

М. И. МОГИЛЬНИЦКИЙ

(Новосибирск)

О ВЛИЯНИИ ПОГРЕШНОСТЕЙ КЛАССИФИКАТОРА НА ПРОЦЕСС ОБУЧЕНИЯ

Основы применения стохастической аппроксимации для задач распознавания образов приводятся в работе*. В результате применения этого метода по отдельным реализациям образов получаем оценки коэффициентов разложения C_1, C_2, \dots, C_j функции, аппроксимирующей заданную $f(\bar{x})$:

$$y = f(\bar{x}) \left\{ \begin{array}{l} + 1; \bar{x} \in A; \\ - 1; \bar{x} \in B, \end{array} \right.$$

где \bar{x} — вектор признаков; А и В — классы. Если ограничиться критерием модуля ошибки, то оптимальными являются те коэффициенты, для которых выражение $I(\bar{C}) = M(|y - C^{*T} x|)$ минимально. Вычисление C_1, C_2, \dots, C_j в соответствии с уравнением

$$\bar{C}_{n+1} = \bar{C}_n + \gamma_{n+1} \text{sign}(y - C_n^T \bar{x}_{n+1}) \bar{x}_{n+1}, \quad (1)$$

где n — номер шага, а γ_{n+1} — величина шага, позволяет получить оценки, имеющие минимальную среднеквадратическую ошибку относительно оптимальных коэффициентов $C_1^*, C_2^*, \dots, C_j^*$.

Описанный алгоритм может быть реализован с помощью специализированного прибора, решающего задачи распознавания (классификатора образов). При этом в процессе работы прибора могут возникнуть ошибки в определении $C_{1n}, C_{2n}, \dots, C_{jn}$, наличие которых связано с аддитивной помехой, возникающей в цепях управления коррекцией коэффициентов, с разбросом параметров элементов этих же цепей от шага к шагу и т. д.

Предположим, что на n -м шаге в вычислении коэффициентов имеется погрешность $\bar{b}_n = \bar{C}_n + \bar{\delta}$, и, следовательно,

$$\bar{b}_{n+1} = \bar{b}_n + \gamma_{n+1} \text{sign}(y - \bar{b}_n^T \bar{x}_{n+1}) \bar{x}_{n+1} + \bar{\delta}. \quad (2)$$

Вектор $\bar{\delta}$ в дальнейшем будем называть вектором случайных уходов; влияние его составляющих $\delta_{1n}, \delta_{2n}, \dots, \delta_{jn}$ заключается в том, что коррекция коэффициентов осуществляется не на величину $\bar{\Delta} = \gamma_{n+1} \times \text{sign}(y - \bar{b}_n^T \bar{x}_{n+1}) \bar{x}_{n+1}$, а на $\bar{\Delta}' = \gamma_{n+1} \text{sign}(y - \bar{b}_n^T \bar{x}_{n+1}) \bar{x}_{n+1} + \bar{\delta}$.

* Я. З. Цыпкин. Адаптация и обучение в автоматических системах. М., «Наука», 1968.

Сами значения случайных уходов $\delta_{1n}, \delta_{2n}, \dots, \delta_{jn}$ не зависят от составляющих вектора признаков $x_{1n}, x_{2n}, \dots, x_{jn}$.

Работа ставит своей целью установление связи случайных уходов со сходимостью и точностью оценок коэффициентов.

Пусть математическое ожидание случайных уходов равно нулю, т. е.

$$M(\bar{\delta}) = 0, \quad (3)$$

а $f(\bar{x})$ представима суммой

$$f(\bar{x}) = \bar{C}^{*T} \bar{x}. \quad (4)$$

Качество оценки на каждом шаге будем определять величиной

$$\Delta \bar{C}_{n+1} = \bar{C}^* - \bar{b}_{n+1} = \Delta \bar{C}_n - \gamma_{n+1} \text{sign}(y - \bar{b}_n^T \bar{x}_{n+1}) \bar{x}_{n+1} - \bar{\delta}. \quad (5)$$

Возведем выражение (5) в квадрат и возьмем математическое ожидание от его правой и левой части. Обозначим разность средних квадратов приближений на $(n+1)$ -м и n -м шаге через $K_{n, n+1}$:

$$\begin{aligned} K_{n, n+1} &= M[\bar{C}^* - \bar{b}_{n+1}]^2 - M[\bar{C}^* - \bar{b}_n]^2 = \\ &= K_{n, n+1} = M(\Delta \bar{C}_{n+1}^T \Delta \bar{C}_{n+1}) - M(\Delta \bar{C}_n^T \Delta \bar{C}_n) = \\ &= \gamma_{n+1}^2 M(\bar{x}_{n+1}^T \bar{x}_{n+1}) - 2\gamma_{n+1} M(|\varepsilon_{n, n+1}|) + M(\bar{\delta}^T \bar{\delta}), \end{aligned} \quad (6)$$

где $\varepsilon_{n, n+1} = y - \bar{b}_n^T \bar{x}_{n+1} = \Delta \bar{C}_n^T \bar{x}_{n+1}$. Для сходимости алгоритма необходимо, чтобы выполнялось условие $K_{n, n+1} < 0$. Очевидно, что существует такая величина шага γ_n и его номер n , при котором $K_{n, n+1}$ становится больше нуля и при дальнейшем обучении сходимости отсутствует. Величина γ_n удовлетворяет известным условиям сходимости [1].

Случайный уход оказывает влияние и на такую важную характеристику классификатора, как разрешающая способность. Под разрешающей способностью классификатора понимается значение наименьшего отличия в признаках, вызывающее пропорциональное отличие в величине коррекции. Этот параметр отражает потенциальную возможность прибора распознавать образы, признаки которых мало различаются.

Пусть величина разрешения равна Δx и на каждый признак накладывается это значение. Тогда величина коррекции для одного и того же коэффициента в двух последовательных шагах будет составлять

$$\begin{aligned} \Delta_{1j} &= \gamma \text{sign}(y_1 - \bar{b}^T \bar{x}) x_j + \delta_{1j}; \\ \Delta_{2j} &= \gamma \text{sign}[y_2 - \bar{b}^T (\bar{x} + \Delta \bar{x})] (x_j + \Delta x) + \delta_{2j}. \end{aligned}$$

Примем так же, что $\gamma = 1$; $\text{sign}(y_1 - \bar{b}^T \bar{x}) = 1$; $\text{sign}[y_2 - \bar{b}^T (\bar{x} + \Delta \bar{x})] = -1$. Тогда уход на один шаг при наличии случайных уходов равен

$$K_{шj} = -\Delta x + \delta_{1j} + \delta_{2j} = -\Delta x + \delta_{шj}, \quad (7)$$

а при отсутствии их

$$K_j = -\Delta x, \quad (8)$$

т. е. величина разрешающей способности определяется значением случайного ухода.

Таким образом, для равенства (7) и (8) необходимо выполнение условия $\delta_{ш} \ll \Delta x$. Так как на практике выполнение требования $\delta_{ш} \ll \Delta x$ не всегда возможно, то для эффективного разрешения двух признаков потребуется продлить время обучения, тем самым выделив

средний уход за ряд шагов. Для связи предельной величины случайных уходов с характеристиками прибора введем понятие чувствительности. Под чувствительностью классификатора понимается такое наименьшее значение $(\gamma_n x_{jn})$, при котором $\Delta_{jn} \neq 0$. Существование отличной от нуля чувствительности связано, например, с дискретностью представления входных данных в цифровом классификаторе, а в аналоговом — с ограниченностью коэффициента усиления элементов.

Очевидно, что для исключения влияния погрешностей на процесс обучения достаточно соблюсти неравенство

$$\sigma_{\text{сл. ух}} \ll m, \quad (9)$$

где $\sigma_{\text{сл. ух}} = \sqrt{M(\delta_j^2)}$, а m — чувствительность.

Так как выражения (6), (7), (9) зависят от дисперсии случайных уходов, то на практике может возникнуть необходимость в ее измерении. Наиболее целесообразным представляется проводить определение $\sigma_{\text{сл. ух}}$ для $x = \text{const}$, чтобы случайный характер величины коррекции определялся только δ . После n шагов по ряду значений $\Delta_{j1}, \Delta_{j2}, \dots, \Delta_{jn}$ находим

$$M(\delta_j^2) = \frac{\sum_{i=1}^n [\Delta_{ji} - M(\Delta_{ji})]^2}{n-1}. \quad (10)$$

Предложенная методика была использована для определения дисперсии случайных уходов и разрешающей способности конкретного прибора.

Эксперименты проводились на электромеханическом классификаторе шума. Схема корректировки коэффициентов разложения для i -го канала приведена на рис. 1. Элементами адаптации являются проволочные потенциометры, один из которых Π_i^1 служит для перемножения признака на коэффициент, а с помощью второго Π_i^2 осуществляется отработка угла поворота. Движки потенциометров механически связаны. Величина напряжения на первом потенциометре равна x_i , а на втором U_k ; причем $U_k \gg x_i$ ($U_k = 100$ в; $x_i = 5$ в). В схему элементов одного канала входят два конденсатора — памяти L_i ,

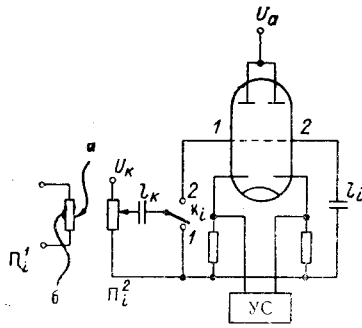


Рис. 1.

величина напряжения на котором равна x_i , и корректировки L_k , а также контакт k_i . Все остальные элементы являются общими, а последовательное подключение схем каналов к лампе с устройством сравнения (УС) происходит с помощью шагового искателя.

При подключении схемы контакт k_i переходит в положение 2. Так как в первый момент сумма напряжений, поступающих на сетку 1, равна нулю, а на сетку 2 x_i , то УС придет в положение «небаланс», электромагнитная муфта срабатывает и от приводного двигателя начнется перемещение движков потенциометров. Перемещение движка Π_i^2 приведет к нарастанию напряжения на сетке 1. Как только приращение напряжения ΔU_k станет равным x_i , так УС придет в положение «баланс», будет подключен следующий канал и k_i перейдет в положение 1.

Для определения $\sigma_{\text{сл. ух}}$ осциллографировалось напряжение на катодной левой половине лампы. Вид полученных кривых показан на рис. 2:

амплитуда их с учетом масштабного множителя равна коррекции коэффициентов. Для одного из каналов по ряду циклов проводились измерения амплитуды, затем эти данные обрабатывались в соответствии с (10). Результаты представлены в таблице.

Напряжение признака, в	$\sigma_{\text{сл. ух.}}$, мв	Точность среднего для доверительной вероятности 0,999, %	$\sigma_{\text{сл. ух.}}$, % среднее значение признака	Число шагов
4	80	± 1	5,3	30
2	100	$\pm 2,95$	6,3	33

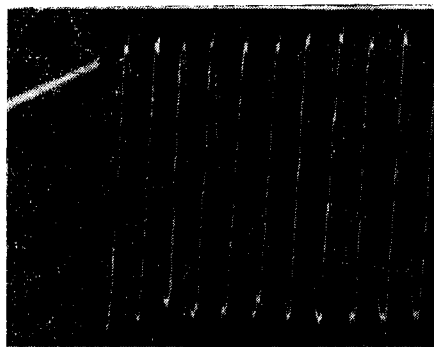


Рис. 2.

Следует отметить, что полученное значение $\sigma_{\text{сл. ух}}$ является завышенным, что связано с дискретностью намотки Π_i^2 (порядка 500 витков). Этот факт подтверждает и величина разрешающей способности, снятая экспериментально и оказавшаяся равной 100 мв. На рис. 3 показаны осциллограммы коррекции, снятые между точками *a* и *b* на одном и том же потенциометре Π_i^1 для двух значений признаков $x_1=2$ в и $x_2=2,1$ в. В области *I* $y=1$ и признак x_2 , а в области *II* наоборот. Смещение осциллограмм вверх, а затем вниз показывает, что эти два значения прибор разрешает. Так как уход мог вызываться и иными причинами, то на

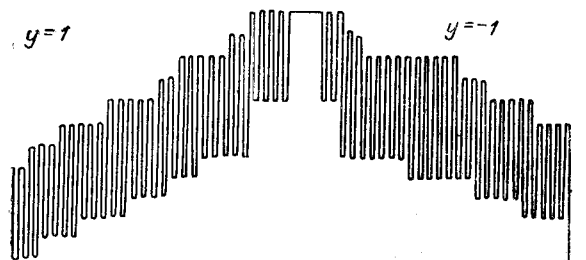


Рис. 3.

рис. 4 показан вид коррекции для $x_1=x_2$. Аналогичная картина получена при уменьшении Δx до 50—80 мв. Случайный характер коррекций не виден из-за дискретности намотки потенциометра.

В заключение следует отметить, что наличие значительных случайных уходов приводит к увеличению среднеквадратиче-

ской ошибки оценок коэффициентов, исключению из обучения признаков, мало отличающихся для различных классов, а все это увеличивает ошибки при распознавании.

Нельзя также не отметить, что оба указанных параметра классификатора могут служить предварительной оценкой при сравнении различных типов приборов.

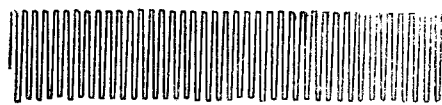


Рис. 4.

Поступила в редакцию
26 декабря 1969 г.,
окончательный вариант —
22 января 1970 г.