

*Моей семье /Д. Форсайт/
Камилле и Оскару /Д. Понс/*

Предисловие

Компьютерное зрение — это пограничная область знаний. И как всякая пограничная область она интересна для изучения и непредсказуема; здесь часто нет авторитетов, на которые можно сослаться, — многие полезные идеи не имеют под собой теоретической основы, а некоторые теории бесполезны на практике; изученные области весьма различаются, и часто кажется, что всякая связь между ними отсутствует. Тем не менее, в этой книге мы попытались представить в некоторой степени упорядоченную картину этой области знаний.

В нашем понимании “компьютерное зрение”, или просто “зрение” (заранее просим прощения у специалистов по зрению людей или животных) — это вид деятельности, в котором для извлечения данных применяются статистические методы и используются модели, построенные с помощью геометрии, физики и теории обучения. Таким образом, на наш взгляд компьютерное зрение основывается на четком представлении о камерах и физическом процессе формирования изображения (часть I), получении простых выводов на основе изучения набора отдельных пикселей (часть II), умении суммировать информацию, полученную из множества изображений (часть III), упорядочении группы пикселей с целью их разделения или получения информации о форме (часть IV), распознавании объектов с помощью геометрической информации (часть V) или вероятностных методов (часть VI). Компьютерное зрение применяется довольно

широко как в относительно старых областях (например, управление мобильными роботами, промышленные средства наблюдения, военные приложения), так и в сравнительно новых (взаимодействие человек/компьютер, поиск изображения в цифровых библиотеках, анализ медицинских изображений и реалистичная передача смоделированных сцен в компьютерной графике). Некоторые из этих приложений обсуждаются в части VII.

ЗАЧЕМ ИЗУЧАТЬ ЗРЕНИЕ

Отличительная черта компьютерного зрения — это извлечение описаний из изображений или последовательности изображений. Это весьма полезная особенность. Процесс снятия изображения обычно недеструктивен, кроме того, он достаточно прост и на сегодняшний момент недорог. Описания, необходимые пользователям, могут в большой мере зависеть от области их применения. Например, такой аспект компьютерного зрения, как определение структуры по движению, позволяет из серии изображений получить представление о том, что изображено на рисунке и как движется камера. В индустрии развлечений подобные методы применяются для отсеивания движения и построения трехмерных компьютерных моделей зданий с сохранением структуры. Эти модели применяются там, где нельзя использовать настоящие здания (их поджигают, взрывают и т.п.). С помощью небольшого числа фотографий можно получить хорошие, простые, точные и удобные модели. Рассмотрим другую ситуацию: люди, желающие контролировать работу мобильных роботов. В этом случае сведения об области, где используется робот, обычно не представляют значительного интереса, важно лишь местонахождение робота в этой области. Таким образом, здесь отсеивается информация о структуре и отслеживается движение, что позволяет определить точное местонахождение робота.

Есть еще целый ряд других важных областей применения компьютерного зрения. Это, например, работа с медицинскими изображениями: создание программных систем, которые могут улучшать набор изображений, выявлять на них важные моменты или события либо визуализировать информацию, полученную из изображений. Другая важная область — различные технические проверки, когда по изображениям объектов определяется, соответствуют ли объекты спецификации. Третья сфера применения компьютерного зрения — интерпретация фотографий, сделанных со спутника, как в военных целях (например, может потребоваться программа, выявляющая интересные, с военной точки зрения, события в указанном регионе или определяющая вред, нанесенный в результате бомбардировки), так и в гражданских (какой урожай кукурузы будет в этом году? сколько осталось тропических лесов?). Четвертая область — это упорядочение и структурирование коллекций картин. Теперь пользователь уже умеет находить нужную информацию в текстовых библиотеках (хотя и здесь имеется ряд сложных, нерешенных вопросов), но еще не всегда знает, что делать с библиотеками статических или движущихся изображений.

Компьютерное зрение сейчас находится в особой точке своего развития. Эта тема стала популярной еще в 1960-х, но только недавно появилась возможность создания полезных компьютерных программ, использующих идеи компьютерного зрения, поскольку компьютеры и программы обработки изображений стали доступны большому количеству пользователей. Не так давно для получения хорошего цветного цифрового изображения нужно было потратить не один десяток тысяч долларов; сейчас для этого нужно не более нескольких сотен. Не так давно цветной принтер можно было найти только в некоторых исследовательских лабораториях; сейчас их используют повсеместно. Таким образом, проводить серьезные исследования и решать многие повседневные задачи (например, упорядочить коллекцию фотографий, создать трехмерную модель окружающего мира, управлять и вносить изменения в коллекцию видеозаписей) теперь можно с помощью методов компьютерного зрения. Наши знания по геометрии и физике зрения и, что еще важнее, умение их применять в нужном направлении значительно развились. Мы уже можем решать задачи, интересующие многих людей, хотя еще не решена ни одна из действительно сложных задач, кроме того, существует множество простых задач, которые можно использовать для поддержания интеллектуальной формы людей, бьющихся над решением трудных задач. Итак, сейчас самое время начать изучение этого предмета!

Что есть в этой книге

Эта книга предназначена как для специалиста по компьютерному зрению, так и для широкого круга читателей. Она окажется полезным источником информации для тех, кто работает с изображениями и робототехникой, а также занимается вычислительной геометрией, компьютерной графикой, обработкой изображений. Эта книга также будет доступна для студентов, поверхностно интересующихся зрением. В каждой главе освещается отдельный аспект данного предмета, и, как видно из табл. 1, главы относительно независимы между собой. Это означает, что книгу можно читать как выборочно, так и “от корки до корки”. Материал в каждой главе подается в порядке возрастания сложности. В конце каждой главы приводятся краткие примечания, содержащие любопытные исторические сведения, и полезные замечания, относящиеся к рассматриваемому вопросу. В книге описаны идеи, представляющие интерес или кажущиеся перспективными, авторы попытались как можно более доступно связать теорию с реальными приложениями. При изложении материала внимание акцентировалось на базовых вопросах геометрии и физики изображений, а также прикладной статистики, поскольку именно эти области оказали огромное влияние на компьютерное зрение.

Читателю, который решил изучить этот труд досконально, следует учесть, что данная книга содержит очень большой набор информации, поэтому усвоить полученные знания за один семестр весьма сложно. Конечно, будущим (или настоящим) профессионалам в области компьютерного зрения стоит прочитать каждое слово, выполнить все упражнения и сообщить обо всех обнаруженных ошибках для исправления второго издания (которое наверняка будет велико-

ТАБЛИЦА 1. Связь между главами: читать главу будет трудно, если вы предварительно не изучите главы, указанные в графе “требуемое”; если при прочтении главы обнаружится один-два непонятных момента, скорее всего они освещены в главах, указанных в графе “полезное”

Часть	Глава	Требуемое	Полезное
I	1: Камеры		
	2: Геометрические модели камер	1	
	3: Геометрическая калибровка камер	2	
	4: Радиометрия — измерение света		
	5: Источники, тени и затенение		4, 1
	6: Цвет		5
II	7: Линейные фильтры		
	8: Определение краев	7	
	9: Текстура	7	8
III	10: Геометрия множественных проекций	3	
	11: Стереозрение	10	
	12: Аффинная структура по движению	10	
	13: Проективная структура по движению	12	
IV	14: Сегментация посредством кластеризации		9, 6, 5
	15: Сегментация посредством выбора модели		14
	16: Сегментация на основе вероятностных методов		15,10
	17: Сопровождение на основе динамических моделей		
V	18: Зрение на основе модели	3	
	19: Гладкие поверхности и их контуры	2	
	20: Аспектные графы	19	
	21: Дистанционные данные		20, 19, 3
VI	22: Поиск шаблонов с помощью классификаторов		9, 8, 7, 6, 5
	23: Распознавание по связям между шаблонами		9, 8, 7, 6, 5
	24: Геометрические шаблоны	2, 1	16, 15, 14
VII	25: Приложение: поиск в цифровых библиотеках		16, 15, 14, 6
	26: Приложение: визуализация на основе изображений	10 13, 12, 11, 6, 5, 3, 2, 1	

лепной книгой, о которой стоит подумать уже сейчас!). Поскольку для изучения компьютерного зрения не требуются глубокие познания в математике, здесь не нужен и слишком обширный математический аппарат. Мы попытались сделать эту книгу самодостаточной, чтобы читателям с математическим образованием на уровне выпускника инженерного вуза не приходилось обращаться к другим источникам. Математические выкладки сведены к необходимому минимуму, — в конце концов, это книга о компьютерном зрении, а не о прикладной матема-

тике! — поэтому требуемые формулы включены в основной текст главы, а не вынесены в отдельные приложения.

Взаимозависимость между главами сведена к минимуму, чтобы читателю, которого интересуют отдельные темы, не пришлось штудировать всю книгу. Разумеется, сделать каждую главу полностью самостоятельной невозможно, поэтому в табл. 1 показаны связи между главами.

Чего нет в этой книге

По компьютерному зрению доступно немало материалов, и было нелегко создать такой труд, который могли бы осилить простые смертные. Чтобы сделать это, пришлось урезать материал, пренебречь некоторыми темами и т.д. В самый последний момент было решено убрать целых две главы: введение в теорию вероятности с описанием соответствующих следствий, а также перечисление методов сопровождения объектов с нелинейной динамикой. Эти главы можно найти на Web-сайте книги <http://www.cs.berkeley.edu/~daf/book.html>.

Некоторые темы мы обошли либо по собственному усмотрению, либо потому, что наши возможности уже были исчерпаны, о некоторых темах мы узнали слишком поздно, чтобы вставить их в книгу, а какие-то главы необходимо было сократить. Вместо подробных рассуждений по вопросам, которые представляют в основном исторический интерес, в конце каждой главы помещены краткие исторические примечания. Никто из нас не считает себя хорошим “археологом мысли”, так что у некоторых идей история может быть намного глубже, чем мы это показали. В книге также отсутствует подробное описание деформируемых таблиц и мозаик — две темы, которые имеют большое практическое значение; мы попытаемся ввести их во второе издание данной книги.

БЛАГОДАРНОСТИ

Выходу этой книги в свет мы обязаны многим людям. Несколько анонимных рецензентов читали черновики этой книги и сделали ценные замечания, за что мы чрезвычайно благодарны. Наш редактор, Алан Апт (Alan Apt), упорядочил сделанные рецензии, в чем ему весьма помог Джейк Вард (Jake Warde). Мы благодарим их обоих. Лесли Гален (Leslie Galen), Джо Альбрехт (Joe Albrecht) и Диана Пэриш (Dianne Parish) из компании Integre Technical Publishing помогли нам во многих вопросах, связанных с иллюстрациями и корректурой. Некоторые приведенные в книге изображения мы получили из Master Photos Collection IMSI, 1895 Francisco Blvd. East, San Rafael, CA 94901-5506, USA. Подготавливая библиографию, мы использовали материал по компьютерному зрению Кейта Прайса (Keith Price), который можно найти по адресу <http://iris.usc.edu/Vision-Notes/bibliography/contents.html>.

Наши коллеги изучали как отдельные главы, так и всю книгу в целом, после чего вносили ценные и детальные предложения. Мы благодарим Кобуса Барнарда (Kobus Barnard), Маргарет Флек (Margaret Fleck), Дэвида Кригмена (David

Kriegman), Джайтендру Малик (Jitendra Malik) и Эндрю Циссермана (Andrew Zisserman). Некоторые наши студенты также вносили предложения, выдвигали идеи по поводу рисунков, давали комментарии по корректуре и выполняли другую ценную работу. Мы благодарим Окан Арикан (Okan Arıkan), Себастьяна Блайнда (Sébastien Blind), Марту Цепеда (Martha Cepeda), Стефана Ченни (Stephen Chenney), Френка Чо (Frank Cho), Якапа Генка (Yakup Genc), Джона Гаддона (John Haddon), Сергея Иоффе (Sergey Ioffe), Светлану Лазебник (Svetlana Lazebnik), Кетти Ли (Cathy Lee), Сунг-иль Пэ (Sung-il Pae), Дэвида Паркса (David Parks), Фреда Ротгангера (Fred Rothganger), Аттавис Садсенг (Attawith Sudsang) и слушателей наших курсов по компьютерному зрению при Калифорнийском университете в Беркли. Многие преподаватели в различных университетах в своих лекциях по зрению использовали наши (часто необработанные) черновики этой книги. В число учебных заведений, где студентам пришлось изучать “сырой” материал, вошли Университет Карнеги-Меллона, Стэнфордский университет, Университет штата Висконсин в Мэдисоне, Калифорнийский университет в Санта-Барбаре, Университет Южной Калифорнии и другие, о которых мы не знаем. Мы благодарны редакторам за все полезные комментарии, особенно хочется выделить Криса Брэглера (Chris Bregler), Чака Дайера (Chuck Dyer), Мартиала Хеберта (Martial Hebert), Дэвида Кригмена (David Kriegman), Б.С. Маньюнат (B.S. Manjunath) и Рэма Неватья (Ram Nevatia), которые посылали нам множество подробных и полезных комментариев и исправлений. Книга во многом выиграла благодаря замечаниям Айдин Алалиоглу (Aydin Alalioglu), Сринивас Акелла (Srinivas Akella), Мэри Банич (Marie Vanich), Сержа Белонги (Serge Belongie), Аджиит М. Чаудхари (Ajit M. Chaudhari), Навнит Далал (Navneet Dalal), Ричарда Хартли (Richard Hartley), Глена Хилеи (Glen Healey), Майка Хеса (Mike Heath), Хейли Ибен (Hayley Iben), Стефани Джонквирес (Stephanie Jonquires), Тони Льюиса (Tony Lewis), Бенсона Лимкеткай (Benson Limketkai), Симона Маскела (Simon Maskell), Брайана Милха (Brian Milch), Тамары Миллер (Tamara Miller), Корделии Шмид (Cordelia Schmid), Бриджит (Brigitte) и Джерри Серлин (Gerry Serlin), Илана Шимшони (Ilan Shimshoni), Эрика де Штурлера (Eric de Sturler), Камилло Дж. Тейлор (Camillo J. Taylor), Джеффа Томпсона (Jeff Thompson), Клер Валлат (Claire Vallat), Даниель С. Вилкерсон (Daniel S. Wilkerson), Джинган Ю (Jingahn Yu), Хао Жанг (Hao Zhang) и Женджиоу Жанг (Zhengyou Zhang). Если вы обнаружите опечатку, пожалуйста, сообщите об этом подробно по адресу daf@cs.berkeley.edu, используя в сообщении фразу “book typo”; во втором издании мы попытаемся отблагодарить всех, кто первым обнаружит каждую опечатку.

Мы благодарим П. Бесл (P. Besl), Б. Боуфама (B. Boufama), Дж. Костерия (J. Costeira), П. Дебевека (P. Debevec), О. Фаугераса (O. Faugeras), Я. Генка (Y. Genc), М. Хеберта (M. Hebert), Д. Хьюбера (D. Huber), К. Икеучи (K. Ikeuchi), А. Е. Джонсона (A. E. Johnson), Т. Кенеды (T. Kanade), К. Кутулакосу (K. Kutulakos), М. Левоя (M. Levoy), С. Махамуда (S. Mahamud), Р. Мора (R. Mohr), Х. Моравека (H. Moravec), Х. Мураса (H. Murase), Й. Ох-

ту (Y. Ohta), М. Окутами (M. Okutami), М. Поллефейса (M. Pollefeys), Х. Саито (H. Saito), К. Шмида (C. Schmid), С. Салливана (S. Sullivan), К. Томаси (C. Tomasi), М. Тека (M. Turk) за предоставление оригиналов некоторых рисунков, приведенных в этой книге.

В число самых важных персон из длинного списка тех, кому должен Д. Форсайт, входят Джеральд Алансвайт (Gerald Alanthwaite), Майк Бреди (Mike Brady), Том Фейр (Tom Fair), Маргарет Флек (Margaret Fleck), Джайтендра Малик (Jitendra Malik), Джо Мунди (Joe Mundy), Майк Родд (Mike Rodd), Чарли Ротвелл (Charlie Rothwell) и Эндрю Циссерман (Andrew Zisserman). Дж. Понс выражает признательность Оливеру Фаугерасу (Olivier Faugeras), Майку Бреди (Mike Brady) и Тому Бинфорду (Tom Binford). Он также хотел бы поблагодарить за помощь Шерон Коллинз (Sharon Collins). Без нее эта книга, как и большинство его работ, возможно никогда не была бы закончена. Оба автора также хотели бы отметить огромное влияние работ Жана Коендеринка (Jan Koenderink).

ВОЗМОЖНЫЕ КУРСЫ ЛЕКЦИЙ

Всю книгу можно изучить за два (довольно насыщенных) семестра. Авторы предлагают изучить одну главу-приложение (возможно, главу о визуализации на основе изображений) в первом семестре, а другую — во втором. Стоит отметить, что полный и подробный курс необходим небольшому числу факультетов, поэтому данная книга построена так, чтобы преподаватель мог выбирать разделы для изучения по своему вкусу. В табл. 2–6 показаны примерные программы курсов, рассчитанных на 15-недельный семестр, хотя, естественно, преподаватели могут перестроить приведенные программы по своему усмотрению.

В табл. 2 предлагается программа односеместрового ознакомительного курса по компьютерному зрению для старшекласников или первокурсников факультетов компьютерных наук, электротехники или иных технических либо научных дисциплин. Студенты получают довольно полное представление о компьютерном зрении, включая такие области его применения, как цифровые библиотеки и визуализация на основе изображений. Основные моменты геометрии и физики формирования изображений рассмотрены полностью, хотя самый сложный теоретический материал опущен. Уровень подготовки студентов может быть очень разным, поэтому на второй или третьей неделе им может понадобиться изучение литературы по теории вероятности (с этой целью на web-сайте данной книги выложена соответствующая глава). Мы предлагаем перенести главы о прикладном применении на конец семестра, но, возможно, главу 20 стоит пройти приблизительно на десятой неделе обучения, а главу 21 — на шестой.

В табл. 3 представлена программа для студентов факультетов компьютерной графики, которые желают ознакомиться с вопросами компьютерного зрения, связанными с их специальностью. Особое внимание уделяется методам, которые позволяют восстановить модели объектов по информации, содержащейся в рисунках; чтобы понять освещаемые вопросы, требуются знания о камерах

ТАБЛИЦА 2. Односеместровый ознакомительный курс по компьютерному зрению для старшеклассников или первокурсников факультетов компьютерных наук, электротехники или иных технических или научных дисциплин

Неделя	Глава	Разделы	Ключевые темы
1	1, 4	1.1, 4 (резюме)	камеры-обскуры, радиометрическая терминология
2	5	5.1–5.5	модели локального затенения; точечные, линейные и плоские источники; фотометрическое стерео;
3	6	все	цвет
4	7, 8	7.1–7.5, 8.1–8.3	линейные фильтры; сглаживание с целью подавления шумов; определение краев
5	9	все	текстура как статистика выходов фильтров; синтез; восстановление формы
6	10, 11	10.1, 11	основы стереометрии; стерео
7	14	все	сегментация как вид кластеризации
8	15	15.1–15.4	подбор линий, кривых; подбор как оценка максимального правдоподобия; устойчивость
9	16	16.1, 16.2	скрытые значения и ОМ-алгоритм
10	17	все	сопровождение с помощью фильтра Кальмана; ассоциация данных
11	2, 3	2.1, 2.2, все из 3	калибровка камер
12	18	все	зрение на основе модели с использованием соответствий и калибровки камер
13	22	все	сопоставление с шаблоном с помощью классификаторов
14	23	все	сопоставление по связям
15	25, 26	все	поиск изображений в цифровых библиотеках; визуализация на основе изображений

и фильтрах, поэтому в программу включены соответствующие главы и разделы. Одним из важных вопросов в современном мире графики стало сопровождение движущихся объектов, описание которого также включено в предлагаемый курс. Мы допускаем, что у студентов может быть разный уровень подготовки, поэтому в книге приводятся некоторые базовые знания из теории вероятности.

В табл. 4 приведена программа для студентов, которые интересуются в основном применением компьютерного зрения. Здесь выделена информация, которая представляет непосредственный практический интерес. Для некоторых студентов на второй или третьей неделе может понадобиться чтение дополнительной литературы по теории вероятности.

В табл. 5 предложена программа для студентов, изучающих познание или искусственный интеллект и желающих получить представление об основных

ТАБЛИЦА 3. Курс для студентов, изучающих компьютерную графику

Неделя	Глава	Разделы	Ключевые темы
1	1, 4	1.1, 4 (резюме)	камеры-обскуры, радиометрическая терминология
2	5	5.1–5.5	модели локального затенения; точечные, линейные и плоские источники; фотометрическое стерео
3	6.1–6.4	все	цвет
4	7, 8	7.1–7.5, 8.1–8.3	линейные фильтры; сглаживание с целью подавления шумов; определение краев
5	9	9.1–9.3	текстура как статистика выходов фильтров; синтез
6	2, 3	2.1, 2.2, все из 3	калибровка камер
7	10, 11	10.1, 11	основы стереометрии; стерео
8	12	все	аффинная структура по движению
9	13	все	проективная структура по движению
10	26	все	визуализация на основе изображений
11	15	все	подбор; устойчивость; алгоритм RANSAC
12	16	все	скрытые значения и ОМ-алгоритм
13	19	все	поверхности и контуры
14	21	все	дистанционные данные
15	17	все	сопровождение, фильтры Кальмана и ассоциация данных

понятиях компьютерного зрения. Эта программа менее нагружена, чем предыдущие, и она допускает более слабый уровень математической подготовки. Примерно на 2–3 неделе обучения студентам придется немного почитать о вероятности (например, главу на web-сайте книги).

В табл. 6 показана программа для студентов факультетов прикладной математики, электротехники или физики. Эта программа предполагает весьма загруженный семестр; некоторые вопросы рассматриваются обзорно (считается, что студенты могут самостоятельно обработать значительный объем математических выкладок). В зависимости от уровня подготовки, некоторым студентам на второй или третьей неделе может понадобиться литература по теории вероятности. Для организации небольшого перерыва в довольно сухой и сложной программе мы предлагаем рассмотреть цифровые библиотеки; вместо этого можно также изучить главу по визуализации на основе изображений, или по дистанционным данным.

ОБОЗНАЧЕНИЯ

В книге использованы следующие обозначения: точки, прямые и плоскости обозначаются греческими или курсивными латинскими буквами (например, P , Δ или Π). Векторы обычно обозначаются латинскими или греческими

ТАБЛИЦА 4. Курс для студентов, интересующихся применением компьютерного зрения

Неделя	Глава	Разделы	Ключевые темы
1	1, 4	1.1, 4 (резюме)	камеры-обскуры, радиометрическая терминология
2	5, 6	5.1,5.3–5.5, 6.1–6.4	модели локального затенения; точечные, линейные и плоские источники; фотометрическое стерео; цвет: физика, человеческое восприятие, цветовые пространства
3	2, 3	все	модели камер и калибровка камер
4	7, 9	все из 7; 9.1–9.3	линейные фильтры; текстура: как статистика выходов фильтров; синтез текстуры
5	10, 11	все	стереометрия
6	12,13	все	аффинная структура по движению; проективная структура по движению
7	13, 26	все	проективная структура по движению; визуализация на основе изображений
8	14	все	сегментация с помощью кластеризации; определение границ и вычитание фона
9	15	все	подбор линий, кривых; устойчивость; алгоритм RANSAC
10	16	все	скрытые значения и ОМ-алгоритм
11	25	все	поиск изображений в цифровых библиотеках
12	17	все	сопровождение, фильтры Кальмана и ассоциация данных
13	18	все	зрение на основе моделей
14	22	все	поиск шаблонов с помощью классификаторов
15	20	все	дистанционные данные

полужирными буквами (например, \mathbf{v} , \mathbf{P} или $\underline{\xi}$), но вектор, соединяющий две точки P и Q , часто будет обозначаться как \overline{PQ} . Строчные буквы обычно употребляются для обозначения геометрических фигур на плоскости изображений (например, p , \mathbf{p} , δ), а прописные — для объектов сцены (например, P , Π). Матрицы обозначаются рукописными латинскими буквами (например, \mathcal{U}).

Привычное трехмерное евклидово пространство будет обозначаться как \mathbb{E}^3 , векторное пространство, образованное наборами из n действительных чисел, которые подчиняются обычным законам сложения и умножения на скаляр, обозначается как \mathbb{R}^n , а для обозначения нулевого вектора в таких пространствах обычно употребляется запись $\mathbf{0}$. Подобным образом, векторное пространство, образованное действительными матрицами $m \times n$, обозначается как $\mathbb{R}^{m \times n}$. При $m = n$ (квадратная матрица) символом Id обозначается единичная матрица, т.е. матрица $n \times n$, диагональные элементы которой равны 1, а все остальные

ТАБЛИЦА 5. Курс для студентов, изучающих познание или искусственный интеллект

Неделя	Глава	Разделы	Ключевые темы
1	1, 4	1, 4 (резюме)	камеры-обскуры; линзы; камеры и глаз; радиометрическая терминология
2	5	все	модели локального затенения; точечные, линейные и плоские источники; фотометрическое стерео; взаимное отражение; определение яркости
3	6	все	цвет: физика, человеческое восприятие, цветовые пространства; постоянство цвета
4	7	7.1–7.5, 7.7	линейные фильтры; выборка; масштаб
5	8	все	определение краев
6	9	все	текстура; представление; синтез; восстановление формы
7	10.1, 10.2	все	основы стереометрии
8	11	все	стереозрение
9	14	все	сегментация с помощью кластеризации
10	15	все	подбор линий, кривых; устойчивость; алгоритм RANSAC
11	16	все	скрытые значения и ОМ-алгоритм
12	18	все	зрение на основе модели
13	22	все	поиск шаблонов с помощью классификаторов
14	23	все	распознавание по связям между шаблонами
15	24	все	геометрические шаблоны по пространственным отношениям

ные — 0. Транспонированная матрица матрицы \mathcal{U} размера $m \times n$ с коэффициентами u_{ij} — это матрица \mathcal{U}^T , имеющая размер $n \times m$ и коэффициенты u_{ji} . Элементы пространства \mathbb{R}^n часто отождествляются с векторами-столбцами или матрицами $n \times 1$, например, $\mathbf{a} = (a_1, a_2, a_3)^T$ — это транспонированная матрица 1×3 (*вектор-строка*), т.е. матрица 3×1 (*вектор-столбец*), или элемент пространства \mathbb{R}^3 .

Скалярное (или внутреннее) произведение двух векторов $\mathbf{a} = (a_1, \dots, a_n)^T$ и $\mathbf{b} = (b_1, \dots, b_n)^T$ в пространстве \mathbb{R}^n определяется следующим образом:

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + \dots + a_n b_n.$$

Его также можно записать в виде произведения двух матриц, т.е. $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$. Под $|\mathbf{a}|^2 = \mathbf{a} \cdot \mathbf{a}$ мы подразумеваем квадрат евклидовой нормы вектора \mathbf{a} , а под d — расстояние, следующее из определения Евклидовой нормы вектора в пространстве \mathbb{E}^n , т.е. $d(P, Q) = |\overrightarrow{PQ}|$. Для матрицы \mathcal{U} в пространстве $\mathbb{R}^{m \times n}$ через $|\mathcal{U}|$ обычно будет обозначаться ее *фробениусова норма*, т.е. корень квадратный из суммы квадратов ее элементов.

ТАБЛИЦА 6. Программа для студентов факультетов прикладной математики, электротехники или физики

Неделя	Глава	Разделы	Ключевые темы
1	1, 4	все	камеры, радиометрия
2	5	все	модели локального затенения; точечные, линейные и плоские источники; фотометрическое стерео; взаимное отражение и примитивы затенения
3	6	все	цвет: физика, человеческое восприятие, цветовые пространства, постоянство цвета
4	2, 3	все	параметры и калибровка камер
5	7, 8	все	линейные фильтры и определение краев
6	8, 9	все	определение краев (окончание); текстура: представление, синтез, восстановление формы
7	10, 11	все	стереометрия
8	12, 13	все	структура по движению
9	14, 15	все	сегментация с помощью кластеризации; подбор линий, кривых; устойчивость; алгоритм RANSAC
10	15, 16	все	подбор (окончание); скрытые значения и ОМ-алгоритм
11	17, 25	все	сопровождение; фильтры Кальмана, ассоциация данных; поиск изображений в цифровых библиотеках
12	18	все	зрение на основе модели
13	19	все	поверхности и их контуры
14	20	все	аспектные графы
15	22	все	сравнение с шаблоном

Если норма вектора \mathbf{a} равна единице, то скалярное произведение $\mathbf{a} \cdot \mathbf{b}$ равно длине (с соответствующим знаком) проекции \mathbf{b} на \mathbf{a} . В общем случае

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos \theta,$$

где θ — угол между двумя векторами. Из приведенного определения следует, что необходимым и достаточным условием ортогональности двух векторов является равенство нулю их скалярного произведения.

Векторное (или *внешнее*) произведение двух векторов $\mathbf{a} = (a_1, a_2, a_3)^T$ и $\mathbf{b} = (b_1, b_2, b_3)^T$ в пространстве \mathbb{R}^3 — это вектор

$$\mathbf{a} \times \mathbf{b} \stackrel{\text{def}}{=} \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

Отметим, что $\mathbf{a} \times \mathbf{b} = [\mathbf{a}_\times] \mathbf{b}$, где

$$[\mathbf{a}_\times] \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}.$$

Векторное произведение двух векторов \mathbf{a} и \mathbf{b} в \mathbb{R}^3 ортогонально к этим двум векторам, а необходимым и достаточным условием коллинеарности \mathbf{a} и \mathbf{b} является равенство нулю векторного произведения $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. Если угол между векторами \mathbf{a} и \mathbf{b} снова обозначить θ , то можно показать, что

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| |\sin \theta|.$$

УПРАЖНЕНИЯ И РЕСУРСЫ

Для выполнения упражнений, предлагаемых на протяжении всей книги, иногда могут понадобиться стандартные программы для численных расчетов по линейной алгебре, разложения сингулярных значений, а также линейных и нелинейных вычислений по методу наименьших квадратов. Значительное число таких средств предлагается в пакете MATLAB, а также в свободно распространяемых библиотеках, таких как LINPACK, LAPACK и MINPACK, которые можно загрузить из архива Netlib (<http://www.netlib.org/>). На веб-сайте книги <http://www.cs.berkeley.edu/~daf/book.html> предлагаются ссылки и на другие программы. По этому адресу также можно найти массивы данных (или ссылки на массивы) для заданий по программированию.