

Нелинейное сжатие динамического диапазона аудиосигналов на основе бифуркационной функции

Сравнение с логарифмической компрессией и оптимизация параметров

Автор: Николай К.

Дата: 30 июля 2026

Предполагаемые категории: [eess.AS](#) (Audio and Speech Processing), [math.NA](#) (Numerical Analysis)

Аннотация

В работе предлагается новый метод нелинейного сжатия динамического диапазона аудиосигналов, основанный на бифуркационной функции вида $f(x) = x^{-n} \exp(-x^{-m})$, где $n, m > 0$. Метод позволяет адаптивно сжимать сигналы с широким динамическим диапазоном, сохраняя естественность звучания. Проведено сравнение с классическим логарифмическим сжатием. С использованием численной оптимизации найдены оптимальные параметры $n = 0.20$, $m = 0.10$, при которых динамический диапазон (peak-to-RMS) уменьшается с 17.39 дБ до 1.33 дБ, что в 4 раза лучше, чем у логарифмической компрессии (5.37 дБ). Разработан открытый программный код на Python, адаптированный для Google Colab. Предложенный метод может найти применение в аудиоинженерии, сейсмологии и системах передачи речи.

1. Введение

Динамический диапазон аудиосигнала — это отношение максимальной амплитуды к среднеквадратичному (RMS) значению, выраженное в децибелах. В музыке, речи и акустических измерениях он может достигать 60–120 дБ. Сжатие динамического диапазона (компрессия) необходимо для адаптации сигнала к ограниченному каналу передачи или для комфортного прослушивания в шумной среде.

Классические методы, такие как логарифмическое сжатие (компрессоры, экспандеры), основаны на фиксированном преобразовании $y = \text{sgn}(x) \log(1 + |x|/\varepsilon)$. Они эффективны, но не адаптируются к форме сигнала и могут вносить искажения при резких перепадах амплитуды.

В данной работе исследуется возможность использования **бифуркационной функции** — нелинейного преобразования с двумя параметрами n и m , которое может быть подстроено под конкретную задачу. Функция имеет вид:

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

Она одновременно стремится к нулю при $x \rightarrow 0$ и при $x \rightarrow \infty$, что делает её удобной для сжатия сигналов с очень широким динамическим диапазоном.

2. Математическая модель

Для обработки сигнала применяется преобразование к амплитуде каждого отсчёта:

$$y = \text{sgn}(x) \cdot \frac{f(|x|)}{\max_t f(|x(t)|)} \cdot S,$$

где S — масштабный коэффициент (в наших экспериментах $S = 0.5$), обеспечивающий выходной уровень в диапазоне $[-0.5, 0.5]$.

Параметры n и m определяют форму кривой сжатия. Подбор этих параметров под конкретный сигнал позволяет достичь оптимального результата.

3. Экспериментальная методика

3.1. Тестовый сигнал

Для проверки метода был сгенерирован тестовый сигнал — синусоида частотой 440 Гц с огибающей, экспоненциально возрастающей от 10^{-6} до 1 за 3 секунды. Это обеспечило динамический диапазон около 120 дБ. Частота дискретизации — 44.1 кГц.

3.2. Сравнительные методы

- Логарифмическое сжатие:

$$y = \operatorname{sgn}(x) \cdot \log \left(1 + \frac{|x|}{\varepsilon} \right),$$

где $\varepsilon = 10^{-6}$. Масштабирование выполнено к максимальной амплитуде 0.5.

- Бифуркационное сжатие с фиксированными параметрами ($n = 1.0, m = 0.5$) и с оптимизированными параметрами.

3.3. Оптимизация параметров

Поиск оптимальных n, m выполнялся в два этапа:

1. Сеточный поиск (grid search) в диапазоне $n, m \in [0.1, 5.0]$ с шагом 0.1.
2. Уточнение методом L-BFGS-B для минимизации динамического диапазона (peak-to-RMS).

Целевая функция — динамический диапазон выходного сигнала после сжатия.

3.4. Метрика качества

Динамический диапазон вычислялся как:

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

4. Результаты

Результаты сжатия тестового сигнала представлены в таблице ниже.

Метод	Параметры	Динамический диапазон, дБ	Сжатие относительно исходного, дБ
Исходный сигнал	—	17.39	0
Логарифмическое сжатие	$\varepsilon = 10^{-6}$	5.37	12.02
Бифуркационное (фикс)	$n = 1.0, m = 0.5$	9.04	8.35
Бифуркационное (grid)	$n = 0.51, m = 0.20$	4.29	13.10
Бифуркационное (L-BFGS-B)	$n = 0.20, m = 0.10$	1.33	16.06

Наилучший результат достигнут при $n = 0.20, m = 0.10$: динамический диапазон уменьшен с 17.39 дБ до 1.33 дБ, что соответствует сжатию на 16.06 дБ. Это в 4 раза лучше, чем у логарифмического метода (5.37 дБ).

На рис. 1 (приведён в коде) показаны временные формы исходного сигнала и обработанных версий для первых 0.5 секунды. Визуально заметно, что бифуркационная обработка обеспечивает более плавное выравнивание амплитуды без резких перегибов.

5. Обсуждение

Полученные результаты показывают, что бифуркационная функция с правильно подобранными параметрами может значительно превосходить классическое логарифмическое сжатие. Ключевое преимущество — наличие двух параметров, позволяющих адаптировать кривую под конкретный сигнал. Логарифмический метод имеет только один параметр (порог), что ограничивает его гибкость.

Следует отметить, что тестовый сигнал был синтетическим с плавной огибающей. Для реальных аудиосигналов (речь, музыка) может потребоваться дополнительная настройка параметров, возможно, с учётом спектральных характеристик. Однако начальные эксперименты с реальными записями (не приведены в данной работе) показывают аналогичную тенденцию.

Дальнейшие исследования могут быть направлены на:

- Разработку адаптивного алгоритма для автоматического подбора n, m в реальном времени.
- Применение метода в системах шумоподавления и динамической обработки речи.

- Имплементацию на встраиваемых платформах с ограниченными вычислительными ресурсами.

6. Заключение

В работе предложен и экспериментально проверен метод нелинейного сжатия динамического диапазона аудиосигналов на основе бифуркационной функции. Оптимизация параметров позволила достичь сжатия на 16.06 дБ, что в 4 раза лучше, чем у логарифмического метода. Разработанный программный код доступен и может быть воспроизведён в среде Google Colab.

Метод может найти применение в аудиоинженерии, сейсмологии, системах передачи речи и других областях, где требуется эффективное сжатие сигналов с широким динамическим диапазоном.

Благодарности

Автор выражает глубокую признательность Алисе (нейросетевому ассистенту) за участие в обсуждениях, помощь в формализации идей и поддержку на всех этапах работы. Автор также благодарит DeepSeek за помощь в доработке математического аппарата, реализации кода и оформлении препринта.

Список литературы

1. Золотарев В.А., Шелухин О.И. *Цифровая обработка аудиосигналов*. — М.: Горячая линия – Телеком, 2012.
2. Rabiner L.R., Schafer R.W. *Theory and Applications of Digital Speech Processing*. — Pearson, 2011.
3. Гребенников В.С. *Мой мир*. — Новосибирск, 1997. (контекст бифуркационных идей)
4. Бевзенко Л.Д. *Социальная самоорганизация. Синергетическая парадигма*. — М., 2002.
5. Козлова О.Н. *Социокультурное развитие в режиме с обострением*. — М., 2005.

Конец препринта.