

«ПЕРВЫЕ ШАГИ В НАУКЕ»

УДК 621.391

ДЕКОМПОЗИЦИЯ НА ЭМПИРИЧЕСКИЕ МОДЫ С ПАРАБОЛИЧЕСКОЙ ИНТЕРПОЛЯЦИЕЙ ОГИБАЮЩИХ В ЗАДАЧАХ ОЧИСТКИ СИГНАЛОВ ОТ ШУМА

Клионский Д.М., аспирант кафедры МО ЭВМ Санкт-Петербургского государственного электротехнического университета “ЛЭТИ”, klio2003@list.ru

Научные руководители:

Орешко Н.И., к.т.н., в.н.с. Научно-инженерного центра Санкт-Петербургского государственного электротехнического университета “ЛЭТИ”, oreshko@nicetu.spb.ru

Геппенер В.В., д.т.н., профессор кафедры МО ЭВМ Санкт-Петербургского государственного электротехнического университета “ЛЭТИ”, geppenerv@mail.ru

Ключевые слова: фильтрация, декомпозиция, эмпирические моды, параболическая интерполяция, сплайн-интерполяция.

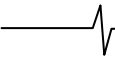
Декомпозиция на эмпирические моды

Большинство реальных сигналов, рассматриваемых в самых разных областях знаний, чаще всего являются нестационарными [6], что проявляется в непостоянстве их отдельных характеристик во времени. Чаще всего на практике рассматривается случай стационарного в широком смысле поведения, подразумевая, что среднее значение и дисперсия сигнала не зависят от времени, а автокорреляционная функция (АКФ) зависит лишь от относительного сдвига пары отсчетов сигнала [6]. У многих из существующих подходов к анализу и обработке сигналов при практическом использовании наблюдается ряд существенных недостатков. Например, широко распространенный классический анализ Фурье [3,4,6] из-за своей относительной простоты вычислений (формулы Парсевала, Планшереля, наличия быстрых вычислительных алгоритмов, наличия свойств о преобразовании Фурье производной и первообразной заданной функции, теоремы запаздывания во временной и частотной областях, теоремы подобия [6] и др.) практически сразу после появления быстрых вычислительных алгоритмов стал доминировать над всеми другими методами анализа сигналов и применяться ко всем типам сигналов. Несмотря на то, что преобразование Фурье выполняется при довольно общих условиях [6] (условиях Дирихле, абсолютной интегрируемости), есть несколько существенных ограничений, накладываемых на сигналы, для которых оно вычисляется. Суть их состоит в том, что сигналы должны быть *строго периодическими* функциями (для разложения в ряд Фурье, т. е. для представления в виде линейной суперпозиции гармонических функций кратных частот) или содержать так называемые повторяющиеся паттерны (под паттернами понимаются фрагменты одной общей природы, при этом не обязательно являющиеся идентичными), так как в противном случае анализ в частотной области может дать неверную интерпретацию результатов. Также необхо-

Рассматривается новый адаптивный метод анализа и обработки нестационарных сигналов – декомпозиция на эмпирические моды (ДЭМ). Метод позволяет эффективно решать задачи предварительной обработки (очистки от шума, выделения тренда и пр.), структурного анализа (мультиразрешающего анализа, частотно-временного анализа) и интеллектуального анализа сигналов (сегментация, кластер-анализ). Приводятся основные преимущества параболической ДЭМ и показываются ее более широкие возможности при решении задачи предварительной обработки сигналов – очистки от шума.

димо, чтобы сигналы обладали свойством стационарности по отношению к ряду характеристик (среднее значение, мгновенная частота [6,7-9]).

Однако анализ Фурье неэффективен при исследовании сигналов с изменяющимся частотным содержанием (например, ЛЧМ-сигнала, из Фурье-спектра которого нельзя сделать вывод о линейном законе изменения частоты), поскольку тригонометрический базис содержит немодулированные гармонические функции с не изменяющейся во времени частотой, в то время как сигналы могут включать в себя компоненты, занимающие разные частотные полосы и с изменяющимися мгновенными значениями частоты (частотная модуляция). Преобразование Фурье определено на всей временной оси, поэтому из-за отсутствия локализационных свойств (при вычислении преобразования используются все значения данных) возникают сложности при необходимости определения основных диапазонов концентрации энергии (на основе классической Фурье-периодограммы и ее модификаций), которая рассеивается по достаточно широкому диапазону частот, создавая порой ложную картину распределения энергии по частотам. Функции тригонометрического базиса (и некоторых других часто используемых базисов) не обеспечивают качественного временного разрешения, что весьма существенно для нестационарных сигналов (так как часто требуется точная локализация отдельных особенностей). Еще одним подводным камнем является то, что в силу своей гладкости (функции тригонометрического базиса дифференцируемы бесконечное число раз) они могут оказаться неэффективными при анализе сигналов, содержащих



участки с резкими изменениями, поскольку может потребоваться очень большое число функций для обеспечения требуемой точности представления. Кроме того, возникает известный эффект Гиббса (характерные осцилляции в окрестности точек разрыва) [6].

Получившие в последнее время широкое распространение частотно-временные методы (например, широкий класс частотно-временных распределений, объединенных в класс Козна) [4], представляющие сигнал на частотно-временной плоскости, дают заметные преимущества по сравнению с классическим анализом Фурье, однако с их помощью невозможно напрямую получить мультимасштабное представление сигнала (в виде совокупности компонент с разным разрешением), которое требуется для лучшего понимания его внутренней структуры. Для повышения точности и достоверности анализа нестационарных сигналов, включая решение важнейшей и первичной задачи предварительной обработки [7-10], необходим специальный подход, обладающий свойством *адаптивности* к каждому конкретному рассматриваемому сигналу. Из известных на сегодняшний день подходов ДВП на основе схемы диадического банка фильтров [3-4] частично удовлетворяет данному требованию, а также многим другим, представляющим исключительную практическую ценность, а именно: имеет хорошо разработанный математический аппарат, быстрые вычислительные алгоритмы, широкий класс решаемых задач. В силу особенностей конструирования вейвлет-базиса [3-4] (на основе масштабных растяжений и сдвигов вдоль временной оси материнской вейвлет-функции) появляется возможность адаптивно обрабатывать сигналы путем довольно точного учета локальных временных особенностей. Но главная проблема, порой затрудняющая их эффективное практическое применение, – наличие большого многообразия и неочевидность выбора вейвлета для решения конкретной задачи, если нет эвристических критериев выбора базиса. В целом следует признать, что вейвлет-преобразование в настоящее время играет одну из ведущих ролей в обработке сигналов в силу наличия большого числа специально разработанных базисов и применимости к различным важным практическим задачам (очистка от шума, сжатие изображений и сигналов, мультиполосный и мультиразрешающий анализ и др.) [3-4].

В данной статье будет подробно рассмотрена новая адаптивная технология анализа и обработки сигналов – *декомпозиция на эмпирические моды (ДЭМ)* [1-2, 7-10]. ДЭМ, обладая подобными вейвлет-разложению свойствами, не требует априорного выбора базиса. Другими словами, разложение по этой системе функций с целью последующего анализа проводится с учетом *локальных особенностей* (таких как экстремумы и нули сигнала) и *внутренней структуры* (наличия в сигнале трех основных типов компонент – шумовых, трендовых и сигнальных) каждого конкретного сигнала. Сами функции формируются, а правильнее сказать, *извлекаются* непосредственно из исходного сигнала, следовательно, такой базис всегда *уникален* (индивидуален, не повторяется в точности для других сигналов), *апостериорен* (т. е. становится полностью известным лишь после разло-

жения данных, а не до него, как, например, в случае с тригонометрическим базисом, где функции известны заранее) и *адаптивен* (т. е. приспосабливается к особенностям и свойствам данных).

В случае разложения по базису Фурье и вейвлет-базису используется следующая формула:

$$s(k) = \sum_i c_i f_{i\text{det}}(k),$$

где $s(k)$ – исходный сигнал; c_i – коэффициенты разложения; $f_{i\text{det}}(k)$ – детерминированные (априорно известные) базисные функции.

Для ДЭМ данная формула приобретает следующий вид:

$$s(k) = \sum_i c_i f_i(k), \quad (1)$$

где c_i – коэффициенты разложения; $f_i(k)$ – априорно неизвестные (*апостериорно* известные) функции, извлекаемые из исходного сигнала. Сами базисные функции формируются апостериорно, что является несомненным преимуществом по сравнению с первыми двумя случаями.

Термин “базис” здесь использован не совсем корректно, поскольку, согласно определению, базис – линейно независимое множество векторов (функций), линейная оболочка [12] которого образует все линейное пространство (т. е. каждый вектор этого пространства можно представить в виде линейной комбинации базисных векторов); здесь, однако, линейная независимость не является строго доказанной и, кроме того, извлекаемые компоненты в большинстве случаев могут применяться лишь для обработки того сигнала, из которого они были извлечены. Поэтому такой базис является *эмпирическим, аппроксимативным* и составляет своего рода совокупность “строительных блоков” для представления сигналов. Далее под “базисом” применительно к ДЭМ будет как раз пониматься такой аппроксимативный базис. В некоторых случаях при рассмотрении нескольких сигналов из одного и того же класса можно провести разложение какого-то одного из них, а затем, используя полученные компоненты и специально введенные веса для каждой из них, представлять другие однотипные сигналы. В этом случае можно утверждать (после проверки линейной независимости компонент), что извлеченные из сигнала компоненты являются базисом.

Технология обработки сигналов на основе ДЭМ в настоящее время получила широкое распространение при решении следующих основных задач предварительной обработки сигналов:

- очистка сигналов от шумов различной природы;
- выделение тренда в сигналах;
- анализ степени хаотичности сигналов на основе оценивания показателя Херста;
- проведение мультиразрешающего анализа (исследование тонкой структуры сигналов и их глобального поведения в целом; анализ свойств отдельных компонент) и мультиполосного анализа (исследование сигналов в различных частотных полосах);
- выделение шаблонов в сигналах и их классификация на основе физической и математической интерпретаций (шумовые, трендовые компоненты, характерные шаблоны и др.).

Процесс формирования базисных функций на основе специальным образом разработанного алгоритма не исчерпывает всю технологию, а является лишь первым (начальным) ее этапом. Далее по найденным базисным функциям формируется итоговое частотно-временное представление сигналов – так называемый *спектр Гильберта–Гуанга* [1-2, 7-10]. Такое представление на частотно-временной плоскости получается путем вычисления мгновенных частот каждой базисной функции, их специальной нормировки и нанесения рассчитанных значений на итоговую трехмерную диаграмму. Такое частотно-временное распределение позволяет также получить информацию об амплитуде, которая отображается соответствующим цветом (поэтому данное распределение, хотя и отображается как двумерное на плоскости, на самом деле является трехмерным: “время – частота – амплитуда” или “время – частота – энергия”). Спектр Гильберта–Гуанга, с одной стороны, аналогичен широко известным в частотно-временном анализе распределениям, но принципиально отличается тем, что он конструируется на основе адаптивно сформированного базиса, что позволяет рассчитывать на более точные результаты при дальнейшем анализе. Вкратце охарактеризовав идею второго этапа данной технологии, приведем основные задачи, решаемые путем исследования построенного спектра Гильберта–Гуанга:

- выявление скрытых в шумах модуляций сигнала (амплитудных, частотных);
- идентификация частотных и временных диапазонов концентрации энергии;
- классификация множества сигналов по спектру Гильберта–Гуанга;
- расчет на основе спектра Гильберта–Гуанга так называемого *маргинального спектра* [1-2,7-10] (путем численного интегрирования по времени спектра Гильберта–Гуанга), являющегося аналогом частотного спектра Фурье для нестационарных сигналов;
- расчет на основе спектра Гильберта–Гуанга так называемой *мгновенной плотности энергии* (путем численного интегрирования по частоте квадрата спектра Гильберта–Гуанга), позволяющей наглядно выявлять изменения энергетических соотношений в сигналах.

После проведения предварительной обработки и анализа самой структуры сигнала следует этап интеллектуального анализа, который осуществляется на основе специальной технологии *Data Mining* [5,8,9]. Интеллектуальный анализ ориентирован на добычу из “сырых” исходных данных (прошедших этап предварительной обработки) *новых, нетривиальных* (скрытых от обнаружения более простыми методами, такими как, например, статистический анализ или визуализация), *практически полезных и поддающихся интерпретации* знаний. Кроме того, задачи интеллектуального анализа должны быть *автоматизированы*, чтобы, по возможности, исключить эксперта из непосредственного процесса извлечения знаний, а привлечь его лишь на стадии интерпретации результатов. *Data Mining* использует представление сигналов в виде некоторых абстрактных упрощенных структур – *моделей*, которые вместе с тем должны обеспечивать необходимую точность и достоверность результатов и обладать свойством физичности

(т. е. отражать основные закономерности, свойственные некоторому явлению или процессу). Задач интеллектуального анализа довольно много, при этом основные из них, решаемые с помощью технологии ДЭМ, – это *сегментация, кластер-анализ* и *секвенциальный анализ* [5,13]. Таким образом, можно сформулировать следующие основные решаемые задачи:

- выявление участков в сигналах (сегментов), обладающих однородными свойствами;
- упрощение описания исходного сигнала, заданного в виде набора пар “временной отсчет – сигнальный отсчет” и устранение избыточности путем объединения отсчетов в однородные сегменты. Описание найденных сегментов с помощью набора информативных признаков (статистических, энергоинформационных, частотных и др.);
- поиск и объединение в группы (кластеры) выделенных сегментов;
- построение временных шаблонов с учетом их упорядоченности во времени (секвенциальный анализ).

Итак, еще раз сформулируем основные этапы технологии ДЭМ:

- 1) предварительная обработка сигналов;
- 2) анализ внутренней структуры сигналов;
- 3) интеллектуальный анализ сигналов.

Понятие эмпирической моды. Алгоритм ДЭМ

Эмпирическая мода [1-2, 7-10] (ЭМ, английское название IMF – Intrinsic Mode Function) – это функция, заданная непрерывно на интервале существования сигнала или дискретно в виде вектора отсчетов, имеющая в общем случае произвольную форму и произвольную аналитическую запись (если таковая существует), но при этом, однако, непременно удовлетворяющая двум необходимым условиям:

1. Общее число максимумов и минимумов такой функции (т. е. общее число экстремумов) должно быть строго равно числу нулей функции (в дискретном варианте задания нули могут быть найдены с использованием различных алгоритмов интерполяции [12]), либо отличаться от числа нулей по модулю не более, чем на единицу:

$$N_{\max} + N_{\min} = N_{\text{zero}} \pm 1 \text{ или } N_{\max} + N_{\min} = N_{\text{zero}}, \quad (2)$$

где N_{\max} , N_{\min} , N_{zero} – число максимумов, минимумов и нулей функции соответственно, не считая краевые отсчеты сигнала, которые в некоторых случаях могут оказаться единственными экстремумами сигнала (при этом имеем случай монотонно возрастающей или монотонно убывающей функции на всей ее области определения);

2. *Локальное (мгновенное) среднее значение функции*, определенное в виде полусуммы двух *огбающих*: верхней, полученной путем интерполяции найденных локальных максимумов, и нижней, полученной путем интерполяции найденных локальных минимумов, – должно быть меньше или равно заранее определенному пороговому значению η , зависящему от машинной точности ε и погрешностей, связанных с получением, преобразованием и передачей сигнальной информации. В качестве средства интерполяции чаще всего используются *кубические сплайны*.



Добиться точного равенства нулю локального среднего значения в каждый момент времени невозможно по ряду объективных причин. К их числу относятся вычислительные погрешности (связанные с особенностями машинной арифметики с плавающей запятой), плохая обусловленность [12] систем уравнений, на основе которых осуществляется расчет коэффициентов сплайнов, а также особенности самого сигнала (например, краевые эффекты – сильные осцилляции интерполирующей функции на краях, благодаря которым вблизи краев локальное среднее практически всегда отличается от нуля). Аналитически данное условие записывается в виде:

$$\frac{U(k) + L(k)}{2} \leq \eta, \quad k = \overline{1, N}, \quad (3)$$

где $U(k)$ и $L(k)$ – значения верхней и нижней огибающих сигнала в k -й момент времени (k – номер отсчета сигнала); N – общее количество сигнальных отсчетов; η – некоторый порог, устанавливаемый экспертом и обычно принимающий близкие к нулю значения.

Равенство числа экстремумов и нулей с точностью до 1 необходимо для того, чтобы ЭМ была *узкополосной* функцией, т. е. удовлетворяла следующему условию:

$$\Delta\omega \ll \omega_0, \quad (4)$$

где $\Delta\omega$ – эффективная ширина Фурье-спектра; ω_0 – несущая частота. Это дает преимущества при частотной локализации, так как мера узкополосности функции, как будет показано далее, связана с числом экстремумов и нулей.

Условие, касающееся равенства нулю полусуммы двух интерполированных огибающих, имеет свою физическую интерпретацию. Из него следует, что ЭМ является стационарной функцией относительно своего локального среднего значения, которое не превосходит некоторого порога в любой момент времени. Кроме того, для достижения этого условия ЭМ должна иметь положительные значения в точках максимумов и отрицательные значения в точках минимумов, поскольку иначе для отдельных моментов времени условие может не выполняться. Наконец, важно отметить, что ЭМ в общем случае обладает одновременно и амплитудной, и частотной модуляциями [1-2,6]. Закон амплитудной модуляции может быть установлен из огибающих, полученных интерполяцией экстремумов или на основе преобразования Гильберта (как квадратный корень из суммы квадратов отсчетов исходного сигнала и сигнала, сопряженного по Гильберту), а закон частотной модуляции устанавливается на основании закона изменения мгновенной частоты. В качестве информативного параметра на спектре Гильберта выступает цвет (цвет определяется значением наносимой абсолютной величины), так как именно по его перепадам можно определить соответствующие зависимости (изменение амплитудных соотношений, амплитудные модуляции, частотные модуляции и др.).

Понятие огибающей, введенное для определения ЭМ, всегда рассматривается с учетом выбранного вида интерполяции. Как было сказано, используется в основном кубическая сплайн-интерполяция, а также эрмитова интерполяция. У сплайнов вообще и кубических в част-

ности, есть ряд преимуществ перед другими функциями, обладающими, как и они, высокой степенью гладкости. Во-первых, по сравнению с полиномиальными функциями у сплайнов нет эффекта “раскачивания” – сильных осцилляций интерполирующей функции в случае, если фрагмент (участок сигнала) имеет заведомо не полиномиальную природу. Кроме того, кубические сплайны являются непрерывными (подразумевается равенство значений на стыках) и дважды дифференцируемыми. Последнее означает, что сам сплайн не имеет острых углов (определенность первой производной), и в любой точке определен радиус его кривизны (определенность второй производной) [12], что дает возможность оценивать интенсивность осцилляций самой ЭМ. Но самым важным является то, что кубический сплайн минимизирует осциллирующее поведение функции – т. е. из всех дважды дифференцируемых, непрерывных на некотором интервале $[a; b]$ функций $f(k)$, интерполирующих заданную совокупность точек $\{(x_k, y_k)\}_{k=1}^N$, кубический сплайн меньше всего осциллирует.

Что касается самого алгоритма ДЭМ (собственно метода извлечения самих ЭМ из исходного сигнала), он подробно изложен в [1-2, 7-10]. Общее число ЭМ M приближенно определяется как $M = \log_2 N \pm 1$, где N – число отсчетов сигнала. Сходимость алгоритма основана на том, что при переходе к следующей компоненте общее число экстремумов уменьшается примерно в 2 раза (согласно интерпретации диадического банка фильтров [1-2, 7-10]), и в конце концов не будет достаточного количества экстремумов для формирования огибающей. Это и будет означать завершение работы алгоритма.

Основные недостатки ДЭМ со сплайновой интерполяцией огибающих. ДЭМ с параболической интерполяцией огибающих

Классический алгоритм ДЭМ, описанный выше, использует кубические сплайны в качестве средства интерполяции огибающих, что гарантирует относительно небольшую вычислительную сложность (при определении коэффициентов полиномов третьей степени на участках между локальными экстремумами) и точное аналитическое представление при расчетах. Основные недостатки – явления “всплесков” (резкое превышение значением функции среднего уровня в промежутке между узлами интерполяции), краевые эффекты (осцилляции в окрестности точек начала и окончания сигнала) и отрицательное влияние ошибок в определении местоположения экстремумов. Именно последние вносят довольно серьезную погрешность при выполнении ДЭМ, поскольку при их неточном определении возможно явление т.н. *избыточности* (возникновение в разложении компонент, являющихся следствием неточностей выполнения алгоритма ДЭМ, а не структурных особенностей сигнала) и искажения структуры самих компонент. *Параболическая интерполяция* огибающих направлена на более точное определение местоположения экстремумов и препятствие появлению ложных компонент, не имеющих физического смысла (т.е. являющихся неинформативными для конкретного сигнала). Поясним сказанное следующим образом. Одним из про-

стейших и наиболее часто используемых алгоритмов определения экстремумов является следующий.

Обозначим $s[k]$ как текущий отсчет сигнала, а $s[k-1]$ и $s[k+1]$ как предыдущий и последующий отсчеты соответственно. Для поиска экстремумов необходимо проверить следующие условия:

1) если $s[k] > s[k-1]$ и $s[k] \geq s[k+1]$ или $s[k] \geq s[k-1]$ и $s[k] > s[k+1]$, то отсчет с номером k – максимум;

2) если $s[k] < s[k-1]$ и $s[k] \leq s[k+1]$ или $s[k] \leq s[k-1]$ и $s[k] < s[k+1]$, то отсчет с номером k – минимум;

3) в противном случае отсчет с номером k не является ни максимумом, ни минимумом.

Таким образом, данный метод поиска экстремумов заключается в нахождении окрестности, соответствующей максимуму или минимуму. Соответственно точка, расположенная посередине окрестности, либо превышает (в случае максимума), либо уступает (в случае минимума) значениям в соседних точках. Очевидно, максимальная ошибка e_{\max} определения экстремума составляет половину интервала дискретизации (если истинное значение экстремума лежит ровно посередине между двумя значениями отсчетов):

$$e_{\max} = T / 2, \quad (5)$$

где T – интервал (период) дискретизации. Данная ошибка может довольно существенно повлиять на точность самого разложения. В связи с этим предлагается новый метод определения экстремумов с использованием *параболической интерполяции*. Сам метод заключается в следующем:

1) Выбираются 3 первых отсчета исходного сигнала $s(k)$.

2) По трем отсчетам строится квадратичный полином вида $ak^2 + bk + c$, где a, b, c – коэффициенты подлежащие определению (для их нахождения необходимо решить систему из трех уравнений с тремя неизвестными).

3) Если коэффициент при старшей (второй) степени отличен от нуля ($a \neq 0$), то среди трех точек есть экстремум, который располагается посередине (вторая точка в последовательности из трех точек). Координата экстремума уточненным способом определяется как координата вершины параболы.

4) Если коэффициент при старшей степени равен нулю, то все 3 точки лежат на одной прямой (с положительным или отрицательным наклоном), и среди них нет

экстремума. Далее необходимо сдвинуть скользящее окно (включающее 3 точки, среди которых ищут экстремальную) на 1 точку вправо (т.е. в новом окне сохраняются 2 точки из предыдущего окна и добавится одна новая) и повторить ту же последовательность действий.

Следует отметить, что на практике коэффициент при старшей степени сравнивается не с нулем, а с некоторым порогом по той причине, что всегда присутствуют погрешности в вычислениях (сами вычисления проводятся не в аналитических, а численных системах, например, в системе MATLAB [12]). Таким образом, условие наличия экстремума среди трех точек выглядит как $|a| \geq \varepsilon$ – наличие экстремума, $|a| < \varepsilon$ – отсутствие экстремума;

В данном методе точность определения экстремумов повышается за счет определения их местоположения как *координаты вершины параболы* x_0 :

$$x_0 = -\frac{b}{2a}. \quad (6)$$

Огибающая в данной модификации метода ДЭМ будет состоять из кусочно-кубических участков, как и в классическом методе ДЭМ, однако определение экстремумов выполняется более точно за счет вычисления координаты вершины параболы.

Одним из базовых положений (своего рода постулатов) метода ДЭМ является следующий. Если ДЭМ применяется к функции, которая заведомо является ЭМ (т.е. удовлетворяет двум необходимым условиям отнесения к классу ЭМ), то в результате разложения не должно получаться иной функции, кроме самой ЭМ. Например, применение ДЭМ к мультигармоническому сигналу не должно порождать ничего другого, кроме содержащихся в нем гармоник. Однако классическая ДЭМ в некоторых случаях приводит к появлению побочных (ложных) компонент в разложении, очень близких по значениям к нулю. Это и есть уже упоминавшаяся избыточность ДЭМ, на борьбу с которой и направлены параболическая интерполяция и уточнение местоположения экстремумов. Ниже сказанное проиллюстрировано на примере моногармонического и мультигармонического сигналов.

Как видно из рис. 1, в результате разложения получаются 4 компоненты вместо одной (при этом значения двух последних компонент и результирующего остатка, как видно из рисунка, очень малы, но избыточность в разложении все равно присутствует).

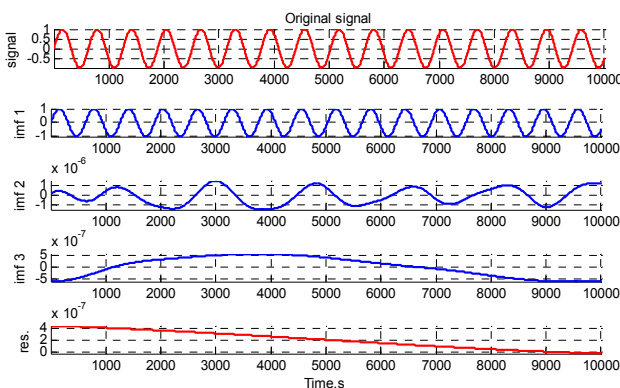


Рис. 1. Пример классической ДЭМ моногармонического сигнала (вверху) с извлечением 3-х ЭМ и результирующего остатка

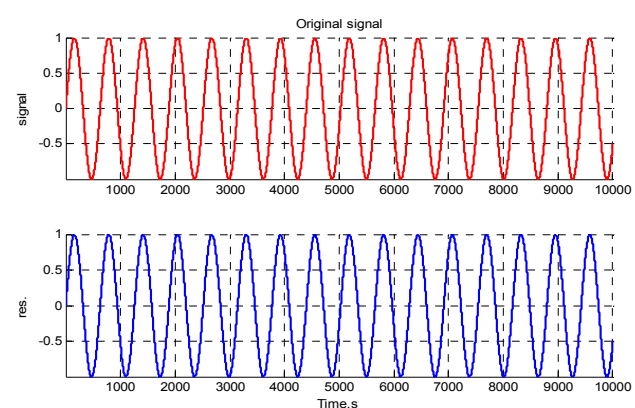


Рис. 2. Пример параболической ДЭМ моногармонического сигнала (вверху) с извлечением 1-й ЭМ

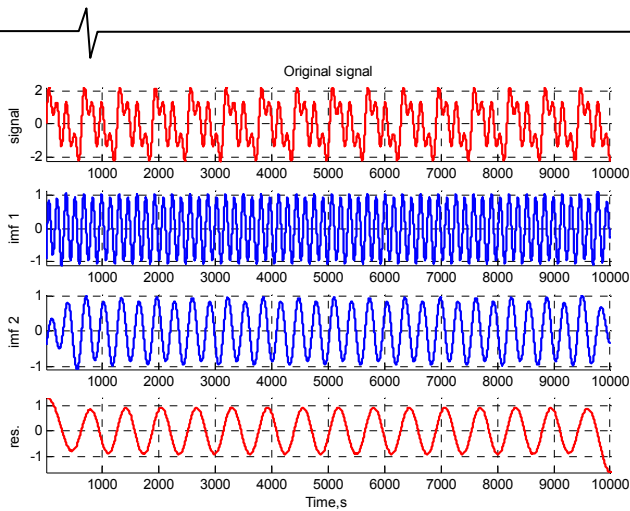


Рис. 3. Пример параболической ДЭМ мультигармонического сигнала (вверху) с извлечением 3-х ЭМ и результирующего остатка

На рис. 2 виден результат параболической ДЭМ и одна компонента, в точности соответствующая исходной гармонике. Далее рассмотрим пример с тремя гармониками.

Из рис. 3 видно, что параболическая ДЭМ порождает только 3 ЭМ, совпадающие с тремя гармониками, что и следовало ожидать. Небольшие искажения в форме вызваны краевыми эффектами, влиянием шага дискретизации и критерием остановки процесса отсеивания [1-2, 7-10].

Вместе с тем классическая ДЭМ (рис. 4) дает 6 компонент (что в 2 раза больше) и наглядно показывает избыточность в разложении гармонического сигнала.

Из этих примеров наглядно видны преимущества параболической ДЭМ при борьбе с избыточностью.

Очистка сигналов от шума на основе ДЭМ

В данном разделе будут продемонстрированы преимущества параболической ДЭМ при решении задачи очистки сигналов от шума. Вся процедура будет основана на параболической ДЭМ исходного сигнала, классификации всех извлеченных ЭМ и применении теории регрессии. Для *абсолютно произвольного* сигнала все ЭМ можно разбить на две большие группы:

- 1) основные ЭМ (включая шумовые ЭМ и ЭМ-шаблоны);
- 2) трендовые ЭМ (включая компенсирующие ЭМ и истинные трендовые ЭМ).

Основные ЭМ в разложении всегда имеют четкий физический смысл и отражают внутреннюю структуру и особенности, свойственные данному конкретному сигналу. К их числу относятся шумовые ЭМ и ЭМ-шаблоны. Появление в разложении первых объясняется наличием в исходном сигнале шума, а вторые связаны непосредственно с самим полезным сигналом и входящими в него компонентами. В качестве поясняющего примера можно привести следующий: при разложении мультигармонического сигнала с аддитивным шумом получается несколько шумовых ЭМ (они располагаются на начальных уровнях, и их число определяется интенсивностью шума, его частотными свойствами и общим числом отсчетов) и ЭМ-шаблоны (в разложении следуют за шумовыми ЭМ), к числу которых относятся соответствующие

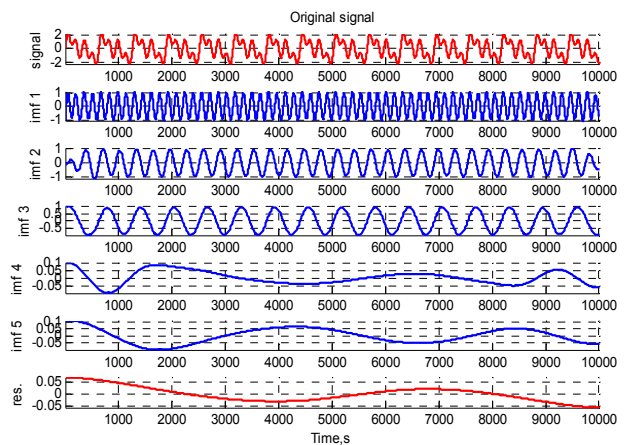


Рис. 4. Пример классической ДЭМ мультигармонического сигнала (вверху) с извлечением 5-и ЭМ и результирующего остатка

элементарные гармоники, образующие незашумленный мультигармонический сигнал. Понятие “ЭМ-шаблон” подразумевает некоторую функцию, характерную для данного типа сигналов, например, любую из элементарных гармоник для мультигармонического сигнала или, например, ЛЧМ-сигнал и элементарную гармонику в случае наличия их аддитивной смеси с шумом.

Трендовые ЭМ являются медленно меняющимися функциями (обычно описываются полиномами невысокой степени или экспоненциальной функцией), характеризующими изменение во времени локального среднего значения сигнала. Среди них, согласно вышеприведенной классификации, выделяют *истинные трендовые ЭМ* (всегда имеющие физический смысл), описывающие истинную динамику среднего значения, и так называемые *компенсирующие ЭМ*, которые могут возникать в разложении в том случае, если тренд в сигнале отсутствует. Истинные трендовые ЭМ появляются, например, при разложении суммы гармонического сигнала и полиномиального тренда (полинома). Компенсирующие (ложные) ЭМ – результат несовершенства самого алгоритма ДЭМ, критериев остановки процесса отсеивания, неточностей при вычислениях (ошибок округления и пр.). Их появление не связано с физическими или математическими особенностями рассматриваемых сигналов, а объясняется только лишь несовершенством вычислительной процедуры. Компенсирующие ЭМ обычно создают избыточность в разложении, а их название объясняется тем, что в сумме они дают функцию, очень близкую к нулю, т. е. компенсируют друг друга. Одной из наиболее типичных причин появления компенсирующих ЭМ являются побочные эффекты интерполяции огибающих, такие как появление “всплесков”, сильные осцилляции вблизи концов сигнала (краевые эффекты). Для частичного преодоления вышеперечисленных проблем существуют специальные методы: подавление краевых эффектов с помощью зеркального отображения экстремумов, наиболее близко расположенных к границам сигнала [1,2], тестирование различных критериев остановки процесса отсеивания [1-2], уточнение местоположения экстремумов для построения огибающих и пр.

Рассмотрим новый подход к выявлению всех рассмотренных типов ЭМ с целью идентификации шумовой составляющей, характерных сигнальных составляющих и тренда. Для этого модель исходного сигнала представим в следующем виде [14]:

$$s = x\beta + e, \quad (7)$$

где s – исходный сигнал; x – матрица регрессоров; β – вектор неизвестных коэффициентов, определяющих удельный вес регрессоров при описании исходного сигнала и подлежащих оцениванию; e – белый гауссовский шум с параметрами $N(0; \sigma^2)$.

Теперь, используя результат разложения с помощью ДЭМ, запишем связь между всеми извлеченными ЭМ и исходным сигналом на основании свойства полноты разложения:

$$s(k) = \sum_{i=1}^{M-1} c_i(k) + r_M(k), \quad (8)$$

где $r_M(k)$ – результирующий остаток в разложении. Из первого слагаемого правой части можно выделить отдельно самую первую ЭМ, которая будет являться приближением шума, присутствующего в сигнале, поскольку вследствие наиболее сильно выраженного по сравнению со всеми остальными ЭМ высокочастотного характера первую ЭМ можно трактовать как приближение исходного шума:

$$s(k) = c_1(k) + \sum_{i=2}^{M-1} c_i(k) + r_M(k) = e(k) + \sum_{i=2}^{M-1} c_i(k) + r_M(k). \quad (9)$$

Теперь сделаем еще одно уточнение. Поскольку точная сходимость суммы всех ЭМ и результирующего остатка к исходному сигналу математически строго не доказана (эта сходимость обычно рассматривается в инженерном смысле, т. е. абсолютная разность между исходным сигналом и суммой всех извлеченных ЭМ есть малое число, обычно порядка $10^{-16} \dots 10^{-10}$), введем некоторые весовые коэффициенты для каждой ЭМ (по умолчанию все веса единичные), которые будем определять в соответствии с методом наименьших квадратов (МНК). При этом должна улучшиться точность представления (восстановления) исходного сигнала полученными из него же компонентами:

$$\begin{aligned} s(k) &= c_1(k) + \sum_{i=2}^{M-1} \beta_i c_i(k) + \beta_M c_M(k) = \\ &= e(k) + \sum_{i=2}^{M-1} \beta_i c_i(k) + \beta_M c_M(k), \end{aligned} \quad (10)$$

где β – вектор весовых коэффициентов. Данное выражение можно переписать в более общем векторно-матричном виде:

$$s = C\beta + e, \quad (11)$$

где C – матрица, в столбцах которой хранятся отсчеты ЭМ (число столбцов на единицу меньше общего числа функций в разложении, так как первая ЭМ выступает как приближение исходного шума и не включается в эту матрицу). МНК-оценка вектора коэффициентов β определяется формулой:

$$\hat{\beta} = (C^T C)^{-1} C^T s. \quad (12)$$

Введение данных весовых коэффициентов и их дальнейшее вычисление (например, по МНК) может использоваться не только для повышения точности при восстановлении сигналов, но и для придания ДЭМ смысла “базиса”. Как было сказано, базис из ЭМ является апостериорным, следовательно, для каждого сигнала он индивидуален и, будучи полученным однажды, не может использоваться для представления другого сигнала. Однако при работе с однотипными сигналами (например, гармоники с разными амплитудами, но одинаковыми частотами) данные весовые коэффициенты могут применяться для получения нового базиса из старого путем умножения ранее извлеченных компонент на найденные веса. Для коэффициентов можно сформировать интервальную оценку, т. е. найти границы интервала, в который конкретный коэффициент регрессионной модели попадает с заданной доверительной вероятностью P . Можно показать, что границы доверительного интервала (ДИ) определяются следующим образом:

$$\beta_i \in \left[\hat{\beta}_i - t_\gamma \hat{\sigma}_e \sqrt{\{CC^T\}^{-1}_{ii}}; \hat{\beta}_i + t_\gamma \hat{\sigma}_e \sqrt{\{CC^T\}^{-1}_{ii}} \right], \quad (13)$$

где $\beta_i, \hat{\beta}_i$ – истинное значение коэффициента и его оценка соответственно, а t_γ – квантиль распределения Стьюдента с $N - M - 1$ степенями свободы.

В классической задаче регрессии после нахождения соответствующих коэффициентов модели часто осуществляется их проверка на значимость, по результатам которой конкретный элемент модели с проверяемым коэффициентом либо сохраняется, либо исчезает. Аналогично эту задачу можно решить для всех найденных весовых коэффициентов. Для этого рассмотрим две альтернативные гипотезы, одна из которых предполагает незначимость данного коэффициента модели (нулевая гипотеза H_0), а вторая, соответственно, его значимость (ненулевая гипотеза H_1), т. е. его отличность от нуля. В данной ситуации нулевая гипотеза отвергается, поскольку все коэффициенты являются априорно значимыми, отличными от нуля, поскольку все ЭМ с теми или иными весами участвуют при восстановлении исходного сигнала. Однако нижеприведенную статистику, изначально предназначенную именно для проверки значимости коэффициентов модели, можно рассматривать как классификационную (т. е. как некоторую функцию, связанную с исходными данными), а рассчитываемые с ее помощью значения распределять по группам (каждая группа будет соответствовать одному из типов ЭМ) на основе некоторой дискриминирующей процедуры, например, кластер-анализа [5, 13]. Сама статистика имеет вид

$$T_{li} = C_H = \frac{|\hat{\beta}_i|}{\sqrt{\{V^{-1}\}_{ii} \hat{\sigma}_e}}, \quad (14)$$

где V – матрица ковариаций ЭМ, определяемая как $V = CC^T$ (при вычислении статистики используются диагональные элементы матрицы, обратной к ковариационной); $\hat{\sigma}_e$ – среднеквадратическое отклонение (СКО) шума, которое можно определить по первой ЭМ с использованием, например, робастной медианной оценки [4], устойчивой к наличию погрешностей и аномальных ошибок в данных.

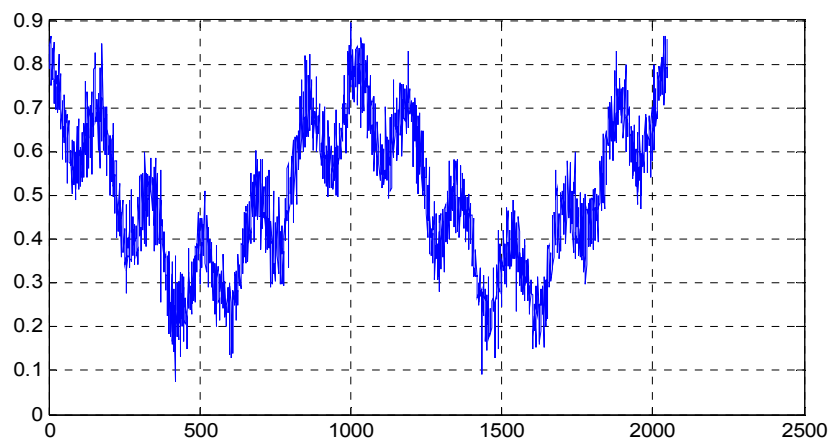


Рис. 5. Зашумленный волноподобный процесс

Таблица 1

Значения статистики T_1 для всех выделенных из сигнала ЭМ с помощью параболической ДЭМ

Номер ЭМ	β	$\{V^{-1}\}_{ii}$	Длина ДИ	$T_1 \cdot 10^{-4}$	<i>range</i>	<i>R</i>	Номер кластера
2	0.9126	0.9569	0.1519	19.7734	0.1669	0.1078	1
3	0.8841	1.8971	0.2139	13.6050	0.1019	0.0755	1
4	0.9506	3.2693	0.2808	11.1436	0.0646	0.0544	1
5	0.9155	6.1275	0.3844	7.8388	0.0513	0.0942	1
6	0.9989	0.0984	0.0487	67.4880	0.2378	0.4315	2
7	0.9977	0.0242	0.0242	135.9106	0.4108	0.8556	2
8	1.0003	0.0019	0.0068	482.9503	0.0273	0.0689	-

Выше (в табл. 1) проиллюстрирован пример расчетов для зашумленного волноподобного процесса с СКО шума, равным 0.05. Сам процесс показан на рис. 5. При использовании классической и параболической ДЭМ было получено 9 и 8 компонент соответственно. Медианная оценка СКО шума по первой ЭМ (для параболической ДЭМ) составляет 0.0472.

Для автоматического проведения классификации всех ЭМ и, соответственно, выявления шумовых ЭМ можно воспользоваться одним из алгоритмов кластер-анализа, например, алгоритмом *k*-средних (*k*-means) [5] с числом кластеров, равным 2. Число 2 выбрано для того, чтобы в итоге получить две независимых группы компонент: шумовые ЭМ, выделенные в табл. 1 жирным шрифтом, и остальные компоненты (ЭМ-шаблоны и трендовые ЭМ). Также можно использовать специальные алгоритмы (например, основанные на критериях качества разбиения [5]), позволяющие определить число кластеров вместо того, чтобы задавать его самим. Такая процедура может быть необходима для более тщательного анализа компонент, например, если требуется более тонкое разделение внутри группы ЭМ-шаблонов.

Помимо значений статистики T_1 , можно назвать и другие признаки для классификации ЭМ, а именно *длину ДИ* для соответствующих регрессионных коэффициентов. Также используются статистические характеристики такие, как *размах* компоненты *range* (разность между ее максимальным и минимальным значениями) и *коэффициент корреляции* с исходным сигналом *R*. Все эти признаки (значения классификационной статистики, размах и коэффициент корреляции с исходным сигналом) являются информативными для задачи кластер-анализа и используются при кластеризации *совместно*. В данном случае при

кластер-анализе использовались все 3 признака. Последняя компонента не включалась в анализ, поскольку она заведомо является трендовой (в силу особенностей ДЭМ) и не играет никакой роли при очистке от шума.

Результаты кластеризации на основе значений статистики соответствуют высказанным предположениям: компоненты с 1-й по 5-ю относятся к одному кластеру (1-я компонента априорно является шумовой), а все остальные – ко второму. В некоторых случаях, для более детальной классификации, можно задать число кластеров, равное 3, с целью выделить три типа ЭМ: шумовые, ЭМ-шаблоны и трендовые ЭМ.

В случае необходимости различения шумовых и трендовых компонент (если значения классификационной статистики оказываются для них близкими и кластер-анализ не позволяет их различить) возможно использование *показателя Херста H* [11], характеризующего степень регулярности функции и ее склонность к хаотическому поведению. Данная величина заключена в пределах [0; 1] и имеет два характерных диапазона значений. При $0 \leq H < 0.5$ компоненты являются антиперсистентными (с неустойчивым трендом, с кратковременной памятью) [11], т. е. наблюдается так называемая тенденция возврата к среднему. Значения из данного диапазона имеют все шумовые ЭМ. При $0.5 < H \leq 1$ компоненты являются персистентными (с устойчивым трендом, с долговременной памятью) [11], т. е. если имелась тенденция к возрастанию (убыванию), то она будет сохранена и в дальнейшем на некотором промежутке. Чем ближе значение *H* к 1, тем сильнее выражена данная тенденция и тем дольше она сохраняется. Трендовые компоненты имеют показатель Херста, заключенный в этом диапазоне. Наконец, еще одним сред-

ством различения компонент является энтропия Шеннона [3-4], которая для шумовых компонент, как для более хаотичных и близких к случайным, дает значения большие, чем для трендовых.

Ниже показан график параболической ДЭМ для волноподобного процесса. Из графика видно, что первые 5 компонент являются шумовыми и в совокупности со-

ставляют исходный шум. 6-я и 7-я компоненты относятся к ЭМ-шаблонам (они представляют собой 2 гармоники), а последняя компонента является трендовой. В сумме 6-я, 7-я и 8-я компоненты составляют исходный сигнал. Из данного разложения также видно, что волноподобный процесс имеет гармоническую природу, т.е. является линейной комбинацией гармоник и тренда.

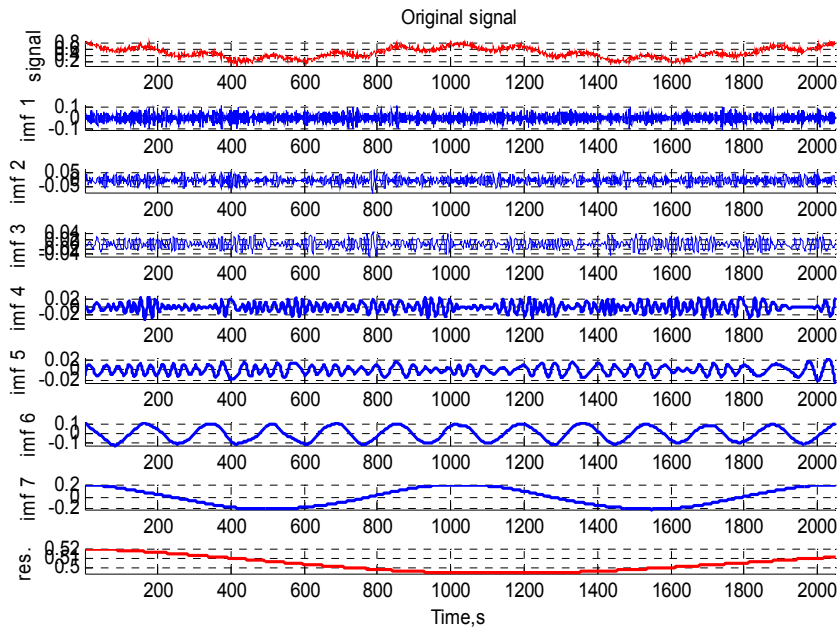


Рис. 6. Параболическая ДЭМ волноподобного процесса

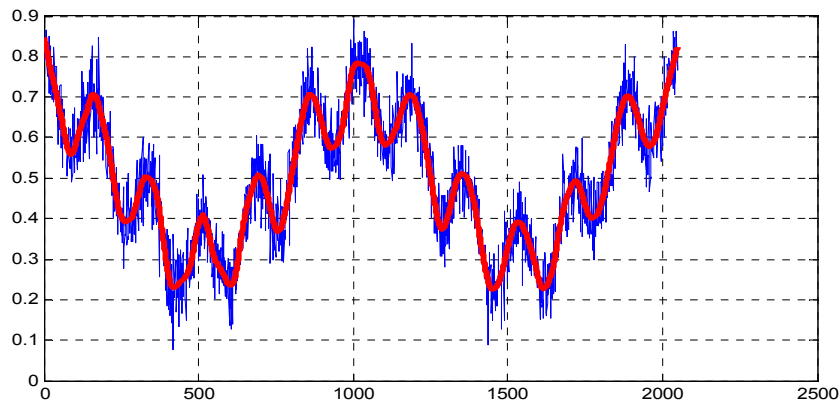


Рис. 7. Зашумленный волноподобный процесс и результат восстановления исходного волноподобного процесса с помощью параболической ДЭМ и теории регрессии

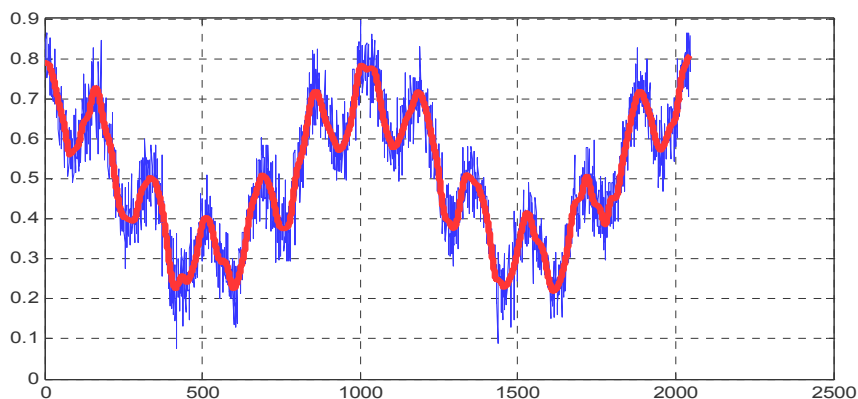


Рис. 8. Зашумленный волноподобный процесс и результат восстановления исходного волноподобного процесса с помощью классической ДЭМ и теории регрессии



Выше, на рис. 7, показан результат восстановления незашумленного волноподобного процесса (жирной линией поверх зашумленного сигнала). На основании данного графика и его сравнения с графиком на рис. 8, где показан результат работы с использованием классической ДЭМ, можно сделать вывод о высокой точности восстановления полезного сигнала. Одним из визуальных критериев является гладкость кривой, а одним из количественных критериев – СКО выделенного шума.

Из сравнения графиков видно, что точность восстановления как в центральной части кривой, так в местах изгибов выше (лучше) при параболической интерполяции, и, соответственно, график восстановленного сигнала не имеет локальных изломов (флуктуаций, нарушений гладкости). В качестве критерия точности восстановления исходного волноподобного процесса будем использовать, как было сказано ранее, СКО выделенного шума. Для случая параболической ДЭМ СКО оказывается равным 0.0490, а в случае классической ДЭМ – 0.0484 (истинное значение СКО шума – 0.05), что подтверждает более высокую точность параболической ДЭМ.

Данная статья и проведенная научная работа поддержаны в рамках федеральной целевой программы “Научные и научно-педагогические кадры инновационной России” на 2009-2013 годы, а также Министерства образования и науки РФ (контракт 13.G25.31.0054)

Литература

1. Huang N. E., et al. The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis // Proc. Royal Soc. London. 1998. Vol. 454. P. 903-995.
2. Huang N. E., Shen S. S. P. Hilbert-Huang Transform and Its Applications. World Scientific, 2005. 350 p.
3. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в Matlab. 3-е изд. М.: ДМК, 2008. 304 с.
4. Малла С. Вейвлеты в обработке сигналов. М.: МИР, 2005. 671 с.
5. Барсегян А.А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Mining. СПб.: БХВ-Петербург, 2004.
6. Сергиенко А. Б. Цифровая обработка сигналов. СПб.: Питер, 2007.
7. Клионский Д. М., Орешко Н. И., Геппенер В. В. Декомпозиция на эмпирические моды и ее использование при анализе дробного броуновского движения // Цифровая Обработка Сигналов. 2008. № 3. С. 37-45.
8. Klionsky D. M. Oreshko N. I., Geppener V. V. Empirical Mode Decomposition in Segmentation and Clustering of Slowly and Fast Changing Non-Stationary Signals // Pattern Recognition and Image Analysis. 2009. Vol. 19. № 1. P. 14-29.
9. Клионский Д. М., Орешко Н. И., Геппенер В. В. Новый

подход к автоматизированному выявлению шаблонов в телеметрических сигналах на основе декомпозиции на эмпирические моды // Научные ведомости Белгородского государственного университета. – 2009. - № 11. С. 118-128.

10. Клионский Д.М., Неунывакин И. В., Орешко Н. И., Геппенер В. В. Декомпозиция на эмпирические моды и ее применение для идентификации информативных компонент и прогнозирования значений сигналов с использованием нейронных сетей // Нейрокомпьютеры. 2010. № 6. С. 43-50.
11. Петерс Э. Э. Фрактальный анализ финансовых рынков // изд. М.: Интернет-Трейддинг, 2004. – 304 с.
12. Мэтьюз Д.Г., Финк К.Д. Численные методы. Использование Matlab // изд. дом “Вильямс”, Москва-Санкт-Петербург-Киев. - 2001.
13. Васильев А.В., Геппенер В.В., Жукова Н.А., Тристанов А.Б., Клионский Д.М. Применение алгоритмов кластеризации и классификации в задачах обработки и интерпретации телеметрической информации / тез. докл. междунар. конф. DSPA-2007, Москва, 25-27 марта, 2007.
14. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976.

EMPIRICAL MODE DECOMPOSITION WITH THE PARABOLIC ENVELOPE INTERPOLATION IN THE TASKS OF SIGNAL DENOISING

Dmitrij Klionskiy, Nikolay Oreshko, Vladimir Geppener

The present paper is concerned with a new adaptive method of non-stationary signal analysis and processing – Empirical mode decomposition (EMD). This method allows one to effectively solve the problems of preprocessing (denoising, detrending, etc.), structural analysis (multiresolution analysis, time-frequency analysis) and intellectual analysis of signals (segmentation, clustering). The investigated modification of EMD – EMD with parabolic envelope interpolation – helps obtain even better accuracy in signal processing than the classical EMD. The main advantages of the classical EMD are provided and a wide range of opportunities for dealing with one of the most significant problems of preprocessing – denoising – are shown. The underlying ideas of the suggested method include the usage of the regression theory and a general classification of components extracted from the original signal. The experiments carried out for multiharmonic signals have confirmed the high accuracy of the given method with regard to signal denoising in comparison with the classical EMD.