

СПОСОБ СЖАТИЯ РЕЧЕВЫХ СИГНАЛОВ, ОСНОВАННЫЙ НА ДИСКРЕТНОМ ПРЕОБРАЗОВАНИИ УОЛША

Ю.В. Данюк, к.т.н. К.А. Бохан, к.в.н. А.Н. Явтушенко
(представил д.т.н., проф. А.В. Королёв)

Предложен способ сжатия речевых сигналов на основе дискретного преобразования Уолша.

Постановка проблемы. Человеческая речь представляет собой исторически первый и наиболее естественный вид связи. Речевое общение между людьми достигло такой степени совершенства, что в произносимых словах и фразах возможно передать гораздо больше информации, чем, например, в печатном тексте. Поэтому столь важное место в технике связи занимает передача речевых сообщений. При этом, однако, требуется разрешить противоречие между ограниченной пропускной способностью каналов связи и высокой информационной емкостью речевых сигналов, предназначенных для передачи по таким каналам. Поэтому в теории и технике связи стала актуальной задача компактного представления речи с целью доведения ее информационного объема до значений, соизмеримых с объемом печатного текста, но без потери для потребителя информационной содержательности.

Известно, что человек может воспринимать смысловую "информацию" со сравнительно малой скоростью порядка десятков или сотен бит в секунду, а частота механических колебаний артикуляционных органов при произношении звуков речи невелика и по грубым оценкам не превышает нескольких десятков герц [1]. Это позволяет предположить, что для обмена "речевой информацией" достаточно располагать пропускной способностью канала, не превышающей 100 бит/с [2]. При таких оценках эффективность, например, цифровой системы передачи со стандартной импульсно-кодовой модуляцией (скорость 64 кбит/с) оказывается весьма низкой. Это свидетельствует об актуальности сжатия речевых сигналов с целью уменьшения объемов данных для передачи и хранения.

Анализ литературы свидетельствует о множестве подходов в области сжатия речевых сигналов [3 – 5]. Наиболее эффективные алгоритмы сжатия речевых сигналов строятся на основе применения аппарата математических преобразований. В результате математического преоб-

разования исходному набору данных, представленному последовательностью во временном пространстве, ставится в однозначное соответствие набор данных в частотном пространстве. Причем свойства полученного набора таковы, что обеспечивают более компактное представление его в цифровом виде [2 – 4].

Как правило, для получения отображения сигнала в частотной области используются ортогональные преобразования, классическим представителем которых является преобразование Фурье. Однако, преобразованию Фурье присущ ряд особенностей, снижающих эффективность применения.

Во-первых, преобразование Фурье обладает высокой вычислительной сложностью. Для выполнения быстрого преобразования Фурье с прореживанием по частоте, для сигнала, имеющего N отсчетов, необходимо выполнить $N \log_2 N$ операций комплексного умножения и сложения [4]. Вычислительная сложность вынуждает разбивать сигнал на блоки небольшой длины, что приводит к разрыву статистических взаимосвязей, имеющихся между элементами сигнала, а также прибегать к округлению с той или иной степенью погрешности полученных в результате преобразования вещественных значений, использованию окон для компенсации потерь энергии.

Во-вторых, число коэффициентов ряда, в который раскладывается исходный сигнал, в общем случае бесконечно. Для многих реальных сигналов аппроксимация по ограниченному числу коэффициентов дает слишком большую погрешность обратного преобразования [4].

Ввиду того, что в литературе о применении дискретного преобразования Уолша именно для сжатия речевых сигналов упоминается редко [5], а также для избежания указанных выше недостатков, предлагается использование дискретного преобразования Уолша. Таким образом, стоит задача разработки метода сжатия речевых сигналов на основе дискретного преобразования Уолша.

Целью статьи является разработка способа сжатия речевых сигналов на основе дискретного преобразования Уолша.

Разработка способа сжатия речи на основе дискретного преобразования Уолша. Дискретное преобразование Уолша позволяет представить исходный сигнал $x(t)$ при длине преобразования N в виде [2]:

$$x(t) = \sum_{k=0}^N d_k \text{wal}_w(k, t), \quad (1)$$

где

$$\text{wal}_w(k, t) = \begin{cases} \text{cal}(s_k, t), & k - \text{четное}; \\ \text{sal}(s_k, t), & k - \text{нечетное}; \end{cases} \quad (2)$$

s_k – частота функции $wal_w(k, t)$, определяемая как

$$s_k = \begin{cases} 0, & k=0; \\ k/2, & k \text{ – четное}; \\ (k+1)/2, & k \text{ – нечетное}. \end{cases}$$

Если $wal_w(k, t)$ выразить через составляющие sal и cal , то выражение (1) принимает вид

$$x(t) = a_0 wal_w(0, t) + \sum_{k=1}^N [a_k cal(k, t) + b_k sal(k, t)], \quad (3)$$

где a_0 – постоянная составляющая; a_k – коэффициенты при cal функциях; b_k – коэффициенты при sal функциях.

Из приведенных соотношений видно, что представление сигнала в виде ряда Уолша аналогично представлению их в виде ряда Фурье, что следовало и ожидать из-за сильного сходства между синусоидами и функциями Уолша (рис. 1).

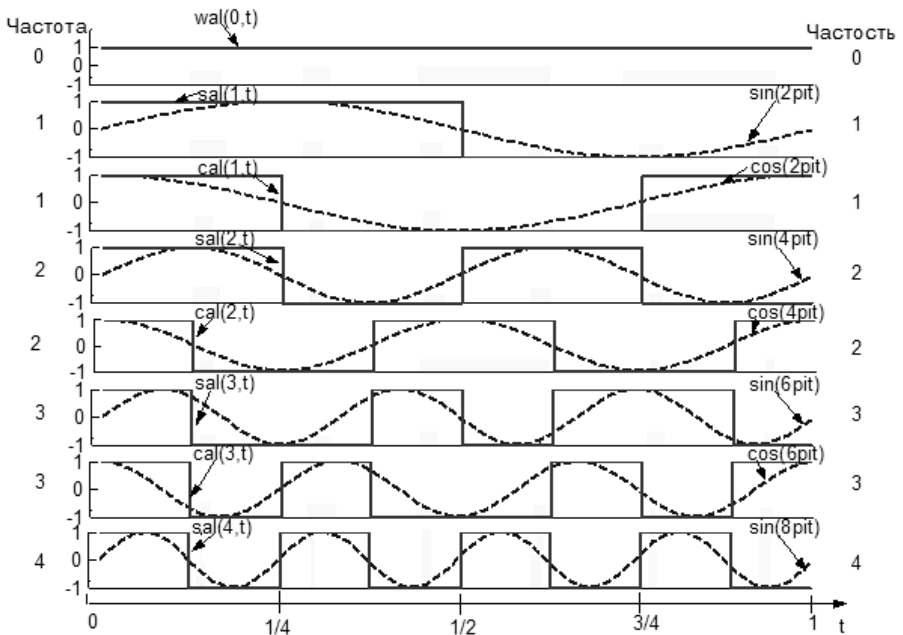


Рис. 1. Функции Уолша и гармоники Фурье

Помимо преобразования Фурье существуют его производные – косинусное и синусное преобразование, которые в качестве базиса исполь-

зуют \sin или \cos соответственно названию. Возникает вопрос о возможности такого подхода в преобразовании Уолша. Наиболее выгодным с точки зрения сжатия является использование только половины отсчетов трансформанты – коэффициентов только при sal или cal функциях.

Для определения возможности такого подхода необходимо определить энергетические потери при использовании только коэффициентов при sal или cal функций. С этой целью произведен подсчет распределения энергии между функциями Уолша. В процентном соотношении указанное распределение представлено на рис. 2.

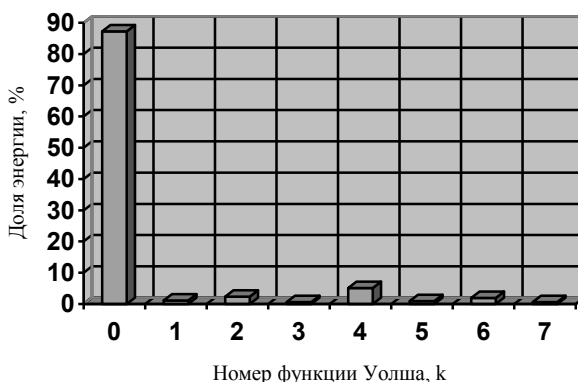


Рис. 2. Распределение энергии между sal и cal составляющими (при $N = 8$)

На рис. 2 по оси абсцисс расположены функции sal или cal , причем нечетным номерам соответствуют cal функции, а четным – sal . Из данного рисунка видно, что большая часть энергии в среднем порядка 91% сконцентрирована в постоянной составляющей и коэффициентах при cal составляющих (все нечетные составляющие на графике). Такое обстоятельство говорит о том, что энергетические потери восстановленного сигнала составят не более 9%. Таким образом, существует возможность использования только cal функций для представления сигнала в частотно-временной области.

При таком подходе разложение в ряд Уолша примет вид

$$x(t) = a_0 \text{wal}_w(0, t) + \sum_{k=1}^N a_k \text{cal}(k, t). \quad (4)$$

При обратном преобразовании в качестве коэффициентов при sal функциях используются нули.

Среднеквадратичное отклонение (СКО) ε определяется соотношением

$$\varepsilon = 1/N \cdot \sqrt{|x(t) - x(t)'|^2}, \quad (5)$$

где $x(t)$ – исходный сигнал; $x(t)'$ – восстановленный сигнал.

Усредненное экспериментальное значение СКО при анализе 120 отрезков речи с использованием преобразования (4) равняется 8,642, что в процентном соотношении от динамического диапазона составляет 3,389%. Значение СКО можно уменьшить за счет передачи информации о поведении коэффициентов при *sal* функциях. С этой целью предлагается в качестве коэффициентов при *sal* функциях использовать среднеарифметическое значение коэффициентов при *sal* функциях в каждом блоке. Тогда СКО согласно выражения (5) составит уже 7,3 (2,86% от динамического диапазона).

Выводы. Таким образом, проведенное исследование свидетельствует о возможности эффективного применения дискретного преобразования Уолша для сжатия речи. Предложенный способ сжатия на основе дискретного преобразования Уолша позволяет произвести сжатие речевого сигнала в два раза со 100% разборчивостью слов.

ЛИТЕРАТУРА

1. Звездин В.С. *Речевое общение человека и ЭВМ.* – М.: Знание, – 1980. – 64 с.
2. Ахмед Н., Рао К.Р. *Ортогональные преобразования при обработке цифровых сигналов.* – М.: Связь, 1980. – 248 с.
3. Залманзон Л.А. *Преобразование Фурье, Уолша, Хаара и их применение в управлении, связи и других областях.* – М.: Наука, 1989. – 496 с.
4. Бондарев В.Н., Трестер Г., Чернега В.С. *Цифровая обработка сигналов: методы и средства.* – Х.: Конус, 2001. – 398 с.
5. *Вокодерная телефония. Методы и проблемы / Под ред. А.А. Пирогова.* – М.: Связь, 1974. – 536 с.

Поступила 18.02.2004

ДАНИЮК Юрий Владимирович, адъюнкт ХВУ. В 2000 году окончил ХВУ. Область научных интересов – обработка и передача информации.

БОХАН Константин Александрович, канд. техн. наук, нач. лаб. ИВЦ ХВУ. В 1999 году окончил ХВУ. Область научных интересов – обработка и передача информации.

ЯВТУШЕНКО Анатолий Николаевич, канд. военных наук, доцент. Окончил ХВВКУ в 1971 году. Область научных интересов – обработка информации.