

УДК 621.391.037.372

## МЕТОД ОЦЕНКИ УРОВНЯ КЛИППИРОВАНИЯ РЕЧЕВОГО СИГНАЛА

С.В. Алейник, Ю.Н. Матвеев, А.Н. Раев

Рассмотрены различные способы оценки уровня клиппирования речевого сигнала. Предлагается и исследуется новый способ оценки уровня клиппирования, обладающий лучшими характеристиками по сравнению с известными.

**Ключевые слова:** клиппирование, речевой сигнал.

## Введение

Клиппирование – один из видов искажения формы сигнала, которое происходит при перегрузке усилителя и при превышении выходным напряжением усилителя его динамического диапазона. На осциллограмме клиппирование обычно выглядит как обрезание сигнала по амплитуде. При этом различают одностороннее (обрезание «только сверху» или «только снизу») и двустороннее клиппирование. В цифровом сигнале клиппирование проявляется тем, что отсчеты сигнала группируются около его максимального и минимального значений («мягкое» клиппирование), либо просто равны соответствующим максимальным и минимальным значениям («жесткое» клиппирование). Математически процесс двустороннего жесткого клиппирования дискретного сигнала  $x(k)$  можно записать следующим образом [1]:

$$x_{cl}(k) = \begin{cases} x(k), & \text{if } |x(k)| < A, \\ Ax(k)/|x(k)|, & \text{otherwise,} \end{cases}$$

где  $k$  – временной индекс;  $x_{cl}(k)$  – клиппированный сигнал;  $||$  – операция взятия модуля;  $A$  – порог клиппирования.

Типичный вид клиппированного речевого сигнала представлен на рис. 1.

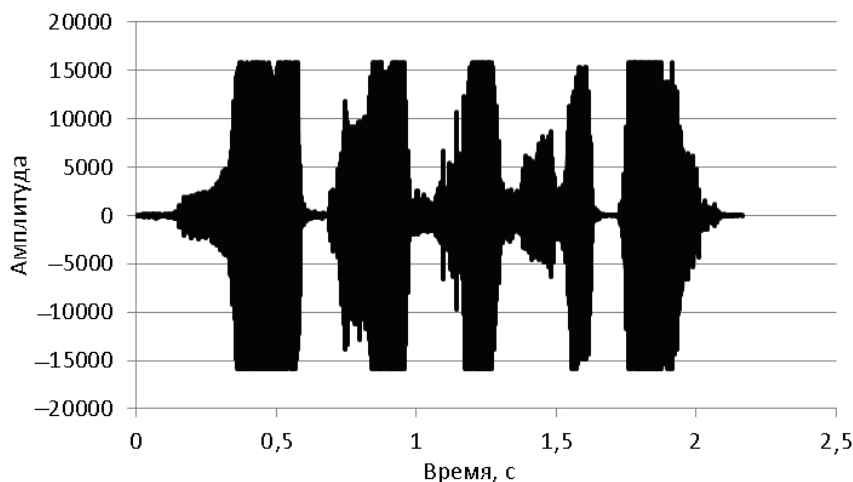


Рис. 1. Клиппированный речевой сигнал для  $A=16000$

Клиппирование приводит к повышению уровня высокочастотных составляющих сигнала, к появлению гармоник высших порядков, что приводит к ухудшению качества звука, и, соответственно, к ухудшению результатов дальнейшей обработки, например, качества распознавания речи или голоса. Таким образом, задача оценки уровня клиппирования акустического сигнала (с целью, например, дальнейшей отбраковки сильно искаженных участков) представляется достаточно актуальной.

Если известен порог клиппирования  $A$  и мощность сигнала  $P_{sig}$ , то можно вычислить так называемое «отношение клиппирования» (clipping ratio, CR), служащее характеристикой того, насколько сильно клиппирован сигнал [1, 2]:

$$CR_{lin} = A / \sqrt{P_{sig}}, \quad (1)$$

$$CR_{db} = 10 \lg(A^2 / P_{sig}). \quad (2)$$

В реальной жизни, однако, чаще всего порог клиппирования неизвестен, а в формулах (1) или (2) вместо теоретических приходится использовать экспериментальные оценки величин  $A$  и  $P_{sig}$ , что приводит к высокой дисперсии рассчитанной  $CR$ .

Часто работы посвящены оценке клиппирования узкого круга известных (например, OFDM [1, 2]) сигналов, и соответственно при оценке требуемых характеристик используются присущие данным сигналам специфические свойства. Некоторые алгоритмы используют знание исходного (неклиппированного) сигнала [3] и, следовательно, ориентированы скорее на оценку качества устройства обработки (усилителя и пр.), а не самого сигнала.

Целью предлагаемого авторами исследования является разработка метода оценки уровня клиппирования речевого сигнала в случае, когда исходный неискаженный сигнал неизвестен, а параметры анализируемого сигнала (частота дискретизации, мощность, способ кодирования и т.п.) варьируются в широких пределах.

### Существующие методы оценки уровня клиппирования неизвестного сигнала

Подход, когда исходный сигнал неизвестен, исследовался в работах [4, 5]. В работе [4] для оценки уровня клиппирования предлагается использовать «взвешенное дифференцирование». Действительно, если на клиппированном участке соседние отсчеты сигнала  $x(k)$  и  $x(k-1)$  равны или близки по амплитуде, то величина  $d(k) = x(k) - x(k-1)$  будет равна или близка к нулю. Соответственно, усредненное абсолютное значение  $D = \langle |d(k)| \rangle$  может служить индикатором уровня клиппирования – чем сильнее клиппирован сигнал, тем ближе к нулю величина  $D$ . Данный метод хорошо работает на медленно меняющихся неосциллирующих сигналах, т.е. тех, в которых присутствуют относительно длинные участки, для которых значение  $x(k)$  равно или близко к значению  $x(k-1)$ . Именно такой вариант и рассмотрен в работе [4]. К сожалению, вышеуказанное условие не выполняется для речевого сигнала. Как звонкие, так и глухие звуки в речевом сигнале содержат быстро меняющиеся компоненты, поэтому даже на участках сильного клиппирования соседние отсчеты могут сильно отличаться друг от друга. Таким образом, величина  $D$  уже не близка к нулю даже на сильно клиппированном сигнале, и, следовательно, точность оценки уровня клиппирования описанным методом невысока.

В работе [5] для оценки уровня клиппирования используется, на наш взгляд, более подходящий, гистограммный подход, т.е. строится и анализируется гистограмма обрабатываемого сигнала. Известно, что плотность распределения амплитуд неклиппированного речевого сигнала может быть достаточно хорошо аппроксимирована симметричными распределениями, такими как гамма-распределение или распределение Лапласа [6]. Общий вид данных распределений – одномодовые, с плавно спадающими хвостами. Совсем другая картина наблюдается в случае клиппированного сигнала (рис. 2).

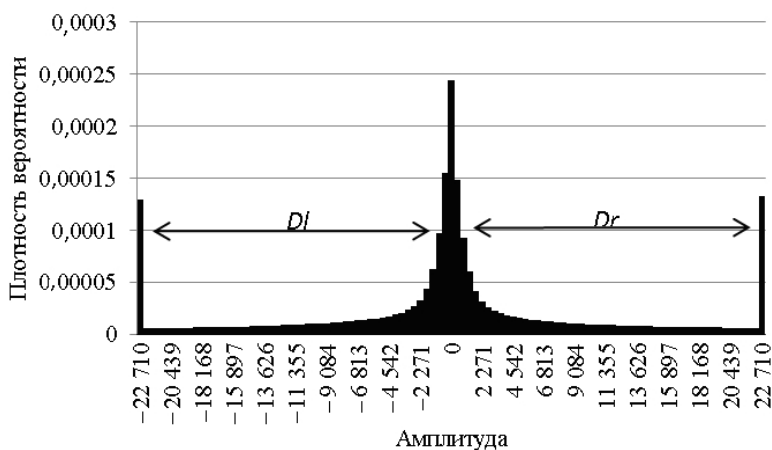


Рис. 2. Гистограмма значений амплитуды клиппированного речевого сигнала

Если центральная часть гистограммы действительно напоминает распределение Лапласа (или симметричное гамма-распределение), то, в отличие от неклиппированного сигнала, в данном случае имеются резкие «всплески» на хвостах, вызванные концентрацией отсчетов у максимального и минимального значений динамического диапазона клиппированного сигнала.

В работе [5] предлагаются два способа гистограммной оценки уровня клиппирования. Первый базируется на расчете степени отклонения  $d(a)$  рассчитанной нормализованной гистограммы  $h_{calc}(a)$  от некоей заранее известной «базовой» плотности распределения  $h_{base}(a)$ :

$$d(a) = h_{calc}(a) - h_{base}(a),$$

где  $a$  – абсцисса гистограмм (амплитуда сигнала). Для расчета окончательного уровня клиппирования используется либо мгновенное значение  $d(a)$ , либо усредненное по выбираемым участкам гистограммы

(исключается центральная часть), но во всех случаях необходимо знание базовой плотности и расчет отклонения. Второй способ основан на подсчете количества локальных максимумов на хвостах рассчитанной гистограммы  $h_{calc}(a)$ . Оба способа имеют существенные недостатки. Во-первых, получаемая оценка уровня клиппирования имеет низкую точность на ограниченном наборе данных. Так, например, при покадровой обработке речевого сигнала с частотой дискретизации 8 кГц, длиной кадра 0,5 с и количеством отсчетов гистограммы 100 количество отсчетов сигнала для построения гистограммы равно 4000. Такой статистики недостаточно для достоверной оценки реальной функции плотности распределения амплитуды сигнала, поэтому подсчет количества максимумов часто дает неверные результаты, а рассчитываемая степень отклонения  $d(a)$  имеет завышенное значение. Во-вторых, в случае кодированного (по А- или  $\mu$ -закону) речевого сигнала его гистограмма по определению имеет гребенчатый характер, что делает оценку уровня клиппирования подобного сигнала обоими способами совершенно неверной.

### Предлагаемый метод оценки уровня клиппирования речевого сигнала

В предлагаемом подходе также применяется гистограммный метод, но, в отличие от работы [5], для повышения надежности оценки не используется ни базовая гистограмма, ни амплитудные значения расчетной гистограммы.

Предлагаемый метод основан на следующих соображениях:

- если речевой сигнал не клиппирован, то хвосты гистограммы плавно повышаются от правой и левой границ гистограммы к ее моде. Естественные флуктуации значений гистограммы при этом незначительные, а расстояния (по оси «Амплитуда») между соседними локальными максимумами невелики;
- если речевой сигнал клиппирован, то на хвостах гистограммы наблюдаются выраженные максимумы, и расстояния (по оси «Амплитуда») от этих максимумов до соседних максимумов, превышающих их по величине (расстояния  $Dl$  и  $Dr$  на рис. 2), значительны.

Следовательно, величина  $\max(Dl, Dr)$ , нормированная на общий размах гистограммы, может служить оценкой уровня (коэффициента) клиппирования сигнала.

Приведем алгоритм расчета предлагаемого коэффициента.

1. Рассчитываем гистограмму амплитуд речевого сигнала.
2. Определяем левый ( $Min\_Index$ ) и правый ( $Max\_Index$ ) индексы ненулевых отсчетов гистограммы.
3. Принимаем:
  - левый индекс  $i=Min\_Index$ ;
  - правый индекс  $j=Max\_Index$ ;
  - максимальная дистанция  $Dist\_Max = 0$ .
4. Цикл: синхронно увеличивая  $i$  и уменьшая  $j$ , находим локальные максимумы в гистограмме и вычисляем расстояния  $Dl$  и  $Dr$  (рис. 2).
5. Находим максимум:  $Dist\_Max = \max(Dist\_Max, Dl, Dr)$ .
6. Если  $i < j$ , то возвращаемся к п. 4, если нет, то выходим из цикла и рассчитываем коэффициент клиппирования:  
$$R = 2Dist\_Max / (Max\_Index - Min\_Index).$$

Следует сразу заметить, что в случае сильно клиппированного сигнала локальные максимумы на хвостах гистограммы оказываются по величине больше центрального максимума. Тогда  $Dl = Dr = 0,5(Max\_Index - Min\_Index)$ , и, следовательно,  $R = 1$ .

Выделим свойства предложенного коэффициента клиппирования  $R$ :

1.  $0 \leq R \leq 1$ ;
2. инвариантен к частоте дискретизации сигнала;
3. инвариантен к амплитуде и амплитудному сдвигу сигнала (зависимость устраняется при нормировке в процессе построения гистограммы);
4. слабо чувствителен к точности оценки гистограммы;
5. слабо чувствителен к кодированию речевого сигнала по А- и  $\mu$ -законам;
6. в случае сильно нестационарного сигнала вида «почти весь кадр – нулевой сигнал и один короткий резкий всплеск на малом количестве отсчетов» возможно получение завышенной оценки уровня клиппирования;
7. простые сигналы, например, гармонический или прямоугольный, имеют  $R = 1$ .

### Экспериментальные исследования предлагаемого коэффициента

Результаты экспериментов по определению плотности распределения коэффициента  $R$  для различных уровней клиппирования речевого сигнала приведены на рис. 3.

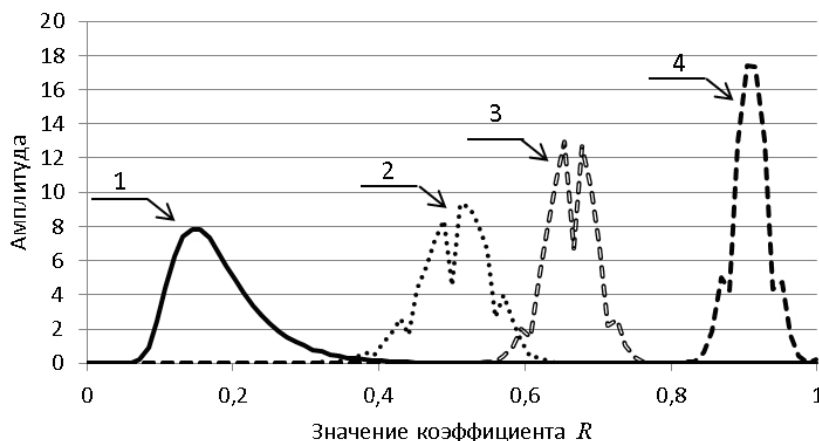


Рис. 3. Плотность распределения коэффициента клиппирования  $R$ : (1) – неклиппированный речевой сигнал; (2), (3) и (4) – клиппированный сигнал на уровне 10%, 30% и 60% от максимума соответственно

На рис. 3 кривая 1 – неклиппированный речевой сигнал; кривые 2, 3 и 4 – клиппированный сигнал на уровне 10%, 30% и 60% от максимума соответственно. Для оценки гистограммы использовались кадры сигнала длительностью в 4000 отсчетов с количеством отсчетов в гистограмме 200. Из рис. 3 видно, что коэффициент клиппирования  $R$  достаточно точно отражает уровень клиппирования сигнала.

Совмещенные осциллограммы речевого сигнала с участками с различным уровнем клиппирования и коэффициента  $R$  приведены на рис. 4 (шкала коэффициента – справа). Длина кадра в данном эксперименте соответствовала 1 с для частоты дискретизации 11025 Гц. Очевидно, что коэффициент  $R$  позволяет достаточно точно осуществлять «мгновенную» оценку уровня клиппирования речевого сигнала на коротких интервалах.

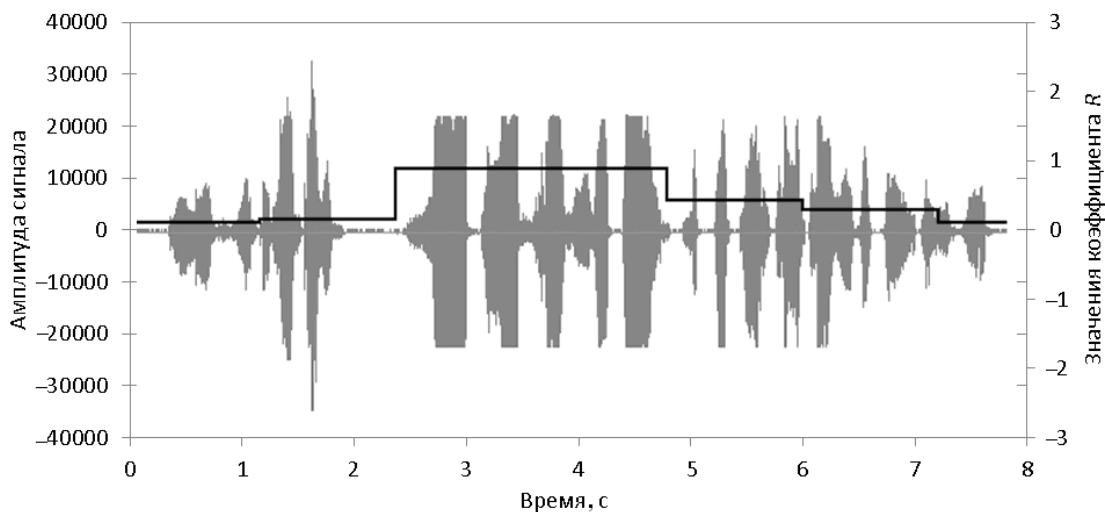


Рис. 4. Осциллограммы речевого сигнала (серый) и соответствующего коэффициента  $R$  (черный); шкала значений коэффициента – справа

### Заключение

Теоретические исследования и проведенные эксперименты показали большую точность предложенного метода оценки уровня клиппирования речевого сигнала по сравнению с известными. Предложенный метод тестировался на различных сигналах, с различной частотой дискретизации, без использования и с использованием кодирования речевого сигнала по А- и  $\mu$ -законам, и продемонстрировал хорошую точность при длительностях кадров анализа 4000 отсчетов и более, при числе отсчетов в гистограмме от 200 до 500. Данный метод может использоваться для оценки качества речевого сигнала в реальном масштабе времени с целью, например, отбраковки фрагментов сигнала низкого качества.

Предложенный алгоритм не требует больших затрат памяти и вычислительных ресурсов и может быть реализован в системах реального времени на базе сверхбольших интегральных схем, программируемых логических интегральных схем и т.п. [7].

Повышение робастности метода при изменении параметров вычисления алгоритма является предметом дальнейших исследований.

### Литература

1. Chen H., Haimovich A.M. Iterative estimation and cancellation of clipping noise for OFDM signals // IEEE Commun. Lett. – 2003. – V. 7. – № 7. – P. 305–307.
2. Zhidkov S.V. Detection of clipped code-division multiplexed signals // Electronics Letters. – 2005. – V. 41. – № 25. – P. 1383–1384.
3. Kim J. Method and apparatus for evaluating audio distortion. – US Patent 005402495, Int.Cl. H04B 15/00, 1995.
4. Riemer T.E., Weiss M.S., Losh M.W. Discrete Clipping Detection by Use of a Signal Matched Exponentially Weighted Differentiator // Proceedings of the IEEE Southeastcon'90. – USA: New Orleans, Louisiana, 1990. – P. 245–248.
5. Otani T., Tanaka M., Ota Y., Ito S. Clipping detection device and method. – US Patent 20100030555 A1, Int.Cl. G10L 21/02, 2010.
6. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов: Пер. с англ. / Под ред. Ю.Н. Прохорова, М.В. Назарова. – М: Радио и связь, 1981. – 496 с.
7. Чураев С.О., Адамова А.Д., Палташев Т.Т. Реализация алгоритма шумоподавления в речевом тракте систем мобильной связи на базе СБИС // Научно-технический вестник СПбГУ ИТМО. – 2011. – № 1 (77). – С. 72–76.

*Алейник Сергей Владимирович* – ООО «ЦРТ-Инновации», научный сотрудник, aleinik@speechpro.com  
*Матвеев Юрий Николаевич* – ООО «ЦРТ-Инновации», доктор технических наук, главный научный сотрудник, НИУ ИТМО, профессор, matveev@speechpro.com  
*Раев Андрей Николаевич* – ООО «ЦРТ», директор научно-исследовательского департамента, raev@speechpro.com