

О МИНИМАЛЬНОМ АЛФАВИТЕ В СЛОЖНОЙ СИСТЕМЕ
РАСПОЗНАВАНИЯ И ТАКСОНОМИИ

Н.Г. Загоруйко

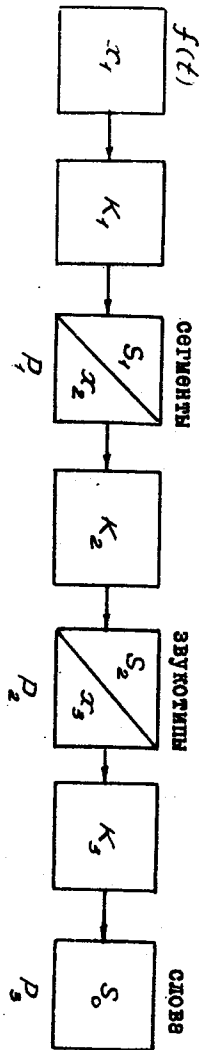
В работе [1] описан алгоритм таксономии мелких элементов (звукотипов), обеспечивающий выбор минимального числа групп этих элементов, достаточного для безошибочного распознавания конкретного набора крупных элементов (устных команд). Рассмотрим место этого алгоритма в сложной (многоуровневой) распознающей системе.

Γ^0 . Распознающее устройство представим в виде трех последовательно работающих элементарных автоматов (рис. 1). В состав каждого автомата входит рецептор, производящий описание поступающего на его вход сигнала, классификатор, осуществляющий процедуру принятия решений в этом пространстве описаний, и эффектор, фиксирующий принятое решение. Эффектор предыдущего автомата совпадает с рецептором последующего: классификатор следующей ступени работает в пространстве решений предыдущего автомата.

Представим себе, что рецептор первого автомата преобразует речевой сигнал в последовательность "спектральных разрезов", усредненных на отрезке сигнала продолжительностью 10-15 мсек.

Пусть алфавитом объектов распознавания первого автомата (S_1) будет конечный набор коротких речевых элементов типа "сегмент".

Рис.1. Блок-схема многоуровневой системы распознавания.



Классификатор (K_1) принимает решение о принадлежности любого из спектральных разрезов к тому или иному сегменту. Будем считать, что некоторая последовательность сегментов образует более крупный речевой элемент - "звукотип". Список звукотипов будет алфавитом второго автомата (S_2). Классификатор (K_2) определяет, к какому из звукотипов принадлежит та или иная последовательность сегментов.

Наконец, третий автомат должен распознавать команду из конкретного словаря (S_0) как некоторую последовательность звукотипов.

Алгоритм [1] позволяет осуществить оптимальную группировку звукотипов, т.е. минимизировать алфавит второго автомата (S_2). Очевидно, что с помощью того же алгоритма можно осуществить оптимальную группировку и элементов алфавита первого автомата (S_1), т.е. минимизировать число групп сегментов. Функция потерь для этой задачи таксономии есть величина допустимых ошибок распознавания групп звукотипов (P_2). Основой для выбора наиболее предпочтительного варианта группировки служит, как и раньше [1], матрица ошибок распознавания мелких элементов (здесь - сегментов).

Можно продолжить применение этого алгоритма и дальше - для уменьшения пространства признаков X_1 . Следует только иметь в виду, что решающее правило классификатора первого автомата будет отличаться от правил других автоматов тем, что здесь нельзя оперировать порядком следования признаков из множества X_1 , т.к. все они появляются одновременно. Нельзя также говорить и о матрице ошибок распознавания мелких элементов (спектральных составляющих).

Критерием для выбора варианта группировки может служить какая-нибудь другая мера близости, например, близость составляющих по частной шкале. Тогда в группу будут объединяться соседние частотные составляющие. Если такой вариант группировки не имеет практических преимуществ перед другими, то тогда единственным управляющим критерием остается критерий совместности в смысле [1].

Может оказаться, что сокращение избыточности описания более целесообразно не за счет группировки, а за счет вычеркивания некоторых мелких элементов. В этом случае можно воспользоваться одним из алгоритмов поиска наиболее информативных подпространств, например, методами [2] или [3].

2⁰. В работе [I] не указывается способ учета влияния надежности сегментации речевого потока и распознавания мелких элементов на надежность распознавания более крупных элементов.

Пусть P_0 - требуемая надежность распознавания слов. Тогда минимальный набор групп звукотипов с учётом надёжности их выделения и распознавания можно будет получить, используя алгоритм [I] со следующими изменениями.

В ходе эксплуатации макета распознающего устройства записываются все варианты результатов распознавания слов заданного набора с указанием частоты встречаемости этих вариантов.

Количество дикторов и разнообразие внешних условий (шумов, характеристик помещений и т.д.) должно быть таким, чтобы эту обучающую выборку можно было считать достаточно представительной. Тогда частоты встречаемости вариантов можно считать априорными вероятностями их появления. Для предварительного сокращения объема исходного материала можно вычеркнуть те варианты слов, суммарная априорная вероятность которых не превышает допустимых для этого слова ошибок распознавания.

Оставшиеся варианты слов записываются в виде строк матрицы, столбцами которой служат все звукотипы исходного их набора, а на пересечении строки и столбца стоит порядковый номер следования данного звукотипа в данном варианте слова. В такой таблице найдут свое отражение все ошибки членения и распознавания мелких элементов. Эта таблица и служит в дальнейшем базой для группировки звукотипов по алгоритму [I]. Ясно, что условие достаточности в этом случае будет, в отличие от [I], выполнено, если не сливаются варианты разных слов. Слияние же вариантов одного и того же слова можно только приветствовать.

Тем же способом можно учесть и ошибки сегментации и распознавания сегментов при формировании групп этих более мелких элементов.

3⁰. Может оказаться, что после окончания процедуры группировки в таблице останется более чем по одному варианту одного и того же слова.

Если решение о принадлежности данной последовательности групп звуков к тому или иному слову классификатор K_3 будет принимать лишь при полном их совпадении, то тогда в качестве эталонов слов придется оставить все эти их варианты. Однако число эталонных вариантов можно уменьшить, если K_3 будет

оперировать не абсолютным совпадением реализации и эталона, а какой-нибудь мерой близости.

В этом случае выбор минимального набора эталонных вариантов слов может быть сделан так же, как в алгоритмах, использующих функции "принадлежности" [4], "сходства" [5] или "соседства" [6]. В результате останутся только те варианты, расстояние до которых от любой реализации этого слова будет меньше, чем до эталонных вариантов других слов (речь идет о каком-либо способом измеренном расстоянии между траекториями слов в пространстве звукотипов).

4⁰. Последовательное применение алгоритма [I] позволит выбрать минимальный алфавит объектов распознавания на всех иерархических уровнях сложной системы распознавания.

Как отмечалось в [I], при группировке элементов выполняются требования, предъявляемые к таксономии: максимальные связи между элементами одной группы и минимальные связи между группами. Но в отличие от ранее описанных в литературе алгоритмов таксономии [7, 8, 9, 10], в [I], кроме этих требований, выполняются ещё и следующие: количество групп (таксонов) должно быть минимальным и достаточным для распознавания конкретного набора более крупных элементов с заданной надежностью. Этими условиями чётко фиксируется цель таксономии, вследствие чего исключается какая бы то ни было возможность для использования терминов типа "самообучение", "обучение без учителя" и т.д. Всё, что требуется от "учителя", - конечная цель процедуры таксономии, методы её достижения и критерии для оценки получаемых результатов - здесь задано однозначно.

"Обучение без учителя" имеет место тогда, когда задан способ группировки (максимальные связи внутри группы и минимальные связи между группами), но не указано, для какой конкретной цели эта группировка делается, т.е. что от чего нужно будет отличать с помощью получаемых групп. Такая таксономия "на все случаи жизни" для каждого конкретного использования может оказаться неоптимальной. Например, естественное для "самообучения" объединение в группу глухих взрывных "п", "т", "к" окажется неприемлемым, если в числе распознаваемых будут слова типа "ток", "кот" и т.д.

Существует только одно (тривиальное) решение задачи таксономии "на все случаи жизни", состоящее в группировке без потери информации, т.е. в объединении в группу объектов абсолютно неразличимых в пространстве интересующих нас свойств.

Л и т е р а т у р а

1. Н.Г. Загоруйко, В.Н. Елкина. Алфавит с минимальной избыточностью.-Данный сборник, стр. 49-57.
2. Г.С. Лбов. Выбор эффективной системы зависимых признаков-Вычислительные системы, г. Новосибирск, Изд-во "Наука", Сиб.отд., 1963, вып. 19.
3. В.Н. Елкина, Н.Г. Загоруйко. Об алфавите объектов распознавания. - Вычислительные системы, г. Новосибирск, Изд-во "Наука", Сиб.отд., 1966, вып. 22.
4. И.Т. Турбович. Об оптимальном методе опознания образов при взаимнокоррелированных признаках. Опознание образов. - Теория передачи информации. АН СССР, ИППЧ, Изд-во "Наука", 1965.
5. А.Г. Француз. Распознавание образов с использованием "функций близости". - Доклад на Всесоюзном симпозиуме по распознаванию образов. Москва, июнь 1965.
6. Г. Себестиан. Процессы принятия решений при распознавании образов, пер.с английского. Изд-во "Техника", Киев, 1965.
7. Е.А. Елкин, В.Н. Елкина, Н.Г. Загоруйко. О применении методики распознавания образов к решению задач палеонтологии. - Доклад на Всесоюзном совещании по применению математики в геологии. Новосибирск, 1965.
8. М.И. Шлезингер. О самопроизвольном различении образов. - Читающие автоматы, Киев, 1965.
9. Э.М. Браверман. Метод потенциальных функций в задаче обучения машины распознаванию образов без учителя. - Автоматика и телемеханика, 1966, № 10.
10. А.А. Дорофеев. Алгоритмы обучения машины распознаванию образов без учителя, основанные на методе потенциальных функций. - Автоматика и телемеханика, 1966, № 10.