

ВНУТРИКАДРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ В РЕАЛЬНОМ ВРЕМЕНИ С ПОМОЩЬЮ ОДНОРОДНЫХ СИСТОЛИЧЕСКИХ АРХИТЕКТУР

Н.А. Лукин

Институт машиноведения УрО РАН
Россия, 620049. Екатеринбург, ул. Комсомольская, 34
E-mail: nicklookin@mail.ru

Ключевые слова: внутрикадровая обработка изображений, систолические архитектуры, побитовая обработка данных, функционально-ориентированные процессоры.

Аннотация: Цифровая обработка изображений в реальном времени в составе мобильных систем требует решения проблем создания высокопроизводительных специализированных процессоров для внутрикадровой обработки изображений. Проводится анализ графов алгоритмов внутрикадровой обработки и оцениваются их параметры, прежде всего степень параллелизма. Описываются архитектуры функционально-ориентированных процессоров на основе однородных вычислительных сред, предназначенные для сверхбыстрой внутрикадровой обработки. Излагаются результаты исследований и разработки микроэлектронной реализации таких процессоров.

1. Введение

Цифровая обработка изображений является одной из наиболее актуальных задач в системах реального времени. Она может требовать значительных аппаратных затрат, что представляет собой научно-техническую проблему, особенно, для мобильных систем. Требование реального времени в ряде применений, например, при решении задач дистанционного зондирования Земли, может сделать неприемлемым использование процессорных платформ общего назначения [1], при этом затруднительным может оказаться применение даже таких продвинутых решений, как мультипроцессорные архитектуры графических процессоров [2]. Такая ситуация характерна для систем с ограничением на энергопотребление, например, на борту малогабаритного космического аппарата.

Еще одна проблема создания эффективных средств обработки данных в видеосистемах связана с развитием оптических сенсоров, которое происходит по двум основным направлениям: а). рост частоты смены кадров, что требует изменения интерфейса между сенсором и системой обработки данных; б). увеличение числа выходных каналов съема пикселей из матрицы чувствительных элементов, что затрудняет использование современной технологии мультиплексирования каналов потоков входных пикселей.

Становится очевидным, что совершенствование датчиков может потребовать существенного изменения архитектуры видеопроцессоров для перспективных встроенных систем обработки изображений.

Таким образом, миниатюризация мобильных систем обработки изображений и развитие микроэлектронных технологий изготовления оптических сенсоров приводит к

необходимости совершенствования архитектур и аппаратной реализации видеопроцессоров.

2. Особенности алгоритмов цифровой обработки изображений в системах реального времени

Цифровая обработка изображений в составе системы реального времени, как правило, представляет собой двухэтапную обработку данных.

Этап 1. Предварительная (первичная, внутрикадровая) обработка, назначением которой является повышение достоверности изображения. На этом этапе обработки реализуются алгоритмы коррекции геометрических искажений, повышения контрастности, усреднения, сглаживания, фильтрации, сегментации [3]. Особенностью данных алгоритмов является значительное количество независимых потоков входных данных (до десятков млн., т.е. число пикселей) и высокая степень внутреннего параллелизма вычислений. В настоящее время на рынке видеосистем отсутствуют как вычислительные средства, которые могут обеспечить ввод и обработку такого числа независимых потоков данных, так и сенсоры, которые могут выдавать все пиксели одновременно.

Этап 2. Вторичная (межкадровая) обработка, назначением которой является идентификация параметров внешних объектов и, в конечном счете, их распознавание. Спецификой подобных алгоритмов является существенно более низкая степень параллелизма, значительное число итераций по данным. В отличие от потокового характера обработки данных, специфичного для внутрикадровой обработки изображений, для межкадровой обработки характерны значительно меньшая скорость вычислений и их последовательный (итерационный) характер [4].

Информационные характеристики алгоритмов обработки данных, важные с точки зрения реализации в вычислительных системах реального времени отражаются в структуре их информационных графов G [5]. Для алгоритмических графов основными параметрами являются следующие:

а) Степень параллелизма (DOP – Degree of Parallelism) [6]. Это число операций или процедур, которые могут выполняться одновременно. Как правило, для корректной оценки производят оценку текущего параметра $DOP(h)$, где h – номер яруса. Можно показать, что $\forall G, \exists m, H \in \mathbb{N}; DOP(\square_m) = \text{Max}(DOP(\square)), \square = 0, 1, \dots, H$. Таким образом, $DOP(h_m)$ есть максимально возможный параллелизм обработки данных для конкретного алгоритма.

б) Ширина (диаметр) ЯПФ графа [5]. Шириной D графа G называется максимальная ширина из всех ярусов: $D = DOP(\square_m)$. Для обеспечения параллелизма степени $DOP(h_m)$ в многопроцессорной системе необходимо выполнение условия: $\text{Min}(k)\{M_k\} \geq D, k \in \mathbb{N}$ где M_k – количество процессоров в k -ом варианте их назначения на обработку данных. Ширина графа D может использоваться в качестве меры аппаратной сложности вычислений (L_h) для алгоритма или классов алгоритмов [7].

в) Высота ЯПФ графа [6]. Высотой графа H называется количество ярусов в ЯПФ. Параметр H может использоваться в качестве меры временной сложности вычислений (L_t) для алгоритма или класса алгоритмов [7].

Главной особенностью практически любого алгоритма внутрикадровой обработки изображений является то, что $D = N1 \cdot N2$, где $N1, N2$ – число пикселей светочувствительной матрицы сенсора по координатам i и j . Для стандартного КМОП-сенсора с матрицей размерностью 4М пикселей ($N1 = N2 = 2K$), $D = 4 \cdot 10^6$ вершин

начального яруса графа G . Причиной этого является функциональная несвязность (независимость) пикселей матрицы сенсора: $p_{i,j} \neq f(p_{(i\pm 1),(j\pm 1)})$. Таким образом, алгоритмы обработки изображений характеризуются максимальным потенциальным параллелизмом.

Пример. Рассмотрим простейший алгоритм двумерной фильтрации изображения в окне 3×3 :

$$(1) \quad (p_f)_{i,j} = \frac{1}{9} (\sum_{i-1}^{i+1} \sum_{j-1}^{j+1} p_{i,j} \cdot c_{i,j}),$$

где $p_{i,j}$ – значение пикселя исходного изображения; $c_{i,j}$ – значение элемента ядра фильтра; $(p_f)_{i,j}$ – значение пикселя отфильтрованного изображения; $i = 0, 1, \dots, N1$; $j = 0, 1, \dots, N2$.

Анализ алгоритма дает $D = 9$, $H = 6$. Это означает, что минимальное время однократной реализации данного фильтра составляет $6t$ (t – время выполнения арифметической операции) при числе процессоров, равным 9, т.е. $L_{\square} = 9$. Для реализации алгоритма необходимо выполнить 18 арифметических операций, т.е. алгоритмическая сложность вычислений $L_a = 18$.

Автономная временная сложность вычислений L_t^{aut} соответствует режиму однократной обработки данных, конвейерная временная сложность L_t^{conv} , соответствует режиму непрерывной обработки потоков данных, поступающих на вход системы. Для рассматриваемого алгоритма:

- при однократной фильтрации интервал времени от поступления массивов $\{p_{i,j}\}$ и $\{c_{i,j}\}$ до выдачи параметра $(p_f)_{i,j}$ составляет $L_t^{aut} = 6t$;
- при непрерывной фильтрации $L_t^{conv} = t$.

В таблице 1 приведены параметры алгоритмического графа $2D$ -фильтрации изображений размерностью 3×3 и оценки сложности вычислений. Алгоритм $2D$ -фильтрации **полного** изображения размерностью $N1 \times N2$ пикселей требует реализации примерно $N1 \cdot N2$ таких фильтров. ЯПФ графа алгоритма обработки полноразмерного изображения имеет максимально возможную ширину $D \approx N1 \cdot N2$, в то же время высота графа минимальна: $H = 6$. Параметры алгоритмических графов приведены в таблице 1.

Таблица 1. Характеристики ЯПФ графов для алгоритма $2D$ -фильтрации изображений и оценки сложности вычислений.

Размерность матрицы изображения	Параметры графа		Оценки сложности вычислений			
	D	H	L_t^{aut}	L_t^{conv}	L_h	L_a
3x3	9	6	6	1	9	18
N1xN2	N1·N2	6	6	1	N1·N2	18 N1·N2

Как видно, при теоретически-возможном максимальном распараллеливании размерность изображения не влияет на временную сложность вычислений.

Алгоритмы внутрикадровой обработки обладают максимальным параллелизмом обработки данных, что дает возможность реализации видеообработки в режиме жесткого реального времени.

3. Однородные вычислительные среды и функционально-ориентированные процессоры

Одним из перспективных и эффективных подходов к построению архитектуры процессора внутрикадровой обработки изображений является использование однородных вычислительных сред (ОВС). Реализация в них массового параллелизма

обработки данных теоретически позволяет обеспечить предельную производительность вычислений. ОВС представляет собой двумерную решетку, в узлах которой располагаются процессорные элементы (ПЭ), каждый из которых соединен с другими только локальными связями, образуя решетчатый граф [5].

Локальная связность ОВС позволяет наращивать число ПЭ методом простой стыковки, что облегчает задачу увеличения числа ПЭ с целью увеличения производительности. Главной особенностью архитектуры ОВС является глобальное распараллеливание обработки данных, уровень которого может достигать до бита.

Архитектура ОВС ориентирована на эффективную реализацию базовых процедур цифровой обработки изображений, поэтому сами процессоры принадлежат к классу функционально-ориентированных процессоров (ФОП), далее - ОВС-ФОП.

Можно выделить два основных типа архитектур ОВС-ФОП:

- Архитектуры с одним потоком команд и множественными потоками данных – SIMD [8]. Этот тип ОВС-ФОП содержит аппаратно реализованный двумерный массив ПЭ, который по-тактно выполняет одну общую программу и реализует конкретный алгоритм под управлением одного потока команд. Все ПЭ выполняют в каждом такте одну и ту же команду. Конфигурация массива ПЭ не зависит от реализуемого алгоритма, а в большей степени определяется структурой массива обрабатываемых данных.

- Архитектуры с множественным потоком команд и множественным потоком данных – MIMD [8]. Эти ОВС-ФОП содержат множество ПЭ, соединенных локально и предварительно настроенных на реализацию конкретного алгоритма. Каждый ПЭ выполняет только одну команду на всем интервале времени решения задачи. Конфигурация массива ПЭ представляет собой результат отображения графа алгоритма на клеточное пространство ОВС.

4. ОВС-ФОП для внутрикадровой обработки изображений

ОВС-ФОП с хранением программ во внешней памяти (SIMD). В середине 90-ых годов интерес исследователей и разработчиков вызвали задачи обработки ансамблей точечных изображений общим числом более 50 тыс. за время 10 мс, включая различные виды фильтрации и алгоритмов типа motion estimation, что было невозможно реализовать ни на одной существующей в то время процессорной платформе и было принято решение о разработке ОВС-ФОП. Такая разработка была проведена коллективом разработчиков, при этом известными решениями уже являлись некоторые известные проекты (например, процессор GAPP разработки фирмы NCR, USA [9]), которые могли служить ориентирами для создания архитектурных и схемотехнических решений.

На основе результатов исследований по параллельной обработке потоков данных было разработано систолическое операционное устройство (СОУ) и на его основе двумерный ОВС-ФОП, в котором были реализованы базовые алгоритмические конструкции, составляющие основу процедур обработки двумерной фильтрации массивов данных. Архитектура ФОП позволяла вводить пиксели в темпе поступления и обрабатывать их таким образом, чтобы завершить обработку к моменту прихода следующего кадра. Данный ОВС-ФОП имеет два независимо работающих интерфейса – один предназначен для работы с датчиком, другой - с центральной вычислительной системой.

Исследования возможностей систолического ОВС-ФОП показали, что по сравнению с другими архитектурами он обеспечивает максимальную скорость внутрикадровой обработки изображений. Например, оконтуривание произвольного

числа протяженных изображений, попавших в плоскость датчика, занимает всего 9 тактов. В этом процессоре реализовано микропрограммное управление, это позволяет программисту задавать и изменять в каждом такте команды и адреса локальных ОЗУ ПЭ.

На этапе создания видеосистемы в целом была проведена разработка СБИС СОУ, архитектура которой представляет собой матрицу 8×8 ПЭ, каждый из которых содержит битовое АЛУ, систему коммутаторов и ОЗУ емкостью 128 бит. Однородность архитектуры СОУ позволила реализовать 64 ПЭ в габаритах одного полупроводникового кристалла СБИС. Для обеспечения реализации уже ОВС-ФОП была разработана однослойная ситалловая микроплата размером 60×48 мм², на которой размещается 8 бескорпусных кристаллов СБИС, что в результате дает 512 ПЭ. Полномасштабный ОВС-ФОП состоял из 32 микроплат, он реализовал внутрикадровую и межкадровую обработку изображений в темпе поступления кадров ($f = 100$ Гц) и имел массу 200 г и объем 0.3 дм³ при мощности потребления не более 10 Вт.

ОВС-ФОП с хранимым алгоритмом (MIMD). Начиная с 1970-ых годов, в России был реализован ряд проектов MIMD ОВС-ФОП, в ряде случаев были разработаны и испытаны экспериментальные образцы БИС ОВС либо ФОП в целом. ОВС-ФОП данного типа отличается от архитектур SIMD режимом функционирования: на предварительном этапе каждый ПЭ настраивается на выполнение конкретной команды, и после того, как настроенными окажутся все ПЭ, на внешние интерфейсы подаются потоки данных, которые сразу же начинают обрабатываться и перемещаться по всей ОВС. Обработка данных в этом случае совмещается с процессом их перемещения по массиву ПЭ. Очевидно, что в данном случае реализуется максимально быстрая обработка информации, так как она ведется в темпе приема входных данных.

В течение ряда лет были проведены работы по разработке ОВС на основе СБИС «МиниТера» с архитектурой MIMD [11]. Особенности этой СБИС:

- в состав каждого ПЭ входит АЛУ и потоковый умножитель, позволяющий умножать пары чисел, поступающих на его входы;
- в каждом ПЭ имеется регистр конфигурации, позволяющий в процессе обработки данных оперативно изменять содержимое служебных регистров;
- система команд каждого ПЭ содержит всего 49 команд.

В ходе проведения работ был разработан макет ОВС-ФОП на базе СБИС «МиниТера» и создана компьютерная рабочая станция, на которой был разработан и отработан ряд прикладных программ цифровой обработки сигналов и изображений.

5. Заключение

Внутрикадровая обработка изображений является одной из наиболее вычислительно трудоемких задач мобильных видеосистем реального времени. Алгоритмы внутрикадровой обработки обладают максимально-возможным уровнем параллелизма, что создает предпосылки для достижения высокой производительности видеопроцессоров. Одним из возможных решений этой проблемы является применение функционально-ориентированных процессоров на основе однородных вычислительных сред – ОВС-ФОП, которые реализуют параллельную обработку на уровне отдельных слов данных. Процессоры с однородной структурой позволяют предельно минимизировать время внутрикадровой обработки изображений за счет организации обработки потоков пикселей на двумерных массивах процессорных элементов.

Список литературы

1. Jensen J. R.. Introductory Digital Image Processing: A Remote Sensing Perspective: Prentice Hall Press, 2015. 623 p.
2. Remote Sensing – Advanced Techniques and Platforms / Edited by B. E. Ramirez, 2012. 462 p.
3. Popescu D., Stoican F., Stamatescu G., et al. // Sensors. 2020. Vol. 20. P. 817. doi: 10.3390/s20030817.
4. Cassinis L.P., Fonod R., et al. // IWSCFF 2019 Workshop. 16-19 July 2019, Glasgow, Scotland.
5. Воеводин В.В. Математические модели и методы в параллельных процессах. М.: Наука, 1986.
6. Цилькер Б.Я., Орлов С.А. Организация ЭВМ и систем. – С.Пб. Питер, 2006. 668 с.
7. Лукин Н.А. Основы теории проектирования архитектур функционально-ориентированных процессоров для систем реального времени // Высокопроизводительные вычислительные системы. Материалы Пятой Международной научной молодежной школы и Пятой Международной молодежной научно-технической конференции. 31 августа - 6 сентября 2008, Таганрог. Таганрог: Изд-во ТТИ ЮФУ, 2008. С. 115-166.
8. Flynn M. J. Some computer organizations and their effectiveness // IEEE Transactions on Computers. 1972. Vol. C-21, No. 9. P. 948-960.
9. Preston K. Detection of Weak, Subpixel Targets Using Mesh-Connected Cellular Automata // IEEE Transactions on Aerospace and Electronic Systems. 1990. Vol. AES-26, No. 3. P. 548-558.
10. Lookin N.A. Function-oriented processors with homogenous architecture for data flow signal processing in the space-born control systems // 2014 24nd Int. Crimean Conference “Microwave & Telecommunication Technology” (CriMiCo’2014). Sevastopol, 2014, Crimea, Russia, Vol. 1. P. 417-418.