

英特尔® 软件学院系列课程

可视化计算应用开发

Visual Computing Application Development

英特尔® 亚太研发有限公司 组编
英特尔® 软件学院教材编写组 编
陈泽宇 徐礼爽 宋志凯 杨青



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

策划：英特尔® 软件学院中国区经理 郎朗



可视化计算应用开发

英特尔® 软件学院教材编写组 编

编写组名单 陈泽宇 徐礼爽
宋志凯 杨 青

上海交通大学出版社

内 容 提 要

本书把介绍图像和视频处理的基础理论知识和英特尔公司提供的可视化计算软件开发库相结合,理论联系实际,强调实践,在每个章节,都配有相关的实例来说明如何利用英特尔集成性能原件来开发相关应用的关键代码。全书分9章,系统讲述了可视化计算的应用开发,可作为软件工程师的参考书,亦可作为大中专院校计算机相关专业的教学参考用书。

图书在版编目(CIP)数据

可视化计算应用开发 / 英特尔®软件学院教材编写组
编. —上海: 上海交通大学出版社, 2011
英特尔®软件学院系列课程
ISBN 978-7-313-07044-9
I. ①可… II. ①英… III. ①计算机网络—程序设计—技术培训—教材 IV. ①TP393.09
中国版本图书馆 CIP 数据核字(2011)第 001897 号

可视化计算应用开发

英特尔®软件学院教材编写组 编

上海交通大学出版社出版发行

(上海市番禺路 951 号 邮政编码 200030)

电话: 64071208 出版人: 韩建民

上海交大印务有限公司印刷 全国新华书店经销

开本: 787 mm×960 mm 1/16 印张: 21.25 字数: 389 千字

2011 年 1 月第 1 版 2011 年 1 月第 1 次印刷

印数: 1~5 030

ISBN 978-7-313-07044-9/TP 定价: 49.50 元

版权所有 侵权必究

进入 21 世纪,信息技术和信息产业在全球范围内迅猛发展的势头更为强劲,如何尽快适应新技术和新应用带来的挑战,及时更新员工知识结构,并动态调整企业人才培养战略,已经成为广大科技公司迫切需要解决的问题之一。对于高等教育、职业教育等专业组织机构来说,则面临着紧跟企业前进步伐,准确接轨社会发展趋势,瞄准世界科技前沿水平,不断进行教育教学创新,提高学生实践能力,开拓学生知识视野的现实需求。在中国,尽管近年来已经在科技人才培养方面取得了长足的进步,但是就整体现状而言,尤其在知识更新和技术创新方面,距离完全满足社会的需求还存在着较大的发展空间。

英特尔公司历来关注技术的发展创新和科技人才的培养。英特尔®软件学院隶属于英特尔软件与服务事业部,作为英特尔公司专业的对外培训机构,为全球的软件开发人员提供了丰富的前沿技术培训课程。多年来,英特尔®软件学院一直致力于培训软件开发人员,与中国的软件开发人员共同发展,帮助其掌握和应用英特尔的最新技术及经验,提高软件开发技术水平,提升产品开发技能。目前,英特尔®软件学院在中国已经发展成为面向软件开发、项目管理及商业运营方向的优秀的一站式培训服务基地。依托英特尔公司强大的师资力量,沿袭英特尔用户需求至上的传统,英特尔®软件学院已经与国内多家知名公司、大学和教育机构建立了长期稳定



的合作关系,迄今已有数万名工程师和大学教师参与了英特尔®软件学院的技术培训,并学以致用。

《英特尔®软件学院系列课程》由英特尔®软件学院牵头,联合国内的顶级高等学府合作编写。相信这本教材作为英特尔®软件学院的重点课程之一,在科技人才培养和知识创新方面必将发挥重要的作用。

英特尔亚太研发有限公司 总经理
英特尔公司软件与服务事业部 中国区总经理
梁兆柱 博士

前言

可视化计算技术的研究和应用对象并不是很清晰、严格,因此可视化计算技术在不同的上下文中可能会涉及多个领域,如数据可视化、计算机图形学、数字图像处理、多媒体技术、计算机视觉等,这些都可以称为可视化计算技术。

图像和视频是可视化计算技术中重要的组成部分,信息时代大量的信息都是通过数字图像和视频展示给每一个人。为了向程序员普及或推广数字图像和视频软件开发技术,降低可视化计算软件开发的门槛,本书把介绍图像和视频处理的基础理论知识和英特尔公司提供的可视化计算软件开发库相结合,注重理论联系实际,强调实践。在每个章节,都配有相关实例来说明如何利用 Intel® IPP 开发相关应用的关键代码。

本书共分 9 章,具体内容如下:

第 1 章 可视化计算软件开发基础。介绍软件开发相关技术,英特尔图像、视频高性能库开发包,以及编程环境。

第 2 章 数字图像基础。详细介绍数字图像的基本概念,图像处理的常用技术,以及 Intel® IPP 图像编程基础。

第 3 章 图像编码与文件格式。详细介绍常用编码技术,以及用 Intel® IPP 实现 JPEG 编码。

第 4 章 灰度、彩色图像增强。详细介绍亮度、对比度增强和彩色增强



相关技术,以及用 Intel® IPP 实现相关增强技术。

第 5 章 图像的平滑和锐化。详细介绍空间域和频域的图像平滑和锐化技术,以及用 Intel® IPP 实现相关平滑和锐化技术。

第 6 章 图像几何变化。详细介绍图像的各种仿射变换,以及用 Intel® IPP 实现相关图像几何变换技术。

第 7 章 数字视频基础。详细介绍视频编解码的基本原理和相关技术,以及用 Intel® IPP 和 Intel® Media SDK 实现相关编解码技术。

第 8 章 图像处理软件开发综合实例。综合前几章的内容,将通用的图像处理技术打包实现,同时添加了一个 GUI 的图像几何变换的实例。

第 9 章 视频处理软件开发综合实例。详细分析音视频应用的一个架构及其实现,以及用 DirectShow 开发视频播放程序。

本书可以作为软件程序工程师的参考书,也可作为大中专院校计算机相关专业的教学参考用书。

由于时间仓促,作者水平有限,书中的缺点和不足之处,敬请读者批评指正。

《可视化计算应用开发》教材编写组

2010 年 4 月

目 录

1 可视化计算软件开发基础	1
1.1 可视化计算概述	1
1.2 软件开发与优化基础	7
1.3 可视化计算软件开发平台	11
1.4 上机实验	22
2 数字图像基础	28
2.1 数字图像概述	28
2.2 图像基本概念	32
2.3 图像变换	47
2.4 图像压缩编码	63
2.5 IPP 图像处理编程基础	64
2.6 上机实验	72
3 图像编码与文件格式	79
3.1 无损压缩	80
3.2 有损压缩	88
3.3 图像文件格式	93
3.4 用 IPP 函数实现 JPEG 图像编码算法	99
3.5 上机实验	109



4 图像增强——灰度、彩色	118
4.1 亮度、对比度增强	118
4.2 彩色增强	129
4.3 上机实验	134
5 图像增强——平滑、锐化	142
5.1 图像平滑(去噪)	142
5.2 图像锐化	148
5.3 频域处理	151
5.4 IPP 函数库中的滤波函数	161
5.5 上机实验	165
6 图像几何变换	176
6.1 图像缩放	177
6.2 图像旋转	178
6.3 图像变形	181
6.4 图像几何变换中的插值算法	184
6.5 上机实验	188
7 数字视频基础	192
7.1 数字视频概述	193
7.2 视频压缩编码	196
7.3 视频编码标准	202
7.4 视频文件格式	217
7.5 用 IPP 开发数字视频应用	221
7.6 用 Media SDK 开发数字视频应用	224
7.7 上机实验	228
8 图像处理软件开发综合实例	231
8.1 通用图像处理程序	231

8.2 用鼠标交互对图像进行几何变换.....	245
9 视频处理软件开发综合实例.....	273
9.1 通用音视频应用架构及实现.....	273
9.2 用 DirectShow 开发视频播放程序	299
附录	314
附录 1 IPP 图像视频处理函数名的命名规则.....	314
附录 2 景深的概念	316
附录 3 位图文件结构定义	318
附录 4 JPEG 中 Huffman 编解码说明	319
附录 5 伪彩色增强中的伪彩色编码表	323
附录 6 数字视频彩色格式	328

1 可视化计算软件开发基础

1.1 可视化计算概述

早期计算机的发明是科学计算的需要,其使用者主要是从事科学计算的科技工作者。随着计算机软硬件技术的发展,计算机的功能也在进一步扩展,开始从科技的象牙塔走向大众,并成功应用于生产(如工业控制、企业生产执行系统 MES)、办公(无纸化办公)、生活(网上冲浪、购物)及娱乐(数字电影、游戏)等各个方面。普适计算(Ubiquitous Computing)是 21 世纪的计算模式,计算将无处不在。

在计算机接近、走向大众的过程中,计算机对图形图像、音视频的捕捉(capture)、设计(design)、存储(storage)、处理(process)、传输(transfer)等关键的可视化计算技术功不可没。

可视化计算(Visual Computing)技术的研究和应用对象并不是很清晰、严格,因此可视化计算技术在不同的上下文中可能会涉及多个领域,如数据可视化(Data Visualization)、计算机图形学(Computer Graphics)、数字图像处理(Image Processing)、多媒体技术(MultiMedia)、计算机视觉(Computer Vision)等,这些都可以称为可视化计算技术。

随着可视化计算技术的发展,对具备设计和开发视觉计算软件能力的工程师的需求也在进一步扩大。由于可视化计算技术涉及较深的数学等理论知识和计算机编程技术,因此,对程序员的要求也会相对较高,程序员从零起步开发可视化计算软件的难度较大。因此一个具备高性能的、支持多种软件平台(多种操作系统)和硬件平台(同时支持嵌入式开发)的可视化计算软件开发库是可视化计算软件开发的必然需求。

英特尔公司在多年的技术积累的基础上,为可视化计算软件开发人员提供了满足上述特性的软件开发库,包括 Intel® IPP(Integrated Performance Primitives, 集成性能元件)和 Intel® Media SDK(Software Development Kit, 软件开发库)。

为了向程序员普及或推广可视化计算软件开发技术,降低可视化计算软件开发的门槛,书中文字、图例、实验、程序代码互为补充,向大家介绍图像和视频处理开发中的技术及应用。全书通过介绍图像和视频处理的基础理论知识和英特尔公司提供的可视化计算软件开发库相结合,注重理论联系实际,强调实践。

在学习本书前,读者应学习过 C/C++ 程序设计语言,并具备一定的计算机编程基础和实践经验,能较为快速地阅读 C/C++ 语言程序代码,并且 C/C++ 程序编程比较熟练,对面向对象技术最好要有一定了解。

1.1.1 图像

1) 数字图像

图像是人类接触世界,获取信息的重要途径。用图像来表达信息直观、形象、丰富,因此,才有“一图胜千言”,“百闻不如一见”等说法。

图像在传达信息方面有其他媒介(文字、声音等)无法替代的优越性,因此,在网络信息时代,计算机能够表示、处理图像是必然的需求。纵观各类网站,其内容多是图文并茂,栩栩如生。通常将计算机能够显示、存储、处理的图像称为数字图像。

同时,在信息时代,计算机能够处理的图像将不限于人类肉眼局限的可见光(人类肉眼感知的图像仅限于比较小的一段电磁波谱),数字图像可以通过各种成像途径获取,这些途径可以覆盖几乎整个电磁波谱。如图 1.1 为用微波获取的金星图像。其中(b)图为(a)图一块 10 像素见方区域的灰度图,(c)图为(b)图的图像灰度值(值的范围为 0~255,有 256 个灰度等级),灰度的值的大小与微波反射能量的级别相对应。

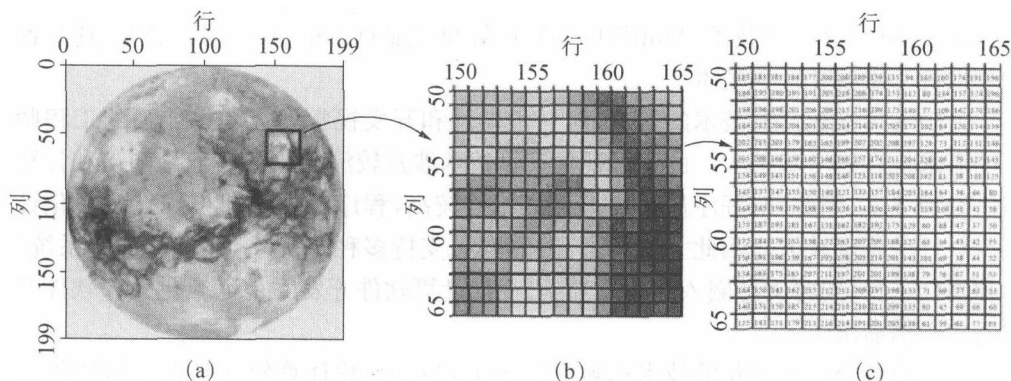


图 1.1 微波反射形成的金星图像

(a) 金星图片 (b) 某一块区域放大后效果 (c) 该区域图像灰度值

2) 数字图像处理

数字图像处理作为一门学科大约形成于 20 世纪 60 年代初期。早期的图像处理的目的是改善图像的质量,它以人为对象,以改善人的视觉效果为目的。但是数字图像处理不仅可以对人类肉眼可观察的图像进行加工处理,还可以对包括超声波、电子显微镜等成像机器产生的图像进行处理。因此,数字图像处理设计的领域非常广泛,如 70 年代生物医学中的 CT 技术(如图 1.2 所示),对人类作出了划时代的贡献。数字图像还广泛应用于航空航天(如图 1.3 所示)、生物医学工程、工业检测、机器人视觉等领域,使图像处理成为一门引人注目、前景远大的新型学科。

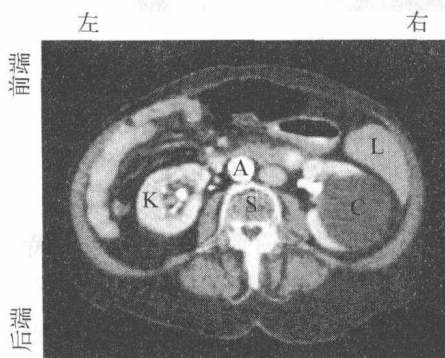


图 1.2 医学 CT 成像

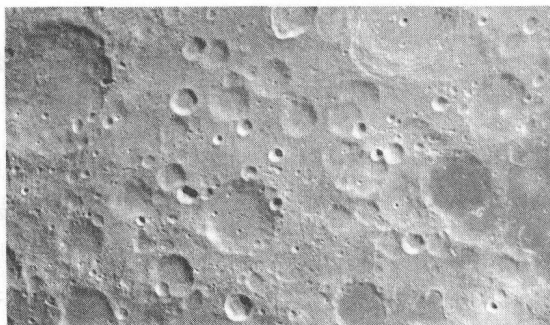


图 1.3 嫦娥一号卫星传回的月面图片

图 1.2 是一个人体肚脐部位的腹部成像的切片。图中可以看到很多器官,如(L)肝脏、(K)肾脏、(A)主动脉、(S)脊柱、(C)膀胱(盖住了右侧的肾脏)。CT 成像比传统的医学 X 射线对人体内部组织有更好的可视效果。

图像处理中,常用的图像处理方法有图像变换、图像编码和压缩、图像增强、图像分割、图像分类(识别)等。下面简要介绍。

(1) 图像增强。图像增强的目的是为了提高图像的质量,如去除噪声,提高图像的清晰度等。图像增强突出图像中所感兴趣的部分,如怎样处理可使图像中物体轮廓清晰,细节明显;怎样进行处理可减少图像中噪声的影响。如图 1.4 所示,在经过图像增强处理后,(b)图比(a)图整体上感觉更清晰。

(2) 图像变换。由于图像阵列很大,直接在空间域中进行处理,计算量将很大(由于数据存在着空间冗余信息)。因此,往往采用各种图像变换的方法,如傅里叶变换、沃尔什变换、离散余弦变换等间接处理技术,将空间域的处理转换为变换域处理,不仅可减少计算量,而且可获得更有效的处理(如傅里叶变换可在频域中进行数字滤波处理)。目前新兴研究的小波变换在时域和频域中都具有良好的局部



图 1.4 图像增强处理效果图

(a) 原图 (b) 图像增强处理后效果图

化特性,它在图像处理中也有着广泛而有效的应用。

(3) 图像编码和压缩。图像编码压缩技术可减少描述图像的数据量(比特数),以便节省图像传输、处理的时间和减少所占用的存储器容量。压缩可以在不失真的前提下获得,也可以在允许失真的条件下进行(在允许失真的情况下压缩比率通常会更高)。编码是压缩技术中最重要的方法,它在图像处理技术中是发展最早且比较成熟的技术。

(4) 图像分割。图像分割是数字图像处理中的关键技术之一。图像分割是将图像中有意义的特征部分提取出来,比如图像中的边缘、区域等,这是进一步进行图像识别、分析和理解的基础,也是计算机视觉中常用到的图像处理方法。虽然目前已研究出不少边缘提取、区域分割的方法,但还没有一种普遍适用于各种图像的有效方法。因此,图像分割是目前图像处理研究的热点之一,目前,其研究还在不断深入。

(5) 图像分类(识别)。图像分类(识别)属于模式识别的范畴,其主要内容是图像经过某些预处理(增强、复原、压缩)后,进行图像分割和特征提取,从而进行判决分类。图像分类常采用经典的模式识别方法,有统计模式分类等。

在本书的数字图像处理部分,将重点强调图像的增强、图像变换、压缩编码部分。

1.1.2 视频

1) 视频

当电影胶片以每秒 24 格画面匀速转动,一系列静态画面就会因“视觉暂留”作

用而造成一种连续的视觉印象,产生逼真的动感。电影就是基于人的视觉暂留作用原理制作成的,视频的基本原理也是一样的,拿电视来说,虽然有两个不同的标准 PAL(25 帧,我国和东南亚及欧洲等地方使用的制式)和 NTSC(30 帧,美国、日本等国的制式),其帧率都是大于 24 帧。这种一个个静态的图像经过快速的播放而形成动态、连续的画面叫做视频。视频是人们最熟悉的名词之一,日常生活当中电视机、录像机等都有视频输出接口。

2) 数字视频

电影、电视(传统模拟电视机)和录像已经属于传统的视听媒体,其视频信号属于模拟信号。随着现代科技的发展,已经出现了数字电视技术。与原来的模拟电视相比,数字电视有清晰的电视画面,可与 DVD 相媲美;有优质的音响效果,由于采用了数字技术,使得数字电视的伴音更趋逼真;有抗干扰功能,数字电视受其他电器的干扰很小,因此画面稳定;扩展功能多,可增加上网、点播等功能。数字视频的编辑通常是通过非线性编辑(for non-linear editing, NLE)系统进行的。几乎所有的电影、电视节目、广告都是用非线性编辑系统制作的。由于计算机网络和多媒体技术的发展,在个人计算机上也可以进行数字视频的编辑,越来越多的人利用计算机制作自己的小电影。同时,在因特网上的流式视频和点对点视频传输也都是新近的热点,“土豆”、“YouTube”等是关于数字视频的互联网网站,用户可以上载分享自己的视频,也能看到别人分享的视频。数字视频(Digital video)信息技术已经渗透到工作、学习、娱乐的各个方面,成为我们生活中不可或缺的组成部分。

有很多不同的数字视频编码方法和文件容器格式,支持不同大小、质量、分辨率、色度精度和编码功能的图像编码。在本书的第 7 章,将详细介绍这些编码原理知识和相关标准。

1.1.3 音频

和视频一样,音频也分为模拟音频和数字音频。任何可以听见的声音经过音频线或话筒的传输都是一系列的模拟信号。也就是说模拟信号是我们可以听见的。而数字信号就是用一堆数字记号(其实只有二进制的 1 和 0)来记录声音,而不是用物理手段来保存信号(用普通磁带录音就是一种物理方式)。当然我们实际上也听不到数字信号。

音频是个专业术语,人类能够听到的所有声音都称之为音频。声音被录制下来以后,无论是说话声、歌声、乐器都可以通过数字音乐软件处理,或是把它制作成 CD,这时候所有的声音没有改变,因为 CD 本来就是音频文件的一种类型。而音频

只是储存在计算机里的声音。如果有计算机再加上相应的音频卡就可以把所有的声音录制下来,声音的声学特性如音的高低等都可以用计算机硬盘文件的方式储存下来。反过来,也可以把储存下来的音频文件用一定的音频程序播放,还原以前录下的声音。

模拟时代的录音制作与数码时代的区别:模拟时代是把原始信号以物理方式录制到磁带上(当然在录音棚里完成了),然后加工、剪接、修改,最后录制到磁带、密纹唱片等广大听众可以欣赏的载体上。这一系列过程全是模拟的,每一步都会损失一些信号,到了听众手里自然是差了好远,更不用说什么高保真(HI-FI)了。数码时代是第一步就把原始信号录成数码音频资料,然后用硬件设备或各种软件进行加工处理,这个过程与模拟方法相比有无比的优越性,因为它几乎不会有任何损耗。对于机器来说这个过程只是处理一下数字而已,当然丢码的可能性也有,但只要操作合理就不会发生。最后把这堆数字信号传输给数字记录设备如 CD 等,损耗自然小很多了!

在计算机内播放或是处理音频文件,也就是要对声音文件进行数、模转换,这个过程同样由采样和量化构成,人耳所能听到的声音,最低的频率是从 20Hz 起一直到最高频率 20 kHz,20 kHz 以上人耳是听不到的,因此音频的最大带宽是 20 kHz,音频有很多种格式,如 CD 格式(接近无损)、wav、MP3 等。

本书不对音频的相关技术进行详细阐述。

1.1.4 游戏

自从人类诞生的时候就有游戏,自古至今人类发明的游戏不计其数,比如各式各样的棋牌类游戏、体育竞技游戏,甚至古罗马的竞技场上残忍的搏斗,在有些人的眼里,也是一种轻松的游戏。

在信息时代,传统的游戏得到了延伸和扩展。电脑游戏现在已经非常流行,在行业者的眼里已经是一个巨大的产业。电脑游戏是指在电子计算机上运行的游戏软件,是一种本身提供娱乐功能的电脑软件。电脑游戏产业与电脑硬件(视频显示卡等)、电影、电脑软件、互联网的发展联系甚密。电脑游戏为游戏参与者提供了一个虚拟的空间,从一定程度上让人可以摆脱现实世界,在另一个世界中扮演真实世界中扮演不了的角色(如 CS、星际等)。同时电脑多媒体技术的发展,使游戏给了人们很多体验和享受。

电脑游戏的发展离不开计算机软硬件的发展,特别是游戏软件开发,涉及游戏创意、网络、音视频、嵌入式软件(如手机游戏等)、人机交互、人工智能等各个领域,是和视觉计算分不开的,也是一项极具挑战性的工作。

1.2 软件开发与优化基础

1.2.1 C 语言概述

计算机的发展离不开软件的发展。在软件的发展历史中,高级语言的出现无疑是一个里程碑。在高级语言出现之前,科学工作者主要通过机器语言(机器执行时的 0,1)或汇编语言来操纵计算机。由于这些语言非常底层,虽然计算机容易理解这些语言,但是对人来说非常抽象、不直观、难度大、效率低等,因此一种更高级的语言是计算机软件发展的必然需求。高级语言能摆脱低级语言的束缚,人们只要关注应用的逻辑,而不用关注这些逻辑在计算机上具体是怎样通过 0 和 1 来实现的(比如要实现一个加法功能,可以很容易用高级语言的表示式来实现,而不用关心加法在处理器中怎样通过寄存器实现的)。

Fortran 是第一个电脑高级语言,它是 1954 年美国的 IBM 的 IT 成果。C 语言是在 20 世纪 70 年代初问世的。1978 年由美国电话电报公司(AT&T)贝尔实验室正式发表了 C 语言。早期的 C 语言主要是用于 UNIX 系统。由于 C 语言的强大功能和各方面的优点逐渐为人们认识,到了 80 年代,C 开始进入其他操作系统,并很快在各类大、中、小和微型计算机上得到了广泛的使用,成为当代最优秀的程序设计语言之一。

C 语言是一种结构化语言。它层次清晰,便于按模块化方式组织程序,易于调试和维护。C 语言的表现能力和处理能力极强。它不仅具有丰富的运算符和数据类型,便于实现各类复杂的数据结构,还可以直接访问内存的物理地址,进行位(bit)一级的操作。由于 C 语言实现了对硬件的编程操作,因此 C 语言集高级语言和低级语言的功能于一体,既可用于系统软件的开发,也适合于应用软件的开发。此外,C 语言还具有效率高、可移植性强等特点。因此广泛地移植到了各类各型计算机上,从而形成了多种版本的 C 语言。

C 语言的数据类型包括基本数据类型(整型、实型等)、构造数据类型(数组类型、结构类型、联合类型)、指针类型(其值用来表示某个量在内存存储器中的地址)。

C 语言的运算符有算术运算符、关系运算符、位操作运算符、条件运算符、指针运算符等。

C 语言的语句分为表达式语句、函数调用语句、控制语句(条件判断语句、循环语句、转向语句)、复合语句等。