I. A. Shcherbakov
Saint Petersburg Electrotechnical University «LETI»

A. V. Ponomarev
*Saint Petersburg Institute of Informatics and automation
of the Russian Academy of Sciences (SPIIRAS)*

SEMANTAGS: SEMANTIC TAGGING OF DOCUMENTS BASED ON CROWDSOURCING

Enriching objects (e.g., Internet resources) with semantic labels (or, tags) characterizing the content of these objects allows to significantly improve the search results quality and, in general, contributes to more efficient processing of information. However, fully automatic semantic tagging does not always give an adequate result due to existing constraints in the algorithmic processing of natural language, while manual tagging is rather laborious, especially if the number of objects is large. The latter can be partly alleviated in some cases by using the crowdsourcing technology. The article proposes SemanTags system (web-service), which allows to leverage crowdsourcing to tag arbitrary objects (scientific articles, Internet pages, etc.) with classes of some problem-oriented ontology written in OWL 2. The paper describes the architecture of the system, as well as the new quality assurance mechanism, adapted for social tagging of objects by classes of an ontology.

Crowdsourcing, crowd computing, ontologies, OWL 2, taxonomy, semantic tagging, semantic search

УДК 681.518.5

А. А. Алексеев, А. М. Спиваковский
*Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В. И. Ульянова (Ленина)*

Применение гибридного вейвлет-спектрального метода для выявления диагностических признаков

Рассматриваются актуальные информационно-вычислительные аспекты в решении задач диагностики сложных объектов и процессов управления. Для оперативного формирования пространства диагностических признаков минимальной размерности предлагается гибридный подход на основе быстрых ортогональных вейвлет-спектральных преобразований, адаптируемых к виду характеристик, описывающих состояния диагностируемого объекта. Данный подход обеспечивает высокую скорость сходимости ряда для анализируемого класса сигналов и характеризуется высокой степенью избирательности преобразования в приспособленном базисе. Достоинством предлагаемого гибридного подхода является однозначность вычисления коэффициентов разложения при любой размерности исходных данных и любом усечении ряда разложения, что позволяет увеличивать точность восстановления аппроксимируемой функции путем добавления компонентов из отсеченной части ряда.

Вейвлет-спектральные преобразования, адаптивные базисы, диагностика объектов управления, диагностические признаки

Нормальное функционирование технической системы управления обычно сопровождается периодическим, экспоненциальным, пилообразным и т. д. характером наблюдаемых колебаний и устойчивым во времени поведением их характеристик. Напротив, аномальное функционирование, обусловленное проявлением каких-либо дефектов объекта, сопровождается возникновением новых составляющих в наблюдаемых сигналах, таких как всплески, линейные тренды, прямоугольные или треугольные импульсы, а также шумовые реализации аддитивного или мультипликативного вида [1], Поэтому для выявления информации, позволяющей оценить текущее состояние диагностируемого объекта, необходимо наличие диагностических признаков как для нормального течения процесса, так и для аномального его поведения [2].

В реальных условиях функционирования технической системы для извлечения диагностической информации использование детерминированных сигналов наталкивается на значительные трудности вследствие воздействия помех и шумов естественного и/или искусственного происхождения. Таким образом, задача выявления диагностической информации существенно усложняется, поскольку необходимо учитывать случайный характер первичной информации, используемой для формирования диагностических признаков [3].

Одним из наиболее эффективных инструментов для оперативного формирования пространства диагностических признаков минимальной размерности является развиваемый авторами гибридный подход. В его математической и алгоритмической основе лежат быстрые вейвлет-спектральные преобразования [4]–[10]. Одним из главных достоинств вейвлет-преобразования является возможность позиционирования местоположения возникшего дефекта, т. е. осуществления его локализации.

При диагностической оценке текущего состояния технической системы определяющим является обеспечение оперативности в решении задач фильтрации, контроля формы сигналов, обнаружения, распознавания образов.

Наилучшим способом настроен на фильтрацию шумов и помех аппарат спектральных преобразований. Последовательное воздействие вейвлет- и спектральных операторов на диагностируемый процесс при реализации алгоритмов быстрых ортогональных преобразований с возможностью их подстройки позволяет получать актуальную и информативную диагностическую информацию.

На кафедре автоматики и процессов управления СПбГЭТУ «ЛЭТИ» разработан и эффективно используется в различных приложениях математический инструментарий, в основе которого лежат обобщенные и адаптируемые ортогональные преобразования с быстрыми схемами вычислений [6]. Единая форма представления соответствующего аппарата позволяет не только формировать многообразие ортогональных базисов (в том числе и известные базисы Фурье, Уолша, Хаара), но и выбирать из них наиболее подходящие требованиям задачи и критериям качества обработки сигналов адаптируемые (или приспособленные) базисы с алгоритмами быстрых преобразований. Отличительной особенностью такого подхода является относительно несложное формирование вейвлет- и спектральных операторов с последующей их адаптацией для решения задачи сравнения эталонных и текущих диагностических признаков путем выявления степени их отличия в режимах нормального и аномального функционирования.

На основании сравнения результатов по точности приближения и скорости сходимости рядов разложений в различных метриках приближения осуществляется выбор наиболее подходящего базиса разложения и создается возможность формирования информативных диагностических признаков, обеспечивающих более достоверное и надежное выявление дефектов.

Отметим что, если известные базисы Фурье, Уолша и Хаара имеют универсальный неизменный характер поведения своих функций, то формируемые приспособленные базисы, напротив, представляют собой наборы функций с варьируемыми мгновенными значениями в зависимости от вида анализируемых сигналов. Очевидно, что такая особенность приспособленных базисов позволяет использовать их только для конкретного класса сигналов. Для других классов сигналов требуется пересчет соответствующих параметров базисных функций – степеней свободы – и формирование новой совокупности функций. Особенно эффективно использование базисных функций для анализа и диагностики сигналов, близких по форме, но незначительно отличающихся на малых локальных отрезках времени или во всех точках интервала наблюдения.

Реализация алгоритмов быстрых преобразований в базисах Фурье, Уолша, Хаара базируется на факторизованном матрично-ядерном представлении соответствующих операторов. Возможность унифицировать структуру факторизованной матрицы позволила сформулировать задачу разработки обобщенного ортогонального оператора, который включает в себя матрицы как известных базисов, так и новые матрицы, принимающие ортогональный вид. Простейшим вещественным ортогональным ядром такого оператора является матрица 2 × 2:

$$[V] = \sqrt{2}\begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}.$$

Для базиса Уолша существует ядро с углом-параметром φ = π/4. Совокупность ортогональных ядер с углами-параметрами π/4 и 0 приводит к базису Хаара.

Обобщенное ортогональное ядро [6], включающее в себя как вещественные, так и комплексные значения элементов, имеет вид:

$$[V] = \sqrt{2}\begin{bmatrix} \cos\varphi_{ij} & w^{-k\gamma}\sin\varphi \\ \sin\varphi_{ij} & -w^{-k\gamma} - \cos\varphi_{ij} \end{bmatrix},$$

$$i = \overline{1, 2^{n-1}},\ \ j = \overline{1, n},\ \ k = 0, 1, 2, \ldots,\ \ \gamma = \pi / 2^{n-1}.$$

Угол φ может быть произвольным как внутри факторизованной матрицы, так и для всех факторизованных матриц. То же относится и к показателям степени для дискретных комплекснозначных экспоненциальных функций. Такое большое разнообразие матриц позволяет выбрать из них наиболее подходящие к форме анализируемого сигнала.

Спектральные методы более эффективны для выявления периодических компонент процессов, но, как правило, мало пригодны для обнаружения их локальных особенностей. С другой стороны, вейвлет-метод целесообразно использовать для получения локальных особенностей сигналов, но не для выявления их периодических составляющих. Кроме этого, необходимо учитывать, что основная проблема в реализации собственно спектральных методов состоит в невысокой информативности спектральных характеристик для диагностирования зашумленных процессов, особенно, если дефекты скрыты в помехах. Напротив, вейвлет-метод эффективно выявляет локальные дефекты в помехах, но не для полигармонических процессов. При этом необходимо учитывать необходимость преобразования больших информационных потоков, которые, как правило, априорно не определены, что, в свою очередь, приводит к большим вычислительным затратам.

Поэтому основной целью вейвлет-спектральной обработки диагностируемых процессов является существенное снижение размерности обрабатываемых информационных массивов при незначительном ухудшении формы, вида и свойств за ограниченное время. Для достижения этой цели наиболее эффективно использование предложенных приспособленных (адаптируемых) базисов двух типов (ПБ-1, ПБ-2). Далее рассмотрены основные положения метода формирования приспособленных базисов, наилучшим образом адаптируемых к классу ритмических случайных процессов.

К ритмическим случайным процессам относятся квазипериодические колебания $x = \{x(l, t)\}$, где l – номер реализации ($l = 1, 2, 3, \ldots$), имеющие периодическое поведение статистических характеристик (среднего по множеству m, дисперсии D, автокорреляционной функции R). В тех случаях, когда информация о корреляционных свойствах случайного процесса отсутствует или ею можно пренебречь, на первый план выходит задача построения системы базисных функций, приспособленной к имеющимся статистическим характеристикам и в первую очередь к математическому ожиданию циклического случайного процесса: $m_x(t) = m_x(t + T)$.

Обычно формальное использование априорных данных осуществляется на основе математической модели изучаемого явления в виде статистических характеристик случайных процессов.

Введем следующие ограничения на исследуемые классы случайных процессов. Пусть анализируемые процессы представлены наборами реализаций с равномерным характером отклонений (рис. 1, *а*) и с локальным характером отклонений (рис. 1, *б*).

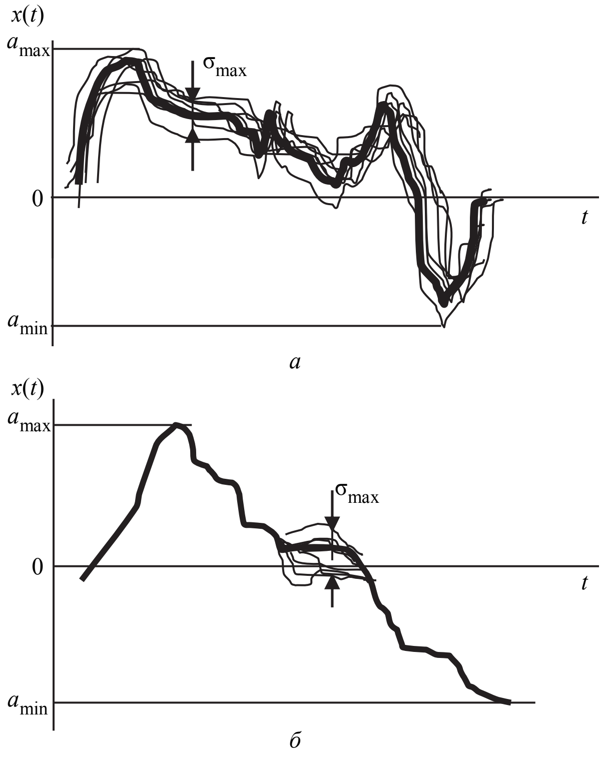

Рис. 1

При этом для данных процессов есть возможность зафиксировать (измерить) максимальное ($a_{\max}$) и минимальное ($a_{\min}$) амплитудные значения, а также вычислить максимальный разброс $\sigma_{\max}$ (среднеквадратичное значение) конкретных реализаций относительно среднего по множеству.

Будем считать компактным множеством реализаций такой их набор, при котором $\sigma_{max} <<$ $<< (a_{max} - a_{min})$.

В отношении рассматриваемых классов диагностируемых процессов с равновероятным (рис. 1, *а*) и локальным (рис. 1, *б*) характером отклонений реализаций от эталона необходимо отметить их нестационарные свойства относительно математического ожидания и/или дисперсии. Также важным является их происхождение, обусловленное физическими особенностями некоторого колебательного механизма с явно выраженной ритмикой колебаний, причем период T колебаний отдельных реализаций не остается фиксированным. Но при этом период эталона (в частности, математического ожидания) является величиной постоянной. Это ограничение необходимо для установления граничных моментов времени при формировании системы функций, приспособленной к данному классу случайных процессов.

Под приспособленным базисом здесь понимается система ортогональных векторов, первый из которых совпадает с эталоном – математическим ожиданием анализируемого случайного процесса, удовлетворяющего приведенным ранее условиям, а остальные векторы формируются исходя из требований по обеспечению наивысшей скорости сходимости ряда разложения в выбранной метрике приближения.

Рассмотрим методы формирования адаптируемых спектральных операторов с целью компактного представления ритмического случайного процесса $\{x(t)\}$ с известным математическим ожиданием $m_x(t) = m_x(t + T) = m$. Адаптируемость (приспособление) состоит в получении возможности спектрального представления некоторого эталона класса случайных процессов, интерпретируемого как реализация процесса в режиме нормального функционирования (норма), в виде его математического ожидания только одним ненулевым коэффициентом. При этом конкретные реализации анализируемого класса случайных процессов описываются минимумом значимых коэффициентов разложения при заданной точности восстановления (аппроксимации) в метриках равномерного C и среднеквадратичного L_2 приближений.

Процедура формирования приспособленного базиса заключается в следующем. На первом этапе для заданного счетного множества реализаций случайного процесса $\{X\} = \{x_{lt}\}$, $l = [1, L]$ отыскивается среднее по множеству, представленное в векторном виде

$$m = M\{x_{lt}\} = [m_i]^T,\ i = [1,N], \tag{1}$$

где T означает транспонирование вектора.

Далее на втором этапе, используя (1) в качестве эталона, ему сопоставляется в области коэффициентов разложения пространство минимальной размерности, т. е. один коэффициент, отличный от нуля.

Согласно теореме Парсеваля, величина данного коэффициента численно равна среднеквадратичному значению, вычисленному по известным отсчетам эталона. При разложении эталона по найденному адаптируемому базису количество нулевых коэффициентов равно $(N - 1)$, что эквивалентно наивысшей скорости сходимости спектра. Очевидно, что отклонения конкретных реализаций x_{it} от эталона $[m_i]^T$ во временной области будут приводить к появлению новых коэффициентов в области преобразования, а также к изменению величины основного ненулевого коэффициента в большую или меньшую сторону, причем степень этого изменения определяется характером и уровнем отличия сопоставляемых реализаций. Учет ненулевых коэффициентов позволяет при выборе величины порога в соответствующей метрике достигать заданной ошибки приближения.

Матричное уравнение для определения необходимых углов-параметров по заданным значениям отсчетов вектора $[m_i]^T = [m]^T$ и одному коэффициенту U_1, равному (в соответствии с теоремой Парсеваля) среднеквадратичному значению:

$$[U]^T = [U_1,\ 0,\ \ldots,\ 0\ ]^T = \\ = [(\sum m_i^2)^{1/2},\ 0,\ \ldots,\ 0\ ]^T. \tag{2}$$

На основе прямого ортогонального преобразования

$$[U]^T = (1/N)\ [\Psi]\ [m]^T,$$

получим аналогичное матричное уравнение для $[m]^T$ в случае вещественного базиса $[\Psi]$:

$$[m]^T = N\ [\Psi]^{-1}\ [U]^T = [\Psi]^{\mathrm{T}}\ [U]^T,$$

поскольку

$$[\Psi]^{-1} = (1/N)[\Psi]^T.$$

Выразим матрицу $[\Psi]^T = [\Phi]$ через обобщенное ортогональное ядро

$$[V_{rk}] = \begin{bmatrix} \cos\varphi_{rk} & \sin\varphi_{rk} \\ \sin\varphi_{rk} & -\cos\varphi_{rk} \end{bmatrix}.$$

При этом для наглядности выкладок и без потери общности будем рассматривать случай $N \times N = 2^3 \times 2^3$:

$$\begin{aligned}[\Phi] = &\left[(\varphi_{11}),(\varphi_{12}),(\varphi_{13}),(\varphi_{14})\right]\times\\ &\times\left[(\varphi_{21}),(\varphi_{22}),(\varphi_{23}),(\varphi_{24})\right]\times\\ &\times\left[(\varphi_{31}),(\varphi_{32}),(\varphi_{33}),(\varphi_{34})\right],\end{aligned} \quad (3)$$

где в матричных микроблоках использована условная запись матрично-ядерного представления вида:

$$\left[V_{ij}\right] = \begin{bmatrix}\cos\varphi_{ij} & \sin\varphi_{ij}\\ \sin\varphi_{ij} & -\cos\varphi_{ij}\end{bmatrix},\ i=[1,3],\ \ j=[1,4].$$

В (3) на основании условия (2) из всей матрицы [Φ] в восстановлении вектора $[m_i]^T$ участвуют только элементы первого столбца. В общем виде их можно получить перемножением факторизованных матриц в соответствии с (3). В результате получим совокупность уравнений для искомых углов-параметров:

$$\begin{aligned}m_1 &= U_1\cos\varphi_{11}\cos\varphi_{21}\cos\varphi_{31},\\ m_2 &= U_1\cos\varphi_{11}\cos\varphi_{21}\sin\varphi_{31},\\ m_3 &= U_1\cos\varphi_{11}\sin\varphi_{21}\cos\varphi_{32},\\ m_4 &= U_1\cos\varphi_{11}\sin\varphi_{21}\sin\varphi_{32},\\ m_5 &= U_1\sin\varphi_{11}\cos\varphi_{22}\cos\varphi_{33},\\ m_6 &= U_1\sin\varphi_{11}\cos\varphi_{22}\sin\varphi_{33},\\ m_7 &= U_1\sin\varphi_{11}\sin\varphi_{22}\cos\varphi_{34},\\ m_8 &= U_1\sin\varphi_{11}\sin\varphi_{22}\sin\varphi_{34}.\end{aligned}$$

Поделив соседние уравнения друг на друга, находим следующие выражения для углов $\{\varphi_{31}\}^T$, определяющие параметры ядер факторизованной матрицы $[G_3]$:

$$\operatorname{tg}\varphi_{31} = m_2/m_1;\ \operatorname{tg}\varphi_{32} = m_4/m_3;$$
$$\operatorname{tg}\varphi_{33} = m_6/m_5;\ \operatorname{tg}\varphi_{34} = m_8/m_7.$$

Далее, поступая аналогичным образом, можно найти для углов-параметров матриц $[G_1]$ и $[G_2]$ следующие зависимости:

$$\operatorname{tg}\varphi_{21} = \left(\frac{m_3^2+m_4^2}{m_1^2+m_2^2}\right)^{1/2},\ \ \operatorname{tg}\varphi_{22} = \left(\frac{m_7^2+m_8^2}{m_5^2+m_6^2}\right)^{1/2},$$

$$\operatorname{tg}\varphi_{11} = \left(\frac{m_5^2+m_6^2+m_7^2+m_8^2}{m_1^2+m_2^2+m_3^2+m_4^2}\right)^{1/2}.$$

В общем случае при $N = 2^n$ справедливы следующие соотношения для углов-параметров, входящих в ядра $[V_{ri}]$ матриц $[G_r]$:

$$\begin{gathered}\operatorname{tg}\varphi_{ni} = \frac{m_{2i}}{m_{2i-1}},\ i = [1, N/2];\\ \operatorname{tg}\varphi_{(n-1)i} = \left(\frac{m_{4i-1}^2+m_{4i}^2}{m_{4i-3}^2+m_{4i-2}^2}\right)^{1/2}, \\ i=[1,N/4];\ \operatorname{tg}\varphi_{1i} = \frac{\sum\limits_{i=N/2+1}^{N} m_i^2}{\sum\limits_{i=1}^{N/2} m_i^2},\ i=\left[1, N/2^j\right].\end{gathered} \quad (4)$$

Полученные численные значения углов-параметров (4) общим количеством $(N - 1)$, дополненные ненулевым коэффициентом разложения $U_1 = \left(\sum\limits_{i=1}^{N} m_i^2\right)^{1/2}$, позволяют по исходным N значениям отсчетов вектора $[m]^T$ однозначно синтезировать первую функцию приспособленного базиса, а остальные формировать целенаправленным образом, исходя из дополнительных степеней свободы.

Общее число степеней свободы для обобщенного ортогонального оператора равно $n2^{n-1}$, поэтому существует возможность варьирования $(n2^{n-1} - 2^n) = 2^{n-1}(n-2)$ степенями свободы, размещение которых в факторизованных матрицах $[G_r]$ позволяет учитывать тот или иной характер отклонений конкретных реализаций от эталона.

Углы-параметры φ_{rk} = var не зависят от отсчетов эталонного вектора $[m]^T$, что позволяет провести дальнейшую адаптацию базиса под те или иные требования (уменьшение объема памяти ЭВМ, сокращение времени вычислений, повышение скорости сходимости в заданной метрике, упрощение вида базисных функций).

В частности, для экономии памяти ЭВМ целесообразно принять условие

$$\varphi_{12} = \varphi_{13} = \varphi_{14} = \varphi_{11};\ \varphi_{23} = \varphi_{21};\ \varphi_{24} = \varphi_{22}. \quad (5)$$

При использовании закономерности (5) в расстановке углов-параметров ядер намного упрощается процедура их определения в случае необходимости уточнения разложения при увеличении размерности $N = 2^n$ исходных данных. При этом ранее вычисленные значения углов остаются без изменения и расчет новых углов-параметров производится только для добавляемых факторизованных матриц $[G_r]$.

Полученный ортонормированный приспособленный базис ПБ-1 адаптирован к равновероятным отклонениям от эталона – математического ожидания, и в записи через обобщенную форму прямого произведения [6] принимает вид

$$[\Phi]=\left[(\varphi_{11})_4\right]\left[(\varphi_{21})_2;(\varphi_{22})_2\right]\times$$
$$\times\left[(\varphi_{31}),(\varphi_{32}),(\varphi_{33}),(\varphi_{34})\right].$$

К числу самых важных требований в задаче адаптации базиса относится обеспечение максимально возможной скорости сходимости ряда разложения в выбранной метрике приближения. Решение этой задачи необходимо связывать с характером отклонений ансамбля реализаций исследуемого процесса от эталона класса.

Для локального отклонения реализаций от эталона на малом интервале наиболее высокую скорость сходимости в метрике C обеспечивает базис, функции которого отличаются локальным поведением на $[0, T]$ и равны нулю на его большей части. Из традиционных систем функций такой особенностью обладает базис Хаара, который формируется из ядер двух типов. Одно из них имеет нулевые значения элементов ядер, т. е. $\varphi = 0$, что и обеспечивает его специфический вид, характеризуемый локализацией базисных векторов с изменяемым интервалом кратности 2.

Свойство наилучшего равномерного приближения, присущее базису Хаара для широкого класса аналитических функций, можно придать и сформированному адаптируемому базису, если после процедуры приспособления приравнять нулю оставшиеся углы-параметры, т. е. добавить только одну степень свободы, удовлетворяющую условию:

$$\varphi_{23}=\varphi_{24}=\varphi_{1k}=0,\ k=2\ldots4.$$

В этом случае формирование базиса, адаптированного к локальным отклонениям от эталона, осуществляется на основе следующего соотношения:

$$[\Phi]=\left[\begin{pmatrix}C_{11} & S_{11}\\ S_{11} & -C_{11}\end{pmatrix},\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}_3\right]\times$$
$$\times\left[\begin{pmatrix}C_{21} & S_{21}\\ S_{21} & -C_{21}\end{pmatrix},\begin{pmatrix}C_{22} & S_{22}\\ S_{22} & -C_{22}\end{pmatrix},\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}_2\right]\times$$
$$\times\left[(\varphi_{31}),(\varphi_{32}),(\varphi_{33}),(\varphi_{34})\right]. \qquad (6)$$

Сформированный на основе (6) приспособленный базис ПБ-2 включает в себя функции с большим числом нулевых значений, т. е. имеет свойства вейвлет-функций, в связи с чем в таком базисе можно осуществлять перестраиваемые вейвлет-преобразования. Однако при его применении для анализа процессов с равновероятными отклонениями отсчетов реализации от эталона уже не обеспечивается минимизация числа коэффициентов разложения, в отличие от ПБ-1, построенного на основе условия (5). При фильтрации диагностируемых процессов в виде реализаций сигналов с локальными изменениями на временных сегментах необходимо учитывать то, что лежащие в основе спектральной фильтрации разложения в базисах Фурье и Уолша по точности приближения уступают разложению Хаара. Это обусловлено тем, что конечный ряд в соответствующем базисе с числом членов $N \le M$ не позволяет точно воспроизвести локальные изменения. Фильтр на основе спектральных коэффициентов в базисе Хаара, напротив, за счет локального характера своих функций позволяет эффективно отфильтровать шумовые компоненты и достаточно точно обеспечить воспроизведение формы диагностируемого процесса с соответствующими импульсными значениями.

Высокая степень изменчивости диагностируемых процессов смешанного типа требует выполнения тщательных экспериментальных исследований по фильтрации их от шумов различного вида и с различными значениями параметров. Адаптируемые базисы позволяют выполнять фильтрацию таких процессов с более высоким уровнем подавления шумов. В связи с большими размерами потоков разнообразной диагностической информации одной из центральных является задача снижения размерности пространства исследуемых диагностических сигналов и процессов. Компрессия данных позволяет выделить наиболее значимую информацию и отсеять второстепенные данные. При этом рассматривается основная характеристика в задаче компрессии данных – скорость сходимости рядов разложения в выбранном базисе разложения в заданной метрике.

В практических задачах компрессии диагностической информации важен не только факт равномерной сходимости той или иной частной суммы обобщенного ряда Фурье, но и скорость этой сходимости, т. е. быстрота стремления к нулю последовательности чисел $\varepsilon_\infty(x, N)$:

$$\varepsilon_\infty(x,N)=$$
$$=\max\left|x(t)-\sum_{n=1}^{N\le M}a_n\varphi_n(t)\right|\le\varepsilon_{\max},\ N<M. \qquad (7)$$

Частная сумма ряда $\sum_{n=1}^{N\le M} a_n \varphi_n(t)$, называемая полиномом наилучшего приближения, обеспечивает, согласно (7), наиболее высокую скорость сходимости в базисе $\{\varphi_n(t)\}$. Несмотря на то, что полиномы Чебышева для широкого класса функций обеспечивают наилучшее равномерное приближение, их структура достаточно сложна для реализации [6]. Поэтому с практической точки зрения более подходящими являются функции базисов Уолша, Хаара, Фурье, позволяющие для некоторых классов случайных процессов получить среднеквадратическое и равномерное приближение и обладающие быстрыми вычислительными процедурами.

При анализе реализаций квазипериодических случайных процессов, когда необходимо получение равномерного приближения анализируемых данных, точность приближения для базисов Фурье, Уолша и Хаара можно определить по следующим формулам, соответственно:

$$\max_{N\le M}\left| x(t) - \sum_{n=1}^{N\le M} a_n w^{nt} \right| \le E'_{\max},$$

$$\max_{N\le M}\left| x(t) - \sum_{n=1}^{N\le M} b_n \operatorname{wal}(n,t) \right| \le E''_{\max},$$

$$\max_{N\le M}\left| x(t) - \sum_{n=1}^{N\le M} c_n \chi(n,t) \right| \le E'''_{\max},$$

где $E'_{\max} \le E''_{\max} \le E'''_{\max}$.

Результаты экспериментальной проверки эффективности гибридного вейвлет-спектрального подхода для диагностируемых сигналов с локальной изменчивостью и зашумленностью на всем интервале определения показали, что по сравнению с другими системами ортогональных функций гибридный вейвлет-спектральный подход на основе приспособленных базисов обладает следующими характерными особенностями:

1) наиболее высокой скоростью сходимости ряда разложения для анализируемого класса сигналов;

2) наиболее высокой степенью избирательности преобразования в приспособленном базисе;

3) однозначностью вычисления коэффициентов разложения при любой длине (размерности) исходных данных и любом усечении ряда разложения;

4) возможностью неограниченного увеличения точности восстановления аппроксимируемой (в равномерном или среднеквадратичном смысле) функции путем добавления компонентов из отсеченной части ряда;

5) высокими сглаживающими и помехоустойчивыми свойствами;

6) относительной простотой вычисления коэффициентов разложения;

7) наличием алгоритмов быстрых преобразований для ряда базисов фиксированного типа (Фурье, Уолша–Адамара, Хаара, Сланта, Хартли);

8) возможностями перестройки и адаптации ортогональных базисов, что особенно важно при решении задач диагностики.

Благодаря этим особенностям с помощью адаптируемых ортогональных базисов могут быть достигнуты существенная экономия времени и минимизация размерности пространства диагностических признаков.

СПИСОК ЛИТЕРАТУРЫ

1. Методическое и программное обеспечение автоматизированного эксперимента в динамике машин/ М. Б. Левин, Д. Е. Одуло, Д. Е. Розенберг и др. М.: Наука, 1989.

2. Автоматизация измерений и обработка данных при испытаниях самолета на прочность / И. Ф. Образцов, А. С. Голубков, А. И. Никитин и др. М.: Машиностроение, 1991.

3. Вибрация и вибродиагностика судового электрооборудования / А. А. Александров, А. В. Барков, Н. А. Баркова и др. Л.: Судостроение, 1986.

4. Алексеев А. А., Спиваковский А. М. Методы цифровой обработки сигналов: учеб. пособие. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2009.

5. Алексеев А. А., Солодовников А. И., Спиваковский А. М. Методы вейвлет-обработки сигналов и изображений: учеб. пособие. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2004.

6. Солодовников А. И., Спиваковский А. М. Основы теории и методы спектральной обработки информации: учеб. пособие. Л.: Изд-во ЛГУ, 1986.

7. Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения // Успехи физ. наук. 1998. Т.186. С. 1145–1170.

8. Дьяконов В. П. Вейвлеты. От теории к практике. М.: Изд-во «СОЛОН-Р», 2002.

9. Дремин И. М., Иванов О. В., Нечитайло В. А. Вейвлеты и их использование // Успехи физ. наук. 2001. Т. 171. С. 465–561.

10. Ватолин Д. С. Методы сжатия данных: М.: ДИАЛОГ-МИФИ, 2002.

A. A. Alekseev, A. M. Spivakovskiy
Saint Petersburg Electrotechnical University «LETI»

APPLICATION OF A HYBRID WAVELET-SPECTRAL METHOD FOR DETECTING DIAGNOSTIC CHARACTERISTICS

The current information and computational aspects are considered in solving problems of diagnosing complex objects and control processes. A hybrid approach based on fast orthogonal wavelet-spectral transform, adapted to the type of characteristics, describing the state of the object being diagnosed, is proposed for the operational formation of the space of diagnostic signs of the minimum dimension. This approach provides high speed of convergence of series for the analyzing class of signals and characterized by high degree of selectivity of transformation in adaptive basis. The advantage of the proposed hybrid approach is the unambiguous calculation of decomposition coefficients in any dimension of the source data and any truncation of the decomposition series. This allows to increase accuracy of the recovery approximated function by adding components from the amputated part of the series.

Wavelet-spectral transformations, adaptive bases, diagnostics of control objects, diagnostic signs