СИСТЕМЫ УПРАВЛЕНИЯ

УДК 621.391:519.28

СИНТЕЗ НЕПАРАМЕТРИЧЕСКИХ АЛГОРИТМОВ РАСПОЗНАВАНИЯ ГРУПП РАДИОИЗЛУЧЕНИЙ НА ОСНОВЕ ЛИНЕЙНОЙ ИНТЕРПОЛЯЦИИ СТАТИСТИЧЕСКИХ УСРЕДНЕННЫХ ФУНКЦИЙ ПРАВДОПОДОБИЯ

Разработан метод синтеза алгоритмов распознавания групп радиоизлучений с обучением путем построения и линейной интерполяции статистических усредненных функций правдоподобия. Полученные результаты развивают на случай непараметрического обучения предложенные ранее подходы к синтезу алгоритмов распознавания образов, заданных сложными эталонными описаниями в виде совокупностей эталонных значений и (или) интервалов эталонных значений признаков (параметров) радиоизлучений.

Г.В. Певцов
Доктор технических наук.

Н.Г. Батурин
Кандидат технических наук.

В.А. Лупандин

При решении задач контроля правильности использования спектра радиочастот излучающими объектами возникает задача идентификации источников радиоизлучений. При этом часто один источник может излучать несколько разных видов радиосигналов. То есть задача идентификации источников радиосигналов (образов) сводится к задаче распознавания групп радиоизлучений. В пространстве признаков (параметров сигналов и их источников) такой образ может описываться одним или несколькими интервалами эталонных значений и (или) одним или несколькими дискретными эталонными значениями признаков. В [1 – 4 и др.] развита методика синтеза алгоритмов распознавания групп радиоизлучений (образов), реализующих проверку сложных статистических гипотез. Методика базируется на введенном сложном эталонном описании образов в виде $\Im$ - мерных совместных априорных условных плотностей вероятности смешанного типа эталонных векторов s независимых признаков s_j, $j\in\{1, 2,\ldots, \Im\}$, для каждого из L образов U_i, $i\in\{1, 2,\ldots, L\}$:

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

$$\sum_{r=1}^{R_{ij}} p_{ijr} + \sum_{d=1}^{D_{ij}} p_{ijd} = 1 \;, \qquad \sum_{r=1}^{R_{ij}} I_{ijr}p_{ijr} + \sum_{d=1}^{D_{ij}} I_{ijd}p_{ijd} = 1$$

$$\forall i\in\{1, 2,\ldots, L\},\ \forall j\in\{1, 2,\ldots,\Im\}$$

где $w_{ijr}(s_j, s'_{ijr}, s''_{ijr})$ - априорные плотности распределения признака s_j на каждом из R_{ij} эталонных интервалов $[s'_{ijr}, s''_{ijr}]$, $r\in\{1, 2,\ldots, R_{ij}\}$; $\delta(s_j-s_{ijd})$ – функции Дирака, как плотности вероятности математических ожиданий s_{ijd} каждого из D_{ij} возможных дискретных эталонных значений признака s_j, $d\in\{1,2,\ldots,D_{ij}\}$; p_{ijr} и p_{ijd} – априорные условные вероятности наблюдения r-го интервала или d-го значения при наблюдении образа 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 U_i.

Решающие правила, получаемые в результате синтеза по разработанной методике, предполагают сравнение с порогом статистик отношений $\Lambda_i(x)=w_i(x)/w_q(x)$ усредненных функций правдоподобия $w_i(x)$ вида [5]

$$w_i(\mathbf{x}) = W(\mathbf{x}|U_i) = \int_{S_i} w_i(\mathbf{s})W(\mathbf{x}\mid\mathbf{s})\mathbf{ds} \quad (2)$$

где $W(x|s)$ – зависящая от значений вектора параметров s функция правдоподобия наблюдаемой выборки x; S_i – область определения образа U_i в пространстве признаков S.

Эталонное описание (1) построено в предположении о том, что виды и параметры априорных распределений известны. Однако в практике создания систем распознавания радиоизлучений чаще возникает необходимость в алгоритмах, которые способны обучаться до или в процессе ведения распознавания. При этом если эталонное описание представляет собой совокупность неизвестных априорных распределений признаков и неизвестны функции правдоподобия $W(x|s)$, обычно применяется непараметрический подход к оцениванию усредненных функций правдоподобия $w_i(x)$. Часто на практике для этого методом гистограмм определяются усредненные статистические функции правдоподобия $w^*_i(x)$. Однако ступенчатый характер получаемых функций приводит к существенным неточностям разделения выборочного пространства на области, соответствующие распознаваемым образам. Простейшим способом выравнивания статистических распределений является применение линейной интерполяции.

Целью статьи является развитие методики синтеза алгоритмов распознавания групп радиоизлучений на случай непараметрического обучения путем построения усредненных линейно интерполированных статистических функций правдоподобия $w^*_i(x)$.

Пусть на множестве U объектов распознавания наблюдается L образов $U_i\subset U$, $i\in\{1,2,\ldots,L\}$, представляющих собой множества (группы) объектов распознавания (видов радиоизлучений). Каждый из образов проявляется в метрике $\Im$-мерного евклидового пространства признаков S. В метрике каждого признака s_j, $j \in \{1,2,\dots,\Im\}$, каждому образу априорно может соответствовать ν_i эталонных значений и (или) интервалов эталонных значений признака, описывающих группу видов радиоизлучений. Неизвестное априорное распределение вектора признаков $s=\{s_1,s_2,\dots,s_j,\dots,s_{\Im}\}$ для каждого i-го образа представляет собой τ-мерную совместную плотность вероятности смешанного типа $W(s|U_i)=w_i(s)$ вида (1) вектора s на множестве U_i, определенную в области S_i пространства признаков. По результатам Ξ испытаний для каждого из образов в метрике каждого признака s_j, $j \in \{1,2,\dots,\Im\}$, получена первичная статистическая совокупность, из которой через группированный статистический ряд в метрике каждого признака могут быть сформированы усредненные линейно интерполированные одномерные статистические функции правдоподобия $w^*_{ij}(x_j)$. Выдвигается L гипотез $H_1,H_2,\dots,H_L$ о том, что наблюдаемая выборка x ζ-кратно измеренных значений $\Im$ признаков принадлежит одному из образов U_i. Задача состоит в установлении до наблюдения дискретно-аналогового нерандомизированного статистически оптимального правила δ, реализующего разделение $\zeta \times \Im$-мерного евклидового пространства выборок X на L непересекающихся областей X^*_q, $q \in \{1,2,\dots,L\}$, $\bigcup_{q=1}^{L} \mathbf{x}^*_q = \mathbf{x}$, и приписывающего каждой из областей одного из L решений γ_q о принятии гипотезы H_q.

Опишем аналитически одномерную эмпирическую линейно интерполированную дифференциальную функцию распределения w^*_{ijn}, соответствующую n-му элемен-

$$w^*_{ijn}(x_j) = \frac{Q_{ijn}}{\Xi_{ijn}} \left[\Xi_{ijn(k-1)} + \left(\Xi_{ijnk} - \Xi_{ijn(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] \quad (3)$$

ту i-й группы (образа) в метрике j-го признака, в виде:

$$k_j = \mathrm{ent}(x_j / \Delta s_j),$$

$$1/Q_{ijn} = \frac{1}{\Xi_{ijn}} \int_{-\infty}^{-\infty} \left[\Xi_{ijn(k-1)} + \left(\Xi_{ijnk} - \Xi_{ijn(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] dx_j$$

$$k_j = \mathrm{ent}(x_j / \boldsymbol{\Delta} s_j);$$

где Q_{ijn} — нормировочный коэффициент,

Ξ_{ijnk} — количество выборочных значений j-го признака, попавших в k-й интервал $[-0{,}5k_j\Delta s_j,\ 0{,}5k_j\Delta s_j]$ при наблюдении (в ходе испытаний) n-го элемента, $n \in \{1,2,\dots,\nu_i\}$, i-й группы; Ξ_{ijn} — количество наблюдений n-го элемента i-й группы в метрике j-го признака.

В соответствии с [1–3] определим статистики i-го элемента вектора оценок отношений $\Lambda^*_i(x)$ усредненных-статистических функций правдоподобия в виде (4)

Для этого, полагая признаки и наблюдаемые выборки независимыми, применяя фильтрующее свойство функ-

$$\Lambda^*_i(\mathbf{x}) = w^*_i(\mathbf{x}) / w^*_1(\mathbf{x}) \quad (4)$$

ции Дирака, из (1), (2) имеем:

$$w_i(\mathbf{x}) = \prod_{j=1}^{\Im} \left[\sum_{r=1}^{R_{ij}} I_{ijr} p_{ijr} \int_{\mathbf{s}_{ij}} w_{ijr}(s_j, s'_{ijr}, s''_{ijr}) W(\mathbf{x}|s_j)\, ds_j + \sum_{d=1}^{D_{ij}} I_{ijd} p_{ijd} W(\mathbf{x}|s_{ijd}) \right] \quad (5)$$

Для получения $w^*_i(x)$ необходимо в (5) вместо $\int_{\mathbf{s}_{ij}} w_{ijr}(s_j, s'_{ijr}, s''_{ijr}) W(\mathbf{x}|s_j)\, ds_j$, $W(\mathbf{x}|s_{ijd})$, p_{ijr} и p_{ijd} подставить их оценки, полученные из обучающей выборки. 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. Основываясь на (3), (5), запишем:

$$w^*_i(\mathbf{x}) = \frac{1}{\Xi_i^{\Im}} \prod_{j=1}^{\Im} \left\{ \sum_{r=1}^{R_{ij}} I_{ijr} Q_{ijr} \left[\Xi_{ijr(k-1)} + \left(\Xi_{ijrk} - \Xi_{ijr(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] + \right.$$

$$\left. + \sum_{d=1}^{D_{ij}} I_{ijd} Q_{ijd} \left[\Xi_{ijd(k-1)} + \left(\Xi_{ijdk} - \Xi_{ijd(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] \right\}$$

$$k_j = \mathrm{ent}(x_j / \Delta s_j), \quad (6)$$

На основе (4), (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 (4) и сравнить его с порогом. В частности, байесовский (Б) алгоритм многоальтернативного распознавания образов, заданных составными эталонными описаниями, имеет вид:

$$\delta_{Б}: \sum_{i=2}^{L} \left(\Pi_{it} - \Pi_{iq} \right) \frac{p^*_i}{p^*_1} \Lambda^*_i(\mathbf{x}) \overset{\gamma_q}{\geq} \Pi_{1q} - \Pi_{1t},$$

$$t=1,2,\dots, L,\ t \neq q,\ q=2,3,\dots,L, \quad (7)$$

где $\Pi_{iq} \geq 0$ — элементы матрицы потерь Π размером $L \times L$; γ_q — решения принять гипотезы H_q; p^*_i — оценки априорных вероятностей наблюдения образов U_i, $p^*_i = \Xi_i / \Xi$, $\sum_{i=1}^{L} p^*_i = 1$. К области $X^*_q, q \in \{2,3,\dots,A\}$, относятся точки выборочного пространства X, удовлетворяющие системе неравенств (7). Область X^*_1 определяется из условия $x^*_1 = x - \bigcup_{q=2}^{L} \mathbf{x}^*_q$.

При применении критерия максимума апостериорной вероятности (MAB) принимается решение γ_q, q=2,3,…,L, если

$$\delta_{MAB}: p^*_q \Lambda^*_q(\mathbf{x}) \overset{\gamma_q}{=} \max_{2 \le i \le L} p^*_i \Lambda^*_i(\mathbf{x}), \quad (8)$$

$$(p^*_i / p^*_1) \Lambda^*_i(x) \geq 1, \quad i=2,3,\dots,L,$$

и решение γ_1, если $(p^*_i / p^*_1) L^*_i(x) < 1$, $\forall i \in \{2,3,\dots,L\}$.

Для реализации стратегии максимального правдоподобия (МП) в (6) необходимо положить априорную равновероятность наблюдения образов, их компонентов, любого из эталонных интервалов или дискретных значений признаков, одинаковую относительную степень их информативности: $p_i = 1/L$, $\forall i \in \{1,2,\dots,L\}$; $p_{ijr} = p_{ijd} = 1/(R_{ij} + D_{ij})$, $I_{ijr(d)} = 1$, $\forall i \in \{1,2,\dots,L\}$, $\forall j \in \{1,2,\dots,\Im\}$, $\forall r \in \{1,2,\dots,R_{ij}\}$; $\forall d \in \{1,2,\dots,D_{ij}\}$, т.е.

$$w^*_i(\mathbf{x}) = \prod_{j=1}^{\Im} \left\{ \frac{1}{R_{ij} + D_{ij}} \sum_{r=1}^{R_{ij}} \frac{Q_{ijr}}{\Xi_{ijr}} \left[\Xi_{ijr(k-1)} + \left(\Xi_{ijrk} - \Xi_{ijr(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] + \right.$$

$$\left. + \frac{1}{R_{ij} + D_{ij}} \sum_{d=1}^{D_{ij}} \frac{Q_{ijd}}{\Xi_{ijd}} \left[\Xi_{ijd(k-1)} + \left(\Xi_{ijdk} - \Xi_{ijd(k-1)} \right) \frac{x_j - (k_j - 1)\Delta s_j}{\Delta s_j} \right] \right\}$$

$$k_j = \mathrm{ent}(x_j / \Delta s_j), \quad (9)$$

Принимается решение γ_q, q=2,3,…,L, если

$$\delta_{МП}: \Lambda^*_q(\mathbf{x}) \overset{\gamma_q}{=} \max_{2 \le i \le L} \Lambda^*_i(\mathbf{x}), \quad \Lambda^*_i(x) \geq 1,\ i=2,3,\dots,L, \quad (10)$$

и решение γ_1, если $\Lambda_i(x)<1$, $\forall i=2,3,\dots,L$.

Для анализа полученных решающих правил обычно определяют полную вероятность ошибки $p_{ош}$ вида

$$p_{ош} = \sum_{i=1}^{L} p_{ош\,i} = \sum_{i=1}^{L} p_i \sum_{\substack{q=1\\ q\neq i}}^{L} P\{\gamma_q|\mathbf{U}_i\} \quad (11)$$

$$\bigcup_{i=1}^{L} \mathbf{X}_i = \mathbf{X}, \quad \bigcap_{i=1}^{L} \mathbf{X}_i = \varnothing$$

В идеальном случае полностью известных компонентов усредненных функций правдоподобия $w_i(x)$ входящие в (11) полные вероятности $P\{\gamma_q|U_i\}$ принятия ошибочных решений γ_q при наблюдении образа U_i определяются как усредненные по эталонным описаниям образов условные вероятности ошибок [3,4]

$$P\{\gamma_q|U_i\}=P\{x\in X_q|s\in S_i\}=\int_{\mathbf{S}_i} w_i(\mathbf{s}) \int_{\mathbf{X}_q} W(\mathbf{x}\,|\,\mathbf{s}\in \mathbf{S}_i)\,\mathbf{dx}\,\mathbf{ds}, \quad (12)$$

или

$$P\{\gamma_q|U_i\}= \int_{\mathbf{X}_q} \int_{\mathbf{S}_i} w_i(\mathbf{s}) W(\mathbf{x}\,|\,\mathbf{s}\in \mathbf{S}_i)\,\mathbf{ds}\,\mathbf{dx} = \int_{\mathbf{X}_q} w_i(\mathbf{x})\,\mathbf{dx}. \quad (13)$$

Из (6), (11) и (13) для оценок $w^*_i(x)$ и областей X^*_q имеем асимптотическую оценку полной вероятности ошибки:

$$p^*_{ош}=\frac{1}{\Xi}\sum_{i=1}^{L}\frac{1}{\Xi_i^{\Im-1}}\sum_{\substack{q=1\\ q\neq i}}^{L}\int_{\mathbf{X}^*_q}\prod_{j=1}^{\Im}\left\{\sum_{r=1}^{R_{ij}} I_{ijr} Q_{ijr}\left[\Xi_{ijr(k-1)}+\left(\Xi_{ijrk}-\Xi_{ijr(k-1)}\right)\frac{x_j-(k_j-1)\Delta s_j}{\Delta s_j}\right]+\right.$$

$$\left.+\sum_{d=1}^{D_{ij}} I_{ijd} Q_{ijd}\left[\Xi_{ijd(k-1)}+\left(\Xi_{ijdk}-\Xi_{ijd(k-1)}\right)\frac{x_j-(k_j-1)\Delta s_j}{\Delta s_j}\right]\right\}\mathbf{dx}$$

$$k_j=\mathrm{ent}(x_j/\Delta s_j). \quad (14)$$

Зачастую может оказаться сложно в явном виде определить (12)–(14) из-за необходимости вычисления кратных интегралов. В этом случае целесообразно отыскать дифференциальные функции распределения статистик (4) оценок отношений правдоподобия $\Lambda^*_i(x)$, например, методом Монте-Карло. Затем по известной методике [5] можно перейти к однократному интегрированию этих функций.

Таким образом, разработанный метод позволяет синтезировать непараметрические алгоритмы распознавания групп радиоизлучений с обучением путем получения для каждой группы первичной статистической совокупности и формирования из них отношений усредненных линейно интерполированных статистических функций правдоподобия.

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. Синтез алгоритмов распознавания образов, заданных сложными эталонными описаниями в метрике азимутов на источники радиоизлучений // Радиоэлектроника.– 2000.– №4. – С.38–45. (Изв. высш. учебн. заведений).

3. 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).

4. Певцов Г.В. Синтез байесовских алгоритмов многоальтернативного распознавания образов, заданных сложными эталонными описаниями // Радиоэлектроника. – 2003. – №1. – С.58–63. (Изв. высш. учебн. заведений).

5. Левин Б.Р. Теоретические основы статистической радиотехники. Книга 2.– М.: Сов радио, 1975.– 392 с.