

УДК 621.391

ИССЛЕДОВАНИЕ ХАРАКТЕРИСТИК РЕЧЕВОГО КОДЕРА ДЛЯ БЫСТРОГО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ НА ОСНОВЕ АЛГОРИТМА МАЛЛА

М.С. АНТОНЕНКО, Т.М. ПЕЧЕНЬ

*Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь**Поступила в редакцию 7 ноября 2020*

Аннотация. Рассмотрен метод компрессии, основанный на применении дискретного вейвлет-преобразования для обработки речевого сигнала. Предложена модель речевого кодера для быстрого вейвлет-преобразования, основанного на кратномасштабном анализе с использованием алгоритма Малла и исследованы его характеристики.

Ключевые слова: вейвлет-преобразование, алгоритм Малла, речевой кодер, компрессия.

Введение

Основной вид коммуникации между людьми – это передача речи. По этой причине актуальным остается совершенствование методов цифровой обработки и передачи речевых сообщений [1]. Новые средства обработки сигналов и соответствующие специализированные инструментальные программные среды создают возможности появления эффективных систем передачи и обработки речевых сообщений [2]. Компрессия и кодирования речи позволяют решать сложные проблемы по минимизации числа бит необходимых для передачи сигнала.

В настоящее время одной из важных задач является создание систем низкоскоростной передачи с высоким качеством восприятия сигнала, способных корректно функционировать в реальных условиях окружающей среды [3]. Решение проблемы компрессии речевого сигнала особую роль играет в таких сферах как Интернет-телефония, сотовая связь, запись и хранение речевых сообщений в специальных и портативных устройствах [4].

При низких скоростях передачи информации параметрическое кодирование более эффективное. В данном виде кодирования в качестве параметров используют различные характеристики представления сигнала в частотной области – спектральные кодеры, использующие вейвлет-преобразование [5].

Одномерное вейвлет-преобразование

Малла предложил эффективный алгоритм реализации дискретного вейвлет-преобразования, который представляет собой классическую схему обработки: входной одномерный сигнал раскладывается на высокочастотные и низкочастотные компоненты с использованием пары ортогональных фильтров.

В основе преобразования заложено, что любую функцию $s(t)$ можно рассматривать на любом m' -уровне разрешения. На этом уровне между ее усредненными значениями и флуктуациями вокруг средних значений для разделения функции используем формулу

$$s(t) = \sum_{k=-\infty}^{\infty} C_{m'}, k(t) + \sum_{m=m'}^{\infty} \sum_{k=-\infty}^{\infty} D_m, k\psi_m, k(t). \quad (1)$$

Как видно из (1) при бесконечных пределах первая сумма стремится к нулю и поэтому может быть опущена. В результате получаем «чистое» вейвлет-преобразование. При решении практических задач цифровые данные представлены в виде конечного набора отсчетов. Таким

образом, наилучший уровень разрешения определяется интервалом, в котором содержится один отсчет. В результате суммирование будет выполняться в конечных пределах. Значение $m = 0$ как правило принимается для наилучшего уровня разрешения. Принятая форма вейвлетов $\varphi_{m,k} = 2m / 2\varphi(mt - k)$ усредняет отсчеты. Это происходит при уменьшении значений $m = 0, -1, -2, \dots$. Для удобства расчетов обычно исключается использование отрицательных индексов масштабирования. В этом случае знак « \leftarrow » обычно вводится непосредственно в функции вейвлетов $\varphi_{m,k} = 2 - m/2\varphi(2 - mt - k)$. После выполнения данной операции вычисляются вейвлет-коэффициенты для значений $m > 0$.

Такой кратномасштабный анализ для случая с последовательным увеличением значений m способствует получению алгоритма, состоящего с быстрых итерационных вычислений:

$$C_{m+1}, k = \sum_n h_n C_m, 2k + n, \quad (2)$$

$$D_{m+1}, k = \sum_n g_n C_m, 2k + n, \quad (3)$$

$$C_0, k = \int_{k\Delta t}^{(k+1)\Delta t} s(t) \varphi(t - k) dt. \quad (4)$$

Алгоритм Малла представляет собой пирамидальный алгоритм вычисления вейвлет-коэффициентов. В этом случае вейвлеты Добеши обеспечивают наиболее приемлемые условия для обработки речевого сигнала. Уравнение (4) используются для случаев если известна аналитическая форма функции $s(t)$.

Поясним суть операций, выполняемых формулами (2) и (3). На первой итерации преобразования цифровой фильтр h_n из сигнала $s_k = C_0$, k выделяет низкие частоты, а октавный фильтр g_n выделяет верхние частоты. На выходе фильтра h_n отсутствует верхняя половина частот. Выполняется децимация выходного сигнала: осуществляется по формуле (2) [4]. На выходе фильтра g_n освобождается место в области низких частот, и аналогичное прореживание выходного сигнала приводит к транспонированию верхних частот на освободившееся место [5].

Таким образом, каждый из выходных сигналов несет информацию о своей половине частот, при этом выходная информация представлена таким же количеством отсчетов, что и входная [3].

Восстановление сигнала

Поскольку в формулах (2), (3) вместо базисных функций используются фильтры, то обратные преобразования, т.е. последовательную сборку сигнала от больших m к малым и реконструкцию сигналов по значениям его вейвлет-коэффициентов с любого уровня разрешения, имеет смысл также выразить через фильтры реконструкции

$$C_{m-1} = \sum_{n \in I} C_{m,n} hr(k) - 2n + \sum_{n \in I} D_{m,n} gr(k) - 2n. \quad (5)$$

Алгоритм вычислений по формуле (5) является обратным алгоритму декомпозиции. Таким образом, аппроксимируются коэффициенты C_m и D_m в 2 раза меньший шаг дискретизации с двукратным увеличением частоты Найквиста и восстановлением спектра коэффициентов C_m в низкочастотную часть нового главного диапазона спектра C_{m-1} , а спектра коэффициентов D_m в высокочастотную часть спектра D_{m-1} . Далее выполняется расстановка нулевых значений между коэффициентами C_m и D_m , фильтрация полученных массивов низкочастотным $hr(k)$ и высокочастотным $gr(k)$ фильтрами реконструкции, и сложением результатов фильтрации. Модули частотных характеристик фильтров $hr(k)$ и $gr(k)$ должны повторять модули

частотных характеристик фильтров $hr(k)$ и $gr(k)$ [3]. Однако, фильтры декомпозиции $hr(k)$ и $gr(k)$ являются односторонними и фазосдвигающими, и при реконструкции коэффициентов $C_m - 1$ этот сдвиг фазы должен быть исключен. Этого можно достичь реверсом значений коэффициентов фильтров декомпозиции по следующей формуле [4]

$$hr(k) = \text{reverse}(h(k)), gr(k) = \text{reverse}(g(k)). \quad (6)$$

Точность преобразования сигналов зависит от потерь информации при выполнении прореживания спектров. Главным образом, эти потери возникают на срезах полос пропускания фильтров низких и высоких частот, крутизна которых зависит от порядка фильтров, их согласованности, и типа вейвлетных функций [2].

Обязательным условием преобразования сигнала является его задание количеством отсчетов, равном $N = 2m$, где значение $m \geq 1$ определяет максимально возможное число уровней декомпозиции сигнала при целочисленных значениях кратности сдвигов операторов фильтров количеству отсчетов вейвлетных коэффициентов на каждом уровне декомпозиции. Требованием является выполнение условия количество отсчетов сигнала равного $N = 2m$ [4].

Структурная схема декомпозиции и реконструкции речевого сигнала

Структурные схемы декомпозиции и реконструкции сигнала алгоритма Малла представлены на рис. 1.

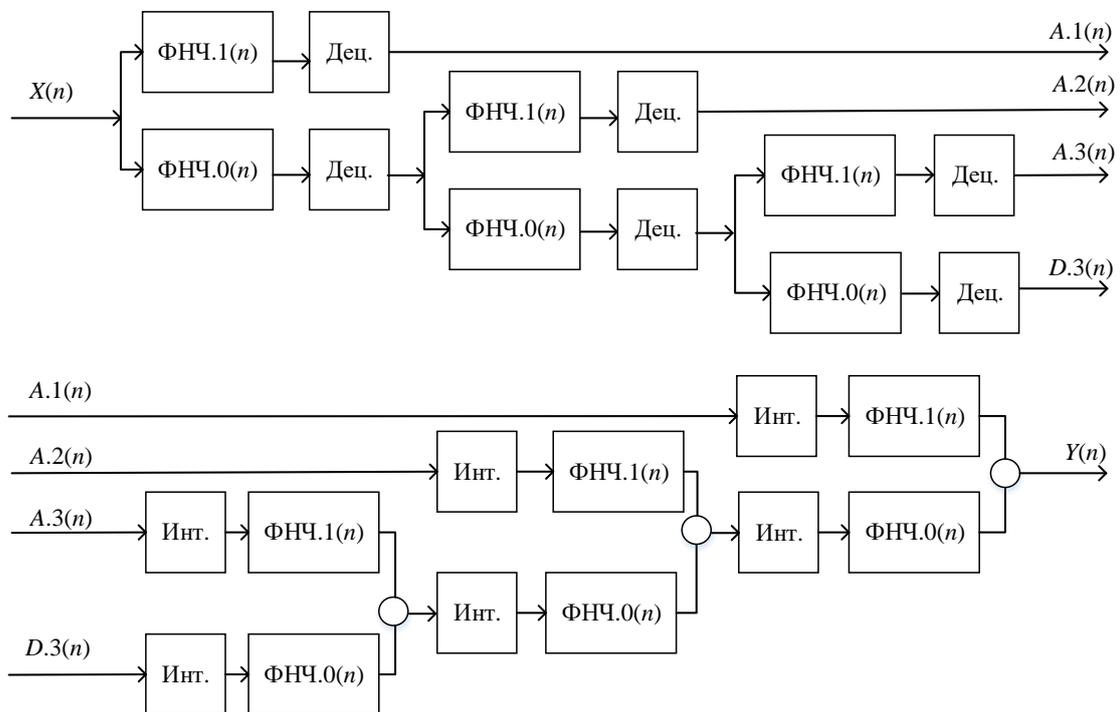


Рис. 1. Структурные схемы декомпозиции и реконструкции сигнала

Алгоритм основан на четырех операциях: фильтрации, децимации, интерполяции и суммирования. Функционально фильтры представляют собой низкочастотные и высокочастотные фильтры с конечной импульсной характеристикой (КИХ-фильтры). Дециматор (Дец.) – это устройство, которое позволяет уменьшать частоту дискретизации в 2 раза. Интерполятор (Инт.) повышает частоту дискретизации в 2 раза.

Моделирование процесса кодирования декодирования речевого сигнала

Моделирование работы системы кодер/декодер производилось в специализированной программе MATLAB R2016b с использованием реальных речевых сигналов, ранее записанных в файл. Схема моделирования работы системы кодер/декодер представлена на рис. 2.

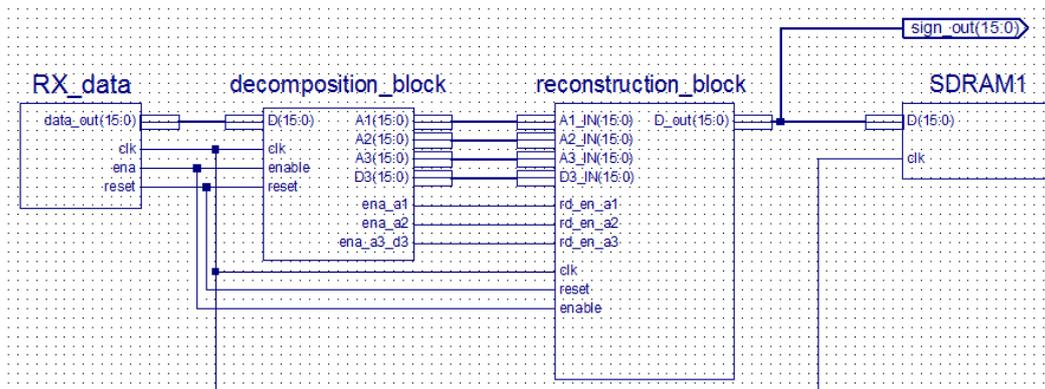


Рис. 2. Схема моделирования работы системы кодер/декодер

Как видно из рис. 2 в состав схемы моделирования работы системы кодер/декодер входят следующие блоки:

- RX_data – блок для генерации управляющих сигналов;
- decomposition_block – блок декомпозиции;
- reconstruction_block – блок реконструкции сигнала;
- SDRAM1 – блок сохранения результата работы сигнала.

Входной блок моделирования генерирует сигналы «clk», «enable», «reset», а также на вход основной схемы подает входные отсчеты. Отсчеты сигнала читаются из файла, который был создан в MATLAB R2016b. Для моделирования был взят wav-аудио файл с записью голоса. С помощью функции «wavread» были считаны значения сигнала. На рис. 3 представлена осциллограмма исходного речевого сигнала.

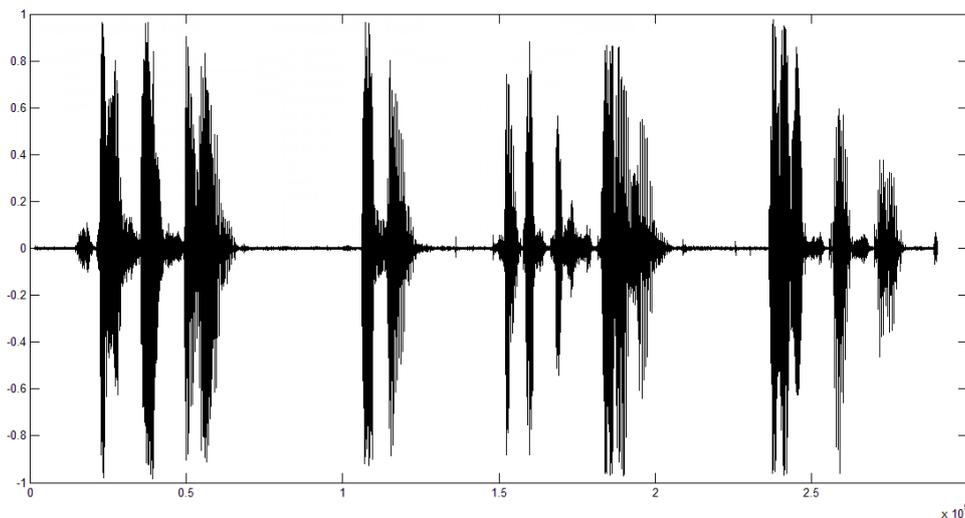


Рис. 3. Осциллограмма исходного речевого сигнала

Файл, содержащий отсчеты сигнала, называется «speech_data.txt». Выходной блок моделирования SDRAM записывает отсчеты восстановленного сигнала в файл. Отсчеты сигнала записываются в файл «SIGNAL_OUT.txt».

Поскольку амплитуда сигнала находится в пределах от -1 до 1 , поэтому необходимо умножить каждый отсчет на 32767 , для перевода в диапазон от -32767 до 32767 . Затем числа преобразовываем в двоичный вид и записываем в файл.

Рассмотрим листинг *m*-файла (wav_ADC) данной процедуры:

```

fSpData = fopen('speech_data.txt','w');
Y = wavread('D:/test.wav');
Z = 0;
for ii = 1:290304
    tmp = round(Y(ii) * 32767);
    if tmp < 0
        Z = dec2bin(abs(tmp),15);
        Z = strcat('1',Z);
        fprintf(fSpData,'%s\n',Z);
    else
        Z = dec2bin(tmp,16);
        fprintf(fSpData,'%s\n',Z);
    end;
    Z = 0;
end;
fclose(fSpData);

```

Затем пропустив сигнал через разработанную схему и, сохранив в файл результаты в блоке SDRAM1, обратно преобразуем в wav-файл. Листинг *m*-файла (wav_DAC) для преобразования обратно в wav-файл.

```

fSpData = fopen('SIGNAL_OUT.txt','r')
dout = 0;
for ii = 1:290304
    strTmp = fscanf(fSpData,'%s',[1]);
    decTmp = bin2dec(strTmp(2:16));
    decTmp1 = decTmp/32767;
    if strTmp(1) == '1'
        dout(ii) = decTmp1 *(-1);
    else
        dout(ii) = decTmp1;
    end;
end;
fclose(fSpData);

```

На рис. 4 представлена осциллограмма восстановленного сигнала.

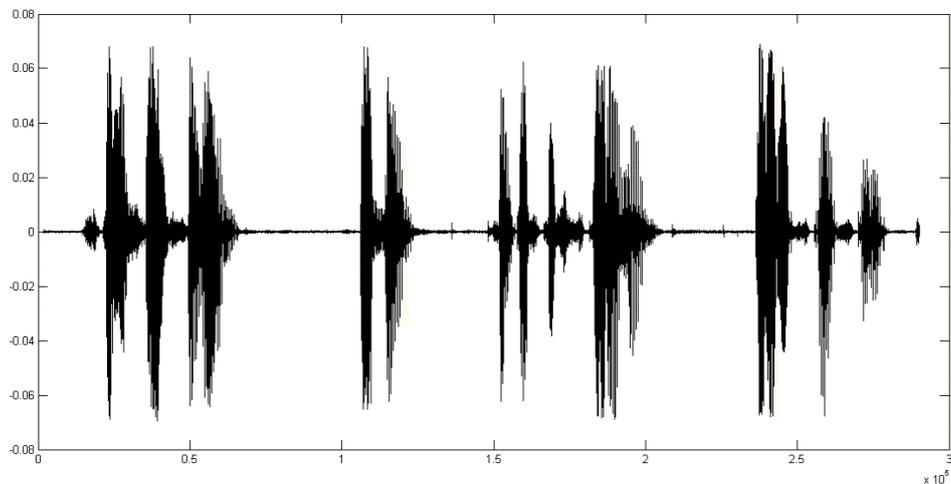


Рис. 4. Осциллограмма восстановленного сигнала

Для оценки качества восстановления сигнала оценим разность вход/выход сигналов и корреляция вход/выход.

Разность исходного сигнала и полученного представлена на рис. 5.

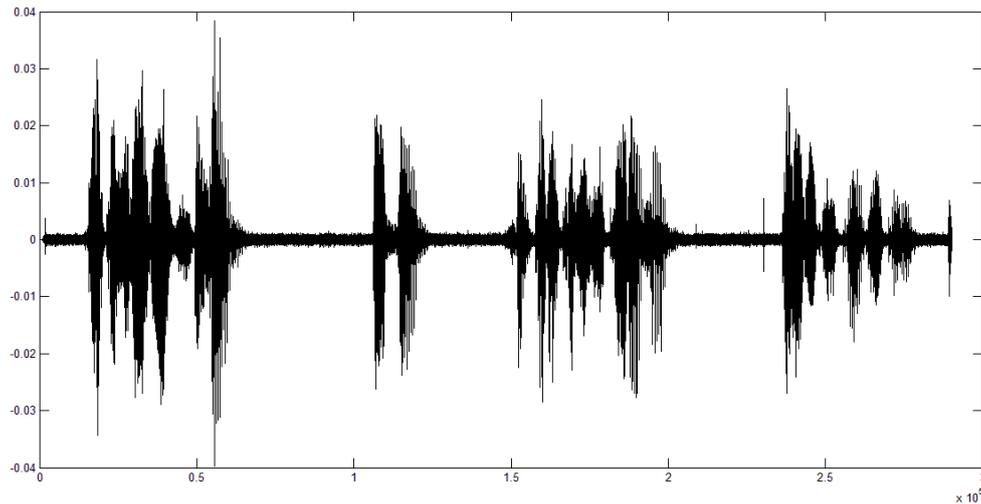


Рис. 5. Разность исходного и полученного сигнала

Из графика видно, что ошибки достигают 0,04 по амплитуде, а относительно исходных сигналов $0,04/0,9 = 0,044 = 4\%$. Эти ошибки нормальны, так как происходило преобразование коэффициентов фильтра к целым двоичным числам до 2^3 , а также при реконструкции происходило деление, отбрасывались младшие разряды.

На рис. 6 показан ненормированный уровень корреляции двух исследуемых сигналов.

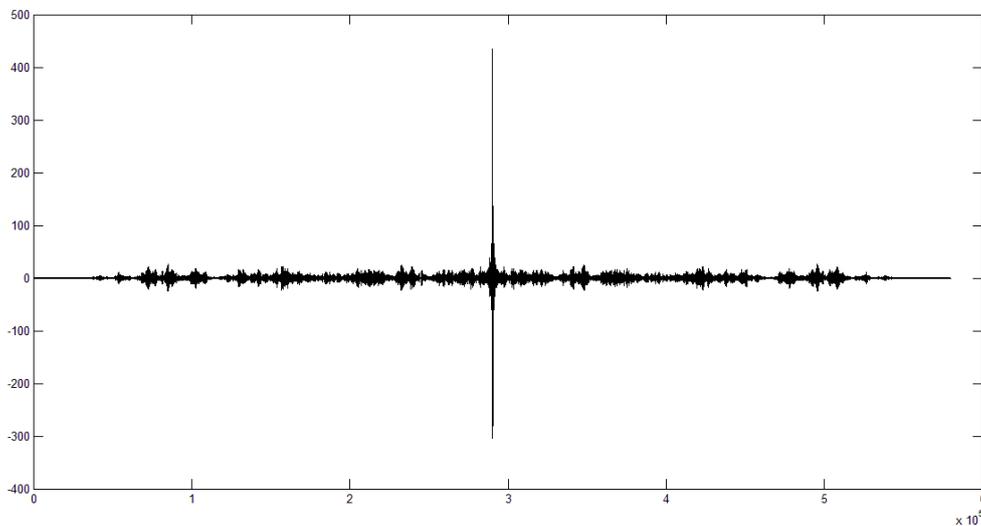


Рис. 6. Уровень корреляции исходного и полученного сигнала

Как видно из представленного графика на рис. 6, имеется пик в точке со значением по оси абсцисс равным $2,9 \cdot 10^5$. Это подтверждает явное сходство сигналов.

Заключение

Проведено моделирование системы кодер/декодер в специализированной программе MATLAB R2016b с использованием реального речевого сигнала. В результате установлена практически полная идентичность преобразованного и восстановленного сигналов, что позволяет судить о высоком качестве восстановления. Точность восстановления составила величину примерно равную 4%. Алгоритм при этом обеспечивает трехуровневую компрессию речевого сигнала. Разработанный модуль кодирования/декодирования может использоваться при создании высокоэффективных кодеков речевого сигнала для низкоскоростных линий коммуникаций.

INVESTIGATION OF THE CHARACTERISTICS OF SPEECH ENCODER FOR FAST WAVELET TRANSFORMATION BASED ON THE MALL ALGORITHM

M.S. ANTONENKO, T.M. PECHEN

Abstract. A compression method based on the use of a discrete wavelet transform for processing a speech signal is considered. A model of a speech encoder for fast wavelet transform based on multiple-scale analysis using the Mall algorithm is proposed and its characteristics are investigated.

Keywords: wavelet transform, the Mall algorithm, speech encoder, compression.

Список литературы

1. Graf M., Truong H.L. // Computer networks. 1999. Vol. 31, Issue 3. P. 273.
2. Куньянь Л., Франц Дж.А., Саймар мл. Р. // Цифровые процессоры обработки сигналов серии TMS320. 1987. Т. 75, № 9. С. 8–27.
3. Kondoz A.M. Digital speech: coding for low bit rate communication systems. NY.: John Wiley & Sons, 1996.
4. Makhoul J., Roucos S., Gish H. // Proc. IEEE 1985 Nov. Vol. 73, 1551–1588.
5. Das A., Gersho A. // Int. J. of Speech Technology. 1999. Vol. 2., P. 317–327.