大家好，本节课我们将学习第一章第二节第三部分内容数据编码与压缩。

数字是我们日常生活中不可缺少的朋友，可是数字除了计算、用于表示物品数量之外，还有很重要的用途。下面，我们来看一下数字的威力吧！

身份证是能够证明一个人身份的证件，身份证号码由18位数字组成，前6位为行政区划代码，第7至14位为出生日期码，第15至17位为顺序码，第18位为校验码。

在我们的身边还有很多使用数据编码的实例，比如车辆的车牌号、电影院的座位号、商品的保质期等。同样，在计算机中数据也需要编码，才能够被处理，无论处理的是字符、图像、声音，还是其他形式的内容，都需要转换成二进制形式的编码。

首先，我们来了解字符编码。字符是人与计算机交互过程中不可或缺的重要内容，它是许多种文字和符号的总称。由于计算机只识别0和1，因此在处理各种字符时，就需要将字符转换为计算机可以识别的二进制数据，以便文本在计算机中存储和通过通信网络的传递。常见的例子包括将拉丁字母表编码成摩斯电码和ASCⅡ。计算机中存储一个“0”或“1”占用一个二进制位（bit），8个二进制位组成1字节（byte）。ASCⅡ编码用1字节表示英文字母、数字和常见字符。计算机中常见的存储单位与换算关系如表所示。

接着，我们来学习声音编码。声音是人们用来表达和传递信息的一种重要形式，从物理学上看，声音是一种波。应用计算机处理声音时，需要将声波的模拟信号转换为数字信号，也就是声音的数字化，声音数字化通过采样、量化和编码来实现。

通常，音频所占的存储容量取决于采样频率、量化位数、声道数和时长，其计算公式为：

音频所占的存储容量=采样频率×量化位数×声道数×时长/8

我们通过一道例题来理解此公式，一首时长为100s的双声道音乐，采样频率为44100Hz，量化位数为16，计算该音乐的音频所占的存储容量。

音频所占的存储容量=44100×16×2×100/8= 17640000B

44100

17640000/1024/1024≈17MB

最后，我们来了解一下图像编码。一幅图像可以看作由许多彩色或各种级别灰度的点组成的，这些点按横纵进行排列，被称为像素。(图)图像中每个像素的颜色值都用一个或多个二进制位来存储。通过扫描仪、数字照相机和智能手机等设备，可以把图像数字化，再经过相关软件处理，就可以把图像转换成相应格式的文件进行保存。图像所占的存储容量可用该图像所有像素的字节数来表示，其计算公式为：

图像所占的存储容量=水平像素数×垂直像素数×颜色深度/8

我们依旧以一个例题来理解该公式。图像尺寸为90×72像素，颜色深度为24位，计算该图像所占的存储容量。

图像所占的存储容量=90×72×24/8=19440B

19440/1024≈19KB

通过前面的学习，我们可以看出，很多信息，尤其是图像、音频和视频等类型的信息，经过数字化后生成的数据量较大，包含了许多冗余信息。例如，一幅图像中的建筑背景、蓝天和绿地，其中许多像素信息是相同的；一段视频的相邻帧之间，除了运动物体有少许变化外，其他静止的像素信息也都是重复的。因此，在不损失有用信息的前提下，可以按照一定的编码规则对数据进行重新组合，以去除数据冗余，也就是对数据进行压缩。

常用的压缩方法分为无损压缩和有损压缩。无损压缩指对压缩后的数据进行还原后，得到的数据与压缩前完全相同。例如，未压缩的数据是8个字母A，经过压缩后可以用A08代替。常用的无损压缩算法有zip压缩算法和7z压缩算法等。压缩后所生成的文件称为压缩包，可能只有原来文件的几分之一，甚至更小。

相对于无损压缩，有损压缩指在压缩过程中会损失一定的信息， 压缩后的数据无法还原成压缩前的样子。例如，音频、图像和视频数字化后存在很多冗余信息，而损失一些冗余信息并不影响视听效果，因此有损压缩被广泛应用于音频、图像和视频文件。

本节课我们学习的内容较多，大家来一起回顾一下吧。首先我们学习了数据编码，包括字符编码、声音编码和图像编码，接着我们学习了数据压缩，包括无损压缩和有损压缩。

本节课的内容就到这里了，谢谢大家。