

## РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ДВУМЕРНОГО ВЕЙВЛЕТ-АНАЛИЗА

*М.С. Бабий, канд. техн. наук, доцент;  
А.П. Чекалов, канд. техн. наук, доцент,  
Сумский государственный университет, г. Сумы*

*Описаний алгоритм обчислення ознак розпізнавання для двовимірних графічних зображень, що ґрунтується на декомпозиції зображення за фільтрами Габора. Розроблено програму розпізнавання зображень. Тестування програми на зразках ORL-бази зображень облич показало високу ефективність алгоритму розпізнавання.*

**Ключові слова:** вейвлети Габора, швидке перетворення Фур'є, графічні зображення, розпізнавання, ORL-база.

*Описан алгоритм вычисления признаков распознавания для двумерных графических изображений, основанный на декомпозиции изображения по фильтрам Габора. Разработана программа распознавания изображений. Тестирование программы на образцах из ORL-базы изображений лиц показало высокую эффективность алгоритма распознавания.*

**Ключевые слова:** вейвлеты Габора, быстрое преобразование Фурье, графические изображения, распознавание, ORL-база.

### ВВЕДЕНИЕ И ПОСТАНОВКА ЗАДАЧИ

Методы распознавания изображений, в которых признаками являются яркости отдельных пикселей, часто не учитывают взаимное расположение этих пикселей. В результате двумерная задача решается по существу как одномерная.

То же самое можно сказать и о большинстве нейросетевых методов, в частности, о широко используемом методе обратного распространения ошибки в многослойной нейронной сети. Двумерный характер изображения учитывается, например, в свёрточных нейронных сетях, содержащих свёрточные слои, в которых фрагменты входного изображения, взятые с некоторым сдвигом, преобразуются в пиксели следующего слоя [1]. Однако такая сеть достаточно сложна в реализации.

Методы, основанные на Фурье-анализе, не получили большого распространения, так как качество распознавания в большой степени зависело от масштаба и ориентации распознаваемого объекта на изображении.

Более перспективным оказался метод, основанный на использовании вейвлетов Габора [2,3]. За счет обнуления постоянной составляющей метод вейвлетов Габора устойчив к изменению яркости изображения. Различия в контрастности изображений могут быть легко устранены нормализацией. Наконец, вейвлеты характеризуют изображение в локальной области вокруг точки, а не только в самой точке.

Анализ литературных источников показывает, что большинство работ по вейвлетам Габора связано с распознаванием изображений лиц. Соответственно в теорию вводятся дополнительные понятия и процедуры: эластичные графы, высокоэнергетичные точки, локализация овала лица. При этом практически не приводятся подробности программной реализации.

Целью настоящей работы является разработка алгоритма и программы распознавания изображений на основе двумерного вейвлет-анализа. При этом алгоритм должен быть общим и не ориентированным на конкретный класс изображений. Объем вычислений и затраты оперативной памяти должны быть минимально необходимыми для обеспечения хорошего

качества распознавания. Программа должна основываться на некоммерческом программном обеспечении.

### РЕАЛИЗАЦИЯ МЕТОДА ВЕЙВЛЕТОВ

Двумерный вейвлет Габора имеет форму плоской волны, ограниченной по амплитуде функцией Гаусса. Каждый вейвлет  $j$  из заданного набора вейвлетов определяется своим волновым вектором  $k_j$  [3]:

$$\psi_j(\vec{r}) = \frac{k_j^2}{\sigma^2} \exp\left(-\frac{k_j^2 r^2}{2\sigma^2}\right) \left[ \exp(i\vec{k}_j \vec{r}) - \exp\left(-\frac{\sigma^2}{2}\right) \right], \quad (1)$$

где  $\vec{r} = \vec{r}(x, y)$ . Вектор  $k_j$  определим для пяти различных частот, индекс  $p = 0, \dots, 4$ , и восьми различных ориентаций, индекс  $q = 0, \dots, 7$ :

$$\begin{aligned} \vec{k}_j &= (k_{jx}, k_{jy}) = (k_p \cos \phi_q, k_p \sin \phi_q); \\ k_p &= 2^{\frac{p+2}{2}} \pi; \quad \phi_q = q \frac{\pi}{8}; \quad j = q + 8p. \end{aligned}$$

Пусть изображение задано массивом  $I(\vec{r})$  яркостей пикселей в точках  $\vec{r}(x, y)$ . Вейвлет-преобразование изображения можно представить как свертку  $R_j(\vec{r}_0)$  изображения с  $j$ -м вейвлетом из набора вейвлетов Габора:

$$R_j(\vec{r}_0) = \int I(\vec{r}) \psi_j(\vec{r} - \vec{r}_0) dx dy. \quad (2)$$

По аналогии с другими работами величины  $R_j$  будем называть коэффициентами отклика фильтров Габора.

Вектор признаков распознавания для изображения будем строить из значений  $R_j$  в узлах решетки с шагом 8 пикселей по горизонтали и вертикали. Такой шаг вполне приемлем для изображений с длиной стороны порядка ста пикселей.

Для ввода метрики условимся считать расстояние между соседними пикселями равным единице. Тогда если мы выберем значение  $\sigma = 2$ , то для вейвлета с индексом  $p=2$  стандартное отклонение функции Гаусса  $\sigma/k_p=8$ , что равно расстоянию до следующей точки решетки. Соответственно при  $p=0$  отклонение будет в два раза меньше, а при  $p=4$  – в два раза больше.

Непосредственное вычисление свертки

$$R(\vec{r}) = I(\vec{r}) * \psi(\vec{r}) \quad (3)$$

достаточно трудоемко. Применим к обеим частям дискретное преобразование Фурье  $F$ , в результате операция свертки преобразуется в умножение спектров

$$F[R(\vec{r})] = F[I(\vec{r})] F[\psi(\vec{r})]. \quad (4)$$

Выполняя обратное преобразование Фурье, получим

$$R(\vec{r}) = F^{-1}\{F[I(\vec{r})] F[\psi(\vec{r})]\}. \quad (5)$$

Дискретное преобразование Фурье для двумерного случая имеет вид

$$f_{k_x k_y} = \sum_{n_x=0}^{N_x-1} \sum_{n_y=0}^{N_y-1} x_{n_x n_y} \exp\left(-\frac{2\pi i}{N_x} k_x n_x\right) \exp\left(-\frac{2\pi i}{N_y} k_y n_y\right).$$

Для его выполнения удобно использовать алгоритм быстрого преобразования Фурье (БПФ).

Обработку графических изображений будем проводить с помощью некоммерческой библиотеки OpenCV. В её состав входит функция cvDFT, реализующая алгоритм прямого и обратного БПФ, в том числе и для двумерных входных данных. Особенностью БПФ является сведение задачи для  $N$  чисел к задаче для  $N_1=N/m$ , где  $m$  – делитель числа  $N$ . В реализации cvDFT оптимальный размер стороны изображения представляет собой произведение целых степеней чисел 2, 3 и 5. Соответственно для экономии памяти целесообразно заранее задать размер стороны изображения равным такому произведению.

Для уменьшения вычислительной сложности ограничимся только действительной (чётной) компонентой вейвлета Габора. В этом случае объем входных данных уменьшается в два раза, а для комплексных выходных данных используется специальный упакованный формат, также позволяющий в два раза уменьшить объем выходных данных.

Пример декомпозиции изображения можно видеть на рис. 1. Слева дано исходное изображение, в центре показаны изображения нескольких вейвлетов Габора: слева направо для значений  $q=0,2,4$ , сверху вниз – для значений  $p=0,2,4$ . Значения сигнала здесь приведены к диапазону яркости рисунка. Справа показаны результаты декомпозиции исходного изображения по этим вейвлетам в виде амплитуды отклика вейвлетов на изображение.

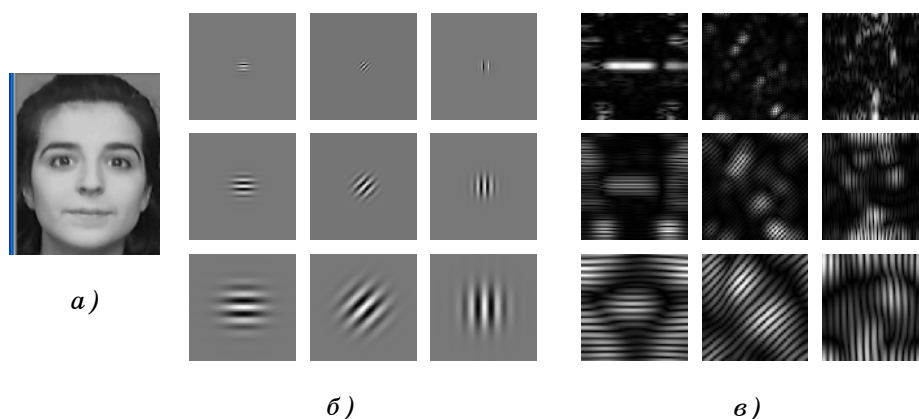


Рисунок 1 – Результаты декомпозиции изображения: а) - исходное изображение; б) вейвлеты Габора; в) амплитуды отклика вейвлетов

### ПРОГРАММА И РЕЗУЛЬТАТЫ ТЕСТИРОВАНИЯ

В соответствии с описанным алгоритмом разработана программа распознавания черно-белых графических изображений. Программа написана на языке C++ для среды Visual Studio 2008. Компьютерная обработка изображений выполняется с помощью дополнительной библиотеки OpenCV v2.0, скомпилированной для Visual Studio.

На первом этапе работы программы создается банк из 40 вейвлетов Габора для пяти различных частот и восьми ориентаций. Их примеры показаны на рис. 1б.

На втором этапе согласно входному списку файлов вводятся графические изображения, для каждого из которых по результатам декомпозиции строится вектор признаков. Допустимыми форматами для изображений являются *pgm*, *bmp*, *dib*, *png*, *tif*, *jpg*, *jpeg*. Для унификации процесса построения вектора признаков исходное изображение масштабируется к стандартному размеру.

На третьем этапе выполняется распознавание изображений. Входными данными здесь являются обучающий и тестовый наборы изображений. В данной реализации программы использован простейший метод распознавания – метод ближайших соседей.

Для тестирования программы была взята база данных ORL [4], содержащая по 10 фотографий лиц в формате *pgm* для 40 человек в соответствующих каталогах от *s1* до *s40*.

В первом тесте обучающий набор состоял из изображений с номерами 1 – 5 в каталогах *s1* – *s5*, а тестовый – из изображений с номерами 6 – 10 в тех же каталогах. Тестирование показало, что все изображения из тестового набора правильно отнесены к соответствующим классам.

Тестирование для 10 классов распознавания, для которого были взяты те же номера изображений, но уже для каталогов *s1* – *s10*, показало следующий результат (табл.1). Каждая запись в таблице представлена в формате “имя тестового изображения – имя изображения из обучающей выборки, ближайшего к тестовому в пространстве признаков”.

Таблица 1

<i>s1/6</i> – <i>s1/4</i>	<i>s3/6</i> – <i>s3/1</i>	<i>s5/6</i> – <i>s5/5</i>	<i>s7/6</i> – <i>s7/3</i>	<i>s9/6</i> – <i>s9/4</i>
<i>s1/7</i> – <i>s1/1</i>	<i>s3/7</i> – <i>s3/5</i>	<i>s5/7</i> – <i>s5/4</i>	<i>s7/7</i> – <i>s7/3</i>	<i>s9/7</i> – <i>s3/4</i>
<i>s1/8</i> – <i>s1/1</i>	<i>s3/8</i> – <i>s3/4</i>	<i>s5/8</i> – <i>s5/2</i>	<i>s7/8</i> – <i>s7/1</i>	<i>s9/8</i> – <i>s9/4</i>
<i>s1/9</i> – <i>s1/1</i>	<i>s3/9</i> – <i>s3/1</i>	<i>s5/9</i> – <i>s5/5</i>	<i>s7/9</i> – <i>s7/1</i>	<i>s9/9</i> – <i>s9/4</i>
<i>s1/10</i> – <i>s1/5</i>	<i>s3/10</i> – <i>s3/3</i>	<i>s5/10</i> – <i>s5/5</i>	<i>s7/10</i> – <i>s7/4</i>	<i>s9/10</i> – <i>s9/2</i>
<i>s2/6</i> – <i>s2/2</i>	<i>s4/6</i> – <i>s4/3</i>	<i>s6/6</i> – <i>s6/2</i>	<i>s8/6</i> – <i>s8/5</i>	<i>s10/6</i> – <i>s10/2</i>
<i>s2/7</i> – <i>s2/3</i>	<i>s4/7</i> – <i>s4/4</i>	<i>s6/7</i> – <i>s6/5</i>	<i>s8/7</i> – <i>s8/2</i>	<i>s10/7</i> – <i>s10/3</i>
<i>s2/8</i> – <i>s2/3</i>	<i>s4/8</i> – <i>s4/3</i>	<i>s6/8</i> – <i>s6/1</i>	<i>s8/8</i> – <i>s8/5</i>	<i>s10/8</i> – <i>s10/2</i>
<i>s2/9</i> – <i>s2/4</i>	<i>s4/9</i> – <i>s4/2</i>	<i>s6/9</i> – <i>s6/4</i>	<i>s8/9</i> – <i>s8/1</i>	<i>s10/9</i> – <i>s10/5</i>
<i>s2/10</i> – <i>s2/3</i>	<i>s4/10</i> – <i>s4/5</i>	<i>s6/10</i> – <i>s6/4</i>	<i>s8/10</i> – <i>s8/3</i>	<i>s10/10</i> – <i>s10/2</i>

Т.о. из 50 тестовых изображений только *s9/7* было неправильно отнесено к классу *s3*. Это подтверждает высокую эффективность разработанной программы. Тем не менее качество распознавания можно повысить за счет учета нечетной компоненты вейвлетов Габора, ввода эталонных векторов для классов распознавания, специальной локализации вейвлетов для конкретного класса распознаваемых изображений.

## ВЫВОДЫ

1 Предложен алгоритм построения вектора признаков распознавания, основанный на использовании вейвлетов Габора и не связанный с конкретным типом распознаваемых объектов.

2 Разработана программа распознавания двумерных изображений, не требующая предварительного масштабирования изображения и допускающая использование различных типов графических файлов.

3 Тестирование программы на образцах из ORL-базы изображений лиц показало высокую эффективность её работы даже при использовании только четной компоненты вейвлетов Габора и при использовании простейшего из методов распознавания – метода ближайших соседей.

## SUMMARY

### RECOGNITION OF IMAGES ON THE BASE OF THE TWODIMENSIONAL WAVELET ANALYSIS

*M.S. Babiy, O.P. Chekalov,  
Sumy State University, Sumy*

*There is described algorithm of calculation of the recognition features for twodimensional graphic images, founded on decomposition of the images on Gabor filters. Program of the image recognition is designed. During testing on samples from ORL-database of the faces, program has shown high efficiency of the recognition algorithm.*

**Key words:** *wavelets of Gabor, fast Fourier transform, graphic images, recognition, ORL-database.*

## СПИСОК ЛІТЕРАТУРИ

1. Lawrence S., Giles C. L., Tsoi A. C. and Back A. D. Face Recognition: A Convolutional Neural Network Approach // IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition. - P. 1-24.
2. B. Kepenekci. Face Recognition Using Gabor Wavelet Transform. A thesis submitted to the graduate school of natural sciens of the Middle East Technical University, September 2001.
3. Wiskott L., Fellous J. M., Kruger N., Malsburg C. Face Recognition by Elastic Bunch Graph Matching, IEEE Trans. on. PAMI. – 1997.- 19(7). - P.775-779.
4. The ORL Database of Faces [електронний ресурс]. – Режим доступа: [http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att\\_faces.tar.Z](http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.tar.Z).

*Поступила в редакцію 29 марта 2012 г.*