

Е. И. Сахарова, А. А. Макашов,
А. Н. Кропотов

ИСПОЛЬЗОВАНИЕ ВЕЙВЛЕТОВ ХААРА ДЛЯ ОБРАБОТКИ И СКЛЕЙКИ ИЗОБРАЖЕНИЙ

Рассмотрена возможность применения вейвлетов Хаара для сжатия изображений. Предложен алгоритм видеосклейки для изображений, разложенных в вейвлет-ряд. Рассмотрены достоинства и недостатки использования вейвлет-преобразований для сжатия и обработки изображений в видеоинформационных системах мобильных роботов.

E-mail: sm42@sm.bmstu.ru

Ключевые слова: вейвлет Хаара, вейвлет-разложение, обработка изображений, алгоритмы сжатия изображений, видеосклеика.

В видеоинформационных системах мобильных робототехнических систем одной из основных проблем является оперативная обработка и хранение изображений. Несжатый двумерный массив имеет очень большой размер, и для работы с ним требуется высокое быстродействие и большой объем памяти. Особо остро чувствуется нехватка ресурсов при обработке большого числа изображений или получении изображения большого размера в процессе составления масштабированных фотопланов поверхности (видеосклеика), над которой движется мобильный аппарат (летающий, подводный или наземный).

Для ускорения процесса видеосклейки применяют сжатие поступающей видеоинформации. Существует множество подходов к сжатию видеоинформации, каждый из которых имеет свои достоинства и недостатки. Помимо вейвлетов для этого могут применяться разреженные матрицы, пространство волновых чисел и преобразования Фурье, сегментация, векторизация, фрактальное представление и другие преобразования [1]. Вейвлеты удобны тем, что древовидная структура хранения данных позволяет избавиться от лишней информации в памяти и проведения лишних вычислений, т.е. хранимую информацию легко оптимизировать по нужному критерию.

В технических приложениях видеосклеика применяется для построения траектории мобильного аппарата, а также для получения карты местности.

Вейвлеты Хаара для обработки изображений определяются как

$$\phi_i^j(x) = \phi(2^j x - i) \quad (i = 0, \dots, 2^j - 1);$$

$$\phi(x) = \begin{cases} 1 & \text{при } 0 \leq x < 1; \\ 0 & \text{в противном случае,} \end{cases}$$

$$\psi_i^j(x) = \psi(2^j x - i) \quad (i = 0, \dots, 2^j - 1);$$

$$\psi(x) = \begin{cases} 1 & \text{при } 0 \leq x < 1/2; \\ -1 & \text{при } 1/2 \leq x < 1; \\ 0 & \text{в противном случае.} \end{cases}$$

Любая функция может быть представлена в виде суммы базисных функций:

$$f(x) = \sum_{ij} c_i^j \phi_i^j(x) + \sum_{ij} d_i^j \psi_i^j(x),$$

где

$$c_i^j = \frac{c_i^{j+1} + c_{i+1}^{j+1}}{2}; \quad d_i^j = \frac{c_i^{j+1} - c_{i+1}^{j+1}}{2}$$

– рекуррентное соотношение дерева вейвлет-разложения.

Для полного изображения необходимо определить двумерное вейвлет-преобразование. Рассмотрим два возможных построения двумерного вейвлета Хаара [2].

Поскольку нестандартный двумерный базис занимает меньший объем памяти и требует $(8/3)(m^2 - 1)$ против $4(m^2 - m)$ операций, будем использовать именно его. Базис состоит из одной масштабирующей функции, соответствующей первому приближению $\phi\phi_{00}^0(x, y) = \phi\phi(x, y)$, а также трех масштабируемых сдвинутых вейвлетов $\phi\psi$, $\psi\phi$, $\psi\psi$:

$$\phi\psi_{kl}^j(x, y) = 2^j \phi\psi(2^j x - k, 2^j y - l);$$

$$\psi\phi_{kl}^j(x, y) = 2^j \psi\phi(2^j x - k, 2^j y - l);$$

$$\psi\psi_{kl}^j(x, y) = 2^j \psi\psi(2^j x - k, 2^j y - l),$$

где j – степень двойки, соответствующая текущему уровню ряда; k – место в текущем ряду по x ; l – место в текущем ряду по y .

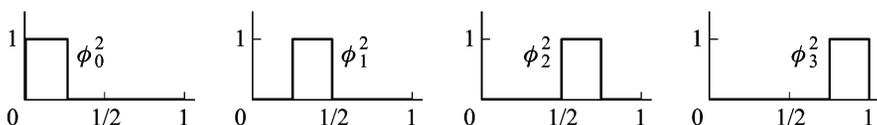


Рис. 1. Базис характеристических функций вейвлета Хаара для второго уровня разложения

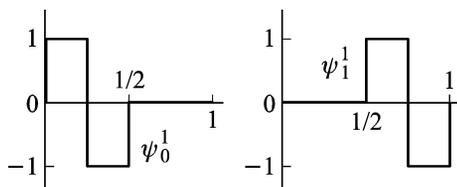


Рис. 2. Вейвлеты Хаара для первого уровня разложения

Используя формулы для построения базиса, строим дерево вейвлет-разложения. Поскольку при построении двумерных вейвлетов Хаара вычисляются вертикальная, горизонтальная и диагональная разности, для сжатия видеoinформации можно использовать граничное пороговое значение модуля разности, меньше которого участок ветви считается монохромным и все потомки узла удаляются. Модуль разности можно вычислять, например, как сумму модулей трех дочерних вейвлетов одного узла (норма-1) или как корень из суммы квадратов модулей (норма-2). При таком сжатии видеoinформации, у дерева в узлах появляются “пустые” потомки, что позволяет значительно сократить объем используемой памяти и производимых вычислений. Во избежание нивелирования больших областей, можно установить пороговое значение для каждого уровня в отдельности или вообще не устанавливать порога для близких к корню вейвлет-дерева уровней.

Данный алгоритм сжатия дает хорошие результаты. Для изучения особенностей сжатия и восстановления изображений взято изображение 720×576 (414720) пикселей (рис. 3). Результаты работы алгоритма при различных уровнях сжатия приведены на рис. 4.

Таким образом, при сохранении всего лишь 2380 значимых чисел из 414720 (0,57%) сохраняется представление о том, что было на исходном изображении. Замена вейвлетов Хаара более гладкими функциями (например, сплайн-вейвлетами или синусно-косинусным базисом) сделает сжатые изображения более гладкими, т.е. артефактов будет меньше.

При построении большого изображения необходимо воспользоваться таким алгоритмом вейвлет-разложения, который будет учитывать относительный сдвиг, поскольку, как следует из формул вейвлет-базиса, данное разложение не инвариантно по отношению к сдвигу.



Рис. 3. Исходное изображение

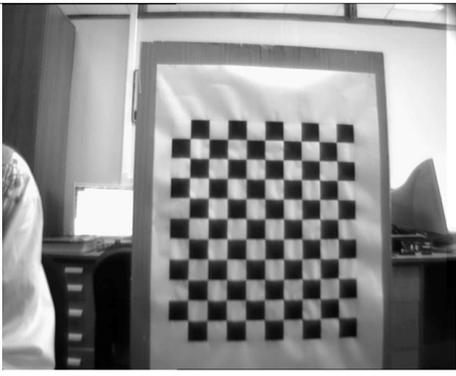
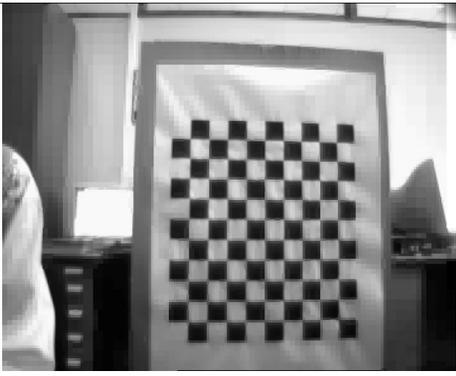
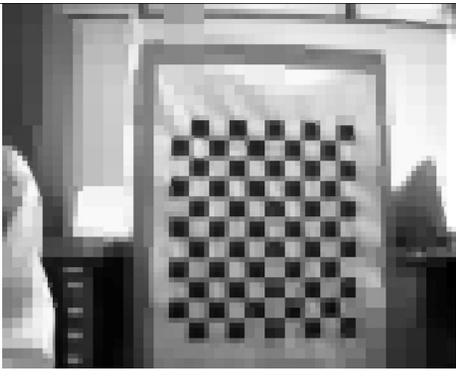
Значение порога (количество вейвлетов -> количество чисел)	Изображение
10 (32230 -> 96691)	
100 (4105-> 12316)	
400 (793 -> 2380)	

Рис. 4. Качество сжатия вейвлетами Хаара в зависимости от установленного порога

При учете сдвига объединение двух деревьев превратилось бы просто в добавление недостающих ветвей. Таким образом, перед добавлением в дерево большого изображения новых ветвей, необходимо рассчитать смещение малого изображения относительно нулевой точки.

Для этих целей можно использовать несколько различных подходов. С точки зрения мультиразрешающего анализа удобно использовать метод оптического потока [3] с последующими уточнени-

ями по уровням. Однако тогда ввиду неинвариантности вейвлет-преобразования относительно сдвига разложение необходимо делать дважды: первый раз — для вычисления относительного смещения между кадрами, второй раз — перед добавлением ветвей в большое изображение. Также можно определить разницу между текущим и предыдущим изображениями любым подходящим для этого методом.

Как было уже указано ранее, в процессе создания видеосклейки существуют две основные проблемы:

- проблема позиционирования и масштабирования изображений перед добавлением;
- проблема оптимизации хранения и обработки данных.

К решению проблемы относительного позиционирования изображений существует два принципиально различных подхода, один из которых основан на использовании в процессе склейки навигационной информации, полученной с других датчиков. Достоинством данного подхода является ускорение процесса видеосклейки ввиду отсутствия необходимости проведения позиционных вычислений. Недостатки — наличие погрешностей из-за невозможности оценить содержание видеoinформации, невозможность в ряде случаев получить доступ к навигационной информации (неуниверсальность метода).

Более универсальным подходом к решению данной проблемы является позиционирование изображений на основе содержащейся в них видеoinформации. Недостатком этого подхода является возможное наличие шумовых помех, движущихся объектов и искажений, вызванных физическими параметрами камеры, — это погрешности видеосъемок. Из искажений, вызванных физическими параметрами камеры, дисторсия убирается заранее, перед началом съемки, в процессе калибровки видеокамеры по паттерну “шахматная доска” (см. рис. 3) с помощью специально разработанной программы.

В основе используемого метода оптического потока лежит предположение о непрерывности изменения яркости в каждой точке изображения:

$$dE = \frac{\partial E}{\partial x} dx + \frac{\partial E}{\partial y} dy + \frac{\partial E}{\partial t} dt = 0.$$

Таким образом, для каждой точки экрана можно записать

$$\frac{\partial E}{\partial x} \Delta x_{\text{экр}} + \frac{\partial E}{\partial y} \Delta y_{\text{экр}} = -\frac{\partial E}{\partial t};$$
$$x = \frac{\lambda X}{Z}; \quad y = \frac{\lambda Y}{Z},$$

где λ — фокусное расстояние в пикселях; X, Y, Z — координаты точки относительно камеры; x, y — координаты пикселя;

$$\Delta x = \tilde{\eta}_1 - x \cdot \tilde{\eta}_3 - y \cdot \tilde{\phi}_3;$$

$$\Delta y = \tilde{\eta}_2 - y \cdot \tilde{\eta}_3 + x \cdot \tilde{\phi}_3;$$

$$\tilde{\eta}_1 = \frac{\eta_1}{1 + \eta_3}, \quad \tilde{\eta}_2 = \frac{\eta_2}{1 + \eta_3}, \quad \tilde{\eta}_3 = \frac{\eta_3}{1 + \eta_3}, \quad \tilde{\phi}_3 = \frac{\theta}{1 + \eta_3};$$

$$\eta_1 = \lambda \cdot \frac{\Delta X}{Z_0}, \quad \eta_2 = \lambda \cdot \frac{\Delta Y}{Z_0}, \quad \eta_3 = \frac{\Delta Z}{Z_0}.$$

После чего приведенные координаты $\tilde{\eta}_1$ и $\tilde{\eta}_2$, $\tilde{\eta}_3$ и $\tilde{\phi}_3$ (горизонтальное и вертикальное смещения кадра, масштаб, угол относительного поворота) находятся методом наименьших квадратов.

Для уменьшения влияния помех, вызванных наличием шумов и движущихся объектов, используется алгоритм предсказателя, основанный на экстраполяции движения, вычисленного на предыдущем кадре. В основе работы предсказателя лежит предположение, что скорость видеокамеры за время, прошедшее между съемкой двух последовательных кадров, изменяется незначительно, и экстремум функции оптического потока будет находиться в окрестности предыдущей точки. Также для решения этой задачи можно использовать поиск взвешенной суммы при помощи метода медианной фильтрации оптического потока, разработанного Ф. Торром (Torr' M-Estimator) [4].

Как уже было отмечено ранее, вейвлет-разложение очень удобно использовать для масштабирования изображения. Древовидная структура позволяет показывать изображения нужного масштаба, обращаясь непосредственно к нужному уровню структуры. Так, при разрастании изображения видеосклейки до размера, превышающего экран монитора, можно выводить изображение, находящееся уровнем выше. Масштаб при уменьшении вейвлет-структуры на один уровень уменьшится в 2 раза (рис. 5, верхнее и нижнее изображения).

Возможность динамической подстройки процедуры разложения и вывода позволяет несколькими способами оптимизировать вычисления как по памяти, так и по быстродействию. Рассмотрим два возможных варианта оптимизации вычислений: отрезание корня и отрезание листьев.

Отрезание листьев — это переход на уровень ближе к корню; оптимизация за счет потери мелких деталей изображения. Отрезание корня — это переопределение верхнего уровня как следующего от корня, со сбросом боковых ветвей; оптимизация за счет сброса дальних участков склейки. При отрезании корня сбрасываемая информация может быть записана на жесткий диск с возможностью последующей загрузки в случае необходимости. Выбор конкретного алгоритма оптимизации

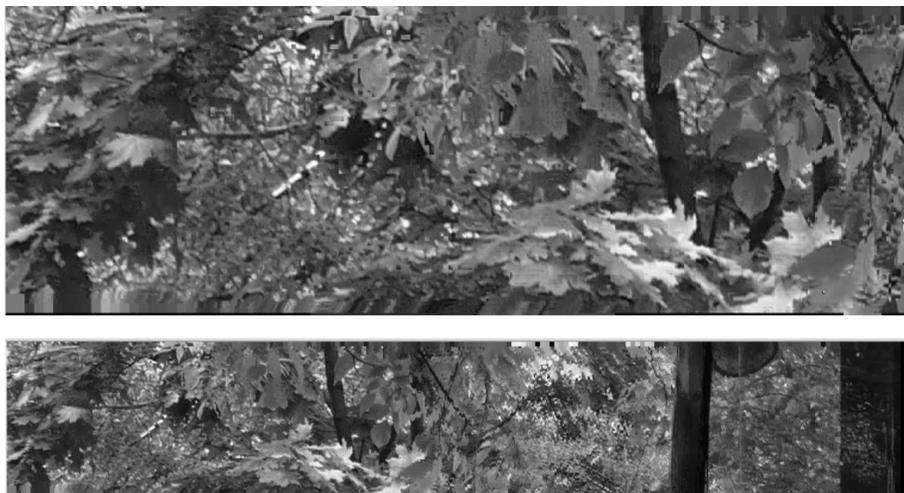


Рис. 5. Результаты работы склейки

зависит от поставленной перед видеосклейкой задачи. Если необходимо получить подробную карту, предпочтительнее отрезание корня; если речь идет лишь о восстановлении траектории движения, можно использовать отрезание листьев.

В заключение можно сделать вывод, что использование вейвлет-разложения для обработки и склейки изображений является перспективным методом, позволяющим оптимизировать вычисления по памяти и быстродействию, что позволит обеспечить более высокие характеристики мобильных робототехнических систем. Недостатком данного метода является неинвариантность вейвлетов, вследствие чего расчет перемещений необходимо проводить отдельно, до выполнения вейвлет-разложения.

СПИСОК ЛИТЕРАТУРЫ

1. Ян е Б. Цифровая обработка изображений. – М.: Техносфера, 2007.
2. Столниц Э., ДеРоуз Т., Салезин Д. Вейвлеты в компьютерной графике. R&C Dynamics. – Москва–Ижевск, 2002.
3. Артюхов Т. Ю., Кропотов А. Н., Макашов А. А., Сахарова Е. И. Опыт создания системы локальной видеонавигации для подводных аппаратов // Труды конференции “Системы технического зрения” ИКИ РАН, 2010.
4. Torr P. H. S. and Murray D. W. The development and comparison of robust methods for estimating the fundamental matrix // International Journal Computer Vision. – 1997. – Vol. 24. No. 3. – P. 271–300.

Статья поступила в редакцию 23.03.2012