

Переход от 2D- к 3D-изображениям: модификация двухмасштабной модели и алгоритмов обработки

П. А. Чочиа

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия

Поступила в редколлегию 15.09.2014

Аннотация—Формулируются особенности трехмерных изображений. Рассматривается применение двухмасштабной модели к 3D-изображениям. Показываются способы модификации различных алгоритмов обработки применительно к трехмерным изображениям.

КЛЮЧЕВЫЕ СЛОВА: обработка изображений, трехмерное изображение, модель изображения.

1. ВВЕДЕНИЕ

Под термином *трехмерное* или *3D-изображение* часто понимаются совершенно различные виды данных [1–5], основные из которых следующие.

1. Данные, задаваемые функцией трех координат, которая гомеоморфно отображает некоторый объемный участок трехмерного пространства, включая все содержащиеся в нем объекты.
2. Стереоскопическое изображение; пара соответствующим образом получаемых двумерных изображений, которые за счет диспаратности дают представление о пространственном расположении объектов, видимых с точки наблюдения.
3. Изображение, являющееся некоторой проекцией (например, аксонометрической) объектов исходной трехмерной сцены; оно дает наблюдателю представление о форме и расположении объектов, однако само по себе остается двумерным.
4. Двумерное изображение, каждая точка которого соответствует некоторым координатам в трехмерном пространстве, например, дальности или рельефу.
5. Особым способом сформированные изображения, создающие образы наблюдаемых объектов, например голограммы.
6. Видеопоследовательности, включающие набор кадров наблюдаемых объектов. Такие данные могут быть записаны в виде трехмерного массива, аналогично первому типу, однако при этом одна из координат является не пространственной, а координатой времени.

В дальнейшем под трехмерным изображением (непрерывным или дискретным) будут пониматься изображения исключительно первого типа. По существу, 3D-изображение представляет собой расширение обычного 2D-изображения добавлением еще одного пространственного измерения. Предполагается, что значения точек исходного пространства с координатами (m, n, k) описываются функцией $f(m, n, k)$, имеющей свойства плотности некоторой выбранной физической характеристики. Таким образом, *трехмерное изображение* (или *3D-изображение*) представляет собой набор данных, записываемых в дискретном представлении в виде трехмерного массива, элементы которого соответствуют точкам отображаемого участка трехмерного пространства, а значение каждого элемента отражает величину выбранной физической характеристики в окрестности соответствующей точки пространства, усредненную согласно

заданной апертурной функции. Иногда такие изображения также называют *объемными изображениями*. По аналогии с двумерным изображением, дискретный элемент которого называют *пиксель*, дискретный элемент 3D-изображения называют *воксель* (англ. *voxel* — volumetric pixel) — элемент объемного изображения.

Способы формирования трехмерного изображения могут быть разными; наиболее известным является томографическое сканирование — рентгеновская компьютерная томография (КТ) или магнитно-резонансная томография (МРТ). Кроме того, возможно получение 3D-изображения в результате сейсморазведки при проведении геологических исследований, в микроскопии при использовании объектива с переменным фокусным расстоянием, при компьютерном моделировании трехмерных объектов и сцен, а также другими путями.

Обработка и анализ трехмерных изображений играют в настоящее время существенную роль во многих областях исследований, особенно в медицине и геологии. В настоящей работе будут рассмотрены вопросы расширения двухмасштабной модели изображения [6, 7] и применения ее к 3D-изображениям, вопросы модификации операций частотной и пространственной фильтрации при переходе в 3D, вопросы сглаживания и декомпозиции изображения [6, 8], фильтрации помех, обнаружения контуров и объектов, вычислительные аспекты реализации некоторых алгоритмов в 3D-пространстве.

2. ОСОБЕННОСТИ ТРЕХМЕРНОГО ИЗОБРАЖЕНИЯ

В дискретном виде трехмерное изображение представляется массивом $\mathbf{X} = [x_{mnk}]$ размерами $M \times N \times K$. Как и в 2D-изображении, значение каждого элемента x_{mnk} есть квантованное на $(X_{\max} + 1)$ градаций значение логарифма яркости (энергии) $0 \leq x_{mnk} \leq X_{\max}$, которое для краткости будем называть просто *яркостью*. Трехмерное изображение, как средство отображения информации о некоторой наблюдаемой сцене или предмете, можно рассматривать состоящим из плотно упакованных связных трехмерных областей (объектов), соответствующих деталям сцены или предмета. *Областью* или *объектом* будем называть максимальное по размеру связное множество элементов изображения, имеющих близкие, возможно плавно меняющиеся значения яркости. Области могут соприкасаться произвольным образом, в том числе одна область может быть полностью окружена другой. На границах соседних областей значения яркости должны заметно различаться. Не соприкасающиеся области могут иметь произвольные, в том числе и совпадающие яркости. Пространственные границы между соседними областями, как объектами различающейся яркости, называют *контурами*.

Особенности трехмерных изображений:

- В трехмерном изображении объект (3D-объект) — пространственная фигура, а само 3D-изображение — пространственно ограниченная совокупность плотно упакованных 3D-объектов, заполняющих отображаемый участок пространства.
- Контуры в 3D-изображении суть пространственные границы между объектами.
- Сечение 3D-изображения плоскостью любого направления, а также проекция 3D-изображения на плоскость любого направления дают двумерный сигнал со всеми свойствами обычного 2D-изображения.

2.1. Локальные операции, окрестности и соседство элементов

Рассматриваемая ниже двухмасштабная модель изображения является по сути *локальной*, т.е. такой, которая для каждого произвольно взятого элемента изображения x_{mnk} описывает его взаимосвязи лишь с некоторым пространственно ограниченным множеством элементов $V_d(x_{mnk})$ или просто $V_d(x)$, которые его окружают и находятся от центрального элемента (в смысле некоторой выбранной метрики межэлементных расстояний) на дистанции, не пре-

вышающей d . Считая, что элемент (воксель) x_{mnk} является центральным, окружающее его множество элементов $V_d(x)$ при $d \leq 2 \div 3$ будем называть *окрестностью* и обозначать далее как V_{mnk} , а при $d \gg 1$ — *фрагментом* и обозначать через W_{mnk} . Отметим также, что в зависимости от выполняемых операций сам центральный элемент x_{mnk} может как принадлежать, так и не принадлежать $V_d(x)$. Соответственно, операции вида

$$y_{mnk} = f\{x_{ijl} \mid x_{ijl} \in V_d(x_{mnk})\}, \quad (1)$$

в которых результат в каждой точке (m, n, k) зависит лишь от значений элементов x_{ijl} , входящих в $V_d(x_{mnk})$, называются *локальными операциями*.

При переходе из 2D в 3D, варианты симметричных окрестностей и соседства элементов претерпевают следующие изменения. Окрестность из 2×2 элементов (4 пикселя) становится окрестностью из $2 \times 2 \times 2$ элементов (8 вокселей), в которой каждый воксель соседствует с каждым. В двумерной окрестности из 3×3 элементов (9 пикселей), как известно, можно рассматривать два варианта соседства элементов: 4-соседство (только по граням пикселей) и 8-соседство (по граням и вершинам пикселей) [12]. Аналогом первого из них в 3D будет окрестность с 6-соседством вокселей (Рис. 1,а). Аналогом второго — окрестность с 26-соседством вокселей (Рис. 1,в). Возможен также промежуточный вариант с 18-соседством вокселей (Рис. 1,б). Таким образом, в случае (Рис. 1,а) соседними с центральным вокселем (который на рисунках не виден) считаются только те воксели, которые имеют с ним общие грани; в случае (Рис. 1,б) — общие грани и ребра; а в случае (Рис. 1,в) — общие грани, ребра и вершины. Выбор варианта окрестности обычно определяется контекстом алгоритма.

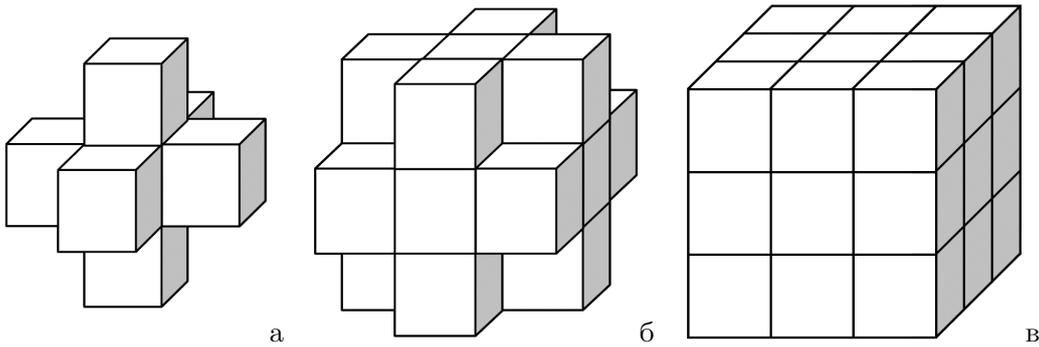


Рис. 1. Варианты окрестностей и соседства элементов (вокселей) в 3D: а) 6-соседство вокселей; б) 18-соседство вокселей; в) 26-соседство вокселей.

3. 3. ДВУХМАСШТАБНАЯ МОДЕЛЬ ТРЕХМЕРНОГО ИЗОБРАЖЕНИЯ

По аналогии с моделью обычного 2D-изображения [7], 3D-изображение также можно описать двухмасштабной моделью, представляющей статистические взаимосвязи элементов как на масштабе малого размера в несколько шагов дискретизации, так и на масштабе большого размера, соразмерного объектам изображения. Значения элементов 3D-изображения $X = [x_{mnk}]$ при этом представляются в виде суммы статистически независимых компонент:

$$x_{mnk} = S_{mnk} + t_{mnk} + \xi_{mnk}. \quad (2)$$

Первый член суммы — кусочно-гладкая компонента S_{mnk} , определяющая средние уровни яркости протяженных областей изображения; t_{mnk} — текстурно-детальная компонента, несущая информацию о текстуре и мелких деталях; ξ_{mnk} — шумовая компонента, определяемая шумами регистратора, аналого-цифрового преобразователя и др. Все компоненты предполагаются независимыми и аддитивными, а t_{mnk} и ξ_{mnk} — нормально распределенными и несмещенными.

3.1. Масштаб малого размера

На масштабе малого размера (масштабе элементов окрестности) рассматривается сравнительно небольшое связное множество элементов, расположенных на расстоянии нескольких шагов дискретизации. Как и в двумерной модели [7], элементы трехмерного изображения разделяются на два непересекающихся множества: попадающие на контурные участки (контурные) и не попадающие (фоновые), составляющие вместе полное изображение. Вводится понятие окрестности V_{mnk} элемента x_{mnk} как группы из R элементов $x_{mnk}^r \in V_{mnk}$, $r = 1, \dots, R$, ближайших к x_{mnk} , и попадающих в то же множество (контурное или фоновое), что и центральный элемент x_{mnk} .

Методом наименьших квадратов проводится гиперплоскость, наиболее близкая значениям элементов из V_{mnk} , составляющая с гиперплоскостью, ориентированной вдоль осей координат **MNK**, некоторый угол, величина и направление которого в точке (m, n, k) характеризуется вектором \mathbf{g}_{mnk} . В точке r окрестности проведенная гиперплоскость отличается от значения x_{mnk}^r на случайную величину γ_{mnk}^r . Такое представление позволяет связать значения элементов окрестности формулой:

$$x_{mnk}^r = \mu_{mnk} + \rho^r g_{mnk}^r + \gamma_{mnk}^r, \quad (3)$$

где μ_{mnk} — значение проведенной плоскости в центральной точке окрестности (m, n, k) , ρ^r — расстояние между центральным элементом x_{mnk} и x_{mnk}^r , g_{mnk}^r — величина проекции \mathbf{g}_{mnk} на вектор из x_{mnk} в x_{mnk}^r , а γ_{mnk}^r — случайная величина.

Вводится понятие контурной маски $\mathbf{E} = [e_{mnk}]$, совпадающей по размерам с изображением: $e_{mnk} = 1$ для контурных и $e_{mnk} = 0$ для фоновых элементов. Обозначая для контурных и фоновых элементов g_{mnk}^r через φ_{mnk}^r и ψ_{mnk}^r , а γ_{mnk}^r через ζ_{mnk}^r и η_{mnk}^r соответственно, представим g_{mnk}^r и γ_{mnk}^r в виде сумм $g_{mnk}^r = e_{mnk}^r \varphi_{mnk}^r + (1 - e_{mnk}^r) \psi_{mnk}^r$ и $\gamma_{mnk}^r = e_{mnk}^r \zeta_{mnk}^r + (1 - e_{mnk}^r) \eta_{mnk}^r$. В результате получим формулу модели трехмерной окрестности, описывающую статистические связи ее элементов:

$$x_{mnk}^r = \mu_{mnk} + e_{mnk}^r (\varphi_{mnk}^r \rho^r + \zeta_{mnk}^r) + (1 - e_{mnk}^r) (\psi_{mnk}^r \rho^r + \eta_{mnk}^r). \quad (4)$$

Здесь ζ_{mnk}^r — стохастическое возбуждение в точке r окрестности для контурных, а η_{mnk}^r — для фоновых элементов. Случайные величины φ_{mnk}^r , ψ_{mnk}^r , ζ_{mnk}^r , и η_{mnk}^r считаются некоррелированными и несмещенными, а шумовые составляющие ζ_{mnk}^r , и η_{mnk}^r — нормально распределенными. Эксперименты, проводившиеся на двумерных изображениях [6, 7], показывают, что значения дисперсий компонент φ и ψ различаются в 10-100 раз.

3.2. Масштаб большого размера

На масштабе большого размера (масштабе объектов фрагмента) предполагается, что гладкие составляющие S^v тех частей областей u^v ($v = 1, \dots, V$), которые попадают во фрагмент W_{mnk} , могут быть представлены полиномом степени не выше, чем ω . Тогда составляющая S_{ijl} , внутри фрагмента W_{mnk} может быть описана формулой:

$$S_{ijl}^v(W_{mnk}) = \sum_{v=1}^V \delta_{u^v} \sum_{p=0}^{\omega} \sum_{q=0}^{\omega-p} \sum_{r=0}^{\omega-p-q} a_{pqr}^v i^p j^q l^r; \quad (5)$$

здесь (i, j, l) — точка области u^v во фрагменте W_{mnk} ; $\delta_{u^v} = 1$, если точка $(i, j, l) \in u^v$ и $\delta_{u^v} = 0$ в остальных случаях. Добавляя в (5) текстурную t_{ijl} и шумовую ξ_{ijl} составляющие, получим выражение для значения элемента изображения внутри фрагмента:

$$x_{ijl}^v = \sum_{v=1}^V \delta_{u^v} \left(\sum_{p=0}^{\omega} \sum_{q=0}^{\omega-p} \sum_{r=0}^{\omega-p-q} a_{pqr}^v i^p j^q l^r + t_{ijl}^v + \xi_{ijl}^v \right). \quad (6)$$

Это общая формула модели, описывающей значения элементов областей внутри фрагмента. Значения t_{ijl}^v и ξ_{ijl}^v считаются нормально распределенными, но значения дисперсий t^v вообще говоря различаются от области к области.

На многих реальных изображениях области, в пределах типичного фрагмента анализа, имеют приблизительно постоянные средние локальные яркости, не меняющиеся заметно. По этой причине во многих случаях допустимо выбрать минимальную степень полинома: $\omega = 0$. Тогда $S_{ijl}^v(W_{mnk}) = S_{mnk}^v$ и (6) преобразуется к следующему виду:

$$x_{ijl}^v = \sum_{u=1}^V \delta_{uv} (S_{mnk}^v + t_{ijl}^v + \xi_{ijl}^v). \quad (7)$$

Это формула кусочно-постоянной модели фрагмента для представления участков областей изображения, попадающих во фрагмент W_{mnk} .

На масштабе большого размера обычно предполагается, что фрагмент W_{mnk} имеет форму прямоугольного параллелепипеда. Такой выбор обусловлен особенностями алгоритмической реализации методов обработки трехмерных изображений.

4. ИЗМЕНЕНИЕ АЛГОРИТМОВ ФИЛЬТРАЦИИ ПРИ ПЕРЕХОДЕ В 3D

Большинство алгоритмов фильтрации сравнительно просто модифицируются при переходе от одномерных к двумерным сигналам. Как правило, также несложно происходит их изменение при переходе и к трехмерным сигналам. Покажем это на примере наиболее распространенных алгоритмов, основанных на частотной и пространственной фильтрациях.

4.1. Алгоритмы частотной фильтрации

Одним из эффективных подходов в обработке сигналов является применение глобальных методов фильтрации, использующих ортогональные преобразования: Фурье, Волша-Адамара, Хаара, косинусное и пр. [10]. Наиболее распространен класс преобразований, обеспечивающих разложение сигналов по гармоническим функциям, из которых важнейшим является преобразование Фурье; на его примере и покажем модификацию преобразования.

При переходе от 2D- к 3D-изображениям трехмерными становятся как пространственная, так и частотная области представления данных, т.е. вместо двумерного получаем трехмерное Фурье-преобразование; все основные свойства преобразования Фурье при этом сохраняются.

Пусть $f(x, y, z)$ — непрерывная функция трех переменных x , y и z . Пара трехмерных непрерывных преобразований Фурье (прямое и обратное) задается следующими выражениями:

$$F(u, v, w) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y, z) e^{-i2\pi(xu+yv+zw)} dx dy dz \quad (8)$$

и

$$f(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v, w) e^{i2\pi(xu+yv+zw)} du dv dw, \quad (9)$$

где u , v и w — непрерывные частотные, а x , y и z — непрерывные пространственные переменные.

Трехмерное дискретное преобразование Фурье (ДПФ) может быть записано в виде:

$$F(u, v, w) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} f(m, n, k) e^{-i2\pi(um/M+vn/N+wk/K)} \quad (10)$$

где $f(m, n, k)$ — трехмерный цифровой массив размерами $M \times N \times K$. Обратное трехмерное дискретное преобразование Фурье будет иметь вид:

$$f(m, n, k) = \frac{1}{MNK} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \sum_{w=0}^{K-1} F(u, v, w) e^{i2\pi(um/M+vn/N+wk/K)} \quad (11)$$

для $0 \leq m < M$, $0 \leq n < N$ и $0 \leq k < K$ и $0 \leq u < M$, $0 \leq v < N$ и $0 \leq w < K$.

Как и в двумерном случае, частотная фильтрация трехмерного сигнала производится выполнением прямого преобразования Фурье (10), модификацией полученного спектра $F(u, v, w)$, например умножением его на амплитуду спектра (трехмерного) того или иного фильтра, и последующим обратным преобразованием (11). По аналогии с преобразованием Фурье, другие ортогональные преобразования при переходе к трехмерному сигналу также несложно видоизменяются добавлением размерности.

4.2. Алгоритмы пространственной фильтрации

Пространственная фильтрация обычно осуществляется локальными операторами согласно формуле (1) с определенными ограничениями на размеры локальной области анализа $V_d(x)$. Это обусловлено тем, что в случае функции $f(.,.)$ в (1) общего вида, когда не существует рекуррентных соотношений для определения следующих значений на основе ранее сделанных, объем необходимых вычислений возрастает как минимум пропорционально числу элементов, попадающих в $V_d(x)$. Окрестности, которые в 2D случае содержат порядка $(2d)^2$ пикселей, в 3D будут содержать уже порядка $(2d)^3$ вокселей. Тем не менее для некоторых случаев, например, когда в основе фильтрации лежит вычисление арифметического среднего или порядковых статистик по $V_d(x)$, в трехмерном случае, как и в двумерном, удается построить быстрые алгоритмы.

Алгоритмы, использующие оценку среднего по фрагменту

Ряд алгоритмов обработки двумерных изображений, связанных со сглаживанием, выделением низко- или высокочастотных составляющих, улучшением изображения, обобщаются следующей формулой [6, 9]:

$$y_{mn} = f(x_{mn} - S_{mn}, D_{mn}) + bS_{mn} + c, \quad (12)$$

где x_{mn} — значение центрального элемента, S_{mn} — оценка среднего по фрагменту (например, значение арифметического среднего или медианы), D_{mn} — оценка дисперсии по фрагменту, $f(u, v)$ — зависимость усиления контраста, b и c — параметры преобразования. Так, при $f(u, v) = 0$ и $c = 0$ получаем выделение низкочастотной составляющей (сглаживание), при $f(u, v) = u$, $b = 0$ и $c = 0,5$ — выделение высокочастотной составляющей, при $f(u, v) > u$, $b \leq 1$ и $c \approx (1 - b)/2$ — улучшение изображения [6, 9].

Очевидно, преобразования вида (12) легко переносятся на 3D-изображения добавлением размерности, но с учетом того, что S_{mnk} , а также u и v в $f(u, v)$ должны определяться уже не по двумерному, а по трехмерному фрагменту (прямоугольному параллелепипеду) с центром в точке (m, n, k) :

$$y_{mnk} = f(x_{mnk} - S_{mnk}, D_{mnk}) + bS_{mnk} + c. \quad (13)$$

Модификации алгоритмов быстрого вычисления арифметического среднего или медианы S_{mnk} , а также дисперсии D_{mnk} по прямоугольному параллелепипеду рассмотрены в конце данной статьи.

Модификация операторов контурных перепадов

В основе большинства алгоритмов обнаружения контуров на изображении так или иначе лежат операторы контурных перепадов, в которых, как правило, используется оценка значений либо первой, либо второй производной [11, 12]. В случае 2D-изображения контуры объектов суть линии, разделяющие плоские области, и в идеальном случае (когда контурные линии не имеют разрывов) получаем обычную плоскую карту изображаемой сцены.

В 3D-пространстве контурное изображение представляет собой множество разделяющих объекты поверхностей произвольной формы, и здесь возможно возникновение самых различных конфигураций контуров, например, становятся допустимыми самопересечения и узлы. Сечение 3D-контурного изображения плоскостью представляет собой двумерную карту контуров. Однако такая карта вообще говоря не обязана совпадать с картой, получаемой проведением контуров по 2D изображению, являющемуся сечением 3D изображения той же плоскостью.

Следует также отметить, что, как показано в [16], обнаружение контуров не по самому изображению, а по выделенной путем декомпозиции сглаженной компоненте S_{mnk} (см. раздел 4.2.4), позволяет достичь более высокой слитности контурных линий, меньшей их толщины, а также существенного снижения количества ложных контуров и помех.

Рассмотрим модификации базовых операторов контурных перепадов на основе первой и второй производных при переходе от 2D- к 3D-изображениям, которые могут быть построены в рамках трехмерных окрестностей, рассмотренных выше.

На основе первой производной: оператор Робертса

Оператор Робертса для двумерного изображения описывается следующей формулой [13]:

$$y_{m,n} = \{|x_{m,n} - x_{m+1,n+1}| + |x_{m+1,n} - x_{m,n+1}|\}/2, \quad (14)$$

т.е. как сумма модулей разностей значений диагональных элементов по квадрату 2×2 пикселей. Для 3D-изображения оператор Робертса будет представлять собой сумму модулей разностей значений элементов в диагоналях куба из $2 \times 2 \times 2$ пикселей:

$$y_{mnk} = \{|x_{mnk} - x_{m+1,n+1,k+1}| + |x_{m+1,n,k} - x_{m,n+1,k+1}| + |x_{m,n+1,k} - x_{m+1,n,k+1}| + |x_{m,n,k+1} - x_{m+1,n+1,k}|\}/4. \quad (15)$$

На основе первой производной: оператор Собела

Оператор Собела для двумерного изображения описывается формулой [12]:

$$y_{m,n} = \{|x_{m-1,n-1} + 2x_{m-1,n} + x_{m-1,n+1} - x_{m+1,n-1} - 2x_{m+1,n} - x_{m+1,n+1}| + |x_{m-1,n-1} + 2x_{m,n-1} + x_{m+1,n-1} - x_{m+1,n-1} - 2x_{m+1,n} - x_{m+1,n+1}|\}/8. \quad (16)$$

Чтобы задать оператор Собела (и аналогичные ему) для 3D-изображения, первоначально определим частные отклики как модули разностей значений элементов для каждого из направлений m , n и k , что соответствует выражению под одним из знаков модуля в (16). Для направления m такую зависимость $y_m(m, n, k)$ с точностью до выбора коэффициентов a , b и c можно выразить следующей формулой:

$$y_m(m, n, k) = |ax_{m-1,n-1,k-1} + bx_{m-1,n-1,k} + ax_{m-1,n-1,k+1} + bx_{m-1,n,k-1} + cx_{m-1,n,k} + bx_{m-1,n,k+1} + ax_{m-1,n+1,k-1} + bx_{m-1,n+1,k} + ax_{m-1,n+1,k+1} - ax_{m+1,n-1,k-1} - bx_{m+1,n-1,k} - ax_{m+1,n-1,k+1} - bx_{m+1,n,k-1} - cx_{m+1,n,k} - bx_{m+1,n,k+1} - ax_{m+1,n+1,k-1} - bx_{m+1,n+1,k} - ax_{m+1,n+1,k+1}|/(4a + 4b + c). \quad (17)$$

Аналогичным образом записываются модули разностей $y_n(m, n, k)$ и $y_k(m, n, k)$ для составляющих n и k . Окончательно трехмерный оператор, аналогичный оператору Собела, выразим через корень из суммы квадратов по трем направлениям:

$$y_{mnk} = \{[y_m(m, n, k)]^2 + [y_n(m, n, k)]^2 + [y_k(m, n, k)]^2\}^{1/2}. \quad (18)$$

Коэффициенты a , b и c рекомендуется выбирать вблизи значений 1, 2 и 3 соответственно.

На основе второй производной: оператор Лапласа

Оператор Лапласа для двумерного изображения описывается следующей формулой [12]:

$$y_{mn} = \{x_{m-1,n-1} + x_{m-1,n} + x_{m-1,n+1} + x_{m,n-1} + x_{m,n+1} + x_{m+1,n-1} + x_{m+1,n} + x_{m+1,n+1}\}/8 - x_{mn}. \quad (19)$$

Модификацию оператора Лапласа для 3D-изображения можно выразить следующей формулой:

$$y_{mnk} = \left(\frac{1}{Q_V} \sum_{x_{ijl} \in V_{mnk}} x_{ijl} \right) - x_{mnk}, \quad (20)$$

где V_{mnk} — окрестность точки x_{mnk} , не включающая саму центральную точку x_{mnk} , Q_V — число точек в такой окрестности, $\{x_{ijl}\}$ — набор элементов окрестности V_{mnk} . В качестве V_{mnk} может выступать одна из окрестностей, показанных на Рис. 1; для окрестности (а) $Q_V = 6$, для окрестности (б) $Q_V = 18$, для окрестности (в) $Q_V = 26$.

Другие контурные операторы

Как и для двумерного случая, возможно построение и других детекторов, позволяющих обнаруживать контурные перепады [14], например, разности максимума и минимума значений элементов окрестности V_{mnk} :

$$y_{mnk} = \max\{x_{ijl} \mid x_{ijl} \in V_{mnk}\} - \min\{x_{ijl} \mid x_{ijl} \in V_{mnk}\}.$$

Фильтрация импульсных помех

Импульсными помехами называют сравнительно редкие искажения отдельных элементов изображения, когда значения помехи значительно отличаются от истинных значений сигнала и не коррелированы с ними. Вопрос их устранения важен с позиции методологии разработки алгоритмов обнаружения шумовых выбросов со статистикой, резко отличающейся от статистики сигнала. На изображении импульсные помехи выглядят как хаотически расположенные точки, хорошо различимые визуально и, как правило, случайные по яркости. Различие статистических свойств помех и изображений позволяет разработать методы фильтрации, использующие межэлементные связи последних [15].

Модель искажения изображения импульсными помехами несложна. Значение каждого из элементов x_{mnk} изображения с вероятностью p заменяется на случайное значение ξ_{mnk} независимо от значений остальных элементов. Обозначим через $\mathbf{X}' = [x'_{mnk}]$ исходное неискаженное изображение, через $\mathbf{X} = [x_{mnk}]$ — искаженное импульсной помехой, а через $\mathbf{Y} = [y_{mnk}]$ — результат фильтрации. Процесс искажения представится в виде:

$$x_{mnk} = \begin{cases} x'_{mnk}, & \text{с вероятностью } (1 - p) \text{ для неискаженного элемента;} \\ \xi_{mnk}, & \text{с вероятностью } p \text{ для элемента, искаженного помехой.} \end{cases} \quad (21)$$

Без существенного ограничения общности можно полагать, что значения импульсных помех ξ_{mnk} распределены равномерно на всем диапазоне яркостей $[0, X_{max}]$. Задача фильтрации импульсных помех состоит в обнаружении помех и последующей коррекции искаженных отсчетов яркости.

Наиболее распространенные алгоритмы фильтрации основываются на предсказании значения элемента \tilde{x}_{mnk} по окружающей его окрестности V_{mnk} . При этом используются локальные корреляционные связи близлежащих элементов изображения и учитывается, что шум пространственно декоррелирован. При обнаружении сравниваются наблюдаемое x_{mnk} и предсказываемое \tilde{x}_{mnk} значения. Если они отличаются более, чем на величину некоторого заданного порога обнаружения δ , считается, что x_{mnk} — помеха и осуществляется ее исправление на значение \tilde{x}_{mnk} . Общая формула вычисления \tilde{x}_{mnk} для большинства методов при этом практически совпадает с формулой (1):

$$\tilde{x}_{mnk} = f\{x_{ijl} \mid x_{ijl} \in V_{mnk}\}. \quad (22)$$

В качестве функции $f\{\cdot\}$ в (22) могут выступать различные линейные или нелинейные операторы, которые могут являться как каузальными (т.е. зависящими от значений уже обработанных элементов), так и некаузальными [15]. Обобщенная формула фильтрации для большинства алгоритмов при этом выглядит следующим образом:

$$y_{mnk} = \begin{cases} x_{mnk}, & \text{если } |x_{mnk} - \tilde{x}_{mnk}| < \delta; \\ \tilde{x}_{mnk}, & \text{если } |x_{mnk} - \tilde{x}_{mnk}| \geq \delta. \end{cases} \quad (23)$$

Как было показано в [15], наилучшие результаты фильтрации импульсных помех достигаются алгоритмами, основанными на использовании в качестве функции $f\{\cdot\}$ в (22) значения медианы по симметричной окрестности обрабатываемого элемента: $f(V_{mnk}) = \text{med}\{x_{ijl} \in V_{mnk}\}$. При больших p , когда вероятность искажения пары соседних вокселей достаточно высока, рекомендуется итеративный процесс фильтрации: первоначально с большим значением порога δ , затем с постепенным его уменьшением. Можно также рекомендовать для первой итерации применять окрестность, показанную на Рис. 1,а для последующих — окрестности на Рис. 1,б и в. Наибольшее сглаживание достигается при использовании последней окрестности.

Декомпозиция изображения

Вопросы декомпозиции, которая в терминах модели (2) означает разделение изображения на сглаженную S и текстурно-детальную компоненты $(t + \xi)$, были рассмотрены в работах [6, 8, 9, 16]. Согласно изложенному в них алгоритму, для каждой точки изображения (m, n) производится последовательный локальный анализ изображения сначала по окрестности V_{mn} , а затем по фрагменту W_{mn} , окружающим центральный элемент x_{mn} . При этом по сути анализируются лишь распределения значений элементов по окрестности V_{mn} и фрагменту W_{mn} . Упрощенный вариант данного алгоритма декомпозиции (сглаживания) изображения впоследствии был опубликован в [17] под названием *билатеральная фильтрация*.

Такой подход позволяет легко перейти от 2D- к 3D-изображению — достаточно лишь вместо двумерных окрестности V_{mn} и фрагмента W_{mn} подставить в формулы трехмерные окрестность V_{mnk} и фрагмент W_{mnk} . Тогда, по аналогии с [16], алгоритм декомпозиции для трехмерного изображения формулируется следующим образом. При заданных размерах $l \times l \times l$ окрестности V_{mnk} и $L \times L \times L$ фрагмента W_{mnk} ($l < L$), центрированных в точке (m, n, k) , ширине яркостных интервалов анализа Δ^V и Δ^W , а также ранговых параметрах $n^V < l^3/2$ и $n^W < L^3/2$ соответственно, значение сглаженной компоненты S_{mnk} в (2) находится при помощи следующих операций, выполняемых для каждой точки (m, n, k) изображения:

1. Подсчитываются нормированные гистограммы (вероятности распределения) по трехмерным окрестности $H_{mnk}^V = \{h_{mnk}^V(i)\}$ и фрагменту $H_{mnk}^W = \{h_{mnk}^W(i)\}$ с центром в точке (m, n, k) .
2. По гистограмме окрестности H_{mnk}^V и заданному значению n^V находятся ранговые параметры $R_1^V = R^V(n^V/l^3)$ и $R_2^V = R^V(1 - n^V/l^3)$; здесь $R(x)$ определяется как решение уравнения $\sum_{i=0}^{R(x)} h_{mnk}^V(i) = x$, где $h_{mnk}^V(i)$ — гистограмма значений элементов в окрестности V_{mnk} . Сравнением значения центрального элемента x_{mnk} с R_1^V и R_2^V определяется промежуточное усеченное значение \tilde{x}^V :

$$\tilde{x}^V = \begin{cases} x_{mnk}, & \text{если } R_1^V \leq x_{mnk} \leq R_2^V; \\ R_1^V, & \text{если } x_{mnk} < R_1^V; \\ R_2^V, & \text{если } x_{mnk} > R_2^V. \end{cases}$$

3. С помощью сигма-фильтра [18], являющегося частным случаем парзеновского окна [19], находится урезанное среднее \bar{x}_{mn} . Для этого из элементов окрестности V_{mnk} выбираются те z значений $x_{mnk}^r \in V_{mn}$ ($r = 1, \dots, z$), которые попадают в интервал $(\tilde{x}^V - \Delta^V, \tilde{x}^V + \Delta^V)$, где Δ^V — полуширина интервала. По значениям x_{mnk}^r , попадающим в данный интервал, подсчитывается среднее:

$$\bar{x}_{mnk} = A(V_{mnk}, x_{mnk}, n^V, \Delta^V) = \frac{1}{z} \sum_{r=1}^z x_{mnk}^r, \quad \tilde{x}^V - \Delta^V \leq x_{mnk}^r \leq \tilde{x}^V + \Delta^V. \quad (24)$$

4. Аналогично п. 2, по гистограмме фрагмента H_{mnk}^W и заданному n^W находятся ранговые значения $R_1^W = R^W(n^W/L^3)$ и $R_2^W = R^W(1 - n^W/L^3)$. Сравнением значения \bar{x}_{mnk} с R_1^W и R_2^W определяется усеченное значение \tilde{x}^W :

$$\tilde{x}^W = \begin{cases} \bar{x}_{mnk}, & \text{если } R_1^W \leq \bar{x}_{mnk} \leq R_2^W; \\ R_1^W, & \text{если } \bar{x}_{mnk} < R_1^W; \\ R_2^W, & \text{если } \bar{x}_{mnk} > R_2^W. \end{cases}$$

5. Сглаженное значение S_{mnk} находится по гистограмме фрагмента H_{mnk}^W как медиана значений H_{mnk}^W , попадающих в интервал $(\tilde{x}^W - \Delta^W, \tilde{x}^W + \Delta^W)$:

$$S_{mnk} = \text{med}(W_{mnk}, \bar{x}_{mnk}, n^W, \Delta^W). \quad (25)$$

Полученное значение S_{mnk} считается искомой сглаженной компонентой.

Обнаружение объектов заданного объема

В [16] показано, что изложенный выше алгоритм декомпозиции можно использовать для обнаружения объектов на изображении. Аналогично 2D-изображениям, для которых решается задача обнаружения объектов по их площади, в трехмерной модификации ставится задача обнаружения объектов по их объему. Данная задача также формулируется в трех вариантах: обнаружение объектов с объемом (т.е. числом элементов) Q^j больше заданного T , меньше заданного T , и обнаружение объектов, имеющих объем в интервале $T_1 < Q^j < T_2$. Алгоритм обнаружения для 3D-изображений принципиально не отличается от алгоритма для 2D, который изложен в [16] (достаточно лишь в обозначения координат добавить третье измерение, а плоский прямоугольный фрагмент W_{mn} заменить на прямоугольный параллелепипед W_{mnk}), поэтому здесь не приводится. Как понятие площади в двумерном случае, так и понятие объема в трехмерном варианте алгоритма используется несколько в необычном смысле — как „локальный“ объем, т.е. объем той части объекта, которая попадает внутрь фрагмента W_{mnk} .

5. НЕКОТОРЫЕ ВЫЧИСЛИТЕЛЬНЫЕ АЛГОРИТМЫ

5.1. Вычисление суммы по прямоугольному параллелепипеду

Введем следующее обозначение для суммы по фрагменту 2D-изображения:

$$S_{(ij)(mn)} = \sum_{u=i}^{m-1} \sum_{v=j}^{n-1} x_{uv},$$

т.е. $S_{(ij)(mn)}$ — сумма значений элементов x_{uv} , попадающих в прямоугольный фрагмент, диагональные точки которого имеют координаты (i, j) и $(m-1, n-1)$. Обратим внимание, что фрагмент не включает точку с координатами (m, n) и соответствующие ей строку и столбец.

Аналогично для трехмерного изображения, $S_{(ijl)(mnk)}$ — сумма значений элементов x_{ijl} в прямоугольном параллелепипеде с угловыми координатами (i, j, l) и $(m-1, n-1, k-1)$:

$$S_{(ijl)(mnk)} = \sum_{u=i}^{m-1} \sum_{v=j}^{n-1} \sum_{w=l}^{k-1} x_{uvw}.$$

Для двумерного изображения классический способ вычисления суммы $S_{(mn)(m+H, n+L)}$ по скользящему прямоугольному фрагменту размерами $H \times L$ элементов при переходе от элемента (m, n) к элементу $(m, n+1)$ сводится к формуле

$$S_{(m, n+1)(m+H, n+L+1)} = S_{(m, n)(m+H, n+L)} - S_{(m, n)(m+H, n+1)} + S_{(m, n+L)(m+H, n+L+1)}, \quad (26)$$

где последние два члена — суммы по левому (удаляемому) и правому (добавляемому) столбцам элементов. Такой алгоритм требует 4 арифметических операции независимо от размеров фрагмента — 2 операции в выражении (26) и две операции на пересчет каждой из сумм по столбцу $S_{(m, n)(m+H, n+1)}$ при переходе от строки m к строке $m+1$. Дополнительно требуется N ячеек для хранения массива сумм по столбцам.

При переходе к трехмерному изображению, формула (26) для скользящего прямоугольного параллелепипеда размерами $H \times L \times J$ будет модифицирована следующим образом:

$$S_{(m, n, k+1)(m+H, n+L, k+J+1)} = S_{(m, n, k)(m+H, n+L, k+J)} - S_{(m, n, k)(m+H, n+L, k+1)} + S_{(m, n, k+J)(m+H, n+L, k+J+1)}, \quad (27)$$

где последние два члена — суммы по левой (удаляемой) и правой (добавляемой) граням элементов. Данный алгоритм требует уже 6 арифметических операции независимо от размеров фрагмента — 2 операции в выражении (27), две операции на пересчет каждой из сумм по граням и две на пересчет сумм по столбцам. Кроме того, требуется J ячеек для хранения массива сумм по граням и $J \times L$ ячеек для хранения сумм по столбцам.

Для случая двумерного изображения известен также другой алгоритм вычисления суммы по прямоугольнику произвольного размера. Пусть для каждой точки (m, n) подсчитаны суммы S_{mn} по прямоугольнику с диагональными элементами x_{00} и $x_{m-1, n-1}$, т.е. $S_{mn} = S_{(0,0)(m, n)}$. Тогда сумма $S_{(ij)(mn)}$ значений элементов внутри прямоугольника с угловыми координатами (i, j) и $(m-1, n-1)$ вычисляется следующим образом:

$$S_{(ij)(mn)} = S_{mn} - S_{m, j} - S_{i, n} + S_{ij}, \quad (28)$$

что в среднем для каждого элемента изображения требует 2 операции для вычисления суммы S_{mn} и 3 операции для вычислений по формуле (28). Однако для хранения сумм S_{mn} требуется уже $M \times N$ ячеек, т.е. столько же, каков размер изображения.

Алгоритм (28) также может быть распространен на трехмерное изображение. Пусть для каждой точки (m, n, k) подсчитаны суммы S_{mnk} по прямоугольному параллелепипеду с диагональными элементами x_{000} и $x_{m-1, n-1, k-1}$, т.е. $S_{mnk} = S_{(000)(mnk)}$. Нетрудно показать, что в таком случае сумма $S_{(ijl)(mnk)}$ значений элементов внутри прямоугольного параллелепипеда с угловыми координатами (i, j, l) и $(m-1, n-1, k-1)$ вычисляется при помощи следующей операции:

$$S_{(ijl)(mnk)} = S_{mnk} - S_{mjk} - S_{mnl} - S_{ink} + S_{mjl} + S_{ijk} + S_{inl} - S_{ijl}. \quad (29)$$

Таким образом с учетом того, что для вычисления каждого из значений S_{mnk} требуется 3 арифметических операции, для вычисления суммы $S_{(ijl)(mnk)}$ для каждого элемента трехмерного изображения потребуется в среднем 10 арифметических операций. Объем дополнительной требуемой памяти составит $M \times N \times K$ ячеек с разрядностью достаточной для хранения значений сумм S_{mnk} .

Аналогичным образом можно находить дисперсии по фрагменту $D_{(ijl)(mnk)}$, вычисляя значения сумм квадратов $S_{(mnk)}(x^2)$ для фрагментов с углом в каждой из точек (m, n, k) , а также по прямоугольному параллелепипеду $S_{(ijl)(mnk)}(x^2)$, и пользуясь затем формулой

$$D_{(ijl)(mnk)} = \{S_{(ijl)(mnk)}(x^2) - (S_{(ijl)(mnk)})^2\} / N_{(ijl)(mnk)}, \quad (30)$$

где $S_{(ijl)(mnk)}(x^2)$ — сумма квадратов значений элементов, попадающих в параллелепипед, а $N_{(ijl)(mnk)} = (m-i) \times (n-j) \times (k-l)$ — число точек в параллелепипеде.

5.2. Вычисление порядковых статистик по прямоугольному параллелепипеду

Основой для вычисления порядковых статистик по фрагменту как 2D- так и 3D-изображения служит гистограмма распределения значений яркости по этому фрагменту $h^W(x)$, а также ее интегральная характеристика $F^W(x)$:

$$F^W(x) = \sum_{i=0}^x h^W(i); \quad F^W(X_{\max}) = N^W, \quad (31)$$

где X_{\max} — максимально возможное значение яркости, а N^W — число точек во фрагменте W . Порядковые статистики вида $R^W(n)$, где $0 \leq n \leq N^W$, представляют собой зависимость:

$$R^W(n) = z, \quad \text{если} \quad F^W(z-1) < n \leq F^W(z). \quad (32)$$

Алгоритм скользящего вычисления гистограммы по фрагменту строится аналогично формулам (26) и (27), т.е. при смещении фрагмента к следующей точке производится удаление точек на одной грани фрагмента и добавление точек на противоположной грани [20]. В двумерном случае алгоритм скользящего вычисления гистограммы по фрагменту при переходе от точки (m, n) к соседней точке $(m, n+1)$ требует в среднем $2H$ числа операций (H — число строк во фрагменте). В трехмерном случае при переходе от точки (m, n, k) к точке $(m, n, k+1)$ потребуется уже $2H \times L$ операций, где $H \times L$ — число точек в грани параллелепипеда, перпендикулярной направлению смещения K .

Как видно, в трехмерном случае число требуемых операций растет пропорционально произведению $H \times L$. Однако, если $(H \times L) > (X_{\max} + 1)$, вместо операций со значениями отдельных точек становится выгодным сформировать предварительные гистограммы $h_{(ijk)(m, n, k+1)}^F(x)$ для соответствующих граней параллелепипеда и осуществлять уже операции вычитания и прибавления таких гистограмм:

$$h_{(i, j, k+1)(m, n, k+J+1)}^W(x) = h_{(ijk)(m, n, k+J)}^W(x) - h_{(ijk)(m, n, k+1)}^F(x) + h_{(ijk+J)(m, n, k+J+1)}^F(x), \quad (33)$$

где J — размер фрагмента в направлении смещения K . Выполнение действий по формуле (33) требует в среднем $2(X_{\max} + 1)$ арифметических операций на одну точку изображения независимо от размера параллелепипеда. Для пересчета гистограмм по граням параллелепипеда $h_{(ijk)(m,n,k+1)}^F(x)$ нужно дополнительно в среднем $2L$ операций на точку и $(X_{\max} + 1) \times K$ ячеек памяти для хранения K гистограмм.

Число операций $(X_{\max} + 1)$, требуемое для прибавления/вычитания каждой из гистограмм по граням параллелепипеда $h_{(ijk)(m,n,k+1)}^F(x)$, можно значительно уменьшить, если воспользоваться свойством пространственной корреляции обрабатываемых данных. На участках с медленным изменением яркости, которых на реальных изображениях обычно большинство, размах значений элементов сравнительно невелик — в несколько раз меньше полного диапазона в $(X_{\max} + 1)$ градаций. Добавив к каждой из гистограмм $h_{(ijk)(m,n,k+1)}^F(x)$ по 2 ячейки для запоминания минимального и максимального значений распределения, и, соответственно, обрабатывая лишь указываемый диапазон градаций, удастся дополнительно в несколько раз сократить общее число операций.

СПИСОК ЛИТЕРАТУРЫ

1. Lohmann G. Volumetric Image Analysis, Teubner, Chichester: Wiley, 1998.
2. Nikolaidis N., Pitas I. 3-D Image Processing Algorithms. NY.: Wiley, 2000.
3. Cyganek B. An Introduction to 3D Computer Vision Techniques and Algorithms. NY.: Wiley, 2009.
4. Toriwaki J., Yoshida H. Fundamentals of Three-Dimensional Digital Image Processing. NY.: Springer, 2009.
5. Красильников Н. Цифровая обработка 2D- и 3D-изображений. С.-Пб.: БХВ-Петербург, 2011.
6. Чочиа П.А. Обработка и анализ изображений на основе двухмасштабной модели: Препринт ИППИ АН СССР. М.: ВИНТИ, 1986.
7. Чочиа П.А. Двухмасштабная модель изображения // Кодирование и обработка изображений. М.: Наука, 1988, С. 69–87.
8. Чочиа П.А. Сглаживание изображения при сохранении контуров // Кодирование и обработка изображений. М.: Наука, 1988, С. 87–98.
9. Chochia P.A. Image Enhancement Using Sliding Histograms // Computer Vision, Graphics, Image Processing, 1988, vol. 44, no. 2, pp. 211–229.
10. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. М.: Связь, 1980.
11. Прэтт У. Цифровая обработка изображений. М.: Мир, 1982. Т. 1, 2.
12. Гонсалес Р., Вудс Р. Цифровая обработка изображений, М.: Техносфера, 2012.
13. Робертс Л. Автоматическое восприятие трехмерных объектов. // Интегральные роботы. М.: Мир, 1973, С. 162–208.
14. Monga O., Deriche R., Rocchisani J. 3D edge detection using recursive filtering: application to scanner images // CVGIP: Image Understanding, 1991, vol. 53, no. 1, pp. 76–87.
15. Чочиа П.А. Цифровая фильтрация импульсных помех на телевизионных изображениях // Техника средств связи: сер. Техника телевидения, 1984, вып. 1, С.26–36.
16. Чочиа П.А. Некоторые алгоритмы обнаружения объектов на основе двухмасштабной модели изображения. // Информационные процессы, 2014, Т. 14, № 2, С. 117–136.
17. Tomasi C., Manduchi R. Bilateral filtering for gray and color images // Proc. IEEE 6th Int. Conf. on Computer Vision, Bombay, India, 1998, pp. 839–846

18. Lee J.-S. Digital Image Smoothing and the Sigma Filter // Computer Vision, Graphics, Image Processing, 1983, vol. 24, no. 2. pp. 255–269.
19. Parzen E. On the estimation of a probability density function and mode // Annals of Mathematical Statistics, 1962, vol. 33, pp. 1065–1076.
20. Чочиа П.А. Параллельный алгоритм вычисления скользящей гистограммы // Автометрия, 1990, № 2, С. 40–44.

Transition from 2D- to 3D-images: modification of the two-scale image model and image processing algorithms

Chochia P. A.

The specificities of three dimensional images are formulated. The adaptation of the two-scale image model to 3D images is studied. The variants for modification of various image processing algorithms to 3D images are demonstrated.

KEYWORDS: image processing, 3D image, image model.