ACELP (Conjugate Structure – Algebraic Code Excited Linear Predictive);
- с возбуждением от регулярных импульсов – RPE-LTP (Residual Pulse Excitation – Long Term Predictive).

В системах связи с подвижными объектами наибольшее применение нашли следующие методы кодирования речи – RPE-LTP, ASELP, VCELP.

**Постановка задачи.** При передаче цифровых речевых сигналов по КВ, УКВ радиосетям информационные последовательности битов подвергаются влиянию помех, в результате чего искажается исходная комбинация.

Исходная комбинация будет искаженной незначительно при наличии единичных некоррелированных ошибок.

Простейший пример – двоичный симметричный канал без памяти (наличие ошибок без группирования). Но при приближении канала к реальному происходит группирование ошибок на приемной стороне, следовательно, снижается достоверность принятой информации [4, 9].

Под помехоустойчивостью кодеков речи, а точнее, параметров речевого сигнала, которые передаются по каналу связи и используются для синтеза исходного речевого сигнала на приемной стороне, понимается способность кодека противостоять вредному действию помех и выдавать значения передаваемых параметров максимально приближенных к переданным.

Помехоустойчивость кодека характеризуется степенью соответствия принятых параметров сигнала (а после восстановления и самого сигнала) переданным при заданной помехе. Таким образом, при сравнении нескольких кодеков речи, более помехоустойчивый будет тот, который при одинаковой помехе продемонстрирует меньшее различие между принятым и переданным сигналами.

Кодирование речи (кодирование источника), заключается в том, чтобы передавать только определенные параметры речевого сигнала, а не весь сигнал [10 – 13].

Таковыми параметрами могут быть:

- параметры, описывающие передаточную функцию речевого тракта человека (коэффициенты линейного предсказания, связанные с ними коэффициенты отражения (КО) или отношение логарифма площадей) речеобразующего фильтра;
- период основного тона для гласных звуков;
- мощность сигнала возбуждения;
- указатель гласный/негласный звук;
- признак типа тон/шум/пауза и др.

Выбранные параметры речевого сигнала в данном случае могут быть представлены следующим образом:

$$\begin{aligned}
 B &= \{b_1, b_2, \dots, b_n\}; \\
 C &= \{c_1, c_2, \dots, c_n\}; \\
 D &= \{d_1, d_2, \dots, d_n\}; \\
 &\quad \vdots \\
 B \supset A, C \supset A, D \supset A, \dots
 \end{aligned}$$

где  $A$  – множество параметров речевого сигнала;  $B, C, D, \dots$  – наборы параметров речевого сигнала (коэффициенты линейного предсказания, коэффициенты отражения, ...) предназначенные для передачи по каналу связи.

Использование этих параметров позволяет синтезировать на приемной стороне исходный сигнал.

Кодированная речь (ансамбль параметров  $A$ ) – сжатое сообщение, является более уязвимым к воздействию помех, чем несжатое сообщение. Традиционно для защиты от помех после кодирования источника применяется кодирование канала, внося избыточность в передаваемое сообщение. Избыточность в этом случае имеет не случайный характер, как это имело место в исходном сообщении, а строго обоснована теоретически и гарантирует на приемной стороне требуемое качество передачи речи.

Искажение определенных параметров речевого сигнала при передаче по каналу связи имеет различную степень влияния на качество восстановленного сигнала, искажения различных бит этих параметров также в разной степени влияют на восстановленный сигнал.

Следовательно, существует возможность разделения всего ансамбля бит параметров с выхода кодера на группы по степени влияния: наиболее чувствительные, умеренно чувствительные, минимально чувствительные к битовым ошибкам и т.д.:

$$H = \{h_1, h_2, \dots, h_n\},$$

где  $H$  – количество бит в одном отсчете речевого кадра;  $h$  – количество бит в определенной группе;  $n$  – количество групп бит в речевом кадре на выходе кодера речи.

Для разделения бит параметров на группы (области) должна быть разработана методика.

### Решение задачи

Одним из наиболее распространенных объективных показателей оценки качества является отношение сигнал/шум (ОСШ), который определяется в соответствии с формулой

$$ОСШ = 10 \log_{10} \left\{ \frac{\sum_{n=0}^M s^2(n)}{\sum_{n=0}^M (s(n) - \hat{s}(n))^2} \right\}, \quad (1)$$

где  $s(n)$  и  $\hat{s}(n)$  – значения исходного и восстановленного (декодированного) речевых сигналов соответственно;  $M$  – общее число выборок в пределах речевого сигнала.

ОСШ является интегральным показателем качества восстановления речи. Более точным показателем, учитывающим присутствие в речевом сигнале низкоамплитудных компонент, является сегментное ОСШ (СЕГОСШ), основанное на вычислении кратковременного ОСШ для каждого  $N$ -точечного сегмента речи:

$$СЕГОСШ = \frac{10}{L} \times \times \sum_{i=0}^{L-1} \log_{10} \left\{ \frac{\sum_{n=0}^{N-1} s^2(iN+n)}{\sum_{n=0}^M (s(iN+n) - \hat{s}(iN+n))^2} \right\}, \quad (2)$$

где  $L$  и  $N$  – число сегментов и отсчетов в сегменте речевого сигнала соответственно;  $i$  – номер сегмента речевого сигнала;  $M = LN$  – число отсчетов речевого сигнала, состоящего из  $L$  сегментов с  $N$  отсчетами.

Так как операция усреднения в приведенном выражении осуществляется после логарифмирования, то СЕГОСШ более точно оценивает качество нестационарного речевого сигнала.

Для оценки степени помехоустойчивости того или иного кодека необходимо произвести сравнение его фактической помехоустойчивости с его потенциальной помехоустойчивостью. Под потенциальной помехоустойчивостью понимают максимально возможную степень соответствия переданного сигнала принятому (т.е. учитываются только помехи вносимые кодером при сжатии и декодером при восстановлении, без учета влияния канала связи). Фактическая помехоустойчивость – помехоустойчивость кодека, определенная с учетом конкретных параметров канала связи (вид канала, вероятность битовой ошибки и т.д.).

Используя рассчитанные значения СЕГОСШ, выделим группу бит, наиболее чувствительных к битовым ошибкам в некую область значений, назовем ее областью грубых ошибок (ГО) и обозначим  $V_1$ . Соответственно группу бит, умеренно чувствительных к битовым ошибкам – в область ошибок средней грубости (ОСГ)  $V_2$ , и минимально чувствительных к битовым ошибкам – в область незначительных ошибок (НО)  $V_3$ .

Деление на группы позволяет осуществить оптимальное кодирование потока параметров сигнала. Наиболее чувствительные группы могут быть закодированы кодом с большей помехоустойчивостью и большей избыточностью, умеренно чувствительные –

менее помехоустойчивым кодом (избыточность снижается), группы, которые относятся к минимально чувствительным, могут передаваться без дополнительного кодирования.

При делении на группы необходимо добиться выполнения следующего условия: искажение бит сигнала в определенной группе в равной степени влияет на качество восстановленного сигнала, а искажение бит сигнала из разных групп имеет различную степень влияния на восстановленный сигнал.

В табл. 1 – 5 и на рис. 1 приведены расчетные значения и графики, соответственно, влияния искаженных параметров речевого сигнала на качество восстановленной речи для кодека речи GSM 06.10 (FR) 13 Кбит/с, при тестировании 10 различных звуковых файлов разговорной речи (16 бит/отсчет, 8 кГц, монофонический сигнал), длительность каждого файла 2000 мс. Таким образом, объем выборки – 1000 наборов параметров сжатой речи на выходе кодера.

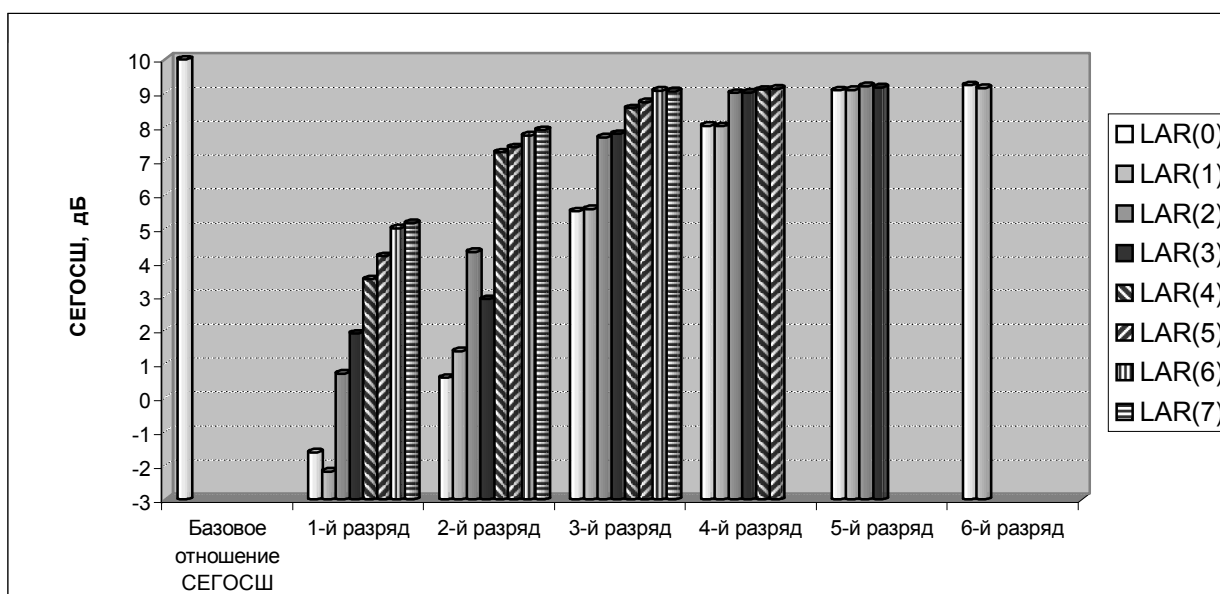


Рис. 1. Зависимость СЕГОСШ от ошибки в фиксированном разряде параметра (логарифм отношения площадей – LAR) сжатого речевого сигнала

В кодеке речи GSM 06.10 (FR) 13 Кбит/с определены следующие параметры кодированной речи:

- логарифм отношения площадей –  $LAR_i$ , где  $i \in [0..7]$ , ( $LAR_0$ ,  $LAR_1$  кодируются 6-ю битами;  $LAR_2$ ,  $LAR_3$  – 5 бит;  $LAR_4$ ,  $LAR_5$  – 4 бита;  $LAR_6$ ,  $LAR_7$  – 3 бита);
- задержка долговременного предсказания –  $N_i$ , где  $i \in [0..3]$ , ( $N_i$  кодируются 7-ю битами);
- усиление долговременного предсказания –  $b_i$ , где  $i \in [0..3]$ , ( $b_i$  кодируются 2-я битами);
- позиция подпоследовательности –  $M_i$ , где  $i \in [0..3]$ , ( $M_i$  кодируются 2-я битами);
- максимальное значение амплитуды импуль-

сов –  $X_{max_i}$ , где  $i \in [0..3]$ , ( $X_{max_i}$  кодируются 6-ю битами);

- амплитуда импульса в одном из 4-х поддиапазонов –  $X_i(j)$ , где  $i$  – номер импульса в поддиапазоне,  $i \in [0..12]$ ;  $j$  – номер поддиапазона,  $j \in [0..3]$ , ( $X_i(j)$  кодируются 3-я битами).

Каждый параметр на выходе кодера был подвергнут искажению – каждый фиксированный разряд поочередно, кратность ошибки  $N = 1$  (канал связи без группирования ошибок, двоичный симметричный канал без памяти).

Анализ влияния искаженных бит параметров речевого сигнала на качество восстановленной речи с

использованием объективной оценки СЕГОСШ позволяет условно разделить шкалу нормированных значений  $y_n$  на две равнозначные области.  $y_n$  – нормированное значение качества речи.

$$y_n = \frac{\bar{x}}{x}, \quad (3)$$

где  $\bar{x}$  – рассчитанное СЕГОСШ, дБ;  $x$  – максимальное (базовое) значение СЕГОСШ, дБ.

Граница условного деления нормированных значений  $y_n$  – критерий  $\mu$  – порог приемлемой помехоустойчивости, характеризует степень влияния передаваемых по линии связи бит параметров сжатого сигнала на качество восстановленного сигнала. Значение данного критерия выбирается экспериментальным путем.

В свою очередь,

$$\mu = f(p_1, p_2, \dots, p_i), \quad (4)$$

где  $p_i$  – вероятность битовой ошибки определенного параметра кодированного речевого сигнала;  $i$  – параметр речевого сигнала.

Вероятность битовой ошибки  $p_i$  – функция зависимости:

$$p_i = f(P_{\text{ошкС}}), \quad (5)$$

где  $P_{\text{ошкС}}$  – вероятность битовой ошибки в канале связи (рассматривается дискретный канал связи).

Для выявления областей распределения разделим шкалу нормированных значений  $y_n$  на две равнозначные области, критерий  $\mu_1$  равен 0,5. Таким образом, все биты, имеющие нормированное значение  $y_n < 0,5$ , распределяются в область ГО  $V_1$ . Разделим оставшуюся область  $y_n (0,5 - 1)$  еще на две равные области. Получим вторую границу  $\mu_2$ , равную 0,75. Биты параметров речевого сигнала в пределах границ  $\mu_1$  и  $\mu_2$  оказываются принадлежащими области  $V_2$  (ОСГ). Биты, распределенные выше границы  $\mu_2$ , – область  $V_3$  (НО).

Следовательно, вычисление определенной области группирования бит сводится к следующему правилу:

$$V_{ГБ} = \begin{cases} y_n \in V_1, \text{ при } y_n < \mu_1; \\ y_n \in V_2, \text{ при } \mu_1 \leq y_n \leq \mu_2; \\ y_n \in V_3, \text{ при } y_n > \mu_2. \end{cases} \quad (6)$$

Учитывая перцепционные свойства слуха человека

$$V_{ГБ} = \begin{cases} y_n \in V_1, \text{ при } \bar{x} < 0, \end{cases} \quad (7)$$

где  $V_{ГБ}$  – все множество бит параметров речевого сигнала, в общем виде получим:

$$V_{ГБ} = \begin{cases} y_n \in V_1, \text{ при } y_n < \mu_1; \\ y_n \in V_2, \text{ при } \mu_1 \leq y_n \leq \mu_2; \\ \vdots \\ y_n \in V_k, \text{ при } \mu_{k-1} \leq y_n \leq \mu_k; \\ \vdots \\ y_n \in V_n, \text{ при } y_n > \mu_n. \end{cases} \quad (8)$$

где

$$\mu_1 = \frac{y_{n\text{max}}}{2}; \quad (9)$$

$$\mu_n = \mu_{n-1} + \frac{\mu_{n-1}}{2}. \quad (10)$$

Диапазон критериев  $\mu$  –  $n$  выбирается из следующих соображений:

$$V_{\text{нд}} \geq ((V_1 + \Delta V_1) + (V_2 + \Delta V_2) + \dots + (V_n + \Delta V_n)) \text{ бит/с}, \quad (11)$$

где  $V_{\text{нд}}$  – скорость передачи данных определенным кодеком речи;  $\Delta V_1, \Delta V_2, \Delta V_n$  – избыточность, вносимая тем или иным помехоустойчивым кодом;

$$\Delta V_n = f(R_m), \quad (12)$$

где  $R_m$  – вид помехоустойчивого кодирования.

Так как все таблицы расчетных значений и графики представляют собой достаточно большой объем данных, приведены лишь результаты расчетов для бит параметров, которые распределены в об-

ласть  $V_1$ , и зависимости СЕГОСШ от ошибки в фиксированном разряде параметра (логарифм отношения площадей – LAR) сжатого речевого сигнала (табл. 1 – 4).

Базовое отношение СЕГОСШ для всех расчетов равно 10 дБ. СЕГОСШ  $i$  – СЕГОСШ для ошибки в  $i$ -м разряде, дБ.

Таблица 1

СЕГОСШ для параметра LAR (логарифм отношения площадей) при ошибке в фиксированном разряде

	СЕГОСШ 1	СЕГОСШ 2	СЕГОСШ 3
LAR(0)	-1,6088	0,5952	5,5190
LAR(1)	-2,1715	1,3836	5,5785
LAR(2)	0,7206	4,3127	
LAR(3)	1,9015	2,9230	
LAR(4)	3,5070		
LAR(5)	4,1934		
LAR(6)	5,0173		
LAR(7)	5,1644		

Таблица 2

СЕГОСШ для параметра  $N$  (задержка) при ошибке в фиксированном разряде

	СЕГОСШ 1	СЕГОСШ 2	СЕГОСШ 3	СЕГОСШ 4
N(0)	4,3503	5,1267	5,8705	5,5718
N(1)	4,4302	4,9638	5,2326	5,3718
N(2)	4,5216	5,7852	5,4937	5,7346
N(3)	4,2588	5,1426	5,8607	5,2191

Таблица 3

СЕГОСШ для параметра  $N$  (задержка) при ошибке в фиксированном разряде (продолжение)

	СЕГОСШ 5	СЕГОСШ 6
N(0)	5,6213	5,7896
N(1)	5,4568	5,8695
N(2)	5,7545	5,9256
N(3)	5,4256	5,9564

Таблица 4

СЕГОСШ для параметра  $X_{max}$  (максимальное значение амплитуды импульсов) при ошибке в фиксированном разряде

	СЕГОСШ 1	СЕГОСШ 2
Xmax0	5,4467	5,6696
Xmax1	5,5454	5,7802
Xmax2	5,8121	5,9033
Xmax3	5,7342	5,9768

Совокупность бит составляющих область  $V_2$ : LAR<sub>0</sub>, LAR<sub>1</sub> – 4-е разряды, LAR<sub>2</sub>, LAR<sub>3</sub> – 3-е разряды, LAR<sub>4</sub>, LAR<sub>5</sub> – 2, 3-е разряды, LAR<sub>6</sub>, LAR<sub>7</sub> – 2-е разряды;  $N_i$  – 7-е разряды;  $b_i - 1$ , 2-е разряды;  $M_i - 1$ , 2-е разряды;  $X_{max_i} - 3$ , 4-е разряды;  $X_i(j) - 1$ -е разряды;  $X_0(0) - X_2(12) - 2$ -е разряды.

Область  $V_3$  включает: LAR<sub>0</sub>, LAR<sub>1</sub> – 5, 6-е разряды, LAR<sub>2</sub>, LAR<sub>3</sub> – 4, 5-е разряды, LAR<sub>4</sub>, LAR<sub>5</sub> – 4-е разряды, LAR<sub>6</sub>, LAR<sub>7</sub> – 3-е разряды;  $X_{max_i} - 5$ , 6-е разряды;  $X_3(0) - X_3(12) - 2$ -е разряды,  $X_i(j) - 3$ -е разряды.

Используя данную методику для кодека речи GSM 06.10 (FR), получим следующее деление бит сжатого сигнала (табл. 5).

Незначительное расхождение результатов распределения бит по областям, при использовании объективных оценок, с делением, осуществленным в системе GSM, с использованием субъективных оценок [14], подтверждает достоверность данной методики.

Таблица 5

Сравнение полученных результатов распределения бит по областям

Данная методика		GSM 06.10 (FR)	
Распределение бит по областям	Количество бит	Распределение бит по классам	Количество бит
VTB	260	Общее число бит	260
V1	46	Ia	50
V2	129	Ib	132
V3	85	II	78

Основные этапы разработанной методики:

- выбор объективного показателя оценки качества речи;
- моделирование параметров канала связи;
- выполнение расчетов для выбранного показателя с учетом характеристик смоделированного канала связи;
- нахождение максимума нормированного коэффициента;
- расчет порогов помехоустойчивости;
- группирование элементарных единиц передаваемого сообщения в соответствии с полученными областями.

### Заключение

Предложенная методика позволяет оценить помехоустойчивость параметров сжатого речевого сигнала, объективно распределить данные параметры по степени влияния (областям) на восстановленный речевой сигнал, благодаря чему может быть снижена канальная скорость передачи при сохранении таких параметров речевого сигнала, как разборчивость и громкость в допустимых пределах.

### Литература

1. Головин О.В., Чистяков Н.Н. Радиосвязь. – М.: Горячая линия – Телеком, 2001. – 288 с.
2. Каганов В.И. Радиопередающие устройства. – К.: Академия, 2002. – 288 с.
3. Маковеева М.М., Шинаков Ю.С. Системы связи с подвижными объектами. – М.: Радио и связь, 2002. – 440 с.
4. Бабак В.П., Наритник Т.М. Обробка сигналів у радіоканалах цифрових систем передавання інформації. – К.: НАУ, 2005. – 475 с.
5. Коротаев Г.А. Методы линейного предсказания // Зарубежная радиоэлектроника. – 1980. – № 10. – С. 45-52.
6. Коротаев Г.А. Анализ и синтез речевого сигнала методом линейного предсказания // Зарубежная радиоэлектроника. – 1990. – № 3. – С. 67-71.
7. Коротаев Г.А. Некоторые аспекты линейного предсказания при анализе и синтезе речевого сигнала // Зарубежная радиоэлектроника. – 1991. – № 7. – С. 25-33.
8. Коротаев Г.А. Эффективный алгоритм кодирования речевого сигнала на скорости 4,8 кбит/с и ниже // Зарубежная радиоэлектроника. – 1996. – № 3. – С. 50-56.
9. Финк Л.М. Теория передачи дискретных сообщений. – М.: Сов. радио, 1970. – 728 с.
10. Маркел Дж.Д., Грей А.Х., Линейное предсказание речи. – М.: Связь, 1980. – 308 с.
11. Шаффер Р.В., Рабинер Л.Р. Цифровая обработка речевых сигналов.: Пер. с англ. / Под ред. М.Б. Назарова, Ю.Н. Прохорова – М.: Радио и связь, 1981. – 784 с.
12. Назаров Х.З., Прохоров Ю.Н. Методы цифровой обработки и передачи речевых сигналов. – М.: Связь, 1982. – 236 с.
13. Прохоров Ю.Н.. Статистические модели и рекуррентное предсказание речевых сигналов. – М.: Радио и связь, 1986. – 316 с.
14. GSM 06.01: "Digital cellular telecommunications system (Phase 2+); Full rate speech; Transcoding".

*Поступила в редакцию 1.03.2006*

**Рецензент:** канд. техн. наук, доцент О.Н. Одарущенко, Полтавский военный институт связи, Полтава.