

УДК 621.39

## РАЗРАБОТКА АЛГОРИТМА АВТОМАТИЧЕСКОЙ РЕГУЛИРОВКИ УСИЛЕНИЯ ДЛЯ РЕЧЕВОГО СИГНАЛА

Филимонов В.А., к.т.н., доцент ВАС им. С.М. Буденного, e-mail: awgn@yandex.ru.

### DEVELOPMENT OF AUTOMATIC GAIN CONTROL ALGORITHM OF FOR SPEECH SIGNAL

Filimonov V.A.

This article discusses AGC automatic gain control algorithms for digital analytical input signal. Derivation of AGC algorithm for digital analytical speech signal is given. Impact of input signal response and algorithm changeable parameters on output signal response is described. Realization of discuss is possible only on the base of signal digital processing. Algorithm efficiency is confirmed in programmable way. Programmed listings are given in the article text. Special feature of discussed algorithms is their potential resistance defined by lack of feedback circuits.

**Key words:** automatic gain control, speech signal, digital signal processing.

**Ключевые слова:** автоматическая регулировка усиления, речевой сигнал, цифровая обработка сигналов.

#### Введение

Основной задачей автоматической регулировки усиления (АРУ) является поддержание постоянной амплитуды выходного сигнала при изменении амплитуды входного сигнала.

В общем виде АРУ является нелинейной, нестационарной, зависящей от сигнала системой с обратной связью. Именно поэтому анализ системы АРУ выполняется скорее эмпирически, чем математически, а сами системы так мало обсуждаются в литературе по ЦОС [1]. Разработка качественной системы АРУ на основе ЦОС является нетривиальной задачей. В зависимости от вида входного сигнала (непрерывный или дискретный цифровой сигнал) системы АРУ могут иметь различные структуры. Представленные в технической литературе алгоритмы АРУ (например, [1]) не используют преимуществ ЦОС, и являются переносом, с незначительными дополнениями, алгоритмов аналоговой схемотехники на цифровой сигнал.

Основным преимуществом ЦОС для систем АРУ является возможность использования аналитического сигнала. В основе любой системы АРУ для аналитического сигнала лежит алгоритм нормализации сигнала, называемый еще мгновенным АРУ. Рассмотрим его более подробно.

Как известно, любой отсчет цифрового аналитического сигнала представляет собой комплексное число:

$$\dot{s}(n) = \operatorname{Re}\{\dot{s}(n)\} + i\operatorname{Im}\{\dot{s}(n)\},$$

которое может быть представлено в показательной форме:

$$\dot{s}(n) = |\dot{s}(n)|e^{i\theta},$$

$$\text{где } \theta = \arctg \frac{\operatorname{Im}\{\dot{s}(n)\}}{\operatorname{Re}\{\dot{s}(n)\}}, \quad |\dot{s}(n)| = \sqrt{[\operatorname{Re}\{\dot{s}(n)\}]^2 + [\operatorname{Im}\{\dot{s}(n)\}]^2}.$$

Рассматриваются алгоритмы автоматической регулировки усиления (АРУ) при использовании цифрового аналитического входного сигнала. Представлен вывод алгоритма АРУ для цифрового аналитического речевого сигнала и описано влияние качества входного сигнала и изменяемых параметров алгоритма на качество выходного сигнала. Работоспособность алгоритмов подтверждена экспериментально, листинги программ приведены в тексте статьи. Особенность рассматриваемых алгоритмов является их потенциальная устойчивость, обусловленная отсутствием цепей обратной связи.

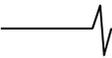
Нормализация цифрового сигнала производится согласно следующего выражения:

$$\dot{s}_n(n) = \frac{\dot{s}(n)}{|\dot{s}(n)|}. \quad (1)$$

Нормализация сигнала, согласно (1), идеально подходит, например для устранения паразитной амплитудной модуляции, в случае если цифровой сигнал  $\dot{s}(n)$  является гармоническим колебанием (комплексной экспонентой). Пример устранения амплитудной модуляции представлен в листинге 1.

Листинг 1

```
clc;
clear;
N = 8000; % количество отсчетов сигнала
n = 0:1:N-1; % нормированное время
f0 = 500; % частота сигнала
fs = 8000; % частота дискретизации
s = sin(2*pi*f0*n/fs); % + sin(pi*f0*n/fs); % действительный сигнал
% *** МОДУЛИРУЕМ ПО АМПЛИТУДЕ
s_m = s.*(1 + 0.8*sin(2*pi*10*n/fs));
% s_m = awgn(s_m, 20); % добавление шума
% *****
s_h = hilbert(s_m); % аналитический сигнал
figure;
plot(real(s_h)); grid on; hold on;
plot(imag(s_h), 'r');
title('Сигнал на входе');
figure;
plot(abs(s_h)); grid on;
title('Амплитудная огибающая сигнала');
% нормализация сигнала
s_n = s_h./(abs(s_h));
```



```
figure;
plot(real(s_n)); grid on; hold on;
plot(imag(s_n), 'r');
title('Сигнал после нормализации');
```

Однако, если сигнал  $\hat{s}(n)$  представляет собой смесь гармонического колебания и шума, то нормализация сигнала приводит к искажению сигнала. Вследствие нелинейности операции (1), искажения так же будут нелинейными. В случае если сигнал  $\hat{s}(n)$  представляет собой сумму гармонических колебаний с различной частотой, нормализация сигнала так же приведет к неустранимым искажениям. Таким образом, мгновенная АРУ (нормирование сигнала) имеет очень ограниченное применение.

Последовательность абсолютных величин комплексных отсчетов представляет собой амплитудную огибающую сигнала (см. листинг 1, усреднение которой позволяет оценивать плавные изменения амплитуды сигнала, при этом временной интервал усреднения зависит от скорости изменения амплитуды сигнала (амплитудной огибающей). Усреднение амплитудной огибающей можно произвести с помощью однородного фильтра (фильтра скользящего среднего) [1], [3].

Для выравнивания амплитуды сигнала применяется та же процедура, что и при нормализации сигнала (1), но в знаменателе используется не модуль абсолютного значения отсчета, а усредненное значение огибающей.

Так как операция усреднения выполняется с помощью КИХ-фильтра, то усредненный сигнал будет задержан относительно входного. Величина задержки равна половине длины импульсной характеристики однородного фильтра [2]. Тогда изменение абсолютного значения текущего отсчета будет производиться по прошедшим отсчетам. Более правильно производить изменения текущего отсчета, исходя из значений прошедших и предшествующих отсчетов. ЦОС позволяет это выполнить путем задержки обрабатываемого сигнала (листинг 2).

Листинг 2

```
clc;
clear;
N = 8000; % количество отсчетов сигнала
n = 0:1:N-1; % нормированное время
f0 = 500; % частота сигнала
fs = 8000; % частота дискретизации
s = sin(2*pi*f0*n/fs) + sin(pi*f0*n/fs); % действительный сигнал
% % *** МОДУЛИРУЕМ ПО АМПЛИТУДЕ (амплитудные искажения)
s_m = s.*(1 + 0.5*sin(2*pi*5*n/fs));
% % *****
s_h = hilbert(s_m); % аналитический сигнал
figure;
plot(real(s_h)); grid on; hold on;
plot(imag(s_h), 'r');
title('Сигнал на входе');
% ***** АРУ *****
K = 50; % длительность усреднения в нормированном времени
h = 1:1:K;
h = h./h; % коэффициенты однородного фильтра
s_sr = (1/K)*filter(h, 1, abs(s_h)); % усредненное значение огибающей
% Задержка обрабатываемого сигнала равна половине ИХ фильтра
h_zad = 1:1:ceil(K/2);
```

```
h_zad = [h_zad*0 1];
s_h_zad = filter(h_zad, 1, s_h);
y_aru = s_h_zad./s_sr; % выравнивание уровня сигнала
figure;
plot(real(y_aru)); grid on; hold on;
plot(imag(y_aru), 'r');
title('Сигнал на выходе АРУ');
```

По результату работы программы можно увидеть первоначальные искажения сигнала, вызванные переходными процессами фильтра. Изменяя амплитуду входного сигнала можно заметить, что амплитуда выходного сигнала не изменяется. Изменяя длительности усреднения в большую и меньшую сторону, можно сделать вывод: увеличение длительности усреднения (подставьте  $K = 1000$ ) приводит к тому, что система АРУ не отслеживает изменения амплитуды сигнала, уменьшение длительности приводит к искажению многочастотного сигнала, а при  $K = 1$  происходит нормализация сигнала.

При неизвестном периоде паразитной амплитудной модуляции поиск оптимальной длины усреднения становится невыполнимой задачей.

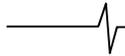
Описанный выше алгоритм АРУ не подходит для речевого сигнала. Сложность создания систем АРУ для речевого сигнала обусловлена наличием пауз между словами, фразами и отдельными слогами. Оцифрованный речевой сигнал (как и исходный аналоговый) всегда имеет незначительный шум в паузах, как минимум шум квантования. Незначительный шум в паузах будет усиливаться системой АРУ до номинального уровня выходного сигнала. Для борьбы с этим эффектом можно выполнять блокирование системы АРУ на время пауз. Однако, решение задач определения порога срабатывания, временных границ пауз, переходных процессов и др. приводит к труднореализуемой системе.

Другой подход к решению данной проблемы представлен в алгоритме, рассмотренном ниже. К сожалению, данный алгоритм разработан эвристическим, а не математическим методом. Для проверки работоспособности алгоритма необходимо воспользоваться студийной записью речи в формате wav, или записать файл данного формата самостоятельно, используя микрофон и звуковую карту компьютера совместно с программой в среде Matlab составленной на основе листинга 3.

Листинг 3

```
clc;
clear;
fs = 22050; % частота дискретизации
s = wavrecord(14*fs, fs); % длительность 14 секунд
s = s'; % перевод в вектор-строку
figure;
plot(s); grid on;
title('Временная форма речевого сигнала');
% *****
wavwrite(s, fs, 'My_voice.wav'); % запись в файл
```

Алгоритм АРУ для речевого сигнала в целом повторяет алгоритм рассмотренный выше. Однако к усредненному значению огибающей сигнала добавляется постоянная составляющая, значение которой больше на порядок максимальных абсолютных значений шума в паузах речевого сигнала. Данную операцию можно вы-



полнить путем добавления к модулю каждого отсчета усредненного значения огибающей константы  $L$ . Тогда сигнал на выходе системы АРУ получается следующим образом:

$$\dot{s}_{\text{ВЫХ}}(n) = \frac{\dot{s}_{\text{ВХ}}(n)}{|\dot{s}_{\text{СП}}(n)| + L} \quad (2)$$

Выбор значения  $L$  зависит от мощности шума в паузах входного сигнала. Найти данное значение можно только эмпирическим путем, т.к. качество входного и выходного сигнала можно оценить только методом субъективной оценки разборчивости речи. Очевидно, что значение  $L$  влияет как на среднее абсолютное значение (уровень) шума в паузах выходного сигнала, так и на степень выравнивания уровня речевого сигнала между паузами.

Степень влияния константы  $L$  на выходной сигнал зависит от уровня входного сигнала. Так мощность шума в паузах речевого сигнала на выходе системы АРУ зависит как от значения  $L$ , так и от мощности шума в паузах на входе системы, точнее от их соотношения. Исключив из выражения (2) усреднение, выведем зависимость абсолютного значения амплитудной огибающей выходного сигнала  $|\dot{s}_{\text{ВЫХ}}(n)|$  в паузах от отношения

$L/|\dot{s}_{\text{ВХ}}(n)|$ , с учетом ряда ограничений. Не будем учитывать переходные процессы возникающие в моменты перехода от сигнала к паузе, кроме того примем, что входной сигнал в паузах (шум) имеет постоянную амплитудную огибающую, следовательно  $|\dot{s}_{\text{СП}}(n)| = |\dot{s}_{\text{ВХ}}(n)|$ .

Тогда, исходя из выражения (2), амплитудная огибающая выходного сигнала равна

$$|\dot{s}_{\text{ВЫХ}}(n)| = \frac{|\dot{s}_{\text{ВХ}}(n)|}{\left| \frac{|\dot{s}_{\text{ВХ}}(n)|}{|\dot{s}_{\text{ВХ}}(n)| + L} + L \right|} = \frac{|\dot{s}_{\text{ВХ}}(n)|}{|\dot{s}_{\text{ВХ}}(n)| + L} \quad (3)$$

Разделив числитель и знаменатель правой части выражения (3) на  $|\dot{s}_{\text{ВХ}}(n)|$ , получаем:

$$|\dot{s}_{\text{ВЫХ}}(n)| = \frac{1}{1 + L/|\dot{s}_{\text{ВХ}}(n)|} \quad (4)$$

Графически данную зависимость (рис. 1) можно получить, используя листинг 4.

Листинг 4

```
clc;
clear;
delta = 0:0.01:10;
S_sh = 1./(1+delta);
figure;
plot(delta, S_sh); grid on;
```

Таким образом, уровень шума в паузах выходного речевого сигнала зависит не столько от уровня входного шума в паузах, сколько от соотношения  $L$  к  $|\dot{s}_{\text{ВХ}}(n)|$ . Отметим, что увеличение этого соотношения приводит к уменьшению выходного уровня сигнала в целом. Кроме того оно влияет на степень выравнивания уровня речевого сигнала между паузами, при этом степень выравнивания улучшается с уменьшением значения  $L$  (отношения  $L/|\dot{s}_{\text{ВХ}}(n)|$ ).

Следовательно, для уменьшения шума в паузах

необходимо увеличивать значение  $L$  (отношение  $L/|\dot{s}_{\text{ВХ}}(n)|$ ), а для улучшения выравнивания сигнала между пауз надо его уменьшать. Данное противоречие приводит к тому, что для получения качественной работы алгоритма АРУ, основанной на выражении (2), необходимо, чтобы входной сигнал имел минимально возможный шум в паузах.

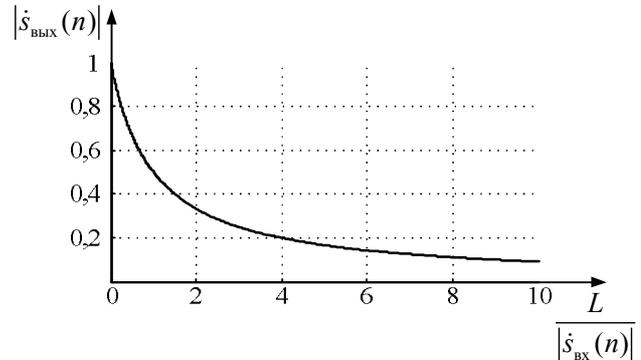


Рис. 1. Зависимость абсолютного значения выходного сигнала от соотношения  $L/|\dot{s}_{\text{ВХ}}(n)|$

Данный алгоритм имеет еще один варьируемый параметр – длительность усреднения, который влияет на степень выравнивания сигнала между паузами, так же данный параметр влияет на длительность переходных процессов на границах пауз. Исследование алгоритма на студийных записях показали, что длительность усреднения должна быть выбрана равной периоду стационарности речевого сигнала составляющего 1–2 мсек.

Полностью описанный алгоритм представлен в листинге 5.

Листинг 5

```
clc;
clear;
[s, Fs] = wavread('My_voice.wav'); % чтение данных из файла
Fs % частота дискретизации
s = s'; % перевод в вектор-строку
n = 0:1:length(s)-1; % нормированное время
figure;
plot(n, s); grid on;
title('Временная форма речевого сигнала до АРУ');
disp('Воспроизведение входного сигнала')
sound(s, Fs); % воспроизведение входного сигнала
s_h = hilbert(s); % аналитический сигнал
% *** АРУ ***
% время стационарности 1-2 мсек, т.е. K =
Fs*0.001(0.002) = 22(44)
% при K = 1 и L = 0 производится нормализация сигнала
K = 22;
h = 1:1:K;
h = h./h; % коэффициенты однородного фильтра
S = (1/K)*filter(h, 1, abs(s_h));
figure;
plot(n, S); grid on;
title('Временная форма усредненной огибающей сигнала');
% ***** Задержка входного сигнала *****
h_zad = 1:1:ceil(K/2); % равна половине длины ИХ
h_zad = [h_zad*0 1];
s_h_zad = filter(h_zad, 1, s_h);
% ***** Выравнивание сигнала *****
L = 0.1; % уровень добавляемой постоянной составляющей
y_aru = s_h_zad./(S + L);
```

```

y_aru = 0.5*y_aru; % изменение амплитуды выходно-
го сигнала
figure;
plot(n, real(y_aru)); grid on; hold on;
plot(n, imag(y_aru), 'r');
title('Временная форма речевого сигнала после
APU');
% *****
disp('Воспроизведение выходного сигнала')
sound(real(y_aru), Fs);

```

Преимуществом данного алгоритма являются отсутствие цепей обратной связи и условных операторов (if, for и т.п.). Основным недостатком алгоритма – повышение уровня шума в паузах речевого сигнала.

К сожалению, в технической литературе отсутствует методика проверки качества систем АРУ для речевых сигналов. Общеизвестным параметром речевого сигнала является пик-фактор (отношение максимального значения к среднеквадратическому отклонению или максимальной мгновенной мощности к средней). Качество работы АРУ можно оценить сравнением вероятностей распределения значений отсчетов случайного цифрового речевого сигнала до и после выравнивания (листинг 6 в дополнении к листингу 5).

Листинг 6

```

pik_faktor_input = (abs(max(s)))/(std(s,1))
pik_faktor_output =
(max(abs(y_aru)))/(std(abs(y_aru),1))
% ***** ГИСТОГРАММЫ *****
setka = -1:1/128:1; % получаем 257 уровней оцени-
вания

```

```

N = length(s); % длина обрабатываемого сигнала
% вычисление вероятности распределения значений
p_u_q = (1/N)*hist(s, setka);
figure;
stem(setka, p_u_q); grid on;
title('Распределение вероятностей записанного го-
лоса');
% вычисление вероятности распределения значений
p_u_q = (1/N)*hist(real(y_aru), setka);
figure;
stem(setka, p_u_q); grid on;
title('Распределение вероятностей голоса после
APU');

```

Результат работы программы показывает, что сигнал после АРУ имеет более прямоугольную форму распределения случайных значений и его пик-фактор меньше. Однако оба эти параметра не учитывают пауз в речевом сигнале, следовательно, данные оценки не являются полностью объективными.

### Литература

1. Ричард Лайонс. Цифровая обработка сигналов: Второе издание. Пер. с англ. - М.: ООО «Бином-Пресс», 2006 г. - 656 с.
2. Солонина А.И., Улахович Д.А., Арбухов С.М., Соловьева Е.Б. Основы цифровой обработки сигналов. Изд. 2-е. - СПб.: БВХ-Петербург, 2005. - 768 с: ил.
3. Юкио Сато Цифровая обработка сигналов. Пер. с яп. Селениной Т.Г. М.: Додэка-XXI, 2010. - 176 с: ил.



Отдел оптовых поставок  
E-mail: [opt@bhv.spb.su](mailto:opt@bhv.spb.su)



### Солонина А.И.

#### Цифровая обработка сигналов в зеркале MATLAB

*Рецензенты:* Е.Б. Соловьева, д-р техн. наук, завкафедрой теоретических основ электротехники Санкт-Петербургского государственного электротехнического университета «ЛЭТИ»; А.А. Монаков, д-р техн. наук, проф. кафедры радиотехнических систем Санкт-Петербургского государственного университета аэрокосмического приборостроения.

Учебное пособие представляет собой базовый курс по цифровой обработке сигналов с традиционными темами: линейные дискретные системы; эффекты квантования в цифровых системах с фиксированной точкой; дискретные сигналы с описанием в частотной области, включая алгоритмы дискретного и быстрого преобразований Фурье; случайные процессы и их статистические характеристики; методы непараметрического и параметрического спектрального анализа; цифровые фильтры общего и специального назначения; многоскоростные системы и их полифазные структуры; банки фильтров; основы адаптивной обработки сигналов.

Теория ЦОС преподносится с позиции моделирования в MATLAB: теоретические аспекты базовых методов и алгоритмов ЦОС увязываются с их моделированием в MATLAB и иллюстрируются соответствующими примерами.

Книга предназначена для студентов и преподавателей вузов, а также специалистов в области цифровой обработки сигналов.

**Солонина Алла Ивановна**, кандидат технических наук, доцент кафедры радиосистем и обработки сигналов Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича с многолетним научно-методическим и педагогическим стажем, автор многочисленных публикаций и книг по цифровой обработке сигналов, в том числе «Алгоритмы и процессоры цифровой обработки сигналов», «Основы цифровой обработки сигналов. Курс лекций», «Цифровая обработка сигналов. Моделирование в MATLAB», «Цифровая обработка сигналов. Моделирование в Simulink», «Цифровая обработка сигналов и MATLAB», выпущенных издательством «БХВ-Петербург».