



中国地质大学(武汉)本科教学工程项目资助
中国地质大学(武汉)实验技术研究项目资助
中国地质大学(武汉)实验教学系列教材

多媒体信息系统

DUOMEITI XINXI XITONG

周 林 ◎主 编

吴 亮
贾孝良 ◎副主编



中国地质大学出版社
ZHONGGUO DIZHI DAXUE CHUBANSHE



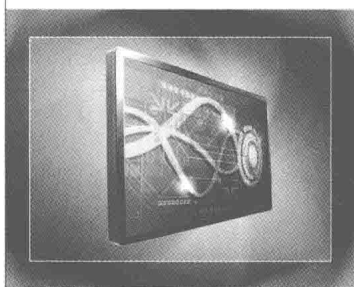
中国地质大学(武汉)本科教学工程项目资助
中国地质大学(武汉)实验技术研究项目资助
中国地质大学(武汉)实验教学系列教材

多媒体信息系统

DUOMEITI XINXI XITONG

周 林 主 编

吴 亮 副主编
贾孝良



中国地质大学出版社
ZHONGGUO DIZHI DAXUE CHUBANSHE

图书在版编目(CIP)数据

多媒体信息系统/周林主编. —武汉:中国地质大学出版社,2016.12
中国地质大学(武汉)实验教学系列教材
ISBN 978-7-5625-3977-3

- I. ①多…
- II. ①周…
- III. ①多媒体技术-高等学校-教材
- IV. ①P37

中国版本图书馆 CIP 数据核字(2016)第 303081 号

多媒体信息系统

责任编辑:彭琳

出版发行:中国地质大学出版社(武汉市洪山区鲁磨路 388 号)
电 话:(027)67883511 传 真:(027)67883580
经 销:全国新华书店

开本:787 毫米×1 092 毫米 1/16
版次:2016 年 12 月第 1 版
印刷:武汉市教文印刷厂

主 编 周 林
副主编 吴 亮 贾孝良

责任校对:周旭

邮 编:430074

E-mail: cbb@cug.edu.cn

Http://www.cugp.cug.edu.cn

字数:371 千字 印张:14.5
印次:2016 年 12 月第 1 次印刷
印数:1—500 册

ISBN 978-7-5625-3977-3

定价:38.00 元

如有印装质量问题请与印刷厂联系调换

中国地质大学(武汉)实验教学系列教材

编委会名单

主 任:唐辉明

副主任:徐四平 殷坤龙

编委会成员:(以姓氏笔画排序)

公衍生 祁士华 毕克成 李鹏飞

李振华 刘仁义 吴 立 吴 柯

杨 喆 张 志 罗勋鹤 罗忠文

金 星 姚光庆 饶建华 章军锋

梁 志 董元兴 程永进 蓝 翔

选题策划:

毕克成 蓝 翔 张晓红 赵颖弘 王凤林

前 言

在数字全球化的影响下,数字产品应用已经广泛影响着我们的生活,传统的纸质印刷媒体在多个领域已经逐渐被数字多媒体所替代。与此同时,数字多媒体的传播与应用需求使得多媒体信息系统应运而生。多媒体信息系统不仅可以实现多媒体资源的组合播放,提供人性化的工作环境和轻松惬意的娱乐休息空间,而且是一条行之有效的宣传途径,相对以往的宣传手段,它能实现信息发布的统一化管理,信息传递快捷、方便,显示效果多样灵活,富有更强的视觉冲击力和印象感染力。借助日益普及的高速信息网,多媒体技术及多媒体信息系统正广泛应用于咨询服务、图书、教育、通信、军事、金融、医疗等诸多行业,并正潜移默化地改变着我们生活的面貌。多媒体技术应用是当今信息技术领域发展最快、最活跃的技术,是新一代电子技术发展和竞争的焦点。因此,在本科教育中广泛开展多媒体信息系统及相关技术的教学非常必要。

笔者首先介绍了多媒体信息系统中涉及的基本概念及基础知识,然后介绍了常见媒体的处理技术与基于内容的检索技术,最后介绍了多媒体同步通信技术以及典型多媒体信息系统实例。全书共分为7章:第一章主要讲述多媒体信息系统相关的一系列概念和音频、图像、视频、颜色模型的基础知识;第二章主要讲述数字图像处理的相关技术,包括图像增强、图像复原、图像分割、形态学图像处理、图像压缩与编码等内容;第三章讲述数字音频编码技术及标准,主要介绍了数字音频压缩编码的基础知识、无损音频编码技术和感知音频编码技术;第四章讲述视频压缩编码技术及标准,主要介绍了视频编码的基础知识、H.264/AVC视频编码技术、HEVC视频编码技术和AVS视频编码技术;第五章主要介绍了媒体图像、音频和视频的基于内容的信息检索技术;第六章主要介绍了多媒体同步的基本知识,并详细介绍了一个音视频同步传输的实现示例;第七章主要介绍了3类常见的多媒体信息系统,即视频会议系统、移动教育系统和移动医疗系统。

笔者近年一直从事本课程的本科教学工作,积累了一定的教学资料和教学经验。笔者采取了由浅入深、循序渐进的方式向读者讲述多媒体信息系统的相关内

容,旨在让初学者对多媒体技术及多媒体信息系统有一个初步了解,为后续的学习及研究打下基础。本书可被定位为一本入门级的书籍,也可作为各院校本科生相关课程的教材或初学者的参考用书。

本书的出版得到了“中国地质大学(武汉)实验技术研究经费”的资助。在本书的筹备过程中,李奔、王淑锦、王雪柔等同学参与了资料的搜集,笔者所在院系的领导及同事给予了大量的关心与帮助。在此,对以上各方面的支持与帮助表示衷心的感谢。

由于篇幅所限,全书内容只涵盖了多媒体信息系统的一小部分基础知识,所涉及内容也难免存在分析不够深入或介绍不够详细的问题。另外,由于笔者的水平有限,书中难免有错误与不足之处,欢迎广大读者及专家批评指正。

周 林

2016年11月30日于武汉

5

目 录

第一章 多媒体信息系统基础	(1)
第一节 基本概念	(1)
第二节 音频基础知识	(3)
第三节 颜色模型	(6)
第四节 图像基础知识	(10)
第五节 视频基础知识	(13)
本章小结	(15)
习 题	(16)
第二章 数字图像处理技术	(17)
第一节 图像增强	(17)
第二节 图像复原	(33)
第三节 形态学图像处理	(39)
第四节 图像分割	(51)
第五节 图像压缩与编码	(58)
本章小结	(70)
习 题	(71)
第三章 数字音频编码技术及标准	(72)
第一节 数字音频压缩编码概述	(72)
第二节 无损音频编码技术	(78)
第三节 感知音频编码技术	(92)
本章小结	(98)
习 题	(98)
第四章 视频压缩编码技术及标准	(99)
第一节 视频压缩编码基础	(99)
第二节 H. 264/AVC 视频编码技术	(102)
第三节 HEVC 视频编码技术	(112)
第四节 AVS 视频编码技术	(120)
本章小结	(129)

习 题	(130)
第五章 基于内容的多媒体信息检索技术	(131)
第一节 多媒体信息检索概述	(131)
第二节 基于内容的图像检索	(133)
第三节 基于内容的音频检索	(142)
第四节 基于内容的视频检索	(155)
本章小结	(166)
习 题	(166)
第六章 多媒体同步技术	(167)
第一节 多媒体同步概述	(167)
第二节 多媒体同步模型	(171)
第三节 多媒体同步控制机制	(175)
第四节 同步性能的测量标准	(178)
第五节 音频/视频同步传输的实现	(179)
本章小结	(193)
习 题	(193)
第七章 多媒体信息系统实例	(194)
第一节 高清视频会议系统	(194)
第二节 移动教育系统	(205)
第三节 基于云计算的移动医疗系统	(210)
本章小结	(220)
习 题	(220)
主要参考文献	(221)

第一章 多媒体信息系统基础

第一节 基本概念

一、媒体的概念及分类

按照传统的说法,媒体(media)是指信息的载体,如报纸、杂志、电视、广播、因特网等。在计算机领域有两层含义:一是信息的存储实体,如磁盘、光盘、磁带、半导体存储器等,中文常译为媒质或媒介;二是传递信息的基本元素,如文字、声音、图形、动画和图像等。国际电话电报咨询委员会(CCITT)把媒体分成5类。

1. 感知媒体(perception medium)

感知媒体是指能直接作用于人的感觉器官,使人直接产生感觉(视、听、嗅、味、触觉)的一类媒体,如语言、音乐、图形、动画、数据、文字、文件等。

2. 表示媒体(representation medium)

表示媒体是为了更有效地加工、处理及传输感知媒体而人为研究和构造出来的一种媒体。它包括上述感知媒体的各种编码,如语音编码、静止和活动图像编码,以及文本编码等。

3. 呈现媒体(presentation medium)

呈现媒体是感知媒体与用于通信的电信号之间在转换中用于信息输入和输出的媒体。可分为输入呈现媒体(如键盘、摄像机、话筒、扫描仪等)和输出呈现媒体(如显示器、扬声器、打印机等)两种。

4. 存储媒体(storage medium)

存储媒体用于存储表示媒体,即存放感知媒体数字化产品(结果)的媒体,如磁盘、光盘、磁带、半导体存储器等。

5. 传输媒体(transmission medium)

传输媒体是指用来将表示媒体从一处传递到另一处的物理传输介质,如同轴电缆、双绞线、光纤及其他通信信道。

二、多媒体概述

多媒体(multimedia),一般理解为多种媒体的综合,是指由两种或两种以上媒体组合成的信息交互和传播的形式。该定义中的媒体是一个宽泛的概念,囊括了以上五大分类。

多媒体技术是指利用计算机技术把文本、图像、图形、动画、音频及视频等多种媒体有机地集成起来,使人们能以更加自然的方式使用信息,并与计算机进行交互,且使表现的信息图、文、声并茂的技术。简言之,多媒体技术就是计算机综合处理声、文、图信息的技术,常被简称为多媒体。

多媒体信息是指以两种或两种以上媒体表达同一信息内容的一种综合信息类型。该定义中的媒体元素主要包括文本、图形、图像、音频、视频、动画等。

1. 文本(text)

“文本”来自英文 text,又译作本文、正文、语篇和课文等。从词源上看,它表示编织的东西。这与中国“文”的概念颇有类似之处。《说文解字·序》中称:“仓颉之初作书,盖依类象形,故谓之文。……文者,物象之本也。”可见,鸟兽身上的花纹彩羽,以及物体的形状、线条、色彩即为“文”。但文本的概念后来变成“任何由书写所固定下来的话语”。文本可为一个单句,也可由一系列句子组成。在计算机中,文本主要指用于记载和储存文字、数字、字母和符号信息,而不含图像、声音等信息的媒体形式。常见的文本类型有 txt、rtf、doc、docx、wps 等。

2. 图形(graphics)

图形又称矢量图,它用一组指令集合来描述事物的颜色、形状、轮廓、大小、位置等属性,常见的图形有直线、圆、矩形、曲线、图表等。每个图形都是一个独立的实体,这使得图形可以自由无限地重新组合。图形主要具有无级缩放、不产生锯齿、高分辨率输出等特点。图形文件的格式较多,常见的有 ai、cdr、dxf、emf、eps、svg 等。

3. 图像(image)

这里主要指静态图像,以二维整数数组形式表示,如照片、画片等。数字图像可由许多不同的输入设备和技术生成,例如数码相机、扫描仪、坐标测量机等。数字图像经常采用一定的压缩格式进行传输和储存。目前,比较常见的图像格式有 BMP、GIF、JPEG、PNG 等。

4. 音频(audio)

音频是人类能够听到的所有声音的统称,包括语音、音乐、噪声等。声音在空气中以一种连续波的形式传播,故又称声波。当声波的频率在 20Hz~20kHz 之间时,可被人耳感知即为音频。常见的音频格式有 cda、wav、mp3、wma、vqf 等。

5. 视频(video)

视频泛指将一系列静态图像以电信号的方式加以捕捉、记录、处理、储存、传送与重现的各种技术。由于人眼看到一幅画或一个物体后,在 0.34s 内不会消失,根据这一视觉暂留原理,当连续的图像变化每秒超过 24 帧(frame)画面时,人眼就无法辨别单幅的静态画面,看上去是

平滑连续的视觉效果,这样连续的画面叫作视频。视频技术最早是为了电视系统而发展,但现在已经发展为各种不同的格式,如mpeg、mov、avi、rm、asf等。

6. 动画(animation)

动画是通过把人物的表情、动作、变化等分解成许多动作瞬间的画幅,再用摄影机连续拍摄成一系列画面,给视觉造成连续变化的图画。它的基本原理与电影、电视一样,都是视觉暂留原理。动画与视频不同之处在于:动画的每帧图像是计算机产生出来或人工绘制的,而视频的每帧图像则是通过实时摄取自然景象或者活动对象获得的。

多媒体信息系统是指在各类软硬件的支持下,能完成对多媒体信息进行收集、存储、检索、管理、处理、传输、表现等任务的技术系统。当前研发的大多数软硬件系统都属于此列,如视频会议系统、地理信息系统软件、移动图书馆系统等。

第二节 音频基础知识

声音是机械振动或气流振动引起周围传播媒质(气体、液体、固体等)发生波动的现象,通常将产生声音的发声体称为声源。当声源产生振动时,引起邻近空气的振动,这样空气就随着声源体所振动幅度的不同,而产生密或稀的振动,空气的这种振动被称为声波。声波可以用一条连续的曲线来表示,如图1-1所示。声波具有普通波所具有的特性,例如反射、折射和衍射等。

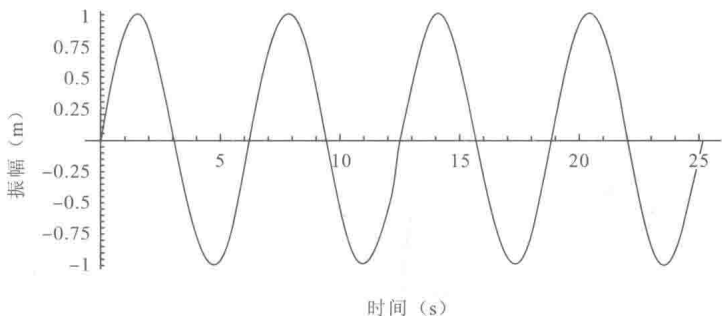


图 1-1 声音是一种连续的波

一、声音的基本特征

1. 频率

频率是单位时间内声波信号振动的次数,单位是赫兹(Hz)。声波信号按频率可以分为:次声波($<20\text{Hz}$)、音频信号($20\text{Hz}\sim 20\text{kHz}$)和超声波($>20\text{kHz}$)。

2. 频谱

声音信号按频率成分可分为:纯音、复音、基音和谐音。纯音是频率单一、振幅随时间按正

弦函数规律变化的声音,复音是由不同频率和不同振幅的声波组合而成的声音,基音是复音中频率最低的成分(分音),谐音是复音中频率与基音成整数倍的分音。声音的频谱结构是用基音、谐音数目、各谐音振幅大小及相位关系来描述的。

3. 声压及声压级

由声波引起的交变压强称为声压,一般用 p 表示,单位是帕(Pa)。声压的大小反映了声音振动的强弱,同时也决定了声波的幅度大小。用声压的相对大小(称声压级或声强)来表示声压的强弱。声压级用符号 SPL 表示,单位是分贝(dB):

$$\text{SPL} = 20 \times \lg \frac{p}{p_{\text{ref}}} \quad (1-1)$$

式中: p 为声压有效值; p_{ref} 为参考声压,一般取 2×10^{-5} Pa,这个数值是人耳所能听到的 1kHz 声音的最低声压,低于这一声压,人耳就无法觉察出声波的存在了。

4. 响度

在客观的度量中,声音的强弱是由声波的振幅(声压)决定的,但响度与声波的振幅并不完全一致。响度不仅取决于振幅的大小,还取决于频率的高低以及与声源的距离。响度用符号 N 表示,单位是宋(sone)。国际上规定,1 宋(sone)是指频率 1kHz 的纯音在声压级 40dB 时的响度。

$$N = 2^{0.1 \times (\text{SPL} - 40)} \quad (1-2)$$

当声音减弱到人耳刚刚可以听见时,此时的声音强度称为最小可听阈值,简称为听阈或闻阈。一般以 1kHz 纯音为准进行测量,人耳刚能听到的声压级为 0dB(通常大于 0.3dB 即有感受)。当声音增强到使人耳感到疼痛时,这个听觉阈值称为痛阈。仍以 1kHz 纯音为准来进行测量,使人耳感到疼痛时的声压级为 130~140dB。

5. 音调

人耳对声音高低的感受称为音调。音调主要与声波的频率有关。声波的频率高,则音调也高。当分别敲击一个小鼓和一个大鼓时,会感受它们所发出的声音不同。小鼓被敲击后振动频率快,发出的声音比较清脆,即音调较高;而大鼓被敲击后振动频率较慢,发出的声音比较低沉,即音调较低。频率越高,振动越快,声音越尖。

6. 音色

音色是人们区别具有同样响度、同样音调的两个声音之所以不同的特性,或者说是人耳对各种频率、各种强度的声音的综合反应。不同材质、形状的发音源,泛音的排列组合也不同,也就构成了这一物体特殊的音色。音色是一个主观量,常用柔和、刺耳、饱满等词描述。

二、音频数字化基础

音频数字化是指通过采样、量化把模拟声波信号转换成由二进制数组成的数字音频文件的过程。采样和量化过程所用的主要硬件是由模拟到数字的转换器(A/D 转换器),在数字音频回放时,再由数字到模拟的转换器(D/A 转换器)将数字声音信号转换成原始的电信号。

1. 采样(sampling)

采样是指每隔一定的时间间隔,抽取信号的一个瞬时幅度值(样本值),即在时间上将模拟信号进行离散化。采样后所得的一系列在时间上离散的样本值称为样值序列。单位时间内采样的次数称为采样频率。显然采样频率越高,所得到的离散幅值的数据点就越逼近于连续的模拟音频信号曲线,同时采样的数据量也越大。

为了保证数字化的音频能够准确(可逆)地还原成模拟音频进行输出,奈奎斯特(Nyquist)采样定理要求:采样频率必须大于或等于模拟信号频谱中最高频率的 2 倍。

常用的音频采样率有 8kHz、11.025kHz、22.05kHz、16kHz、37.8kHz、44.1kHz、48kHz 等。例如:话音信号频率在 0.3~3.4kHz 范围内,用 8kHz 的抽样频率,就可获得能取代原来连续话音信号的抽样信号,而一般 CD(代表小型镭射盘,是一个用于所有 CD 媒体格式的一般术语)的采样频率为 44.1kHz。

2. 量化(quantization)

量化,就是把采样得到的声音信号幅度转换成数字值,用于表示信号强度。

量化精度:用多少个二进位来表示每一个采样值,也称为量化位数。声音信号的量化位数一般是 4bits、6bits、8bits、12bits 或 16bits。

3. 声道数

声音通道的个数称为声道数,是指一次采样所记录产生的声音波形个数。单声道信号一次产生一组声波数据。如果一次产生两组声波数据,则称其为双声道或立体声。随着声道数的增加,所占用的存储容量将成倍增长。数据量计算公式为:

数据量(MB)=[采样频率(kHz)×量化精度×声道数×时间(s)]/[8(B/s)×1.024]

例如:1min 声音,单声道、8 位采样位、采样频率为 11.025kHz,数据量为 0.66MB/min;若采样频率为 22.05kHz,则数据量为 1.32MB/min;若是立体声,则为 2.6MB/min。

4. 声音质量与数据率

根据声音的频带,通常把声音的质量分成 5 个等级,由低到高分别是数字电话、调幅广播(AM)、调频广播(FM)、CD 和数字录音带(DAT)的声音。在这 5 个等级中,使用的采样频率、样本精度、通道数和数据率如表 1-1 所示。

表 1-1 声音质量和数据率

质量	采样频率 (kHz)	样本精度 (B/s)	单道声 (立体声)	数据率(未压缩) (kB/s)	频率范围 (Hz)
电话	8	8	单声道	64.0	200~3 400
AM	11.025	8	单声道	88.2	50~7 000
FM	22.050	16	立体声	705.6	20~15 000
CD	44.1	16	立体声	1 411.2	20~20 000
DAT	48	16	立体声	1 536.0	20~20 000

备注:电话使用 μ 率编码,动态范围为 13 位,而不是 8 位。

第三节 颜色模型

国际照明委员会(International Commission on Illumination,简称 CIE)对颜色的描述作了一个通用的定义,用颜色的 3 个特性来区分颜色。这些特性是色调、饱和度和明度,它们是颜色所固有的并且是截然不同的特性。

色调(hue)又称为色相,指颜色的外观,用于区别颜色的名称或颜色的种类。色调用红、橙、黄、绿、青、蓝、靛、紫等术语来刻画。用于描述感知色调的一个术语是色彩(colorfulness)。

饱和度(saturation)是相对于明度的一个区域的色彩,是指颜色的纯洁性,它可用来区别颜色明暗的程度。完全饱和的颜色是指没有渗入白光所呈现的颜色,例如仅由单一波长组成的光谱色就是完全饱和的颜色。

明度(brightness)是视觉系统对可见物体辐射或者发光多少的感知属性。它和人的感知有关。由于明度很难度量,因此国际照明委员会定义了一个比较容易度量的物理量来度量明度,称为亮度(luminance),亮度即辐射的能量。明度的一个极端是黑色(没有光),另一个极端是白色,在这两个极端之间是灰色。

光亮度(lightness)是人的视觉系统对亮度的感知响应值,光亮度可用作颜色空间的一个维,而明度(brightness)则仅限用于发光体,该术语用来描述反射表面或者透射表面。

颜色模型是表示颜色的一种数学方法,人们用它来指定和产生颜色,使颜色形象化。颜色模型选用一些参数来描述颜色在颜色模型中的位置,但并没有说明是什么颜色,其颜色取决于采用的坐标。从技术角度上区分,颜色模型可分为 3 类:①计算机图形颜色模型,这类模型主要用于计算机、摄像机等的颜色显示系统,包括 RGB、HSI、HSL 和 HSV 等颜色模型;②CIE 颜色模型,这类颜色模型是由国际照明委员会定义的,通常作为国际性的颜色模型标准,用作颜色的基本度量方法,包括 CIE 1931 XYZ、Lab、Luv 和 LCH 等可作为转换过渡的颜色模型;③电视系统颜色模型,由广播电视需求推动而开发的颜色模型,主要目的是通过压缩色度信息以有效地播送彩色电视图像,包括 YUV、YIQ、YCbCr 等颜色模型。

以下详细介绍几种常见的颜色模型。

一、RGB 颜色模型

在 RGB 模型中,颜色空间里所有的颜色都是由 R、G、B(红、绿、蓝)3 种光依不同的比例相加而成。如在 24 位图像中,每一种颜色成分都由 0~255 之间的数值表示。RGB 中的每一色光均含有亮度成分,例如 R 的成分越多,表示越红越亮。各色光混合后,会比原来单独的色光还亮,称为相加混色。当红色、蓝色和绿色的光以最大强度组合在一起时,眼睛看到的颜色就是白色。当每一种颜色成分的值为 0 时,即表示没有任何颜色的光,因此,眼睛看到的颜色就为黑色。

RGB 颜色模型适用于以主动光源显示影像的场合,如电视、电脑、投影仪等。

二、CMY/CMYK 颜色模型

CMY/CMYK 颜色模型主要用于彩色印刷、彩色胶片和绘画中,属于相减混色。当颜色

(即油墨)被添加到一种表面如白纸时,颜色会“减少”表面的亮度。当每一种颜色成分(C、M、Y)的值都最大时,所得到的颜色即为黑色。当每一种颜色成分的值都为0时,即表示表面没有添加任何颜色,因此,表面本身就会显露出来——在这个例子中白纸就会显露出来。

从理论上来看,任何一种由颜料呈现的颜色都可以用青(cyan,简写为C)、品红(magenta,简写为M)和黄(yellow,简写为Y)这3种基色按不同的比例混合而成,这就是CMY颜色模型。但由于彩色墨水和颜料的化学特性,用等量的CMY三基色得到的黑色不是真正的黑色,因此,在印刷术中常加一种真正的黑色墨水(black ink,简写为K),于是CMY颜色模型就变成了CMYK颜色模型。

由于RGB颜色模型比CMYK颜色模型大,因此,简单的代数转换不能获得满意的结果,故在此不讨论两者的转换问题。

三、YUV 颜色模型

YUV是PAL(Phase Alternating Line,逐行倒相)制彩色电视系统所采用的一种颜色模型。其中,Y表示明亮度(luminance或luma),也就是灰阶值;U和V表示的则是色度(chrominance或chroma),作用是描述影像色彩及饱和度,用于指定像素的颜色;U表示蓝色色差(即 $B-Y$),V代表红色色差($R-Y$)。

亮度信号Y和色度信号U、V是分离的。如果只有Y信号分量而没有U、V分量,表示的图为黑白灰度图。这就解决了彩色电视和黑白电视兼容的问题。同时,由于人眼对彩色的敏感度低于对亮度的敏感度的视觉特性,并基于“大面积着色原理”,可用较窄的频带传送U、V信号,优化彩色电视信号的传输。

RGB颜色模型转YUV颜色模型的公式如下:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1-3)$$

YUV颜色模型转RGB颜色模型的公式如下:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.395 & -0.581 \\ 1 & 2.032 & 0 \end{bmatrix} \cdot \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \quad (1-4)$$

四、YIQ 颜色模型

YIQ颜色模型是由YUV颜色模型推导而来,是国家电视系统委员会(National Television System Committee,简称NTSC)制彩色电视系统所采用的一种颜色模型。YIQ系统中,Y分量代表图像的亮度信息,I、Q两个分量则携带颜色信息,I分量代表从橙色到青色的颜色变化,而Q分量则代表从紫色到黄绿色的颜色变化。I、Q是通过将U、V轴逆时针旋转 33° 获得的。

$$I = V \times \cos 33^\circ - U \times \sin 33^\circ \quad (1-5)$$

$$Q = V \times \sin 33^\circ + U \times \cos 33^\circ \quad (1-6)$$

人眼彩色视觉的特性表明,人眼分辨红、黄之间颜色变化的能力最强,而分辨蓝、紫之间颜

色变化的能力最弱。通过一定的变化, I 对应于人眼最敏感的色度, 而 Q 对应于人眼最不敏感的色度。这样, 传送 Q 可以用较窄的频带, 而传送分辨率较强的 I 信号时, 可以用较宽的频带。

RGB 颜色模型转 YIQ 颜色模型的公式如下:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1-7)$$

RGB 颜色模型转 YIQ 颜色模型的公式如下:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.956 & 0.621 \\ 1 & -0.272 & -0.647 \\ 1 & -1.106 & 1.703 \end{bmatrix} \cdot \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (1-8)$$

五、 YC_bC_r 颜色模型

YC_bC_r 颜色模型也是由 YUV 颜色模型派生的一种颜色模型, 其中, Y 是指亮度分量, C_b 指蓝色色度分量, 而 C_r 指红色色度分量。 YC_bC_r 颜色模型主要用于数字电视系统, 在数字多功能光盘(DVD)、摄像机、数字电视等消费类视频产品中, 有时简称为 YCC。

YC_bC_r 不是一种绝对颜色模型, 是 YUV 压缩和偏移的版本。

RGB 颜色模型转 YC_bC_r 颜色模型的公式如下:

$$\begin{bmatrix} Y \\ C_b \\ C_r \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 & 0 \\ -0.1687 & -0.3313 & 0.5000 & 128 \\ 0.5000 & -0.4187 & -0.0813 & 128 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \\ 1 \end{bmatrix} \quad (1-9)$$

RGB 颜色模型转 YIQ 颜色模型的公式如下:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1.40200 & 0 \\ 1 & -0.34414 & -0.71414 \\ 1 & 1.77200 & 0 \end{bmatrix} \cdot \begin{bmatrix} Y \\ C_b - 128 \\ C_r - 128 \end{bmatrix} \quad (1-10)$$

六、HSV 颜色模型

HSV(Hue, Saturation and Value), 由 A. R. Smith 根据颜色的直观特性于 1978 年创建, 也称六角锥体模型(hexcone model)。HSV 的表示方法如下。

色调(H): 用角度度量, $0^\circ \sim 360^\circ$ 。红色为 0° , 按逆时针方向计算; 绿色为 120° ; 蓝色为 240° 。

饱和度(S): 取值范围为 $0.0 \sim 1.0$ 。

亮度值(V): 取值范围为 0.0 (黑色) ~ 1.0 (白色)。

HSV 和 RGB 之间没有转换矩阵, 但可对它们之间的转换算法进行描述。RGB 颜色模型转 HSV 颜色模型的描述如下:

$$\begin{aligned} \max &= \max(R, G, B) \\ \min &= \min(R, G, B) \\ \text{if } R &= \max, H = (G - B) / (\max - \min) \end{aligned}$$

```

if  $G = \max$ ,  $H = 2 + (B - R) / (\max - \min)$ 
if  $B = \max$ ,  $H = 4 + (R - G) / (\max - \min)$ 
 $H = H * 60$ 
if  $H < 0$ ,  $H = H + 360$ 
 $V = \max(R, G, B)$ 
 $S = (\max - \min) / \max$ 

```

HSV 颜色模型转 RGB 颜色模型的描述如下:

```

if  $s = 0$ ,  $R = G = B = V$ 
else  $H \neq 60$ 
 $i = \text{INTEGER}(H)$ 
 $f = H - i$ 
 $a = V * (1 - s)$ 
 $b = V * (1 - s * f)$ 
 $c = V * (1 - s * (1 - f))$ 
switch( $i$ )
case 0:  $R = V$ ;  $G = c$ ;  $B = a$ ;
case 1:  $R = b$ ;  $G = v$ ;  $B = a$ ;
 $\star$  case 2:  $R = a$ ;  $G = v$ ;  $B = c$ ;
case 3:  $R = a$ ;  $G = b$ ;  $B = v$ ;
case 4:  $R = c$ ;  $G = a$ ;  $B = v$ ;
case 5:  $R = v$ ;  $G = a$ ;  $B = b$ ;

```

七、HSI 颜色模型

HSI(Hue, Saturation and Intensity), 是美国色彩学家孟塞尔(H. A. Munseu)于 1915 年提出的, 它反映了人的视觉系统感知彩色的方式, 以色调、饱和度和强度 3 种基本特征量来感知颜色。

色调(H): 与光波的波长有关, 它表示人的感官对不同颜色的感受, 如红色、绿色、蓝色等; 也可表示一定范围的颜色, 如暖色、冷色等。

饱和度(S): 表示颜色的纯度, 纯光谱色是完全饱和的, 加入白光会稀释饱和度。饱和度越大, 颜色看起来就会越鲜艳, 反之亦然。

强度(I): 表示光的强度。

对任何 3 个 $[0, 1]$ 范围内的 R, G, B 值, 其对应 HSI 模型中的 H, I, S 分量的计算公式为:

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \quad (1-11)$$

$$\text{式中: } \theta = \arccos \left\{ \frac{\frac{1}{2}((R-G) + (R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}$$

$$S = 1 - \frac{3}{(R+G+B)} \min(R, G, B) \quad (1-12)$$