

АВТОМАТИЧЕСКОЕ РАЗЛИЧЕНИЕ СЛОВ «ПУСК» И «СТОП»

Б. Н. ЕПИФАНЦЕВ, Н. Ф. КУЧЕР, А. Н. ОСОКИН, В. М. РАЗИН

(Представлена кафедрой вычислительной техники)

Проблема объективного распознавания слуховых образов вызывает в последнее время большой интерес. Речь идет о возможности восприятия, или вернее, распознавания специальными устройствами речи, до последнего времени воспринимаемой только человеком [1]. Тенденция к автоматизации различного рода процессов вполне закономерно начинает проникать и в эту область человеческой деятельности.

Уже отмечалось [2], что решение проблемы построения устройств, способных распознавать связную речь (отдельные слова, фразы), в общем случае чрезвычайно осложняется различиями в степени ясности и отчетливости произношения. Как указывал академик Л. В. Щерба [3], возможно бесконечное количество переходных ступеней в произношении, начиная от абсолютной ясности и четкости до небрежной скороговорки, когда все неударные слоги наполовину «съедаются».

Как известно [2], к настоящему времени для большинства фонем не найдено достаточно четких инвариантных признаков, обеспечивающих хорошее распознавание этих звуков. Поэтому ближайшей задачей по проблеме управления голосом принято считать [6, 7] разработку и использование устройств, распознающих ограниченный набор фонем и их сочетаний, что, по существу, сводится к распознаванию ограниченного набора слов. По принятой в литературе [2] терминологии под распознаванием ограниченного набора понимается узнавание фонемы или слова, входящих в набор, причем на любые другие звуки устройство не реагирует. При различении ограниченного набора правильное узнавание осуществляется только при произнесении слов или фонем набора. На любые другие звуки устройство или не реагирует, или выдает какое-либо слово набора.

Устройство, различающее ограниченный набор слов, получается наиболее простым и надежным при использовании для распознавания так называемых «дифференциальных» признаков речевых сигналов, таких как наличие или отсутствие основного тона, шума, смычек и т. д. [4, 5, 6].

Нами был проведен поиск признаков, простых в анализе схемными методами и не зависящих от качеств произношения и смены дикторов. Из опубликованных источников [4] известно следующее:

а) при произнесении звука «С» — шумового, переднеязычного, отсутствует основной тон независимо от того, кто произносит этот звук — мужчина или женщина;

б) звуки «П», «Т», «К» — глухие согласные, взрывные — произносятся без участия голосовых связок;

в) звуки «О», «У» — гласные, губные; главный источник при произнесении этих звуков — колеблющиеся голосовые связки, т. е. при произнесении «О», «У» присутствует основной тон.

Анализ осциллограмм слов «пуск» и «стоп», произносимых различными дикторами, показал следующее:

1. В обоих исследуемых словах присутствует звонкий участок.
2. От начала речевого сигнала до появления основного тона, что соответствует началу звонкого участка в словах «пуск» и «стоп», имеется некоторый интервал времени.
3. Разность во времени от начала речевого сигнала до появления основного тона в слове «пуск» всегда меньше аналогичной разности в слове «стоп».

На основании вышеизложенного для различения выбраны следующие дифференциальные признаки слов «пуск» и «стоп», инвариантные по отношению к диктору:

1. Наличие или отсутствие звонкого участка в слове.
2. Наличие или отсутствие шумового участка в начале слова.
3. Величина временного интервала от начала слова до начала звонкого участка.

Авторами разработан алгоритм различения русских слов «пуск» и «стоп» и построено устройство, автоматически различающее указанные слова, функциональная схема которого приведена на рис. 1.

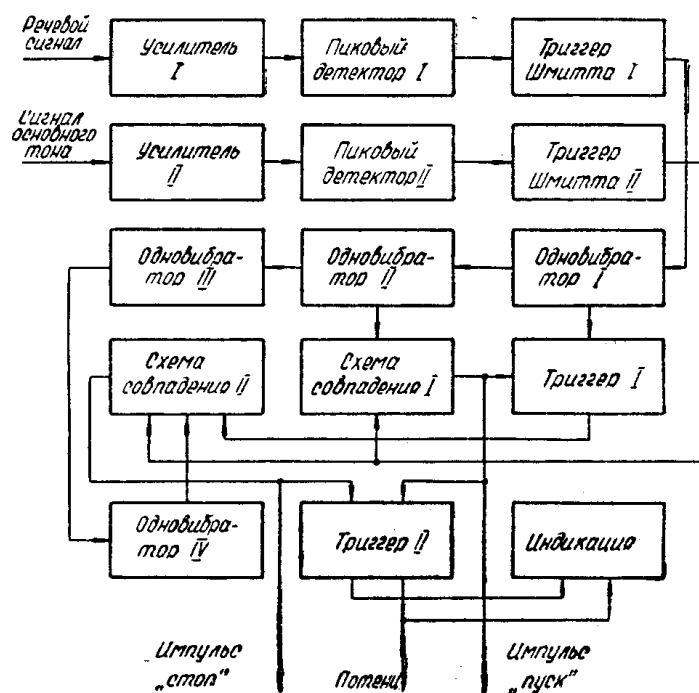


Рис. 1. Функциональная схема устройства, различающего слова «пуск» и «стоп»

С целью упрощения аппаратуры предлагается сигнал основного тона снимать с помощью ларингофона, помещенного на определенном участке шеи диктора. С целью увеличения помехозащищенности от посторонних звуков, что особенно важно при работе устройства в реальной обстановке, и обеспечения однотипности применяемых элементов речевой сигнал предлагается снимать с другого ларингофона, также помещенного на другом участке шеи.

Работа устройства происходит следующим образом:

Речевой сигнал усиливается усилителем 1. Пиковый детектор 1 выделяет огибающую речевого сигнала, из которой триггер Шмитта формирует прямоугольный импульс, равный по длительности произнесенному слову. Передним фронтом этого импульса запускается одновибратор 1, генерирующий импульс длительностью T_1 , несколько большей, чем длительность взрывного «П» в начале слова «пуск». Задним фронтом импульса одновибратора 1 запускается одновибратор 2, который генерирует импульс длительностью T_2 . Задним фронтом этого импульса запускается одновибратор 3, который генерирует импульс длительностью T_3 , причем сумма $T_1 + T_2 + T_3$ должна быть меньше, чем длительность участка «СТ» в слове «стоп». Задним фронтом этого импульса запускается одновибратор 4, генерирующий импульс длительностью T_4 .

Сигнал основного тона усиливается усилителем 2. Пиковый детектор 2 выделяет огибающую этого сигнала, а триггер Шмитта 2 формирует из нее импульсы, возникающие при наличии звонкого участка в слове и равные по длительности звонким участкам слов «пуск» и «стоп».

При произнесении слова «пуск» начало взрывного «П» соответствует переднему фронту импульсов триггера Шмитта 1, импульсы которого поступают на одновибратор 1. Одновибратор 1 генерирует импульс, которым запускается одновибратор 2. Импульс одновибратора 2 поступает на схему совпадения 1, на второй вход которой подается импульс триггера Шмитта 2. Если произнесено слово «пуск», то импульс одновибратора 2 совпадает во времени со звонким участком и на выходе схемы совпадения 1 появится импульс, соответствующий произнесенному слову «пуск», который переведет триггер 2 в единичное состояние, загорится лампочка световой сигнализации «пуск». Импульс с выхода схемы совпадения 1 переведет триггер 1 в состояние «0» и схема совпадения 2 окажется в закрытом состоянии. При произнесении слова «стоп» начало шумового «С» соответствует переднему фронту импульса триггера Шмитта 1, который запускает одновибратор 1. Импульсом этого одновибратора производится перевод триггера 1 в состояние «1» и схема совпадения 2 окажется подготовленной к дальнейшей работе. Этим же импульсом запускается одновибратор 2, который в свою очередь запускает одновибратор 3, запускающий одновибратор 4. Так как $T_1 + T_2 + T_3$ меньше, чем длительность участка «СТ» в слове «стоп», на выходе схемы совпадения 1 импульса не будет вследствие того, что импульс на выходе триггера Шмитта 2 появится только после окончания интервала времени $T_1 + T_2 + T_3$.

Совпадение импульсов одновибратора 4 и триггера Шмитта 2 в схеме совпадения 2 происходит во время звучания звонкого «О» в слове «стоп», на выходе схемы совпадения 2 появляется импульс, что соответствует наличию произнесенного слова «стоп». С выхода схемы совпадения 2 сигнал поступает на нулевой вход триггера 2, на единичном выходе которого потенциал будет близок к нулю, а потенциал, снимаемый с нулевого выхода, зажжет лампочку «стоп» световой сигнализации.

Длительности импульсов одновибраторов были подобраны экспериментально и составляют соответственно: $T_1 = 15$ мсек, $T_2 = 30$ мсек, $T_3 = 60$ мсек, $T_4 = 15$ мсек.

Для оценки работоспособности устройства при использовании его различными операторами были проведены следующие испытания. Произвольно была выбрана бригада дикторов (5 мужчин и 5 женщин). Они произносили 30 раз слово «пуск», затем 30 раз слово «стоп», затем 30 раз «пуск-стоп». Результат различения фиксировался по загораю-

щимся лампочкам и по импульсным выходам устройства с помощью двухлучевого осциллографа.

Результаты испытаний свидетельствуют о 100%-ной различаемости устройством указанных слов.

Использование двухканальной схемы устройства, а также применение одновибраторов 1, 3 с целью задержки позволяет получить повышенную помехоустойчивость работы устройства, так как устройство может сработать только при наличии сигнала с обоих ларингофонов, что возможно лишь при произнесении звонких звуков, имеющих в словах «пуск» и «стоп», а задержка, создаваемая одновибраторами 1 и 3, позволяет исключить ложное срабатывание устройства от кратковременных резких звуков, например, кашля. Принцип, положенный в основу работы устройства, как уже отмечалось, основан на свойствах структуры слов «пуск» и «стоп», которая является инвариантной по отношению к диктору. Поэтому результат различения не зависит от того, мужчина или женщина, один или несколько человек воспользуется устройством, так как индивидуальности голоса данное устройство не учитывает.

ЛИТЕРАТУРА

1. А. А. Харкевич. Оpozнание образов. «Радиотехника», № 5, 1959.
 2. Г. И. Цемель. Автоматическое различие органического набора слов. Известия АН СССР, ОТН, «Энергетика и автоматика», № 5, 1960.
 3. Л. В. Щерба. Избранные работы по русскому языку. Учпедгиз, 1957.
 4. Е. Т. Семенова, И. М. Тетельбаум. Признаки распознавания ограниченного набора слов. Сб. «Вопросы бионики», под редакцией М. Г. Гаазе-Рапопорта, «Наука», 1967.
 5. Л. А. Чистовиц и др. Речь, артикуляция и восприятие. «Наука», 1965.
 6. Е. Н. Мясникова. Объективное распознавание звуков речи. «Энергия», 1967.
 7. М. А. Сапожков. Речевой сигнал в кибернетике и связи. «Связьиздат», 1963.
-