

ШУМОПОДАВЛЕНИЕ ПРИ ОБРАБОТКЕ РЕЧЕВОГО СИГНАЛА С ИСПОЛЬЗОВАНИЕМ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

NOISE REDUCTION FOR SPEECH SIGNAL PROCESSING BY USING WAVELET TRANSFORM

**M. Hein
V. Dovgal
V. Kudinov**

Summary. Speech plays an important role in multimedia system. Speech enhancement is to remove noise from speech for multimedia systems. Noise is an unwanted interference in any form of communication generally degrades the quality of the information signal. During transmission and reception signals are often corrupted by noise which can cause severe problems for downstream processing and user perception. Therefore an automated means of removing the noise would be an invaluable first stage for many signal processing tasks. Denoising has long been a focus of research and yet there always remains room for improvement. There are so many techniques to improve the signal quality or to regenerate the signal. In this paper we present a method for speech denoising using wavelet transform. It is often necessary to perform denoising in speech processing system operating in highly noisy environment. Wavelet transform is one of the most promising techniques used in signal processing, due to its ability to decompose signals and to reduce noise having non-stationary characteristics.

Keywords: Speech signals, Wavelet Transform, Discrete Wavelet Transform (DWT).

Хейн Мин Зо

аспирант, Курский Государственный Университет,
г. Курск
heinminnzaw13@gmail.com

Довгаль Виктор Митрофанович

Д.т.н., профессор, Курский государственный
университет
vmdovgal@yandex.ru

Кудинов Виталий Алексеевич

Д.п.н., профессор, Курская государственная
сельскохозяйственная академия
kudinovva@yandex.ru

Аннотация. Речь играет важную роль в мультимедийной системе. Усиление речи — это удаление шума из речи для мультимедийных систем. Шум — это нежелательная интерференция в любую форму связи, как правило, деградирует качество информационного сигнала. Во время передачи и приема сигналы часто повреждаются шумом, который может вызвать серьезные проблемы для нисходящего потока обработки и восприятия пользователя. Поэтому автоматизированное средство снятия шума было бы бесценным первым шагом для многих задач обработки сигналов. Де-шумирование уже давно является объектом исследований, и все же всегда остается место для его улучшения. Существует так много техник для улучшения качества сигнала или для регенерации сигнала. В этой статье мы представляем метод шумоподавления речи с использованием вейвлет-преобразования. Часто бывает необходимо выполнять шумоподавление в системе обработки речи, работающей в высоко шумной среде. Вейвлет-преобразование является одним из наиболее перспективных техник, используемых при обработке сигналов, благодаря его способности разлагать сигналы и уменьшать шум, имеющий нестационарные характеристики.

Ключевые слова: речевые сигналы, вейвлет-преобразование, дискретное вейвлет-преобразование (DWT).

Введение

Сигналы, служат для передачи информации из одного места в другое. Они могут быть в форме света, звука, изображения и т.д. Сигналы, используемые в качестве носителя в виде беспроводной связи, звуковой сигнал, используемый человеком для общения с другим человеком, изображение, используемое для наблюдения и понимания. Если вклады сигналов важны в нашей жизни, то их обработка также важна. Сигнал повреждается любым типом нежелательного сигнала, называемого мешающим сигналом, что приводит к снижению качества информации. Чтобы сохранить сигнал для дальнейшей обработки в не причинной операции, такой как хранение, сравнение, идентификация, проверка, цель исследования.

Речевое шумоподавление — это область инженерии, в которой изучаются методы, используемые для восстановления исходной речи из шумных сигналов, поврежденных различными типами шумов. Шум может быть в виде белого шума, розового шума, шумового шума и многих других типов шума, присутствующих в окружающей среде. За последние десятилетия удаление шума из речевых сигналов является областью, представляющей интерес для исследователей при обработке речи. Вейвлет-методы в основном используются для шумоподавления речи. Основная идея вейвлетов — проанализировать по шкале. Вейвлет-преобразования могут разлагать сигнал на несколько шкал, которые представляют разные диапазоны частот, и в каждом масштабе, можно приближенно определить положение мгновенных структур сигнала. Такое свойство можно использовать

для шумоподавления. Хотя без информации о сигнале, подлежащем анализу, вейвлет выбирает информацию, сильно уменьшая ее количество.

Вейвлет и мульти-разрешение

Вейвлет — это малая волна, а вейвлет-преобразования преобразуют сигнал в серию всплесков и обеспечивают способ анализа осциллограмм, ограниченных как по частоте, так и по длительности. Это позволяет сохранять сигнал более эффективно, чем преобразование Фурье. Вейвлет-преобразование предпочтительнее преобразования Фурье (FT) и короткого временного преобразования Фурье (STFT), поскольку оно обеспечивает мульти-разрешение.

В сигнале временной области, независимой переменной является время, а зависимая переменная — амплитуда. Большая часть информации скрыта в частотном контенте. Используя вейвлет-преобразование, мы можем получить информацию о частоте, которая невозможна, работая во временной области. Анализ нестационарного сигнала с использованием преобразования Фурье и короткого преобразования Фурье не дает удовлетворительных результатов. Более эффективные результаты могут быть получены с помощью анализа вейвлет-преобразования. В коротком временном преобразовании Фурье используется фиксированное временное частотное разрешение, где, как и в вейвлет-преобразовании, используется технология мульти-разрешения. Одним из преимуществ анализа вейвлет-преобразования является способность выполнять локальный анализ. Вейвлет-анализ позволяет выразить внешний вид сигнала, что другие методы анализа пропускают, такие как точки пробоя, отсутствие непрерывности и т.д.

В мульти-разрешении анализе сигнал имеет хорошее временное разрешение и низкое частотное разрешение на высоких частотах, и другим способом хорошее разрешение по частоте и низкое временное разрешение на низких частотах [1]. Он более подходит для короткой продолжительности более высокой частоты и более длинная длительность низкочастотных компонентов. Предполагается, что низкие частоты появляются на всю длительность сигнала тогда как высокие частоты появляются время от времени как короткий интервал. Это часто бывает в практических приложениях.

Вейвлет-преобразование

В этой работе, мы изложили лишь некоторые ключи и понятия вейвлет-преобразования, более строгая математическая обработка этого вопроса можно найти в [2,

3, 4, 5]. Непрерывное время вейвлет-преобразование (CWT) $f(t)$ определяется как:

$$\begin{aligned} CWT_{\psi} f(a, b) &= W_f(b, a) = \\ &= |a|^{-\frac{1}{2}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t-b}{a} \right) dt \end{aligned} \quad (1)$$

Здесь $a, b \in \mathbb{R}, a \neq 0$, и они расширяют и переводят коэффициенты, соответственно. Это умножение $|a|^{-\frac{1}{2}}$ для целей нормировки энергии, так что преобразованный сигнал будет иметь одинаковую энергию в каждом масштабе. Функция анализа $\psi(t)$, так называемый материнский вейвлет, должен удовлетворять, что она имеет нулевую чистую зону, которые предполагают, что преобразование ядра вейвлет-преобразования является функцией поддержки компактно (локализованное во времени).

Одним из недостатков CWT является то, что представление сигнала часто является избыточным, так как a и b непрерывны над \mathbb{R} (действительное число). Исходный сигнал может быть полностью реконструирован выборочной версией $W_f(b, a)$. Типично, мы образец в диадической сетке, т.е. $a = 2^{-m}$ и $b = n2^{-m}, m, n \in \mathbb{Z}^+$. Подставляя последнее в уравнение (1):

$$DWT_{\psi} f(a, b) = \int_{-\infty}^{\infty} f(t) \psi^*(t) dt \quad (2)$$

где $\psi_{m,n}(t) = 2^{-m} \psi(2^m t - n)$ представляет собой расширенную и переведенную версию материнского вейвлета $\psi(t)$.

Из-за ортонормированных свойств в дискретном вейвлет-преобразовании отсутствует избыточность информации. Кроме того, при таком выборе a и b существует алгоритм анализа, который разлагает сигнал на шкалы с различным временным и частотным разрешением.

Различия между различными вейвлет-функциями матери (например, Haar, Daubechies, Coiflets, Symlet, Biorthogonal и т.д.) заключаются в том, как определяются эти сигналы масштабирования и вейвлеты. Выбор вейвлета определяет окончательную облик формы волны; также, для преобразования Фурье разложенные формы сигналов всегда являются синусоидными.

Вейвлет-разложение приводит к уровням аппроксимированных и детальных коэффициентов. Алгоритм разложения и восстановления сигнала вейвлет-сигнала



Рис. 1. Алгоритм де-шумирования

от вейвлет-преобразования иллюстрируется многочисленными источниками [2, 3, 4].

Вейвлет-фильтр

Вейвлет — это термин, используемый для описания короткого временного ряда (обычно менее 100 выборок), который может использоваться для представления, например, функции источника. Как было показано ранее, вейвлет может быть изучен как временной ряд во временной области или в частотной области в виде амплитудного или фазового спектра. Для любого спектра амплитуды имеется бесконечное число вейвлетов временной области, которые могут быть построены путем изменения фазового спектра. Существует два специальных типа фазовых спектров, представляющих особый интерес.

Минимальный фазовый вейвлет имеет короткую продолжительность времени и концентрацию энергии в начале вейвлета. Он равен нулю до нуля (причинный). Идеальным сейсмическим источником был бы всплеск (максимальная амплитуда на каждой частоте), но лучшим практическим одним был бы минимальным фазе. Весьма распространено преобразование данного вей-

влет-источника вейвлет-памяти в его минимальный эквивалент фазы, поскольку несколько этапов обработки (например, предсказательная деконволюция) работают лучше всего, предполагая, что входные данные являются минимальной фазой.

Максимальный фазовый вейвлет — это обратное время минимальной фазы, и в каждой точке фаза больше для максимума, чем минимум. Все остальные каузальные вейвлеты строго говоря, смешанные фазы и будут иметь более длительный период времени. Свертка двух минимальных фазовых вейвлетов является минимальной фазой. Вейвлет с нулевой фазой имеет меньшую длительность, чем минимальный эквивалент фазы. Вейвлет симметричен с максимумом в момент времени (не причинный). Тот факт, что энергия доходит до нулевого времени, физически не реализуется, но вейвлет полезен для увеличения разрешающей способности и простоты выбора событий отражения (пика или желоба). Свертка нулевого и минимального фазовых вейвлетов является смешанной фазой (потому что фазовый спектр исходного минимального фазового вейвлета не является единственным минимальным фазовым спектром для нового модифицированного вейвлета), и его следует избегать.[6]

Мягкое и жесткое пороговые значения

В последние годы вейвлет-преобразование [7] стало мощным инструментом для многомасштабного представления и анализа сигналов. Вейвлет-преобразование локализует информацию на частотно-временной плоскости, в частности, они способны торговать одним типом разрешения для другого, что делает их особенно подходящими для анализа речевого сигнала. Идея вейвлет-преобразований заключается в том, что сдвиг должен допускать только изменение во времени, но не в форме.

Недавно были предложены различные методы, основанные на вейвлетах, для целей де-шумирования речи. Метод коэффициента вейвлет-разделения — это процедура де-шумирования речи для удаления шума путем сокращения вейвлет-коэффициентов в домене вейвлета. Метод основан на пороговом сигнале, что каждый вейвлет-коэффициент сигнала сравнивается с заданным порогом; если коэффициент меньше порога, то он устанавливается на ноль, в противном случае он сохраняется или слегка уменьшен по амплитуде. Мягкое и жесткое пороговое значение используются для де-шумирования сигналов. Использование вейвлетов для удаления шума из сигнала требует идентифицирования, какие компоненты содержат шум, а затем восстановления сигнала без этих компонентов [8].

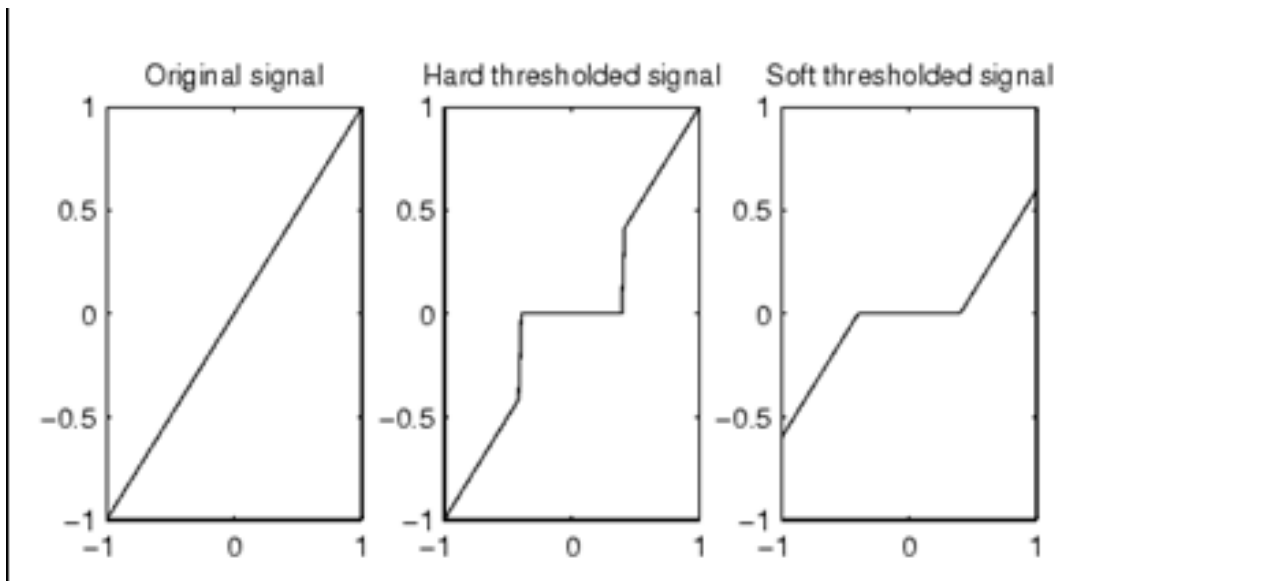


Рис. 2. Сравнение между мягким и жестким порогом

Принцип, под которым действует порог вейвлетов, заключается в улучшении коэффициентов вейвлетов, зависит от определенного значения, называемого порогом. Алгоритм де-шумирования суммируется следующим образом:

- i) Вычислить дискретное вейвлет-преобразование для шумового сигнала.
- ii) На основе алгоритма, называемого алгоритмом порогового значения и пороговым значением, сжимаются некоторые вейвлет-коэффициенты детали.
- iii) Вычислить обратное дискретное вейвлет-преобразование.

Пороговое значение используется в вейвлет-области для сглаживания или для удаления некоторых коэффициентов подсигналов вейвлет-преобразования измеренного сигнала. Это снижает уровень шума сигнала в нестационарной среде. Существует два распространенных способа порога результирующих вейвлет-коэффициентов. В первом случае значения коэффициентов установлены на ноль, абсолютное значение которых ниже порогового значения. Это обычно называют жестким порогом. Второй, называемый мягким порогом, идет еще на один шаг и уменьшает величину остальных коэффициентов на пороговое значение. Жесткий порог поддерживает масштаб сигнала, но вводит звон после реконструкции из-за прерывности в вейвлет-коэффициентах. Мягкий порог устраняет эту проблему, что приводит к более плавному сигналу. Методы мягкого и жесткого порога используются для

оценки вейвлет-коэффициентов при пороговом шумопонижении.

Жесткий порог можно описать как обычный процесс установки нуля элементов, абсолютные значения которых ниже порога.

$$T_{Hard}(x) = \begin{cases} x & |x| \geq thr \\ 0 & |x| < thr \end{cases} \quad (3)$$

Мягкий порог — это расширение жесткого порога, сначала устанавливающее на ноль элементы, абсолютные значения которых ниже порога, а затем уменьшают ненулевые коэффициенты в направлении 0. [9]

$$T_{Soft}(x) = \begin{cases} Sign(x)(x - thr) & x \geq thr \\ 0 & -thr \leq x < thr \\ Sing(x)(x + thr) & x < -thr \end{cases} \quad (4)$$

Заключение

Речевое шумоподавление (де-шумирования) — это область инженерии (техники), в которой изучаются методы, используемые для восстановления исходной речи из шумных сигналов, поврежденных различными типами шумов. Более высокий порог хорошо удаляет шум, но часть исходного сигнала также удаляется шумом. Мягкое пороговое значение лучше жесткого порога, поскольку мягкое пороговое значение дает лучшие результаты, чем жесткое пороговое значение. Жесткий порог — это самый простой способ, но мягкий порог обладает хорошими математическими свойствами и дает лучшую производительность шумоподавления.

ЛИТЕРАТУРА

1. Prof. Dr. Ir. M. Steinbuch, Dr. Ir. M.J.G. van de Molengraft, June 7 (2005), Eindhoven University of Technology, Control Systems Technology Group Eindhoven, «Wavelet Theory and Applications», a literature study, R.J.E. Merry, DCT 2005.53.
2. C. Gargour, M. Gabrea, V. Ramachandran, J. Lina. «A Short Introduction to Wavelets and Their Applications», IEEE Circuits and Systems Magazine, ISSN: 1531–636X, vol. 2, pp. 57–67, 2009.
3. C. Taswell. «The What, How and Why of Wavelet Shrinkage Denoising», Computing in Science and Engineering, ISSN: 1521–9615, vol. 2, no. 3, pp. 12–19, June 2000.
4. S. Tsai, «Wavelet Transform and Denoising», Master's Thesis, URN: etd-12062002–152858, Chapter 4, pp. 35–42.
5. <http://taco.poly.edu/WaveletSoftware/denoise2.html>
6. <http://www.xsgeo.com/course/basic.htm>
7. James S. Walker, A primer on Wavelets and their scientific applications, 2nd Edition, Chapman & Hall/CRC USA, 2008.
8. Robi Polikar «The Wavelet Tutorial» Rowan University College of Engineering.
9. D. L. Donoho. «Denoising by Soft thresholding» IEEE Trans on Information Theory. 1995, vol. 41, no. 3, pp. 613–627.

© Хеин Мин Зо (heinminnzaw13@gmail.com),

Довгаль Виктор Митрофанович (vmdovgal@yandex.ru), Кудинов Виталий Алексеевич (kudinovva@yandex.ru).

Журнал «Современная наука: актуальные проблемы теории и практики»

