

§ 3.6 JPEG

<http://media.hust.edu.cn>

- JPEG — Joint Photographic Experts Group
- 其中“联合”的含意是指国际电子学委员会（IEC）和国际标准化协会（ISO）联合组成的一个图像专家小组
- JPEG多年来一直致力于标准化工作，他们开发研制出连续色调、多级灰度、静止图像的数字图像压缩编码标准，因此又称为JPEG标准



数字媒体实验室

110

JPEG

- JPEG标准是一个适用范围广泛的通用标准
 - 静态图像的压缩
 - 视频图像序列的帧内图像的压缩
- JPEG专家组开发了两种基本的压缩算法
 - 采用以DCT为基础的有损压缩算法
 - 采用以预测技术为基础的无损压缩算法



数字媒体实验室

111

§ 3.6.1 JPEG压缩原理

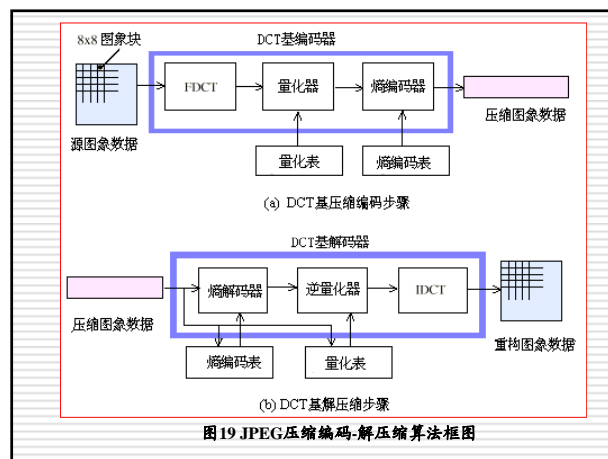
<http://media.hust.edu.cn>

- 滤掉高频信息以减少数据量，然后通过一个无损耗的压缩算法对得到的结果数据进行压缩
- 低频信息在定义一个图像的特性方面贡献更大，丢失一些高频信息未必会影响图像质量
- 利用人的视觉系统特性，使用量化和无损压缩编码相结合去掉视觉冗余和数据本身的冗余



数字媒体实验室

112



§ 3.6.2 主要步骤

<http://media.hust.edu.cn>

- **FDCT**（Forward DCT）把空间域表示的图像变换到频率域表示的图像
- 使用加权函数对DCT系数进行**量化**，这个加权函数对于人的视觉系统是最佳的
- **Z字形编码**（zigzag scan）
- 使用**DPCM**对直流系数（**DC**）进行编码



数字媒体实验室

114

主要步骤（续）

<http://media.hust.edu.cn>

- **RLE**（run-length encoding）对交流系数（**AC**）进行编码
- **熵编码**：使用霍夫曼可变字长编码器对量化系数进行编码。
- 译码或解压缩的过程与压缩码过程正好相反



数字媒体实验室

115

(1) 离散余弦变换

<http://media.hust.edu.cn>

- JPEG采用的是 8×8 大小的子块的二维DCT
- 8×8 数据块输入分解成64个正交基信号，每个正交基信号对应一个频率中的一个，这 $F(u, v)$ “谱”组成
- $F(u, v)$ 即DCT变换系数，即DCT系数值
- 64个变换系数中包括一个代表直流分量的“DC系数”和63个代表交流分量的“AC系数”



数字媒体实验室

116

- IDCT是FDCT的逆过程，它把64个DCT变换系数经逆变换运算，重建一个64点的输出图像
- 如果FDCT和IDCT变换计算所使用的设备的计算精度足够高，且系数未经过量化，那么原始的64点信号就能精确地恢复

135 144 149 153 155 155 155 155	235.6 -1.0 -12.1 -5.2 2.1 -1.7 -2.7 1.3
144 151 153 156 159 156 156 156	-22.6 -17.5 -6.2 -3.2 -2.9 -0.1 0.4 -1.2
150 155 160 163 158 156 156 156	-10.9 -9.3 -1.6 1.5 0.2 -0.9 -0.6 -0.1
155 161 162 160 160 159 159 159	-7.1 -1.9 0.2 1.5 0.9 -0.1 0.0 0.3
155 160 161 162 162 155 155 155	-0.6 -0.8 1.5 1.6 -0.1 -0.7 0.6 1.3
161 161 161 161 160 157 157 157	1.8 -0.2 1.6 -0.2 -0.8 1.5 1.0 -1.0
162 162 161 163 162 157 157 157	-1.3 -0.4 -0.3 -1.5 -0.5 1.7 1.1 -0.8
162 162 161 161 163 158 158 158	-2.6 1.6 -3.8 -1.8 1.9 1.2 -0.6 -0.4

源图像数据

FDCT系数

(2) 量化

<http://media.hust.edu.cn>

- 对DCT系数 $F(u, v)$ 量化处理是造成DCT编解码信息损失的根源
- 不同频率的余弦函数对视觉影响不同，所以可以根据不同的视觉阈值来选择量化表中的元素值的大小
- 这样通过心理视觉实验可确定对应于不同频率的视觉阈值，以确定不同频率的量化器步长



数字媒体实验室

119

亮度量化表

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

色度量化表

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

235.6 -1.0 -12.1 -5.2 2.1 -1.7 -2.7 1.3	15 0 -1 0 0 0 0 0
-22.6 -17.5 -6.2 -3.2 -2.9 -0.1 0.4 -1.2	-2 -1 0 0 0 0 0 0
-10.9 -9.3 -1.6 1.5 0.2 -0.9 -0.6 -0.1	-1 -1 0 0 0 0 0 0
-7.1 -1.9 0.2 1.5 0.9 -0.1 0.0 0.3	0 0 0 0 0 0 0 0
-0.6 -0.8 1.5 1.6 -0.1 -0.7 0.6 1.3	0 0 0 0 0 0 0 0
1.8 -0.2 1.6 -0.2 -0.8 1.5 1.0 -1.0	0 0 0 0 0 0 0 0
-1.3 -0.4 -0.3 -1.5 -0.5 1.7 1.1 -0.8	0 0 0 0 0 0 0 0
-2.6 1.6 -3.8 -1.8 1.9 1.2 -0.6 -0.4	0 0 0 0 0 0 0 0

FDCT系数

规格化量化系数

80	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

235.6	-1.6	12.1	5.2	2.1	1.7	2.7	1.3	240	0	-10	0	0	0	0	0	0
-22.6	-17.5	-5.2	-3.2	-2.9	-0.1	0.4	-1.2	-24	-12	0	0	0	0	0	0	0
-10.9	-9.3	-1.6	1.5	0.2	-0.9	-0.6	-0.1	-14	-13	0	0	0	0	0	0	0
-7.1	-1.9	0.2	1.5	0.9	-0.1	0.0	0.3	0	0	0	0	0	0	0	0	0
-0.6	-0.8	1.5	1.6	-0.1	-0.7	0.6	1.3	0	0	0	0	0	0	0	0	0
1.8	-0.2	1.6	-0.3	-0.8	1.5	1.0	-1.0	0	0	0	0	0	0	0	0	0
-1.3	-0.4	-0.3	-1.5	-0.5	1.7	1.1	-0.8	0	0	0	0	0	0	0	0	0
-0.0	1.6	-0.8	-1.8	1.9	1.2	-0.6	-0.4	0	0	0	0	0	0	0	0	0

FDCT系数 逆量化后的系数

139	144	149	153	155	155	155	155	144	146	149	152	154	156	156	156	156
144	151	153	156	159	156	155	156	148	150	152	154	156	156	156	156	156
150	155	160	163	158	156	155	156	155	156	157	158	158	157	156	155	155
159	161	162	160	160	159	159	159	160	161	161	162	164	159	157	155	155
159	160	161	162	162	155	155	155	163	163	164	163	162	160	158	156	156
161	161	161	161	160	157	157	157	163	164	164	164	162	160	158	157	157
162	162	161	163	162	157	157	157	160	161	162	162	162	161	159	158	158
162	162	161	161	163	158	158	158	158	159	161	161	162	161	159	158	158

源图像数据 重构图像数据

(3) Z字型编码

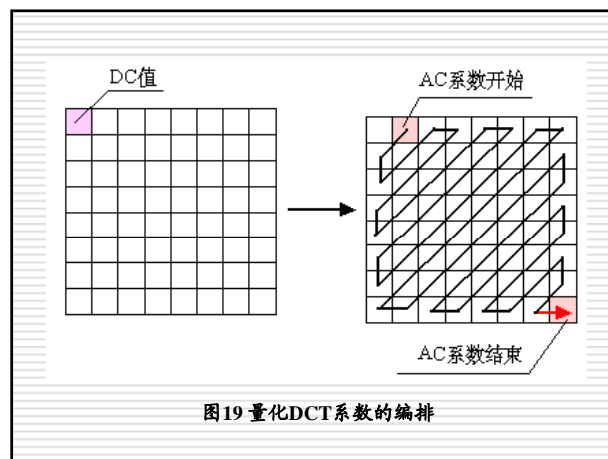
<http://media.hust.edu.cn>

- 量化后的DCT系数要重新编排，目的是为了增加连续的“0”系数的个数，就是“0”的行程长度
- 方法是按照Z字形的式样编排
- 这样就把一个 8×8 的矩阵变成一个 1×64 的向量，频率低的系数放在矢量的顶部



数字媒体实验室

124



0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

图20 量化DCT系数的序号

15	0	-1	0	0	0	0	0
-2	-1	0	0	0	0	0	0
-1	-1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

规格化量化系数

- 按行排列: 15 0 -1 0 0 0 0 0 -2 -1 0 0 0 0 0 (46个连续的0)
- Z字型编码: 15 0 -2 -1 -1 -1 0 0 -1 0 0 0 0 0 0 0 (55个连续的0)

(4) 直流系数的编码

<http://media.hust.edu.cn>

□ DC系数有两个特点:

- 一是系数的数值比较大
- 二是相邻图像块的DC系数数值变化不大

□ 根据这个特点, JPEG算法使用了差分脉冲编码调制(DPCM)技术, 对相邻图像块之间的DC系数的差值(Delta)进行编码



数字媒体实验室

128

(5) 交流系数的编码

<http://media.hust.edu.cn>

- 量化后的AC系数的特点是1*63矢量中包含有许多“0”系数, 并且许多“0”是连续的
- 因此使用RLE进行编码
- JPEG使用了1个字节的高4位来表示连续“0”的个数
- 使用低4位来表示编码下个非“0”系数所需要的位数, 跟在它后面的是量化AC系数的数值

数字媒体实验室

129

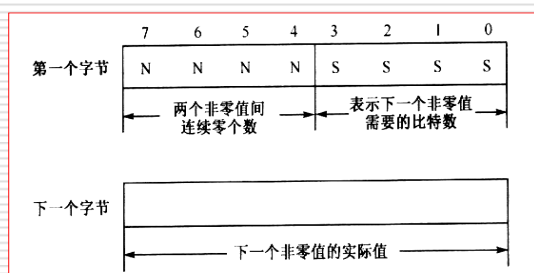


图21 AC系数行程编码码字

(6) 熵编码

<http://media.hust.edu.cn>

- 使用熵编码还可以对DPCM编码后的直流DC系数和RLE编码后的交流AC系数作进一步的压缩
- 压缩数据符号时, 霍夫曼编码器对出现比较高的符号分配比较短的代码, 而对出现频度较低的符号分配比较长的代码
- 这种可变长度的霍夫曼码表可以事先进行定义

数字媒体实验室

131

JPEG标准提供的亮度DC系数差的Huffman编码表

Table for luminance DC coefficient differences

Category	Code length	Code word
0	2	00
1	3	010
2	3	011
3	3	100
4	3	101
5	3	110
6	4	1110
7	5	11110
8	6	111110
9	7	1111110
10	8	11111110
11	9	111111110

JPEG标准提供的色度DC系数差的Huffman编码表

Table for chrominance DC coefficient differences

Category	Code length	Code word
0	2	00
1	2	01
2	2	10
3	3	110
4	4	1110
5	5	11110
6	6	111110
7	7	1111110
8	8	11111110
9	9	111111110
10	10	1111111110
11	11	11111111110

JPEG标准提供的亮度AC系数的Huffman编码表的开始部分

Table for luminance AC coefficients

Run Size	Code length	Code word
0/0 (EOB)	4	1010
0/1	2	00
0/2	2	01
0/3	3	100
0/4	4	1011
0/5	5	11010
0/6	7	1111000
0/7	8	11111000
0/8	10	1111110110
0/9	16	111111110000010
0/A	16	1111111110000011
1/1	4	1100
1/2	5	11011
1/3	7	1111001
1/4	9	111110110
1/5	11	11111110110

JPEG标准提供的色差AC系数的Huffman编码表的开始部分

Table for chrominance AC coefficients

Run Size	Code length	Code word
0/0 (EOB)	2	00
0/1	2	01
0/2	3	100
0/3	4	1010
0/4	5	11000
0/5	5	11001
0/6	6	111000
0/7	7	1111000
0/8	9	111110100
0/9	10	1111110110
0/A	12	1111111101100
1/1	4	1011
1/2	6	111001
1/3	8	11110110
1/4	9	111110101
1/5	11	11111110110

(7) 组成位数据流

<http://media.hust.edu.cn>

- JPEG编码的最后一个步骤是把各种标记代码和编码后的图像数据组成一帧一帧的数据
- 这样做的目的是为了便于传输、存储和译码器进行译码，这样的组织的数据通常称为数据流（JPEG bitstream）



数字媒体实验室

136

思考题：

<http://media.hust.edu.cn>

- 给出JPEG压缩编码算法的主要计算步骤。其中使图像质量下降的是哪一步？
- DC系数和AC系数的含义是什么？它们各有什么特点？
- 在JPEG中为什么要进行Z字形编码和RLE编码？
- 在JPEG中使用了哪些熵编码？



数字媒体实验室

137

作业题：

- 编写一段程序实现JPEG算法中的8*8的二维DCT变换、量化、逆量化和逆二维DCT变换。
- 具体要求：逐个读入下页的4个8*8的十六进制整数串（量化采用标准亮度量化表，输出内容（ASCII码），输出内容包括原始数据、变换后的数据、量化后的数据、逆量化的数据和反变换的数据。



数字媒体实验室

138

98 9C 96 99 9C A1 A1 A6 B2 C9 EA E4 C9 B8 D3 E2
94 95 95 96 98 A0 A1 A7 B3 CB E2 EA D3 CD E6 E2
95 94 91 94 9D A3 A9 A6 A7 A9 B7 BC D4 D8 C0 B2
8D 92 8F 94 8F 8F 8C 87 84 83 92 91 9D A2 98 90
7F 7C 7B 74 72 73 72 6F 6F 6B 7A 7F 8B 85 5D 4E
5A 61 6A 5D 58 54 4D 49 51 5F 6D 72 77 67 5C 54
6A 72 74 73 74 74 6F 70 72 72 7F 89 8D 94 8B 7E
77 7F 85 89 87 9A A2 A6 AE AF BE C9 CC C7 A2 89

76 7A 7C 87 91 A3 B3 C3 C3 C0 C5 CF D5 C7 99 89
7F 83 7F 7E 89 96 9A A2 A7 A3 9C 9E A6 A2 89 91
7A 7F 81 7F 7F 8C 90 90 99 96 92 90 90 8C 8C 96
7F 7B 77 77 7A 81 84 87 90 88 87 81 7F 8B 98 9A
84 7E 7C 76 74 70 72 74 74 70 74 78 8C A2 9D 94
85 81 87 88 83 7C 78 7C 80 85 8D 99 A0 A1 94 8D
88 89 92 96 96 9A 9D 9D 9C 9E A1 A1 A1 9D 8D 94
94 94 9C A1 A5 AB B2 AE A6 A5 A5 A6 A1 99 96 95

§ 3.7 MPEG <http://www.chiariglione.org/mpeg/standards/>

<http://media.hust.edu.cn>

- MPEG标准概述
- MPEG-1
- MPEG-2
- MPEG声音压缩编码
- MPEG视频压缩编码
- MPEG-4多媒体应用标准
- MPEG-7多媒体内容描述接口
- MPEG-21

数字媒体实验室

140

§ 3.7.1 MPEG标准概述

<http://media.hust.edu.cn>

- MPEG (Moving Picture Experts Group) 运动图像专家小组的活动开始于1988年, 其目的是要在1990年建立一个标准的草案
- JPEG的目标是专门集中于静止图像压缩, MPEG的目标是针对活动图像的数据压缩, 但是静止图像与活动图像之间有密切关系

数字媒体实验室

141

MPEG标准概述

- 一个视频序列图像可以看作是独立编码的静止图像序列, 只是以视频速率顺序地显示
- MPEG专家小组的研究内容
 - 视频压缩
 - 音频压缩
 - 音频和视频的同步

数字媒体实验室

142

MPEG标准概述

- MPEG标准阐明了
 - 声音和电视图像的编码和解码过程,
 - 规定了声音和图像数据编码后组成位数据流的句法
 - 提供了解码的测试方法等
- 但是, MPEG标准没有对所有内容都作严格规定, 尤其是对压缩和解压缩的算法
- 人们可以不断改进编码和解码算法, 提高声音和电视图像的质量以及编码效率

数字媒体实验室

143

已经开发和正在开发的MPEG标准有:

<http://media.hust.edu.cn>

- MPEG-1: 数字电视标准, 1992年正式发布
- MPEG-2: 数字电视标准
- MPEG-3: 已于1992年7月合并到高清晰度电视 (High-Definition TV, HDTV) 工作组
- MPEG-4: 多媒体应用标准(1999年发布)
- MPEG-7: 多媒体内容描述接口标准(2001年发布)
- MPEG-21: 数字节目的网上实时交换协议

数字媒体实验室

144

§ 3.7.2 MPEG-1

<http://media.hust.edu.cn>

- MPEG-1视频是面向比特率大约为1.5Mbit的视频压缩
- MPEG-1音频是面向每通道速率为64Kbit、128Kbit/s和192Kbit/s的数字音频信号的压缩
- MPEG-1的最终目标是解决数字视频和数字音频等多样压缩数据流的复合和同步
 - MPEG视频
 - MPEG音频
 - MPEG系统

数字媒体实验室

145

MPEG-1标准是VCD工业标准的核心，现在已经走入千家万户；利用MPEG-1音频第三层的MP3音乐格式也倍受青睐。

§ 3.7.3 MPEG-2

<http://media.hust.edu.cn>

- MPEG-2标准克服并解决了MPEG-1不能满足日益增长的多媒体技术、数字电视技术对分辨率和传输率等方面的技术要求的缺陷
- MPEG-2系统支持五项基本功能
 - 解码时多压缩流的同步
 - 将多个压缩流交织成单个的数据流
 - 解码时缓冲器初始化
 - 缓冲区管理
 - 时间识别



数字媒体实验室

147

MPEG-2视频解码器的芯片级产品已被世界上许多著名的大公司纷纷推出，是工业标准DVD的核心标准。

§ 3.7.4 MPEG-Audio编码

<http://media.hust.edu.cn>

- MPEG声音数据压缩编码不是依据波形本身的相关性和模拟人的发音器官的特性，而是利用人的**听觉系统的特性**来达到压缩声音数据的目的
- 这种压缩编码称为**感知声音编码**（Perceptual Coding）



数字媒体实验室

149

听觉系统的感知特性

<http://media.hust.edu.cn>

- 在MPEG Audio压缩编码算法中的三个特性：
 - 对响度的感知
 - 对音高的感知
 - 掩蔽效应



数字媒体实验室

150

(1) 对响度的感知

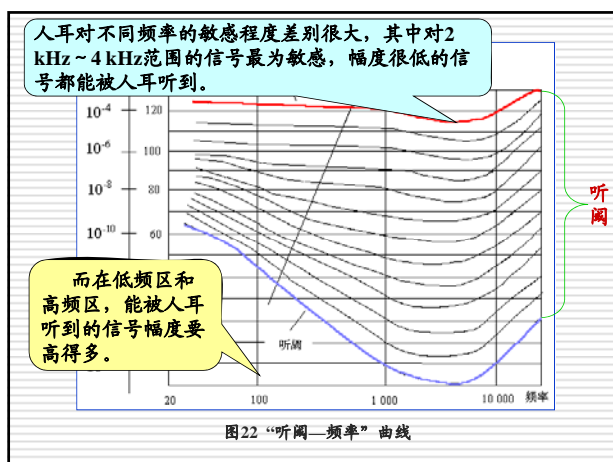
<http://media.hust.edu.cn>

- 声音的响度就是声音的强弱
- 当声音弱到人的耳朵刚刚可以听见时，我们称此时的声音强度为“**听阈**”
- 当声音强到使人耳感到疼痛时，我们称此时的声音强度为“**痛阈**”
- 实验表明，听阈是随频率变化的。测出的“听阈—频率”曲线如图22所示



数字媒体实验室

151



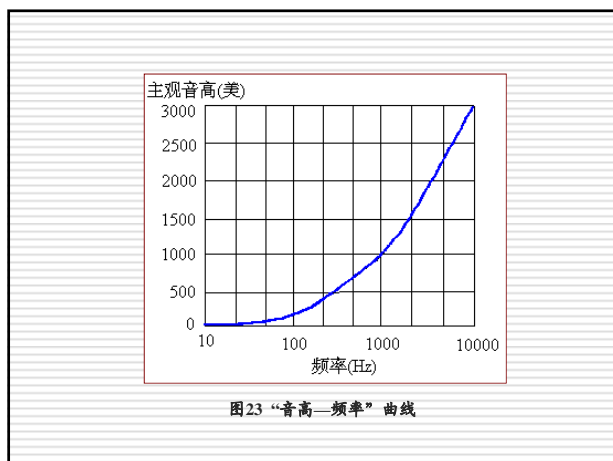
(2) 对音高的感知

<http://media.hust.edu.cn>

- 客观上用频率来表示声音的音高，其单位是Hz
- 而主观感觉的音高单位则是“美(Mel)”，主观音高与客观音高的关系是 $Mel = 1000 \log_2 (1 + f)$
- 人耳对频率的感觉也有一个范围
- 人耳可以听到的最低频率约20 Hz，最高频率约18000 Hz

数字媒体实验室

153



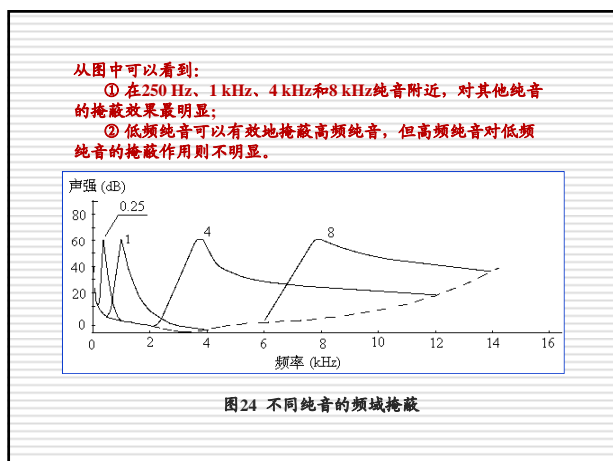
(3) 掩蔽效应

<http://media.hust.edu.cn>

- 一种频率的声音阻碍听觉系统感受另一种频率的声音的现象称为**掩蔽效应**
- 前者称为掩蔽声音(masking tone)，后者称为被掩蔽声音(masked tone)
- 掩蔽可分成频域掩蔽和时域掩蔽
- 频域掩蔽：一个强纯音会掩蔽在其附近同时发声的弱纯音，这种特性称为频域掩蔽，也称同时掩蔽(simultaneous masking)

数字媒体实验室

155



- 在时间上相邻的声音之间也有掩蔽现象，并且称为**时域掩蔽**
- 时域掩蔽又分为超前掩蔽(pre-masking)和滞后掩蔽(post-masking)
- 产生时域掩蔽的主要原因是人的大脑处理信息需要花费一定的时间
- 一般来说，超前掩蔽很短，只有大约5 ~ 20 ms，而滞后掩蔽可以持续50 ~ 200 ms

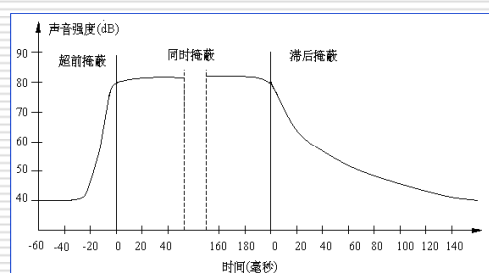


图25 时域掩蔽

压缩策略

<http://media.hust.edu.cn>

□ MPEG音频数据压缩主要依据是人耳朵的听觉感知特性，使用“**心理声学模型**”来达到压缩声音数据的目的

□ MPEG Audio采用两种感知编码：

子带感知编码（Perceptual subband coding）
Dolby AC-3（Audio Code number 3）编码

简称AC-3，由杜比实验室开发。

数字媒体实验室

159

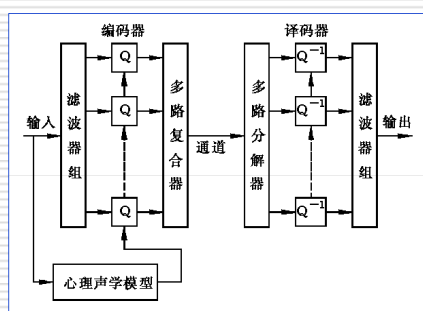


图26 感知子带压缩算法框图

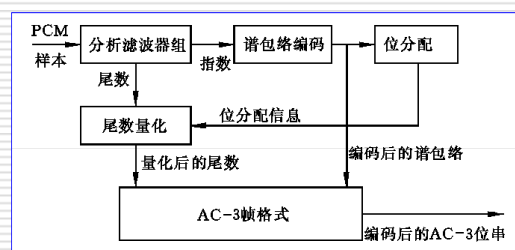


图27 Dolby AC-3压缩编码算法框图

- MPEG-1 Audio的编码对象是20~20000Hz的宽带声音，采用了感知子带编码
- MPEG-2标准定义了两种声音数据的压缩格式
 - 一种称为MPEG-2 Audio，与MPEG-1 Audio是兼容的，称为**MPEG-2 BC**（Backward Compatible）
 - 另一种称为**MPEG-2 AAC**（Advanced Audio Coding），与MPEG-1声音格式不兼容，称为非向后兼容MPEG-2 NBC（Non-Backward-Compatible）标准

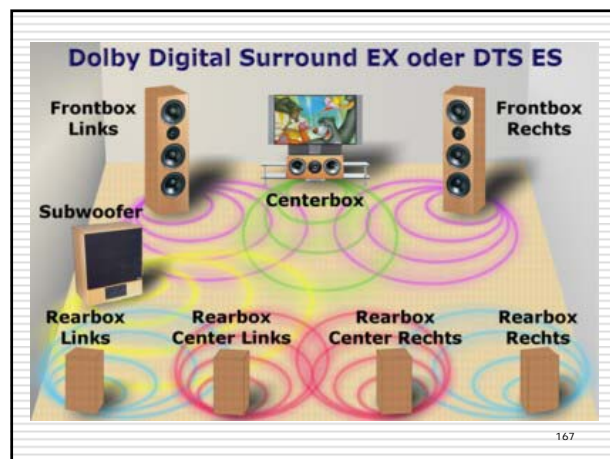
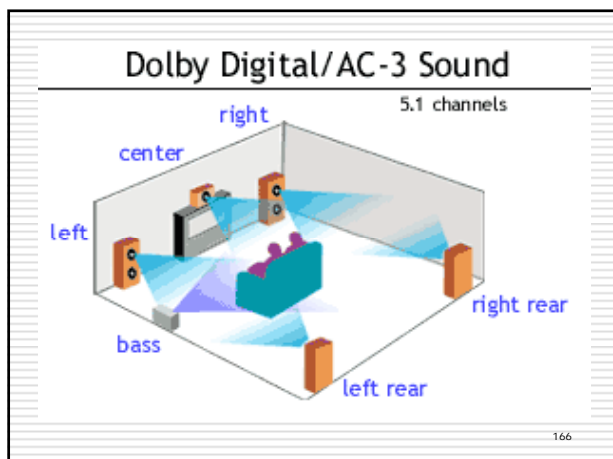
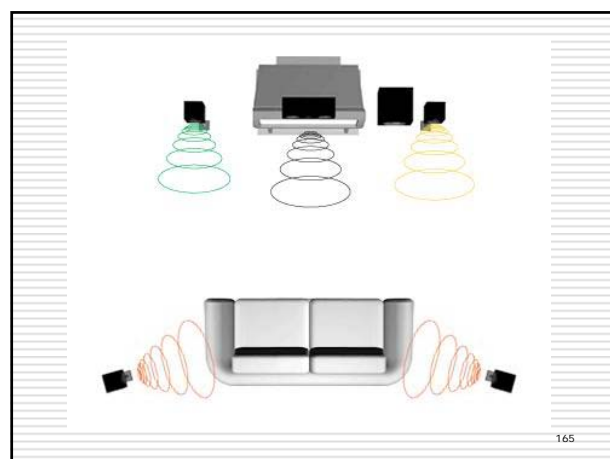
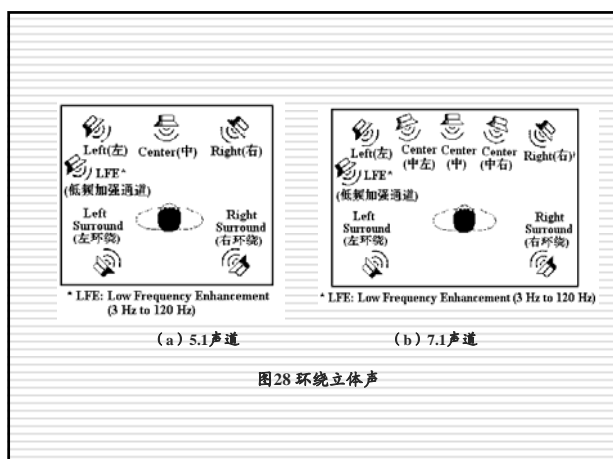
环绕立体声

<http://media.hust.edu.cn>

- MPEG-2 Audio的“5.1环绕声”也称为“3/2-立体声加LFE”，其中的“1”就是指LFE（Low Frequency Effects）声道
- 它的含义是播音现场的前面可有3个喇叭声道（左、中、右），后面可有2个环绕声喇叭声道

数字媒体实验室

163



§ 3.7.5 MPEG-Video编码

<http://media.hust.edu.cn>

MPEG-Video图像压缩技术基本思想和方法可以归纳成两个要点:

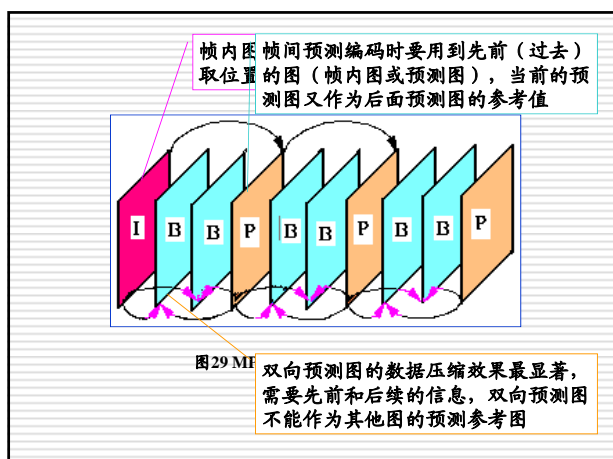
- ① 在空间方向上, 图像数据压缩采用JPEG压缩算法来去掉冗余信息
- ② 在时间方向上, 图像数据压缩采用运动补偿(Motion Compensation)算法来去掉冗余信息

MPEG的三种图像

<http://media.hust.edu.cn>

□ 为了在保证图像质量基本不降低而又能够获得高的压缩比, MPEG专家组定义了三种图像:

- 帧内图 (Intra pictures, I)
- 预测图 (Predicted Pictures, P)
- 插补图, 即双向预测图 (Bidirectional Prediction, B)



- 帧内图（I）和预测图（P）及双向预测图（B）沿时间轴上的排列顺序：

I B B P B B P B B P B B P B B P B B P B B P B B P B B

- 图的组织结构是十分灵活，它们的组合可由应用规定的参数决定，如随机存取和编码延迟等

说明：

<http://media.hust.edu.cn>

- 电视图像的帧序列中，不能全部是插补图B，B图必须由参考图进行插补，参考图可以是帧内图（I）或预测图（P），B图不能作参考图
- 在两个参考图之间出现双向预测图B的频度是可选择的
- 当增加参考图之间B图的数目时，将会减少B图与参考图之间的相关性

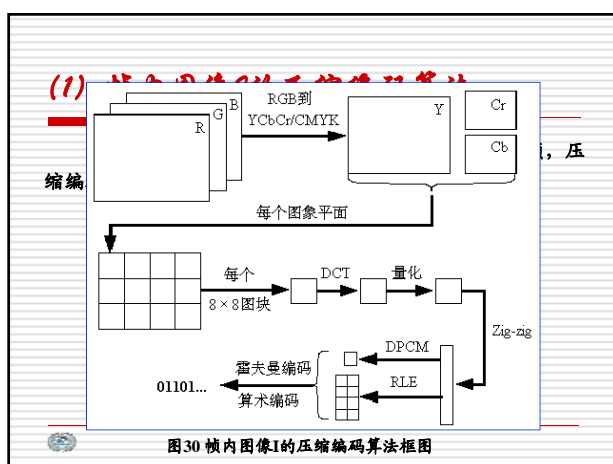
数字媒体实验室

172

I B B P B B P B B P B B P B B P B B P B B P B B P B B

- I图、P图、B图三者之间存在因果关系
- 如第4帧的P图是由第1帧的I图预测，第1帧I图第4帧P图共同预测出它们之间的双向预测B图
- 因此，接收端解码器的输入（发送端编码器的输出），不能按照时间的顺序，而是按照以下的排列顺序：

I P B B P B B P B B P B B I B B



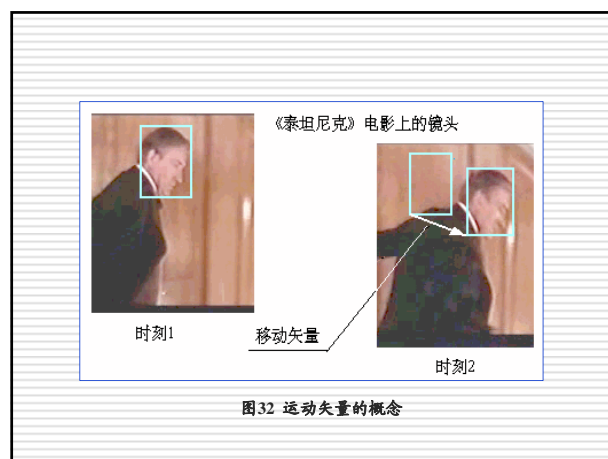
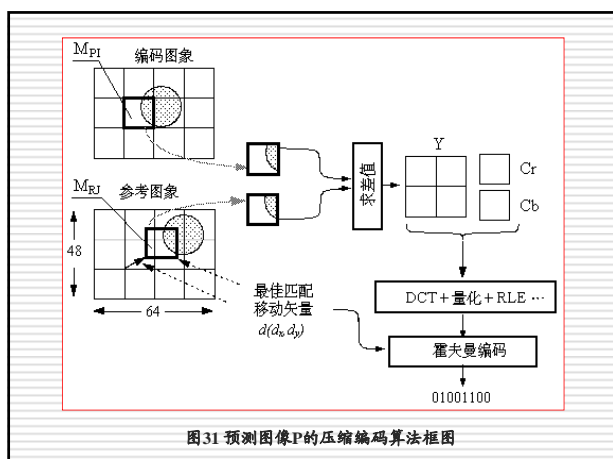
(2) 预测图像P的压缩编码算法

<http://media.hust.edu.cn>

- 预测图像的编码也是以图像宏块(macroblock)为基本编码单元，一个宏块定义为I×J像素的图像块，一般取16×16
- 预测图像P使用两种类型的参数来表示：
 - 一种参数是当前要编码的图像宏块与参考图像的宏块之间的差值
 - 另一种参数是宏块的运动矢量

数字媒体实验室

175



运动矢量的求解

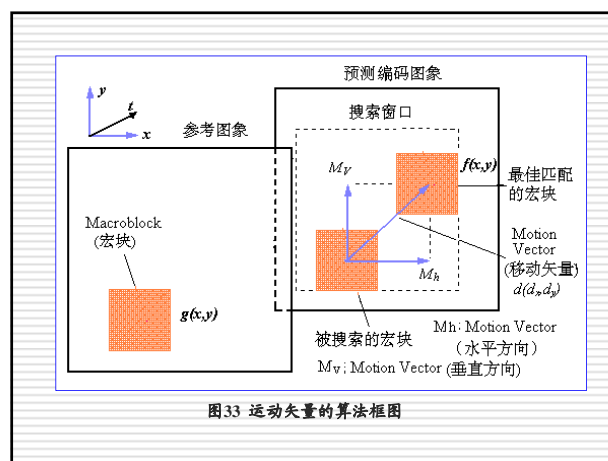
<http://media.hust.edu.cn>

- 在求两个宏块差值之前，需要找出编码图像中的预测图像编码宏块MPI相对于参考图像中的参考宏块MRJ所移动的距离和方向
- 这就是运动（移动）矢量(motion vector)



数字媒体实验室

178



最佳匹配

<http://media.hust.edu.cn>

- 要使预测图像更精确，就要求找到与参考宏块MRJ最佳匹配的预测图像编码宏块MPI
- 所谓最佳匹配是指这两个宏块之间的差值最小
- 通常以绝对值AE(absolute difference)最小作为匹配判据

$$AE = \sum_{i=0}^{15} \sum_{j=0}^{15} |f(i, j) - g(i - d_x, j - d_y)| \quad (i = j = 16)$$



数字媒体实验室

dx和dy分别是参考宏块MRJ的移动矢量d(dx, dy)在X和Y方向上的分量。

180

- 有些学者提出了以均方误差MSE(mean-square error)最小作为匹配判据

$$MSE = \frac{1}{64} \sum_{i=0}^{15} \sum_{j=0}^{15} [f(i, j) - g(i - d_x, j - d_y)]^2 \quad (i = j = 16)$$

- 也有些学者提出以平均绝对帧差MAD(mean of the absolute frame difference)最小作为匹配判据

$$MAD = \frac{1}{64} \sum_{i=0}^{15} \sum_{j=0}^{15} |f(i, j) - g(i - d_x, j - d_y)| \quad (i = j = 16)$$

分析可知:

<http://media.hust.edu.cn>

- 对预测图像的编码实际上就是寻找最佳匹配图像宏块
- 找到最佳宏块之后就找到了最佳运动矢量 $d(dx, dy)$



数字媒体实验室

182

最佳宏块搜索算法

<http://media.hust.edu.cn>

- 为减少搜索次数, 现在已开发出许多简化算法用来寻找最佳宏块, 下面介绍其中的三种:
 - 二维对数搜索法 (2D-logarithmic search)
 - 三步搜索法 (three-step search)
 - 对偶搜索法 (conjugate search)



数字媒体实验室

183

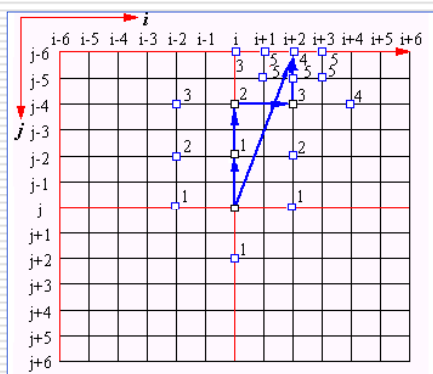


图34 二维对数搜索法

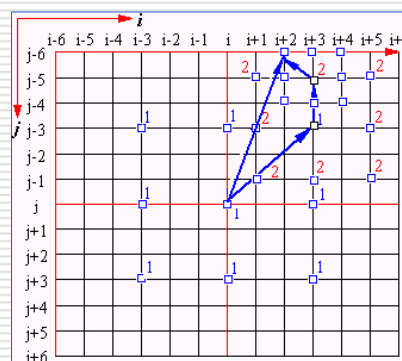


图35 三步搜索法

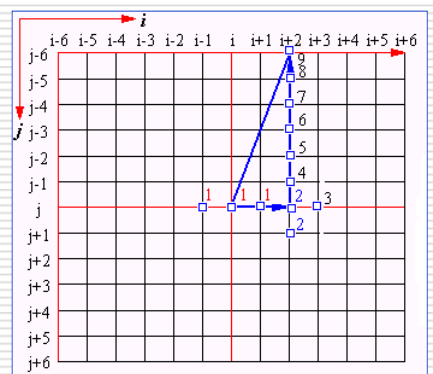


图36 对偶搜索法

(3) 双向预测图像B的压缩编码算法

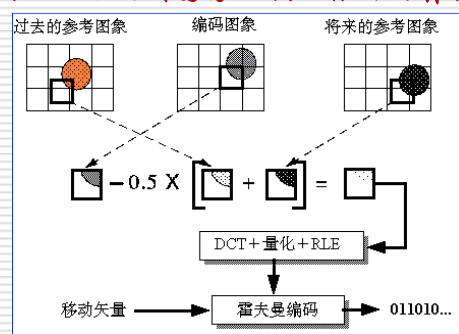


图37 双向预测图像B的压缩编码算法框图

(4) MPEG电视图像的结构

<http://media.hust.edu.cn>

- MPEG编码允许选择I图像的频率和位置，I图像的频率是指每秒钟出现I图像的次数，位置是指时间方向上I帧所在的位置
- 一般情况下，I图像的频率为2
- MPEG编码也允许在一对I图像或者P图像之间选择B图像的数目
- I图像、P图像和B图像数目的选择依据主要是根据节目的内容



数字媒体实验室

188

举例说明：

<http://media.hust.edu.cn>

- 对于快速运动的图像，I图像的频率可以选择高一些，B图像的数目可以选择少一点
- 对于慢速运动的图像I图像的频率可以低一点，而B图像的数目可以选择多一点
- 编码参数为：帧内图像I的距离为N=15，预测图像(P)的距离为M=3
- 一个典型的I、P、B图像安排如图38所示



数字媒体实验室

189

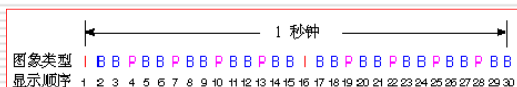


图38 MPEG电视帧编排

§ 3.7.6 MPEG-4多媒体应用标准

<http://media.hust.edu.cn>

- MPEG-4从1994年开始工作，它是为视听(audio-visual)数据的编码和交互播放开发算法和工具
- MPEG-4是一个数据速率很低的多媒体通信标准
- MPEG-4的目标是要在异构网络环境下能够高度可靠地工作，并且具有很强的交互功能



数字媒体实验室

191

MPEG-4

<http://media.hust.edu.cn>

MPEG-4的标准名是Very-low bitrate audio-visual coding (甚低速率视听编码)。MPEG-4文件有6个部分

- MPEG-4系统标准
- MPEG-4电视图像标准
- MPEG-4声音标准
- MPEG-4一致性测试标准
- MPEG-4参考软件
- MPEG-4传输多媒体集成框架



数字媒体实验室

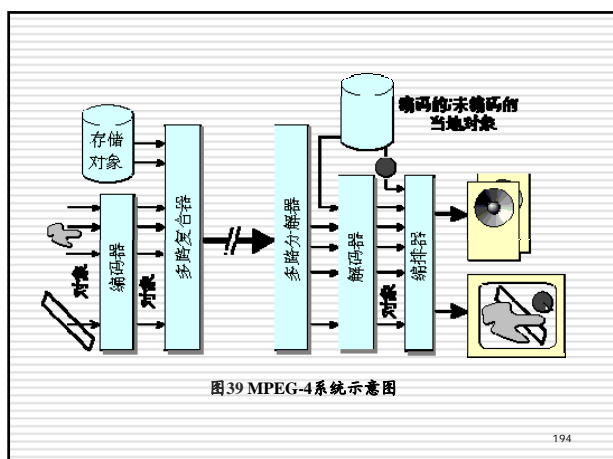
192

- MPEG-4引入了基于对象表达(object-based representation)的概念，用来表达视听对象(audio/visual objects, AVO)
- MPEG-4扩充了编码的数据类型，由自然数据对象扩展到计算机生成的合成数据对象，采用合成对象/自然对象混合编码(Synthetic/Natural Hybrid Coding, SNHC)算法
- 在实现交互功能和重用对象中引入了组合、合成和编排等重要概念

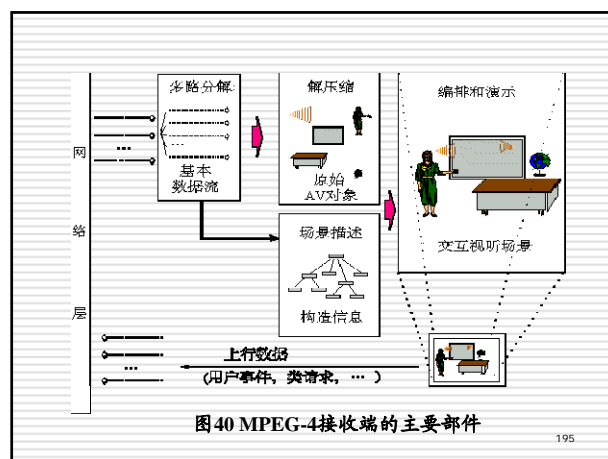


数字媒体实验室

193



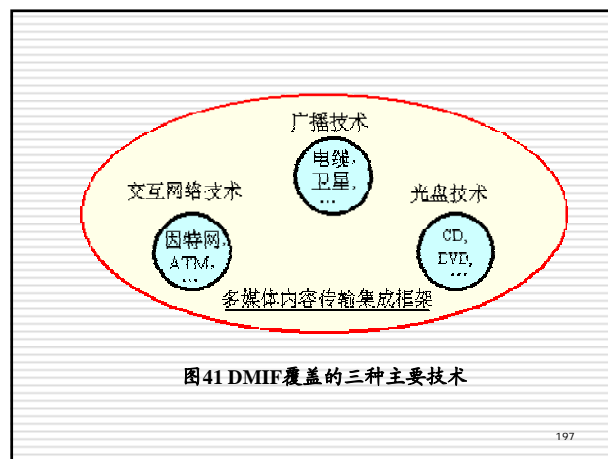
194



195

- MPEG-4中制定了一个称为**传输多媒体集成框架** (Delivery Multimedia Integration Framework, DMIF)的会话协议，用来管理多媒体数据流
- 该协议在原则上与文件传输协议FTP(File Transfer Protocol)类似，其差别是：FTP返回的是数据，而DMIF返回的是指向到何处获取数据流的指针
- DMIF覆盖了三种主要技术：广播技术，交互网络技术和光盘技术，如图41所示

196



197

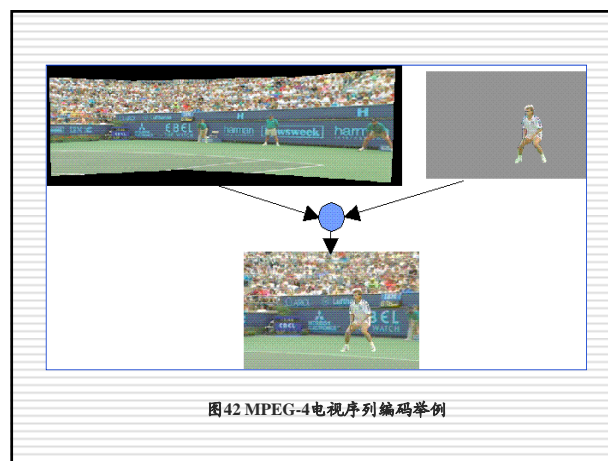
MPEG-4应用

<http://media.hust.edu.cn>

- MPEG-4将应用在移动通信和公用电话交换网（public switched telephone network, PSTN）上
- 支持可视电话（videophone）、电视邮件（video mail）、电子报纸（electronic newspapers）和其他低数据传输速率场合下的应用

数字媒体实验室

198



§ 3.7.7 MPEG-7 多媒体内容描述接口

<http://media.hust.edu.cn>

- MPEG-7的工作于1996年启动，名称叫做多媒体内容描述接口（Multimedia Content Description Interface）
- 目的是制定一套描述符标准，用来描述各种类型的多媒体信息及它们之间的关系，以便更快更有效地检索信息



数字媒体实验室

200

§ 3.7.8 MPEG21: Multimedia Framework

<http://media.hust.edu.cn>

- MPEG - 21由MPEG - 7发展而来，2001年开始启动
- MPEG - 21主要规定数字节目的网上实时交换协议



数字媒体实验室

201

MPEG-21

1	Vision, Technologies and Strategy
2	Digital Item Declaration
3	Digital Item Identification
4	Intellectual Property Management and Protection
5	Rights Expression Language
6	Rights Data Dictionary
7	Digital Item Adaptation
8	Reference Software
9	File Format

§ 3.7.4 MPEG的新计划

<http://media.hust.edu.cn>

New Standards	
MPEG-A	Multimedia Application Formats
MPEG-B	MPEG Systems Technologies
MPEG-C	MPEG Visual Technologies
MPEG-D	MPEG Audio Technologies
MPEG-E	MPEG Multimedia Middleware
MPEG-F	Universal 3D file format



数字媒体实验室

203

课后作业

<http://media.hust.edu.cn>

- 请简要论述MPEG-Video图像压缩技术的基本思想和方法，详细讨论三种图像I、P、B的具体含义及其压缩编码算法。



数字媒体实验室

204

The End