

УДК 534.87

## ИССЛЕДОВАНИЕ ВЛИЯНИЯ ШУМА НА РЕЧЕВОЙ СИГНАЛ

© 2011 г.

*В.Е. Гай*

Нижегородский государственный технический университет им. Р.Е. Алексева

iamuser@inbox.ru

Поступила в редакцию 07.02.2011

Исследуется возможность использования теории активного восприятия для оценки степени искажения речевого сигнала.

*Ключевые слова:* цифровая обработка сигналов, теория активного восприятия, отношение сигнал / шум.

### Введение

Одна из задач обработки речевых сигналов – оценка уровня искажения сигнала. Задача оценки качества сигнала обычно возникает при тестировании алгоритмов кодирования речевого сигнала, исследовании разнообразных передающих систем.

Целью работы является разработка метода оценки качества искажённого сигнала на основе теории активного восприятия [1, 2].

### Постановка задачи

Процесс искажения сигнала может быть записан в виде:

$$g = \text{DEG} [f],$$

где  $f$  – исходный сигнал,  $g$  – искажённый сигнал,  $\text{DEG} [\cdot]$  – оператор, вносящий искажение.

К методу оценки качества можно предъявить следующие требования:

1) значение вычисляемого показателя должно изменяться с изменением уровня искажения сигнала;

2) оценка значения показателя должна выполняться за приемлемое время (определяется решаемой задачей).

Одним из используемых методов оценки уровня искажений сигнала является вычисление отношения сигнал / шум (signal-to-noise ration, SNR) [3]:

$$\text{SNR} = 10 \log_{10} \left( \frac{\sum_{n=1}^{N_s} s^2(n)}{\sum_{n=1}^{N_v} v^2(n)} \right),$$

где  $s$  – исходный сигнал,  $v$  – шум,  $N_s$  и  $N_v$  – количество отсчётов в исходном и шумовом сигналах. Для использования данного отношения должна быть доступна информация о сигнале и шуме.

### Предлагаемый алгоритм

1) Построение спектрального представления сигнала с использованием  $U$ -преобразования для одномерных сигналов [1, 2, 4, 5]. Результат применения алгоритма к исходному сигналу – набор из  $M$  мгновенных спектров, где  $M$  – число сегментов сигнала. Каждый спектр представляет собой набор коэффициентов  $M = \{m_1, \dots, m_N\}$ ,  $N$  – число фильтров, используемых при формировании спектра. Спектр вычисляется по отрезку сигнала длиной  $len$  отсчётов.

Фильтры Уолша системы Хармута [4, 5], используемые при реализации алгоритма, показаны на рис. 1.

2) Поиск похожих цепочек спектров:

2.1) из  $M$  выбираются ранее не использованные цепочки спектров  $L_i$  и  $L_j$  ( $i \neq j$ )  $L_i = \{M_i, \dots, M_{i+c_{len}}\}$ , где  $M_i$  –  $i$ -й спектр набора  $M$ ,  $c_{len}$  – длина цепочки спектров;

2.2) если абсолютная величина разности между соответствующими спектрами, входящими в  $L_1$  и  $L_2$ , меньше некоторой пороговой величины ( $thr$ ), считается, что эти две цепочки похожи, после чего цепочка  $L_2$  отмечается как использованная (повторяющаяся) и в дальнейших сравнениях участия не принимает;

2.3) подсчитывается количество повторяющихся ( $d$ ) и уникальных (неповторяющихся) ( $nd$ ) цепочек спектров сигнала.

3) Вычисляется показатель, характеризующий искажение сигнала.

Для оценки степени искажения сигнала могут использоваться два величины:

1) отношение числа удалённых спектров к числу спектров в сигнале:

$$err_1 = d / (nd + d),$$

где  $err_1 \in [0; 1]$ ;

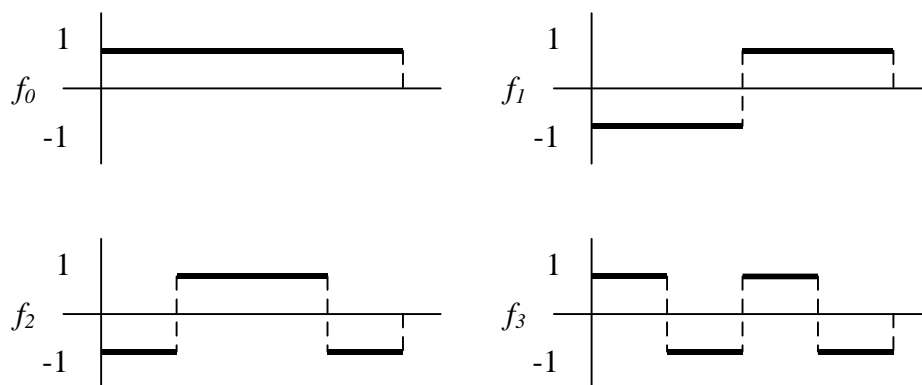


Рис. 1. Фильтры Уолша системы Хармута

2) отношение числа удалённых спектров к числу удалённых спектров в том же сигнале, но без шума:

$$err_2 = d / d_{wn},$$

где  $d_{wn}$  – количество удалённых спектров после обработки сигнала без шума,  $err_2 \in [0; 1)$ .

В обоих случаях меньшие значения показателя соответствуют большему уровню искажения сигнала.

При оценке степени искажения предпочтительнее использовать второй показатель, так как он позволяет учитывать особенности исходного сигнала. Недостатком данного способа является то, что исходный (незашумленный) сигнал не всегда доступен. В результатах экспериментов для сравнения приведены результаты вычисления обоих показателей.

### Вычислительный эксперимент

При проведении эксперимента используются следующие шумы:

- 1) аддитивный нормальный шум;
- 2) аддитивный равномерный шум;
- 3) аддитивный коррелированный шум;
- 4) мультипликативный шум.

Тестовые сигналы:

1) записи речи 10 дикторов, амплитуда каждой записи нормализована на отрезке  $[0; 1]$ . Если не выполнить указанную нормализацию, шум будет по-разному оказывать влияние на сигнал и сравнивать результаты работы алгоритма для разных сигналов будет нельзя;

2) синус с частотой 50 Гц.

Результаты исследования предложенного метода:

1) Существует минимальная длина сигнала  $len_f > len$ , необходимая для выполнения оценки качества сигнала. Если анализируемый сигнал не отвечает указанному условию, то вычисляемое значение показателя вне зависимости от

уровня искажения всегда будет равно 0. Как показывают эксперименты, для речевого сигнала, в среднем,  $len_f$  составляет 0.5 с.

2) Допустим, имеются два искажённых сигнала  $g_1$  и  $g_2$ , каждый из которых получен на основе сигнала  $f$  с помощью оператора DEG  $[\cdot]$ . Для формирования  $g_1$  и  $g_2$  использовался один и тот же шум, но с разным среднеквадратическим отклонением (СКО,  $\sigma_1 < \sigma_2$ ). Применяя предложенный алгоритм к каждому сигналу, получим оценки степени искажения каждого сигнала  $err_{g1}$  и  $err_{g2}$ , причём  $err_{g1} > err_{g2}$ .

3) При тестировании алгоритма использовались различные сочетания значений параметров. В результате экспериментов для практического использования алгоритма предлагаются следующие значения параметров:  $c_{len} = 1$ ,  $N = 4$ ,  $thr = 0.1$ ,  $len = 32$ .

Длина обрабатываемого сегмента ( $len$ ) выбрана равной 32 отсчётам, так как слуховая система оценивает изменения в звуковом импульсе, происходящие на интервале времени меньше 1 мс [6], а при частоте дискретизации 16 кГц 2 мс соответствуют 32 отсчётам.

Длина анализируемого сигнала ( $c_{len}$ ) может меняться в зависимости от требуемой точности оценки степени искажения.

4) Для всех видов шумов, описанных выше, получено, что значение показателя  $err_1$  ( $err_2$ ) уменьшается с увеличением дисперсии шума, искажающего сигнал.

5) Время вычисления показателя зависит от длины сигнала и выбранных значений параметров. При рекомендуемых параметрах для сигнала длиной 50000 отсчётов время оценки составляет 54 секунды.

На рис. 2а показан периодический сигнал ( $SNR = 15$  дБ), искажённый нормальным шумом, на рис. 2б – результат его восстановления с помощью метода, описанного в [7, 8].

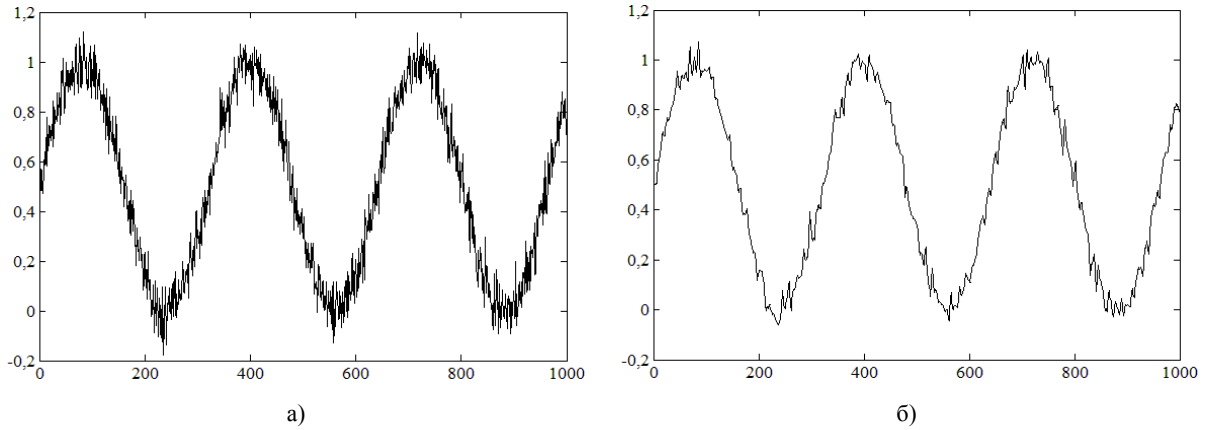


Рис. 2. Искажённый (а) и восстановленный (б) сигналы

Таблица 1

Периодический сигнал (синус, 50 Гц), нормальный шум

СКО шума	0.034	0.063	0.109	0.199	0.340	0.609
Искажённый сигнал						
дБ	25	20	15	10	5	0
$err_1$	0.303	0.090	0.026	0.012	0.001	0
$err_2$	0.305	0.091	0.027	0.013	0.001	0
Восстановленный сигнал						
дБ	26.850	22.348	16.301	12.853	7.811	1.947
$err_1$	0.419	0.182	0.106	0.054	0.042	0.028
$err_2$	0.422	0.183	0.107	0.055	0.043	0.029

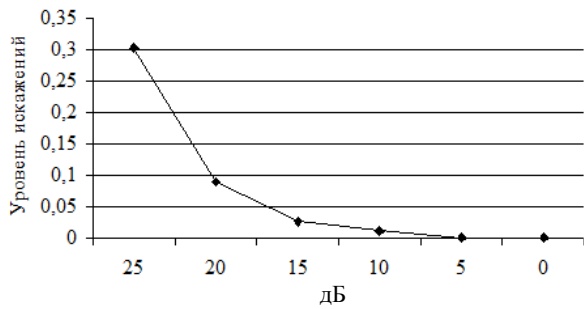


Рис. 3. Зависимость значения показателя  $err_1$  от уровня искажения сигнала

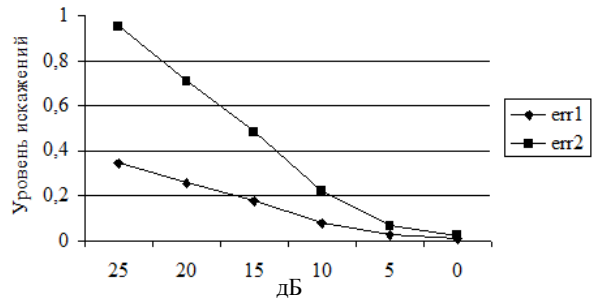


Рис. 4. Зависимость значений показателей  $err_1$  и  $err_2$  от уровня искажения сигнала

В табл. 1 приведены результаты вычисления показателей  $err_1$  и  $err_2$  для искажённого и восстановленного периодического сигнала (см. рис. 3). В табл. 2 можно видеть результаты исследования речевого сигнала до и после искажения (см. рис. 4).

На графике (см. рис. 3) по оси ОХ отложено отношение сигнал / шум (в дБ), по оси ОУ – уровень искажений, вычисленный в соответствии с предложенным алгоритмом.

Таким образом, предлагаемый метод в действительности демонстрирует снижение уровня

искажения сигнала после выполнения фильтрации, что согласуется с результатами визуальной оценки (см. табл. 1 и 2).

Значения показателей  $err_1$  и  $err_2$  демонстрируют одинаковое поведение, поэтому при практическом применении данного алгоритма должен использоваться показатель  $err_1$ , так как для его расчёта не требуется наличие информации о незашумленном сигнале. Значения показателей уровня искажения сигнала, вычисленные на основе предложенного метода, согласуются со значениями отношения сигнал / шум для тех же сигналов.

Таблица 2

## Речевой сигнал, равномерный шум

СКО шума	0.012	0.024	0.040	0.069	0.121	0.222
Искажённый сигнал						
дБ	25	20	15	10	5	0
$err_1$	0.345	0.257	0.175	0.079	0.024	0.007
$err_2$	0.952	0.710	0.484	0.219	0.067	0.021
Восстановленный сигнал						
дБ	30.321	25.483	18.278	14.495	5.966	1.618
$err_1$	0.361	0.304	0.224	0.119	0.060	0.033
$err_2$	0.994	0.837	0.538	0.330	0.166	0.091

**Выводы**

В данной работе описан алгоритм оценки степени искажения сигнала. Результаты исследования подтверждают возможность его использования для решения поставленной задачи.

Достоинствами предложенного алгоритма являются низкая вычислительная сложность, простота реализации, а также возможность оценки уровня искажений только на основе имеющегося сигнала.

Дальнейшее направление исследований – доработка предложенного метода при решении задачи разборчивости речи.

*Список литературы*

1. Утробин В.А. Информационные модели системы зрительного восприятия для задач компьютерной

обработки изображений. Н. Новгород: НГТУ, 2001. 234 с.

2. Утробин В.А. Компьютерная обработка изображений. Принятие решений в пространстве эталонов. Н. Новгород: НГТУ, 2004. 221 с.

3. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. М.: Мир, 1978. 860 с.

4. Хармут Х.Ф. Применение методов теории информации в физике: Пер. с англ. М.: Мир, 1989. 344 с.

5. Хармут Х.Ф. Передача информации ортогональными функциями: Пер. с англ. М.: Связь, 1975. 272 с.

6. Элементы теории биологических анализаторов / Под ред. Н.В. Позина. М.: Наука. Главная редакция физико-математической литературы, 1978. 360 с.

7. Donoho D.L. De-noising by soft-thresholding // IEEE Transactions on Information Theory. 1995. Vol. 41, No. 3. P. 613–627.

8. Donoho D.L., Johnstone I. M. Ideal denoising in an orthonormal basis chosen from a library of bases // C.R.A.S. Paris. 1996. T. 319. Ser. I. P. 1317–1322.

**INVESTIGATION OF NOISE INFLUENCE ON SPEECH SIGNAL**

V.E. Gai

The possibility to use the theory of active perception in evaluating the distortion degree of a speech signal is investigated.

*Keywords:* digital signal processing, theory of active perception, signal-to-noise ratio.