

УДК 004.932.1

УЛУЧШЕНИЕ КАЧЕСТВА РАСПОЗНАВАНИЯ РЕЧИ С ПОМОЩЬЮ ПРЕДВАРИТЕЛЬНОЙ ВЕЙВЛЕТ-ФИЛЬТРАЦИИ

К. т. н. А. В. Садченко, О. А. Кушниренко, Д. В. Стрижов

Одесский национальный политехнический университет
Украина, г. Одесса
anjand@mail.ru, kuuk@mail.ru

Проведен сравнительный анализ качества распознавания речи на фоне белого шума при использовании предварительной низкочастотной фильтрации и вейвлет-анализа, показаны преимущества нового метода.

Ключевые слова: преобразование Фурье, вейвлет-анализ, распознавание речи, голосовое управление.

Существенное увеличение производительности сигнальных процессоров при снижении их стоимости позволяет строить эффективные системы голосового управления. Актуальность таких систем не вызывает сомнений, так как они представляют собой один из самых удобных интерфейсов взаимодействия с системой автоматического управления. Примером систем автоматического управления служат справочные системы, системы типа «умный дом» и т. д. Основным требованием к таким системам является высокая вероятность правильного распознавания речи.

Целью работы была разработка адаптивного алгоритма предварительной фильтрации речевого сигнала от шума для улучшения качества распознавания речи.

Вне зависимости от алгоритма распознавания речи на качество полученного результата существенное влияние оказывают шумы. Для очистки звукового (речевого) сигнала от шума обычно применяют фильтры низкой частоты во временной области либо Фурье-анализ (в частотной области). Типовая схема системы распознавания речи с предварительным цифровым фильтром низкой частоты с конечной импульсной характеристикой (КИХ) показана на рис. 1.

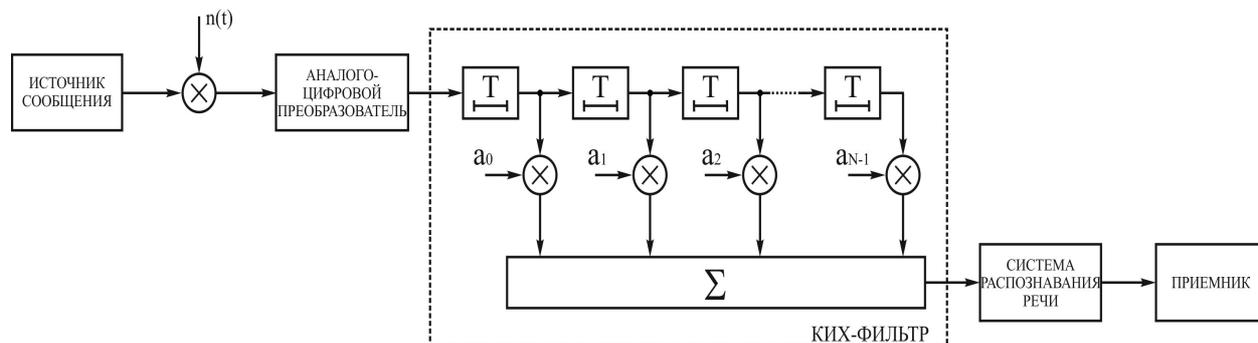


Рис. 1. Структурная схема системы распознавания речи с предварительным цифровым КИХ-фильтром: Т — линия задержки; X — умножитель на коэффициент импульсной характеристики КИХ-фильтра

На рис. 2. показано также, что на сигнал источника сообщений (рис. 2, а) воздействует белый гауссов шум (БГШ) $n(t)$ (рис. 2, б).

На рис. 3 приведена структурная схема шумоподавителя с блоком вейвлет-анализа.

В соответствии с измеренной дисперсией шума устанавливается порог подавления спектральных коэффициентов. Результаты моделирования в системе MATLAB представлены на рис. 4.

Видно, что при одинаковом подавлении амплитуды шумовой составляющей (критерием оценки эффективности алгоритмов тут является одинаковая дисперсия БГШ на выходе этих двух схем) сигнал на выходе цифрового ФНЧ (КИХ-фильтра) (рис. 4, а) потерял высокочастотные ос-

циляции, т. е. искавился по форме относительно исходного сильнее, чем в схеме с блоком вейвлет-анализа (рис. 4, б).

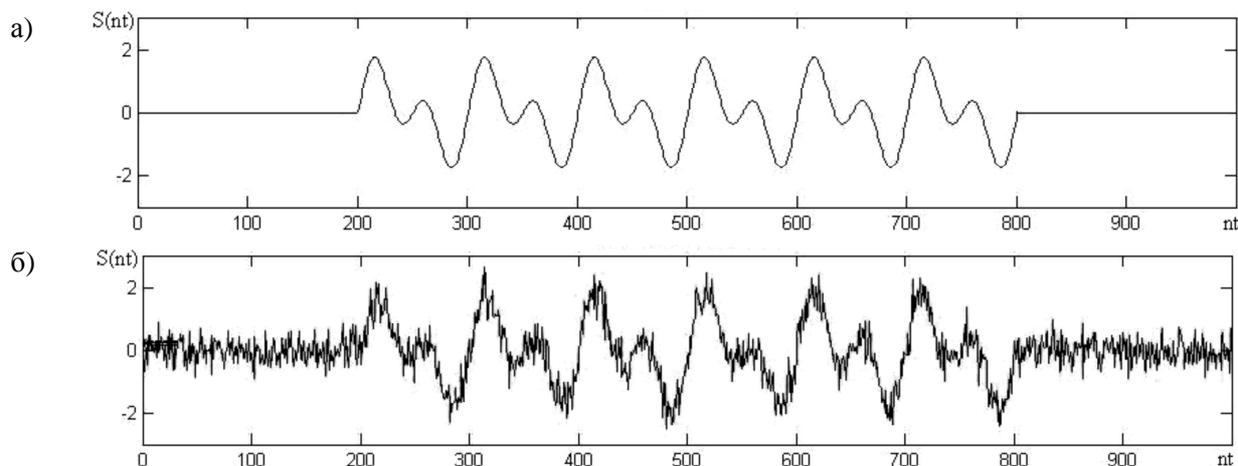


Рис. 2. Полезный сигнал (а) и «зашумленный» сигнал (б) источника сообщений

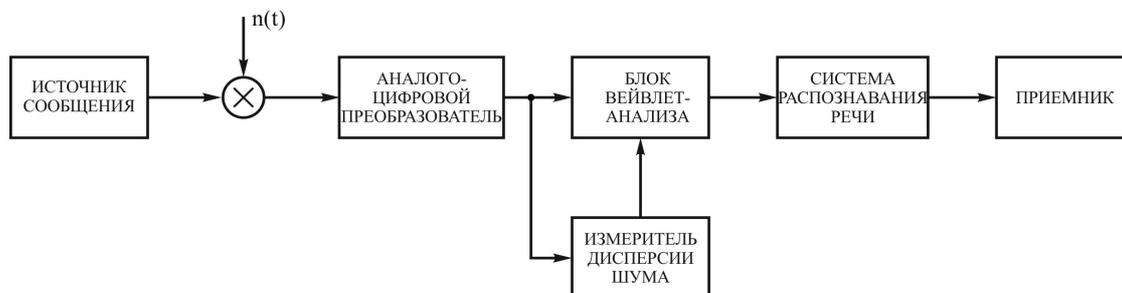


Рис. 3. Структурная схема системы распознавания речи с блоком вейвлет-анализа

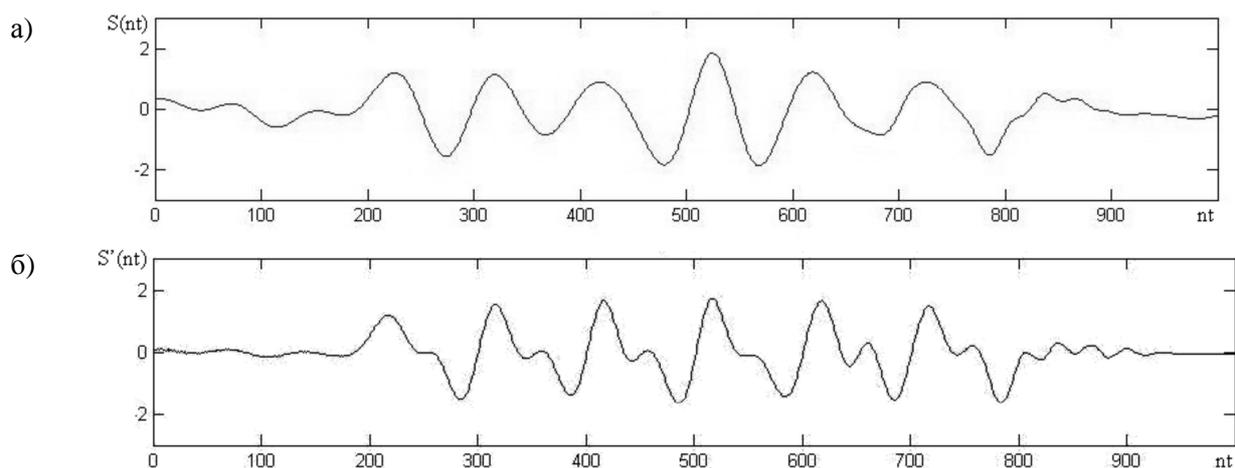


Рис. 4. Сигнал, восстановленный после КИХ- фильтра (а) и с помощью вейвлет преобразования (б)

Следует отметить, что выбор функции материнского вейвлета и числа уровней декомпозиции сигнала также влияет на эффективность шумоподавления. В качестве сравнительного критерия оценки качества восстановления сигнала можно выбрать критерий минимума среднеквадратического отклонения (СКО) [2].

Экспериментальным путем было установлено, что наилучшим СКО обладает вейвлет-функция Добеши длины 19 при числе уровней декомпозиции 4. При меньшем числе уровней разложения не происходит хорошей очистки от шума, а при большем — теряются локальные особенности сигнала.

Результаты распознавания речи (посимвольное сравнение) с помощью сервиса «google» приведены в таблице.

Результаты распознавания речи

Дисперсия, БГШ	Посимвольная вероятность правильного распознавания, %		
	Без предварительной обработки	С цифровым ФНЧ	С блоком вейвлет- анализа
0,1	90	94	99
0,2	88	92	94
0,5	66	77	85

Установлено, что при восстановлении сигнала с помощью ФНЧ мы теряем высокочастотную составляющую спектра, что может привести к потере полезной составляющей сигнала, снижению вероятности правильного распознавания, тогда как при восстановлении сигнала с помощью вейвлет-преобразования сохраняется высокочастотная составляющая спектра сигнала. Благодаря этому удастся улучшить качество распознавания речи.

ИСПОЛЬЗОВАННЫЕ ИСТОЧНИКИ

1. Садченко. А. В., Кушниренко О. А., Савчук Ю. А., Шаргородский С. И. Система оптического распознавания типов радиокомпонентов на печатных платах // Труды 15-й МНПК «СИЭТ-2014». Т. 1. — Украина, г. Одесса. — 2014.— С. 193-194.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений.— Москва: изд. «Техносфера», 2005.
3. Штарк Г. Применение вейвлетов для ЦОС.— Москва: изд. «Техносфера», 2007.
4. Rose R. Robust speech recognition techniques applied to a speech in noise task // European Conference on Speech Communication and Technology.— Aalborg, Denmark — 2001.
5. Гольденберг Л. М. и др. Цифровая обработка сигналов: Справочник.— Москва: Радио и связь, 2007.
6. Кушниренко О. А., Садченко А. В., Троянский А. В. Повышение помехоустойчивости «скользящего» корреляционного алгоритма распознавания печатных символов // Восточно-Европейский журнал передовых технологий.— 2014.— Т. 4, № 2 (70).— С. 32–36.
7. Гребнов С. В. Методы шумоочистки в задачах распознавания речи // Тезисы докладов Международной научно-технической конференции “Состояние и перспективы развития электротехнологии” (XV Бенардосовские чтения).— Россия, г. Иваново.— 2009. — Т.1.— С. 60–61.

A. V. Sadchenko, O. A. Kushnirenko, D. V. Strizhov

Improving the quality of speech recognition using wavelet pre-filtering.

The authors conduct a comparative analysis of the quality of speech recognition in white noise using a low-pass filtering and pre-wavelet analysis. The advantages of the new method are shown in the paper.

Keywords: *Fourier transform, wavelet analysis, speech recognition, voice control.*