

УДК 004.931

АЛГОРИТМИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ КЛАССИФИКАЦИИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ ВЕЙВЛЕТ-ПРЕБРАЗОВАНИЯ ХААРА И НЕЙРОННЫХ СЕТЕЙ

Тхи Тху Чанг Буй, Нгок Хоанг Фан, В.Г. Спицын

Томский политехнический университет

E-mail: trangbt.084@gmail.com

Предложен новый алгоритм для классификации цифровых изображений, основанный на их цветовой информации, форме и текстуре. Разработана и реализована на языке объектно-ориентированного программирования C# программа для классификации цифровых изображений с использованием вейвлет-преобразования Хаара и многослойной нейронной сети. Сделан вывод о возможности использования коэффициентов прямого вейвлет-преобразования Хаара в качестве входных данных для нейронной сети при классификации цифровых изображений. Показано, что использование предложенного оригинального алгоритма, основанного на вейвлет-преобразовании Хаара и нейронной сети, дает возможность эффективной классификации цифровых изображений.

Ключевые слова:

Вейвлет-преобразование Хаара, нейронная сеть, моменты цвета, классификация цифровых изображений, обработка цифровых изображений.

Key words:

Haar wavelet transform, neural network, color moments, image classification, image processing.

Введение

Классификация цифровых изображений из баз данных характеризуется высокой трудоемкостью для традиционных алгоритмов машинного обучения из-за большого числа изображений и множества деталей, которыми описываются изображения. Указанные алгоритмы характеризуются невысокой стабильностью при классификации изображений из больших баз данных. Кроме того, эти алгоритмы требуют больших временных затрат для классификации. Существующие системы хранения изображений, такие как QBIC [1] и VisualSEEK [2] ограничивают методы классификации способами описания изображений, основанными на форме, текстуре и цветовой информации [3].

Одним из существующих методов для распознавания, классификации и восстановления изображений является метод, основанный на нейронных сетях. Чтобы уменьшить число входных нейронов сети, система классификации изображений обычно располагается на шаге предобработки. Одним из шагов предобработки цифровых изображений является вейвлет-преобразование. В настоящее время, вейвлет-преобразование является популярным методом для анализа изображений и получения информации об изображении, таких как форма и текстура.

Классификация цифровых изображений на основе вейвлет-преобразования и нейронной сети была изучена в работах [4–6]. В работе [4] использовано вейвлет-преобразование Хаара для классификации цифровых изображений из базы данных 600 изображений (300 для обучения и 300 для тестирования нейронной сети). В результате работы показано, что 81,7 % изображений обучающего множества и 76,7 % изображений тестируемого множества были правильно классифицированы. В работе [5] для классификации изображений при-

менялось вейвлет-преобразование Добеши (*Daubechies*). Использовались 120 цветных изображений самолетов для обучения нейронной сети и 240 изображений для ее тестирования. Наилучшим результатом работы является 88 % правильно классифицированных изображений. В работе [6] использовано вейвлет-преобразование Добеши для классификации изображений из базы данных 400 цветных изображений самолетов (150 для обучения и 250 для тестирования нейронной сети). В указанной работе показано, что 98 % изображений обучающего множества и 90 % изображений тестируемого множества были правильно классифицированы.

В данной работе предложен новый алгоритм, основанный на комбинации вейвлет-преобразования Хаара и нейронной сети для классификации цифровых изображений из базы данных. Цветное изображение самолета делится на три RGB компонента. Моменты цвета первого порядка и коэффициенты разложения вейвлет-преобразования Хаара [7] трех RGB компонентов изображения являются входными векторами многослойной нейронной сети, обученной алгоритмом обратного распространения ошибки. Предложенная нейронная сеть определяет принадлежность изображения самолета одной из шести категорий, рис. 1.

Представление содержания цифровых изображений

Содержание и черты цифровых изображений обычно используются в классификации изображений. В этой работе используются моменты цвета и вейвлет-коэффициенты разложения для представления содержания цифровых изображений.

Моменты цвета

Моменты цвета успешно использованы во многих системах восстановления цветного изображения [8], особенно когда изображение содержит

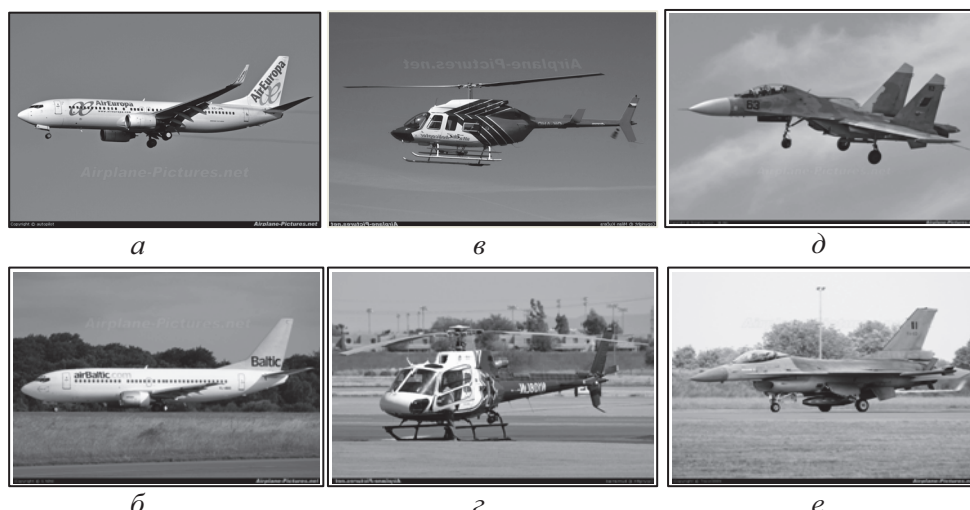


Рис. 1. Категории изображений самолетов для классификации: а) коммерческий самолет в полете; б) коммерческий самолет на земле; в) вертолет в полете; г) вертолет на земле; д) военный самолет в полете; е) военный самолет на земле

только один объект. Моменты цвета первого, второго и третьего порядков являются эффективными для представления распределения цвета изображений. В данной работе для подачи на вход нейронной сети используется момент цвета первого порядка. Моменты цвета определяются по следующим формулам:

$$\mu_k = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M p_{i,j}^k,$$

$$\sigma_k = \left(\frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (p_{i,j}^k - \mu_k)^2 \right)^{1/2},$$

$$s_k = \left(\frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M (p_{i,j}^k - \mu_k)^3 \right)^{1/3}.$$

где $p_{i,j}^k$ — значение k -го цветного компонента изображения ij -х пикселей; M и N — высота и ширина изображения; μ_k , σ_k , s_k — моменты цвета первого, второго и третьего порядков.

Вейвлет-преобразование

Вейвлет-преобразование обычно используется в системах восстановления содержания изображения. На каждом уровне вейвлет-преобразования, сигнал разлагается на 4 поддиапазона частот (*subbands*, квадранта), (LL_n , LH_n , HL_n , HH_n), где буква L — низкая частота (*low frequency*), H — высокая частота (*high frequency*) и n — уровень разложения. На рис. 2 представлены стандартные обозначения квадрантов преобразованного изображения: LL , LH , HL , HH . Квадрант LL_n представляет изображение с низким разрешением (cA_n), HL_n , LH_n , HH_n — вертикальные детали (cV_n), горизонтальные детали (cH_n), диагональные детали (cD_n) изображения. В данной работе используется вейвлет-преобразование Хаара [8] по 6 уровням разложения. Полученные вейвлет-коэффициенты подаются на входы нейронной сети.

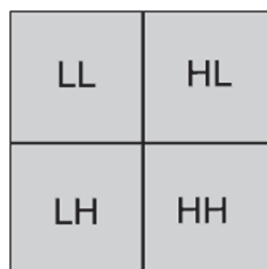


Рис. 2. Однократное применение двумерного вейвлет-преобразования к квадратному изображению

Алгоритм классификации цифровых изображений

Целью данной работы является классификация большого числа цифровых изображений на основе информации о цвете и форме изображения. Предложенный алгоритм классификации основан на совместном применении момента цвета первого порядка, вейвлет-преобразования Хаара и многослойной нейронной сети.

Сначала размер исходного изображения уменьшается до 256×256 пикселей. После этого, полученное изображение делится на 3 RGB компонента. Каждый компонент далее делится на 6 частей с одинаковым размером 128×128 пикселей (рис. 3). Входной вектор нейронной сети определяется следующими шагами:

- Вычисляется момент цвета первого порядка каждой из 6 частей 3 базовых RGB компонентов и в результате получаются 18 входов для нейронной сети, которые содержат информацию о цвете изображения.
- Применяется вейвлет-преобразование Хаара 6-го уровня разложения к каждой из 6 частей 3 базовых компонентов для получения горизонтальной, вертикальной и диагональной деталей с размером 2×2 (cH_6 , cV_6 , cD_6). Полученная информация подается на входы нейронной сети (при этом не используется изображение с низкой инфор-

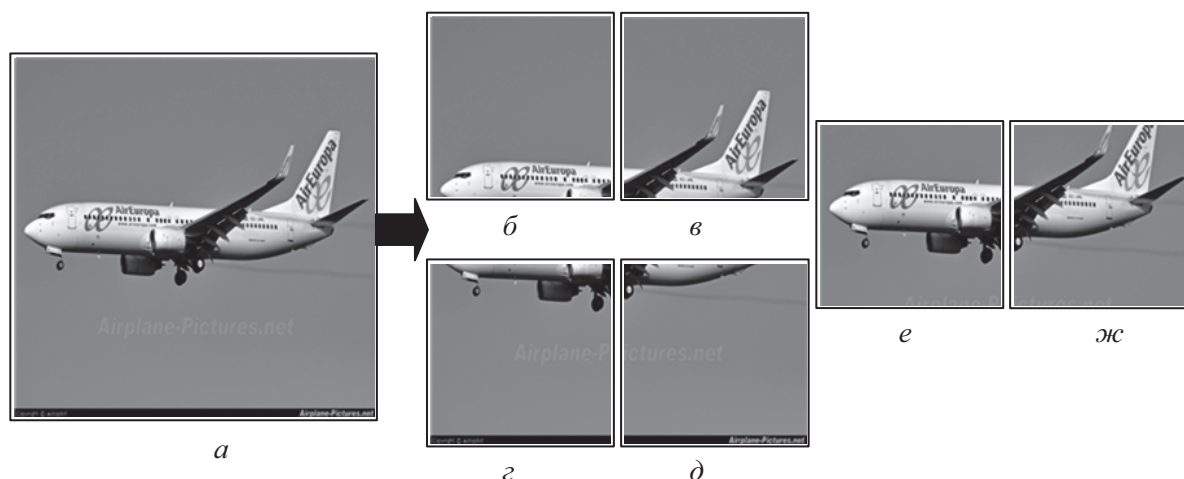


Рис. 3. Изображения самолета: а) изображение после уменьшения размера до 256×256 пикселей; б–ж) 6 частей с одинаковым размером 128×128 пикселей

мацией, полученное после применения вейвлет-преобразования (cA_6). Чтобы уменьшить число входов нейронной сети, используется только горизонтальная информация о 6 частях красного компонента (cH_6-R), вертикальная информация о 6 частях зеленого компонента (cV_6-G) и диагональная информация о 6 частях синего компонента (cD_6-B). Всего на данном этапе получаем 72 входа нейронной сети, которые содержат информацию о форме и текстуре изображения.

- Применяется вейвлет-преобразование Хаара первого уровня разложения к каждому из 3 базовых RGB компонентов. В результате определяются коэффициенты разложения Eh , Ev и Ed каждого компонента, определяющие вклад горизонтальной, вертикальной и диагональной деталей. Таким образом, получаем еще 9 входов нейронной сети. В результате, после 3 шага нейронная сеть имеет 99 входов.

Структура предложенной нейронной сети

В данной работе для классификации цифровых изображений используется многослойная нейронная сеть с одним скрытым слоем, обученная алгоритмом обратного распространения ошибки. Предложенная сеть имеет 99 входных нейронов (x_1-x_{99}), 150 скрытых нейронов (h_1-h_{150}) и 6 выходных нейронов (y_1-y_6).

Сигмоидальная функция $\sigma(z)=(1+e^{-z})^{-1}$ часто применяется при решении задач классификации. Ее преимуществом является возможность оценки вероятности принадлежности объекта классу, эффективность вычисления производной, ограниченность выходного значения. Таким образом, в данной работе в качестве функции активации используется биполярная сигмоидальная функция. Для повышения скорости сходимости сети применяются алгоритм *Nguyen-Widrow* для инициализации весов нейронной сети и моментум (импульсный параметр) при обучении сети [9]. Значение моментума равно 0,05.

Различие между предложенным алгоритмом и алгоритмами [4–6] классификации изображений показано в табл. 1. Во второй строке табл. 1 указан способ представления изображения на входном слое нейронной сети.

Таблица 1. Различие алгоритмов классификации изображений

Алгоритм	Источник информации			Предложенный алгоритм
	[4]	[5]	[6]	
Методология	• Хаар • Диагональный момент	• Добеши • Гистограмма	• Добеши • Моменты цвета	• Хаар • Моменты цвета
Структура нейронной сети (Входной – Скрытый – Выходной)	49–49–30	48–49–6	99–64–6	99–150–6

Результаты экспериментов

Программа была реализована на языке объектно-ориентированного программирования C# (*Visual studio* 2008). Она предоставляет возможность создания и обучения многослойной нейронной сети, формирования базы данных для обучения и тестирования сети. Программа имеет отдельный интерфейс для проверки и тестирования сети.

В работе использованы 150 цветных изображений самолетов для обучения и 240 изображений для тестирования предложенной нейронной сети (обрабатывались изображения из базы URL: <http://www.airplane-pictures.net>). На этапе обучения сеть сходится после 150000 эпох со среднеквадратической ошибкой 0,001. Программа была тестирована 30 раз. Результаты численных экспериментов показали, что 59...81 % изображений самолетов из тестируемого множества и устойчиво 98...100 % из обучающего множества были правильно классифицированы. Наилучший результат работы программы показал, что 100 % изображений самолетов

из обучающего множества и 81 % изображений самолетов из тестируемого множества были правильно классифицированы. Таким образом, для наилучшего результата средний процент правильно распознанных изображений составил 90 %. Классификация одного изображения размером 1024×787 пикселей происходит в среднем за 425 мс на компьютере с процессором Intel(R) Core(TM)2 Duo 2.00 ГГц и ОЗУ объемом 2 Гб).

Таблица 2. Сравнение результатов работы алгоритмов классификации изображений

Алгоритм	Источник информации			Предложенный алгоритм
	[4]	[5]	[6]	
Количество изображений для обучения	300	120	150	150
Количество изображений для тестирования	300	240	250	240
Процент правильно классифицированных изображений, %	79	88	93	90

В табл. 2 представлены результаты сравнения работы предложенного алгоритма с другими алго-

ритмами классификации изображений. Полученные результаты показывают, что использование предложенного алгоритма, основанного на комбинированном применении момента цвета первого порядка, вейвлет-преобразования Хаара и многослойной нейронной сети, дает возможность эффективной классификации цифровых изображений из базы данных.

Заключение

1. Предложен новый алгоритм решения задачи классификации цифровых изображений на основе совместного применения момента цвета первого порядка, вейвлет-преобразования Хаара и многослойной нейронной сети.
2. Разработана программа, реализующая предложенный алгоритм классификации цифровых изображений на языке объектно-ориентированного программирования C#. Показано, что использование алгоритма дает возможность эффективной классификации цифровых изображений.

Работа выполнена при финансовой поддержке гранта РФФИ № 09-08-00309.

СПИСОК ЛИТЕРАТУРЫ

1. Niblack W., Barber R., Equitz W., Flickner M., Glasman E., Petkovic D., Yanker P., Faloutsos C., Taubin G. The QBIC Project: Querying Images by Content Using Color, Texture, and Shape // Proc. Int. Conf. on Storage and Retrieval for Image and Video Databases. – Bellingham, Washington, USA, 1993. – P. 173–187.
2. Smith J.B., Chang S.F. Tools and Techniques for Color Image Retrieval // Proc. Int. Conf. on Symposium on Electronic Imaging: Science and Technology Storage and Retrieval for Image and Video Databases. – San Jose, CA, USA, 1996. – P. 426–437.
3. Swain M.J., Ballard D.H. Color indexing // International Journal of Computer Vision. – 1991. – V. 7. – № 1. – P. 11–32.
4. Park S.B., Lee J.W., Kim S.K. Content based image classification using a neural network // Pattern Recognition Letters. – 2004. – V. 25. – № 3. – P. 287–300.
5. Gonzalez A.C., Sossa J.H., Felipe E.M. Wavelet transforms and neural networks applied to image retrieval // Proc. Int. Conf. on Pattern Recognition. – Hong Kong, 2006. – P. 909–912.
6. Lofti M., Solimani A., Dargazany A., Afzal H., Bandarabadi M. Combining wavelet transforms and neural networks for image classification // 41st Southeastern Symp. on System Theory. – Tennessee, USA, 2009. – P. 44–48.
7. Буй Тхи Тху Чанг; Спицын В.Г. Разложение цифровых изображений с помощью двумерного дискретного вейвлет-преобразования и быстрого преобразования // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 73–76.
8. Flickner M., Sawhney H., Niblack W., Ashley J., Huang Q., Dom B., Gorkani M., Hafner J., Lee D., Petkovic D., Steele D., Yanker P. Query by image and video content: The QBIC system // IEEE Computer. – 1995. – V. 28. – № 9. – P. 23–32.
9. Fausett L.V. Fundamentals of Neural Networks. – Architectures, Algorithms, and Applications. – Upper Saddle River: Prentice Hall, 1994. – 476 p.

Поступила 05.09.2011 г.