

**МЕТОДЫ ОБНАРУЖЕНИЯ
ЭМПИРИЧЕСКИХ ЗАКОНОМЕРНОСТЕЙ
(Вычислительные системы)**

2001 год

Выпуск 167

УДК 519.816

**РАСПОЗНАВАНИЕ С ИСПОЛЬЗОВАНИЕМ
ЗНАНИЙ О ПРОСТРАНСТВЕ ОБРАЗОВ**

Н.Г. Загоруйко

В в е д е н и е

Распознавание большого числа образов обычно делается в некотором общем для них признаковом пространстве. Между тем очевидно, что для надежного распознавания пар образов (A, B) , (A, C) и (B, C) было бы целесообразно использовать такие индивидуально подобранные для каждой пары характеристики, по которым образы в каждой паре наиболее сильно отличаются друг от друга. Так, при распознавании устных слов "пятнадцать", "шестнадцать" и "пятьдесят" для первой пары слов следовало бы учитывать характеристики, связанные с началами этих слов, а для второй пары — с концами слов. Если распознается K образов, то при этом подходе требуется построить эталоны для всех их парных сочетаний и затем в процессе распознавания организовать экономичную процедуру попарного сравнения конкурирующих образов. Начнем с выбора наиболее информативного ("компетентного") подпространства характеристик для каждой пары образов.

Работа поддержана грантом РФФИ № 99-01-00582.

Выбор компетентных подпространств

Будем рассматривать случай распознавания по признакам, измеренным в сильных шкалах и в предположении справедливости гипотезы глобальной компактности [1]. Это позволяет ограничиться одним эталоном на образ и использовать в качестве меры близости Евклидово расстояние. Метод несложно распространить и на случай разнотипных признаков, а также на предположение о полимодальной или локальной компактности, когда в качестве эталонов используется несколько прецедентов или даже все обучающие реализации.

Каждая реализация любого из K образов описывается исходным набором из N признаков. Если признак X_i принимает одно и то же значение и для образа A , и для образа B , то его информативность для этой пары следует считать равной 0, вне зависимости от того, находится ли он в какой-нибудь зависимости от других признаков или нет. Если же расстояние между средними значениями (μ) этого признака для данных образов велико, то он может нести полезную информацию.

Дополнительно нужно учитывать еще и дисперсию (δ) значений этого признака на всех обучающих реализациях каждого образа. Такой подход к оценке потенциальной информативности признака хорошо отражает критерий Фишера:

$$J(iab) = \frac{|\mu(ia) - \mu(ib)|}{\delta(ia) + \delta(ib)}.$$

Примем информативность $J(iab)$ в качестве весового коэффициента i -го признака при определении взвешенного расстояния между объектами в признаковом пространстве. В результате для каждой пары образов N -мерное пространство рассматривается через свой вариант весовых коэффициентов. Каждый такой вариант взвешенного пространства будем называть компетентным подпространством для данной пары образов. В общем случае расстояние от реализации Y до эталона образа A в этом

компетентном подпространстве будет равно

$$R(aY) = \sqrt{\frac{\sum_{i=1}^{i=N} J(iab) * (X(ia) - X(iY))^2}{\sum_{i=1}^{i=N} J(iab)}}.$$

По аналогичной формуле оценивается и расстояние $R(ab)$ между эталонами образов A и B в этом подпространстве.

Распознавание методом попарного сравнения

Рассмотрим процесс распознавания принадлежности реализации Y к одному из трех образов: A, B и C . Сравнивая распознаваемую реализацию Y с эталонами образов A и B , мы определяем на какой из этих двух образов более похожа реализация в подпространстве, оптимальном именно для этого случая. То же относится к сравнению между другими парами (A, C) и (B, C) .

Для сокращения необходимого числа сравнений сформулируем

УТВЕРЖДЕНИЕ 1 (о транзитивности). Если $R(aY) < R(bY)/AB$ и $R(aY) < R(cY)/AC$, то $R(aY)/AB < R(bY)/BC$ и $R(aY)/AB < R(cY)/BC$.

То есть, если в двух парных сравнениях (A, B) и (A, C) победителем оказывается один и тот же образ A , то победитель среди проигравших образов $(B$ и $C)$ не может победить образ A . Так что, сравнение между проигравшими можно и не проводить.

Доказательством этого утверждения является тот факт, что при формировании компетентного подпространства для всех пар образов с наибольшим весом учитываются те признаки, которые наилучшим способом подчеркивают различия именно между этими двумя образами. То есть, для каждого из них создаются наилучшие условия для победы над каждым из конкурентов в отдельности. И если распознаваемая реализация является представителем образа A , то при первом же попадании этого образа в состав сравниваемой пары он выиграет соревнование у своего конкурента в этой паре, а затем и у всех остальных конкурентов.

На основании сказанного процедура распознавания строится так. На этапе обучения формируется по одному эталону на каждый из K образов и C_K^2 строк с N весовыми коэффициентами для каждой пары образов. N -мерный вектор распознаваемой реализации сравнивается с эталонными векторами любой пары образов с использованием строки коэффициентов для этой пары. Определяется образ, расстояние от эталона которого до реализации оказалось меньшим. Он ставится на первую позицию и к нему на вторую позицию ставится любой следующий образ из числа не участвующих в сравнении. Победитель этой пары снова занимает первую позицию, а на вторую ставится следующий новый конкурент. И такая процедура повторяется $K - 1$ раз. Победитель последней пары и является победителем над всеми конкурентами.

В итоге решение о принадлежности реализации к одному из K образов принимается за число шагов, равное $K - 1$. При этом судьба каждого образа решается в наилучших для него условиях. В традиционной схеме распознавания в признаковом пространстве, общем для всех образов, это важное и естественное требование не выполняется. Если среди большого числа образов есть пара образов, похожих друг на друга, но сильно отличающихся от остальных, то в качестве информативных будут выбраны такие признаки, которые являются лучшими в среднем, и тонкие различия внутри этой пары не будут подчеркиваться. Расстояние от реализации до эталонов этих образов будет измеряться по второстепенным для них признакам, шумы в которых могут замаскировать их существенные отличия. Известно, что качество распознавания тем выше, чем больше информативность признакового пространства. На основании приведенных выше аргументов можно без дополнительных доказательств сделать следующее

УТВЕРЖДЕНИЕ 2 (о преимуществах описанного метода). Если $P(N)$ — ошибки при распознавании в пространстве характеристик, общих для всех образов, а $P(n)$ — ошибки при распознавании путем попарного сравнения в компетентных подпространствах, то $P(n) \leq P(N)$.

Метод вычеркивания

В случае большого количества образов для дальнейшего сокращения времени на принятие решений при контролируемом увеличении риска ошибок можно воспользоваться методом вычеркивания [1]. Он состоит в проецировании вектора распознаваемого образа на координатную ось и вычеркивании из числа конкурентов тех образов, проекции которых на эту же ось удалены от проекции реализации на расстояние больше некоторого порога.

Одним из недостатков этого метода является его слабая помехоустойчивость. Случайные изменения величины отдельного признака могут привести к непоправимым ошибкам. Желательно было бы оперировать расстояниями не по отдельным признакам, а по всей их совокупности.

Версия метода вычеркивания с учетом этого требования может выглядеть так. На основании предварительно найденной матрицы парных взвешенных расстояний между эталонами для каждого образа составляется список остальных образов, упорядоченных по этому расстоянию. Для образа *A* на первом месте будет находиться его самый сильный конкурент, затем конкурент послабее и так далее до самого непохожего образа. Если на некотором шаге парного сравнения образ *A* окажется победителем, то для дальнейшего сравнения целесообразно выбрать его сильнейшего конкурента. Если образ *A* победит и его, то высока вероятность того, что он победит в соревновании и с теми конкурентами, которые находятся в конце его списка. Будем на каждом шаге приписывать образам штрафные баллы, равные порядковому номеру образа в списке образа победителя. При смене лидера порядковые номера образов меняются. Но если какой-то образ будет долго оставаться в числе наиболее слабых конкурентов меняющимся лидером, то он будет набирать большую сумму штрафных баллов. Если эта сумма превысит некоторый порог, то соответствующий образ можно вычеркнуть из числа распознаваемых. Таким путем общее число конкурирующих образов будет сокращаться, и решение будет получено за число шагов, меньшее, чем $K - 1$.

Чем выше порог, при достижении которого образ вычеркивается из списка, тем меньше риск вычеркнуть нужный образ. Величину порога следует устанавливать в процессе распознавания тестового материала при настройке распознающей системы.

Описанный метод требует предварительного вычисления и запоминания N -мерных весовых векторов для всех пар образов, что при большом их числе может потребовать слишком больших затрат памяти. Если же весовые коэффициенты вычислять для каждой пары образов непосредственно в процессе распознавания, это будет сказываться на времени принятия решений. Компромисс между требованиями экономии и времени и памяти можно обеспечить, если применить следующую стратегию распознавания.

Распознавание образов в пространстве образов

Этот способ предназначен для быстрого выхода на самую вероятную гипотезу без затрат памяти на запоминания весовых коэффициентов и основан на использовании "пространства образов" Z , которое строится так. В качестве эталонов для каждого образа используются средние значения и дисперсии исходных признаков X . Введем переменные $Z_i, i = 1, 2, \dots, K$, на которых отложим расстояния от образа i до всех образов в исходном пространстве признаков X . В сформированном в итоге пространстве образов Z каждый образ будет представлен K -мерным вектором. Если теперь в пространстве признаков X измерить расстояние $R(iY)$ от распознаваемой реализации Y до эталона i -го образа и отложить его на оси Z_i , то можно найти образ j , проекция которого на эту ось ближе всего находится к отложенной точке. Следующее сравнение реализации Y нужно делать с эталоном именно этого образа. Получив расстояние $R(jY)$, мы на плоскости (Z_i, Z_j) найдем следующий образ, наиболее близкий к Y . На t -м шаге мы в t -мерном подпространстве можем выделить K' образов, наиболее похожих на реализацию Y . Вероятность того, что отобранные K' образов случайно оказались похожими на реализацию Y , сокращается с ростом числа шагов t по степенному закону.

Дальнейшее соревнование между этими образами делается в режиме описанного выше попарного сравнения в компетентных пространствах. Но поскольку этих образов немного, то весовые коэффициенты компетентности признаков можно вычислять прямо в процессе сравнения реализации Y с каждой парой из K' образов.

Затраты памяти в этом методе с ростом количества образов растут линейно. Что касается затрат времени, то можно предположить

УТВЕРЖДЕНИЕ 3. *Для успешного распознавания по этой схеме будет достаточно провести небольшое фиксированное число t предварительных шагов и фиксированное число парных сравнений, суммарное количество которых будет существенно меньше количества образов и будет слабо зависеть от этого количества.*

Выводы

Экспериментальная проверка описанного метода делалась в процессе решения задачи распознавания десяти зрительных образов, реализации которых были поражены шумами высокого уровня [2]. Задачу удалось решить успешно. При этом была подтверждена справедливость сформулированных выше утверждений 1-3. При попарном сравнении правильная гипотеза оказывалась на первом месте обычно уже на третьем шаге. При использовании пространства образов было достаточно на первом этапе использовать три шага, и оставлять для парного сравнения три сильнейших конкурента.

Автор выражает глубокую благодарность С.Е.Гулявскому за экспериментальную проверку метода и А.С.Нудельману за его участие в обсуждении.

Л и т е р а т у р а

1. ЗАГОРУЙКО Н.Г. Прикладные методы анализа данных и знаний. — Новосибирск: Изд. ИМ СО РАН, 1999.

2. ЗАГОРУЙКО Н.Г., ГУЛЯЕВСКИЙ С.Е. Использование пространства образов для распознавания маркеров на тепловыделяющих элементах атомных реакторов //Труды VI Всесоюз. конф. с участием стран СНГ "Методы и средства обработки сложной графической информации". — Н.Новгород, 2001. — С. 87-89.

Поступила в редакцию
27 марта 2001 года