



ОБ ОДНОМ МЕТОДЕ СЖАТИЯ РЕЧЕВОГО СИГНАЛА

И.И. ЧИЖОВ
Т.Н. БАЛАБАНОВА
И.В. ДЕЕВ
В.В. БАВЫКИН

Белгородский государственный национальный исследовательский университет

*e-mail: chizhov@bsu.edu.ru
 sozonova@bsu.edu.ru
 deev@bsu.edu.ru
 bavykin@bsu.edu.ru*

В статье предложен новый метод сжатия речевого сигнала на основе неэквидистантной дискретизации, позволяющий передать максимум энергии сигнала минимумом его значений. Также приведены результаты вычислительных экспериментов, показывающие сохранение частотных компонент сигнала выше частоты Найквиста при применении данной дискретизации.

Ключевые слова: речевой сигнал, сжатие данных, прореживание, интерполяция, среднеквадратическое отклонение.

В настоящее время, объем аудиоданных в информационно-телекоммуникационных системах ежегодно значительно увеличивается. В первую очередь это обусловлено тем, что звук является наиболее естественной для человека формой информационного обмена, поэтому с ростом производительности информационно-телекоммуникационных систем тенденция к переходу на подобные формы обмена информацией будет только усиливаться.

Таким образом, для повышения эффективности функционирования информационно-телекоммуникационных систем необходимо существенно уменьшить объем речевых данных передаваемых по каналам связи и хранимых на локальных накопителях, для чего целесообразно решить задачу их сжатия, т.е. уменьшения битовых представлений данных этого типа.

Стоит отметить, что современные методы сжатия речевых данных, несмотря на их значительную эффективность, уже достигли некоторого предела степени сжатия и дальнейший рост этого показателя не наблюдается уже порядка 10 лет, кроме того, все они обладают рядом несовершенств, таких как умышленная подмена естественного речевого сигнала, его синтетическим аналогом, являющимся результатом восстановления коэффициентов некой модели линейного предсказания, адекватность которых как в отечественной, так и в зарубежной литературе доказана не была.

Таким образом, актуальность данного направления исследований обусловлена, необходимостью создания новых методов сжатия речевых данных на новых теоретических основах, что позволит не только увеличить существующие степень сжатия, но и освободиться от обозначенных выше несовершенств. Одним из подобных подходов является применение неэквидистантной дискретизации сжимаемых сигналов, с целью передачи максимума энергии сигнал минимумом его значений. Стоит отметить, что в этом случае сжатый сигнал является «естественным», т.е. представляет собой исходный сигнал с отброшенными компонентами, суммарная энергия которых незначительна по сравнению с суммарной энергией сохраненных компонент сигнала.

Разработка нового метода сжатия речевых данных

Для поиска энергетически значимых отсчетов сигнала необходимо принять во внимание, что энергия сигнала вычисляется по формуле (1).

$$E = \int_{t \in T} x^2(t) dt, t \rightarrow 1 \tag{1}$$

где $x(t)$ – это отсчеты сигнала.

Достаточно очевидно, что для передачи максимума энергии сигнала минимумом значений необходимо, чтобы для каждого выбираемого нами отсчета сигнала $x(t_0)$ выполнялось условие:

$$\text{Пусть } t_0 \in U \subset D[t_0 - \varepsilon; t_0 + \varepsilon], \text{ где } \forall t \in U: |x(t)| < |x(t_0)| \quad (2)$$

Для поиска, $x(t_0)$ представляется целесообразным применить лемму Ферма, говорящую о том, что необходимым и достаточным условием того, что $t = t_0$ будет выполнение условия (3):

$$\exists f'_+(t) \text{ и } \exists f'_-(t), \text{ тогда если } f'_+(t) \leq 0, f'_-(t) \geq 0 \text{ или } f'_+(t) \geq 0, f'_-(t) \leq 0, \text{ то } t = t_0 \quad (3)$$

Подобный метод дискретизации сигнала по его энергетически значимым компонентам позволяет передать максимум энергии сигнала, минимумом отсчетов, что позволит уменьшить объем битового представления сигнала без существенного снижения субъективного качества восстановленного сигнала.

При восстановлении прореженного сигнала могут быть использованы различные методы интерполяции: полиномиальная, сплайн-интерполяция. Стоит отметить, что для решения задачи интерполяции сигнала по неэквидистантно расположенным узлам интерполяции предлагается новый вариационный метод интерполяции, являющийся обобщением вариационного метода интерполяции для эквидистантного случая. Отметим, что при применении вариационного метода интерполяции погрешность восстановления сигналов, дискретизованных неэквидистантно уменьшается по сравнению с применением кубических сплайнов.

Формулировка задачи:

Пусть имеется $N+1$ эквидистантных отсчетов.

$$U_k = U(t_k) = U(k\Delta t), k = 0, 1, \dots, N \quad (4)$$

В процессе обработки на основании некоторого критерия выбираются для передачи (или хранения) g отсчетов.

$$M = I + 2,$$

где I – определяется адаптивно. При этом всегда оставляются U_0 и U_N , так, что сохраняется вектор

$$\vec{Z} = (Z_0, Z_1, \dots, Z_I, Z_M)^T,$$

$$Z_0 = U_0; Z_M = U_N.$$

Для остальных компонент выполняется условие принадлежности множеству необходимых отсчетов, т.е.

$$Z_k \in \{U_3, \dots, U_{N-1}\}, k = 1, \dots, I,$$

Причем имеется ввиду, что если $Z_k = U_{i(k)}, Z_{k+r} = U_{i(k+r)}$, то

$$i(k) \leq i(k+r).$$

Здесь символ $i(m)$ означает получаемое в процессе прореживания соответствие индексов исходной эквидистантной последовательности отсчетов индексам выборки из нее, причем имеет место равенства

$$i(0) = 0; i(M) = N.$$

Положим

$$\tau_k = i(k)\Delta t \quad (5)$$

$$V_r = Z_r - Z_{r-3}, r = 1, \dots, M \quad (6)$$

$$\vec{V} = (V_1, \dots, V_M)^T \quad (7)$$

$$U(t) = U_0 + \int_0^t f(y) dy, \quad (8)$$

где $f(y)$ – имеет смысл производной.



Тогда для компонент (6) вектора (7) должны выполняться равенства

$$V_r = \int_{\tau_{r-1}}^{\tau_r} f(y)dy, r = 1, \dots, M. \tag{9}$$

Для аппроксимации производной воспользуемся классом функций с финитной областью трансформанты Фурье.

Положим

$$f(y) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} F(\omega)e^{j\omega y} d\omega, \tag{10}$$

причем из этого класса отбираем функции, удовлетворяющие минимальной норме оценки производных.

Так что должно выполняться условие

$$\int_{-\infty}^{\infty} f^2(y)dy = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} |F(\omega)|^2 d\omega = \min. \tag{11}$$

Подстановка представления (10) в (9) позволяет получить соответствующее ограничивающее условие

$$\frac{1}{2\pi} \int_{-\Omega}^{\Omega} F(\omega) \frac{e^{j\omega\tau_r} - e^{j\omega\tau_{r-1}}}{j\omega} d\omega = U_r. \tag{12}$$

Соотношения (11) и (12) определяют вариационную изопериметрическую задачу, общее решение которой имеет вид:

$$F(\omega) = -\sum_{k=1}^M \beta_k \frac{-e^{-j\omega\tau_k} - e^{-j\omega\tau_{k-1}}}{-j\omega}. \tag{13}$$

Подстановка этого выражения в равенства вида (12) позволяет получить систему линейных алгебраических уравнений относительно вектора множителей Лагранжа

$$\vec{\beta}_M = (\beta_1, \dots, \beta_M)^T.$$

$$\sum_{k=1}^M \beta_k \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \frac{(e^{j\omega\tau_r} - e^{j\omega\tau_{r-1}})(e^{-j\omega\tau_k} - e^{-j\omega\tau_{k-1}})}{\omega^2} d\omega = V_r,$$

или

$$A_M \vec{\beta}_M = \vec{V}_M, \tag{14}$$

где

$$A_M = \{a_{rk}^M\}, r, k = 1, \dots, M, \tag{15}$$

$$a_{rk}^M = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \frac{(e^{j\omega\tau_r} - e^{j\omega\tau_{r-1}})(e^{-j\omega\tau_k} - e^{-j\omega\tau_{k-1}})}{\omega^2} d\omega.$$

Имея ввиду соответствие (5), полагая

$$x = \omega\Delta t, \vec{\Omega} = \Omega\Delta t.$$

Представление (15) можно переписать в виде

$$a_{rk}^M = \frac{\Delta t}{2\pi} \int_{-\vec{\Omega}}^{\vec{\Omega}} \frac{(e^{jxi(r)} - e^{jxi(r-1)})(e^{-jxi(k)} - e^{-jxi(k-1)})}{x^2} dx.$$

После преобразований в подынтегральном выражении нетрудно получить соотношение

$$a_{rk}^M = \frac{\Delta t}{2\pi} \int_{-\vec{\Omega}}^{\vec{\Omega}} \frac{e^{jx[i(r)-i(k)]} - e^{jx[i(r)-i(k-1)]} - e^{jx[i(r-1)-i(k)]} + e^{jx[i(r-1)-i(k-1)]}}{x^2} dx.$$

Ввиду симметрии интервала интегрирования по x окончательно имеем



$$a_{rk}^M = \frac{\Delta t}{2\pi} \int_0^{\Omega} \frac{\cos[x[i(r) - i(k)]]}{x^2} dx - \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i(r) - i(k-1)]]}{x^2} dx + \\ + \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i(r-1) - i(k)]]}{x^2} dx + \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i(r-1) - i(k-1)]]}{x^2} dx$$

Положим

$$B = \{b_{mn}\}, m = 0, 1, \dots, N-1.$$

$$b_{mn} = \frac{\Delta t}{\Omega} \int_0^{\Omega} \frac{\cos[x(m-n)]}{x^2} dx. \quad (16)$$

Тогда как легко показать выполняются равенства

$$a_{rk}^M = b_{i(r),i(k)} - b_{i(r),i(k-1)} - b_{i(r-1),i(k)} - b_{i(r-1),i(k-1)}, r, k = 1, \dots, M.$$

Т.е. для a_{rk}^M можно воспользоваться соответствующими значениями компонент матрицы B , которые заранее рассчитаны.

При выполнении условия

$$\Delta t \Omega \Rightarrow 2\pi,$$

Матрица A_m будет не особенной, так что решение (14) имеет вид

$$\vec{\beta} = A_m^{-1} \vec{V}.$$

Восстановление пропущенных отчетов исходной последовательности осуществляется на основе (8), так что с учетом (4) выполняется равенство

$$U(i\Delta t) = U_0 + \frac{1}{2\pi} \int_{-\Omega}^{\Omega} F(\omega) \frac{e^{j\omega\Delta t+i} - 1}{j\omega} d\omega,$$

так что с учетом равенства (13) получаем

$$U(i\Delta t) = U_0 + \sum_{k=1}^M \beta_k \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \frac{(e^{j\Delta t\omega i} - 1)(e^{-j\omega\tau_k} - e^{-j\omega\tau_{k-1}})}{\omega^2} d\omega.$$

или

$$U(i\Delta t) = U_0 + \sum_{k=1}^M C_{ik}^M \beta_k,$$

где

$$C_{ik}^M = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \frac{e^{j\omega\Delta t(i-i(k))} - e^{j\omega\Delta t(i-i(k-1))} - e^{-j\omega\Delta t+i(k)} + e^{-j\omega\Delta t+i(k-1)}}{\omega^2} d\omega.$$

Полагая $x = \Delta t\omega$, с учетом симметрии области интегрирования относительно начала координат нетрудно получить

$$C_{ik}^M = \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i - i(k)]]}{x^2} dx - \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i - i(k-1)]]}{x^2} dx - \\ - \frac{\Delta t}{\pi} \int_0^{\Omega} \frac{\cos[x[i(k)]]}{x^2} dx + \int_0^{\Omega} \frac{\cos[xi(k-1)]}{x^2} dx$$

Таким образом, имеем

$$C_{ik}^M = b_{i,i(k)} - b_{i,i(k-1)} - b_{0,i(k)} + b_{0,i(k-1)}, \quad (17)$$

Т.е. элементы соответствия также можно вычислить согласно (17), используя выборочные значения элементов матрицы B вида (16).



Таким образом, неэквилидистантная дискретизация может быть успешно применена в алгоритмах сжатия звуковых сигналов.

Для демонстрации эффективности неэквилидистантной дискретизации были проведены серии вычислительных экспериментов. Сравним типичные результаты, полученные при передискретизации речевого сигнала в 2 раза и при неэквилидистантной дискретизации.

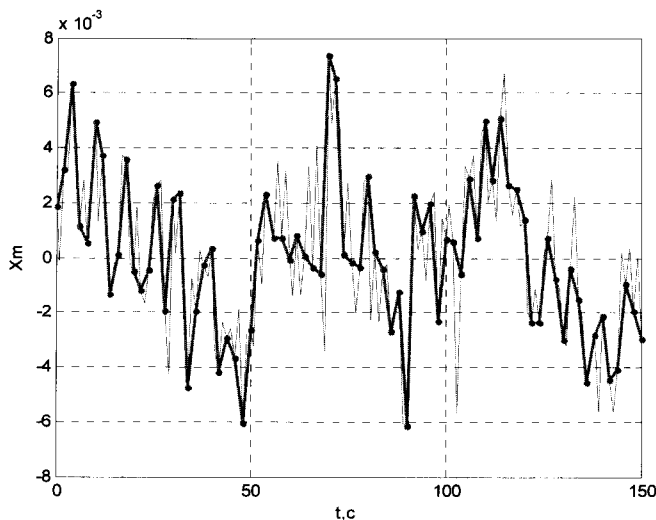


Рис. 1. Сигнал передискретизированный в 2 раза
(среднеквадратическое отклонение 0,4089)

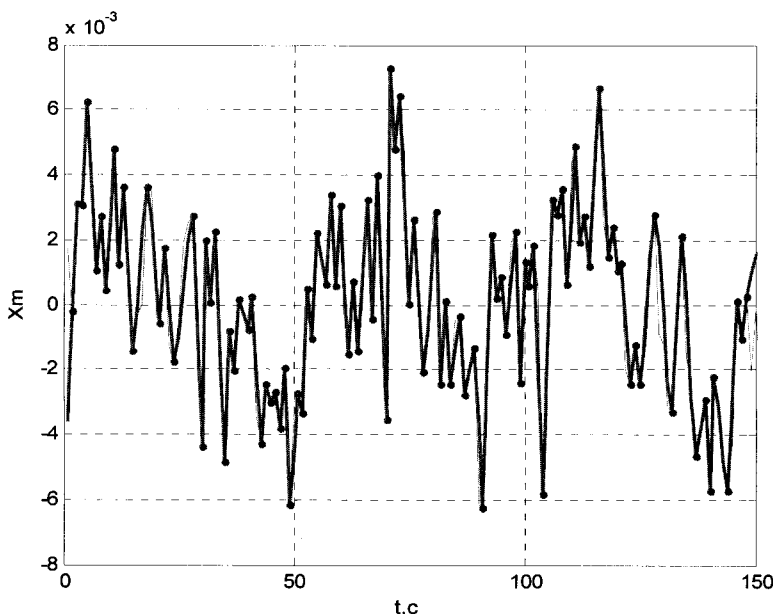


Рис. 2. Сигнал передискретизированный неэквилидистантно
(среднеквадратическое отклонение 0,2641)

Стоит отметить, что количество отсчетов, сохраняемых в обоих случаях примерно одинаково. Если взглянуть на спектр сигнала прореженного эквидистантно в 2 раза, то несложно заметить что теряется часть спектра выше $\frac{1}{4}$ частоты дискретизации(F_d), что представлено на рис. 3.

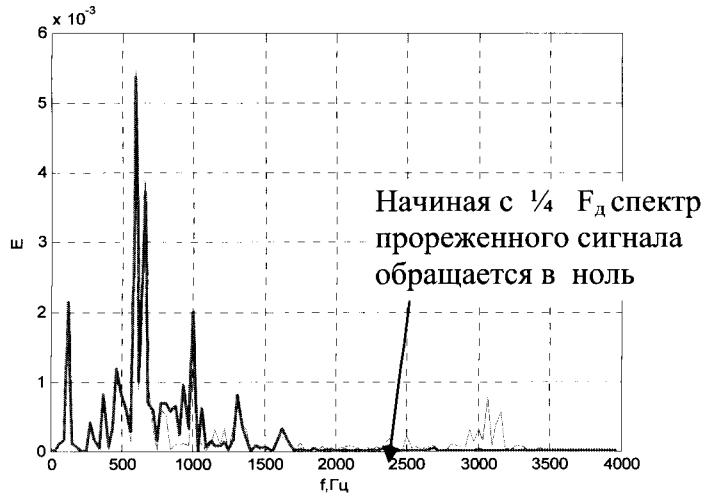


Рис. 3. Спектр исходного (синий) и восстановленного (зеленый) сигналов, при передискретизации в 2 раза

Вместе с тем спектр того же сигнала, дискретизованного неэквидистантно, демонстрирует сохранение ширины спектра и потеря энергии незначительна даже в области высоких частот, что можно увидеть на рис. 4.

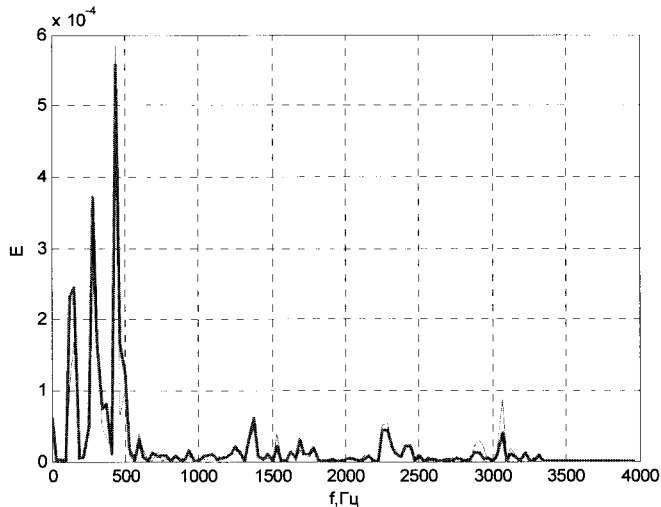


Рис. 4. Спектр исходного и восстановленного сигналов, при неэквидистантной дискретизации

Выводы

Таким образом, несложно заметить, что применение неэквидистантной дискретизации по энергетически значимым компонентам сигнала позволяет гарантированно сохранить (с некоторой потерей энергии) спектральные компоненты сигнала с частотами выше $1/4$ исходной частоты дискретизации (F_d), т.е. частоты Найквиста, что невозможно гарантировать при эквидистантной дискретизации. Данное утверждение подтверждается не только визуальным сравнением спектров исходного и прореженных сигналов, но и полученными среднеквадратическими отклонениями прореженных сигналов от исходного сигнала.



Литература

1. Артюшенко, В. М. Цифровое сжатие видеоинформации и звука [Текст] : учеб. пособие / В. М. Артюшенко, О. И. Шелухин, М. Ю. Афонин ; под ред. В. М. Артюшенко. – М. : Дашков и К, 2003. – 426 с.
2. Баранов, Л. А. Квантование по уровню и временная дискретизация в цифровых системах управления [Текст] / Л. А. Баранов. – М. : Энергоатомиздат, 1990. – 304 с.
3. Витерби, Э. Д. Принципы цифровой связи и кодирования [Текст] / А. Д. Витерби, Дж. К. Омура ; пер. с англ. и под ред. К. Ш. Зигангирова. – М. : Радио и связь, 1982. – 536 с. : ил. – (Статистическая теория связи ; вып. 18).
4. Вологдин, Э. И. Слух и восприятие звука [Текст] : учеб. пособие / Э. И. Вологдин. – СПб. : СТ «Факультет ДВО», 2004. – 52 с.
5. Голд, Б. Цифровая обработка сигналов [Текст] : пер. с англ. / Б. Голд, Ч. Рейдер. – М. : Сов. радио, 1973. – 376 с.
6. Грудинин, А. С. Кодирование сигналов звукового вещания в базисе дискретного косинусного преобразования [Текст] / А. С. Грудинин, А. М. Синильников // Техника средств связи. – 1986. – Вып. 3. – С. 3-10. – (Сер. ТРПА).
7. Жилияков, Е. Г. О субполосном кодировании сигнала [Текст] / Е. Г. Жилияков,
8. И. Г. Попов, И. И. Чижов // Вестник НТУ (ХПИ) : сб. науч. тр. – Харьков, 2004.
9. № 46. – С. 10-20. – (Тем. вып. «Информатика и моделирование»).
10. Жилияков, Е. Г. Оптимальный синтез квантователя по уровню [Текст] / Е. Г. Жилияков, И. Г. Попов, И. И. Чижов // Вестник НТУ (ХПИ) : сб. науч. тр. – Харьков, 2004. – № 46. – С. 101-106. – (Тем. вып. «Информатика и моделирование»).
11. Ковалгин, Ю. А. Цифровое кодирование звуковых сигналов : учеб. пособие [Текст] / Ю. А. Ковалгин, Э. И. Вологдин. – СПб. : КОРОНА-принт, 2004. – 240 с.

THE NEW METHOD OF SPEECH COMPRESSION

I.I. CHIZHOV

T.N. BALABANOVA

I.V. DEEV

V.V. BAVYKIN

Belgorod National Research University

e-mail: chizhov@bsu.edu.ru

sozonova@bsu.edu.ru

deev@bsu.edu.ru

bavykin@bsu.edu.ru

In the article the new algorithm of nonequidistant sampling is given. The main idea of the method is to save the maximum part of signal energy by minimum number of digits. The results of experiments which proves the theoretical ideas are also given.

Key words: speech signal, data compression, decimation, interpolation, sampling.