

УДК 621.383

Ю. В. Маргышевский

(Томск)

**ДЕКОМПОЗИЦИЯ ВЕЙВЛЕТ-СПЕКТРА
СИГНАЛА МИКРОДЕНСИТОМЕТРА
В ЗАДАЧЕ ИДЕНТИФИКАЦИИ ФИЛИГРАНИ ДОКУМЕНТОВ ***

Предложено для решения задачи идентификации филигрانی документов использовать декомпозицию вейвлет-спектра сигнала микроденситометра. Показано преимущество предложенного метода обработки сигнала перед традиционно используемыми.

Введение. Историки, архивисты, библиотекари, работающие с бумажными документами XIII – начала XIX вв., постоянно сталкиваются с проблемой их датировки. В случае, если дата не устанавливается непосредственно из содержания документа, необходимо прибегать к процедуре многостадийного анализа рассматриваемого источника, учитывающего целый комплекс признаков. Большое значение при этом имеют данные палеографии и ее составной части – филигранологии.

На бумаге документов, написанных в XIII – начале XIX вв., при просмотре ее на просвет видны так называемые филигрانی – водяные знаки, появление которых связано с процессом ручного производства бумаги. Технология этого процесса выглядела следующим образом. Подготовленную бумажную массу разливали в специальные формы с сетчатым дном. Сетка содержала функциональную часть – систему частых продольных проволок и редких поперечных, а также маркерную в виде сюжетного изображения некоторых литер и цифр. В результате уплотнения бумажной массы слой ее становился более толстым в тех местах, где он не соприкасался с проволокой, и более тонким в местах соприкосновения. В соответствии с такой технологией видимые на готовой бумаге водяные знаки представляют собой следующие элементы: сетку из частых (вержеры) и редких (понтюзо) линий, изображение сюжетной части, литерное сопровождение (обычно инициалы владельца или мастера бумажной мельницы или их полные имена, название фирмы и другое) и «белую дату» (год отливки бумаги).

В настоящее время обработка сигнала микроденситометра заключается в предварительной фильтрации сигнала, а процесс идентификации филигрانی в контроле периодичности вержеров и понтюзо путем ручной калькуляции или на основании анализа Фурье спектра сигнала [1]. Однако это не дает желаемого результата, поскольку расстояние между вержерами и понтюзо не

* Работа выполнена при поддержке Министерства образования Российской Федерации (грант № 13-02, 1999 г.).

постоянно по полю документа, поэтому достоверное определение периода вержеров и понтюзо затруднено.

Целью работы является исследование (поиск) достаточно простого метода обработки сигнала микроденситометра, позволяющего повысить достоверность идентификации филигранны на основании анализа периодичности вержеров и понтюзо документа. Одним из таких методов является вейвлет-преобразование (ВП).

Вейвлет-преобразование. Известно, что главное отличие ВП от преобразования Фурье состоит в более информативном представлении частотно-временных и масштабно-временных свойств сигнала, а именно локализованность спектра ВП во времени позволяет увидеть изменение во времени спектральных свойств сигнала [2, 3].

Способность ВП анализировать сигнал на разных масштабах часто сравнивают с математическим «микроскопом», что позволяет успешно применять ВП для анализа сигналов самой различной природы. Ниже приводится метод обработки сигнала микроденситометра для идентификации филигранны документа на основе анализа вейвлет-спектра сигнала.

Вейвлет-преобразование основано на том, что сигнал любой длительности с конечной энергией может быть представлен в виде комбинации элементарных функций вейвлет-базиса, представляющих собой варианты сдвига, расширения (сжатия) одной и той же базисной функции – вейвлета. Интегральное ВП от квадратично-интегрируемой функции $Z(t)$ представляется в виде

$$W(a, b) = \langle Z(t), \psi_{a, b}(t) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} Z(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (1)$$

где $\psi(t)$ – базисная функция вейвлета; $a, b \in R$ и $a \geq 0$: a – масштабный коэффициент ВП, b – параметр сдвига.

Существует обратное вейвлет-преобразование, восстанавливающее из вейвлет-спектра $W(a, b)$ исходную функцию

$$Z(t) = \frac{1}{C_{\psi}} \int_{-\infty}^{\infty} \int_0^{\infty} \langle Z(t), \psi_{a, b}(t) \rangle \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \frac{dadb}{a^2}, \quad (2)$$

где $C_{\psi} = \int_{-\infty}^{\infty} |\Psi(\omega)|^2 |\omega|^{-1} d\omega < \infty$ – нормирующий коэффициент, а $\Psi(\omega)$ – фурье-образ вейвлета.

Если продолжить уже упоминавшуюся аналогию ВП с математическим микроскопом, то сдвиг b фиксирует точку фокусировки микроскопа, масштабный коэффициент a – увеличение, а выбором базисного вейвлета определяются оптические качества микроскопа.

Характеристиками вейвлет-спектра сигнала являются мера перемежаемости и контраста [2].

Мера локальной перемежаемости спектра

$$I_W(a, b) = \frac{W^2(a, b)}{\langle W^2(a, b) \rangle_b}, \quad (3)$$

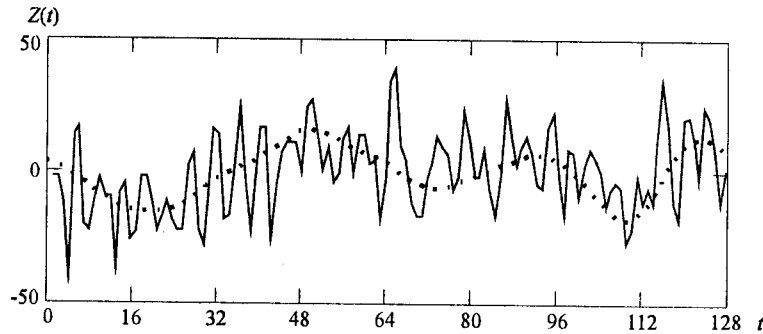


Рис. 1. Вид фрагмента строки сигнала микроденситометра: исходный сигнал (сплошная кривая), результат вейвлет-обработки (пунктирная кривая)

где $\langle \rangle_b$ – усреднение по параметру сдвига b . Мера локальной перемежаемости показывает отклонение от среднего значения спектров на каждом масштабе ВП и позволяет определить степень неравномерности распределения энергии сигнала по масштабам.

Мера контраста

$$C_w(a, b) = \frac{W^2(a, b)}{E'_w(a, b)}, \quad (4)$$

где $E'_w(a, b) = \int_{a'=0}^{a'=a} W^2(a', b) da'$, позволяет определить в исследуемом сигнале

малые отклонения в его структуре.

Исследование вейвлет-спектра. Вид фрагмента строки сигнала, полученного с микроденситометра, приведен на рис. 1 (сплошная линия). В его структуре отражено наличие как информативной составляющей (вержеров, понтюзо, водяных знаков), так и помех, обусловленных неоднородностью плотности бумаги и ее структуры, деформацией и дефектами бумаги, а также загрязнением документа.

Для идентификации филигрании документа в качестве вейвлет-функции использован Mhat-вейвлет, являющийся второй производной функции Гаусса:

$$\psi(t) = (-1)^{m+1} \frac{d^m}{dt^m} \exp\left(-\frac{t^2}{2}\right), \quad m=2. \quad (5)$$

Результатом ВП сигнала по базису $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$ является двумер-

ный массив коэффициентов $W(a, b)$ (вейвлет-спектр, приведенный на рис. 2). На этом рисунке за начало отсчета принят верхний левый угол, по оси ординат отложен масштаб a базисной функции, а по оси абсцисс – время b .

Анализ вейвлет-спектра показывает, что во фрагменте строки сигнала присутствует периодическая структура, соответствующая информативной части, а именно максимумы (три светлые области) в месте присутствия сетки и минимумы (три темные области), соответствующие областям между элементами сетки, линии локальных минимумов, разделяющих эти области, а в

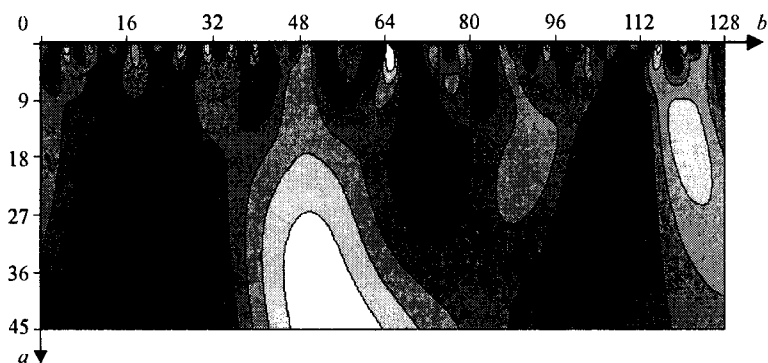


Рис. 2. Вейвлет-спектр сигнала

области малых масштабов имеется помеха, обусловленная неоднородностью бумажной массы. Для выделения информативной составляющей сигнала от помех воспользуемся мерой контраста вейвлет-спектра.

Для этого усредним меру контраста по параметру сдвига b и получим масштабное изменение контраста вейвлет-спектра $C_W(a, b)$ в зависимости от масштаба a (рис. 3). Минимум характеристики $C_W(a, b)$ на этом рисунке разделяет вейвлет-спектр на высокочастотную и низкочастотную составляющие. Выполнив декомпозицию вейвлет-спектра путем приравнивания к нулю высокочастотной части спектра, а затем обратное вейвлет-преобразование, получим сглаженную версию исходного сигнала (см. рис. 1, штриховая линия), которая характеризует редкую сетку (понтюзо). Очевидно, что если обработать массив строк сигналов, то можно выделить сигналы, соответствующие вержер-сетке.

Сравнение полученного результата с результатами традиционно применяемых методов фильтрации сигнала микроденситометра показывает, что метод декомпозиции вейвлет-спектра имеет существенное преимущество при решении задачи идентификации филигранны документов перед используемыми для этого методами фильтрации (рис. 4). Отметим также хорошее совпадение экстремальных точек (соответствующих понтюзо в области 16, 48, 80 и 112 отсчетов по горизонтальной оси) полученного сигнала с вейвлет-спектром, приведенным на рис. 2.

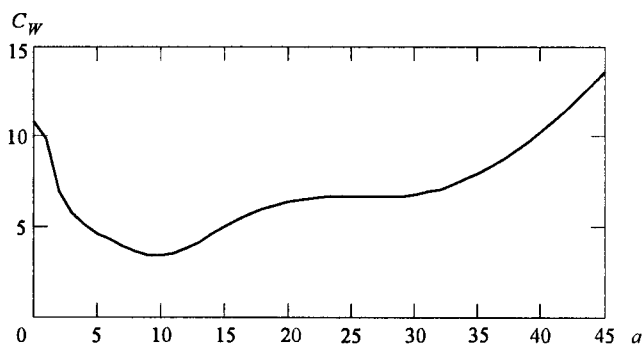


Рис. 3. Изменение контраста C_W вейвлет-спектра

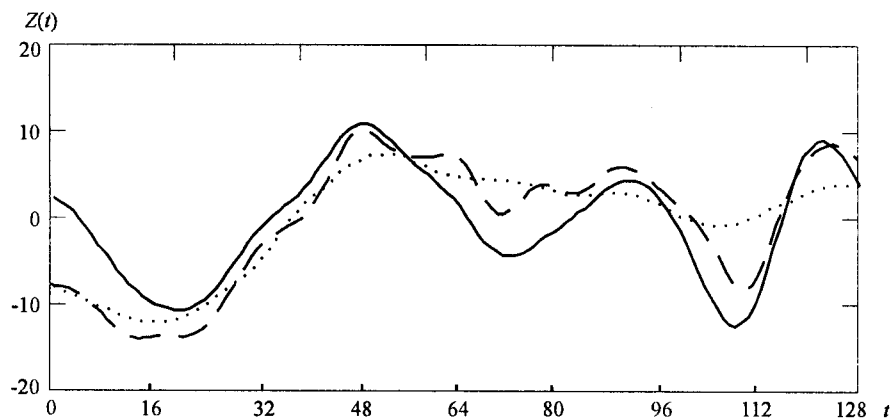


Рис. 4. Результаты выделения информативной составляющей сигнала микроденситометра различными методами: вейвлет-обработка (сплошная кривая), сглаживание методом наименьших квадратов (пунктирная), полосовой фильтр (штриховая)

Заключение. Результаты обработки сигналов микроденситометра показывают, что при использовании ВП информативная составляющая сигнала, обусловленная присутствием понтюзо в филигранях документа, выделяется более эффективно в сравнении с другими альтернативными методами.

Двумерный характер вейвлет-спектра, обладающий большим набором информативных признаков (линии локальных минимумов), позволяет проводить идентификацию документов простым сравнением «картинок» спектров или применять при этом хорошо разработанные методы распознавания изображений.

Расчеты проводились средствами пакета MathCAD 7.0, при этом все используемые функции написаны на языке C++ и оформлены в виде DLL приложений MathCAD 7.0.

В дальнейшем планируется использовать вейвлет-преобразование для обработки двумерных изображений документов в задаче идентификации филиграней.

Автор выражает благодарность канд. физ.-мат. наук В. М. Климкину за постановку задачи и обсуждение полученных результатов.

СПИСОК ЛИТЕРАТУРЫ

1. Белов В. В., Борисов Б. Д., Гриднев Ю. В. и др. Проблемы воспроизведения, компьютерной обработки и идентификации филиграней // История и компьютер. 2000. 25, № 3. С. 118.
2. Астафьева Н. М. Вейвлет-анализ: основы теории и примеры применения // УФН. 1996. 166, № 11. С. 1145.
3. Daubechies I. Ten lectures on wavelets // Lecture Notes Series. Philadelphia. SIAM. 1991.