СОДЕРЖАНИЕ

Введение 10

▼ 1

Машинное зрение, LabVIEW и NI Vision 14
1.1. Проблематика машинного зрения 14
1.1.1. Задачи машинного зрения 15
1.1.2. Уровни и методы машинного зрения 16
1.1.3. Сопряженные технические дисциплины 17
1.1.4. Требования к алгоритмам машинного зрения 19
1.1.5. Роль специальных программных средств в разработке приложений машинного зрения 22
1.2. Знакомство с LabVIEW 24
1.3. Знакомство с NI Vision 44
1.3.1. Платформа NI Vision: захват, обработка и анализ изображений в LabVIEW 44
1.3.2. Функции NI Vision 45
1.4. Знакомство с Vision Assistant 50
1.4.1. Начало работы с Vision Assistant 50
1.4.2. Пример работы в Vision Assistant 52
1.4.3. Экспорт проекта в LabVIEW 56

▼ 2

Цифровые изображения 59
2.1. Растровое изображение 59
2.1.1. Изображение как двумерный массив данных 59
2.1.2. Алгебраические операции над изображениями 62

2.2. Виды изображений ... 75
2.2.1. Физическая природа изображений ... 76
2.2.2. Тип пикселя ... 81
2.3. Устройства оцифровки и ввода изображений ... 84
2.3.1. Линейки и матрицы, сканеры и камеры ... 84
2.3.2. Геометрия изображения ... 87
2.3.3. Цифровые и аналоговые устройства ... 88
2.3.4. Пространственное разрешение ... 89
2.3.5. Программное обеспечение ... 91
2.4. Форматы хранения и передачи цифровых изображений ... 92
2.4.1. Методы сжатия цифровых изображений ... 94
2.4.2. Формат BMP ... 96
2.4.3. Формат PCX ... 99
2.4.4. Формат GIF ... 99
2.4.5. Формат TIFF ... 100
2.4.6. Формат JPEG ... 101
2.4.7. Формат DICOM ... 105
2.5. Цифровые видеопоследовательности ... 107
2.5.1. Скорость съемки ... 107
2.5.2. «Смаз» изображения ... 108
2.5.3. Этапы проектирования системы видеосъемки ... 109
2.5.4. Быстрая съемка и съемка быстропротекающих процессов ... 111
2.5.5. Форматы хранения и передачи цифровых видеопоследовательностей ... 112

▼ 3

Методы обработки изображений ... 118
3.1. Гистограмма и гистограммная обработка. Бинаризация и сегментация. Профили и проекции ... 119
3.1.1. Гистограмма и гистограммная обработка изображений ... 121
3.1.2. Бинаризация полутоновых изображений ... 137
3.1.3. Адаптивная бинаризация ... 137
3.1.4. Сегментация многомодальных изображений ... 141
3.1.5. Обработка цветных изображений ... 141
3.1.6. Профиль вдоль линии и анализ профиля ... 151
3.1.7. Проекция и анализ проекции ... 160

3.2. Фильтрация изображений. Ранговая нелинейная фильтрация. Выделение объектов 163
3.2.1. Задача фильтрации изображений 163
3.2.2. Фильтрация бинарных изображений 168
3.2.3. Нелинейная фильтрация полутоновых изображений 175
3.2.4. Задача выделения объектов интереса 180
3.3. Линейная фильтрация изображений. Линейная фильтрация в пространственной области. Преобразование Фурье. Линейная фильтрация в частотной области 184
3.3.1. Линейная фильтрация изображений. Линейная фильтрация в пространственной области 184
3.3.2. Преобразование Фурье. Линейная фильтрация в частотной области 192
3.3.3. Вейвлет-анализ 204
3.4. Выделение контуров на полутоновых изображениях 216
3.4.1. Задача выделения контуров 216
3.4.2. Операторы выделения контуров в IMAQ Vision 222
3.5. Математическая морфология Серра (ММ) 228
3.5.1. Теоретические основы математической морфологии 229
3.5.2. Операции математической морфологии в IMAQ Vision 239

▼ 4

Методы анализа изображений 251
4.1. Выделение и анализ связных областей 251
4.1.1. Выделение связных областей на бинарных изображениях 251
4.1.2. Методы сегментации полутоновых изображений 256
4.1.3. Геометрическое описание выделенных областей 261
4.1.4. Выделение и анализ связных областей в IMAQ Vision 265
4.2. Выделение геометрических примитивов 280
4.2.1. Методы выделения геометрических примитивов на основе преобразования Хафа 280
4.2.2. Выделение геометрических примитивов в IMAQ Vision 285
4.3. Обнаружение объектов, заданных эталонами 296
4.3.1. Методы обнаружения объектов, заданных эталонами 296
4.3.2. Функции привязки эталонов в IMAQ Vision 302
4.4. Измерения на изображениях 306

4.4.1. Функции геометрических измерений в IMAQ Vision 306
4.4.2. Функции яркостных измерений в IMAQ Vision 311
4.5. Считывание символьной информации 312
4.5.1. Задача оптического распознавания символов (OCR) 312
4.5.2. Считывание штриховых кодов 323
4.5.3. Считывание информации технических индикаторов 332
4.6. Служебные функции пакета IMAQ Vision 333
4.6.1. Ручные геометрические измерения и построения 333
4.6.2. Работа с растровым изображением и областями интереса 343

▼ 5

Цифровая фотограмметрия и бесконтактные измерения 347
5.1. Методы цифровой фотограмметрии 347
5.1.1. Фотограмметрический метод бесконтактных трехмерных измерений 347
5.1.2. Математическая модель камеры 348
5.1.3. Калибровка видеокамер 350
5.1.4. Внешнее ориентирование камер 351
5.1.5. Решение задачи стереосоответствия при построении модели поверхности. Структурированный подсвет 352
5.1.6. Приведение координат точек снимков к нормальному случаю съемки. Определение элементов взаимного ориентирования по внешнему 354
5.1.7. Вычисление пространственных координат точек поверхности 355
5.2. Фотограмметрический комплекс для бесконтактных измерений на базе LabVIEW и PXI 356
5.2.1. Аппаратное обеспечение комплекса 356
5.2.2. Программное обеспечение комплекса 360

▼ 6

Compact Vision System – новая промышленная платформа для систем технического зрения 369
6.1. Compact Vision System (CVS) 368
6.1.1. Технические характеристики CVS 368
6.1.2. Подключение и настройка CVS 369

6.1.3. Создание приложения для CVS 374
6.2. Пример приложения: фотограмметрический комплекс на базе CVS 377
6.2.1. Аппаратное обеспечение комплекса 377
6.2.2. Программное обеспечение комплекса 378
6.2.3. Пример сканирования детали 388

7

Примеры практических систем машинного зрения 391
7.1. Автоматизация измерений и технический контроль 391
7.1.1. Система автоматического выделения и фильтрации следа частиц 391
7.1.2. Система автоматизированного бесконтактного измерения объема круглых лесоматериалов 392
7.1.3. Система определения гранулометрического состава рудной массы (разработка ИИТ и ООО «НВП Центр-ЭСТАгео») 394
7.1.4. Система автоматического измерения угла схождения сварного шва 395
7.1.5. Система автоматического распознавания и подсчета некондиционных кристаллов на круглой пластине 396
7.1.6. Система компенсации геометрических искажений и бесшовной сшивки изображений, получаемых от многокамерных систем видеоввода 397
7.1.7. Система автоматизированного контроля качества внутренней поверхности труб 398
7.2. Зрение роботов 399
7.2.1. Область применения: автомобильные системы 399
7.2.2. Область применения: мобильные роботы 401
7.2.3. Система обнаружения препятствий на дороге перед движущимся транспортным средством 404
7.2.4. Система автоматической привязки телефрагментов к ортофотоизображению 404
7.3. Видеонаблюдение 405
7.3.1. Область применения: видеонаблюдение 405
7.3.2. Система обнаружения и сопровождения движущихся объектов по признаку их движения 406

7.3.3. Система стереообнаружения движения в зоне видеонаблюдения 411
7.3.4. Система считывания регистрационных номеров автомобилей 412
7.3.5. Система считывания номеров железнодорожных вагонов и цистерн 413
7.4. Биометрия 414
7.4.1. Область применения: биометрия 414
7.4.2. Система обнаружения и распознавания лиц 421
7.4.3. Система трехмерной реконструкции и формирования строго фронтального изображения лица человека 423
7.4.4. Система автоматического выделения человеческого лица и слежения за его чертами 425
7.4.5. Система распознавания жестов руки человека 426
7.4.6. Система для биомеханических исследований на основе высокоскоростной стереосъемки движений человека 426
7.4.7. Система слежения за положением головы и направлением взгляда ребенка 428
7.5. Обработка документов, распознавание текста и штриховых кодов 429
7.5.1. Система автоматического поиска и считывания штриховых кодов 429
7.5.2. Система автоматического распознавания машиночитаемых документов 430
7.5.3. Система считывания номеров денежных банкнот 431
7.5.4. Система оценки подлинности денежных банкнот 432
7.5.5. Система для оценки ветхости денежных банкнот 433
7.6. Медицинские приложения 434
7.6.1. Проекты в области анализа и обработки медицинских изображений 434
7.6.2. Системы для компьютерного анализа томографических изображений 435
7.6.3. Система компьютерного анализа томографических изображений для диагностики воспалительных заболеваний пазух и полости носа (синуитов) 436
7.6.4. Система компьютерного анализа томографических изображений для оценки степени ожирения у мужчин 436

7.6.5. Система компьютерного анализа медицинской рентгенографической информации для ранней диагностики остеопороза 438
7.6.6. Автоматизированное рабочее место врача-рентгенолога 438
7.6.7. Система телемедицины в области радиологических обследований 440
7.6.8. Модуль управления данными медицинского обследования на основе технологии DICOM 441

Список литературы 442
Литература по машинному зрению 443
Литература по машинному зрению на русском языке 443
Литература по машинному зрению на английском языке 445
Литература по LabVIEW и NI Vision 452
Литература по LabVIEW и NI Vision на русском языке 452
Литература по LabVIEW и NI Vision на английском языке 453

Предметный указатель 454

Введение

Настоящая книга написана и издана при поддержке академического гранта фирмы National Instruments. Первоначально она задумывалась как русский аналог замечательного англоязычного пособия [184] (Klinger T. Image processing with LabVIEW and IMAQ Vision), однако в ходе работы исходный замысел претерпел существенные изменения, и теперь книга представляет собой не только практическое пособие по построению приложений машинного зрения в среде LabVIEW с использованием библиотеки IMAQ Vision, но и полный учебный курс по тематике машинного зрения и цифровой обработки изображений. В теоретических разделах рассматриваются основные аспекты получения, хранения, обработки и анализа цифровых изображений, а также автоматического выделения и распознавания на изображениях различного рода объектов. В практических разделах описываемые методы подробно иллюстрируются примерами программ и схем обработки, созданных в LabVIEW на базе IMAQ Vision. Впервые на русском языке дается подробное описание новой промышленной платформы для систем технического зрения Compact Vision System (NI CVS), методов и приемов работы с ней. Помимо учебных и методических примеров, книга также содержит краткое описание значительного количества практических приложений.

Книга имеет следующую структуру.

В главе 1 обзорно рассматриваются основные темы книги. Вводятся основные понятия, связанные с проблематикой машинного зрения. Рассматриваются основные задачи машинного зрения, уровни анализа изображения, методы машинного зрения, место машинного зрения в ряду сопряженных технических дисциплин, основные требования к алгоритмам машинного зрения. Происходит первое знакомство со специализированными программными средствами LabVIEW, NI Vision и Vision Assistant, при помощи которых далее будет происходить изучение и рассмотрение на примерах различных процедур обработки и анализа изображений.

Глава 2 посвящена растровым цифровым изображениям. Рассматривается представление изображений как двумерных массивов данных. Описаны алгебраические операции над изображениями. Обсуждается физическая природа изображений, в том числе – изображения различных диапазонов длин волн и изображения различной физической природы. Описаны различные типы программных структур изображений, определяемых различным типом пикселя. Сделан краткий обзор устройств, используемых для оцифровки и ввода изображений в компь-

ютер. Обсуждаются геометрия изображения, цифровые и аналоговые устройства ввода, пространственное разрешение источников изображения. Описаны современные форматы хранения и передачи цифровых изображений. Кратко рассмотрена также связанная с этим проблема сжатия цифровых изображений. В заключении главы рассматриваются вопросы, связанные с вводом, хранением и передачей уже не отдельных кадров, а протяженных во времени цифровых видеопоследовательностей. Рассмотрены: скорость съемки, проблема «смаза» изображения, выбор экспозиции, выбор камеры, проблемы освещенности и выбор подсветки. Описаны устройства для съемки быстропротекающих процессов. В заключении раздела кратко описаны современные форматы хранения и передачи цифровых видеопоследовательностей.

Глава 3 посвящена методам обработки изображений. При этом под обработкой изображений «в узком смысле» понимается такая «низкоуровневая» обработка изображения, результатом которой также является растровое изображение или другой растровый объект (например, одномерный массив). В разделе 3.1 описаны гистограмма изображения, гистограммная обработка, бинаризация и многомодальная гистограммная сегментация цифровых изображений, адаптивные методы определения порогов бинаризации. Рассмотрены также профили и проекции изображений на оси координат. В разделе 3.2 рассмотрены нелинейные методы помеховой фильтрации изображений, а также выделение мало- и среднеразмерных объектов. Описана задача фильтрации изображений от шума. Рассмотрены следующие процедуры оконной фильтрации изображения: логическая фильтрация помех, бинарная медианная фильтрация, бинарная ранговая фильтрация, взвешенные ранговые фильтры, нелинейная фильтрация полутоновых изображений, ранговая оконная фильтрация. Рассматривается также задача выделения объектов интереса и метод нормализации фона. Раздел 3.3 посвящен линейной фильтрации изображений. Рассматриваются: линейная фильтрация в пространственной области, преобразование Фурье, линейная фильтрация в частотной области, связь линейной фильтрации в частотной и пространственных областях. В разделе 3.4 описаны методы выделения контуров (краев) на полутоновых изображениях. Рассмотрены различные дискретные операторы вычисления производных, операторы вычисления векторов градиентов, оператор Марра, оператор Кани. Раздел 3.5 посвящен описанию математической морфологии Серра. Излагаются теоретические основы математической морфологии. Описаны морфологические операции на бинарных изображениях и их обобщение на случай полутонового изображения. Приводятся алгоритмы морфологического выделение «черт» и объектов (областей) на изображениях.

Глава 4 посвящена методам анализа изображений. При этом под «анализом изображений в узком смысле» понимается такая «средне-» или «высокоуровневая» обработка изображения, результатом которой является уже не растровое изображение, а какой-либо векторный объект (список объектов) или семантическая информация, выражаемая в текстовом или любом другом символьном виде. В разделе 4.1 рассмотрены методы выделения и анализа связных областей. Описаны основные методы выделения связных областей на бинарных изображениях. Рассматриваются также алгоритмы сегментации полутоновых изображений.

Приводится список наиболее распространенных систем признаков, используемых для яркостно-геометрического описания выделенных областей. Раздел 4.2 посвящен методам выделения геометрических примитивов. Описано преобразование Хафа для поиска прямых линий, преобразование Хафа для поиска окружностей, рассмотрены различные стратегии анализа аккумулятора Хафа при поиске геометрических примитивов. Далее описаны реализованные в системе NI Vision эвристические процедуры поиска и выделения прямых линий, окружностей, а также линий, близких к прямым, и линий, близких к окружностям. Раздел 4.3 посвящен рассмотрению методов обнаружения на изображении объектов, заданных эталонными изображениями (образцами). Описаны методы корреляционного обнаружения, морфологический подход к сравнению изображений ⇒. П. Пытьева, методы согласованной фильтрации, обобщенное преобразование Хафа, функции привязки эталонов в NI Vision. В разделе 4.4 рассматриваются проблемы калибровки датчиков и последующих метрических измерений на изображениях. Описаны общие методы калибровки изображений, а также процедуры калибровки изображений, доступные в IMAQ Vision. Рассмотрены также функции геометрических и яркостных измерений на изображениях, поддерживаемые NI Vision. Раздел 4.5 посвящен методам считывания символьной информации. Рассмотрена задача оптического распознавания символов (OCR), структура современных систем OCR и различные методы распознавания текстовых символов. Рассмотрена также задача обнаружения и считывания штриховых кодов на изображениях. Описаны различные типы штриховых кодов, а также поддерживаемые системой NI Vision специализированные методы считывания информации аналоговых и цифровых технических индикаторов. Раздел 4.6 содержит описание служебных функций NI Vision, позволяющих осуществлять ручные геометрические измерения и построения, а также работу с растровым изображением и областями интереса.

Глава 5 посвящена цифровой фотограмметрии и бесконтактным измерениям. В разделе 5.1 описаны теоретические основы фотограмметрических измерений. Рассмотрены: математическая модель камеры, калибровка видеокамер, внешнее ориентирование камер, решение задачи стереосоответствия при построении модели поверхности, формирование структурированного подсвета, приведение координат точек снимков к нормальному случаю съемки, определение элементов взаимного ориентирования по внешнему, вычисление пространственных координат точек поверхности, ортотрансформирование цифрового снимка. В разделе 5.2 описан фотограмметрический комплекс для бесконтактных измерений на базе LabVIEW и промышленного компьютера PXI.

Глава 6 посвящена описанию новой промышленной платформы для систем технического зрения Compact Vision System (CVS) фирмы National Instruments. В разделе 6.1 описаны технические характеристики CVS, подключение и настройка CVS, а также алгоритм создания приложений для CVS. В разделе 6.2 приводится пример приложения для данной вычислительной системы: фотограмметрический комплекс на базе CVS.

В главе 7 рассматриваются многочисленные примеры практических систем машинного и компьютерного зрения, разработанных в последние годы в лабораториях компьютерного зрения ГосНИИ Авиационных систем и ЗАО «Институт

информационных технологий». Некоторые из этих систем были созданы с использованием средств LabVIEW и IMAQ Vision, другие – с использованием иных программных средств. В контексте данной главы принципиально то, что во всех описанных примерах практических приложений решались достаточно типовые задачи машинного и компьютерного зрения, а также то, что технические характеристики описанных прикладных систем являются типовыми и удовлетворительными на современном этапе развития технологии, программных и аппаратных средств машинного зрения. Примеры приложений сгруппированы по следующим крупным тематическим блокам: автоматизация измерений и технический контроль; машинное зрение, видеонаблюдение и биометрия; обработка документов, распознавание текста и штриховых кодов; медицинские приложения.

Данная книга появилась в результате совместных усилий большого коллектива авторов, представленных ниже:

- Д.т.н., чл.-корр. РАН, проф. Желтовым С. Ю. подготовлены теоретические разделы главы 1 и список литературы по тематике машинного зрения.
- К.т.н., с.н.с. Визильтером Ю. В. написаны введение и основные теоретические разделы глав 2, 3, 4, 7.
- Князем В. А. написаны теоретические разделы главы 5.
- Ходаревым А. Н. написан ряд разделов, посвященных работе с платформой NI Vision, в частности разделы 6.1, 6.2, раздел 1.4 и ряд примеров использования функций IMAQ Vision в главах 2, 3, 4.
- Моржиным А. В. описан ряд примеров использования функций IMAQ Vision в главах 2, 3, 4.
- Максимовым А. А. написан раздел 5.2.
- Каратеевым С. Л. написан ряд параграфов раздела 3.3.
- К.ф-м.н. Ортюковым С. И. представлен обзор штриховых кодов в разделе 4.5.
- К.т.н. Морзеевым Ю.В. написан обзор биометрических методов в разделе 7.2.
- Раздел 1.2 подготовлен к.т.н., доц. АлтГУ А. Я. Сурановым.
- Раздел 1.3 подготовлен сотрудником представительства фирмы National Instruments в России В. Гурьевым.
- Большую помощь в создании книги в целом оказали сотрудники представительства фирмы National Instruments в России во главе с А. Спиридоновым, а впоследствии – с П. Сепояном.
- В создании и описании практических приложений, описанных в главе 7, принял участие большой коллектив сотрудников ГосНИИ Авиационных систем (ГосНИИАС) и ЗАО «Институт информационных технологий».
- На протяжение 2003–2006 годов материалы, вошедшие впоследствии в состав данной книги, широко использовались в качестве курсов лекций, учебных и методических пособий в учебном процессе базовой кафедры №539 при ГосНИИАС «Авиационно-космические системы обработки информации и управления» Московского государственного института радиотехники, электроники и автоматики (МИРЭА) под руководством заведующего кафедры, д.т.н., чл.-корр. РАН, Себрякова Г. Г.

Авторы выражают глубокую признательность и благодарность всем, кто помогал им в работе над этой книгой.

Машинное зрение, LabVIEW и NI Vision

1

1.1. Проблематика машинного зрения

В последние годы цифровая обработка и цифровой анализ изображений находят все большее применение в различных областях науки и техники, таких как интеллектуальные робототехнические комплексы, системы промышленного контроля, системы управления движущимися аппаратами, обработка данных дистанционного зондирования, биомедицинские исследования, новые технологии обработки документов и множество других. Далее мы будем использовать обобщающий термин «*машинное зрение*» (Machine vision) как понятие, наиболее полно объемлющее круг инженерных технологий, методов и алгоритмов, связанных с задачей интерпретации сцены наблюдения по ее двумерным проекциям (изображениям), а также как практическое использование результатов этой интерпретации.

Проблематика машинного зрения столь привлекательна для современных исследователей по той причине, что аппаратные возможности, предоставляемые в данной области последними достижениями электроники и вычислительной техники, достигли такого уровня, что они уже во многом приближаются к «техническим характеристикам» человека. Разрешение многих сенсоров для получения видеоинформации практически соответствует числу элементов сетчатки глаза человека, а возможности ЭВМ и специальных процессоров близки к характеристикам «вычислительных мощностей», используемых для обработки изображений в мозгу. Со своей стороны, биомеханика вплотную подошла к разработке сложных механических манипуляторов, достаточных для имитации моторной деятельности человека по управлению различными техническими системами. И таким образом на пути к осуществлению заветной фантастической мечты будущего – созданию сложных автономных робототехнических комплексов, «интеллектуальных машин», функционирующих в реальном масштабе времени, – стоит необходимость решения большей частью лишь одной принципиальной задачи – разра-

ботки методов и алгоритмов «понимания» изображений. Однако именно эта задача во многих случаях оказывается и наиболее трудной.

1.1.1. Задачи машинного зрения

Удивительная сложность проблемы «понимания изображений» может быть проиллюстрирована тем обстоятельством, что ее интеллектуальная (алгоритмическая) составляющая оказалась более трудным «орешком», чем традиционные задачи типа компьютерной игры в шашки или шахматы, которые долгие годы служили полем приложения методов «искусственного интеллекта». Компьютер, на равных играющий в шахматы с чемпионами мира, уже создан, а компьютерной программы, «понимающей» любую видимую сцену, пока нет. Более того, создание такой программы в перспективе ближайших десятилетий даже не предвидится научным сообществом. Это связано, по-видимому, со сложностью основного предмета, находящегося в центре внимания данной дисциплины, а именно – двумерного изображения.

Скорее правилом, чем исключением, является отсутствие у информационного семантического содержания изображения какой-либо «причинной» или динамической модели формирования, в том смысле, что это информационное семантическое содержание возникает не под действием каких-либо физических законов, описывающихся математическими уравнениями. Информационное наполнение изображения проявляется в виде бесконечного разнообразия *яркостно-геометрических структур*, модели порождения которых могут просто отсутствовать. Особенно сложной задачей является «понимание» объектов, присутствующих в сцене наблюдения. Обнаружение и идентификация многих типов таких объектов, например зданий и дорог на аэрофотоснимках, превратились даже в отдельные направления исследований. Так, только проблеме выделения зданий на изображениях были посвящены в последние годы несколько крупных международных конференций, поставивших лишь ряд новых проблем в дополнение к существующим [72, 107, 171]. Таким образом, следует признать, что общая теория «понимания изображений» за последние 30–40 лет еще не вышла из юношеского возраста, и то ее состояние, которое может быть зафиксировано сегодня, – это сочетание ряда нерешенных к настоящему моменту теоретических задач, с одной стороны, и большого числа идей и подходов, далеких от окончательного вида хорошо разработанной теории, – с другой.

В то же время для решения многих практически важных задач машинного зрения общая проблема «понимания изображений» может быть редуцирована к гораздо более простой и ясной проблеме *обнаружения* и *распознавания* или *измерения* по одному или нескольким изображениям объектов, удовлетворяющих некоторому заранее известному *модельному* описанию. Теория машинного зрения предлагает целый ряд различных модельных описаний наблюдаемых объектов, которые могут быть использованы для их обнаружения и измерения. В литературе описан широкий спектр таких моделей – от простейших признаковых описаний до высоко-специализированных и изощренных структурных моделей.

Однако общий метод составления работоспособных моделей по вышеупомянутым причинам отсутствует. Отсюда следует, что разработка и использование моделей, пригодных для эффективного решения задачи обнаружения соответствующих объектов, в значительной степени остается на грани науки и искусства, то есть требует особого «know-how» или, другими словами, знания предметной области, отражающего многолетний опыт исследований по решению частных задач.

Решение задачи автоматического выделения сложных объектов открывает перед системами «машинного зрения» огромное число потенциальных областей применения, таких как промышленная инспекция и контроль качества, робототехника, навигация и транспортировка, медицина и биомеханика, инженерный труд, автоматизация проектирования и множество других. При этом сама задача обнаружения объектов является, безусловно, базовой, но частной технологической задачей по отношению ко всему комплексу основных целевых задач, которые в общих чертах могут быть сформулированы следующим образом:

- обнаружение объектов и изменений в сцене наблюдения;
- высокоточные измерения элементов сцены;
- слежение за объектами;
- самоориентация и самопозиционирование;
- реконструкция поверхностей и обнаружение трехмерных структур;
- описание сцены и идентификация объектов;
- организация зрительной обратной связи при работе управляемых устройств, манипуляторов или мобильных роботов в изменчивой среде.

1.1.2. Уровни и методы машинного зрения

На протяжении десяти-пятнадцати последних лет в алгоритмическом аспекте последовательность действий по обработке изображения принято рассматривать в согласии с так называемой *модульной парадигмой* [18]. Эта парадигма, предложенная Д. Марром на основе длительного изучения механизмов зрительного восприятия человека, утверждает, что обработка изображений должна опираться на несколько последовательных уровней восходящей информационной линии: от «иконического» представления объектов (растровое изображение, неструктурированная информация) – к их символическому представлению (векторные и атрибутивные данные в структурированной форме, реляционные структуры и т. п.). Исходя из этого, в области машинного зрения принято выделять следующие основные этапы обработки данных:

- предобработка изображений;
- *сегментация*;
- выделение геометрической структуры;
- определение относительной структуры и семантики.

Связанные с этими этапами уровни обработки обычно называются соответственно: обработка нижнего уровня, среднего уровня, высокого уровня. В то время как алгоритмы обработки нижнего уровня (фильтрация простых шумов, гис-

тограммная обработка) могут рассматриваться как хорошо проработанные и детально изученные, алгоритмы среднего уровня (сегментация) продолжают сегодня оставаться центральным полем приложения инженерных и исследовательских усилий. За последние годы значительный прогресс был достигнут по отношению к проблемам сопоставления точек и фрагментов изображений (matching) [89, 153], выделения признаков внутри малых фрагментов [85, 126, 137], высокой точности 3D-позиционирования точек [88, 91], что подразумевает соответствующее моделирование и калибровку датчиков и их комбинаций, выделение простых яркостно-геометрических структур типа «точка», «край», «пятно», «прямая линия», «угол» [49, 50, 117, 126, 149].

Методы обработки высокого уровня, относящиеся собственно к «пониманию изображений», находятся еще в начальной фазе развития и по-прежнему представляют собой «вызов» для сообщества исследователей в области компьютерного зрения и искусственного интеллекта. Безусловно, перспектива создания будущих поколений «интеллектуальных машин» в основном зависит от дальнейшей разработки именно этого круга алгоритмов.

В настоящее время известно несколько основных алгоритмических подходов и математических формализмов, используемых при разработке практических систем анализа изображений. Это гистограммные преобразования, анализ проекций, линейная и нелинейная фильтрация изображений, яркостная и текстурная сегментация, корреляционное обнаружение и согласованная фильтрация, морфологический подход Ю. П. Пытьева, математическая морфология Серра, метод «нормализации фона», преобразование Хафа, структурно-лингвистический подход и ряд других. Большинство этих методов будут рассмотрены в данной книге.

Значительный вклад в разработку методов и алгоритмов обработки изображений и машинного зрения внесли работы Ярославского М. П., Бакута П. А., Злобина В. К., Баклицкого В. К., Лабунца В. Г., Левшина В. Л., Пытьева Ю. П., Серра, Р. Харалика, Е. Дэвиса, У. Гренандера, К. Ту и др. За последние десятилетия создано множество успешных систем машинного зрения, в которых в тех или иных сочетаниях реализованы упомянутые подходы и парадигмы. Однако единого математического формализма и единой общепризнанной методики разработки алгоритмов анализа изображений по-прежнему не существует, и, следовательно, наука об обработке изображений все еще находится на одной из начальных стадий своего развития, переживая период бурного роста, чреватый возможностью появления в любой момент новых, самых неожиданных и революционных методик и теорий.

1.1.3. Сопряженные технические дисциплины

Наряду с термином «машинное зрение» в литературе часто употребляются такие понятия, как зрение роботов (robot vision), компьютерное зрение (computer vision), обработка изображений (image processing), понимание изображений (image understanding). Рассмотрим взаимоотношения этих понятий, с тем чтобы установить смысл терминов, которые будут использоваться здесь в дальнейшем.

Компьютерное зрение представляет собой научную дисциплину, изучающую теорию и базовые алгоритмы анализа изображений и сцен.

Машинное зрение следует рассматривать как гораздо более комплексную и технологическую область научных и инженерных знаний, охватывающую все проблемы разработки практических систем: выбор схем освещения исследуемой сцены, выбор характеристик датчиков, их количества и геометрии расположения, вопросы калибровки и ориентирования, выбор или разработка оборудования для оцифровки и процессорной обработки, разработка собственно алгоритмов и их компьютерная реализация – то есть весь круг сопутствующих задач.

Зрение роботов мы предлагаем трактовать как более узкую область технологий машинного зрения, а именно ту их часть, которая обеспечивает функционирование систем машинного зрения в условиях жестких временных ограничений. К этому понятию, безусловно, относятся проблемы разработки основанных на изображениях информационных систем, входящих в состав систем управления сложными динамическими объектами (самолет, автомобиль, системы контроля технических и технологических процессов на производстве), так как необходимость формирования обратных связей по результатам обработки входных изображений в системах управления, очевидно, требует их быстрого анализа в реальном масштабе времени.

Традиционный термин «*обработка изображений*» чаще употребляется в последние годы не как обозначение научной дисциплины, а как указатель на предметную область. Наметилась также тенденция использования этого термина для обозначения обработки нижнего уровня, когда результатом обработки изображения снова является изображение. В то же время термин «*понимание изображений*» употребляется для обработки верхнего уровня, часто в контексте применения методов искусственного интеллекта [42, 52].

Машинное зрение тесно связано с целым рядом смежных дисциплин. Это – уже упоминавшиеся ранее компьютерное зрение, искусственный интеллект, а также распознавание образов (pattern recognition) и цифровая фотограмметрия (digital photogrammetry).

Цифровая фотограмметрия, бурно развивающаяся в последние годы дисциплина, пришла на смену аналитической фотограмметрии. В то время как классическая аналитическая фотограмметрия изучала в основном метрические соотношения между точками снимков и реальной сцены, современная цифровая фотограмметрия, особенно так называемая close-range (не очень удачный перевод «короткобазисная», но другого русскоязычного термина пока нет) ставит самые сложные задачи анализа и 3D-описания сцены по видеоданным оптических сенсоров. В широком смысле многие ведущие западные ученые в настоящее время используют термины digital close-range photogrammetry и machine vision практически как синонимы. В узком смысле считается, что цифровая фотограмметрия концентрируется прежде всего не на вопросах «понимания» сцены или обнаружения объектов, а на вопросах высокоточного измерения различных ее элементов и реконструкции форм трехмерных поверхностей с использованием стерео и многокамерной съемки, а также специального структурированного подсвета.

1.1.4. Требования к алгоритмам машинного зрения

Приступая к изучению предмета «машинное зрение», необходимо помнить о том, что это, прежде всего, инженерная прикладная дисциплина. Всякая прикладная техническая дисциплина отличается от фундаментальной теоретической дисциплины наличием конкретных практических ограничений, налагаемых на развиваемые в ее рамках методы, и на порождаемые ею инструменты решения практических задач.

Рассмотрим требования к методам и алгоритмам машинного зрения на примере наиболее специфической группы алгоритмов – алгоритмов обнаружения объектов на изображениях. При этом будем рассматривать три основных типа требований:

- робастность;
- точность;
- вычислительная реализуемость.

Робастность

Разрабатываемые нами алгоритмы, решающие прикладные задачи машинного зрения, должны работать на реальных изображениях. В то же время каждый, кто на практике пытался конструировать алгоритмы обработки реальных изображений, хорошо знает, насколько изменчивы и не формализуемы могут быть факторы, влияющие на процесс их обработки и анализа. Вот только основные из этих факторов:

- *помехи и «шум»* – имеют десятки видов источников возникновения, к числу которых можно отнести несовершенство сенсоров приемо-передающей аппаратуры и аппаратуры оцифровки изображений, трудные условия съемки, недостаток освещения и ряд других;
- *сложный текстурированный фон*, на котором должно происходить обнаружение объектов, например обнаружение штриховой наклейки на схожей с ней по структуре газетной странице и т. п.;
- *эффекты загораживания (заслонения)* одних объектов другими объектами, как правило, не определенной заранее формы, например – облако на космофотоснимке и т. п., загораживающие помехи;
- искажающие *оптические эффекты* в виде различных расфокусировок, дисторсий объективов, ракурсных искажений и др.;
- *эффекты резкой смены освещения*, блики, тени, особенно в динамически меняющихся сценах;
- *разнообразие и изменчивость самих объектов* – переменная структура (как у текстовых строк, автомобильных номеров или штриховых кодов), возможные дефекты, временные изменения формы (сгибание-разгибание конечностей, движение механических частей машин), вегетационные циклы для растительности и т. п.;
- *эффекты изменения среды между сенсорами и объектами наблюдения* – задымление, атмосферные осадки, пыль, искусственные помехи и многое другое;

- *несинхронность регистрации и обработки данных* в динамических задачах, связанная с ограничениями быстродействия компьютерных средств хранения и анализа изображений – особенно критична для промышленных приложений с заданным временем реакции на событие. Сюда же можно отнести также сбои в компьютерных программах обработки.

Даже беглый анализ приведенных факторов легко установит практическую невозможность их полного формального математического описания – вероятностного, радиометрического или геометрического.

Отсутствие формализованного описания ключевых факторов, вносящих неопределенность в процесс обработки, приводит к тому, что, в отличие от многих других областей обработки сигналов, говорить о существовании единственного оптимального алгоритма для решения той или иной задачи обработки изображений практически невозможно. Так, например, только для задачи сегментации графики и текста на цифровых изображениях документов придуманы сотни принципиально различных алгоритмов. Представим себе, что существует несколько алгоритмов, достигающих примерно одинаковых результатов на «идеальных» изображениях. Тогда возникает естественный вопрос, как сравнить эти алгоритмы по качеству их работы. При разработке реальных алгоритмов в настоящее время стандарт «де-факто» состоит в проверке эффективности работы сконструированных алгоритмов на больших выборках реальных данных или изображениях, содержащих, по возможности, все неприятные ситуации. Такие алгоритмы, которые обладают устойчивостью к значительным искажениям и меняющимся факторам, принято называть устойчивыми, или *робастными*. Робастность следует считать первым основным практическим требованием при разработке алгоритмов машинного зрения.

Точность

Второе важнейшее свойство, которым должны, как правило, обладать алгоритмы обнаружения объектов на изображениях, можно определить как *точную локализацию* объектов, подлежащих обнаружению, или контуров объектов, подлежащих измерению. Это означает, что необходимо не просто выделить объект, но и точно указать в системе координат изображения (или сцены) его положение и размеры в каком-либо смысле. Несколько неясное толкование «локализации», данное выше, связано с тем, что по сравнению со своей эталонной моделью объект может быть заметно искажен геометрически, причем аналитическая модель искажения может отсутствовать. Характерные примеры – штриховой код, находящийся на смятой упаковке, пятно на коже и т. п. В этих случаях локализация объекта является нетривиальной задачей. В более простой ситуации, при аналитически заданной с точностью до параметров геометрии искажений, под точной локализацией можно понимать знание о положении какой-либо характерной точки объекта и параметрах геометрии искажения (поворот, аффинные коэффициенты, изгибы и др.). При этом встречающиеся случаи ошибок локализации целесообразно разделить на две группы – нормальные и аномальные ошибки.

Нормальная ошибка – это правильная локализация объекта с некоторой позиционной или параметрической неточностью, характеризуемой количественными оценками. Для объектов, характеризуемых габаритными размерами, большими чем 3×3÷5×5, элементов изображения, позиционные нормальные ошибки могут быть значительно меньше размера элемента изображения, уменьшаясь с величиной объекта. В этом случае принято говорить о *субпиксельной локализации,* или *субпиксельных измерениях.* Это особенно важно для задач стереообнаружения; так, при малых параллаксах 3D-объектов субпиксельная привязка самым существенным образом определяет точность измерения их пространственного положения.

К аномальным ошибкам следует отнести ситуацию перепутывания объектов или возникновение артефактов, что связано с фатальными количественными ошибками позиционирования или просто ложным обнаружением. Требования по исключению или ограничению уровня аномальных ошибок составляют очень важную часть требований к алгоритмам обнаружения.

Заметим также, что требования по точности предъявляются и к алгоритмам обработки изображений низкого уровня. Например, во многих измерительных задачах фильтры, устраняющие помехи, не должны при этом существенно изменять видимое положение контуров объекта. Требование «сохранения краев» или, в более общем виде, сохранения неизменными измеряемых характеристик полезного сигнала является принципиальным требованием «по точности», предъявляемым в машинном зрении к алгоритмам нижнего уровня, поскольку, если данные были «испорчены» уже на нижнем уровне обработки, то на более высоких уровнях анализа никакие самые совершенные математические методы помочь не смогут, и задача в целом не будет решена с необходимой точностью.

Вычислительная реализуемость

Несмотря на колоссальный путь, который прошла в целом вычислительная техника за последние декады двадцатого и первые годы двадцать первого веков, несмотря на создание обширной специализированной процессорной базы для обработки изображений, в основной массе приложений, особенно промышленных приложений реального времени, характеристики вычислителей и их свойства все еще далеки от желаемых.

Важное отличие, присущее процедурам обработки и анализа изображений по сравнению с задачами распознавания или интерпретации уже сегментированного образа, заключается в том, что обнаружение и измерение в практических задачах всегда связано с процедурой *поиска* объекта. Именно реализация процедуры поиска объекта связана с угрозой лавинообразного роста потребного числа вычислений. Проиллюстрируем это на примере простой задачи поиска объекта на основе сравнения с растровым эталоном или шаблоном (template). Если построить какой-либо функционал соответствия между объектом размером $M \times M$ и фрагментом $M \times M$ из изображения $N \times N$, то простой перебор фрагментов требует количества вычислений не менее чем $M^2 \times N^2$ операций, что составляет, например,

10 миллиардов операций при размере объекта 50×50, а изображения – 2000×2000 элементов. Даже принимая во внимание значительное увеличение возможностей современных компьютеров, такие объемы вычислений занимают секунды машинного времени, что далеко выходит за пределы требований систем обнаружения реального времени, предназначенных для таких задач, как навигация и наведение, инспекционный контроль, обработка машиночитаемых документов и т. п.

Более того, реальные задачи обработки визуальной информации изобилуют дополнительными степенями свободы, когда искомая яркостно-геометрическая структура на изображении может иметь не только произвольные положение, угловую ориентацию и масштаб, но и подвергаться разным преобразованиям, не только аффинным или проективным, но и гораздо более сложным, таким как различные «резиновые» модели или «коробление» (warping – искажение, напоминающее смятую бумагу). Все это многократно увеличивает потребное для перебора время расчетов и требует применения качественно новых идей по организации процесса анализа изображения.

Оценивая размер изображения, предназначенного для обработки, например, как 1000 на 1000 элементов, что не является чем-то необычным для современных видеодатчиков (можно вспомнить, что бытовые цифровые фотоаппараты и даже мобильные телефоны давно перешли отметку 2 мегапикселя в ПЗС матрице), мы получим даже для простейших операций характерную оценку количества потребных вычислений порядка нескольких гигабайтов операций на кадр. Причем приложениям реального времени необходимо выполнять эти вычисления в темпе кадровой развертки, что приводит к оценке потребного быстродействия около 50 Gflop/sec. Сами по себе эти оценки не являются запредельными для мощных современных компьютеров, однако следует учесть, что основные приложения машинного зрения находятся в тех областях, где массогабаритные и стоимостные характеристики конструируемых устройств по определению ограничены и должны быть весьма невелики.

Таким образом, вычислительная реализуемость алгоритмов машинного зрения по-прежнему относится к числу наиболее важных факторов, которые необходимо учитывать при их разработке.

Для более углубленного знакомства с методами и алгоритмами машинного зрения мы рекомендуем список литературы, приведенный в конце данной книги.

1.1.5. Роль специальных программных средств в разработке приложений машинного зрения

Как мы уже отмечали выше, единого математического формализма и единой общепризнанной методики разработки алгоритмов в области машинного зрения пока не существует, и вряд ли они будут разработаны в ближайшие десятилетия.

Это, с одной стороны, является серьезной проблемой, затрудняющей и замедляющей разработку новых приложений и практических систем машинного зрения, а с другой стороны, делает сам процесс разработки алгоритмических и про-

Конец ознакомительного фрагмента.
Приобрести книгу можно
в интернет-магазине «Электронный универс»
(e-Univers.ru)