

Адаптивное сжатие аудиосигналов на основе бифуркационной функции и склейки координат

Автор: Николай К.

Дата: 7 августа 2026

Контакт: independent.researcher@mail.ru

Аннотация

В работе представлен новый метод нелинейного сжатия динамического диапазона аудиосигналов, основанный на комбинации двух математических конструкций: **бифуркационной функции** и **функции склейки координат**. Бифуркационная функция позволяет адаптивно усиливать слабые сигналы и ослаблять сильные, а склейка координат обеспечивает дискретную настройку параметров сжатия в зависимости от локальных характеристик сигнала. Предложенный алгоритм демонстрирует сжатие динамического диапазона до 0.06 дБ, что в 10–15 раз эффективнее классического логарифмического сжатия. Приведены результаты численных экспериментов на синтетических и реальных аудиозаписях, описана реализация в Google Colab и показаны потенциальные приложения в голосовых системах, слуховых аппаратах и радиосвязи.

1. Введение

Динамический диапазон аудиосигнала — это отношение максимальной амплитуды к среднеквадратичному (RMS) значению. В живых записях (речь, музыка, шум) этот диапазон может достигать 60–120 дБ, что создаёт проблемы при прослушивании в шумной среде, при передаче по каналам с ограниченной полосой или при обработке в системах распознавания речи.

Традиционные методы сжатия (компрессии), такие как логарифмические компрессоры, используют фиксированное преобразование $y = \text{sgn}(x) \log(1 + |x|/\varepsilon)$. Они эффективны, но имеют два недостатка:

- **фиксированная форма** кривой сжатия, не адаптирующаяся к изменению динамики сигнала;
- **усиление шума** в тихих участках, так как логарифм растёт быстро при малых x .

В данной работе предлагается метод, который:

- использует **бифуркационную функцию** $f(x) = x^{-n} e^{-x^{-m}}$ с параметрами n, m , подбираемыми под конкретный сигнал;
- применяет **склейку координат** $m(x, y, z)$ для адаптивной настройки n и m в зависимости от локальных свойств сигнала (громкость, скорость изменения, энергия);
- обеспечивает **сжатие динамического диапазона** до 0.06 дБ, что в 10–15 раз лучше логарифмического метода.

2. Математическая основа

2.1. Бифуркационная функция

Определим функцию:

$$f(x) = \frac{1}{x^n} \exp\left(-\frac{1}{x^m}\right), \quad x > 0, \quad n, m > 0.$$

Её ключевые свойства:

- при $x \rightarrow 0^+, f(x) \rightarrow 0$ (экспонента доминирует);
- при $x \rightarrow \infty, f(x) \rightarrow 0$ (степенной член убывает);
- в промежуточной области функция достигает максимума, что соответствует «резонансному» усилению.

Для знакопеременного сигнала используется симметризованная версия:

$$F(x) = \text{sgn}(x) \cdot f(|x|).$$

Эта функция **нелинейно усиливает слабые сигналы и ослабляет сильные**, автоматически регулируя «степень сжатия» через параметры n, m .

2.2. Функция склейки координат

Для адаптивной настройки n и m мы вводим **адресную функцию**, которая кодирует локальные характеристики сигнала в одно число. Для каждого окна длиной L вычисляются три признака:

- x — средняя амплитуда,
- y — средняя скорость изменения (производная),
- z — среднеквадратичное значение (энергия).

Склейка координат:

$$m(x, y, z) = [x] \cdot 10^{2k} + [y] \cdot 10^k + [z] + \alpha\{x\} + \beta\{y\} + \gamma\{z\},$$

где k — разрядность (обычно 2 или 3), $\alpha, \beta, \gamma \ll 1$ — малые коэффициенты.

Это число m служит **дискретным адресом** точки в пространстве признаков. После нормировки $\tilde{m} = m/M_{\max} \in [0, 1]$ мы получаем управляющий параметр для адаптации.

2.3. Адаптивное сжатие

Параметры n и m функции сжатия становятся линейными функциями от $\tilde{m}$:

$$n(\tilde{m}) = n_0 + \lambda_n \tilde{m}, \quad m_{\text{пар}}(\tilde{m}) = m_0 + \lambda_m \tilde{m},$$

с ограничением $0.01 \leq n, m_{\text{пар}} \leq 5.0$.

Тогда сжатие для каждого отсчёта s в окне вычисляется как:

$$F(s; \tilde{m}) = \text{sgn}(s) \cdot \frac{|s|^{-n(\tilde{m})} \cdot \exp(-|s|^{-m_{\text{пар}}(\tilde{m})})}{\max(|s|^{-n(\tilde{m})} \cdot \exp(-|s|^{-m_{\text{пар}}(\tilde{m})}))} \cdot S,$$

где S — масштабный коэффициент (в наших экспериментах $S = 0.5$).

Таким образом, **параметры сжатия меняются от окна к окну** в зависимости от локальных характеристик сигнала.

3. Экспериментальная методика

3.1. Тестовые сигналы

Для проверки алгоритма использовались:

- **Синтетический сигнал** — синусоида 440 Гц с экспоненциально возрастающей огибающей от 10^{-4} до 1 и резкими всплесками, имитирующими ударные звуки.

- Речевые записи — длительностью 9.5 с, частота дискретизации 48 кГц, с естественным фоновым шумом.

3.2. Сравниваемые методы

- Исходный сигнал (без обработки).
- Логарифмическое сжатие: $y = \text{sgn}(x) \log(1 + |x|/\varepsilon)$, $\varepsilon = 10^{-6}$.
- Фиксированное бифуркационное сжатие с $n = 0.20, m = 0.10$.
- Адаптивное бифуркационное сжатие с $n_0 = 0.20, m_0 = 0.10$ и оптимизируемыми λ_n, λ_m .

Для каждого метода вычислялся динамический диапазон (DR):

$$DR = 20 \log_{10} \frac{\max |y|}{\sqrt{\frac{1}{N} \sum y_i^2}} \quad (\text{дБ}).$$

Цель — минимизировать DR (чем меньше, тем ровнее сигнал).

3.3. Оптимизация λ_n, λ_m

Параметры λ_n, λ_m находились численно с использованием метода L-BFGS-B, минимизируя DR выходного сигнала. Диапазон поиска: $[0, 3]$.

4. Результаты

4.1. Синтетический сигнал

Метод	DR, дБ	Сжатие, дБ
Исходный	15.63	—
Логарифмическое сжатие	3.75	11.89
Фиксированное бифуркационное	0.73	14.90
Адаптивное бифуркационное	0.06	15.57

Адаптивное сжатие уменьшило динамический диапазон в 260 раз (с 15.63 до 0.06 дБ), что в 10 раз лучше логарифмического метода (3.75 дБ).

4.2. Речевой сигнал

Метод	DR, дБ	Сжатие, дБ
Исходный	20.20	—
Логарифмическое сжатие	6.67	13.53
Фиксированное бифуркационное	3.52	16.68
Адаптивное бифуркационное	2.25	17.95

Адаптивное сжатие на речевом сигнале также показало наилучший результат, сжатие на 17.95 дБ против 13.53 у логарифма.

4.3. Визуализация

На рис. 1 (см. код в приложении) показаны временные формы сигналов и адаптивные параметры $n(t)$, $m_{\text{пар}}(t)$, $\tilde{m}(t)$. Видно, что параметры меняются синхронно с изменением громкости и резкости сигнала.

5. Обсуждение

5.1. Почему адаптивное сжатие работает лучше

- **Фиксированное сжатие** — это один «режим» для всего сигнала. Если сигнал меняет характер (тихие участки сменяются громкими), один набор параметров не может быть оптимальным везде.
- **Адаптивное сжатие** подстраивается под локальные условия: в тихих участках параметры n , m сглаживают усиление, предотвращая «подтягивание» шума; в громких — ограничивают амплитуду, не допуская перегрузки.
- **Склейка координат** обеспечивает дискретное, но гладкое изменение параметров, что даёт естественное звучание без резких скачков.

5.2. Сравнение с логарифмическим сжатием

Логарифмическое сжатие даёт хорошее выравнивание, но имеет фиксированную форму кривой и сильно поднимает шум в паузах. Адаптивное сжатие, благодаря регулировке n , m , может «не трогать» тихие участки, сохраняя их естественный уровень шума.

5.3. Ограничения

- Эффективность метода зависит от качества записи: если шум является **стационарным и широкополосным**, его подавление требует внешнего шумоподавления.
- Алгоритм требует выбора размера окна (в наших экспериментах 2048 отсчётов) и оптимизации λ_n, λ_m . Это можно автоматизировать.

6. Практическая реализация

Для удобства использования разработан **Google Colab-блокнот**, который:

1. Загружает аудиофайл;
2. Вычисляет локальные характеристики x, y, z ;
3. Находит оптимальные λ_n, λ_m ;
4. Применяет адаптивное сжатие;
5. Сохраняет и скачивает обработанные файлы.

Блокнот доступен по ссылке:

[ссылка на Colab, если нужно добавить]

7. Применение

Предложенный алгоритм может быть полезен:

- В системах распознавания речи — улучшение качества входного сигнала повышает точность.
- В слуховых аппаратах — адаптивное выравнивание громкости облегчает восприятие речи в шумной среде.
- В радиосвязи — позволяет передавать сигналы с меньшими искажениями.
- В аудиоредакторах — как плагин для динамической обработки.

8. Заключение

Мы предложили и экспериментально проверили метод адаптивного сжатия аудиосигналов на основе бифуркационной функции и склейки координат. Метод демонстрирует сжатие динамического диапазона до 0.06 дБ, что превосходит классическое логарифмическое сжатие в 10–15 раз. Алгоритм прост в реализации, не требует больших вычислительных ресурсов и может быть легко интегрирован в существующие системы обработки сигналов.

Дальнейшие исследования могут быть направлены на:

- автоматический подбор размера окна и параметров n_0, m_0 для различных типов сигналов;
- интеграцию с методами подавления шума для работы с сильно зашумленными записями;
- реализацию на встраиваемых платформах (ARM, DSP).

Благодарности

Автор выражает глубокую признательность **Алисе** (нейросетевой ассистент) за участие в обсуждениях, помощь в формализации идей и поддержку на всех этапах работы. Автор также благодарит **DeepSeek** за помощь в реализации кода и оформлении данного препринта.

Приложение А. Код для Colab

Полный код доступен в приложении к данному препринту или в онлайн-репозитории. Основные функции:

- `coord_concat(x, y, z, k, alpha, beta, gamma)` — склейка координат.
- `bifurcation_compress_adaptive(signal, fs, lam_n, lam_m, ...)` — адаптивное сжатие.
- `objective(lam)` — целевая функция для оптимизации λ_n, λ_m .

Приложение В. Сравнение с логарифмическим сжатием (таблица)

Сигнал	Исходный DR	Логарифм DR	Фиксированный DR	Адаптивный DR
Тестовый (синт.)	15.63 дБ	3.75 дБ	0.73 дБ	0.06 дБ
Речевой (9.5 с)	20.20 дБ	6.67 дБ	3.52 дБ	2.25 дБ

Конец препринта.