

УДК 621.391.037.372

ОПРЕДЕЛЕНИЕ КЛИППИРОВАННЫХ ФРАГМЕНТОВ В АКУСТИЧЕСКИХ СИГНАЛАХ

С.В. Алейник^a, Ю.Н. Матвеев^{a, b}, А.В. Шолохов^c^a ООО «ЦРТ-Инновации», Санкт-Петербург, Россия^b Университет ИТМО, Санкт-Петербург, Россия, matveev@mail.ifmo.ru^c Университет Восточной Финляндии, Йоенсуу

Аннотация. Исследован способ определения клиппированных участков в акустических сигналах, обладающий лучшими характеристиками по сравнению с другими известными способами. Данный способ основан на построении гистограммы амплитуд анализируемого сигнала и вычислении расстояний между локальными максимумами гистограммы на ее хвостах и в центральной части. Отличие гистограмм неклиппированного и клиппированного сигналов заключается в том, что гистограмма неклиппированного сигнала имеет плавно спадающие хвосты, в то время как гистограмма клиппированного сигнала имеет на хвостах заметные и легко обнаруживаемые всплески. Величина данных всплесков и качество детектирования клиппированных фрагментов соответственно зависят от параметров исследуемого способа. Основной целью работы является нахождение оптимальных параметров исследуемого способа. Путем математического моделирования детально исследованы характеристики способа: построены плотности распределения целевой величины для различных длин анализируемого кадра сигнала, количества отсчетов в гистограмме и уровней клиппирования акустических сигналов. Показано, что при длине кадра в 6000–8000 отсчетов и количестве отсчетов в гистограмме, равном 200–300, достигается хорошее различие клиппированных и неклиппированных участков акустического сигнала. При этом порог разделения может варьироваться в пределах 0,45–0,55. Приведены примеры работы детектора клиппирования, основанного на исследованном способе, на реальных акустических сигналах при различных уровнях клиппирования.

Ключевые слова: акустический сигнал, клиппирование, коэффициент клиппирования.

Благодарности. Работа выполнена при государственной финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01).

DETECTION OF CLIPPED FRAGMENTS IN ACOUSTIC SIGNALS

S.V. Aleinik^a, Yu.N. Matveev^{a, b}, A.V. Sholokhov^c^a STC-Innovation, Ltd., Saint Petersburg, Russia^b ITMO University, Saint Petersburg, Russia, matveev@mail.ifmo.ru^c University of Eastern Finland, Joensuu, Finland

The paper deals with investigation of the method for detecting clipped fragments in acoustic signals with better characteristics as compared with the other known methods. This method is based on the histogram construction for the analyzed signal amplitudes and calculating the distances between the local peaks of the histogram on its tails and in the central part. The difference between histograms of non-clipped and clipped signals is that the histogram of a non-clipped signal has smoothly decaying tails while the histogram of a clipped signal has visible and easily detectable outbursts on its tails. The value of these outbursts and consequently the quality of detection of clipped fragments depends on the parameters of the method under investigation. The main aim of this paper is finding the optimal parameters of the method. Characteristics of the method are studied in detail by mathematical modeling; density functions of target values for different lengths of a studied signal frame and the number of histogram counts and levels of clipping of acoustic signals are built. It is shown that good separation between clipped and non-clipped signal fragments of acoustic signals can be achieved for the frame length between 6000 and 8000 samples and the number of histogram bins between 200 and 300. In this case the threshold level for the best separation can vary between 0.45–0.55. Examples of clipping detector operation based on the proposed method and on real acoustic signals are shown for the case of different clipping levels.

Keywords: acoustic signal, clipping, clipping coefficient.

Acknowledgements. This work was partially financially supported by the Government of the Russian Federation, Grant 074-U01.

Введение

Клиппирование – один из видов искажения формы сигнала [1]. На осциллограмме клиппирование проявляется как обрезание сигнала по амплитуде [2]. При этом возможно одностороннее (обрезание «только сверху» или «только снизу») и двустороннее клиппирование. В цифровом сигнале клиппирование выглядит как группировка отсчетов сигнала около его максимального и минимального значений (мягкое клиппирование), либо они просто равны соответствующим максимальным и минимальным значениям (жесткое клиппирование) [2]. Математически процесс двустороннего жесткого клиппирования дискретного сигнала $x(k)$ можно записать следующим образом [2, 3]:

$$x_{cl}(k) = \begin{cases} x(k), & \text{if: } |x(k)| < A, \\ A * (x(k) / |x(k)|) & \text{otherwise,} \end{cases} \quad (1)$$

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

В случае мягкого клиппирования формула (1) неверна. Мягкое клиппирование может быть описано как воздействие на сигнал некоей нелинейной функции, например, сигмоидной:

$$x_{cl}(k) = A \left(\frac{2}{1 + e^{\beta x(k)}} - 1 \right), \quad (2)$$

где A и β – параметры функции: A – максимально возможная амплитуда; β – крутизна, характеризующая степень жесткости (чем больше β , тем ближе клиппирование к жесткому).

Например, типичный вид исходного неклиппированного и клиппированного (мягкое клиппирование) речевых сигналов для $A=20000$ и $\beta=0,00045$ представлен на рис. 1. Видно, что в случае мягкого клиппирования ограничение сигнала происходит более плавно, чем при жестком – нет резких изломов графика сигнала в районе максимальных и минимальных значений сигнала.

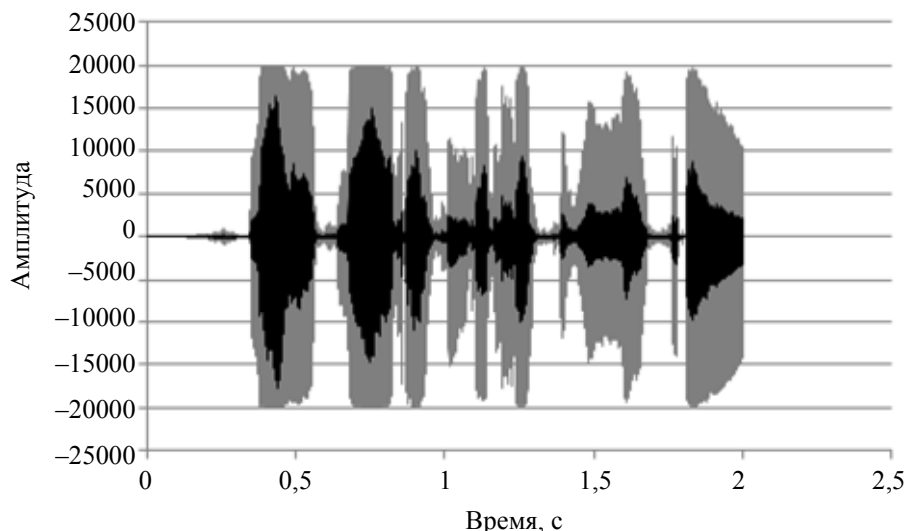


Рис. 1. Неклиппированный (черный цвет) и клиппированный (серый цвет) речевые сигналы для $A=20000$ и $\beta=0,00045$, мягкое клиппирование в соответствии с (2)

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]:

$$CR_{lin} = A / \sqrt{P_{sig}}, \quad (3)$$

$$CR_{dB} = 10 \log_{10}(A^2 / P_{sig}). \quad (4)$$

Однако в [2] отмечено, что в реальной жизни чаще всего порог клиппирования неизвестен, что делает невозможным использование формул (3) и (4) для оценки уровня клиппирования анализируемого сигнала.

Большинство исследований посвящено оценке уровня клиппирования узкого круга известных (например, OFDM [3–5], сейсмических [6], и пр.) сигналов; соответственно, при оценке требуемых характеристик используются присущие данным сигналам специфические свойства [2]. Некоторые алгоритмы используют знание исходного (неклиппированного) сигнала [7–9] и, следовательно, ориентированы скорее на оценку качества устройства обработки (усилителя и пр.), а не самого сигнала.

В работе [2] был исследован новый метод 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. Целью предлагаемой работы является дальнейшее детальное исследование метода [2], определение границ его применимости, а также определение оптимальных параметров метода.

Гистограммный метод оценки уровня клиппирования неизвестного сигнала

Гистограммный метод оценки уровня клиппирования неизвестного сигнала исследован в работах [2, 9, 10]. В данном методе 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 [11, 12]. Общий вид этих распределений – одномодовые, с плавно спадающими хвостами. В работе [2] приведена гистограмма жестко клиппированного сигнала, имеющая резкие всплески на хвостах. На рис. 2 приведена гистограмма клиппированного сигнала в случае мягкого клиппирования.

Видно, что в случае мягкого клиппирования гистограмма сигнала также имеет всплески на хвостах, но в данном случае всплески более «размазаны». Это свойство характерно не только для речевого, но для любых клиппированных сигналов.

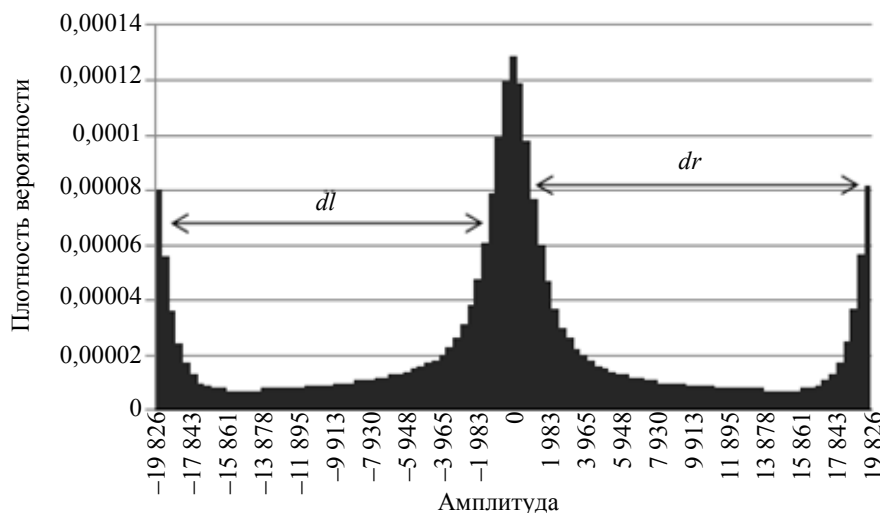


Рис. 2. Гистограмма значений амплитуды клиппированного речевого сигнала (мягкое клиппирование). Величины dl и dr – расстояния между локальными максимумами на хвостах и в области медианы гистограммы (см. Алгоритм 1)

Наши исследования [2] показали, что амплитуда всплесков и их количество варьируются в широких пределах, поэтому методы, основанные на замере данных показателей [9, 10], дают неустойчивые результаты.

Предложенный в [2] метод основан на вычислении и анализе расположения максимумов на оси X , что оказывается более робастным к типу и величине клиппирования, чем анализ амплитуд гистограммы.

Модифицированный алгоритм выглядит следующим образом.

Алгоритм 1. Вычисление коэффициента клиппирования.

- Вычислить гистограмму сигнала: $H(k)$, $k = 0, K-1$.
- Найти самый левый kl и самый правый kr индекс по оси X , где гистограмма не равна нулю.
- Вычислить $Denom = kr - kl$.
- Установить $yl_0 = H(kl)$; $yr_0 = H(kr)$; $dl = dr = 0$; $Dmax = 0$.
- Цикл: Пока ($kr > kl$) выполнять:
 - увеличить: $kl = kl + 1$;
 - уменьшить: $kr = kr - 1$;
 - вычислить:
 - Если $H(kl) \leq yl_0$, то: $\{ dl = dl + 1; \}$
 - Иначе: $\{ yl_0 = H(kl); dl = 0; \}$
 - Если $H(kr) \leq yr_0$, то: $\{ dr = dr + 1; \}$
 - Иначе: $\{ yr_0 = H(kr); dr = 0; \}$
 - $Dmax = \max\{Dmax, dl, dr\}$.

- Вычислить коэффициент клиппирования: $Rcl = 2Dmax / Denom$.

Гистограмма рассчитывается по следующему алгоритму.

Алгоритм 2. Вычисление гистограммы.

Пусть $x(n)$, $n = 0, N-1$ – дискретный временной сигнал, K – количество отсчетов в гистограмме.

Тогда:

- для всех $n = 0, N-1$, найти минимальное $xmin$ и максимальное $xmax$ значения амплитуды сигнала;
- установить нулевые значения всех отсчетов гистограммы: $H(k) = 0$, $k = 0, K-1$;
- для всех $n = 0, N-1$ выполнить:
 - вычислить значение: $y = (x(n) - xmin) / (xmax - xmin)$;
 - вычислить индекс отсчета гистограммы: $k = (\text{int})\{Ky\}$;
 - увеличить значение отсчета гистограммы:
 - Если $k < N$, то: $H(k) = H(k) + 1$;
 - Иначе: $H(k-1) = H(k-1) + 1$.

Следует заметить, что в случае сильно клиппированного сигнала локальные максимумы на хвостах гистограммы оказываются больше центрального максимума. Тогда $dl = dr = 0,5 * (Max_Index - Min_Index)$, и, следовательно, $R = 1$.

Теоретические свойства коэффициента клиппирования R ($0 \leq R \leq 1$) подробно описаны в [2]. В настоящей работе авторы сконцентрировались на экспериментальных исследованиях поведения коэффициента клиппирования на различных сигналах и при различных параметрах.

Экспериментальные исследования предлагаемого коэффициента

Авторами проведены детальные исследования плотностей распределения коэффициента R для разных уровней клиппирования сигнала и различных параметров алгоритма, а именно, варьировались длина кадра N и количество отсчетов в гистограмме K . Результаты экспериментов по определению плотности распределения коэффициента R для различного количества K отсчетов в гистограмме приведены на рис. 3 (неклиппированный сигнал) и рис. 4 (клиппированный 50% сигнал).

Объединенные кривые для клиппированного и неклиппированного сигналов изображены на рис. 5. Анализ рис. 5 показывает, что установка порога обнаружения в пределах 0,45–0,55 позволяет достаточно хорошо разделять клиппированный и неклиппированный сигналы.

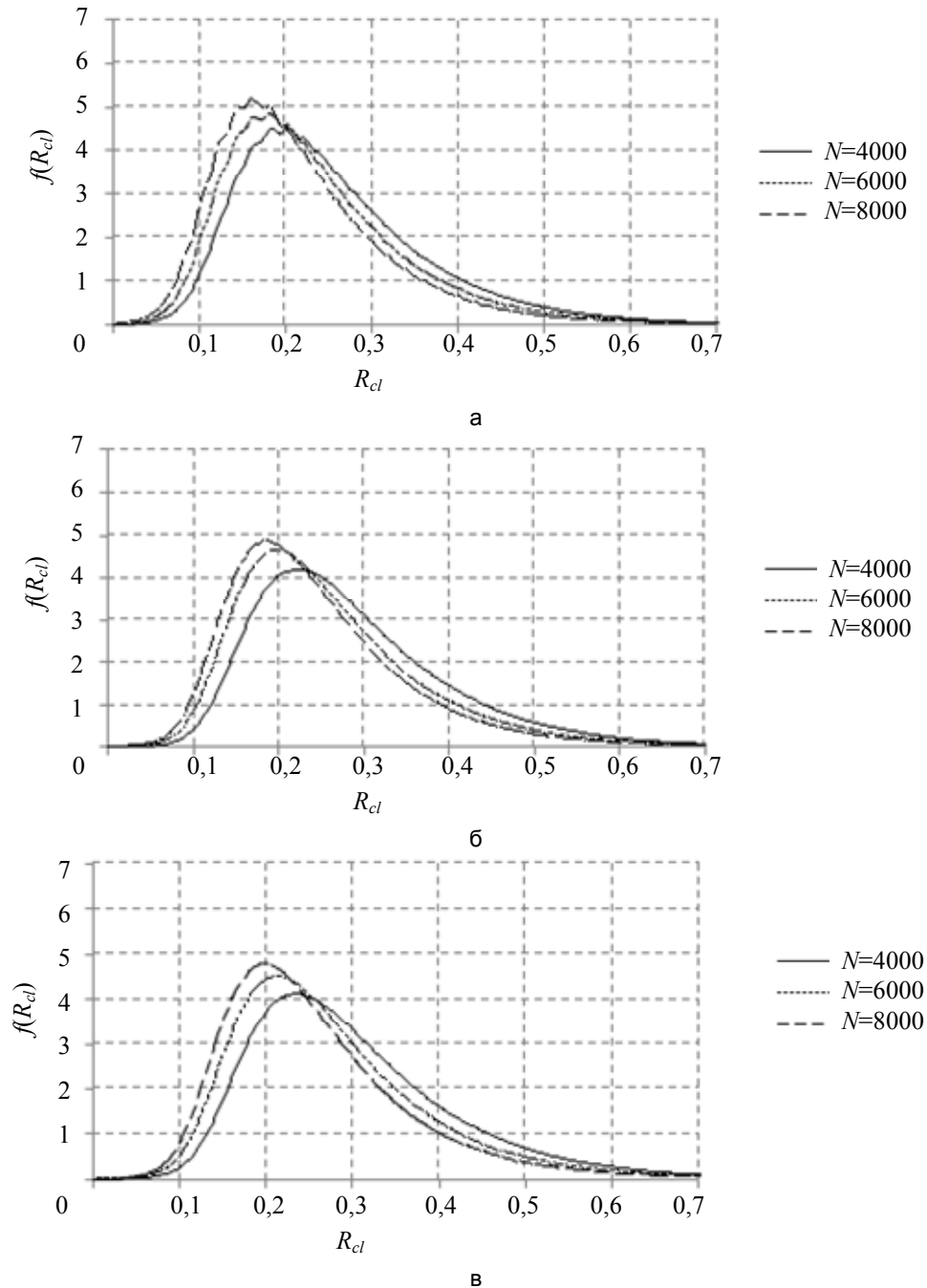


Рис. 3. Плотность распределения f коэффициента клиппирования R_{cl} для различного количества отсчетов в гистограмме K для неклиппированного речевого сигнала: $K = 101$ (а); $K = 201$ (б); $K = 301$ (в)

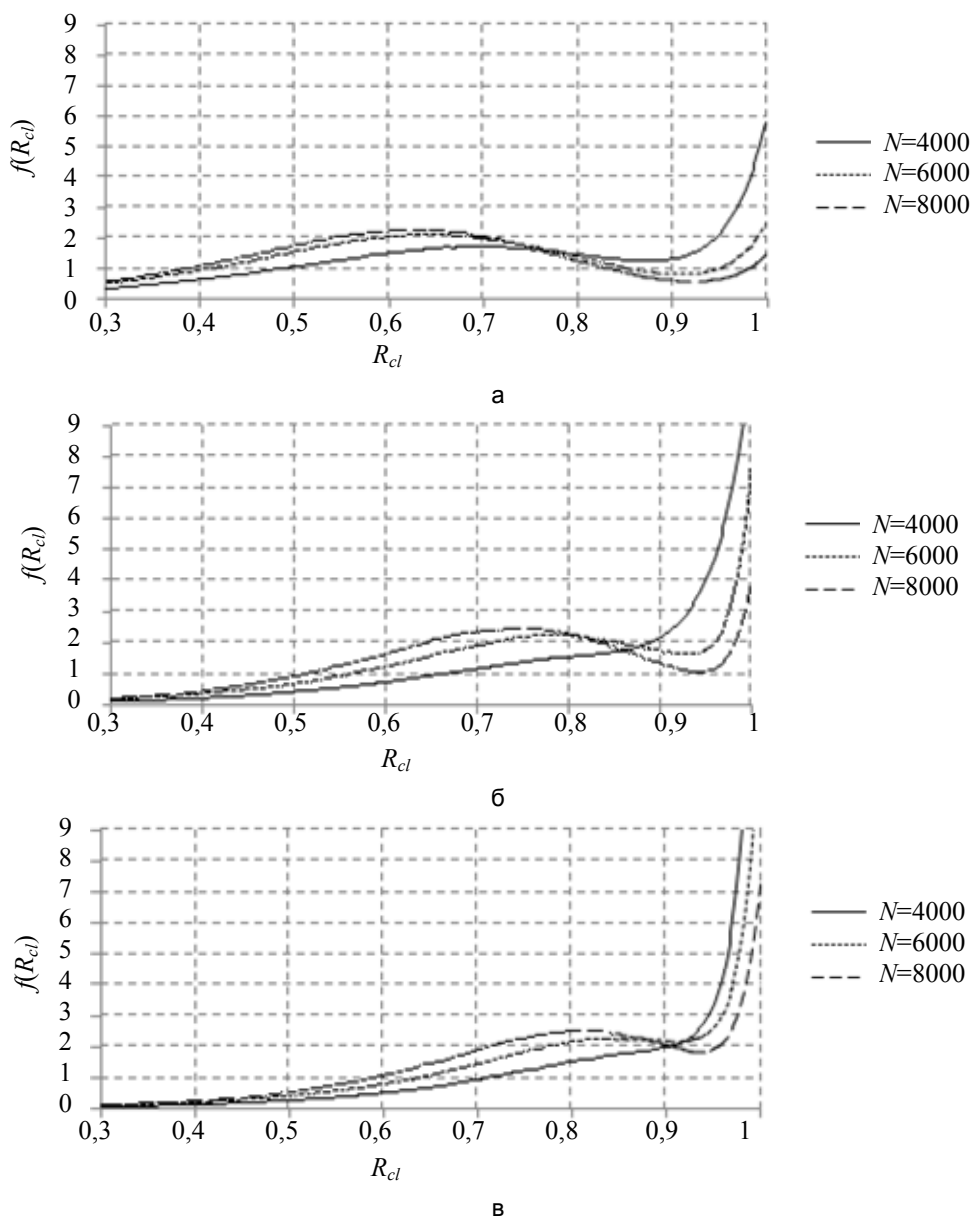


Рис. 4. Плотность распределения f коэффициента клиппирования R_{cl} для различного количества отсчетов в гистограмме K для клиппированного (50%) речевого сигнала: $K = 101$ (а); $K = 201$ (б); $K = 301$ (в)

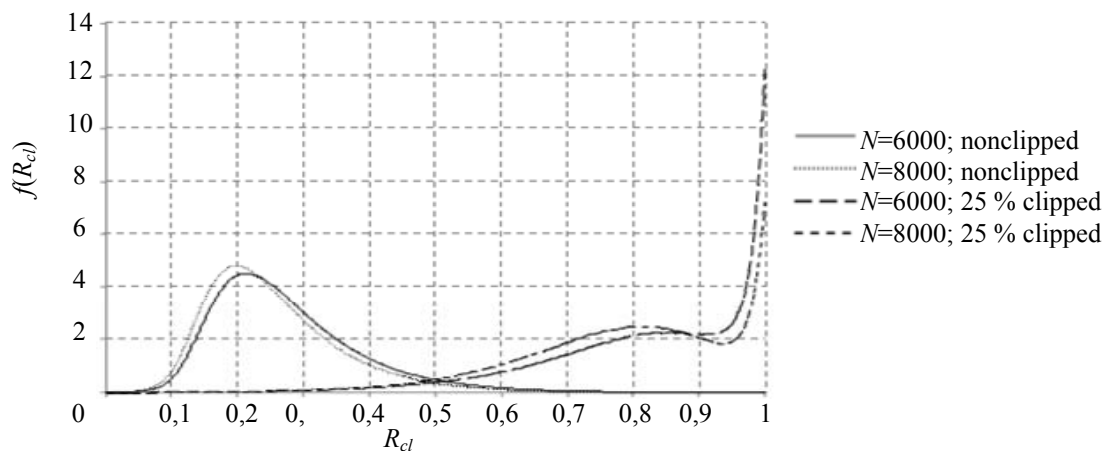


Рис. 5. Плотность распределения коэффициента клиппирования R для $K=31$ в случаях неклиппированного и клиппированного (25%) речевых сигналов

Примеры обработки речевых сигналов

Примеры обработки реальных речевых сигналов изображены на рис. 6–8. Из этих рисунков видно, что порог, равный 0,55, позволяет полностью разделить неклипированные и клипированные участки сигнала.

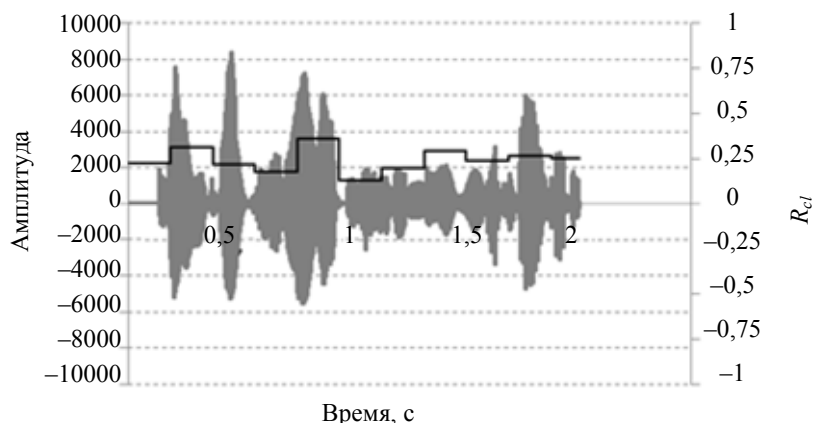


Рис. 6. Осциллограммы речевого сигнала (серый цвет) и соответствующего коэффициента R (черный цвет, шкала значений коэффициента – справа); неклипированный сигнал

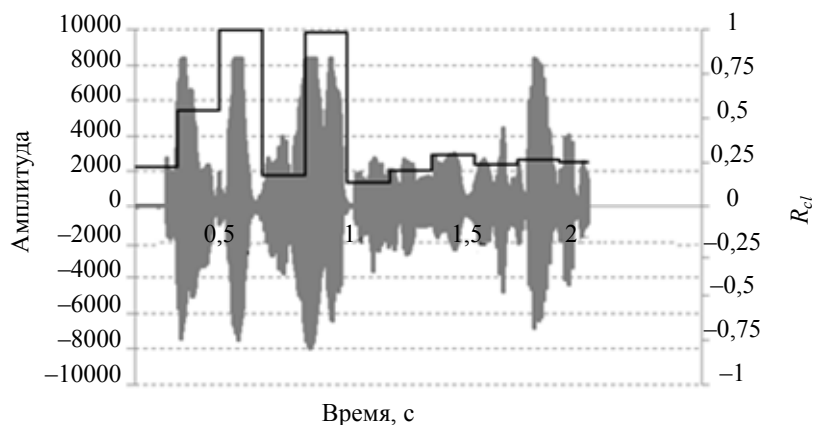


Рис. 7. Осциллограммы речевого сигнала (серый цвет) и соответствующего коэффициента R (черный цвет, шкала значений коэффициента – справа); коэффициент клипирования 25%

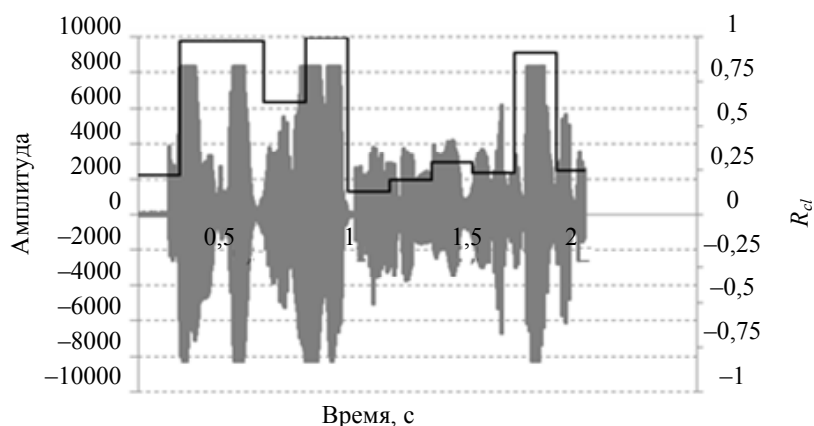


Рис. 8. Осциллограммы речевого сигнала (серый цвет) и соответствующего коэффициента R (черный цвет, шкала значений коэффициента – справа); коэффициент клипирования 50%

Заключение

Проведенные теоретические и экспериментальные исследования показали работоспособность предложенного метода детектирования клипированных фрагментов акустического сигнала и позволили определить его оптимальные параметры. Предложенный метод детектирования клипированных участ-

ков успешно используется в системах верификации и идентификации личностей по голосу, описанных в работах [13–16]. Следует, однако, указать, что описанный метод имеет определенные ограничения.

1. Резкое увеличение коэффициента клиппирования до величин, близких к единице, в случае, когда более 50% сигнала клиппировано, делает его более подходящим для задачи обнаружения факта клиппирования, чем для задачи измерения уровня клиппирования.
2. Необходимо помнить, что на простых гармонических сигналах метод дает значение, близкое к 1, даже при отсутствии клиппирования.
3. Исследования показали, что метод может давать завышенные оценки в случае, когда анализируемый кадр речевого сигнала содержит значительное количество нулевых отсчетов.

Литература

1. Алейник С.В., Симончик К.К. Алгоритмы выделения типовых помех и искажений в речевых сигналах // Изв. вузов. Приборостроение. 2013. Т. 56. № 2. С. 18–24.
2. Алейник С.В., Матвеев Ю.Н., Раев А.Н. Метод оценки уровня клиппирования речевых сигналов // Научно-технический вестник информационных технологий, механики и оптики. 2012. № 3 (79). С. 79–83.
3. Chen H., Haimovich A.M. An iterative method to restore the performance of clipped and filtered OFDM signals // IEEE International Conference on Communications. 2003. V. 5. P. 3438–3442.
4. Zhidkov S.V. Detection of clipped code-division multiplexed signals // Electronics Letters. 2005. V. 41. N 25. P. 33–34.
5. Zillmann P., Rave W., Fettweis G. Soft detection and decoding of clipped and filtered COFDM signals // Proc. IEEE Vehicular Technology Conference. 2007. P. 1598–1602.
6. Yang W., Ben-Zion Y. An algorithm for detecting clipped waveforms and suggested correction procedures // Seismological Research Letters. 2010. V. 81. N 1. P. 53–62.
7. Kim J. Method and apparatus for evaluating audio distortion. Patent US 5402495, 1995.
8. Riemer T.E., Weiss M.S., Losh M.W. Discrete clipping detection by use of a signal matched exponentially weighted differentiator // Proc. IEEE Southeastcon. New Orleans, USA, 1990. P. 245–248.
9. Otani T., Tanaka M., Ota Y., Ito S. Clipping detection device and method. Patent US 20100030555, 2010.
10. Liu X., Jia J., Cai L. SNR estimation for clipped audio based amplitude distribution // Proc. International Conference on Natural Computation (ICNC). Shenyang, China, 2013. P. 1434–1438.
11. Rabiner L.R., Schafer R.W. Introduction to Digital Speech Processing. Hanover, NOWPress, 2007. 194 p.
12. Матвеев Ю.Н. Оценка доверительного интервала общего решения ансамбля классификаторов // Изв. вузов. Приборостроение. 2013. Т. 56. № 2. С. 74–79.
13. Матвеев Ю.Н., Симончик К.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2010 // Труды 20 Международной конференции по компьютерной графике и зрению ГрафиКон'2010. Санкт-Петербург, 2010. С. 315–319.
14. Белых И.Н., Капустин А.И., Козлов А.В., Лоханова А.И., Матвеев Ю.Н., Пеховский Т.С., Симончик К.К., Шулипа А.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2010 // Информатика и ее применение. 2012. Т. 6. № 1. С. 91–98.
15. Козлов А.В., Кудашев О.Ю., Матвеев Ю.Н., Пеховский Т.С., Симончик К.К., Шулипа А.К. Система идентификации дикторов по голосу для конкурса NIST SRE 2012 // Труды СПИИРАН. 2013. № 2 (25). С. 350–370.
16. Kozlov A., Kudashev O., Matveev Y., Pekhovsky T., Simonchik K., Shulipa A. SVID speaker recognition system for the NIST SRE 2012 // Lecture Notes in Computer Science. 2013. V. 8113 LNAI. P. 278–285.

- Алейник Сергей Владимирович** – научный сотрудник, ООО «ЦРТ-Инновации», Санкт-Петербург, Россия, aleinik@speechpro.com
- Матвеев Юрий Николаевич** – доктор технических наук, профессор, Университет ИТМО, Санкт-Петербург, Россия; главный научный сотрудник, ООО «ЦРТ-инновации», Санкт-Петербург, Россия, matveev@mail.ifmo.ru
- Шолохов Алексей Владимирович** – аспирант, Университет Восточной Финляндии, Йоенсуу, Финляндия, sholokhovalexey@gmail.com
- Sergei A. Aleinik** – research scientist, STC-Innovation, Ltd., Saint Petersburg, Russia, aleinik@speechpro.com
- Yuri N. Matveev** – D.Sc., Professor, ITMO University, Saint Petersburg, Russia, matveev@mail.ifmo.ru; Chief research scientist, STC-Innovation, Ltd., Saint Petersburg, Russia, matveev@speechpro.com
- Alexei V. Sholokhov** – postgraduate, University of Eastern Finland, Joensuu, Finland, sholokhovalexey@gmail.com

Принято к печати 21.05.14

Accepted 21.05.14