

ФОРМИРОВАНИЕ ЭТАЛОНОВ ДЛЯ МОРФЕМНОГО РАСПОЗНАВАНИЯ РЕЧИ

Б.В. ПАНЧЕНКО

г. Минск, 220068, СССР

Рассматривается алгоритм формирования эталонных образов в системе морфемного распознавания слитных фраз, составленных из словоформ заданного словаря. В процессе обучения осуществляется символическое сравнение текстов слов словаря и с использованием морфологических правил формируется каталог морфем. На следующем этапе осуществляется ДП-сравнение речевых образов пар слов, имеющих в своем составе одинаковую морфему, с целью определения границ морфемных фрагментов в речевых образах слов, и образуется массив речевых образов морфем. Каталог эталонов морфем формируется с использованием процедуры кластеризации речевых образов одинаковых морфем.

ВВЕДЕНИЕ

В последнее время благодаря достижениям в области распознавания речи созданы достаточно эффективные системы, пригодные для практического применения. Они способны распознавать сравнительно небольшое количество акустически изолированных речевых команд, произнесенных известным диктором. В алгоритмах этих систем используется так называемый глобальный метод [1], основанный на универсальных алгоритмах распознавания сравнением распознаваемого образа с эталонными, как правило, без учета особенностей речевого сигнала. В качестве речевых элементов эти системы используют речевые команды, состоящие из одного или нескольких слов, рассматриваемых как единое целое.

Практическое использование этих систем ограничивается присутствием им недостатками: работа с одним диктором, ограниченность словаря команд, необходимость строгого соблюдения дикторской дисциплины, недостаточная комфортность работы с системой, необходимость дословного произнесения при распознавании фраз.

Стремление устранить отдельные недостатки глобального метода привело к разработкам алгоритмов распознавания слитнопроизнесенных команд и ключевых слов в речевом сигнале фраз. Практическое использование

этих алгоритмов значительно приблизило системы распознавания к пользователю. Однако распознавание слитнопроизнесенных фраз по эталонам, построенным по отдельным словам, осуществляется с недостаточной надежностью из-за эффекта слияния слов, требует больших вычислительных затрат и поэтому эффективно только на небольших подобранных словарях.

Ограниченность глобального метода сделало необходимым развитие работ по распознаванию речи в направлении разработки другого метода, называемого аналитическим, в котором используется представление речевого сообщения как композиции относительно небольшого числа базовых речевых элементов. Суть метода заключается в принятии на первом этапе решения о минимальных базовых речевых элементах сообщения и реконструкции сообщения по этим элементам на втором этапе с использованием правил грамматики и синтаксиса.

Таким образом, система распознавания, использующая этот метод, состоит из акустического процессора, в котором принимается решение о речевых элементах, и соединенного последовательно с ним лингвистического процессора, осуществляющего реконструкцию сообщения по последовательности элементов.

В разное время разными исследователями разрабатывались и проводились исследования речераспознающих систем с использованием трехуровневых лингвистических иерархий на основе эталонных слогов, полуслогов, фонем [2], [3], [4]. Для реконструкции слов (лексический уровень) из этих субсловесных единиц и предложений (третий уровень) применяется сложный грамматический и синтаксический анализ. Привлекательным в смысле минимальности алфавита используемых субсловесных единиц является пофонемное распознавание, однако необходимо отметить его чрезвычайную сложность, связанную с надежностью определения фонемных сегментов и вариативностью параметров фонем в речевом потоке.

МОРФЕМНОЕ РАСПОЗНАВАНИЕ

Не ставя вопрос о принципиальной достижимости высоконадежного фонемного распознавания речи, работающего в реальном масшта-

бе времени, мы считаем, что более скорый практический выход дадут системы, использующие более крупные субсловесные единицы, такие, например, как морфемы, являющиеся минимальными значимыми частями слов.

Выбор морфемы в качестве минимальной субсловесной единицы основан на следующих соображениях.

При использовании эталонов морфем нет необходимости в явном моделировании эффектов коартикуляции, поскольку они естественным образом присутствуют в относительно длинном отрезке речевого сигнала, соответствующего морфеме, т.е. акустический образ морфемы более устойчив к аллофоническим вариациям. Использование морфемного уровня позволяет значительно расширить лексические возможности распознающей системы, поскольку многообразие парадигматических форм слов может быть получено из ограниченного каталога морфем. По данным, приведенным в словаре морфем русского языка [5], для составления приблизительно 52000 слов, содержащихся в словаре, использовано около 5 тыс. морфем. Интересно отметить, что из них 260 наиболее продуктивных корневых морфем дают почти половину слов данного словаря (около 20 тысяч). Показателем эффективности использования морфем на первом уровне лингвистической иерархии является также индекс синтетичности языка, который определяет отношение числа морфем к числу слов в тексте, и для русского языка составляет 1,90 (для английского - 1,68).

Получаемый при работе распознающей системы на выходе акустического процессора кортеж основных морфем речевого сообщения позволяет эффективно использовать в лингвистическом процессоре хорошо разработанный набор относительно простых морфологических правил (правила словообразования, структурные закономерности между морфемами, фактор ветвления, показывающий число морфем, которые могут следовать за конкретной морфемой и т.д.).

Работа акустического процессора основана на разработанном Н.П. Дегтяревым [6] алгоритме поиска в речевом сигнале фразы ключевых слов заданного словаря. В нашей системе в качестве ключевых слов выступают морфемы и морфемные блоки. К используемому в акустическом процессоре каталогу морфем предъявляются требования лингвистической достаточности. Это требование должно удовлетворять используемые в лингвистическом процессоре морфологические правила, по которым осуществляется реконструкция слов. С другой стороны, к морфемам, включаемым в состав каталога, предъявляются требования достаточной акустической предельности - каждая из этих морфем должна содержать в себе не менее одного полного слога.

Акустический процессор включает в себя предпроцессор обработки речевого сигнала, формирующего вектор-параметры, значения которых регистрируются каждые секунду

(отсчеты речевого сигнала), запоминающее устройство для хранения эталонов морфем и классификатор, который определяет входение морфем словаря в распознаваемое сообщение. Классификация морфем сообщения осуществляется средствами динамического программирования.

Для уменьшения влияния эффекта границ на надежность распознавания в процедуре классификации используется весовая функция, учитывающая тип терминальных фонем морфемы и расстояние конкретного отсчета морфемного фрагмента от его границ.

На выходе акустического процессора образуются две последовательности: 1) морфемы в символическом представлении и 2) номера первых отсчетов речевых образов распознанных морфем и нераспознанных фрагментов речевого сигнала в порядке возрастания.

Надежность работы морфемного классификатора в значительной степени зависит от качества эталонов и именно алгоритму автоматического формирования эталонов морфем уделяется основное внимание в данной работе.

АЛГОРИТМ ОБУЧЕНИЯ

Процесс получения морфем ω_j^* , включаемых в каталог (словарь) W^* , связан с определением границ конкретных морфем в речевой реализации слова (здесь и далее знаком астериска обозначаются речевые образы слов и морфем). В процессе обучения решается задача определения временных границ t_k, t_m отрезка речевого сигнала при известном фонемном содержании морфемы $\omega \in W$, ограниченной этими границами. Общая структура алгоритма приведена на рис. 1.

Процедура обучения, целью которой является формирование каталога W^* эталонов морфем, происходит в несколько этапов и представляет собой суперпозицию алгоритмов $A = (A_1(A_2(A_3)))$.

На первом этапе, реализуемом алгоритмом A_1 , осуществляется фонемное транскрибирование введенных текстов слов $v_i \in V$ словаря пользователя и на основе морфемного анализа с использованием морфологических правил формируется каталог $W = \{\omega_j\}$. При морфемном анализе используются списки префиксальных W_p и суффиксальных W_s морфем (а также списки словообразующих морфемных блоков). Алгоритм A_1 осуществляет вычленение в каждом $v_i \in V$ морфем на основе символического попарного сравнения слов и сравнения с элементами $\omega_p \in W_p, \omega_s \in W_s$, в результате которого v_i представляется как цепочка

$$v_i = \omega(1), \omega(2), \dots, \omega(j), \dots, \omega(x)$$

морфем, где x может принимать значения от 1 до 11 [7]. Затем производится анализ морфемного состава словаря V и методом исключения формируется каталог W^* , в

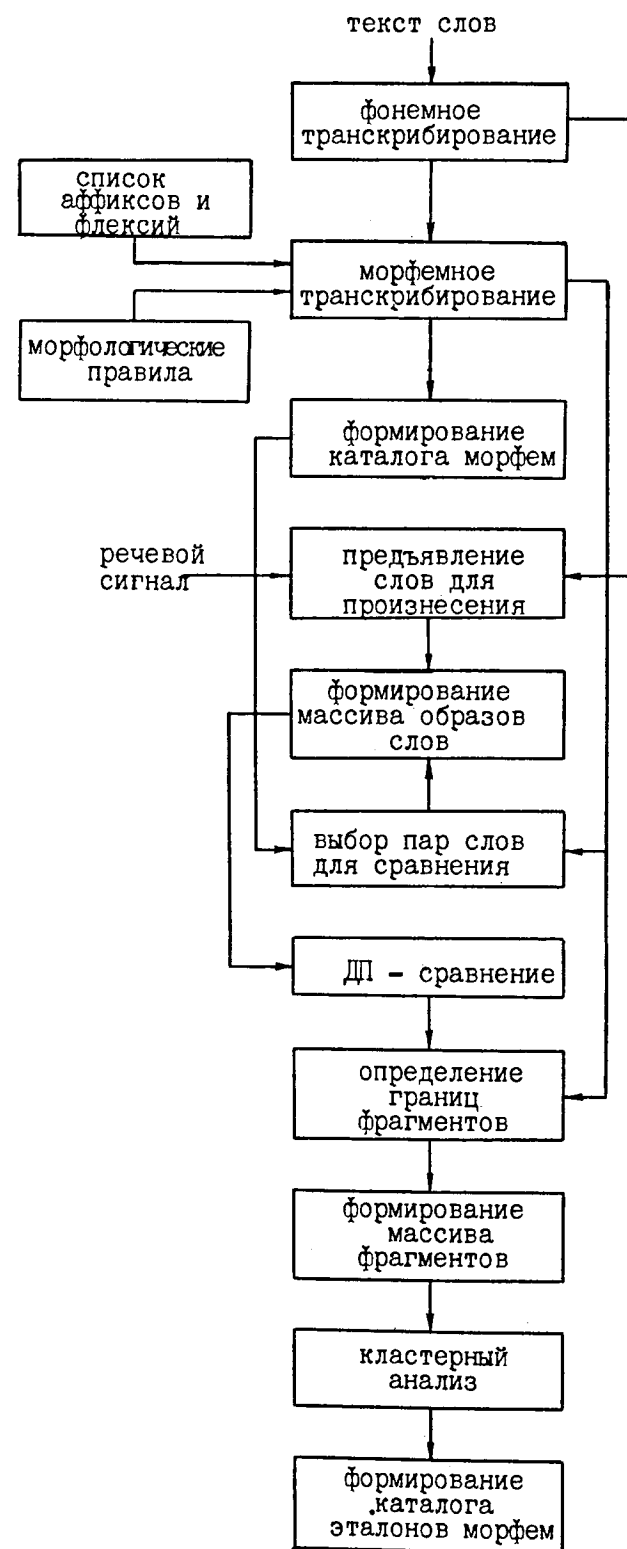


Рис. 1. Общий алгоритм формирования эталонов морфем

который включаются только морфемы, удовлетворяющие требованию акустической представительности и лингвистической достаточности.

На втором этапе (алгоритм А2) процедуры обучения в систему заносятся в речевые реализации v_i^* слов словаря пользователя, т.е. формируется словарь V^* . Как обычно, речевой сигнал слова описывается последовательностью

$$v_i^* = (X_{i1}, X_{i2}, \dots, X_{ij}, \dots, X_{Li})$$

из элементов X_j , представляющих собой вектор в пространстве принятого параметрического описания, L - длина сигнала с некоторым шагом ΔT . Итак, каждому слову v_i^* поставлен в соответствие речевой образ v_i^* . Это соответствие используется для поиска в образе v_i^* фрагментов $w_k^* = (X_k, \dots, X_m)$, соответствующим морфемам. Если слово v_i^* состоит из одной морфемы, то $v_i^* = w_j$ и его образ v_i^* заносится в массив W^{**} под именем w_j . Для определения границ морфем в речевом сигнале многоморфемного слова используется процедура ДП-сравнения, реализуемая теми же средствами, что и в процессе классификации. Из элементов $\{v_i^*\}$ выбираются пары слов v_n^* и v_q^* , отличающихся только одной морфемой и к соответствующим элементам

$$v_n^*, v_q^* \in V^*$$

применяется процедура ДП-сравнения для определения функционала меры сходства речевых сигналов v_n^* и v_q^* . При выборе пар возможны три варианта:

- 1) $w_n(i) \equiv w_q(i)$
- 2) $w_n(x) \equiv w_q(x)$
- 3) $w_n(j) \equiv w_q(j)$

Для первого случая используется прямое ДП-сравнение, для второго - обратное ДП-сравнение, а в третьем случае для поиска границ нетерминального морфемного фрагмента используется, как и при классификации, алгоритм поиска ключевых элементов в речевом сигнале по эталонам правой и левой терминальных морфем из каталога W^* . Поэтому процедура поиска границ морфемных фрагментов начинается с выбора и обработки пар v_n^* и v_q^* , имеющих одинаковые терминальные морфемы. При сравнении образов

$$v_n^* = (X_{n1}, X_{n2}, \dots, X_{ni}, \dots, X_{nn}) \quad \text{и}$$

$$v_q^* = (X_{q1}, X_{q2}, \dots, X_{qj}, \dots, X_{qq})$$

определяется интегральная мера сходства $D(i, q)$ на каждом шаге i и интегральная мера сходства $D(j, n)$, на каждом шаге j . Экстремум каждого из функционалов $D(i, q)$ и $D(j, n)$ меры сходства пар (v_n^*, v_q^*) и (v_q^*, v_n^*) , определенный с учетом вероятностной оценки его положения по фонемному содержанию сравниваемых образов, указывает номер граничного отсчета морфемного фрагмента

$w_n^*(k)$ и $w_q^*(k)$ в речевом сигнале v_n^* и v_q^* , соответственно:

$$i = m = \arg \min D(i, q)$$

$$j = r = \arg \min D(j, n)$$

Понятно, что фрагменты $w_n^*(k)$ и $w_q^*(k)$ имеют одинаковые имена в $\{w\}$. Полученные при обработке всех возможных пар речевых образов слов морфемные фрагменты образуют массив W^{**} . Формирование каталога эталонов морфем осуществляется алгоритмом А3, реализующим кластерный анализ.

В массиве W^{**} , выделяются подмножества образов морфем $\{w_k^*\}_1^N$, имеющих одинаковые имена w_k . Все N элементов подмножества $\{w_k^*\}$ поступают на ДП-сравнение для вычисления матрицы расстояний

$$\{D_{ij}\} \quad i = 1, N, \quad j = 1, N$$

между образами. По полученной матрице строится обобщенный эталон для всех образов, т.е. кластер представляется одним эталоном, полученным в результате усреднения образов, образующих данное подмножество.

ЗАКЛЮЧЕНИЕ

Алгоритм формирования эталонов морфем разрабатывался для использования в системе распознавания слитнопроизносимых фраз, основанной на алгоритме принятия решений о речевых единицах в текущем речевом сигнале Н.П.Дегтярева [6]. Выбор в качестве базовых элементов морфем, которые являются носителями лексических и грамматических значений значительно расширяет лексические возможности распознающей системы, позволяет строить произносимые фразы с использованием широкого разнообразия словоформ. Однако необходимо отметить, что использование сложного алгоритма обучения может быть целесообразно и оправдано только при работе с достаточно большими словарями.

Практическое применение систем морфемного распознавания зависит не только от надежности классификации морфем, но и в значительной степени от разработанности правил реконструкции слов и предложений.

ЛИТЕРАТУРА

1. J.M.Pierrel. Use of linguistic constraints for automatic continuous speech understanding: the Myrtille II system, Technology and Science of Informatics, vol. 1, N 5, 1983.
2. S.B.Davis and P.Mermelstain. Comparison of parametric representation for monosyllabic word recognition in continuously spoken sentences, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-28, 1980, pp. 357-366.
3. A.E.Rosenberg, L.R.Rabiner, J.G.Wilpon and D.Kahn. Demisyllable based isolated word recognition system. IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-31, 1983, pp. 713-726.
4. K.Shicano and M.Kohla. A Linguistic processor in a convergational speech recognition system. Rev. ECL, vol. 26, 1978, pp. 1486-1504.
5. А.Н.Кузнецова, Т.Ф.Ефремова. Словарь морфем русского языка. М., "Русский язык", 1986.
6. Н.П.Дегтярев. Алгоритм распознавания слов в непрерывном сигнале. Доклад на Всесоюзном симпозиуме "Бионика интеллекта". Харьков, 1987.
7. А.И.Кузнецова. Морфемная глубина слов в русском языке. ПСЛ, 1984.