

УДК 621.391

Я.В. ЯНСОНС*Военный институт телекоммуникаций и информатизации
Национального технического университета Украины „КПИ”, Украина***МЕТОД ПОВЫШЕНИЯ КАЧЕСТВА ВОССТАНОВЛЕННОГО СИГНАЛА
НА ОСНОВЕ ДИФФЕРЕНЦИРОВАННОЙ ЗАЩИТЫ КАДРА РЕЧЕВЫХ ДАННЫХ**

Предлагается новый метод повышения качества восстановленного речевого сигнала. Данный метод основан на создании структуры кадра речи и дифференцированной защиты структурированного кадра, что позволяет максимизировать значение качества принимаемого речевого сигнала. Метод позволяет: разместить, объективно оцененные биты кадра речи на основе действующих рекомендаций ITU-T, в передаваемом кадре в порядке снижения степени их важности в процессе восстановления принимаемого сигнала; распределить биты параметров по областям группирования с различным уровнем защиты; выбрать уровень защиты каждой области, из имеющегося семейства помехоустойчивых кодов, с учетом пропускной способности канала связи. Приведены результаты применения метода для низкоскоростного кодека MELP на скорости 2.4 Кбит/с.

помехоустойчивость, сжатие речи, кодирование речи, канальное кодирование**Введение**

Передача речевого трафика в системах коммуникации составляет более 50 % всей информации. Значительное снижение доли речевой информации в ближайшее время, по прогнозам специалистов не предполагается.

За последнее время значительно повысились требования предъявляемые к качеству передаваемой речевой информации. Качество речи - параметр, характеризующий оценку звучания речи в испытываемой системе низкоскоростной передачи речи, выраженную в баллах по пятибалльной шкале или в процентах предпочтения при сравнении с эталонным трактом [1].

Качество речи напрямую зависит от достоверности принятых данных.

Достоверность в свою очередь изменяется под влиянием помех действующих в канале передачи.

Предназначение

Предлагаемый метод предназначен для повышения качества восстановленного речевого сигнала путем создания канальной структуры кадра речи и

дифференцированной защиты полученной структуры. Под созданием канальной структуры кадра речи понимается следующее:

- определение количества групп, по степени влияния на качество восстановленного сигнала, в рассматриваемом кадре;
- определение наполнения групп (распределение определенных бит параметров по соответствующим группам);
- определение номиналов скоростей выбранного семейства помехоустойчивых кодов для каждой из групп.

**Математическая постановка
и решение задачи**

Система кодирования речи рассматривается на этапах как кодирования источника так и канального кодирования и семейство кодов C_1, \dots, C_z , имеющих свои исправляющую способность и вероятность ошибки декодирования.

При заданной пропускной способности канала связи и длине кадра речи осуществляется определение при заданных условиях, количества облас-

тей групування, порядку розподілення бит по областям, а также набора помехоустойчивых кодов для возможности максимизации значения качества речи восстановленного сигнала на приемной стороне.

Рассматриваемый канал связи – дискретный симметричный без памяти (ДСБП).

Помехи – независимые с равномерным распределением.

Каждый код из семейства характеризуется своей скоростью $\hat{R} = \{r_1, \dots, r_z\}$, так что $(r_1 < r_2 < \dots < r_z)$, в соответствии с выбранной группой кодов C_1, \dots, C_z .

В случае разбиения кадра речи на fr подкадров, установим множество кодовых скоростей $R = (r_i^1, \dots, r_i^{fr}) \in \hat{R}^{fr}$, где к каждому подкадру fr применен каналный код C_i , со скоростью $r_i^{fr} \in \hat{R}$, так что $r_i^I < r_i^{II} < \dots < r_i^{fr}$.

Выбор количества областей группирования (подкадров) ограничен значением $fr \leq 3$, так как, дальнейшее увеличение числа областей приведет к незначительному выигрышу по качеству в сравнении с увеличением сложности реализации. Каждая область требует использования отличного от других помехоустойчивого кода.

Установим $p(r_i^{fr})$ – вероятность ошибки декодирования подкадра fr , защищенного кодом C_i . Так как семейство кодов C_1, \dots, C_z взято следующим образом со скоростями $\left\{ \frac{c}{b}, \frac{c+1}{b}, \dots, \frac{b-1}{b} \right\}$, где c и b – положительные целые числа, при этом $c < b$, то $p(r_1^{fr}) < \dots < p(r_z^{fr}) < 1$.

Допущением является такое применение помехоустойчивых кодов, которое обеспечивает снижение уровня защиты бит параметров исходя из снижения степени влияния на качество восстановленного сигнала.

Для рассматриваемой системы вероятность

ошибки декодирования кратности t равна

$$P_{n_{fr}}^m(R, t) = C_{n_{fr}}^t \cdot p(r_i^{fr})^t \cdot (1 - p(r_i^{fr}))^{n_{fr}-t}, \quad (1)$$

где m – количество возможных вариантов размещения N бит по fr областям:

$$m(N) = N + 1 \Big|_{fr=2}; \quad (2)$$

$$m(N) = \sum_{j=1}^{N+1} j \Big|_{fr=3}; \quad (3)$$

$C_{n_{fr}}^t$ – число сочетаний из n_{fr} по t ;

$$C_{n_{fr}}^t = \frac{n_{fr}!}{t!(n_{fr}-t)!}; \quad (4)$$

fr – количество подкадров основного кадра;

n_{fr} – количество бит в подкадре fr ;

Обозначим через Q^i – множество значений качества речи при кратности ошибки i , так что

$$Q^i = \{Q_{1,\dots,i}; \dots; Q_{1,\dots,N}; \dots; Q_{1,\dots,N-1,N}; \dots; Q_{N-(i-1), N-(i-2), \dots, N-1, N}\}. \quad (5)$$

$Q^i \in \widehat{Q^{t_{max}}}$, где $\widehat{Q^{t_{max}}}$ – множество всех возможных значений качества речи при максимальной кратности ошибки t_{max} .

Q_{max} – максимально возможное значение качества речи кодека при условии отсутствия помех в канале связи.

Значения множества $\widehat{Q^{t_{max}}}$ и Q_{max} находятся экспериментально [2, 3].

Множество значений степени снижения качества речи при искажении i бит e^i , $i=1, \dots, t_{max}$, определяется следующим образом

$$e^i = \{e_{1,\dots,i} | e_h = Q_{max} - Q_h, h = (1, \dots, i), \dots, (N-(i-1), N-(i-2), \dots, N-1, N)\}. \quad (5)$$

В случае рассмотрения подкадра fr в роли множества значений степени снижения качества речи при искажении i бит e^i рассматривается множество e_{fr}^i , которое аналогично e^i . Отличие состоит в том что индексы 1 и N в группе индексов $(1, \dots, i), \dots, (N-(i-1), N-(i-2), \dots, N-1, N)$ при e в

множестве e_{fr}^i , являются соответственно первым и последним индексами в подкадре fr .

Обозначим через Z_{fr}^i – сумму значений группы элементов множества e_{fr}^i при кратности ошибки i подкадра fr .

Тогда для e_{fr}^i

$$Z_{fr}^1 = \sum_{i=1}^{n_{fr}} e_i ; \quad (6.1)$$

$$Z_{fr}^2 = \sum_{i=1}^{n_{fr}} \sum_{j=i+1}^{n_{fr}} e_{i,j} ; \quad (6.2)$$

$$Z_{fr}^3 = \sum_{i=1}^{n_{fr}} \sum_{j=i+1}^{n_{fr}} \sum_{g=j+1}^{n_{fr}} e_{i,j,g} . \quad (6.3)$$

При увеличении i вычисление Z_{fr}^i аналогично.

Значение усредненной ошибки качества речи подкадра fr $e_{fr}^m(t)$.

$$e_{n_{fr}}^m(t) = Z_{fr}^t / C_{n_{fr}}^t . \quad (7)$$

Таким образом ожидаемое искажение для определенного множества скоростей R , и варианта размещения бит m будет иметь следующий вид:

$$F^m(e, R) = \sum_{j=1}^{fr} \sum_{i=1}^t P_{n_j}^m(R, t) \cdot e_{fr}^m(t) . \quad (8)$$

Исходные данные и результаты

Исходя из поставленной задачи и рассмотренного математического аппарата исходными данными для использования методики будут следующие:

- N – количество бит в кадре речи;
- fr – максимальное количество подкадров основного кадра;
- \widehat{Q}^t – множество всех возможных значений качества речи при максимальной кратности ошибки t ;
- t – максимальное значение рассматриваемой кратности ошибки в канале связи;
- P_{KC} – вероятность ошибки в канале связи;

– C_1, \dots, C_z – семейство канальных кодов, характеризующиеся своими исправляющей способностью и скоростью;

– $p(r_z)$ – вероятность ошибки декодирования канального кода со скоростью r_z ;

– V_{kc} – пропускная способность канала связи.

Результат применения метода – канальная структура кадра речи, дифференцированная защита полученной структуры. Полученная структура включает в себя количество областей группирования бит параметров в соответствии со степенью влияния на восстановленный сигнал, распределение конкретных бит по областям, параметры корректирующего кода каждой области. Итоговая структура позволяет получить максимальное качество передаваемого речевого сигнала.

Пример использования методики

Рассмотрим использование метода на примере кодека речи типа MELP (Mixed Excitation Linear Prediction – линейное предсказание со смешанным возбуждением) на скорости 2,4 Кбит/с [4].

Исходными данными для использования метода являются биты параметров кодека MELP объективно ранжированные в зависимости от степени влияния на качество восстановленного сигнала [1], а также значения качества речевого сигнала при различной степени кратности ошибки канала связи t_{max} . Кратность рассматриваемой ошибки t учитывается в пределах двух исходя из снижения значения вероятности ошибки большей кратности для возможных значений ошибки в радиоканале [5]. Примем $t_{max}=2$. Канал связи дискретный симметричный без памяти с заданной вероятностью ошибки P_{KC} . Так как рассматриваются коротковолновые и ультракоротковолновые радиосети примем $P_{KC} = 0,01 \div 0,02$.

В качестве семейства кодов рассматриваются сверточные коды со скоростями 1/4, 1/3, 1/2, а также коды полученные путем выкалывания из

базовых кодов со скоростями 1/3 и 1/2 соответственно 3/5, 5/7 и 2/3, 3/4, 4/5.

Были выбраны коды со следующими полиномами для 1/4 – (135 135 147 163), 1/3 – (117 127 155), 1/2 – (133 171). Для кодов с выкалыванием были

взяты матрицы выкалывания: для 2/3 – $\begin{bmatrix} 11 \\ 10 \end{bmatrix}$, для 3/4

– $\begin{bmatrix} 110 \\ 101 \end{bmatrix}$, для 4/5 – $\begin{bmatrix} 1101 \\ 1010 \end{bmatrix}$, для 3/5 – $\begin{bmatrix} 11 \\ 11 \end{bmatrix}$, для 5/7 –

$\begin{bmatrix} 101 \\ 110 \\ 111 \end{bmatrix}$ [5 – 7].

Практическая часть заключается в том, чтобы сравнить при заданной пропускной способности канала связи значения качества речевого сигнала в случае применения единого помехоустойчивого кода ко всем передаваемым данным и качества речи полученного при разделении бит параметров на группы с использованием выбранных кодов из заданного семейства. Критерием отбора той или иной группы является минимальное значение ожидаемой ошибки.

Исследования проводились путем имитационного и математического моделирования на ПЭВМ.

На рис. 1 и 2 представлены результаты применения разработанного метода.

Рис. 1 соответствует ситуации при вероятности ошибки в канале связи $P_{кс}=10^{-2}$, а рис. 2 – $P_{кс}=2*10^{-2}$. Каждому номиналу пропускной способности соответствует значение кодовой скорости при равномерном кодировании.

Диапазон изменения усредненных значений качества сигнала для кодека MELP находится между 0,7 и 2,5 в единицах MOS. С целью упрощения восприятия результатов данный диапазон был масштабирован.

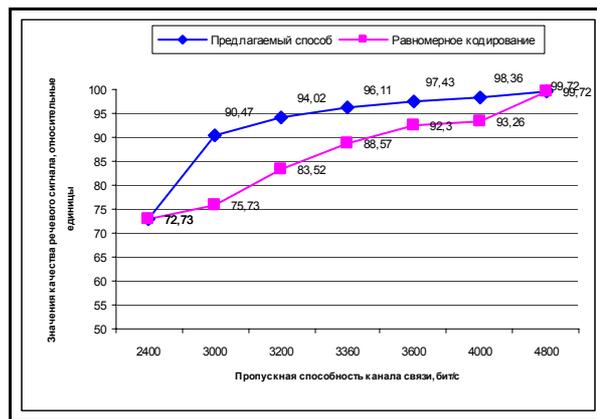


Рис. 1. Сравнение результатов предложенного метода и метода равномерной защиты речевых кадров. $P_{кс}=1*10^{-2}$

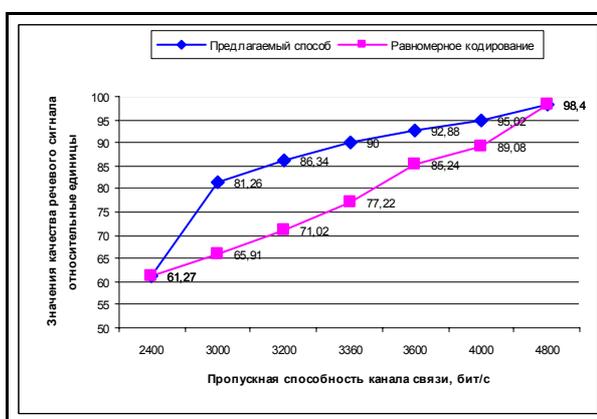


Рис. 2. Сравнение результатов предложенного метода и метода равномерной защиты речевых кадров. $P_{кс}=2*10^{-2}$

Значения от 0,7 до 2,5 заменены шкалой от 0 до 100.

В табл. 1 представлены для обоих методов кодирования:

- значения качества речевого сигнала в единицах MOS;
- полученный выигрыш, в процентах;
- полоса пропускания канала связи, бит/с;
- варианты размещения бит параметров передаваемого сигнала по областям группирования;
- скорости помехоустойчивых кодов.

Как видно из приведенных результатов предложенный метод позволяет повысить качество принимаемого сигнала на 10-15 % в сравнении с применением единого помехоустойчивого кода ко всей совокупности передаваемых данных.

Таблица 1

Значения для обоих методов кодирования

Пропускная способность канала связи	Значения качества речевого сигнала в единицах MOS				Размещение бит параметров	Кодовые скорости	Размещение бит параметров	Кодовые скорости	Выигрыш	
	$P_{кс}=1*10^{-2}$		$P_{кс}=2*10^{-2}$						$P_{кс}=1*10^{-2}$	$P_{кс}=2*10^{-2}$
	Предлагаемый метод	Равномерное кодирование	Предлагаемый метод	Равномерное кодирование	Предлагаемый метод	Равномерное кодирование				
2400	2,0093	2,0093	1,803	1,803	54	-	54	-	-	-
3000	2,3285	2,0632	2,1628	1,8865	2 30 22	2/3 5/7 -	54	4/5	14,73%	15,35%
3200	2,3924	2,2034	2,2542	1,9784	45 9	5/7 -	54	3/4	10,5%	15,32%
3360	2,43	2,2944	2,3201	2,09006	2 50 2	2/3 5/7 -	54	5/7	7,53%	12,78%
3600	2,4539	2,4539	2,372	2,2344	9 45	1/2 5/7	54	2/3	5,13%	7,64%
4000	2,4705	2,3788	2,4104	2,3036	24 30	1/2 5/7	54	3/5	5,09%	5,93%
4800	2,4951	2,4951	2,4712	2,4712	54	1/2	54	1/2	-	-

Заклучение

Таким образом, предложен метод создания канальной структуры кадра речи, позволяющей максимизировать качество передаваемых речевых данных. Данная структура в отличие от существующих, позволяет определить количество областей группирования передаваемых бит, автоматизировано разместить, объективно оцененные биты кадра речи в передаваемом кадре. Размещение осуществлено в порядке снижения степени важности бит на качество восстановленного сигнала. Выбрать необходимую степень защиты каждого уровня с учетом пропускной способности канала связи.

Применение такой структуры обеспечивает повышение помехоустойчивости принятого речевого сигнала на 10-15 % в отличии от метода равномерной защиты передаваемых данных.

Литература

- ГОСТ Р 51061-97 Системы низкоскоростной передачи речи по цифровым каналам параметры качества речи и методы измерений.
- Стрюк А.Ю., Янсонс Я.В. Методика исследова-

ования помехоустойчивости параметров кодеков речи с использованием оценок PESQ // Информационные технологии и компьютерная инженерия. – 2007. – № 1(8). – С. 130.

3. Стрюк А.Ю., Янсонс Я.В. Методика исследования помехоустойчивости бит параметров низкоскоростных кодеков речи // Радиоэлектронные и компьютерные системы. – 2007. – № 5. – С. 160 - 165.

4. Supplee L., Cohn R., Collura J., McGree A. A 2,4 kbit/s MELP- New U.S. Federal Standard Proceedings of ICASSP, IEEE. –1997. – P.1591-1594.

5. Харкевич А.А. Борьба с помехами. – М.: Наука, 1965. – 276 с.

6. Todd K. Moon, Error correction codes (Mathematical methods and algorithms). – Hoboken, New Jersey. – John Wiley & Sons, Inc., 2005. – 804 p.

7. Блейхут Р., Теория и практика кодов контролируемых ошибки. – М.: Мир, 1986. – 576 с.

Поступила в редакцию 12.02.2008

Рецензент: канд. техн. наук К.О. Польщиков. Военный институт телекоммуникаций и информатизации Национального технического университета Украины «КПИ», Полтава.