

УДК 621.391

Я.В. Янсон

Военный институт телекоммуникаций и информатизации НТУУ «КПИ», Полтава

ОЦЕНКА КОЛИЧЕСТВА ОБЛАСТЕЙ ДИФФЕРЕНЦИРОВАНИЯ ЗАЩИТЫ ПАРАМЕТРОВ КАДРОВ КОДЕРА РЕЧИ

Показана актуальность основных проблем, затрагиваемых в исследованиях. Произведено обоснование выбора количества областей дифференцирования защиты элементов параметров кадров выходных речевых сигналов кодеков речи. Показаны преимущества использования определенного количества областей дифференцирования защиты элементов сжатого речевого сигнала. Выбор областей осуществляется в методе повышения качества восстановленного речевого сигнала на основе критерия минимального ожидаемого искажения кадра речи и дифференцируемой защиты объективно распределенных элементов параметров кадров выходных речевых сигналов кодеков речи. Осуществлена оценка вычислительной сложности применяемых помехоустойчивых кодов из семейства доступных для различных вариантов выбора областей защиты элементов. Произведена оценка объемов выделяемой оперативной памяти для осуществления процедур декодирования помехоустойчивых кодов, используемых в методе.

Ключевые слова: кодирование речи, объективная оценка качества речи, канальное кодирование, качество речевого сигнала.

Введение

В настоящее время наблюдается тенденция дифференцирования защиты элементов данных от ошибок в области сжатия и передачи различных типов данных [2]. Выбор данных методов поясняется тем, что они позволяют осуществить необходимую степень защиты передаваемых данных исходя из требований, предъявляемых к качеству восстановленных сигналов и ширине полосы пропускания.

Постановка задачи. Для обеспечения высокого качества восстановленных на приемной стороне речевых сигналов существует множество различных методов.

Повышение качества осуществляется на различных этапах, как формирования, так и преобразования сообщений при передаче [3, 4]. Каждому этапу соответствуют свои методы учитывающие характеристики сигнала на данном этапе.

Предложенный в работе [1] метод повышения качества восстановленного речевого сигнала кодеков речи путем дифференцирования защиты элементов сигнала решает данную задачу на промежуточном этапе между кодером источника и канальным кодером.

При использовании метода повышения качества восстановленного речевого сигнала кодеков речи путем дифференцирования защиты элементов сигнала одним из главных является вопрос о количестве групп разбиения основного кадра речи. Снижение количества градаций защиты приводит к снижению качества результирующего потока. Излишнее повышение числа областей дифференцирования защиты элементов влечет за собой возрастание сложности реализации на этапе защиты.

Требования максимизации качества речевых сигналов и минимизации вычислительной сложности выдвигает актуальную проблему выбора рационального числа областей дифференцирования защиты элементов параметров кадров кодеков речи.

Путем использования рационального числа областей можно добиться повышения качества без неоправданного повышения вычислительной сложности и объемов занимаемой вычислительной памяти.

Приведенные в работе исследования являются составной частью технологии повышения качества восстановленного речевого сигнала кодеков речи путем дифференцирования защиты элементов сигнала.

Цель работы – продолжить исследования, начатые в [1, 5, 6]: определить количество областей дифференцирования защиты элементов параметров кадров кодеков речи.

Решение задачи

Использование метода повышения качества речевых сигналов позволяет достичь значений показателя качества речевых сигналов, превышающих результаты, полученные с применением существующих методов. Суть метода заключается в определении структуры кадра речевого сигнала. Под определением 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.

Результирующая структура кадра зависит от начальных условий, среди которых различают:

- тип кодека речевого сигнала;
- модель канала связи;
- семейство помехоустойчивых кодов;
- число фиксированных групп дифференцированной защиты элементов и др.

В данном случае рассмотрим влияние числа

фиксированных групп дифференцированной защиты элементов параметров кадров выходных речевых сигналов кодеров речи.

Моделирование осуществлялось с использованием кодека MELP 2.4 кбит/с, [7].

Канал связи – дискретный симметричный канал без памяти.

Ошибки – независимые, с равномерным распределением. $BER = 1 \cdot 10^{-2}$.

Результаты применения метода при фиксированном числе групп представлены на рис. 1 и в табл. 1 – 3.

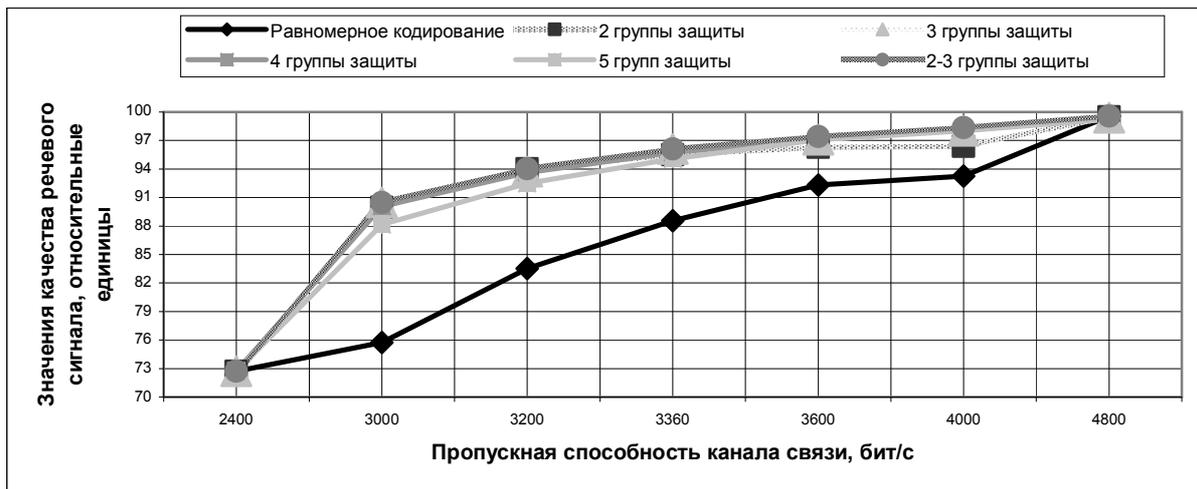


Рис. 1. Качество речевого сигнала для различных вариантов (1 – 5) дифференцирования защиты элементов параметров. $BER = 1 \cdot 10^{-2}$

Данные, полученные в процессе исследований, показывают, что при использовании двух групп наблюдается повышение качества. Значения качества при использовании трех групп имеет различный характер, как в сторону повышения, так и в сторону снижения, относительно вариантов использования структур с фиксированным количеством групп равных двум.

Таблица 1

Результаты применения метода при фиксированном числе групп

Пропускная способность канала связи	Значения качества речевого сигнала в единицах MOS					
	1 гр	2 гр	3 гр	4 гр	5 гр	2-3 гр
2400	2,0093	2,0093	2,0093	2,0093	2,0093	2,0093
3000	2,0632	2,3192	2,3285	2,3214	2,2869	2,3285
3200	2,2034	2,3924	2,3873	2,3851	2,3644	2,3924
3360	2,2944	2,4192	2,4300	2,4233	2,4110	2,43
3600	2,4539	2,4327	2,4474	2,4486	2,4472	2,4539
4000	2,3788	2,4349	2,4642	2,4653	2,4640	2,4705
4800	2,4951	2,4910	2,4936	2,4937	2,4860	2,4951

Фиксированное количество групп, равное четырем и пяти, приводит к снижению или же к незначительному повышению итогового качества восстановленного речевого сигнала в сравнении с более низким количеством групп защиты, при равных значениях полосы пропускания. Данные результаты свидетельствуют о неэффективном использовании доступной полосы пропускания.

В рассматриваемом методе повышения качества восстановленного речевого сигнала кодеров речи условием дифференциации защиты элементов параметров кадров речевого сигнала кодеров речи является снижение степени защищенности элементов параметров при снижении величины влияния на качество восстановленного речевого сигнала.

Учитывая данные условия, при повышении количества групп будет повышаться степень избыточности включенных в рассмотрение помехоустойчивых кодов.

Таким образом, повышение защиты каждого наиболее важного элемента, повлечет за собой снижение защиты совокупности менее важных элементов (табл. 2).

Пример. Фиксированное значение полосы пропускания X бит/с; $1, 2, \dots, i$ – группы дифференцирования защиты; r_1, r_2, \dots, r_i – скорости помехоустойчивых кодов; $1/r_1, 1/r_2, \dots, 1/r_i$ – избыточность,

вносимая помехоустойчивым кодом в соответствии с определенной скоростью; n_1, n_2, \dots, n_i – количество элементов в группе i ; t_k – длительность кадра речи.

Таблица 2

Анализ повышения защиты каждого наиболее важного элемента

Пропускная способность канала связи	1 группа		2 группы		3 группы	
	Размещение бит параметров	Кодовые скорости	Размещение бит параметров	Кодовые скорости	Размещение бит параметров	Кодовые скорости
2400	54	-	54	-	54	-
3000	54	4/5	30 24	5/7	2 30 22	2 3 5 7
3200	54	3/4	45 9	5/7	4 40 10	2 3 5 7
3360	54	5/7	50 4	5/7	2 50 2	2 3 5 7
3600	54	2/3	52 2	2/3	8 45 1	1 2 5 7
4000	54	3/5	51 3	3/5	23 30 1	1 2 5 7
4800	54	1/2	53 1	1/2	1 52 1	1 3 1 2

$X_1 = (n_1/r_1 + n_2/r_2 + n_3/r_3)/t_k$ – для трех групп.

В случае увеличения числа групп также увеличивается количество помехоустойчивых кодов.

$X_2 = (n_n/r_n + n_1/r_1 + n_2/r_2 + n_3/r_3)/t_k$ – для четырех групп (индекс n обозначает новый по отношению к предыдущему варианту структуры с меньшим количеством групп).

Таблица 3

Используемые помехоустойчивые коды

Пропускная способность канала связи	3 группы	4 группы	5 групп
	Размещение бит параметров (Кодовые скорости)	Размещение бит параметров (Кодовые скорости)	Размещение бит параметров (Кодовые скорости)
2400	54	54	54
3000	1 4 25 24	1 3 4 20 26	2 30 22
	(1 2 2 3 5 7)	(1 2 3 5 2 3 5 7)	(2 3 5 7 -)
3200	2 40 3 9	1 3 2 35 13	45 9
	(2 3 5 7 3 4)	(1 2 3 5 2 3 5 7)	(5 7 -)
3360	4 45 3 2	1 2 45 3 3	2 50 2
	(2 3 5 7 3 4)	(1 2 2 3 5 7 3 4)	(2 3 5 7 -)
3600	9 4 40 1	8 3 2 40 1	9 45
	(1 2 2 3 5 7)	(1 2 3 5 2 3 5 7)	(1 2 5 7)
4000	24 4 25 1	23 3 2 25 1	24 30
	(1 2 2 3 5 7)	(1 2 3 5 2 3 5 7)	(1 2 5 7)
4800	2 49 2 1	1 1 47 4 1	54
	(1 3 1 2 2 3)	(1 4 1 3 1 2 2 3)	(1/2)

Как правило (результаты экспериментальных исследований) $r_n < r_1 < r_2 < r_3$. В данном случае происходит перераспределение элементов параметров речевого сигнала по областям дифференцирования элементов параметров кадров речевых сигналов кодексов речи.

Определенное количество элементов перемещается в область n_n . Так как $1/r_n > 1/r_1 > 1/r_2 > 1/r_3$, выделение избыточности для каждого из элементов группы n_n вынуждает переместить в область с более низкой степенью защиты некоторое количество элементов. Количество элементов, для которых уровень защиты снижается, вплоть до нуля, превосходит число элементов группы n_n на величину, пропорциональную избыточности вносимой помехоустойчивым кодом r_n .

В некоторых случаях это приводит к снижению качества восстановленных речевых сигналов – значимость совокупности элементов, для которых уровень защиты уменьшился, превысила степень значимости элементов, уровень защиты которых возрос. А в некоторых случаях качество восстановленных речевых сигналов возрастает – элементы, попавшие под защиту кода с большей избыточностью, имеют больший вес для итогового качества речевого сигнала, чем элементы, степень защиты которых снижена.

Оценка вычислительной сложности

Рассмотрим вычислительную сложность используемого метода при различных вариантах группирования.

Тип помехоустойчивого кодирования – сверточное.

Семейство помехоустойчивых 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 $c < b$.

Используемые помехоустойчивые коды приведены в табл. 3 [8, 9].

Так как акцент в данном случае сделан на выбор вариантов возможных скоростей, анализ литературы [8 – 10] позволил выделить по одному коду по скорости из группы наиболее эффективных.

Для снижения скоростей используют так называемое «выкалывание» или «перфорирование» [10]. Используемые перфорированные коды с матрицами выкалывания представлены в табл. 4, 5.

Таблица 4

Характеристики кодов

№	Порождающий многочлен	Скорость кода R	d_{min}	k
1	[117 127 155]	1/3	15	7
2	[133 171]	1/2	10	7

Таблица 5
Использованные перфорированные коды

№	Порождающий многочлен	Исходная скорость кода R1	Матрица выкальвания	Скорость перфорированного кода R2
1	[133 171]	1/2	$\begin{bmatrix} 11 \\ 10 \end{bmatrix}$	2/3
2	[133 171]	1/2	$\begin{bmatrix} 110 \\ 101 \end{bmatrix}$	3/4
3	[133 171]	1/2	$\begin{bmatrix} 1101 \\ 1010 \end{bmatrix}$	4/5
4	[117 127 155]	1/3	$\begin{bmatrix} 10 \\ 11 \\ 11 \end{bmatrix}$	3/5
5	[117 127 155]	1/3	$\begin{bmatrix} 101 \\ 110 \\ 111 \end{bmatrix}$	5/7

Данные коды получаются из стандартных путем наложения матрицы выкальвания на последовательность символов с выхода кодера, что приводит к потере определенных бит. На приемной стороне случайным образом происходит восстановление потерянных бит с использованием обратной матрицы выкальвания. Сложность таких кодов, как и вероятность правильного приема, снижается, вследствие уменьшения минимального кодового расстояния. Преимуществом является возможность использования различных вариантов скоростей без реализации кодов с иными порождающими многочленами.

Сложность сверточного кодера определяется сложностью кодера и декодера. Сложность кодера, в свою очередь, определяется количеством его простейших элементов (числом разрядов в регистре сдвига, сумматоров по модулю 2 и связей сумматоров с разрядами сдвигов). Длина регистра сдвига – нескольких десятков единиц, а каждый сумматор имеет связи приблизительно с половиной разрядов. Сложность кодера линейно зависит от длины кодового ограничения k . Практическая реализация таких устройств не представляет труда.

Сложность декодеров определяется методом декодирования. Оптимальным с точки зрения достоверности восстановленных данных является алгоритм декодирования Витерби [8, 9].

Сложность реализации декодера по количеству выполняемых операций для декодирования одного информационного символа представлена в табл. 6 [11]. В этой таблице K определяет конструктивную

длину кодов (длину кодового ограничения). Зная сложность реализации метода декодирования N_c можно оценить сложность декодирования всего кода:

$$N_{\text{сверт}} = I N_c, \quad (1)$$
где I – число итераций декодирования.

Таблица 6
Сложность алгоритма декодирования Витерби на один информационный элемент

Вид операции	
Сложение	$2 \cdot 2^{k-1} + 5$
Сравнение	2^{k-1}

Оценка объема выделяемой памяти

Основная область применения декодера Витерби – каналы с мягкими решениями, т.е. каналы, в которых на значения входе кодера – вещественные числа. Для дальнейшей обработки они должны быть представлены в цифровой форме. От точности представления зависит сложность аналого-цифрового преобразования, сложность устройств, выполняющих арифметические операции над метриками ребер и путей, объем памяти для хранения метрик.

Если длина пути составляет N символов канала, то максимальное значение метрики путей может достигнуть величины N^2 и для хранения такой метрики нужно отвести $(8 + \log_2 N)$ двоичных ячеек памяти.

Для любого узла решетки существует путь, ведущий в этот узел и отличающийся от пути с наименьшей метрикой не больше чем в nk символах.

Если на некотором ярусе из всех значений метрик вычесть минимальное, это не повлияет на выбор путей в будущем, и, значит, вероятность ошибки декодирования не изменится. Эта операция может выполняться на каждом ярусе или один раз после обработки нескольких ярусов. В результате число бит для хранения метрик каждого пути приблизительно равно $(8 + \log_2 nk)$ и при реальных значениях параметров кодера не превышает 16 бит [8 – 11].

Объем памяти для хранения путей составляет [8 – 11]:

$$V = 10 \cdot k \cdot 2^k \quad (2)$$

Таким образом, рассматривая сложность декодирования и объемы выделяемой на данную процедуру памяти, относительно одного информационного символа, необходимо отметить, что вычислительная сложность и объем выделяемой памяти для определенных кодов и производных от них перфорированных кодов равны. Данные выводы следуют из того, что основной переменной является длина кодового ограничения.

Из вышесказанного также следует, что коды, имеющие различные порождающие многочлены и различные номиналы скоростей, но при этом с равным значением кодового ограничения, также будут

иметь одинаковые вычислительные сложности и затрачивать один и тот же объем памяти.

Для всего семейства выбранных кодов длина кодового ограничения равна 7. Следовательно, искомые характеристики, такие как вычислительная сложность и объем выделяемой памяти, равны.

Исходя из этого использование вместо одного помехоустойчивого кода двух, трех, четырех и более приводит к повышению вычислительной сложности и объемов выделяемой памяти соответственно в два, три, четыре и более раз.

Так же в случае увеличения областей дифференцирования элементов возрастают вычислительные затраты, связанные с разделением общего потока на соответствующее количество групп, их корректного представления, упорядочивания для заданной работы определенного помехоустойчивого кода и т.д. Аналогичные процедуры необходимы и на приемной стороне. Соответственно увеличение общего числа групп защиты приводит к пропорциональному возрастанию сложности на этапе выполнения описанных процедур.

Заключение

В данной работе обоснован выбор количества групп дифференцирования защиты элементов параметров кадров выходных речевых сигналов кодеров речи. Использование ограничения в пределах двух – трех групп позволяет получить значения качества восстановленных речевых сигналов выше аналогичных с использованием иного количества групп.

Повышение количества групп дифференцирования защиты элементов приводит также к повышению вычислительной сложности и объемов выделяемой памяти проектируемой аппаратуры связи пропорционально количеству групп дифференцирования.

Список литературы

1. Янсонс Я.В. Метод повышения качества восстановленного сигнала на основе дифференцированной защиты кадра речевых данных // Радиоэлектронные и компьютерные системы. – 2008. – № 5(32). – С. 116-120.
2. Hamzaoui R., Stankovic V. Rate-based versus distortion-based optimal error protection of embedded codes // Konstanzer Schriften in Mathematik und Informatik. – 2003. – №194.
3. Jebali I., Loizou P., Spanias A. Speech Processing Using Higher Order Statistics // Proc. ISCAS-93, May 1993. – P. 160-163.
4. Spanias A. A Hybrid Model for Speech Synthesis // Proc. ISCAS-90, New Orleans, May 1990. – P. 1521.
5. Стрюк А.Ю., Янсонс Я.В. Методика оценки помехоустойчивости параметров сжатого сигнала кодеков речи // Радиоэлектронные и компьютерные системы. – 2006. – № 5. – С. 198.
6. Стрюк А.Ю., Янсонс Я.В. Методика исследования помехоустойчивости параметров кодеков речи с использованием оценок PESQ // Информационные технологии и компьютерная инженерия. – 2007. – № 1(8). – С. 130.
7. 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.
8. Блейхут П. Теория и практика кодов контролируемых ошибок. – М.: Мир, 1986. – 576 с.
9. Todd K. Moon. Error correction codes (Mathematical methods and algorithms). – Hoboken, New Jersey. – John Wiley & Sons, Inc., 2005. – 804 p.
10. Robert H. Morelos-Zaragoza. The art of error correcting coding. – John Wiley & Sons, Ltd., West Sussex, PO191 UD, England. – 2002. – 233 p.
11. Золотарев В.В., Овечкин Г.В. Сложность реализации эффективных методов декодирования помехоустойчивых кодов // Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: Тез. докл. 12-й Междунар. науч.-техн. конф. – Рязань: РГРТА, 2004. – С. 12-14.

Поступила в редколлегию 11.07.2008

Рецензент: канд. техн. наук, ст. научн. сотр. П.Н. Гроза, Военный институт телекоммуникаций и информатизации НТУУ «КПИ», Полтава.

ОЦІНКА КІЛЬКОСТІ ОБЛАСТЕЙ ДИФЕРЕНЦІЮВАННЯ ЗАХИСТУ ПАРАМЕТРІВ КАДРІВ КОДЕРА МОВИ

Я.В. Янсонс

Показана актуальність основних проблем що розглядаються в дослідженнях. Проведено обґрунтування вибору кількості областей диференціювання захисту елементів параметрів кадрів вихідних мовних сигналів кодерів мови. Показані переваги використання певної кількості областей диференціювання захисту елементів стислого мовного сигналу. Вибір областей здійснюється в методі підвищення якості відновленого мовного сигналу на основі критерію мінімального очікуваного спотворення кадру мови і диференційованого захисту об'єктивно розподілених елементів параметрів кадрів вихідних мовних сигналів кодерів мови. Здійснена оцінка обчислювальної складності застосовуваних завадостійких кодів з сімейства доступних для різних варіантів вибору областей захисту елементів. Проведена оцінка об'ємів оперативної пам'яті, що виділяється, для здійснення процедур декодування завадостійких кодів, використовуємих в методі.

Ключові слова: кодування мови, об'єктивна оцінка якості мови, каналне кодування, якість мовного сигналу.

AMOUNT ESTIMATION OF SPEECH CODER FRAME PARAMETERS DIFFERENTIATION DEFENSE AREAS

Y.V. Yansons

Actuality of basic problems affected in researches is demonstrated. The ground of choice speech coder frame parameters differentiation defense areas is made. Advantages of the use certain amount of compressed speech signal differentiation defense areas are shown. The areas choice is carried out in the method of quality increase the recovered speech signal. This method is based on determination of speech frame structure and differentiated defense of the structured frame, that allows to maximize the quality of received speech signal. The computational complexity estimation of the applied noise-immunity codes from accessible family for the different choice defense areas is carried out. The on-line storage volume estimation for decoding procedures of noise-immunity codes in-use in a method is carried out.

Keywords: speech encoding, objective estimation of speech quality, channel encoding, speech signal quality.

