

УДК 62-5:621.391:410.51.

О СЕТЕВОМ АЛГОРИТМЕ РАСПОЗНАВАНИЯ ФОНЕМ ПО ПОСЛЕДОВАТЕЛЬНОСТИ СЕГМЕНТОВ

Г.Я.Волошин, И.В.Бахмутова, А.А.Прокопенко

В работе [1] предложена методика использования языковой избыточности для повышения надежности автоматического распознавания речи. Эта методика сводит задачу к поиску кратчайшего пути на ориентированном графе. В той же работе показана применимость этой методики для распознавания фонем по последовательности ненадежно распознанных сегментов. В последнее время предложен другой [2], на наш взгляд, интересный алгоритм распознавания фонем по последовательности сегментов. В нем, как и в [1], используются данные о временных и лингвистических закономерностях русской речи, но методика их использования в части, касающейся временных данных, существенно иная, чем в [1]. Аналогичные методы с успехом используются в других конкретных постановках задач распознавания (в том числе для распознавания устных команд [3,4]). По всей вероятности, работа в данном направлении ведется и другими исследователями.

В связи с этим мы считаем целесообразным опубликовать результаты экспериментальной проверки сетевой методики распознавания речи по последовательности ненадежно распознанных сегментов. Под сегментом понимается участок речевого сигнала фиксированной длительности (мы выбирали его равным 20 мсек). Показания исходного автомата, распознающего сегменты, моделировались следующим способом (см. [1]).

Пусть на вход автомата подается непрерывный речевой сигнал. Он делится на участки длительностью 20 мсек и о каждом

из них выдается вероятное решение. Предполагалось, что каждый звук автомат может "перепутать" не более, чем с четырьмя другими. Каждому из конкурирующих пяти звуко-типов с помощью датчика случайных чисел приписывалась одна из следующих априорных вероятностей: 0,5 ; 0,3 ; 0,1 ; 0,07 ; 0,03.

Таким образом, на каждом сегменте получился столбец вероятности решения, а на достаточно длинном речевом сигнале - матрица, в которой с помощью алгоритма Форда искалась наиболее правдоподобная последовательность сегментов с учетом априорных вероятностей сочетаний двух соседних сегментов. Эти вероятности были вычислены по данным о средней длительности фонем русской речи и априорным вероятностям фонем и их парных сочетаний. После выделения наиболее вероятной последовательности производилось объединение рядом стоящих одинаковых элементов в один.

Группы близких звуко-типов (возможные ошибки исходного распознающего автомата) выбирались из фонетических соображений. Представление об этих группах можно получить по примеру матрицы (табл. I), вводимой в ЭЦВМ для обработки по сетевому алгоритму (элементы в столбцах стоят в порядке убывания априорных вероятностей)

Т а б л и ц а I

С	С	С	С	Ц	С	С	Ш	Ы	Й	Й	Й	Й	Э	Й	Й	Г	К	Г	Г	Н	М	М	М	Л	Б	Ь	Λ
Λ	Λ	С	Ц	Т	С	Т	Ш	Ы	Й	Й	Й	Й	А	А	А	Й	Й	К	Г	Б	Д	М	М	Н	Б	Ь	Λ
Λ	Б	А	Ь	Ц	С	Ц	Т	С	С	Ш	Ы	Й	А	Й	Э	А	Э	Б	В	Д	Б	М	Н	Н	Н	Н	Λ
А	Б	Ь	Б	А	Ш	Т	С	Ш	Ш	С	Т	Э	А	Э	А	Ы	Ы	Ы	Ы	Э	Ы	В	Д	К	В	Н	Н
Н	М	Б	А	А	А	Э	Б	Т	Ш	Ш	С	Т	Т	Ц	А	Й	А	Э	А	А	А	Й	Н	А	Д	Б	В
Б	В	Б	В	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А	А

"Сигма" (С'Й Г М Λ)

Предполагалось, что на вход автомата подаются произносимые вслух русские слова (словарь не ограничен).

Ниже приводятся примеры результатов экспериментов. В первой строке приведена искаженная последовательность (верхняя строка исходной матрицы), во второй - исправленная по сетевому алгоритму (и в первом, и во втором случае объединены стоящие рядом одинаковые сегменты)

I. Распад (Р Λ С П А Т):

Р Л Р Л Λ Б Ь Λ С С'Ц С С'С П А Э А А А Э А А А Т П

Р А С Т А Т

2. Игр (И Г Р):

И И И И И И И И И И Г К Г Р Л Р Л

И Г Р

3. Тип (Т И П):

Т Т И И И И И И И И П

Т И П

4. Адрес (А Д Р Е С):

А Э А А А Э А А А Д Б Г Д Р И Б Л Б Б С С С С С С С

А Д В И С

5. Член (Ч Л Е Н):

Ч Ш Ч Л Р Л Л Л А Э А И Э А Э А Э Н Н Н М

Ч Л Э Н

6. Транслятор (Т Р А Н С Л А Т Ъ Р):

Т П Р Л Р Л Л Б Л А Н Н Н М С С С С С С С Л Р Л Л А Э А Л А А
Э А Э А Т П Э Б Л Б Б Р Л Р Л

Т Р В Л М С Л А Т П Л Б Р

7. Остаток (А С Т А Т Ъ К):

Л Б Б А С С С С С С С Т П А Э А А А Э А А Т П Э Б Л Б Б Б Т К
А С Т А Т Ъ К

8. Регистр (Р И Е Г И С Т Р):

Р И Б Л Б Б Б Г Б Г Д Г И И И И И И Э И И И С С С С С С С Т П Р
Л Р Л

Р И Е Г И С Т Р

9. Мантисса (М А Н Т И С С А):

Н М М М Л Б Б Л Н М Н Н Н Т Т И И И И И И Э И И И С С С С С С С
С С С С С С Л Б Б Л

М А Н Т И С С А

Анализ результатов эксперимента показал следующее. Гипотетический распознающий автомат распознавал неправильно более 40% сегментов. После обработки матриц сетевым алгоритмом оказалось, что если принять истинное количество фонем за 100%, то появилось 7% "лишних" фонем, 0,4% фонем пропущено и 8% распознаны неправильно. Таким образом, оказались справедливыми предположения, высказанные в [1], а именно: пропусков границ между фонемами практически нет, зато достаточно много (7%) лишних границ членения. Надежность распознавания существенно возросла по сравнению с исходной матрицей (в пересчете на сегменты ошибки стали составлять менее 6% вместо 40% исходных). Теперь с большим основанием можно утверждать, что саль второго уровня (исправление ошибок распознавания фонем)

должна предусматривать возможность удаления "лишних" фонем (что легко делается) и не обязательно должна предусматривать восстановление пропущенных фонем.

Результаты экспериментальной проверки безусловно зависят от того, какого вида матрицы предъявляются для обработки, то есть от того, какой распознающий автомат будет работать на входе. Но модельные эксперименты, описанные здесь и в [1], показывают достаточно высокую потенциальную эффективность предложенного метода.

Л и т е р а т у р а

1. Г.Я.ВОЛЮНИН. Об использовании языковой избыточности для повышения надежности автоматического распознавания речи. - Вычислительные системы. Труды ИМ СО АН СССР, вып. 28, 1967.
2. Н.Г.ЗАГОРУЙКО. Распознавание фонем по последовательности сегментных решений. Доклад на Всесоюзной летней школе-семинаре "АРСО-1У", Канев, 1968.
3. В.А.КОВАЛЕВСКИЙ. Распознавание как оценка параметров процесса, порождающего сигналы. Конференция по теории автоматов и искусственному мышлению. (Аннотация к докладу) Ташкент, 1968.
4. Т.К.ВИНЦУК. Распознавание некоторых классов речевых сигналов. Автореферат, Киев, 1967.

Поступила в редакцию
15 января 1969 г.