

王相海 宋传鸣 著

图像及视频可分级编码



 科学出版社
www.sciencep.com

图像及视频可分级编码

王相海 宋传鸣 著

科学出版社

北 京

内 容 简 介

本书是关于图像及视频可分级编码的一本学术专著,介绍了图像和视频信息的冗余特性、基本编码方法以及图像和视频可分级编码的含义和分类情况,对 DCT 和小波变换的基本理论以及图像和视频的编码技术、图像和视频的可分级编码技术进行了论述,对图像和视频可分级编码的未来发展方向进行了展望。

本书可供计算机应用、通信与电子系统、信号与信息处理等相关专业的研究人员、工程技术人员、高校教师、研究生和高年级本科生学习参考。

图书在版编目(CIP)数据

图像及视频可分级编码 / 王相海,宋传鸣著. —北京:科学出版社,2009
ISBN 978-7-03-023121-5

I. 图… II. ①王…②宋… III. 图像编码 IV. TN919.81

中国版本图书馆 CIP 数据核字(2008)第 152731 号

责任编辑:任 静 魏英杰 / 责任校对:李奕萱

责任印制:赵 博 / 封面设计:嘉华永盛

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2009 年 1 月第 一 版 开本: B5(720×1000)

2009 年 1 月第一次印刷 印张: 22 1/2

印数: 1—3 000 字数: 438 000

定价: 55.00 元

(如有印装质量问题,我社负责调换〈新蕾〉)

本书的研究工作及出版得到如下项目的资助：

国家自然科学基金项目(60372071)

辽宁省自然科学基金项目(20032105,20072156)

辽宁省高等学校优秀人才支持计划项目(RC-04-11)

辽宁省教育厅科学技术研究项目(20060486)

辽宁省高等学校重点学科建设专项项目

江苏省普通高校研究生科研创新计划项目(CX07B-121z)

计算机软件新技术国家重点实验室(南京大学)开放基金项目
(A200506,KFKT2008B14)

视觉与听觉信息处理国家重点实验室(北京大学)开放基金项目

南京邮电大学图像处理与图像通信江苏省重点实验室开放基金项
目(ZK207008)

辽宁省教育厅学术专著出版基金

大连市人民政府学术专著出版基金

辽宁师范大学学术专著出版基金

前 言

随着网络和多媒体技术的发展,人们对图像和视频编码的要求越来越高,不仅要求编码技术具有较好的压缩效果,而且要求它能适应网上渐进传输的需求。在许多实际应用中,由于用户需求不同、终端能力不同、异构网络的不同支路所能提供的 QoS 不同或网络传输条件的变化,比如噪声、拥塞等原因,需要提供不同质量的图像和视频信号,比如视频会议、视频点播、多媒体数据库浏览等。解决此类问题的最好方法是用单个编码器产生分层次的压缩码流,对不同层次的码流解码可以获得不同的图像质量,例如,处理能力低的终端只对码流的一部分进行解码,获得低分辨率的图像,而处理能力高的终端对整个码流进行解码获得高分辨率的图像,这种编码机制通常被称为可分级编码机制。

实际上,图像和视频的可分级编码思想在早期的国际标准 JPEG 和 MPEG-2 中便有所体现。近年来,随着流媒体和无线信道多媒体信息传输的蓬勃发展,图像和视频的可分级编码技术已经成为信息编码技术领域的研究热点,同时被许多现行的国际编码标准,诸如 JPEG-2000 和 MPEG-4 等所接纳并得以不断发展。

本书是作者在所承担的国家自然科学基金项目、辽宁省自然科学基金项目、辽宁省高等学校优秀人才支持计划项目、辽宁省高等学校重点学科建设专项项目、计算机软件新技术国家重点实验室(南京大学)开放基金项目、视觉与听觉信息处理国家重点实验室(北京大学)基金项目、江苏省普通高校研究生科研创新计划项目、南京邮电大学图像处理与图像通信江苏省重点实验室开放基金项目等期间的研究成果,以及第一作者完成的南京大学博士后出站报告的基础上,参考国内外最新的文献撰写而成的一本学术专著。全书共分五章:第一章介绍了图像和视频信息的冗余特性,以及图像和视频编码技术的总体发展情况,同时对目前常见的图像和视频国际编码标准进行了概述,重点介绍了图像和视频可分级编码的含义以及分类情况;第二章介绍了 DCT 和小波变换的基本理论,以及传统的基于 DCT 和小波变换的图像及视频编码技术;第三章着重论述了图像的可分级编码技术,包括国际图像压缩标准中的可分级编码,基于 DCT、小波变换和 MP 的图像可分级编码;第四章重点阐述了视频可分级编码技术,包括目前国际视频压缩标准中的可分级编码,基于 DCT、2D 小波和 3D 小波的视频可分级编码;第五章对图像和视频可分级编码的未来发展方向进行了展望。

目前国内有关多媒体信息编码的论著已出版了不少,但专门论述可分级编码的著作还鲜有见到,希望本书的出版能起到抛砖引玉的作用,对这一领域的研究有

所帮助。

衷心感谢南京大学计算机系的张福炎教授和吉林大学数学系的周蕴时教授,是他们指引作者进入了一个全新的研究领域,同时张福炎老师在百忙中审阅了全稿,并提出了许多宝贵的意见和建议;非常感谢刘丹、毕迎春、肖羽、孙强、刘男等研究生,他们在不同程度上参与了国家自然科学基金项目和其他基金项目的研究,并取得了可喜的成绩,这些成绩为本书增色不少;特别感谢方玲玲、刘丹、刘男、张宏为、肖羽、夏春宇等研究生为本书校对所付出的努力;衷心感谢辽宁师范大学计算机与信息技术学院多媒体信息处理与图形学研究室的各位老师和同学在讨论班上对本书相关内容的热烈讨论。另外,本书在撰写过程中,参阅了大量相关书籍和文献,在此向这些作者致以最诚挚的谢意!

本书的出版得到辽宁省教育厅学术专著出版基金、大连市人民政府学术专著出版基金、辽宁师范大学学术专著出版基金和辽宁师范大学计算机与信息技术学院的资助,在此深表谢意!

由于图像和视频的可分级编码技术是一个较新的研究领域,限于作者的水平,书中不妥或错误之处在所难免,敬请读者指正。

作 者

2008年7月于大连

目 录

前言

第一章 图像和视频编码技术概述	1
1.1 图像和视频信息的冗余特性	1
1.1.1 统计冗余	1
1.1.2 心理可视冗余	4
1.2 信源编码的信息论结果	6
1.2.1 信息熵	7
1.2.2 Shannon 的两个信源编码定理	7
1.3 图像的变换编码	9
1.3.1 图像的变换编码特性	9
1.3.2 图像的正交变换	9
1.4 量化	10
1.4.1 均匀量化	11
1.4.2 非均匀量化	12
1.4.3 矢量量化	13
1.5 运动估计与预测	14
1.6 静态图像编码技术的发展	15
1.6.1 基于波形的编码技术	15
1.6.2 第二代编码技术	17
1.6.3 基于分形的编码技术	18
1.6.4 面向 Internet 的图像编码技术讨论	19
1.7 视频编码技术的发展	21
1.7.1 基于波形的编码技术	21
1.7.2 基于对象的编码技术	21
1.7.3 基于模型的编码技术	22
1.7.4 基于分形的编码技术	22
1.8 图像和视频可分级编码含义及分类	22
1.8.1 图像可分级编码技术	23
1.8.2 视频可分级编码技术	24
1.9 总结	25

参考文献	26
第二章 离散余弦变换与小波变换编码基础	33
2.1 离散余弦变换	34
2.1.1 一维离散余弦变换的定义	34
2.1.2 二维离散余弦变换的定义	35
2.1.3 离散余弦变换的矩阵表示	35
2.1.4 快速离散余弦变换	36
2.1.5 基于 DCT 的整数变换	48
2.1.6 编码标准选择离散余弦变换的原因	53
2.2 基于离散余弦变换的图像编码	55
2.2.1 JPEG 的基线系统	56
2.2.2 JPEG 的无损模式	63
2.3 基于离散余弦变换的视频编码	64
2.3.1 MPEG-1 视频编码标准	64
2.3.2 MPEG-2 视频编码标准	68
2.3.3 MPBG-4 视频编码标准	70
2.3.4 H. 264/MPEG-4 AVC 视频编码标准	76
2.3.5 AVS 视频编码标准	81
2.4 小波分析基础	82
2.4.1 从傅里叶变换到小波变换	83
2.4.2 连续小波变换	86
2.4.3 离散小波变换	91
2.4.4 多分辨率分析	94
2.4.5 系数分解快速算法	96
2.4.6 正交小波的构造	99
2.4.7 紧支集正交小波的构造	104
2.4.8 紧支集双正交小波的构造	104
2.4.9 二维小波的多分辨率分析及 Mallat 算法	114
2.4.10 小波图像编码的一般构架	116
2.5 提升方案小波变换	118
2.5.1 提升方案的定义	119
2.5.2 基于提升方案的传统小波构造方法	122
2.5.3 提升小波变换的基本步骤	131
2.5.4 整数小波变换	132
2.6 小波图像编码	135

2.6.1 概述	135
2.6.2 JPEG-2000 图像编码标准	136
2.7 小波基和图像分解层数对图像编码的性能分析	142
2.7.1 采用的实验图像及分类	142
2.7.2 小波基对不同类型图像 EZW 算法性能的影响	143
2.7.3 小波分解层数对不同类型图像 EZW 算法性能的影响	144
2.7.4 结果分析	146
2.8 小波视频编码	148
2.8.1 基于空间域运动估计的小波视频编码	148
2.8.2 基于小波域运动估计的视频编码	149
2.9 总结	150
参考文献	151
附录	157
第三章 图像可分级编码	164
3.1 概述	164
3.2 图像压缩标准中的可分级编码概述	164
3.2.1 JPEG 中的图像可分级编码分析	164
3.2.2 PNG 中的可分级编码分析	167
3.2.3 JPEG-2000 中的可分级编码分析	168
3.2.4 MPEG-4 VTC 中的图像可分级编码分析	170
3.2.5 比较与讨论	171
3.3 基于 DCT 的图像可分级编码	172
3.4 基于小波的图像可分级编码	173
3.4.1 空间可分级编码	173
3.4.2 数率可分级编码	173
3.4.3 空间和数率混合可分级编码	202
3.4.4 一种基于内容的数率可分级编码算法	202
3.5 基于 MP 的图像可分级编码	208
3.6 总结	209
参考文献	210
第四章 视频可分级编码	213
4.1 概述	213
4.2 视频压缩标准中的可分级编码概述	214
4.2.1 MPEG-2 中的可分级编码分析	217
4.2.2 H.263+ 中的可分级编码分析	219

4.2.3	MPEG-4 中的可分级编码分析	220
4.2.4	H.264 中的可分级编码分析	224
4.3	基于 DCT 的视频可分级编码技术	230
4.3.1	MPEG-4 中 FGS 编码的进一步讨论	231
4.3.2	细粒度可分级视频编码研究进展	235
4.4	基于二维小波的视频可分级编码	241
4.4.1	小波域视频对象的二维运动估计	242
4.4.2	小波域视频对象可分级编码	286
4.4.3	帧差图像的可分级编码	291
4.4.4	适应于小波视频可分级编码的码率控制技术	299
4.4.5	可分级编码的预测结构	306
4.5	基于三维小波的视频可分级编码	317
4.5.1	不进行运动估计和补偿的三维小波视频编码	318
4.5.2	进行运动估计和补偿的三维小波视频编码	320
4.5.3	一种典型的基于三维小波的视频可分级编码	326
4.6	总结	327
	参考文献	328
第五章	图像和视频编码技术展望	341
5.1	图像编码技术展望	341
5.2	视频编码技术展望	342
5.2.1	视频可分级编码技术方面	342
5.2.2	小波域运动估计技术方面	344
5.2.3	视频编码码率控制方面	345
5.3	总结	346
	参考文献	346

第一章 图像和视频编码技术概述

1.1 图像和视频信息的冗余特性

图像和视频压缩对多媒体信息存储、网上传输等应用是必须的,也是可行的。这主要表现在图像和视频信息中通常包含着一定的冗余(redundance),这些冗余要占用额外的存储空间,通过削减这些冗余势必会使原始的图像和视频信息得以压缩。图像和视频的这种冗余一般可分为统计冗余(statistical redundancy)和心理可视冗余(psychovisual redundancy)^[1]。

1.1.1 统计冗余

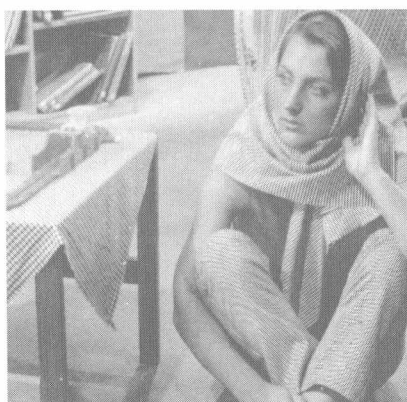
图像和视频的统计冗余又可分为空间冗余(spatial redundancy)、时间冗余(temporal redundancy)和编码冗余(coding redundancy)。前两种冗余是图像和视频帧中像素之间的冗余,即图像和视频帧中像素间的光亮度保持着一定程度的相关性,它们之间不是统计独立的;编码冗余则是指从编码(压缩)技术的角度,图像和视频信息中含有多余的信息。下面分别对这几种冗余进行讨论。

1. 空间冗余

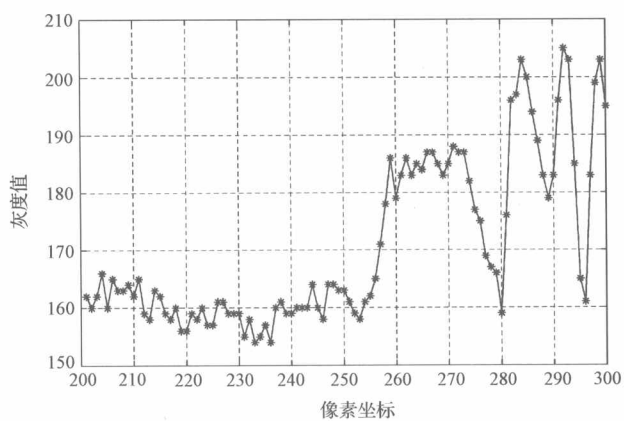
空间冗余表示一个图像或视频帧内像素间光亮度的统计相关性。在一个自然图像或视频帧内,除了边缘区域外,一般每一行(列)上大多数相邻像素光亮度的变化会保持着一种趋势,相邻像素的光亮度之间存在着很高的自相关特性,这种自相关性即空间冗余特性。图 1.1(a)给出了一幅包含人物和背景的自然图像(512×512);图 1.1(b)和图 1.1(c)分别给出了该图第 213 行和 125 列上各像素光亮度值的情况。从图中可以看出,沿着行或列像素的光亮度值的变化是渐进的,这种空间冗余特性意味着一个像素的光亮度值可以通过其相邻像素的光亮度值预测得到,这样对图像中的像素,可以通过其已编码的相邻像素对该像素进行预测,而没有必要对每一个像素都独立地表示。

2. 时间冗余

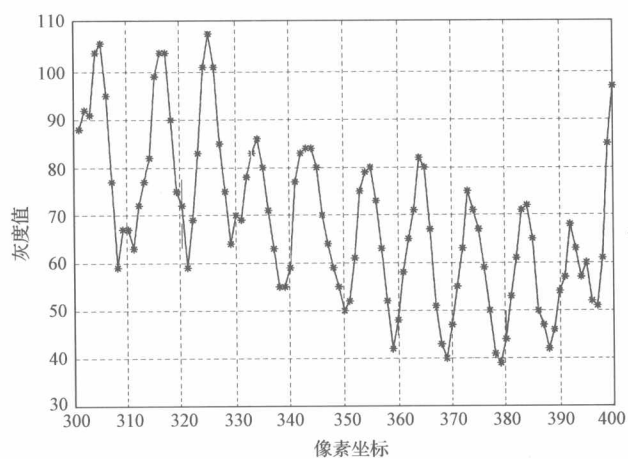
时间冗余一般仅包含在视频序列信息中,所以也被称为帧间冗余,是指视频序列中连续帧像素间的统计相关性。对于一个视频序列,只要两个连续帧间的时间



(a) Barbara (512×512)



(b) Barbara图像第213行中第201到第300列的灰度值



(c) Barbara图像第125列中第301到第400行的灰度值

图 1.1 自然图像的空间冗余特性

间隔足够小,读者可以想象相邻两帧的相似度是很高的,如图 1.2(a)和图 1.2(b)分别给出了 Stefan 序列第 4 帧和第 5 帧,从中可以看出二者的相似度。利用视频序列的这种时间冗余特性,可以沿着时间轴通过一帧对其相邻帧进行预测,从而大大削减视频信息中的这些相关信息。

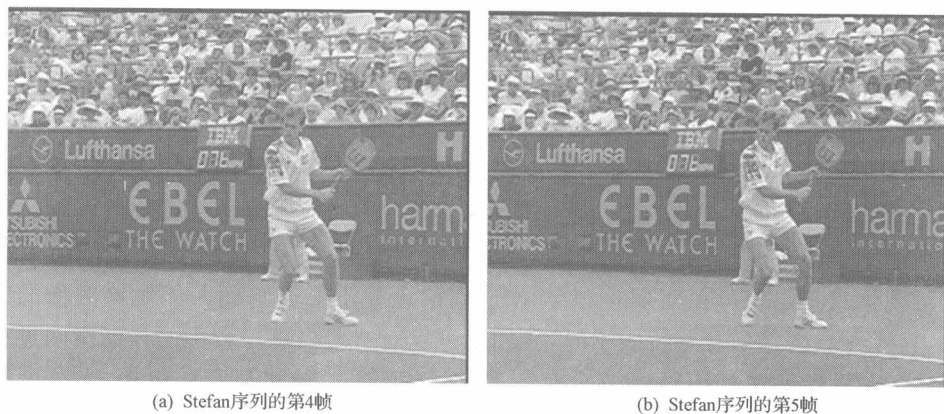


图 1.2 视频信息的时间冗余特性

3. 编码冗余

编码冗余是指信息的表示,即编码本身相关的一种冗余。表 1.1 给出了对通信联络中 8 个字符分别采用等长编码和 Huffman 编码^[2]方法的编码情况。其中 Code1 列表示等长编码的码字;Code2 列表示 Huffman 编码的变长码字。

表 1.1 8 个字符的等长编码与 Huffman 编码实例

符号	出现的概率	Code1	Code2
a1	0.05	000	0110
a2	0.29	001	10
a3	0.07	010	1110
a4	0.08	011	1111
a5	0.14	100	110
a6	0.23	101	00
a7	0.03	110	0111
a8	0.11	111	010

表 1.1 中,码字 Code1 的平均长度为 3 bit;码字 Code2 的平均长度 L_{avg} 可通过下式计算

$$L_{\text{avg}} = 4 \times 0.05 + 2 \times 0.29 + 4 \times 0.07 + 4 \times 0.08 + 3 \times 0.14 + 2 \times 0.23 + 4 \times 0.03 + 3 \times 0.11 \\ = 2.71 \text{ bit}$$

可以看出,变长编码的 Code2 较等长编码的 Code1 更加经济。可见,对于相同的符号集,不同的编码方法可能会导致编码效率的不同,比如 Code1 中就含有一些冗余,去除这些冗余不会影响到编码效率。

1.1.2 心理可视冗余

心理冗余来自人类视觉系统(human visual system, HVS)对可视图像和视频所产生的一种心理特性。对于一幅图像, HVS 并不会对其场景中的所有信息均同等对待,其中一些信息对 HVS 而言可能表现的更加重要,而对另一些信息则可能表现的不是很重要,有时甚至可能会完全忽略它们的存在。从这个意义上说,有一些可视信息是心理可视冗余的,削减掉这些信息并不会影响到人们对该图像的感知,这样在对图像信息进行表示时,我们就可以对这些信息用较小的存储空间进行表示。与图像和视频编码相关的来自 HVS 的心理视觉冗余通常包括:光亮度屏蔽(luminance masking)、纹理屏蔽(texture masking)、频率屏蔽(frequency masking)、时间屏蔽(temporal masking)和颜色屏蔽(color masking)。

1. 光亮度屏蔽

对于一幅仅包含均匀前景和背景的灰度图像,如果二者的灰度落差为零,即二者的灰度相同,则 HVS 无法将二者区分;若二者的灰度落差特别大,比如前景灰度为 0,背景的灰度为 255,则 HVS 可以非常容易地将二者区分开来。在图像和视频编码中为了使量化更加精细,更值得关心的是将前景区域从背影区域中辨别出来的临界阈值(critical threshold, or just noticeable difference)。根据 Weber 定律,有

$$\frac{\Delta I}{I} = K$$

其中, ΔI 表示临界阈值; I 表示背景刺激强度; K 表示二者的比值,该值为一近似常数值。简单地说, Weber 定律表示将前景区域从背影区域中辨别出来的临界值与背景光亮度的比例是一常数,该常数通常被称为 Weber 分数。例如 HVS 对图像前景辨别的 Weber 分数近似为 0.02^[1],若背景的光亮度为 200,则从该背景中辨别出前景的临界阈值则近似为 4,从而 HVS 能从该背景中辨别出前景的光亮度不应高于 196。从 Weber 定律可以看出,当背景很亮时, HVS 要从背景中区分出前景对象,则需要有较大的灰度级差,而对于较暗的背景,则需要较小的灰度级差。另外,从 Weber 定律可得 $\Delta I = K \times I$,可见 Weber 定律预测临界阈值 ΔI 与背景强度 I 呈线性关系。当然, Weber 定律并不总是对的,进一步的研究发现,实际临

界阈值 ΔI 的增加通常要比 Weber 定律所预测的要慢一些, 比如 Watson 等提出了下列的指数表示模型^[1]

$$\Delta I = I_0 \cdot \max \left\{ 1, \left(\frac{I}{I_0} \right)^\alpha \right\}$$

其中, I_0 为背景光亮度为 0 时预测临界阈值; α 为一常数, 近似等于 0.7。

2. 纹理屏蔽

对于一幅纹理图像, 纹理越强人眼能够识别出图像中对象的临界阈值就越大, 即人眼临界阈值随着图像细节的增加而增加, 此即为图像的纹理屏蔽效应。图像的这种纹理屏蔽效应表明人眼对图像中光滑区域较纹理区域更加敏感, 这样在图像和视频编码中, 可根据图像纹理区域的强弱变化来自适应地调节量化的级别。

3. 频率屏蔽

图像中某一区域随着频率的增加, 如增加高频随机噪声, 该区域对人眼的临界阈值也会增加, 这种效应即为频率屏蔽。HVS 之所以具有这种频率屏蔽效应主要是因为人眼更像是低通滤波器^[3], 它对高频内容的敏感程度相对较弱。利用 HVS 的这种频率屏蔽效应在对图像进行编码时可以先通过一些变换对图像进行分频处理, 比如 DCT 变换或小波变换等, 然后再对不同频率的变换系数进行不等长量化处理。

4. 时间屏蔽

当视频场景突然变化时, HVS 通常需要经历一些时间才能适应这种场景变化, 而在这过渡期间, HVS 对播放的场景细节则不是很敏感, 这种效应即为 HVS 的时间屏蔽特性。利用这种时间屏蔽特性, 在对图像和视频编码时可以对突然发生变化场景的前后若干视频帧分配相对较少的数率。

5. 颜色屏蔽

颜色作为可视光的一种感觉, 它是一个具有光亮度和一组与电磁波频谱相关的波长能量。光亮度(intensity)是可见光的一个分布, 而与电磁波频谱相关的波长能量则表现为色度(chrominance)。色度属性又包括色调(hue)和饱和度(saturation)。一个颜色的色调由颜色成分中的主波长来确定; 饱和度则是对颜色纯度的一个度量, 比如纯颜色光的饱和度为 100%, 而白光的饱和度则为 0。

在计算机图像处理中, 通常采用的颜色模型包括: RGB 模型、HIS 模型、YUV 模型、YIQ 模型、YDbDr 模型和 YCbCr 模型^[1,4]。在这几个颜色模型中, RGB 模

型是最著名的颜色模型,这是由于 HVS 中颜色的感应区域是由三个不同的人眼圆锥细胞组成,它们分别对 R、G、B 三个主颜色光进行感应,而绝大多数颜色均可分解为 R、G、B 三种色光。RGB 模型主要用于颜色图像的获取和显示。然而,对于包括图像和视频压缩的彩色信号处理,HIS、YUV、YIQ、YDbDr 和 YCbCr 等亮度-色度模型更加有效并被广泛使用,当然它们均可与 RGB 模型进行相互转换,具体转换公式如下

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} Y \\ Db \\ Dr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.450 & -0.883 & 1.333 \\ -1.333 & 1.116 & -0.217 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix},$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

$$\begin{cases} H = \frac{1}{360} \left[90 - \arctan \left(\frac{2R - G - B}{\sqrt{3}(G - B)} \right) + \{0, G > B, 180, G < B\} \right] \\ S = 1 - \left[\frac{3\min(R, G, B)}{R + G + B} \right] \\ I = \frac{R + G + B}{3} \end{cases}$$

在 RGB 颜色模型中,HVS 对绿色的敏感程度要强于红色,而对红色的敏感程度又要强于蓝色,这样如果对图像的代表仅用于人类的视觉效果,那么对 R、G、B 三个分量的同等表示并不是必要的,仅仅对 HVS 能够接收到的信息来分配较多的存储空间显然会使视频编码更加有效。此外,在亮度-色度模型中,HVS 对亮度分量的敏感程度要高于色度分量^[5,6],利用这种心理特性,在图像和视频编码中可以对亮度分量分配较色度分量更多的比特,如在图像和视频的国际编码标准中通常对图像的亮度分量采用全分辨率表示,而对两个色度分量则通过对水平和垂直方向进行采样而获得不同的表示形式。

1.2 信源编码的信息论结果

由 Shannon 所创立的信息论为图像和视频编码的研究提供了重要的指导作用。

1.2.1 信息熵

1. 信息量

信息被定义为知识、事实和消息,它的载体是一个有次序的符号。信息是可以被度量的。

设一个符号所出现的概率为 p ,则该符号所携带的信息量 I 被定义为

$$I = \log_2 \frac{1}{p} = -\log_2 p$$

其单位为比特(bit)。事实上,定义中的对数也可取其他底数,但此时信息量的单位有所不同,比如当取以 e 为底的对数时,信息量的单位为奈特(Nat),而当取以 10 为底的对数时,信息量的单位为笛特(Det),在通信中最常用的是以 2 为底的对数。它们之间具有如下换算关系

$$1\text{bit} = 0.693\text{Nat} = 0.301\text{Det}$$

2. 离散信源

信息论中信源被看成是产生消息(符号)、消息序列和连续消息的来源。在信源编码理论中,把随机过程作为信源,即在任何信号编码中,将给定信号看成是随机过程的一个实现。这里只考虑离散信源的情况。

所谓离散信源是指信源所产生的符号来自某一离散集合。设离散信源 X 所产生的符号来自离散集合 $S: \{S_i, i = 1, 2, \dots, m\}$, 离散集合 S 中各符号出现的概率为 $\{p(s_i), i = 1, 2, \dots, m\}$, 则离散信源 X 可用下式来描述

$$X = \left\{ \begin{array}{cccc} s_1 & s_2 & \cdots & s_m \\ p(s_1) & p(s_2) & \cdots & p(s_m) \end{array} \right\}, \sum_{i=1}^m p(s_i) = 1.$$

3. 信息熵

对于上述离散信源 X , 定义其中每个符号 $s_i (i = 1, 2, \dots, m)$ 所携带的平均信息量为 X 的信息熵 $H: H(X) = -\sum_{i=1}^m p(s_i) \log_2 [p(s_i)]$ 。

信源 X 的信息熵是在平均意义下表征信源总体特征的量,其单位为比特(bit)。它是信源中符号出现概率的函数,容易验证当信源中的所有符号出现的概率相同时,信源的熵达到最大。

1.2.2 Shannon 的两个信源编码定理

信源编码是指把信源中的取样符号转换成二进制比特流的过程。Shannon 信