

第3章 图像的采集及压缩编码

3.1 相机概述

数码相机是在照相机的基础上发展起来的,最早出现在美国。尽管发展的历史只有 40 多年,但它从一开始就显示了强大的生命力,当时主要利用它通过卫星向地面传送照片。进入 20 世纪 90 年代以后,数码相机有了较大发展,许多世界著名的照相机生产厂家纷纷致力于数码相机的研究、开发。从 20 世纪 90 年代中期开始,实用型的数码相机开始出现在全球市场。1997 年全球销量达到 200 万台,1999 年出货量为 490 万台。经过 10 多年的发展,2010 年全球数码相机出货量已达 1.41 亿台,已基本达到最大销量。2014 年,著名的数码相机市场分析公司 Agency GFK 公布了在最近一年内数码相机类产品在全球市场的销售状况。数据显示,目前普通便携数码相机的销售继续萎缩,市场占有率下降得比较厉害;具备高性能以及高价格的数码相机,销量却在不断攀升中,其中,具备可更换镜头特性的数码相机,比如微单、单电相机在全球范围内的销量继续增加,而且会给厂商带来丰厚的利润。我国数码相机市场起步较晚,1996 年年底卡西欧和柯达公司的数码相机进入中国市场,1998 年数码相机市场销量只有 4 万台。面对国外知名厂商的大举进攻,国内传统照相机生产厂商,如海鸥、凤凰等厂商也开始生产数码相机,如 2000 年中国摄影行业博览会上,海鸥 DSC 一 1100 型百万像素的数码相机亮相,引起了业界的广泛关注。目前,中国是全球最大的数码相机产地,2010 年数码相机出口 1.09 亿台,出口额达 70.7 亿美元,同比增长了 12%。2011 年 1~3 月,数码相机进出口总额达 19.04 亿美元,进口额达 5.69 亿美元,出口额达 13.35 亿美元,共出口了 2011 万台。

数码相机也称为数字照相机,它是继扫描仪之后发展起来的计算机外部输入设备。它集成了图像系统的转换、存储、传输等部件,具有数字化存储模式、与计算机交互处理、实时拍摄等特点,它是机、光、电一体化的产品。它可将瞬息万变的缤纷世界轻而易举地拍摄成影像,并以数字方式输入给计算机,同时可根据需要进行修改、处理,近年来数码相机已经开始畅销起来,随着网络的应用,它将越来越多地受到人们的关注。

数码相机由光学镜头、光电传感器 CCD 电荷耦合器件、CMOS 互补型金属氧化物半导体)、模/数转换器(ADC)、图像处理器(MPU 微处理器)、影像存储器、显示屏和输出控制单元等基本部件组成。它的工作原理与扫描仪的工作原理是类似的,都是利用电荷耦合器件 CCD(Charge Coupled Device)作为感光器件。数码相机利用一个光学镜头让影像聚焦成像在感光元件 CCD 上,然后通过模/数转换器将每一个像素上的光信号转换成数字信号,传输给相机中的缓存,再通过图像处理器读出缓存中的数字信号,判读成图像信号后将其压缩存放在数码相机的存储器中。与扫描仪不同的是,它是通过光学镜头在光圈和快门的控制下,实现在电子存储设备上的“曝光”,完成被拍摄影像的记录。

数码相机的外形酷似传统照相机,因为它是在照相机的基础上发展起来的,它的成像机

理与传统的相机完全相同，都是利用了光学镜头的透镜成像原理，但在影像处理上却有很大的不同。数码相机处理快捷，不需要胶卷，免去了冲洗胶卷的等待时间。如果相机带有液晶屏，在拍摄的同时就可以看到拍摄效果。同时数码相机具有灵活性，如增加音频记录和视频输出，就可以将照片在电视机上播放，连接到计算机上，可作为计算机的输入设备，适应能力强，使用方便。因此，数码相机具有普通照相机无可比拟的优越性。

由于数码相机的工作原理与扫描仪的工作原理类似，因此有许多指标，如分辨率、色彩位数、色调调整、接口等与扫描仪类似。同时数码相机属于照相机的范畴，因此具有与普通相机相同的性能，如曝光方式、测光方式、快门时间范围、曝光补偿方式、曝光补偿范围，对焦方式等。因而衡量数码相机的性能指标主要有以下几点。

(1) 分辨率

数码相机生成图像的质量在一定程度上取决于图像传感器像素数量。目前所用的传感器有 CCD、Super CCD 和 CMOS。所包含的像素越多，记录的细节就越多，分辨率越高，生成的图像质量就越好，普通数码相机的分辨率有 $640 \times 480\text{dpi}$ 、 $800 \times 600\text{dpi}$ 、 $1024 \times 768\text{dpi}$ 等。一般情况下，影像传感器的像素数应大于所拍摄物体的最大像素数，这个数量可以由分辨率计算出来，如柯达公司的 DC4800 数码相机，分辨率为 $2240 \times 1500\text{dpi}$ ，其包含的像素为 $2240 \times 1500 = 33600$ ，即 336 万像素，实际标称为 330 万像素，这就意味着图像传感器的最大像素数量和用于感光的像素数量以及最后所能生成的最大分辨率图像的像素数量并不完全相同。这是因为数码相机的图像传感器中并不是所有的像素都能够用于生成图像，CCD 边缘一部分像素被遮挡住了，这些像素被用于检测没有光线照射时的电荷值。

另外，数码相机也有光学分辨率与插值分辨率之分。光学分辨率是根据纵横像素数测定的，像素数越多，分辨率越高，拍摄效果就越好。插值分辨率是采用软件插值的方法增加图像的像素数，实际意义不是很大。

(2) 光学镜头

光学镜头是数码相机的重要部件，与传统的相机类似，主要指标有镜头大小、镜头焦距和变焦范围。数码相机的光学镜头主要分为 3 种类型，即固定聚焦镜头、自动聚焦镜头和手动聚焦镜头。从一般的常识来讲，镜头越大越好。同时也与镜头的口径大小——光圈有关。光圈越大，进光量越多，通常用光圈系数来表示（光圈系数 $f = \text{焦距} / \text{光孔直径}$ ）。镜头的焦距，即焦点与光线离开镜头位置之间的距离，通常用 mm 表示。如果被拍摄景物距离不变，镜头的焦距越大，所拍摄景物放大倍率越大。固定聚焦镜头拍摄质量比较好，使用方便，不用对焦，但是拍摄的景物受限制，它正逐步被自动聚焦镜头和手动聚焦镜头所取代。变焦范围决定了拍摄物体的选择范围，变焦镜头的焦距变化倍数越大，说明数码相机的性能越好。光学变焦倍数指的是在相同距离下，从最小焦距变化到最大焦距时拍摄同一景物所能放大的倍数。通常有些数码相机的镜头上会标出一些数据，如奥林巴斯的 C-2500L 数码相机的镜头上就标有“OLYMPUSLENSAF ZOOM9.2-28mm 1: 2.8-3.9”，这就表示“奥林巴斯变焦镜头，最小焦距为 9.2mm，此时的最大光圈为 2.8；最大焦距为 28mm，此时的最大光圈为 3.9”，说明该镜头具有 3 倍变焦。

(3) 存储装置

数码相机的图像以数字方式存储在磁或光介质上。存储器可分为内置式存储器和可移动式存储器。早期的数码相机均使用可移动式存储器，分为 I 型、II 型、III 型 PCMCIA 卡（简

称 PC 卡) 和 SSFDC 卡 (简称智能卡) 等多种。目前常见的存储介质有 Compact Flash 卡、Smart Media 卡、Memory Stick 以及软盘, 一些新型数码相机甚至使用了 IBM 的 MicroDrive 硬盘。各种介质的存储容量不同, 而且它们之间互不兼容。

(4) 色彩位数

色彩位数与扫描仪的色彩位数相同, 它反映出数码相机所拍摄图像所含色彩的丰富程度, 位数越多, 所表现的色彩越丰富。目前数码相机多为 24 位或 36 位。

(5) 连拍速度

拍摄活动的景物或连续拍摄时, 要求相机能提供较短的曝光时间和较高的连拍速度。目前数码相机中最快的连拍速度可达 5 幅/秒, 但连拍 3 秒左右后必须等待 10 秒左右才能继续拍摄。

(6) 通信接口

数码相机最常见的传输接口是 RS-232C 串口, 但是传输速度很慢, 最快传输速率也仅为 912 600bps。一般高档数码相机通过 SCSI 或 SCSI-2 接口与计算机相连。目前也有一些数码相机采用 USB 接口, 同时也出现了一些 IEEE 1394 接口的数码相机, 使传输速率有了很大的提高。此外还有一些数码相机支持红外线接口, 只需将电脑与数码相机的红外线对准, 就可以进行档案传输, 不用再连接数据线, 十分方便。

(7) 白平衡

有的数码相机有白平衡调整, 可分为自动调整和手动调整两类方式。

(8) 相当感光度

数码相机与普通相机不同, 对曝光量有一定的要求, 也就有了感光灵敏度高低的问题。部分数码相机的相当感光度有一范围, 通常应将它置于最佳感光度上。这个数值越大, 相机感光器要正常成像需要的光线就越少, 有助于在光线较暗而闪光灯有可能效果不佳 (例如在较大的室内) 的环境下拍摄照片。

3.2 光敏传感器

3.2.1 光敏传感器概述及原理

光敏传感器是采用光电元件作为检测元件的传感器。它首先把被测量的变化转换成光信号的变化, 然后借助光电元件进一步将光信号转换成电信号。光敏传感器一般由光源、光学通路和光电元件三部分组成。光电检测方法具有精度高、反应快、非接触等优点, 而且可测参数多, 传感器的结构简单, 形式灵活多样, 因此, 光电式传感器在检测和控制中应用非常广泛。光电传感器是各种光电检测系统中实现光电转换的关键元件, 它是把光信号 (红外、可见及紫外光辐射) 转变成为电信号的器件。

光敏传感器是利用光敏元件将光信号转换为电信号的传感器, 它的敏感波长在可见光波长附近, 包括红外线波长和紫外线波长。光传感器不只局限于对光的探测, 它还可以作为探测元件组成其他传感器, 对许多非电量进行检测, 只要将这些非电量转换为光信号的变化即可。光敏传感器是目前产量最多、应用最广的传感器之一, 它在自动控制和非电量电测技术中占有非常重要的地位。光敏传感器的种类繁多, 主要有: 光电管、光电倍增管、光敏电阻、

光敏三极管、光电耦合器、太阳能电池、红外线传感器、紫外线传感器、光纤式光电传感器、色彩传感器、CCD 和 CMOS 图像传感器等。

光敏传感器是以光电器件作为转换元件的传感器。它可用于检测直接引起光量变化的非电量，如光强、光照度、辐射测温、气体成分分析等；也可用来检测能转换成光量变化的其他非电量，如零件直径、表面粗糙度、应变、位移、振动、速度、加速度，以及物体的形状、工作状态的识别等。光电式传感器具有非接触、响应快、性能可靠等特点，因此在工业自动化装置和机器人研究中获得广泛应用。近年来，新的光电器件不断涌现，特别是 CCD 图像传感器的诞生，为光电传感器的进一步应用开创了新的一页。

3.2.2 光电效应

光电效应是指当具有一定能量 E 的光子投射到某些物质的表面时，具有辐射能量的微粒将透过受光的表面层，赋予这些物质的电子以附加能量，或者改变物质的电阻大小，或者使其产生电动势，导致与其相连接的闭合回路中电流的变化，从而实现了光—电转换过程。光电效应一般分为下述两种。

(1) 外光电效应

光照射在光电材料上，材料表面的电子吸收能量，若电子吸收的能量足够大，电子会克服束缚脱离材料表面而进入外界空间，从而改变光电子材料的导电性，这种现象称为外光电效应。

根据爱因斯坦的光电子效应，光子是运动着的粒子流，每种光子的能量为 $h\nu$ (ν 为光波频率， h 为普朗克常数， $h = 6.63 \times 10^{-34} \text{ J/Hz}$)，由此可见不同频率的光子具有不同的能量，光波频率越高，光子能量越大。假设光子的全部能量交给光子，电子能量将会增加，增加的能量一部分用于克服正离子的束缚，另一部分转换成电子能量。根据能量守恒定律：

$$\frac{1}{2}mv^2 = h\nu - A \tag{3-1}$$

式中， m 为电子质量， v 为电子逸出的初速度， A 微电子所做的功。

由上式可知，要使光电子逸出阴极表面的必要条件是 $h\nu > A$ 。由于不同材料具有不同的逸出功，因此对每一种阴极材料，入射光都有一个确定的频率限，当入射光的频率低于此频率限时，不论发光强度有多大，都不会产生光电子发射，此频率限值称为“红限”。相应的波长为

$$\lambda_x = \frac{hc}{A} \tag{3-2}$$

式中， c 为光速， A 为逸出功。

(2) 内光电效应

在光线作用下能使物体电阻率改变的称为内光电效应。内光电效应一般有光导效应和光生伏特效应。

① 当受到光照射时，吸收电子能量，其电阻率降低的导电现象称为光导效应。它属于内光电效应。当光照在半导体上时，若电子的能量大于半导体禁带的能级宽度，则电子从价带跃迁到导带，形成电子，同时，价带留下相应的空穴。电子、空穴仍留在半导体内，并参与导电在外电场作用下形成电流。

除金属外，多数绝缘体和半导体都有光电效应，半导体尤为显著，根据光导效应制造的光电元件有固有入射光频率，当光照在光电阻上，其导电性增强，电阻值下降。光强度越强，

其阻值越小，若停止光照，其阻值将恢复到原阻值。

② 在光线作用下能够使物体产生一定方向的电动势的现象叫做光生伏特效应。据此效应制造的光电器件有光电池、光电二极管、三极管、管控晶闸管和光耦合器等。

3.2.3 光电转换元件

根据光电效应种类的不同，光电转换元件分为外光电转换元件和内光电转换元件。根据外光电效应制造的光电元件有光电管和光电倍增管；根据内光电效应制造的光电元件有光敏电阻、光电池等。

(1) 光电管

光电管的种类繁多，典型的产品有真空光电管和充气光电管，它的外形和结构如图 3-1 所示，半圆筒形金属片制成的阴极 K 和位于阴极轴心的金属丝制成的阳极 A 封装在抽成真空的玻璃壳内，当入射光照射在阴极上时，单个光子就把它的全部能量传递给阴极材料中的一个自由电子，从而使自由电子的能量增加 $h\nu$ 。当电子获得的能量大于阴极材料的逸出功 A 时，它就可以克服金属表面束缚而逸出，形成电子发射。这种电子称为光电子，光电子逸出金属表面后的初始动能为 $\frac{1}{2}mv^2$ 。

光电管正常工作时，阳极电位高于阴极，如图 3-2 所示。在入射光频率大于“红限”的前提下，从阴极表面逸出的光电子被具有正电位的阳极所吸引，在光电管内形成空间电子流，称为光电流。此时若光强增大，轰击阴极的光子数增多，单位时间内发射的光电子数也就增多，光电流变大。在图 3-2 中所示的电路中，电流和电阻上的电压降就和光强成函数关系，从而实现光电转换。当光线照射到光电阴极 K 上时，电子从阴极表面逸出，并被光电阳极的正电厂吸收，外电路产生电流 I ，在负载电阻 R 上的电压 U 。

光电管的光电特性如图 3-3 所示，从图中可知，在光通量不太大时，光电特性基本是一条直线。

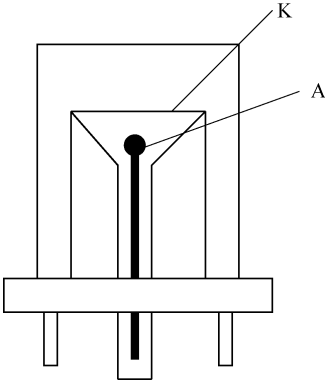


图 3-1 光电管结构示意图

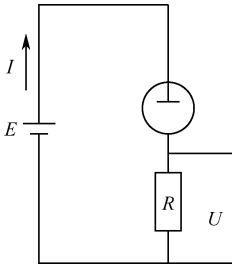


图 3-2 光电管测量电路

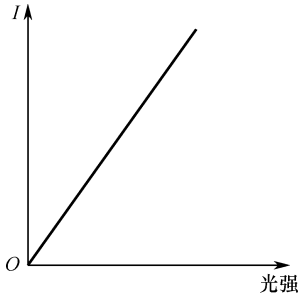


图 3-3 光电管的光电特性

光电管的特性主要取决于光电极的材料，其基本的特性是光谱特性、光电特性和伏安特性。

① 光谱特性

用单位辐射通量、不同波长的光分别照射光电管，在光电管上产生大小不同的光电流。这里，光电流 I 与光波波长 λ 的关系曲线称为光谱特性曲线，又称频谱特性。

对于不同波长区域的光，应选用不同光电阴极的光电管。此外在测量与控制技术中，光

电管可以担负人眼不能胜任的工作。

② 光电特性

光电管在固定阳极电压下，光通量与光电流（阳极电流）之间的关系称为光电特性。图 3-4 为光电管的光电特性曲线。从图 3-4 可知，光电管的光电特性基本上呈线性关系，直线的斜率为其灵敏度。

③ 伏安特性

光电管在光通量一定的情况下，阳极电压与阳极电流的关系称为伏安特性。图 3-5 为光电管的伏安特性曲线。在阳极电压大于 50V 时，光电流开始饱和，阳极电流近于常数，而与电压无关。真空光电管一般工作于伏安特性的饱和部分，内阻达几百兆欧。

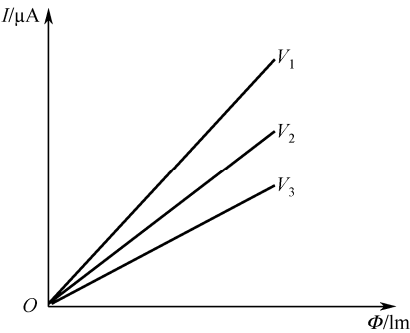


图 3-4 光电管的光电特性曲线

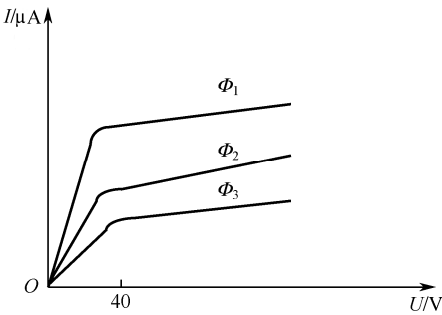


图 3-5 光电管的伏安特性曲线

(2) 光电倍增管

由于真空光电管的灵敏度低，因此人们研制了具有放大光电流能力的光电倍增管。图 3-6 是光电倍增管结构示意图。

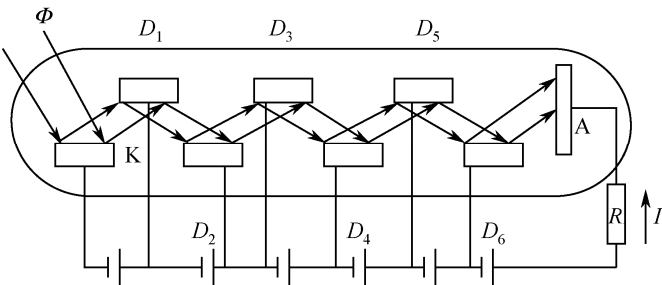


图 3-6 光电倍增结构示意图

从图中可以看到光电倍增管也有一个阴极 K 和一个阳极 A ，与光电管不同的是在它的阴极和阳极间设置了若干个二次发射电极， $D_1, D_2, D_3, \dots$ 称为第一倍增电极，第二倍增电极，第三倍增电极， $\dots$ ，倍增电极通常为 $10\sim15$ 级。光电倍增管工作时，相邻电极之间保持一定电位差，其中阴极电位最低，各倍增电极电位逐级升高，阳极电位最高。当入射光照照射阴极 K 时，从阴极逸出的光电子被第一倍增电极 D_1 加速，以高速轰击 D_1 ，引起二次电子发射，一个入射的光电子可以产生多个二次电子， D_1 发射出的二次电子又被 $D_1、D_2$ 间的电场加速，射向 D_2 并再次产生二次电子发射……，这样逐级产生的二次电子发射，使电子数量迅速增加，这些电子最后到达阳极，形成较大的阳极电流。若倍增电极有 n 级，各级的倍增率为 σ ，则光电倍增管的倍增率可以认为是 σ_N ，因此，光电倍增管有极高的灵敏度。在

输出电流小于 1mA 的情况下，它的光电特性在很宽的范围内具有良好的线性关系。光电倍增管的这个特点，使它多用于微光测量。

(3) 光敏电阻

当光照射到半导体材料上时，价带中的电子受到能量大于或等于禁带宽度的光子轰击，并使其由价带越过禁带跃入导带，如图 3-7 所示，使材料中导带内的电子和价带内的空穴浓度增加，从而使电导率变大。

光敏电阻的结构如图 3-8 所示。管芯是一块安装在绝缘衬底上带有两个欧姆接触电极的光电导体。光导体吸收光子而产生的光电效应，只限于光照的表面薄层，虽然产生的载流子也有少数扩散到内部去，但扩散深度有限，因此光电导体一般都做成薄层。为了获得高的灵敏度，光敏电阻的电极一般采用梳状图案。

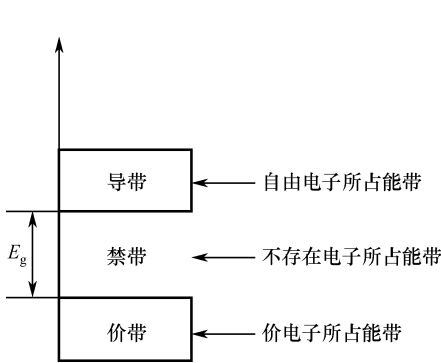


图 3-7 光敏电阻工作原理

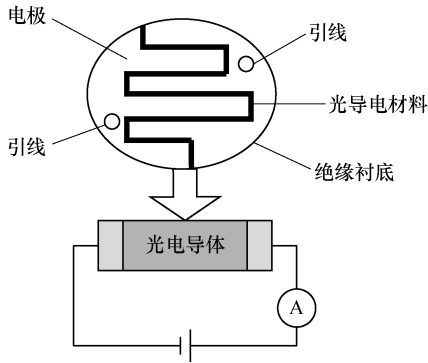


图 3-8 金属封装的硫化镉光敏电阻结构

光敏电阻的工作原理是基于内光电效应。在半导体光敏材料的两端装上电极引线，将其封在带有透明窗的管壳里就构成了光敏电阻。

(4) 光电池

光电池是利用光生伏特效应将光直接转变成电能的器件。由于它可把太阳能直接变电能，因此又称为太阳能电池。它是基于光生伏特效应制成的，是发电式有源元件。它有较大面积的 PN 结，当光照射在 PN 结上时，在结的两端出现电动势。

命名方式：把光电池的半导体材料的名称置于光电池（或太阳能电池）之前。如硒光电池、砷化镓光电池、硅光电池等。目前，应用最广、最有发展前途的是硅光电池。

各种光电池的特性如下：

- ①硅光电池价格便宜，转换效率高，寿命长，适于接收红外线。
- ②硒光电池光电转换效率低（0.02%）、寿命短，适于接收可见光（响应峰值波长 0.56 μm ），最适宜制造照度计。
- ③砷化镓光电池转换效率比硅光电池稍高，光谱响应特性则与太阳光谱最吻合。其工作温度最高，更耐受宇宙射线的辐射。因此，它在宇宙飞船、卫星、太空探测器等电源方面的应用是有发展前途的。

硅光电池的结构如图 3-9 所示。它是在一块 N 型硅片上用扩散的办法掺入一些 P 型杂质（如硼）形成 PN 结。当光照到 PN 结区时，如果光子能量足够大，将在结区附近激发出电子-空穴对，在 N 区聚积负电荷，P 区聚积正电荷，这样 N 区和 P 区之间将出现电位差。若将 PN 结两端用导线连起来，则电路中有电流流过，电流的方向由 P 区流经外电路至 N 区。若将外电路断开，就可测出光生电动势。

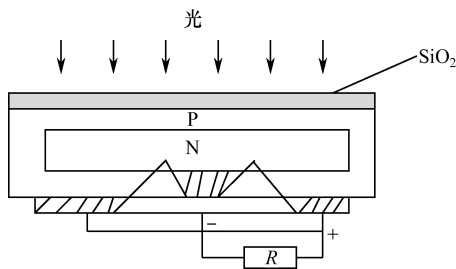


图 3-9 硅光电池的结构

3.2.4 光敏传感器的应用

(1) 包装充填物高度检测

用容积法计量包装的成品，除了对重量有一定误差范围要求外，一般还对充填高度有一定的要求，以保证商品的外观质量，不符合充填高度的成品将不许出厂。图 3-10 所示为借助光电检测技术控制充填高度的原理。当充填高度 h 偏差太大时，光电接头将失去电信号，即由执行机构将包装物品推出进行处理。

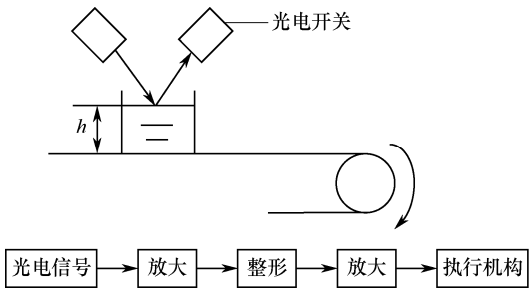


图 3-10 利用光电检测技术控制充填高度

(2) 光电色质检测

图 3-11 所示为包装物料的光电色质检测原理。若包装物品规定底色为白色，因质量不佳，有的出现泛黄，则在产品包装前应先由光电检测色质，物品泛黄时就有比较电压差输出，接通电磁阀，由压缩空气将泛黄物品吹出。

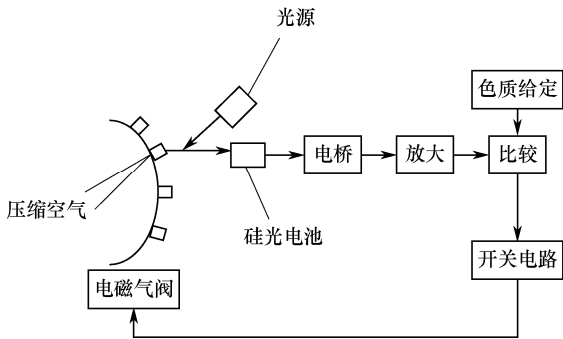


图 3-11 包装物料的光电色质检测原理

(3) 烟尘浊度监测仪

防止工业烟尘污染是环保的重要任务之一。为了消除工业烟尘污染，首先要知道烟尘排

放量，因此必须对烟尘源进行监测、自动显示和超标报警。烟道里的烟尘浊度是用通过光在烟道里传输过程中的变化大小来检测的。如果烟道浊度增加，光源发出的光被烟尘颗粒的吸收和折射增加，那么到达光检测器的光将减少，因而光检测器输出信号的强弱便可反映烟道浊度的变化。

(4) 光敏传感器其他方面的应用

利用光电开关还可以进行产品流水线上的产量统计、对装配件是否到位及装配质量进行检测，例如灌装时瓶盖是否压上、商标是否漏贴等（如图 3-12 所示）。

此外，利用反射式光电传感器还可以检测布料的有无和宽度。利用遮挡式光电传感器检测布料的下垂度，其结果可用于调整布料在传送中的张力，利用安装在框架上的反射式光电传感器可以发现漏装产品的空箱，并利用油缸将空箱推出。

光电传感器采用光电元件作为检测元件，首先把被测量的变化转变为信号的变化，然后借助光电元件进一步将光信号转换成电信号。光电传感器一般由光源、光学通路和光电元件 3 部分组成。光电检测方法具有精度高、反应快、非接触等优点，而且可测参数多，传感器的结构简单，形式灵活多样，体积小。近年来，随着光电技术的发展，光电传感器已成为系列产品，其品种及产量日益增加，用户可根据需要选用各种规格产品，其在各种轻工自动机上获得了广泛的应用。

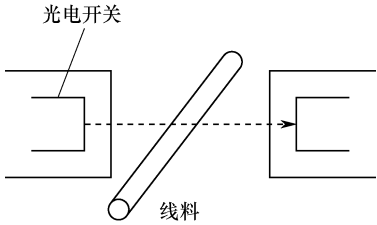


图 3-12 瓶子灌装检测示意图

3.3 图像编码技术

3.3.1 图像编码概述

第 1 章中我们已经提到，视觉物联网的带宽资源和能耗非常有限，所以，在对采集的数据进行传输前必须进行编码压缩。

图像编码也称为图像压缩，是指在满足一定保真度的要求下，对图像数据的进行变换、编码和压缩，去除多余数据减少表示数字图像时需要的数据量，以便于图像的存储和传输，即以较少的数据量有损或无损地表示原来的像素矩阵的技术。

图像压缩编码可分为两类：一类压缩是可逆的，即从压缩后的数据中可以完全恢复原来的图像，信息没有损失，称为无损压缩编码；另一类压缩是不可逆的，即从压缩后的数据中无法完全恢复原来的图像，信息有一定损失，称为有损压缩编码。

随着人们对传统压缩编码方法的深入研究和应用，逐渐发现了这些传统方法的许多缺点。如高压缩比时恢复图像会出现方块效应，人眼视觉系统（HVS）的特性不易被引入到算法中等。为了克服这些缺点，1985 年 M. Kuni 等人提出了第二代图像压缩编码的概念。经过 30 多年的发展，在这一框架下，人们提出了几种新的编码方法：分形编码、小波变换编码和基于模型的编码方法等。于是，对数据压缩技术的研究就突破了传统 Shannon 理论的框架，使得压缩效率得以极大提高。

数字图像的冗余主要表现为以下几种形式：空间冗余、时间冗余、视觉冗余、信息熵冗余、结构冗余和知识冗余。图像数据的这些冗余信息为图像压缩编码提供了依据。图像编码的目的就是充分利用图像中存在的各种冗余信息，特别是空间冗余、时间冗余以及视觉冗余，

以尽量少的比特数来表示图像。利用各种冗余信息，压缩编码技术能够很好地解决在将模拟信号转换为数字信号后所产生的带宽需求增加的问题，它是使数字信号走上实用化的关键技术之一。虽然表示图像需要大量的数据，但是图像数据是高度相关的，或者说存在冗余信息，去掉这些信息后可以有效压缩图像，同时不会损害图像的有效信息。

3.3.2 图像编码方法

经典图像编码技术根据编码原理可分为：熵编码、预测编码、变换编码和混合编码等。

1. 熵编码

熵编码即编码过程中按熵原理不丢失任何信息的编码。信息熵为信源的平均信息量（不确定性的度量）。常见的熵编码有：香农（Shannon）编码、哈夫曼（Huffman）编码和算术编码（Arithmetic Coding）。在视频编码中，熵编码把一系列用来表示视频序列的元素符号转变为一个用来传输或存储的压缩码流。输入的符号可能包括量化后的变换系数、运动向量、头信息（宏块头、图像头、序列的头等）以及附加信息（对于正确解码来说重要的标记位信息）。

(1) 行程编码

熵编码即编码过程中按熵原理不丢失任何信息的编码。信息熵为信源的平均信息量（不确定性的度量）。常见的熵编码有：香农（Shannon）编码、哈夫曼（Huffman）编码和算术编码（Arithmetic Coding）。在视频编码中，熵编码把一系列用来表示视频序列的元素符号转变为一个用来传输或存储的压缩码流。输入的符号可能包括量化后的变换系数、运动向量、头信息（宏块头、图像头、序列的头等）以及附加信息（对于正确解码来说重要的标记位信息）。

例如：555555777773332222111111

行程编码为：(5, 6) (7, 5) (3, 3) (2, 4) (1, 7)。可见，行程编码的位数远远少于原始字符串的位数。

行程编码分为定长和不定长（变长）编码两种。定长编码是指编码的行程长度所用的二进制位数固定，而变长行程编码是指对不同范围的行程长度使用不同位数的二进制位数进行编码。使用变长行程编码需要增加标志位来表明所使用的二进制位数。

行程编码比较适合于二值图像的编码，一般用于量化后出现大量零系数连续的场合，用行程来表示连零码。行程编码对传输差错很敏感，一位符号出错就会改变行程编码的长度，使整个图像出现偏移，因此，一般要用行同步、列同步的方法，把差错控制在一行一列之内。

(2) 哈夫曼编码

哈夫曼编码（Huffman Coding），又称霍夫曼编码，是一种编码方式，哈夫曼编码是可变字长编码（VLC）的一种。Huffman 于 1952 年提出一种编码方法，该方法完全依据字符出现概率来构造异字头的平均长度最短的码字，有时称之为最佳编码，一般就叫做 Huffman 编码。

实际应用中，除采用定时清洗以消除误差扩散和采用缓冲存储以解决速率匹配以外，主要问题是解决小符号集合的统计匹配，例如黑（1）、白（0）传真信源的统计匹配，采用 0 和 1 不同长度游程组成扩大的符号集合信源。游程，指相同码元的长度（如二进制码中连续的一串 0 或一串 1 的长度或个数）。按照 CCITT 标准，需要统计 2×1728 种游程（长度），这样，实现时的存储量太大。事实上长游程的概率很小，故 CCITT 还规定：若 l 表示游程长度，则 $l=64q+r$ 。其中 q 称主码，r 为基码。编码时，不小于 64 的游程长度由主码和基码组成。

而当 l 为64的整数倍时,只用主码的代码,已不存在基码的代码。

哈夫曼编码具有一些明显的特点:

- ① 编出来的码都是异字头码,保证了码的唯一可译性。
- ② 由于编码长度可变,因此译码时间较长,使得哈夫曼编码的压缩与还原相当费时。
- ③ 编码长度不统一,硬件实现有难度。
- ④ 对不同信号源的编码效率不同,当信号源的符号概率为2的负幂次方时,达到100%的编码效率;若信号源符号的概率相等,则编码效率最低。
- ⑤ 由于“0”与“1”的指定是任意的,故由上述过程编出的最佳码不是唯一的,但其平均码长是一样的,故不影响编码效率与数据压缩性能。

(3) 算术编码

算术编码是在20世纪80年代发展起来的一种熵编码方法,其基本原理是将被编码的数据序列表示成0和1之间的一个间隔(也就是一个小数范围),该间隔的位置与输入数据的概率分布有关。信息越长,表示间隔就越小。因而表示这一间隔所需的二进制位数就越多。算术编码有两种模式:一种是基于信源概率统计特性的固定编码模式,另一种是针对未知信源概率模型的自适应模式。

算术编码适合于由相同的重复序列组成的文件,算术编码接近压缩的理论极限。这种方法,是将不同的序列映像到0到1之间的区域内,该区域表示成可变精度(位数)的二进制小数,越不常见的数据要的精度越高(更多的位数),这种方法比较复杂,因而不太常用。

2. 预测编码

预测编码是根据离散信号之间存在着一定关联性的特点,利用前面一个或多个信号预测下一个信号,然后对实际值和预测值的差(预测误差)进行编码。如果预测比较准确,误差就会很小。在同等精度要求的条件下,就可以用比较少的比特进行编码,达到压缩数据的目的。

预测编码中典型的压缩方法有脉冲编码调制(PCM, Pulse Code Modulation)、差分脉冲编码调制(DPCM, Differential Pulse Code Modulation)、自适应差分脉冲编码调制(ADPCM, Adaptive Differential Pulse Code Modulation)等。

(1) 差分脉冲编码调制

在PCM(Pulse Code Modulation, 脉冲编码调制)系统中,原始的模拟信号经过采样后得到的每一个样值都被量化成为数字信号。为了压缩数据,可以不对每一样值都进行量化,而是预测下一样值,并量化实际值与预测值之间的差值,这就是差分脉冲编码调制(DPCM)。1952年,贝尔(Bell)实验室的C.C.Cutler取得了差分脉冲编码调制系统的专利,奠定了真正实用的预测编码系统的基础。

在图像信号中应用DPCM时,用作预测的像素和被预测的像素可以在同一行,也可以在不同行(同一帧),甚至在不同帧,分别称为一维预测、二维预测和三维预测。

DPCM的优点是算法简单,容易硬件实现,缺点是对信道噪声很敏感,会产生误差扩散。即某一位码出错,对图像一维预测来说,将使该像素以后的同一行各个像素都产生误差;而对二维预测,该编码引起的误差还将扩散到以下的各行。这样,将使图像质量大大下降。同时,DPCM的压缩率也比较低。随着变换编码的广泛应用,DPCM的作用已很有限。

(2) 自适应差分脉冲编码调制

进一步改善量化性能或压缩数据率的方法是采用自适应量化或自适应预测,即自适应脉冲编码调制(ADPCM)。它的核心想法是:利用自适应的思想改变量化阶的大小,即使用小

的量化阶去编码小的差值,使用大的量化阶去编码大的差值;使用过去的样本值估算下一个输入样本的预测值,使实际样本值和预测值之间的差值总是最小。

① 自适应量化

在一定量化级数下减少量化误差或在同样的误差条件下压缩数据,根据信号分布不均匀的特点,希望系统具有随输入信号的变化区间足以保持输入量化器的信号基本均匀的能力,这种能力叫自适应量化。

自适应量化必须有对输入信号的幅值进行估值的能力,有了估值才能确定相应的改变量。若估值在信号的输入端进行,称为前馈自适应;若在量化输出端进行,称为反馈自适应。信号的估值必须简单,占用时间短,才能达到实时处理的目的。

② 自适应预测

预测参数的最佳化依赖信源的特征,要得到最佳预测参数显然是一件烦琐的工作,而采用固定的预测参数往往又得不到较好的性能。为了能使性能较佳,又不至于有太大的工作量,可以采用自适应预测。

为了减少计算工作量,预测参数仍采用固定值,但此时有多组预测参数可供选择,这些预测参数根据常见的信源特征求得。编码时具体采用哪组预测参数需根据特征来自适应地确定。为了自适应地选择最佳参数,通常将信源数据分区间编码,编码时自动地选择一组预测参数,使该实际值与预测值的均方误差最小。随着编码区间的不同,预测参数自适应地变化,以达到准最佳预测。

3. 变换编码

变换编码是将空间域里描述的图像,通过某种变换(常用的是二位正交变换,如离散余弦变换、K-L 变换等),映射到另一变换域中,使变换后的系数之间的相关性降低。图像变换本身并不能压缩数据,但变换后图像的大部分能量只集中到少数几个变换系数上,采用适当的量化和熵编码才可以有效地压缩图像。

(1) K-L 变换

K-L 变换是一种最佳正交变换。它是用数据本身的相关矩阵对角化后完成的,这种变换将产生完全不相关的变换系数。如果图像数据之间是高度相关的,经过 K-L 变换后的系数将出现多个零值,同时某些系数的值会很小。

K-L 变换的变换矩阵是由图像数据本身求得的,不同的图像数据有不同的变换矩阵,由此造成反变换矩阵的不唯一性;另外 K-L 变换矩阵的构造计算量很大,因而它不是一种实用的变换方法,通常作为评价其他线性变换的比较基准。

(2) 离散余弦变换

K-L 变换算法复杂度较高,所以在实际编码工作中,人们常用离散余弦变换。对大多数图像信源来说,DCT (Discrete Cosine Transform, 离散余弦变换)是现行编码方法中最接近 K-L 变换的方法。

DCT 是先根据变换系数的能量分布,将整个图像分成 $N \times N$ 像素块,然后对这 $N \times N$ 像素块逐一进行变换。其中变换后幅值较大的图像系数大多集中在图像块的左上角。与其他系数相比,这些低频系数包括了图像的大部分内容,具有的能量最大,在变换图像中的地位最重要,因此应使它们的量化误差最小。另一方面,大多数图像的高频分量较小,对图像质量影响甚微,加上人眼对高频成分的失真不太敏感,因此可以使用更粗的量化,一般采用设定

阈值的方法,置小于阈值的变换系数为零,由此传送变换系数所用的码率要大大小于传送图像像素所用的码率,从而大大提高了编码效率。经区域编码和阈值编码后,变换图像的系数大部分为零,因此必须采用有效的方法将非零系数和零系数组织起来,在带有最少冗余的同时保证最大的连零系数出现概率,在 DCT 图像编码中,可以对变换系数采用 Z 形扫描。

3.3.3 图像编码新技术

1) 分形编码

分形编码是在数学家 Mandelbro 建立的分型几何理论的基础上发展起来的一种编码方法。分形编码最大限度地利用了图像在空间域上的自相似性,通过消除图像的几何冗余来压缩数据。1988 年, M.Barnsley 成功地将迭代函数系统 (IFS) 用于描述图像的自相似性,并将其用于图像编码。

对分形定义的一般描述:

- ① 分形应有精细的结构,有任意小比例的细节。
- ② 它是不规则的,以至其局部和整体都不能用传统的几何语言来描述。
- ③ 分形通常有某种自相似的形式,可能是近似的或是统计的。
- ④ 其“分形维数”一般大于其拓扑维数,并且通常能以非常简单的方法定义,由迭代方法产生。

分形编码的最显著的特点是自相似性 (self-similarity)。与经典方法相比,它不但去除了数据之间局部的相关性,而且去除了整体与局部之间的相关性,所以有望达到经典编码方法所达不到的压缩比,是一种思想全新、很有潜力的编码技术。分形编码的主要特点有:

- (1) 分形编码的图像压缩比经典编码方法的压缩比高出许多。
- (2) 由于分形编码可把图像划分成大得多、形状复杂得多的区分,故压缩所得的文件的大小不会随着图像像素数目的增加即分辨率的提高而变大。而且,分形压缩还能依据压缩时确定的分形模型给出高分辨率的清晰的边缘线,而不是将其作为高频分量加以抑制。
- (3) 分形压缩和解压缩不对称,压缩速度较慢,而解压缩速度很快。这是由于对每块确定仿射变换时,要对整幅图像进行相似性搜索,因而速度较慢。而恢复时只需简单的反复迭代过程,因而速度较快。

2) 模型编码

基于模型的图像编码技术是近几年发展起来的一种很有前景的低比特率编码方法。它利用了计算机视觉和计算机图形学中的方法和理论,其基本出发点是在编、解码两端分别建立起相同的模型,针对输入的图像提取模型参数,或根据模型参数重建图像。模型编码方法的核心是建模和提取模型参数,其中模型的选取、描述和建立是决定模型编码质量的关键因素。为了对图像数据建模,一般要求对输入图像具有某些先验知识。

根据使用的模型的不同,模型编码可以分为语义基编码和物体基编码。

基于模型的图像编码方法是利用先验模型来抽取图像中的主要信息,并以模型参数的形式表示它们,因此可以获得很高的压缩比。然而在模型编码方法的研究中还存在很多问题,例如:①模型法需要先验知识,不适合一般的应用;②对不同的应用所建模型是不一样的;③在线框模型中控制点的个数不易确定,还未找到有效的方法能根据图像内容来选取;④由于利用模型法压缩后复原图像的大部分是用图形学的方法产生的,因此看起来不够自然;⑤传统的误差评估准则不适合用于对模型编码的评价。

3) 小波变换编码

小波变换编码是随着小波变换理论的研究而提出的一种编码方式。小波变换的本质是多分辨率或多尺度地分析信号,非常适合视觉系统对频率感知的对数特性,因此,它很适合于图像信号的处理。

小波变换编码一方面具有传统编码方法的一些优点,能够很好地消除统计冗余,另一方面它的多分辨率特性提供了利用人眼视觉特性的很好机制,而且变换后的图像数据能够保持原图像在各种分辨率下的精细结构,为进一步去除其他形式的冗余提供了便利。

小波变换编码的核心问题是要对子带图像进行小波分解系数的量化和编码。低频子带图像包含原图像的大部分能量,即包含图像的基本特性。它在图像重构算法中起主导作用,对重建图像的质量有很大影响,因此这部分信号应精确保留。

图像信号经过小波多分辨率分解之后成为若干不同频带的信号,这些频带信号具有不同的特点,因此为压缩编码提供了很好的依据。

(1) DCT+矢量量化的压缩

经过小波多分辨率分解之后,所得到的低频部分仍然保持原图像的概貌,因此可以使用 JPEG 的 ADCT 编码对它进行压缩。对于高频部分,可以发现三个方向的子带各自包含了原图像中在水平、垂直和对角线方向的高频分量,因而其中的大部分区域变化幅度较低,且能量较低,而能量高、变化大的区域集中在图像内物体的边缘部分,它保持着物体的结构特征。因此可以根据高频带内的这些分布特点采用高效的矢量量化方法进行压缩。

(2) 嵌入式零树编码 EZW

EZW 算法是由 J.M.Shapiro 提出的小波变换图像压缩编码方法。它主要包括 4 个步骤:

① 小波多分辨率分解

利用小波多分辨率分解,将原图像变换为一个低频带 LL_n 和 $3n$ 个高频带部分 LH_j 、 HL_j 和 HH_j , $j=n, \dots, 2, 1$, n 为分解的级数。

② 零树编码

根据图像分解后系数在不同尺度之间的自相似性预测无效系数出现的位置。

首先用四叉树建立在不同分辨率下子带系数的父子关系。如果从某个系数起,它的后代全为零,并且该系数的父节点非零,则该系数称为一个零树根,它和它的所有后代节点构成一个零树。

由于高频部分对应于图像中边缘部分,因此在不同尺度下系数的分布在空间上存在对应关系,即自相似性。根据这一性质,对变换系数量化后,出现零树的可能性很大,对所有子带进行扫描时,可以分辨 4 类不同的系数:零树根、正值、负值和孤立零值。于是可以有效地记录所有非零系数的位置,即所谓有效图,这就是所谓的零树编码得以压缩的主要原因。

③ 比特面逐步近似量化

为了得到嵌入式码流,使用类似于比特面的编码传输,按 2 的幂次递减的次序对各子带的变换系数进行扫描量化。根据作用的不同,将扫描分为主扫描和从扫描,其中主扫描用以搜索量化步长改变后出现的新的零树根,从扫描则用于对已经找到的非零系数(有效系数)进行更细一层量化修正。当从扫描完成后将量化步长进行降幂再进行下一轮主扫描,依次反复进行,最后达到所希望的量化水平。

④ 自适应算术编码

为了提高压缩效率,采用自适应算术编码对扫描输出符号进行编码。

由于 EZW 算法具有良好的编码性能,因而受到人们的广泛关注,先后出现了许多改进的算法,例如 SPIHT 算法改进了逐渐传输方案,使得即使不采用算术编码也能达到与 EZW 相当的压缩效果;此外还有一些利用彩色分量的相关性进行 EZW 压缩以及对任意形状的对象进行 EZW 压缩的算法。EZW 算法的主要思想已经被 MPEG-4 的静止图像编码部分以及 JPEG2000 等国际标准采用。由此可见其重要地位。

从现在的研究结果可看到,小波变换编码已获得了较好的编码效果,是现代图像压缩技术研究的热点之一,也是十分有前途的一种方法。

3.3.4 图像编码评价

随着众多图像压缩算法的出现,如何评价图像压缩算法就成为重要的课题。一般来说,评价图像压缩算法的优劣有以下 4 个参数。

(1) 算法的编码效率

算法的编码效率通常有几种表现形式:平均码字长度 (R),图像的压缩比 (rate, r),每秒钟所需的传输比特数 (bps),图像熵与平均码长之比 (η)。

(2) 编码图像的质量

图像质量评价可分为客观质量评价和主观质量评价。最常用的客观质量评价指标是均方误差 (MSE) 和峰值信噪比 (PSNR)。主观质量评价是指由一批观察者对编码图像进行观察并打分,然后综合所有人的批评结果,给出图像的质量评价。客观质量评价能够快速有效地评价编码图像的质量,但符合客观质量评价指标的图像不一定具有较好的主观质量。主观质量能够与人的视觉效果相匹配,但其评判过程缓慢费时。

(3) 算法的适用范围

特定的图像编码算法具有其相应的适用范围,并不是对所有图像都有效。一般来说,大多数基于图像信息统计特性的压缩算法具有较广的适用范围,而一些特定的编码算法的适用范围较窄,如分形编码主要用于自相似性高的图像。

(4) 算法的复杂度

算法的复杂度指完成图像压缩和解压缩所需的运算量和硬件实现该算法的难易程度。优秀的压缩算法要求有较高的压缩比,压缩和解压缩速度快,算法简单,易于硬件实现,还要求解压缩后的图像质量较好。选用编码方法时一定要考虑图像信源本身的统计特性、多媒体系统的适应能力、应用环境以及技术标准。

3.4 图像存储

数字图像有多种存储格式,每种格式一般由不同的软件所支持。随着信息技术的发展和图像应用领域的不断拓宽,还会出现新的图像格式。因此,要进行图像存储与处理,必须了解图像文件的存储格式,即图像文件的数据结构。

每一种图像文件均有一个头文件,在头文件之后才是图像数据。文件头的内容由制作该图像文件的公司决定,一般包括文件类型、文件制作者、制作时间、版本号、文件大小等内容。各种图像文件的制作还涉及图像文件的压缩方式和存储效率。下面介绍几种常见的图像存储格式。

3.4.1 BMP 图像存储格式

BMP (BitMap Picture) 文件格式是 Windows 系统交换图形、图像数据的一种标准格式, 已成为 PC 上 Windows 系统中事实上的工业标准, 有压缩和非压缩两种形式。它以独立于设备的方法描述位图, 可用非压缩格式存储图像数据, 解码速度快, 支持多种图像的存储, 常见的各种 PC 图形图像软件都能对其进行处理。该格式支持 1~24 位颜色深度, 使用的颜色模式可为 RGB、索引颜色、灰度和位图等, 与设备无关。

3.4.2 TIFF 图像存储格式

标记图像文件格式 TIFF 是由 Aldus 为 Macintosh 机开发的一种图像文件格式, 最早流行于 Macintosh, 现在 Windows 上主流的图像应用程序都支持该格式。目前, 它是 Macintosh 和 PC 上使用最广泛的位图格式, 在这两种硬件平台上移植 TIFF 图像十分便捷, 大多数扫描仪也都可以输出 TIFF 格式的图像文件。该格式支持的色彩数最高可达 16M (M 在英文中经常被用来表示数量单位, 作 Million (100 万) 讲) 种。其特点是: 存储的图像质量高, 但占用的存储空间也非常大, 其大小是相应 GIF 图像的 3 倍, JPEG 图像的 10 倍; 表现图像细微层次的信息较多, 有利于原稿阶调与色彩的复制。该格式有压缩和非压缩两种形式, 其中压缩形式使用的是 LZW (Lempel-Ziv-Welch) 无损压缩方案。

3.4.3 GIF 图像存储格式

GIF 图像存储格式由 CompuServer 公司创建, 存储色彩最高只能达到 256 种, 仅支持 8 位图像文件。在颜色深度和图像大小上, GIF 类似于 PCX; 在结构上, GIF 类似于 TIFF。该格式采取 LZW 压缩算法, 存储效率高, 正因为它是经过压缩的图像文件格式, 所以大多用在网络传输上和 Internet 的 HTML 网页文档中, 速度要比传输其他图像文件格式快得多。它的最大缺点是最多只能处理 256 种色彩, 故不能用于存储真彩色的图像文件, 但其 GIF89a 格式能够存储成背景透明的形式, 并且可以将数张图存成一个文件, 从而形成动画效果。

3.4.4 PCX 图像存储格式

PCX 图像格式最早是由 ZSoft 公司的 PC Paintbrush 图形软件所支持的一种经过压缩的 PC 位图文件格式。后来, Microsoft 将 PC Paintbrush 移植到 Windows 环境中, PCX 图像格式也就得到了更多的图形图像处理软件的支持。该格式支持的颜色数从最早的 16 色发展到目前的 1677 万色。它采用行程编码方案进行压缩, 带有一个 128 字节的文件头。该格式比较简单, 适合保存索引和线画稿模式图像。其不足之处是它只有一个颜色通道。PC 格式支持 1~24 位格式颜色深度以及 RGB、索引颜色、灰度和位图颜色模式。

3.4.5 JPEG 图像存储格式

JPEG (Joint Photographer's Experts Group, 联合图像专家组) 格式是由 ISO 和 CCITT 为静态图像所建立的第一个国际数字图像压缩标准, 主要是为了解决专业摄影师所遇到的图像信息过于庞大的问题。由于 JPEG 的高压缩比和良好的图像质量, 使得它广泛应用于多媒体和网络程序中。JPEG 和 GIF 成为 HTML 语法选用的图像格式。

JPEG 格式支持 24 位颜色, 并保留照片和其他连续色调图像中存在的亮度和色相显著和细微变化。

JPEG 一般基于 DCT 变换的顺序型模式压缩图像。JPEG 通过有选择地减少数据来压缩文件大小。因为它会弃用数据，故 JPEG 压缩为有损压缩。较高品质设置导致弃用的数据较少，但是 JPEG 压缩方法会降低图像中细节的清晰度，尤其是包含文字或矢量图形的图像。

思考题

1. 衡量数码相机的性能指标主要有哪些？
2. 简述光敏传感器组成结构、原理和主要应用。
3. 简述图像编码技术的发展和代表方法的原理。
4. 图像编码技术如何评价？
5. 数字图像存储形式有哪些？各有什么特点？