



Государственный комитет  
СССР  
по делам изобретений  
и открытий

патентно-техническая  
библиотека  
**О П И С А Н И Е**  
**ИЗОБРЕТЕНИЯ**

**К АВТОРСКОМУ СВИДЕТЕЛЬСТВУ**

(61) Дополнительное к авт. свид-ву —  
(22) Заявлено 25.12.77 (21) 2562179/18-24  
с присоединением заявки № —  
(23) Приоритет —  
Опубликовано 30.12.80. Бюллетень № 48  
Дата опубликования описания: 30.12.80

(11) 792247

(51) М. Кл.<sup>3</sup>  
G 06 F 3/16

(53) УДК 681.327.  
.11(088.8)

(72) Авторы  
изобретения:

В. А. Свириденко и В. С. Звездин

(71) Заявитель

Московский институт инженеров гражданской авиации

(54) УСТРОЙСТВО ДЛЯ ВВОДА И ВЫВОДА РЕЧЕВОЙ  
ИНФОРМАЦИИ

Изобретение относится к области речевой связи человека с машинами.

Известны устройства для передачи и приема речевых сигналов на основе идентификации параметров речеобразующего тракта (1, 2) методами линейного предсказания по критерию минимума дисперсии ошибки предсказания [1] и [2].

Наиболее близким по технической сущности является вокодерное устройство, в котором предполагается, что модель речеобразующего тракта, структура которой одновременно является и структурой синтезатора речи на приемной стороне, содержит генератор возбуждающей функции  $e(t)$ , которая может быть или шумовой ( $e(t)=e_w(t)$ ) для невокализованных участков речи или квазипериодической импульсной последовательностью ( $e(t)=e_n(t)$ ) с частотой основного тона  $f_{от}$  для невокализованных участков речи [3]. В соответствии с этим такой синтезатор состоит из генератора белого шума с равномерным спектром и генератора квазипериодических импульсов. В зависимости от вида возбуждающей функции к одному из указанных ге-

нераторов подключается с помощью переключателя "тон-шум" усилитель, регулируемый коэффициент  $A$  усиления которого определяет интенсивность соответствующих участков речи. Выходной сигнал усилителя  $E(t)=A e(t)$  подается на первый вход сумматора, выход которого через линейный фильтр с регулируемыми параметрами  $\alpha = (\alpha_1, \dots, \alpha_p)$  подключен ко входу второго сумматора. Указанные сумматор и фильтр образуют линейную систему с переменными параметрами, модулирующую речеобразующий голосовой тракт. Выходной сигнал этой линейной системы является речевым сигналом  $S(t)$ .

На передающей стороне анализатор речевого сигнала по его дискретизированной копии  $\{S_i\}$  дает оценки ряда параметров, что выполняет специальные оценщики, функционирующие по известному алгоритму. В их число входит устройство оценки соотношения "тон-шум", устройство оценки интенсивности  $A$  участка речи, выделитель основного тона, идентификатор вектора параметров. Последний определяет оценку  $\bar{\alpha}$  вектора  $\bar{\alpha}$  в соот-

ветствии с алгоритмом минимума среднеквадратической ошибки.

С выхода анализатора через уплотнитель в канал (далее он предполагается дискретным) передаются сигнал классификации "тон-шум" (например 0 — тон, 1 — шум) оценка  $A^*$  интенсивности анализируемого участка речи, оценка  $f_{от}$  на выходе блока выделения основного тона, оценка  $\vec{d}^*$  вектора  $\vec{d}$  с выхода блока идентификации параметров, на вход которого подается сигнал  $\{s_i\}$  и ошибка  $\{e_i\}$ .

На приемной стороне принятые оценки тон-шум  $A^*$ ,  $f_{от}^*$ ,  $\vec{d}^*$  после их разделения в блоке разделения сигналов подаются в соответствующие блоки синтезатора: на переключатель "тон-шум", усилитель с регулируемым коэффициентом усиления, генератор квазипериодических импульсов, линейный фильтр с регулируемыми параметрами. Структура синтезатора, как указывалось, полностью повторяет модель речеобразующего тракта.

Такой вокодер не позволяет добиться высокого качества системы синтетической телефонии в силу отсутствия в ней компонента моделирующих участков речевого сигнала, соответствующих носовым звукам. Это снижает узнаваемость говорящего, речевое сообщение которого передается по такой системе синтетической телефонии.

Целью изобретения является улучшение качества вводимой и выводимой информации путем повышения степени идентификации сигналов.

Поставленная цель достигается тем, что в устройство для ввода и вывода речевой информации, содержащее первый линейный фильтр, подключенный к первому вычитателю, и последовательно соединенные преобразователь речевых сигналов, микрофон, аналого-цифровой преобразователь, первый дискриминатор, блок уплотнения сигналов, параллельно подключенный через блоки измерения интенсивности речевого сигнала и выделения основного тона к аналого-цифровому преобразователю, блок сопряжения, блок разделения сигналов, подключенный через второй линейный фильтр к первому сумматору, первый генератор, первый переключатель, подключенный ко второму генератору и блоку разделения сигналов, и усилитель, подключенный к блоку разделения сигналов, дополнительно введены последовательно соединенные второй переключатель, подключенный к аналого-цифровому преобразователю и блоку уплотнения сигналов, второй вычитатель, третий вычитатель, третий линейный фильтр, первый блок идентификации параметров, подключенный к третьему вычитателю и первому блоку идентификации параметров, второй дискриминатор, подключенный к

аналого-цифровому преобразователю и блоку уплотнения сигналов, третий переключатель, подключенный ко второму переключателю, второй блок идентификации параметров, подключенный к первому линейному фильтру, блоку уплотнения сигналов и второму вычитателю, четвертый линейный фильтр, подключенный к первому вычитателю, четвертый вычитатель, подключенный ко второму переключателю, второму вычитателю и первому блоку идентификации параметров, второй сумматор, подключенный к третьему переключателю, и пятый линейный фильтр, подключенный ко второму дискриминатору, и последовательно соединенные четвертый переключатель, подключенный к усилителю, блоку разделения сигналов и первому сумматору, третий сумматор, шестой линейный фильтр, подключенный к блоку разделения сигналов, четвертый сумматор, подключенный к выходной шине, и седьмой линейный фильтр, подключенный к блоку разделения сигналов, первому сумматору и второму линейному фильтру.

Предлагаемое устройство, блок-схема которого представлена на чертеже, содержит аналого-цифровой преобразователь 1, второй переключатель 2, четвертый вычитатель 3, четвертый линейный фильтр 4, первый вычитатель 5, первый линейный фильтр 6, третий переключатель 7, пятый линейный фильтр 8, второй блок 9 идентификации параметров, второй вычитатель 10, третий вычитатель 11, третий линейный фильтр 12, первый блок 13 идентификации параметров, первый дискриминатор 14, блок 15 измерения интенсивности сигнала, блок 16 выделения основного тона, второй дискриминатор 17, блок 18 сопряжения, второй генератор 19, первый генератор 20, первый переключатель 21, усилитель 22, четвертый переключатель 23, первый сумматор 24, второй линейный фильтр 25, седьмой линейный фильтр 26, четвертый сумматор 27, шестой линейный фильтр 28, блок 29 уплотнения сигналов, блок 30 разделения сигналов, второй сумматор 31, третий сумматор 32, выходная шина 33 и микрофон 34.

Предложенное устройство работает следующим образом.

Речевой сигнал  $S(t)$  подвергается обработке на приемной стороне в анализаторе с целью оценки по нему параметров модели речеобразующего тракта. При цифровой реализации предлагаемого устройства сигнал  $S(t)$  проходит блок 1, где дискредитируется с частотой  $f_d$  (обычно 6–20 кГц) и квантуется. Полученная последовательность  $\{s_i\}$  проходит через схему образования оценки функции возбуждения  $E(t)$  для случая прохождения  $E(t)$  че-

рез модель голосового тракта или  $E(t)$  для случая прохождения  $E(t)$  через модель носового тракта.

Рассмотрим  $\tilde{E}(t)$ . Переключатель 2 стоит в положении голосовой тракт. В вычитателе 3 образуется  $\tilde{Y}(t)$  — оценка  $\tilde{Y}(t) = S(t) - \tilde{Z}(t)$ , а  $\tilde{Z}(t)$  — оценка  $Z(t)$ , образуемая на выходе сумматора 31. Вычитание  $\tilde{Z}(t)$  необходимо для устранения влияния отключенного носового тракта. Сигнал  $\tilde{Z}(t)$  проходит линейный фильтр 4. Далее полученный на выходе фильтра сигнал  $\tilde{X}(t)$  проходит вычитатель 5 и линейный фильтр 6 цепи обратной связи. Вычитатель 5 и линейный фильтр 6 выполняют роль линейного предсказателя. Точно такую же роль играют вычитатель 11 и линейный фильтр 12. В вычитателе 10 образуется сигнал  $\tilde{Z}(t) = S(t) - \tilde{Y}(t)$ , где устраняется влияние отключенного голосового тракта.

Блоки 9 и 13 идентификации параметров на основе анализа соответственно сигналов  $\tilde{Y}(t)$  и  $\tilde{Z}(t)$ , а также  $\tilde{E}(t)$  и  $E(t)$  производят оценки параметров линейных фильтров  $\tilde{\alpha}^*$ ,  $\tilde{\beta}^*$  и  $\tilde{\gamma}^*$ . Эти оценки управляют параметрами линейных фильтров 4 и 6 (для  $\tilde{\alpha}^*$  и  $\tilde{\beta}^*$ ) и линейных фильтров 8 и 12 (для  $\tilde{\gamma}^*$ ). Новым является второй дискриминатор 17. Его работа в простейшем случае: может быть основана на том, что непосредственно перед лицом говорящего ставят специальный микрофон с двумя раздельными секциями, направленными на нос и рот. Более интенсивный сигнал в одной из секций и определяет положение переключателя. Другое решение — включение избирательного фильтра в схему переключателя (такой случай изображен на чертеже), характеристика которого моделирует носовую полость («фильтр, согласованный с линейной системой, образующий сигнал  $y(t)$  из  $E(t)$ ). Превышение за интервал анализа уровнем сигнала заданного порога показывает, что  $E(t)$  проходит в течение этого интервала через носовую область. Выходной сигнал дискриминатора 17 подается на спаренные переключатели 2 и 7.

Оценки векторов  $\tilde{\alpha}^*$ ,  $\tilde{\beta}^*$  и  $\tilde{\gamma}^*$ , а также выходные сигналы дискриминаторов 14 и 17, а также блоков 15 и 16 подаются в блок 18 сопряжения через блок 29 уплотнения сигналов.

Рассмотрим работу синтезатора на приемной стороне, параметры которого изменяются под воздействием соответствующих сигналов, принимаемых из канала.

Функция возбуждения  $e_{ш}(t)$ , соответствующая вокализованным участкам речи, генерируется блоком 20. Частота следования импульсов  $f_{от}$  соответствует частоте основного тона и может меняться в требуемых пределах.

Функция возбуждения  $e_{ш}(t)$  для невокализованных участков речи формируется генератором 19, на выходе которого появляется случайный процесс с равномерным спектром. Переключатель 21 выбирает в соответствии с участком речи тип функции возбуждения  $e(t)$  путем подключения к формирователю речевого сигнала генераторов 19 или 20. После усиления в блоке 22 функция  $E(t) = A e(t)$  подается на переключатель 23. Он выполняет роль мягкого неба, отключая блоки, моделирующие голосовой тракт в момент произнесения носовых звуков. Функция  $E(t)$  с выхода сумматора 24 подается на линейный фильтр 8, вектор параметров  $\tilde{\beta} = (\beta_1, \dots, \beta_M)$  которого подбирается так, чтобы скорректировать требуемым образом сквозную характеристику, моделирующую форму голосового тракта. Параметры  $\tilde{\alpha}$  и  $\tilde{\beta}$  отражают положение артикуляционных органов, причем вектор  $\tilde{\beta}$  определяет в некоторой мере характер коартикуляции и взаимодействие голосового и носового трактов. В другом положении переключателя 23 (стрелка вниз) сигнал  $E(t)$  подается на линейный фильтр образованный сумматором 32 и фильтром с параметрами  $\tilde{\gamma} = (\gamma_1, \dots, \gamma_L)$  в петле обратной связи (фильтр 28). В простейшем случае параметры  $\tilde{\gamma}_i$  постоянны, если считать, что форма носовой полости неизменна. На выходе сумматора 27 образуется речевой сигнал  $S(t) = Z(t) + y(t)$ , где  $Z(t)$  — сигнал на выходе линейного тракта, моделирующего носовой тракт, а  $y(t)$  — выходной сигнал фильтра 26.

Параметрами генератора 20, положением переключателей 21 и 23, усилением блока 22, фильтров 26, 27 и 28 управляют соответственно переданные через блок 18 выходные сигналы блока 16, дискриминаторов 14 и 17, блока 15, блоков идентификации параметров 9 (оценки  $\tilde{\alpha}^*$  и  $\tilde{\beta}^*$ ) и 13 (оценки  $\tilde{\gamma}^*$ ). Указанные сигналы подаются в соответствующие блоки после их разделения в блоке 30 разделения сигналов.

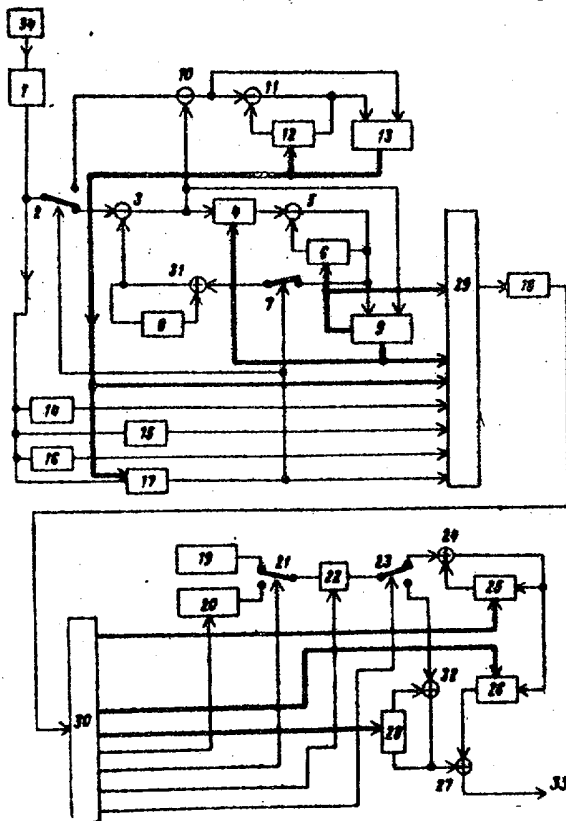
Предложенное устройство для ввода и вывода речевых сообщений позволяет существенно повысить качество системы при незначительном увеличении скорости передачи цифровой информации об оценках параметров модели речеобразующего тракта, что обеспечит более высокую степень идентификации говорящего. При использовании предложенного устройства в системах речевого общения человека и ЭВМ значительно расширяются возможности систем путем улучшения характеристик идентификации дикторов или их верификации.

## Формула изобретения.

Устройство для ввода и вывода речевой информации, содержащее первый линейный фильтр, подключенный к первому вычитателю, и последовательно соединенные преобразователь речевых сигналов, микрофон, аналого-цифровой преобразователь, первый дискриминатор, блок уплотнения сигналов, параллельно подключенный через блоки измерения интенсивности речевого сигнала и выделения основного тона к аналого-цифровому преобразователю, блок сопряжения, блок разделения сигналов, подключенный через второй линейный фильтр к первому сумматору, первый генератор, первый переключатель, подключенный ко второму генератору и блоку разделения сигналов, и усилитель, подключенный к блоку разделения сигналов, отличающееся тем, что, с целью улучшения качества вводимой и выводимой информации путем повышения степени идентификации сигналов, в устройство введены последовательно соединенные второй переключатель, подключенный к аналого-цифровому преобразователю и блоку уплотнения сигналов, второй вычитатель, третий вычитатель, третий линейный фильтр, первый блок идентификации параметров, подключенный к третьему вычитателю и первому блоку идентификации параметров, второй дискриминатор, подключенный к аналого-цифровому преобразователю и блоку уплотнения сигналов, третий

переключатель, подключенный ко второму переключателю, второй блок идентификации параметров, подключенный к первому линейному фильтру, блоку уплотнения сигналов и второму вычитателю, четвертый линейный фильтр, подключенный к первому вычитателю, четвертый вычитатель, подключенный ко второму переключателю, второму вычитателю и первому блоку идентификации параметров, второй сумматор, подключенный к третьему переключателю, и пятый линейный фильтр, подключенный ко второму дискриминатору, и последовательно соединенные четвертый переключатель, подключенный к усилителю, блоку разделения сигналов и первому сумматору, третий сумматор, шестой линейный фильтр, подключенный к блоку разделения сигналов, четвертый сумматор, подключенный к выходной шине, и седьмой линейный фильтр, подключенный к блоку разделения сигналов, первому сумматору и второму линейному фильтру.

- Источники информации, принятые во внимание при экспертизе
1. Шефер, Рабинер. Цифровое представление речевых сигналов. - "ТИИЭР", т.63, № 4, 1975. с.141-159.
  2. Шредер. Вокодеры, анализ и синтез речи. - "ТИИЭР", т. 54, № 5, 1966.
  3. Flanagan I.L. Speech Analysis, Synthesis and Perception, 2-nd ed. Springer-Verlog. Serlin, 1972 (прототип).



ВНИИПИ Заказ 9434/48  
Тираж 751 Подписное

Филиал ППП "Патент",  
г. Ужгород, ул. Проектная, 4