

Модели Характеристики	Архитектурная модель	Абстрактные ФАС	Компьютерные ФАС
Временная сложность	+	+	+
Аппаратная сложность	+	-	+
Емкостная сложность	+	+	+
Программная сложность	-	-	+
Структурная сложность	-	-	+
Дискретность	+	+	+
Детерминированность	+	+	+
Элементарность	+,-	+,-	+
Массовость	+	+	+
Иерархичность	+	-	+

Как видно из данной таблицы, компьютерные ФАС имеют наиболее полные характеристики.

Выводы

1. Архитектурные методы проектирования компьютеров основаны на опыте разработчика, умении находить новые схематехнические и системотехнические решения, которые только частично базируются на использовании характеристик сложности абстрактных алгоритмов.
2. Модели абстрактных формальных алгоритмических систем, которые созданы для исследования проблем разрешимости, не эффективны для разработки реальных аппаратно-программных средств, так как не учитывают влияние аппаратных средств на сложность вычислительного процесса.
3. Модель аппаратно-программных алгоритмов (SH-модель) с выявлением в своём определении аппаратно-программных средств является основой новой технологии проектирования компьютерных систем.
4. Центральным понятием аппаратных средств является "элементарный преобразователь", который позволяет уточнить понятие теории алгоритмов "элементарность" и "иерархичность".
5. Использование множества элементарных преобразователей позволило сформулировать понятие характеристики "аппарат-

ная сложность" и уточнить понятие характеристики "временная сложность".

6. Исследование компьютерных систем посредством теории сложности SH-модели алгоритма имеет преимущества в сравнении с методами архитектуры и методами анализа абстрактных формальных алгоритмических систем.
7. Следует различать программную и структурную сложность.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Трахтенброт Б.А. Алгоритмы и вычислительные автоматы. – М.: Сов. радио, 1974. – 200 с.
2. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ, 2-е издание. : Пер.с англ. – М.: Изд. дом «Вильямс», 2005. – 1296 с.
3. Черкасский М.В. SH-модель алгоритму // *Комп'ютерна інженерія та інформаційні технології. Вісник НУ "Львівська політехніка"* – Львів, 2001, №433, с.127-134.
4. Даль В.И. Толковый словарь живого великорусского языка в четырёх томах. – М.: изд "Русский язык" 1999.
5. Математическая энциклопедия. Гл. ред. И.М. Виноградов. – М.: Советская Энциклопедия, 1977. – в 5-и томах.
6. Мухаммед ибн Муса ал-Хорезми. Книга об индийском счете. – Ташкент, изд-во "Фан", 1973.
7. Успенский В.А., Семенов А.Л. Теория алгоритмов: основные открытия и приложения. – М.: Наука Гл. ред. физ.-мат. лит., 1987. – 288 с.
8. Блюм М. Машино-независимая теория сложности рекурсивных функций// Библиотека "Кибернетического сборника". Проблемы математической логики. М.: Мир, 1970.
9. Cherkaskyy M.V. Computer algorithms: axioms and consequences // *Advanced Computer Systems and Networks: Design and Application. Proceedings of the 2nd International Conference ACSN – 2005. September 21-23, 2005, Lviv – Ukraine.* – Publishing House of Lviv Polytechnic National University, 2005, p.147.
10. Anatoly Sachenko, Mykola Cherkaskyy, Oleksandr Osolinskiy. Algorithm Conception Clarification Based on the SH-Model. Proceedings of IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 6-8 September 2007, Dortmund, Germany. – Pp. 300-303.

Материал поступил в редакцию 31.10.08

CHERKASSKIY M., SACHENKO A. Technology of designing on the basis of the theory of complexity of computer algorithms

The lacks of designs of abstract models of algorithms are considered. The necessity of specification of concept "algorithm" is shown. The new model of algorithm (SH-model) intended for construction of effective technology of designing of computer means is described.

УДК 681.4

Дудкин А.А.

ОСНОВНЫЕ ПОДХОДЫ К РЕАЛИЗАЦИИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ ТОПОЛОГИЧЕСКИХ СЛОЕВ ИНТЕГРАЛЬНЫХ СХЕМ В СИСТЕМАХ КОМПЬЮТЕРНОГО ЗРЕНИЯ ПРОЕКТИРОВАНИЯ И ПРОИЗВОДСТВА ЭЛЕКТРОННЫХ ИЗДЕЛИЙ

Введение. Системы компьютерного зрения (СКЗ) относятся к промышленным программным продуктам, существенной чертой которых является высокий уровень сложности: один разработчик практически не в состоянии охватить все аспекты такой системы. Системы подобного типа обычно имеют большое время жизни, большое количество пользователей оказывается в зависимости от их нормального функционирования. Это обуславливает необходимость индустриализации производства программной продукции – применения высокоэффективных технологий создания программного обеспечения (ПО).

Основные подходы к реализации алгоритмов обработки.

Ниже приведены основные подходы, которые нашли воплощение в проектировании систем обработки изображений топологии инте-

гральных схем (ИС), которые в совокупности обеспечивают эффективную обработку топологии ИС.

Объектно-ориентированное проектирование (ООП). В основе ООП лежит представление о том, что программную систему необходимо проектировать как совокупность взаимодействующих друг с другом объектов, рассматривая каждый объект как экземпляр определенного класса, причем классы образуют иерархию. Объектно-ориентированный подход отражает топологию языков высокого уровня, таких как Smalltalk, Object Pascal, C++, CLOS и Ada. Объектная модель имеет ряд преимуществ: позволяет в полной мере использовать возможности объектных и объектно-ориентированных языков программирования; существенно повышает уровень унификации разработки и пригодность для повторного использования не только программ, но и проектов, что ведет к созданию среды разра-

ботки: приводит к построению систем на основе стабильных промежуточных описаний, что упрощает процесс внесения изменений и др.

Методы и алгоритмы обработки реализуются в виде завершённых самостоятельных модулей. Объектно-ориентированный подход подразумевает также разработку проблемно-ориентированных алгоритмов, т. е. алгоритмов, учитывающих специфику изображений объектов: специальные алгоритмы обработки слоев, специализированный контроль при восстановлении топологии и т.п.

Параллельная обработка. Изображения топологических слоев ИС (металлизации, диффузионного или поликремневого и др.) получают с помощью специальных или стандартных устройств ввода данных (сканеров) и представляются множеством растровых кадров, которое формирует поток данных на входе СКЗ. Задача СКЗ – обеспечить обработку потока за минимальное время с применением различных алгоритмов для кадров разных слоев.

В настоящее время параллельная обработка информации является одной из наиболее востребованных технологий проектирования ПО. Во многих случаях разработка параллельного приложения осуществляется на базе имеющихся последовательных алгоритмов или их композиции. Перенос существующего кода и его адаптация для средств организации параллельных вычислений требует существенных затрат. Вычислительные операции, входящие в состав параллельного алгоритма, часто имеют универсальный характер и могут применяться в различных задачах обработки информации. Реализация элементов алгоритма в виде переносимых операций позволяет перейти к компонентному проектированию, когда программа конструируется из крупных блоков. В этом случае сокращается время разработки и отладки программы. В настоящее время подходы компонентного проектирования активно применяются в параллельном программировании.

Наконец, процесс выполнения параллельной программы требует применения современных методов планирования и оптимизации нагрузки. В связи с широким распространением параллельных компьютеров их архитектуры и характеристики различаются между собой. Во многих случаях узлы параллельных систем имеют гетерогенные характеристики, как пространственные, так и временные. Для эффективного выполнения параллельных программ в таких системах необходимо снабжать средства организации параллельных вычислений возможностями контроля и динамического изменения конфигурации алгоритма.

Помимо названных проблем, которые приходится решать при проектировании параллельной обработки, немаловажным является обеспечение эффективного взаимодействия режима работы на ПЭВМ с кластерной обработкой, включая мультипроцессорные (суперкомпьютерные) кластерные архитектуры.

Сочетание ручной и автоматической обработок. Первая проблема, которую необходимо решить при проектировании СКЗ, это понять, какого вида информация требуется системе и как представить ее в виде совокупности признаков, извлеченных из изображения. Т.е. важно определить заранее объект обработки в терминах измерений и правил, чтобы затем реализовать программно или аппаратно. Важно также выделить недопустимые случаи, чтобы исключить ошибочное отнесение области к объекту.

Далее, анализ качества исходного изображения невозможен без участия компетентного пользователя, равно как построение схемы обработки и анализ результата.

Система должна быть робастной, т.е. адаптировать себя автоматически и достигать постоянно высокой производительности, несмотря на нерегулярность в освещении и условиях фона, приспосабливаться к неопределенностям в углах, позициях и т.п.

Интерфейс системы должен обеспечивать: автоматизированную настройку системы для работы в автоматическом режиме; автоматическое подключение новых алгоритмов; ввод и информационную поддержку нескольких видов представления данных и их визуализацию; автоматический выбор наиболее подходящего алгоритма обработки для конкретного топологического слоя.

Использование внутреннего представления. Внутреннее представление вырабатывается для удобства последующих оптимизационных преобразований и ускорения вычисления. Оно отражает

структуру информации (объекта), специфическую для рассматриваемого преобразования. Поскольку основным объектом в СКЗ является изображение, то в системе используются различные его представления: цветные и полутоновые, бинарные, векторные, матричные, а также различные цветовые координаты.

Возможность использования внешнего ПО. Реализация данного принципа достигается за счет использования форматов-стандартов проектирования ИС (GDSII, Source, Gerber).

Кроссплатформенность. Данный принцип важен, т.к. разрабатываемые системы являются многооперационными: основная система создана как приложение для ПЭВМ в среде Windows, подсистема параллельной обработки – в Linux-подобной ОС.

Основные сведения о системе обработки. Система обработки изображений топологических слоев ИС [1] предназначена для автоматизации этапа восстановления топологии ИС при решении задачи обратного проектирования или автоматизированном контроле топологии при изготовлении ИС или оригиналов топологии (фотошаблонов, ФШ). Топологический слой представлен в виде матрицы перекрывающихся кадров, поскольку для достижения требуемого разрешения съемка кристалла производится отдельными частями. Кадры представлены в формате BMP. Параметры совмещения (склейки) кадров и слоев задаются в текстовом файле настройки системы (*.PAR). Данный файл содержит также конструкторско-технологические ограничения (КТО), необходимые для работы алгоритмов восстановления топологии, а также описание аппаратной вычислительной системы для параллельной обработки (рис. 1):

Результатом работы системы являются описание формы и положения объектов топологии слоя в векторном формате SOURCE или GDSII и текстовый файл склейки (*.CTW), с которым работает редактор топологии GLEW (разработка НТК "БЕЛМИКРОСИСТЕМЫ" УП «Завод полупроводниковых приборов» НПО «Интеграл»).

Система состоит из трех основных подсистем: склейки кадров изображения; предварительной обработки кадров; сегментации и векторного описания объекта.

Каждая из подсистем может работать в трех режимах (рис. 2): пошаговом; пакетном на одной ПЭВМ; пакетном на кластере.

Редактор топологии и ПО для выделения опорных точек и совмещения изображений топологических слоев ИС и ФШ является внешним программным средством.

Обработка изображений топологии в системе начинается с настройки подсистем на обработку в интерактивном режиме. Затем осуществляется автоматическое повторение схемы обработки для каждого исходного изображения слоя на одной ПЭВМ или вычислительном кластере.

Основные подсистемы. В подсистеме склейки кадров реализован ряд эвристических алгоритмов склейки. На рис. 3 представлено главное окно подсистемы, где меню Файл содержит опции для выбора файлов для склейки; меню Настройки позволяет задать параметры, с которыми будет производиться склейка, и выбрать один из алгоритмов склейки. К примеру, 4С означает что склеиваются четверки кадров, а затем склеиваются полученные кадры и т.д. Выбор того или иного алгоритма склейки зависит от степени и характера погрешностей, вносимых устройством ввода. Нахлест по вертикали в пикселях определяет величину S зоны перекрытия кадров (задается в окне Постоянный) и величину ΔS возможного смещения кадров в этом направлении из-за погрешностей сканирования (задается в окне Переменный). Таким образом, выбор оптимального положения кадров возможен в диапазоне $[S-\Delta S, S+\Delta S]$. Нахлест по горизонтали в пикселях по аналогии с предыдущим пунктом меню определяет величину зоны перекрытия кадров и величину возможного смещения кадров по горизонтали. Меню Помощь содержит справочную информацию о подсистеме.

В главном окне размещены следующие информационные панели: Фрагменты отображает список файлов, выбранных для склейки. Последняя склейка в графическом виде показывает результат склейки группы кадров, а Прогресс – состояние процесса склейки.

Данная подсистема может работать параллельно с другими подсистемами на свободных вычислительных мощностях.

из на
связа
же с
ным
циал
отдел
риц с
товке
Г
подго
го оп
выпол
ния, д
произ
Такая
делен
смотри
обраб
ботку

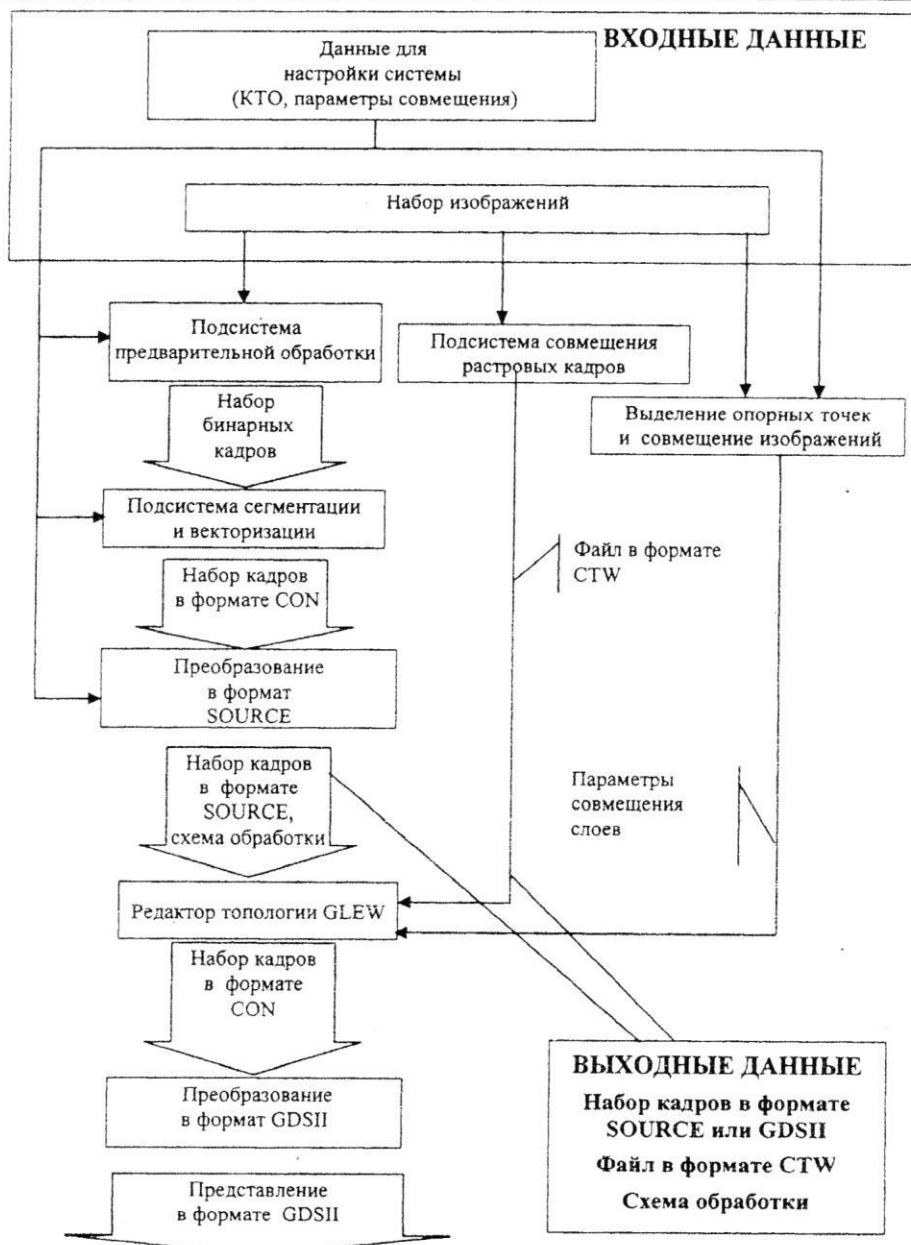


Рис. 1. Схема обработки изображений топологических слоев ИС на одной ПЭВМ

Построение матриц совпадения соседних кадров является одной из наиболее продолжительных стадий общего алгоритма склейки. Это связано с тем, что матрица строится для каждого кадра по одной и той же схеме, но с использованием различных исходных данных. Основным недостатком при этом является необходимость повторной инициализации данных перед каждой итерацией, связанной с обработкой отдельного кадра. Параллельная версия алгоритма построения матриц совпадения лишена этого недостатка при соответствующей подготовке данных для их эффективного однократного использования.

Подсистема предварительной обработки предназначена для подготовки изображений к этапу сегментации и получения векторного описания объекта. Оператор может корректировать параметры выполняемых операций, менять последовательность их выполнения, добавлять или удалять операции из цепочки обработки, т. е.

производить так называемую ручную подстройку схемы обработки. Такая подстройка необходима для настройки подсистемы на определенный класс изображений. В подсистеме по умолчанию предусмотрены два основных сценария обработки. Один ориентирован на обработку изображений слоев металлизации ИС, второй – на обработку шаблонов этих слоев.

Главное окно подсистемы представлено на рис. 4, где меню Файл содержит опции для выбора файла или набора файлов для обработки и сохранения результата; меню Операция и Обработка пакета содержат следующие операции для работы с файлом или набором файлов соответственно: Бинаризация, Площадной фильтр, Восстановление элементов, Выделение контуров. Все операции. Параметры для обработки задаются в окне диалога Настройка. Переключатель шаблон – кристалл позволяет выбрать и сохранить параметры (рис. 4) по умолчанию для изображений ФШ или кристалла.

Пошаговая обработка – позволяет поэтапно выполнить обработку изображения в активном окне, а также используется для настройки параметров обработки на текущее множество изображений (в случае недостаточного качества обработки с параметрами, установленными по умолчанию), чтобы далее выполнить пакетную обработку всего набора изображений в автоматическом режиме.

Пакетная обработка – выполняет автоматически обработку всего множества кадров по выбранной схеме, работает в фоновом режиме и позволяет оператору параллельно выполнять другие задачи.

Результатом является набор бинарных изображений в формате DIB.

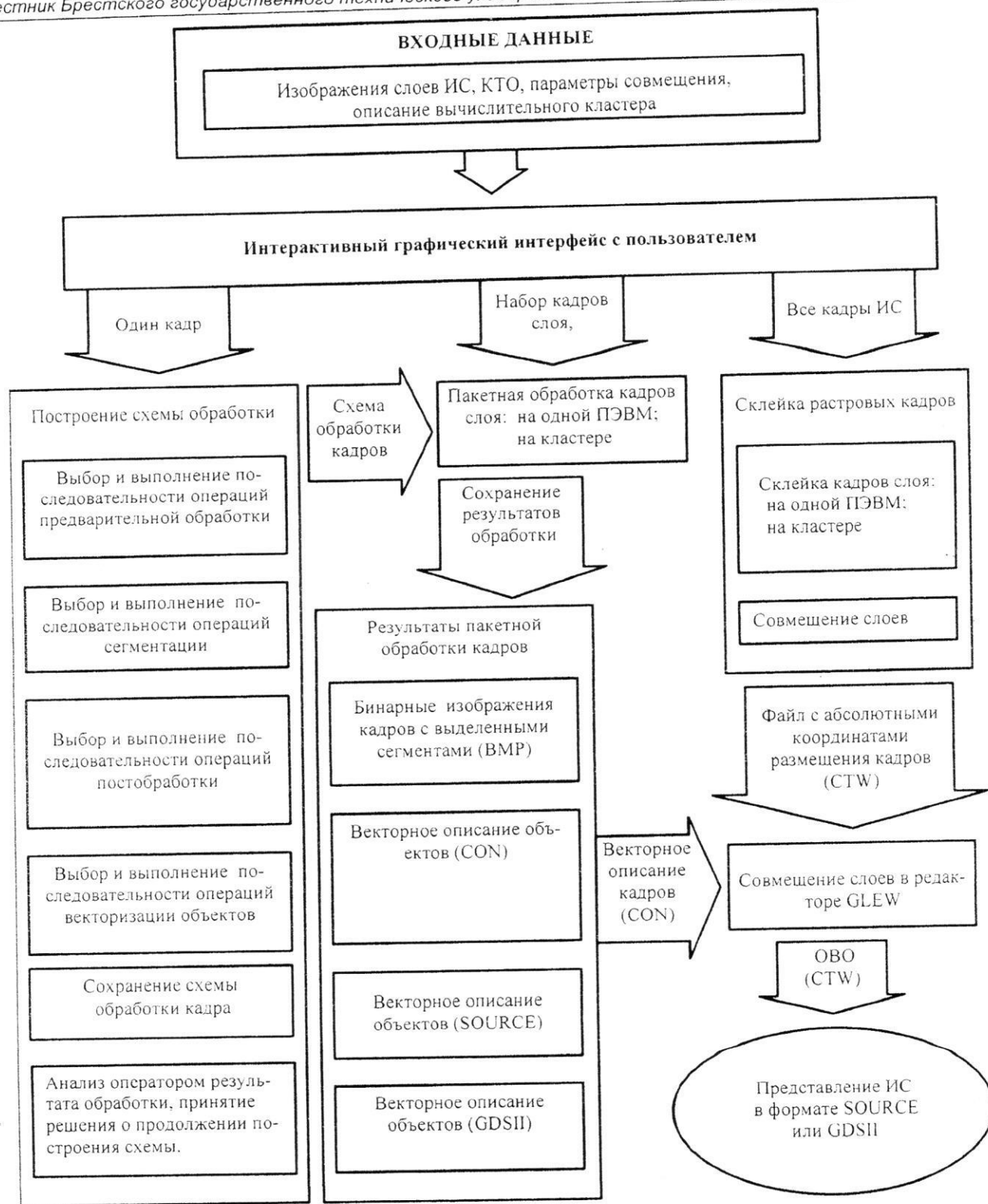


Рис. 2. Режимы работы с системой

Подсистема сегментации и векторного описания объектов решает задачу получения векторного описания объектов. Главное окно программы представлено на рис. 5, где меню Файл содержит опции для выбора файла или набора файлов для обработки и сохранения результата.

Меню Обработка позволяет выбрать следующие режимы обработки изображений.

1. Обработка одного файла – такой тип обработки объекта на изображении, находящемся в активном окне, который используется

для настройки параметров алгоритма сегментации на обработку группы файлов, чтобы далее выполнить обработку этой группы в автоматическом режиме.

2. Автоматическая обработка пакета файлов, с помощью которой выделяется объект из группы файлов, хранящихся в отдельной директории пакетной обработки. При такой обработке выбранные параметры настройки алгоритма распространяются на все файлы из группы. Файлы с результатами хранятся в той же директории.



Рис. 3. Главное окно подсистемы склейки фрагментов изображений

Список файлов для пакетной обработки:

M_0000.bmp
M_0001.bmp

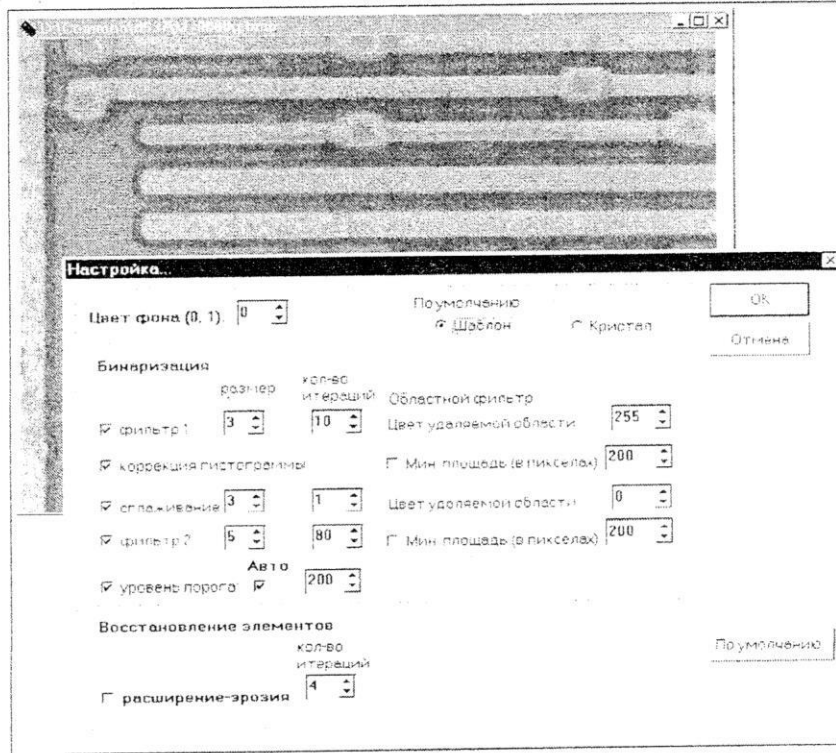


Рис. 4. Главное окно подсистемы предварительной обработки

3. Обработка пакета файлов в полуавтоматическом режиме предназначена для обработки группы файлов в случае, когда не удается успешно произвести выделение объекта на всех кадрах из группы при одинаковых настройках алгоритма сегментации. Выходом из этой ситуации является разбиение этого пакета на части и подстройка алгоритма сегментации на каждую из новых частей. Если же изображения в группе сильно варьируются по своим характеристикам и подстройку алгоритма необходимо проводить почти для каждого изображения, то удобнее вести обработку такого пакета в полуавтоматическом режиме.

После обработки каждого кадра выводится окно подтверждения. Параллельно эксперт может визуально наблюдать результат обработки в отдельном окне. Если результат его удовлетворяет – можно перейти к обработке следующего кадра из пакета либо повторить обработку с другими параметрами. В последнем случае автоматически выводится окно настройки (опции), что дает возможность быстро произвести подстройку алгоритма сегментации. Имеется три группы параметров:

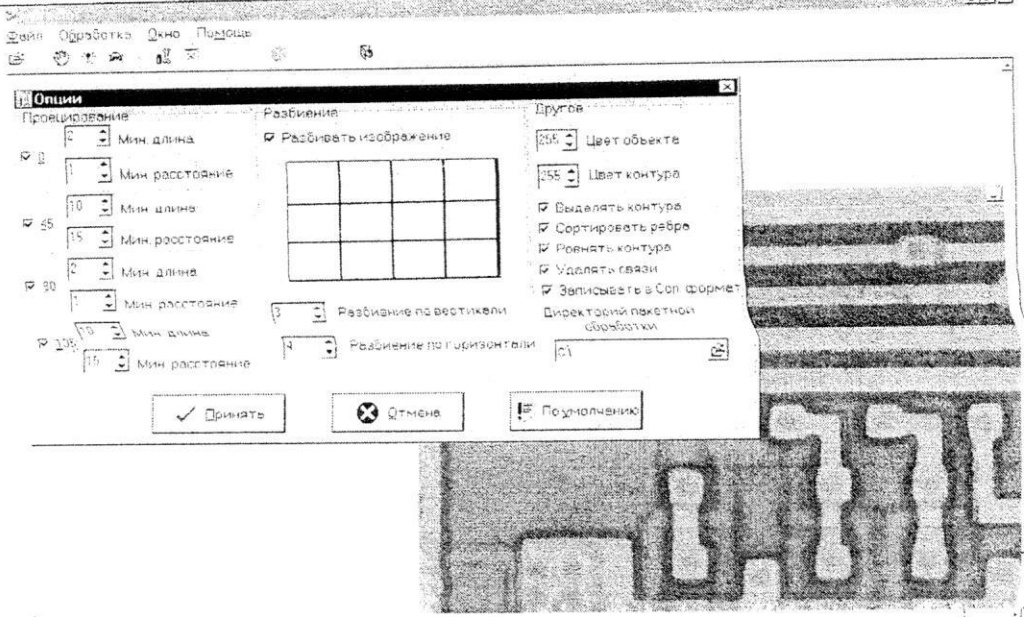


Рис. 5. Главное окно подсистемы векторизации

Проецирование позволяет выбирать предполагаемую ориентацию сегментов, аппроксимирующих выделяемый объект (0, 45, 90 и 135°), и для каждой ориентации соответственно задать в пикселях минимальную длину сегмента, который может быть выделен на изображении, и минимальное расстояние между сегментами.

Разбиение позволяет выбирать количество фрагментов, на которое разбивается изображение при обработке.

Другое позволяет задавать цвет объекта и контура для визуализации, операции обработки и форматы сохранения результата.

Подсистема работает с файлами типа *.bmp (бинарный файл) и *.c.bmp (контурный файл), которые создаются подсистемой предварительной обработки.

Векторное представление выделенных объектов формируется во внутреннем формате CON (в виде бинарного файла).

Организация параллельной обработки. Сценарий алгоритма, ориентированный на применение в системе пакетной обработки на одной ПЭВМ, представляется во внутреннем текстовом формате. Для организации выполнения на кластере используется представление на языке XML. Сценарий содержит описание операций с параметрами вызова и типов кадров, обрабатываемых сценарием, очередь типов, а также распределение операций по процессорам. Он подается на вход интерпретатора сценария и выполняется для всего множества входных данных. Интерпретатор осуществляет разбор файла сценария, предоставляет исполнителю необходимую для формирования контекста вызова информацию, управляет созданием новых вычислительных гранул и постановкой их в очередь планирования.

В состав архитектуры (рис. 6) входят клиент (на основе платформы Windows) и параллельный сервер (на основе платформы Linux). Взаимодействие между клиентом и сервером осуществляется через файловую систему. Клиент на основе ОС Windows осуществляет выполнение задач, обеспечивающих интерфейс системы с пользователем. Он содержит хранилище данных и ПО системы, позволяющее пользователю выполнить все операции для обработки слоев и склейки кадров. Здесь пользователь может задать параметры, которые будут использованы при работе параллельной системы. Клиент также содержит средства удаленного вызова процедур параллельного сервера [2].

Клиент имеет доступ к общему каталогу на сервере с помощью средств сервера Samba. Сервер настраивает определенный каталог для общего доступа, и клиент может монтировать его при загрузке Windows. В данный каталог помещаются файлы для обработки (например, кадры), а также сценарии обработки. Сценарий представляет

собой сгенерированный средствами ПО клиента текстовый файл в синтаксисе Linux shell, содержащий команды запуска процедур параллельного сервера с требуемыми параметрами командной строки, указанными в описании процедур. Сгенерированный сценарий помещается в общий каталог в виде исполняемого командного файла.



Рис. 6. Архитектура вычислительного комплекса

Выполнение сценария со стороны клиента осуществляется посредством протокола SSH (Secure Shell). Реализация этого протокола существует как для платформ Linux, так и для Windows. Со стороны клиента Windows используется программа-клиент Putty, позволяющая настроить доступ к удаленной машине и выполнять на ней команды и сценарии. Программа поддерживает настройку сеанса с указанием требуемого сервера и сценария и может быть вызвана с помощью средств запуска приложений Windows API.

Со стороны сервера существует возможность настройки сервера SSH на беспарольный доступ, чтобы выполнение сценариев с помощью клиента не требовало участия пользователя. Для этого ключи клиента передаются на сервер и указываются как «авторизованные».

Сервер представляет собой параллельный вычислительный кластер на базе ОС Linux, который использует библиотеку поддержки параллельных вычислений MPICH и настраивается на использование SSH без пароля в качестве команды удаленного запуска.

Заключение. В статье описана система, позволяющая автоматизировать этап восстановления топологии ИС при обратном проектировании и контроле изготовления ИС или их фотошаблонов. В отличие от [1], где описаны лишь основные функции системы, представлена как структура системы, так и режимы работы, способы настройки системы и разработки сценариев обработки (подготовки пакета для обработки), а также архитектура вычислительных

средств параллельной обработки. В основе системы лежат подходы автоматической сегментации, бинаризации и векторизации объектов, автоматизированное построение схем обработки и анализ результата оператором, после чего осуществляется корректировка вычислительной схемы. Использование объектно-ориентированного подхода, параллельной обработки, внутреннего представления, сочетание ручной и автоматической обработок в совокупности обеспечивают эффективную обработку топологии ИС.

СПИСОК ЦИТИРОВАННЫХ ИСТОЧНИКОВ

1. Система цифровой обработки изображений слоев интегральных микросхем / А.А. Дудкин [и др.] // Идентификация образов; под ред. Р.Х. Садыхова. – Минск, 2001. – Вып. 2. – С. 72-87.
2. Дудкин, А.А. Реализация параллельной обработки изображений в системах технического зрения на базе МРП и многоагентной архитектуры / А.А. Дудкин, А.В. Отвагин // Вестник Брестского государственного технического университета. Физика, математика, информатика. – 2006. – № 5 (41). – С. 2-8.

Материал поступил в редакцию 20.09.08

DOUDKIN A.A. Main approaches to implementation integrated circuits layout processing in computer vision systems for design and production of electronic devices

The system is described to automate integrated circuits layout restoring during integrated circuits re-design and production inspection. An automatic segmentation, binarization and vectorization, computer-aided development of processing scenarios and analyses procedures are in the basis of the system. An effective layout processing is provided by implementation of object-oriented approach, parallel realization, using inner representation and combination of manual and automatic processing.

УДК 004.8.032.26

Головко В.А., Безобразов С.В., Меленчук В.В.

НЕЙРОСЕТЕВОЙ ПОДХОД ДЛЯ ОБНАРУЖЕНИЯ ЭЛЕКТРОННОГО СПАМА

Введение. С развитием компьютерных наук и компьютерной техники общество столкнулось с проблемой развития киберпреступности. Сегодня киберпреступники используют не только средства угрозы компьютерной информации, такие как компьютерные вирусы и сетевые атаки, но и рассылку спама [1]. Спамом принято называть — массовую рассылку писем любого вида корреспондентам, не желающим их получать. К спаму причисляют рассылки коммерческой, политической рекламы и антирекламы, а также мошеннические письма и письма, содержащие в себе угрозу информации, например вирусы. Спам приобрёл широкое распространение в системах электронной почты, т.к. стоимость отправки электронных писем крайне низкая, в отличие от обычной бумажной корреспонденции. Например, за один день, по некоторым данным рассылается более 55 миллиардов писем [2], классифицированных как спам.

На сегодняшний день разработано много методов обнаружения и блокировки спама, однако это не спасет ситуацию. Киберпреступники постоянно разрабатывают новые методы обхода существующей защиты, и пользователи по всему миру продолжают получать огромное количество нежелательной почты каждый день.

В данной статье рассмотрен нейросетевой подход для классификации электронной почты. В первом разделе статьи представлен обзор наиболее популярных методов обнаружения электронного спама. Второй раздел содержит описание нейросетевого подхода классификации электронного спама. В третьем разделе представлены результаты проведенных экспериментов, выявлены недостатки и преимущества разработанного метода.

1. Эволюция методов борьбы с нежелательной корреспонденцией. Задача любой системы фильтрации спама — разделение потока входящей корреспонденции на спам и нормальную почту. Эта задача может решаться на разных этапах доставки писем: на почтовом сервере или на компьютере клиента.

Для анализа письма годятся любые его признаки, но на практике чаще всего используют следующие:

- IP-адрес отправителя;
- e-mail адрес отправителя и адрес для ответа;
- тема и текст сообщения;
- совокупность прочих признаков, таких как почтовый клиент, время отправки, отметки промежуточных узлов доставки и др.

Рассмотрим подробнее каждый из перечисленных методов.

Фильтрация по IP-адресам. Первые методы фильтрации спама были весьма простыми. Например, ведение «чёрных списков» IP-адресов, откуда запрещена доставка корреспонденции [1]. Т.е. на почтовом сервере проверяются все адресаты, от которых идет корреспонденция. Если адрес отправителя находится в «черном списке», то такое письмо блокируется и не доходит до конечного адресата. Составлением «черного списка» может заниматься как отдельная всемирная организация, так и создатели почтового сервера. К плюсам данного подхода можно отнести надёжность. Надёжность блокирования нежелательных писем обуславливается определением на основе статистики «мест» распространения спама, и в итоге письма блокируются еще до доставки конечному пользователю. К недостаткам данного метода можно отнести слабую эффективность в настоящий момент. Для отправки спама сейчас широко используются бесплатные открытые почтовые службы, а также заражённые вредоносными программами компьютеры обычных пользователей, так называемые зомби-сети [3]. Таким образом, блокировка IP-адресов, с которых рассылается большой поток нежелательных писем, скажется на обычных пользователях, не подозревающих, что их компьютеры заражены.

Таким образом, способ с использованием «белого списка», антипод «черного списка», годится лишь в исключительных случаях закрытых сообществ, например, для организации корпоративной почты.

Фильтрация по ключевым словам. Тема письма и его текст являются важным признаком для анализа спама. Как правило, текст спама отличается от текста нормальной корреспонденции, так как спам несет в себе отличную от простой переписки функцию: заманивание пользователей посетить Интернет-сайт или купить какую-нибудь вещь, оглашение коммерческой или политической информации и др. Фильтрация по ключевым словам и фразам является еще одним методом обнаружения спама [1]. Метод имеет ряд преимуществ, связанных с простотой его реализации — достаточно занести в базу данных запрещенные ключевые слова или фразы и в дальнейшем проверять наличие этих ключевых слов в тексте письма. Однако метод имеет и ряд существенных недостатков:

- правильно настроить подходящие ключевые слова достаточно

Безобразов Сергей Валерьевич, старший преподаватель кафедры «Интеллектуальные информационные технологии» Брестского государственного технического университета.

Беларусь, БрГТУ, 224017, г. Брест, ул. Московская, 267.

Меленчук Василий Васильевич, инженер-программист представительства компании «СIB Software GmbH» (ФРГ) в Республике Беларусь.

Физика, математика, информатика