УДК 621.391:519.28

СИНТЕЗ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ГРУПП РАДИОИЗЛУЧЕНИЙ С ОБУЧЕНИЕМ НА ОСНОВЕ НЕПАРАМЕТРИЧЕСКОГО ОЦЕНИВАНИЯ УСРЕДНЕННЫХ ФУНКЦИЙ ПРАВДОПОДОБИЯ МЕТОДОМ НОРМАЛЬНЫХ ВКЛАДОВ

д.т.н. Г.В. Певцов, В.А. Лупандин, к.в.н. Д.И. Уманец

Разработан метод синтеза непараметрических алгоритмов распознавания групп радиоизлучений с обучением на основе сглаживания эмпирических усредненных функций правдоподобия нормальными вкладами. Метод является развитием предложенных ранее подходов к синтезу алгоритмов распознавания образов, заданных сложными эталонными описаниями в виде совокупностей эталонных значений и (или) интервалов эталонных значений признаков (параметров) радиоизлучений.

Постановка проблемы и анализ литературы. При идентификации объектов, излучающих разные виды сигналов, часто решаемую задачу можно рассматривать в форме распознавания групп радиоизлучений (образов) по выборочным значениям их признаков. В пространстве признаков (параметров радиоизлучений и их источников) такой образ может описываться одним или несколькими интервалами эталонных значений и (или) одним или несколькими дискретными эталонными значениями признаков. В [1 – 3] и др. развита методика синтеза алгоритмов распознавания групп радиоизлучений, реализующих проверку сложных статистических гипотез. Методика базируется на методах проверки сложных статистических гипотез и введенном сложном эталонном описании образов в виде $\Im$-мерных совместных априорных условных плотностей вероятности смешанного типа эталонных векторов $\mathbf{s}$ независимых признаков s_j для каждого из L образов $\mathbf{U}_i$:

$$w_i(s) = W(s|U_i) = \prod_{j=1}^{\Im}\left[\sum_{r=1}^{R_{ij}} I_{ijr}p_{ijr}w_{ijr}(s_j, s'_{ijr}, s''_{ijr}) + \sum_{d=1}^{D_{ij}} I_{ijd}p_{ijd}\delta(s_j - s_{ijd})\right], \quad (1)$$

$$\sum_{r=1}^{R_{ij}} p_{ijr} + \sum_{d=1}^{D_{ij}} p_{ijd} = 1,\ \sum_{r=1}^{R_{ij}} I_{ijr}p_{ijr} + \sum_{d=1}^{D_{ij}} I_{ijd}p_{ijd} = 1,\ \forall I \in \overline{1,L},\ \forall j \in \overline{1,\Im},$$

где $w_{ijr}(s_j, s'_{ijr}, s''_{ijr})$ – априорные плотности распределения признака s_j на каждом из R_{ij} эталонных интервалов $[s'_{ijr}, s''_{ijr}]$, $r \in \{1, 2, ..., R_{ij}\}$; $\delta(s_j - s_{ijd})$ – функции

Дирака, как плотности вероятности математических ожиданий s_{ijd} каждого из D_{ij} возможных дискретных эталонных значений признака s_j, $d \in \overline{1,D}$; p_{ijr} и p_{ijd} – априорные условные вероятности наблюдения r-го интервала или d-го значения при наблюдении образа $\mathbf{U}_i$ в метрике признака s_j; $I_{ijr(d)} \in [0, 1]$ – коэффициенты, характеризующие относительную степень информативности r-го интервала или d-го значения признака s_j при распознавании образа $\mathbf{U}_i$.

В практике создания систем распознавания радиоизлучений часто эталонные описания образов априори известны с точностью до количества элементов, входящих в (1). Если эталонное описание представляет собой совокупность неизвестных априорных распределений признаков и неизвестны функции правдоподобия $W(\mathbf{x} \mid \mathbf{s})$, для обучения применяются непараметрические методы. Для непараметрического оценивания распределений случайных величин широко применяется метод нормальных вкладов [4]. Однако на случай распознавания образов, заданных сложными эталонными описаниями вида (1), этот метод до настоящего времени не обобщался.

Целью статьи является развитие методики синтеза алгоритмов распознавания групп радиоизлучений на случай непараметрического обучения методом нормальных вкладов.

Пусть на множестве $\mathbf{U}$ объектов распознавания наблюдается L образов $\mathbf{U}_i \subset \mathbf{U}$, $i \in \overline{1,L}$, представляющих собой множества (группы) объектов распознавания (видов радиоизлучений). Каждый из образов наблюдается в $\mathfrak{I}$-мерном евклидовом пространстве признаков $\mathbf{S}$. В метрике каждого признака s_j, $j \in \overline{1,\mathfrak{I}}$, каждому образу априорно может соответствовать ν_{ij} эталонных значений и (или) интервалов эталонных значений признака, описывающих группу видов радиоизлучений. Неизвестное априорное распределение вектора признаков $\mathbf{s} = \{s_1, s_2, \dots, s_j, \dots, s_{\mathfrak{I}}\}$ для каждого i-го образа представляет собой $\mathfrak{I}$-мерную совместную плотность вероятности смешанного типа $W(\mathbf{s} \mid \mathbf{U}_i) = w_i(\mathbf{s})$ вида (1) вектора $\mathbf{s}$ на множестве $\mathbf{U}_i$, определенную в области $\mathbf{S}_i$ пространства признаков. По результатам Ξ испытаний получена обучающая выборка – первичная статистическая совокупность $\mathbf{x}^*$, включающая в себя результаты независимых наблюдений всех элементов каждого из образов в метрике каждого признака s_j, $j \in \overline{1,\mathfrak{I}}$. Выдвигается L гипотез $H_1, H_2, \dots, H_L$ о том, что наблюдаемая выборка $\mathbf{x}$ ζ-кратно измеренных значений $\mathfrak{I}$ признаков принадлежит одному из образов $\mathbf{U}_i$. Задача состоит в получении методом нормальных вкладов оценок усредненных функций правдоподобия $w^*_i(\mathbf{x})$ и установлении на их основе дискретно-аналогового нерандомизированного статистически оптимального правила δ, реализующего до наблюдения разделение $\zeta\times\Im$-мерного евклидового пространства выборок **X** на L непересекающихся областей $\mathbf{X}^*_q$, $q\in\overline{1,L}$, $\bigcup_{q=1}^{L} X^*_q = X$, соответствующих каждому из L решений γ_q о принятии гипотезы H_q.

Опишем аналитически одномерную эмпирическую дифференциальную функцию распределения w^*_{ijn}, соответствующую n-му элементу i-й группы (образа) в метрике j-го признака, в виде суммы Ξ_{ijn} нормальных (гауссовских) вкладов [4]:

$$w^*_{ijn}(x_j) = \frac{1}{\Xi_{ijn}\Delta_j\sqrt{2\pi}}\sum_{z=1}^{\Xi_{ijn}}\exp\left[-\frac{\left(x_j - x^*_{ijnz}\right)^2}{2\Delta_j^2}\right], \qquad (2)$$

где Ξ_{ijn} – количество в выборке $\mathbf{x}^*$ наблюдений n-го элемента i-й группы в метрике j-го признака; Δ_j – ширина нормальных вкладов.

Для определения ширины вкладов Δ_j элементы $\mathbf{x}^*_{ijn} = \{x_{ijnz}\}$ обучающей выборки $\mathbf{x}^*$, соответствующие эмпирической функции w^*_{ijn}, ранжируются по величине и представляются в виде вариационного ряда $\{x_{ijnh}\}$, в котором $x_{ijn(h+1)} \geq x_{ijnh}$, $h = 1, 2, \ldots, \Xi_{ijn}$. Ширина вкладов находится по правилу [4]:

$$\Delta_j = 0{,}5\max_h\left(x_{ijn(h+1)} - x_{ijnh}\right).$$

В соответствии с [1 – 3] определим статистики i-го элемента вектора оценок отношений $\Lambda^*_i(\mathbf{x})$ эмпирических усредненных функций правдоподобия $w^*_i(x)$ в виде

$$\Lambda^*_i(x) = w^*_i(x)/w^*_1(x), \qquad (3)$$

где
$$w^*_i(x) = W^*(x\mid U_i) = \int_{S^*_i} w^*_i(s)W^*(x\mid s)ds; \qquad (4)$$

$W^*(\mathbf{x}\mid\mathbf{s})$ – зависящая от значений вектора параметров **s** оценка функции правдоподобия наблюдаемой выборки **x**; $\mathbf{S}^*_i$ – оценка области определения образа $\mathbf{U}_i$ в пространстве признаков **S**.

Полагая признаки и наблюдаемые выборки независимыми, применяя фильтрующее свойство функции Дирака, из (1), (4) имеем

$$w^*_i(\mathbf{x}) = \prod_{j=1}^{\Im}\left[\sum_{r=1}^{R_{ij}} I_{ijr}p^*_{ijr}\int_{S^*_{ij}} w^*_{ijr}(s_j, s'_{ijr}, s''_{ijr})W^*(\mathbf{x}\mid s_j)\,ds_j + \sum_{d=1}^{D_{ij}} I_{ijd}p^*_{ijd}\,W^*(\mathbf{x}\mid s_{ijd})\right]. \qquad (5)$$

Для определения $w^*_i(\mathbf{x})$ необходимо из обучающей выборки получить оценки функций $\int_{S_{ij}} w_{ijr}(s_j, s'_{ijr}, s''_{ijr})W(\mathbf{x}\mid s_j)\,ds_j$, $W(x\mid s_{ijd})$ и вероятностей p_{ijr}, p_{ijd}. Первые два элемента можно определить из (3) заменой индексов n на r и d, соответственно. Априорные вероятности p_{ijr} и p_{ijd} определим как:

$$p^*_{ijr} = \Xi_{ijr}/\Xi_i \; ; \quad p^*_{ijd} = \Xi_{ijd}/\Xi_i \; ,$$

где Ξ_{ijr}, Ξ_{ijd} – количество проявлений r-го интервала и d-го значения i-го образа в метрике j-го признака; Ξ_i – количество наблюдений i-го образа. Из (2), (5), имеем

$$w^*_i(x) = \frac{1}{\left(\sqrt{2\pi}\,\Xi_i\right)^{\Im}} \times$$

$$\times \prod_{j=1}^{\Im} \frac{1}{\Delta_j} \left\{ \sum_{r=1}^{R_{ij}} I_{ijr} \sum_{z=1}^{\Xi_{ijr}} \exp\left[-\frac{\left(x_j - x^*_{ijrz}\right)^2}{2\Delta_j^2} \right] + \sum_{d=1}^{D_{ij}} I_{ijd} \sum_{z=1}^{\Xi_{ijd}} \exp\left[-\frac{\left(x_j - x^*_{ijdz}\right)^2}{2\Delta_j^2} \right] \right\}. \quad (6)$$

На основе (6) могут быть получены алгоритмы, оптимальные относительно наиболее часто применяемых в практике распознавания образов критериев оптимальности. Для этого необходимо в соответствии с применяемым критерием определить оценку отношения правдоподобия (3) и сравнить его с порогом в соответствии с [1 – 3].

Заметим особенность применения стратегии максимального правдоподобия: в (6) необходимо положить априорные равные вероятности наблюдения образов, их компонентов, любого из эталонных интервалов или дискретных значений признаков:

$$p_i = 1/L; \; p_{ijr} = p_{ijd} = 1/(R_{ij} + D_{ij}) = 1/\nu_{ij}; \; I_{ijr(d)} = 1,$$

$$\forall\, i \in \overline{1,L}\,,\; \forall\, j \in \overline{1,\Im}\,,\; \forall\, r \in \overline{1,R_{ij}}\,,\; \forall\, d \in \overline{1,D_{ij}}\,, \text{ т.е.}$$

$$w^*_i(x) = \frac{1}{(2\pi)^{\Im/2}} \times$$

$$\times \prod_{j=1}^{\Im} \frac{1}{\Delta_j \nu_{ij}} \left\{ \sum_{r=1}^{R_{ij}} \frac{1}{\Xi_{ijr}} \sum_{z=1}^{\Xi_{ijr}} \exp\left[-\frac{\left(x_j - x^*_{ijrz}\right)^2}{2\Delta_j^2} \right] + \sum_{d=1}^{D_{ij}} \frac{1}{\Xi_{ijd}} \sum_{z=1}^{\Xi_{ijd}} \exp\left[-\frac{\left(x_j - x^*_{ijdz}\right)^2}{2\Delta_j^2} \right] \right\}.$$

Для прогнозирования качества алгоритма (6) целесообразно использовать полную вероятность ошибки

$$p_{ош} = \sum_{i=1}^{L} p_{ош\,i} = \sum_{i=1}^{L} p_i \sum_{q=1,\, q \neq i}^{L} P\{\gamma_q | \mathbf{U}_i\}\,, \quad \bigcup_{i=1}^{L} X_i = X\,, \quad \bigcap_{i=1}^{L} X_i = \varnothing\,, \quad (7)$$

где p_i – априорные вероятности наблюдения образов $\mathbf{U}_i$, $\sum_{i=1}^{L} p_i = 1$, оценки которых определим в виде $p^*_i = \Xi_i/\Xi$.

Входящие в (7) полные вероятности $P\{\gamma_q|\mathbf{U}_i\}$ принятия ошибочных решений γ_q при наблюдении образа $\mathbf{U}_i$ определяются как усредненные по эталонным описаниям образов условные вероятности ошибок [3]:

$$P\{\gamma_q|\mathbf{U}_i\} = P\{\mathbf{x}\in\mathbf{X}_q|\mathbf{s}\in\mathbf{S}_i\} =$$

$$= \int\limits_{S_i} w_i(s) \int\limits_{X_q} W(x\,|\,s\in S_i)\,dx\,ds = \int\limits_{X_q} w_i(x)\,dx\,. \qquad (8)$$

Из (6) – (8) для оценок $w^*_i(\mathbf{x})$ и областей $\mathbf{X}^*_q$ имеем асимптотическую оценку полной вероятности ошибки

$$p^*_{\text{ош}} = \frac{1}{(2\pi)^{\Im/2}\,\Xi}\sum_{i=1}^{L}\Xi_i^{1-\Im}\times$$

$$\times\sum_{\substack{q=1\\ q\neq i}}^{L}\int\limits_{X^*_q}\prod_{j=1}^{\Im}\frac{1}{\Delta_j}\left\{\sum_{r=1}^{R_{ij}}I_{ijr}\sum_{z=1}^{\Xi_{ijr}}\exp\left[-\frac{\left(x_j-x^*_{ijrz}\right)^2}{2\Delta_j^2}\right]+\sum_{d=1}^{D_{ij}}I_{ijd}\sum_{z=1}^{\Xi_{ijd}}\exp\left[-\frac{\left(x_j-x^*_{ijdz}\right)^2}{2\Delta_j^2}\right]\right\}dx\,.$$

Выводы. Таким образом, разработанный метод позволяет синтезировать непараметрические алгоритмы распознавания с обучением групп радиоизлучений путем непараметрического оценивания отношений эмпирических усредненных функций правдоподобия, сглаженных нормальными вкладами.

ЛИТЕРАТУРА

1. *Певцов Г.В. Синтез алгоритма распознавания радиоизлучений на основе байесовского правила проверки сложных гипотез // Радиоэлектроника. – 1998. – № 4. – С. 49 – 57 (Изв. высш. учебн. заведений).*
2. *Певцов Г.В., Лупандин В.А. Синтез алгоритмов многоальтернативного распознавания образов на основе проверки сложных статистических гипотез по критерию максимума апостериорной вероятности // Радиоэлектроника.– 2001. – № 11. – С. 77 – 80 (Изв. высш. учебн. заведений).*
3. *Певцов Г.В. Синтез байесовских алгоритмов многоальтернативного распознавания образов, заданных сложными эталонными описаниями // Радиоэлектроника. – 2003. – № 1. – С. 58 – 63 (Изв. высш. учебн. заведений).*
4. *Кутин Г.И., Кузнецов А.С. О непараметрической оценке плотности распределения сигналов методом нормальных вкладов // Приборостроение. – 1983. – № 11. – С. 7 – 11 (Изв. высш. учебн. заведений).*

Поступила 11.06.2003

***ПЕВЦОВ Геннадий Владимирович**, доктор техн. наук, ст. научн. сотр., зам. начальника Научного центра Войск ПВО по научной работе. В 1978 году окончил КВИРТУ. Область научных интересов – обработка радиотехнических сигналов и информации.*

***ЛУПАНДИН Владимир Анатольевич**, адъюнкт ХВУ. В 1992 году окончил Харьковское ВВКИУ РВ. Область научных интересов – обработка радиотехнических сигналов и информации.*

***УМАНЕЦ Дмитрий Илларионович**, кандидат военных наук, доцент, сотрудник Укрспецэкспорт. В 1969 году окончил Житомирское ВЗРККУ. Область научных интересов –*

военная кибернетика.