

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет»
(Астраханский государственный университет)

Кафедра английской филологии

Письменный перевод

по книге Computer vision

выходные данные The University of Washington, Mar. 2000

перевод стр. с 10 по 22

для сдачи кандидатского экзамена
по иностранному языку
(английский)
указать конкретный язык

Выполнил:

ФИО Свищев Н. Д.

Аспирант 1го года обучения,

*Кафедра информационной безопасности и
цифровых технологий*

Астрахань – 2021 г.

Chapter 1 Introduction.

This book is an introduction to the broad field of computer vision.

Without a doubt machines can be built to see; for example, machines inspect millions of light bulb filaments and miles of fabric each day.

Automatic teller machines (ATMS) have been built to scan the human eye for user identification and cars have been driven by a computer using camera input.

This chapter introduces several important problem areas where computer vision provides solutions.

After reading this chapter, you should have a broad view of some problems and methods of computer vision.

1 DEFINITION The goal of computer vision is to make useful decisions about real physical objects and scenes based on sensed images.

In order to make decisions about real objects, it is almost always necessary to construct some description or model of them from the image.

Because of this, many experts will say that the goal of computer vision is the construction of scene descriptions from images.

Although our study of computer vision is problem-oriented, fundamental issues will be addressed.

Critical issues raised in this chapter and studied in the remainder of the text include the following.

In order to make decisions about real objects, it is almost always necessary to construct some description or model of them from the image.

Because of this, many experts will say that the goal of computer vision is the construction of scene descriptions from images.

Although our study of computer vision is problem-oriented, fundamental issues will be addressed.

Critical issues raised in this chapter and studied in the remainder of the text include the following.

Sensing: How do sensors obtain images of the world?

How do the images encode properties of

Глава 1 Вступление.

Данная книга служит введением в широкий круг вопросов компьютерного зрения.

Без сомнения, можно сконструировать машины для обработки зрительных данных: например, на фабриках машины ежедневно контролируют миллионы осветительных ламп накаливания.

Разработаны системы безопасности, сканирующие человеческий глаз для идентификации личности и известны эксперименты по управлению автомобилем на основе компьютерной обработки данных, поступающих с видеокамеры.

В данной главе представлен ряд важных прикладных областей, в которых возможно применение систем компьютерного зрения.

После прочтения данной главы у читателя сложится общее представление о некоторых задачах и методах компьютерного зрения.

1 Определение. Цель компьютерного зрения заключается в формировании полезных выводов относительно объектов и сцен реального мира на основе анализа изображений, полученных с помощью датчиков.

Для формирования выводов относительно объектов реального мира почти всегда необходимо построить некоторое описание или модель этих объектов на основе изображения.

Поэтому многие эксперты приходят к выводу, что цель компьютерного зрения заключается в формировании описаний сцен по изображениям.

Хотя наш учебник по компьютерному зрению имеет прикладной характер, в нем также обсуждается ряд фундаментальных вопросов.

Перечислим важнейшие темы, которые будут обозначены в данной главе и затем будут изучаться в оставшейся части книги.

Каким образом при помощи датчиков можно получить изображения сцен реального мира?

Каким образом в изображениях представлены свойства реальных объектов, например материал, форма, освещенность и пространственные отношения?

Кодирование информации: Как изображения передают информацию для понимания трехмерного мира, в том числе информацию для идентификации объектов, об их геометрических и текстурных свойствах и движении?

Представление: В какой форме можно хранить описания объектов, их частей, свойств и взаимосвязей?

Алгоритмы: Какие существуют методы для обработки информации изображений и для построения описаний реальных объектов и сцен?

В последующих главах будут рассматриваться эти и

the world, such as material, shape, illumination and spatial relationships?

Encoded Information: How do images yield information for understanding the 3D world, including the geometry, texture, motion, and identity of objects in it?

Representations: What representations should be used for stored descriptions of objects, their parts, properties and relationships?

Algorithms: What methods are there to process image information and construct descriptions of the world and its objects?

These issues and others will be studied in the following chapters.

We now introduce various applications and some important issues that arise in their context.

Machines that see? Scientists and science fiction writers have been fascinated by the possibility of building intelligent machines, and the capability of understanding the visual world is a prerequisite that some would require of such a machine.

Much of the human brain is dedicated to vision. Humans solve many visual problems effortlessly, yet most have little analytical understanding of visual cognition as a process. Allan Turing, one of the fathers of both the modern digital computer and the field of artificial intelligence, believed that a digital computer would achieve intelligence and the ability to understand scenes.

Such lofty goals have proved difficult to achieve and the richness of human imagination is not yet matched by our engineering.

However, there has been surprising progress along some lines of research.

While building practical systems is a primary theme of this text and artificial intelligence is not, we will sometimes ponder the deeper questions, and, where we can, make some assessment of progress.

Consider, for example, the following scenario, which could be realized within the next few years.

A TV camera at your door provides images to your home computer which you have trained to recognize some faces of people important to you.

When you call in to your home message

some other questions. Now let's briefly list various practical problems and a set of important issues, arising in the context of these problems.

Can machines see? Authors of fantastic and popular science literature were fascinated by the idea of constructing intelligent machines, and the ability of visual perception of the world became an unquestionable requirement, which some of them demanded of such machines.

Vision is a very important function of the human brain. Many visual perception problems people solve unconsciously and without strain, but our understanding of this process is very distant.

Alan Turing, one of the inventors of the modern digital computer and the founder of the field of artificial intelligence, believed that a digital computer would be able to understand scenes.

These high goals are difficult to achieve, and the richness of human imagination is far beyond engineering possibilities.

However, a number of research directions have achieved impressive progress.

Although the main theme of this book is the development of practical systems, and not the study of artificial intelligence problems, we will sometimes study more profound questions, and, where possible, give some evaluation of achieved results.

Let's consider, for example, the following scenario, which can be realized in a few years.

A television camera in your door sends images to your home computer, which has been trained to recognize faces of people important to you.

When you call your home computer, it not only reports on telephone calls, but also informs you of possible visits of your sister or neighbor.

Similar directions of modern research and development will be considered in various parts of the book.

Practical problems.

Possibilities of application of computers for processing images are very large, in this book we will list only a few different problems, but they will also be of interest in further study of the subject and help to orient oneself in this field.

Preliminary consideration of the structure of digital images. Digital images can represent various objects, for example, a picture, a page of text, a human face, a map

center, your computer not only reports the phone messages.

Application problems

The applications of computers in image analysis are virtually limit less only a small sample of applications can be included here, but these will serve us well for both motivation and orientation to the field of study.

A preview of the digital image A digital image might represent a cartoon, a page of text, a person's face, a map of Katmandu, or a product for purchase from a catalog.

A digital image contains a fixed number of rows and columns of pixels, short for picture elements.

Pixels are like little tiles holding quantized values - small numbers, often between 0 and 255, that represent the brightness at the points of the image.

Depending on the coding scheme, 0 could be the darkest and 255 the brightest, or visa-versa.

At the top left in Figure 11 is a printed digital image of a face that is 257 rows high by 172 columns wide.

At the top center is an 8 x 8 subimage extracted from the right eye of the left image at the bottom of the figure are the 64 numbers representing the brightness of the pixels in that subimage.

The numbers below 100 in the upper right of the subimage represent the lower reflection from the dark of the eye (iris), while the higher numbers represent the brighter white of the eye.

A color image would have three numbers for each pixel, perhaps one value for red, one for blue, and one for green.

Digital images are most commonly displayed on a monitor, which is basically a television screen with a digital image memory.

A color image that has 500 rows and 500 columns is roughly equivalent to what you see at one instant of time on your TV A pixel is displayed by energizing a small spot of luminescent material, displaying color requires energizing 3 neighboring spots of different materials.

A high resolution computer display has roughly 1200 by 1000 pixels.

Катманду, или товар для покупки по каталогу.

Цифровое изображение состоит из фиксированного количества строк и столбцов пикселей (pixels), этот термин является сокращением от слов «элемент изображения» (picture element).

Пиксели напоминают маленькие квадратики, в которых хранятся дискретные значения-небольшие числа, часто от 0 до 255, представляющие яркость точек изображения.

В зависимости от схемы кодирования 0 может соответствовать самой малой (темной), а 255— самой большой (светлой) по яркости, или наоборот.

На рис слева сверху напечатано цифровое изображение лица. Изображение содержит 257 строк пикселей в высоту и 172 столбца в ширину.

Вверху в центре показан фрагмент изображения размерами 8 x 8 пикселей, взятый из окрестности изображения правого глаза, на левом изображении в нижней части рис приведена таблица с 64 числами-значениями яркости пикселей фрагмента изображения.

Числа меньше 100 в правом верхнем углу фрагмента представляют собой слабо отражающую свет темную область глаза (зрачок), а большие значения соответствуют более светлому белому цвету глаза.

В цветных изображениях для каждого пикселя может храниться три числа, например одно число для красной, одно для синей и одно для зеленой составляющей цвета.

Цифровые изображения чаще всего отображаются на мониторе, который можно представить себе как телевизионный экран с цифровой памятью для изображения.

Цветное изображение размерами 500 строк и 500 столбцов примерно соответствует изображениям, которые вы видите на экране своего телевизора.

Пиксел высвечивается путем передачи энергии маленькой области люминесцентного материала.

Воспроизведение цвета требует передачи энергии трем соседним областям, сделанным из различных материалов. Размеры компьютерного дисплея высокого разрешения составляют примерно 1200 x 1000 пикселей.

Более подробно структура цифровых изображений рассматривается во второй главе, а кодирование и интерпретация цвета рассматривается в главе 6.

Поиск в базе данных изображений разработка баз данных изображений стала возможной благодаря доступности цифровой памяти большого объема, линий связи с большой пропускной способностью и персональных компьютеров с мультимедиа-возможностями.

Для эффективного использования большого количества существующих изображений необходимо разработать удобные средства доступа. Обычные

The next chapter discusses digital images in more detail, while coding and interpretation of color in digital images is treated in Chapter 6.

Image Database Query Huge digital memories, high bandwidth transmission and multimedia personal computers have facilitated the development of image databases.

Good use of the many existing images requires good retrieval methods.

Standard database techniques apply to images that have been augmented with text keys; however, content-based retrieval is needed and is a topic of much current research.

Suppose that a newly formed company wants to design and protect a new logo and that an artist has created several candidates for the company to consider a logo cannot be used if it is too similar to one of an existing company, so a database of existing logos must be searched this operation is analogous to patent search and is done by humans, but could be greatly aided by machine vision methods. See Figure 1.2.

There are many similar problems Suppose an architect or an art historian wants to search for buildings with a particular kind of entryway.

It would be desirable to just provide a picture, perhaps fetched from the database itself, and request the system to produce other similar pictures.

In a later chapter, you will see how geometric, color, and texture features can be used to aid in answering such an image database query Suppose that an advertising agency wants to search for existing images of young children enjoying eating.

This semantic requirement, which is simple for humans to understand, presents a very high level of difficulty for machine vision. Characterizing "children", "enjoyment", and "eating" would require complex use of color, texture, and geometric features.

We note in passing that a computer algorithm has been devised that decides whether or not a color image contains a naked person.

This could be useful for parents who want

технологии баз данных можно применять для доступа к изображениям, снабженным текстовыми метками.

Однако также необходимы средства поиска изображений на основе содержания, которые в настоящее время являются предметом интенсивных исследований.

Представим, что недавно организованная фирма хочет разработать и зарегистрировать эмблему. Дизайнер представил на рассмотрение компании несколько вариантов эмблем.

Эмблему нельзя использовать, если она слишком похожа на эмблему какой-либо существующей компании, поэтому надо проверить базу данных существующих эмблем.

Эта процедура аналогична патентному поиску и выполняется вручную, но могла бы быть существенно упрощена с помощью методов машинного зрения.

На рис. показано несколько похожих эмблем.

Еще один пример: допустим, архитектору или историку-искусствоведу требуется найти изображения зданий, у которых вход оформлен в определенном стиле.

Для поиска они могут указать изображение-образец, возможно, полученное из той же базы данных, и запросить систему найти остальные похожие изображения.

В гл. 8 будет показано, каким образом можно использовать геометрические, цветовые и текстурные признаки для обслуживания подобных запросов к базе данных изображений.

Предположим, что рекламное агентство хочет найти существующие изображения маленьких детей, которые что-то едят с радостным видом.

Это семантическое описание, легко воспринимаемое людьми, очень сложно представить для системы машинного зрения.

Понятия «дети», «радостно» и «есть» могут потребовать сложного комбинированного учета цветовых, текстурных и геометрических признаков.

Попутно отметим, что был разработан компьютерный алгоритм для обнаружения обнаженных людей на цветных изображениях.

Данный алгоритм может быть полезен для родителей, желающих проконтролировать изображения, которые их дети загружают из Интернета.

Методы извлечения изображений из баз данных подробнее рассматриваются в гл. 8.

Контроль отверстий в поперечных балках.

В конце 1970-х гг инженер из Милуоки разработал систему машинного зрения для подсчета количества болтовых отверстий в поперечных балках, изготовленных для компаний-производителей грузовых

to screen images that their children retrieve from the web. Image database retrieval methods are treated in Chapter 8.

Inspecting crossbars for holes In the late 1970s an engineer in Milwaukee implemented a machine vision system that successfully counted the number of bolt holes in crossbars made for truck companies.

The truck companies demanded that every cross bar be inspected before being shipped to them, because a missing bolt hole on a partly assembled truck was a very costly defect either the assembly line would have to be stopped while the needed hole was drilled, or worse, a worker might ignore placing a required bolt in order to keep the production line running to create a digital image of the truck crossbar lights were placed beneath the existing transfer line and a digital camera above it.

When a crossbar came into the field of view, an image was taken dark pixels inside the shadow of the crossbar were represented as 1's indicating steel and pixels in the bright holes were represented as 0's, indicating that the hole was drilled.

The number of holes can be computed as the number of external corners minus the number of internal corners all divided by four Figure 1.3 shows three bright holes ('0's) in a background of '1's an external corner is just a 2 x 2 set of neighboring pixels containing exactly 3 ones while an internal corner is a 2 x 2 set of neighboring pixels containing exactly 3 zeroes example processing of an image with 7 rows and 33 columns is shown in the figure and a skeleton algorithm is also shown.

Holecounting is only one example of many simple, but powerful operations possible with digital images. (As the exercises below show, the holecounting algorithm is correct only if the holes are "4-connected" and "simply connected" - that is, they have no background pixels inside them. discussed further in Chapter 3 and in more detail in the text by Rosenfeld.)

These concepts are Examining the inside of a human head.

Magnetic resonance imaging (MRI) devices can sense materials in the interior of 3D objects Figure 1.4 shows a section

автомобилей.

Данные компании требовали контроля каждой балки, так как отсутствие болтового отверстия на частично собранном грузовике является дорогостоящим дефектом.

Для его исправления либо надо останавливать конвейер и сверлить отверстие, либо, что хуже, рабочий может пропустить установку болта, чтобы не останавливать конвейер.

Для получения цифрового изображения поперечной балки осветители были помещены под транспортировочной линией, а над линией была расположена видеокамера.

При поступлении в поле зрения очередной балки выполнялся съем изображения.

Темные пиксели в области тени от балки на изображении имели значение 1 и обозначали сталь, а пиксели в области светлых просверленных отверстий имели значение 0.

Количество отверстий вычислялось как четверть разности количества внешних и внутренних углов.

На рис показаны три светлых отверстия (значения 0) на темном фоне (значения 1) внешний угол (external corner) -это окрестность 2 x 2 соседних пикселей, содержащая 3 единицы, а внутренний угол (internal corner)- окрестность 2 x 2 соседних пикселей, содержащая 3 нулевых значения.

На рис. 1.3 показан пример обработки изображения из 7 строк и 16 столбцов.

Ниже также схематично описан алгоритм обработки.

Подсчет отверстий является примером простой, но практически полезной операции над цифровыми изображениями (Как показано в приведенных далее упражнениях, алгоритм подсчета отверстий корректно работает только с «4-связными» и «просто связными» отверстиями - т. е. внутри отверстий не должно быть фоновых пикселей.

Эти понятия будут рассматриваться в гл. 3, а более детальное описание можно найти в книге Розенфельда.)

Контроль медицинских изображений головы человека.

Датчики изображений на эффекте ядерного магнитного резонанса (ЯМР) могут воспринимать вещества внутри трехмерных объектов.

На рис. 1.4 показано поперечное сечение человеческой головы яркие области соответствуют движущемуся веществу, так что в действительности это изображение потока крови.

Специалист может узнать на изображении важные кровеносные сосуды.

Нечетко видимые кометоподобные структуры связаны с глазами.

through a human head: brightness is related to movement of material, so this is actually a picture of blood flow.

One can "see" important blood vessels.

The wispy comet-like structures are associated with the eyes.

MRI images are used by doctors to check for tumors or blood flow problems such as abnormal vessel constrictions or expansions.

The image at the right in Figure 1.4 was made from a copy of the one on the left by making every pixel of value 208 or more bright (255) and those below 208 dark (0).

Most pixels correctly show blood vessels versus background, but there are many incorrectly "colored" pixels of both types.

Machine vision techniques are often used in medical image analysis, although usually to aid in data presentation and measurement rather than diagnosis itself.

Wouldn't it be great if we could "see" thoughts occurring in the brain!

Well, it turns out that MRI can sense organic activity related to thought processes and this is a very exciting current area of research.

Processing scanned text pages a common problem is to convert information from paper documents into digital form for information systems.

For example, we might want to make an old book available on the Internet, or we might need to convert a blueprint of some object into a geometry file so that the part can be made by a numerically controlled machine tool.

Figure 1.5 shows the same message in both Chinese and English.

The Chinese characters were written on paper and scanned into an image of 482 rows and 405 columns.

The postscript file encoding the graphics and printed in the figure has a size of 68,464 bytes.

The English version is stored in a file of 115 bytes, each holding one ASCII character.

There is an entire range of important applications in processing documents.

Recognizing individual characters from the dots of the scanner or FAX files is one such application that is done fairly well today, provided that the characters conform

YMR-images are used in medicine for detection of tumors and disturbances of blood circulation, such as anomalous narrowing or expansion of vessels.

On the right is a copy of the image on the left, in which all pixels with a value of 208 or more are made maximally bright (255), and those with a value less than 208 are made dark (0).

The majority of pixels correctly show blood vessels on the background, but there are also incorrectly "colored" pixels of both types.

Methods of machine vision are often used in the analysis of medical images, although they are usually applied for visualization of information and data measurements, and not for diagnosis.

Of course, it would be very good, if we could "see" thought processes, occurring in the brain!

It turns out, that the method of YMR can transmit organic activity, connected with thought processes, and this is a very perspective area of modern research. Digital image of 100 rows and 500 columns of pixels.

Suppose, that the diameter of one bolt hole in a beam is 1.2 cm.

Determine the radius and area of this hole in the image in pixels.

Processing of scanned pages of text.

Transformation of information, contained in paper documents, into digital form for information systems is a very widespread task.

For example, it is necessary to provide access to an old book through the Internet, or it is necessary to transform a paper drawing of some detail into a file with data on its geometry for the subsequent manufacture of this detail on a machine with program control.

On Fig. 1.5 is shown the writing of one and the same phrase in Chinese and English.

Chinese ideograms were written on paper and scanned as an image of 482 rows and 405 columns.

The file in the Postscript language with a description of the graphic image (and printed on Fig. 1.5) occupies 68464 bytes.

The English version occupies in the file 115 bytes, one byte for each symbol of the ASCII table.

Applicable to this case, it is possible to list a whole set of important tasks of document processing.

Recognition of individual symbols from the dots of the scanner or fax image is one such application that is done fairly well today, provided that the characters conform

to standard patterns.

Providing a semantic interpretation of the information, possibly to be used for indexing in a large database, is a harder problem.

Accounting for snow cover using a satellite image.

Much of the earth's surface is scanned regularly from satellites, which transmit their images to earth in digital form these images can then be processed to extract a wealth of information for example, inventory of the amount of snow in the watershed of a river may be critical for regulating a dam for flood control, water supply, or wildlife habitat.

Estimates of snow mass can be made by accounting for the number of pixels in the image that appear as snow a pixel from a satellite image might result from sensing a 10 meter by 10 meter spot of earth, but some satellites reportedly can see much smaller spots than that.

Often, the satellite image must be compared to a map or other image to determine which pixels are in a particular area or watershed.

This operation is usually manually-aided by a human user interacting with the image processing software and will be discussed more in Chapter 11 where image matching is covered.

Figure 16 is a photograph taken on a space shuttle flight managed by the Johnson Space Center in Houston, Texas It shows the town of Wenatchie, Washington, where the Wenatchie River flows into the Columbia River.

Computers are known for their ability to handle large amounts of data, certainly the earth scanning satellites produce a tremendous amount of data useful for many purposes

For example, counts and locations of snow pixels might be input to a computer program that simulates the hydrology for that region (Temperature information for the region must be input to the program as well.) Another related application is taking inventory of crops and predicting harvests.

Yet another is taking inventory of buildings for tax purposes this is usually done manually with pictures taken from airplanes.

Understanding a scene of parts At many

время реализовано на уровне коммерческих продуктов, при условии, что символы имеют стандартное начертание.

Семантическая интерпретация информации, возможно, необходимая для индексации в большой базе данных, представляет собой намного более сложную задачу.

Оценка снежного покрова по спутниковым изображениям.

Большая часть земной поверхности регулярно сканируется с помощью спутников, передающих на Землю цифровые изображения.

Последующая обработка этих изображений позволяет получить много полезной информации.

Например, оценка количества снега в водоразделе реки может быть критически важна для управления плотинами, системами водоснабжения, для охраны окружающей среды.

Для оценки массы снега можно подсчитать количество, которые выглядят, как снег.

Пиксел спутникового пикселей на изображении может представлять довольно крупный участок поверхности, например 10х 10 метров, но у некоторых спутников разрешение намного выше.

Часто спутниковое изображение требуется сравнить с картой или другими изображениями и определить, какие пиксели попадают в определенную область или водную поверхность.

Это сравнение обычно проводится вручную с помощью специальных программ обработки изображений (более подробно сопоставление изображений описано в гл. 11).

На рис. 1.6 показана фотография, сделанная во время полета космического корабля «Спейс-Шаттл».

На ней снят город Уэнатчи, штат Вашингтон, около которого река Уэнатчи впадает в реку Колумбия.

Возможности компьютеров по обработке больших объемов данных хорошо известны.

Спутниковая фотосъемка земной поверхности дает огромное количество данных, полезных для многих целей. Например, количество и местоположение снега могут быть входными данными для компьютерной программы, моделирующей гидросферу данного региона (также ей требуется информация о температуре).

Другим возможным приложением является оценка посевов и расчет урожая.

Еще одна задача - учет городских зданий для налогообложения: это обычно делается вручную по изображениям, полученным при аэрофотосъемке.

Анализ местности, содержащей объекты промышленного производства во многих

points of manufacturing processes, parts are transferred on conveyors or in baxes.

Parts must be individually placed in machines, packed, inspected, etc.

If the operation is dull or dangerous, a vision-guided robot might provide a solution.

The underlying image of Figure 17 shows three workpieces in a robot's workspace.

By recognizing edges and holes, the robot vision system is able to guess at both the identity of a part and its position in the workspace.

Using a 3D model made by computer-aided-design (CAD) for each guessed part and its guessed position, the vision system then compares the sensed image data with a computer graphic generated from the model and its position in space. Bad matches are rejected while good matches cause the guess to be refined.

The bright lines in Figure 17 show three such refined matches between the image and models of the objects it contains. Finally, the robot eye-brain can tell the robot aim how to pick up a part and where to put it. The problems and techniques of 3D vision are covered in the later chapters of this text.

Operations on Images This book presents a large variety of image operations.

Operations can be grouped into different categories depending on their structure, level, or purpose.

Some operations are for the purpose of improving the image solely for human consumption, while others are for extracting information for downstream automatic processing.

Some operations create new output images, while others output non-image descriptions.

A few important categories of image operations follow. Changing pixels in small neighborhoods. Pixel values can be changed according to how they relate to a small number of neighboring pixels, for example, neighbors in adjacent rows or columns.

Frequently, isolated 1's or 0's in a binary image will be reversed in order to make them the same as their neighbors. The purpose of this operation could be to remove likely noise from the digitization process or, it could be just to simplify image content, for example,

производственных процессах есть этапы, на которых отдельные детали перемещаются на конвейере или в упаковке.

Детали требуется по отдельности помещать в механизмы, в упаковку, контролировать и т. п.

Если подобная операция ресурсозатратна или опасна, то может быть целесообразно применение роботизированного комплекса с системой машинного зрения.

На рис. 1.7 показано изображение с тремя деталями, расположенными в рабочем пространстве робота.

Распознавая ребра и отверстия, зрительная система робота способна решить, какие именно детали видны и каково их местоположение и ориентация в рабочем пространстве.

В процессе распознавания каждой детали используется трехмерная модель, предварительно сформированная в системе автоматизированного проектирования (САПР).

Зрительная система сравнивает входное изображение с компьютерными изображениями, сгенерированными по данным моделей и данным о положении объекта в пространстве.

Плохо совпадающие предположения отбрасываются, а правдоподобные уточняются.

Яркими линиями на рис. 1.7 показаны три уточненных соответствия между изображением и моделями объектов, находящимися на этом изображении.

После распознавания система управления роботом может сообщить манипулятору, как надо захватить деталь и куда ее поместить.

Задачи и методы трехмерного компьютерного зрения в данной книге рассматриваются в гл. 13 и 14.

Операции обработки изображений.

В данной книге рассмотрено много различных операций обработки изображений.

Их можно разделить на несколько категорий в зависимости от структуры, уровня или назначения.

Некоторые операции предназначены исключительно для повышения качества изображений при человеческом восприятии, в то время как остальные предназначены для извлечения информации, пригодной для автоматической обработки.

Некоторые операции создают новые выходные изображения, а некоторые в качестве выходных данных генерируют описания неграфической формы.

Ниже перечислены несколько важных категорий операций обработки изображений.

Модификация пикселей в малых окрестностях. Значения пикселей можно изменять с учетом их взаимосвязи с небольшим числом близлежащих

to ignore tiny islands in a lake or imperfections in a sheet of paper another common operation is to change border pixels to be background pixels as shown in Figure 1.8.

The images of bacteria have fuzzy borders and often fuse together by changing the black border pixels to white, the bacteria images, although smaller, have clearer borders and some formerly fusing pairs are separated. These operations are treated in Chapter 3.

Enhancing an entire image Some operations treat the entire image in a uniform manner the image might be too dark - say its maximum brightness value is 120 - so all brightness values can be scaled up by a factor of 2 to improve its displayed appearance.

Noise or unnecessary detail can be removed by replacing the value of every input pixel with the average of all nine pixels in its immediate neighborhood.

Alternatively, details can be enhanced by replacing each pixel value by the contrast between it and its neighbors.

Figure 19 shows a simple contrast computation applied at all pixels of an input image.

Note how the boundaries of most objects are well detected.

The output image results from computations made only on the local 3x3 neighborhoods of the input image Chapter 5 describes several of these kinds of operations perhaps an image is taken using a fish eye lens and we want to create an output image with less distortion: in this case, we have to "move" the pixel values to other locations in the image to move them closer to the image center such an operation is called image warping and is covered in Chapter 11.

Combining multiple images An image can be created by adding or subtracting two input images. Image subtraction is commonly used to detect change over time.

Figure 1.10 shows two images of a moving part and the difference image resulting from subtracting the corresponding pixel values of the second image from those of the first image.

Image subtraction captures the boundary of the moving object, but not perfectly.

пикселей, например, пикселей из соседних строк или столбцов.

Часто на бинарных изображениях изолированные значения 1 и 0 меняются на противоположные значения, чтобы они совпадали со своими соседями.

Назначением этой операции может быть удаление шума, появившегося в процессе оцифровки.

Или, это делается для упрощения изображения; например, для игнорирования мелких островов на изображении озера или дефектов на изображении листа бумаги.

Еще одна распространенная операция — изменение значений граничных пикселей на значения фоновых пикселей (рис. 1.8).

Изображения отдельных бактерий на рис. 1.8 имеют нечеткие границы и часто сливаются вместе.

После замены черных граничных пикселей на белые, изображения бактерий хотя и становятся меньше, но получают более явные границы.

Некоторые ранее сливавшиеся пары разделяются. Подобные операции будут рассмотрены в гл. 3.

Глобальное улучшение качества изображения некоторые операции выполняют однородную обработку целого изображения.

Изображение может быть слишком темным, например, максимальная яркость равна 120, так что все значения яркости можно увеличить в два раза для улучшения вида изображения при выводе на экран.

Для удаления шума или несущественных деталей можно заменить значение каждого входного пиксела средним значением 9 пикселей его окрестности.

С другой стороны, детали можно подчеркнуть, если заменить каждый пиксел значением контраста между ним и соседними пикселями.

На рис. 1.9 показан результат вычисления контраста для всех пикселей входного изображения.

Отметим, что границы большинства объектов оказались хорошо выделенными.

Выходное изображение получено вычислением контраста в окрестностях размером 3 x 3 на входном изображении.

В гл. 5 описано несколько типов подобных операций. Возможно, изображение было получено с помощью широкоугольного объектива («рыбий глаз») и требуется получить выходное изображение без искажений объектива: в данном случае можно «переместить» значения пикселей в другие положения на изображении.

Подобная операция называется деформацией изображения (см. гл. 11).

Комбинация нескольких изображений.

Изображение можно сформировать путем сложения или вычитания двух исходных изображений.

(Since negative pixel values were not used, not all changes were saved in the output image.)

In another application, urban development might be more easily seen by subtracting an aerial image of a city taken five years ago from a current image of the city. Image addition is also useful.

Figure 1.11 shows an image of Thomas Jefferson "added" to an image of the great arch opening onto the lands of the Louisiana Purchase; more work is needed in this case to blend the images better. Computing features from an image we have already seen the example of counting holes.

More generally, the regions of 0's corresponding to holes in the crossbar inspection problem could be images of objects, often called blobs - perhaps these are microbes in a water sample.

Important features might be average object area, perimeter, direction, etc.

We might want to output these important features separately for every detected object.

Chapter 3 describes such processing.

Chapters 6 and 7 discuss means of quantitatively summarizing the color or texture content of regions of an image.

Chapter 4 shows how to classify objects according to these features; for example, is the extracted region the image of microbe A or B?

Figure 1.12 shows output from a well-known algorithm applied to the bacteria image of Figure 1.8 giving features of separate regions identified in the image, including the region area and location. Regions with area of a few hundred pixels correspond to isolated bacteria while the large region is due to several touching bacteria.

Extracting non-iconic representations. Higher-level operations usually extract representations of the image that are non-iconic, that is, data structures that are not like an image. (Recall that extraction of such descriptions is often defined to be the goal of computer vision.) Figure 1.12 shows a non-iconic description derived from the bacteria image.

In addition to examples already mentioned, consider a report of the count of microbes of type A and B in a slide from a

Вычитание изображений часто применяется для обнаружения изменений в течение некоторого промежутка времени.

На рис. 1.10 показаны два изображения движущейся детали и разностное изображение, полученное путем вычитания пикселей второго изображения из соответствующих пикселей первого изображения.

Вычитание изображений сохраняет границу движущегося объекта, хотя и не идеальным образом. (Так как отрицательные значения пикселей не используются, то в выходном изображении представлены не все изменения.)

В другом примере новые городские постройки могут быть выявлены путем вычитания изображения с аэрофотосъемкой города пятилетней давности из изображения, полученного недавно.

Сложение изображений также является полезной операцией.

На рис. 1.11 показано изображение Томаса Джефферсона, «добавленное» к изображению арки в Сент-Луисе.

В данном случае чуть больше усилий требуется приложить для лучшего смешения изображений.

Вычисление характерных признаков изображения выше была рассмотрена задача подсчета отверстий.

В общем случае, области значений 0 или 1, которые могут являться изображениями объектов (например, отверстия в поперечной балке или микробы в образце воды), часто называются пятнами.

Важными характерными признаками являются средняя площадь объектов, периметр, направление и др.

Подобные признаки описываются в гл. 3.

В главах 6 и 7 описываются способы количественного учета цвета или текстуры во фрагментах изображений.

В гл. 4 показано, как по этим признакам выполняется классификация объектов; например, решается, является ли обнаруженная на изображении область микробом вида A или B?

На рис. 1.12 показаны выходные данные известного алгоритма, примененного к изображению бактерий на рис. 1.8, который позволяет вычислить характерные признаки отдельных областей, найденных на изображении, в том числе площадь и местоположение каждой области.

Области с площадью несколько сотен пикселей соответствуют отдельным бактериям, а большая область соответствует нескольким соприкасающимся бактериям.

Формирование неграфических описаний. Высокоуровневые операции обычно формируют неграфические описания изображений, т. е. структуры данных, не являющиеся изображениями. (Напомним,

microscope or the volume of traffic flow between two intersections of a city computed from a video taken from a utility pole.

In another important application, the (iconic) input might be a scanned magazine article and the output a hypertext structure containing sections of recognized ASCII text and sections of raw images for the figures.

As a final example, in the application illustrated in Figure 1.7, the machine vision system would output a set of three detections, each encoding a part number, three parameters of part position and three parameters of the orientation of the part.

This scene description could then be turned over to the motion-planning system, which would decide on how to manipulate the three parts.

что формирование таких описаний часто считается основной целью компьютерного зрения.) на рис. 1.12 показано неграфическое описание изображения бактерий.

Кроме уже описанных примеров, можно привести отчет о подсчете микробов видов А и В на слайде с микроскопа или об объеме транспортного потока между двумя частями города, вычисленного по видеоданным.

В другой важной задаче графическое изображение может быть сканированной журнальной статьей, а выходным представлением — гипертекстовая структура, содержащая части распознанного текста и рисунки из статьи.

Еще одним примером (рис. 1.7) данных является формирование набора из трех распознанных объектов, для каждого — система машинного зрения, которая должна в качестве выходных указать идентификационный код, три параметра положения и три параметра ориентации.

Это описание сцены, в свою очередь, может быть представлено в виде подсистемы планирования движения для осуществления взаимодействия робота-манипулятора с данными тремя деталями.