

Chapter1: Introduction

- 多媒體是融合兩種以上媒體的人機交互式信息交流和傳播媒體
 - 多媒體是信息交流和傳播媒體
 - 多媒體是人機交互式媒體
 - 多媒體信息都是以數字形式存儲和傳輸的
 - 多媒體是融合兩種以上的媒體
- 多媒體數據量大, 帶來傳輸和存儲問題
 - 解決辦法: 數據壓縮與解壓縮
- 媒體種類:
 - preception medium(感知媒體): 能直接作用於人的感官, 使人產生感覺的媒體. 感覺媒體包括人類語言, 音樂和自然界各種聲音, 活動圖像, 圖形, 曲線, 動畫, 文本等.
 - representation medium(表示媒體): 為傳輸感知信息而研究出來的中間手段, 包括語音編碼, 音樂編碼, 圖像編碼, 文本編碼, 活動圖像編碼和靜止圖像編碼等.
 - presentation medium(顯示媒體): 用於通信中電信號與感官媒體之間轉換所用的媒體, 分輸入顯示媒體和輸出顯示媒體兩種.
 - intercahnge medium
 - storage medium(存儲媒體): 用於存儲表示媒體(如硬盤, 軟盤, 光盤, 磁帶和存儲器等)
 - transmission medium(傳輸媒體): 用於傳輸表示媒體的物理實體(如電話線, 雙絞線, 銅軸電纜, 光鈎, 無線電和紅外線等)
- 4 characteristics of a multimedia system
 - system must be computer controlled
 - 在計算機的控制下, 以交互方式表示信息; 捕獲, 處理和產生信息; 提供共享存儲能力; 傳輸多媒體信息.
 - system must be integrated
 - 盡可能用集成化方法完成產生, 存儲, 傳遞和顯示信息的功能
 - information must be represented digitally
 - 以二進制序列描述數據
 - interface is usually interactive
 - 第一級別交互: 用戶僅能選擇顯示開始時間, 次序, 速度或信息項
 - 第二級別交互: 用戶的輸入能夠被記錄, 以便加注, 修改或豐富信息
 - 第三級別交互: 對用戶輸入準確處理, 然後產生應答
- size of 1 minute digital audio

数字音频格式	频带 (Hz)	带宽 (KHz)	取样率 (KHz)	量化位数	存储容量 (MB)
电话	200-3400	3.2	8	8	0.46
会议电视伴音	50-7000	7	16	14	1.6
CD-DA	20-20000	20	44.1	16	5×2
DAT	20-20000	20	48	16	5.49×2
数字音频广播	20-20000	20	48	16	5.49×6

- size of 1 minute digital vedio

数字电视 格 式	空间×时间 ×分辨率	取样率 (MHz)	量化位数	存储容量 (MB)
公用中间 格式 (CIF)	352×288 ×30	亮度 13.5; 4:1:1	亮度、色差 共 12	261
CCIR 601 号建议	PAL 720×480 ×30 NTSC 720×576 ×25	亮度 13.5 4:2:2	亮度、色差 共 16	1186.5 1186.5
HDTV 亮度信号	1280×720 ×60	4:2:2	16	6328

Chapter2: Basic of Digital Audio

- 聲音是一種像光的波, 但在某些物理設備上能被壓縮和擴張
- 人的聽覺範圍: **20~20000Hz**
- 次聲波: **<20Hz**
- 超聲波: **>20000Hz**
- 聲音是时基媒体
 - 時基媒體(Time-based Media): 時間依賴的連續媒體

- 泛音: 物體局部震動產生的次要頻率
- 基音: 物體整體震動產生的最大頻率
- 樂音: 泛音是基音的**整數**倍
- 噪音: 泛音是基音的**非整數**倍

$$SNR = 10\log_{10} \frac{V_{signal}^2}{V_{noise}^2} = 20\log_{10} \frac{V_{signal}}{V_{noise}} (dB)$$

- 信噪比(SNR):

$$SQNR = 20\log_{10} \frac{V_{signal}}{V_{quan-noise}} \approx 20\log_{10} \frac{2^{N-1}}{\frac{1}{2}} = 20 \times N \times \log_{10} 2 = 6.02N (dB)$$

- 信號量化噪聲比(SQNR):

N is quantization levels

0dB – 听觉阈

Threshold of hearing	0
Rustle of leaves	10
Very quiet room	20
Average room	40
Conversation	60
Busy street	70
Loud radio	80
Train through station	90
Riveter	100
Threshold of discomfort	120
Threshold of pain	140
Damage to ear drum	160

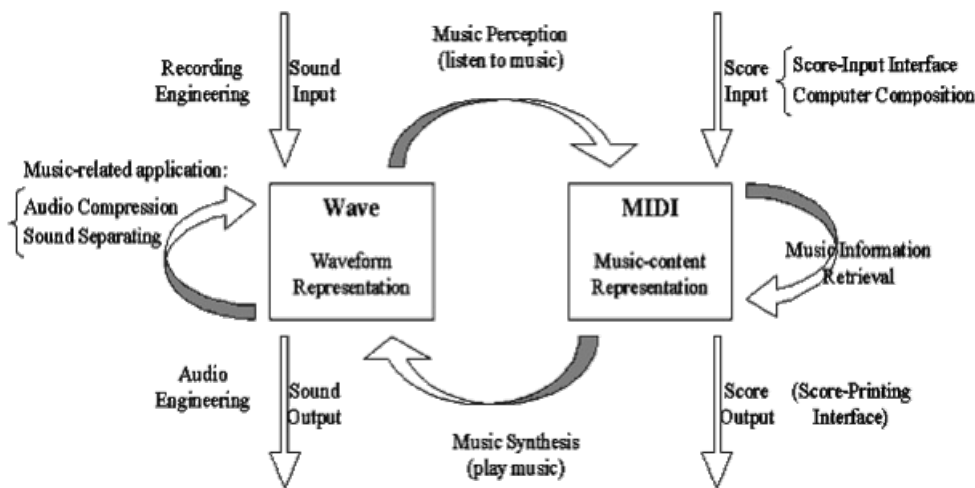
120dB – 不适阈

140dB – 痛阈

160dB – 损伤阈

- 聲音等級(dB):
- 雙耳定位: 通過雙耳判斷聲源的方向和遠近
- HRTF(Head-Related Transfer Function): 聲波從自由場傳到鼓膜處的變換函數[\[more details\]](#)
- SRS(Sound Retrieval System): 利用兩個揚聲器獲得環繞立體聲的3D立體聲系統
- 數據量計算: 數據量(Byte)=採樣頻率(Hz)(採樣位數/8)聲道數*時間(s)
- 音樂基本組成部分
 - pitch(音高): 基音的頻率決定
 - loudness(響度): 聲波的振幅決定
 - timbre(音色): 基音與泛音的比例決定
 - time(時值): 震動的持續時間
- 計算機音樂研究
 - 合成
 - FM Synthesis(頻率調製合成)
 - Wavetable Synthesis(波表合成)

- 檢索
- 感知
- 處理



- MIDI(Musical Instrument Digital Interfaces): 用於音樂合成, 樂器和計算機之間交換音樂信息的一種標準協議
- MIDI信息以**MIDI MESSAGE**來傳輸, 是控制設備執行某個動作從而產生聲音的指令

MIDI messages

channel messages

voice messages: 攜帶演奏數據

mode messages: 表示合成響應MIDI數據的方式

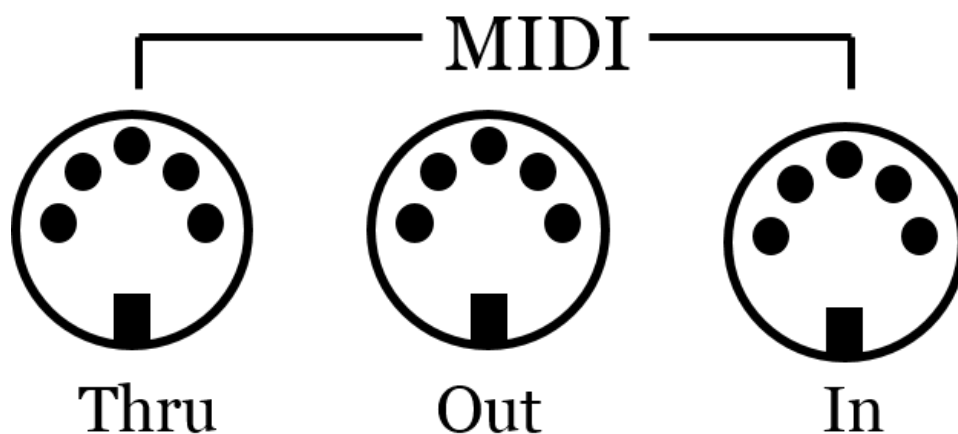
system messages

common messages: 標識系統中的所有接收器

real-time messages: MIDI部件之間同步

exclusive messages: 廠商的標識代碼

- MIDI數據流是單向異步, 每字節10bits(8bits **data**+1bit **start**+1bit **end**). 數據率為31.25kbps



- MIDI串行接口

Chapter3: Basics of Color and Image

- 顏色僅是視覺系統的一種屬性(或腦物質的一種屬性), 並不是光的屬性. 光僅僅具有波長與頻率這種物理性質, 並不具有顏色性質
- 條件等色: 光譜分佈不同但看上去相同的顏色
- 顏色三要素:
 - 色調: 反映顏色的類別(對應光譜中的主波長)
 - 飽和度: 彩色光所呈現顏色的深淺或純潔程度(100%飽和度指沒有混入白光的純色光)
 - 亮度: 光波振幅, 即光能量(人眼對**亮度**非常敏感)
- 顏色模版(Color Models): 精確標定和生成各種顏色的一套規則和定義
 - HSL: 色調, 飽和度和亮度(對於人)

- RGB: 紅, 綠, 藍磷光體(對於顯示設備)

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = 1 - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- CMYK: 青色, 品紅, 黃色和黑色(對於打印設備)
- L*a*b*: 光亮度分量, 綠到紅分量, 藍到黃分量

$$K \equiv \min\{C, M, Y\}$$

- 視頻顏色空間模型:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.212 & -0.523 & 0.311 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- YIQ: 亮度, 兩個彩色分量(NTSC彩色電視制式)
- YUV: 亮度, 兩個彩色分量(PAL, SECAM彩色電視制式)

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- YCbCr: JPEG, MPEG 壓縮, 高質量彩色視頻

- YIQ/YUV利用人眼對亮度較敏感的特性, 降低數字彩色圖像的存儲容量
- γ校正: 由於光電轉換的非線性特性, 因此引入γ校正來改善圖片效果

- 圖

- 矢量圖像(Graphics)
- 位圖圖像(Image)

- 屬性

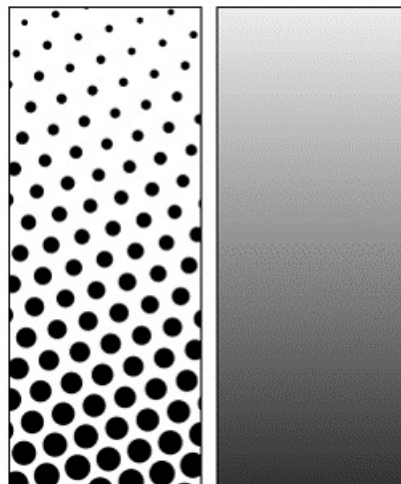
- 分辨率(resolution):

- 顯示分辨率: 顯示設備的分辨率
- 圖像分辨率: 圖像的像素密度
- 打印/掃描分辨率: 打印/掃描像素密度(dpi: dots per inch)

- 像素深度(pixel depth): 存儲所用位數(或灰度圖像的灰度級數)
- 調色板(顏色查找表): 通過調色板的索引降低存儲量, 用於只包含24位RGB彩色的N個離散點(偽彩色圖)

- 位圖圖像類型

- 真彩色: 每個像素的顏色直接由基色分量RGB決定
- 偽彩色: 每個像素的顏色要通過調色板來查找RGB決定
- 直彩色: 每個像素的顏色通過查找基色強度, 然後變換得到的RGB決定



- 抖動顯示(dithering): 通過半色調點(Halftone Spots)模擬產生灰度色調
- 流行的圖片文件格式

- BMP: the major system standard graphics file format for Microsoft Windows, used in Microsoft Paint and other programs. Many sub-variants within the BMP standard.

- Windows採用的圖像文件存儲格式
- 4 components
 - 位圖文件頭(bitmap-file header)
 - 位圖信息頭(bitmap-information header)
 - 彩色表(color table)
 - 定義位圖的字節陣列

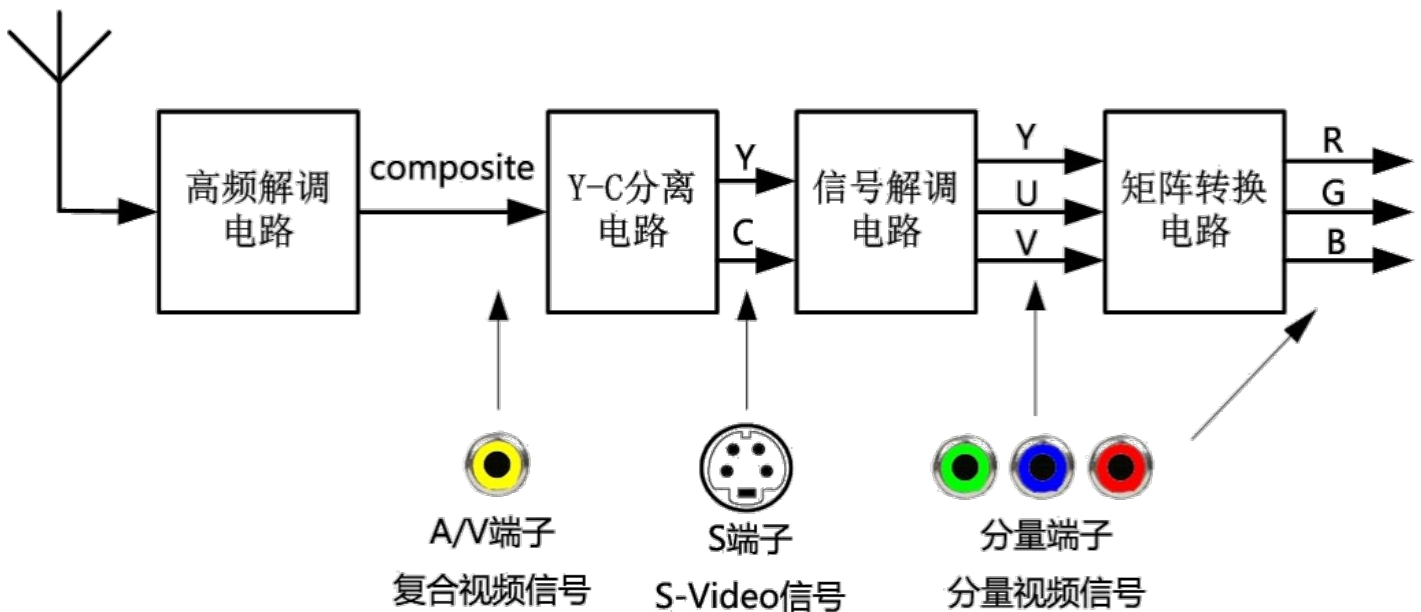
- **8-bit GIF**: one of the most important format because of its historical connection to the WWW and HTML markup language as the first image type recognized by net browsers.
 - **LZW**無損壓縮存儲圖像
 - 允許背景透明
 - 可在文件中存放多夫彩色圖像並製作出幻燈片或動畫效果
 - 支持圖片交叉存儲(縮略圖)
 - 可在圖像中加入文本
 - 以塊存儲
 - **GIF數據流組成**:
 - GIF文件頭
 - 數據塊, 數據子塊: 表示圖形/圖像
 - 控制信息塊: 顯示圖形/圖像
 - GIF文件結束塊
- **PNG**: standing for Portable Network Graphics, meant to supersede the GIF standard, and extends it in important ways.
 - 灰度圖像深度可達16位, 彩色圖像深度可達48位
 - 從**LZ77**派生的無損壓縮算法
 - **PNG數據流組成**:
 - 8 bits PNG文件署名域(PNG file signature)
 - 三個以上數據塊
- **JPEG**: currently the most important common le format.
 - 是一種連續色調, 多級灰度, 靜止圖像的數字圖像壓縮編碼方法
 - 基於離散余弦變換編碼
 - 每個像素每個分量量化為255級, 用8 bits表示
- **TIFF**: stands for Tagged Image File Format.
- **EXIF** (Exchange Image File): an image format for digital cameras, Compressed EXIF files use the baseline JPEG format.

Chapter4: Fundamental Concepts in Video

- 彩色電視系統採用亮度色度信號混合傳送的方式
- 全球彩色電視標準制式
 - **NTSC(National Television Systems Committee)**
 - 525行/幀, 30幀/秒
 - 高寬比: 4:3(電視) 3:2(電影) 16:9(高清晰度電視)
 - 隔行掃描, 2場/幀, 262.5行/場
 - 每場開始部分保留20掃描線作為控制信息, 485條線作為可視數據(Laser disc約420, S-VHS約320)
 - 彩色空間: YIQ
 - 信道帶寬: 6MHz, 亮度: 4.2MHz, 色度I:1.6MHz, Q: 0.6MHz, 色度副載頻3.58MHz
 - 美國, 加拿大, 日本, 韓國, 菲律賓和中國台灣等採用
 - **PAL(Phase-Alternative Line)**
 - 625行/幀, 25幀/秒
 - 高寬比: 4:3
 - 隔行掃描, 2場/幀, 312.5行/場
 - 彩色空間: YUV
 - 信道帶寬: 8MHz, 亮度: 5.5MHz, 色度U, V: 各1.8MHz, 色度副載頻4.43MHz
 - 逐行倒相: 每行的色度信號與上一行倒相
 - 西歐, 中國, 印度和朝鮮等採用
 - 相對於NTSC的優勢
 - 逐行倒相可以盡可能減少相位錯誤
 - NTSC沒有顏色校正
 - **SECAM(Séquential Couleur Avec Mémoire)**
 - 625行/幀, 25幀/秒
 - 高寬比: 4:3
 - 隔行掃描, 2場/幀, 312.5行/場
 - 彩色空間: YUV

- 信道帶寬: 8MHz, 亮度: 5.5MHz, 色度U, V(交替行發送): 4.25MHz, 色度副載頻4.41MHz

- 保持兼容性
 - Three RGB signals are converted into Luminance Y, and color difference signals using simple mathematical relation.
 - Retain the same signal timings of BW(黑白電視).
- 大面積著色及高頻混合原理
 - 彩色細節失明: 人眼對黑白圖像的細節有較高的分辨力, 而對彩色圖像的細節分辨力較低.
 - 大面積著色原理: 對著色面積大的各種顏色, 對細節部分不需要顯示色度的區別.
 - 高頻混合原理: 顏色的細節部分(對應於信號的高頻部分), 可以用亮度信號取代.
- 掃描方式
 - 逐行掃描
 - 隔行掃描
- 模擬視頻傳輸信號分類
 - 分量視頻: 採用三根信號線連接視頻設備傳送信號
 - 複合視頻: 只需要一根信號線來傳送信號(亮度信號與色度信號混合)
 - S-Video: 採用稱為S端子的四芯連接器



- 顏色子採樣: 對顏色信號使用的採樣頻率低於對亮度信號的採樣頻率 - 數字電視分類 - 按圖像清晰度分類 - HDTV - 垂直方向720條掃描線或更多 - 常用: (p: progressive scan, i: interlaced scan) - 720p: 1280*720p - 1080p: 1920*1080p - 1080i: 1920*1080i - 寬高比: 16:9 - SDTV - LDTV - 按信號傳輸方式分類 - 地面數字電視 - 衛星數字電視 - 有線數字電視 - 按產品類型分類 - 數字電視顯示器 - 數字電視機頂盒 - 一體化數字電視接收機 - 按顯示屏幕幅型比分類 - 4:3 - 16:9

Chapter5: Lossless Compression Algorithms

- 無損編碼: 指使用壓縮後的數據進程數據重構(解壓縮), 得到的數據與原始數據完全相同
- 有損編碼: 重構後得到的數據與原始數據不同, 但不影響人對原始數據表達的信息造成誤解
- 壓縮比: 壓縮前數據大小/壓縮後數據大小

- 信息量: $I(s_i) = \log_2 \frac{1}{p_i} = -\log_2 p_i$ (bits), p_i 為第*i*個事件的概率

$$\eta = H(S) = \sum_i p_i \log_2 \frac{1}{p_i} = - \sum_i p_i \log_2 p_i$$

- 熵(Entropy): , 表示信息源S中的信息量屬性, 也表示最小平均編碼位數

- 熵編碼:

- Huffman Coding[\[more details\]](#)

- Adaptive Huffman Coding: 隨著頻率增大, 符號重新編碼

- Arithmetic Coding[\[more details\]](#)

- Run-Length Coding(RLC): 行程編碼(適用於編碼出現連續片段的信息)

- 00000000111100000000 => 804180

- Dictionary-based Coding

- LZ77
- LZSS

- LZ78
- LZW[\[more details\]](#)

位置	1	2	3	4	5	6	7	8	9
字符	A	B	B	A	B	A	B	A	C

步骤	位置	词典	输出
		(1) A	
		(2) B	
		(3) C	
1	1	(4) A B	(1)
2	2	(5) B B	(2)
3	3	(6) B A	(2)
4	4	(7) A B A	(4)
5	6	(8) A B A C	(7)
6	--	--	(3)

Chapter6: Lossy Compression Algorithmns

- 有損主要產生於量化過程
- Quantization(量化)
 - Digitization(數字化): 在AD轉換過程, 對信號幅度量化
 - Coding(編碼): 在熵編碼前, 對正交系數, 差值或預測誤差量化
- 量化器(Quantizer)設計標準 (兩標準有限制, 需要權衡)
 - 給定量化分級層數, 滿足量化誤差最小
 - 限定量化誤差, 滿足以儘量小的平均比特數, 表示量化輸出
- 量化分類:
 - 標量量化
 - 均勻標量量化 (Uniform Scalar Quantization)
 - 非均勻標量量化 (Non-uniform Scalar Quantization)
 - 自适应标量量化 (Adaptive Scalar Quantization)
 - 矢量化
- Predictive Coding(預測編碼): 根據相關性特點考慮, 利用一個已經出現的信號對下一個信號進行預測, 得到一個預測值. 然後對實際值和預測值求差, 對這個差值信號進行編碼. (預測編碼是否有損由是否使用量化器決定)
- Predictive Coding of digital image
 - 預測類型:
 - 一維預測(行內預測): 利用同一行上相鄰的樣值進行預測
 - 二維預測(幀內預測): 利用同一行和前面幾行的數據進行預測
 - 三維預測(幀間預測): 利用相鄰幀的樣值進行預測
 - 幀間預測編碼中視頻信號的冗餘度主要體現在空間相關性(幀內), 時間相關性(幀間)和色度相關性
 - 幀間預測編碼被廣泛應用到MPEG系列視頻壓縮標準
 - 活動補償幀間預測步驟
 - 圖像分割: 把圖像劃分為靜止的背景和若干運動的物體, 由於構成同一物體的所有像素位移相同
 - 運動矢量估值: 考察前後兩個動畫, 利用運動估值算法得到兩個物體的位移矢量
 - 運動補償: 用運動矢量補償物體的運動效果, 再進行預測
 - 編碼: 除了差值編碼外, 對位移矢量以及區域分割信息進行編碼
- Differential PCM(DPCM): 量化表需要設計
- Adaptive Differential PCM(ADPCM): 預測系數自適應調正, 量化器參數自適應調整
 - 優勢
 - 有效降低信號失真
 - 斜率過載
 - 顆粒噪聲
 - 假輪廓噪聲
 - 邊沿繁忙
- Transform Coding(變換編碼): 將原始信號經過數學上的正交變換後, 對變化系數進行量化, 編碼和傳輸
 - (Karhunen-Loève Transform)K-L變換: 可逆線性變換(去相關性好)
 - DCT(離散余弦變換): 用DCT的變換矩陣求解代替K-L的特徵向量求解, 可以節省大量計算, 廣泛應用於圖像壓縮編碼. IDCT用於重構信號

- DFT(離散傅立葉變換)
- FFT(快速傅立葉變換)
- 小波(wavelet)變換

- 1D Discrete Cosine Transform(DCT):

$$F(u) = \frac{C(u)}{\sqrt{\frac{1}{2}N}} \sum_{i=0}^{N-1} \cos \frac{(2i+1)u\pi}{2N} f(i)$$

- 2D Discrete Cosine Transform(DCT):

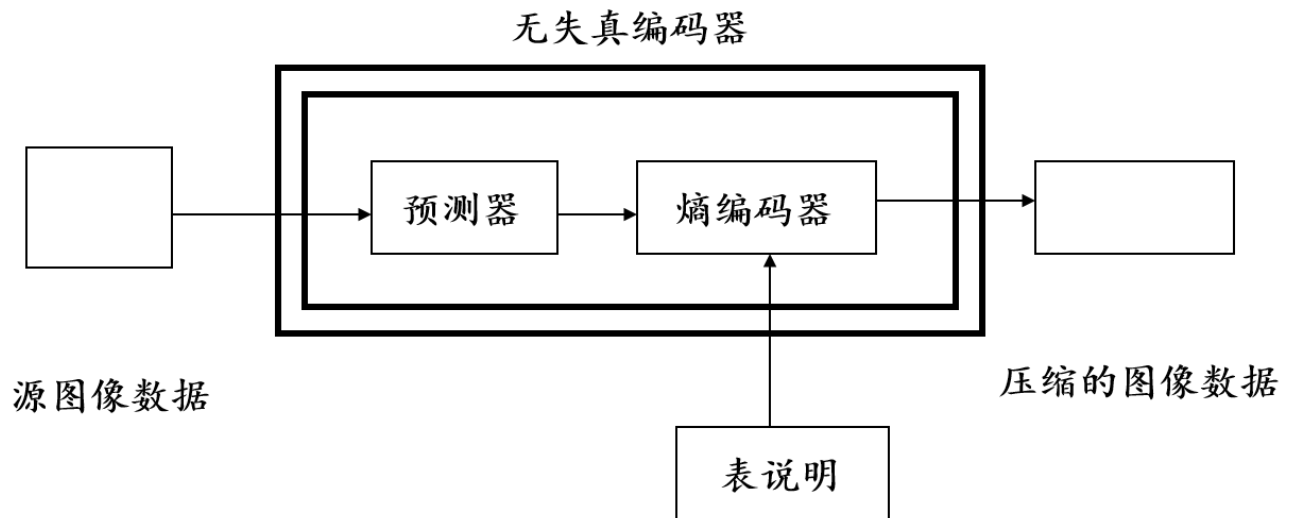
$$F(u) = \frac{C(u)C(v)}{\sqrt{2N}} \sum_{i=0}^7 \sum_{j=0}^{N-1} \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{2N} f(i,j)$$

- 3D Discrete Cosine Transform(DCT):

$$F(u, v, w) = \frac{C(u)C(v)C(w)}{\sqrt{8N}} \sum_{i=0}^7 \sum_{j=0}^7 \sum_{k=0}^{N-1} \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} \cos \frac{(2k+1)w\pi}{2N} f(i,j,k)$$

Chapter7: Image Compression Algorithmns

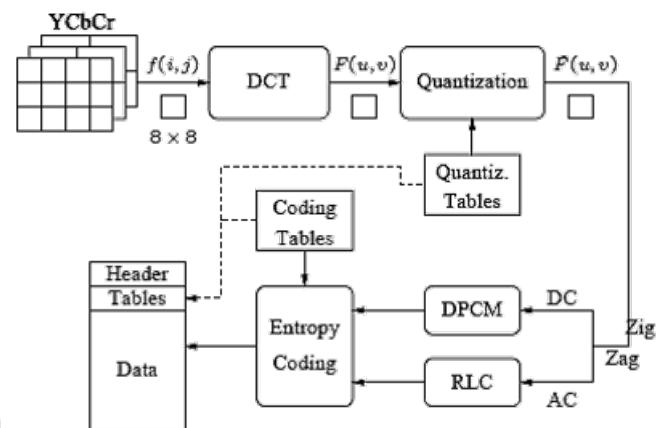
- JPEG(Joint Photographic Experts Group)
- JPEG壓縮算法:
 - 採用DCT(離散余弦變換)為基礎的有損壓縮算法
 - 採用預測編碼技術為基礎的DPCM無損壓縮算法
- JPEG模式:
 - Lossless Mode: 基於DPCM的無損編碼模式(三鄰域二維預測編碼和熵編碼), 壓縮比可達2:1



DPCM预测编码框图

		c	b	d
		a	x	

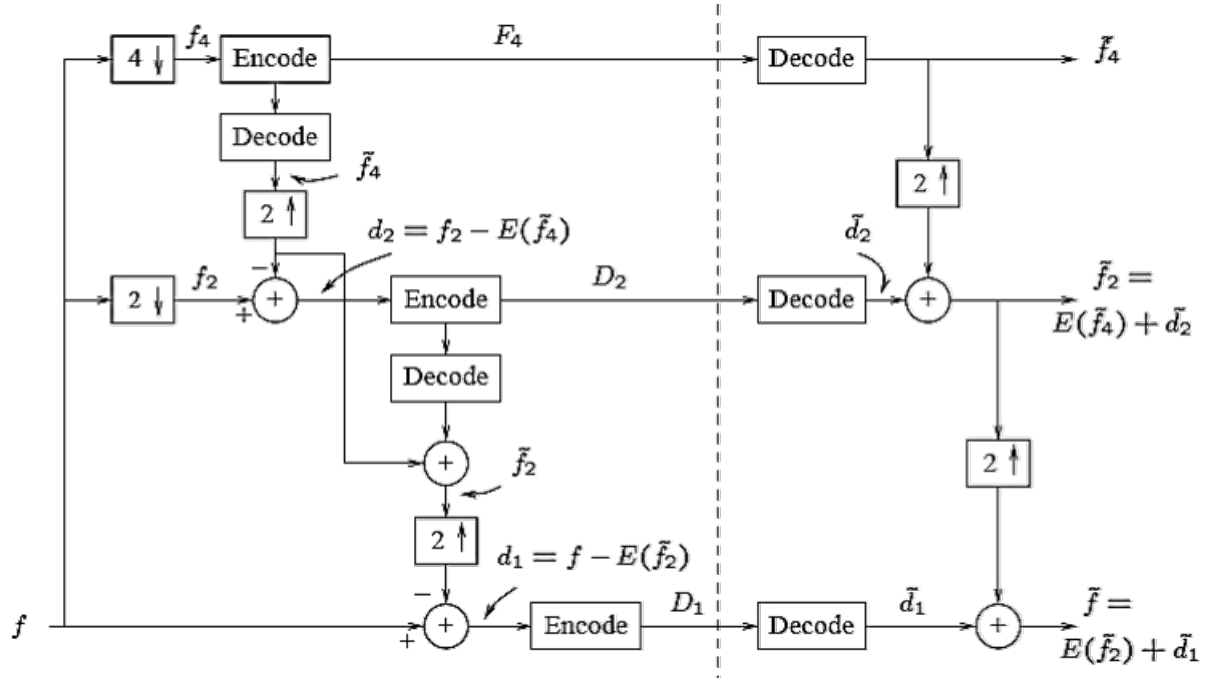
选择值	可能的预测值
0	非预测
1	a
2	b
3	c
4	a+b-c
5	a+(b-c)/2
6	b+(a-c)/2
7	(a+b)/2



◦ Sequential Mode: 基於DCT的有損順序編碼模式, 壓縮比可達10:1

- 核心壓縮編碼步驟：
 - 使用FDCT (Forward Discrete Cosine Transform) 正向離散余弦變換把空間域表示的圖變成頻率域表示的圖。
 - 使用加權函數對DCT系數進行量化。(加權函數對於人的視覺是最佳的)
 - 對直流系數採用DPCM, 對交流系數採用RLE, 然後用Huffman或者Arithmetic進行熵編碼。
- 量化: JPEG採用均勻量化器進行量化, 由於人眼對亮度信號比對色差信號更敏感, 因此採用兩種根據心理學視覺加權函數得到的量化表
 - 亮度量化值
 - 色差量化值

- Progressive Mode: 基於DCT的遞增編碼模式
 - 模式與Sequential Mode相似, 不同之處在於編碼要經過多次掃描, 每次掃描的粗糙程度不一樣
 - 分類:
 - Spectral Selection: 按頻段遞增
 - Successive Approximation: 按位遞增
- Hierarchical Mode: 基於DCT的分層編碼模式



- 步驟:
 1. 降低原始圖像的空間分辨率
 2. 對已降低分辨率的圖像按照順序編碼進行壓縮存儲
 3. 對低分辨率圖像解碼, 用插值法提高圖像分辨率
 4. 對高分率圖像, 把原圖像的預測值跟原圖像進行差值運算, 然後基於DCT編碼
 5. 反復c, d

- JPEG2000 Standard
 - 編碼方式
 - DCT-based
 - wavelet-based
 - JPEG2000可以處理多達256個通道的信息, 而JPEG只能處理三個顏色通道
- JPEG-LS Standard
 - 該標準主要特點是減低複雜度
- JBIG(Joint Bi-level Image Processing Group)
 - 該標準為了允許用電子錶處理文檔
 - 無損壓縮
 - JBIG-2提出了基於模型的編碼方式

Chapter8: Basic Audio Compression Techniques

- 聲音編碼方式
 - Waveform Coding: 基於音頻數據的統計特性進行編碼, 直接對波形信號進行採樣和量化
 - 脈衝編碼調製(PCM)
 - 增量調製(DM)
 - 自適應脈衝編碼調製(APCM)
 - 差分脈衝編碼調製(DPCM)
 - 自適應差分脈衝編碼調製(ADPCM)
 - Source Coding: 基於話音的聲學參數進行編碼, 提取話音生成模型的參數, 進行編碼
 - Hybrid Coding: 綜合上述兩種編碼技術
 - Perceptual Coding: 基於人的聽覺特性進行編碼

- 兩種基於對數關係的非均勻量化方法

- μ 律壓擴法:
$$F_{\mu}(x) = \text{sgn}(x) \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)}$$
, μ 一般取值100或255

- A律壓擴法:
$$F_A(x) = \begin{cases} \text{sgn}(x) \frac{A|x|}{1 + \ln A} & , |x| \leq \frac{1}{A} \\ \text{sgn}(x) \frac{1 + \ln(A|x|)}{1 + \ln A} & , \frac{1}{A} \leq |x| \leq 1 \end{cases}$$
, A一般取值87.6

- DM(Δ M), 增量調製: 是一種預測編碼技術, 當 $\Delta > 0$ 時, 上升; 當 $\Delta < 0$ 時, 下降.

- **ADM**, 自適應增量調製: 自動調整量化階 Δ 的大小
 - 調整方法:
 - 前向自適應: 根據量化樣本值的均方根來估計
 - 後向自適應: 根據過去樣本, 提取量化階信息
 - **APCM**(自適應脈衝編碼調製): 根據輸入信號的幅度大小來調整量化階
 - **DPCM**(差分脈衝編碼調製): 根據過去的樣本去估算下一個樣本信號的幅度大小
 - **ADPCM**: 綜合了APCM和DPCM
- **話音(Voice)**的形成: 肺中空氣受到擠壓形成氣流, 通過聲帶沿著聲道釋放出去, 從而形成**話音**
 - 濁音: 聲道打開, 聲帶再先打開後關閉, 氣流經過使聲帶發生張弛振動, 變為准週期震動氣流 (其激勵源可等效為准週期的脈衝信號)
 - 清音: 聲帶不振動, 而在某處保持收縮, 氣流在聲道裏收縮後告訴通過產生湍流, 再經過主聲道的調整形成 (其激勵源可等效為一種白噪聲信號)
 - 爆破音: 聲道關閉之後產生壓縮空氣, 然後突然打開聲道形成
- 相位不敏感特性: 只要能量頻譜分佈正確, 無須擔心波形的確切形式
- 聲音質量:
 - 廣播質量: 帶寬為**7000Hz**的中等質量話音
 - 長途電話質量: 帶寬為**3400Hz**, 信噪比為**30db**, 有失真
 - 通信質量: 完全可以聽懂, 但與長途電話質量相比有明顯失真
 - 合成質量: **80%~90%**的可懂度, 聽起來像機器聲音, 失去講話者的特徵
- 聲音冗餘:
 - 幅度非均勻分佈
 - 樣本之間相關性
 - 週期之間相關性
 - 基音之間相關性
 - 靜止系數(話音間隙)
 - 長期相關性
- **通道聲音合成器(Channel Vocodor)**: 使用一組濾波器將信號分解為不同的頻率分量, 同時進行分析, 以決定音調和語音的激勵源. 它產生一個描述聲音模型的激勵參數向量.
- **共振峰聲碼器(Format Vocodor)**: 語音中部分頻率強度較大, 稱為重要頻率(共振峰).
- **線性預測編碼(LPC)**: 通過分析話音波形來產生聲道激勵和轉移函數的參數, 然後對參數進行編碼.
- **碼激勵線性預測(CELP)**: 試圖通過複雜的激勵描述機制來彌補簡單的LPC模型在語音質量上的缺陷. 首先使用一個完整的激勵向量集合, 即一個碼本, 和真實語音匹配, 然後把最佳匹配項序號發送出去.
- **混合激勵合成器(Hybrid Excitation Vocoder)**
 - **MBE**(Multi-Band Excitation)
 - **MELP**(Multi-Band Excitation Linear Predictive)
- **MPEG音頻壓縮**: 沿用心理測量(Psychoacoustics)模型的優點, 構建一個大的多維查找表, 用很少的位數來傳送掩蔽的頻率部分
 - **Auditory Masking**(聲音掩蔽效應): 一部分聲音阻礙聽覺系統感受另一部分聲音的現象
 - **Psychoacoustics**
 - (Equal-Loudness Relations)等響度關係
 - (Frequency Masking)頻率掩蔽效應
 - (Temporal Masking)時間掩蔽效應
 - **MPEG-1**: 先對輸入應用一個濾波器組, 把輸入分解得到頻率成分, 同時對數據應用心理聲學模型, 用於位分配塊; 使用被分配的位數, 量化從濾波器組得到的信息.
 - 採用**感知子帶**編碼:
 - 將音頻信號用濾波器分成**32**個子帶
 - 用**FFT**將子帶變換到頻率域
 - 根據心理聲學模型估計各個子帶的感知閾值
 - 根據對感知閾值的估計, 對各個子帶進行比特分配和量化
 - **MPEG-1音頻**使用**3**個向下兼容的音頻壓縮層
 - **Layer1**: 子帶是頻帶相等的子帶, 其心理聲學模型僅使用頻域掩蔽特性. 主要用於小型數字盒式磁帶(digital compact cassette, DCC), 採用192kbps.
 - **Layer2**: 除了使用頻域掩蔽特性, 還利用時間掩蔽特性, 並且在低, 中和高頻段對比特分配作了一些限制. 其應用包括數字廣播聲音(digital broadcast audio, DBA), CD-I(compact disc-interactive), 數字音樂和VCD(video compact disc)等.
 - **Layer3**: 使用臨界頻帶濾波器, 把聲音頻帶分成非等寬帶的子帶, 心理聲學模型除了使用頻域掩蔽特性和時間掩蔽特性外, 還考慮到立體聲數據的冗餘, 編碼方面使用到**Huffman Coding**.
 - 層數越高, 心理声学模型應用更複雜

- 層數越高, 壓縮質量更好
- MP3壓縮性能:

Sound Quality	Bandwidth	Mode	Compression Ratio
Telephony	3.0 kHz	Mono	96:1
Better than Short-wave	4.5 kHz	Mono	48:1
Better than AM radio	7.5 kHz	Mono	24:1
Similar to FM radio	11 kHz	Stereo	26 - 24:1
Near-CD	15 kHz	Stereo	16:1
CD	> 15 kHz	Stereo	14 - 12:1

◦ MPEG-2

- BC Audio標準與MPEG-1標準相同
- 三層壓縮層與MPEG-1標準相同
- 擴充部分:
 - 增加了16kHz, 22.05kHz和24kHz的採樣頻率
 - 擴展了編碼器的輸出速率範圍, 由32~328kbps到8~640kbps
 - 增加了聲道數, 支持5.1聲道和7.1聲道的環繞聲
 - 支持Liner PCM和Dolby AC-3編碼

- AAC(Advanced Audio Coding)標準

◦ MPEG-4

- 將多種不同音頻成分整合成一種標準: 語音壓縮, 基於感知的編碼器, 文語轉換, MIDI.
- 可集成從話音到高質量的多通道聲音, 從自然聲音到合成聲音, 編碼方式還包括參數編碼(parametric coding), 碼激勵線性預測(code excited linear predictive)編碼, 時間/頻率編碼, 結構化聲音(structured audio)編碼和文本-語音TTS(text-to-speech)系統的合成聲音等

● 音頻壓縮標準(Audio Compression Standards)

- G711: Pulse Code Modulation (PCM) of Voice Frequencies
- G722: 7 kHz Audio Coding within 64 kbit/s
- G723.1: Dual Rate Speech Coder for Multimedia Telecommunication Transmitting at 5.3 & 6.3 kbit/s
- G725: System Aspect for the Use of the 7 kHz Audio Codec within 64 kbit/s
- G726: 40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)
- G727: 5-,4-,3-,and 2-bits/sample embedded adaptive differential pulse code modulation
- G728: Coding of Speech at 16 kbit/s Using Low-Delay Code Excited Linear Prediction (LD-CELP)
- G729: Coding of Speech at 8kbit/s using Conjugate-Structure Algebraic-Code-Excited Linear Prediction(CS-ACELP)
- G764: Voice Packetization - Packetized Voice Protocol
- MPEG Audio
- ...

- GSM(Global System for Mobile communications, 全球數字移動通信系統)音頻編碼, 用於數字蜂窩電話通信協議.

Chapter9: Basic Video Compression Techniques

- 視頻是由時間上連續的一系列幀(圖像)組成
- 解決視頻壓縮的可行方法: 基於前面的幀進行預測編碼, 首先按順序提取圖像, 然後進行預測, 最後對差值進行編碼
- 運動補償: 由於幀間圖像差別由攝像頭或物體運動造成, 因此可以通過在這些幀裏探測相應像素或區域的移動, 並測量其差值來進行運動補償
 - 運動估計(查找運動矢量)
 - 基於運動補償的預測
 - 預測誤差生成
- 宏塊: 每幀圖像分成大小為N*N的宏塊
- 運動矢量: 參考宏塊到目標宏塊的位移
 - 其求解是一個匹配問題, 即相關性判定問題

- 由於矢量搜索計算複雜度較高, 因此限制在一個較小的相鄰區域內進行
- 獲取一個宏塊的運動矢量複雜度為: $(2p+1) \times (2p+1) \times N^2 \times 3 \implies O(p^2 N^2)$

$$O(\log_2 p \times N^2)$$

- 2D Logarithmic Search(二維對數搜索)複雜度為

- 基於運動補償的壓縮編碼, 只需要對運動矢量和差值宏塊編碼

- 數字視(音)頻編解碼標準

- H.26x系列 (ITU標準)

- H.261

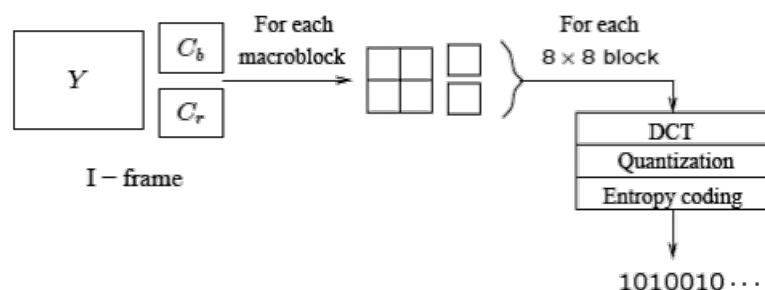
- 冗餘降低策略:

- 利用二維DCT減少圖像的空間域冗餘度
- 利用運動補償預測減少圖像的時間域冗餘度
- 利用視覺加權量化減少圖像的"感知域"冗餘度
- 利用熵編碼來減少圖像的"統計域"冗餘度

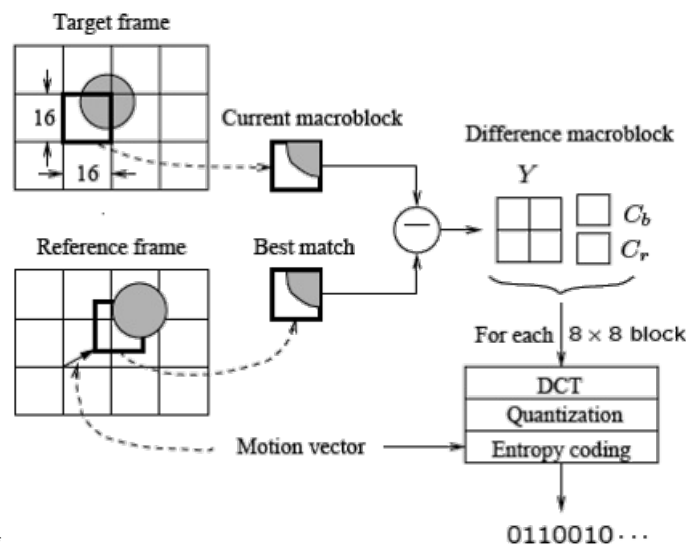
- 信源編碼算法:

- 將預測誤差或輸入圖像劃分為 8×8 的像素塊, 然後將四個亮度像素塊和兩個在空間位置上與之重疊的色差像素塊符合成一個 16×16 的宏塊
- 對於幀序列的第一副圖像, 採用幀內變換編碼
- 幀間預測採用混合方法

- 定義了兩種類型的幀



- I幀



- P幀

- 應用: 可視電話, 視頻會議等對稱應用

- H.262(MPEG-2)

- H.263

- 支持CIF和QCIF格式等
- 運動向量不僅從當前宏塊產生, 水平分量和垂直分量分別由左上, 上方和右上的宏塊預測得出
- 應用: 通用電話交換網, 局域網視頻通信

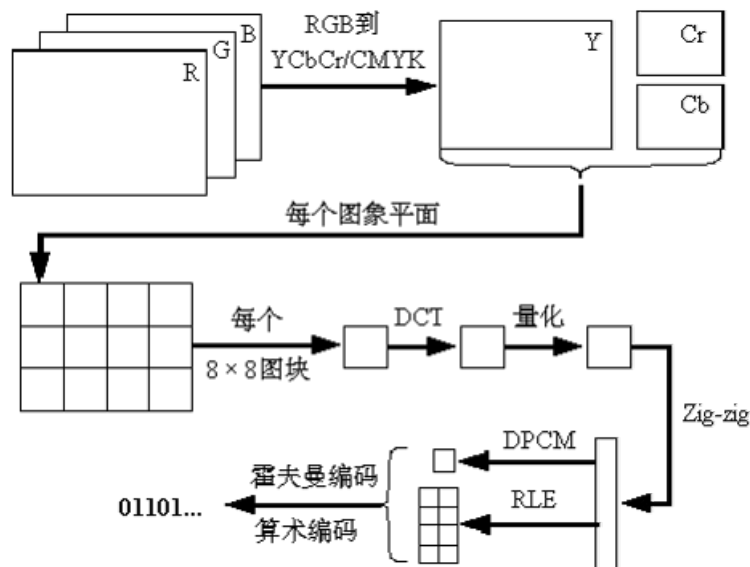
- H.264(MPEG4 part10/AVC)

- 基於VLC的熵編碼: UVLC, CAVLC
- 運動補償: 分割成 4×4 的精準運動補償
- 幀內預測: H263+更為複雜的預測方式
- 轉換, 掃描和量化: 整數精確 4×4 DCT, 非線性量化步長
- 含有環路去塊效應濾波器
- 增加了差錯恢復能力
- 應用: 視頻廣播, 視頻通信和存儲媒體

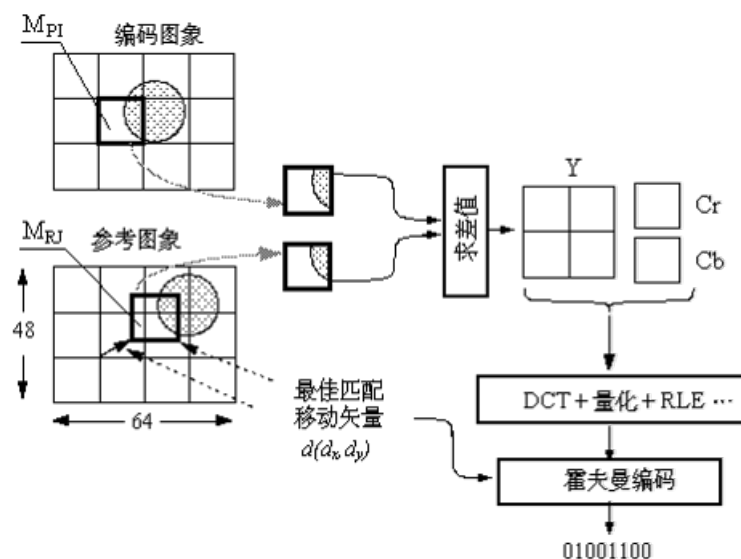
◦ MPEG系列 (ISO/IEC標準)

■ MPEG-1: 低於1.5Mb/s的數字化運動圖像及伴音編碼(1992)

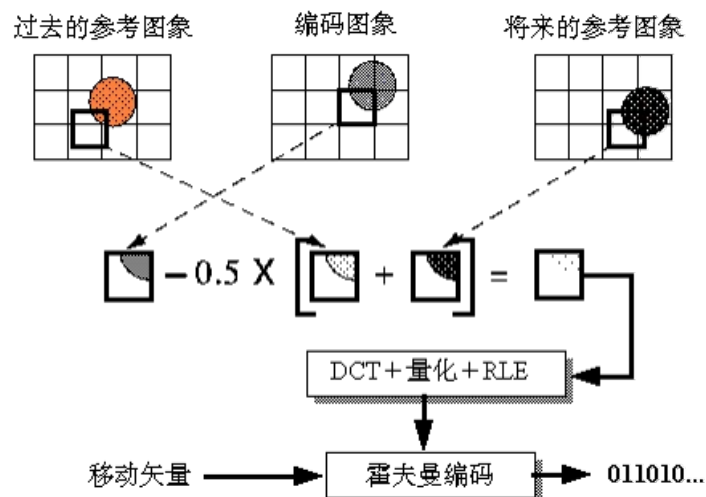
- 基於幀的壓縮算法
- MPEG-1標準只支持非交錯式視頻
- MPEG-1系統: 規定電視圖像數據, 聲音數據及其他相關數據的同步
- MPEG-1視頻: 規定電視數據的編碼和解碼
- MPEG-1音頻: 規定聲音數據的編碼和解碼
- MPEG-1一致性測試: 說明如何測試比特數據流和幾碼是否滿足MPEG-1規定
- MPEG-1軟件模擬



■ I幀: 內幀



■ P幀: 單向預測幀



■ B幀: 雙向預測幀

- 應用: VCD, MP3, 局域網視頻傳輸

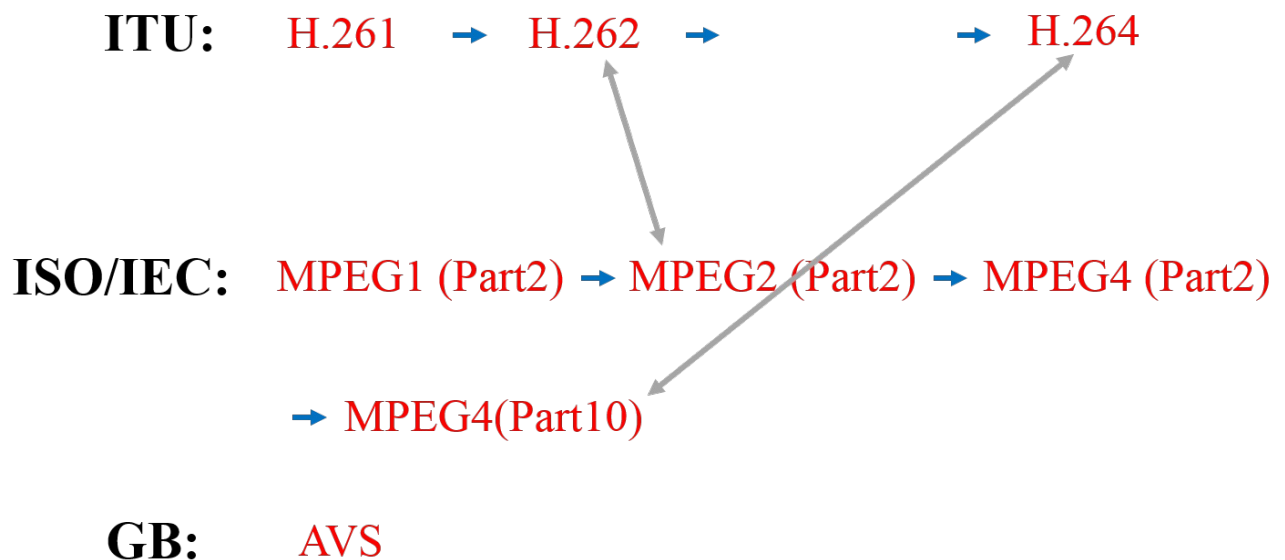
■ MPEG-2: 運動圖像及伴音通用編碼(1994)

- 基於塊的壓縮算法, 使用圖層的方法
- 可以說是MPEG-1的一個擴充, 相比MPEG-1新添加的功能:

- 能夠在很寬的範圍內對不同分辨率和不同輸出比特率的圖像信號有效的進行壓縮
 - 處理隔行掃描的視頻信號能力
 - 多樣化的取樣模式: 4:2:0, 4:2:2, 4:4:4
 - 可伸縮的視頻編碼模式:
 - 編碼時可以在圖像質量和數據速率之間作出調整
 - 解碼時只對一部分進行解碼, 從而獲得不同質量的重構圖像
 - 引入配置和等級的概念
 - 應用: DVD, DVB, HDTV
 - MPEG-3: 合併到HDTV(1992)
 - MPEG-4: 甚低速率音視頻編碼(1999)
 - 基於面向對象的編碼
 - 碼率覆蓋: 5kbps~10Mbps
 - 目的是為視聽(audio-visual)數據編碼和交互播放開發算法和工具
 - 基本碼流: 視聽對象, 場景描述信息, 對象描述子等
 - 合成: 將目標放置在同一顯示空間的行為
 - 重現: 將來自同一顯示空間的對象傳輸到專門的播放設備的行為
 - 場景描述語言: BIFS(Binary Format for Scenes)
 - BIFS使用的是參數方法, 場景描述由一些帶有屬性和其他信息(包括事件源和目標)的節點構成的編碼層次(樹)組成
 - 樹中節點表示特定AV對象(媒體對象)
 - 算法核心是基於內容的編碼和解碼, 即抽取單獨的物理對象, VO (Video Object)
 - Texture Coding: 用於編碼VOP中灰度(或色度)的變化及樣式
 - Shape Coding: 用於編碼VOP的形狀
 - Sprite Coding: 用於區分前景物體和背景
 - 2D Mesh Geometry Coding: 將圖形分成若干網格
 - 3D Model-Based Coding:
 - Face Object Coding and Animation
 - Body Object Coding and Animation
- | | |
|----------------------------|--|
| Video-object Sequence (VS) | 1. Video-object Sequence (VS) : delivers the complete MPEG-4 visual scene, which may contain 2-D or 3-D natural or synthetic objects. |
| Video Object (VO) | 2. Video Object (VO) : a particular object in the scene, which can be of arbitrary (non-rectangular) shape corresponding to an object or background of the scene. |
| Video Object Layer (VOL) | 3. Video Object Layer (VOL) : facilitates a way to support (multi-layered) scalable coding. A VO can have multiple VOLs under scalable coding, or have a single VOL under non-scalable coding. |
| Group of VOPs (GOV) | 4. Group of Video Object Planes (GOV) : groups Video Object Planes together (optional level). |
| Video Object Plane (VOP) | 5. Video Object Plane (VOP) : a snapshot of a VO at a particular moment. |
- 基於運動補償的VOP編碼步驟:
 - 運動估計
 - 基於運動補償的預測
 - 預測誤差的編碼
 - 宏塊分類:
 - 內部宏塊: 完全處於VOP內的宏塊
 - 邊界宏塊: 跨越了邊界的宏塊
 - 外部宏塊: 完全處於VOP外的宏塊
 - 應用: 多用於因特網視頻傳輸, 流媒體應用, 交互式視聽服務及遠程控制等
 - MPEG-7: 多媒體內容描述接口標準(2001)
 - 使用MPEG-7描述工具生成的描述與數據本身結合在一起, 從而允許對數據進行快速有效的搜索, 過濾以得到用戶感興趣的信息

- MPEG-21: 多媒體應用框架標準(2003)
 - 規定大範圍的網絡及設備中的多媒體定義, 標識, 描述, 管理及保護等.
- MPEG-A: A suite of standards specifying application formats that involve multiple MPEG and, where required, non MPEG standards(2007)
- MPEG-B: MPEG systems technologies. (ISO/IEC 23001) (e.g., Binary MPEG format for XML, Fragment Request Units, BSDI and others)(2006)
- MPEG-C: MPEG video technologies. (ISO/IEC 23002) (e.g., Accuracy requirements for implementation of integer-output 8x8 inverse discrete cosine transform and others)(2006)
- MPEG-D: MPEG audio technologies. (ISO/IEC 23003) (e.g., MPEG Surround, SAOC-Spatial Audio Object Coding and USAC- Unified Speech and Audio Coding)(2007)
- MPEG-E: Multimedia Middleware. (ISO/IEC 23004) (a.k.a. M3W) (e.g., Architecture, Multimedia application programming interface (API), Component model and others)(2007)
- MPEG-V: Media context and control. (ISO/IEC 23005) (a.k.a. Information exchange with Virtual Worlds) (e.g., Avatar characteristics, Sensor information, Architecture and others)(2011)
- MPEG-M: MPEG eXtensible Middleware (MXM). (ISO/IEC 23006) (e.g., MXM architecture and technologies, API, MPEG extensible middleware (MXM) protocols)(2010)
- MPEG-U: Rich media user interfaces. (ISO/IEC 23007) (e.g., Widgets)(2010)
- MPEG-H: High Efficiency Coding and Media Delivery in Heterogeneous Environments. (ISO/IEC 23008) Part 1 – MPEG Media Transport; Part 2 – High Efficiency Video Coding; Part 3 – 3D Audio.(2013)
- MPEG-DASH: Information technology – Dynamic adaptive streaming over HTTP (DASH). (ISO/IEC 23009) Part 1 – Media presentation description and segment formats(2012)

◦ AVS (中國標準)



Chapter10: Visual Coding in MPEG-4, Other MPEG Standards

Chapter11: Computer and Multimedia Networks

- OSI和TCP/IP的層次協議:

OSI	TCP / IP	
Application		FTP, Telnet, SMTP, HTTP, SNMP, etc.
Presentation	Application	
Session		
Transport	Transport	TCP (connection-oriented) UDP (connectionless)
Network	Internet	IPv4, IPv6, RSVP
Data link	Network access (LLC and MAC)	X.25, Ethernet, Token ring, FDDI, PPP/SLIP, etc.
Physical	Physical	10/100Base-T, 1000Base-T, Fiber Channel, etc.

- TCP(Transmission Control Protocol) - 基於有鏈接的 - 只適用於包交換網絡 - 依賴於IP層來把消息傳送到指定IP地址的電腦 - 提供消息報文分組, 錯誤檢測, 重傳, 包重新序列化及複用的功能 - UDP(User Datagram Protocol) - 面向無連接的, 數據通過一個數據報來傳遞 - 只提供服用和錯誤檢測功能(Checksum) - 源端端口是一個可選項, 因為沒有ACK - 比TCP要快得多, 但是並不可靠 - 適用於實時多媒體應用 - 要自己實現流量控制, 擁塞避免 - IP(Internet Protocol) - 兩個主要服務 - 包的尋址(Packet Addressing) - 包的分片(Packet Fragmentation) - 版本: - IPV4: 32 bit numbers - IPV6: 128 bit numbers - Multiplexing - FDM(Frequency Division Multiplexing) - 多個通道是通過頻率來區分的 - 模擬信號必須得調製, 使得信號攜帶單一頻率的通道 - 主要調製技術: - AM(Amplitude Modulation) - FM(Frequency Modulation) - PM(Phase Modulation) - QAM(Quadrature Amplitude Modulation) - WDM(Wavelength Division Multiplexing) - 通過波長來區分 - WDM的容量非常巨大, 數據率可達到每秒兆位級別 - 三種WDM: - CWDM(Coarse WDM): Multiplexor can put up to 8 Fiber channel connections, uses 1200~1700 nm, interval >20nm. - DWDM(Dense WDM): Employs densely spaced wavelengths so as to allow a larger number of channels than WDM. - WWDM(Wideband WDM): Allows the transmission of color lights with a wider range of wavelengths. - TDM(Time Division Multiplexing) - 通過時間片來區分 - 如果源數據是模擬信號, 必須先數字化然後轉換成PCM(Pulse Code Modulation, 脈衝編碼調製) - 兩種TDM: - Synchronous TDM: 每個信道都分配時間, 不管有沒有數據 - Asynchronous TDM: 只分配給有數據的信道 - TDM載波標準: - T1: a Synchronous TDM for 24 voice channels(23 for data, 1 for synchronization) - E1: ITU-T標準

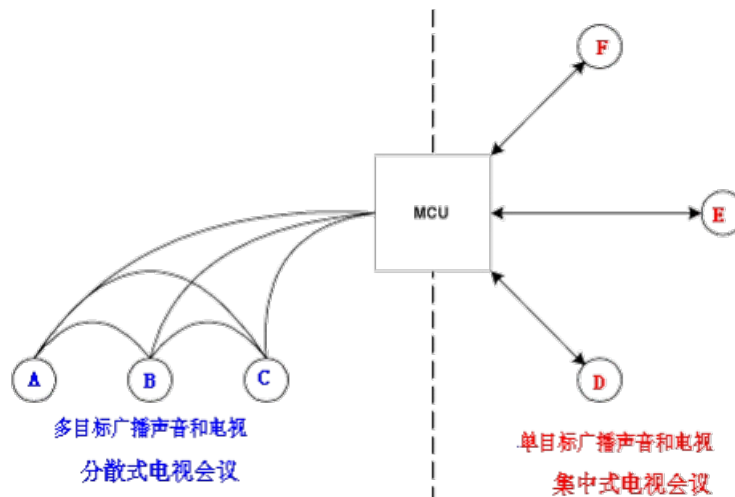
Format	Num of channels	Data Rate (Mbps)	Format	Num of channels	Data Rate (Mbps)
T1	24	1.544	E1	32	2.048
T2	96	6.312	E2	128	8.448
T3	672	44.736	E3	512	34.368
T4	4032	274.176	E4	2048	139.264
			E5	8192	565.148

- ISDN(Integrated Services Digital Network) - ITU-T制定(1980s) - 默認情況下, ISDN指(Narrowband ISDN)窄帶ISDN, 另外還有一個(Broadband ISDN)寬帶ISDN. 默認的包轉換技術是ATM(Asynchronous Transfer Mode) - 全雙工通道類型: - B(Bearer) channel: 64kbps, 用於數據傳送, 大部分情況用於電路交換網絡 - D(Delta) channel: 16kbps or 64kbps, 用於搭建, 控制和維護網絡 - ISDN採用Synchronous TDM來進行複用 - 可用的兩種介面: - Basic Rate Interface - Primary Rate Interface - SONET(Synchronous Optical NETwork) - Bellcore指定, 一個光釐標準 - SONET採用Synchronous TDM來進行複用 - SONET-SDH(Synchronous Digital Hierarchy) - ITU-T指定的與SONET相似的標準 - ADSL(Asymmetric Digital Subscriber Line, 非對稱用戶線路) - 採用離散多音頻技術, 將電話線路0~1.1MHz頻段劃分成256個頻寬為4.3kHz的子頻帶(採用頻分複用FDM, 可以避免相信道之間相互的干擾) - 4kHz以下的頻段用於傳送POTS(傳統電話業務) - 20kHz~138kHz的頻段用於傳送上行信號 - 138kHz到1.1MHz的頻段用於傳送下行信號 - 非對稱指的是上行帶寬和下行帶寬的不對稱 - LAN(Local Area Network) - 局域網指局限於一個地理位置的網絡, 通常是相關的幾個基站之間 - LAN的數據鏈路層含有兩子層 - MAC(Medium Access Control) Layer - 綁定IP地址和物理地址 - LLC(Logical Link Control) Layer - 流量和差錯控制 - Ethernet: - 包交換網絡, 一種著名的局域網 - CSMA/CD(Carrier Sense Multiple Access with Collision Detection, 載波偵聽多路訪問帶衝突檢測) 解決了MAC的問題 - 當多站同時發消息到信道時則會產生衝突 - 通過監聽等待沒有站發消息時才發消息, 從而解決衝突問題 - WAN(Wide Area Network) - 廣域網指穿越國家或城市的網絡 - 使用到的包交換技術: - Circuit Switching(電路交換) - 端到端的電路 - 通過採用FDM或Synchronous TDM來解決多用戶或不同數據率的問題 - Packet Switching(包交換) - 數據被分成包傳遞 - 兩種交換和定位包的方式: - datagram(數據報) - virtual circuit(虛電路) - MAN(Metropolitan Area Network) - 城域網有時指介於局域網和廣域網之間的網絡 - Access Network

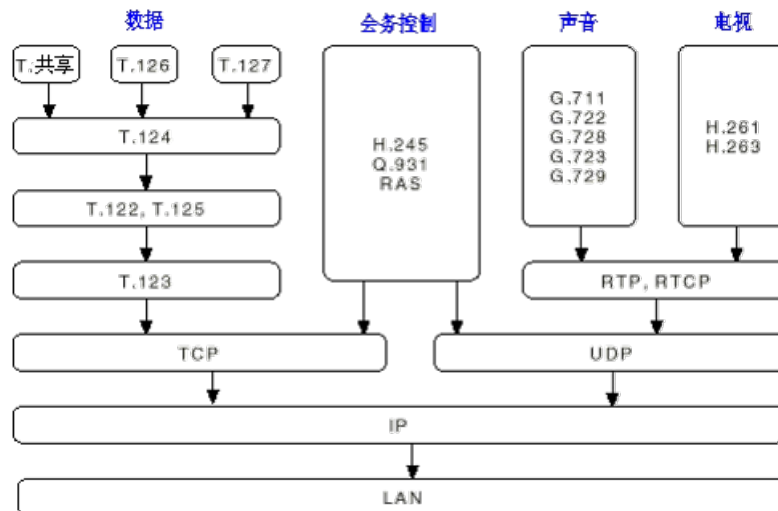
- 指的是終端用戶連接到核心網絡之間的部分 - **Hybrid Fiber-Coax(HFC) Cabel Network** - 通過**Optical Network Units(ONU)**來連接到核心網絡 - 通過銅軸電纜服務終端用戶 - 潛在問題在於銅軸電纜部分的噪音干涉, 而且上行信道存在隱私和安全問題 - **Fiber To The Crub(FTTC)** - 通過雙絞線或銅軸電纜服務終端用戶 - **Fiber To The Home(FTTH)** - 直接連接到核心網絡, 提供有高帶寬 - 存在成本較高的問題 - **Terrestrial Distribution - Satellite Distribution** - 多媒體應用的分佈式要求: - 公共信息服務的要求 - 各種網絡綜合統一化 - 娛樂事業的發展 - 全球信息共享與協同工作 - 多媒體的網絡要求: - 帶寬: 數據量大, 數據率變化大 - 多媒體信息必須以壓縮形式傳送 - 帶寬需: **64kbps~2Mbps** - 多媒體通信量在上行和下行兩個方向可能是非對稱的 - 對連續媒體流的支持能力 - 面向連續比特流的通信 - 時基媒體與連續比特流**CBO(Continuous Bit Straeam Oriented)** - 恒定比特率**CBR(Constant Bit Rate)**媒體流 - 變比特率**VBR(Variable Bit Rate)**媒體流 - 基於延遲的通信類型 - 異步: 對報文分組的傳輸延遲沒有限制 - 同步: 對報文分組的傳輸延遲有一定限制 - 等時: 每個報文分組具有恒定的傳輸延遲 - 具有很強的實時性 - 具有很強的交互能力 - 多媒體的時空約束性 - **QoS(Quality of Service)**的決定因素: - **Data rate(數據率)** - **Latency(延遲)** - **Packet loss or error(包延遲/包錯誤)** - **Jitter(抖動)** - **Sync skew(同步扭曲)** - **Multimedia over IP** - 傳播方式 - **Multicast 組播(多播)**: 傳遞給特定的多個人 - 用**TTL(time-to-live)**來解決包長時間在網絡上流動的問題 - 組播生成樹: 用於描述服務器和接收機之間構成的播放與接收關係 - 組播步驟: - 組播源把數據或組播通知發送給所有路由 - 不想參加組播的終端逆向發一個刪除消息 - 刪除沒有成員的分支和不在最短路徑樹上的分支 - 使用聯結和刪除功能改變成員之間關係 - **Broadcast 廣播**: 傳遞給其他所有人 - **Unicast 單波**: 傳遞給特定的一個人 - **RTP(Real-time Transport Protocol, 實時傳輸協議)** - 提供端對端傳輸服務的實時傳輸協議 - 用於支持在單播和組播網絡服務中傳輸實時數據 - 通常運用於**UDP**上, 通過創建時間戳和排序機制來保證數據有序性 - 組成: - 有效載荷類型域 - 順序號域 - 時間戳域 - 同步源標識符域 - 其它 - **RTCP(Real-time Control Protocol, 實時傳輸控制協議)** - 用於監視服務質量和傳輸有關與會者的信息 - 通過向服務器發送數據傳輸質量的反饋來監控**QoS**變化 - **RTP**和**RTCP**結合用 - **RTP**和**RTCP**數據包通過不同端口發送到相同的**IP**地址 - **RSVP(Resource Reservation Protocol, 資源保留協議)** - 允許應用程序保留因特網上的資源 - **RTSP(Real-time Streaming Prtocol, 實時流協議)** - 為單播和組播上的流式多媒体应用提供播放性能 - 支持不同廠家提供的客戶機和服務器之間的協同工作能力 - **RTSP**和**RSVP**結合用

Chapter12: Multimedia Communication

- 多媒體通信是在位於不同地理位置的參與者之間召開的一種會議或者進行的交流
- 多媒體通信通過局域網, 廣域網, 內聯網, 因特網或者電話網來傳輸壓縮的數字圖像和聲音信號
- 視頻會議系統標準
 - **H.320**
 - 定義通信的建立, 數字電視圖像和聲音壓縮算法
 - 運行在綜合業務數字網
 - 在**56kbps**傳輸率的通信信道上支持幀速率較低的電視圖像
 - 在**1.544Mbps**的信道(**T1**信道)上可以傳輸**CIF**格式的滿幀速率電視圖像
 - 包含了5各部分, 15個標準
 - **H.323**: 描述的是無**QoS**保證的**LAN**多媒體通信終端, 設備和服務
 - 用於局域網上的桌面電視會議
 - 基於信息包交換的多媒體通信系統
 - 組成:
 - 視頻會議終端
 - 網關
 - 信道(網絡)
 - 多點控制單元(MCU)
 - 控制管理軟件
 - 安全保密系統
 - 定義了一個叫做多點控制單元的部件
 - 把通信終端, 網關, 會務器或者多點控制單元叫做終端
 - 網關
 - 提供面向媒體的功能, 如傳送聲音和電視圖像數據以及接收數據包等
 - 擔負線路交換網絡和信息包交換網絡之間的通信
 - 提供異種網絡之間的連通性
 - 會務器
 - 基本功能:
 - 地址轉換: 使用一種可由註冊消息更新的轉換表, 把別名地址轉換成傳輸地址.
 - 准入控制: 使用准入請求/准入確認/准入拒絕消息, 對訪問局域網進行授權.
 - 帶寬控制: 使用帶寬請求/帶寬確認/帶寬拒絕消息, 以強制執行帶寬控制.
 - 區域管理: 管理所有已註冊的**H.323**端點
 - 選擇功能:
 - 呼叫控制信號傳輸方法
 - 呼叫授權
 - 帶寬管理
 - 呼叫管理
 - 提供面向服務的功能, 如身份驗證, 呼叫路由選擇和地址轉換等



- 多點控制器(MCU)



- 協議棧(Protocol Stack)

- 必須支持H.245標準

- 定義流程控制, 加密和抖動管理, 啟動呼叫信號
- 磋商要使用的終端特性和終止呼叫等過程

- 還需要支持Q.391標準, RAS標準和RTP/RTCP標準

◦ H.324

- 用於在公眾交換電話網上的桌面電視會議

● 因特網電話通訊(Internet Telephony)

◦ 因特網電話受到青睞的原因:

- 商業考慮
- IP普遍存在
- 技術成熟
- 向數據網轉移的趨勢

◦ VoIP(Voice on IP)

- 需要使用RTP, RTCP, MGCP(媒體網關控制協議), RSVP

■ SIP:

- 支持基於HTTP IP網絡的超輕量協議標準
- 借鑒了互聯網標準和協議的設計思想, 在風格上遵循簡練, 開放, 兼容和可擴展的原則, 在擴展性, 可伸縮性和複雜性方面都有一些優勢

■ H.323:

- 基於ISDN的Q.931和早期推薦的H系列協議的傳統電路交換方法
- 功能管理和可靠性方面比較完善
- 沿用傳統的電話信令模式, 其集中式控制模式便於管理

● 常用的無線通信網絡

◦ 無線通信網: 2G/3G

	ANSI-41 Core Network	Peak Data Rate R	Carrier Spectrum W
2G	cdmaOne (IS-95A)	14.4 kbps	1.25 MHz
2.5G	cdmaOne (IS-95B)	115 kbps	1.25 MHz
3G	cdma2000 1X	307 kbps	1.25 MHz
3G	cdma2000 1xEV-DO	2.4 Mbps	1.25 MHz
3G	cdma2000 1xEV-DV	4.8 Mbps	1.25 MHz
3G	cdma2000 3X	> 2 Mbps	5 MHz
	GSM MAP Core Network	Peak Data Rate R	Carrier Spectrum W
2G	GSM (TDMA)	14.4 kbps	1.25 MHz
2.5G	GPRS (TDMA)	170 kbps	1.25 MHz
3G	EDGE (TDMA)	384 kbps	1.25 MHz
3G	WCDMA	2 Mbps	5 MHz

- 無限局域網: IEEE 802.11
- 藍牙: 短距離無線通信

 [Go Top](#)

Aleen
more than a coder, more than a designer

