

**ВЫБОР СИСТЕМЫ ПРИЗНАКОВ ДЛЯ РАЗЛИЧЕНИЯ
ОГРАНИЧЕННОГО НАБОРА СЛОВ**

Н. Ф. КУЧЕР, А. Н. ОСОКИН

(Представлена научным семинаром кафедры вычислительной техники)

Устройства ввода информации, использующие системы оптического считывания, выводные быстродействующие устройства, выводные быстродействующие устройства с представлением данных на ЭЛТ и микрофильмах и средства, обеспечивающие операции одновременно по вводу и выводу с использованием ЭЛТ в комбинации со световым пером, графопостроителей и речевой связи с ЭЦВМ считаются весьма перспективными в дальнейшем разрешении проблем связи оператора с машиной [1].

Наиболее близко к практическому выходу в настоящее время стоит задача различения ограниченного набора слов, разделяемых при произношении паузами. Различать можно также слитные фразы, рассматривая их как длинное слово.

Известно более 30 приборов, различающих ограниченный набор слов. К сожалению, ни один из них не нашел применения, так как не удовлетворяет указанным требованиям. Причина неудач обусловлена проблемой признаков различения, которые должны быть инвариантны к диктору, не зависеть от темпа произношения, надежно и просто выделяться аппаратурными методами, характеризовать фонетическую структуру произносимого слова.

В нашей работе [2] предложено для различения ограниченного набора слов использовать следующие признаки:

- 1) глухие смычки — C_{M1} ;
- 2) звонкие смычки — C_{M3} ;
- 3) глухие согласные — C ;
- 4) звонкие согласные и гласные — G .

Данные признаки в определенной степени характеризуют процесс речеобразования, что позволяет использовать для их выделения простые схемы. Для формирования данных признаков можно воспользоваться следующими сигналами, характеризующими процесс речеобразования:

1. Предельно-усиленной (клиппированной) огибающей сигналов основного тона (колебания голосовых связок при произнесении слов) — $x(t)$. Данная огибающая будет равна нулю при произнесении слова на участке глухих согласных, взрывных и фрикативных фонем, а именно: с, ш, ф, х, т, к, п, ч, ц. При произнесении всех звонких фонем обязательно колеблются голосовые связки. Поэтому огибающая сигнала основного тона будет отлична от нуля.

2. Клиппированной амплитудной огибающей $A(t)$ речевого сигнала $S(t)$, представленного в виде $S(t) = A(t) \cos \varphi(t) - y(t)$. Клип-

пированная амплитудная огибающая речевого сигнала будет равна нулю перед произнесением фонем п, т, к, ч, ц (глухая смычка) и фонем б, д, г (звонкая смычка). Следует отметить, что звонкая смычка в начале слова характеризуется отсутствием речевого сигнала и наличием колебаний гортани [3].

3. Огибающей всего слова — $z(t)$. Применение данного сигнала позволяет повысить надежность выделения признаков.

Указанные огибающие приведены на рис. 3 работы [4].

В принятых обозначениях исходных выделяемых сигналов признаки можно записать следующим образом:

$$\text{глухие смычки: } \text{См}_\Gamma = \overline{x(t)} \cdot \overline{y(t)} \cdot z(t);$$

$$\text{звонкие смычки: } \text{См}_3 = x(t) \cdot \overline{y(t)} \cdot z(t);$$

$$\text{звонкие смычки в начале слова: } \text{См}_3 = x(t) \cdot \overline{y(t)} \cdot \overline{z(t)};$$

$$\text{глухие согласные: } \text{С} = \overline{x(t)} \cdot y(t) \cdot z(t);$$

$$\text{гласные и звонкие согласные: } \Gamma = x(t) \cdot y(t) \cdot z(t).$$

Наибольшее количество информации о произнесенном слове будем получать, если в процессе произнесения слова с помощью выделяемых признаков составлять приближенную фонетическую структуру слова. Для этого при произнесении слова нужно не только выделять признаки, но и фиксировать последовательность их появления. Приведем фонетические структуры некоторых слов при предлагаемой системе признаков:

$$\text{ноль} \equiv \Gamma;$$

$$\text{один} \equiv \Gamma \text{См}_3 \Gamma;$$

$$\text{два} \equiv \text{См}_3 \Gamma;$$

$$\text{девять} \equiv \text{См}_3 \Gamma \text{См}_\Gamma \text{С};$$

$$\text{приоритет} \equiv \text{С} \Gamma \text{См}_\Gamma \text{С} \Gamma \text{См}_\Gamma \text{С};$$

$$\text{запятая} \equiv \Gamma \text{См}_\Gamma \text{С} \Gamma \text{См}_\Gamma \text{С} \Gamma;$$

$$\text{готовфототовводконец} \equiv \text{См}_3 \Gamma \text{См}_\Gamma \text{С} \Gamma \text{См}_\Gamma \text{С} \Gamma \text{С} \text{См}_3 \Gamma \text{См}_\Gamma \Gamma \text{См}_\Gamma \text{С}.$$

Из приведенных примеров видно, что некоторые слова характеризуются всего одним признаком. Это объясняется отсутствием предварительного членения речевого потока на отдельные фонемы. Так как в слове «ноль» все звуки звонкие, то признак звонкости «Г» будет единственным. Весьма желательно иметь с целью расширения набора различаемых слов устройство для предварительного (до выделения признаков) членения речи на фонемы.

Используя данные признаки, можно различать 25—30 слов. Если набор слов подобрать специально, то можно различать до 50 слов. Если произносить фразы как набор слитнопроизнесенных слов, то их можно рассматривать как длинные слова. В таком случае совокупность различаемых фраз гораздо больше.

ЛИТЕРАТУРА

1. Новые направления в технике ввода и вывода информации (обзор по материалам зарубежной печати). Сб.: «Средства вычислительной техники и оргтехники». Серия ТС-2, выпуск 2, Москва, 1972.

2. Н. Ф. Кучер, А. Н. Осокин, В. М. Разин. Об одном подходе к распознаванию ограниченного набора слов. Материалы I итоговой конференции НИИ автоматики и электромеханики, Изд-во Томского ЦНТИ, 1972.

3. М. А. Сапожков. Речевой сигнал в кибернетике и связи. Связьиздат, 1963.

4. Н. Ф. Кучер, А. Н. Осокин. Выделение признаков для различения ограниченного набора слов (данный сборник).