

А. Я. ДРЮЧЕНКО, А. И. КАНТЕМИР, С. Ф. КОРЯК, В. Ю. СОКОЛОВ

### НОРМИРОВАНИЕ РЕЧЕВЫХ СИГНАЛОВ

---

Известно, что речевой сигнал характеризуется многократной избыточностью как семантической (что говорится), так и эктосемантической (как говорится) информации.

Как показали исследования [1, с. 136], в динамическом диапазоне сигнала содержится информация о качестве звучания (как говорится). Компрессированный по уровню речевой сигнал имеет разборчивость, которая мало отличается от разборчивости некомпьютеризированного сигнала. Более того, компрессия динамического диапазона речи позволяет повысить помехоустойчивость, т. е. получить большую разборчивость в условиях помех.

При дискретизации некомпьютеризированного (ненормированного) речевого сигнала в системах автоматического распознавания речи параметры, характеризующие динамический диапазон (порядка 50 дБ, т. е. изменения по уровню более чем в 300 раз [1, с. 370]), также становятся признаками, порождая тем самым чрезвычайно громоздкую систему признаков описания одного и того же звукового образа, но имеющего различный масштаб (к тому же нелинейный для разных звуков в слове!). Так как, например, при фонемном распознавании число фоноидов в русском языке составляет по меньшей мере 140 [1, с. 16], то исследователь сталкивается со знаменитым «проклятием размерности» [2, с. 65]. К тому же еще не разработана даже в общем виде законченная в смысле математической строгости и практической реализуемости методика выбора системы эффективных признаков, и решение задачи становится практически неприемлемым.

В этой связи возникает необходимость минимизации пространства признаков. Одним из эффективных методов достижения этой

цели может быть нормирование динамического диапазона исходного речевого сигнала.

В известных устройствах нормирование сопряжено с вырождением речевого сигнала, т. е. не учтены свойства речевых сигналов: квазипериодичность, проявление фонетического качества звука на временном интервале, равном периоду основного тона голоса. Под нормированием же речевого сигнала следует понимать приведение исходного сигнала на каждом заданном временном интервале нормирования к установленному фиксированному значению в точке максимального амплитудного значения сигнала при сохранении всех прочих амплитудных соотношений на интервале нормирования.

Целью разработки предлагаемого устройства является повышение точности воспроизведения формы нормализуемых речевых сигналов за счет учета особенностей речевых сигналов и применения цифровых методов обработки.

В нашем случае нормирование речевых сигналов основано на измерении максимального значения амплитуды исходного сигнала на фиксированном интервале времени, который выбирают равным или несколько превышающим длину элементарного сегмента аллофона вокализованных звуков. Исходный сигнал представляют в виде цифровых отсчетов, которые запоминают на интервале времени нормирования, вырабатывают коэффициент нормирования на данном интервале как частное от деления заданного фиксированного значения (нормы) на максимальное значение сигнала на данном интервале нормирования для регулирования каждого отсчета, посредством умножения хранящихся отсчетов на полученный коэффициент нормирования.

В процессе исследования данного технического решения выявлено, что оно обладает новой совокупностью признаков. До сих

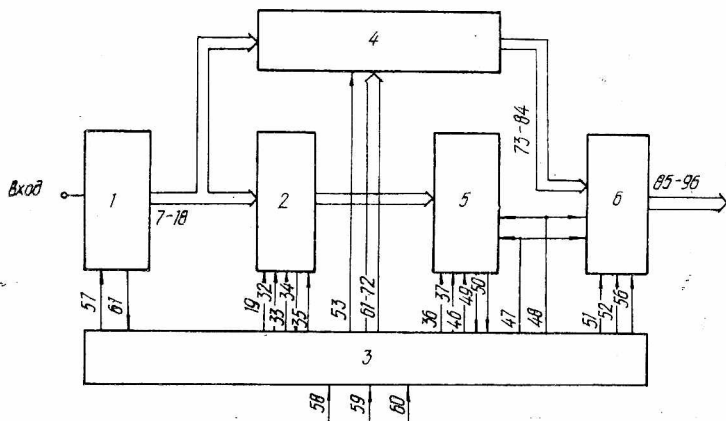


Рис. 1. Функциональная схема устройства для нормирования речевых сигналов

вень плотности в каком-нибудь интервале частот, или что-нибудь иное. От этого выбора может зависеть в конечном итоге результат распознавания [3]. Авторы в качестве репера используют максимальное амплитудное значение исходного речевого сигнала на интервале 10 мсек, т. е. равном или превышающем максимально воз-

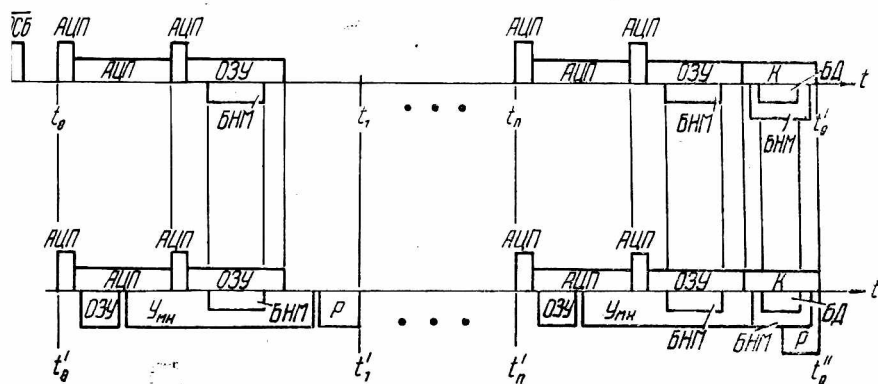


Рис. 2. Временная диаграмма работы устройства

можную длительность одного элементарного сегмента аллофона локализованных звуков для произвольного диктора.

Устройство для нормирования речевых сигналов (рис. 1) содержит аналого-цифровой преобразователь 1 (АЦП), блок нахождения максимума 2 (БНМ), блок управления 3 (БУ), оперативно-запоминающее устройство 4 (ОЗУ), блок деления 5 (БД), блок умножения 6 (БУМ).

Нормирование речевых сигналов заключается в следующем: речевой сигнал преобразуют в последовательность  $n$ -разрядных двоичных чисел и запоминают. Емкость памяти определяют по формуле

$$P = \frac{T}{\tau} n(\text{бит}),$$

где  $T$  — время нормализации (интервал времени, на протяжении которого вырабатывается коэффициент нормирования);  $\tau$  — такт преобразования. Одновременно с записью в память определяют максимальное амплитудное значение входного сигнала на заданном временном интервале (рис. 2). Временной же интервал выбирают, исходя из свойств речевого сигнала, так как на таком интервале может проявиться элементарный сегмент аллофона, который представляет собой период колебания сложной формы, несущий всю информацию о фонетическом качестве звука [4]. По окончании времени нормализации  $T$  находят коэффициент нормирования:

$$k = \frac{N_{\text{норма}}}{N_{\text{тек.мах}}}.$$

Здесь  $N_{\text{норма}}$  — заданное фиксированное (максимальное) значение нормированного сигнала, причем коэффициент нормирования  $K$  находят за время  $t \leq \tau$ . После определения коэффициента  $K$  каждое предыдущее мгновенное значение считывается из памяти и умножается на коэффициент  $K$ . В результате получим пронормированный сигнал.

Устройство работает следующим образом. На вход АЦП 1 подается сигнал, который представляет собой сумму речевого сигнала и положительного постоянного напряжения, которое выбирают равным половине рабочего диапазона АЦП, т. е. речевой сигнал вынесен в положительную область и АЦП 1 работает с однополярным сигналом.

Импульсом сброса по цепи 60 устройство приводят в исходное состояние. В начальный момент времени импульс «Внешний запуск АЦП» по цепи 58 разрешает работу блока управления 3. Блок управления 3 формирует по цепи 57 импульс запуска АЦП 1. АЦП 1 преобразует входной сигнал и выдает на информационную шину АЦП результат преобразования. АЦП 1 формирует по цепи 61 импульс «Конец АЦП» в блок управления 3, который формирует:

— текущий адрес ОЗУ 4 по адресной шине ОЗУ;

— импульс записи ОЗУ 4 по цепи 53, при этом с информационной шины АЦП цифровой отсчет речевого сигнала записывается в соответствующую ячейку ОЗУ 4;

— управляющие сигналы по цепям 19, 33, 35 для обеспечения функционирования блока нахождения максимума 2.

Блок нахождения максимума 2 сравнивает текущий цифровой отсчет речевого сигнала на информационной шине АЦП с предыдущим отсчетом. Если текущий отсчет меньше или равен предыдущему, то на выходе блока нахождения максимума 2 на его информационной шине сохранится дополнительный код (делитель) предыдущего отсчета. И если текущий отсчет больше предыдущего, то на информационной шине блока нахождения максимума 2 формируется дополнительный код текущего отсчета. Описанный цикл повторяется на протяжении выбранного времени нормирования, которое задают величиной максимального адреса адресной шины ОЗУ.

Когда последний отсчет речевого сигнала на заданном времени нормирования записан в ОЗУ 4 и на информационной шине блока нахождения максимума 2 сформирован дополнительный код максимального значения отсчета речевого сигнала на участке нормирования, блок управления 3 разрешает работу блока деления 5, формируя сигналы по цепям 36, 37, 46, 48. При этом блок деления 5 вычисляет коэффициент нормирования, который получается делением максимально возможного отсчета АЦП (разрядность информационной шины АЦП) на дополнительный код максимального значения отсчета речевого сигнала, полученный на протяжении предыдущего времени нормирования. Одновременно с вычислением

ем коэффициента нормирования блок управления 3 устанавливает блок нахождения максимума 2 по цепи 32 в исходное состояние.

Следующий интервал нормирования начинается с приходом импульса запуска АЦП с блока управления 3, который разрешает преобразование АЦП. Одновременно блок управления 3 устанавливает по адресной шине ОЗУ начальный адрес. При этом на информационной шине ОЗУ находится первый отсчет речевого сигнала предыдущего интервала нормирования. Блок управления 3 разрешает работу блока умножения 6 по цепям 47, 48, 51, 52, 56. Блок умножения 6 вычисляет произведение данного отсчета ОЗУ на коэффициент нормирования, полученного в блоке деления 5. По окончании операции умножения на информационной шине блока умножения 6 получаем отсчет пронормированного речевого сигнала.

По окончании работы АЦП 1 импульс «Конец работы» по цепи 61 поступает на блок управления 3. Блок управления 3 разрешает работу блока нахождения максимума 2 и формирует по цепи 53 импульс записи в ОЗУ 4. Таким образом, в начальную ячейку ОЗУ 4 запишется первый отсчет речевого сигнала на новом интервале нормирования. С приходом следующего импульса «Запуск АЦП» по цепи 57 описанный цикл повторяется с увеличением адреса ОЗУ 4 на единицу. По приходу последнего импульса «Запуск АЦП» с блока управления 3 по цепи 57 на втором интервале нормирования повторяется описанный цикл, но в конце которого блок управления 3 разрешает блоку деления 5 нахождение коэффициента нормирования для данного времени нормирования, и устанавливает блок нахождения максимума 2 в исходное состояние.

Дальнейшая работа устройства циклически повторяется по описанной работе устройства на втором интервале нормирования.

Предложенное техническое решение позволяет повысить качество и надежность работы устройств анализа и синтеза речи, так как учитывает свойства речевого сигнала: квазипериодичность, проявление фонетического качества звука на временном интервале, равном периоду основного тона голоса. Тем самым при нормировании речевых сигналов сохраняется информация о фонетической принадлежности речевого сигнала, а также информация, необходимая для идентификации диктора по голосу.

**Список литературы:** 1. Сапожков М. А. Речевой сигнал в кибернетике и связи. М., 1963. 564 с. 2. Вайншток А. П., Книппер А. В., Махонин В. А., Турбович И. Т. Анализ речи режекторными фильтрами//Речевое общение в автоматических системах. М., 1975. С. 142. 3. Мясников Л. Л., Мясникова Е. Н. Автоматическое распознавание звуковых образов. Л., 1970. 160 с. 4. А. с. СССР. № 1030840. Устройство для распознавания речевых сигналов//М. Ф. Бондаренко, А. Я. Дрюченко, А. И. Кантемир, В. Ю. Соколов. Бюл. изобрет. 1985. № 27. С. 62.

Поступила в редколлегию 27.10.86