

УДК 621.391

ОПЫТ РАЗРАБОТКИ НЕАДАПТИВНОЙ СИСТЕМЫ РЕЧЕВОГО ВВОДА  
ДЛЯ УПРАВЛЕНИЯ БЫТОВЫМ МАГНИТОФОНОМ

В.В.Евграфова, С.Н.Кринов, В.Ф.Панин, В.Н.Трунин-Донской

Система речевого управления магнитофоном состоит из переносного пульта управления, инфракрасной линии связи фотоприемного устройства и обрабатывающей ЭВМ. В качестве элементной базы предлагается применить специализированные БИС однокристалльной микро-ЭВМ типа I-8748 и жесткой логики. В пульте управления находится кнопочная панель, устройство выделения параметров речевого сигнала (РС), преобразователь электрического сигнала в оптический с излучающей антенной. Устройство выделения параметров РС предполагается перевести на один кристалл, что определило четыре выделяемых параметра (первая и вторая форманты, число нулевых пересечений и общая интенсивность). Разработка устройства проводилась в следующие этапы: 1) разработка языка управления; 2) разработка устройства выделения параметров РС; 3) снятие статистических характеристик; 4) разработка программного обеспечения; 5) математическое моделирование.

1. В словарь речевого управления магнитофоном входит 20 управляющих слов, разделенных по смысловому принципу на несколько групп. Максимальный коэффициент ветвления равен 10. Слова "Сброс" и "Исправить" являются служебными. Они позволяют соответственно вернуться к началу работы с магнитофоном или изменить предыдущую команду. В пульте управления предусмотрены кнопочные дискретные команды для управления громкостью, тембром, стереобалансом и выключением магнитофона.

2. Разработка устройства выделения параметров РС велась с учетом перевода устройства на один кристалл и создания специализированной БИС. Об этом подробно изложено в [1].

3. Снятие статистических характеристик проводилось в ИППИ АН СССР десятью дикторами (7 мужчин и 3 женщины) и были получены 400 реализаций для 20 управляющих слов. Обработка статистических данных по параметрам интенсивности  $I$ , числа переходов через ноль  $NO$  и формантных частот  $F1$  и  $F2$  позволила выделить пять вторичных признаков  $PC$ , обозначенных следующими кодами:

код 1 -  $I \geq 2$ ,  $NO < 25$ ,  $F2 - F1 \geq 25^*$ ) - соответствует гласной "и",

код 2 -  $I \geq 2$ ,  $NO < 25$ ,  $F2 - F1 < 25$  - соответствует гласной "а",

код 3 -  $I < 2$ ,  $NO \geq 10$  - соответствует звонким и сонорным звукам,

код 4 -  $I \leq 2$ ,  $NO \geq 25$  и  $I > 2$ ,  $NO \geq 100$  - соответствует шипящим звукам,

код 5 -  $I < 2$ ,  $NO < 10$  - соответствует паузе.

4. По результатам статистических данных были разработаны две взаимосвязанные программы:

- программа автоматической сегментации  $PC$ ,

- программа распознавания управляющих слов по группам.

Программа автоматической сегментации производит анализ выделенных параметров  $I$ ,  $NO$ ,  $F1$ ,  $F2$  и формирование цепочки кодов (1-5) в порядке поступления исходной информации, т.е. каждые 20 мсек. Достоверным для формирования одного из кодов является наличие одного и того же признака в течение не менее 40 мсек, т.е. два значения подряд. Программа производит расчет количества однородных признаков, суммарного числа признаков в слове и количества слогов. Начало и конец слова определяются также автоматически. Началом слова является наличие одного из признаков в течение 60 мсек. Концом слова является наличие признаков паузы (код 5) в течение 160 мсек.

Программа распознавания управляющих слов построена по принципу логического дерева. Каждый блок программы предусмотрен для распознавания слов, относящихся к одной из 7 групп. Управление из одного блока в другой передается в соответствии с построением логического дерева. Объем разработанного программного обеспечения составил не более 1000x8 бит. Объем ОЗУ не более 60x8 бит. Время обработки - реальное.

5. Математическая модель программного обеспечения была разработана на языке "Ассемблер" ЭВМ "Электроника-60". Исходными

---

\* Значения  $I$ ,  $NO$ ,  $F1$ ,  $F2$  даны в условных единицах.

данными для работы программы являлись тесты, представляющие собой массивы данных, содержащих последовательную информацию со значениями I, NO, F1, F2 за каждые 20 мсек для каждого слова из словаря. Порядок следования слов друг за другом соответствует вал логическому дереву. В результате работы программы на дисплей выводились распознанное слово и цепочка кодов данного слова, сформированная программой автоматической сегментации. Управление передавалось следующему логическому блоку. В случае нераспознавания слова на дисплей выводилось слово "повторите" и управление передавалось на начало данного блока.

#### Результаты эксперимента и выводы.

1. На математической модели были проверены 200 реализаций 20 слов словаря, полученные пятью дикторами, не принимавшими участие в снятии статистических характеристик (3 мужчин, 2 женщины). Средний процент правильно опознанных слов составил 95.

2. Затраченный вычислительный ресурс обрабатывающей ЭВМ "Электроника-60" не превышает вычислительный ресурс I-8748 (объем ПЗУ - 1024x8, объем ОЗУ - 128x16 бит), что подтверждает возможность применения в системе однокристалльной микро-ЭВМ.

3. Аппаратные затраты на устройство выделения параметров РС составили не более 2000 активных элементов, что делает реальным возможность создания специализированной БИС жесткой логики для выделения параметров РС на КНОП технологии.

#### Л и т е р а т у р а

1. Панин В.Ф., Трунин-Донской В.Н., Чупров В.И. Речевое управление бытовой РЭА с использованием БИС. - В кн.: Автоматическое распознавание слуховых образов (АРСО-13). Тезисы докл. и сообщ. 13-й Всесоюз. школы-семинара. Ч.2. Новосибирск, 1984, с.169-171.

Поступила в ред.-изд.отд.  
12 апреля 1984 года