

ОДИН ИЗ ПОДХОДОВ К СОЗДАНИЮ ПРОГРАММНО-АППАРАТНЫХ СИСТЕМ ОБРАБОТКИ ВИДЕОИНФОРМАЦИИ

© 2004 г. Ю. И. Борисов, канд. техн. наук

Научно-технический центр "Модуль", Москва
E-mail: depart@module.ru

Рассмотрен один из возможных подходов к созданию систем обработки видеoinформации в реальном масштабе времени с взаимоувязанными решениями по конструированию аппаратных систем и программной реализации алгоритмов.

Коды OCIS: 110.2960, 100.1160, 3000.30040.

Поступила в редакцию 13.11.2003.

Одной из актуальных технических задач в настоящее время является обработка в реальном времени потока видеоданных, поступающих от видеокамеры. В данном контексте диапазон используемых длин волн, в котором работает камера, не принципиален. Цель обработки – обнаружение и распознавание объектов определенного типа.

Подобные задачи решают, например, системы технического зрения. Следует, однако, учитывать, что различия в условиях функционирования определяют существенные отличия таких систем друг от друга. В первую очередь по сложности. Можно выделить несколько групп таких систем.

Одна группа систем технического зрения предназначена для работы в лабораторных (производственных) и близких к ним условиях. Это подразумевает: специальное освещение, постоянные ракурсы и масштаб, детерминированные объекты и простые фоны. В зависимости от конкретных условий такие системы реализуются либо на базе персональных компьютеров, либо с использованием специальных аппаратных средств поддержки вычислений.

Другую группу систем технического зрения образуют системы, предназначенные для работы с изображениями естественных сцен. Эта группа принципиально отличается от упомянутой выше повышенной сложностью фона, пониженным качеством освещения, повышенной зашумленностью и значительной неопределенностью в описании объектов. Это определяет существенное усложнение задачи. Именно эта группа – предмет настоящего рассмотрения.

Таким образом, задача обработки видеoinформации может быть сформулирована в следующей постановке. Требуется принять от камеры в реальном масштабе времени поток изображений естественных сцен, обнаружить на изображении объект определенного типа, осуществить его сопровождение на изображениях, классифицировать объект и выдать во внешние устройства информацию о результатах обнаружения и классификации.

Отдельные приемы обработки видеoinформации достаточно хорошо изучены и систематизированы, например, в [1]. Тем не менее, создание законченных систем, способных решать конкретные прикладные задачи, остается большой проблемой.

Ниже приведены основные этапы обработки отдельного кадра изображения:

- дискретизация и квантование кадров непрерывных видеоизображений;

- сжатие изображения (по необходимости);

- нормализация изображения (адаптация к условиям наблюдения);

- отделение объектов от фона (обнаружение объектов);

- сегментация изображений (выделение участков, обладающих общими признаками);

- распознавание объектов.

При обработке последовательности кадров к упомянутым этапам добавляются идентификация объектов на различных кадрах и определение информации о характеристиках объектов, в том числе динамических и статистических.

Почти все из перечисленных задач приводят к необходимости обработки большого объема информации, так как кадр изображения содержит 100 000 и более пикселей, а обрабатывать кадры необходимо с частотой формирования кадров (обычно это 25–30 Гц). Обработка осуществляется методами фильтрации, унитарных преобразований, геометрических преобразований, морфологического анализа, нейросетевого анализа.

Каждая из этих операций в применении к 100 000 и более пикселей и с учетом частоты поступления видеок кадров приводит к большой вычислительной емкости алгоритма обработки изображений (миллиарды элементарных операций в секунду).

Таким образом, основной проблемой, которую приходится решать при разработке систем обработки видеoinформации, является реализуемость вычислений в реальном масштабе времени, которую можно разделить на следующие частные проблемы:

вычислительные — как организовать вычислительный процесс, чтобы максимально использовать возможности аппаратных средств;

аппаратурные — какое количество и какой аппаратуры из реально доступной использовать;

программные — как разрабатывать программное обеспечение задачи, чтобы максимально использовать особенности архитектуры программных средств и существующий программный задел.

Успешное решение такого комплекса проблем требует соответствующей методологии. Принципиальной методической особенностью подхода к созданию аппаратно-программных систем, принятого в НТЦ “Модуль”, является его комплексность. Под этим понимается, что выбираемые алгоритмические подходы, с одной стороны, определяют используемые аппаратные решения, а с другой — сами зависят от используемых аппаратных решений.

Проиллюстрируем это положение. Имеющийся в НТЦ “Модуль” опыт разработки систем обработки видеoinформации позволил условно разделить вычисления, проводимые в ходе работы таких систем, на две большие группы:

массовые операции — их число велико — пропорционально числу элементов изображения — пикселей;

локальные вычисления — их число невелико — пропорционально числу объектов, присутствующих на изображении.

Если опираться на этапы обработки, указанные выше, то область массовых операций распространяется на этапы от сжатия до частичного распознавания объектов. Анализ собственных разработок НТЦ “Модуль” показал, что вычислительные затраты, связанные с массовыми операциями, на порядки превышают затраты, обусловленные локальными вычислениями. Другой практический вывод из имеющегося опыта работ заключается в том, что реализацию массовых операций зачастую можно свести к выполнению операций умножения с накоплением.

Этот вывод в значительной мере определил облик процессора Л1879ВМ1 [2], разработанного в НТЦ “Модуль”, в том числе и для решения задач, связанных с обработкой изображений. Принципиальные особенности данного процессора (его векторной части) — способность за один такт работы выполнять операцию умножения с накоплением и возможность программного деления его разрядной сетки, что позволяет выполнять операции над данными с разрядностью от 1 до 64 разрядов. Такая возможность обеспечивает высокую эффективность использования 64-разрядного вычислительного узла. Особенно это наглядно проявляется при обработке данных малой разрядности (как правило, разрядность одного пиксела равна 8–16). Так, в [3] отме-

чается, что при реализации некоторых цифровых фильтров, сводящихся в конечном счете к умножению с накоплением, этот процессор оказывается на порядок быстрее универсальных процессоров, работающих к тому же на тактовых частотах, на порядок превышающих тактовую частоту Л1879ВМ1.

Элементы векторов в процессоре Л1879ВМ1 на проходе (без затрат дополнительного времени) могут быть подвергнуты пороговому преобразованию. Поскольку операция умножения матрицы на вектор с ограничением по порогу является основной для нейронных сетей, процессор Л1879ВМ1 имеет и другое наименование — NeuroMatrix® процессор.

Сказанное выше позволяет заключить, что при реализации этапов обработки видеoinформации, связанных с массовыми операциями, целесообразно использовать процессоры, аналогичные процессору Л1879ВМ1. А массовые операции в системах обработки видеoinформации, конструируемых на основе такого процессора, в свою очередь целесообразно сводить к эффективно реализуемым данным процессором операциям — операциям умножения с накоплением — особенно при малой разрядности данных.

Максимальное использование операции умножения с накоплением при решении задач обработки изображений является одним из положений рассматриваемой методики. Не все алгоритмы, применяемые при обработке видеoinформации, естественно сводятся к такой операции. Выход тут в разработке нетрадиционных алгоритмических решений. В НТЦ “Модуль” в этом направлении накоплен определенный опыт. Применение нейросетевых решений на этапе распознавания естественно. Нам удалось свести к операциям умножения с накоплением и задачу сегментации.

Высокая вычислительная сложность задач обработки видеoinформации приводит к необходимости использования многопроцессорных систем для реализации алгоритма обработки изображения в реальном времени. Поэтому следующее положение используемой методики заключается в распределении вычислительного процесса между несколькими процессорами. Основными подходами тут являются “разложение” вычислительного процесса на отдельные параллельные процессы, организация конвейера, комбинация указанных способов.

Выбор конкретного способа определяется прикладной задачей и возможностями аппаратной реализации многопроцессорного вычислителя.

Следующее положение используемой методики заключается в учете возможностей распараллеливания вычислительного процесса при разработке или выборе конфигурации аппаратных средств. В НТЦ “Модуль” отработаны подходы к созданию

аппаратных средств, обладающих гибкостью к организации "разложения" вычислительного процесса.

Еще одна из особенностей описываемой методики создания систем обработки видеoinформации заключается в учете особенностей используемых аппаратных средств, в частности сигнального процессора L1879BM1, при разработке программ алгоритмов обработки.

В заключение представляется целесообразным проиллюстрировать сформулированные положения на примере одной из аппаратно-программных систем, созданных в НТЦ "Модуль" для обработки видеoinформации в реальном масштабе времени. В этом качестве выбран комплекс "Трафик-Монитор", созданный на базе процессора L1879BM1 и предназначенный для контроля загруженности автомобильных дорог [4]. Комплекс "Трафик-Монитор" рассчитан на работу с неподвижной камерой, закрепленной над дорожным полотном и наблюдающей за выделенным участком дороги.

Программное обеспечение вычислительного комплекса "Трафик-Монитор" включает: управление вычислительным процессом, обработку оцифрованных видеок кадров изображений, формирование статистической информации о транспортных потоках.

Управление вычислительным процессом заключается в программной поддержке внешних интерфейсов комплекса "Трафик-Монитор", управлении контроллером видеокамер, синхронизации работы процессоров и осуществлении обмена данными между процессорами, обработке сообщений от персонального компьютера диспетчерского пункта, автоматическом перезапуске программы комплекса "Трафик-Монитор" в случае сбойных ситуаций.

Цифровая обработка изображений программно реализует алгоритмы обнаружения, сопровождения и распознавания транспортных средств.

При обнаружении в рассматриваемом комплексе используются:

проективное преобразование, реализованное на векторной части процессора и обеспечивающее нормализацию обрабатываемого в последующем изображения;

бинаризация изображения с использованием фильтра Собела (апертура 3×3) и низкочастотного фильтра (апертура 5×5) и последующей пороговой обработки, реализованных на векторной части процессора;

сегментация на основе анализа геометрической структуры объекта по бинаризованному изображению.

При сопровождении в качестве признаков используется описание формы объекта.

Распознавание в рассматриваемом комплексе реализовано в виде двух последовательных опера-

ций: снижения размерности вектора признаков и собственно распознавания. Снижение размерности вектора признаков достигается за счет использования преобразования Адамара, эффективно реализуемого на векторной части процессора. На этом этапе размерность вектора признаков снижается в десятки раз.

Собственно распознавание класса транспортного средства выполнено на базе использования предварительно обучаемой нейросети. Такой алгоритм эффективно реализуется на векторной части используемого процессора.

Формирование статистической информации программно реализует алгоритмы определения характеристик транспортных потоков. Формирование (сбор) статистической информации о транспортных средствах, прошедших через поле зрения комплекса, составляет задачу, имеющую относительно малую вычислительную сложность. Поэтому указанная часть алгоритмов комплекса реализована традиционно – без использования специфических свойств процессора.

Проведенные расчеты показали, что для решения перечисленных задач в реальном времени необходимо не менее 2 процессоров L1879BM1.

Решение задач вычислительного комплекса "Трафик-Монитор" было разделено на 2 параллельных вычислительных процесса. Первый процесс включает управление вычислениями в целом, фильтрацию, выделение контуров, обнаружение объектов в ночное время (по фарам). Второй процесс включает обнаружение объектов днем (по геометрическим параметрам), распознавание, объединение результатов, расчет статистических характеристик.

Для реализации параллельной работы двух процессов было синтезировано устройство видеообработки BM1, включающее 2 процессора L1879BM1 и оперативную память объемом 16 Мбайт, которое позволило эффективно реализовать обработку 25 кадров в секунду (стандарт PAL) или 30 кадров в секунду (стандарт NTSC) при размере кадра $\sim 10^5$ пикс.

Оптимизация программ, с точки зрения обеспечения минимального времени вычислений, проводилась посредством учета особенностей архитектуры процессора L1879BM1 и являлась одной из наиболее сложных задач разработки. Поэтому исходные тексты программы комплекса "Трафик-Монитор" написаны на языке ассемблера процессора L1879BM1.

В результате применения комплексного подхода был создан программно-аппаратный комплекс "Трафик-Монитор", обеспечивающий в реальном времени обнаружение, сопровождение, классификацию и определение характеристик транспортных потоков.

Выводы

Рассмотрена методология создания аппаратно-программных систем анализа видеоинформации в реальном масштабе времени, заключающаяся в комплексном подходе к разработке систем обработки информации, позволяющем оптимально использовать возможности аппаратной реализации системы и учитывать свойства аппаратуры при создании алгоритма системы. Комплексный подход реализуется путем:

учета особенностей используемых аппаратных средств, в том числе максимального использования операций умножения с накоплением при программной реализации алгоритмов;

использования рациональных способов "разложения" вычислительного процесса на параллельные процессы;

применения для реализации массовых операций эффективных аппаратных решений, в частности процессора Л1879ВМ1;

применения для вычисления параллельных процессов эффективных аппаратных решений, в частности многопроцессорных систем.

Указанная методология позволила создать образцы систем, обеспечивающих решение в реальном времени задач обработки видеоинформации и иной цифровой информации.

ЛИТЕРАТУРА

1. Денисов Д.А. Компьютерные методы анализа видеоинформации. Красноярск: Изд-во Красноярского университета, 1993. 188 с.
2. <http://www.module.ru/rupproducts/proc/nm6403.shtml#an>.
3. Фомин Д., Черников В., Вискне П., Яфраков М., Шевченко П. DSP и RISC объединились // Открытые системы. 1999. № 5-6. С. 14-18.
4. Борисов Ю. Комплекс "Трафик-Монитор" на базе процессора Л1879ВМ1. Особенности разработки // Электроника: НТБ. 2003. № 6.