

## НЕЙРОСЕТЕВОЙ МЕТОД ЦИФРОВОЙ ОБРАБОТКИ НЕСТАЦИОНАРНОГО СИГНАЛА (РЕЧИ) В ВЕЙВЛЕТ-ОБЛАСТИ

*А.В. Меркушева*

### Основные принципы

**Введение.** Детектирование момента изменения свойств нестационарного сигнала является необходимым условием фильтрации с адаптивной подстройкой порога, величина которого зависит от уровня шума. Такая воз-

можность предоставляется в информационно-измерительной системе (ИИС), где зашумленный сигнал, несущий полезную информацию, регистрируется на временных интервалах, которые перемежаются с интервалами, свободными от сигнала. Типичный случай этого варианта контроля и обработки реализуется в информационной системе анализа речевого сигнала. Система служит моделью достаточной общности для нестационарных сигналов в ИИС со спорадически появляющимися информационно значимыми интервалами различной длительности. Другим примером системы такого вида является ИИС, предназначенная для контроля гетерогенного потока товарной нефти [1], где выполняется выделение временных сегментов, на которых отсутствует свободный газ, определяется средняя плотность потока при наличии и отсутствии газа, а затем на основе этих данных вычисляется средняя относительная доля свободного газа.

Метод детектирования изменения свойств нестационарного сигнала основан на вейвлет-преобразовании и нейросетевом алгоритме и рассматривается применительно к информационной системе многоцелевой обработки речевого сигнала (РС). Адаптивная фильтрация шума реализуется путем динамического изменения порога, вычисляемого на интервалах (микропаузах) отсутствия РС, поэтому правильное определение наличия или отсутствия речевого сигнала на анализируемом интервале оказывает существенное влияние на качество фильтрации. Эта задача является актуальной в различных системах обработки РС:

- в цифровой телефонии для удаления пауз при сжатии речевых сигналов;
- в криптографии, где удаление пауз сокращает избыточность, уменьшающую криптостойкость алгоритма шифрования;
- в устройствах фильтрации речевого сигнала с адаптацией к виду и уровню шума, где во время пауз должны быть получены кратковременные характеристики шума.

Таким образом, в описанных ИИС и других приложениях необходимо детектирование (отслеживание и идентификация) изменения свойств нестационарных процессов, а также адаптация к изменению уровня шума. Это особенно важно для обработки сигнала в реальном времени, при

Рассмотрены основные принципы нейросетевого метода детектирования изменения свойств нестационарного речевого сигнала (РС) в области вейвлет-разложения. Представлены результаты разработки и исследования нейросетевого метода цифровой обработки РС: выбор сложности многослойного персептрона на нейронной сети (НС); определение лучшего алгоритма нелинейной оптимизации для обучения НС; способ снижения размерности входных векторов при обучении НС, основанный на принципе главных компонент.

которой алгоритм обработки должен адаптироваться к состоянию процесса.

В разработанном методе использовано вейвлет-разложение речевого сигнала по биортогональному базису Добеши [2], а для принятия решения о типе интервала (сегмента) сигнала – нейронная сеть на многослойном персептроне. Сравнительно с традиционными спектральными методами вейвлет-преобразование дает более точную локализацию сигнала по времени и по частоте (в субполосах разложения), имеет быстрый алгоритм реализации. Биортогональный базис сохраняет фазовые соотношения частотных компонент сигнала после его восстановления обратным вейвлет-преобразованием. Свойство наличия у базиса преобразования конечного числа нулевых моментов гарантирует точную аппроксимацию полиномиальной части сигнала, что ведет к дополнительному снижению объема данных [3].

**Вейвлет-разложение речевого сигнала в соответствии с персептуальной моделью.** С целью адаптации цифровой обработки сигнала к переменному во времени шуму, при решении задачи распознавания речи и пауз предложен подход на основе вейвлет-преобразования и нейронной сети [7, 8]<sup>1</sup>. Вейвлет-преобразование позволяет более точно локализовать частотные свойства сигнала во времени. Биортогональные вейвлеты Добеши позволяют уменьшить объем вычислений при разложении за счет использования коротких фильтров. Вейвлет-функция и масштабированная (скейлинг) функция, используемые для разложения сигнала, представлены на рис.1. Вейвлет-разложение не приводит к увеличению объема данных при переходе от временного представления сигнала к его представлению в вейвлет-области [9].

Алгоритм разделения речи и пауз построен с учетом особенностей восприятия звука человеком [10,11,12], которые описываются персептуальной моделью [13,14,15]. Модель разделяет спектр речевого сигнала на частотные полосы, называемые критическими. Каждая полоса частотного диапазона речи по компоненте шума в ней воспринимается как единое целое, и по слуховому ощущению важна лишь мощность шума в полосе [16]. Персептуальная модель успешно используется в алгоритмах MPEG (Moving

Picture Expert Group по стандартизации аудио- и видеоинформации) для кодирования акустических сигналов.

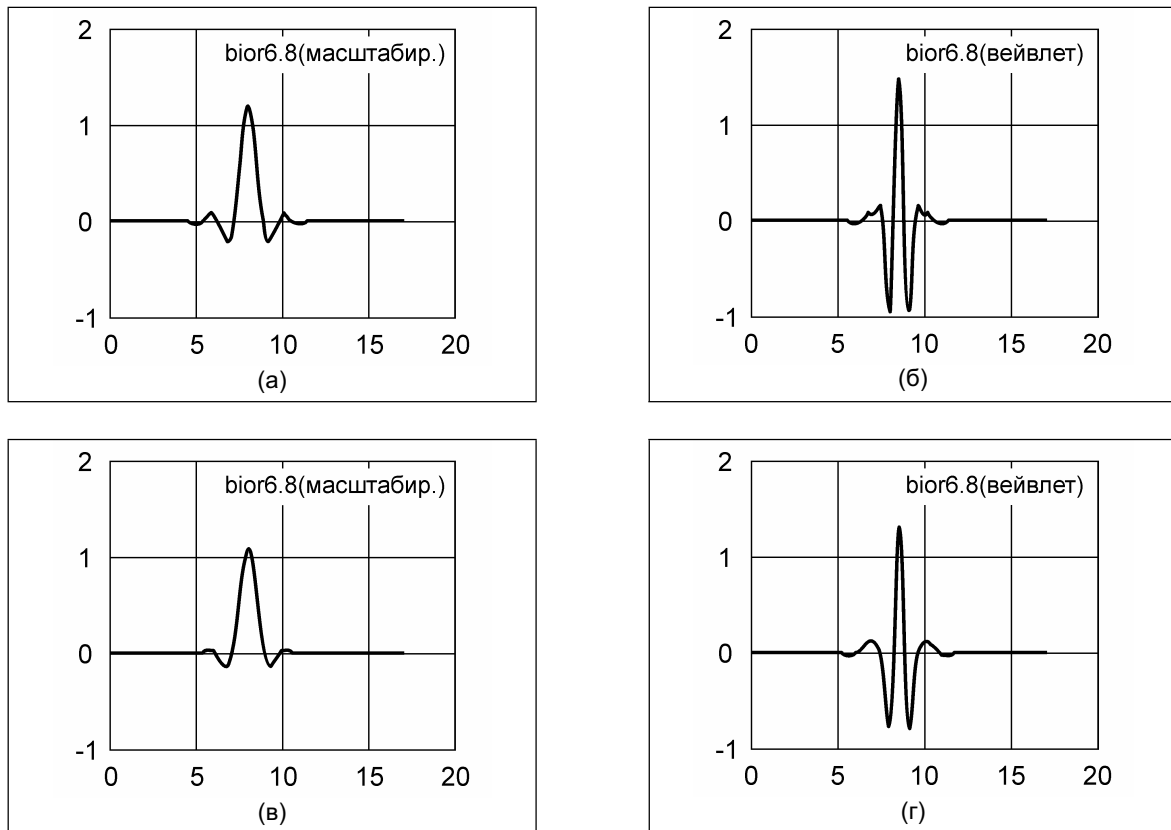


Рис.1. Базисные функции, используемые при разложении сигнала с помощью вейвлет-пакетов: а – масштабирующая функция анализа; б – вейвлет-функция анализа; в – масштабирующая функция синтеза; г – вейвлет-функция синтеза

По перцептуальной модели частотному диапазону 100–12000 Гц соответствует 22 критические полосы, ширина которых постепенно возрастает от 100 до 2500 Гц (табл.1). Диапазон РС ограничен частотой 8 кГц, однако в компьютерных системах обработки РС для уменьшения шума, связанного с дискретизацией и квантованием, используют повышенную частоту дискретизации, в частности 22 кГц. Даже при 10%-й неточности оценки частотного диапазона речи, начиная от самых низких частот, – в пределах 50–9500 Гц, его полностью перекрывают 22 полосы перцептуальной модели. Поэтому в системе обработки РС следует использовать 22 полосы перцептуальной модели.

Для определения речевой активности предложен метод [17], который включает:

- получение сегмента речевого сигнала длительностью 10 – 20 мс;
- использование вейвлет-пакетов для получения разло-

жения сегмента в соответствии с перцептуальной моделью;

- вычисление мощности вейвлет-коэффициентов в каждой области разложения;
- нормирование вектора мощности вейвлет-коэффициентов, полученного на сегменте речевого сигнала;
- использование нормированного вектора в качестве входного вектора нейронной сети.

Вейвлет-преобразование сигнала производит октавополосное разбиение спектра, которое подходит для большинства (но не для всех) задач обработки сигнала. В частности, для получения перцептуальной модели необходимо изменить структуру разбиения частотно-временной плоскости. Достичь гибкого разбиения можно с помощью каскадного соединения блоков вейвлет-фильтров, которое реализуется в методе вейвлет-пакетов. Метод вейвлет-пакетов позволяет на каждом уровне разложения РС производить

<sup>1</sup> Изменение свойств речевого сигнала на интервалах речь/пауза обычно определяется изменением спектральных характеристик. В частности, алгоритм Рабинера (L. Rabiner) [4], разработанный для разделения временных интервалов речевой активности и пауз, построен на принципе анализа энергии сигнала и числа его переходов через нулевой уровень. Алгоритм не учитывает особенностей частотного спектра сигнала и шума и поэтому не всегда позволяет правильно классифицировать интервалы, особенно в условиях достаточно мощного узкополосного шума или музыкального фона. Применение спектрального анализа для детектирования свойств сигнала также не является эффективным, поскольку получение спектра Фурье связано с использованием всего интервала анализируемого сигнала, а кратковременный спектр Фурье имеет либо плохое разрешение по частоте, либо занимает большой временной интервал [5, 6]. Метод Фурье недостаточен также в связи с тем, что после фильтрации восстановление обратным преобразованием может привести к комплексной форме сигнала.

наиболее выгодное разбиение частотного диапазона [18–23]. Полученное разложение РС в соответствии с перцептуальной моделью обеспечено специальным выбором двоичного дерева, определяющего структуру вейвлет–преобразования.

Таблица 1

Частотные диапазоны перцептуальной модели

| Индекс диапазона | Граничные частоты (Гц) | Средняя частота (Гц) | Ширина диапазона (Гц) |
|------------------|------------------------|----------------------|-----------------------|
| 1                | 0–100                  | 50                   | 100                   |
| 2                | 100–200                | 150                  | 100                   |
| 3                | 200–300                | 250                  | 100                   |
| 4                | 300–400                | 350                  | 100                   |
| 5                | 400–510                | 455                  | 110                   |
| 6                | 510–630                | 570                  | 120                   |
| 7                | 630–770                | 700                  | 140                   |
| 8                | 770–920                | 845                  | 150                   |
| 9                | 920–1080               | 1000                 | 160                   |
| 10               | 1080–1270              | 1175                 | 190                   |
| 11               | 1270–1480              | 1355                 | 210                   |
| 12               | 1480–1720              | 1600                 | 240                   |
| 13               | 1720–2000              | 1860                 | 280                   |
| 14               | 2000–2320              | 2160                 | 320                   |
| 15               | 2320–2700              | 2510                 | 380                   |
| 16               | 2700–3150              | 2925                 | 450                   |
| 17               | 3150–3700              | 3425                 | 550                   |
| 18               | 3700–4400              | 4050                 | 700                   |
| 19               | 4400–5300              | 4850                 | 900                   |
| 20               | 5300–6400              | 5850                 | 1100                  |
| 21               | 6400–7700              | 7050                 | 1300                  |
| 22               | 7700–9500              | 8750                 | 1800                  |
| 23               | 9500–12000             | 10750                | 2500                  |
| 24               | 12000–15500            | 13750                | 3500                  |
| 25               | 15500–23500            | 1950                 | 8000                  |

Критерием качества аппроксимации перцептуальной модели служит показатель среднеквадратичной относительной погрешности соответствия границ критических полос и границ субполос вейвлет–пакета [7]. Этот показатель (Q) представлен соотношением (1):

$$Q = \sqrt{\frac{1}{p} \sum_{i=1}^p \frac{1}{2} \left\{ \left( \frac{\hat{f}_i^1 - f_i^1}{\bar{f}_i} \right)^2 + \left( \frac{\hat{f}_i^2 - f_i^2}{\bar{f}_i} \right)^2 \right\}}, \quad (1)$$

где  $f_i^1, f_i^2$  – нижняя и верхняя границы частотной полосы в перцептуальной модели,

$\bar{f}_i$  – положение центра полосы;

$\hat{f}_i^1, \hat{f}_i^2$  – то же для субполос сформированного вейвлет–пакета;

p – количество частотных диапазонов модели, p = 22. Несмотря на то что бинарная структура формирования дерева дает ограниченное количество вариантов, полученное разложение удовлетворительно аппроксимирует частотные диапазоны перцептуальной модели. При первоначальной

оценке показателя Q=0.08, небольшое укрупнение трех полос в частотных диапазонах перцептуальной модели (табл.2) улучшил общее соответствие представления речевого сигнала вейвлет–пакетом и при этом показатель качества Q снижен до 0.03.

Таблица 2

Сопоставление частотных субполос вейвлет–представления РС с перцептуальной моделью

| Верхние границы субполос и индексы частотных диапазонов вейвлет–пакета и перцептуальной модели (Гц) |  |                              |                             |
|---|--|------------------------------|-----------------------------|
| Вейвлет–представление пакетом   |  | Перцептуальная модель        |                             |
| Верхние границы субполос вейвлет–пакета   | Индексы вершин дерева вейвлет–пакета <sup>1)</sup> | Верхние границы полос модели | Индексы полос <sup>2)</sup> |
| 250   | 1  | 300                          | 1–3                         |
| 375   | 2  | 400                          | 4                           |
| 500   | 3  | 510                          | 5                           |
| 625   | 4  | 630                          | 6                           |
| 750   | 5  | 770                          | 7                           |
| 1000  | 6  | 1080                         | 8–9                         |
| 1250  | 7  | 1270                         | 10                          |
| 1750  | 8–9  | 1720                         | 11–12                       |
| 2000  | 10   | 2000                         | 13                          |
| 2250  | 11   | 2320                         | 14                          |
| 2750  | 12–13  | 2700                         | 15                          |
| 3000  | 14   | 3150                         | 16                          |
| 3500  | 15–16  | 3700                         | 17                          |
| 4500  | 17–18  | 4400                         | 18                          |
| 5500  | 19   | 5300                         | 19                          |
| 6500  | 20   | 6400                         | 20                          |
| 7500  | 21   | 7700                         | 21                          |
| 9750  | 22   | 9500                         | 22                          |
| 12025   | 23   | 12000                        | 23                          |

<sup>1)</sup> Индексы 8–9, 12–13, 15–16 и 17–18 указывают на объединение субполос пакета на вершинах дерева разложения сигнала. Верхняя граница субполосы в первом столбце соответствует второму индексу из указанных пар.

<sup>2)</sup> Индексы 1–3, 8–9, 11–12 соответствуют сгруппированным полосам перцептуальной модели. Верхние границы полос соответствуют вторым индексам из указанных пар.

Само вейвлет–преобразование не дает требуемого уменьшения размерности (полное число коэффициентов разложения равно числу отсчетов сигнала), поэтому для нейросетевого алгоритма использован вектор с компонентами субполосной мощности сигнала  $\mathbf{P} = (\sigma_1^2, \sigma_2^2 \dots \sigma_p^2)$ , где  $\sigma_i^2$  – сумма квадратов коэффициентов, принадлежащих i–й полосе разложения.

Таким образом, с использованием вейвлет–пакетов для каждого сегмента РС получены векторы мощности вейвлет–коэффициентов в субполосах, соответствующих перцептуальной модели. Эти векторы отражают спектральные свойства сигнала приблизительно с теми же градациями в частотной области, что и слуховой анализатор человека.

Нормирование векторов ( $\mathbf{p} = \mathbf{P} / \|\mathbf{P}\|$ )

позволяет устранить влияние общей мощности сигнала на возможные решения о его свойствах и тем самым устранить основной недостаток алгоритмов детектирования РС, ориентированных в значительной степени на силу сигнала.

**Применение перцептрона для обнаружения свойств нестационарных сигналов.** Для принятия решения о наличии или отсутствии речевого сигнала на анализируемом сегменте предложено применение нейронной сети [7],[8]. Входным вектором сети является нормированный вектор мощности вейвлет-коэффициентов. Такое решение имеет следующие преимущества:

- система, использующая нейронную сеть, является самонастраивающейся адаптивной системой, для которой нет необходимости предварительно выполнять точный расчет параметров;
- система имеет этап обучения, позволяющий выполнить адаптацию параметров для настройки на конкретные условия сигнала и шума;
- в области вейвлет-разложения не обязательно иметь линейную функцию для разделения вейвлет-образов

сигнала и шума, иными словами свойство линейной разделимости образов не является обязательным;

- нелинейная активационная функция нейрона позволяет принимать нежесткие решения в каждом нейроне сети;
- структура нейронной сети может быть адаптирована к задаче разделения сегментов сигнала и шума.

Нейронной сетью достаточно простой структуры является перцептрон [24–27]. В простейшем случае нейроперцептрон состоит из одного нейрона (рис.2), у которого суммирующий узел дает линейную комбинацию входов  $x_i$ , приложенных к синапсам  $w_i$ , и учитывает внешний порог  $\theta$ . К выходу суммирующего узла  $v$ ,

$$v = \sum_{i=1}^p w_i x_i - \theta, \text{ применяется нелинейная функция}$$

$y = \varphi(v)$ , называемая функцией активации.

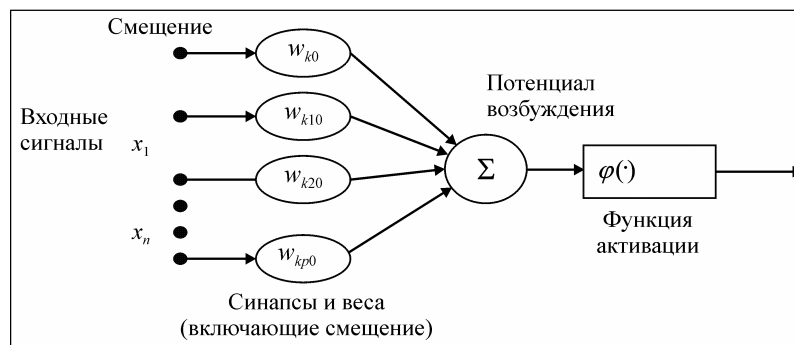


Рис.2. Структурная схема элементарного нейронного перцептрона

Нейроперцептрон с одним слоем не имеет скрытых нейронов и не может классифицировать входные векторы, не являющиеся линейно разделимыми. Эта проблема решается введением скрытых слоев в многослойном перцептроне (МСП), который состоит из входных узлов (составляющих входной слой), одного или более скрытых слоев и выходного слоя вычисляющих узлов. Входной сигнал распространяется через сеть слой за слоем. Каждый нейрон сети содержит нелинейную функцию активации в форме логистической кривой:

$$y_j = \varphi(v_j) = [1 + \exp(-v_j)]^{-1}.$$

Нейроперцептрон выполняет отображение входных  $p$ -мерных векторов

$\mathbf{P} = (\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2)$  на выходное  $q$ -мерное пространство, в задаче обнаружения РС  $q = 1$ .

Согласно Сайбенко–Фунахаши [28],[29], нейроперцептрон способен с любой точностью аппроксимировать желаемое отображение вход–выход, – обучаться сопоставлять множеству входных  $p$ -векторов заданное множество выходных  $q$ -векторов. Так что, существует  $\varepsilon$ -аппроксимация  $F$  любого непрерывного отображения вход–выход  $f$ , которую можно представить в математической форме (2), и эта форма эквивалентна двухслойному перцептрону (3) с одним выходным нейроном и одним скрытым слоем из  $M$  нейронов ( $M$  зависит от величины  $\varepsilon$ ):

$$|F(x_1, x_2, \dots, x_p) - f(x_1, x_2, \dots, x_p)| \leq \varepsilon, \quad (2)$$

$$F(x_1, x_2, \dots, x_p) = \sum_{j=0}^M a_j \varphi \left( \sum_{i=1}^p w_{ji} x_i - \theta_j \right) \quad (3)$$

Здесь  $a_j$  – коэффициенты аппроксимации отображения  $f$  обобщенным рядом Фурье, которые в перцептроне представлены синаптическими связями (весами) выходного нейрона;  $\{w_{ji}\}_{j=1..M}$  – синаптические веса нейронной сети от  $i$ -го к  $j$ -му нейрону, структурно связывающие входные узлы с нейронами скрытого слоя;  $\theta_j$  – потенциал смещения (порог чувствительности)  $j$ -го нейрона;  $\varphi$  – функция преобразования потенциала возбуждения нейрона

$$v_j = \left( \sum_{i=1}^p w_{ji} x_i - \theta_j \right) \text{ в его выходное значение } y_j.$$

Изложенное подтверждает правильность выбора типа нейронной сети в целях разработки метода для детектирования изменения свойств нестационарного сигнала.

**Заключение.** Предложенный метод и модель анализа речевого сигнала (РС) дают основу для решения ряда задач детектирования изменения свойств нестационарного сигнала:

1. Для анализа нестационарного сигнала (прототипом которого служил РС) целесообразно использование нейросетевых алгоритмов на вейвлет-отображении сигнала.

2. Вейвлет-разложение может быть адаптировано к частотной структуре сигнала на основе вейвлет-пакета. Для РС выбор вида пакета определен в соответствии с перцептуальной моделью восприятия звука, которая построена на базе обобщения биофизических исследований процессов фильтрации органа слуха человека.



3. Для реализации нейросетевого алгоритма необходимо решение ряда вопросов:

- Обоснование метода обучения, имеющего лучшую скорость сходимости для векторов аудио-сигнала и позволяющего получить меньшие по величине ошибки распознавания свойств сигнала.
  - Получение ограничения на требуемое время обучения персептрона.
  - Обоснование оптимальной структуры персептрона, которая не является избыточной и позволяет эффективно решать задачи распознавания.
- Эти вопросы отражены во втором разделе этой статьи.

### Исследование и оптимизация нейронной сети

**Характеристики методов обучения нейроперсептрона.** Для алгоритма обучения нейроперсептрона задачи детектирования речевой активности важно иметь быструю сходимость и малые вероятности ошибочных решений на этапе распознавания. Для этой цели применены различные алгоритмы, используемые при обучении нейронных сетей, и исследованы их характеристики для задачи распознавания типа сегментов речь/пауза. Алгоритмы построены на принципе минимизации суммарной ошибки распознавания и основаны на методах нелинейной оптимизации. Проанализированы метод сопряженных градиентов; простое и обобщенное правила дельта-дельта; метод случайного поиска; метод Бройдена – Флетчера и метод параллельных касательных (метод партан).

При обучении нейросети с настройкой ее параметров, – синаптических весов (СВ), применяется «алгоритм с обратным распространением ошибки» (алгоритм ОРО) [29 – 35]. Алгоритм использует понятия:

- ошибки  $e_j(n)$  на выходе нейрона  $j$  на итерации  $n$ :

$$e_j(n) = d_j(n) - y_j(n),$$

где  $d_j$  – желаемый отклик,  $y_j$  – выход нейрона  $j$ , являющегося выходным узлом сети;

- текущего значения суммы квадратов ошибок:

$$E(n) = 0.5 \sum_{j \in C} e_j^2(n),$$

где  $C$  – множество выходных нейронов сети.

Цель обучения состоит в подстройке СВ сети, обеспечивающей минимум средней ошибки:

$$E_{cp} = (N)^{-1} \sum_{n=1}^N E(n),$$

где  $N$  – количество обучающих векторов.

Алгоритм ОРО реализует обучение итеративно с двумя проходами по сети – прямым и обратным, при которых соответственно вычисляются функциональные преобразования вектора входного сигнала  $\mathbf{x} = (x_1, \dots, x_p)^T$  и корректировка вектора параметров  $\mathbf{w}$ . При обратном проходе на каждой итерации  $n$  определяется локальный градиент  $\delta_j(n)$ , для чего служит выражение (4), если нейрон находится в выходном слое, и выражение (5), если нейрон в скрытом слое:

$$\delta_j(n) = e_j(n) \varphi'(v_j(n)) \quad (4)$$

$$\delta_j(n) = \varphi'(v_j(n)) \sum_k \delta_k(n) w_{jk}(n), \quad (5)$$

где  $v_j(n)$  – возбуждение  $j$ -го нейрона,  $y_j(n)$  – выход  $j$ -го ней-

рона,  $e_j(n)$  – ошибка на выходе  $j$ -го нейрона,

$$\varphi'(v_j(n)) = y_j(n) [1 - y_j(n)] -$$

производная функции нелинейности,  $w_{jk}$  – синаптические веса. Локальный градиент ошибки  $\delta_j(n)$  используется для корректировки весов в соответствии с  $\delta$ -правилом по (6) или в соответствии с модифицированным  $\delta$ -правилом по выражению (7):

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (6)$$

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n), \quad (7)$$

где  $\alpha$  и  $\eta$  – параметры, определяющие скорость обучения.

Непосредственное применение метода обратного распространения при реализации алгоритма обучения предполагает задание свободных параметров  $\alpha$  и  $\eta$ .

Более продуктивный подход к обучению основан на нелинейной минимизации функции ошибок, или оптимизации нелинейной функции отображения вход-выход [36–39].

**Метод сопряженных градиентов.** В методе сопряженных градиентов [40] коррекция вектора весов персептрона  $\mathbf{w}(n)$  осуществляется не вдоль вектора градиента ошибки  $\mathbf{g}(n)$ , а вдоль вектора сопряженного градиента  $\mathbf{p}(n)$ , который обновляется на каждом шаге итерации:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta(n) \mathbf{p}(n); \quad (8)$$

$$\mathbf{p}(n+1) = -\mathbf{g}(n+1) + \beta(n) \mathbf{p}(n),$$

где  $\mathbf{g}(n)$  – вектор градиента функции ошибки  $E(\mathbf{w}_n)$ , а параметр  $\beta(n)$  определяется по (9) для метода Флетчера-Ривса и по (10) – для метода Полака – Рибера [37]:

$$\beta(n) = \frac{\mathbf{g}^T(n+1) \mathbf{g}(n+1)}{\mathbf{g}^T(n) \mathbf{g}(n)} \quad (9)$$

$$\beta(n) = \frac{\mathbf{g}^T(n+1) [\mathbf{g}(n+1) - \mathbf{g}(n)]}{\mathbf{g}^T(n) \mathbf{g}(n)} \quad (10)$$

Оценка параметра  $\eta(n)$  включает линейный поиск, состоящий в определении  $\eta$ , для которого критерий  $E_{cp}[\mathbf{w}(n) + \eta \mathbf{p}(n)]$  имеет минимальную величину при данных значениях  $\mathbf{w}(n)$  и  $\mathbf{p}(n)$ :

$$\eta(n) = \arg \left( \min_{\eta} \{E_{cp}[\mathbf{w}(n) + \eta \mathbf{p}(n)]\} \right) \quad (11)$$

Таким образом, при непосредственном использовании алгоритма ОРО параметр скорости обучения  $\eta$  определяется на основе оптимизации критерия минимума функции ошибок.

### Метод дельта – дельта и обобщенное правило

**дельта – дельта.** В качестве более совершенного метода выбора  $\eta(n)$  применена форма адаптации параметра скорости обучения на основе правил Джекобса: правила дельта-дельта адаптации и обобщенного правила дельта-дельта [41]. Оба правила основаны на принципе быстрого спуска по поверхности функции ошибок, в качестве которой служит суммарная квадратичная ошибка нейроперсептрона  $E(n)$  на текущем шаге процесса обучения (12), которая в этом методе рассматривается в пространстве  $\{\eta_i(n)\}_{i=1,2,\dots}$  параметров скорости обучения для синаптических весов  $w_{ji}(n)$  нейронной сети:

$$E(n) = 0.5 \sum_j e_j^2(n) = 0.5 \sum_j [d_j(n) - y_j(n)]^2 \quad (12)$$

Определение градиента

$$\frac{\partial E}{\partial \eta_i(n)}$$

основано на правиле (13) дифференцирования сложной функции и общих соотношениях (14), (15):

$$\frac{\partial E}{\partial \eta_i(n)} = \frac{\partial E(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial \eta_i(n)} \quad (13)$$

$$v_j(n) = \sum_i w_{ji}(n) y_i(n) \quad (14)$$

$$w_{ji}(n) = w_{ji}(n-1) - \eta_i(n) \frac{\partial E(n-1)}{\partial w_{ji}(n-1)} \quad (15)$$

Соотношение (14) выражает возбуждение  $j$ -нейрона выходами  $y_i$  нейронов предшествующего слоя, а (15) – правило градиента для изменения параметров персептрона  $w_{ji}(n)$ . Соотношения (14) и (15) позволяют выразить  $v_j(n)$  в виде:

$$v_j(n) = \sum y_i(n) [w_{ji}(n-1) - \eta_i(n)] \quad (16)$$

Это приводит к определению компонент (17), (18) и (19), необходимых для получения формы выражения (13), пригодной для вычислительного алгоритма:

$$\frac{\partial v_j(n)}{\partial \eta_i(n)} = -y_i(n) \frac{\partial E(n-1)}{\partial w_{ji}(n-1)} \quad (17)$$

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \phi'_j(v_j(n)) \quad (18)$$

$$\frac{\partial E(n)}{\partial y_j(n)} = -[d_j(n) - y_j(n)] = -e_j(n) \quad (19)$$

Полученная таким образом окончательная форма выражения

$$\frac{\partial E}{\partial \eta_i(n)} \text{ имеет вид:}$$

$$\frac{\partial E(n)}{\partial \eta_i(n)} = -\phi'_j(v_j(n-1)) e_j(n) y_i(n) \left[ -\frac{\partial E(n-1)}{\partial w_{ji}(n-1)} \right] \quad (20)$$

Если учесть, что для градиента функции ошибки относительно весов справедливо (21):

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -\phi'_j(v_j(n)) e_j(n) y_i(n), \quad (21)$$

то получается конечная форма градиента в пространстве параметров скорости обучения:

$$\frac{\partial E}{\partial \eta_i(n)} = -\frac{\partial E(n)}{\partial w_{ji}(n)} \cdot \frac{\partial E(n-1)}{\partial w_{ji}(n-1)} \quad (22)$$

Таким образом, значение параметра скорости обучения  $\eta$ , обеспечивающее наискорейший спуск к минимуму поверхности ошибок в пространстве параметров  $\{\eta_i(n)\}$ , адаптивно обновляется по алгоритму в форме соотношения:

$$\Delta \eta_i(n+1) = -\gamma \frac{\partial E(n)}{\partial \eta_i(n)} = \gamma \frac{\partial E(n)}{\partial w_{ji}(n)} \frac{\partial E(n-1)}{\partial w_{ji}(n-1)}, \quad (23)$$

где  $\gamma$  – константа, контролирующая шаг адаптации параметра скорости обучения  $\eta$ .

$$\text{Значения } \frac{\partial E(n)}{\partial w_{ji}(n)} \text{ и } \frac{\partial E(n-1)}{\partial w_{ji}(n-1)}$$

(компоненты градиента функции ошибок на шаге  $n$  и  $(n-1)$  в итеративной процедуре обучения) служат основой следующих правил, сформулированных Джекобсом:

– Если компонента градиента

$$\frac{\partial E(n)}{\partial w_{ji}(n)}$$

имеет одинаковый знак на двух итерациях, то  $\Delta \eta_i(n+1) > 0$ , т.е. адаптация увеличивает скорость обучения для веса  $w_{ji}$ .

– Если у градиента меняется знак на соседних итерациях,  $\Delta \eta_i(n+1) < 0$ , то процедура адаптации уменьшает параметр скорости обучения для веса  $w_{ji}$ .

Приведенная форма алгоритма на основе правила дельта–дельта проста в вычислениях, но может дать неуверенную адаптацию при очень малых и больших значениях градиента. Лучшие результаты дает более сложная модификация алгоритма – обобщенное правило дельта–дельта [41], которое требует большего объема вычислений. Адаптация параметра  $\eta_i(n)$  для синаптической связи  $w_{ji}$  определяется выражениями (24) – (26):

$$\Delta \eta_{ji}(n+1) = \begin{cases} k & \text{при } S_{ji}(n-1)D_{ji}(n) > 0, \\ -\beta \eta_{ji}(n) & \text{при } S_{ji}(n-1)D_{ji}(n) < 0, \\ 0 & \text{иначе,} \end{cases} \quad (24)$$

где  $D_{ji}(n)$  и  $S_{ji}(n)$  определяются по (25) и (26):

$$D_{ji}(n) = \frac{\partial E(n)}{\partial w_{ji}} \quad (25)$$

$$S_{ji}(n) = (1 - \xi)D_{ji}(n-1) + \xi \cdot S_{ji}(n-1) \quad (26)$$

В выражениях (25) и (26)  $\xi$  – положительная константа;  $D_{ji}(n)$  – текущее значение градиента;  $S_{ji}(n)$  – определяется соотношением типа авторегрессии, которое использует параметр веса  $\xi$  и компоненту градиента  $D_{ji}(n-1)$  в качестве базы. Процедура адаптации параметра скорости обучения (24)–(26) составляет основу метода обучения по обобщенному правилу дельта–дельта. Адаптация основана на контроле совпадения или различия знаков на последовательных итерациях у соответствующих компонент  $S_{ji}(n-1)$  и градиента  $D_{ji}(n)$ . При совпадении знаков  $\Delta \eta_{ji}(n)$  составляет  $k$ , а при несовпадении уменьшается экспоненциально:  $\Delta \eta_{ji}(n+1) = -\beta \eta_{ji}(n)$ .

**Метод случайного поиска.** Метод случайного поиска основан на случайном выборе вектора приращения весов по принципу Монте-Карло. Выбор вектора приращения весов  $\Delta w_{ji}$  реализуется поиском в пространстве весов, при этом принимаются только те изменения  $\Delta w_{ji}$ , которые ведут к уменьшению квадратичной ошибки на каждой итерации. Метод достаточно прост по вычислениям, но недостаточно эффективен и медленно сходится.

**Метод Бroyдена – Флетчера.** Метод нелинейной минимизации Бройдена – Флетчера основан на методе Ньютона [36], в котором используется разложение функции ошибок, включающее вторые производные

$$\frac{\partial^2 E(\mathbf{w})}{\partial w_i \partial w_j}, \text{ т.е. матрицу Гессе } \mathbf{H} = \left[ H_{ij} = \frac{\partial^2 E(\mathbf{w})}{\partial w_i \partial w_j} \right]:$$

$$\begin{aligned} E(\mathbf{w}(n) + \Delta \mathbf{w}(n)) &\cong \\ &\cong E(n) + \mathbf{g}^T(n) \cdot \Delta \mathbf{w}(n) + \frac{1}{2} \Delta \mathbf{w}^T(n) \cdot \mathbf{H}^T, \end{aligned} \quad (27)$$

$$\begin{aligned} \text{где } E(n) &\equiv E(\mathbf{w}(n)); \quad \{\mathbf{g}(n)\}_i \equiv \frac{\partial E(n)}{\partial w_i(n)}; \\ \{H(n)\}_{ij} &= \frac{\partial^2 E(\mathbf{w}(n))}{\partial w_i(n) \partial w_j(n)}; \quad \mathbf{g}(n) - \text{градиент ошибки.} \end{aligned}$$

Изменение весов  $\Delta \mathbf{w}(n)$  вектора параметров перцептрона  $\mathbf{w}(n)$  на  $n$ -й итерации определяется условием

$$\Delta \mathbf{w}(n) = \arg \min_{\Delta \mathbf{w}} \{E(\mathbf{w}(n) + \Delta \mathbf{w}(n))\} \quad (28)$$

Соотношение (28) с учетом (27) эквивалентно выражению (29), которое лежит в основе получения  $\Delta \mathbf{w}(n)$  по методу Ньютона:

$$\mathbf{H}(n) \cdot \Delta \mathbf{w}(n) = -\mathbf{g}(n) \quad (29)$$

При решении матричного уравнения (29) относительно  $\Delta \mathbf{w}(n)$  необходимо, чтобы матрица Гессе была положительно определенной и хорошо обусловленной. Но эти условия не всегда выполняются на практике.

Модификация Бройдена – Флетчера для метода Ньютона связана с построением положительно определенной матрицы  $\tilde{\mathbf{H}}(n)$ , которая является аппроксимацией матрицы Гессе на основе конечных разностей градиента функции ошибок [42–46]. Матрицей  $\tilde{\mathbf{H}}(n)$  представляется разложение (30) градиента рядом Тэйлора, которое в более короткой форме эквивалентно (31):

$$\mathbf{g}(\mathbf{w}(n) + \Delta \mathbf{w}(n)) \cong \mathbf{g}(\mathbf{w}(n)) + \tilde{\mathbf{H}}(n) \cdot \Delta \mathbf{w}(n) \quad (30)$$

$$\tilde{\mathbf{H}}(n) \cdot \Delta \mathbf{w}(n) = \mathbf{y}(n) \quad (31)$$

где через  $\mathbf{y}(n) \equiv \mathbf{g}(n+1) - \mathbf{g}(n)$  обозначено приращение градиента на  $n$ -й итерации.

В методе Бройдена-Флетчера аппроксимация матрицы Гессе в алгоритме обучения вычисляется итеративно с помощью соотношения (32):

$$\begin{aligned} \tilde{\mathbf{H}}(n+1) &= \tilde{\mathbf{H}}(n) - \\ &- \frac{1}{\Delta \mathbf{w}^T(n) \cdot \tilde{\mathbf{H}}(n) \cdot \Delta \mathbf{w}(n)} \tilde{\mathbf{H}}(n) \cdot \Delta \mathbf{w}(n) \cdot \Delta \mathbf{w}^T(n) + \\ &+ \frac{1}{\mathbf{y}^T(n) \cdot \Delta \mathbf{w}(n)} \mathbf{y}(n) \cdot \mathbf{y}^T(n) \end{aligned} \quad (32)$$

При этом вектор  $\mathbf{p}(n)$  направления изменения параметров перцептрона  $\Delta \mathbf{w}(n)$  определяется соотношением (33), которое по форме аналогично выражению характерному для метода Ньютона [36]:

$$\tilde{\mathbf{H}}(n) \cdot \mathbf{p}(n) = -\mathbf{g}(n) \quad (33)$$

Изменение весов перцептрона  $\Delta \mathbf{w}(n)$  вдоль этого направления с параметром скорости обучения  $\eta$  определяется соотношением (34):

$$\Delta \mathbf{w}(n) = \eta \cdot \mathbf{p}(n) \quad (34)$$

Таким образом, выражение (32) преобразуется к ви-

ду (35):

$$\begin{aligned} \tilde{\mathbf{H}}(n+1) &= \tilde{\mathbf{H}}(n) + \frac{1}{\mathbf{g}^T(n) \cdot \mathbf{p}(n)} \mathbf{g}(n) \cdot \mathbf{g}^T(n) + \\ &+ \frac{1}{\eta(n) \mathbf{y}(n) \cdot \mathbf{p}(n)} \mathbf{y}(n) \cdot \mathbf{y}^T(n) \end{aligned} \quad (35)$$

Выражения (33) – (35) служат основой алгоритма обучения перцептрона по методу Бройдена – Флетчера. Алгоритм включает оценку градиента  $\mathbf{g}(n)$  функции ошибок  $E(\mathbf{w}(n))$ , определение вектора направления  $\mathbf{p}(n)$  и схему обновления матрицы  $\tilde{\mathbf{H}}(n)$  [36].

Метод Бройдена – Флетчера использует не только первую, но и вторую производную функции ошибок (в форме ее разностной аппроксимации) и более точно определяет направление при поиске минимума.

#### **Метод параллельных касательных (метод partan).**

Метод реализует движение в направлении, которое параллельно касательной к линии равного уровня поверхности ошибок, найденной градиентным методом для предшествующей точки [36, 47]. Метод использует матрицу Гессе (36) и алгоритм сопряженных направлений  $\mathbf{S}_n$  и  $\mathbf{S}_{n+1}$  по соотношению (37):

$$\begin{aligned} \mathbf{H}[E(\mathbf{w}_n)] &= \nabla^2 \{E[\mathbf{w}(n)]\}, \\ \text{т.е. } H_{ij} &= \frac{\partial^2 E(\mathbf{w}_n)}{\partial w_i \partial w_j} \end{aligned} \quad (36)$$

$$(\mathbf{S}_{n+1})^T \cdot [\nabla^2 \{E[\mathbf{w}(n)]\}] \mathbf{S}_n = 0 \quad (37)$$

Условие (37) – это условие ортогональности вектора направления  $\mathbf{S}_{n+1}$  на  $(n+1)$  шаге относительно вектора  $[\nabla^2 \{E[\mathbf{w}(n)]\}] \mathbf{S}_n$ , где  $\nabla^2 \{E[\mathbf{w}(n)]\}$  – матрица Гессе (36).

Корректировка вектора весов нейроперцептрона выполняется по соотношению:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \lambda_n \mathbf{S}_n \quad (38)$$

Значение параметра  $\lambda$ , соответствующее минимуму  $E(\mathbf{w}_{n+1})$  в направлении вектора  $\mathbf{S}_n$ , определяется выражением

$$\lambda_n = - \frac{[\nabla E(\mathbf{w}_n)]^T \cdot \mathbf{S}_n}{\mathbf{S}_n^T \cdot [\nabla^2 \{E(\mathbf{w}_n)\}] \cdot \mathbf{S}_n}, \quad (39)$$

где  $\nabla E(\mathbf{w}_n)$  – вектор градиента

$$[\nabla E(\mathbf{w}(n))]_i = [\text{grad} \{E(\mathbf{w}_n)\}]_i \equiv \frac{\partial E(\mathbf{w}(n))}{\partial w_i(n)}.$$

Таким образом, метод параллельных касательных требует вычисления гиперплоскости, касательной к функции ошибок, на каждой итерации. Поэтому метод связан с большими вычислительными затратами, особенно в задачах большой размерности. В двухмерном варианте метод основан на движении по направлению, параллельному касательной к линии постоянного уровня функции ошибок, однако при большой размерности аргумента (СВ перцептрона) используются касательные гиперплоскости и движение в направлении, параллельном этим гиперплоскостям. Требование вычисления касательных гиперплоскостей и определения  $\min_{\mathbf{w}_n} \{E(\mathbf{w}_n)\}$  вдоль направления, параллельного этим гиперплоскостям, является недостатком

метода. При значительной размерности СВ (в функции ошибок персептрона) метод постепенно теряет эффективность, снижая точность и увеличивая длительность обучения.

**Выбор метода обучения нейроперсептрона для решаемой задачи.** При использовании любого алгоритма обучения важна инициализация СВ и порогов нейронов сети. В качестве начальных значений весов и уровней порога нейронов для создания им равноценных условий подстройки в процессе обучения целесообразно выбирать равномерно распределенные числа. Значения параметров синаптической связи для нейрона  $i$  в сети рекомендуется обычно ограничивать интервалом  $[-2.4(F_i)^{-1}, 2.4(F_i)^{-1}]$ , где  $F_i$  – общее количество входов  $i$ -го нейрона [48].

Обобщая представленный выше анализ, можно утверждать, что в алгоритмах обучения персептрона для задачи детектирования речевой активности может быть использован один из рассмотренных методов, хотя наиболее совершенными представляются два из них:

- обобщенный метод дельта–дельта (правило Джексона), который включает адаптивный выбор параметра скорости обучения;
- метод Бройдена–Флетчера, который использует приближение для производных второго порядка (матрицу Гессе) и позволяет определять величину параметра скорости обучения.

Однако только на основе теоретического анализа нельзя сделать окончательный выбор метода обучения персептрона задаче распознавания речевого сигнала, поскольку невозможно учесть поведение алгоритмов в условиях многообразия характеристик сигнала и шума, величин пауз и наличия переходных сегментов. Исходя из этого, для обоснованного выбора метода обучения, на достаточных данных проведена экспериментальная проверка алгоритмов при различных видах шумов и различных отношениях сигнал/шум.

**Экспериментальное исследование характеристик алгоритмов обучения нейроперсептрона.** В целях объективной оценки преимуществ алгоритмов обучения для детектирования речевой активности проведен компьютерный эксперимент, в котором на речевой сигнал, включающий интервалы активности и микропаузы, наложены различные виды шума:

- белый шум, имеющий равномерный спектр мощности в звуковом диапазоне частот;
- окрашенный шум с постепенным спадом спектра при росте частоты (для розового шума – 6 дБ/октаву), представителем которого может служить шум пылесоса или шум льющейся воды;
- музыкальный фон;
- узкополосный шум, состоящий из детерминированных или случайных компонент, постоянных по своему положению и амплитуде (шум вентилятора).

Эксперимент проведен с использованием нейросетевого эмулятора Essence 1.0 [49], выполненного на языке Java и позволяющего программно моделировать многослойный нейроперсептрон с любым количеством слоев, устанавливать отдельно количество и вид активационной функции нейронов для каждого слоя, выполнять все перечисленные

алгоритмы обучения и использовать различные виды ошибок.

Данными служили оцифрованные записи речевого сигнала и сопутствующего шума с частотой дискретизации 22.5 кГц. На каждом сегменте записи длительностью 20 мс выполнено вейвлет–разложение в соответствии с персептуальной моделью, вычислена средняя мощность вейвлет–коэффициентов в каждой области разложения и полученные векторы мощности вейвлет–коэффициентов нормированы.

На основе анализируемого зашумленного речевого сигнала получено 2000 векторов, которые составили два множества (обучающее и тестовое) по 1000 векторов. Использован супервизорный принцип обучения (с учителем), при котором сети сообщается желаемый отклик на каждый входной обучающий вектор. Программа представляет динамику изменения ошибки в процессе обучения и на тестирующей выборке. При сходимости обучения ошибка почти не изменяется, но если она недостаточно мала, то сеть усложнялась и повторялось ее обучение и тестирование. Напротив, если ошибка по обучающей выборке стремится к нулю, а при тестировании сохраняет недопустимо большое значение, то размер сети уменьшался.

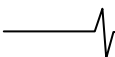
После ряда экспериментов с изменением размера сети для исследования различных алгоритмов обучения принята двухслойная сеть с числом нейронов десять и один (10-1) и с логистической активационной функцией нейронов. Для всех алгоритмов при равном времени обучения (2 мин) зафиксировано число циклов обучения и процент ошибочных решений по обучающей и тестовой выборкам. Эксперимент проведен для всех видов шума и отношениях среднеквадратического отклонения шума к сигналу  $\zeta = 0.02, 0.1$  и  $0.3$ .

Результаты компьютерного эксперимента, позволившие сопоставить различные алгоритмы обучения, показывает таблица 3, где приведено число циклов обучения и процент ошибочных решений при обучении и тестировании для нескольких уровней сложности нейронной сети, типов и уровней сопутствующего шума. Число циклов различается в зависимости от алгоритма и определяется его вычислительной сложностью. Для всех методов обучения процент ошибочных решений на обучающей выборке меньше, чем на тестовой выборке. Однако это отличие не превышает 20 %, что подтверждает адекватность выбранной модели нейронной сети.

При сопоставлении алгоритмов по оценке вероятности ошибочных решений наилучшие характеристики получены для метода Бройдена–Флетчера. Вероятность ошибочных решений зависит от интенсивности шума: для  $\zeta = 0.02$  вероятность ошибочных решений по тестовой выборке не превышает 3.7 %, для  $\zeta = 0.1$  – 4.3 %, а для  $\zeta = 0.3$  – 5.5 %. Вероятность ошибочных решений для узкополосного шума и музыкального фона выше, поскольку спектральный состав шума меньше отличается от речевого сигнала.

Различаются ошибочные решения первого и второго рода [50, 51]. Пусть  $\omega_1$  – событие, состоящее в том, что текущий сегмент или фрейм содержит речевой сигнал;  $\omega_2$  – событие, состоящее в том, что текущий фрейм содержит шум;  $\Omega_1$  – решение о том, что текущий фрейм содержит речевой сигнал;  $\Omega_2$  – решение о том, что текущий фрейм содержит шум. Тогда вероятность ошибочного обнаруже-





ния речевого сигнала во время паузы определяется соотношением:

$$P(\Omega_1 / \omega_2) = \frac{P(\Omega_1, \omega_2)}{P(\omega_2)} \quad (40)$$

Вероятность ошибочного обнаружения паузы при наличии речевого сигнала – соотношением:

$$P(\Omega_2 / \omega_1) = \frac{P(\Omega_2, \omega_1)}{P(\omega_1)} \quad (41)$$

Компьютерный эксперимент показал, что вероятность ошибочного обнаружения паузы при наличии речевого сигнала ниже, чем вероятность ошибочного обнаружения речевого сигнала.

Таблица 3

Сравнение методов обучения персептрона

| Вид шума        | Относительный вклад шума | Метод обучения       | Обучение     |            |          | Тестирование |
|-----------------|--------------------------|----------------------|--------------|------------|----------|--------------|
|                 |                          |                      | Число циклов | Время, мин | Ошибка % | Ошибка %     |
| 1               | 2                        | 3                    | 4            | 5          | 6        | 7            |
| Белый шум       | 0.02                     | Случайный поиск      | 500          | 5          | 10       | 13           |
|                 |                          | Сопряжен. градиенты  | 800          | 3.2        | 3.5      | 5.7          |
|                 |                          | Partan               | 600          | 3.3        | 4        | 4.7          |
|                 |                          | Дельта-дельта        | 600          | 3.3        | 4.5      | 5.2          |
|                 |                          | Обобщ. дельта-дельта | 650          | 4          | 2.5      | 3.8          |
|                 |                          | Бройдена – Флетчера  | 700          | 2          | 2.7      | 3.1          |
|                 | 0.1                      | Случайный поиск      | 500          | 6.5        | 12       | 15           |
|                 |                          | Сопряжен. градиенты  | 900          | 4.5        | 5        | 8            |
|                 |                          | Partan               | 700          | 4          | 6.5      | 6.9          |
|                 |                          | Дельта-дельта        | 700          | 4          | 6.5      | 7            |
|                 |                          | Обобщ. дельта-дельта | 800          | 4.4        | 5        | 5.3          |
|                 |                          | Бройдена – Флетчера  | 800          | 2.2        | 3.4      | 3.6          |
| Белый шум       | 0.3                      | Случайный поиск      | 900          | 7          | 14       | 15           |
|                 |                          | Сопряжен. градиенты  | 900          | 5          | 7        | 8.2          |
|                 |                          | Partan               | 1000         | 4.8        | 9.5      | 9.3          |
|                 |                          | Дельта-дельта        | 700          | 4          | 9        | 5.5          |
|                 |                          | Обобщ. дельта-дельта | 900          | 4.5        | 7        | 7.3          |
|                 |                          | Бройдена – Флетчера  | 1000         | 2.5        | 4.5      | 4.7          |
| Окрашенный шум  | 0.02                     | Случайный поиск      | 600          | 5          | 11       | 12           |
|                 |                          | Сопряжен. градиенты  | 800          | 3.3        | 3.7      | 6.0          |
|                 |                          | Partan               | 600          | 3.5        | 4.1      | 4.9          |
|                 |                          | Дельта-дельта        | 600          | 3.5        | 4.6      | 5.5          |
|                 |                          | Обобщ. дельта-дельта | 700          | 4.5        | 2.4      | 4            |
|                 |                          | Бройдена – Флетчера  | 800          | 2.1        | 2.6      | 3.3          |
|                 | 0.1                      | Случайный поиск      | 500          | 6.4        | 12.5     | 13           |
|                 |                          | Сопряжен. градиенты  | 850          | 4.6        | 4.8      | 4.8          |
|                 |                          | Partan               | 800          | 4.3        | 6        | 6.8          |
|                 |                          | Дельта-дельта        | 600          | 4.1        | 7        | 7            |
|                 |                          | Обобщ. дельта-дельта | 850          | 4          | 4.2      | 4.2          |
|                 |                          | Бройдена – Флетчера  | 900          | 2.0        | 3.2      | 3.4          |
| 0.3             | Случайный поиск          | 600                  | 5            | 14.5       | 15       |              |
|                 | Сопряжен. градиенты      | 800                  | 4            | 6.5        | 6.8      |              |
|                 | Partan                   | 650                  | 3.5          | 8          | 8.5      |              |
|                 | Дельта-дельта            | 700                  | 3.0          | 7          | 8.3      |              |
|                 | Обобщ. дельта-дельта     | 900                  | 3.6          | 6.5        | 6        |              |
|                 | Бройдена – Флетчера      | 800                  | 2            | 4.8        | 5        |              |
| Музыкальный фон | 0.02                     | Случайный поиск      | 500          | 4          | 10       | 10.5         |
|                 |                          | Сопряжен. градиенты  | 750          | 3.1        | 3.8      | 4.0          |
|                 |                          | Partan               | 650          | 3.5        | 4        | 4.2          |
|                 |                          | Дельта-дельта        | 600          | 3.3        | 4.5      | 4.6          |
|                 |                          | Обобщ. дельта-дельта | 750          | 4.0        | 2.6      | 2.8          |
|                 |                          | Бройдена – Флетчера  | 900          | 2.8        | 2.9      | 3.0          |
|                 | 0.1                      | Случайный поиск      | 400          | 7          | 12       | 12.5         |
|                 |                          | Сопряжен. градиенты  | 800          | 5          | 5        | 5.2          |
|                 |                          | Partan               | 850          | 4          | 5.8      | 6            |
|                 |                          | Дельта-дельта        | 550          | 4.2        | 7        | 7.1          |
|                 |                          | Обобщ. дельта-дельта | 800          | 3.9        | 4.1      | 4.2          |
|                 |                          | Бройдена – Флетчера  | 1000         | 2.3        | 3.1      | 3.3          |

Таблица 3 (продолжение 1)

| 1                | 2   | 3  | 4  | 5                                      | 6                                     | 7                                       |
|------------------|-----|--|--|--|---------------------------------------|---|
|                  | 0.3 | Случайный поиск<br>Сопряжен. градиенты<br>Partan<br>Дельта-дельта<br>Обобщ. дельта-дельта<br>Бройдена – Флетчера | 600<br>700<br>600<br>650<br>800<br>900   | 4.5<br>4<br>3.2<br>3.2<br>3.5<br>1.8   | 14<br>6.7<br>8<br>7.3<br>5.3<br>6     | 15.2<br>6.9<br>8.2<br>8<br>5.6<br>5.5   |
| Узкополосный шум | 0.2 | Случайный поиск<br>Сопряжен. градиенты<br>Partan<br>Дельта-дельта<br>Обобщ. дельта-дельта<br>Бройдена – Флетчера | 500<br>700<br>600<br>600<br>600<br>600   | 5<br>3.2<br>3.3<br>3.6<br>4<br>2.6     | 10.5<br>4.5<br>5<br>5.8<br>2.6<br>3.5 | 10<br>4.7<br>5.4<br>6.2<br>2.8<br>3.7   |
| Узкополосный шум | 0.1 | Случайный поиск<br>Сопряжен. градиенты<br>Partan<br>Дельта-дельта<br>Обобщ. дельта-дельта<br>Бройдена – Флетчера | 600<br>700<br>550<br>700<br>800<br>800   | 4.5<br>3.8<br>3.5<br>3.8<br>4.2<br>4.5 | 12<br>6.0<br>6<br>7.0<br>3.5<br>4.1   | 11.8<br>6.4<br>6.6<br>7.3<br>3.8<br>4.3 |
| Узкополосный шум | 0.3 | Случайный поиск<br>Сопряжен. градиенты<br>Partan<br>Дельта-дельта<br>Обобщ. дельта-дельта<br>Бройдена – Флетчера | 750<br>900<br>800<br>900<br>1000<br>1000 | 4<br>5<br>4.3<br>4.5<br>5<br>2.8       | 14<br>8<br>8.2<br>9<br>4.9<br>3.2     | 15<br>9<br>8.1<br>9.5<br>5.3<br>3.6     |

**Выбор структуры персептона.** Выбор структуры многослойного персептона (МСП) важен, поскольку повышенная сложность сети приводит к увеличению объема вычислений и длительности обучения, возникновению неустойчивости и переобученности, которая ухудшает показатели точности на материале тестирования. Чрезмерное упрощение в свою очередь не позволяет получить хорошее качество детектирования свойств сигнала. Поэтому построение нейронной сети для системы детектирования состояния процесса связано с выбором минимальной конфигурации, обеспечивающей хорошее функционирование.

Для сети минимального размера маловероятно обучиться несущественным деталям в обучающих данных, поэтому она может давать лучшие результаты на материале тестирования. Способ реализации такого подхода – это выбор сети малой сложности и наращивание ее до получения приемлемых показателей алгоритма обучения и правильности работы на тестовом материале. Другой подход состоит в начальном выборе сети достаточно сложной структуры, обеспечивающей приемлемую ошибку обучения, после чего производится упрощение структуры сети путем удаления избыточных нейронов. Дополнительным ресурсом является метод регуляризации структуры персептрона, в котором вместо простого критерия среднеквадратической ошибки функция риска учитывает также меру сложности структуры персептрона.

Выбор структуры нейронной сети осуществлен на основе совместного использования методов упрощения и наращивания и метода регуляризации.

Для метода наращивания в ходе проведения серии циклов обучения, основываясь на критерии ошибки (12), определялось место размещения в сети дополнительного нейрона. Одновременно из сети удалялся нейрон, веса которого флуктуировали после сходимости обучения, т.к. этот нейрон не является существенным для задачи обуче-

ния. Результаты моделирования получены на ЭВМ с использованием программы Essence 1.0 для речевого сигнала с белым шумом при отношении среднеквадратического отклонения шума и сигнала 0.1 и 0.3. Метод наращивания сложности МСП проанализирован для структур: 3-1; 5-1; 5-1-1; 5-2-1; 5-3-1 и двух лучших алгоритмов обучения: Бройдена-Флетчера и обобщенного правила дельта-дельта (табл.4). Введение третьего нейрона во втором скрытом слое МСП несущественно меняет ошибку, но обнаруживает флуктуации вектора весов дополнительного нейрона в конце цикла обучения. Поэтому метод указывает на предпочтительность структур 5-2-1.

Таблица 4  
Оптимальная структура МСП по методу наращивания

| Структура МСП | Отношение шум/сигнал | Доля ошибочных решений персептрона (%) |                            |
|---------------|----------------------|--|----------------------------|
|               |                      | Обобщенное правило дельта-дельта       | Алгоритм Бройдена-Флетчера |
| 3-1           | 0.1                  | 8.1                                    | 6.5                        |
|               | 0.3                  | 9.7                                    | 7.0                        |
| 5-1           | 0.1                  | 6.5                                    | 5.0                        |
|               | 0.3                  | 8.3                                    | 5.5                        |
| 5-1-1         | 0.1                  | 4.9                                    | 4.0                        |
|               | 0.3                  | 6.9                                    | 4.5                        |
| 5-2-1         | 0.1                  | 3.7                                    | 3.2                        |
|               | 0.3                  | 5.5                                    | 4.0                        |
| 5-3-1         | 0.1                  | 3.9                                    | 3.4                        |
|               | 0.3                  | 5.8                                    | 4.2                        |

Метод упрощения, где после каждого цикла обучения удалялся нейрон, отсутствие которого дает минимальное увеличение ошибки обучения, применен к структурам персептрона 10-3-1; 10-2-1; 8-2-1; 5-2-1 (табл.5). Результаты анализа согласуются с выводом, полученным по методу наращивания.

Метод регуляризации [33] вместо критерия среднеквадратической ошибки использует функцию риска  $R(\mathbf{w})$ :

$$R(\mathbf{w}) = E(\mathbf{w}) + \lambda \cdot E_c(\mathbf{w}), \quad (42)$$

где  $E(\mathbf{w})$  – среднеквадратическая ошибка,  $E_c(\mathbf{w})$  – мера сложности структуры персептрона, а параметр  $\lambda$  определяет относительную значимость компоненты  $E_c(\mathbf{w})$  и служит параметром регуляризации (поскольку форма критерия  $R(\mathbf{w})$  опирается на теорию регуляризации А.Н. Тихонова). В качестве меры сложности структуры персептрона использован квадрат нормы вектора синаптических весов МСП:

$$E_c(\mathbf{w}) = \|\mathbf{w}\|^2 = \sum_{i \in C} w_i^2, \quad (43)$$

где  $C$  – множество индексов всех весов сети. Эксперимент проведен на структурах МСП 5-2-1 и 5-3-1 с изменением параметра регуляризации от 0 (в исходном состоянии сети) до 0.01 с шагом 0.001 и показал, что один из нейронов второго скрытого слоя в 5-3-1 имеет 1.5 %-й вклад в  $E_c(\mathbf{w})$  и может считаться несущественным. Это подтверждает результат, полученный методом наращивания сложности персептрона.

Таблица 5  
Результаты определения оптимальной структуры МСП методом упрощения

| Структура персептрона | Отношение шум/сигнал | Доля ошибочных решений персептрона (%) |                            |
|-----------------------|----------------------|--|----------------------------|
|                       |                      | Обобщенное правило дельта-дельта       | Алгоритм Бройдена-Флетчера |
| 10-3-1                | 0.1                  | 2.1                                    | 2.3                        |
|                       | 0.3                  | 4.2                                    | 3.1                        |
| 10-2-1                | 0.1                  | 2.3                                    | 2.5                        |
|                       | 0.3                  | 4.5                                    | 3.6                        |
| 8-2-1                 | 0.1                  | 2.9                                    | 2.8                        |
|                       | 0.3                  | 5.0                                    | 3.8                        |
| 5-2-1                 | 0.1                  | 3.7                                    | 3.2                        |
|                       | 0.3                  | 5.5                                    | 4.0                        |

**Использование главных компонент сигнала для обучения распознаванию.** Дальнейшее улучшение алгоритма детектирования речевой активности достигнуто уменьшением размерности входного вектора персептрона, которое позволило уменьшить число синаптических весов, уменьшить объем обучающей выборки и ускорить процесс обучения. Уменьшение размерности входных векторов при условии сохранения содержащейся в них информации о речевом сигнале основано на методе главных компонент [52, 53].

Входной вектор МСП  $\mathbf{x}_j = [x_{1j}, x_{2j}, \dots, x_{pj}]$  имеет  $p = 22$  компоненты, число которых соответствует персептальной модели. Идея метода состоит в получении  $p$  линейных комбинаций из компонент вектора – главных компонент, каждая из которых линейно независима от других. Главная компонента  $\mathbf{y}$  выражается соотношением:

$$\mathbf{y} = \mathbf{X} \hat{\boldsymbol{\beta}}, \quad (44)$$

где матрица  $\mathbf{X}$  имеет вид:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{p1} \\ x_{12} & x_{22} & \dots & x_{p2} \\ \dots & \dots & \dots & \dots \\ x_{1N} & x_{2N} & \dots & x_{pN} \end{bmatrix}.$$

Строками матрицы  $\mathbf{X}$  служат  $p$ -мерные обучающие векторы, поступающие на вход персептрона, ее размерность равна  $N \times p$  ( $N$  – объем обучающей выборки). Вектор

$$\hat{\boldsymbol{\beta}} = [\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p]^T -$$

вектор неизвестных коэффициентов. Каждый такой вектор по (44) определяет соответствующую ему главную компоненту.

Выборочная дисперсия  $\mathbf{y}$  определяется выражением:

$$\text{var}(\mathbf{y}) = \hat{\boldsymbol{\beta}}^T \mathbf{S} \hat{\boldsymbol{\beta}}, \quad (45)$$

где  $\mathbf{S}$  – выборочная ковариационная матрица обучающих векторов, составляющих матрицу  $\mathbf{X}$ :

$$\mathbf{S} = \frac{1}{N-1} \begin{bmatrix} \sum_{i=1}^N x_{1i}^2 & \sum_{i=1}^N x_{1i}x_{2i} & \dots & \sum_{i=1}^N x_{1i}x_{pi} \\ \sum_{i=1}^N x_{2i}x_{1i} & \sum_{i=1}^N x_{2i}^2 & \dots & \sum_{i=1}^N x_{2i}x_{pi} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^N x_{pi}x_{1i} & \sum_{i=1}^N x_{pi}x_{2i} & \dots & \sum_{i=1}^N x_{pi}^2 \end{bmatrix} \quad (46)$$

Главные компоненты, начиная с первой  $\mathbf{y}_1$ , определяются по критерию максимальной дисперсии  $\hat{\boldsymbol{\beta}}^T \mathbf{S} \hat{\boldsymbol{\beta}}$  при условии, что вектор коэффициентов  $\hat{\boldsymbol{\beta}}$  нормирован, т.е.  $\hat{\boldsymbol{\beta}}^T \hat{\boldsymbol{\beta}} = 1$ . Решение условной максимизации использует функцию  $\varphi = \hat{\boldsymbol{\beta}}^T \mathbf{S} \hat{\boldsymbol{\beta}} - \hat{\lambda} (\hat{\boldsymbol{\beta}}^T \hat{\boldsymbol{\beta}} - 1)$  ( $\hat{\lambda}$  – множитель Лагранжа), градиент которой  $d\varphi/d\hat{\boldsymbol{\beta}} = 2\mathbf{S} \hat{\boldsymbol{\beta}} - 2\hat{\lambda} \hat{\boldsymbol{\beta}}$  после приравнивания к нулю дает уравнение  $(\mathbf{S} - \hat{\lambda} \mathbf{I}) \hat{\boldsymbol{\beta}} = 0$  для определения компонент вектора  $\hat{\boldsymbol{\beta}}$ . Это матричное уравнение имеет решение, если равен нулю определитель  $|\mathbf{S} - \hat{\lambda} \mathbf{I}| = 0$ . Определитель представляет собой алгебраическое уравнение степени  $p$  относительно  $\hat{\lambda}$  и позволяет найти  $p$  корней, являющихся характеристическими числами  $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p$  ковариационной матрицы  $\mathbf{S}$ . Наибольшее  $\hat{\lambda}_1$  максимизирует дисперсию  $\mathbf{y}_1$ , а решение системы  $(\mathbf{S} - \hat{\lambda}_1 \mathbf{I}) \hat{\boldsymbol{\beta}} = 0$  дает вектор  $\hat{\boldsymbol{\beta}}_1$ , определяющий первую главную компоненту.

Следующие по величине характеристические числа в качестве решения системы  $(\mathbf{S} - \hat{\lambda}_i \mathbf{I}) \hat{\boldsymbol{\beta}} = 0$  дают векторы  $\hat{\boldsymbol{\beta}}_i$ , которые определяют остальные главные компоненты по соотношению (44). Таким образом, для  $p$ -мерных данных первая главная компонента  $\mathbf{y}_1 = \mathbf{X} \cdot \hat{\boldsymbol{\beta}}_1$  представляет собой линейную комбинацию  $p$  компонент с коэффициентами, равными нормированному собственному вектору корреляционной матрицы  $\mathbf{S}$ , соответствующему максимальному характеристическому числу этой матрицы. Аналогично остальные главные компоненты соответствуют следующим по величине характеристическим числам матрицы  $\mathbf{S}$ . Главные компоненты линейно независимы, и дисперсия каждой компоненты равна соответствующему характеристическому числу матрицы  $\mathbf{S}$ .

Общая дисперсия  $p$  компонент равна следу ( $tr$ ) выборочной ковариационной матрицы:

$$tr(\mathbf{S}) \equiv var(\mathbf{y}_1) + var(\mathbf{y}_2) + \dots + var(\mathbf{y}_p).$$

Относительная значимость  $j$ -й главной компоненты определяется соотношением

$$\frac{var(\mathbf{y}_j)}{\sum_{j=1}^p var(\mathbf{y}_j)} = \frac{\hat{\lambda}_j}{tr(\mathbf{S})} = \frac{\hat{\lambda}_j}{\sum_{j=1}^p \hat{\lambda}_j} \quad (47)$$

Для уменьшения размерности выполнен анализ и выбраны  $q$  наиболее значимых компонент  $q \leq p$ , которые определяют  $\eta$  % от общей дисперсии всех  $p$  компонент. Для  $\eta = 95 \dots 98$  % вклад оставшихся  $p - q$  компонент может считаться несущественным.

При проведении исследования программа Essence 1.0 дополнена методом главных компонент. Исходными данными служили 100 векторов, представляющих нормированную мощность коэффициентов вейвлет-разложения речевого сигнала в соответствии с перцептуальной моделью.

Отношение сигнал/шум варьировалось от 3 до 50. Для каждого из четырех уровней шума речевой сигнал, преобразованный вейвлет-пакетом и имеющий первоначально 22 компоненты, обработан программой главных компонент, в результате чего получены 9 компонент, содержащих не менее 97.5 % энергии исследуемого сигнала. При этом допустимый уровень потери энергии составил не более  $\eta = 2.5$  %. Использование главных компонент в структуре перцептрона показано на рисунке. Обучение перцептрона проведено с использованием алгоритмов Бройдена – Флетчера и обобщенного правила дельта-дельта, т. к. они дали лучшие результаты обучения на исходных обучающих векторах. При уменьшении отношения сигнал/шум (в исследованном диапазоне 3 – 50) ошибка обучения возрастает на 60 % для метода Бройдена – Флетчера и вдвое – для обобщенного правила дельта-дельта Джекобса. Для белого и окрашенного шумов метод обобщенного правила дельта-дельта имеет чуть меньшую ошибку обучения (на 0.1 %) при малом отношении шум/сигнал (до 0.05) и уступает при больших уровнях шума.

Для условий узкополосного шума и музыкального фона метод обобщенного правила дельта-дельта дает меньшую ошибку обучения и тестирования на всем диапазоне уровней шума от 0.5 до 0.2 %.

Для проверки этой закономерности проанализированы алгоритмы обучения по методу сопряженных градиентов и методу параллельных касательных (метод *paran*), которые в аналогичных условиях давали в среднем на 1 % большую ошибку обучения, чем первые два метода. Для белого и окрашенного шумов метод параллельных касательных уступает методу сопряженных градиентов – ошибка на 1 % больше во всем диапазоне отношений шум/речь. Для музыкального фона это различие несколько меньше, а для узкополосного шума метод параллельных касательных дает уже на 0.5 % меньшую ошибку обучения.

Методы Бройдена – Флетчера и обобщенное правило дельта-дельта дают близкие показатели относительной ошибки и количество циклов обучения.

Метод главных компонент дает возможность снизить количество циклов на 30–35 % (в зависимости от метода и ви-

да шума) и уменьшить время обучения на 40–50 %. В то же время ошибка обучения увеличивается на 0.4–0.6 %. Для узкополосного шума и музыкального фона это снижение меньше, чем для белого и окрашенного шумов, во всем диапазоне уровней шума.

Применение метода главных компонент позволило:

- уменьшить размерность входного вектора нейроперцептрона с 22 до 9 компонент;
- уменьшить число оцениваемых в процессе обучения свободных параметров сети (синаптических весов) с 122 до 52 – для нейронной сети 5-2-1;
- соответственно увеличить скорость обучения и уменьшить объем обучающей выборки.

**Заключение.** Для детектирования речевой активности предложен метод, включающий вейвлет-преобразование сигнала в соответствии с перцептуальной моделью и нейросетевой алгоритм (см. раздел I). Аппроксимация перцептуальной модели реализована на основе вейвлет-пакета, разделяющего частотный диапазон сигнала на 22 субполосы.

Принятие решения о наличии речевой активности на анализируемом сегменте осуществлено на основе нейронной сети в форме многослойного перцептрона (МСП). На входы МСП поступает нормированный вектор с компонентами, равными мощности вейвлет-коэффициентов в субполосах разложения сигнала. Исследование показало, что для обучения МСП целесообразно использовать метод Бройдена – Флетчера, который по результатам компьютерного эксперимента имеет наилучшие характеристики по скорости обучения, величине ошибок (на обучающей и тестовой выборке) и по вероятности ошибочной классификации сегмента сигнала.

Проведены экспериментальные исследования по регуляризации структуры перцептрона, основанные на методах наращивания, упрощения и регуляризации, которые показали, что оптимальной для решения данной задачи является структура перцептрона 5-2-1.

Применение метода главных компонент позволило уменьшить размерность входного вектора нейронной сети, уменьшить число оцениваемых в процессе обучения синаптических весов и увеличить скорость обучения при уменьшенном объеме обучающей выборки.

#### Литература

1. Кратиров Д.В., Меркушева А.В. Алгоритм, основанный на вейвлет-преобразовании и нейронной сети, для бесконтактного измерения параметров газожидкостного потока / Сборник докладов Международной конференции «Датчики и системы». СПб.: Изд-во СПбГТУ, 2002. Т. 3. С.51–55.
2. Daubechies I. Painless No orthogonal Expansions / Journal of Mathematical. Physics, 1986, v. 27, p. 1271 – 1283.
3. Villemois L.F. Energy Moments in Time and Frequency for 2-Scale Equations and Wavelets / SIAM Journal of Mathematical Analysis, 1992, v. 23, p. 1119 – 1153.
4. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов. М.: Радио и связь, 1981. – 492 с.
5. Allen J.B., Rabiner L.R. A Unified Approach to Short-Time Fourier Analyses and Synthesis / Proceedings of IEEE, 1977, v. 65, n. 11, p. 1558.
6. Portnoff M.R. Time-Frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analyses / IEEE



- Transactions on Signal Processing, 1980, v. 28, n. 2, p. 55.
7. Исмаилов Ш.Ю., Меркушева А.В. Нейросетевой алгоритм на вейвлет-преобразовании нестационарного сигнала в ИИС / Сб. докладов Международной конференции по мягким вычислениям и измерениям SCM-2001. СПб., 2001, том 1, с. 251 – 256.
  8. Мальхина Г.Ф., Меркушева А.В. Вейвлет-фильтрация нестационарного сигнала с адаптацией на основе нейронной сети / Сб. докладов Международной конференции по мягким вычислениям и измерениям, SCM-2001. СПб: ГЭТУ (ЛЭТИ), 2001, т. 1, с. 239 – 242.
  9. Lang M., Guo H. Noise Reduction Using Undecimated Discrete Wavelet Transform / IEEE Signal Processing Letters, 1996, v. 3, n1, p. 8 – 10.
  10. Atal B.S. Optimal Noise Specter Form / IEEE Transactions on Signal Processing, 1979, n. 6, p. 247 – 254.
  11. Fletcher N. Auditory Patterns / Review of Modern Physics, 1940, p. 47–65.
  12. Zwicker E., Fastl H. Psychoacoustics, Facts and Models. Berlin: Springer-Verlag, 1984. – 420 с.
  13. Scharf B. Critical Bands // Foundation of Modern Auditory Theory / Ed. Tobias J. N.Y.: Acad. Press, 1970, p. 159–222.
  14. Schroeder M.R., Atlas B.C., Hall J.L. Optimizing Digital Speech Coders by Exploiting the Masking Property / Journal of Audio Engineering Society, 1995, v. 43, n 11, p. 914.
  15. Zelinski R., Noll P. Adaptive Bit Allocation with Thin Structure (on Formants) of Optimal Distortion or Noise Specter / IEEE Transactions on Signal Processing, 1977, n. 8, p. 299 – 309.
  16. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов (перев. с англ.). М.: Мир, 1978. – 848 с.
  17. Мальхина Г.Ф., Меркушева А.В. Детектирование речевого сигнала и фильтрация с адаптивным порогом / Сборник трудов факультета техн. кибернетики “Микропроцессорные средства измерений”. СПб: СПбГТУ, 2001, Вып. 2. С. 26–35.
  18. Berger J, Coifman R.R., Goldberg M.J. Removing of Noise from Music Using Local Trigonometric Bases and Wavelet-Packets / Journal of Audio Eng. Society, 1994, v. 42, n. 9, p. 808.
  19. Chui C.K., Li C. Nonorthogonal Wavelet Packets / SIAM Journal of Mathematical Analysis, 1993, v. 24, p. 712 – 738.
  20. Coifman R.R., Wickerhauser M.L. Entropy-Based Algorithms for Best Bases Selection / IEEE Transactions on Information Theory, 1992, v. 38, p. 713–718.
  21. Devis G.M., Mallat S., Zhang Z. Adaptive Time-Frequency Decompositions / Optical Engineering, 1994, v. 33, n. 7, p. 2183.
  22. Feichtinger H.G. Irregular Sampling Theories and Series Expansions of Band-Limited Functions // SIAM Journal of Mathematical Analysis. 1992. V. 23. P. 530.
  23. Ramchadran K., Vetterli M., Herley C. Wavelets, sub-band coding and best bases / Proceedings of IEEE, 1996, v. 84, n. 4, p. 353.
  24. Уоссермен Ф. Нейрокомпьютер и его применение (перев. с англ. под ред. А.И. Галушкина). М.: Изд-во Мир, 1992 – 236 с.
  25. Цыганков В.Д. Нейрокомпьютер и его применение. М., 1993. – 117 с.
  26. Drodlie K.W. Unconstrained Optimization / Numerical Analyses, Ed. D. Jackobs,. London: Academic Press, 1977. p. 229 – 388.
  27. Russo A.R. Tutorial n. 8 / Proceedings of IEEE Conference on Neural Networks. Washington, 1990, p. 12.
  28. Cybenko G. Approximation by Superposition of a Sigmoidal Function / Mathematics of Control, Signals and Systems, 1989, n. 2, p. 303 – 314.
  29. Battiti R. First and Second Order Methods for Learning: Between Steepest Descent and Newtons' Method / Neural Computations, 1992, n. 4, p. 141 – 166
  30. Baum E.B. Neural Net Algorithms that Learn in Polynoial Time from Examples and Qyeries / IEEE Transactions on Neural Networks, 1991, n. 2, p. 5–19.
  31. Baum E.B., Wilczek F. Supervised learning of probablility distributions by neural networks / Ed. D.Z. Anderson. N.Y.: American Institute of Physics, 1988, p. 52–61.
  32. Gallant A.R., White H. / Neural Networks, 1990, n. 5, p. 129–138.
  33. Hinton G.E., Nowban S.J. How Learning Can Guide Evolution / Complex systems, 1987, n. 1, p. 495–512.
  34. Hopfield J.J. The Effectiveness of Analogue Neural Network.Hardware / Networks, 1990, n. 1, p. 27–40.
  35. Saارينen S. et al. Neural networks, back-propagation and automatic differentiation / Automatic Differentiation of Algorithm: Theory, Implementation and Application / Eds. Grievank A., Corless G.F. Philadelphia, MA: SIAM, 1991, p. 31–42.
  36. Гилл Ф., Мюррей У., Райт М. Прикладная оптимизация. М.: Мир, 1985 - 209 с.
  37. Химмельблау Р. Прикладное нелинейное программирование. М: Мир, 1975, с. 98.
  38. Ariel M. Nonlinear Programming: Analyses and Methods. N.Y.: Prentice Hall, 1976.
  39. Dorny C.N. A Vector Space Approach to Models and Optimization. N.Y.: Wiley (Interscience Publishing), 1975. - 289 p.
  40. Hestenes M.G. Conjugate Direction Method in Optimization. Berlin – Heidelberg – N.Y., 1980 -48 p.
  41. Jackobs R.A. Increased Rates of Convergence through Learning Rate Adaptation / Neural Networks, 1989, v. 1, p. 295–307.
  42. Broyden C.G. A Class of Methods for Solving Nonlinear Simultaneous Equations / Mathematics of Computation, 1965, n. 19, p. 577–593.
  43. Broyden C.G. Quasi-Newton Methods and their Application to Function Minimization / Mathematics of Computation, 1967, n. 21, p. 368–381.
  44. Broyden C.G. The Convergence of Minimization Algorithms / Journal of Institute of Mathematical Applications, 1970, n. 6, p. 76–90.
  45. Broyden C.G., Dennis J.E., Moro J.J. On the Local and Superlinear Convergence of Quasi-Newton Methods / Journal of Institute of Mathematical Applications, 1973, n. 12, p. 223–245.
  46. Shanno D.F. Conditioning of quasi-Newton methods for function initialization / Mathematics of Computation, 1970, n. 24, p. 647–657.
  47. Shah B.V., Buchler R.J., Kempthorne O. / Journal Society of Industrial and Applied Mathematics, 1964, v. 12, p. 74.
  48. Haykin S. Neural Networks. A Comprehensive Foundation. N.Y.: Prentice Hall, 1994. 680 p.
  49. Власов Л.В., Мальхина Г.Ф., Тархов Д.А. Нейронный эмулятор «ESSENCE» / Датчики и системы: Сборник докладов Международной конференции. СПб.: Изд-во СПбГТУ, 2002, т. 3, с. 149–153.
  50. Загруднинов Г. М. Достоверность автоматизированного контроля. Казань: Изд-во Казанского университета, 1980. - 280 с.
  51. Любатов Ю.В. Теоретические основы моделирования цифровых систем. М.: МАИ, 1989. - 77 с.
  52. Болч Б., Хуань К. Дж. Многомерные статистические методы для экономики. М.: Статистика, 1979. - 316 с.
  53. Кендал М. Дж., Стьюарт А. Многомерный статистический анализ и временные ряды (Пер.с англ. под ред. А.Н. Колмогорова и Ю.В. Прохорова). М.: Изд-во Наука, 1976. - 736 с.