

АЛГОРИТМЫ ВТОРИЧНОГО СЖАТИЯ РЕЧЕВЫХ СИГНАЛОВ

В данной статье дано описание алгоритмов вторичного сжатия речевых сигналов и приведены результаты их исследования. Первичное сжатие речевого сигнала осуществляется с использованием вейвлет-преобразования.

Ключевые слова: речевые сигналы, вейвлет-преобразование, сжатие.

Введение. При передаче речи по цифровым каналам связи, будь то сотовая или Интернет-телефония, самый важный вопрос – это сколько информации (число бит в единицу времени) придется передавать по каналам, чтобы снабдить пользователя качественной голосовой связью. Ответ на него в каком-то смысле определяет все - стоимость и качество предоставляемых пользователям услуг и аппаратуры, емкость и масштабируемость сети передачи данных и многое другое.

Сжатие речи при ее передаче сокращает объем передаваемых данных, затраты и, благодаря этому, позволяет снижать цены на услуги и привлекать новых пользователей. Именно поэтому повышение эффективности сжатия речевых сигналов является важным вопросом.

Основная часть. С целью повышения эффективности сжатия речевых сигналов необходимо разработать алгоритмы обработки сжатых данных (алгоритмы вторичного сжатия). Для оценки качества восстановленного сигнала используется оценка максимума коэффициента взаимной корреляции между исходным и восстановленным сигналами

$$K_k = \max K_k(j);$$

$$K_j(j) = \frac{\sum_{i=1}^N (S_{i+j}^* - a^*) \cdot (S_i - a)}{\sqrt{\sum_{i=1}^N (S_{i+j}^* - a^*)^2} \cdot \sqrt{\sum_{i=1}^N (S_i - a)^2}}, \quad j = \overline{0, N_2}, \quad N_2 \ll N, \quad (1)$$

где S_i и S_i^* - отсчеты сигналов до и после обработки; a и a^* - среднеарифметические значения; j - параметр, учитывающий задержку восстановленного сигнала, N – объем выборки.

Алгоритмы вторичного сжатия. В алгоритме первичного сжатия речевых сигналов, представленных в цифровой форме, после вейвлет-преобразования, сжатие сигнала осуществляется исключением из последовательности вейвлет-коэффициентов составляющих, значения которых не превысили заданный пороговый уровень. При восстановлении сигнала по усеченному набору данных используется адресная информация об исключенных из рассмотрения коэффициентах. Так как для блоков, состоящих из нескольких последовательно расположенных исключаемых элементов, указываются лишь начальные и конечные адреса, то объем данных после обработки уменьшается. Степень уменьшения объема данных определяет значение коэффициента сжатия.

$$K = \frac{S_c}{S_0}, \quad (2)$$

где S_0 и S_c - объем исходных и сжатых данных.

Экспериментальные исследования алгоритма сжатия, использующего вейвлет-преобразование сигнала, показали высокую эффективность этого алгоритма.

Дальнейшего снижения объема цифровых данных можно достичь компактным описанием (с некоторой погрешностью) последовательности информативных вейвлет-коэффициентов. Такое дополнительное сжатие данных названо вторичным.

Один из способов вторичного сжатия состоит в снижении разрядности описания значений вейвлет-коэффициентов. Для этих целей можно использовать широко известные нелинейные μ – и A – законы компресии, созданные для снижения динамического диапазона отсчетов речевого сигнала.

Нелинейное μ – преобразование отсчетов, полученных после линейного квантования, имеет вид

$$y_{\mu} = \frac{\text{sign}(m)}{\ln(1 + \mu)} \ln \left(1 + \mu \left| \frac{m}{m_p} \right| \right), \quad (3)$$

где m_p - максимальное входное значение оцифрованного звука; m – входное значение оцифрованного звука; μ - константа, обычно 255.

Нелинейное преобразование (3) используется в Северной Америке и Японии. В Европе используется в основном так называемый A – закон преобразования

$$y_A = \begin{cases} \frac{A}{1 + \ln(A)} \cdot \left(\frac{m}{m_p} \right), & \left| \frac{m}{m_p} \right| \leq \frac{1}{A}; \\ \frac{\text{sign}(m)}{1 + \ln(A)} \left(1 + \ln \left(A \cdot \left| \frac{m}{m_p} \right| \right) \right), & \frac{1}{A} \leq \left| \frac{m}{m_p} \right| \leq 1, \end{cases} \quad (4)$$

где A - константа 87,6.

По сути, A – , μ – преобразования – это кодирование, обеспечивающее сжатие отсчетов с учетом психофизических характеристик слуха человека. Таким кодированием можно достичь сжатия около 30% без существенных потерь в качестве сигнала.

Ко второму варианту снижения объема цифровых данных можно отнести следующие методы:

- приближенного описания всей последовательности значащих вейвлет-коэффициентов с помощью аппроксимирующей функции;
- повторного вейвлет-разложения и сжатия последовательности значащих вейвлет-коэффициентов, сохранившихся после первого этапа сжатия.

Оценки эффективности алгоритмов сжатия. Экспериментально исследованы оба варианта вторичного сжатия цифровых данных.

Оценка эффективности сжатия при использовании нелинейных преобразований вейвлет-коэффициентов производилась по следующей схеме.

1. Осуществлялось первичное сжатие отсчетов сигнала в K_{c1} раз.
2. Выполнялась процедура вторичного сжатия со снижением разрядности в K_{c2} раз (при этом общий коэффициент сжатия $K_{c0} = K_{c1} K_{c2}$).
3. Оценивалось значение взаимного коэффициента корреляции K_k (2) восстановленного сигнала с исходным.
4. Определялось значение K_c (без применения вторичного сжатия), обеспечивающее полученное в предыдущем пункте значение K_k .
5. Рассчитывалась оценка выигрыша (проигрыша) от применения вторичного сжатия

$$B, \% = \frac{K_{c0} - K_c}{K_c} \cdot 100. \quad (5)$$

Результаты экспериментов для различных законов компрессии вейвлет-коэффициентов приведены в табл. 1.

Следует отметить, что в таблице приведены данные только для А – закона (4). При использовании μ – закона (3) результаты практически мало отличаются от приведенных для А – закона. Проведено также исследование более простой зависимости $y = \text{sign}(w) \sqrt{|w|}$, которая также обеспечивает дополнительное сжатие вейвлет-коэффициентов w .

Результаты приведенные в табл.1, получены при значении $K_{c2} = 1,6$, что соответствует случаю перекодирования восьмибитных значений в пятибитные. Значения K_k определялись после вторичного сжатия. Значения $K_c(K_k)$ - это коэффициенты первичного сжатия, при которых достигались значения K_k , полученные при вторичном сжатии.

Таблиця 1

K_{C1}	A – закон (μ – закон)				$y = \text{sign}(w) \sqrt{ w }$		
	$K_{C0} = K_{C1}K_{C2}$	K_k	$K_C(K_k)$	$B, \%$	K_k	$K_C(K_k)$	$B, \%$
2	3,200	0,9972	3,64	-12,3	0,9979	4,86	-34,1
4	6,400	0,9957	5,45	17,4	0,9969	5,92	8,1
8	12,800	0,9850	9,15	39,8	0,9902	8,45	51,4
12	19,200	0,9545	13,50	42,2	0,9648	12,20	57,3
16	25,600	0,9118	17,88	43,1	0,9298	16,12	58,8

Оценки $B, \%$ (5) при использовании для снижения разрядности вейвлет-коэффициентов нелинейных преобразований A и $\text{sign}(w) \sqrt{|w|}$ представлены на рис. 1.

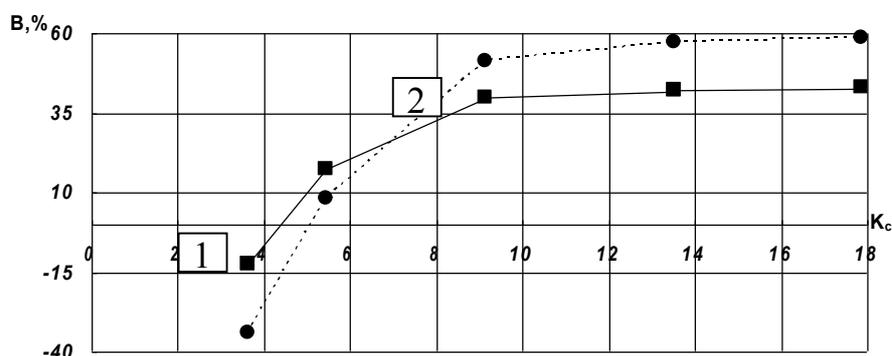


Рис. 1. Зависимости $B, \%$ для A (кривая 1) и $\text{sign}(w) \sqrt{|w|}$ (кривая 2) законов преобразования

Приведенные в табл.1 и на рис.1 результаты получены усреднением оценок по 64 выборкам сигнала с объемом каждой выборки $N=1024$.

Применение сплайн-аппроксимации вейвлет-коэффициентов оправдано при больших значениях коэффициентов сжатия, где выигрыш составляет 15-30%. При малых уровнях сжатия вследствие необходимости сплайн-аппроксимации вейвлет-коэффициентов в большом динамическом диапазоне погрешности сплайн-аппроксимации снижают значение коэффициента взаимной корреляции между восстановленным и исходным сигналами.

Результаты, полученные при использовании вторичного вейвлет-преобразования последовательности вейвлет-коэффициентов после процедуры первичного сжатия, приведены в табл.2. Алгоритм вторичного преобразования полностью повторяет алгоритм первичного преобразования.

Таблиця 2

K_{C2}	K_k	$K_{C0} = K_{C1}K_{C2}$	$K_C(K_k)$	$B, \%$
1,25	0,988	10	8,90	12,36
1,5	0,956	12	10,40	15,53
1,75	0,922	14	11,51	21,63
2	0,896	16	13,43	19,13
2,25	0,854	18	17,33	3,87
2,5	0,847	20	21,14	-5,39
2,75	0,838	22	24,68	-10,89
3	0,821	24	27,01	-11,14

Данные в табл. 2 получены с использованием предложенной выше методики. Речевой сигнал подвергается вейвлет-преобразованию и сжатию (первичному) с $K_{c1} = 8$.

К сохраненным после первичного сжатия вейвлет-коэффициентам была применена процедура вторичного вейвлет-преобразования и сжатия с коэффициентом K_{c2} . После двух циклов обратных вейвлет-преобразований оценивался коэффициент взаимной корреляции восстановленного речевого сигнала с исходным. Показатель $B, \%$ в табл. 2 вычисляется в соответствии с выражением (5).

Выводы

Анализ результатов для нелинейных преобразований A и $\text{sign}(w)\sqrt{|w|}$ показывает, что при малых значениях коэффициента сжатия эффективнее A – закон, а с увеличением коэффициента сжатия предпочтительнее предложенное преобразование $\text{sign}(w)\sqrt{|w|}$. Данный вывод объясняется тем, что при больших коэффициентах сжати сохраняются лишь наибольшие вейвлет-коэффициенты, а более точное представление больших значений аргументов обеспечивается использованием предложенного закона преобразования.

Анализ алгоритмов, включающих процедуру вторичного вейвлет-преобразования, показал, что эффект вторичного сжатия проявляется при $K_{c1} > 6$, при этом коэффициент вторичного сжатия K_{c2} не должен превышать значения 2.

Известные программы архиваторы (Zip, Rar, LZH) позволяют снизить объем данных еще на 15-25%, однако следует учитывать то обстоятельство, что применение этих архиваторов исключает возможность обработки сигнала в реальном времени.

Список литературы

1. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. - М.: ДИАЛОГ-МИФИ, 2003. - 384 с.
2. Storer J. A. Data compression: Methods and Theory / Storer J. A. – Rockville, MD: Computer Science Press. – 1988.
3. Вольфовиц Дж. Теоремы кодирования теории информации / Вольфовиц Дж.; [пер. с англ. Л.Е. Филипповой] – М.: Мир, 1967. – 12 с.
4. Фано Р. Передача информации. Статистическая теория связи / Фано Р.; [пер. с англ.] – М.: Мир, 1965. – 288 с.
5. Бондарев В.Н., Трестер Г., Чернега В.С. Цифровая обработка сигналов: методы и средства: Учеб. пособие для вузов. — Севастополь: Изд-во СевГТУ, 1999.-398с.: ил.
6. Jawerth B., Sweldens W. An overview of wavelet based multiresolution analyses//SIAM Rev.-1994. — 36(3).
7. Zurera M.R., Ferreras F.L., Amores M.P., Bascon S.M., Reyes N.R. New algorithm for translating psycho-acoustic information to the wavelet domain // Signal Processing 81(2001).

Надійшла: 01.06.2011 р.

Рецензент: д.т.н., проф. Давлет'янц О.І.