

ЭНТРОПИЙНЫЕ МЕТОДЫ ОБРАБОТКИ СИГНАЛЬНОЙ ИНФОРМАЦИИ В АДАПТИВНЫХ СИСТЕМАХ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ

Куренков Н.И.¹, Ананьев С.Н.¹, Варакин А.Н.²

¹Центральный научно-исследовательский институт МО РФ, ²НИЦ распознавания образов

Одной из наиболее сложных исследовательских задач, с которой сталкиваются разработчики систем распознавания объектов по их сигналам, является создание адаптивных к условиям неопределенной фоновой и объектовой обстановки алгоритмов обработки сигнальной информации.

Сигнальная информация поступает на вход таких алгоритмов, как правило, в виде многомерных спектральных признаков сигналов (например, радиолокационных доплеровских, акустических, сейсмических) объектов. В подавляющем большинстве случаев такие сигналы не могут быть представлены в виде стационарного процесса, и поэтому применение к ним статистических методов обработки данных, как правило, приводит к ошибкам классификации.

Предлагается один из возможных подходов к построению адаптивных алгоритмов обработки сигнальной информации для систем автоматического распознавания объектов по их акустическим сигналам, основанный на использовании энтропийных свойств обратносимметричных квадратных матриц [2]. Рассмотрим две важнейшие задачи, которые должны решаться в системах автоматического распознавания: 1 - обнаружения полезного сигнала и 2 - оценки информативности признаков объектов распознавания.

1. При решении задачи обнаружения полезного сигнала алгоритм должен обеспечивать непрерывный анализ отличий структур спектральных компонент многомерных входных признаков (везде далее будем называть их входными векторами), принятие решения об обнаружении полезного сигнала при его появлении (появлении существенных отличий в структурах значений спектральных признаков), а после того, как этот сигнал исчезнет – о его отсутствии. Принятие решения об обнаружении полезного сигнала должно осуществляться по факту одновременного превышения значений нескольких частотных компонент входного вектора адаптивного порога обнаружения в таких векторах в течение определенного интервала времени.

Алгоритм обнаружения должен оценивать отличия спектральных признаков полезного сигнала и фона (помехи) и определять начало t_0 и конец $t_0 + \Delta t$ интервала времени, в течение которого наблюдается полезный сигнал. При этом, исходя из общих требований к качеству решения задачи обнаружения, необходимо выполнение ряда условий:

а) Моменты t_0 и $t_0 + \Delta t$ должны быть как можно ближе к истинным значениям, что позволит более точно определить длительность Δt нахождения объекта в зоне обнаружения устройства и избежать потерь информации об объекте.

б) Формируемый критерий обнаружения полезного сигнала должен быть чувствителен к большим и не чувствителен к малым значениям амплитуд сигнала. Соблюдение этого условия позволит добиться существенного превышения значений указанного критерия над фоновыми значениями и, следовательно, выбрать порог принятия решения, обеспечивающий сохранить полезную информацию о классах распознавания в анализируемых признаках.

в) Критерий должен быть максимально простым, а его вычисления должны проводиться в масштабе реального времени.

Анализ известных методов обнаружения полезных сигналов показал, что они не в полной мере отвечают предъявляемым требованиям. Предлагается новый метод, основанный на сопоставлении оценок амплитуд сигналов в частотных полосах гребенки полосовых фильтров (в пяти частотных полосах), из которых три соответствуют акустическому информационному каналу и две - сейсмическому. Диапазон частот акустического канала был разбит следующим образом: от 5 Гц до 60 Гц – первая полоса, от 60 Гц до 120 Гц - вторая полоса и от 120 Гц до 900 Гц – третья полоса. Аналогично для сейсмического канала, - от 5 Гц до 120 Гц – первая полоса, от 120 Гц до 220 Гц с перекрытием по уровню 0,7- вторая полоса.

Таким образом, в каждый момент времени входной сигнал индуцирует матрицу T , состоящую из пяти строк и n столбцов, где число столбцов соответствует времени наблюдения сигнала. Для выявления изменений в структуре сигнала будем передвигаться по столбцам матрицы окном размерности c . Результатом такого сканирования массива входных векторов будет матрица $X_j = T(1:5, j: j + c - 1)$, состоящая из c столбцов матрицы T .

Для вектор-столбцов с положительными компонентами и одинаковой размерности \mathbf{u} , \mathbf{v} и \mathbf{a} определим функционал $F(\mathbf{a}, \mathbf{u}, \mathbf{v}) = \mathbf{a}^{-T} \mathbf{u} + \mathbf{a}^T \mathbf{v}$.

(1)

Сформулируем следующие утверждения: $\forall \mathbf{a} \in \mathbf{R}_+^n = \{\mathbf{x} = (x_1, x_2, \dots, x_n) \mid x_i > 0 \forall i = \overline{1, n}\}$

$$1. F(\mathbf{a}, \mathbf{u}, \mathbf{v}) \geq 2\sqrt{\mathbf{u}^T \mathbf{v}} \quad (2)$$

$$2. F(\mathbf{a}, \mathbf{u}, \mathbf{v}) = 2\sqrt{\mathbf{u}^T \mathbf{v}} \Leftrightarrow a_i = \sqrt{u_i/v_i} \quad (3)$$

Первое утверждение следует из неравенства $\xi + \eta \geq 2\sqrt{\xi\eta} \forall \xi, \eta \in \mathbf{R}_+$, второе – из равенства $F(\sqrt{\mathbf{u}/\mathbf{v}}, \mathbf{u}, \mathbf{v}) = 2\sqrt{\mathbf{u}^T \mathbf{v}}$ (оператор $\sqrt{\cdot}$ применяется к каждой компоненте вектора).

Из (3) следует, $\mathbf{u} = \mathbf{v} \Leftrightarrow \alpha_i = 1 \forall i = \overline{1, n}$, что эквивалентно условию $\tau(\mathbf{a}) = \max\left(\sum \alpha_i, \sum \frac{1}{\alpha_i}\right) = n$ (4)

Применительно к задаче обнаружения (4) трансформируется в критерий следующего вида:

Если $\mathbf{e} = (1, 1, \dots, 1)^T$ и t_j - момент времени, соответствующий j -ому отсчету, то считаем, что обнаружение произошло в момент t_j , если $\tau(\mathbf{a}_j^*) \geq \pi$ для $\mathbf{a}_j^* = \arg \min_{\mathbf{a}} F(\mathbf{a}, X_j \mathbf{e}, X_{j+1} \mathbf{e})$ и $\tau(\mathbf{a}_{j-1}^*) < \pi$ - в противном случае. Момент t_{j+k} выхода объекта из зоны обнаружителя определяется из условия $\tau(\mathbf{a}_s^*) \geq \pi \forall s = \overline{j, j+k}$ и $\tau(\mathbf{a}_{j+k+1}^*) < \pi$. Значение порога π , как показали исследования, целесообразно выбрать равным $\pi = 6$.

На рис. 1 представлены графики изменения величины $y = \tau(\mathbf{a}_j^*)$ для окна $c = 10$ и компонент вектора \mathbf{a}_j^* . Как видно, максимальное и минимальное значения y отличаются друг от друга более чем в три раза, тогда как фоновые значения y практически не отличаются от минимального значения равного пяти. Наш опыт исследований показывает, что значения функции (4) практически не зависят от величины окна $c = 3, \dots, 15$, что является несомненным достоинством предлагаемого метода. Компоненты вектора \mathbf{a} функционала (1) указывают наиболее информативные частотные каналы, т.е. каналы с наибольшим отношением сигнал/шум.

В частности, на рис. 1б) показаны изменения компонент вектора \mathbf{a}_j^* для каждого момента времени j . Из графиков хорошо видно, что наибольшие отклонения соответствуют двум последним каналам - сейсмическим.

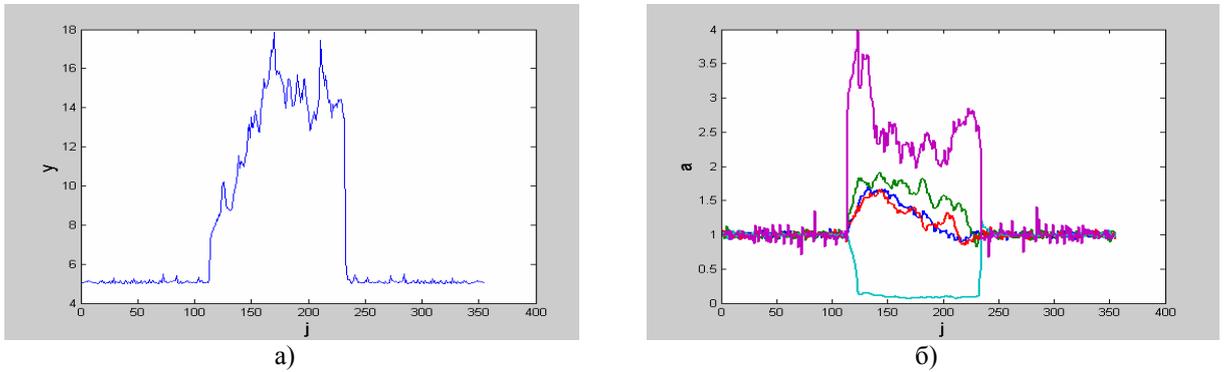


Рис.1 - Графики изменения величины а) $y = \tau(\mathbf{a}_j^*)$ и б) компонент вектора \mathbf{a}_j^*

2. Типовая функция, которую должен выполнять алгоритм автоматической классификации при обработке сигнальной информации после обнаружения полезного сигнала, является измерение отличий между многомерными признаками. Правильный выбор меры отличий имеет принципиальное значение для построения алгоритма выявления и выбора информативных признаков в сигнале, поскольку известно, что использование неинформативных признаков снижает эффективность классификации [1].

Сравнение вектора признаков $\mathbf{u} = (u_1, \dots, u_n)$ распознаваемого объекта с эталонами осуществляется с использованием различных мер близости. Наиболее часто для построения процедур сравнения используют аддитивные меры, например,

$$\rho(\mathbf{u}, \mathbf{v}) = \sqrt{\sum_{i=1}^n r_i (u_i - v_i)^2}, \quad (5)$$

где $\mathbf{r} \in \Delta_n = \{\mathbf{p} = (p_1, \dots, p_n) \mid \sum_{i=1}^n p_i = 1, p_i \geq 0 \forall i = \overline{1, n}\}$.

Широкое применение аддитивных мер обусловлено их инвариантностью по отношению к сдвиговым преобразованиям $\delta(\cdot) : \exists \mathbf{d} \in \mathbb{R}_+^n : \delta(\mathbf{a}) = \mathbf{a} + \mathbf{d} \forall \mathbf{a} \in \mathbb{R}_+^n$. (6)

Однако для систем распознавания объектов по их сигналам большую ценность представляют меры, обладающие инвариантностью к мультипликативным преобразованиям $\gamma(\cdot)$:

$$\exists \mathbf{d} \in \mathbb{R}_+ : \gamma(\mathbf{a}) = \mathbf{a} \cdot \mathbf{d} \forall \mathbf{a} \in \mathbb{R}_+^n. \quad (7)$$

Это обусловлено тем, что принципиально важным для рассматриваемых систем является оценка степени сходства структур значений признаков входного вектора и эталона независимо от значений самих признаков. Это является выражением требования распознавания входных спектров независимо от их мощности («громкости»). Легко увидеть, что мера (5) не инвариантна к преобразованиям (7) и поэтому не может быть эффективной при построении алгоритмов классификации в таких условиях. Кроме того, необходимо, чтобы используемая мера была чувствительна не только к большим, но и к малым значениям амплитуд частотных составляющих спектра, поскольку, информации о классе объекта содержится также в структуре именно этих значений (близких к уровню шумов).

$$\text{Определим [2] для } \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n \text{ значение } \mu(\mathbf{x}, \mathbf{y}) = \ln \left(\sum_{i=1}^n \frac{x_i}{y_i} \cdot \sum_{i=1}^n \frac{y_i}{x_i} \right) \quad (8)$$

Можно показать, что (8) обладает всеми свойствами меры, за исключением одного: $\mu(\mathbf{x}, \mathbf{y})$ совпадает с минимально возможным значением, равным $2 \ln n$, только тогда, когда вектора \mathbf{x} и \mathbf{y} коллинеарные и имеют одинаковое направление $\min_{\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n} \mu(\mathbf{x}, \mathbf{y}) = 2 \ln n \Leftrightarrow \exists \gamma > 0 : \mathbf{x} = \gamma \cdot \mathbf{y}$ (9)

Указанное отличие свойства (9) от традиционного $\mu(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$ не является обременительным для применения меры близости (8) в задачах распознавания, поскольку условие $\exists \gamma > 0 : \mathbf{x} = \gamma \cdot \mathbf{y}$ достигается при ее минимальном значении.

$$\text{Для } \forall \mathbf{x} \in \mathbb{R}_+^n \text{ можно также определить значение } \eta(\mathbf{x}) = \ln \left(\sum_{i=1}^n x_i \cdot \sum_{i=1}^n \frac{1}{x_i} \right), \quad (10)$$

которое характеризует удаление вектора \mathbf{x} от единичного вектора $\mathbf{e} = (1, 1, \dots, 1)$ с точки зрения меры (8) и, следовательно, эту величину можно использовать при оценке информативности признаков распознавания.

Обобщением предложенной меры близости является следующая мера

$$\mu(\mathbf{p}, \mathbf{x}, \mathbf{y}) = \ln \left(\sum_{i=1}^n \sqrt{p_i} \frac{x_i}{y_i} \cdot \sum_{i=1}^n \sqrt{p_i} \frac{y_i}{x_i} \right), \quad (11)$$

для которой справедлива нижняя оценка, совпадающая с энтропией Реньи [4]

$$\mu(\mathbf{p}, \mathbf{x}, \mathbf{y}) \geq 2 \ln \left(\sum_{i=1}^n \sqrt{p_i} \right) \forall \mathbf{p} \in \Delta_n = \{ \mathbf{p} = (p_1, p_2, \dots, p_n) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1 \} \quad (12)$$

Меру (12) примем за основу при построении меры различий между векторами \mathbf{x} и \mathbf{y} :

$$\delta(\mathbf{q}, \mathbf{x}, \mathbf{y}) = \ln \left(\sum_{i=1}^n q_i \frac{x_i}{y_i} \cdot \sum_{i=1}^n q_i \frac{y_i}{x_i} \right) = \mu(\mathbf{p}, \mathbf{x}, \mathbf{y}) - 2 \ln \left(\sum_{i=1}^n \sqrt{p_i} \right) \quad (13)$$

$$\text{при } q_i = \sqrt{p_i} / \left(\sum_{i=1}^n \sqrt{p_i} \right) \quad (14)$$

Поскольку имеет место (14), то, очевидно, $\mathbf{q} \in \Delta_n$ и, следовательно, $\delta(\mathbf{q}, \mathbf{x}, \mathbf{y})$ есть отношение средней арифметической и средней гармонической оценок.

Для оценки информативности частотных признаков объектов предлагается метод, базирующийся на использовании энтропийных свойств многомерных данных, представленных в виде матриц связи специального вида [2].

Пусть \mathbf{X} -матрица, составленная из признаков объектов распознаваемых классов по столбцам, строки соответствуют элементам обучающей выборки всех классов распознавания и представляют собой наборы их спектров.

Каждому j -му признаку, представляемого \mathbf{X}_j j -ым столбцом матрицы \mathbf{X} поставим в соответствие величину

$$H(\mathbf{X}_j) = \ln(\mathbf{e}^T \mathbf{X}_j \mathbf{X}_j^{-T} \mathbf{e}) - \min_{W \in \Omega} \ln(W^{-T} \mathbf{X}_j \mathbf{X}_j^{-T} W), \quad (15)$$

где $\Omega = \{W = (w_1, w_2, \dots, w_n) \mid w_i > 0, w_i = w_j, \forall i = \overline{1, n}; \forall r = \overline{1, k}\}$, w_i - i -я компонента вектора W соответствующая i -ой строке матрицы \mathbf{X} или i -му элементу обучающей выборки принадлежащему r -ому классу распознавания, \mathbf{e} - вектор-столбец, составленный из n единиц по числу всех элементов обучающей выборки.

Непосредственными вычислениями можно убедиться, что (15) принимает максимально возможное значение для тех признаков, которые имеют близкие значения на элементах принадлежащих обучающей выборке одного класса и существенно разные - на элементах выборок других классов. Практически это было подтверждено при проверке работоспособности предлагаемого метода, реализованного в специальном алгоритме оценки информативности частотных признаков. Исходными данными при этом была совокупность массивов входных векторов частотных признаков акустических сигналов движущихся объектов техники различных классов (гусеничный тягач - ГТ, автомобиль КамАЗ и двухвальный гусеничный тягач - МЗКТ), полученных с использованием 120-ти элементной гребенки цифровых полосовых фильтров. На рисунке 2 приведен график информативности частотных признаков акустических сигналов этих объектов. Как следует из рисунка наиболее информативные частотные признаки (полосы гребенки фильтров) объектов лежат в диапазоне от 20 до 400 Гц (2-40 частотные полосы).

Для последующего выбора информативных признаков можно поступить следующим образом. Выбирается некоторый порог информативности π . Для всех признаков $\{\mathbf{X}_j\}$, для которых выполняется условие $H(\mathbf{X}_j) > \pi$, на элементах обучающей выборки строится матрица перепутывания $\mathbf{P} = (p_{ij})$ с вероятностями отнесения i -го элемента к j -ому классу распознавания (в случае отсутствия представительной статистики под вероятностью понимаем частоту). Тогда вероятность распознавания всех объектов обучающей выборки (т.е. отнесения их к своему классу), очевидно, есть $p_o = \sum_{i=1}^k p_{ii}$. Указанная вероятность есть функция параметра π . Путем максимизации этой функции можно выбрать оптимальный порог $\pi_0 = \arg \min_{\pi} p_o(\pi)$, при котором будет обеспечено максимальное значение вероятности $p_o(\pi_0)$ и определены все признаки $\{j\}$, удовлетворяющие условию $H(\mathbf{X}_j) > \pi_0$. Эти признаки и следует использовать при распознавании объектов. Их число, как правило, будет намного меньше исходных. На реальных примерах было показано, что число признаков можно сократить более чем в два раза при увеличении вероятности правильного распознавания вышеуказанных объектов с 0.8 до 0.86.

К оценке информативности признаков можно подойти и по другому, - с использованием формулы (13). Для выявления отличий распространим действие функционала (13) на матрицы с положительными вещественными элементами. Пусть \mathbf{X}_j j -ый столбец матрицы \mathbf{X} , строки которой анализируемые признаки, столбцы - объекты им соответствующие. Мету различий строк матрицы \mathbf{X} определим выражением

$$\delta(\mathbf{q}, \mathbf{X}) = \sum_{j=1}^m \ln(\mathbf{q}^{-T} \cdot \mathbf{X}_j \mathbf{X}_j^{-T} \cdot \mathbf{q}) \quad \mathbf{q} \in \Delta_n, \quad (16)$$

где \mathbf{X}_j^{-T} - транспонированный j -ый столбец матрицы \mathbf{X} с обратными значениями ее элементов x_{ij} . Путем максимизации формулы (16) определим те строки, которые в наибольшей степени отличаются от других. Именно в них (в признаках) содержится полезная информация.

С использованием (16) была оценена информативность частотных признаков в речевых сигналах мужских гласных фонем. Исходными данными были спектры гласных фонем а, о, у, ы, и, е и э, произнесенные одним диктором мужчиной со 128 компонентами во входном векторе.

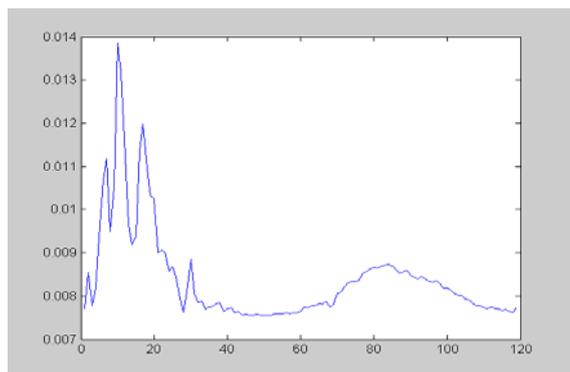
Как видно из рисунка значения информативности частотных признаков ясно характеризуют четыре фанранты речевого сигнала, что хорошо согласуется с известной моделью его генерации Фланагана.

Исследования выполнены при поддержке РФФИ (грант № 05-08-65501).

Литература

1. Айвазян С.А., Енюков И.С., Мешалкин Л. Д. Прикладная статистика: Классификация и снижение размерности. - М.: Финансы и статистика. 1989.-608 с.
2. Куренков Н.И., Лебедев Б.Д. Энтропийный анализ многомерных данных. //Современные проблемы механики гетерогенных сред. Сборник трудов Института прикладной механики РАН к 10-летию его основания. М., РАН, 2000.
3. Лбов Г.С. Методы обработки разнотипных экспериментальных данных — Новосибирск: Наука, 1981. — 158 с.

4. Трухаев Р.И. Модели принятия решений в условиях неопределенности — М.: Наука, 1981. — 258 с.



Номер частотной полосы, $N (f = N \times 10, \text{Гц})$

Рис. 2 Информативность частотных признаков объектов техники (по оси ординат)

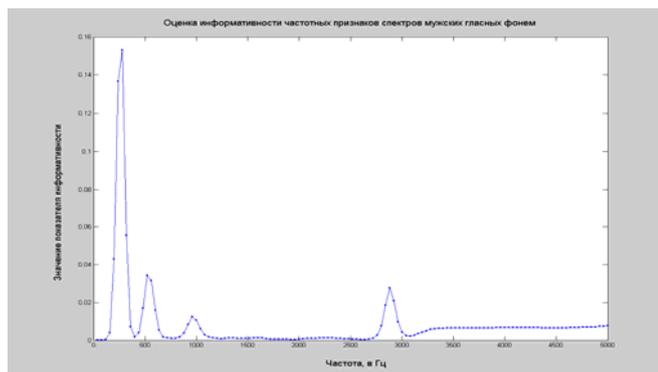


Рис.3 Информативность частотных признаков мужских гласных фонем одного диктора с использованием формулы (16).

ENTROPIC METHODS FOR SIGNAL PROCESSING IN ADAPTIVE PATTERN RECOGNITION SYSTEMS

Kurenkov N.I.¹, Ananiev S.N.¹, Varaksin A.N.²

¹3 Central Scientific Research Institute of Department of Defence of Russian Federation, ²Scientific Research Center for Pattern Recognition

Adaptive signal processing methods for automatic object signal recognition systems are proposed. The basic idea is founded on the utilization of the entropic properties of inversely symmetric matrices.

The methods are implemented in algorithms for detection of friendly signals and estimation of information value for most important features in the frequency domain for recognition tasks to make sure that multidimensional signal data processing will operate in real-time operation mode. An operational capability of proposed algorithms in processing of temporary spectra obtained from acoustic speech signals and technical units is demonstrated. The results of information value estimation for 120 frequency features in vowel phoneme signals and signals obtained from technical units. The selection of the most informative features for technical units enables the probability to be increased (from 0,8 to 0,86) while reducing the number of features more than twice.

