

УДК [519+681.5]:550.3

**В. И. БУЛАЕВ, Р. А. МУНАСЫПОВ****СЖАТИЕ ГЕОФИЗИЧЕСКИХ ДАННЫХ С ПРИМЕНЕНИЕМ  
ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И НЕЙРОННЫХ СЕТЕЙ**

В работе представлен алгоритм выбора параметров вейвлет-преобразования для сжатия геофизических данных с учетом заданного уровня искажений. Предложена методика построения нейронных сетей минимальной сложности для сжатия вейвлет-спектра геофизических данных, обеспечивающая уменьшение количества синаптических весов и ускорение процесса обучения сети за счет декомпозиции для каждого класса геофизических данных множества входов и первого слоя нейронной сети с учетом параметров вейвлет-преобразования. Разработан комплексный алгоритм сжатия геофизических данных с применением вейвлет-преобразования и нейронных сетей. *Вейвлет-преобразования; геофизические данные; нейронные сети*

**ВВЕДЕНИЕ**

В последнее время на многих предприятиях геофизической отрасли все большее применение находят распределенные системы управления и обработки информации. При этом возникает необходимость передачи различного рода данных между элементами этой системы. Зачастую источник и приемник информации оказываются на очень большом расстоянии друг от друга, и единственной возможностью является использование различных средств связи. При большом объеме передаваемой информации затраты составляют значительную сумму, и для сокращения расходов необходимо производить сжатие данных [1].

Существует два класса алгоритмов компрессии: сжатие без потерь и сжатие с потерей некоторого количества информации. Применение первой группы алгоритмов при сжатии геофизических данных малоэффективно, поскольку коэффициент компрессии редко превышает 2, и наиболее перспективным является использование алгоритмов сжатия с потерями. Эти алгоритмы основаны на применении линейных преобразований к исходным данным и последующей обработке полученных коэффициентов.

Анализ существующих алгоритмов сжатия с потерями [2], использующих децимацию-интерполяцию, быстрое преобразование Фурье, дискретное косинусное преобразование показывает, что они не применимы к геофизическим данным, поскольку не обеспечива-

ют требуемой степени компрессии с приемлемым уровнем искажений. Значительное повышение эффективности сжатия можно ожидать от применения методов, использующих новые системы базисных функций — вейвлеты [3]. Вейвлет-преобразование производится в плоскости частота–время, т. е. производит, в отличие от спектральных методов, частотно-временной анализ. Дискретное вейвлет-преобразование (ДВП) на сегодняшний день обеспечивает наиболее компактное представление информации при возможности использования быстрого алгоритма вычислений.

При анализе проблемы было выяснено, что невозможно создать единую модель для всего множества геофизических сигналов. В такой ситуации оправданным является применение именно вейвлет-преобразования, поскольку вейвлет-анализ изначально предполагает оперирование нестационарными данными и позволяет адаптироваться к структуре сигналов путем выбора подходящих параметров преобразования.

Для получения эффекта сжатия коэффициенты, полученные в результате применения преобразования к исходным данным, необходимо подвергнуть операции квантования. Анализ наиболее распространенных алгоритмов квантования показывает, что их применение к коэффициентам вейвлет-преобразования не всегда обеспечивает необходимую степень сжатия. Для этой цели необходимо применить аппарат нейронных сетей,

так как НС, в отличие от традиционных алгоритмов скалярного и векторного квантования, могут производить эффективное нелинейное сжатие полученных коэффициентов вейвлет-преобразования. Однако необходимо использовать НС такой структуры, которая снизит требования к аппаратным ресурсам и ускорит процесс обучения при большой размерности входных данных.

В данной статье предложен алгоритм сжатия геофизических данных с учетом заданного уровня искажений на основе комплексного использования вейвлет-преобразования и нейронных сетей.

### 1. АЛГОРИТМ СЖАТИЯ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

В ходе исследований был разработан алгоритм выбора параметров вейвлет-преобразования для сжатия геофизических данных. Алгоритм состоит из пяти этапов.

1) Формирование банка вейвлет-фильтров.

Подобрать универсальный вейвлет для сжатия геофизических данных, обеспечивающего наибольший коэффициент сжатия, невозможно, поскольку геофизические сигналы — это очень обширный класс сигналов различной природы, характеризующихся разной совокупностью свойств и информационных параметров. Однако все множество геофизических данных можно разбить на несколько классов по каким-либо общим свойствам. Например, можно выделить в отдельную группу данные волнового акустического каротажа, данные скважинной сейсмической разведки, наземной сейсмической разведки, гамма-каротажа и т.д. Все эти сигналы в пределах одного класса характеризуются сходством параметров, поэтому можно подобрать оптимальные параметры вейвлет-преобразования для сжатия геофизических данных какого-либо класса и таким образом произвести эффективную компрессию.

Таким образом, на данном этапе происходит формирование банка вейвлет-фильтров, которые впоследствии будут использованы для сжатия разнотипных геофизических данных [3]. При этом учитываются следующие свойства: компактность носителя, симметрия, гладкость базисных функций, а также количество нулевых моментов.

2) Выбор вейвлета.

В ходе экспериментов была проведена оценка влияния выбора вейвлета на степень компрессии. Геофизические данные подвер-

гались разложению с помощью различных вейвлетов из сформированного банка фильтров. Все вейвлет-фильтры (кроме Хаара) имели одинаковую длину. Результаты приведены в табл. 1.

Таблица 1

Влияние выбора вейвлета на коэффициент сжатия данных акустического каротажа

Вейвлет	Haar	Db9	Sym9	Coif3	Bior3.9	Rbio3.9
$K_{\text{сжатия}}$	2,7	3,8	4,3	4,1	3,8	4,2

Как видно из табл. 1, выбор вейвлета существенно влияет на результат: коэффициент сжатия изменяется почти в полтора раза. Причем больший коэффициент сжатия необязательно ведет к большим искажениям при реконструкции. В частности, вейвлет симплета *sym9* и обратный биортогональный вейвлет *rbio3.9* показали наибольшие коэффициенты компрессии, в то же время при использовании симплета *sym9* погрешность восстановления была наименьшая.

В результате экспериментов с более чем 1000 сигналов в качестве критерия выбора самого эффективного базиса было предложено использовать критерий максимальной компрессии:

$$K = \max_i \{k_i\}, \quad (1)$$

где  $k_i$  — коэффициент сжатия, получаемый при использовании  $i$ -го вейвлета из банка фильтров.

3) Переход к ненормализованному ДВП.

В ходе исследований было установлено, что использование для сжатия геофизических данных вейвлет-преобразования, сохраняющего норму, увеличивает диапазон значений получаемых коэффициентов, что снижает коэффициент компрессии, поскольку для хранения коэффициентов вейвлет-разложения требуется все большее количество бит, и весь выигрыш, получаемый за счет более компактного представления сигнала, теряется. Для устранения этого недостатка было предложено использовать ненормализованное вейвлет-преобразование [4]. Для ненормализованных преобразований все выражения, используемые в стандартном вейвлет-преобразовании с сохранением нормы, справедливы с точностью до нормализующего коэффициента. Прямое вейвлет-преобразование осуще-

ствляется в соответствии с формулами:

$$\begin{aligned} a_{j,k} &= \sum_n a_{j-1,n} h_{n+2k}, \\ d_{j,k} &= \sum_n a_{j-1,n} g_{n+2k}. \end{aligned} \quad (2)$$

Обратное вейвлет-преобразование осуществляется в соответствии с выражением:

$$a_{j-1,n} = \sum_k a_{j,k} h'_{n+2k} + \sum_k d_{j,k} g'_{n+2k}, \quad (3)$$

где  $a_j$  и  $d_j$  — коэффициенты аппроксимации и детализации соответственно.

Исходные вейвлет-фильтры разбиваются на две группы: фильтры разложения ( $h, g$ ) и фильтры восстановления ( $h', g'$ ). Фильтры разложения совпадают с соответствующими исходными вейвлет-фильтрами, используемыми в вейвлет-преобразовании с сохранением нормы. Коэффициенты фильтров восстановления получаются из исходных с помощью следующих соотношений:

$$h' = h2, g' = g2. \quad (4)$$

Таким образом, изначально ортогональный базис разбивается на биортогональную пару (рис. 1). Ненормализованное преобразование образует подкласс биортогональных преобразований — ортогональное ненормализованное преобразование.

4) Расчет уровня декомпозиции вейвлет-преобразования.

В ходе экспериментов было установлено, что при увеличении уровня вейвлет-разложения по базису, не сохраняющему норму, степень компрессии данных возрастает. Такая четкая тенденция сохраняется до определенной степени декомпозиции, определяемой энтропией и частотным наполнением спектра исходных данных. В примере, приведенном в табл. 2, коэффициент сжатия практически не увеличивался после превышения значения уровня разложения  $M = 4$ . Аналогичная закономерность была обнаружена и при использовании других ненормированных базисных функций разложения из сформированного банка вейвлетов.

Таблица 2

Уровень разложения	1	2	3	4	5	6	7	8	9
$K_{\text{сжатия}}$	4,9	5,6	6,2	6,4	6,4	6,5	6,5	6,5	6,5

Это объясняется тем, что с увеличением уровня разложения вейвлет-коэффициенты большой амплитуды концентрируются в низкочастотной области, и получаемый спектр становится компактней. При этом получается большое количество коэффициентов, амплитуда которых близка к нулю, означающих низкую энергетiku соответствующих частотных составляющих сигнала. Их можно отбросить без значительных потерь информации. Однако после достижения определенного уровня декомпозиции спектр более не сжимается.

В результате исследований для расчета оптимального уровня декомпозиции вейвлет-преобразования предложены следующие выражения:

$$\begin{aligned} M &= \left[ \log_2 \frac{F}{F_m} \right] + 1 \\ \text{или } M &= - \lceil \log_2 (2 \cdot \Delta t \cdot F_m) \rceil + 1, \end{aligned} \quad (5)$$

где  $F$  — максимальная частота Фурье-спектра сигнала;

$F_m$  — граничная частота полосы Фурье-спектра сигнала, в которой сосредоточена его основная энергия;

$\Delta t$  — период дискретизации.

Таблица 3

Сравнение различных алгоритмов квантования

Алгоритм	Векторное (нуль-дерево)	Децимация	Трешолдинг	Равномерное квантование	ДКП
$K_{\text{сжатия}}$	4,0	3,9	4,7	6,4	4,9
СКО	0,29	0,13	0,09	0,05	0,05

5) Выбор параметров квантования.

Для получения эффективного сжатия полученные после преобразования вейвлет-коэффициенты необходимо подвергнуть квантованию. При использовании равномерного квантователя сложным является выбор шага квантования, поскольку от его величины зависит уровень искажений, возникающих при сжатии. Наиболее часто используемая мера расчета этих искажений — величина среднего квадрата отклонения (СКО), определяемая как средняя сумма квадратов разностей значений отсчетов исходного  $x[i]$  и восстановленного  $x'[i]$  сигналов.

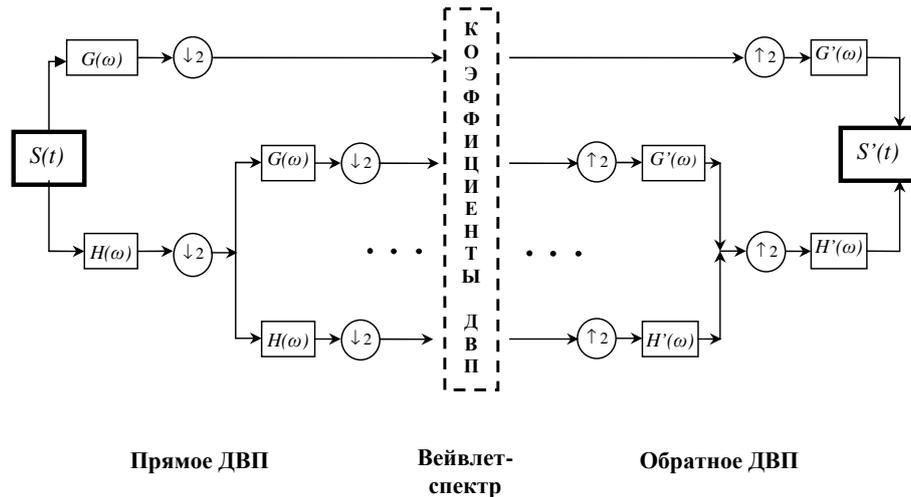


Рис. 1. Схема прямого и обратного вейвлет-преобразования

В ходе экспериментов были исследованы наиболее распространенные алгоритмы: скалярное квантование, алгоритм нуля-дерева, децимация, пороговое ограничение (табл. 3). При этом был использован вейвлет  $\text{sym9}$ , уровень разложения определялся по формуле (4), сжатию подвергались более 1000 сигналов.

В результате анализа было предложено использование равномерного скалярного квантователя.

## 2. СЖАТИЕ ВЕЙВЛЕТ-СПЕКТРА С ПОМОЩЬЮ НЕЙРОННОЙ СЕТИ

В ходе исследований была разработана методика построения нейронных сетей для дальнейшего сжатия вейвлет-спектра геофизических данных. Для этого предложено использовать четырехслойную нейронную сеть с узким горлом (рис. 2).

В работе [5] показано, что такая сеть является универсальным средством для решения задачи редукции размерности как для линейно структурированных, так и для варианта нелинейно структурированных данных. Однако применение нейронной сети данной структуры применительно к задаче сжатия вейвлет-спектра геофизических данных сопряжено с рядом сложностей:

- 1) большая размерность сети;
- 2) длительное время обучения;
- 3) длительное время работы.

С целью увеличения быстродействия нейронной сети и уменьшения требований к аппаратным ресурсам компьютера был предложен следующий подход к построению НС минимальной сложности [6]. Исходное множество входов разбивается на  $N$  подмно-

жеств, входной слой НС также декомпозируется на  $N$  подмножеств. При этом значения входных сигналов каждого подмножества подаются на нейроны только соответствующего подмножества (подсети) входного слоя НС (рис. 3), а на остальные нейроны не подаются.

Таким образом, каждому подмножеству входного сигнала ставится в соответствие своя единственная подсеть фиксированного размера, производящая предварительную обработку, число подсетей равно числу подмножеств, и все подсети имеют размер, соответствующий размеру подмножества входов. Совокупность подсетей образует входной слой НС. Данные с выхода первого слоя поступают на второй слой (узкое горло), при этом все нейроны первого слоя, как и в случае сети стандартной архитектуры, соединены со всеми нейронами выходного слоя.

Такая организация НС позволяет существенно уменьшить количество синаптических весов, ускорить обучение и снизить требования к аппаратным ресурсам, что показал проведенный сравнительный анализ структур и процессов обучения исходной и модифицированной НС (табл. 4–5).

Количество входов было равным  $n = 1024$ , количество нейронов в узком горле  $m = 12$ . В модифицированной сети количество подмножеств входного слоя равно 4, входной сигнал был декомпозирован соответственно на 4 одинаковых по размеру подмножества, по 256 входов в каждом. Во всех слоях в каждой сети использовались сигмоидальные функции активации. Каждая нейронная сеть обучалась по одной и той же обучающей странице, состоящей из 100 примеров сигналов.

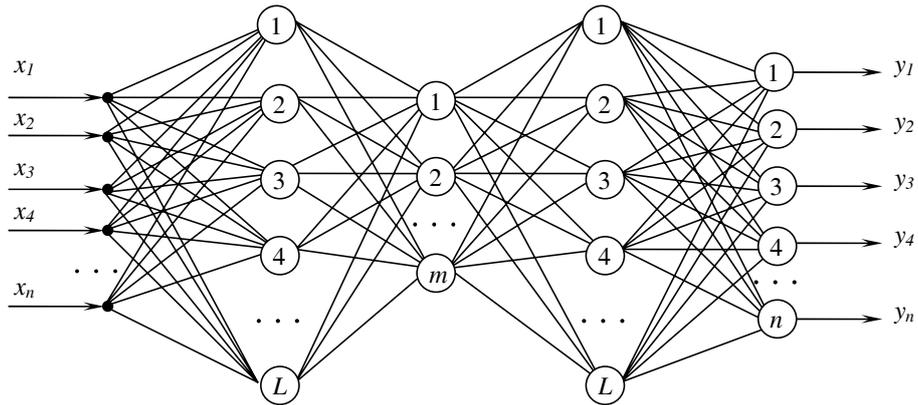


Рис. 2. Стандартная нейронная сеть для сжатия данных

Таблица 4

Структуры сравниваемых сетей

Тип НС	Число нейронов подсети ( $N_s$ )	Число нейронов первого слоя ( $N_1$ )	Число весов связей первого слоя ( $w_1$ )	Число нейронов выходного слоя ( $N_2$ )	Число весов второго слоя ( $w_2$ )
Стандартная	-	1228	1 257 472	12	14 736
Модифицированная	307	1228	314 368	12	14 736

Таблица 5

Результаты процессов обучения

Тип НС	Ошибка обучения	Количество итераций алгоритма обучения	Время обучения
Стандартная	0,008	100 000	1 час 26 мин
Модифицированная	0,010	100 000	25 мин

В ходе исследований была разработана методика синтеза НС для сжатия вейвлет-спектра, которая состоит из следующих этапов [6]:

1. Производится построение структуры нейронной сети. В соответствии с максимальным уровнем вейвлет-разложения производится декомпозиция входного множества значений коэффициентов вейвлет-спектра, а также первого слоя НС (рис. 4).

2. Задается передаточная функция нейронов сети. Для первого и третьего слоя НС устанавливаются линейные функции активации нейронов, так как зависимости между элементарными подмножествами в пределах

коэффициентов преобразования одного уровня имеют линейные компоненты и очень важно использовать эту зависимость. Для нейронов второго и четвертого слоев используются сигмоидальные функции активации.

3. Для ускорения процесса обучения каждому нейрону НС ставится в соответствие дополнительный настраиваемый параметр — смещение. Это позволяет сдвигать начало отсчета логистической функции, давая эффект, аналогичный подстройке порога персептронного нейрона. Смещение настраивается так же, как и все остальные веса. Как показала практика, применение смещений позволяет ускорить обучение в 3–4 раза.

4. Производится инициализация сети. Перед первым обучением синаптическим весам связей присваиваются небольшие начальные значения. При этом для предотвращения параллелизма весам присваиваются как положительные, так и отрицательные малые случайные значения из диапазона  $[-0,1; 0,1]$ .

5. Производится масштабирование обучающих образов, поскольку диапазоны значений коэффициентов вейвлет-преобразования  $X_i$  и данных, которыми оперирует нейронная сеть  $x_i$ , существенно отличаются. Для сохранения исходных соотношений между элементами вейвлет-спектра используется линейное масштабирование, которое осуществляется в соответствии с выражением:

$$x_i = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)}. \quad (6)$$

6. Формирование обучающих страниц. Все множество обучающих примеров разбивается по обучающим страницам, модификации параметров сети осуществляется исходя из всех примеров страницы. Содержимое первой страницы формируется из нескольких

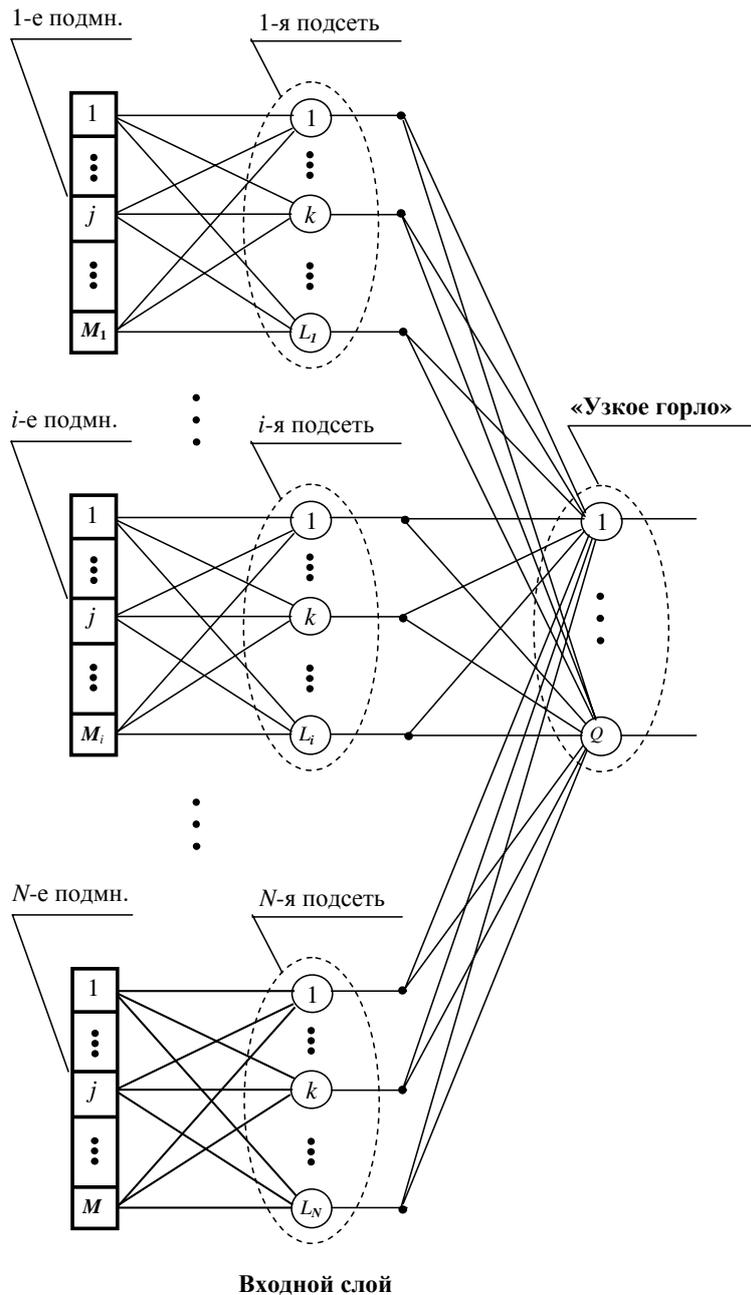


Рис. 3. Схема декомпозиции нейронной сети для уменьшения сложности

опорных примеров вейвлет-спектров. В ходе обучения объем страницы и разнообразие примеров на ней увеличивается.

7. Обучение нейронной сети [7].

### 3. КОМПЛЕКСНЫЙ АЛГОРИТМ СЖАТИЯ ГЕОФИЗИЧЕСКИХ ДАННЫХ С ПРИМЕНЕНИЕМ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ И НЕЙРОННЫХ СЕТЕЙ

В результате исследований был разработан комплексный алгоритм сжатия геофизических данных с применением вейвлет-преоб-

разования и нейронных сетей, который состоит из трех этапов (рис. 5).

1. Подготовительный этап:

1.1) осуществляется выбор параметров ДВП для сжатия геофизических данных в соответствии с разработанным алгоритмом;

1.2) производится синтез НС в соответствии с разработанной методикой построения нейронных сетей минимальной сложности для сжатия вейвлет-спектра геофизических данных.

2. Компрессия:

2.1) поступление исходных данных на вход системы;

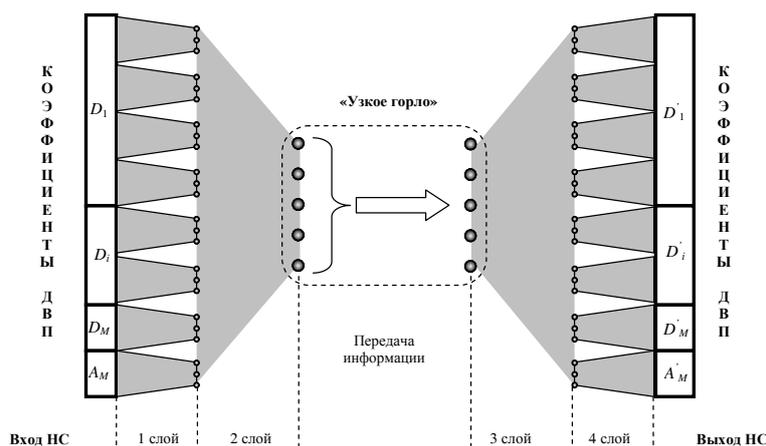


Рис. 4. Структура нейронной сети для сжатия вейвлет-спектра геофизических данных

2.2) осуществление ненормализованного дискретного вейвлет-преобразования с выбранными параметрами;

2.3) квантование полученных вейвлет-коэффициентов;

2.4) масштабирование вейвлет-спектра;

2.5) активизация соответствующей нейронной сети и подача данных на ее вход;

2.6) расчет выходных значений нейронов сжимающего слоя;

2.7) сохранение служебной информации и значений нейронов «узкого горла» в выходной файл для передачи по линии связи.

3. Декомпрессия:

3.1) поступление сжатой версии исходных геофизических данных на вход системы;

3.2) активизация соответствующей нейронной сети и подача данных на выход сжимающего слоя;

3.3) расчет значений нейронов выходного слоя, что соответствует получению масштабированного квантованного вейвлет-спектра;

3.4) обратное масштабирование вейвлет-спектра;

3.5) осуществление процедуры, обратной квантованию;

3.6) осуществление обратного ненормализованного дискретного вейвлет-преобразования;

3.7) сохранение полученных геофизических данных.

В ходе исследований было разработано программное обеспечение [8] и проведен анализ эффективности предложенного метода сжатия для компрессии данных скважинной сейсмической разведки и акустического каротажа. Коэффициент сжатия составил соответственно 17–28 и 8–16 раз. Программное обеспечение зарегистрировано и внедрено в про-

мышленную эксплуатацию в ОАО «Башнефтегеофизика» и ОАО «Когалымнефтегеофизика». Внедрение ПО в производственный режим позволило существенно сократить расходы на передачу данных, экономия составила около 50'000\$ в месяц.

## ВЫВОДЫ

Разработан алгоритм выбора параметров вейвлет-преобразования для сжатия геофизических данных, обеспечивающий эффективную компрессию с необходимой точностью. На первом этапе алгоритма формируется банк вейвлет-фильтров, на втором производится выбор вейвлета, на третьем осуществляется переход к ненормализованному вейвлет-преобразованию, на четвертом рассчитывается уровень декомпозиции и на пятом – выбор шага квантования вейвлет-коэффициентов.

Разработана методика построения нейронных сетей минимальной сложности для сжатия вейвлет-спектра геофизических данных, обеспечивающая уменьшение количества синоптических весов и ускорение процесса обучения сети за счет декомпозиции для каждого класса геофизических данных множества входов и первого слоя НС с учетом параметров вейвлет-преобразования.

Разработан комплексный алгоритм сжатия геофизических данных с применением вейвлет-преобразования и нейронных сетей, состоящий из трех этапов. На первом производятся подготовительные операции: выделение классов данных, выбор параметров вейвлет-преобразования и построение нейронных сетей минимальной сложности. На втором этапе происходит компрессия геофизических данных, на третьем – декомпрессия.

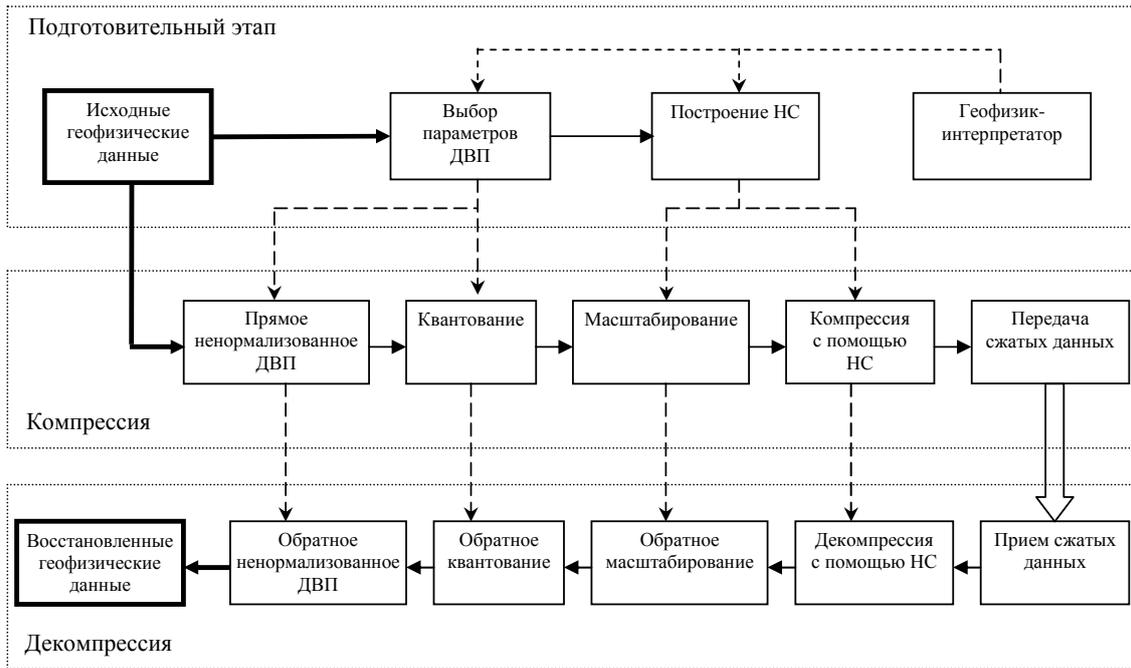


Рис. 5. Комплексный алгоритм сжатия геофизических данных с применением вейвлет-преобразования и нейронных сетей

#### СПИСОК ЛИТЕРАТУРЫ

- Булаев, В. И. Сжатие геофизических данных для передачи по спутниковым каналам связи / В. И. Булаев // Проблемы техники и технологии телекоммуникаций : мат. 3-й междунар. науч.-техн. конф. Уфа : УГАТУ, 2002. С. 141.
- Фомин, А. А. Основы сжатия информации / А. А. Фомин. СПб. : СПбГУ, 1998. 81 с.
- Добеш, И. Десять лекций по вейвлетам / И. Добеш. Ижевск : НИЦ «Регулярная и хаотическая динамика», 2001. 464 с.
- Новиков, И. Я. Основы теории всплесков / И. Я. Новиков, С. Б. Стечкин // Успехи математических наук. 1998. Т. 53, № 6 (324). С. 53–128.
- Боян, Дж. Применение автоассоциативных искусственных нейронных сетей для сжатия информации : дис... на соиск. ученой степ. канд. техн. наук / Дж. Боян. М., 2003. 169 с.
- Ильясов, Б. Г. Минимизация структуры нейронной сети для решения задачи сжатия данных / Б. Г. Ильясов, Р. А. Мунасыпов, В. И. Булаев // Нейрокомпьютеры: разработка, применение. 2007. № 10. С. 31–35.
- Горбань, А. Н. Обучение нейронных сетей / А. Н. Горбань. М. : СП ПараГраф, 1990. 159 с.
- Булаев, В. И. Свид. об офиц. рег. программы для ЭВМ № 2005611758. Программное обес-

печение сжатия данных акустического каротажа FKDCompress / В. И. Булаев, Р. Я. Адиев. М. : Роспатент, 2005.

#### ОБ АВТОРАХ

**Булаев Владимир Иванович**, гл. геофизик ОАО «Таймыр-нефтегеофизика». Дипл. инженер по АСОИ (УГАТУ, 2001). Канд. техн. наук по сист. анализу, управл. и обр. информации (УГАТУ, 2005). Иссл. в обл. искусств. нейрон. сетей, вейвлет-преобразований, интеллект. алгоритмов обработки информации, геофизики.



**Мунасыпов Рустэм Анварович**, проф. каф. техн. кибернетики. Дипл. инженер электрон. техн. (УАИ, 1982). Д-р техн. наук по сист. анализу, управл. и обр. информации (УГАТУ, 2003). Иссл. в обл. интеллект. и адаптивн. систем управления сложн. динамич. объектами.

