

10

Фильтрация речевых сигналов с применением комплексного вейвлет-преобразования двойной плотности

© А.С. Ясин^{1,2}, О.Н. Павлова¹, А.Н. Павлов^{1,3,4}

¹ Саратовский государственный университет им. Н.Г. Чернышевского, Саратов, Россия

² Университет технологий, ул. Аль-Сина, Багдад, Ирак

³ Саратовский государственный технический университет им. Гагарина Ю.А., Саратов, Россия

⁴ Саратовский филиал Института радиотехники и электроники им. В.А. Котельникова РАН, Саратов, Россия
E-mail: pavlov.alexey@gmail.com

Поступило в Редакцию 23 марта 2016 г.

Решается задача повышения качества очистки от помех речевых сигналов при использовании комплексного вейвлет-преобразования двойной плотности по сравнению со стандартным методом вейвлет-фильтрации, основанным на многомасштабном анализе с базисами Добеши. Продемонстрированы значительное увеличение средней оценки разборчивости речи при высокой интенсивности помех и возможность ограничиться меньшим числом уровней разложения для последующей коррекции вейвлет-коэффициентов.

Вейвлет-фильтрация зашумленных сигналов и изображений относится к числу широко применяемых методов очистки информационных сообщений от помех [1–3]. Традиционные схемы фильтрации, основанные на многомасштабном анализе с использованием дискретного вейвлет-преобразования (ДВП), характеризуются высоким быстродействием и являются относительно простыми с точки зрения их программной реализации [4]. Однако они имеют ряд ограничений, для устранения которых предложены методы комплексного вейвлет-преобразования [5–9], применяющие базисные функции, у которых действительная и мнимая части являются сопряженными по Гильберту. Значительное число работ в этой области основывается на методе

дуального комплексного вейвлет-преобразования [5–7] который продемонстрировал высокую эффективность при фильтрации зашумленных изображений. Этот метод применяет ортонормированные базисы, что обеспечивает быстрое действие и отсутствие избыточности разложения сигнала. В то же время использование критической выборки делает данный метод более зависимым от коррекции информативных вейвлет-коэффициентов, которая сопровождается искажениями восстановленного сигнала. По этой причине фильтрацию сигнала, передаваемого по каналу связи, часто проводят с использованием фреймов — неортонормированных (избыточных) базисов. За счет некоторого увеличения времени обработки информационных сообщений, которая тем не менее может проводиться в онлайн-режиме, приобретает возможность сохранения необходимой информации о сигнале в случае удаления части „нужных“ вейвлет-коэффициентов или в случае, когда прямое разложение проводится с недостаточно высокой точностью из-за наличия помех.

В настоящее время большое внимание уделяется созданию комбинированных алгоритмов очистки информационных сообщений от шумов и случайных искажений, применяющих различные приемы вейвлет-фильтрации. В последние годы были предложены перспективные разработки, такие как комплексное вейвлет-преобразование двойной плотности (КВПДП) [8,9]. Алгоритмически этот подход является существенно более сложным, по сути, представляя собой качественно новый уровень в задачах цифровой фильтрации сигналов и изображений. В данной работе исследуются возможности метода КВПДП при фильтрации речевых сигналов и показывается, что этот метод позволяет ограничиться меньшим числом уровней разложения сигнала в базисе вейвлет-функций по сравнению с ДВП для достижения максимального качества очистки информационных сообщений от помех.

Метод КВПДП предусматривает разложение сигнала с применением одной комплексной скейлинг-функции и двух комплексных вейвлет-функций, у которых вещественные и мнимые части являются сопряженными по Гильберту. Для этих функций записываются те же условия, что и для вейвлетов Добеши, применяемых в рамках ДВП [10–12]. Так, для вещественных частей функций справедливы соотношения

$$\varphi(t) = \sqrt{2} \sum_n h_0(n) \varphi(2t - n),$$

$$\psi_{1,2}(t) = \sqrt{2} \sum_n h_{1,2}(n) \psi_{1,2}(2t - n), \quad h_2(n) = h_1(n - 1),$$

которые по аналогии записываются и для мнимых частей. При таком выборе фильтров ψ_i , на новом уровне разрешения осуществляется прореживание только аппроксимирующих коэффициентов (коэффициентов разложения по скейлинг-функциям φ), тогда как детализирующие коэффициенты (коэффициенты разложения по вейвлет-функциям ψ_i) сохраняются в полном объеме. Общая схема разложения сигнала в рамках метода КВПДП представлена на рис. 1. В данной работе применялись фильтры, разработанные группой I.W. Selesnick [8]. Проводилось сопоставление качества фильтрации, осуществляемой для зашумленных речевых сигналов с использованием методов ДВП и КВПДП.

Для количественного описания качества фильтрации речевых сообщений применялась средняя оценка разборчивости речи (MOS) [13], которая по 5-балльной шкале характеризует показатели работы системы связи, используемой для разговора или восприятия речевого материала. По определению MOS предусматривает наличие большого числа субъективных оценок качества системы связи и их последующее усреднение. Для практических целей проводится аппроксимация MOS с применением модели PESQ (perceptual evaluation of speech quality), которая была разработана для объективной оценки MOS в соответствии с общепринятыми стандартами международного союза электросвязи [14,15].

В целях сравнения эффективности методов фильтрации на основе ДВП и КВПДП был рассмотрен ряд примеров речевых сообщений (рис. 2, а) с наложенным аддитивным шумом разной интенсивности и статистики. Для каждого примера вначале применялся многомасштабный анализ на основе ДВП и проводились оценки MOS для нескольких вариантов базисов семейства Добеши (от D^4 до D^{20}) и двух вариантов задания пороговой функции при коррекции вейвлет-коэффициентов — мягкого и жесткого. Соответствующие расчеты осуществлялись при коррекции вейвлет-коэффициентов на разных уровнях разложения. В результате определялись оптимальный уровень разложения и вейвлет-базис, обеспечивающие наилучшее качество очистки голосового сообщения от помех. Было установлено, что применение мягкого варианта задания пороговой функции при коррекции вейвлет-коэффициентов позволяет обеспечивать существенно более высокое значение MOS (увеличение до 50% по сравнению с жестким заданием пороговой функции при отношении сигнал/шум, равным 0 дБ). Это связано с тем, что разрывы пороговой функции при жестком варианте коррекции вейвлет-коэффициентов приводят к более существенным искажениям

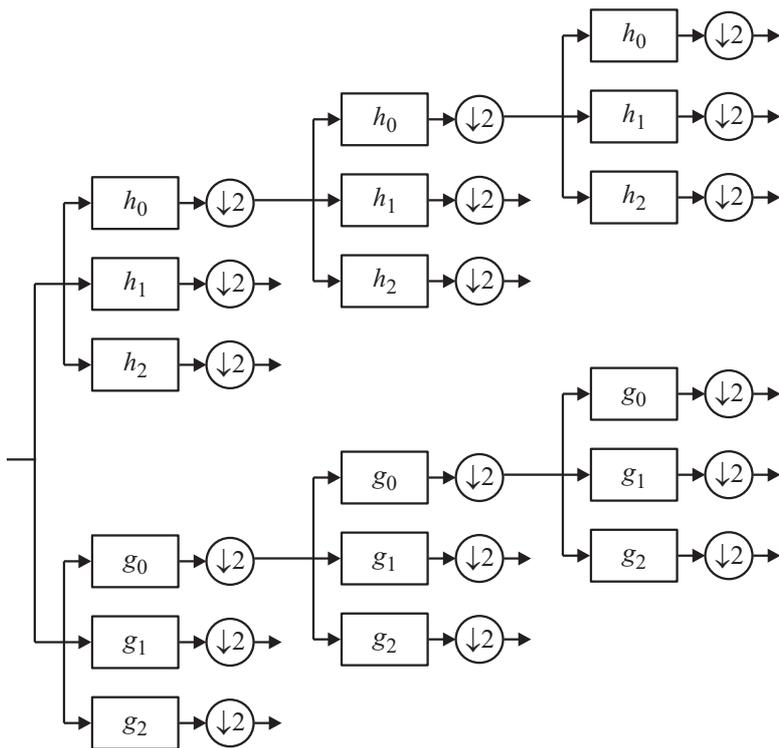


Рис. 1. Схема разложения сигнала при использовании метода КВПДП (приведены 3 уровня разложения). Фильтры h_0 и g_0 соответствуют действительной и мнимой частям скейлинг-функции, а фильтры $h_{1,2}$ и $g_{1,2}$ — действительным и мнимым частям вейвлет-функций.

реконструированного сигнала. Выбор оптимального базиса зависит от анализируемого сигнала, однако для увеличения MOS целесообразно применять сравнительно гладкие вейвлеты, характеризующиеся большей областью задания ($D^{10}-D^{20}$). Качество фильтрации также зависит от числа используемых уровней разложения. Для примера, представленного на рис. 2, *a*, значение MOS, полученное при использовании 2–3 уровней разложения, превосходит значение, полученное при использовании только одного уровня (рис. 2, *b*, черные точки).

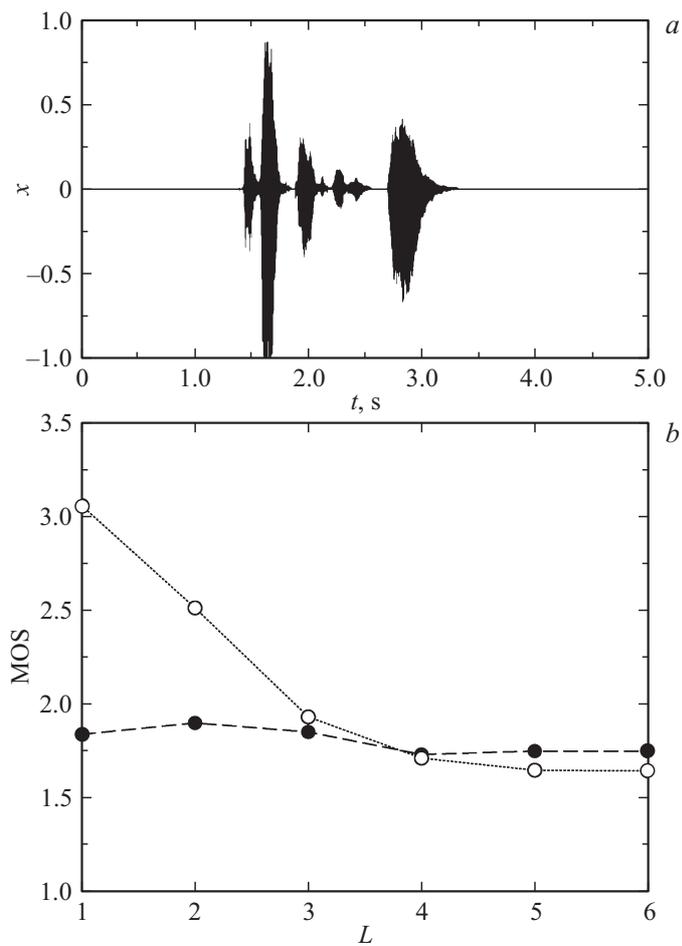


Рис. 2. Пример анализируемого речевого сообщения „Как дела“ (а) и результаты вейвлет-фильтрации добавленного шума на основе методов, применяющих ДВП (черные точки) и КВПДП (белые точки)(b).

Аналогичные результаты были получены для других рассмотренных примеров экспериментальных данных. Отметим, что достигнутое максимальное значение MOS для метода ДВП на рис. 2, b составляет 1.9, что

свидетельствует о невысоком качестве фильтрации речевого сигнала при выбранной интенсивности шума.

Аналогичные расчеты были проведены для метода КВПДП. В данном случае применялся один вариант вейвлет-базиса [8], и оценка качества фильтрации проводилась при рассмотрении разного числа уровней разложения зашумленных речевых сигналов. Характерная зависимость MOS от номера уровня приведена на рис. 2, *b* (белые точки). В отличие от метода, основанного на ДВП, достигнутое максимальное значение MOS составляет 3.1, что позволяет сделать вывод о значительном улучшении данной характеристики качества фильтрации. В других рассмотренных примерах увеличение MOS также было существенным, свидетельствуя о несомненном преимуществе применения метода КВПДП. Важное обстоятельство состоит в том, что в приведенном примере максимум MOS достигается на первом уровне разложения, а не на втором-третьем, как в случае ДВП. В других примерах отмеченная закономерность была подтверждена. Номер уровня, при котором обеспечивается максимум MOS, варьируется при изменении интенсивности шума, но во всех проведенных расчетах максимум для метода КВПДП достигался при меньшем уровне разложения, чем для метода, применяющего ДВП. Данное обстоятельство позволяет уменьшить число уровней разложения для проведения вейвлет-фильтрации, частично скомпенсировав увеличение затрат времени для очистки информационных сообщений от помех, связанное с применением избыточных базисов и комплексных вейвлет-функций.

Исследование выполнено при поддержке гранта Российского научного фонда (проект № 14-12-00324).

Список литературы

- [1] *Meyer Y.* Wavelets: Algorithms and applications. Philadelphia: S.I.A.M., 1993.
- [2] *Jansen M.* Noise reduction by wavelet thresholding. New York: Springer-Verlag, 2001.
- [3] *Donoho D.L.* // IEEE Trans. Inf. Theory. 1995. V. 41. P. 613.
- [4] *Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P.* Numerical recipes in C: the art of scientific computing. Cambridge: Cambridge University Press, 1992.
- [5] *Kingsbury N.G.* // Appl. Comput. Harmon. Anal. 2001. V. 10. P. 234.
- [6] *Selesnick I.W., Baraniuk R.G., Kingsbury N.G.* // IEEE Signal Proc. Mag. 2005. V. 22. P. 123.

- [7] *Selesnick I.W.* // IEEE Signal Proc. Lett. 2001. V. 8. P. 170.
- [8] *Selesnick I.W.* // IEEE Trans. Signal Proc. 2004. V. 52 (5). P. 1304.
- [9] *Bhonsle D., Dewangan S.* // Int. J. Sci. Res. Publ. 2012. V. 2 (7). P. 1.
- [10] *Daubechies I.* Ten lectures on wavelets. Philadelphia: S.I.A.M., 1992.
- [11] *Дремин И.М., Иванов О.В., Нечитайло В.А.* // УФН. 2001. Т. 171. С. 465;
Dremin I.M., Ivanov O.V., Nechitailo V.A. // Phys. Usp. 2001. V. 44. P. 447.
- [12] *Павлов А.Н., Храмов А.Е., Короновский А.А., Ситникова Е.Ю., Макаров В.А., Овчинников А.А.* // УФН. 2012. Т. 182. С. 905;
Pavlov A.N., Hramov A.E., Koronovskii A.A., Sitnikova E.Yu., Makarov V.A., Ovchinnikov A.A. // Phys. Usp. 2012. V. 55. P. 845.
- [13] *Loizou P.C.* Speech enhancement. Theory and Practice (2-nd ed.). Boca Raton: CRC Press, 2013.
- [14] *Beerends J.G., Stemerink J.A.* // J. Audio Eng. Society. 1992. V. 40 (12). P. 963.
- [15] *Thiede T., Treurniet W.C., Bitto R., Schmidmer C., Sporer T., Beerends J.G., Colomes C., Keyhl M., Stoll G., Brandenburg K., Feiten B.* // J. Audio Eng. Society. 2000. V. 48(1/2). P. 3.