

Многомерный вейвлет-анализ изображений*

В. З. Рахманкулов, А. А. Ахрем,
В. В. Герасимов, А. А. Барсегов

В статье приводятся основные свойства многомерных вейвлет-пространств. Рассматриваются также вопросы применения многомерных вейвлет-функций для анализа и распознавания объемных промышленных деталей. Дается описание разработанного авторами программного комплекса обработки и распознавания специальных классов таких деталей с применением аппарата многомерных вейвлет-функций.

Введение

Обычно под системами вейвлетов понимаются сжатия и сдвиги одной функции, образующие систему представления в каком-либо смысле (например, ортонормированный базис в гильбертовом пространстве $L_2(\mathbb{R})$, где $\mathbb{R} = (-\infty, +\infty)$ — числовая ось). В некоторых ситуациях вейвлет-системы состоят из сдвигов и сжатий нескольких функций или целой последовательности.

Интерес к изучению систем вейвлетов возник задолго до появления терминологии и основ теории вейвлетов (пространств вейвлетов) и был обусловлен главным образом потребностью в их использовании для обработки сигналов и изображений. Базисы вейвлетов имеют ряд преимуществ по сравнению с другими базисами, используемыми в качестве аппарата приближения функций. Они обладают так называемой время-частотной локализацией, т. е. быстро убывают на бесконечности как сами базисные

* Работа выполнена при поддержке программы Президиума РАН «Фундаментальные проблемы информатики и информационных технологий» (проект № 2.44).

функции, так и их преобразования Фурье [1–3]. Благодаря этому свойству при разложении по базису сигналов, частотные характеристики которых меняются по времени или по пространству (такowymi являются, в частности, речевые или музыкальные сигналы, сейсмические сигналы, а также изображения), много коэффициентов разложения при ненужных на данном пространственном или временном участке гармониках оказываются малыми и могут быть отброшены, что обеспечивает тем самым сжатие информации.

Допустимость такого отбрасывания объясняется другим важным свойством вейвлетов: вейвлет-разложения являются безусловно сходящимися рядами. Кроме того, существуют эффективные алгоритмы, позволяющие быстро вычислять коэффициенты вейвлет-разложений [1–7]. В работах авторов [8, 9] исследовались возможности применения одномерных вейвлет-функций для решения задач фильтрации, сжатия, оконтурирования изображений специальных классов промышленных деталей.

1. Многомерные вейвлет-функции

В настоящем разделе мы обобщим понятия вейвлет-функции и вейвлет-пространства на многомерный случай. Для этого нам потребуются некоторые определения и факты из теории одномерных вейвлетов.

Обозначим через l_2 гильбертово пространство последовательностей комплексных чисел $c \equiv \{c_n\} (n = 1, 2, \dots)$ таких, что

$$\sum_{n=1}^{\infty} |c_n|^2 < \infty.$$

Скалярное произведение элементов $c', c'' \in l_2$ задается следующей формулой

$$(c', c'') \equiv \sum_{n=1}^{\infty} c'_n \cdot \bar{c}''_n,$$

где $c' = (c'_1, c'_2, \dots)$, $c'' = (c''_1, c''_2, \dots)$, \bar{c}''_n — число, комплексно сопряженное с c''_n .

Определение 1. Пусть H — гильбертово пространство. Система его элементов $\{f_n\} (n = 1, 2, \dots)$ называется системой Рисса с постоянными $\alpha, \beta > 0$; если для любого $c \equiv \{c_n\}$, $c \in l_2$, ряд $\sum_{n=1}^{\infty} c_n f_n$ сходится в H и выполняются неравенства

$$\alpha \|c\|_{l_2}^2 \leq \left\| \sum_{n=1}^{\infty} c_n f_n \right\|_H^2 \leq \beta \|c\|_{l_2}^2.$$

Если система Рисса является базисом, то ее называют базисом Рисса. Имеет место следующее утверждение [1–4].

Теорема 1. Пусть H — гильбертово пространство, $\{f_n\}$ — система Рисса в H с константами α, β . Тогда

- 1) $\{f_n\} (n = 1, 2, \dots)$ является базисом Рисса в пространстве

$$V \equiv \left\{ f = \sum_{n=1}^{\infty} c_n f_n, \sum_{n=1}^{\infty} |c_n|^2 < \infty \right\};$$

- 2) $V \equiv sp\{f_n, n \in \{1, 2, \dots\}\}$, где $sp\{f_n\}$ — множество конечных линейных комбинаций множества $\{f_n\}$; $\overline{sp\{f_n\}}$ — замыкание множества $sp\{f_n\}$;
- 3) для любого элемента $f \in V$ выполняются соотношения

$$\alpha \|f\|^2 \leq \sum_{n=1}^{\infty} |(f, f_n)|^2 \leq \beta \|f\|^2.$$

Определение 2. Совокупность замкнутых пространств $V_j \subset L_2(\mathbb{R})$, $j \in Z$, где Z — множество целых чисел называется кратномасштабным анализом в $L_2(\mathbb{R})$, если выполнены следующие условия:

- I. $V_j \subset V_{j+1}$ для всех $j \in Z$;
- II. $\bigcup_{j \in Z} V_j$ плотно в $L_2(\mathbb{R})$;
- III. $\bigcap_{j \in Z} V_j = \{0\}$;
- IV. $f \in V_j \Leftrightarrow f(2^{-j} \cdot) \in V_0$ для всех $j \in Z$;
- V. Существует такая функция $\varphi \in V_0$, что последовательность целочисленных сдвигов $\{\varphi(\cdot + n)\} (n \in Z)$ образует базис Рисса в V_0 .

Функция φ из условия V называется масштабирующей для данного кратномасштабного анализа $\{V_j\} (j \in Z)$.

Заметим, что главное предназначение кратномасштабного анализа состоит в следующем. На его основе можно строить ортогональные базисы, полученные посредством сжатий и сдвигов одной или нескольких функций.

Отметим, что вопросы существования кратномасштабного анализа $\{V_j\} (j \in Z)$ и методы его построения подробно рассматриваются в работах [1–7].

Вернемся теперь к построению многомерного аналога кратномасштабного анализа (КМА).

Пусть даны многомерные КМА $\{V_j^{(r)}\}, j \in Z, r = 1, \dots, p$.

Определим подпространства $\{V_j\}, j \in Z$, гильбертова пространства $L_2(\mathbb{R}^p)$:

$$V_j \equiv \bigotimes_{r=1}^p V_j^{(r)} = \overline{sp\{f = f_1 \otimes \dots \otimes f_p, f_r \in V_j^{(r)}\}}.$$

Имеет место следующее утверждение [3, 4].

Теорема 2. Пусть КМА $\{V_j^{(r)}\}$, $j \in Z$, $r = 1, \dots, p$ — КМА в гильбертовом пространстве $L_2(R)$. Тогда подпространства V_j пространства $L_2(R^P)$, определенные формулами (1), обладают следующими свойствами:

- I. $V_j \subset V_{j+1}$ для всех $j \in Z$;
- II. $\bigcup_{j \in Z} V_j$ плотно в $L_2(R^P)$;
- III. $\bigcap_{j \in Z} V_j = \{0\}$;
- IV. $f \in V_0 \Leftrightarrow f(2^j \cdot) \in V_j$ для всех $j \in Z$;
- V. Существует функция $\varphi \in V_0$ такая, что последовательность $\{\varphi(\cdot + n)\} (n \in Z^P)$, образует базис Рисса в V_0 .

Совокупность пространств V_j , $j \in Z$, удовлетворяющих условиям I–V теоремы 2, называется КМА в пространстве $L_2(R^P)$.

Пространства вейвлетов W_j определяются теперь, как и в одномерном случае, формулами:

$$W_j \equiv V_{j+1} \ominus V_j, (j \in Z).$$

Остановимся подробнее на двумерном КМА, так как именно двумерный КМА связан с обработкой изображений [3, 4]. В этом случае сигнал представляет собой функцию $f(x, y) \in L_2(R^2)$. Пусть заданы: $\varphi(x)$ — масштабирующая вейвлет-функция и $\psi(x)$ — вейвлет. Они порождают базисные функции $\varphi_{j,n}(x)$ и $\psi_{j,k}(x)$. Двумерный сигнал $f(x, y)$ раскладывается по базисным в $L_2(R^2)$ функциям $\varphi_{j,n}(x) \cdot \varphi_{j,m}(y)$, $\varphi_{j,n}(x) \cdot \psi_{j,m}(y)$, $\psi_{j,n}(x) \cdot \varphi_{j,m}(y)$, $\psi_{j,n}(x) \cdot \psi_{j,m}(y)$. Соответствующие коэффициенты называются следующим образом.

Аппроксимирующие коэффициенты c_j получаются как коэффициенты разложения по вейвлет-базису $\varphi_{j,n}(x) \cdot \varphi_{j,m}(y)$. Горизонтальные детализирующие коэффициенты c'_j получаются как коэффициенты разложения по вейвлет-базису $\varphi_{j,n}(x) \cdot \psi_{j,m}(y)$. Вертикальные детализирующие коэффициенты c''_j получаются как коэффициенты разложения по вейвлет-базису $\psi_{j,n}(x) \cdot \varphi_{j,m}(y)$. Диагональные детализирующие коэффициенты c'''_j получаются как коэффициенты разложения по вейвлет-базису $\psi_{j,n}(x) \cdot \psi_{j,m}(y)$.

Отметим, что горизонтальные коэффициенты больше отражают горизонтальные элементы изображения, вертикальные коэффициенты больше отражают вертикальные элементы изображения, а диагональные коэффициенты больше отражают диагональные элементы распознаваемого изображения.

2. Использование функций Wavelet Toolbox для распознавания изображений

Задача распознавания изображений заключается в сравнении коэффициентов, получившихся в результате вейвлет-разложения тестового изображения, с коэффициентами, получившимися при разложении нескольких эталонных изображений. При совпадении максимального числа коэффициентов тестовое изображение считается идентичным одному из эталонов.

Для написания программы распознавания образов использовалась среда математического моделирования MatLab версии 7.0 и функции, входящие в набор инструментов Wavelet Toolbox.

Для примера в качестве тестовых и эталонных изображений были использованы абстрактные детали, смоделированные в среде 3D Studio MAX (см. рис. 1). Далее на тестовое изображение накладывался шум, а в случае сильного зашумления разрывались контуры объекта. Таким образом, для теста были использованы 4 эталонных изображения и 8 тестовых (4 — со слабым уровнем зашумления и 4 — с сильным).

На первом этапе происходит загрузка в программу тестового изображения. Чтобы загрузить изображение, необходимо указать его имя. Файл, содержащий изображение должен иметь расширение .jpg или .bmp и лежать в той же директории, что и исполняемый модуль программы (файл с расширением .m).

```
test_image = imread('test011.jpg');
```

На втором этапе производится выбор типа вейвлета. Эксперименты проводились с различными вейвлетами и показали, что для точного рас-

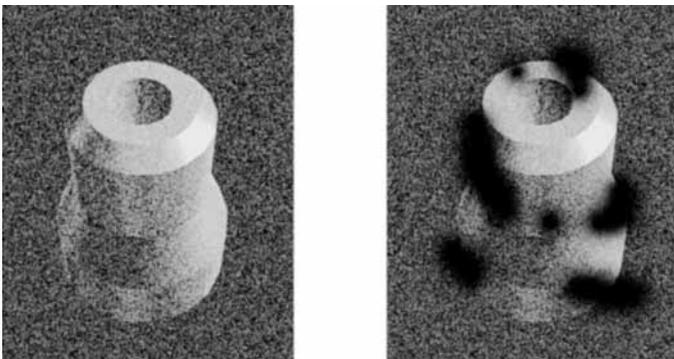


Рис. 1. Пример слабого (а) и сильного (б) зашумления тестового изображения

познавания вполне подходит вейвлет Добеши второго порядка. Вейвлеты Хаара и Добеши первого порядка не могут в достаточной мере описать изображение, а использование вейвлетов более высоких порядков не дает лучшего результата, а иногда даже не позволяет решить задачу.

```
wvlt = 'db2';
```

На третьем этапе происходит загрузка в программу эталонных изображений и выбор цветовой палитры, в которой будет отображаться изображение при обработке (см. рис. 2).

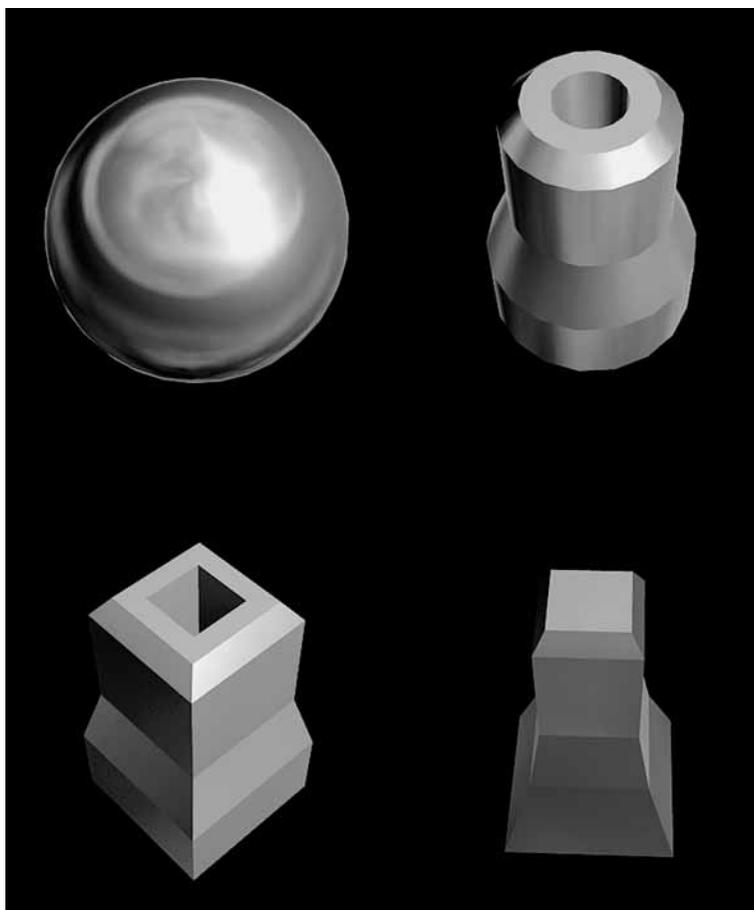


Рис. 2. Эталонные изображения

На четвертом этапе идет масштабирование изображений до стандартного размера (в данном случае 256×256 пикселей) и преобразование исходных изображений к монохромному. Параметр 100 влияет на яркость исходного изображения и позволяет автоматически усреднить все шумы до приемлемого уровня.

```
gray_image = rgb2gray(imresize(test_image,[256 256])+100)
cod_image = wcodemat(gray_image,nbcol);
```

Функция `wcodemat` возвращает закодированную версию изображения `picture`. При этом цветовая гамма принимается автоматически, если значение входного параметра `modul` равно 0, и устанавливается с использованием параметра `integer`, при любом ненулевом значении параметра `modul`.

Способ кодирования может быть задан параметром `column`, допустимыми значениями которого являются `'row'` или `'r'`, соответствующие кодированию построчно, `'col'` или `\verb'c'`, соответствующие кодированию по столбцам, и, наконец, `'mat'` или `\verb'm'`, соответствующие кодированию изображения, как по строкам, так и по столбцам. Необходимо отметить, что запись функции

```
wcodemat (picture, integer, column)
```

является эквивалентной записи

```
wcodemat (picture, integer, column, 1).
```

В то же время, запись

```
wcodemat (picture, integer)
```

является эквивалентной записи

```
wcodemat (picture, integer, 'mat', 1).
```

Наконец, запись

```
picture = wcodemat (picture)
```

эквивалентна записи

```
wcodemat (picture, 16, 'mat', 1).
```

На пятом этапе происходит декомпозиция изображений с использованием выбранного типа вейвлет функции.

```
[ta,th,tv,td] = dwt2(cod_image,wvlt);
```

В результате для каждого изображения мы получаем по 3 матрицы размерностью 129×129 , содержащие коэффициенты вейвлет-преобразования для данного изображения: по вертикали, по горизонтали и по диагонали.

В данной версии программы происходит сложение трех матриц декомпозиции, принадлежащих одному изображению, и ведется сравнение суммарных матриц. В будущих версиях планируется сравнение отдельно каждой из трех декомпозиционных матриц тестового изображения

с соответствующими матрицами эталонных изображений, что позволит улучшить качество распознавания.

На шестом этапе происходит непосредственно сравнение матрицы вейвлет-коэффициентов тестового изображения с матрицами эталонных изображений. Сравнение ведется поэлементно, поэтому в данной версии программы для корректного распознавания объект на тестовом изображении и на эталонном изображении должен располагаться четко на одном и том же месте.

Программа считает количество совпадений вейвлет-коэффициентов и выдает на экран то эталонное изображение, у которого больше всего совпадений с тестовым (см. рис. 3).

Также программа выводит на экран результаты декомпозиции тестового и эталонного изображения (см. рис. 4).

Метод вейвлет-преобразования с использованием вейвлета Добеши второго порядка дает хорошие результаты распознавания даже в случае

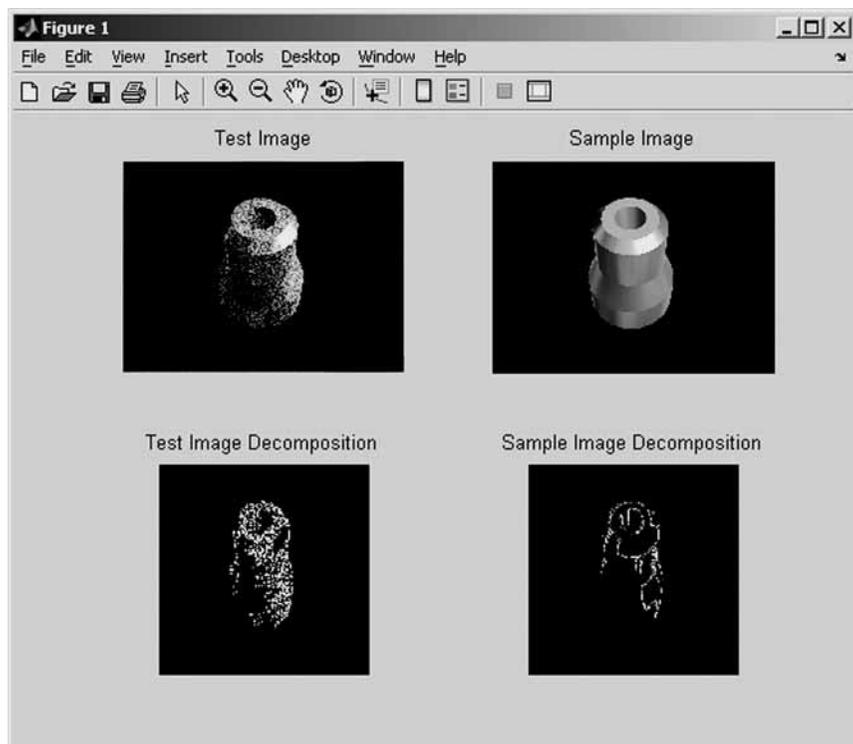


Рис. 3

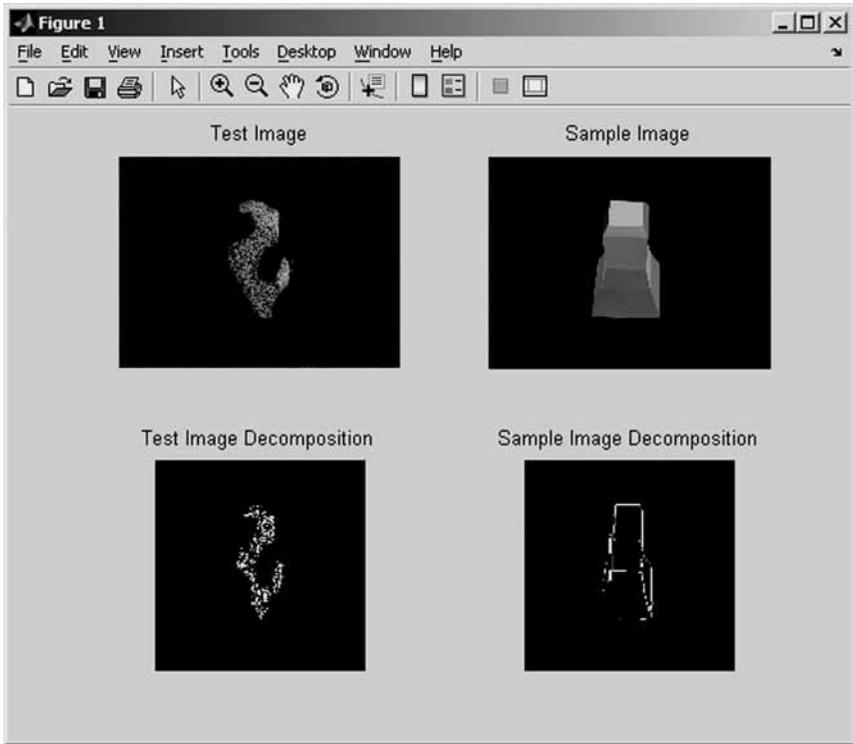


Рис. 4

сильно зашумленного изображения с разорванными границами, там, где даже человеку трудно различить, какой именно объект изображен на картинке.

3. Полный исходный текст программы

```
test_image = imread('test011.jpg'); % Тестовая картинка
wvlt = 'db2'; %Тип вейвлета
```

%Образцы, с которыми сравниваем тестовое изображение

```
samp_arr{1} = imread('sample01.jpg');
samp_arr{2} = imread('sample02.jpg');
samp_arr{3} = imread('sample03.jpg');
samp_arr{4} = imread('sample04.jpg');
```

```
colormap (gray); %Цветовая палитра отображения
nbc = 128; %Количество цветов в палитре

gray_image = rgb2gray(imresize(test_image,[256 256])+100)
cod_image = wcodemat(gray_image,nbc);

[ta,th,tv,td] = dwt2(cod_image,wvlt);
dec_image = th+tv+td;

for k = 1:numel(samp_arr)}

cod_arr{k} = wcodemat(rgb2gray(imresize(samp_arr{k},[256
256])+64),nbc);

[ca{k},ch{k},cv{k},cd{k}] =
dwt2(cod_arr{k},wvlt);

dec_arr{k} = ch{k}+cv{k}+cd{k};

A = dec_image;
B = dec_arr{k};

match_arr{k} = A==B; %Сравнение двух матриц
match_sum = sum(sum(match_arr{k}));

if (match_sum > new_sum)
new_sum = match_sum;
I=k; %Индекс матрицы с максимальным числом совпадений
end
end

subplot(2,2,1);
image(test_image);
axis off equal tight
title('Test Image');

subplot(2,2,2);
image(samp_arr{I});
axis off equal tight
```

```
title('Sample Image');

subplot(2,2,3);
image(dec_image);
axis off equal tight
title('Test Image Decomposition');

subplot(2,2,4);
image(dec_arr{1});
axis off equal tight
title('Sample Image Decomposition');
```

Заключение

Таким образом, в настоящей статье рассмотрены математические методы, алгоритмы и программы распознавания изображений геометрических объектов с помощью аппарата многомерных вейвлет-функций. Даны основные свойства пространств многомерных вейвлетов. Приведены примеры автоматизированного распознавания специальных классов промышленных деталей с помощью различных многомерных вейвлет-функций (вейвлетов Хаара, Добеши и др.)

Литература

1. *Добеши И.* Десять лекций по вейвлетам. М., Ижевск: РХД, 2001.
2. *Чуи К.* Введение в вейвлеты. М.: Мир, 2001.
3. *Петухов А. П.* Введение в теорию базисов всплесков. СПб.: СПбГТУ, 1999.
4. *Блаттер К.* Вейвлет-анализ. Основы теории. М.: Техносфера, 2004.
5. *Новиков И. Я., Протасов В. Ю., Скопина М. А.* Теория всплесков. М.: Физматлит, 2006.
6. *Астафьева Н. М.* Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. 1998. Т. 155. № 11. С. 1145–1170.
7. *Новиков И. Я., Стечкин С. Б.* Основы теории всплесков // Успехи математических наук. 1998. Т. 53. № 6. С. 53–128.
8. *Рахманкулов В. З., Ахрем А. А., Герасимов В. В., Лебедев В. В.* Обработка и распознавание изображений промышленных деталей: Труды ИСА РАН// Интеллектуальные информационные технологии. Концепции и инструментарий. Т. 16. М.: URSS, 2005. С. 99–129.
9. *Рахманкулов В. З., Ахрем А. А., Герасимов В. В., Лебедев В. В.* Вейвлет-анализ изображений промышленных деталей // в печати.