

**Ю.И. СЕНКЕВИЧ, Ю.В. МАРАПУЛЕЦ, О.О. ЛУКОВЕНКОВА, А.А. СОЛОДЧУК
МЕТОДИКА ВЫДЕЛЕНИЯ ИНФОРМАТИВНЫХ ПРИЗНАКОВ В
СИГНАЛАХ ГЕОАКУСТИЧЕСКОЙ ЭМИССИИ**

Сенкевич Ю.И., Марапудец Ю.В., Луковенкова О.О., Солодчук А.А. Методика выделения информативных признаков в сигналах геоакустической эмиссии.

Аннотация. Исследования геоакустической эмиссии в сейсмоактивном регионе на Камчатке показывают, что при подготовке землетрясений и последующей релаксации поля локальных напряжений в пункте наблюдений в геоакустических сигналах возникают ярко выраженные импульсные аномалии. Качественному выделению таких аномалий препятствуют сильное искажение и ослабление амплитуды сигнала. Обзор существующих методов анализа акустической эмиссии показывает, что чаще всего исследователи обращаются к анализу энергетических и статистических свойств сигналов, как более доступных для изучения. Отличительными особенностями предлагаемого авторами подхода являются выделение информативных признаков на основе анализа временной и частотно-временной структур геоакустических сигналов и описание многообразных форм распознаваемых импульсов ограниченным набором паттернов. Настоящее исследование открывает перспективу разработки новых методов выявления аномального поведения геоакустических сигналов, в том числе и перед землетрясениями.

В работе описана методика извлечения информации из потоков импульсов геоакустической эмиссии звукового частотного диапазона. Представлена математическая модель геоакустического импульса, отражающая процесс генерации сигнала от множества элементарных источников. Приводится решение задачи выделения информативных признаков в геоакустических сигналах путем описания фрагментов сигнала матрицами отношений амплитуд локальных экстремумов и интервалов между ними. Приводится результат применения разработанного алгоритма для автоматического описания структуры выделяемых импульсов и для образования множества паттернов, характеризующих особенности сигналов геоакустической эмиссии, наблюдаемых на полевых станциях ИКИР ДВО РАН. Представлена методика сокращения размерности множества выделенных импульсов, позволяющая найти близкие по структуре паттерны. Предложено решение проблемы обработки большого потока данных путем унификации описания импульсов и их систематизации. Результаты проведенных научных работ позволили создать инструмент для исследования динамических свойств сигналов геоакустической эмиссии в интересах разработки детекторов предсказания землетрясений.

Ключевые слова: геоакустическая эмиссия, модель геоакустического импульса, разреженная аппроксимация, информативные признаки, паттерны сигнала.

1. Введение. Геофизические процессы, связанные с подвижками горных пород и грунтов, сопровождаются акустическими колебаниями. Такие сигналы отличаются многообразием форм, широким динамическим диапазоном частот и амплитуд. Многолетние исследования [1-4] позволили выявить закономерные изменения характеристик этих сигналов, которые связаны с процессами развития и подготовки землетрясений.

Наблюдения на Камчатке, проводящиеся более 15 лет, показывают, что в течение нескольких суток перед сильными землетрясениями на

расстоянии сотен километров от эпицентра возникают аномалии сигналов геоакустической эмиссии в звуковом диапазоне частот (ГАЭ), выражающиеся в виде резкого и продолжительного повышения интенсивности акустических сигналов. Впервые наличие таких аномалий в диапазоне частот от единиц до десятков килогерц обсуждалось в работе А.В. Купцова [5]. В дальнейших исследованиях [6-9] эти результаты нашли подтверждение и были дополнены. Непрерывная регистрация геоакустической эмиссии на полигонах ИКИР ДВО РАН приводит к стремительному накоплению огромных объемов необработанных данных. В связи с чем возникает проблема обработки большого потока данных, требующая применения специальных методов анализа, хранения и передачи информации. Одним из способов решения данной проблемы является унификация описания и систематизация импульсов с целью выявления характерных и аномальных паттернов и их дальнейшего анализа в привязке к ряду сейсмических событий.

2. Постановка задачи исследования. Сигналы ГАЭ представляют собой упругие колебания, возникающие в процессе локальной динамической перестройки внутренней структуры приповерхностных осадочных пород, которые характеризуются малой прочностью и высокой пластичностью. В них акустическая эмиссия возникает даже при слабых деформационных изменениях, соизмеримых по уровню с приливными.

Распространение упругих волн в осадочных породах имеет характерные особенности и существенно отличается от распространения в однородных изотропных твердых средах. Модель среды для осадочных пород из-за ее сложности в настоящее время полностью не сформирована. Наиболее близкой является модель пористо-упругой, насыщенной жидкостью или газом среды. Ее основы были заложены в работах М.А. Био [10] и Я.И. Френкеля [11]. Теория Био—Френкеля базируется на предположении, что пористо-упругая среда состоит из упругого однородного скелета с сообщающимися между собой порами, которые заполнены вязкой сжимаемой жидкостью. В отличие от однородных изотропных сред возмущенное состояние такой среды характеризуется двумя полями осредненных векторов смещений: для жесткого пористого каркаса и для жидкости в поровом пространстве относительно каркаса. При анализе распространения упругих волн в осадочных породах с частотами до 20-30 кГц можно воспользоваться более простой моделью пористо-упругой среды Ф. Гассмана, согласно которой каркас и флюид движутся вместе, без проскальзывания [12]. Кроме того, в случае размещения акустического приемника в водоеме у дна [5] можно ограничиваться исследованием только продольных волн, так как поперечные волны в воде

не распространяются, а поверхностные волны на границе между жидкостью и донными породами сильно затухают.

Измерения в осадочных породах, проводимые на полевых станциях ИКИР ДВО РАН, показывают, что типичный сигнал ГАЭ представляет собой последовательность импульсов различной амплитуды и длительности с ударным возбуждением и частотным заполнением от сотен герц до первых десятков килогерц [13]. Форма импульсов геоакустической эмиссии зависит от природы процесса и свойств среды. Так, при уплотнении и разуплотнении осадочных пород, сопровождающихся трением и взаимным соударением входящих в них неоднородных фракций, характерны импульсы с быстро нарастающим фронтом и затяжным спадом. Разнообразие акустических импульсов продиктовано наложением в точке приема сигналов от множества источников, возникающих в объеме пространства, определяемом чувствительностью приемной аппаратуры, а также влиянием возможных шумов техногенного и естественного происхождения.

В настоящее время ведется поиск информации, содержащейся в многообразных динамически изменяющихся структурных формах сигналов ГАЭ. Разработке методов анализа сигналов ГАЭ посвящено достаточно большое количество исследований [13-16]. Достижению значимых результатов в этой области знаний мешают некоторые ограничения. Так, качественному выделению импульсов в сигналах ГАЭ препятствуют их сильное искажение и ослабление амплитуды, связанные с интерференцией сигналов от множества распределенных в зоне приема источников, а также с наличием в ней разнородных отличающихся по плотности фракций. Кроме того, при удалении от точки приема в среде начинают сказываться эффекты неравномерности фазовых скоростей в импульсах на разных частотах, искажающие их форму. Анализ известных методов показывает, что в большинстве случаев исследователи обращаются к анализу энергетических и статистических свойств сигналов ГАЭ, как более доступных для изучения. Отличительной особенностью настоящего исследования является разработка эффективного инструмента, позволяющего изучать многообразие форм распознаваемых импульсов ГАЭ. Целесообразность же проводимых исследований связана с перспективой разработки новых методов выявления аномального поведения геоакустических сигналов, в том числе и перед землетрясениями.

3. Представление модели. Учитывая, что геоакустические импульсы представляют собой аддитивные сигналы, которые образуются наложением элементарных моночастотных импульсов друг на друга, предлагается следующая модель геоакустического сигнала:

$$\begin{cases} s(t) = \sum_{m=1}^{N_1} \alpha_m g_m(t) + \sum_{n=1}^{N_2} \beta_n g_n(t) + R(t), \\ \|R(t)\| \rightarrow \min, \end{cases} \quad (1)$$

где $s(t)$ — фрагмент сигнала, включающий импульс; $g_m(t)$ — функции, аппроксимирующие импульс; $g_n(t)$ — функции, аппроксимирующие шумовую составляющую импульса; α_m и β_n — коэффициенты аппроксимации; N_1 — количество функций, на которые раскладывается импульс; N_2 — количество функций, на которые раскладывается шумовая составляющая фрагмента сигнала. Невязка $R(t)$ определяет отклонение аппроксимации от исходного сигнала. Функциональные зависимости (1) с достаточной точностью описывают сигналы от элементарных источников, а сама модель включает ограниченное число компонент, содержащих информацию о количественных и качественных изменениях состояний среды.

4. Методика исследования. Согласно принятой модели (1) для обнаружения изменений состояний среды необходимо идентифицировать информативные признаки сигналов ГАЭ. Для этого была разработана методика, состоящая из четырех этапов:

- выделение участков сигнала, содержащих полезные данные (импульсы);
- анализ амплитудно-фазовой структуры сигналов;
- выделение базового множества импульсов ГАЭ;
- уточнение модели геоакустического импульса.

4.1. Выделение информативных участков сигнала. Для выделения одиночных импульсов в сигналах геоакустической эмиссии предлагается алгоритм, основанный на вычислении адаптивного порога. Расчет пороговых значений ведется в непересекающихся окнах фиксированной длины n на основе среднеквадратического отклонения (СКО) по формуле:

$$S_k = S_{k-1} + A \cdot (\sigma_k - \sigma_{k-1}),$$

где S_k — пороговое значение в окне от $(k \cdot n)$ -го до $(k \cdot n + n - 1)$ -го отсчета, σ_i — СКО, вычисленное в окне от $((i - 1) \cdot n)$ -го до $(i \cdot n - 1)$ -го отсчета, A — экспериментально определяемый параметр. Поскольку вычисляемый порог должен зависеть лишь от уровня фонового шума, участки сигнала, содержащие импульсы, были исключены из процесса вычисления пороговой функции.

Для определения границ импульса сигнал просматривается скользящим окном длиной Δ в прямом и обратном направлениях, начиная с точки превышения порогового значения. Регистрация начала и конца

импульса производится в момент, когда все отсчеты в окне Δ окажутся меньше заданного значения S_0 :

$$S_0 = B \cdot \sigma_{\Delta},$$

где σ_{Δ} — СКО на данном участке сигнала, B — экспериментально определяемый параметр. Длина окна Δ также определяется экспериментально.

Выбор значения параметра B порогового детектора в условиях стационарного характера и нормального распределения шума осуществляется по известному в статистике правилу «трех сигм». Согласно данному правилу вероятность попадания значений шума в интервал $\mu \pm 3\sigma$ составляет 0.9973. Однако в условиях нестационарных помех, с которыми приходится иметь дело на практике, выбор фиксированного значения эмпирического параметра B становится нетривиальной задачей. Вычислительные эксперименты показали, что при заниженных значениях параметра B ($0 \leq B \leq 1.5$) и сильных отклонениях (выбросах) значений шума резко возрастает вероятность ошибок второго рода — выявление ложных импульсов, а при завышенных значениях B ($B > 1.5$) увеличивается вероятность ошибок первого рода, начинаются пропуски импульсов с небольшими значениями амплитуды.

Изучение специфических свойств геоакустических сигналов позволило снизить пороговые значения детектора импульсов на 3-5 дБ при фиксированной вероятности корректного выделения импульсов. Разработанный алгоритм детектора назван логической фильтрацией. Его главная идея состоит в том, что выделенный пороговой схемой детектора фрагмент сигнала проверяется на целостность цуга локальных максимумов и минимумов, зафиксированных на его протяжении. Если длина непрерывной последовательности обнаруженных в сигнале локальных экстремумов будет превышать заданную величину N , то данный фрагмент сигнала принимается за искомый импульс. И наоборот, если число непрерывно следующих друг за другом локальных экстремумов меньше установленного порога N , отсчеты фрагмента сигнала приравниваются нулю. Этот несложный алгоритм оказывается весьма эффективным в силу физических особенностей генерации импульсов ГАЭ, поскольку неоднородные по плотности фракции осадочных пород под внешним силовым воздействием, как уже отмечалось ранее, становятся источниками колебаний различной частоты, продолжительности и амплитуды, а специфика предложенного алгоритма не влияет на подсчет числа экстремумов в выделяемых фрагментах сигнала. Таким образом, алгоритм инвариантен трансформациям

сигнала в порождающей неоднородной среде. Более детально алгоритм логического фильтра представлен в виде блок-схемы на рисунке 1.

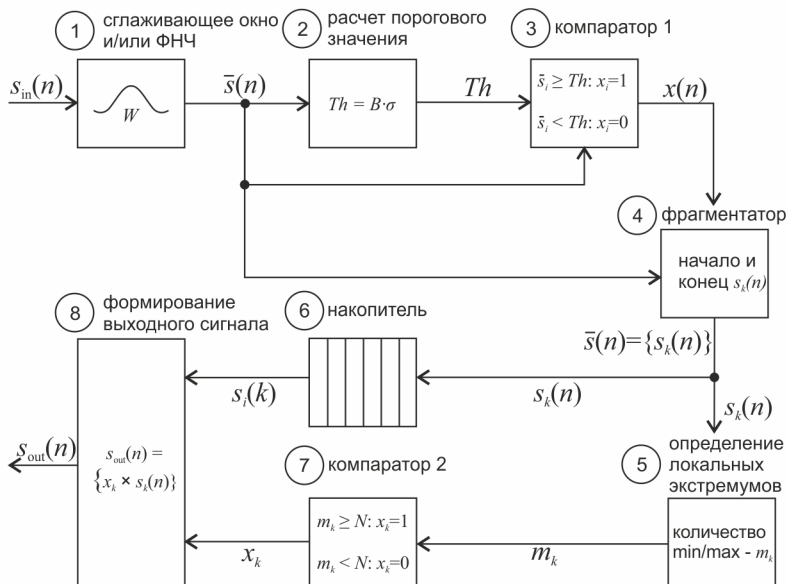


Рис. 1. Блок-схема алгоритма логического фильтра

Алгоритм работы фильтра следующий:

- оцифрованный сигнал ГАЭ $s_{in}(n)$ с частотой дискретизации 48000 Гц обрабатывается сглаживающим временным окном протяженностью от 8 до 16 отсчетов и/или низкочастотным фильтром с граничной частотой, выбираемой в зависимости от задачи исследования (от 500 Гц до 16000 Гц). Данная операция позволяет подавить высокочастотную составляющую шума и уменьшить влияние его одиночных выбросов (блок 1);

- для отфильтрованного сигнала $\bar{s}(n)$ непрерывно рассчитывается порог Th на основе СКО сигнала (блок 2) во временном окне длиной, на порядок превышающей протяженность самых длительных ожидаемых импульсов (как правило, от 3000 до 5000 отсчетов оцифрованного сигнала);

- пороговое значение Th сравнивается со значениями отсчетов отфильтрованного сигнала, затем на основе результатов сравнений формируется управляющий сигнал $x(n)$ (блок 3), при этом $x_i = 1$, если значения

соответствующих отсчетов сигнала $\bar{s}(n)$ превышают пороговое значение B , и $x_i = 0$ в противном случае;

– сигналы $x(n)$ и $\bar{s}(n)$ с выхода линии задержки поступают в блок 4, выполняющий функцию вырезания фрагментов сигнала ГАЭ в соответствии со значениями управляющего сигнала: если $x_i \neq x_{i+1}$, то, начиная с \bar{s}_{i+1} , формируется новый фрагмент;

– полученные фрагменты сигнала ГАЭ $\{s_k(n)\}$ поочередно поступают на вход блока 5, где осуществляется определение локальных экстремумов (максимумов и минимумов) каждого фрагмента $s_k(n)$ с подсчетом их количества m_k ;

– количество локальных экстремумов m_k сравнивается с установленным пороговым значением N длины цуга локальных экстремумов (блок 7), таким образом принимается решение об идентификации фрагмента как импульса ГАЭ или о ложном срабатывании детектора;

– в первом случае фрагмент $s_k(n)$ с накопителя, где фрагменты сигнала сохраняются до момента принятия решения (блок 6), остается без изменения, во втором случае формируется нулевой сигнал, совпадающий по длительности с пропускаемым фрагментом, чем достигается непрерывность последовательности отсчетов выходного сигнала $s_{out}(k)$ (блок 8).

На рисунке 2 проиллюстрированы ключевые моменты логической фильтрации тестового сигнала, представляющего собой последовательность импульсов Берлаге с наложенным белым шумом. Параметры компьютерного моделирования:

- фильтрация сигнала не применялась;
- длительность генерируемых импульсов Берлаге — 208 отсчетов;
- максимальная амплитуда импульсов Берлаге — 100;
- относительная частота модуляции импульсов Берлаге — 0.18;
- коэффициент нарастания импульса Берлаге — 2.3;
- коэффициент затухания импульсов Берлаге — 0.16;
- заданная скважность импульсов — 5.8;
- СКО шума — 100;
- ширина окна расчета СКО — 2048 отсчетов;
- коэффициент порогового значения — $1.2 (Th = 1.2\sigma)$;
- расчетное отношение сигнал/шум (SNR) — 0.47 дБ;
- длина цуга локальных экстремумов $N = 10$.

Для тестирования алгоритма логической фильтрации, представленного на блок-схеме (рисунок 2), был проведен вычислительный эксперимент, результаты которого проиллюстрированы на рисунке 3. По вертикальной оси в логарифмическом масштабе отложено число коррект-

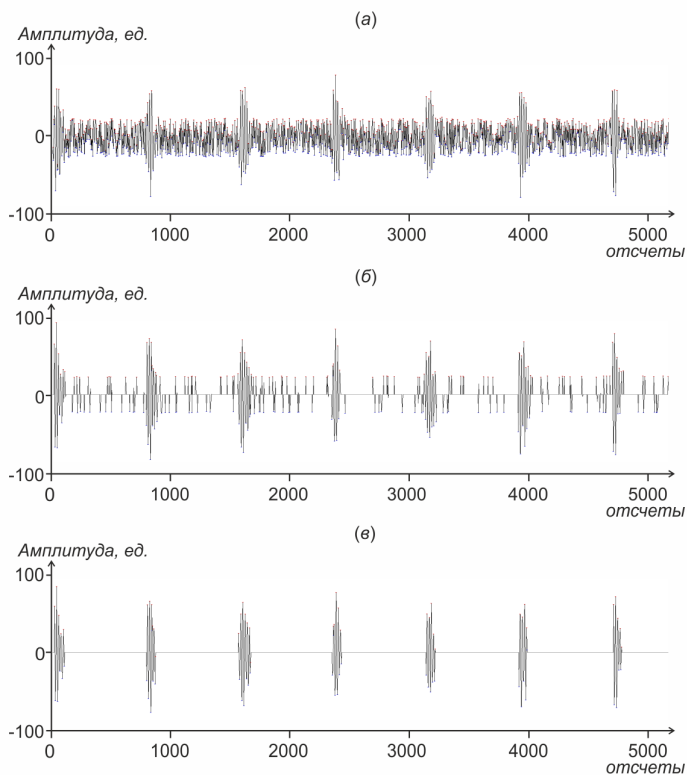


Рис. 2. Логическая фильтрация тестового сигнала: исходный сигнал (а); фрагментированный сигнал, выход блока 5 (б); результат обработки (в)

но выделенных импульсов ГАЭ, а по горизонтальной оси — значения SNR в дБ, рассчитанные во временном окне длиной 208 отсчетов.

Эксперимент проводился с помощью оригинальной компьютерной программы, имитирующей работу логического фильтра. Для эксперимента было сформировано 50 сигналов, содержащих по 100 импульсов Берлаге с характеристиками, описанными выше. Для получения сигналов с фиксированными значениями SNR от -0.5 до 20 дБ к импульсным сигналам добавлялся белый шум различной амплитуды. Далее подсчитывалось количество выделенных импульсов при применении различных значений коэффициента B . Порог Th менялся ступенчато от 0.5σ до 5.0σ , где σ — СКО накладываемого на сигнал белого шума.

На рисунке 3 представлены два типа зависимостей. Графики первого типа отражают результат применения простого порогового детектора (далее Д1) без логического фильтра к сигналу с неизвестными параметрами (на рисунке 3 обозначены сплошными линиями). Графики второго типа — результат включения в схему детектора логического фильтра (далее Д2, графики обозначены пунктирными линиями).

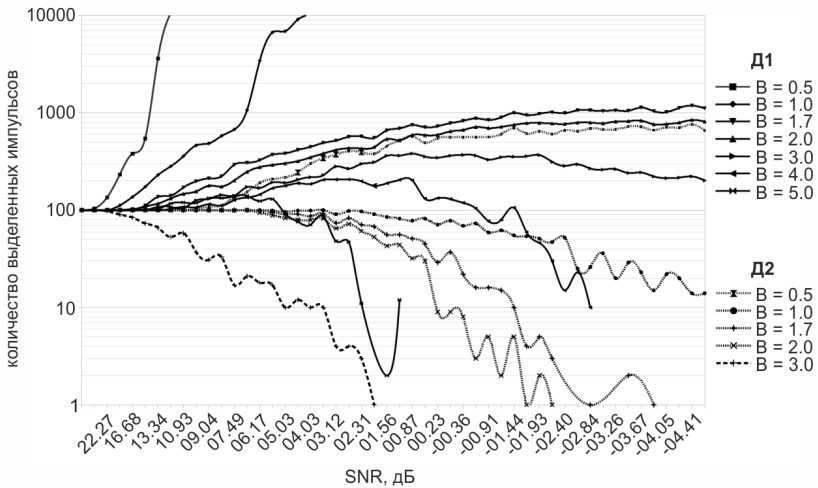


Рис. 3. Результаты сравнения пороговых детекторов без логических фильтров (Д1) и с логическими фильтрами (Д2)

Анализ графиков первой группы показывает, что детектор без логического фильтра (Д1) обеспечивает максимальную вероятность правильного обнаружения 0.95 при $B \geq 1.7$, что соответствует правилу «трех сигм». Дальнейшее уменьшение значения параметра B приводит к лавинному увеличению числа пропусков импульсов и к появлению значительного числа ложных срабатываний — ошибок первого и второго рода соответственно. Количество ошибок для обоих детекторов при различных значениях B представлено в таблице 1.

Детектор с логическим фильтром (Д2) тестировался при двух значениях $N = 12$ и $N = 3$ (минимальное количество локальных экстремумов в цуге). Данный подход позволил оценить чувствительность алгоритма к варьированию его внутренних параметров B и N . При повышении значения B все меньшее число локальных экстремумов импульса превышает порог Th , и следовательно, их число не превышает заданную величину N , тогда вероятность пропуска импульса увеличивается. Когда значение B

уменьшается длина последовательности локальных экстремумов шума может превысить параметр N логического фильтра, тогда вероятность ложной идентификации возрастает и возникают ошибки второго рода. В тех случаях, когда в ходе эксперимента размерность множества ложно выделенных импульсов превышала в три раза фактическое количество импульсов (100), возникала ситуация неопределенности в вычислении ошибки измерений, что и отмечено в таблице 1 (символом « \leftrightarrow »).

Таблица 1. Ошибки идентификации импульсов в вычислительном эксперименте для четырех значений SNR

Параметр B	Ошибки I рода		Ошибки II рода		SNR, дБ	σ_s^2/σ_n^2 , дБ
	D1	D2 ($N = 12/3$)	D1	D2 ($N = 12/3$)		
0.5	–	0/0	15283	0/–	10	5.3
1.0	–	0/0	363	0/63		
1.7	0	0/0	198	0/18		
2.0	0	0/0	156	0/0		
3.0	0	47/0	126	0/0		
0.5	–	0/0	15395	91/–	6.2	1.5
1.0	–	0/0	3325	0/224		
1.7	–	1/0	326	0/26		
2.0	0	6/0	277	0/14		
3.0	0	82/0	169	0/0		
0.5	–	0/0	14212	273/–	3.1	-1.5
1.0	–	9/0	12936	0/–		
1.7	–	26/0	519	0/74		
2.0	–	35/0	415	0/52		
3.0	3	96/3	281	0/0		
0.5	–	5/5	13031	461/461	0	-1.4
1.0	–	22/22	15871	2/2		
1.7	–	63/63	782	0/0		
2.0	–	81/81	645	0/0		
3.0	–	100/100	363	0/0		

Применение логического фильтра позволяет существенно снизить количество ошибок второго рода для значений параметра $B \leq 1.7$ при отношениях SNR от 3 дБ до 0 дБ. Цена такого результата — повышение вероятности ошибки первого рода. Однако на практике исключение ложных срабатываний позволяет отбирать для дальнейшего анализа именно участки полезного сигнала, что облегчает формирование структурных шаблонов импульсов ГАЭ.

Результаты вычислительного эксперимента доказывают применимость алгоритма логической фильтрации в задачах обнаружения импульсов и очистки импульсных сигналов от шума. Следует отметить, что на практике протяженность выделяемого цуга локальных экстремумов

колеблется от 8 до 200 и более. Этот показатель может служить мерой надежности при идентификации импульсов ГАЭ, поскольку появление фрагментов шума с протяженностью цуга локальных экстремумов более N становится маловероятно.

Перед дальнейшим анализом проводится предварительная обработка данных, включающая процедуру нормирования и сглаживания данных, которая позволяет уменьшить влияние высокочастотных помех на амплитудно-фазовые характеристики сигнала.

4.2. Анализ амплитудно-фазовой структуры сигналов. Для описания амплитудно-фазовой структуры выделенных импульсов ГАЭ было решено применить аппарат структурно-лингвистической обработки и анализа сигналов [17], адаптированный к особенностям геоакустических сигналов. Амплитудно-фазовую структуру выделенных импульсов ГАЭ было решено описать на основе отношений локальных экстремумов импульса. Данный подход позволяет представить сигнал в рамках замкнутого множества без использования пороговых значений уровней квантования амплитуды и периода дискретизации, возникающих в процессе оцифровки аналогового сигнала. Повышение точности измерений требует увеличения указанных характеристик оцифровки и порождает комбинаторно зависимое разнообразие допустимых значений сигнала. Избежать такой ситуации возможно, если трансформировать сигнал в последовательность его локальных экстремумов и интервалов между ними. Такое преобразование не кажется критичным, если абсолютные значения сигнала между его соседними локальными экстремумами можно не учитывать, что вполне оправдано на практике измерений большинства натуральных сигналов. Сущность преобразования продемонстрирована на рисунке 4.

Имеем исходное множество последовательных значений локальных экстремумов $\{x_i\}$, где $i = 1 \dots N$. Дополним это множество еще одним множеством с последовательными значениями интервалов времени τ_i , где $i = 1 \dots N - 1$, между локальными экстремумами. Вычислим отношения для каждого из экстремумов следующим образом:

$$r_{i,j} = \begin{cases} 1, & x_i > x_j \\ 0, & x_i \leq x_j \end{cases} \quad \omega_{i,j} = \begin{cases} 1, & \tau_i > \tau_j \\ 0, & \tau_i \leq \tau_j \end{cases},$$

где $r_{i,j}$ — результат логического сравнения i -го и j -го значений амплитуд экстремумов; $\omega_{i,j}$ — результат логического сравнения i -го и j -го значений интервалов между экстремумами. Упорядочим ряды таких отношений в виде квадратных матриц для отношений $r_{i,j}$ и $\omega_{i,j}$.

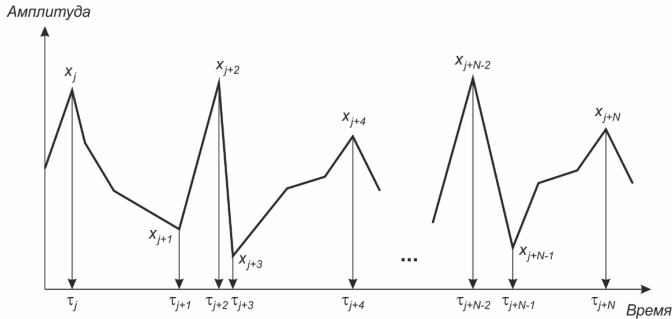


Рис. 4. Пример описания оцифрованного сигнала значениями локальных экстремумов и интервалов между ними

$$\mathbf{R} = \begin{pmatrix} 0 & r_{1,2} & \dots & r_{1,N-1} & r_{1,N} \\ r_{2,1} & 0 & & & r_{2,N} \\ \vdots & & \ddots & & \vdots \\ r_{N-1,1} & & & 0 & r_{N-1,N} \\ r_{N,1} & r_{N,2} & \dots & r_{N,N-1} & 0 \end{pmatrix}, \quad (2)$$

$$\mathbf{W} = \begin{pmatrix} 0 & \omega_{1,2} & \dots & \omega_{1,N-2} & \omega_{1,N-1} \\ \omega_{2,1} & 0 & & & \omega_{2,N-1} \\ \vdots & & \ddots & & \vdots \\ \omega_{N-2,1} & & & 0 & \omega_{N-2,N-1} \\ \omega_{N-1,1} & \omega_{N-1,2} & \dots & \omega_{N-1,N-2} & 0 \end{pmatrix}. \quad (3)$$

Полученные матрицы (2) и (3) являются избыточными в силу алгебраического свойства асимметрии неравенств (если $a > b$, то невозможно $b > a$), поэтому в дальнейшем будем использовать только верхние треугольные части каждой из построенных матриц, объединение которых образует одну квадратную матрицу порядка $N - 1$

$$\mathbf{RW} = \begin{pmatrix} r_{1,2} & r_{1,3} & \dots & r_{1,N-1} & r_{1,N} \\ \omega_{1,2} & r_{2,3} & & & r_{2,N} \\ \omega_{1,3} & \omega_{2,3} & \ddots & & \vdots \\ \vdots & & \ddots & r_{N-2,N-1} & r_{N-2,N} \\ \omega_{1,N-1} & \omega_{2,N-1} & \dots & \omega_{N-2,N-1} & r_{N-1,N} \end{pmatrix}. \quad (4)$$

Матрица (4) представляет описание участка сигнала (импульса), характеризующее взаимное положение локальных экстремумов. Преобразование сигнала, содержащего N экстремумов, в матрицу (4) обладает свойством инвариантности к операциям сдвига, а также амплитудного и временного транспонирования исходного сигнала. Инвариантность матрицы к изменениям сигнала вытекает из свойства неравенств: если $a > b$ и c – любое, то $a + c > b + c$ — для операции сдвига сигнала; если $a > b$ и $c > 0$, то $a \cdot c > b \cdot c$ — для операции транспонирования сигнала. В результате каждой матрице (4) можно сопоставить графический инвариант формы сигнала (паттерн).

Используя представленное преобразование, можно выполнить описание любого фрагмента сигнала. Пример такого структурного описания представлен на рисунке 5. Вверху слева показан эпизод оцифрованного сигнала с выделенным фрагментом; справа — матрица отношений выделенного фрагмента в символах «больше», «меньше», «равно»; внизу слева — паттерн выделенного фрагмента, построенный на основании матрицы отношений.

4.3. Методика выделения базового множества импульсов геоакустической эмиссии. Анализ импульсных сигналов ГАЭ проводится с целью получения информации о процессах, происходящих в литосфере. Классификация геоакустических сигналов, по существу, решает задачу классификации источников этих сигналов. Критерием классификации в предлагаемой авторами методике является структурное описание на основе локальных отношений экстремумов сигнала, которое как было показано ранее, может быть представлено матрицей амплитудно-фазовых отношений (4).

Согласно пункту 4.1 для выделения фрагментов сигнала, содержащих полезную информацию, используются пороговые схемы. Это приводит к накоплению огромного числа импульсов, описываемых разнообразными паттернами, которые образуют алфавит рассматриваемого фрагмента сигнала. Для выделения базового множества наиболее ти-

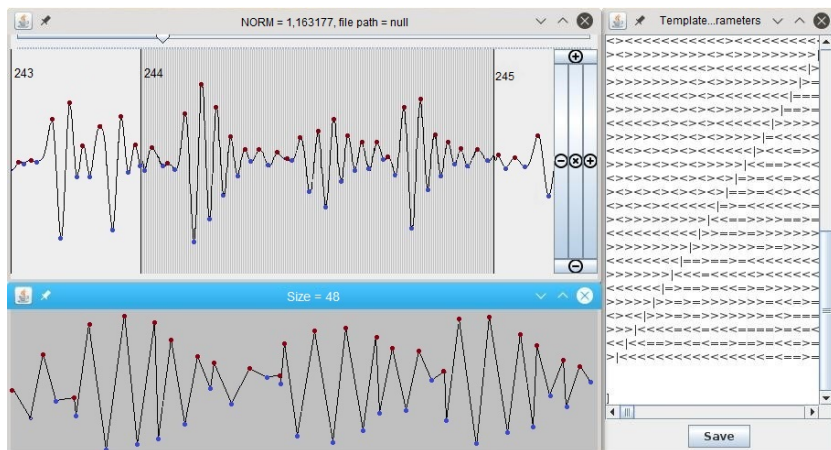


Рис. 5. Пример структурного описания фрагмента сигнала, полученного с помощью авторского программного обеспечения Registry

пичных форм импульсов ГАЭ необходимо сокращать объем алфавита. О причинах получаемого разнообразия импульсных сигналов сказано выше.

В основе механизма выделения базового множества импульсов ГАЭ лежит операция определения степени совпадения матриц структурного описания, соизмеримых по значениям порядка. Матрицы $\mathbf{A}_{n \times n}$ и $\mathbf{B}_{m \times m}$ ($n \geq m$), будут сравнимы, если $0 \leq (n - m)/n \leq P$, где P — эмпирически подбираемый порог.

Сравнение матриц одинакового порядка осуществляется путем подсчета совпадений их элементов при наложении матриц друг на друга. Если значения порядков отличны ($n > m$), осуществляется $n - m + 1$ сравнений, при этом на первом сравнении матрица порядка m вкладывается в матрицу порядка n таким образом, чтобы первые элементы первых строк совпадали, а для дальнейших сравнений сдвиг осуществляется вдоль диагонали большей матрицы, то есть матрица меньшего порядка смещается на один элемент вправо и вниз. По итогам $n - m + 1$ сравнения выбирается наибольшее количество совпадений. Данное значение сравнивается с пороговым, рассчитанным на основании эмпирически составленной формулы $S = g \cdot m^2$.

При превышении наибольшим количеством совпадений порога S принимается решение о допустимой близости структур сравниваемых импульсов. При этом статистика частоты появления импульса со структур-

ной матрицей большего порядка увеличивается на значение статистики частоты появления матрицы меньшего порядка, а паттерн с матрицей меньшего порядка удаляется из алфавита проанализированного фрагмента сигнала. Поясняющий пример сравнения двух матриц порядков 5 и 3 показан на рисунке 6.

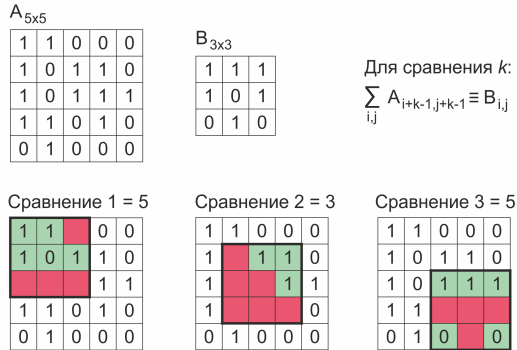


Рис. 6. Представление последовательности операций при выполнении алгоритма сравнения матриц и порядков 5 и 3 соответственно (пояснения в тексте)

4.4. Уточнение модели геоакустического импульса. На последнем этапе для полученного базового множества типичных геоакустических импульсов оцениваются параметры предлагаемой модели (1). В основу метода уточнения модели были положены идеи разреженной аппроксимации. Во-первых, разреженное представление строится по избыточной системе функций, принадлежащих разным семействам, что существенно повышает адаптивность модели импульса. Во-вторых, разреженные представления компактны и неизбыточны. Наконец, методы разреженной аппроксимации хорошо показали себя при решении аналогичных задач в смежных областях науки [18-20].

Решение задачи разреженной аппроксимации предполагает представление сигнала $s(t)$ в виде линейной комбинации минимально возможного числа функций $g_m(t)$ из предварительно выбранного базиса, что согласуется с требованиями к модели:

$$s(t) = \sum_{m=1}^N c_m g_m(t), \|c\|_0 \rightarrow \min, \quad (5)$$

где c — вектор коэффициентов разложения.

Точность и разреженность получаемых аппроксимаций напрямую зависит от выбранной системы функций (в терминах разреженной аппроксимации систему функций называют словарем, а каждую функцию в отдельности — частотно-временным атомом). Для анализа сигналов ГАЭ подходит словарь, построенный путем вариации параметров модулированных и сдвинутых функций Гаусса и Берлаге. Импульс Гаусса задается аналитическим выражением:

$$g(t) = A \cdot \exp(-B(t_{end}) \cdot \Delta \cdot t^2) \cdot \sin(2\pi ft),$$

где A — амплитуда, выбирается таким образом, чтобы $\|g(t)\| = 1$; t_{end} — длина атома; f — частота от 200 до 20000 Гц (диапазон регистрируемых частот); $B(t_{end})$ — предельное значение параметра B , вычисляемое по формуле:

$$B(t_{end}) = -\frac{4 \cdot \ln 0.05}{t_{end}^2},$$

Δ — коэффициент варьирования параметра B относительно предельного значения.

Импульс Берлаге задается аналитическим выражением:

$$g(t) = A \cdot t^{n(p_{max}) \cdot \Delta} \exp\left(-\frac{n(p_{max}) \cdot \Delta}{p_{max} \cdot t_{end}} \cdot t\right) \cdot \sin(2\pi ft),$$

где A — амплитуда, выбирается таким образом, чтобы $\|g(t)\| = 1$; t_{end} — длина атома; p_{max} — положение максимума относительно длины атома, $p_{max} \in [0.01, 0.4]$; f — частота от 200 до 20000 Гц; $n(p_{max})$ — предельное значение параметра n , вычисляемое по формуле:

$$n(p_{max}) = \frac{\ln 0.05}{\ln \frac{1}{p_{max}} - \frac{1}{p_{max}} + 1},$$

Δ — коэффициент варьирования параметра n относительно предельного значения.

Проведенные исследования показали, что импульсы Берлаге подходят для описания участков сигнала, содержащих импульсы ГАЭ, а импульсы Гаусса — для аппроксимации шумовой составляющей сигналов [13].

Задача разреженной аппроксимации (5) не разрешима за полиномиальное время. Алгоритм точного решения требует полного перебора всех возможных комбинаций функций из словаря, то есть имеет факториальную сложность $O(N!)$. Одним из наиболее часто используемых алгоритмов приближенного решения задачи разреженной аппроксимации является алгоритм согласованного преследования (Matching Pursuit, MP). Данный алгоритм был предложен Mallat и Zhang в 1993 году [21], он имеет квадратично-логарифмическую вычислительную сложность $O(N^2 \log N)$ (если для вычисления ковариации используется быстрое преобразование Фурье, в противном случае — кубическую вычислительную сложность $O(N^3)$). Алгоритм MP можно описать в виде итерационной процедуры:

$$\begin{cases} R_0(t) = s(t), \\ (m, h) = \arg \left[\max_{k,j} \langle g_k(t - \tau_j), R_i(t) \rangle \right], \\ R_{i+1}(t) = R_i(t) - \langle g_m(t - \tau_h), R_i(t) \rangle \cdot g_m(t - \tau_h), \end{cases}$$

где $R(t)$ — невязка, τ — сдвиг функции $g(t)$ относительно сигнала $s(t)$.

У алгоритма согласованного преследования имеется несколько существенных недостатков. Во-первых, для обеспечения достаточной точности разложений требуется использование словарей больших размеров, что в свою очередь вызывает степенной рост скорости выполнения алгоритма. Во-вторых, поскольку выбор атомов осуществляется из неизменяющегося словаря, полученные разложения отличаются «грубой» дискретизацией в пространстве параметров. Для решения вышеперечисленных проблем авторами было предложено улучшить классический алгоритм таким образом, чтобы на словарях ограниченного размера можно было строить разложения требуемой точности. Поскольку на каждой итерации алгоритма определяются параметры атома, имеющего наибольшее скалярное произведение с сигналом, то итерацию согласованного преследования можно описать в виде задачи поиска максимума функции многих переменных:

$$F(\tau, \mathbf{p}) = \langle g(t - \tau, \mathbf{p}), R(t) \rangle \rightarrow \max_{\mathbf{p}}$$

Главная идея предлагаемых улучшений — применять методы оптимизации для уточнения параметров \mathbf{p} функции, имеющей максимальное скалярное произведение с сигналом. Разработанный алгоритм

был назван адаптивное согласованное преследование (Adaptive Matching Pursuit, AMP) [13].

На основе изложенных выше соображений авторами были сформулированы уточненные требования к модели ГАЭ (1):

- функции $g_i(t)$ принадлежат линейному нормированному пространству $L^2(R)$;
- словарь $\mathbf{D} = \{g_i(t)\}_{1 \leq i \leq K}$ избыточен;
- $g_i(t)$ представляют собой сдвинутые по времени, модулированные функции Гаусса и Берлаге. Каждый атом однозначно определяется набором параметров \mathbf{p} : частотой f и параметрами, влияющими на форму огибающей импульса;
- атомы нормированы $\|g_i(t)\| = 1$.

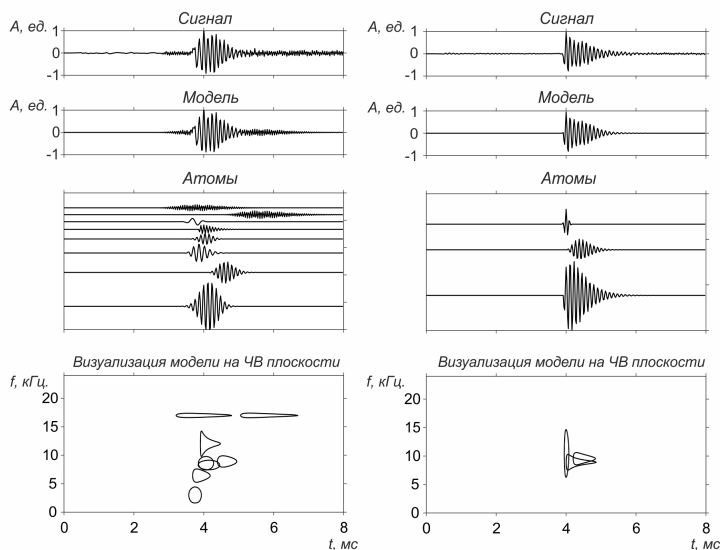


Рис. 7. Примеры моделей двух импульсов ГАЭ (контуры преобразования Вигнера—Вилля вычислены по уровню 0.2)

На рисунке 7 изображены модели двух импульсов ГАЭ. Количество атомов, обеспечивающее достаточную точность модели, определяется с помощью величины ERR , вычисляемой отношением нормы невязки к норме исходного сигнала:

$$ERR = \frac{\|R\|}{\|s\|} \cdot 100\%.$$

Разложение останавливается при достижении *ERR* порогового значения в 5%.

Для визуализации частотно-временной структуры моделей было выбрано преобразование Вигнера—Вилля, поскольку его частотно-временное разрешение минимально зависит от длительности анализируемого сигнала.

$$P_s(\tau, \omega) = \int_{-\infty}^{+\infty} s\left(\tau + \frac{t}{2}\right) \cdot s^*\left(\tau - \frac{t}{2}\right) \exp(-j\omega t) dt.$$

Однако поскольку преобразование нелинейно и строится на основе квадратичной функции от исходного сигнала, то для аддитивных сигналов в отображении проявляется сильная интерференционная компонента. Поэтому было решено вычислять преобразования Вигнера—Вилля отдельно для каждого из атомов, входящих в состав модели сигнала, и накладывать их отображения на частотно-временной плоскости (ЧВ плоскости) друг на друга.

5. Обсуждение результатов. Изложенная методика была реализована в прикладной программе [22] и применена для обработки и анализа данных наблюдений геоакустической эмиссии ИКИР ДВО РАН. Записи сигналов ГАЭ представлены в виде 15-минутных звуковых файлов. В ходе предварительной обработки каждого файла выделяются символы алфавита и ранжируются по значениям их размерности. Пример графического окна программы с результатами предобработки представлен на рисунке 8. Высота столбца отражает нормированное значение частоты появления каждого символа на интервале наблюдения в логарифмическом масштабе. Размерность символов обозначена цифрами справа сверху каждого столбца.

На рисунке 9 представлена динамика количественного состава множеств обнаруженных импульсов ГАЭ, а также множеств выделенных символов алфавитов в сигналах ГАЭ, зарегистрированных на станции «Карымшина» 01.01.2018. По рисунку 9а видно, что количество импульсов ГАЭ и символов алфавита, которые были выявлены в ходе предобработки 40 файлов, соответствующих временному периоду с 02:00:00 до 12:00:00, существенно не изменялось. В то время как на рисунке 9б видны довольно резкие изменения количества соответствующих параметров (временной период с 12:00:00 до 22:00:00). Обнаружение подобных аномалий вероятнее всего характеризует изменение условий генерации импульсов

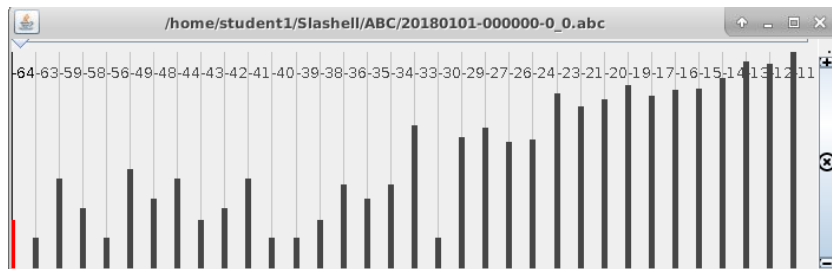


Рис. 8. Ранговое распределения символов алфавита 15-минутной записи сигнала ГАЭ

ГАЭ в среде, а, следовательно, может быть использовано как показатель изменения состояния среды под действием внешних факторов.

Предлагаемая методика кодирования импульсов с последующим выделением алфавита на выбранном интервале наблюдений позволяет рассчитать информационные показатели сигнала ГАЭ. В частности, оценки вероятности появления символов в выделенном алфавите и вычисленные на их основе показатели информационной энтропии. Для этого предусмотрена функция расчета частной энтропии каждого из символов выделенного алфавита по результатам оценки вероятности его появления согласно формуле (6), а также рассчитывается значение информационной энтропии для каждого алфавита по формуле (7):

$$H_i = -\log_2(p_i); \quad (6)$$

$$H(x) = -\sum_{i=1}^n p_i \log_2(p_i), \quad (7)$$

где H_i — частная энтропия для i -го символа алфавита; $p_i = m_i/M$ — оценка вероятности появления i -го символа; m_i — количество обнаруженных импульсов, описываемых паттерном символа i ; M — общее число импульсов, закодированных в n -символов ($\sum p_i = 1$); $H(x)$ — информационная двоичная энтропия для x независимых событий обнаружения импульсов, преобразуемых в n символов.

Частная энтропия (6) позволяет оценить информационный вклад (значимость) каждого символа, входящего в алфавит. Значение же информационной энтропии (7) отражает степень непредсказуемости состава выделяемого алфавита.

Указанные энтропийные показатели в настоящее время применяются для поиска закономерностей в поведении локальной системы

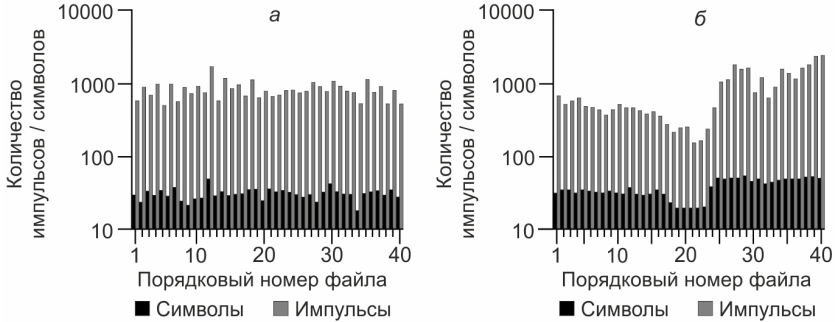


Рис. 9. Результаты предобработки сигнала ГАЭ: (а) — нормальная и (б) — anomальная картина динамики обнаруженного количества импульсов и количества символов

приповерхностных пород в зависимости от влияния внешних факторов. На рисунке 10 представлены примеры нормальной (а) и anomальной (б) картин динамики значений информационной энтропии, рассчитанной для последовательных 15-минутных интервалов сигнала ГАЭ в указанные выше временные периоды (с 02:00:00 до 12:00:00 (рисунок 10а) и с 12:00:00 до 22:00:00 (рисунок 10б) 01.01.2018). Anomalia указывает на заметное понижение неопределенности состава алфавита, связанное с некоторыми внутренними процессами в приповерхностных породах в районе измерений. Интересен тот факт, что картины anomалий (рисунки 9б и 10б) по-разному отражают наблюдаемую динамику. Данное явление еще не нашло своего объяснения.

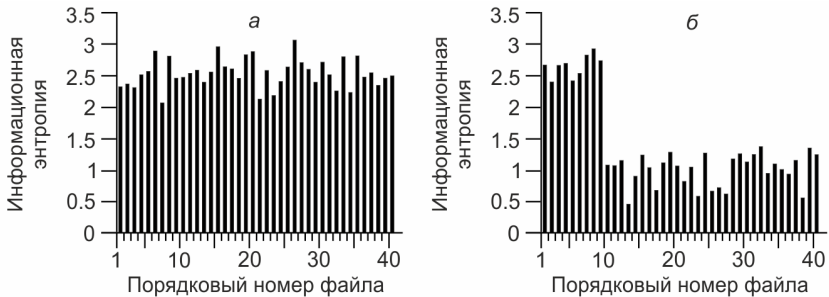


Рис. 10. Результаты предобработки сигнала ГАЭ: (а) — нормальная и (б) — anomальная картины динамики значений информационной энтропии

Таким образом, можно полагать, что представленные здесь показатели рангового распределения символов алфавита, динамика оценки числа выделяемых импульсов и размерности алфавитов, а также энтропия сигнала являются чувствительным инструментом проведения исследований изменчивости состояния системы приповерхностных слоев литосферы под действием внешних факторов в районе наблюдений, в том числе землетрясений, с использованием акустических средств регистрации.

6. Заключение. Результаты проведенных исследований показали, что представленная методика позволяет:

- выделять информативные признаки в сигналах геоакустической эмиссии по их качественным особенностям, а в перспективе и по качественно-количественным особенностям;
- определять подобные по структуре паттерны импульсных сигналов ГАЭ, сокращая тем самым размерность множества выделенных импульсов, порожденных множеством источников, а также интерференцией принимаемых сигналов;
- эффективно решать проблему обработки большого потока данных путем унификации описания импульсов и их систематизации.

Кроме того, на базе методики с применением информационного подхода к описанию сигналов удалось создать инструменты обработки и анализа сигналов ГАЭ, позволяющие проводить исследования по контролю состояния приповерхностных слоев литосферы под воздействием внешних факторов. Также разработанная методика успешно применяется для обработки и анализа импульсных электромагнитных сигналов. Таким образом, практика применения настоящей методики показала возможность ее использования при проведении исследований любых потоков импульсных сигналов. Дальнейшие исследования коллектива авторов предполагают разработку методов выявления закономерностей в поведении геоакустических и электромагнитных сигналов, в том числе и перед землетрясениями.

Литература

1. *Gregori G.P., Poscolieri M., Paparo G. et al.* “Storms of crustal stress” and AE earthquake precursors // *Natural Hazards and Earth System Sciences*. 2010. no. 10. pp. 319–337.
2. *Салтыков В.А.* О возможности использования приливной модуляции сейсмических шумов в целях прогноза землетрясений // *Физика Земли*. 2017. №2. С. 84–96.
3. *Ларионов И.А., Маратулец Ю.В, Мищенко М.А. и др.* Исследования акустической эмиссии приповерхностных осадочных пород на Камчатке // *Геосистемы переходных зон*. 2017. Т. 1, №3. С. 57–63.
4. *Giovanni G.P., Paparo G., Poscolieri M. et al.* The seismic sequence in Central Italy (August–November 2016) Acoustic Emission (AE) monitoring and analysis // *New Concepts in Global Tectonics Journal*. 2016. vol. 4. no. 4. pp. 637–663.

5. *Купцов А.В.* Изменение характера геоакустической эмиссии в связи с землетрясением на Камчатке // *Физика Земли*. 2005. №10. С. 59–65.
6. *Купцов А.В., Ларионов И.А., Шевцов Б.М.* Особенности геоакустической эмиссии при подготовке камчатских землетрясений // *Вулканология и сейсмология*. 2005. №5. С. 45–58.
7. *Гордиенко В.А., Гордиенко Т.В., Краснописцев Н.В. и др.* Аномалия высокочастотных сигналов геоакустической эмиссии как оперативный предвестник землетрясения // *Акустический журнал*. 2008. Т. 54, №1. С. 97–109.
8. *Марapultец Ю.В., Щербина А.О.* Методы исследования пространственной анизотропии геоакустической эмиссии // *Электронный журнал «Техническая акустика»*. 2008. URL: <http://ejta.org/ru/marapultec>. (дата обращения: 11.03.2019).
9. *Пережогин А.С., Шевцов Б.М.* Модели напряженно-деформированного состояния горных пород при подготовке землетрясений и их связь с геоакустическими наблюдениями // *Информационные технологии*. 2009. Т. 14. №3. С. 48–57.
10. *Biot M.A.* Theory of propagation of elastic waves in fluid-saturated porous solid. Path II. Higher frequency range // *J. Acoust. Soc. Amer.* 1956. vol. 28. no. 2. pp. 179–191.
11. *Френкель Я.И.* К теории сейсмических и сейсмoeлектрических явлений во влажной почве // *Известия АН СССР. Серия география и геофизика*. 1944. Т. 8. №4. С. 133–149.
12. *Gassmann F.* Elastic waves through a parking of spheres // *Geophysics*. 1951. vol. 16. pp. 673–685.
13. *Марapultец Ю.В., Луковенкова О.О., Тристанов А.Б., Ким А.А.* Методы регистрации и частотно-временного анализа сигналов геоакустической эмиссии. Владивосток: Дальнаука, 2017. 148 с.
14. *Shcherbina A.O.* On some features of geoacoustic emission signals before earthquakes // *E3S Web Conferences*. 2017. vol. 20. 6 p. URL: https://www.e3s-conferences.org/articles/e3sconf/abs/2017/08/e3sconf_strpep2017_03005e3sconf_strpep2017_03005.html. (дата обращения: 12.03.2019).
15. *Сычев В.Н., Имашев С.А.* Оценка параметра Херста сейсмического сигнала // *Геосистемы переходных зон*. 2017. №2. С. 50–61.
16. *Мищенко М.А.* Статистика возникновения предсейсмических аномалий в геоакустической эмиссии и атмосферном электрическом поле // *Вестник КРАУНЦ. Физико-математические науки*. 2016. №3(14). С. 47–52.
17. *Senkevich Yu.I., Duke V.A., Mishchenko M.A., Solodchuk A.A.* Information approach to the analysis of acoustic and electromagnetic signals // *E3S Web of Conferences*. 2017. vol. 20. 02012. 9 p.
18. *Yang B.Y., Liu R.N., Chen X.F.* Sparse Time-Frequency Representation for Incipient Fault Diagnosis of Wind Turbine Drive Train // *IEEE Transactions on Instrumentation and Measurement*. 2018. vol. 67. no. 11. pp. 2616–2627.
19. *Raj S., Ray K.C.* Sparse representation of ECG signals for automated recognition of cardiac arrhythmias // *Expert systems with applications*. 2018. vol. 105. pp. 49–64.
20. *Huai S., Zhang S.* A novel sparse representation algorithm for AIS real-time signals // *Eurasip journal on wireless communications and networking*. 2018. URL: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1244-9>. (дата обращения: 11.03.2019).
21. *Mallat S.G., Zhang Z.* Matching pursuits with time-frequency dictionaries // *IEEE Transactions Signal Processing*. 1993. vol. 41. pp. 3397–3415.
22. *Сенкевич Ю.И.* Структурно-лингвистическая обработка геофизических сигналов и рядов данных // *Свидетельство о государственной регистрации программы для ЭВМ №2019617637*. 18.06.2019.

Сенкевич Юрий Игоревич — д-р техн. наук., доцент, ведущий научный сотрудник лаборатории акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН. Область научных интересов: прикладные аспекты цифровой обработки сигналов, разработка методов обработки и анализа нестационарных сигналов. Число научных публикаций — 70. senkevich@ikir.ru, www.ikir.ru; ул. Мирная, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Марапулец Юрий Валентинович — д-р физ.-мат. наук., доцент, директор, Институт космофизических исследований и распространения радиоволн ДВО РАН; профессор, кафедры информатики физико-математического факультета, Камчатский государственный университет им. Витуса Беринга(КамГУ); профессор, кафедра систем управления факультета информационных технологий, Камчатский государственный технический университет(КамчатГТУ). Область научных интересов: цифровая обработка сигналов, физическая акустика, геоакустика. Число научных публикаций — 122. marpl@ikir.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Луковенкова Ольга Олеговна — канд. техн. наук., старший научный сотрудник, лаборатория акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН, доцент, кафедра информатики физико-математического факультета, Камчатский государственный университет им. Витуса Беринга(КамГУ), доцент, кафедра систем управления факультета информационных технологий, Камчатский государственный технический университет(КамчатГТУ). Область научных интересов: цифровая обработка сигналов, интеллектуальный анализ временных рядов. Число научных публикаций — 34. o.o.lukovenkova@yandex.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Солодчук Александра Андреевна — канд. физ.-мат. наук., научный сотрудник, лаборатория акустических исследований, Институт космофизических исследований и распространения радиоволн ДВО РАН. Область научных интересов: физическая акустика, геоакустика, разработка методов анализа природных сигналов. Число научных публикаций — 42. aleksandra@ikir.ru, www.ikir.ru; ул. Мирная, 7, 684034, с. Паратунка, Елизовский район, Камчатский край, Российская Федерация; р.т. +7(41531)33193; факс +7(41531)33718.

Поддержка исследований. Работа выполнена при финансовой поддержке Российского научного фонда (проект № 18-11-00087).

Yu.I. SENKEVICH, Yu.V. MARAPULETS, O.O. LUKOVENKOVA,
A.A. SOLODCHUK

TECHNIQUE OF INFORMATIVE FEATURES SELECTION IN GEOACOUSTIC EMISSION SIGNALS

Senkevich Yu.I., Marapulets Yu.V., Lukovenkova O.O., Solodchuk A.A. Technique of Informative Features Selection in Geoacoustic Emission Signals.

Abstract. Studies of geoacoustic emission in a seismically active region in Kamchatka show that geoacoustic signals produce pronounced pulse anomalies during the earthquake preparation and post-seismic relaxation of the local stresses field at the observation point. The qualitative selection of such anomalies is complicated by a strong distortion and weakening of the signal amplitude. A review of existing acoustic emission analysis methods shows that most often researchers turn to the analysis of more accessible to study statistical properties and energy of signals. The distinctive features of the approach proposed by the authors are the extraction of informative features based on the analysis of time and frequency-time structures of geoacoustic signals and the description of various forms of recognizable pulses by a limited pattern set. This study opens up new ideas to develop methods for detecting anomalous behavior of geoacoustic signals, including anomalies before earthquakes.

The paper describes a technique of information extraction from geoacoustic emission pulse streams of sound frequency range. A geoacoustic pulse mathematical model, representing the signal generation process from a variety of elementary sources, is presented. A solution to the problem of detection of geoacoustic signal informative features is presented by the means of description of signal fragments by the matrixes of local extrema amplitude ratios and of interval ratios between them. The result of applying the developed algorithm to describe automatically the structure of the detected pulses and to form a pattern set is shown. The patterns characterize the features of geoacoustic emission signals observed at IKIR FEB RAS field stations. A technique of reduction of the detected pulse set dimensions is presented. It allows us to find patterns similar in structure. A solution to the problem of processing of a large data flow by unifying pulses description and their systematisation is proposed. The results of the research allowed the authors to create a tool to investigate the dynamic properties of geoacoustic emission signal in order to develop earthquake prediction detectors.

Keywords: Geoacoustic Emission, Geoacoustic Pulse Model, Sparse Approximation, Informative Features, Signal Patterns.

Senkevich Yuri Igorevich — Ph.D., Dr.Sci., Associate Professor, Leading Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS. Research interests are applied aspects of digital signal processing, development of processing and analysis methods of non-stationary signals. The number of publications is 70. senkevich@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Marapulets Yuri Valentinovich — Ph.D., Dr.Sci., Associate Professor; Director, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS; Professor, Department of Informatics of Faculty of Physics and Mathematics, KamGU Vitus Bering; Professor, Department of control systems of Faculty of Information Technology, KamchatSTU. Research interests are digital signal processing, physical acoustics, geoacoustics. The number of publications is 122.

marpl@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Lukovenkova Olga Olegovna — Ph.D., Senior Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS; Associate Professor; Department of Informatics of Faculty of Physics and Mathematics, KamGU Vitus Bering; Associate Professor, Department of Control Systems of Faculty of Information Technology, KamchatSTU. Research interests are digital signal processing, time series data mining. The number of publications is 34. o.o.lukovenkova@yandex.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Solodchuk Alexandra Andreevna — Ph.D., Researcher, Laboratory of Acoustic Research, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS. Research interests are physical acoustics, geoaoustics, development of analysis methods of natural signals. The number of publications is 42. aleksandra@ikir.ru, www.ikir.ru; 7, Mirnaya, 684034, Paratunka, Elizovskiy district, Kamchatka region, Russian Federation; office phone +7(41531)33193; fax +7(41531)33718.

Acknowledgements. This research is supported by Russian Science Foundation (project 18-11-00087).

References

1. Gregori G.P., Poscolieri M., Paparo G. et al. [“Storms of crustal stress” and AE earthquake precursors]. *Natural Hazards and Earth System Sciences*. 2010. no. 10. pp. 319–337.
2. Saltykov V.A. [On the possibility of using the tidal modulation of seismic waves for forecasting earthquakes]. *Izvestiya. Physics of the Solid Earth*. 2017. vol. 53. no. 2. pp. 250–261.
3. Larionov I.A., Marapulets Yu.V., Mishchenko M.A. et al. [Issledovaniia akusticheskoi emissii pripoverkhnostnykh osadochnykh porod na Kamchatke]. *Geosistemy perekhodnykh zon — Geosystems of Transition Zones*. 2017. vol. 1, no. 3. pp. 57–63. (In Russ.).
4. Giovanni G.P., Paparo G., Poscolieri M. et al. [The seismic sequence in Central Italy (August–November 2016) Acoustic Emission (AE) monitoring and analysis]. *New Concepts in Global Tectonics Journal*. 2016. vol. 4. no. 4. pp. 637–663.
5. Kuptsov A.V. [Variations in the geoaoustic emission pattern related to earthquakes on Kamchatka]. *Izvestiya. Physics of the Solid Earth*. 2005. vol. 41. no. 10. pp. 825–831.
6. Kuptsov A.V., Larionov I.A., Shevtsov B.M. [Osobennosti geoausticheskoi emissii pri podgotovke kamchatskikh zemletriasenii]. *Vulkanologiya i seismologiya — Volcanology & Seismology*. 2005. no. 5. pp. 45–58. (In Russ.).
7. Gordienko V.A., Gordienko T.V., Krasnopistsev N.V. et al. [Anomaly in high-frequency geoaoustic emission as a close earthquake precursor]. *Acoustical physics*. 2008. vol. 54, no. 1. pp. 82–93.
8. Marapulets Yu.V., Shcherbina A.O. [Metody issledovaniia prostranstvennoi anizotropii geoausticheskoi emissii]. *Elektronnyi zhurnal “Tekhnicheskaya akustika” — Electronic Journal Technical Acoustics*. 2008. Available at: <http://ejta.org/ru/marapuletz>. (accessed 11.03.2019). (In Russ.).
9. Perezhogin A.S., Shevtsov B.M. [Modeli napriazhenno-deformirovannogo sostoianii gornyykh porod pri podgotovke zemletriasenii i ikh svyaz s geoakusticheskimi nabliudeniiami]. *Vychislitelnye tekhnologii — Computational Technologies*. 2009. vol. 14, no. 3. pp. 48–57. (In Russ.).
10. Biot M.A. [Theory of propagation of elastic waves in fluid-saturated porous solid. Path II. Higher frequency range]. *J. Acoust. Soc. Amer*. 1956. vol. 28. no. 2. pp. 179–191.

11. Frenkel Ya.I. [K teorii seismicheskikh i seismoelektricheskikh iavlenii vo vlazhnoi pochve]. *Izvestia AN SSSR. Seriya geografii i geofizika*. 1944. vol. 8. no. 4. pp. 133–149. (In Russ.).
12. Gassmann F. [Elastic waves through a parking of spheres]. *Geophysics*. 1951. vol. 16. pp. 673–685.
13. Marapulets Iu.V., Lukovenkova O.O., Tristanov A.B., Kim A.A. *Metody registratsii i chastotno-vremennogo analiza signalov geoakusticheskoi emissii* [Methods for recording and for time-frequency analysis of geoacoustic emission signals]. Vladivostok: Dalnauka, 2017. 148 p. (In Russ.)
14. Shcherbina A.O. [On some features of geoacoustic emission signals before earthquakes]. *E3S Web Conferences*. 2017. vol. 20. 6 p. Available at: https://www.e3s-conferences.org/articles/e3sconf/abs/2017/08/e3sconf_strpep2017_03005/e3sconf_strpep2017_03005.html. (accessed 12.03.2019).
15. Sychev V.N., Imashev S.A. [Otsenka parametra Khersta seismicheskogo signala]. *Geosistemy perekhodnykh zon — Geosystems of Transition Zones*. 2017. no. 2. pp. 50–61.
16. Mishchenko M.A. [Statistika vozniknoveniia predeismicheskikh anomalii v geoakusticheskoi emissii i atmosfernom elektricheskom pole]. *Vestnik KRAUNTs. Fiziko-matematicheskie nauki — Bulletin KRASEC. Physical & Mathematical Sciences*. 2016. no. 3(14). pp. 47–52. (In Russ.).
17. Senkevich Yu.I., Duke V.A., Mishchenko M.A., Solodchuk A.A. Information approach to the analysis of acoustic and electromagnetic signals. *E3S Web of Conferences*. 2017. vol. 20. 02012. 9 p.
18. Yang B.Y., Liu R.N., Chen X.F. [Sparse Time-Frequency Representation for Incipient Fault Diagnosis of Wind Turbine Drive Train]. *IEEE Transactions on Instrumentation and Measurement*. 2018. vol. 67. no. 11. pp. 2616–2627.
19. Raj S., Ray K.C. [Sparse representation of ECG signals for automated recognition of cardiac arrhythmias]. *Expert systems with applications*. 2018. vol. 105. pp. 49–64.
20. Huai S., Zhang S. [A novel sparse representation algorithm for AIS real-time signals]. *Eurasip journal on wireless communications and networking*. 2018. Available at: <https://jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-018-1244-9>. (accessed 11.03.2019).
21. Mallat S.G., Zhang Z. [Matching pursuits with time-frequency dictionaries]. *IEEE Transactions Signal Processing*. 1993. vol. 41. pp. 3397–3415.
22. Senkevich Yu.I. [Structural and linguistic processing of geophysical signals and data series]. Certificate of state registration of computer programs №2019617637. 18.06.2019. (In Russ.)