

ВЫДЕЛЕНИЕ ПРИЗНАКОВ ДЛЯ РАЗЛИЧЕНИЯ ОГРАНИЧЕННОГО НАБОРА СЛОВ

Н. Ф. КУЧЕР, А. Н. ОСОКИН

(Представлена научным семинаром кафедры вычислительной техники)

В работе [1] показано, что для различения ограниченного набора слов (около 50), достаточного для устного управления ЭЦВМ, достаточно выделять признаки: звонкие смычки, глухие смычки, звонкие согласные и гласные, глухие согласные. Для формирования указанных признаков необходимо сформировать из речевого сигнала клиппированную огибающую речевого сигнала и клиппированную огибающую всего слова.

Для формирования клиппированной огибающей речевого сигнала необходимо выделить огибающую речевого сигнала, представляемого в форме

$$S(t) = A(t) \cos \varphi(t). \quad (1)$$

Огибающая $A(t)$ может быть выражена через сопряженный речевому сигнал $\hat{S}(t)$ следующим образом:

$$A(t) = \sqrt{S^2(t) + \hat{S}^2(t)}, \quad (2)$$

где

$$\hat{S}(t) = H[S(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{S(\tau)}{t - \tau} d\tau - \quad (3)$$

— преобразование Гильберта сигнала $S(t)$. Однако определение $A(t)$, согласно (2), затруднено вследствие неоднозначности, возникающей при проведении операции извлечения квадратного корня. Кроме того, аппаратная реализация операций (2) и (3) оказывается весьма сложной или даже вообще невозможной для широкополосных сигналов.

Речевой сигнал, представленный в форме (1), представляет собой процесс с амплитудной и фазовой модуляцией, поэтому огибающая $A(t)$ могла бы быть выделена с помощью амплитудного детектирования, причем операция детектирования легко реализуется аппаратными методами.

Покажем применимость детектирования к речевым сигналам.

В теории и практике радиоприема и измерительной техники для выделения огибающих сигналов используются линейные одно и двухполупериодные и квадратичные детекторы [2]. В большинстве случаев качественное выделение огибающей обеспечивает линейный однополупериодный детектор. Рассмотрим применение такого детектора для выделения огибающей $A(t)$ речевого сигнала $S(t)$.

Определим отклик $S_1(t)$ линейного детектора на речевой сигнал $S(t)$. Линейный однополупериодный детектор является нелинейным элементом с характеристикой, представленной на рис. 1.

Для решения задачи нахождения отклика воспользуемся методом преобразований [3]. Введем следующие обозначения:

$$\begin{aligned} S(t) &= x \\ S_1(t) &= y \\ A(t) &= A \\ \varphi(t) &= \varphi. \end{aligned} \quad (4)$$

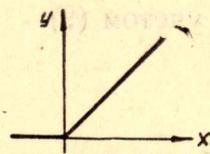


Рис. 1

Тогда характеристика нелинейного элемента-детектора примет вид:

$$y = g(x), \quad (5)$$

где в соответствии с рис 1.

$$\begin{aligned} g_+(x) &= y = x \text{ при } x > 0 \\ g_-(x) &= y = 0 \text{ при } x \leq 0. \end{aligned} \quad (6)$$

В соответствии с методом преобразований, найдем изображение $f(p)$ характеристики линейного детектора

$$\begin{aligned} f(p) &= f_+(p) + f_-(p) = \int_0^{\infty} g_+(x) e^{-px} dx + \int_{-\infty}^0 g_-(x) e^{-px} dx = \\ &= \int_0^{\infty} x e^{-px} dx + \int_{-\infty}^0 0 \cdot e^{-px} dx = \int_0^{\infty} x e^{-px} dx. \end{aligned}$$

Обозначим:

$$px = t; \quad dx = \frac{dt}{p}.$$

Тогда

$$f(p) = \int_0^{\infty} \frac{t}{p} e^{-pt} \frac{1}{p} dt = \frac{1}{p^2} \int_0^{\infty} t e^{-t} dt = \frac{1}{p^2} \Gamma_{(2)} = \frac{1}{p^2} \cdot 1!, \quad (7)$$

где $\Gamma_{(2)}$ — гамма-функция.

Отклик однополупериодного линейного детектора

$$y = g(x) = \frac{1}{2\pi j} \int_c f(p) e^{px} dp = \frac{1}{2\pi j} \left[\int_{c_+} f_+(p) e^{px} dp + \int_{c_-} f_-(p) e^{px} dp \right].$$

Так как $f_-(p) = 0$, то

$$y = g(x) = \frac{1}{2\pi j} \int_{c_+} f_+(p) e^{px} dp,$$

где c_+ — контур интегрирования.

Учитывая, что из (4) $x = A \cos \varphi$, получим

$$y = g(x) = \frac{1}{2\pi j} \int_{c_+} \frac{1}{p^2} e^{pA \cos \varphi} dp. \quad (8)$$

Интеграл вида (8) вычисляется приближенно. Для этого экспонента раскладывается в ряд по формуле Якоби — Энгера

$$e^{(pA) \cdot \cos \varphi} = \sum_{m=0}^{\infty} \varepsilon_m I_m(pA) \cos m\varphi, \quad (9)$$

где $I_m(pA)$ — модифицированная функция Бесселя первого рода m -го порядка;

ε_m — множитель Неймана;

$$\varepsilon_m = \begin{cases} 1 & \text{при } m = 0 \\ 2 & \text{при } m = 1, 2, 3, \dots \end{cases} \quad (10)$$

С учетом (9)

$$y = \frac{1}{2\pi j} \int_{c_+} \frac{1}{p^2} \sum_{m=0}^{\infty} I_m(pA) \cos m\varphi dp.$$

Обозначим:

$$pA = \omega, \quad p = \frac{\omega}{A}, \quad dp = \frac{1}{A} d\omega,$$

то

$$\begin{aligned} y &= \sum_{m=0}^{\infty} \cos m\varphi \frac{\varepsilon_m}{2\pi j} \int_{c_+} \frac{A^2}{\omega^2} I_m(\omega) \frac{d\omega}{A} = \sum_{m=0}^{\infty} A \cos m\varphi \frac{\varepsilon_m}{2\pi j} \int_{c_+} \frac{I_m(\omega)}{\omega^2} d\omega = \\ &= \sum_{m=0}^{\infty} c_m A \cos m\varphi, \end{aligned} \quad (11)$$

где

$$c_m = \frac{\varepsilon_m}{2\pi j} \int_{c_+} \frac{I_m(\omega)}{\omega^2} d\omega \text{ — константа.} \quad (12)$$

Таким образом, отклик $S_1(t)$ линейного однополупериодного детектора на речевой сигнал $S(t)$ выглядит следующим образом:

$$S_1(t) = y = \sum_{m=0}^{\infty} c_m A(t) \cos m\varphi(t) = c_0 A(t) + \sum_{m=1}^{\infty} c_m A(t) \cos m\varphi(t). \quad (13)$$

В работе [3] дано общее решение интеграла (12) в виде

$$c_m = \frac{\varepsilon_m}{\pi} \cdot \frac{\cos m\pi/2}{1 - m^2}. \quad (14)$$

В частности, при $m = 0$, $\varepsilon_m = 1$

$$c_0 = \frac{1}{\pi} \frac{\cos 0}{1} = \frac{1}{\pi}.$$

Из (13) следует, что для получения искомой огибающей $A(t)$ спектры $A(t)$ и $S(t)$ не должны перекрываться, то есть сигнал $S(t)$ должен быть узкополосным. Только в этом случае будет возможно отфильтровать составляющую отклика $C_0 A(t)$ с помощью фильтра низких частот. Однако речевой сигнал является широкополосным, так как для него не выполняется условие узкополосности

$$\frac{\Delta F}{f_0} \ll 1,$$

где ΔF — ширина спектра сигнала;

f_0 — центральная частота спектра сигнала.

Следовательно, успешное применение линейного однополупериодного детектора для выделения огибающей амплитуды речевого сигнала требует предварительного преобразования исходного широкополосного речевого сигнала в узкополосный. Такое преобразование осуществляется с помощью однополосной модуляции речевого сигнала.

Структурная схема устройства для выделения клипированной огибающей приведена на рис. 2 [4]. Речевой сигнал воспринимается микрофоном. С помощью блока однополосной модуляции исходный речевой сигнал, поступающий с микрофона, преобразуется в однополосный

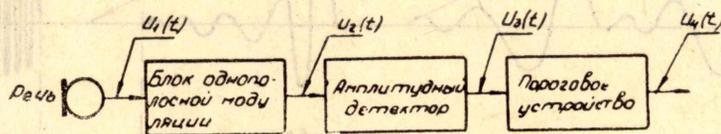


Рис. 2

(узкополосный), огибающая которого пропадает во время появления смычки в произносимой перед микрофоном речевой последовательности. Эта огибающая выделяется амплитудным детектором, постоянная времени которого определяется уже не длительностью смычки и участка наличия звука, а главным образом только величиной несущей частоты однополосного сигнала. Наличие или отсутствие огибающей фиксируется пороговым устройством, на выходе которого имеем прямоугольные импульсы с амплитудой, равной нулю в момент наличия смычки в речевом сигнале, т. е. получаем клипированную огибающую речевого сигнала.

На временной диаграмме (рис. 3) показаны

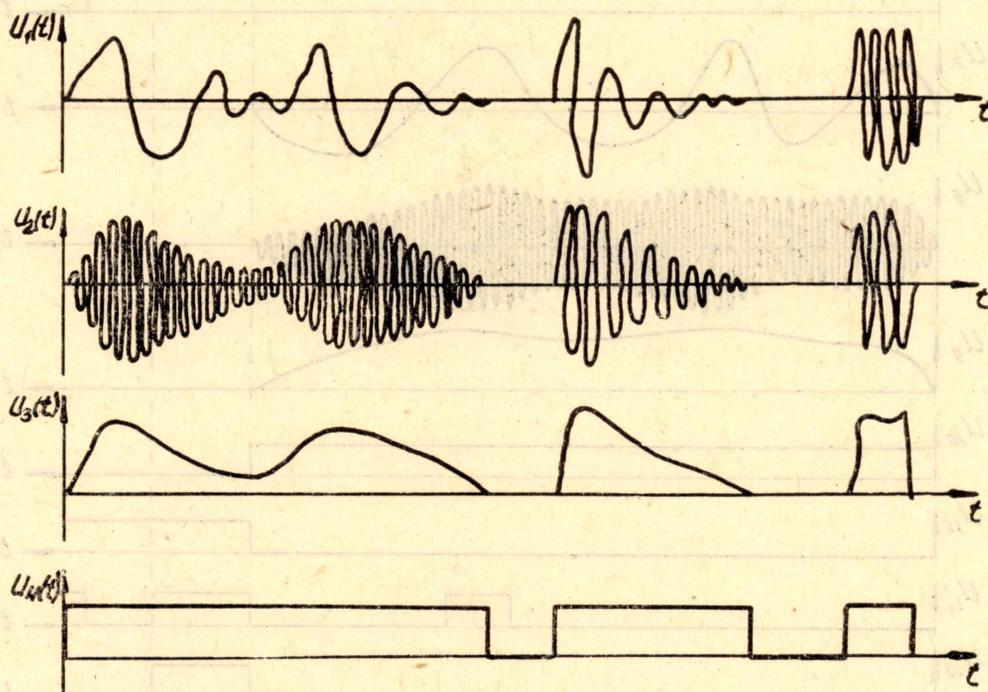


Рис. 3.

- $U_1(t)$ — речевой сигнал;
- $U_2(t)$ — речевой сигнал после однополосной модуляции;
- $U_3(t)$ — огибающая $A(t)$ речевого сигнала;
- $U_4(t)$ — клипированная огибающая речевого сигнала.

Огибающая всего слова выделяется аналогично огибающей речевого сигнала, только постоянная времени амплитудного детектора выбирается большей.

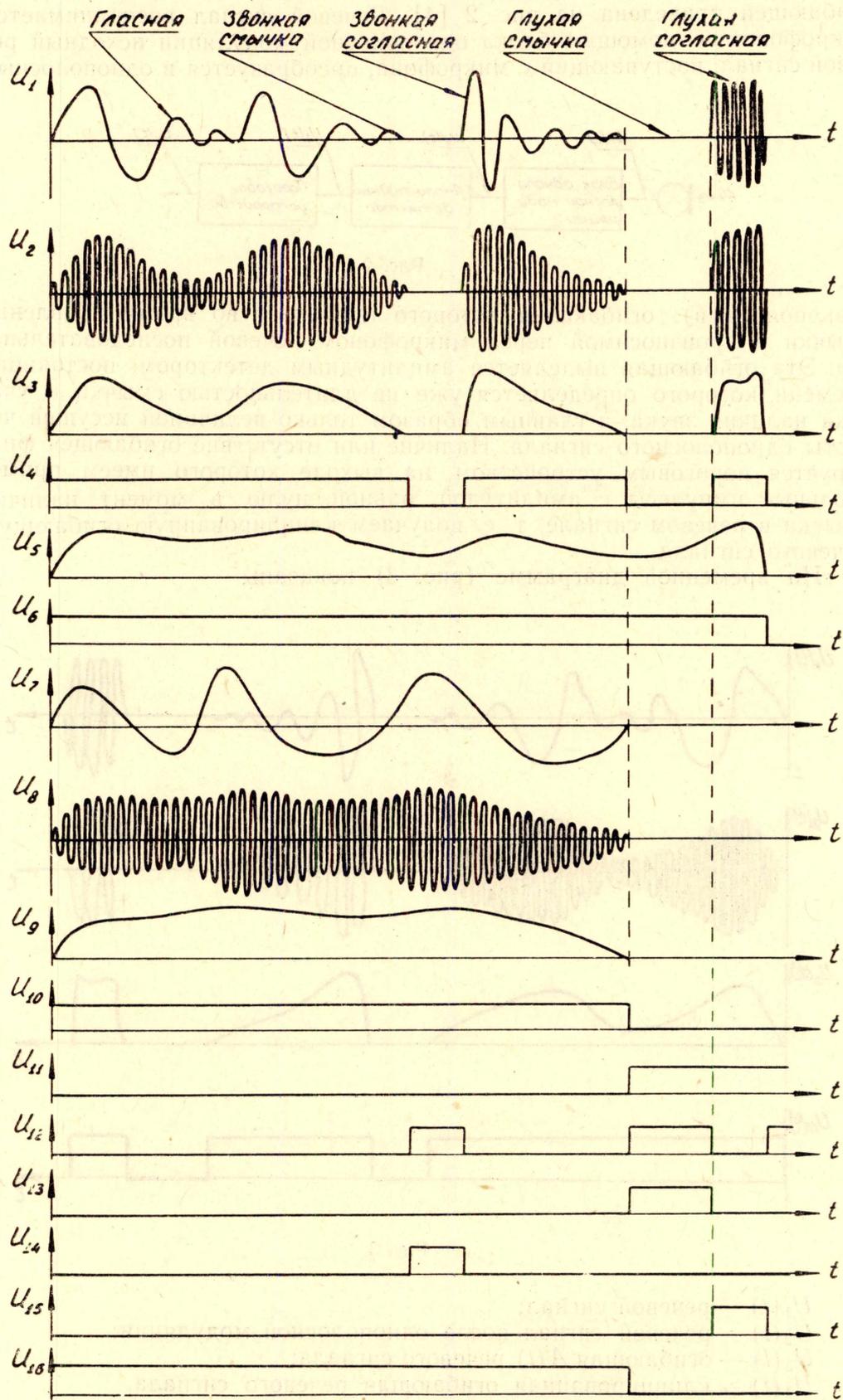


Рис. 4

Для съема сигналов основного тона воспользуемся методом, описанным в [5].

Выделение клиппированной огибающей основного тона схоже с описанным выше выделением клиппированной огибающей речевого сигнала.

Структурная схема устройства выделения признаков приведена на рис. 4. Видно, что схема состоит из формирователей трех ранее указанных сигналов и логической части, которая работает в соответствии с выражениями, приведенными в [1].

Из ряда источников известно [6], что звонкие смычки перед звуками б, д, г характеризуются пропаданием речевого сигнала. Однако, в отличие от глухой смычки перед звуками п, т, к, ц, ч при произнесении звонкой смычки присутствует основной тон (колеблются голосовые связки), что является отличительным признаком звонкой смычки. Звонкие согласные и гласные характеризуются наличием основного тона и огибающей речевого сигнала во время произнесения слова. Глухие согласные характеризуются наличием огибающей, но отсутствием основного тона при произнесении.

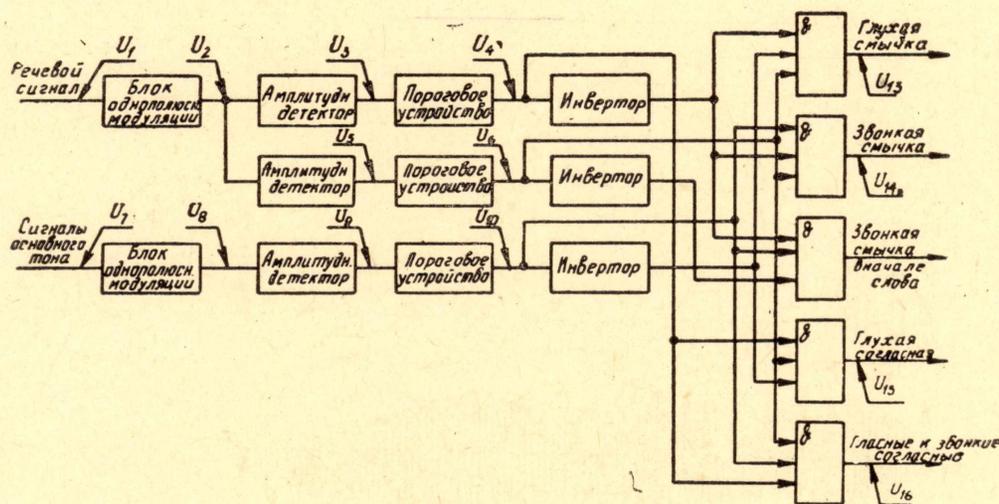


Рис. 5

Временная диаграмма работы устройства приведена на рис. 4, где обозначено:

- U_1 — речевой сигнал;
- U_2 — речевой сигнал после однополосной модуляции;
- U_3 — огибающая речевого сигнала;
- U_4 — клиппированная огибающая речевого сигнала;
- U_5 — огибающая всего слова;
- U_6 — клиппированная огибающая всего слова;
- U_7 — сигнал основного тона;
- U_8 — сигналы основного тона после однополосной модуляции;
- U_9 — огибающая основного тона;
- U_{10} — клиппированная огибающая основного тона;
- U_{11} — инвертированная клиппированная огибающая основного тона;
- U_{12} — инвертированная клиппированная огибающая речевого сигнала;
- U_{13} — признак глухой смычки;
- U_{14} — признак звонкой смычки;
- U_{15} — признак глухой согласной;
- U_{16} — признак звонкой согласной или гласной.

Устройство реализовано на интегральных микросхемах, кроме блока однополосной модуляции.

ЛИТЕРАТУРА

1. Н. Ф. Кучер, А. Н. Осокин. Выбор системы признаков для различения ограниченного набора слов. В данном сборнике.
2. И. С. Гоноровский. Радиотехнические цепи и сигналы. «Советское радио», 1971.
3. В. Давенпорт, В. Рут. Введение в теорию случайных сигналов и шумов. М., ИИЛ, 1962.
4. Н. Ф. Кучер, А. Н. Осокин. Авт. свидетельство № 534782. — Бюллетень «Изобретения, открытия, промышленные образцы, товарные знаки», 1976, № 41.
5. Б. Н. Елифанцев, Н. Ф. Кучер, А. Н. Осокин, В. М. Разин. Автоматическое различение ограниченного набора команд для управления станками. В сб. трудов Всесоюзного 6-летнего школы-семинара (АРСО-6). Таллин, изд-во АН ЭССР, 1972.
6. М. А. Сапожков. Речевой сигнал в кибернетике и связи. Связьиздат, 1963.

