ЭНЕРГЕТИЧЕСКИЕ ПРИЗНАКИ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ

М.П. Шлеймович, С.А. Ляшева

Казанский национальный исследовательский технический университет имени А.Н. Туполева-КАИ, Казань, Россия

Рассмотрены основные процедуры обработки изображений, модели представления изображений, признаки изображений. Описаны одномерное и двумерное вейвлет-преобразования. Приведены модели представления изображений на основе энергетических признаков изображений, вычисленных с применением вейвлет-преобразований.

Ключевые слова: обработка изображений, модели представления изображений, признаки изображений, вейвлет-преобразование, энергетические признаки изображений.

Введение

Современные системы обработки информации и управления базируются на различных технологиях. В качестве одних из часто применяемых можно указать технологии обработки и анализа изображений. Например, они используются в системах навигации движущихся объектов, геоинформационных системах, телекоммуникационных системах, охранных системах, системах контроля дорожного движения, системах управления технологическими процессами и т.д.

В системах, основанных на применении технологий обработки и анализа изображений, в общем случае необходимо обеспечить выполнение следующих процедур:

1. Получение изображений;

2. Преобразование изображений;

3. Выделение признаков изображений;

4. Анализ признаков изображений.

Процедура получения изображений осуществляет ввод изображений в систему в виде определенных структур данных. Следующая процедура изменяет цветовые и/или геометрические характеристики изображений для приведения их к виду, требуемому для дальнейшей обработки. Процедура выделения признаков предназначена для формирования представления изображения в виде набора признаков, значимых с точки зрения решения функциональных задач системы. Последняя процедура выдает информацию о семантическом содержании изображений, например, об объектах, их параметрах и связях на изображениях.

При реализации указанных процедур необходимо выбрать модели представления изображений на базе некоторых признаковых описаний, оптимальных согласно определенным критериям. Наилучшими в данном случае являются модели, обеспечивающие инвариантное и компактное представление изображений, которое можно применять для решения сразу нескольких задач, например, предварительной обработки, сжатия, сегментации, распознавания и др.

Таким образом, определение оптимальных признаков для построения моделей изображений является актуальной и практически значимой задачей.

Модели представления изображений

При проектировании процедур обработки и анализа изображений рассматривают модели низкого, среднего и высокого уровней представления изображений [1].

На низком уровне для представления изображений часто используют функциональные модели, вероятностные модели и иерархические модели. Функциональные модели представляют изображения в виде некоторых функций. Примером такой модели является описание изображения в виде функции пространственных координат, значения которой есть скалярные (для бинарных и полутоновых изображений) или векторные (для цветных и многоспектральных изображений) величины. Вероятностные модели описывают изображения в виде реализаций случайных процессов. Для такого описания применяются функции плотности вероятности и статистические моменты (математическое ожидание, дисперсия и т.д.). Иерархические модели представляют изображения в виде множеств изображений различных масштабов. Примером иерархической модели является гауссова пирамида изображений.

На среднем уровне для представления изображений используют описания их характерных особенностей. Здесь широко применяются контурные модели, модели областей интереса, модели точек интереса, модели структурных элементов. Формирование модели изображения в данном случае осуществляется в два этапа. На первом этапе выполняется сегментация изображения (обнаружение и выделение контуров, областей интереса, точек интереса, структурных элементов), а на втором – его описание в виде векторов признаков, характеризующих соответствующие элементы.

Высокий уровень представления изображений базируется на моделях явного и неявного использования знаний. Примером модели неявного использования знаний является модель, в которой применяются шаблонные изображения, т.е. знания об объектах содержатся в изображениях этих объектов. В моделях явного использования знаний применяются описания в виде правил интерпретации информации, содержащейся в изображении.

Большинство используемых на практике подходов к анализу изображений основано на применении моделей неявного использования знаний. При этом описание изображений базируется на признаках цвета, текстуры, формы и структуры [2].

Признаки цвета предназначены для представления изображений с точки зрения их цветового содержания. Среди различных цветовых признаков пользуются популярностью гистограмма цветов [3], вектор цветовой связности [4], коррелограмма цветов [5], цветовые моменты [6], дескриптор доминантного цвета [7].

Признаки текстуры определяют пространственное распределение цветов (или яркостей) пикселей изображений. Часто применяют статистические текстурные признаки [2], локальные бинарные шаблоны [8], спектральные признаки [9], признаки Тамуры [10] и др.

Признаки формы относятся к областям изображений. К ним относятся, например, округлость области или ее прямоугольность, периметр, площадь, ориентация главных осей и др. [2, 11].

Признаки структуры позволяют учесть наличие на изображениях определенных объектов и их взаимного расположения [2].

В настоящее время активно разрабатываются также модели, основанные на вейвлет-преобразованиях.

Оценки энергии точек изображения на основе вейвлет-преобразования

Вейвлет-преобразование позволяет выполнить многомасштабный анализ сигнала, представленного в виде некоторой функции [12]. Оно выражается следующим образом:

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{s}} \psi^*\left(\frac{x-u}{s}\right) dx, \quad (1)$$

где Wf – результат преобразования; f – исходная функция; ψ^* – комплексное сопряжение сдвинутой и масштабированной вейвлет-функции; u, s – параметры сдвига и масштаба.

Вейвлет-преобразование раскладывает сигнал по функциям $\psi_{u,s}(x)$:

$$Wf(u,s) = \langle f, \psi_{u,s} \rangle, \quad (2)$$

$$\psi_{u,s}(x) = \frac{1}{\sqrt{s}} \psi\left(\frac{x-u}{s}\right), \quad (3)$$

что позволяет выявить его особенности в локальной области (определяется параметром u) при некотором масштабе (определяется параметром s). Функцию ψ называют материнским вейвлетом. Она имеет нулевое среднее значение, центр в нулевой точке и единичную норму.

Выражения (1) - (3) определяют одномерное вейвлет-преобразование. Его обобщение для многомерного сигнала имеет вид [13]:

$$\psi_{u,s}(\mathbf{x}) = \frac{1}{s^{D/2}} \psi\left(\frac{\mathbf{x}-\mathbf{u}}{s}\right), \quad (4)$$

$$Wf(\mathbf{u},s) = \langle f, \psi_{u,s} \rangle, \quad (5)$$

$$Wf(\mathbf{u},s) = \int_{-\infty}^{+\infty} f(\mathbf{x}) \frac{1}{s^{D/2}} \psi^*\left(\frac{\mathbf{x}-\mathbf{u}}{s}\right) d\mathbf{x}, \quad (6)$$

где D – размерность сигнала. Для изображений $D = 2$, $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{u} = (u_1, u_2)^T$.

Вейвлет-преобразования, представленные выражениями (1) - (6) являются непрерывными. На их основе строятся различные эффективные признаковые описания изображений, например, контурные и текстурные [12, 14, 15]. Однако на практике их применение подразумевает приближенный характер результатов и относительно медленные процедуры вычислений.

Для дискретных сигналов, к которым относятся цифровые изображения, можно применить дискретные ортогональные кратно-масштабные вейвлет-преобразования (далее – дискретные вейвлет-преобразования), свободные от указанных недостатков [12]. В их основе лежит представление дискретной функции $f(x)$ в виде суммы

$$f(x) = f_a(x) + f_d(x) \tag{6}$$

аппроксимирующей

$$f_a(x) = \frac{1}{\sqrt{N}} \sum_{k=-\infty}^{+\infty} a_{j_0,k} \varphi_{j_0,k}(x) \tag{7}$$

и детализирующей

$$f_d(x) = \frac{1}{\sqrt{N}} \sum_{j=j_0}^{+\infty} \sum_{k=-\infty}^{+\infty} d_{j,k} \psi_{j,k}(x) \tag{8}$$

составляющих, где N – число значений дискретной последовательности, представляющей исходную функцию; j_0 – уровень разложения; $a_{j,k}$ и $d_{j,k}$ – коэффициенты аппроксимации и детализации соответственно; $\varphi_{j,k}(x)$ и $\psi_{j,k}(x)$ – базисные функции. Процесс преобразования можно представить в итерационном виде:

$$J = \log_2 N\,, \tag{9}$$

$$f(x) = f_a^J(x)\,, \tag{10}$$

$$f_a^j(x) = f_a^{j-1}(x) + f_d^{j-1}(x)\,, \tag{11}$$

где $j = J, \ldots, j_0 + 1$.

В случае конечных дискретных сигналов дискретное вейвлет-преобразование позволяет получить их представление в виде конечного множества коэффициентов аппроксимации и детализации.

Одномерное дискретное вейвлет-преобразование можно обобщить на двумерный случай для изображений. На практике обычно применяют разделимые преобразования, которые выполняются в два этапа: сначала выполняется преобразование для строк, а затем – для столбцов либо в обратном порядке. В случае если сначала преобразование выполняется по x, а затем – по y, применяются следующие базисные функции:

$$\varphi_{j,m,n}^{LL}(x,y) = \varphi_{j,m}(x)\varphi_{j,n}(y)\,, \tag{12}$$

$$\psi_{j,m,n}^{LH}(x,y) = \varphi_{j,m}(x)\psi_{j,n}(y)\,, \tag{13}$$

$$\psi^{HL}_{j,m,n}(x,y)=\psi_{j,m}(x)\varphi_{j,n}(y)\,, \qquad (14)$$

$$\psi^{HH}_{j,m,n}(x,y)=\psi_{j,m}(x)\psi_{j,n}(y)\,. \qquad (15)$$

Соответствующие коэффициенты аппроксимации и детализации обозначим $LL_{j,m,n}$, $LH_{j,m,n}$, $HL_{j,m,n}$, $HH_{j,m,n}$. Отметим, что здесь и далее описывается преобразование для одноканальных изображений. Для многоканальных изображений каждый канал подвергается преобразованию по отдельности.

Для анализа изображений на основе дискретного вейвлет-преобразования эффективно применение признакового описания, построенного на основе коэффициентов детализации различных уровней [16]. Устойчивое к шуму представление изображений можно получить на основе энергетических признаков. Для одномерного дискретного вейвлет-преобразования с ортонормированными базисными функциями и дискретного сигнала $\{f_0, f_1, \ldots, f_{N-1}\}$ справедливо равенство:

$$\sum_{k=0}^{N-1} f_k^2 = \sum_{l=0}^{2^{j_0}-1} a_{j_0,l}^2 + \sum_{j=j_0}^{J-1}\sum_{l=0}^{2^j-1} d_{j,l}^2\,, \qquad (16)$$

представляющее собой аналог известного равенства Парсеваля. В случае изображений равенство (16) принимает следующий вид:

$$\begin{aligned}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} f_{k,l}^2 &= \sum_{m=0}^{2^{j_0}-1}\sum_{n=0}^{2^{j_0}-1} LL_{j_0,m,n}^2 + \\ &+\sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} LH_{j,m,n}^2 + \\ &+\sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} HL_{j,m,n}^2 + \\ &+\sum_{j=j_0}^{J-1}\sum_{m=0}^{2^j-1}\sum_{n=0}^{2^j-1} HH_{j,m,n}^2\,,\end{aligned} \qquad (17)$$

где изображение представляет собой квадратную матрицу $\left[f_{k,l}\right]_{k,l=0}^{N-1}$, размер которой есть степень 2, а результаты преобразования на j-м уровне группируются в матрицу аппроксимирующих коэффициентов $\left[LL_{j,m,n}\right]_{m,n=0}^{2^j-1}$ и матрицы детализирующих горизонтальных $\left[LH_{j,m,n}\right]_{m,n=0}^{2^j-1}$, вертикальных $\left[HL_{j,m,n}\right]_{m,n=0}^{2^j-1}$, диагональных $\left[HH_{j,m,n}\right]_{m,n=0}^{2^j-1}$ коэффициентов.

На основе равенства (17) можно получить оценки энергии для каждой точки изображения посредством выполнения следующего алгоритма:

1. Выполнить преобразование до уровня j_0;

2. Положить:

$$w_{j_0-1,m,n}^2 = LL_{j_0,m,n}^2\,; \qquad (18)$$

3. Последовательно для $j = j_0, \ldots, J-1 (J = \log_2 N)$, $m, n = 0, 1, \ldots, 2^{j+1}-1$ вычислить оценки энергии:

$$w^2_{j,m,n} = \frac{1}{4} w^2_{j-1,m,n} + \\ + LH^2_{j,m/2,n/2} + HL^2_{j,m/2,n/2} + HH^2_{j,m/2,n/2}. \tag{19}$$

Можно убедиться в том, что полученные значения сохраняют энергетическое равенство (17):

$$\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} f^2_{k,l} = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} w^2_{k,l}, \tag{20}$$

где

$$w^2_{k,l} = w^2_{J-1,k,l}. \tag{21}$$

Каждое из множества значений $\{w^2_{k,l}\}_{k,l=0}^{N-1}$ или $\{w_{k,l}\}_{k,l=0}^{N-1}$ может служить весом соответствующей точки, характеризующим ее вклад в полную энергию изображения.

Наиболее быстрым и простым вейвлет-преобразованием изображений является двумерное преобразование Хаара, которое выполняется по формулам:

$$LL_{j-1,m,n} = \frac{1}{4}[LL_{j,2m,2n} + LL_{j,2m+1,2n} + \\ + LL_{j,2m,2n+1} + LL_{j,2m+1,2n+1}], \tag{22}$$

$$LH_{j-1,m,n} = \frac{1}{4}[LL_{j,2m,2n} + LL_{j,2m+1,2n} - \\ - LL_{j,2m,2n+1} - LL_{j,2m+1,2n+1}], \tag{23}$$

$$HL_{j-1,m,n} = \frac{1}{4}[LL_{j,2m,2n} - LL_{j,2m+1,2n} + \\ + LL_{j,2m,2n+1} - LL_{j,2m+1,2n+1}], \tag{24}$$

$$HH_{j-1,m,n} = \frac{1}{4}[LL_{j,2m,2n} - LL_{j,2m+1,2n} - \\ - LL_{j,2m,2n+1} + LL_{j,2m+1,2n+1}], \tag{25}$$

где $j = J, \ldots, j_0 + 1$; $m, n = 0, 1, \ldots, 2^{j+1}$; $LL_{J,m,n}$ – пиксели изображения. Однако базисные функции данного преобразования не являются ортонормированными (хотя они и являются ортогональными). Поэтому для вычисления оценок энергии на его основе описанный выше алгоритм необходимо модифицировать следующим образом:

1. Выполнить преобразование до уровня j_0;

2. Положить:

$$w^2_{j_0-1,m,n} = 2^J LL^2_{j_0,m,n}; \tag{26}$$

3. Последовательно для $j = j_0, \ldots, J-1$ ($J = \log_2 N$), m, $n = 0, 1, \ldots, 2^{j+1} - 1$ вычислить оценки энергии:

$$w^2_{j,m,n} = \frac{1}{4} w^2_{j-1,m,n} + 2^{2(J-j)}[LH^2_{j,m/2,n/2} + \\ + HL^2_{j,m/2,n/2} + HH^2_{j,m/2,n/2}]. \quad (27)$$

При текстурном и контурном анализе часто необходимо получить распределение интенсивности без учета средней яркости или энергии изображений. В этом случае на первом шаге алгоритма можно положить

$$w^2_{j_0-1,m,n} = 0. \quad (28)$$

Обобщить алгоритм с целью получения большей гибкости можно следующим образом:

1) изменить выражения (18), (26), (28):

$$w^2_{j_0-1,m,n} = K_{j_0-1} LL^2_{j_0,m,n}, \quad (29)$$

2) изменить выражения (19) и (27):

$$w^2_{j,m,n} = \frac{1}{4} w^2_{j-1,m,n} + K_j [LH^2_{j,m/2,n/2} + \\ + HL^2_{j,m/2,n/2} + HH^2_{j,m/2,n/2}]. \quad (30)$$

При этом значения нормировочных коэффициентов можно считать настроечными параметрами, специфичными для конкретной задачи.

Веса точек изображений более устойчивы к шуму по сравнению со значениями интенсивностей. В таблице ниже показаны среднеквадратические разности между полутоновыми текстурными изображениями ($D1$ и $D2$ из базы Brodatz) и их зашумленными версиями (шум имеет нормальное распределение с нулевым средним и стандартным отклонением σ). Вычисления выполнены для интенсивностей и энергетических весов изображений, приведенных к размерам 512 × 512. Интенсивности точек изображений имеют значения из диапазона от 0 до 255.

Можно увидеть, что с ростом стандартного отклонения шума, среднеквадратическая разность при использовании изображений в исходной форме увеличивается на большую величину по сравнению с использованием изображений, представленных энергетическими признаками.

Табл. 1. Воздействие шума на изображения

σ	Среднеквадратические разности			
	Интенсивности		Веса	
	*D*1	*D*2	*D*1	*D*2
5	1714,75	1713,56	450,96	510,967
10	3372,63	3372,25	827,04	882,078
15	5035,82	5029,61	1268,19	1361,88
20	6682,05	6659,92	1789,63	1865,28
25	8297,3	8251,36	2051,83	2395,89

Заключение

Описанный подход может служить основой для построения моделей представления изображений, применяемых для широкого спектра задач обработки и анализа изображений. Например, он был применен для поиска и идентификации объектов на изображениях, полученных в бортовой системе беспилотного летательного аппарата [17].

Литература

1. Потапов, А.С. Распознавание образов и машинное восприятие: общий подход на основе принципа минимальной длины описания/ А.С. Потапов. – СПб.: Политехника, 2007. – 548 с.
2. Шапиро, Л. Компьютерное зрение/ Л. Шапиро, Дж. Стокман: пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
3. Long, F. Fundamentals of content-based image retrieval/ F. Long, H. Zhang, D. Feng // Multimedia Information Retrieval and Management – Technological Fundamentals and Applications. - Springer-Verlag, 2003. - P. 1-26.
4. Pass, G. Histogram refinement for content-based image retrieval/ G. Pass, R. Zabih // IEEE Workshop on Applications of Computer Vision. – 1996. - P. 96-102.
5. Huang, J. Spatial Color Indexing and Applications/ J. Huang, S.R. Kumar, M. Mitra, W.-J. Zhu, R. Zabih // International Journal of Computer Vision. – 1999. – Vol. 35(3). - P. 245–268.
6. Stricker, M. Similarity of Color Images/ M. Stricker, M. Orengo.// Proceedings of the SPIE Conference. – 1995. – Vol. 2420. – P. 381-392.
7. Deng, Y. An efficient color representation for image retrieval/ Y. Deng, B.S. Manjunath, Ch. Kenney, M.S. Moore, H. Shin // IEEE Transactions on image processing. – 2001. - Vol. 10, № 1. – P. 140-147.
8. Pietikinen, M. Computer Vision Using Local Binary Patterns/ Matti Pietikinen, Abdenour Hadid, Guoying Zhao, Timo Ahonen. – Springer, 2011.
9. Tuceryan, M. Texture analysis/ M. Tuceryan, A.K. Jain // The Handbook of Pattern Recognition and Computer Vision (2nd edition) / C.Chen, L.F.Pau, P.S.P.Wang (Eds.). - World Scientic Publishing Co., 1998. - P. 207-248.
10. Tamura, H. Texture features corresponding to visual perception/ H. Tamura, S. Mori, T. Yamawaki // IEEE Transactions on Systems, Man, and Cybernetics. – 1978. - Vol. SMC-8, № 6. – P. 460 - 473.
11. Image Databases: Search and Retrieval of Digital Imagery/ V.Castelli, L.D.Bergman (Eds.). - Wiley: New York, 2002.
12. Малла, С. Вейвлеты в обработке сигналов./ С. Малла: пер. с англ. - М.: Мир, 2005. – 671 с.
13. Addison, P.S. The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance/ P.S. Addison. - Institute of Physics Publishing, 2002.
14. Ma, W.Y. A comparison of wavelet features for texture annotation / W.Y. Ma, B.S. Manjunath// Proc. of IEEE Int. Conf. on Image Processing. – 1995. - Vol. II. - P. 256-259.
15. Tang, Y.Y. Wavelet Theory and Its Application to Pattern Recognition/ Y.Y. Tang. - World Scientific Publishing Company, 2009.
16. Sebe, N. Robust Computer Vision – Theory and Applications/ N. Sebe, M.S. Lew. – Kluwer Academic Publishers, 2003.
17. Ляшева, С.А. Распознавание объектов на местности в системах управления беспилотных летательных аппаратов/ С.А. Ляшева, М.В. Медведев, М.П. Шлеймович// Изв. вузов. Авиационная техника. – 2014. - № 3. - С. 64-66.