

УДК 007 : 519.712.2

**И. Б. Гуревич, Ю. И. Журавлев, Д. М. Мурашов,
Ю. Г. Сметанин, А. В. Хилков**

(Москва)

**СИСТЕМА АВТОМАТИЗАЦИИ НАУЧНЫХ ИССЛЕДОВАНИЙ
В ОБЛАСТИ АНАЛИЗА И ПОНИМАНИЯ ИЗОБРАЖЕНИЙ
НА ОСНОВЕ НАКОПЛЕНИЯ И ИСПОЛЬЗОВАНИЯ ЗНАНИЙ. Ч. I***

Описывается система «Черный квадрат. Версия 1.0» – инструментально-программный комплекс для автоматизации научных исследований и обучения в области обработки, анализа, распознавания и понимания изображений. Система представляет собой среду для разработки алгоритмов обработки, анализа, распознавания и понимания изображений со справочными и информационно-поисковыми функциями. Система включает базу данных, содержащую алгоритмы обработки, анализа, распознавания и понимания изображений, инструментальные средства для их выбора, применения и объединения, комбинирования в многозадачном режиме; базу знаний, содержащую стандартизированные блоки задач, средства для визуализации и редактирования используемых данных, электронный многоуровневый справочник по системе, архив тестовых и обработанных изображений и оболочку, связывающую все компоненты системы в единый инструментально-программный комплекс. Основными объектами хранения являются алгоритмы, структуры данных – классы объектов, необходимые для анализа и оценивания изображений, и изображения. Функционирование системы обеспечивается единой интегрированной программной средой, построенной на общей информационной и методической базе и обладающей общей навигационной оболочкой. Функциональные особенности системы: открытость по обрабатываемым системой объектам, библиотекам алгоритмов, наполнению лексикографической и экспертной информацией; стандартная диалоговая оболочка и инструментальная среда разработчика; функциональная полнота алгоритмического наполнения для задач из области компетентности системы; наличие набора баз данных, обеспечивающих полное сопровождение всех решаемых задач; возможность модернизации системы посредством включения новых и исключения устаревших модулей (функционально различные части системы размещены в различных модулях, имеющих легко поддерживаемые связи с остальными частями системы). Использование системы «Черный квадрат. Версия 1.0» и ее последующих версий позволяет стандартизировать и резко увеличивать эффективность разработки специализированных систем обработки, анализа и распознавания изображений, отработки методов создания и передачи технологий извлечения информации из изображений, отработки стандартных решений для систем анализа и оценивания изображений, а также качество решений и информационных технологий, закладываемых в эти системы.

* Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (проекты № 97-07-90334, № 98-07-90180, № 96-15-96085).

Введение. Данная работа посвящена описанию системы «Черный квадрат. Версия 1.0» – инструментально-программного комплекса для автоматизации научных исследований и обучения в области обработки, анализа, распознавания и понимания изображений.

Эта система является последней в серии инструментально-программных комплексов, разработанных на протяжении 1990-х годов в лаборатории «Кибернетические методы в информатике» Научного совета по комплексной проблеме «Кибернетика» Российской академии наук [1–8].

Система представляет собой среду для разработки алгоритмов обработки, анализа, распознавания и понимания изображений со справочными и информационно-поисковыми функциями. Она включает: 1) базу данных, содержащую алгоритмы обработки, анализа, распознавания изображений; 2) инструментальную среду для их выбора, применения и комбинирования и слияния в многозадачном режиме; 3) базу знаний, содержащую блоки задач, средства для визуализации и редактирования используемых данных; 4) электронный многоуровневый справочник по системе; 5) архив тестовых и обработанных изображений.

Функционирование системы обеспечивается единой интегрированной программной средой, построенной на общей информационной и методической базе и обладающей общей навигационной оболочкой.

Минимально необходимые ресурсы: операционная система Windows 95/98/NT, оперативная память 16 Мбайт, память на жестком диске 20 Мбайт.

Система предназначена: а) для автоматизации (разработки, исследования и применения алгоритмов обработки, анализа, распознавания и понимания изображений; обучения методам обработки и анализа изображений); б) для хранения в хорошо структурированном и удобном для пользователя виде информации о методах обработки, анализа, распознавания и понимания изображений; об алгоритмах и самих алгоритмах обработки, анализа, распознавания и понимания изображений; об изображениях и самих изображениях.

Основные области применения: а) автоматизация научных исследований и разработок; б) разработка информационных технологий; в) решение прикладных задач (медицинская диагностика, экологический мониторинг, неразрушающий контроль, техническая диагностика, прогноз и мониторинг чрезвычайных ситуаций).

Инструментально-программные комплексы разрабатывались как инструмент для поддержки исследований, посвященных автоматизации извлечения информации из изображений, что включает автоматизацию выбора метода решения предъявленной задачи, выдачи рекомендаций по выбору метода решения для класса задач, к которому относится предъявленная задача, синтеза алгоритмических процедур решения данной задачи, выдачи рекомендаций по синтезу алгоритмических процедур ее решения.

В процессе наших исследований довольно быстро выяснилось, что автоматизация извлечения информации из изображений предполагает наличие в распоряжении исследователей и пользователей, обладающих разной квалификацией, не только стандартизированной технологии автоматизации, но и поддерживающей ее системы, накапливающей и использующей знания по анализу и оцениванию изображений и обеспечивающей адекватные структурные и функциональные возможности для поддержки «интеллектуализации» выбора и синтеза методов и алгоритмов обработки и анализа изображений.

1. Эволюция концепции автоматизированных систем для обработки и анализа изображений. Разработка автоматизированных систем для обработки и анализа изображений ведется во всем мире со все возрастающей интенсивностью начиная с конца 1970-х годов. Развитие этих систем шло и продолжает идти по нескольким направлениям, часть которых является естественным продолжением предшествующих, часть существует параллельно.

Первоначально это были довольно узкоспециализированные автоматизированные системы обработки изображений (АСОИЗ) (другое название – системы машинного зрения). Общепринято определение АСОИЗ как системы, позволяющей работать с одним или несколькими изображениями некоторого объекта, обрабатывать эти изображения, измерять и анализировать характеристики обрабатываемых изображений и интерпретировать результаты измерений с тем, чтобы обеспечить возможность принятия полезного в некотором смысле решения относительно объекта, воспроизведенного на изображении. В функции АСОИЗ входят отыскание объекта, контроль качества объекта, измерение характеристик объекта, идентификация, распознавание, подсчет и оценка характеристик движущихся объектов [9].

В системах АСОИЗ, предназначенных для решения задач в узкой прикладной области и позволяющих работать с изображениями определенного класса (системы медицинской и технической диагностики, неразрушающего контроля, обработки печатных документов и т. д.), последовательность этапов обработки изображений и применяемые на каждом этапе методы выбираются разработчиком системы. В них используется малое количество алгоритмов обработки и анализа изображений, причем применяются они в строго определенной последовательности.

В процессе развития АСОИЗ возникла концепция систем машинного зрения, основанных на использовании знаний. Такие системы включают подсистемы обработки изображений и вывода и базу знаний. Подсистема обработки изображений обеспечивает выделение из изображения непроецируемых элементов (точек, отрезков, кривых) и информации о его фрагментах. Подсистема вывода обычно основана на правилах. Ее функцией является объединение информации, получаемой подсистемой обработки изображений, с информацией, содержащейся в базе знаний, и выработка решений о том, какую гипотезу следует выдвинуть, какую проверить, какую новую информацию можно получить (вывести) из того, что уже было установлено, и какие еще непроецируемые элементы должна выделить подсистема обработки изображений. В базе знаний содержится информация о физической реальности, изображение которой обрабатывается и анализируется [9].

Появление следующего поколения АСОИЗ связано с концепцией понимания изображений, трактуемой как комбинация методов обработки изображений, распознавания образов и искусственного интеллекта, которая используется для выполнения на ЭВМ анализа изображений, полученных от одно- или многоканального датчика либо временной последовательности изображений. Этот анализ включает обнаружение, определение местоположения и ориентации вызывающих интерес объектов трехмерной реальной среды, воспроизведенных на изображении, и обеспечивает получение достаточно подробных символьных описаний или результата распознавания таких объектов. Понимание изображений часто предусматривает использование геометрического моделирования при сопоставлении объектов на основе моделей и при организации поиска. Последний может быть основан на исполь-

зовании стратегий управления типа «сверху вниз», «снизу вверх», «доска объявлений», а также иерархических и гетероархических [9].

Реализации методологии понимания изображений в полном объеме препятствовали два обстоятельства: значительная стоимость разработки программного обеспечения, составляющая до 95 % стоимости всей разработки, и недостаток персонала, обладающего квалификацией, необходимой для решения прикладных задач обработки и анализа изображений. Аналогичная ситуация имела место в свое время в связи с внедрением экспертных систем и преодолена благодаря разработке оболочек экспертных систем. Создание программного обеспечения для понимания изображений с функциями, аналогичными функциям оболочек экспертных систем, позволяет как ускорить разработку систем понимания изображений, так и разработать прикладные программы для персонала с невысокой квалификацией в области понимания изображений. Хотя эта проблема существует уже более 10 лет, сегодня еще довольно далеко до понимания изображений, сопоставимых по возможностям по отношению к пользователю с существующими оболочками экспертных систем [10].

Стандартные оболочки экспертных систем непригодны для разработки прикладных программ понимания изображений.

Во-первых, операции над изображениями отличаются сложностью и требуют значительных вычислительных затрат. Их реализация не ограничивается лишь заданием правил и фреймов, что является стандартным назначением оболочек. В частности, проверка условий применимости правил (или иных структур представления знаний) при работе с изображениями не может быть осуществлена в существующих оболочках в силу вычислительной сложности, поскольку изображение обычно представляется сотнями тысяч значений отдельных пикселей. Для реализации методов, основанных на использовании правил, необходимо предусмотреть оперативную память второго уровня для хранения характеристик изображения, полученных в результате применения к нему соответствующих операций обработки. В этом случае структуры представления знаний могли бы использовать промежуточные результаты обработки изображения для управления процессом его обработки и извлечения из него информации.

Во-вторых, даже в тех случаях, когда оболочка обеспечивает доступ к некоторым интегральным характеристикам изображения, полученным с помощью какой-либо библиотеки функций обработки изображений, неспециалист в области обработки изображений не сможет разрабатывать прикладные программы, поскольку не обладает опытом, необходимым для определения правил и фреймов.

Все это вкупе означает, что оболочки для понимания изображений должны, помимо механизмов, обеспечивающих возможность включения структур знаний, быть оснащены знанием того, какие операции следует применять к изображениям в различных случаях. Это усугубляется тем обстоятельством, что в обработке изображений и распознавании образов использование эффективных алгоритмических процедур и тщательная разработка методов сокращения перебора являются традиционными средствами компенсации недостатка вычислительных ресурсов.

Таким образом, система понимания изображений должна включать две ключевые подсистемы. Одна из них предназначается для знаний общего характера (универсальных, не зависящих от прикладной области), другая – для получения и использования знаний, относящихся к специфической приклад-

ной области. В этой связи возникло разделение на системы понимания изображений и экспертные системы обработки изображений. Первые обладают знаниями о прикладной области, роль же вторых ограничивается использованием всех известных методов обработки изображений для выделения информации из некоторой сцены, причем осуществляется это, как правило, под управлением системы понимания изображений [11].

Считается, что современная система понимания изображений должна включать следующие модули знаний и средства для их получения и обработки [10]:

1. Модуль знаний о решении задач с помощью программных систем. Это наименее специализированный модуль, содержащий наиболее общие сведения о решении задач с помощью программного обеспечения ЭВМ.

2. Модули общих знаний по обработке изображений. Они должны обеспечивать упрощение разработки прикладных программ в различных прикладных областях в следующих отношениях: а) планирование операций обработки; б) управление выполнением обработки в процессе анализа (обработка методом «сверху вниз» или «снизу вверх», параллельная или последовательная обработка, распределенное или централизованное управление); в) преобразование численных результатов анализа в символьную форму; г) представление и использование знаний об обработке изображений, в том числе знаний, не зависящих от предметной области, и знаний о сегментации, основанной на процедурах распознавания образов; д) реализация методов оценивания качества процесса обработки изображений; е) распознавание объектов на основе их сопоставления и выбор из конкурирующих гипотез.

3. Модуль знаний об используемой библиотеке обработки изображений. Он содержит отображения, связывающие внутренние представления операций обработки изображений в модуле общих знаний по обработке изображений и конкретные функции из библиотеки обработки изображений, выполняющие эти операции; имеет доступ к базе данных, содержащей промежуточные результаты обработки.

4. Модуль знаний о реальном мире.

5. Модуль знаний о прикладной области. Эти знания тесно связаны со знаниями о способах получения изображений, поскольку в прикладной области объекты часто характеризуются с помощью процедур, используемых для их обнаружения на изображении.

6. Интерфейс пользователя. Он содержит знания о способах взаимодействия системы с пользователем; его основные функции – обеспечить пользователю возможность задать цели анализа изображения и использовать знания о прикладной области.

7. Модуль автоматизации обучения. Основное содержание связано с использованием методов распознавания образов в обработке, анализе и понимании изображений.

Одной из наиболее известных систем для разработки прикладных программ обработки и визуализации изображений является инструментально-программная среда "Khoros" [12]. Ее разработка была начата в университете Нью-Мехико в 1987 г. с целью создания средства поддержки исследований в области обработки изображений, однако впоследствии среда "Khoros" стала широко использоваться для автоматизации научных исследований в различных областях, в частности, в распознавании образов, геоинформационных системах, обработке медицинских изображений и распределенной обработке информации.

Библиотека системы включает процедуры для многих прикладных задач обработки изображений. Эти процедуры используются как автономные программы или как библиотечные процедуры, к которым пользователь может обращаться при разработке новой программы.

Библиотечные процедуры составляют три основных класса: арифметические операции, обработка изображений и анализ изображений. В класс арифметических операций входят все арифметические операции, применяемые к пикселям, и логические операции. Вторым классом образуют основные операции обработки изображений – фильтрация, геометрические операции и преобразования. Класс анализа изображений включает процедуры сегментации, выделения признаков и классификации.

Среда оснащена визуальным языком программирования Cantata, блоками генерации программ для расширения визуального языка и включения в систему новых пакетов прикладных программ, диалоговым редактором пользователя, программами диалогового воспроизведения изображения, средствами визуализации, обширной библиотекой, включающей свыше 260 процедур обработки изображений, численного анализа, обработки сигналов и пакетами вывода двух- и трехмерной графики. При работе в сети "Khoros" она обеспечивает возможность распределенных вычислений и позволяет многим пользователям одновременно использовать одни и те же изображения и рабочую область памяти. В ряде университетов система применяется в качестве средства обучения в области обработки сигналов и изображений.

Важным отличием и преимуществом системы относительно других средств разработки прикладных программ является использование единой рабочей области памяти для управления потоком данных и исполнением программных модулей. В системе применяется единый формат данных для всех модулей, что упрощает обмен данными между процедурами системы и разработку новых модулей.

Одна из основных задач проекта "Khoros" – объединение двух относительно независимых парадигм визуального программирования: непосредственной разработки графического интерфейса пользователя и визуального программирования на основе концепции потока данных. Подобное объединение позволяет расширить возможности программной среды и обеспечить пользователю разработку программы, не используя текстовый код (например, С-код).

Ключевой проблемой разработки и использования систем понимания изображений является стандартизация. Исследователи пользуются множеством различных библиотек обработки изображений и вводят собственные представления знаний о получении изображений и промежуточных результатах обработки. Возникает необходимость перехода к общепринятым вариантам имен, аргументов, результатов выполнения элементарных операций обработки изображений, представления и методов доступа к промежуточным результатам обработки изображения.

К началу девяностых годов ряд исследовательских групп разработали, действуя независимо друг от друга, системы, отвечающие целям и задачам их направлений исследований. Далее эти системы продолжали совершенствоваться силами разработавших их организаций с тем, чтобы удовлетворять постоянно возрастающие требования к обработке информации. Эта практика существенно затронула область машинного зрения и понимания изображений в целом и особенно в следующих отношениях: а) исследования, выполненные в одной организации, трудно воспроизводить за их пределами

из-за несовместимости программного обеспечения и структур данных; б) средства, обеспечивающие стандартную реализацию общепринятых алгоритмов обработки изображений, нельзя использовать во многих исследовательских центрах и для передачи исследовательских технологий разработчикам приложений.

Значительный вклад в разрешение проблемы стандартизации вносит крупный исследовательский проект «Программная среда для понимания изображений» (СПИ), выполняемый с 1990 г. по инициативе и при финансовой поддержке Агентства перспективных исследовательских проектов (ARPA), США. Основной целью проекта является увеличение продуктивности научных исследований в области обработки, анализа, распознавания и понимания изображений при помощи стандартной платформы для поддержки прикладных программ и инструментальных средств (языки C и LISP), разрабатываемых по заказам ARPA, главным образом, поддержка обмена результатами исследований в области понимания изображений, тестирование и сравнительный анализ качества алгоритмов [13]. Для этого в рамках проекта разрабатывается среда, включающая хорошо документированные, модульные, стандартизированные объектно-ориентированные интерфейсы. Эта среда обеспечивает реализацию стандартных алгоритмов понимания изображений и позволяет объединять возможности и вычислительные ресурсы существующих инструментально-программных средств и систем понимания изображений.

Структуры данных и операции СПИ обеспечивают основу для формального определения алгоритмов понимания изображений. СПИ задает непротиворечивый стандартизированный формат для программирования алгоритмов понимания изображений, обучения и развития технологий понимания изображений. СПИ может также использоваться в качестве вычислительной модели при разработке новых архитектур для систем понимания изображений, поскольку ее спецификация определяет базовые структуры данных и операции, которые должны поддерживаться любой архитектурой, специализированной для задач понимания изображений. СПИ можно также рассматривать в качестве языкового интерфейса для любой интересной архитектуры понимания изображений.

Одной из целей разработки СПИ являлась поддержка создания приложений. К числу основных прикладных направлений, на которые ориентировалась разработка СПИ, относятся следующие.

1. Интерпретация аэрофотоснимков. Поддержка анализа больших аэрофотоснимков, полученных с помощью различных датчиков, в том числе электрооптических, радиолокаторов с синтезированной апертурой и многоспектральных изображений. Основные задачи: регистрация и объединение изображений, распознавание объектов, выявление изменений и измерения.

2. Интеллектуальное оружие. Применение методов понимания изображений в наведении ракет (алгоритмы идентификации целей и алгоритмы слежения). Основные прикладные задачи: обработка и анализ последовательностей изображений, применение методов активного зрения и распознавание объектов.

3. Картография. Разработка инструментально-программных средств для создания картографических баз данных по изображениям. Основные прикладные задачи: построение и анализ стереоизображений, модели камер и автоматизация выделения признаков изображения.

4. Навигация на основе использования видеоинформации. Поддержка разработки наземных аппаратов-разведчиков. Основные прикладные задачи: слежение за дорогой, анализ поверхности, обход препятствий и распознавание объектов.

5. Техническое зрение. Основные прикладные задачи: разработка датчиков-дальномеров, анализ дальностной информации, сегментация изображений, автоматизация построения моделей и обратная связь на основе видеоинформации.

В большинстве этих приложений система должна иметь возможность оперировать большими изображениями (более чем 10×10 К пикселей). СПИ обеспечивает доступ к таким изображениям в непрерывной системе координат изображения и поддерживает их визуализацию в диалоговом режиме, причем число изображений, используемых в каждом процессе обработки, не должно ограничиваться.

СПИ поддерживает работу со многими типами видеоданных, включая длинные и короткие целые числа, числа с плавающей запятой, комплексные числа, цветные (*RGB*) изображения и многоспектральные изображения. Поддерживается также работа с последовательностями изображений произвольной длины и обеспечивается мультипликационное воспроизведение таких последовательностей. СПИ включает представления для следующих типов видеодатчиков: камеры с точечной диафрагмой, камеры с тонким объективом, стереоскопические камеры, спутниковые камеры и дальномеры. Поддерживается объемное представление трехмерных моделей и их представление поверхностными границами. СПИ позволяет импортировать и экспортировать описания трехмерных объектов в формате IGES.

СПИ основана на объектно-ориентированном подходе. При разработке в максимальной мере использовались готовые компоненты рабочих станций. СПИ работает под управлением ОС UNIX, интерфейсы пользователя выполнены в стандарте X Windows. Предусматривается использование коммерческих интерфейсных пакетов разработчика, программных сред и средств отладки. СПИ поддерживает языки LISP и C++ посредством параллельных классов иерархии объектов и специального механизма для связи между обеими языковыми средами.

В СПИ широко используется взаимодействие на графическом уровне для поддержки работы с признаками и обеспечения удобных средств для построения моделей и распознавания. Эти средства основаны на единой методологии построения интерфейса пользователя и обеспечивают удобный выбор и модификацию графических объектов.

Основные подсистемы СПИ снабжены стандартными интерфейсными протоколами, что позволяет развивать систему независимо от изменений аппаратной части.

СПИ должна отвечать следующим требованиям.

1. Совместимость по данным (изображениям). Обеспечение считывания и записи различных внешних форматов представления изображений помимо собственных внутренних блоков данных.

2. Совместимость представлений символьных данных. Надежное обеспечение отображения внешних представлений символьных конструкций (области, линии и т. д.) в иерархию классов СПИ и обратно.

3. Единая модель процесса обработки. Обеспечение модульного представления алгоритмов с тем, чтобы существующие программные модули

могли использоваться как «черные ящики», вызываемые с конкретными входными и выходными параметрами.

4. Стандартный интерфейс управления процессом обработки. Обеспечение поддержки стандартных интерфейсов для вызова процессов понимания изображений как на программном уровне, так и при использовании существующих СВП. Эти управляющие интерфейсы позволяют пользователю объединять и сопоставлять доступные наборы алгоритмов и инструментально-программные средства.

Разработка СПИ разделена на три этапа.

На первом этапе разработана предварительная версия системы обмена данными СПИ, включающая стандарт организации файлов и библиотеку программиста, которая позволяет пользователям других систем считывать и записывать файлы обмена данными. Также реализован базовый вариант системы, используемый в качестве среды разработки полного варианта СПИ и средства оценки и тестирования решений, закладываемых в СПИ.

На втором этапе реализовано ядро СПИ, предназначенное для разработки сложных алгоритмов понимания изображений. Ядро включает расширенную иерархию классов СПИ, средства графического интерфейса, допускающие расширение пользователем, ряд стандартных средств обработки, дополнительные возможности для обмена данными и механизмы управления для использования сложных алгоритмов в исследовательских задачах.

На третьем этапе разработаны библиотеки СПИ.

СПИ представляет собой открытую структуру, которая может быть специализирована применительно к исследовательским задачам пользователя без потери полной совместимости и возможности взаимодействия со всеми остальными составляющими иерархии классов СПИ. Последняя в значительной мере основывается на математических понятиях, в частности на понятии множества. Именно из этого понятия непосредственно выводятся такие классы СПИ, как отношения, упорядоченные отношения, функции и бинарные отношения. Через эти классы определяются традиционные для программирования структуры данных – массивы, последовательности, матрицы, списки, деревья и графы. Большая часть остальных классов СПИ выводится из этих базовых классов и имеет их свойства.

К основным разделам иерархии классов СПИ относятся пространственные объекты, системы координат и преобразования, изображения, признаки изображений и датчики.

Пространственные объекты являются одним из центральных понятий иерархии классов СПИ. Пространственный объект есть некоторое множество точек, заданное в n -мерном пространстве. Этот объект связан с некоторой системой координат, что обеспечивает возможность применять к нему преобразования и строить проекции. Произвольные сложные объекты могут быть композициями пространственных объектов. Иерархия пространственных объектов включает геометрические и топологические пространственные объекты, представляющие соответственно метрические свойства объектов и характеристики связности объекта или множества объектов независимо от их метрических свойств.

В СПИ геометрические отношения между физическими объектами, датчиками и сценами, пикселями и реальным миром представляются с помощью систем координат и координатных преобразований. При этом каждый геометрический объект связан явным образом с некоторой системой координат, различные объекты могут задаваться в одной и той же либо в разных

системах координат. Координатные преобразования содержат информацию, необходимую для перевода некоторой точки из одной системы координат в другую.

Классы изображений СПИ поддерживают эффективную обработку многих разновидностей видеоинформации: тоновых и цветных изображений, строк, последовательностей, пирамид. Изображения большого размера (более чем $1 \times 1\text{К}$) поддерживаются с помощью механизма страничной организации (значениями пикселей могут быть обычные скалярные величины или многозначные величины, например, *RGB*-изображения или сложные изображения). В СПИ поддерживается также обработка композиционных изображений – наборов простых изображений, пиксели которых могут быть не зарегистрированы. Гибкие классы интерфейсов обеспечивают эффективный доступ, не зависящий от организации данных, к специализированным классам, оптимизированным применительно к соответствующему формату хранения.

Классы признаков изображений представляют объекты, обнаруживаемые на изображении или из него выделяемые. В силу важности метрических характеристик и характеристик связности часть этих классов соответствует геометрическим, а часть – топологическим пространственным объектам. Поскольку классы порождаются изображением, то они содержат дополнительные характеристики, связывающие их с соответствующим исходным изображением и отражающие свойства изображения как сигнала. Одно-, двух- и трехмерные объекты изображения могут быть представлены соответствующими множествами классов признаков изображения. Характеристики признаков изображений позволяют представлять уровни неопределенности, соответствующие специфическим признакам, отношения между признаками.

Собственно датчики и информация, воспроизводимая ими, представляются в СПИ двумя группами классов иерархии. В первой группе классы определяют физические механизмы воспроизведения информации и включают один или несколько датчиков. Классы второй группы имитируют процесс получения информации и включают указатели для сцены, датчика и получаемых данных. Остальные классы представляют информацию, характеризующую фильтры, объективы, излучение и собственно сцену.

Для поддержки многократного использования алгоритмов, разработанных с применением иерархии классов, в СПИ введено понятие задания – пакета программ, предназначенного для решения задач понимания изображений большого или среднего объема (например, реализация алгоритма выделения яркостного перехода либо вызываемая процедура, поддерживающая визуализацию и редактирование пользователем пространственных объектов).

Задания располагают хорошо определенным интерфейсом, описывающим требования к входной и выходной информации заданий в терминах классов и типов СПИ. Интерфейсы программиста для заданий СПИ должны включать: СВП, позволяющие разрабатывать программу в интерактивном режиме посредством комбинирования существующих заданий и быстро оценивать характеристики разрабатываемых алгоритмов; непосредственное программирование на уровне исходного текста программы с тем, чтобы обеспечить максимальную эффективность ее исполнения.

Поскольку в настоящее время в понимании изображений используются различные СВП, задания СПИ определяются таким образом, чтобы с ними

можно было работать в различных средах. Это обеспечивает следующие преимущества:

- расширение набора средств, которыми можно пользоваться при разработке и исполнении заданий СПИ; при соответствующем объединении заданий СПИ все возможности базовой среды могут использоваться для программирования задания СПИ;

- расширение набора операций, которыми можно пользоваться при решении задач понимания изображений; при наличии соответствующих средств поддержки обмена данными задания СПИ можно комбинировать с процедурами, заложенными в базовые среды.

СПИ поддерживает обмен данными между различными системами понимания изображений, в том числе между СПИ и другими системами. Этот обмен обеспечивается на уровнях: а) абстрактного представления стандартизацией объектов понимания изображений, подлежащих обмену, и минимизацией числа компонент, необходимых для определения этих объектов; б) физического представления стандартизацией кодирования абстрактного представления; в) программной поддержки обеспечением передачи данных системы в представление и от представления.

Минимальное представление объектов понимания изображений расширяется с помощью дополнительных признаков. Для того чтобы не возникала несовместимость различных систем, СОД предусматривает имена, типы и значения для таких дополнительных признаков.

Состояние объекта понимания изображений представляется не только его типом, но и отношениями с другими объектами понимания изображений. СОД предусматривает понятие такого отношения и допускает представление подобных отношений между объектами.

Поддержка обмена данными на программном уровне включает два этапа, причем оба могут реализовываться как при использовании СПИ, так и вне ее. Первый этап предусматривает применение интерфейса для разработки прикладных программ для считывания и записи таблиц, представляющих типы и объекты в системе. Этот механизм может быть использован любой объектно-ориентированной программой для операций ввода-вывода при обмене данными. Использование этого интерфейса на втором этапе обеспечивает реализацию объектов вне СПИ (СПИ поддерживает реализацию вспомогательных процедур для создания собственных объектов в своих системах).

Для реального осуществления передачи данных между модулями, использующими различные представления данных, необходимо определить конкретные отображения между совмещаемыми представлениями, т. е. механизмы, при помощи которых обеспечивается точность считывания и записи при преобразовании данных в форматах СПИ в форматы, не используемые в СПИ. Необходимо также определить, могут ли существующие представления нести дополнительную информацию, не используемую в СПИ, но позволяющую восстановить входные данные по выходным.

Отображение данных, представляющих пространственные объекты и признаки изображений, из одной системы в другую может вызывать затруднения из-за использования различных эквивалентных представлений данных.

Для преодоления этих трудностей предусмотрено использование СПИ-спецификации обмена данными в качестве некоторого канонического представления. При этом пользователи других систем должны сами обеспечивать преобразование внутренних представлений своей системы в пред-

ставления, предусмотренные СОД СПИ, если им необходимо переносить свои данные в другие системы. Использование такого механизма исключает необходимость «свертки» схем преобразования данных, которую пришлось бы выполнять, если бы требовалось для каждой системы отдельно определять собственные преобразования для преобразования данных, поступающих в нее из любой другой системы или передаваемых из нее в любую другую систему.

Известно, что для представления видеоданных применяется более ста различных форматов файлов. Изображения, используемые в научных исследованиях, не являются просто массивами скалярных величин: они часто содержат дополнительные данные типа указателей на датчики, калибровочные коэффициенты, сцену. Спецификация обмена данными предусматривает передачу такого рода данных на основе разделения представления собственно изображения и данных об изображении. СОД СПИ определяет, что представление изображения содержит соответствующие дополнительные поля вкупе с данными изображения, которое хранится в виде массива или как указатель на файл данных одного из общепринятых форматов (KBVision, TIFF, GIF, СПИ-формат).

Совместимость на уровне данных для множеств, списков и последовательностей в понимании изображений означает, что два различных представления набора данных должны быть эквивалентны семантически и не обязательно должны иметь одинаковые реализации. Этот вид отображения поддерживается спецификацией обмена данными на основе спецификаций для абстрактных классов. Описанный набор механизмов обеспечения совместимости на уровне данных поддерживает основные функции интегрированной СПИ.

Необходимость использования заданий СПИ как в средах визуального программирования, так и на уровне исходных текстов требует, чтобы спецификация задания СПИ обеспечивала следующие два формата: спецификацию характеристик задания, соответствующих требованиям СВП, в том числе перечень входных и выходных параметров, включая их типы, значения, используемые по умолчанию, и ограничения (для всех характеристик должна предоставляться документация); стандартный прототип языкового уровня функции задания, обеспечивающий возможность вызывать функцию непосредственно в исходном тексте программы.

Естественно, что максимальную пользу исследователям дает возможность использовать задания СПИ совместно с заданиями соответствующих систем в максимально возможном количестве СВП. Необходимо также предусмотреть механизмы, позволяющие применять в некоторой СВП не только задания СПИ, но и задания из как можно большего числа разновидностей других систем.

В этом случае интерфейсы этих систем должны обладать совместимостью на уровне обработки с соответствующей СВП. Для этого на верхнем уровне задания предусматривается согласующий механизм, обеспечивающий задание интерфейсом, согласованным с базовой средой. Реализация такого согласующего механизма зависит как от типа интерфейса базовой среды, так и от вида задания. Опыт использования заданий СПИ в системе "Khoros" показывает, что практически полностью процесс адаптации задания для использования в какой-либо среде может быть автоматизирован.

В системе "Khoros" требуется, чтобы задание выполнялось в СВП Cantata и под управлением интерфейса с командной строкой. Необходимость

этих двух режимов определяется возможностью доступа к описанию задания, находящемуся в файле спецификации интерфейса пользователя. В системе "Khoros" имеются средства поддержки пользователя, позволяющие автоматически создавать такие файлы спецификации интерфейса на основе резидентного компоновщика графического интерфейса пользователя.

При этом файл спецификации интерфейса обрабатывается автоматически с помощью других средств для построения исходного текста-заглушки. После того как пользователь добавляет реализацию задания к этой программе-заглушке, задание становится доступным для использования.

Подобным же образом задания системы KBVision могут быть включены в интерфейс среды Cantata, а задания СПИ и системы "Khoros" – в СВП KBVision. В настоящее время уровень совместного использования заданий определяется в основном совместимостью данных.

Важнейшей причиной несовместимости данных является использование во взаимодействующих системах несовместимых наборов признаков для представления объектов, которые на абстрактном уровне выглядят для пользователя как одни и те же данные. Так, например, кроме данных, характеризующих пиксели, изображения СПИ содержат ссылки на систему координат. В системе "Khoros" модель изображения не включает понятие «системы координат», однако может включать характеристики, не определенные для СПИ-изображения.

Спецификация СОД задает на уровне, абстрагированном от реализации, подобъекты, которые должны присутствовать (для организации обмена данными) в определении объекта в задачах понимания изображений. Если разработчики СВП будут предусматривать в своих системах поддержку любого объекта задач понимания изображений, удовлетворяющего требованиям спецификации СОД, возможности совместного использования различных типов заданий в любой СВП будут существенно увеличены. При этом предполагается, что если система использует некоторый объект задач понимания изображений и если этот объект определяется СОД, то данная система должна обеспечить поддержку СОД для этого объекта.

Менее существенна совместимость представлений данных на физическом уровне, используемых для передачи объектов между различными заданиями. В разных системах обмен объектами задач понимания изображений может осуществляться с использованием различных физических форматов. В то же время согласование со спецификацией СОД обеспечивает независимо от выбранного физического формата значимость информации, содержащейся в данных, для заданий всех систем понимания изображений, соответствующих требованиям СОД.

В рамках этого стандарта «собственные» задания базовой СВП смогут обмениваться данными с использованием естественного физического представления этой среды. Если же необходимо передать данные в «чужие» задания, то они могут экспортироваться в стандартное физическое представление СОД. В этом случае любой тип задания сможет импортировать и интерпретировать эти данные.

Для того чтобы этот механизм работал при использовании заданий смешанных типов, базовая СВП должна играть более активную роль в обмене данными между заданиями, выполнением которых она управляет. Необходимо, чтобы некоторая система (очевидно, СВП) либо обеспечивала преобразование представлений данных в стандартную форму, либо поддерживала «переговоры» между заданиями относительно физического представления, ко-

торое будет использовано при обмене данными в задачах понимания изображений. В обоих случаях базовая СВП должна поддерживать при необходимости передачу данных задач понимания изображений в стандартном физическом представлении СОД.

Существенным вкладом в стандартизацию алгоритмического наполнения систем понимания изображений должно явиться использование результатов алгебраической теории алгоритмов Журавлева, предложившего общую модель алгоритма распознавания и регулярный способ построения комбинированных алгоритмов – полинома над алгоритмами [14, 15].

В значительной мере попытки создания формального аппарата, обеспечивающего единообразное, компактное и эффективное представление процедур обработки и анализа изображений, связаны с разработкой алгебр изображений, обеспечивающих создание языка для стандартного описания преобразований изображений; описание операций над изображениями в виде компактных наборов простых преобразований; описание действий над изображениями как в машинно-независимой форме, так и в форме, адаптированной к конкретным архитектурам.

Расширенная (стандартная) алгебра изображений Риттера [16] объединяет алгоритмы линейной алгебры и математической морфологии и порождает более сложные структуры, частным случаем которых являются морфологические алгоритмы. В рамках стандартной алгебры получены разложения по этому базису основных распространенных операций над изображениями, в том числе и для случаев, когда выход зависит не только от входного изображения, но и от дополнительных параметров.

Перспективными в отношении стандартизации алгоритмического наполнения систем понимания изображений представляются также дескриптивные алгебры изображений [17]. Их алгебраическая специфика определяется тем, что элементами кольца являются и модели изображений, и операции над изображениями. В качестве операций могут использоваться как стандартные алгебраические операции, так и специализированные операции обработки и преобразования изображений, представленные в алгебраической форме.

Дескриптивные алгебры изображений обеспечивают возможность: а) представления операций формирования и преобразования формальных описаний изображений и операций над преобразованиями в задачах обработки, анализа и распознавания изображений; б) приведения представления входной информации в соответствие требованиям применяемых преобразований; в) синтеза унифицированных алгоритмических схем с учетом функциональных возможностей, области компетентности, математических свойств составляющих преобразований, а также характера входной и выходной информации.

Работы У. Гренандера являются основой для создания наиболее общего алгебраического аппарата для работы с образами произвольной природы и их представления как алгебраических структур [18]. Существенным практическим вкладом У. Гренандера следует считать объединение алгебраического аппарата с аппаратом марковских случайных процессов за счет определения вероятностных мер на произвольных алгебраических структурах. Основная идея, лежащая в основе теории Гренандера, заключается в том, что знания об образах можно выразить в терминах регулярных структур.

Следует также отметить результаты Г. Ниманна и Д. Паулуса, касающиеся структур промежуточных данных для процессов сегментации изображений [19].

Со значительно большими сложностями сопряжена стандартизация представления знаний на высших уровнях обработки изображений. Основная сложность возникает в связи с различиями между процедурами построения описаний объектов и определением стратегий обнаружения этих объектов на изображении. Процедурное определение объектов посредством стратегий идентификации увеличивает эффективность исполнения программ. С другой стороны, природа декларативных описаний с помощью физических признаков требует использования баз знаний более общего характера. В настоящее время преимущественно используются процедурные описания просто в силу их вычислительной реализуемости. Учитывая двойственность, связывающую процессы описания объектов и их распознавания в анализе изображений, можно считать, что оба подхода должны использоваться совместно и процессы описания и обнаружения должны иметь возможность взаимодействия.

Еще одна проблема – разработка средств для доступа к знаниям. Созданы первые образцы соответствующих интерфейсов, в том числе графических, позволяющих обращаться к фрагментам изображений и базам данных изображений для получения статистических характеристик признаков и свойств объектов.

Заметим, что потенциальные возможности систем понимания изображений начнут реализовываться в полном объеме лишь после включения в них методов автоматического обучения как для описания, так и для распознавания изображений. С сожалением приходится констатировать, что до сих пор современные возможности математической теории распознавания образов используются в анализе изображений лишь в самой ограниченной степени и обе дисциплины существуют и развиваются почти независимо друг от друга.

Очередной этап развития АСОИЗ – информационно-поисковые системы для информации, представленной в виде изображений (ИПСИ) [20, 21]. Возможность появления этого нового класса систем для работы с изображениями возникла в результате разработки новых технических средств и информационных технологий. Это в первую очередь дисплеи с высоким разрешением, новые стандарты сжатия изображений, средства связи для передачи комплексной (мультимедиа) информации по широкополосным каналам, специализированные сопроцессоры графических устройств, новые методы просмотра изображений, новые методы информационного поиска, использование параллельных ЭВМ, совершенствование методов разработки и наполнения баз знаний. В сочетании с новыми возможностями технологий хранения изображений и новыми информационными технологиями распознавания образов и обработки и анализа изображений это развитие обеспечило представление и вывод изображений чрезвычайно высокого качества, существенное ускорение процессов обработки, анализа и передачи изображений и привело к разработке сверхбольших баз данных изображений.

Непосредственное использование информации, представленной в виде изображений, в сочетании с традиционными возможностями текстовых информационно-поисковых систем дает существенные преимущества в прикладных областях, в частности, описание деталей изображений при помощи признаков, выделяемых на изображении на нижнем уровне обработки, значительно более адекватно, чем их описание с помощью текстовых конструк-

ций, использующих понятия высшего уровня; построение общего описания изображения (концепта изображения) с помощью признаков изображения нижнего уровня нереально; с другой стороны, нетрудно построить концепт изображения с помощью текстового описания; непосредственный доступ к изображению обеспечивает возможность его быстрого просмотра и является естественным способом взаимодействия пользователя с ИПСИ; визуальный интерфейс позволяет увеличить «точность» поиска за счет представления в зрительной форме как спецификаций запроса, так и результата поиска.

В настоящее время ИПСИ широко используются в различных прикладных областях, в том числе: географических информационных системах, автоматизированных системах организационного управления, системах архивирования и передачи медицинских изображений, системах автоматического проектирования, системах автоматического управления производственными процессами, робототехнических системах и базах данных для поддержки научных исследований.

ИПСИ обычно состоит из подсистемы ввода изображений, системы обработки изображений, подсистемы вывода изображений, базы данных изображений и подсистемы передачи изображений.

ИПСИ часто используются совместно или объединяются с другими информационными системами, в частности, системы архивирования и передачи медицинских изображений объединяются с радиологической и больничной информационной системой. Подобное объединение увеличивает прикладную полезность и гибкость использования ИПСИ и одновременно количество квалифицированных пользователей. В результате функции ИПСИ расширяются, решение задач требует доступа ко все большему объему знаний о предметной области и их обработки и ИПСИ во все большей степени становятся системами, основанными на знаниях. В то же время специализированные ИПСИ, основанные на знаниях, должны обеспечивать интегрированную среду для представления, описания, структурирования и пространственного анализа изображений.

Основными составляющими ИПСИ являются система обработки изображений и база данных изображений. Помимо стандартных для систем обработки изображений задач анализа и улучшения качества изображений и распознавания, эти две составляющие ИПСИ должны обеспечивать выполнение следующих функций (их можно рассматривать также как три этапа обработки изображений на основе использования знаний).

1. *Анализ изображений и распознавание образов.* Проводится анализ исходного изображения и распознавание присутствующих на нем объектов. Результат распознавания – объекты – представляются с помощью какой-либо структуры данных с тем, чтобы облегчить в дальнейшем доступ к ним и их преобразование. Роль системы на этом этапе заключается в выборе алгоритмов обработки изображений и структур данных.

2. *Структурирование и понимание изображений.* Объекты, выделенные на изображении, преобразовываются в структуры знаний об изображении с тем, чтобы обеспечить возможность анализа пространственной организации изображения и информационного поиска по изображениям. Роль системы на данном этапе заключается в выборе структур знаний.

3. *Пространственный анализ и информационный поиск по изображениям.* Решение прикладных задач требует объединения разнородной информации, использования собственной процедур решения и базы знаний о прикладной области. Пространственный анализ и информационный поиск по изоб-

ражениям могут взаимно дополнять друг друга. Для упрощения доступа к знаниям об изображении, их визуализации и обработки используются преобразования структур знаний об изображении, в том числе поворот, перенос, изменение точки наблюдения, построение плоских проекций трехмерных объектов, наложение на изображение или удаление с него символьных объектов. Результатом данного этапа является структура знаний, специализированная для задачи пользователя (план навигации в базе, путь, набор изображений, полученных в результате поиска, индексное описание изображения).

Указанные три этапа обработки есть последовательность преобразований, обеспечивающих перевод исходного изображения в некоторую структуру данных, представляющую изображение, затем в структуру знаний об изображении и в конечном счете в специализированную структуру знаний, соответствующую прикладной задаче пользователя.

Эта схема обработки изображений порождает пятиуровневую архитектуру базы данных изображений [22]:

а) уровень пользователя (соответствует анализу пространственных отношений; пользователь сообщает, каким образом база данных изображений должна быть использована применительно к факторам высшего уровня);

б) уровень семантических признаков (соответствует структурированию знаний об изображении; пользователь интерпретирует признаки изображения в контексте прикладной области);

в) уровень признаков изображения (соответствует пониманию изображений; основывается на анализе содержания изображения (пространственные отношения, формы и т. п.));

г) уровень представления признаков (соответствует структурированию данных, содержащихся в изображении; обеспечивает возможность многообразного представления одних и тех же признаков);

д) уровень организации признаков (соответствует хранению и поиску данных, содержащихся в изображении; аналогичен уровню организации памяти в традиционных базах данных).

Разделение на уровень пользователя и уровень семантических признаков позволяет ИПСИ поддерживать использование объектов изображения в различных случаях и предотвращает потерю информации. Пользователь должен иметь возможность обращаться к первым трем уровням, т. е. формулировать запросы, соответствующие уровням пользователя, семантических признаков и признаков изображения. Данный тип архитектуры предполагает также, что модель данных, содержащихся в изображении, должна включать различные аспекты этих данных.

При разработке ИПСИ необходимо учесть следующие требования, касающиеся организации хранения и поиска изображений [21].

Идентификаторы изображений. Все изображения, хранящиеся в ИПСИ, должны быть снабжены каким-либо индексом или идентификатором, причем желательно, чтобы они были разными. Новый индекс для каждого изображения будет очень сложным и громоздким, что приведет к необходимости резервирования значительного объема памяти для хранения индексов. Естественными индексами могут служить текстовые описания изображения либо его признаки. Искусственным индексом является некоторый формальный показатель (например, номер). Используется несколько типов идентификаторов изображений. В ИПСИ, содержащих только изображения, идентификатор может представлять собой упорядоченный список важнейших признаков изображения, иногда с указанием отношений между

признаками в некоторых подписках, относящихся к определенным группам изображений. В ИПСИ, содержащих изображения вместе с некоторым сопроводительным текстом, идентификатор является комбинацией признаков изображения и текста.

Критерии подобия изображений. Поиск в ИПСИ основан, как правило, на частичном совпадении идентификаторов и индексов изображений. В связи с этим необходимо использовать некие критерии подобия изображений, характеризующие «расстояния» между изображениями. Очевидно, что универсального критерия подобия изображений не существует, поэтому формируются такие критерии применительно к предметной области.

Механизмы поиска изображений. Они используются для поиска и выборки изображений, хранящихся в ИПСИ и снабженных идентификаторами. Эти механизмы используют критерии подобия изображений для установления полного или частичного соответствия описаний. В случае номерного индекса может применяться реляционный механизм поиска. При индексе изображения, являющемся его описанием, можно использовать механизм поиска, основанный на статистических характеристиках текста. В случае индекса изображения, являющегося списком признаков изображения, следует использовать более сложный механизм поиска, позволяющий идентифицировать отношения элементов некоторых подписков.

Среда формирования запросов. Среда запросов ИПСИ должна обеспечивать возможность формировать запросы, например, чисто текстовый запрос, запрос, включающий только изображения, или комбинированный запрос, включающий как текст, так и изображения. Среда запросов должна также обеспечивать возможность пошагового уточнения запроса, причем снова с использованием любых или всех типов данных, поддерживаемых соответствующим приложением.

База знаний об изображениях. Эффективность поиска в ИПСИ широкого назначения может быть существенно увеличена при использовании знаний о базе данных изображений. Эти знания могут быть представлены с помощью семантической сети, структурного графа или базы знаний об изображениях. Последние обеспечивают эффективный доступ к изображениям на промежуточном, символьном уровне. Символьные представления изображений позволяют организовать эффективное взаимодействие с процессами вывода на высшем уровне обработки, отвечающими за получение интерпретации изображения. Примером подобного символьного представления знаний может служить иерархический механизм поиска по признакам, основанный на анализе отношений признаков изображения, объекта изображения, компонент объекта изображения и фрагментов изображения.

Сжатие изображений. Поскольку изображения в полном виде требуют для хранения очень большого объема памяти, они сжимаются с помощью одного из действующих стандартов (например, JPEG или MPEG). В некоторых приложениях требуются изображения с полным разрешением и со всеми деталями, что вызывает необходимость использовать методы сжатия без потери информации.

2. Общая характеристика системы «Черный квадрат. Версия 1.0». Нашей целью является разработка инструментально-программного комплекса для автоматизации научных исследований и обучения в области обработки, анализа, распознавания и понимания изображений.

Эта система предназначена для исследователей и разработчиков, в первую очередь специализирующихся в области обработки, анализа, распозна-

вания и понимания изображений или сталкивающихся с необходимостью анализировать и оценивать информацию, представленную в виде изображений.

Приведенный в разд. 1 анализ АСОИЗ позволил сформулировать представление о том, какой должна быть исследовательская система указанного назначения. Ниже следует изложение основных требований, которым она должна отвечать.

Система «Черный квадрат. Версия 1.0» относится к ИПСИ широкого назначения и не связана с узкой предметной областью. Такие системы работают с изображениями, полученными в широких диапазонах условий наблюдения, поэтому, как правило, невозможно заранее сформировать последовательность операций обработки и анализа, которые позволят получить требуемый результат. В этом случае необходимо применение знаний о классе анализируемых сцен, процессе обработки изображения и имеющихся в распоряжении вычислительных методах. Знания закладываются в систему в явном виде, а затем система «проектирует сама себя», выбирая соответствующую последовательность методов (или предлагая ее пользователю). Система должна уметь модифицировать последовательность используемых ею методов и выбирать средства для очередного этапа на основе предыдущих и применять разные последовательности методов для различных участков изображения. Перечисленные операции могут осуществляться в режиме интерактивного взаимодействия с пользователем.

Для решения широкого класса задач, связанных с анализом и оценением информации, представленной в виде изображений, необходима свобода действий, позволяющая дополнять и расширять функциональные возможности системы с помощью подключения новых библиотек алгоритмов, создания макрокоманд или сценариев, добавления новых типов данных и способов обращения с ними, добавления и замены оконных интерфейсов для редактирования данных. Это означает, что система должна представлять собой пополняемое множество программ для работы с данными различных типов и иметь блочную структуру.

Следовательно, система должна сочетать возможности инструментальной среды обработки и анализа изображений и системы, основанной на использовании знаний. Поэтому одним из основных структурных элементов системы является база знаний. Состав базы знаний определяется структурой процессов понимания изображений [23] и спецификой задач автоматизации разработки, исследования и тестирования алгоритмов обработки, анализа и распознавания изображений. Как отмечалось выше, в системы понимания изображений базы знаний обычно включают модули универсального знания, не связанного с какой-либо прикладной областью (знания, необходимые для планирования и управления процессами обработки, отображения результатов, оценки качества обработки, распознавания объектов и разрешения конфликтов, знания о методах обработки и анализа), и модули знания, связанного с конкретной предметной областью (стратегии сегментации, описания объектов, специализированные стратегии для извлечения признаков и идентификации объектов).

Система содержит программную реализацию иерархий классов основных объектов, используемых в анализе изображений, располагает специализированным интерфейсом пользователя, включает семейство алгоритмов распознавания, позволяющее решать основные задачи анализа и понимания изображений с помощью эффективных вычислительных процедур, и обеспе-

чивает накопление и структуризацию знаний и опыта в области анализа изображений.

Система должна содержать пополняемый («открытый») набор программ для работы с данными различных типов; объединять эти программы с помощью оболочки, имеющей встроенный диспетчер управления данными, связь с базой данных системы, где зарегистрированы эти программы и структуры данных, загрузчик оконных интерфейсов, предназначенных для визуализации и редактирования данных; обеспечивать возможность применения зарегистрированных программ и быстро настраиваться на работу с данными практически любого типа (включая данные большого объема).

Среди аналогичных наиболее известных систем, разработанных и развиваемых в последнее время, следует отметить Halcon и MatLab.

Система Halcon (коммерческая версия системы Horus) [24] – сложный пакет анализа изображений, предназначенный для разработки приложений, исследований и обучения. Система имеет следующие компоненты: ядро системы, которое состоит из модулей, вызывающих операторы, управляющие операциями ввода/вывода и переводящие объекты данных во внутренний язык системы; средства управления данными, которые отвечают за управление основной памятью (оптимизированной для обработки изображений), а также за создание, управление и уничтожение графических объектов и кортежей; библиотеку, включающую более 750 операторов для анализа изображений и визуализации данных; средство HDevelop для ускоренной разработки программ анализа изображений (Computer Aided Vision Engineering CAVE-tool); средства отладки; базу знаний операторов, которая содержит справочную информацию по каждому оператору системы (имена и количество операторов, типы параметров, предлагаемые значения параметров).

Halcon – открытая система, в которой поддерживаются два основных метода программирования: интерактивная разработка программ с помощью HDevelop и с использованием языков С и С++ (Halcon предоставляет интерфейс для этих языков).

Данная система является мощным инструментальным средством для обработки и анализа изображений, однако в базе знаний содержится только справочная информация об алгоритмах. Система не допускает визуализации новых типов данных и не поддерживает технологию "Plug-in", что не позволяет подключать пользовательские интерфейсы. К числу недостатков системы следует отнести отсутствие возможности запускать одновременно несколько алгоритмов при работе под операционной системой Windows 9x.

Универсальный пакет программ MatLab [25] включает инструментальные средства обработки изображений. Универсальность пакета приводит к затруднениям при решении задач понимания изображений, обладающих определенной спецификой: в MatLab наиболее развиты операции с плавающей запятой, что приводит к замедлению обработки изображений, для которых характерны операции с целыми числами; MatLab, как и Halcon, не позволяет подключать новые пользовательские интерфейсы и не имеет базы знаний.

Итак, система «Черный квадрат. Версия 1.0» является средой для разработки алгоритмов обработки, анализа и распознавания изображений, обладающей справочными и информационно-поисковыми функциями.

Система включает базу данных, содержащую алгоритмы обработки, анализа, распознавания и понимания изображений, инструментальные средства для их выбора, применения и объединения, комбинирования и слияния в

многозадачном режиме, базу знаний, содержащую стандартизированные блоки задач, средства для визуализации и редактирования используемых данных, электронный многоуровневый справочник по системе, архив тестовых и обработанных изображений и оболочку, связывающую все компоненты системы в единый инструментально-программный комплекс. Основными объектами хранения являются алгоритмы, структуры данных – классы объектов, необходимые для анализа и оценивания изображений, и изображения. Функционирование системы обеспечивается единой интегрированной программной средой, построенной на общей информационной и методической базе и обладающей общей навигационной оболочкой.

При построении проблемно-ориентированных версий системы используются схемы классификации алгоритмов, ориентированные на решение определенного класса задач и базирующиеся на концепциях, методологии и терминологии соответствующей области знаний.

Выбор подобной реализации позволяет облегчить разработку и модернизацию алгоритмов для работы с изображениями. Модульная организация дает возможность обновлять компоненты системы по частям, облегчает локализацию и исправление ошибок. Новый или модифицированный алгоритм может быть включен в работу сразу после регистрации. Поэтому разработка нового алгоритма разбивается на несколько шагов:

- подготовка базовых алгоритмов, обычно простейших операций над изображением или над числовой матрицей, если они не присутствуют в системе;
- составление макрокоманды, использующей базовые алгоритмы, решающие поставленную задачу;
- регистрация нового алгоритма и использование его далее в качестве базового.

При большом числе базовых алгоритмов возникает задача поиска нужного алгоритма или выбора его из нескольких однотипных. Эта задача решается с использованием возможностей поиска по базе данных, содержащей информацию о зарегистрированных алгоритмах.

3. Функциональные характеристики и архитектура системы «Черный квадрат. Версия 1.0». Решение задач анализа и понимания изображений включает принятие сложных решений на разных уровнях обработки и требует наличия у пользователя высокой квалификации в данной предметной области. Для получения искомого решения обычно требуется соединение ряда процессов и этапов обработки. На каждом из таких этапов может возникать проблема выбора наиболее подходящего метода и задания значимых его параметров.

Известно, что чисто синтаксическое комплексирование недостаточно для эффективного решения задач анализа и понимания изображений. Необходимо обеспечить возможность семантического комплексирования методов и процедур.

Следовательно, система должна обеспечивать формальное и точное представление «квалификации» специалиста по анализу и пониманию изображений и включать средства для эмуляции стратегий выбора и применения известных методов обработки, используемых специалистами при решении задач анализа и понимания изображений.

При разработке системы, обеспечивающей искомую автоматизацию, была определена модель стратегий выбора и применения известных методов

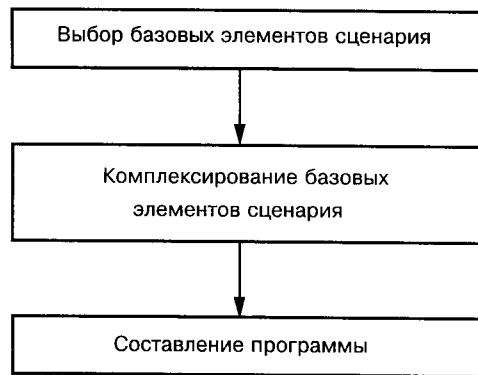


Рис. 1

обработки, используемых специалистами при решении задач анализа и понимания изображений, и архитектура системы.

Модель включает планирование процесса решения задачи и управление исполнением этого процесса. В планирование входят выбор базовых элементов процесса (методы), комплексирование его базовых элементов (составление собственно плана) и составление программы (рис. 1). В управление исполнением входят исполнение программ, регистрация и анализ результатов и оптимизация процесса обработки (рис. 2).

Архитектура, основанная на этой модели, должна обеспечивать представление знаний по анализу и пониманию изображений, описание задач, планирование, управление исполнением, функционирование базы знаний и комплексирование алгоритмов при помощи таких объектов и понятий, как запросы, операторы, цели, контекст и правила.

Описание задачи. Знания о методах анализа и понимания изображений представляются с помощью формальных конструкций на двух уровнях. Цель представляет собой абстрактное выражение функций системы (задач обработки), реализацией которых служит один или несколько операторов, поставленных им в соответствие. Оператор может быть простым (соответствует исполняемой программе) или сложным, для которого задается разбиение на элементы. Запрос включает исходные данные и информацию о характере обработки, которой они должны быть подвержены. Контекст включает всю доступную потенциально полезную информацию о задаче (технические условия, характеристики датчиков и т. д.). Цели, операторы, запросы и контекст представляются в виде фреймов.

Планирование. Основным механизмом служит иерархическое формирование плана на основе сценариев. План представляет собой совокупность этапов обработки, обеспечивающих в целом решение задачи анализа и пони-

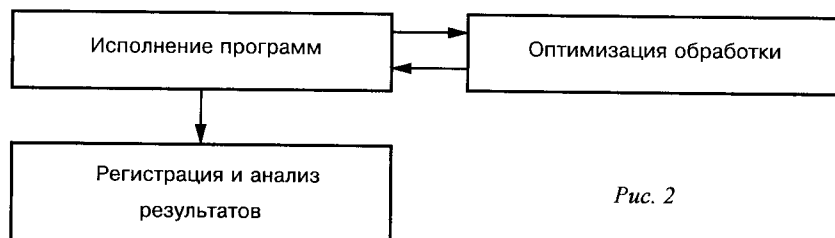


Рис. 2

мания изображений. Каждый этап плана может предусматривать решение сложной задачи анализа и понимания изображений в соответствии с собственным планом.

Элементы плана представляют «абстрактные» задачи анализа и понимания изображений (цели), а не конкретные алгоритмы. В процессе исполнения для выбора оператора, наиболее подходящего для решения соответствующей задачи, используются специальные правила выбора. Элементарные планы включены в составные операторы в виде диаграмм разбиения, содержащих одну или несколько подзадач, соединенных последовательно или параллельно. Эти подзадачи рассматриваются как запросы, поскольку потоки данных между подзадачами составного оператора известны.

Управление исполнением. Эта функция заключается в исполнении плана на конкретных исходных данных и в конкретном контексте. Правила выбора обеспечивают выбор оператора из нескольких возможных на основе значений контекстных полей, характера исходных данных, ограничений, присутствующих запросам, и т. п. Для задания начальных значений параметров процедур используются правила инициализации, выражающие грубое начальное приближение. При итеративном процессе обработки операторы, имеющие настраиваемые параметры, корректируются с помощью правил настройки. Для оценки результатов обработки запроса используются правила оценки.

База знаний. Статическая часть базы состоит из совокупности стандартных блоков задач, каждый из которых включает цели и поставленные им в соответствие операторы. Для каждого типа задачи определяется соответствующая цель. Поскольку способов достижения цели может быть несколько, для каждого выбирается оператор, который ставится в соответствие данной цели. В случае составного оператора блоки задач должны быть сформированы для каждого типа подзадачи. Процедура формирования блоков задач выполняется для всех задач и подзадач до тех пор, пока для всех подзадач всех составных операторов не будут определены все блоки задач и каждому простому оператору не будет поставлена в соответствие программа. Кроме того, в базу вводятся правила четырех указанных выше типов, имитирующих стратегию решения задач анализа и понимания изображений.

Комбинирование, объединение и слияние алгоритмов. Решение задачи начинается с формирования запроса, который определяет искомую цель, данные, посредством обработки которых эта цель должна быть достигнута, и контекст решения задачи. С помощью правил выбора проводится ранжирование возможных операторов и выбирается наилучший, который после использования правил инициализации применяется к данным. В случае простого оператора это означает исполнение соответствующей программы, в случае составного – формируются запросы для его подзадач и дерево запросов. Запросы, включенные в дерево, обрабатываются в естественной последовательности. Результат обработки запроса анализируется с помощью правил оценки, и при отсутствии искомого результата повторно применяется тот же оператор с модифицированными параметрами либо следующий по рангу оператор.

В функции системы входит решение следующих классов задач анализа и распознавания изображений:

- 1) сопоставление двух изображений в целом для установления их принадлежности к одному классу (определяется, представляют ли изображения один и тот же объект или сцену);

2) сопоставление изображения в целом с набором или серией последовательных (по времени) изображений, представляющих некоторый класс изображений, т. е. объектов или сцен (цель та же, что и в задаче 1);

3) задач 1 и 2 для случая нескольких классов;

4) поиск на предъявленном для распознавания изображении некоторой регулярности/нерегулярности (объекта, ситуации), на которую следует обратить внимание, хотя она и не задавалась в априорном перечне эталонов (ассоциативный поиск; ограниченно детерминированный набор классов – задачи логическо-семантической фильтрации в сочетании с самообучением);

5) поиск на предъявленном для распознавания изображении регулярности/нерегулярности/фрагмента заданного вида;

6) разбиение множества изображений на непересекающиеся подмножества (задача автоматической классификации);

7) автоматическая классификация на одном изображении (разбиение изображения на однородные области, группы объектов, сегментация области, выделение признаков объектов);

8) совместно задач 6 и 7;

9) автоматическое выделение непроектируемых элементов, характерных объектов изображения, признаков-объектов, пространственных и логических отношений для синтеза формализованных представлений и описаний изображения;

10) приведение изображения к виду, удобному для распознавания; автоматический синтез формализованных представлений/описаний изображения;

11) задач 9 и 10 в диалоговом режиме;

12) задачи восстановления (пропущенных кадров в последовательности изображений; изображений в целом по их фрагментам; фрагментов изображения (и объектов) на основе непроектируемых элементов, признаков и порождающих процедур с учетом контекста изображения в целом; траектории задачи по ее фрагментам и неизвестных фрагментов траекторий);

13) выбор и формирование траектории задачи распознавания изображений (в смысле задачи распознавания со стандартной обучающей информацией);

14) задач 1–13 при наличии на изображениях динамических объектов, сложной фоновой обстановки (в том числе динамических и статических помех) и с учетом способов получения, формирования и представления изображений.

Система (с помощью объектных библиотек) поддерживает операции чтения, распределения в памяти, сохранения, копирования и т. д. со следующими типами объектов: векторы (целые числа, действительные числа с плавающей запятой); матрицы (целые числа, действительные числа с плавающей запятой, структурные); изображения (1, 4, 8, 16, 24, 32 бит/пиксел) в форматах BMP, TIFF, PCX, JPEG (чтение и сохранение для TIFF, PCX, JPEG); фрагменты (векторов и матриц) прямоугольные, произвольной формы.

В версии 1.0 системы реализованы следующие базовые классы операций: стандартные математические операции; арифметические операции; измерения на матрицах (метрики); логические операции; операции над матрицами (поворот, растяжение, определение следа, решение линейных систем, вычисление собственных чисел и т. д.); преобразование типов матриц (бинаризация, расширение и т. д.); стандартные операции обработки изображений – матричная фильтрация, свертка (выделение границ, сглаживание, ран-

говая фильтрация, матричная фильтрация Фурье, вейвлеты и т. д.); методы математической морфологии (эрозия, дилатация, открытие, закрытие и т. д.); сегментация (разделение на зоны влияния, сегментация по локальной информации, построение остова бинарного изображения); улучшение качества изображения (регулировка контраста, яркости).

Система поддерживает обработку тоновых, бинарных и цветных изображений.

Инструментально-программный комплекс состоит из следующих основных частей: таблицы баз данных, составляющие базу знаний; программные средства поддержки базы знаний; оболочка и пользовательские интерфейсы; набор объектно-ориентированных библиотек для работы с данными – считывание, сохранение, размещение в памяти, преобразования; алгоритмические библиотеки и макрокоманды; система помощи.

База знаний имеет следующий состав: таблицы алгоритмов, сценариев, решающих правил анализа изображений (используются также для накопления знаний о решенных задачах); таблицы используемых структур данных; таблицы каталогов и описаний эталонных изображений, решенных задач, методик, алгоритмов и сценариев.

Набор интерфейсов включает: системный интерфейс (оболочка), редакторы таблиц базы знаний, редактор изображений, редакторы числовых матриц, интерфейсы исполняющей системы, информационные и справочные интерфейсы.

Системный интерфейс (оболочка) поддерживает динамическое системное меню, обеспечивающее доступ к функциям системы, и линейки быстрого доступа. Основные системные функции: работа с файлами данных, печать данных, вызов других интерфейсов.

Редакторы таблиц базы знаний обеспечивают возможность заполнения и редактирования соответствующих таблиц базы знаний, в том числе формирование и изменение процедур обработки и распознавания, создание и редактирование классификаций алгоритмов и сценариев.

Редактор изображений предназначен для просмотра и коррекции целочисленных матриц изображений с 1-, 4-, 8-, 16- и 32-битовыми полями.

Редактор числовых матриц и векторов предназначен для просмотра и коррекции целочисленных двумерных матриц или векторов с 1-, 4-, 8-, 16- и 32-битовыми полями.

Интерфейсы исполняющей системы обеспечивают поиск и выбор алгоритмов или сценариев, ввод параметров алгоритмов, запуск на исполнение, управление процессами обработки изображений.

Информационные и справочные интерфейсы предназначены для отображения системных сообщений, вывода промежуточных и результирующих данных в процессе выполнения алгоритмов и процедур, ввода ключей поиска и отображения справочной информации.

Программные средства поддержки базы знаний включают: системные программы, программы поддержки и обслуживания таблиц базы знаний, справочную систему, программы экспорта и импорта данных.

Системные программы обеспечивают организацию вычислительных процессов в системе и обмен данными между функциональными блоками.

Программы поддержки и обслуживания таблиц базы знаний предназначены для формирования строки запроса, получения и обработки результатов запросов, поиска, чтения, изменения, удаления и добавления новых записей, индексирования и проверки целостности данных в таблицах.

Программы справочной системы выполняют поиск справочных материалов по темам и/или контексту.

Программы экспорта и импорта данных осуществляют чтение и вывод данных в различных форматах.

Основные функциональные блоки системы и их взаимодействие иллюстрируются рис. 3.

Итак, перечислим основные функциональные характеристики системы.

1. Классы задач, решаемых с помощью системы: предварительная обработка изображений; приведение изображений к виду, удобному для распознавания; обнаружение, выделение и вычисление признаков на изображениях; распознавание; получение информации о способе решения задачи, решенных задачах и об алгоритмах, включенных в систему; тестирование и сравнение математических и функциональных характеристик алгоритмов.

2. Базовое алгоритмическое наполнение системы: арифметические и логические операции над матрицами, векторами и изображениями; аффинные преобразования; векторная и матричная алгебра; преобразования Фурье, Габора, вейвлеты; матричная фильтрация, выделение яркостных переходов и контуров; сопоставление фрагментов изображений и текстур – морфологические операции; методы сегментации; методы дискретной оптимизации;

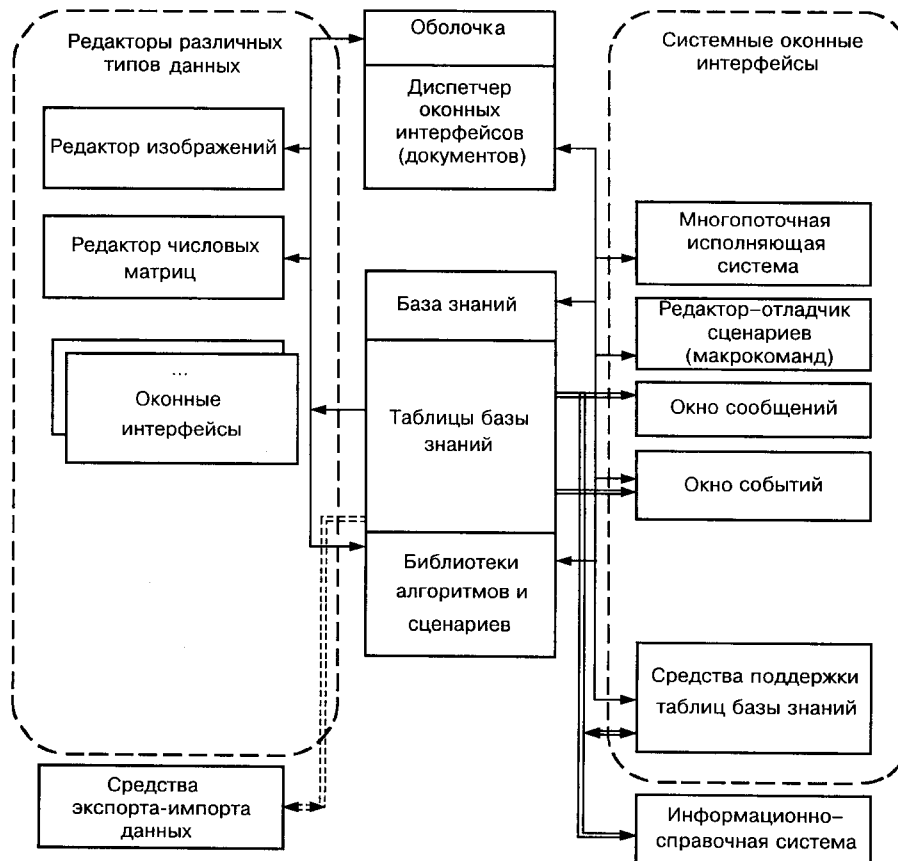


Рис. 3

вычисление метрик для матриц и векторов; представление изображений в виде бинарных деревьев; алгоритмы принятия решения.

3. Функциональные особенности системы: открытость по обрабатываемым системой объектам, библиотекам алгоритмов, наполнению лексикографической и экспертной информацией; стандартная диалоговая оболочка и инструментальная среда разработчика; функциональная полнота алгоритмического наполнения для задач из области компетентности системы; наличие набора баз данных, обеспечивающих полное сопровождение всех решаемых задач; возможность модернизации системы посредством включения новых и исключения устаревших модулей (функционально различные части системы размещены в различных модулях, имеющих легко поддерживаемые связи с остальными частями системы).

СПИСОК ЛИТЕРАТУРЫ

1. Журавлев Ю. И., Гуревич И. Б. и др. Банк алгоритмов обработки изображений // Распознавание образов и анализ изображений: новые информационные технологии (РОАИ-1-91) (Минск, 14–18 октября 1991 г.): в 4-х частях. Ч. II. Минск: Институт техн. кибернетики АН БССР, 1991. С. 3.
2. Zhuravlev Yu., Gurevitch I. B. et al. Development and investigation of the mathematical and computational basis for a system of information technologies of pattern recognition and image understanding // Pattern Recogn. and Image Analysis: Advances in Mathematical Theory and Applications. 1993. 3, N 3. P. 266.
3. Zhuravlev Yu., Gurevitch I. B. et al. Development and Investigation of the Mathematical and Computational Basis for a System of Information Technologies of Pattern Recognition and Image Understanding Tech. Rep. VA, USA: George Mason Univ., 1994.
4. Журавлев Ю. И., Гуревич И. Б. и др. Алгоритмическая база знаний «Анализ и распознавание изображений 8.95» // Тез. докл. 2-й Всерос. с участием стран СНГ конф. «Распознавание образов и анализ изображений: новые информационные технологии». Ч. 4. Ульяновск: УГТУ, 1995. С. 3.
5. Zhuravlev Yu., Gurevitch I. B. et al. Algorithmic knowledge base "Analysis and Recognition of Image 8.95" // Pattern Recogn. and Image Analysis Advances in Mathematical Theory and Applications. 1996. 6, N 2. P. 332.
6. Gurevitch I. B., Polikarpova N. S., Zhuravlev Y. I. Logical models of images and recognition operators for image understanding environment // Proc. of the 10th Scandinavian Conf. on Image Analysis, Lappeenranta, Finland, June 9–11, 1997. Pattern Recogn. Soc. of Finland. Lappeenranta, V. II. 1997. P. 809.
7. Gurevitch I. B., Khilkov A. V., Polikarpova N. S. et al. A prototype of a system for developing and testing methods of analysis and evaluation of the information in the form of images (OS Windows'95) // Pattern Recogn. and Image Analysis: Advances in Mathematical Theory and Applications. 1998. 8, N 3. P. 354.
8. Gurevitch I. B., Khilkov A. V., Polikarpova N. S. et al. An open system with database functions for solving image analyses and processing problems // Collection of Abstracts. The 5th Open German-Russian Workshop on Pattern Recognition and Image Understanding, 21–25 September 1998. Herrsching, Germany: FORWISS, 1998.
9. Haralick R. M., Shapiro L. G. Computer and Robot Vision. V. II. Reading. Massachusetts: Addison-Wesley Publ. Company, Inc., 1993.
10. Crevier D., Lepage R. Knowledge-based image understanding systems: a survey // Computer Vision and Image Understanding. 1997. 67, N 2. P. 161.
11. Matsuyama T. Expert systems for image processing, analysis, and recognition: declarative knowledge representation for computer vision // Advances in Electronics and Electron Phys. 1993. 86. P. 81.
12. Konstantinides K., Rasure J. R. The Khoros software development environment for image and signal processing // IEEE Trans. on Image Process. 1994. 3, N 3. P. 243.

13. **Mundy J. L. and IUE Committee.** The image understanding environment: overview // Image Understanding Workshop: Proc. of a Workshop held in Washington, D. C., April 18–21, 1993. San Mateo, CA, USA: Morgan Kaufmann Publ., Inc., 1993. P. 283.
14. **Zhuravlev Yu. I.** Correct algebras over sets of incorrect (heuristic) algorithms // Kibernetika (Kiev). 1977. N 4. P. 14; N 6. P. 21; 1978. N 2. P. 35 (in Russian).
15. **Zhuravlev Yu. I.** An algebraic approach to recognition or classification problems // Pattern Recogn. and Image Analysis: Advances in Mathematical Theory and Applications. 1998. 8, N 1. P. 59.
16. **Ritter G. X., Wilson J. N.** Handbook of Computer Vision Algorithms in Image Algebra. Boca Raton, Florida: CRC Press, Inc., 1996.
17. **Gurevitch I. B., Smetanin Yu. G., Zhuravlev Yu. I.** Descriptive image algebras: Determination of the base structures // Pattern Recogn. and Image Analysis: Advances in Mathematical Theory and Applications. 1999. 9, N 4. P. 635.
18. **Grenander U.** General Pattern Theory. A Mathematical Study of Regular Structures. Oxford: Clarendon Press, 1993.
19. **Paulus D., Niemann H.** Iconic-symbolic interfaces // Proc. SPIE. 1992. **1659**. P. 204.
20. **Chang S. K., Hsu A.** Image information systems: where do we go from here // IEEE Trans. on Knowledge and Data Eng. 1992. 4, N 5. P. 431.
21. **Ang H., Narasimhalu A. D., Al-Hawamdeh S.** Image information retrieval systems // Handbook of Pattern Recogn. and Computer Vision /Eds. C. H. Chen, L. F. Pau, P. S. Wang. Singapore: World Sci. Publ. Co. Pte. Ltd., 1993.
22. **Fierens F., Van Cleynenbreugle J., Suetens P., Oosterlinck A.** A software environment for image database research // Journ. of Visual Languages Comput. 1992. 3. P. 49.
23. **Marr, Vision.** A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco: W. H. Freeman and Company, 1982.
24. **Halcon** Version 5.1. User's Manual. Munich, Germany: MVTec Software GmbH, 1998.
25. **Matlab** Version 5.2 User's Manual. Natick, MA, USA: The MathWorks Inc., 1998.

Поступила в редакцию 19 августа 1999 г.
