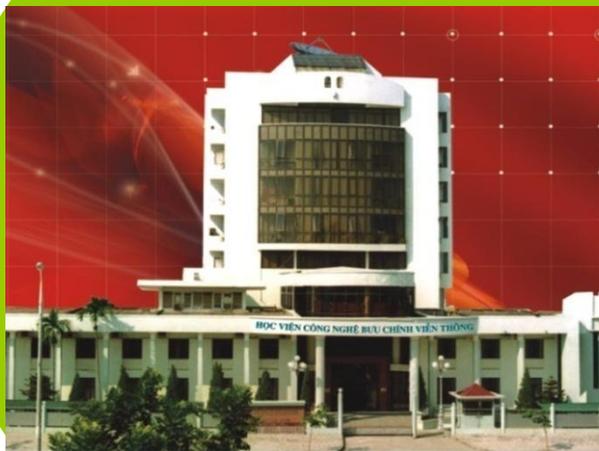




HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



BÀI GIẢNG MÔN

Xử lý âm thanh và hình ảnh

Giảng viên: LÊ NHẬT THĂNG

Điện thoại/ E-mail: thangln@ptit.edu.vn/0904342557

Bộ môn: Chuyên Mạch

Học kỳ 1: Năm 2010-2011

CHƯƠNG TRÌNH MÔN HỌC

- THỜI LƯỢNG: 60/48LT/6BT/6 thảo luận

- NỘI DUNG:

Chương 1: Giới thiệu chung 4 LT/ 0 BT/ 0 KT/0 TL

Chương 2: Kỹ thuật xử lý âm thanh 12 LT/ 02 BT/ 0 KT/02 TL

Chương 3: Kỹ thuật xử lý ảnh 20 LT/ 02BT/ 0 KT/02TL

Chương 4: Các chuẩn mã hóa âm thanh hình ảnh trong truyền thông đa phương tiện 12 LT/ 02BT/ 0 KT/ 02 TL

TÀI LIỆU THAM KHẢO

- [1] Wai C. Chu, *Speech Coding Algorithms- Foundation and Evolution of Standardized Coders*, John Wiley & Sons, 2003.
- [2] R. C. Gonzalez, R. E. Woods , *Digital Image Processing*, Prentice Hall, 2nd Edition, 2001.
- [3] Bài giảng xử lý âm thanh và hình ảnh

Chương 1: Giới thiệu chung

1.1 Các khái niệm lý thuyết cơ sở

1.2 Vai trò của xử lý âm thanh hình ảnh trong truyền thông đa phương tiện

1.1 Các khái niệm lý thuyết cơ sở

- Một số kiến thức cơ bản cần chuẩn bị: hệ tuyến tính, ma trận và vectơ, xác suất; lập trình Matlab
- Đặc trưng cơ bản của tín hiệu: tương tự, số; Nguyên tắc biến đổi A/D; Ưu điểm của tín hiệu số;
- Cần phân biệt giữa khái niệm xử lý (processing) tín hiệu âm thanh và hình ảnh liên quan đến rất nhiều lĩnh vực, kỹ thuật và nén (compression). Nén chỉ là một phần của xử lý tín hiệu
- Phân biệt rõ khái niệm nén hay còn gọi là mã hóa nguồn và khái niệm mã hóa kênh; khái niệm entropy; mã hóa không có tổn thất (lossless) và mã hóa có tổn thất (lossy)

Ưu điểm của tín hiệu số

- Tín hiệu chỉ nhận hai giá trị 0,1
- Đơn giản hóa việc hợp nhất các dịch vụ viễn thông vào một mạng hợp nhất
- Các loại dịch vụ khác nhau có thể sử dụng một phương tiện chung nhất
- Giá thành thiết bị kỹ thuật số ngày càng giảm , chất lượng tốt, giá thành rẻ
- Nhu cầu về dịch vụ thông tin ngày càng tăng nhanh và đa dạng hóa: ví dụ Điện thoại thấy hình, hội nghị video tốc độ thấp, ảnh tĩnh, ảnh động, âm thanh chất lượng cao, đa phương tiện, truyền hình chất lượng cao

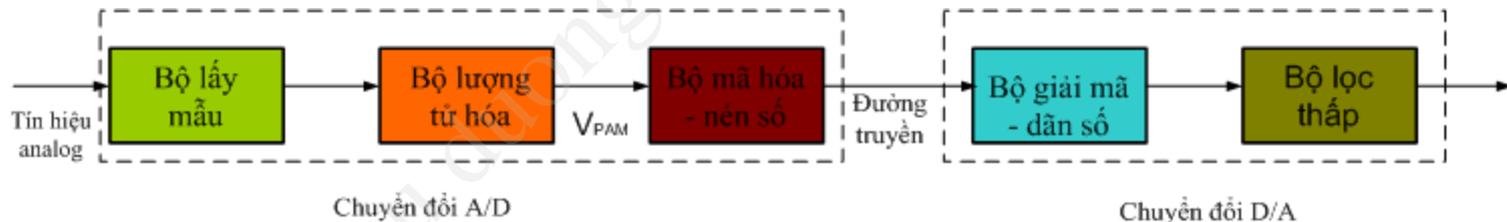
Nguyên tắc biến đổi A/D

- Điều xung mã PCM
- Điều xung mã vi sai DPCM
- Điều chế Delta
- Điều chế Delta thích ứng ADM

Điều xung mã PCM

Điều xung mã PCM được đặc trưng bởi 3 quá trình :

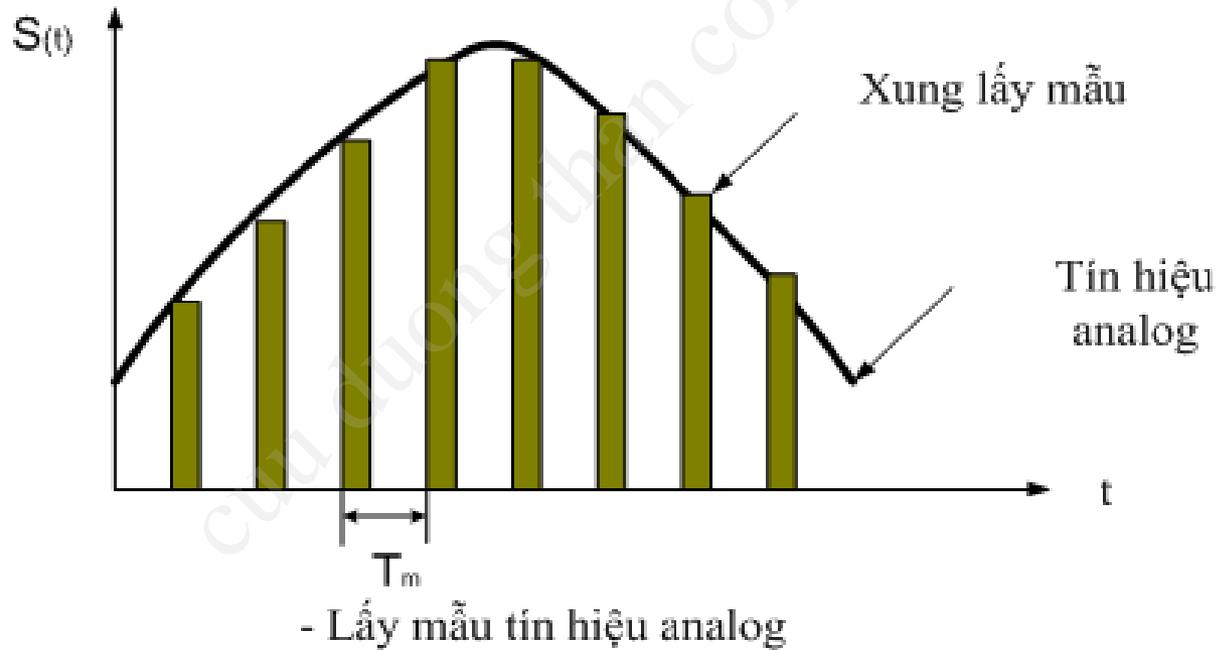
- Lấy mẫu
- Lượng tử hóa
- Mã hóa



- Sơ đồ khối quá trình chuyển đổi A/D và D/A trong hệ thống PCM

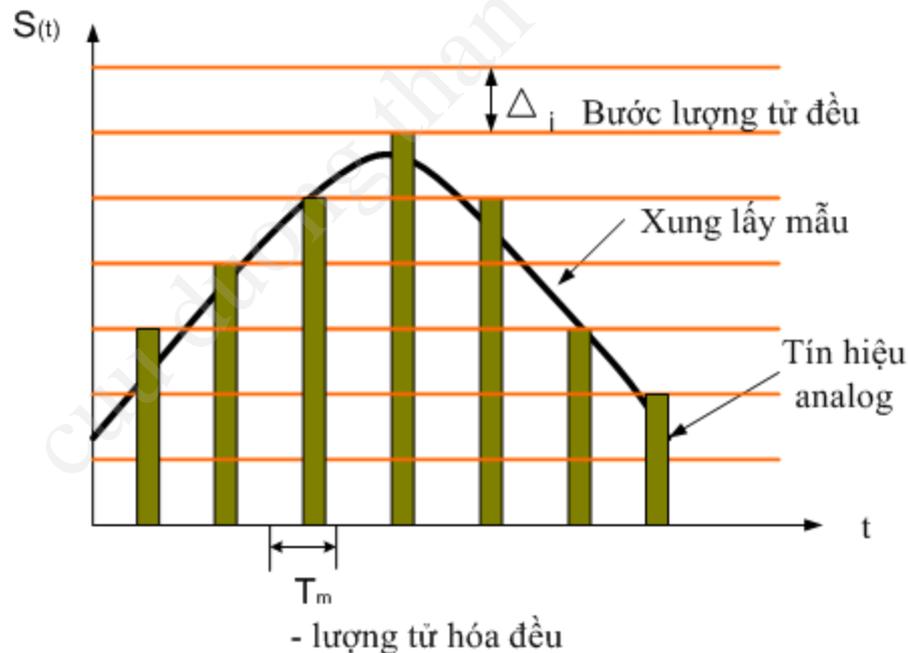
Lấy mẫu

Khái niệm : Lấy mẫu là quá trình rời rạc hóa tín hiệu theo thời gian



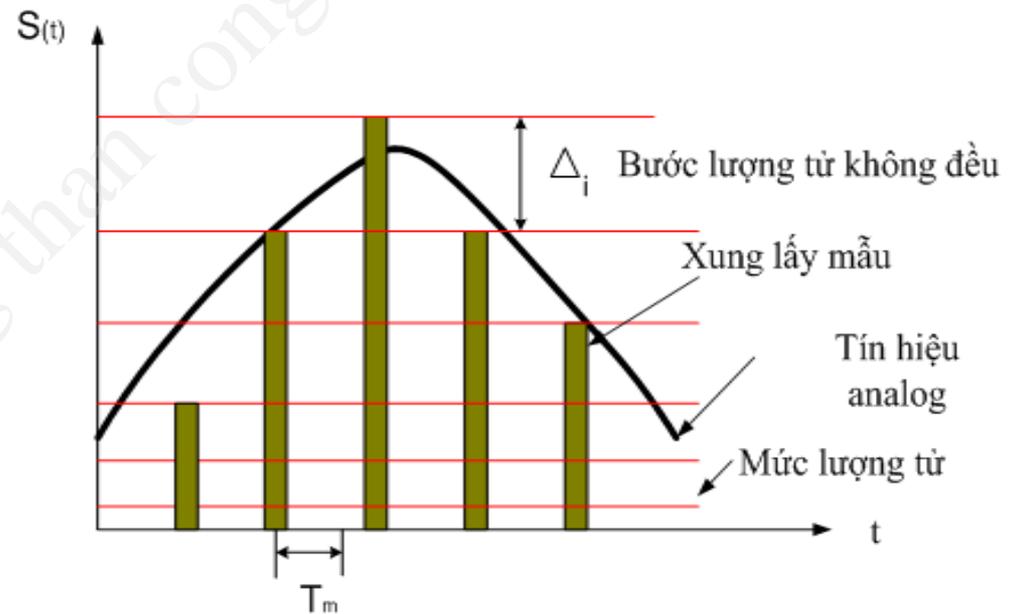
Lượng tử hóa

1. Lượng tử hóa đều: Chia biên độ xung lấy mẫu thành các khoảng đều nhau, mỗi khoảng là một bước lượng tử đều Δ



Lượng tử hóa

2. Lượng tử hóa không đều : Chia biên độ xung lấy mẫu thành các khoảng không đều nhau theo nguyên tắc khi biên độ xung lấy mẫu càng lớn thì độ dài bước lượng tử càng lớn.

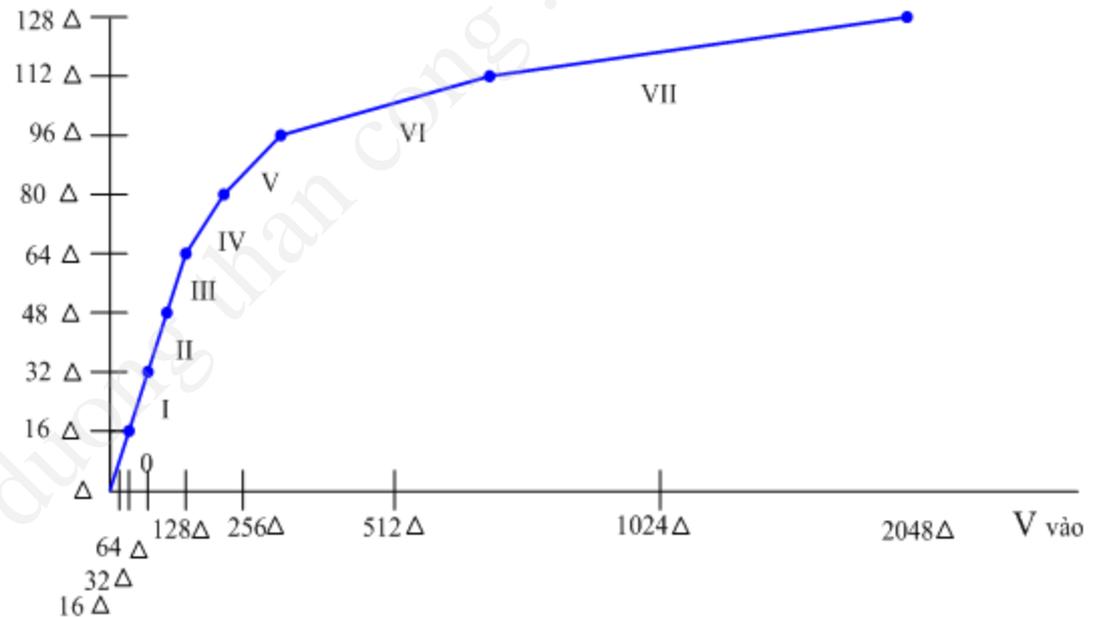


- Lượng tử hóa không đều

Mã hóa

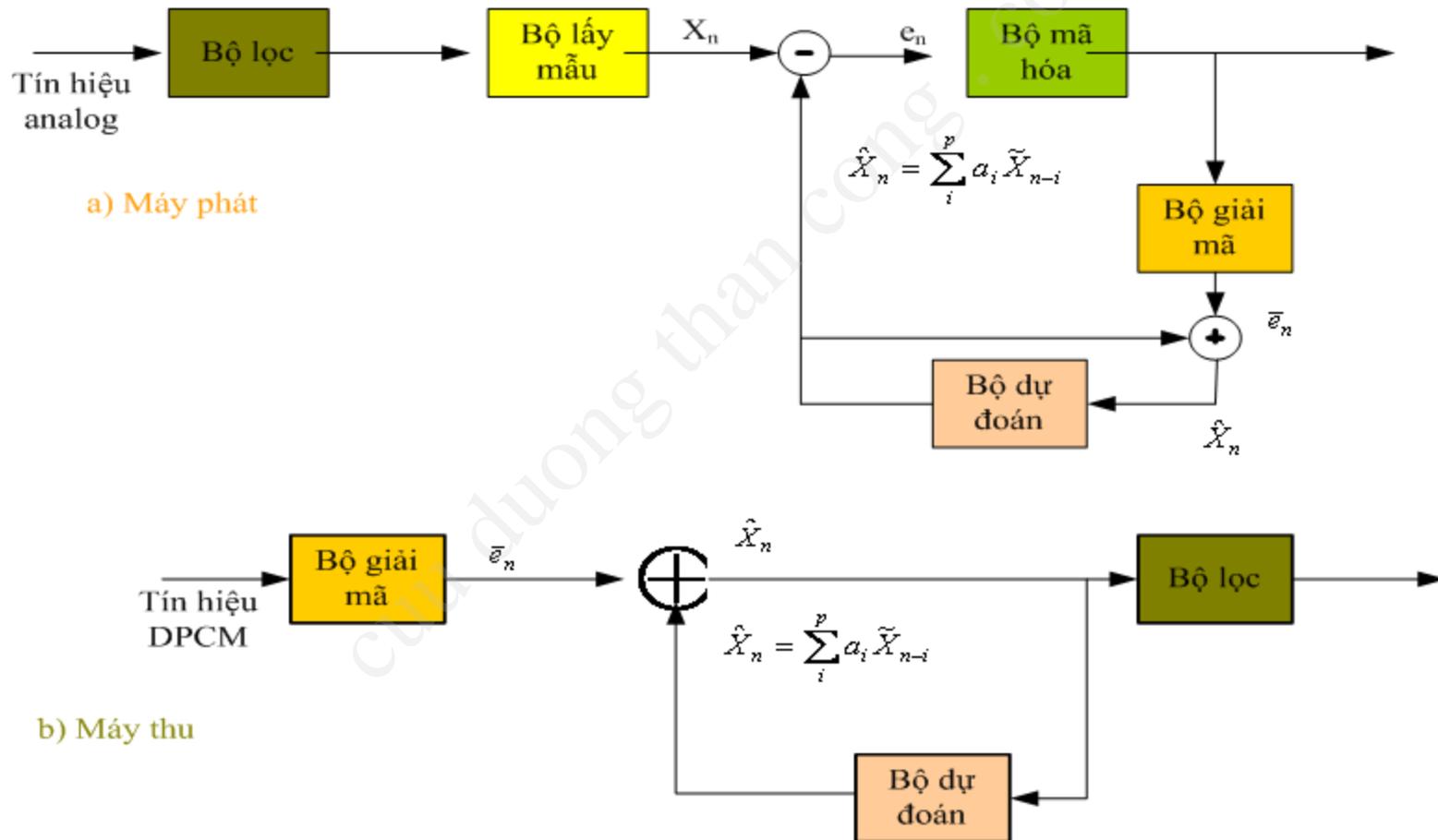
1. Chức năng:
Chuyển đổi biên độ
xung lượng tử
thành một từ mã 8
bít.

2. Đặc tính bộ mã
hóa $A=87,6/13$



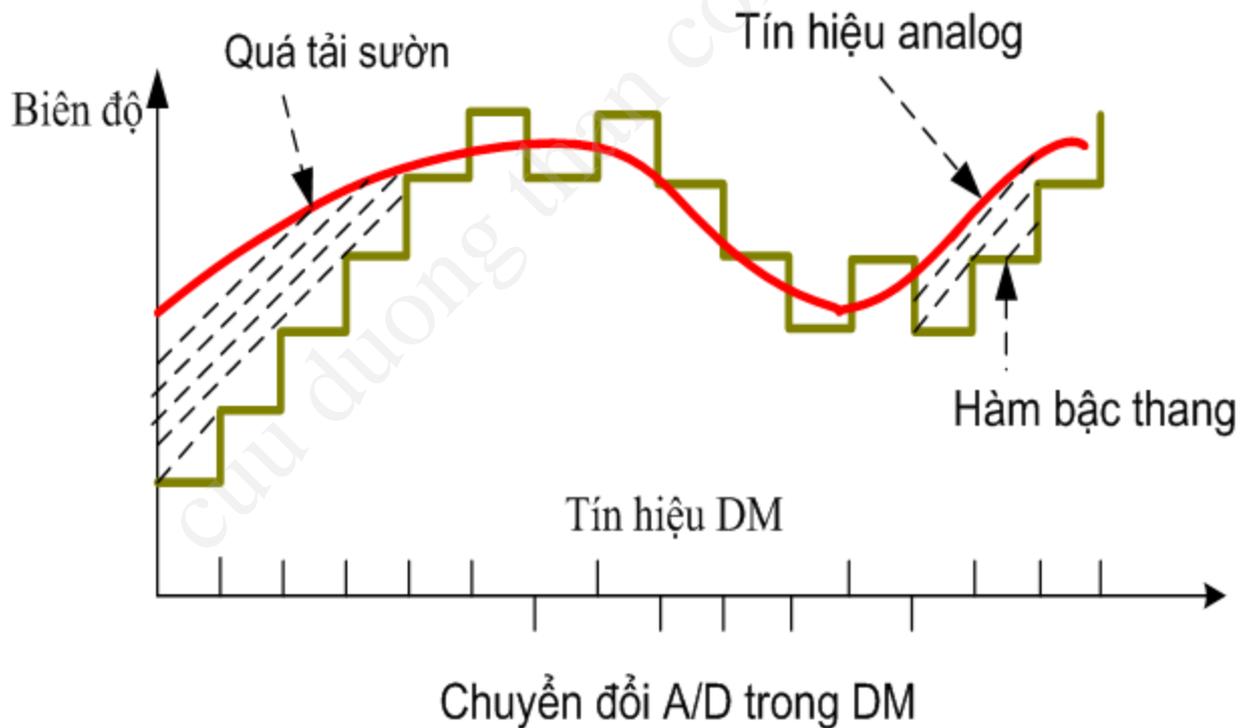
- Nhánh dương đặc tính biên độ bộ mã hóa - nén số $A = 87,6/13$

Điều xung mã vi sai DPCM

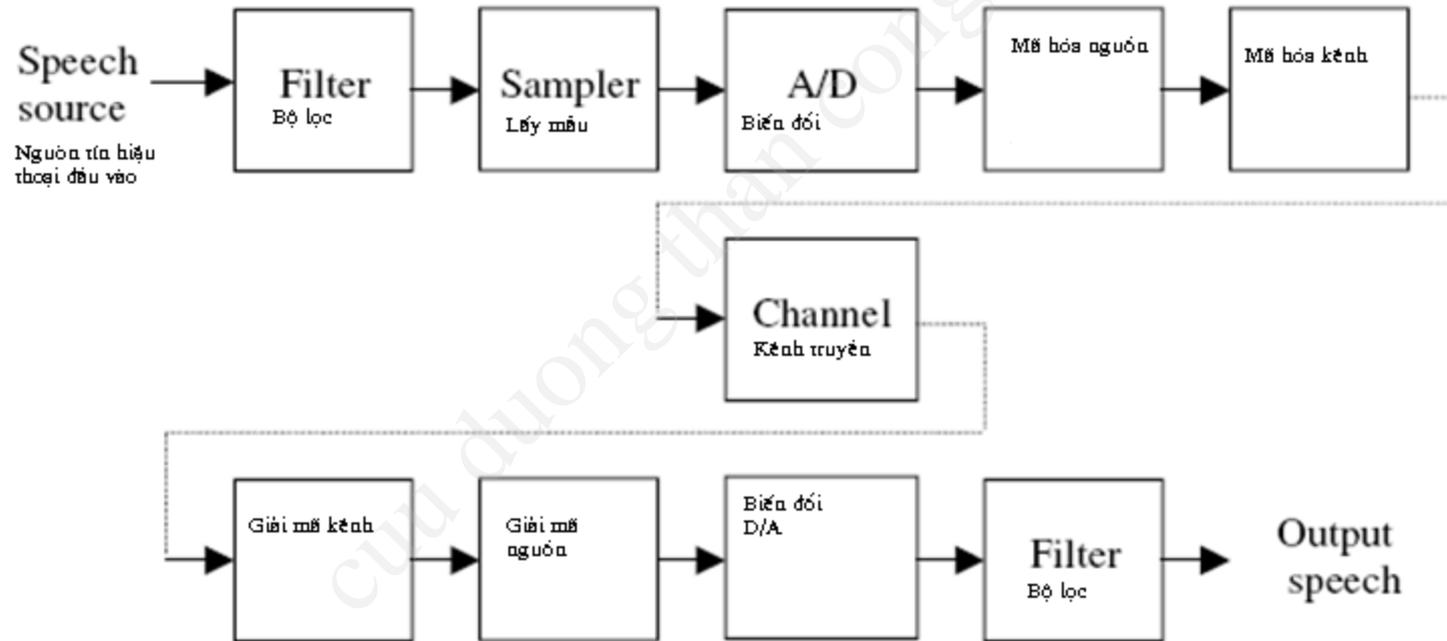


- Sơ đồ khối máy phát (a) và máy thu (b) DPCM

Điều chế Delta



Cấu trúc hệ thống mã hóa thoại



Thoại đầu ra

Tiếp

- Mã hóa nguồn: Mục đích là lấy dữ liệu nguồn và thu nhỏ chúng bằng cách loại bỏ những phần dư thừa không cần thiết còn tồn tại trong nguồn, để lại phần nguồn với số lượng bít ít hơn, nhưng nhiều tin tức.
- Mã hóa kênh: Là tìm ra những mã có thể truyền thông nhanh chóng chứa đựng nhiều mã ký hợp lệ và có thể sửa lỗi hoặc ít nhất phát hiện lỗi (Mã khối tuyến tính: mã chẵn lẻ, mã tuần hoàn, Mã kết hợp...).

1.2 Vai trò của xử lý âm thanh hình ảnh trong truyền thông đa phương tiện

- Xu hướng phát triển của viễn thông: các nhu cầu về dịch vụ và các hạn chế của công nghệ truyền dẫn, chuyển mạch liên quan để thấy được vai trò của xử lý âm thanh và hình ảnh;
- Các chuẩn nén âm thanh và hình ảnh đang được sử dụng phổ biến trong truyền thông: G711, G729; JPEG; MPEG; H264...

ĐÁNH GIÁ VỀ HIỆN TRẠNG MẠNG VIỄN THÔNG VIỆT NAM

- Môi trường và khả năng phát triển mạng Viễn thông Việt Nam là rất rộng mở
- Gồm nhiều mạng cung cấp dịch vụ riêng lẻ
- Các mạng này gần như độc lập với nhau

ĐÁNH GIÁ VỀ HIỆN TRẠNG MẠNG VIỄN THÔNG VIỆT NAM

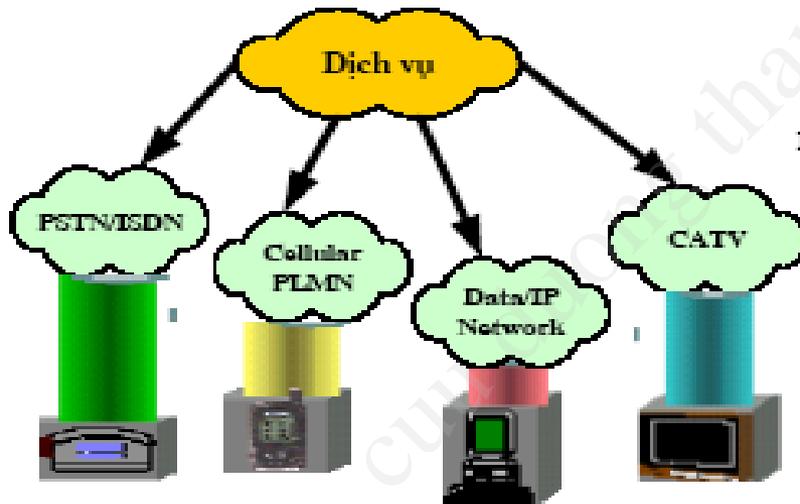
- ❖ Chỉ truyền được các dịch vụ độc lập tương ứng với từng mạng.
- ❖ Thiếu mềm dẻo, kém hiệu quả trong việc vận hành, bảo dưỡng cũng như sử dụng tài nguyên trong mạng
- ❖ Kiến trúc tổng đài độc quyền làm cho các nhà khai thác gần như phụ thuộc hoàn toàn vào nhà cung cấp tổng đài
- ❖ Các tổng đài chuyên mạch kênh đã khai thác hết năng lực và lạc hậu với nhu cầu của khách hàng
- ❖ Sự bùng nổ lưu lượng thông tin đã bộc lộ sự kém hiệu quả của chuyên mạch kênh TDM.

Động lực thúc đẩy sự phát triển của viễn thông

- Công nghệ điện tử phát triển, với sự tích hợp cao, giá thành thấp
- Sự phát triển của kỹ thuật số
- Sự cạnh tranh giữa các nhà cung cấp dịch vụ viễn thông toàn cầu
- Công nghệ truyền dẫn quang và chuyển mạch có những tiến bộ vượt bậc
- Mạng Internet phát triển
- Nhu cầu trao đổi thông tin ngày càng tăng và đa dạng hóa đặc biệt là các dịch vụ đa phương tiện.

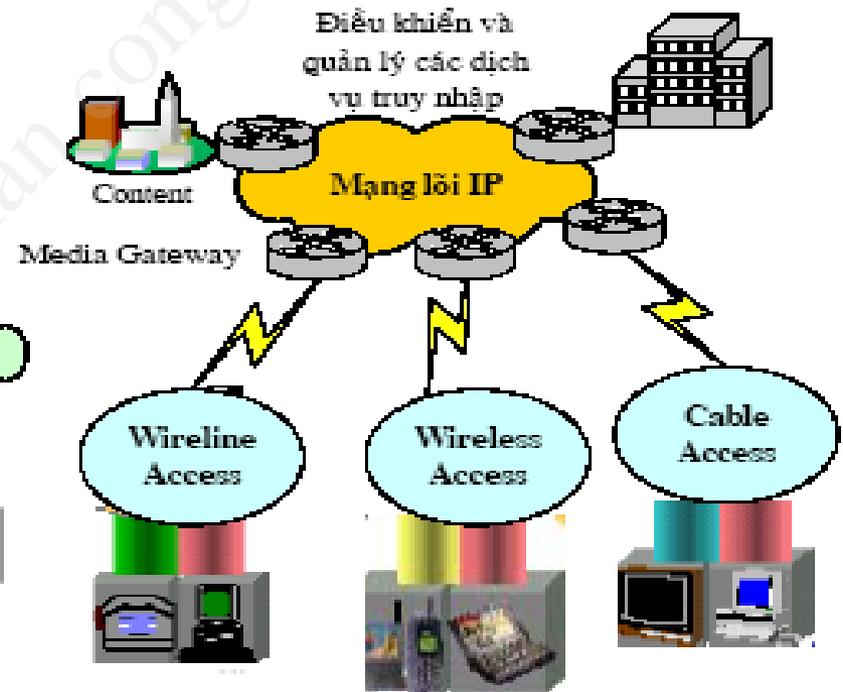
Xu hướng phát triển của mạng

Hiện tại
Các mạng dịch vụ riêng lẻ



Các mạng truy nhập, truyền dẫn, chuyển mạch riêng lẻ

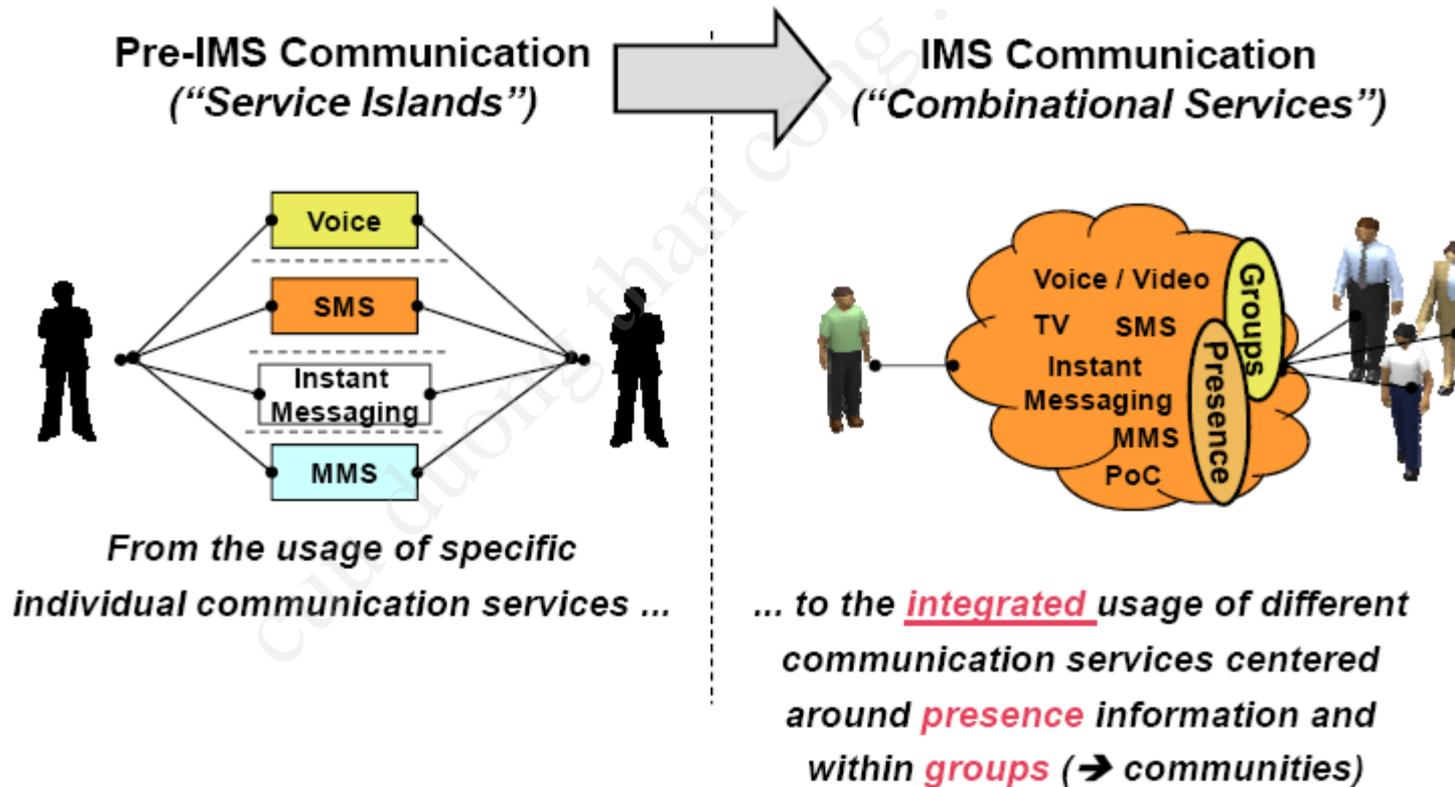
Tương lai
Mạng đa dịch vụ



Liên mạng trên cơ sở IP

Xu hướng phát triển của mạng

IMS integrates different Communication Services



Yêu cầu của khách hàng

- Muốn được cung cấp các dịch vụ có tốc độ bất thay đổi
- Dịch vụ có chất lượng cao, giá thành thấp
- Dịch vụ đáp ứng tính thời gian thực
- Dịch vụ đa phương tiện
- Để đáp ứng yêu cầu của khách hàng đối với các dịch vụ đa phương tiện thì xử lý âm thanh và hình ảnh đóng một vai trò rất quan trọng

Mã hóa thoại tốc độ thấp

- Dịch vụ thoại hiện chiếm một thị phần lớn trong các dịch vụ viễn thông.
- Mặc dù với sự phát triển của công nghệ truyền thông qua cáp quang đã làm cho băng thông không còn là vấn đề lớn trong giá thành của các cuộc gọi truyền thống. Tuy nhiên, băng thông trong các cuộc gọi đường dài, các cuộc gọi quốc tế, các cuộc gọi qua vệ tinh hay các cuộc gọi di động thì cần phải duy trì băng thông ở một mức nhất định.
- Vì vậy việc giảm băng thông thoại xuống dưới 64kbps là cách tốt nhất. Do đó việc phát triển các bộ mã hóa thoại tốc độ thấp là rất cần thiết, giúp giảm thiểu số lượng tín hiệu cần truyền đi trên đường truyền nhưng vẫn đảm bảo chất lượng của cuộc gọi.

Hướng dẫn ôn tập chương 1

1. Phân biệt các loại tín hiệu (liên tục, rời rạc).
2. Quá trình số hóa tín hiệu tương tự
3. Những lợi ích và ứng dụng của nén dữ liệu trong truyền thông đa phương tiện
4. Một tín hiệu hình Sin có biên độ 5V cần được biến đổi thành dạng số sao cho nhận được tỷ số tín hiệu trên tạp âm lượng tử hóa không thấp hơn 25 dB. Yêu cầu cần bao nhiêu bước lượng tử hóa như nhau và cần có bao nhiêu bit để mã hóa mỗi thành phần rời rạc.
5. Giả sử một tín hiệu có phân bố đều (uniform), được lượng tử hóa đều 256 mức, có tỷ số S/N là 18dB. Nếu muốn tăng tỷ số S/N của tín hiệu thành 30dB thì số mức lượng tử sẽ phải là bao nhiêu?

Chương 2

Xử lý âm thanh

Đặc tính âm thanh

- Âm thanh trong thế giới tự nhiên về bản chất là những sóng âm được tạo ra từ dao động của vật thể và được truyền đi trong một môi trường truyền âm nhất định.
- Âm thanh của tiếng nói, tương tự, là những sóng âm được tạo ra từ dao động của các bộ phận trong bộ máy phát âm và được truyền đi trong môi trường truyền âm.

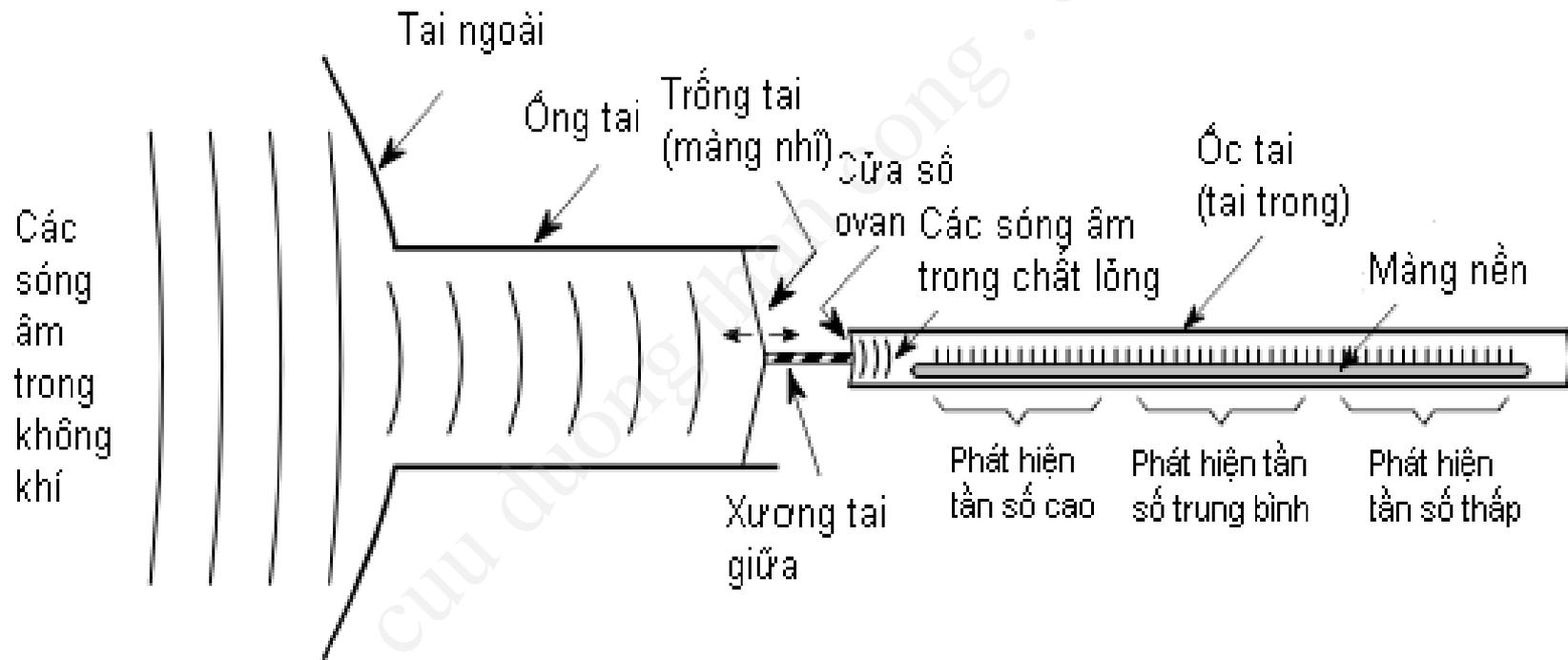
2.1 Các đặc trưng vật lý của âm thanh

- **Độ cao (hauteur /pitch):** do tần số dao động của dây thanh và/hoặc của các bộ phận khác trong bộ máy phát âm quyết định. Tần số dao động (số chu kì dao động trong một giây) càng lớn thì âm thanh càng cao và ngược lại. Đơn vị để đo độ cao của âm thanh là Hertz (viết tắt là Hz).
- **Độ mạnh (intensité/intensity):** do biên độ dao động của vật thể quyết định. Biên độ dao động là trị số lớn nhất mà dao động đạt tới trong một nửa chu kì. Biên độ dao động càng lớn, âm thanh càng vang to và ngược lại. Đơn vị đo độ mạnh của âm thanh là decibel (viết tắt là dB).

Các đặc trưng vật lý của âm thanh

- **Độ dài (durée/length):** do thời gian dao động của vật thể quyết định.
- **Âm sắc (timbre):** phụ thuộc vào độ cao, độ dài và độ mạnh tham gia bổ sung vào các thành phần kết cấu của âm.

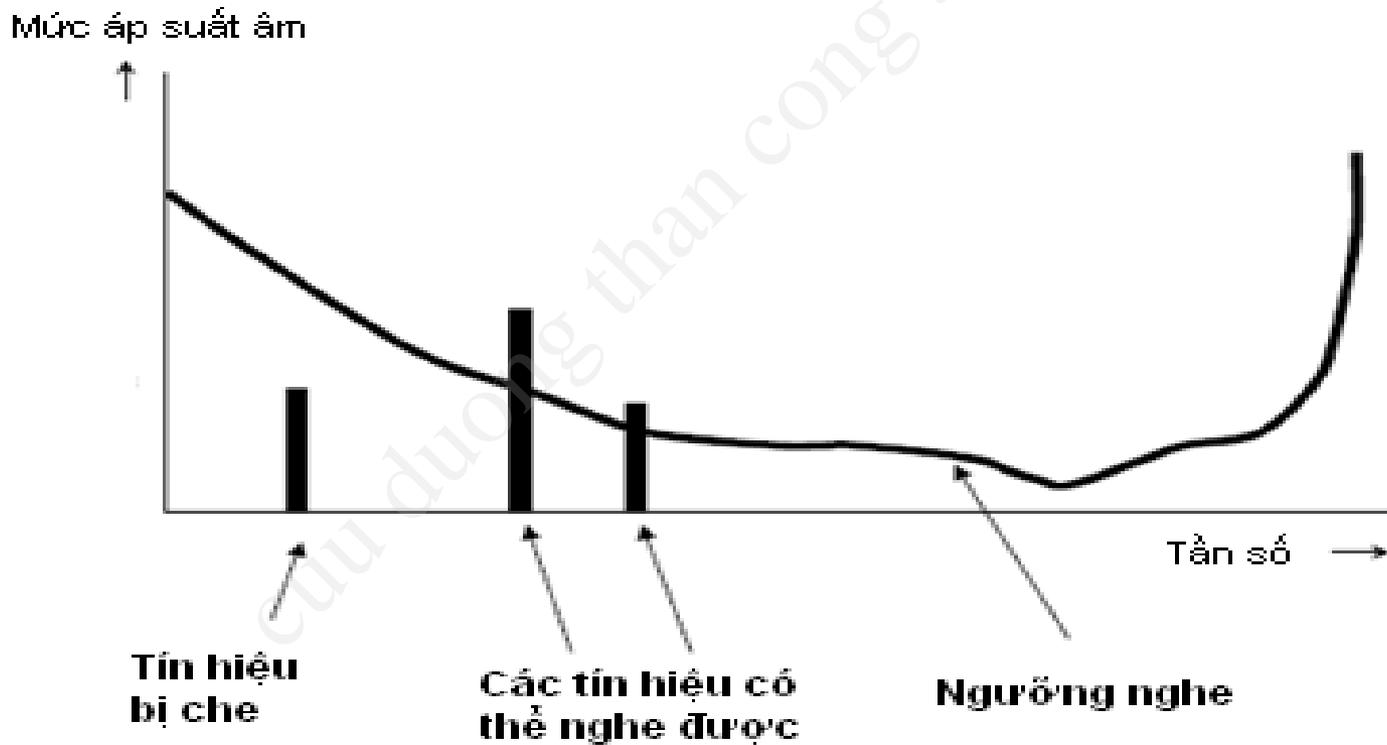
Các đặc điểm của hệ thống thính giác con người



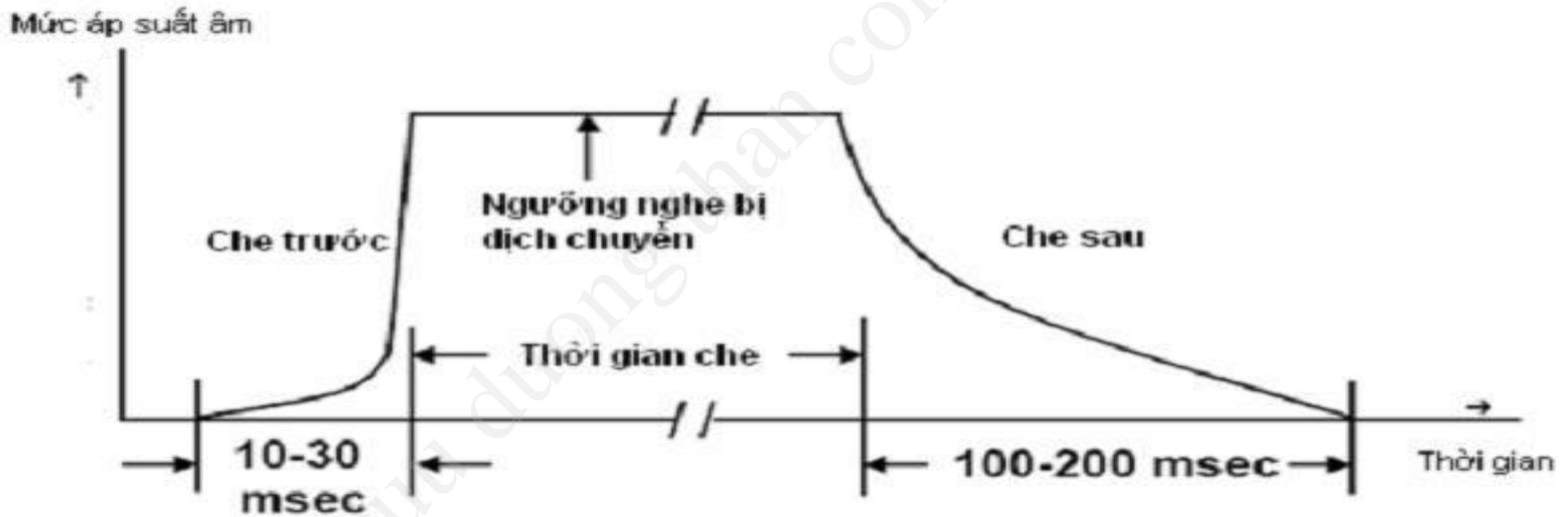
Sự cảm thụ của tai người đối với âm thanh

- Khoảng cách dải âm lớn nhất và yếu nhất mà tai con người có thể nghe là 120dB, tức là dải 1 triệu lần biên độ. Người nghe có thể phát hiện sự thay đổi độ ồn âm thanh khi tín hiệu bị thay đổi khoảng 1dB (biên độ thay đổi 12%)
- Nhận biết mức độ ồn liên quan mật thiết với công suất âm thanh theo bậc mũ 1/3.
 - Nếu tăng công suất âm lên 10 lần, người nghe nhận được rằng độ ồn tăng lên tầm 2 lần ($10^{1/3} \approx 2$ lần)
 - Dải nghe của con người thông thường từ 20Hz đến 20kHz, độ nhạy âm lớn nhất từ 1kHz đến 4kHz.
 - Khả năng xác định hướng nguồn âm tốt nhưng xác định khoảng cách đến nguồn âm kém.

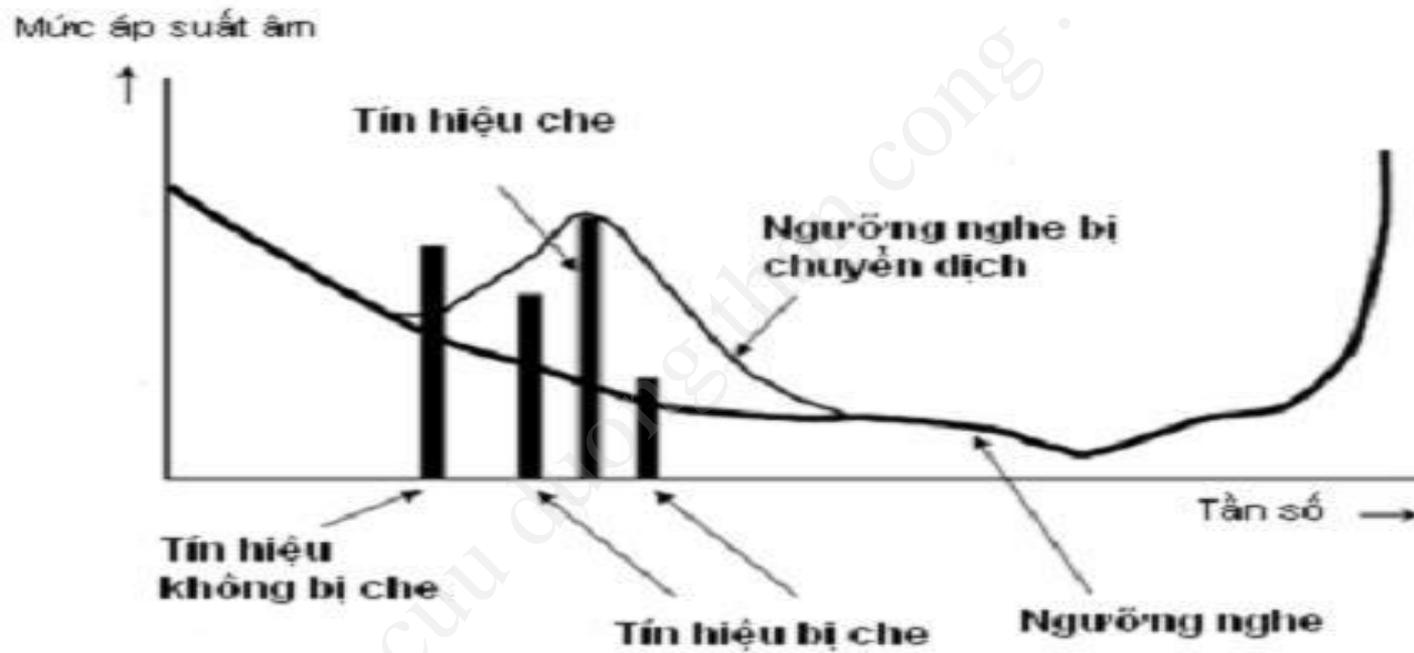
Ngưỡng nghe



Mặt nạ thời gian



Mặt nạ tần số



Chuẩn nén audio

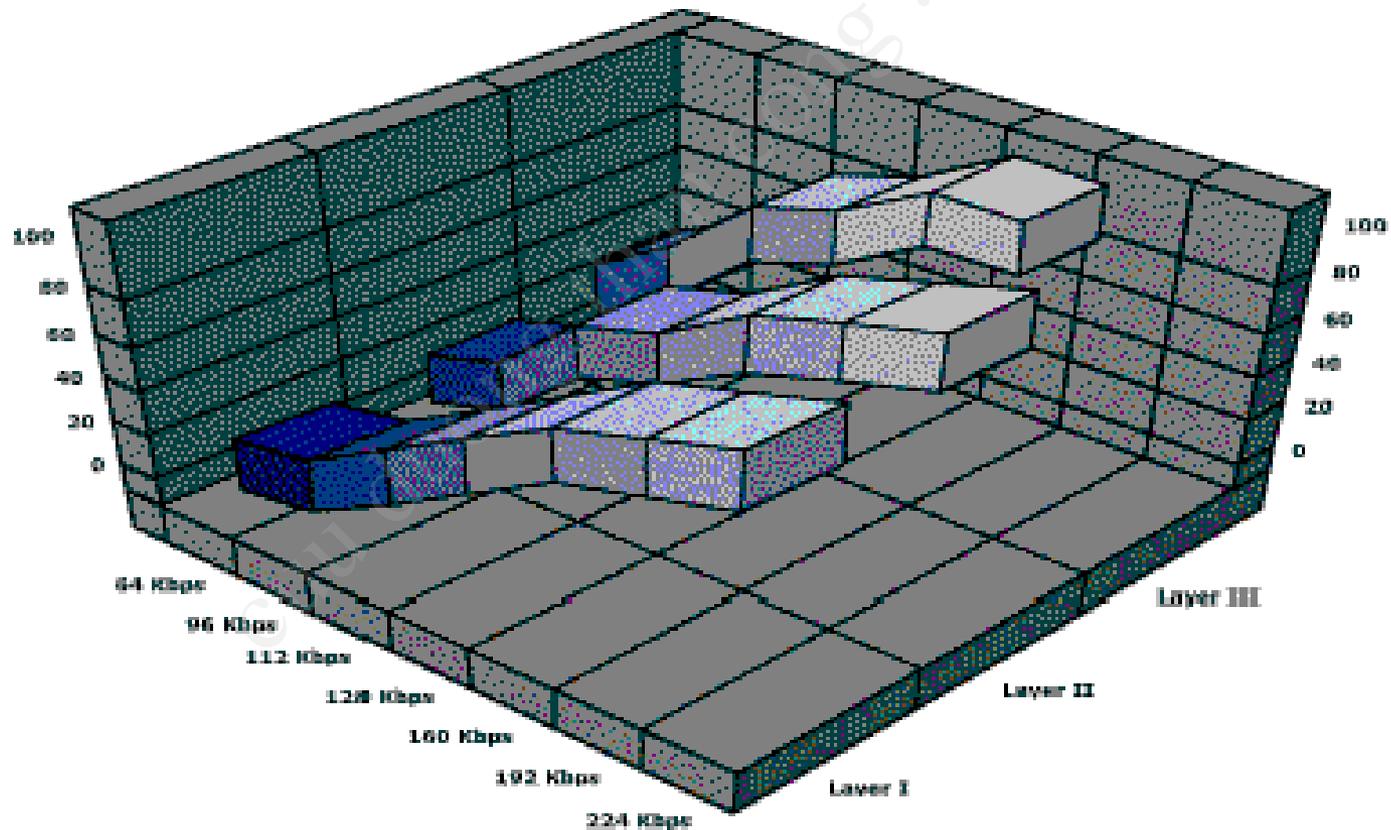
- **MP3** (MPEG 1 layer 3): ra đời năm 1980 từ viện nghiên cứu Fraunhofer Institute (Đức).
- **ACC**: Ra đời năm 1997 từ Fraunhofer Institute (Đức) kết hợp với một số công ty như AT&T, Sony, Dolby, là định dạng cải tiến của MP3.
- **OGG**: Là định dạng nguồn mở được Xiph.org Foundation đề xuất năm 1993, nén tốt và có chất lượng ở tốc độ bit thấp.

Chuẩn nén audio

- **Realaudio:** Định dạng của công ty RealNetworks, chủ yếu dùng cho phát nhạc trực tuyến, định dạng đầu tiên ra đời năm 1995, đến nay đã có RealAudio 10
- **WMA:** Định dạng âm thanh của Microsoft, ra mắt năm 1999, trên lý thuyết có thể nén 96 kbps với chất lượng của MP3 128 kbps. WMA cũng phổ biến trong thế giới âm thanh phát trực tuyến.

Các lớp MPEG/audio

Có 3 lớp:



MPEG-1

- **Đặc tính**

Lớp I	Lớp II	Lớp III
Dùng cho thiết bị dân dụng	Dùng cho thiết bị chuyên dụng, đa môi trường	Dùng cho thiết bị chuyên dụng, đa môi trường
Tốc độ dòng số liệu từ 32-448kbps	Tốc độ dòng số liệu từ 32-384kbps	Tốc độ dòng số liệu từ 32-320kbps
384mẫu/khung/kênh	1152mẫu/khung/kênh	1152mẫu/khung/kênh
32 băng con đều nhau, mỗi băng con gồm block 12 mẫu	32 băng con đều nhau, mỗi băng con gồm block 36 mẫu	32 băng con tới hạn thành 18 MDCT
Chu kỳ một khung 8ms cho kênh có $f_s=48\text{kHz}$	Chu kỳ một khung 24ms cho kênh có $f_s=48\text{kHz}$	Chu kỳ một khung 24ms cho kênh có $f_s=48\text{kHz}$
Hệ số tỷ lệ 6 bits/băng, phân phối bit theo phương thức ứng trước.	Hệ số tỷ lệ 6 bits/băng, phân phối bit theo phương thức ứng trước.	Hệ số tỷ lệ 6 bits/băng, phân phối bit theo phương thức ứng trước.

Các giải thuật nén âm thanh

Nén không tổn thất

- Mã hóa Huffman
- Mã hóa Huffman sửa đổi
- Mã hóa số học
- Giải thuật Lempel – Ziv – Welch (LZW)

Các giải thuật nén âm thanh

Nén có tổn thất

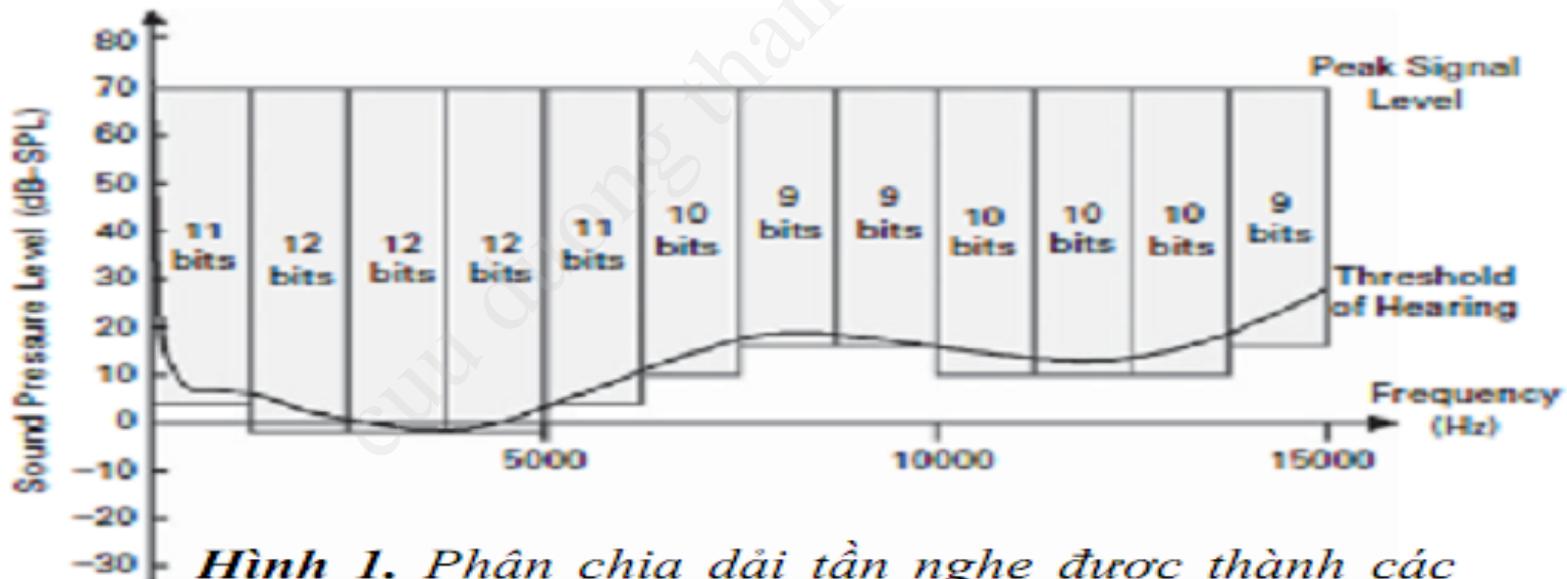
- Các phương pháp nén âm thanh đơn giản:
 - LCP(Linear Predictive Coding)
 - CELP (Code Excited Linear Predictor)
- Nén âm thanh dùng mô hình âm – tâm lý (Psychoacoustics):
 - Hệ thống nghe và phát âm của con người
 - Che tần số
 - Băng giới hạn
 - Che nhất thời
- Nén âm thanh MPEG

Nén audio MP3

- MP3 là nhóm MPEG-1 lớp 3 cung cấp chất lượng audio gần giống với chất lượng CD ở tốc độ bit thấp
- MP3 hỗ trợ các tần số lấy mẫu khác nhau như: 32kHz; 44,1kHz; 48kHz; tốc độ bit có thể thay đổi từ 32 đến 448kbps

Nén audio MP3

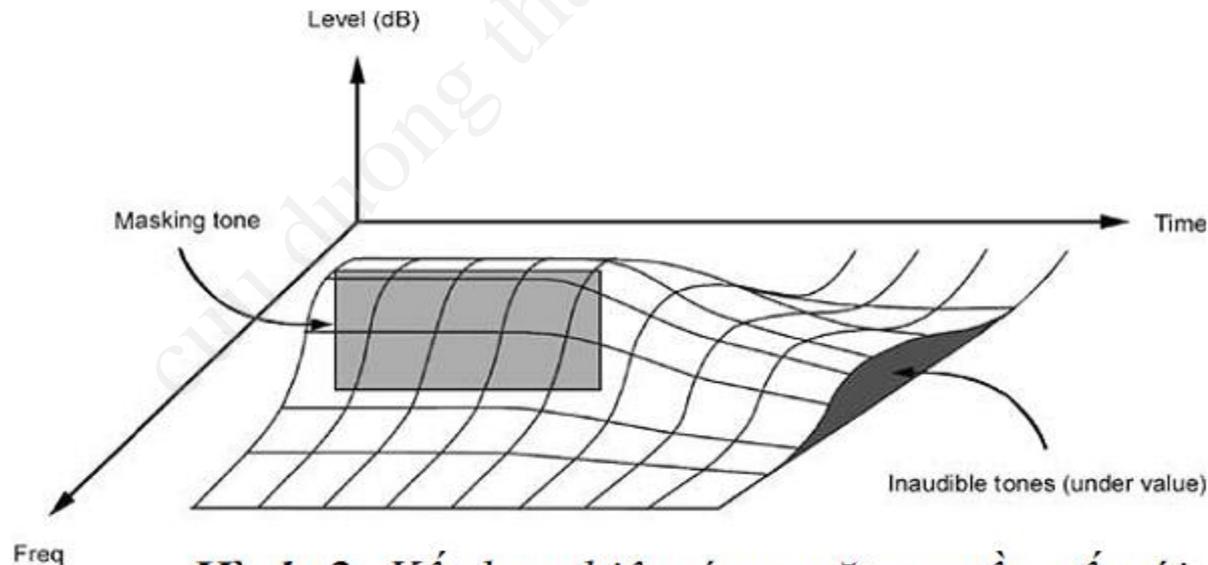
- Mã hóa audio cảm quan là kỹ thuật lợi dụng những đặc điểm cảm quan của tai người để đạt được tỉ lệ nén cao với chất lượng tốt



Hình 1. Phân chia dải tần nghe được thành các băng con và lượng tử hóa các mẫu trong từng băng với số bit khác nhau [1]

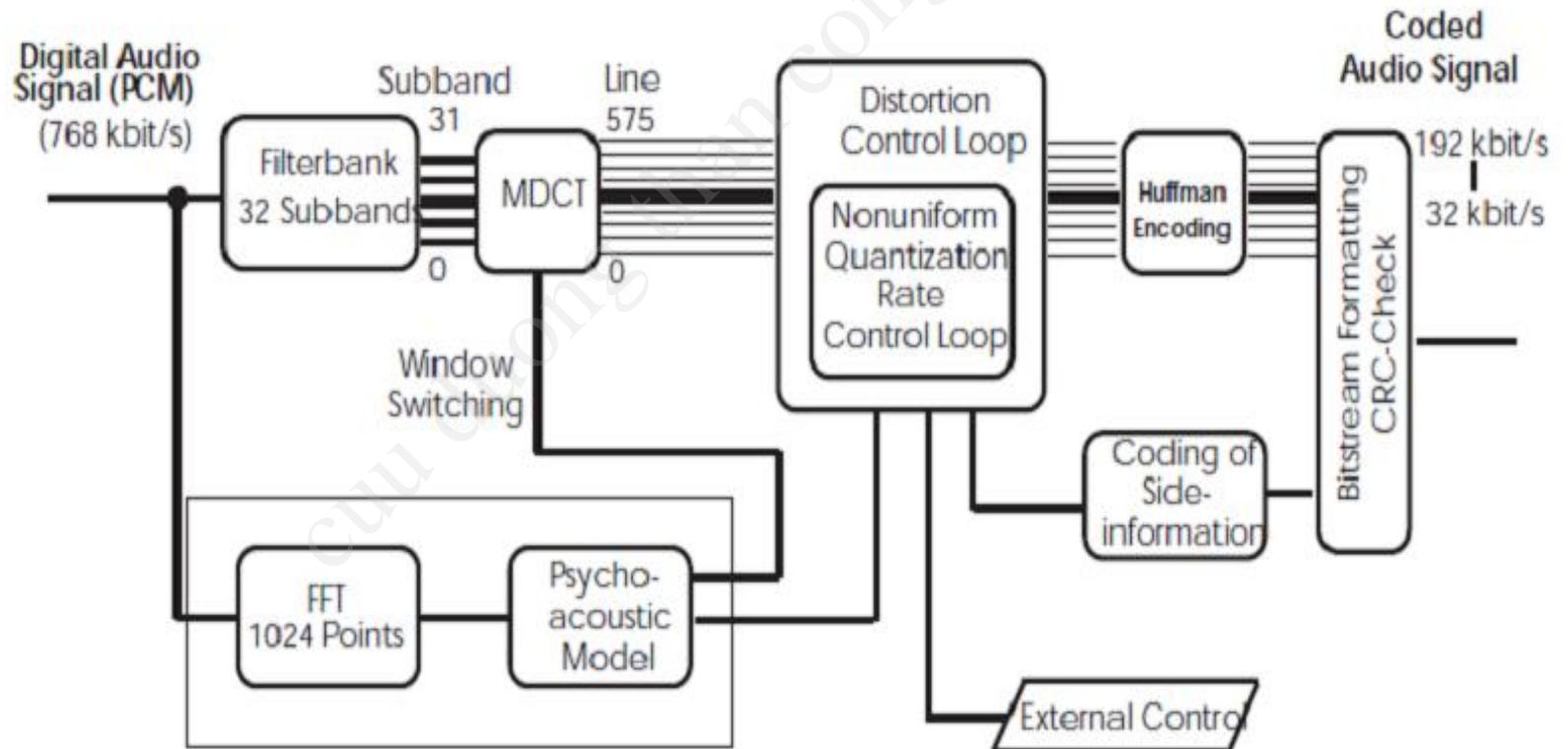
Nén audio MP3

- **Hiệu ứng mặt nạ tần số:** Hai âm thanh mạnh yếu khác nhau với tần số khác nhau xảy ra cùng 1 lúc
- **Hiệu ứng mặt nạ thời gian:** Âm thanh yếu hơn phát ra ngay trước hoặc ngay sau âm thanh mạnh



Hình 2. Kết hợp hiệu ứng mặt nạ tần số với mặt nạ thời gian [1]

Bộ mã hóa MP3



Kết quả so sánh chất lượng các file MP3

Tốc độ bit	Nhạc cổ điển lấy mẫu ở 32kHz	Nhạc cổ điển lấy mẫu ở 44.1kHz	Nhạc Rap lấy mẫu ở 32kHz	Nhạc Rap lấy mẫu ở 44.1kHz
32kbps	Rất méo, rất ồn và những nốt dài bị đứt đoạn	Không cải thiện mấy so với tần số lấy mẫu 32kHz	Rất méo, rất ồn và những đoạn nói dài bị đứt đoạn	Không cải thiện mấy so với tần số lấy mẫu 32kHz
64kbps	Vẫn méo và ồn, những nốt dài ít bị đứt đoạn hơn	Méo và bị đứt đoạn một ít ở các nốt dài, tốt hơn so với tần số lấy mẫu 32kHz, tốc độ 64kbps	Vẫn còn méo và ồn, những đoạn nói dài bớt bị đứt đoạn	Tốt hơn so với rap lấy mẫu ở tần số 32kHz, nhưng vẫn còn nhận ra méo, ồn và đứt đoạn
128kbps	Còn méo một ít, nhưng hầu như các nốt dài không còn bị đứt đoạn	Chất lượng gần với file gốc, khó phân biệt với file wav	Còn méo, ồn và bị đứt đoạn nhưng có thể chấp nhận được	Chất lượng gần với file gốc, khó phân biệt với file gốc

Giải thuật

1. Dùng bộ lọc thông dải để chia tín hiệu âm thanh thành các sub-band theo tần số, tương ứng với 32 băng giới hạn → lọc sub-band.
2. Xác định số lượng che của mỗi band gây bởi các band lân cận bằng các kết quả bước 1 → mô hình âm - tâm lý.
3. Nếu mức to của một băng mà nhỏ hơn ngưỡng che thì không mã hóa nó.
4. Ngược lại, xác định số bit cần thiết để mã hóa sao cho nhiễu sinh ra bởi việc lượng tử hóa này thấp hơn đường cong che.
5. Định dạng dòng dữ liệu bit

Ví dụ

- Sau khi phân tích, 16 band đầu tiên trong số 32 band như sau:

Band	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Level(dB)	0	8	12	10	6	2	10	60	35	20	15	2	3	5	3	1

- Nếu mức to của âm thứ 8 là 60dB, nó sẽ che band thứ 7 ở mức 12dB và band thứ 9 ở mức 15dB.
- Mức to ở band 7 là 10dB ($< 12\text{dB}$) nên bị bỏ qua, không mã hóa.
Mức to band 9 là 35 ($> 15\text{dB}$) nên được tiếp tục xử lý.
- **Layer I:** bộ lọc loại DCT với 1 frame và độ rộng tần số như nhau trên mỗi sub-band. Mô hình âm-tâm lý chỉ sử dụng hiệu quả che tần số (Frequency masking).
- **Layer II:** sử dụng 3 frame trong bộ lọc (trước, hiện tại và kế tiếp, tổng cộng 1152 mẫu). Mô hình âm-tâm lý có sử dụng hiệu quả che nhất thời (Temporal masking).
- **Layer III:** dùng bộ lọc băng giới hạn tốt hơn, mô hình âm-tâm lý có sử dụng hiệu quả che nhất thời, và có dùng bộ mã hoá Huffman.

Quá trình tạo ra tiếng nói

Cơ chế tạo âm thanh:

- Để phát âm, dưới sự điều khiển của hệ thần kinh, nói chung không khí từ phổi được đẩy qua khí quản, vào thanh hầu rồi thoát qua các cộng minh trường phía trên thanh hầu để thoát ra ngoài. Có thể thấy hai trường hợp.
- **Trường hợp 1:** Không khí thoát ra làm rung dây thanh với một tần số nào đó để tạo nên một âm với một thanh điệu nhất định. Ta có âm hữu **thanh** (sons sonores/voiced sounds).
- **Trường hợp 2:** Không khí thoát ra không làm rung dây thanh được gọi là các âm vô **thanh** (sons sourds/ voiceless sounds).

Cấu trúc bộ máy phát âm

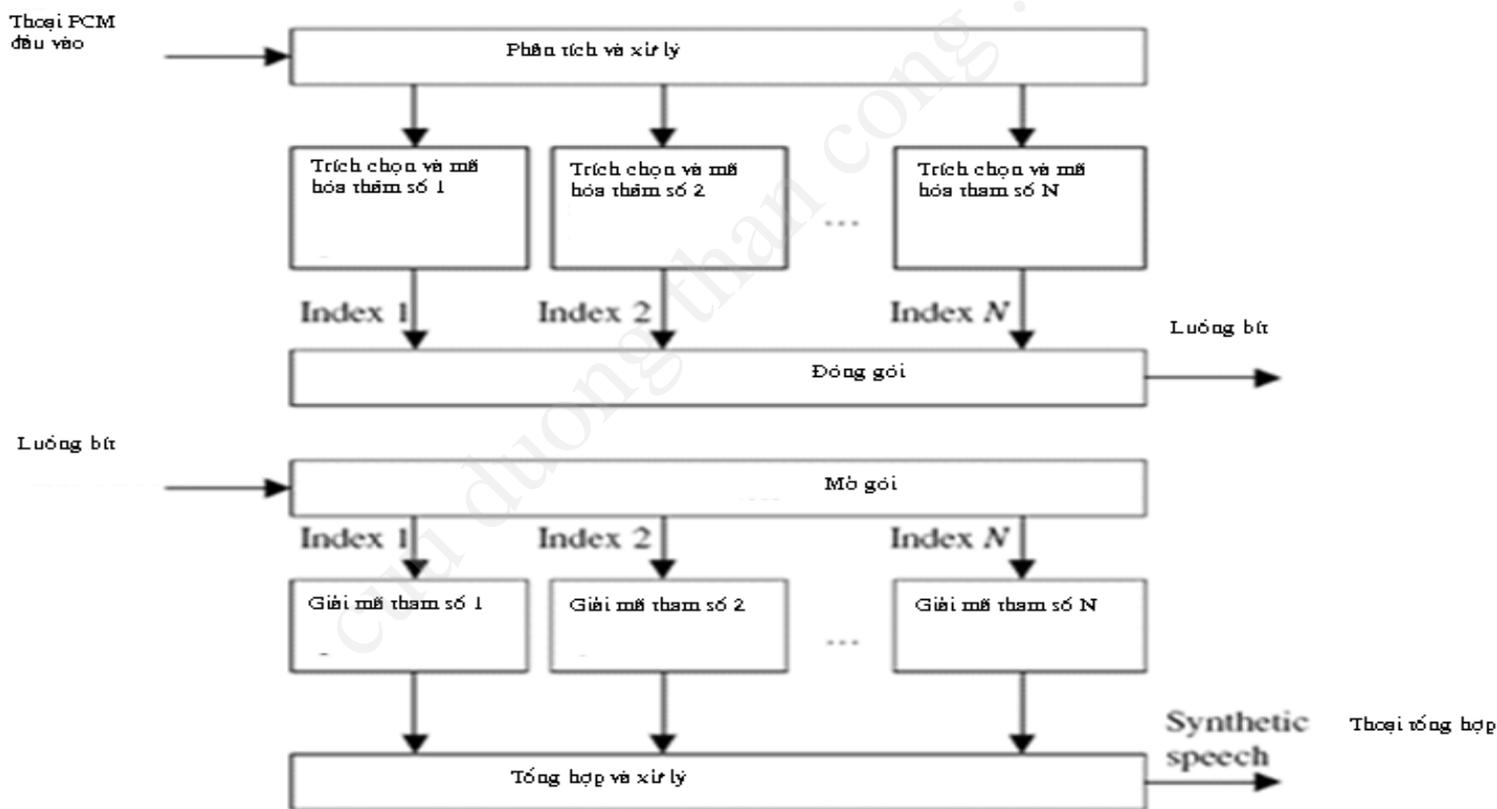
- 1. Các cơ quan tạo năng lượng cho hoạt động phát âm là phổi, khí quản.
- 2. Các cơ quan tạo lập, khuếch đại và phát ra âm thanh là các bộ phận trong thanh hầu, trong khoang hầu, khoang miệng và khoang mũi.

2.2 Nén thoại

Yêu cầu của bộ mã hoá thoại

- Tốc độ bit thấp.
- Chất lượng thoại cao.
- Nhận dạng tiếng nói / ngôn ngữ khác nhau
- Cường độ mạnh ở trong kênh truyền nhiều
- Hiệu suất cao đối với các tín hiệu phi thoại
- Kích thước bộ nhớ thấp và độ phức tạp tính toán thấp
- Độ trễ mã hóa thấp

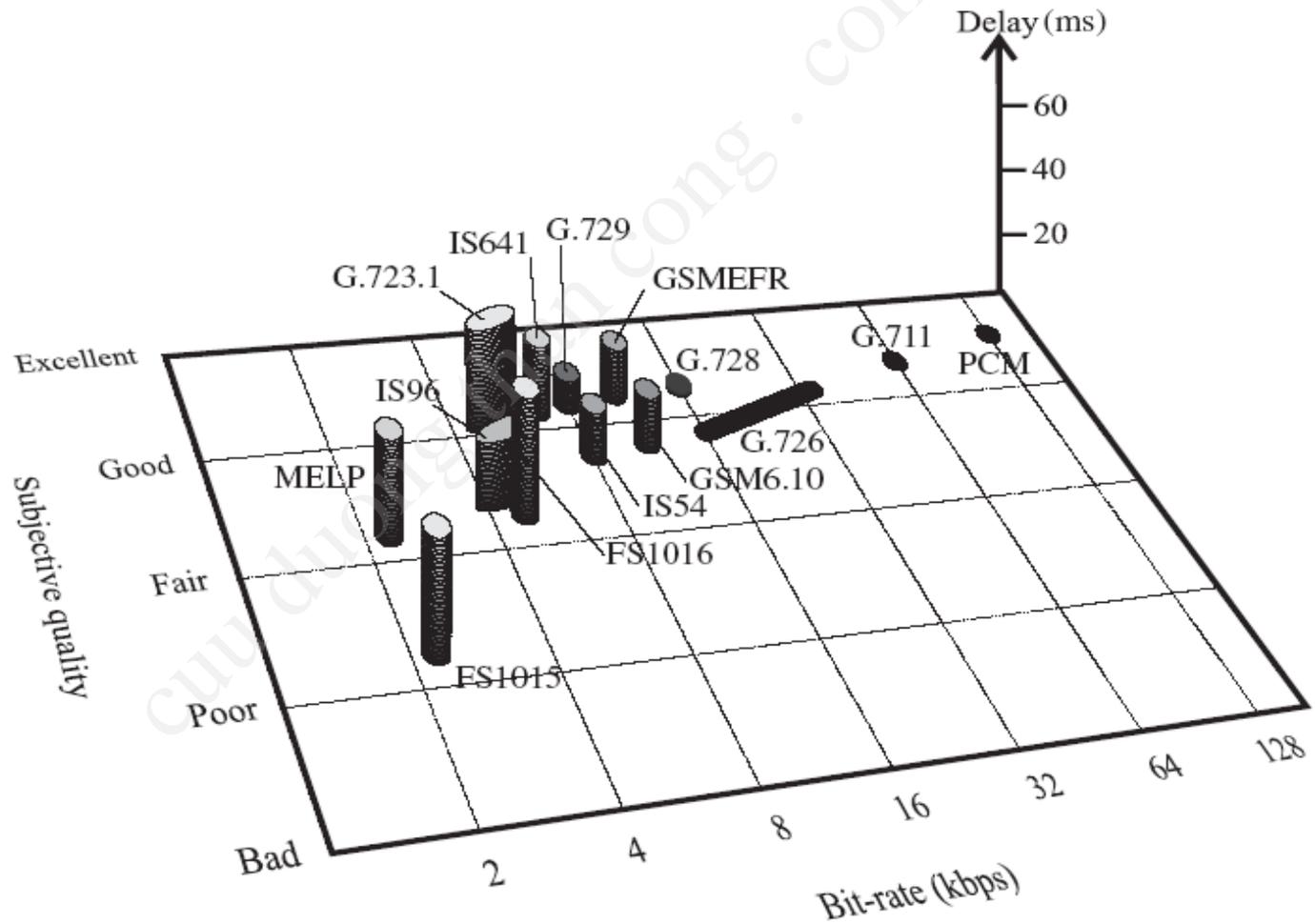
Kiến trúc tổng quát của bộ mã hóa và giải mã thoại tốc độ thấp



Chuẩn mã hóa thoại cơ bản

Ứng dụng	Tốc độ bit (kbps)	Băng thông (kHz)	Tổ chức đặt ra chuẩn	Kí hiệu chuẩn	Thuật toán dùng cho chuẩn
Điện thoại thông thường	64	3,2	Itu	g.711	μ -law hoặc A-law PCM
	32	3,2	Itu	g.726	adpcm
Hội nghị qua điện thoại	48-64	7	Itu	g.722	Subband-adpcm
	16	3,2	Itu	g.728	Low delay Celp
Điện thoại di động số	13	3,2	Gsm	Full-rate	Rpe-ltp
	12,2	3,2	Gsm	Efr	Acelp
	8,0	3,2	Tia	IS-54	Vselp
	6,5	3,2	Gsm	Half-rate	Vselp
	8,0	3,2	Itu	g.729	Acelp
Các ứng dụng có tốc độ bit thay đổi	16-40	3,2	Itu	g.727	Adpcm
	5,3-6,3	3,2	Itu	g.723.1	Mplpc, celp
	4,75-12,2	3,2	Gsm	Amr	Acelp
	1-8	3,2	Cdma	IS-96	Qcelp
Liên lạc cần bảo mật	2,4	3,2	Ddvpc	Fs1015	Lpc-10
	2,4	3,2	Ddvpc	Melp	Melp
	4,8	3,2	Ddvpc	Fs1016	Celp
Điện thoại qua vệ tinh	4,15	3,2	Inmarsat	M	Imbe
	3,6	3,2	Inmarsat	Mini-m	Ambe

So sánh giữa các chuẩn



Dự đoán tuyến tính-LP

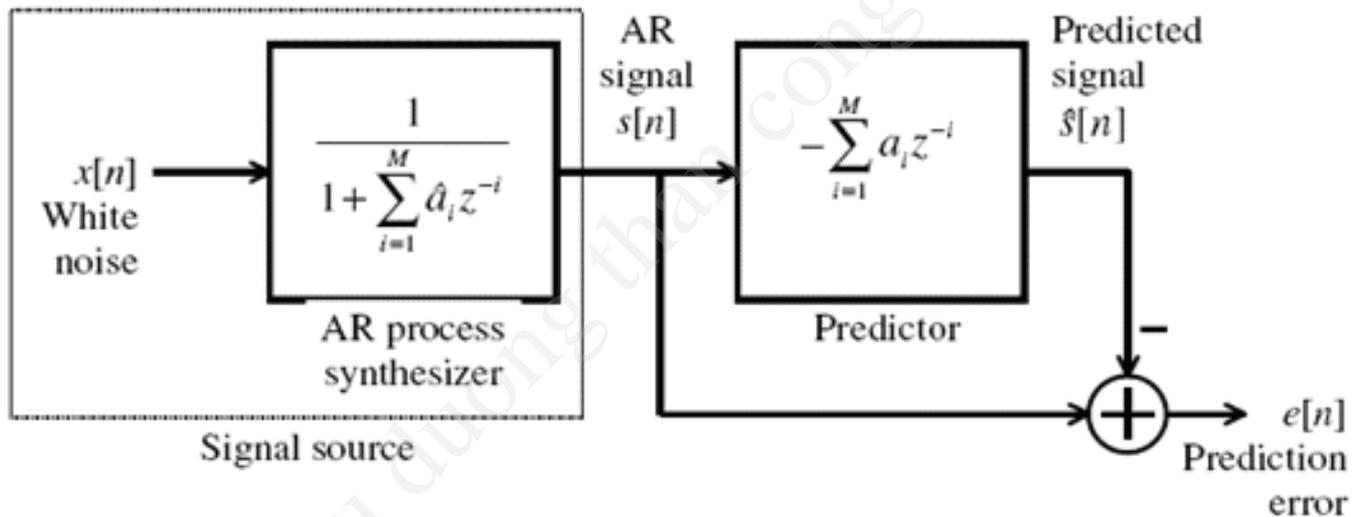
- Dự đoán tuyến tính (Linear prediction, viết tắt là LP) là một phần không thể thiếu của hầu hết tất cả giải thuật mã hóa thoại hiện đại ngày nay.

- Ý tưởng cơ bản là mẫu tiếng nói tại thời điểm n có thể được xấp xỉ bởi một tổ hợp tuyến tính M mẫu trước nó:

$$X(n) = a_1X(n-1) + a_2X(n-2) + \dots + a_MX(n-M)$$

- Trong một khung tín hiệu, các trọng số dùng để tính toán kết hợp tuyến tính được tìm bằng cách tối thiểu hóa bình phương trung bình lỗi dự đoán; các trọng số tổng hợp, hoặc các hệ số dự đoán tuyến tính (LPC) được dùng đại diện cho một khung cụ thể.

Hệ thống nhận dạng dự đoán tuyến tính



AR: Autoregressive

Dự đoán tuyến tính

- Dự đoán tuyến tính thực hiện ước đoán $\hat{s}[n]$ dựa vào M mẫu trong quá khứ:

$$\hat{s}[n] = - \sum_{i=1}^M a_i s[n-i]$$

Lỗi dự đoán được tính bằng công thức:

$$e[n] = \hat{s}[n] - s[n]$$

a_i là các ước đoán của các thông số AR được xem là các hệ số dự đoán tuyến tính (LPC).

M là bậc dự đoán.

Tiếp

- Để tìm tập các hệ số dự đoán: a_i , $i = 1, 2, \dots, M$ trên khung được phân tích, cách tiếp cận cơ bản là ta cực tiểu hóa sai số bình phương trung bình.

$$J = E\{e^2[n]\} = E\left\{\left(s[n] + \sum_{i=1}^M a_i s[n-i]\right)^2\right\}$$

Thông số LPC tối ưu có thể được tìm bằng cách thiết lập đạo hàm riêng phần của J khi a_i tiến tới zero

$$\frac{\partial J}{\partial a_k} = 2E\left\{\left(s[n] + \sum_{i=1}^M a_i s[n-i]\right)s[n-k]\right\} = 0$$

- Dùng giải thuật Levinson-durbin và Leoux-Gueguen tính LPC

Tiếp

$$E\{s[n]s[n-k]\} + \sum_{i=1}^M a_i E\{s[n-i]s[n-k]\} = 0$$

or

$$\sum_{i=1}^M a_i R_s[i-k] = -R_s[k]$$

for $k = 1, 2, \dots, M$, where

$$R_s[i-k] = E\{s[n-i]s[n-k]\},$$

$$R_s[k] = E\{s[n]s[n-k]\}.$$

Tiếp

$$\mathbf{R}_s \mathbf{a} = -\mathbf{r}_s,$$

$$\mathbf{R}_s = \begin{pmatrix} R_s[0] & R_s[1] & \cdots & R_s[M-1] \\ R_s[1] & R_s[0] & \cdots & R_s[M-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_s[M-1] & R_s[M-2] & \cdots & R_s[0] \end{pmatrix},$$

$$\mathbf{a} = [a_1 \quad a_2 \quad \cdots \quad a_M]^T,$$

$$\mathbf{r}_s = [R_s[1] \quad R_s[2] \quad \cdots \quad R_s[M]]^T.$$

$$\mathbf{a} = -\mathbf{R}_s^{-1} \mathbf{r}_s.$$

Tiếp

$$\mathbf{S} = (s(0), s(1), \dots, s(159))$$

$$\begin{bmatrix} R(0) & R(1) & R(2) & R(3) & R(4) & R(5) & R(6) & R(7) & R(8) & R(9) \\ R(1) & R(0) & R(1) & R(2) & R(3) & R(4) & R(5) & R(6) & R(7) & R(8) \\ R(2) & R(1) & R(0) & R(1) & R(2) & R(3) & R(4) & R(5) & R(6) & R(7) \\ R(3) & R(2) & R(1) & R(0) & R(1) & R(2) & R(3) & R(4) & R(5) & R(6) \\ R(4) & R(3) & R(2) & R(1) & R(0) & R(1) & R(2) & R(3) & R(4) & R(5) \\ R(5) & R(4) & R(3) & R(2) & R(1) & R(0) & R(1) & R(2) & R(3) & R(4) \\ R(6) & R(5) & R(4) & R(3) & R(2) & R(1) & R(0) & R(1) & R(2) & R(3) \\ R(7) & R(6) & R(5) & R(4) & R(3) & R(2) & R(1) & R(0) & R(1) & R(2) \\ R(8) & R(7) & R(6) & R(5) & R(4) & R(3) & R(2) & R(1) & R(0) & R(1) \\ R(9) & R(8) & R(7) & R(6) & R(5) & R(4) & R(3) & R(2) & R(1) & R(0) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \end{bmatrix} = \begin{bmatrix} -R(1) \\ -R(2) \\ -R(3) \\ -R(4) \\ -R(5) \\ -R(6) \\ -R(7) \\ -R(8) \\ -R(9) \\ -R(10) \end{bmatrix}$$

$$\begin{aligned} R(k) &= \sum_{n=0}^{159-k} s(n)s(n+k) \\ &= \text{autocorrelation of } s(n) \end{aligned}$$

Giải thuật Levison-Durbin

- Giải thuật Levison-Durbin thực hiện việc tìm bộ dự đoán bậc thứ M từ bộ dự đoán bậc thứ $M-1$. Đây là quá trình lặp đệ quy cho đến khi tìm được lần đầu tiên bộ dự đoán bậc zero, sau đó sẽ dùng bộ bậc zero để tính bộ dự đoán bậc 1 và quá trình tiếp tục cho đến khi tính toán được bộ dự đoán có bậc cần tìm.

Giải thuật Levinson-Durbin

The Levinson–Durbin algorithm is summarized as follows. Inputs to the algorithm are the autocorrelation coefficients $R[l]$, with the LPCs and RCs the outputs.

- Initialization: $l = 0$, set

$$J_0 = R[0].$$

- Recursion: for $l = 1, 2, \dots, M$

Step 1. Compute the l th RC

$$k_l = \frac{1}{J_{l-1}} \left(R[l] + \sum_{i=1}^{l-1} a_i^{(l-1)} R[l-i] \right).$$

Step 2. Calculate LPCs for the l th-order predictor:

$$\begin{aligned} a_l^{(l)} &= -k_l, \\ a_i^{(l)} &= a_i^{(l-1)} - k_l a_{l-i}^{(l-1)}; \quad i = 1, 2, \dots, l-1. \end{aligned}$$

Stop if $l = M$.

Step 3. Compute the minimum mean-squared prediction error associated with the l th-order solution

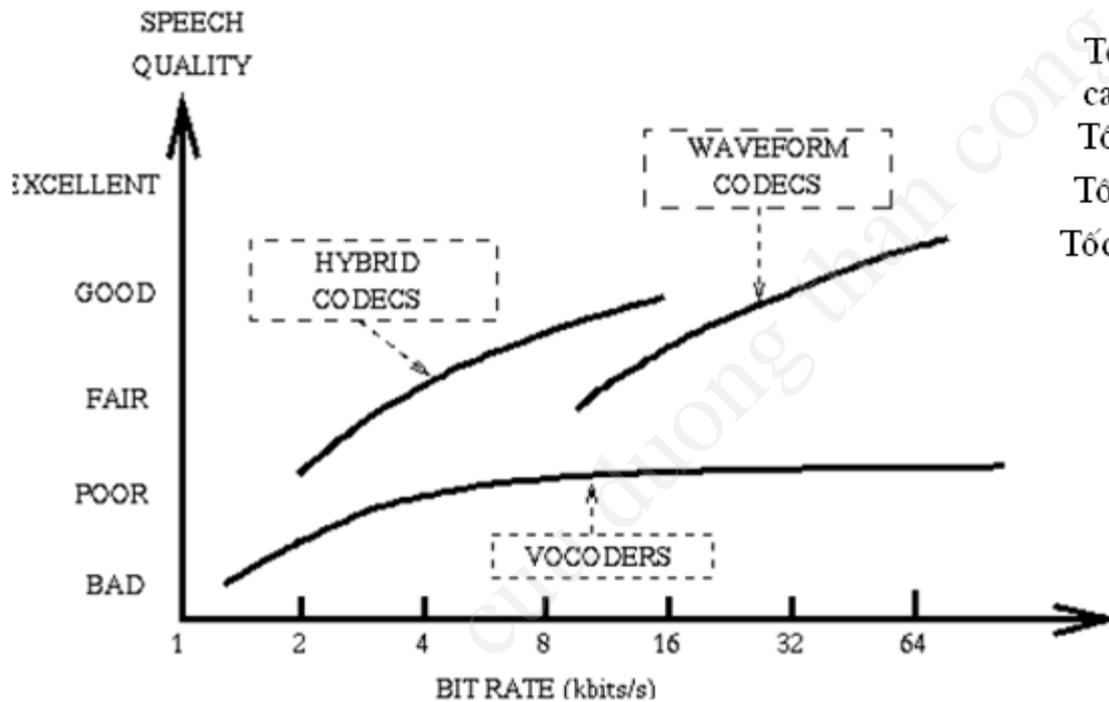
$$J_l = J_{l-1} (1 - k_l^2).$$

Set $l \leftarrow l + 1$; return to Step 1.

The final LPCs are

$$a_i = a_i^{(M)}; \quad i = 1, 2, \dots, M.$$

Phân loại mã hóa thoại



Phân loại bộ mã hóa thoại

Đặc tính

- Tốc độ bit cao
- Tốc độ bit TB
- Tốc độ thấp
- Tốc độ rất thấp

Tốc độ

- >15 kbps
- 5 to 15 kbps
- 2 to 5 kbps
- <2 kbps

Mã hoá sóng

- *Mã hoá dạng sóng*: người ta chia mã hoá dạng sóng ra làm hai loại chính
- Trong miền thời gian: mã hoá điều xung mã (PCM), điều biến xung mã vi sai (DPCM) và điều biến xung mã vi sai thích nghi (ADPCM).
- Trong miền tần số: mã hoá băng con SBC (subband coding) và mã hoá biến đổi thích nghi ATC (Adaptive Transform Coding).

Mã hoá sóng

- Tại phía phát: Bộ mã hóa nhận các tín hiệu tiếng nói tương tự và mã hóa thành tín hiệu số trước khi truyền đi
- Tại phía thu: Làm ngược lại để khôi phục tiếng nói

Ví dụ: PCM, DPCM, ADPCM..vv

Mã hoá sóng

- Khôi phục được tín hiệu sóng giống như tín hiệu gốc
- Độ phức tạp, giá thành, độ trễ công suất tiêu thụ thấp
- Chỉ tạo được tiếng nói chất lượng cao tại các tốc độ lớn hơn 16kbps
- Không tạo được tiếng nói chất lượng cao tại tốc độ nhỏ hơn 16kbps

Mã hóa Vocoder

- Mã hoá Vocoder sử dụng mô hình tạo tín hiệu thoại và khai thác các thông số của mô hình này để mã hoá tín hiệu.
- Các vocoder hoạt động dựa trên mô hình cơ quan phát âm
- **Thông tin được gửi đến bộ giải mã bao gồm: tham số của bộ lọc, tín hiệu kích thích V/UV, chu kỳ pitch ..vv.**
- Ví dụ : LPC, mã hoá dự đoán tuyến tính có sự kích thích kết hợp MELP và mã hoá dự đoán tuyến tính kích thích bằng tín hiệu sau dự đoán RELP..vv

Mã hóa Vocoder (mã hóa kiểu phát âm)

- Có nhiều kỹ thuật mã hoá Vocoder như: mã hoá formant, mã hoá tham số và mã hoá đồng hình. Tuy nhiên, hiện nay chủ yếu tập trung vào nghiên cứu và phát triển các bộ mã hoá tham số như mã hoá dự đoán tuyến tính kích thích bằng hai trạng thái (mã hoá LPC),
- Mã hoá dự đoán tuyến tính có sự kích thích kết hợp MELP và mã hoá dự đoán tuyến tính kích thích bằng tín hiệu sau dự đoán RELP. Các bộ mã hoá tham số này thường dùng cho điện thoại qua vệ tinh và trong quân đội.

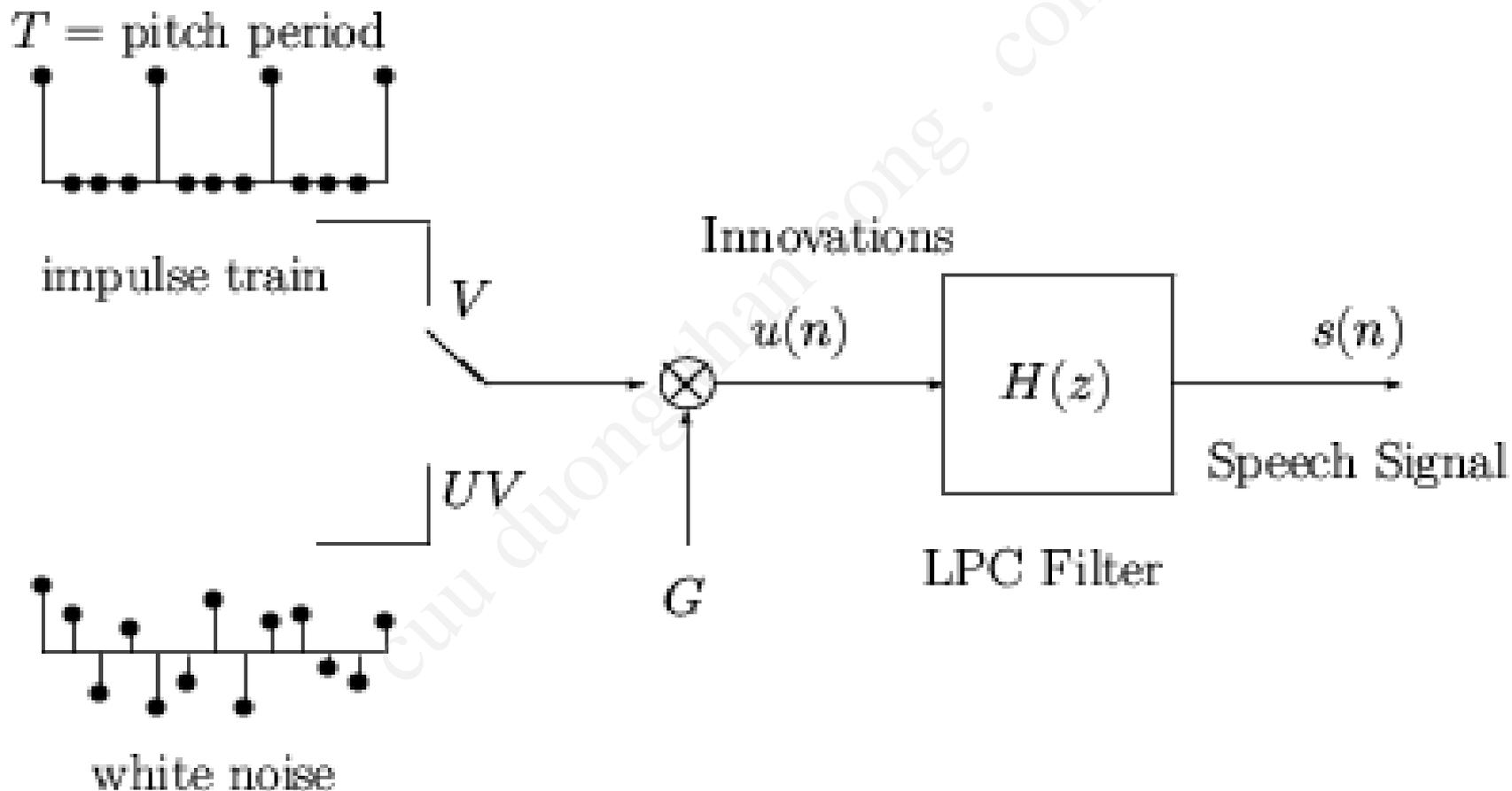
Ưu nhược điểm của Vocoder

- Chất lượng phụ thuộc nhiều vào mô hình thoại
- Các Vocoder có thể phát âm khá giả tạo
- Chất lượng kém các vocoder rất nhạy cảm với lỗi.
- Có thể cung cấp thoại số với tốc độ nhỏ hơn 2kbps

Mã hóa lai

- Sử dụng cả 2 công nghệ mã hoá sóng và mã hoá Vocoder
- Có thể đạt được chất lượng thoại tốt tại các tốc độ bit 2-16kbps.
- Mã hóa lai phổ biến nhất là mã hóa phân tích bằng cách tổng hợp AbS (Analysis-by-Synthesis), RPE-LTP, CELP, ACELP, CS-CELP...

Mã hóa dự đoán tuyến tính LPC



Mô hình LPC tổng hợp tiếng nói

Mối quan giữa mô hình LPC với mô hình phát âm

Cường hộng $\iff H(z)$ (Bộ lọc LPC)

Không khí $\iff u(n)$

Dao động của dây thanh âm $\iff V$ (Các âm hữu thanh)

Chu kỳ dao động của dây thanh âm $\iff T$ (Độ cao âm thanh)

Các âm vô thanh $\iff UV$ (Vô thanh)

Lượng không khí $\iff G$ (Độ tăng ích)

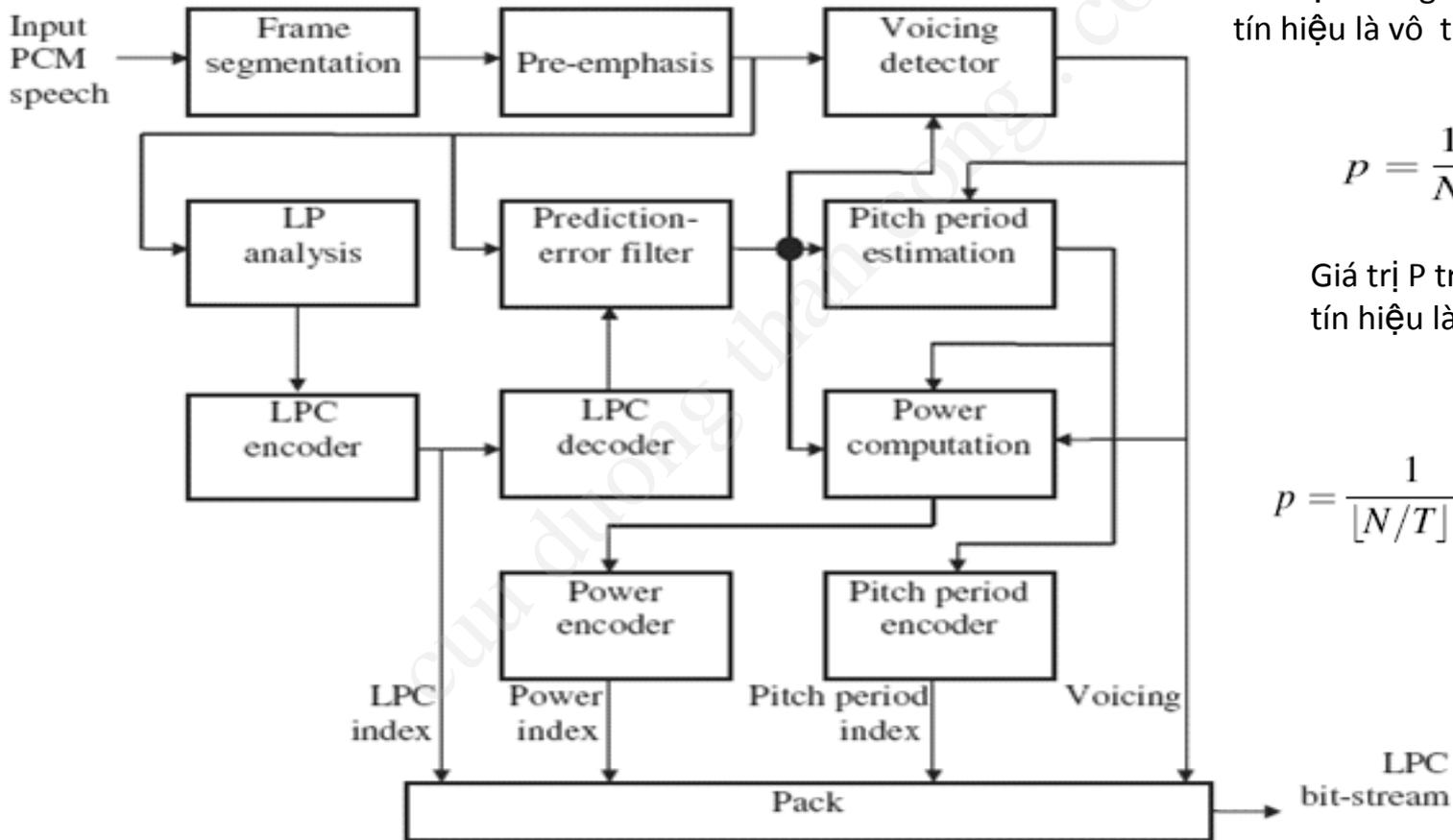
Mã hóa dự đoán tuyến tính LPC

$$\mathbf{S} = (s(0), s(1), \dots, s(159))$$

- Tín hiệu thoại số khi đưa vào mô hình sẽ được chia thành các frames 20ms , mô hình sẽ phân tích và trích trộn và Vector A đại diện cho 160 mẫu thoại đó và nó sẽ được mã hóa và gửi tới đầu thu:

$$\mathbf{A} = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9, a_{10}, G, V/UV, T)$$

Sơ đồ khối của bộ mã hóa LPC



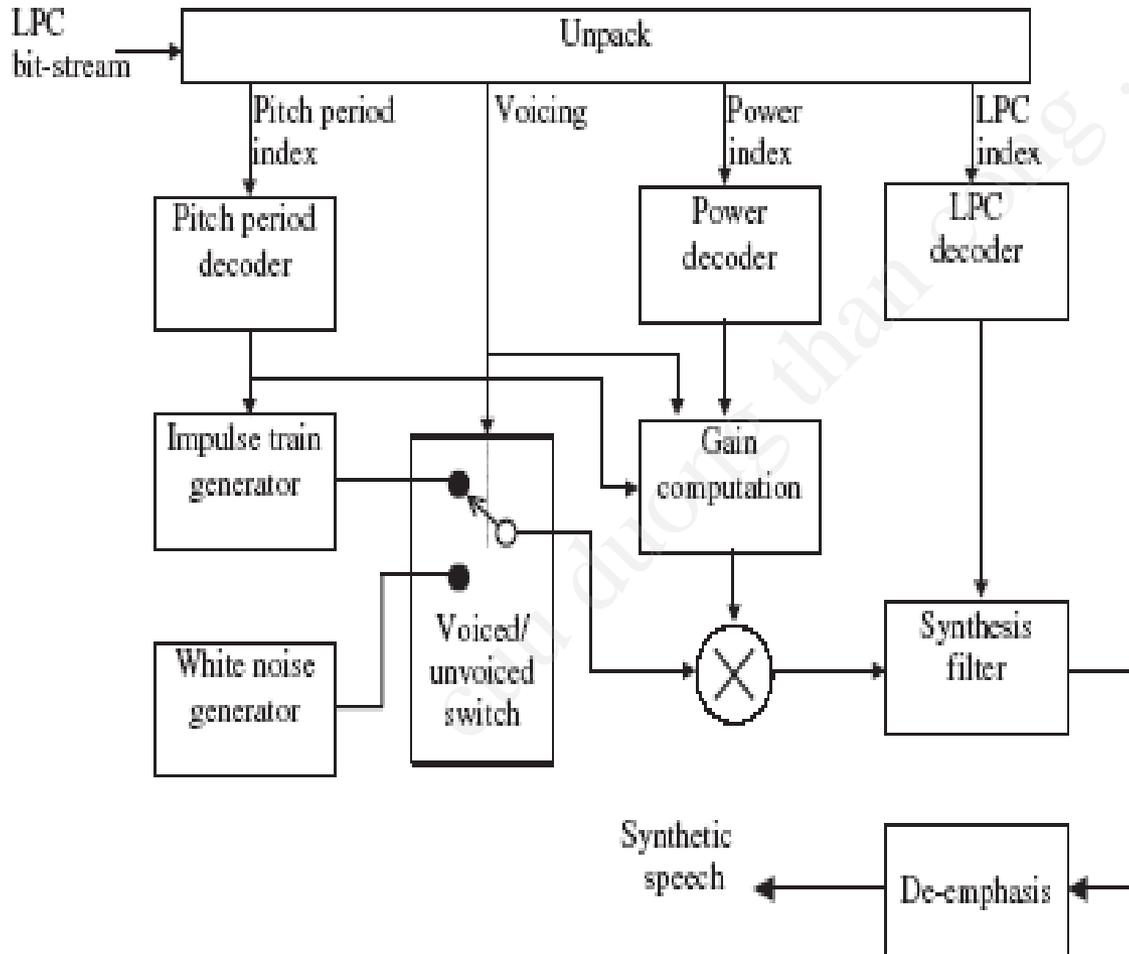
Giá trị P trong trường hợp tín hiệu là vô thanh

$$p = \frac{1}{N} \sum_{n=0}^{N-1} e^2[n].$$

Giá trị P trong trường hợp tín hiệu là hữu thanh

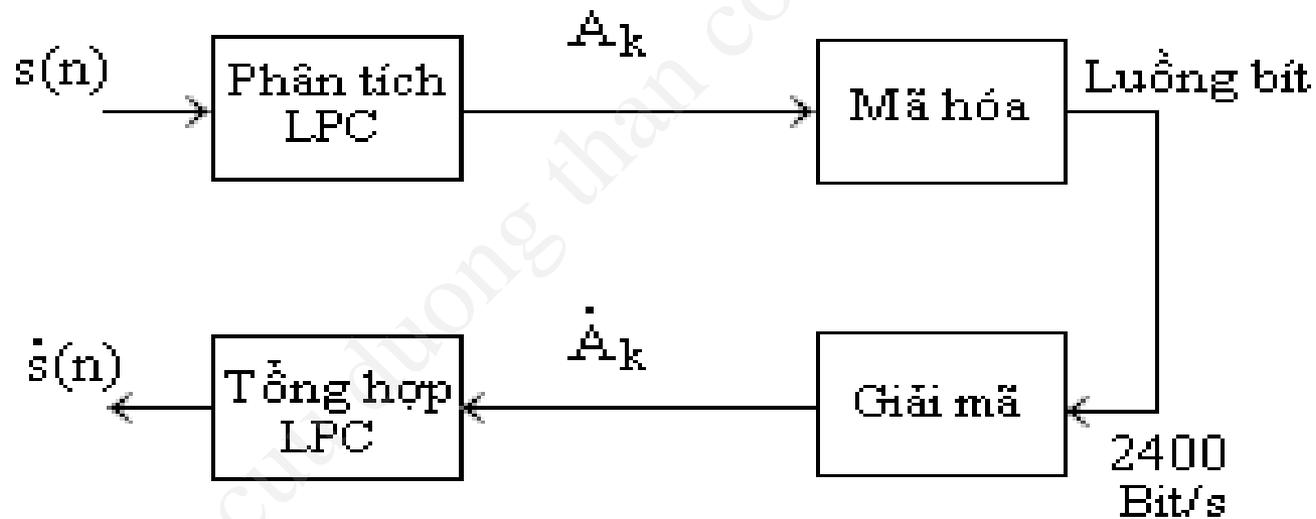
$$p = \frac{1}{[N/T]T} \sum_{n=0}^{[N/T]T-1} e^2[n]$$

Sơ đồ khối của bộ mã hóa LPC



$$g = \sqrt{Tp}.$$

2.4kbps LPC Vocoder



2.4kbps LPC Vocoder

Parameter Name	Parameter Notation	Rate (bits/frame)
LPC (LSP)	$\{a_k\}_{k=1}^{10}$ ($\{\omega_k\}_{k=1}^{10}$)	34
Gain	G	7
Voiced/Unvoiced & Period	$V/UV, T$	7
Total		48

Nhược điểm của LPC

- Giới hạn 1: Trong một số trường hợp, một khung âm thanh không phải chỉ được phân loại thành hữu thanh và vô thanh
- Giới hạn 2: Việc sử dụng hoàn toàn nhiễu ngẫu nhiên hoặc hoàn toàn chuỗi xung có chu kỳ tạo kích thích không phù hợp với thực tế
- Giới hạn 3: Thông tin về pha của tín hiệu nguyên thủy không được xem xét.
- Giới hạn 4: Phương pháp thực hiện việc tổng hợp các khung thoại, trong khi một chuỗi xung dùng để kích thích bộ lọc tổng hợp với các hệ số có được từ việc phân tích LP vi phạm nền tảng của mô hình AR.

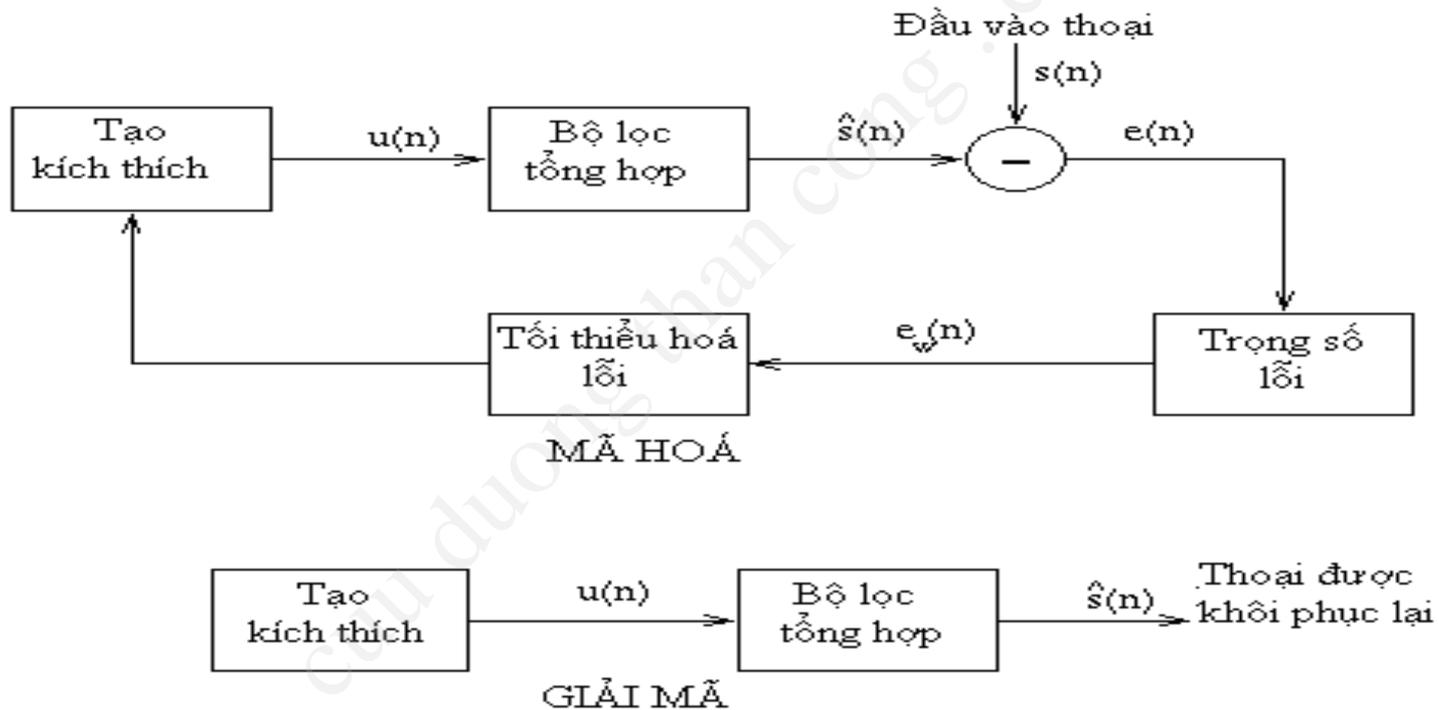
Mã hóa LPC

- Dự đoán tuyến tính là một phần không thể thiếu của hầu hết các giải thuật mã hóa thoại tốc độ thấp
- Việc phát triển và cải tiến các bộ mã hóa dựa trên mô hình LPC vẫn đang được xem xét để đưa ra các thuật toán mã hóa thoại tốc độ thấp thậm chí dưới 2kbps mà chất lượng thoại vẫn đảm bảo yêu cầu

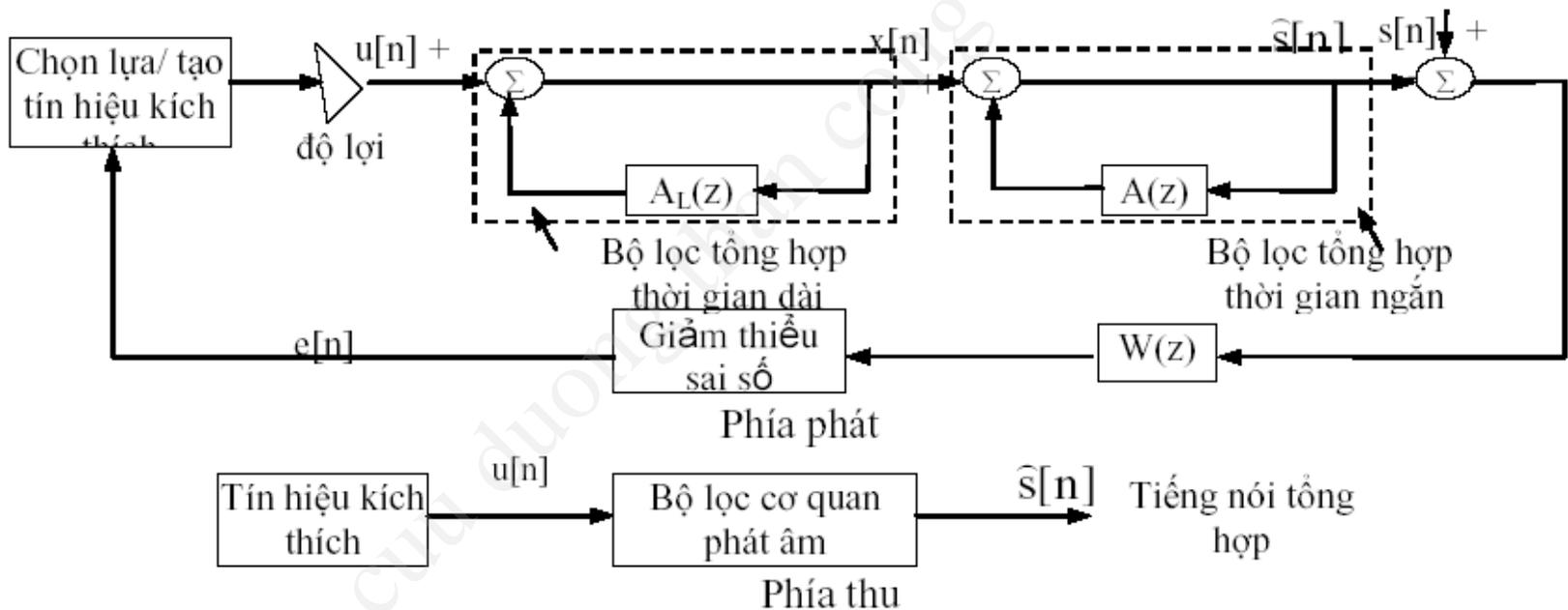
Mã hoá phân tích bằng tổng hợp AbS

- Mã hóa lai có nhiều phương pháp nhưng phương pháp phổ biến nhất là mã hoá phân tích bằng cách tổng hợp AbS (Analysis-by-Synthesis). Bộ mã hoá này cũng sử dụng mô hình cơ quan phát âm của người giống như mã hoá nguồn. Tuy nhiên, thay vì sử dụng các mô hình tín hiệu kích thích đơn giản như mã hoá nguồn thì ở đây tín hiệu kích thích được chọn sao cho cố gắng đạt được dạng sóng tiếng nói tái tạo càng giống với dạng sóng tiếng nói ban đầu càng tốt. Đây chính là đặc tính phân biệt sự khác nhau giữa các bộ mã hoá kiểu AbS. Thuật toán tìm ra dạng sóng kích thích này quyết định tới độ phức tạp của bộ mã hoá.

LPC phân tích bằng tổng hợp



Mã hoá phân tích AbS (Analysis-by-Synthesis)



Mã hoá phân tích bằng tổng hợp AbS (Analysis-by-Synthesis)

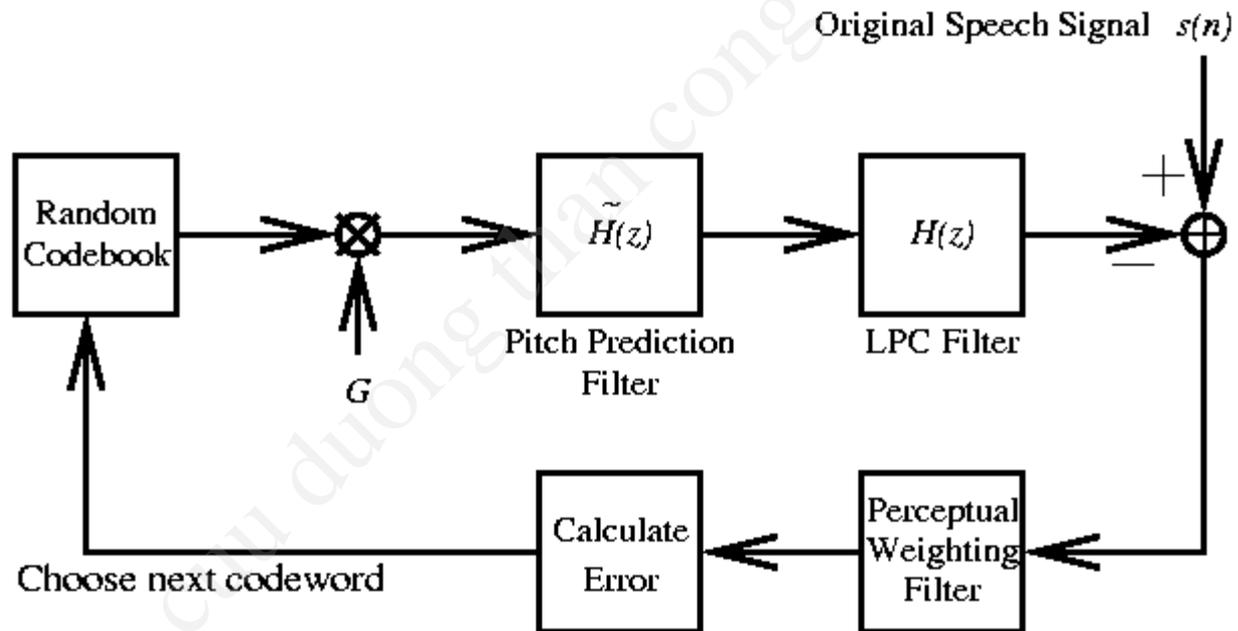
Bộ dự đoán thời gian ngắn thực chất là bộ lọc tổng hợp tiếng nói. Bộ lọc này sẽ thực hiện việc tổng hợp tiếng nói khi có tín hiệu kích thích đưa đến ngõ vào của nó. Các hệ số của bộ lọc này sẽ được xác định bằng phương pháp dự đoán tuyến tính như đã đề cập ở trên.

Các đoạn tiếng nói hữu thanh có dạng sóng tuần hoàn và sự tuần hoàn này có thể được khai thác để trợ giúp cho quá trình dự đoán tiếng nói. Từ điều này người ta đã đưa ra khái niệm về dự đoán thời gian dài hay dự đoán pitch. Cũng giống như các bộ STP, các bộ LTP cũng là các bộ dự đoán tuyến tính nhưng trong khi STP thực hiện việc dự đoán dựa trên các mẫu kề nhau thì LTP dựa trên các mẫu từ một hay nhiều chu kì pitch trước đó. Đây là lí do gọi nó là dự đoán thời gian dài.

Tiếp

- Ngoài việc khai thác các tính chất tiếng nói để mã hoá, người ta còn khai thác sự cảm nhận âm thanh của tai người (tai người không cảm nhận được những âm thanh bị che đi bởi các âm thanh khác có năng lượng lớn hơn một mức nhất định) trong mã hoá tiếng nói bằng khái niệm bộ lọc nhân cảm nhận.

4.8 kbps CELP Coder



4.8 kbps CELP Coder

- CELP Dự đoán tuyến tính kích thích mã Code-Excited Linear Prediction.
- Nguyên tắc giống LPC Vocoder chỉ có một số điểm khác :
 - Kích thước Frame là 30 msec (240 mẫu)
 - Mã hóa trực tiếp
 - Cần nhiều bit mã hóa hơn
 - Tính toán phức tạp hơn
 - Sử dụng thêm bộ lọc dự đoán chu kỳ pitch
 - Sử dụng lượng tử hóa Vector

Nhận xét

Hầu hết tất cả các bộ mã hóa đều dựa trên mô hình LPC, tùy theo cách tạo ra tín hiệu kích thích mà người ta đưa ra các loại mã hoá lai khác nhau như:

- Mã hoá đa xung MPE-LTP
- Mã hoá xung đều RPE-LTP
- Mã hoá kích thích bằng mã CELP, ACELP, CS-ACELP..
- Mã hoá kích thích vectơ tổng VSELP....vv

Các bộ mã hóa trên đã khắc phục nhược điểm của LPC và cung cấp dịch vụ thoại tốc độ thấp và chất lượng tương đối tốt

Một số phương pháp đánh giá chất lượng thoại

- Phương pháp đánh giá theo thang điểm MOS (Mean Opinion Score) dựa trên khuyến nghị ITU-T P.800
- Phương pháp đánh giá dựa trên mô hình giác quan PSQM (Perceptual Speech Quality Measurement) theo khuyến nghị ITU-T P.861
- PESQ (Perceptual Evaluation of Speech Quality) theo khuyến nghị ITU-T P.862
- Phương pháp dựa trên mô hình đánh giá truyền dẫn E-model theo tiêu chuẩn ETR 250 của ETSI.

Các yếu tố ảnh hưởng tới chất lượng thoại trong VoIP

- Độ ổn định
- Băng thông
- Tiếng vọng
- Trễ: Trễ xử lý, trễ do mã hoá, trễ đệm ở thiết bị đầu cuối IP, trễ gói hoá H.323, trễ truyền dẫn mạng
- Biến động trễ
- Tổn thất gói

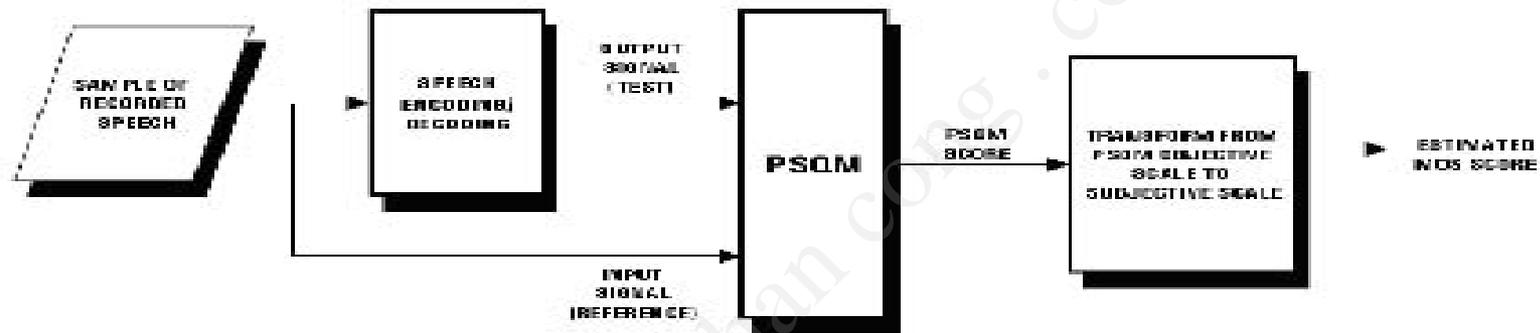
Phương pháp đánh giá chủ quan (MOS)

- Bài kiểm tra hội thoại (Conversation Opinion Test).
- Đánh giá phân loại tuyệt đối (Absolute Category Rating (ACR) Test).
- Phương thức phân loại theo suy hao (Degradation Category Rating (DCR)).
- Phương thức phân loại so sánh (Comparison Category Rating (CCR)).

Nhược điểm MOS

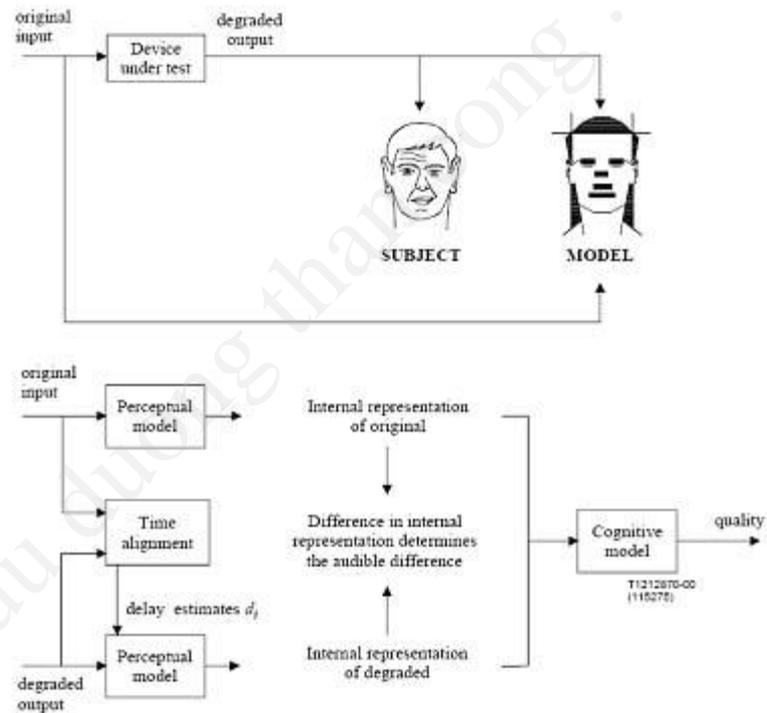
- Phương thức này mang tính chất chủ quan vì kết quả phụ thuộc vào nhiều yếu tố không thể kiểm soát của chủ thể như: trạng thái tâm lý, thái độ đối với bài kiểm tra và trình độ văn hóa. Trên thực tế, phương thức đánh giá chất lượng thoạt theo thang điểm MOS không phải là phương thức nhất quán.
- Phương thức này rất tốn kém, đòi hỏi nhiều người tham gia và thiết lập phức tạp.
- Khi cần thực hiện đo thường xuyên các tham số chất lượng thì việc sử dụng phương pháp đánh giá chất lượng này là không thực tế.

Phương thức đánh giá chất lượng thoại PSQM

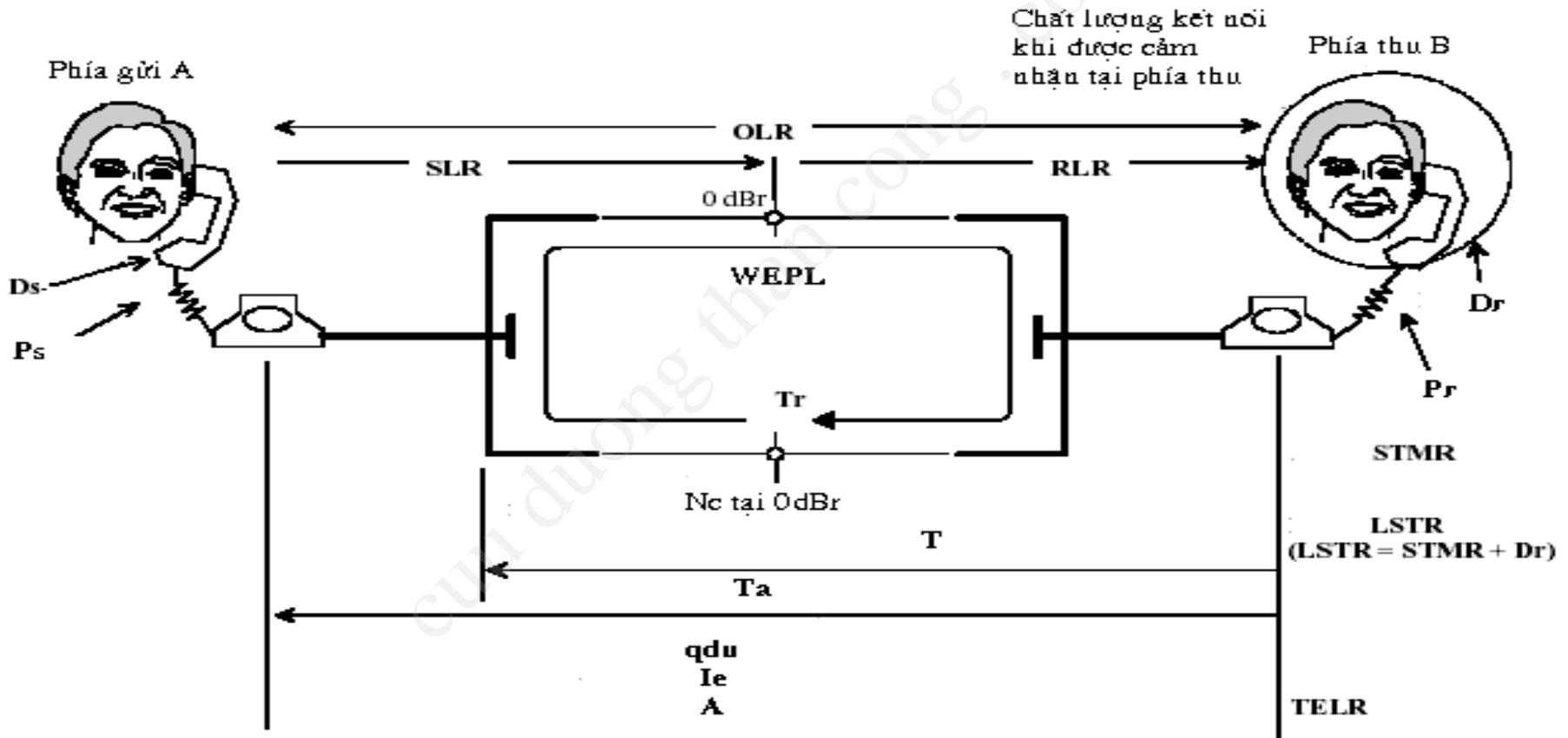


Để thực hiện phép đo PSQM, một mẫu tiếng nói được đưa vào hệ thống và được xử lý bởi một bộ mã hóa thoại bất kỳ. Những tính chất của tín hiệu vào giống như của các tín hiệu sử dụng cho phép đánh giá MOS được định nghĩa trong chuẩn ITU P.830.

Phương pháp PESQ



Mô hình đánh giá truyền dẫn E-Model



Cấu hình tham khảo của mô hình E

Các tham số dùng để đánh giá chất lượng thoại của mô hình E-Model

- Giá trị truyền dẫn R
- Tỷ lệ tín hiệu trên nhiễu
- Tham số suy hao I_s
- Tham số suy hao liên quan đến trễ I_d
- Tham số suy hao thiết bị I_e
- Tham số tích cực A

Kết quả đánh giá chất lượng thoại



Các phương pháp cải thiện QoS trong mạng VoIP

- Tốc độ truy nhập cam kết
- Xếp hàng trên cơ sở lớp
- Lớp dịch vụ
- Các dịch vụ phân biệt
- Quyền ưu tiên IP
- Chuyển mạch nhãn đa giao thức MPLS
- Xếp hàng theo VC
- Định tuyến theo chính sách
- Các hàng QoS
- Loại bỏ sớm ngẫu nhiên
- Giao thức dữ trữ tài nguyên
- Định hình lưu lượng
- Xếp hàng hợp lý theo trọng số

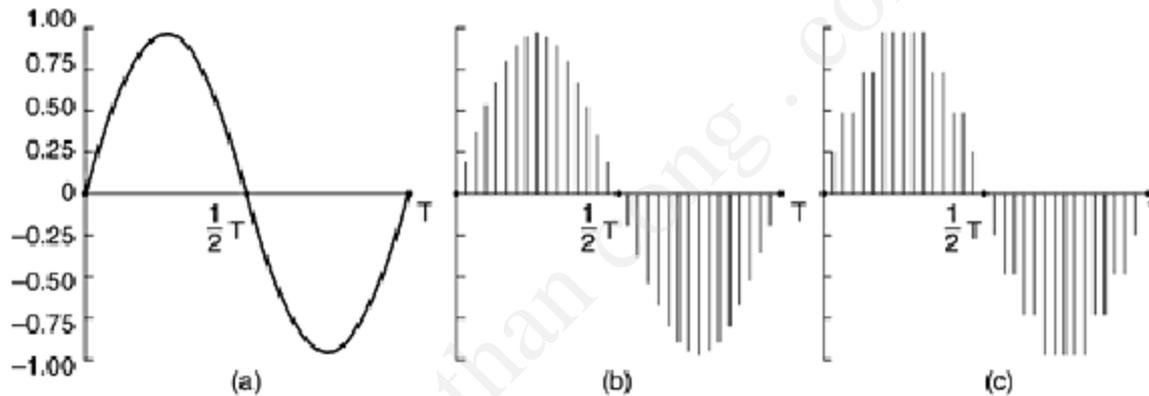
Nhận xét

- Sử dụng PESQ để đánh giá chất lượng thoại một chiều từ đầu cuối đến đầu cuối.
- Mô hình đánh giá E-Model có thể được sử dụng để phân tích hệ thống nhằm xác định các yếu tố ảnh hưởng đến chất lượng thoại.
- Ngoài ra, nếu có điều kiện có thể sử dụng kết hợp phương pháp đánh giá chủ quan để kiểm chứng lại việc đánh giá theo PESQ.

Hướng dẫn ôn tập chương 2

- Các tham số đánh giá đặc trưng cho âm thanh.
- Sơ đồ chức năng của tai người.
- Sự cảm thụ của tai người đối với âm thanh (Ngưỡng nghe, mặt nạ tần số, mặt nạ thời gian, các dải băng tần tới hạn) và ứng dụng trong mã hóa âm thanh.
- Quá trình tạo tiếng nói và ứng dụng trong mã hóa tham số tín hiệu thoại.
- Mô hình chung của bộ mã hóa thoại.
- Phân loại các phương pháp mã hóa tín hiệu thoại (mã hóa dạng sóng, mã hóa tham số và mã hóa lai).
- Phương pháp mã hóa PCM, DPCM, ADPCM.
- Mã hóa tham số - LPC.
- Mã hóa lai – AbS, MPE, RPE và CELP.
- Mã hóa âm thanh.
- So sánh giữa nén thoại và nén âm thanh.

Hướng dẫn ôn tập chương 2



- Trong các hình vẽ trên : (a) Tín hiệu hình âm thanh hình Sin; (b) Lấy mẫu tín hiệu; (c) Lượng tử hóa các xung mẫu.
- Ở hình vẽ (c) chúng ta có thể thấy nhiều lượng tử xuất hiện do việc sử dụng các từ mã 4 bit để biểu diễn cho 9 mức giá trị khác nhau. Xung mẫu đầu tiên tại thời điểm 0 là chính xác, nhưng một số các xung mẫu tiếp theo thì không. Tính toán sai số (theo %) cho các xung mẫu tại các thời điểm $1/32$, $2/32$ và $3/32$ của chu kỳ lấy mẫu?

Chương 3

Xử lý ảnh

Xử lý ảnh số

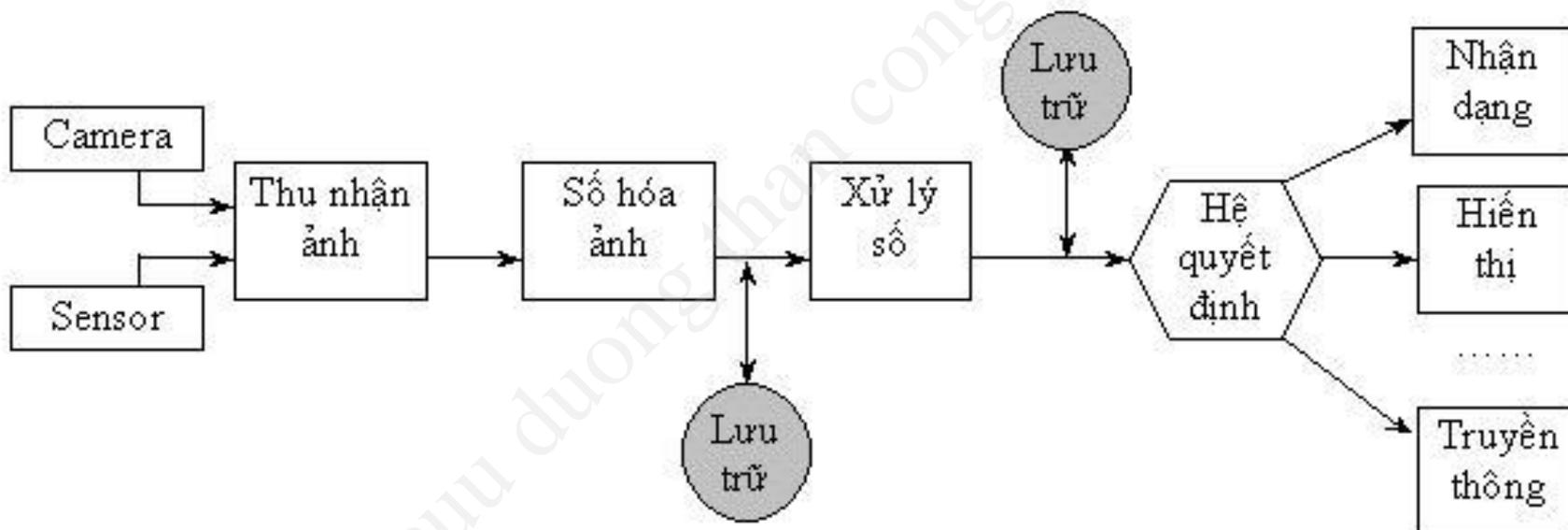
Xử lý ảnh số là quá trình biến đổi ảnh số trên máy tính (PC).

Ảnh số được tạo ra bởi một số hữu hạn các điểm ảnh, mỗi điểm ảnh nằm tại một vị trí nhất định và có 1 giá trị nhất định. Một điểm ảnh trong một ảnh còn được gọi là một pixel

Tại sao cần xử lý ảnh số?

- Tại sao phải nén tín hiệu?
 - Ảnh được sử dụng mọi lúc, mọi nơi
 - Hạn chế về không gian lưu trữ và tốc độ đường truyền
- Lý do phải xử lý ảnh
 - Ảnh có thể bị lỗi trong quá trình thu ảnh, truyền dẫn và hiển thị (hồi phục, nâng cao chất lượng ảnh, nội suy)
 - Ảnh có thể mang các nội dung nhạy cảm (vd, chống lại copy không hợp pháp, giả mạo và lừa đảo)
 - Tạo các bức ảnh có hiệu ứng nghệ thuật
- Lý do phải phân tích ảnh
 - Dạy máy tính có khả năng “nhìn” được (nhận dạng)

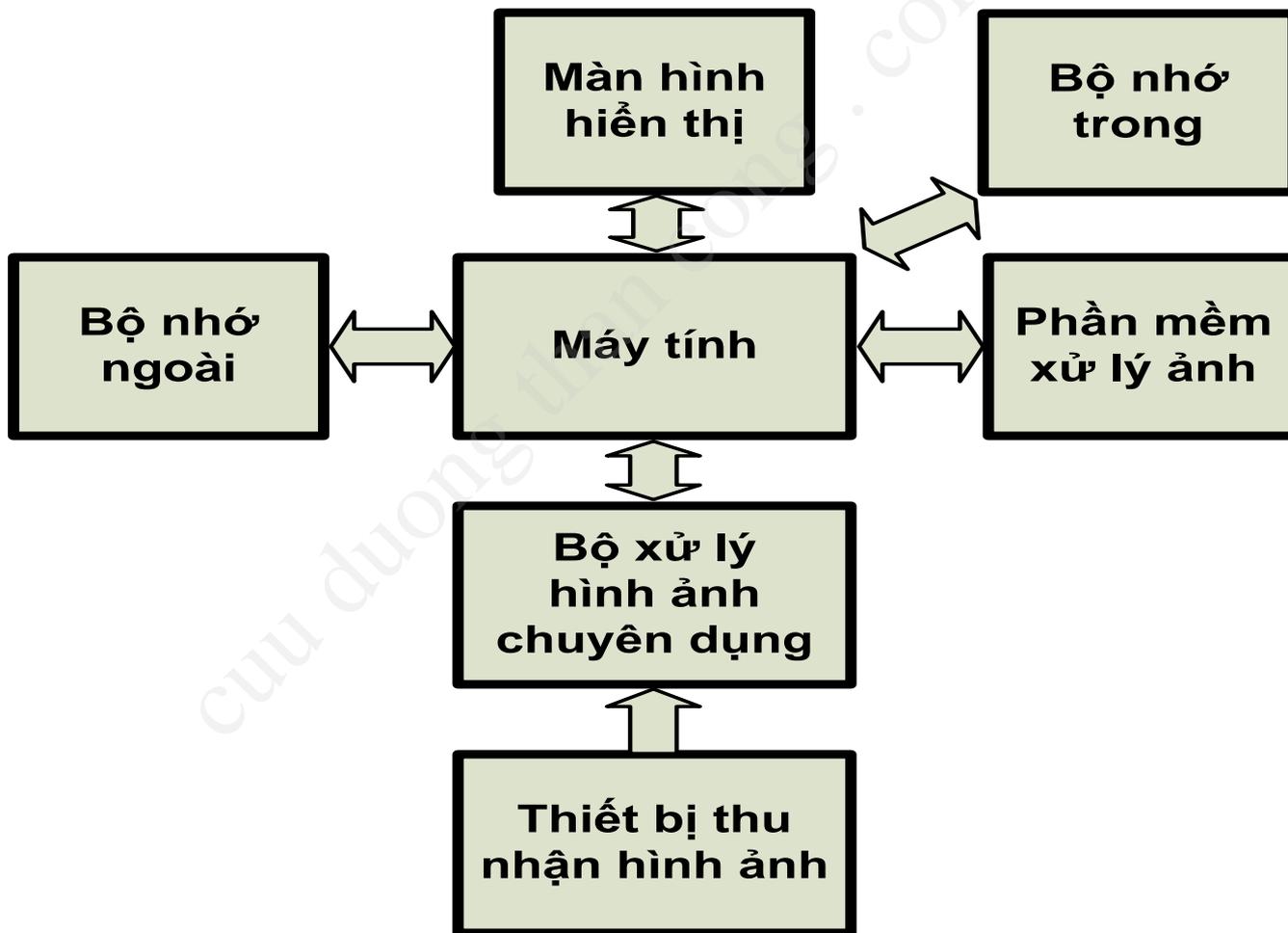
Các giai đoạn chính trong hệ thống xử lý ảnh



Các giai đoạn chính trong hệ thống xử lý ảnh

- + Thu nhận ảnh: - Qua các camera (tương tự, số).
 - Từ vệ tinh qua các bộ cảm ứng (Sensors).
 - Qua các máy quét ảnh (Scanners).
- + Số hóa ảnh: Biến đổi ảnh tương tự thành ảnh rời rạc để xử lý bằng máy tính: Thông qua quá trình lấy mẫu (rời rạc về mặt không gian) và lượng tử hóa (rời rạc về mặt biên độ).
- + Xử lý số: là một tiến trình gồm nhiều công đoạn nhỏ: Tăng cường ảnh (Enhancement), khôi phục ảnh (Restoration), phát hiện biên (Edge Detection), phân vùng ảnh (Segmentation), trích chọn các đặc tính (Feature Extraction)...
- + Hệ quyết định: Tùy mục đích của ứng dụng mà chuyển sang giai đoạn khác là hiển thị, nhận dạng, phân lớp, truyền thông...

Các thành phần chính của hệ thống xử lý ảnh

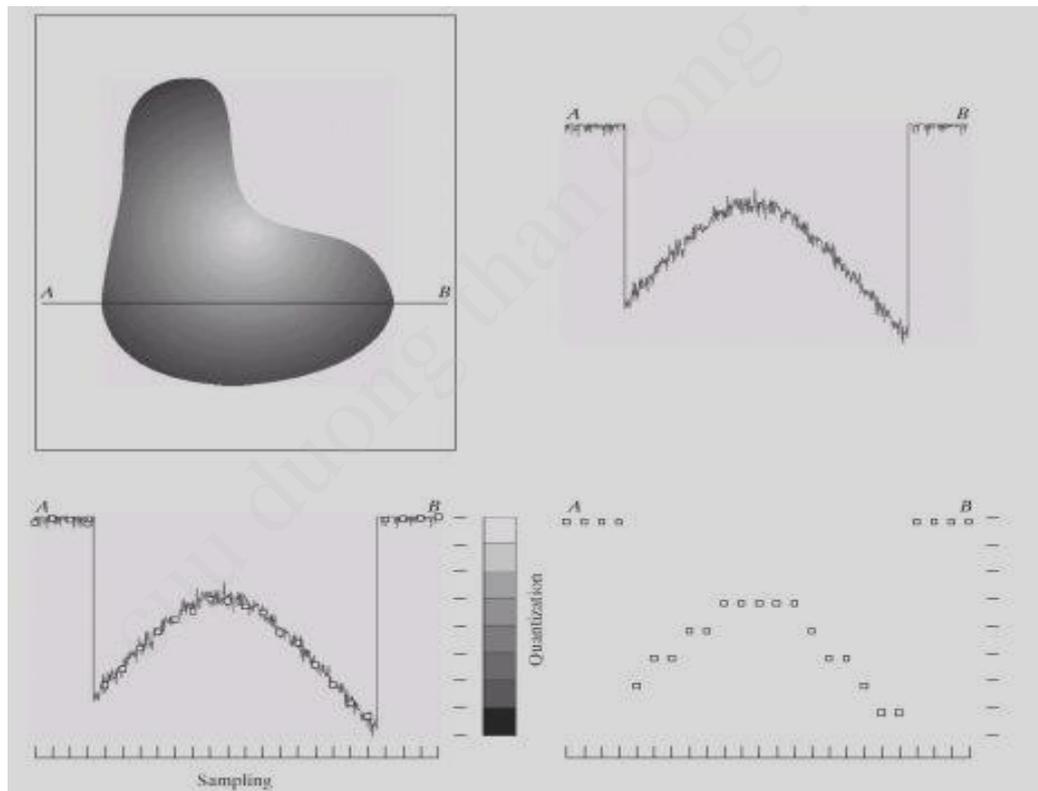


Ứng dụng của xử lý ảnh

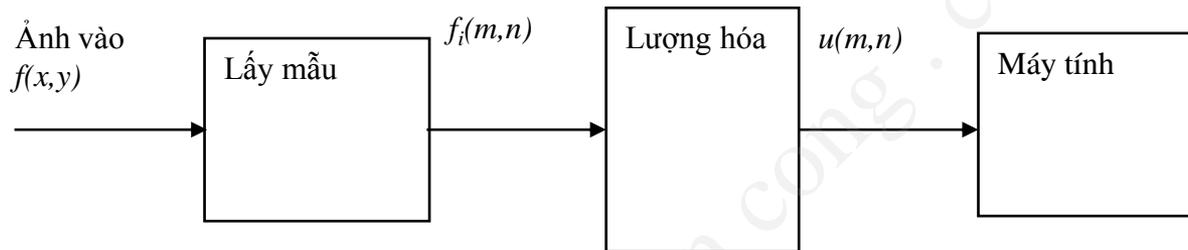
- Trong y học
- Trong lĩnh vực địa chất, hình ảnh nhận được từ vệ tinh có thể được phân tích để xác định cấu trúc bề mặt trái đất. Kỹ thuật làm nổi đường biên (image enhancement) và khôi phục hình ảnh (image restoration) cho phép nâng cao chất lượng ảnh vệ tinh và tạo ra các bản đồ địa hình 3-D với độ chính xác cao.
- Trong ngành khí tượng học, ảnh nhận được từ hệ thống vệ tinh theo dõi thời tiết cũng được xử lý, nâng cao chất lượng và ghép hình để tạo ra ảnh bề mặt trái đất trên một vùng rộng lớn, qua đó có thể thực hiện việc dự báo thời tiết một cách chính xác hơn.
- Xử lý ảnh còn được sử dụng nhiều trong các hệ thống quản lý chất lượng và số lượng hàng hóa trong các dây chuyền tự động, ví dụ như hệ thống phân tích ảnh để phát hiện bọt khí bên vật thể đúc bằng nhựa, phát hiện các linh kiện không đạt tiêu chuẩn (bị biến dạng) trong quá trình sản xuất hoặc hệ thống đếm sản phẩm thông qua hình ảnh nhận được từ camera quan sát.
- Xử lý ảnh còn được sử dụng rộng rãi trong lĩnh vực hình sự và các hệ thống bảo mật hoặc kiểm soát truy cập
- Ngoài ra, có thể kể đến các ứng dụng quan trọng khác của kỹ thuật xử lý ảnh tĩnh cũng như ảnh động trong đời sống như tự động nhận dạng, nhận dạng mục tiêu quân sự, máy nhìn công nghiệp trong các hệ thống điều khiển tự động, nén ảnh tĩnh, ảnh động để lưu và truyền trong mạng viễn thông v.v.

Số hóa ảnh

- Phương pháp chung để số hóa ảnh là lấy mẫu theo hàng và mã hóa từng hàng.



Nguyên tắc số hóa ảnh



- Ảnh vào là ảnh tương tự.
- Tiến trình lấy mẫu thực hiện các công việc sau: Quét ảnh theo hàng, và lấy mẫu theo hàng. Đầu ra là rời rạc về mặt không gian, nhưng liên tục về mặt biên độ.
- Tiến trình lượng hóa: lượng tử hóa về mặt biên độ (độ sáng) cho dòng ảnh vừa được rời rạc hóa.

Lấy mẫu

- Yêu cầu tín hiệu có dải phổ hữu hạn $|f_x| \leq f_{x \max}$
- Ảnh thỏa mãn điều kiện trên, và được lấy mẫu đều trên một lưới hình chữ nhật, với bước nhảy(chu kỳ lấy mẫu) $\Delta x, \Delta y$ sao cho

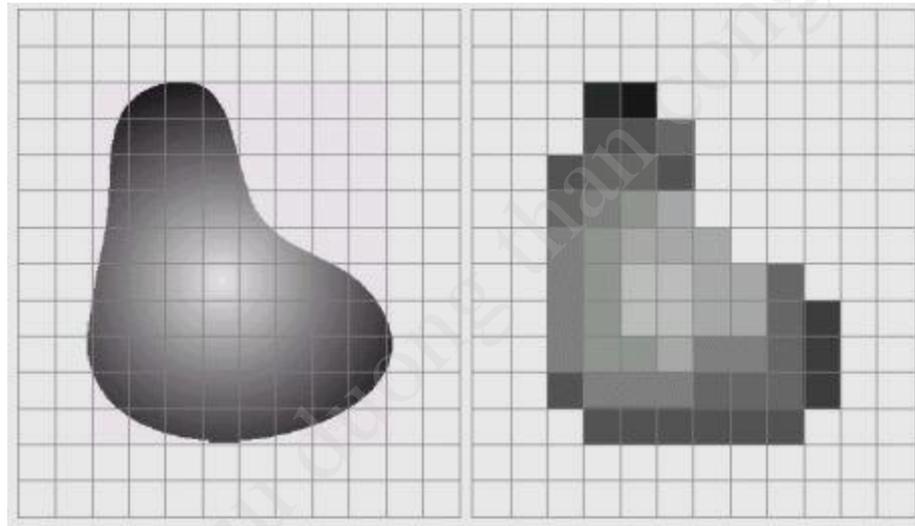
$$\frac{1}{\Delta x} \geq 2 f_{x \max}, \quad \frac{1}{\Delta y} \geq 2 f_{y \max}$$

- Thực tế luôn tồn tại nhiễu ngẫu nhiên trong ảnh, nên có một số kỹ thuật khác được dùng đó là: lưới không vuông, lưới bát giác.

Lượng tử hóa

- Lượng hóa ảnh nhằm ánh xạ từ một biến liên tục u (biểu diễn giá trị độ sáng) sang một biến rời rạc u^* với các giá trị thuộc tập hữu hạn
- Cơ sở lý thuyết của lượng hóa là chia dải độ sáng biến thiên từ L_{min} đến L_{max} thành một số mức (rời rạc và nguyên) - Phải thỏa mãn tiêu chí về độ nhạy của mắt. Thường $L_{min}=0$, L_{max} là số nguyên dạng 2^B (Thường chọn $B=8$, mỗi điểm ảnh sẽ được mã hóa 8 bit).

Ảnh tương tự và Ảnh số hóa



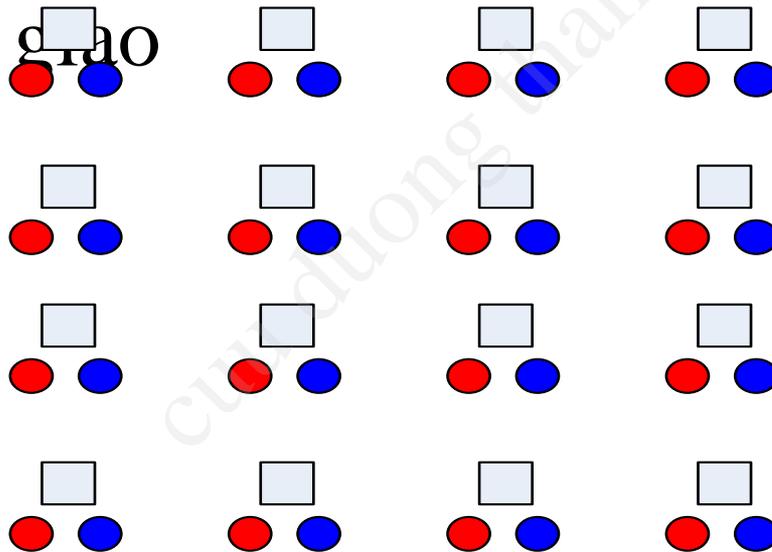
Các tiêu chuẩn lấy mẫu

- Các tiêu chuẩn lấy mẫu video thành phần: có nhiều tiêu chuẩn lấy mẫu theo thành phần, điểm khác nhau chủ yếu ở tỷ lệ giữa tần số lấy mẫu và phương pháp lấy mẫu tín hiệu chói và tín hiệu màu (hoặc hiệu màu):
- đó là các tiêu chuẩn 4:4:4, 4:2:2, 4:2:0, 4:1:1.

Tiêu chuẩn 4:4:4

- Tiêu chuẩn 4:4:4: Tín hiệu chói và màu được lấy mẫu tại tất cả các điểm lấy mẫu trên dòng tích cực của tín hiệu video. Cấu trúc lấy mẫu trực giao

trực giao



Tiêu chuẩn
4:4:4

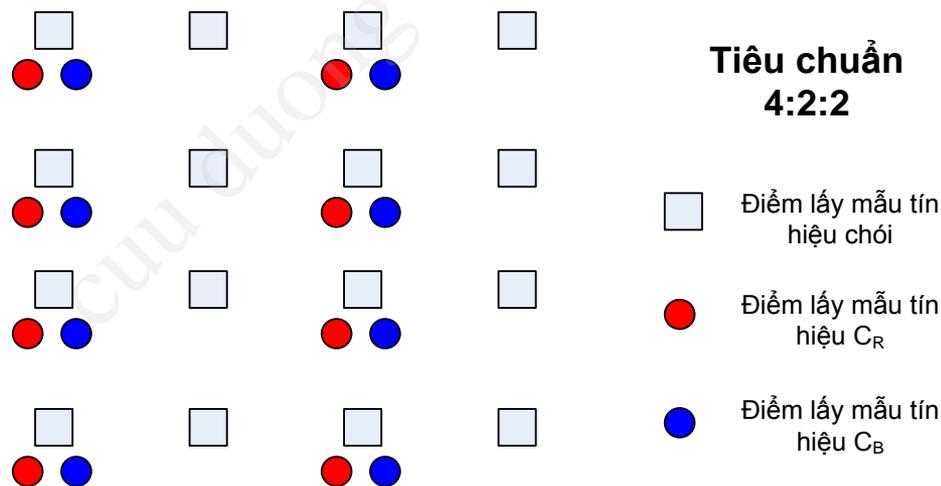
- Điểm lấy mẫu tín hiệu chói
- Điểm lấy mẫu tín hiệu C_R
- Điểm lấy mẫu tín hiệu C_B

Tiêu chuẩn 4:4:4

- ví dụ khi số hóa tín hiệu video có độ phân giải 720x576 (hệ PAL), 8 bit lượng tử /điểm ảnh, 25 ảnh/s luồng dữ liệu số nhận được sẽ có tốc độ : $3 \times 720 \times 576 \times 8 \times 25 = 249 \text{Mbits/s}$

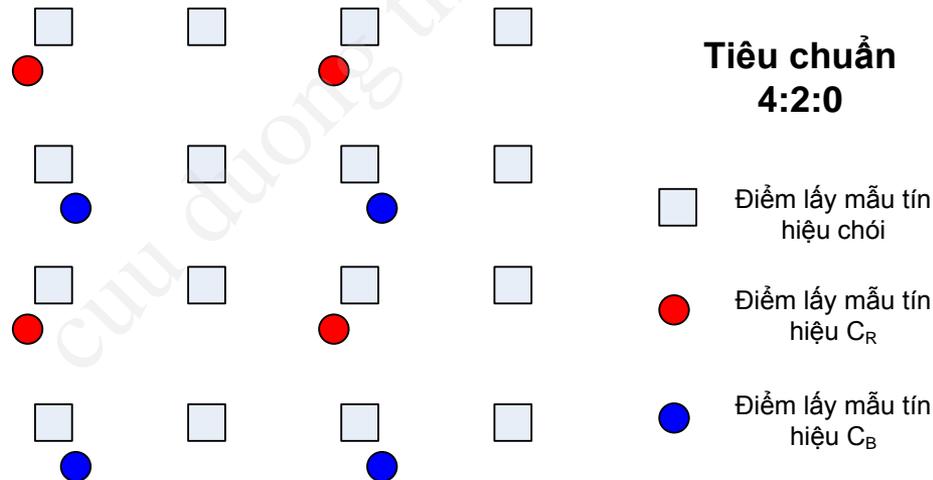
Tiêu chuẩn 4:2:2

- Tín hiệu chói được lấy mẫu tại tất cả các điểm lấy mẫu trên dòng tích cực của tín hiệu video. Tín hiệu màu trên mỗi dòng được lấy mẫu với tần số bằng nửa tần số lấy mẫu tín hiệu chói



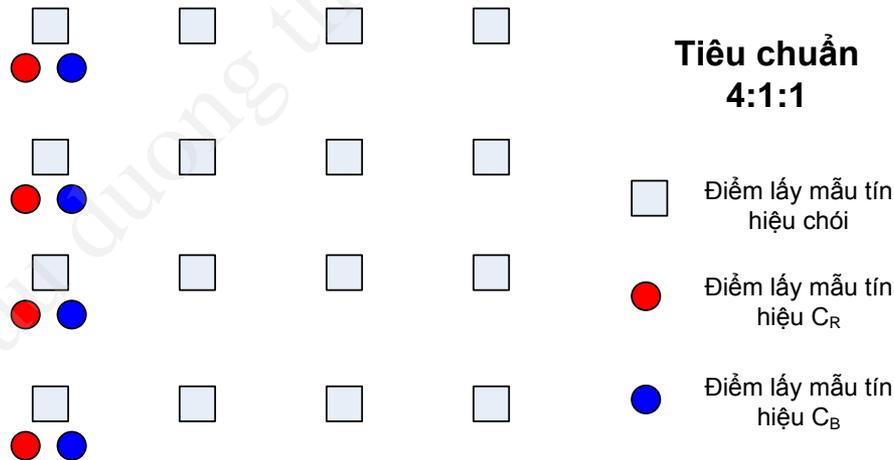
Tiêu chuẩn 4:2:0

- Tín hiệu chói được lấy mẫu tại tất cả các điểm lấy mẫu trên dòng tích cực của tín hiệu video. Cách một điểm lấy mẫu một tín hiệu màu. Tại dòng chẵn chỉ lấy mẫu tín hiệu màu CR, tại dòng lẻ lấy mẫu tín hiệu màu CB. Như vậy, nếu tần số lấy mẫu tín hiệu chói là f_D , Thì tần số lấy mẫu tín hiệu màu sẽ là $f_D/2$.

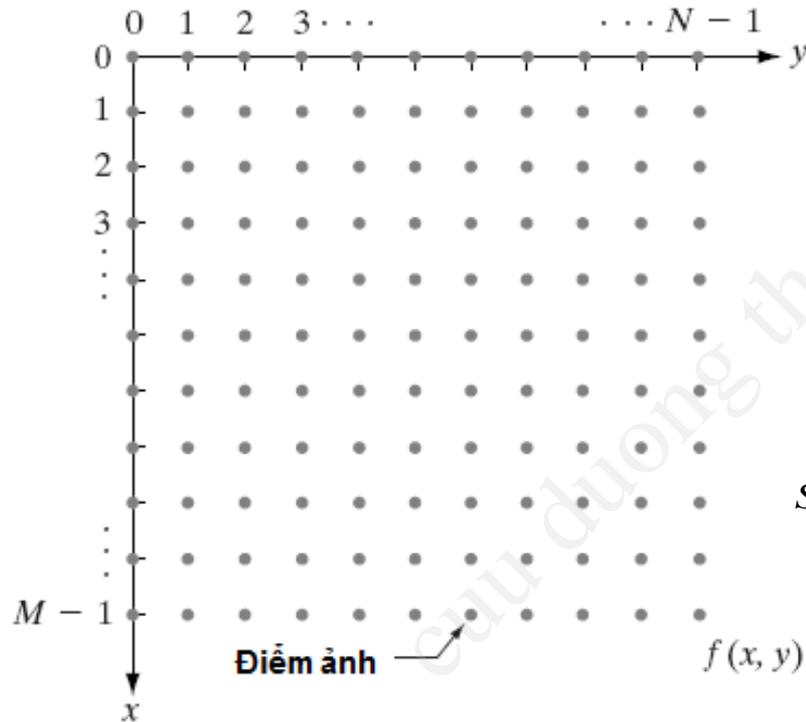


Tiêu chuẩn 4:1:1

- Tín hiệu chói được lấy mẫu tại tất cả các điểm lấy mẫu trên dòng tích cực của tín hiệu video. Tín hiệu màu trên mỗi dòng được lấy mẫu với tần số bằng một phần tư tần số lấy mẫu tín hiệu chói. Như vậy, nếu tần số lấy mẫu tín hiệu chói là f_D , thì tần số lấy mẫu tín hiệu màu CR và CB sẽ là $f_D/4$.



Biểu diễn tín hiệu ảnh số



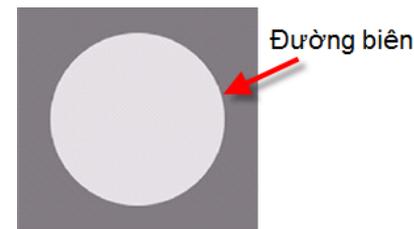
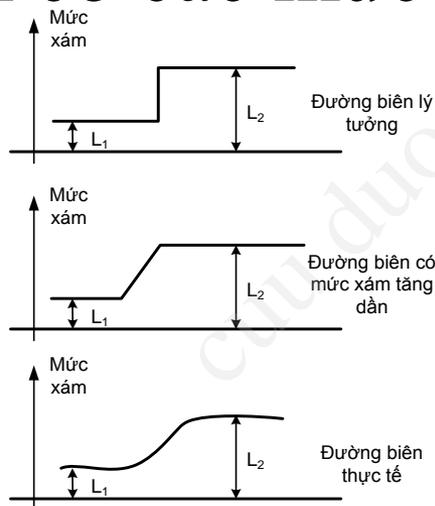
$$f_{x,y} = \begin{bmatrix} f_{0,0} & f_{0,1} & \dots & f_{0,N-1} \\ f_{1,0} & f_{1,1} & \dots & f_{1,N-1} \\ \vdots & \vdots & & \vdots \\ f_{M-1,0} & f_{M-1,1} & \dots & f_{M-1,N-1} \end{bmatrix}$$

$$S(m, n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} S(k, l) \delta(m-k, n-l)$$

Với $0 \leq m, k \leq M-1, 0 \leq n, l \leq N-1$

Các phương pháp xác định và dự đoán biên ảnh

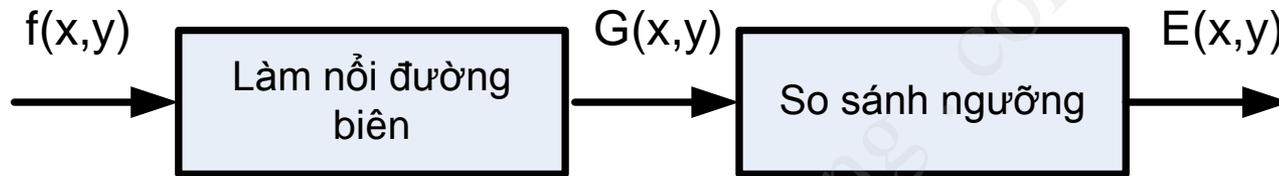
- Đường biên là đường nối các điểm ảnh nằm trong khu vực ảnh có thay đổi đột ngột về độ chói, đường biên thường ngăn cách hai vùng ảnh có các mức xám gần như không đổi.



Phương pháp phát hiện đường biên

- 1- Phương pháp phát hiện đường biên trực tiếp dựa trên các phân tích về sự thay đổi độ chói của ảnh. Kỹ thuật chủ yếu dùng để phát hiện biên là dùng đạo hàm. Khi lấy đạo hàm bậc nhất của ảnh ta có phương pháp gradient, khi lấy đạo hàm bậc hai ta có kỹ thuật Laplace.
- 2- Phương pháp phát hiện đường biên trong ảnh màu: phân tích ảnh màu thành các ảnh đơn sắc (R,G,B) và xác định đường biên trên cơ sở sự thay đổi màu sắc trong các ảnh đơn sắc nói trên.
- 3- Phân tích ảnh thành vùng theo các đặc điểm đặc trưng (thí dụ kết cấu bề mặt (texture)), ranh giới giữa các vùng chính là đường biên của ảnh.

Phương pháp Gradient



- Ảnh gốc $f(x,y)$ được đưa vào khối làm nổi đường biên. Ở đây, bằng phương pháp xử lý tuyến tính hoặc phi tuyến ảnh $F(x,y)$ được làm tăng mức chênh lệch độ chói giữa các vùng ảnh. Ảnh $G(x,y)$ là ảnh gốc đã được tăng cường biên độ đường biên giữa các vùng ảnh. Sau đó, tại khối so sánh, người ta so sánh giá trị các điểm ảnh $G(x,y)$ với mức ngưỡng T để xác định vị trí các điểm có mức thay đổi độ chói lớn.

Nén ảnh

- Nén Ảnh là gì?
- Mục đích của nén ảnh
- Các khái niệm
- Nén Ảnh
- Các chuẩn ảnh nén hiện nay

NÉN ẢNH LÀ GÌ?

- Nén là quá trình làm giảm thông tin dư thừa trong dữ liệu.

1600x1200



5,7MB

Nén



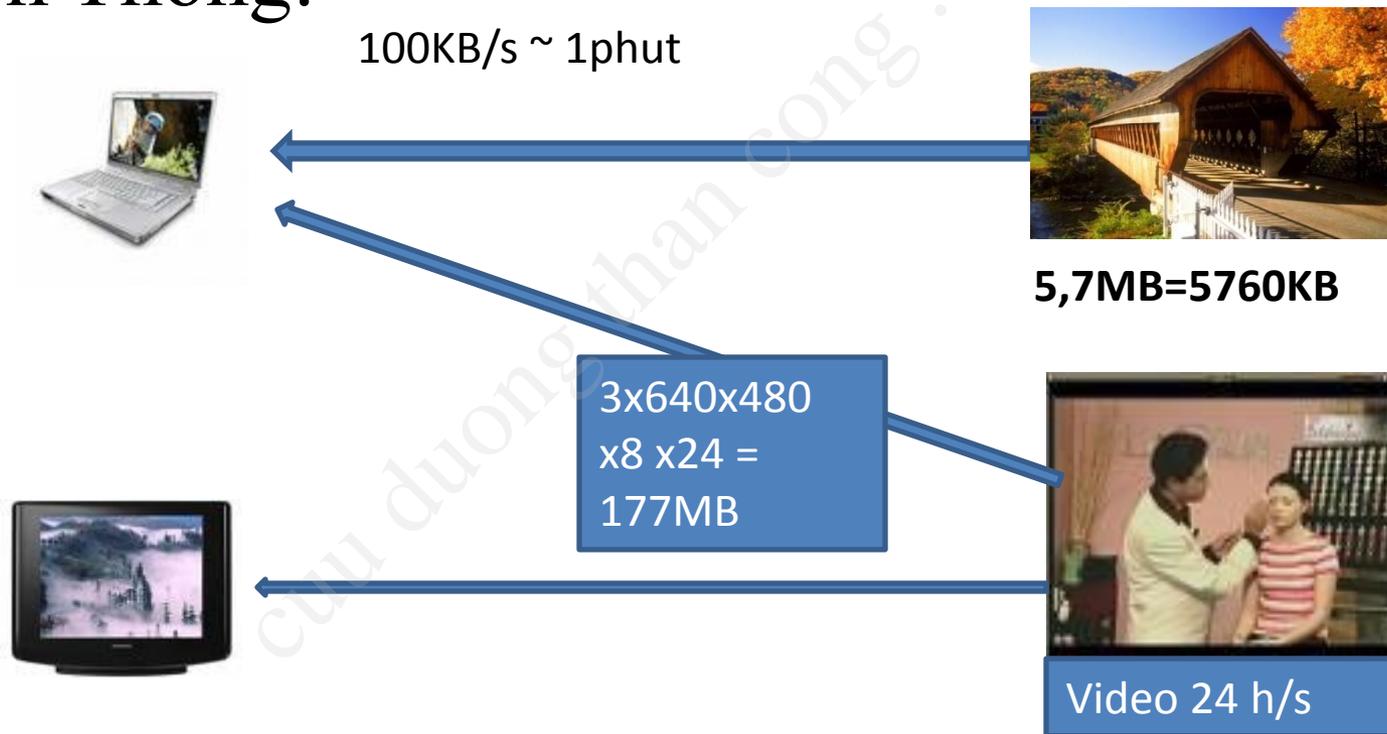
1600x1200



406KB

MỤC ĐÍCH CỦA VIỆC NÉN ẢNH

- Truyền Thông:



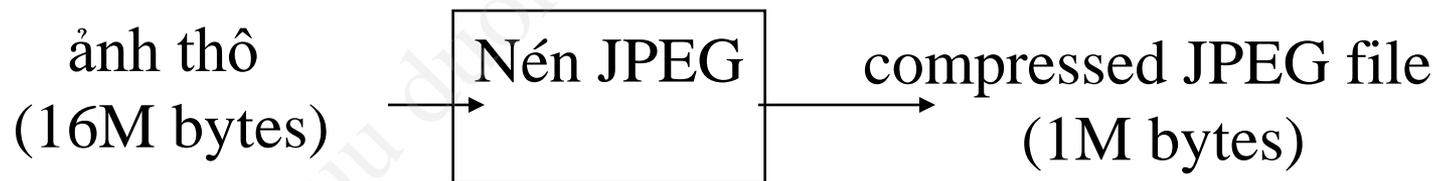
Nén ảnh

- Tại sao lại cần nén ảnh?

- Ví dụ: camera số (4Mpixel)

Dữ liệu ban đầu – 24bits, 5.3M pixels → 16M bytes

256M memory card (\$30-50) → 16 pictures



compression ratio=16 → 256 pictures

Ví dụ

High-definition television (HDTV) generates images with a resolution of 1125 horizontal TV lines interlaced (where every other line is painted on the tube face in each of two fields, each field being $1/60$ th of a second in duration). The width-to-height aspect ratio of the images is 16:9. The fact that the horizontal lines are distinct fixes the vertical resolution of the images. A company has designed an image capture system that generates digital images from HDTV images. The resolution of each TV (horizontal) line in their system is in proportion to vertical resolution, with the proportion being the width-to-height ratio of the images. Each pixel in the color image has 24 bits of intensity resolution, 8 pixels each for a red, a green, and a blue image. These three “primary” images form a color image. How many bits would it take to store a 2-hour HDTV program?

Ví dụ

The width-to-height ratio is $16/9$ and the resolution in the vertical direction is 1125 lines (or, what is the same thing, 1125 pixels in the vertical direction). It is given that the resolution in the horizontal direction is in the $16/9$ proportion, so the resolution in the vertical direction is $(1125) \times (16/9) = 2000$ pixels per line. The system “paints” a full 1125×2000 , 8-bit image every $1/30$ sec for each of the red, green, and blue component images. There are 7200 sec in two hours, so the total digital data generated in this time interval is $(1125)(2000)(8)(30)(3)(7200) = 1.166 \times 10^{13}$ bits, or 1.458×10^{12} bytes (i.e., about 1.5 terrabytes).



Nén để giảm không gian lưu trữ, tiết kiệm băng thông

CÁC KHÁI NIỆM

Một số tham số được dùng để đánh giá chất lượng của ảnh nén

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} ||I(i, j) - K(i, j)||^2$$

Peak Signal-to-Noise Ratio(PSNR):

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right)$$

CÁC KHÁI NIỆM

- Tỷ số nén: Là đặc trưng của mọi phương pháp nén.

$$C_R = \frac{n_1}{n_2}.$$

- Dư thừa dữ liệu: $R_D = 1 - \frac{1}{C_R}$
 - Sự phân bố kí tự
 - Sự lặp lại ký tự.
 - Sự phân bố của các chuỗi ký tự .
 - Dư thừa trong pixel (Interpixel Redundancy)
 - Dư thừa tâm lý thị giác

DƯ THỪA DỮ LIỆU

- Sự phân bố ký tự
 - Một số ký tự (pixel) xuất hiện với tần suất lớn hơn so với các ký tự khác trong dữ liệu gốc (ảnh). Ta có thể thay thế các ký tự này bằng từ mã nhị phân ít bit hơn và các ký tự xuất hiện nhiều hơn bằng từ mã nhị phân có nhiều bit hơn.
 - Dùng mã hóa Huffman để mã hóa loại dư thừa này.

DƯ THỪA DỮ LIỆU

- Sự lặp lại các ký tự
 - Một chuỗi các ký tự (bit 1 hoặc 0) được lặp lại nhiều lần. Ta có thể mã hóa chuỗi lặp đó bằng ít bit hơn. Đây chính là nguyên tắc hoạt động của mã hóa RLC (Run-Length Coding)

DƯ THỪA DỮ LIỆU

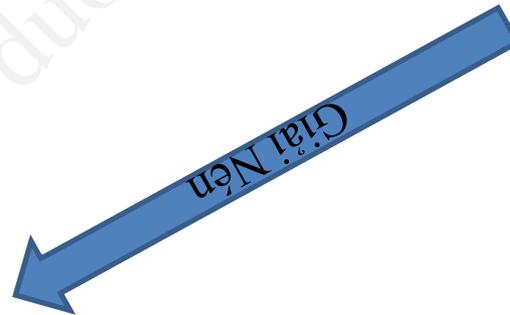
- Sự lặp lại của các chuỗi ký tự:
 - Một số chuỗi ký tự (pixel) có tần suất xuất hiện tương đối cao.
 - Có thể mã hóa các chuỗi đó bằng từ mã ít bit hơn .
 - Để xử lý loại dư thừa này ta sử dụng phương pháp mã hóa LWZ (mã hóa kiểu từ điển) .

Nén tổn thất và nén không tổn thất

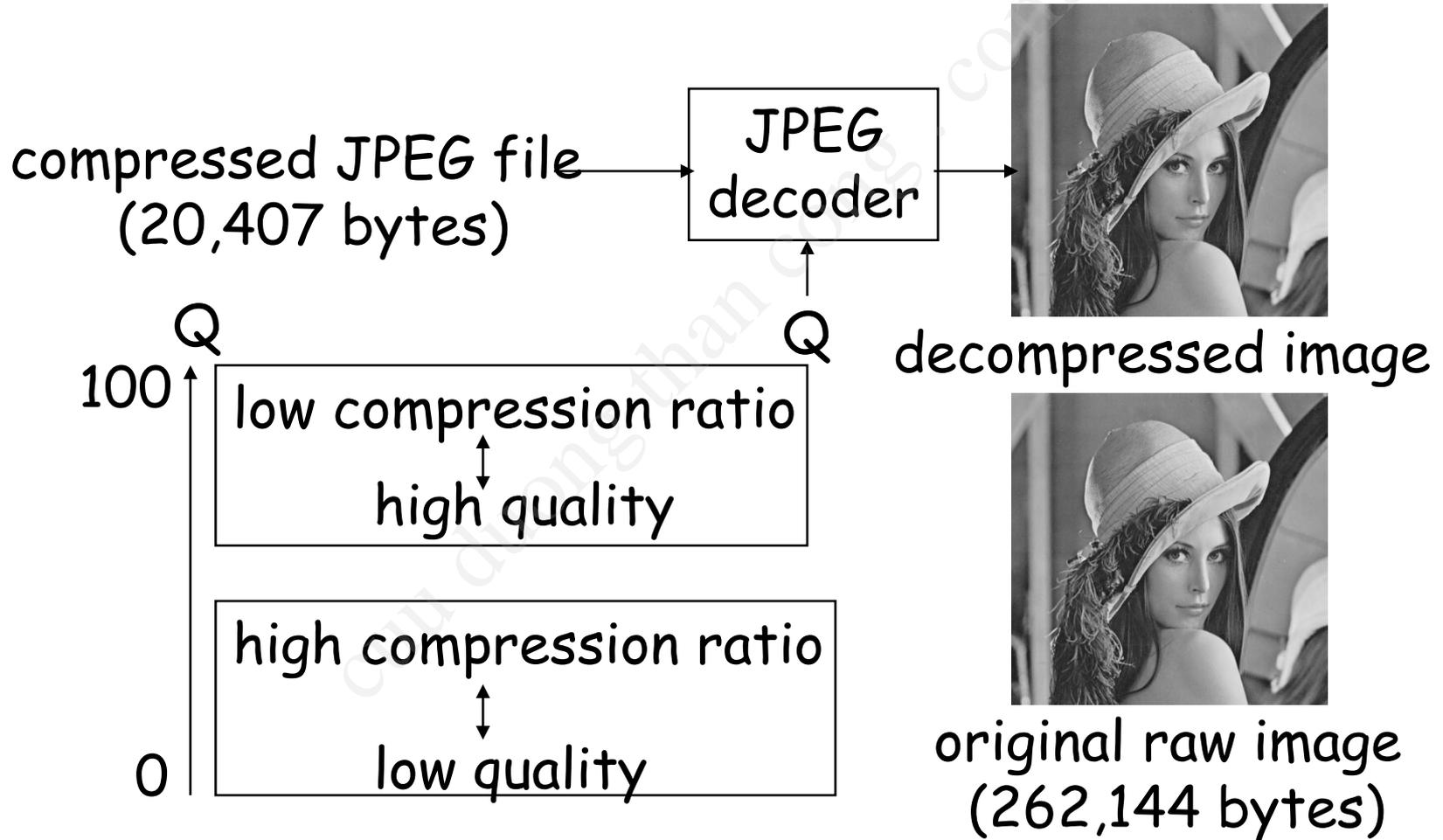
- Nén mất thông tin và không mất thông tin
 - Nén mất thông tin:



ẢNH NÉN

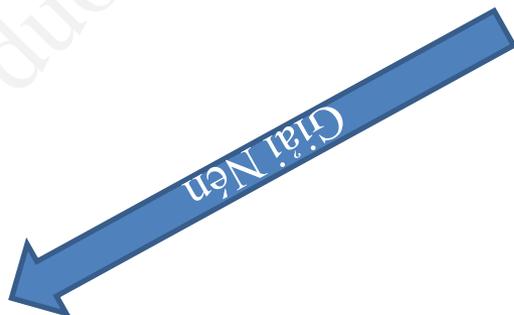


Nén ảnh có tổn thất



Nén tổn thất và nén không tổn thất

- Nén mất thông tin và không mất thông tin
 - Nén không mất thông tin:



ẢNH NÉN

Nén ảnh không tổn thất

- Định nghĩa
 - Ảnh sau khi giải nén giống hoàn toàn với ảnh gốc (zero error)
- Hệ số nén (Compression ratio)
 - Phụ thuộc lớn vào kiểu ảnh và nội dung ảnh

ảnh nhân tạo	ảnh tự nhiên
>10	1~3
- Ứng dụng
 - Lưu trữ và truyền các ảnh y học

Các kỹ thuật nén ảnh không tổn thất phổ biến

- WinZip
 - Dựa trên thuật toán Lempel-Ziv được phát minh cách đây 30 năm
- GIF (Graphic Interchange Format)
 - Dựa trên thuật toán LZ nâng cao, tạo ra bởi Welch năm 1983
- PNG (Portable Network Graphics)

u



□

y.

□

ch.

□

u.

entropy (mã hóa không tổn thất)

- Mã hóa loại dài chạy (RLC-Run Length Coding):
(độ dài của chuỗi mã hóa thông tin (độ dài, vị trí)).
- Mã hóa nhị phân: các bit thay thế bằng cơ sở 2.
- Mã hóa nhị phân thay thế: đây là mã hóa nhị phân thay thế các bit bằng các bit.
- Mã hóa với độ dài (của từ mã) thay đổi (VLC-Variable-Length Coding)

n

(mã hóa tồn thất)

-
-
-
-
-

ch nghi.

C.

nh

- a ITU-TS (the International Telecommunications Union-Telecommunication Sector)
ISO (International Standards Organization)
- - - - -
audio.
- nx 64Kbps.
- - i
64Kbps.

entropy (mã hóa không tổn thất)

- Run length Coding (RLC)
- Huffman
- Lempel Ziv – Wench (LZW)

RUN LENGTH CODING- RLC

- Tư tưởng của phương pháp này là dựa trên sự lặp lại các bit.
- Thay thế các bit đó bởi chiều dài chuỗi và bit lặp .
- Để phân biệt với các ký tự khác ta có thể thêm 1 từ mã đặc biệt trước 2 thông tin chiều dài chuỗi và bit lặp .
- **Ví Dụ: Ta có 1 dãy các giá trị mức xám như sau**
 - 55 22 22 22 22 22 22 22 22 22 51 52 52 52 60 ...
 - Ta có thể thay đoạn mã trên bằng
 - 55 E 8 22 51 E 3 52 60 ...
 - Với E là ký tự đặc biệt , giá trị sau E là chiều dài ký tự lặp và ký tự lặp .

HUFFMAN

- Dựa vào mô hình thống kê tính tần suất xuất hiện các ký tự.
- Gán cho các ký tự có tần suất cao bằng một từ mã ngắn, các ký tự có tần suất thấp bằng một từ mã dài.

□ Thuật toán:

□ Bước 1:

- Tính tần suất xuất hiện các ký tự trong dữ liệu gốc
- Xây dựng bảng mã
- Sắp xếp lại bảng mã theo thứ tự tần suất giảm dần

□ Bước 2:

- Tạo cây huffman

HUFFMAN

Bảng tần xuất sắp xếp theo thứ tự giảm dần

<i>Ký tự</i>	<i>Tần suất</i>	<i>Ký tự</i>	<i>Tần suất</i>	<i>xác suất</i>
"1"	152	"0"	1532	0.2770
"2"	323	"6"	602	0.1088
"3"	412	","	536	0.0969
"4"	226	" "	535	0.0967
"5"	385	"3"	112	0.0746
"6"	602	"5 "	385	0.0696
"7"	92	"2"	323	0.0585
"8"	112	"_ "	315	0.0569
"9"	87	"4"	226	0.0409
"0"	1532	"+"	220	0.0396
","	536	"1"	152	0.0275
"+"	220	"8"	112	0.0203
"_ "	315	"7"	92	0.0167
" "	535	"0"	87	0.0158

Mã LZW

- Được Jacob Braham Ziv đưa ra lần đầu tiên năm 1977, sau đó phát triển thành một họ giải thuật nén từ điển là **LZ**.
- Năm 1984, Terry Welch cải tiến giải thuật LZ thành một giải thuật tốt hơn :LZW
- Dùng để giảm dư thừa trong pixel
- Không cần biết trước xác suất phân bố của các pixel
- Thường được dùng để nén các loại văn bản, ảnh đen trắng, ảnh màu, ảnh đa mức xám... Và là chuẩn nén cho các dạng ảnh GIF và TIFF.

Mã LZW

- Phương pháp :
- Xây dựng 1 từ điển

Cấu trúc
từ điển

0	0
1	1
...	...
...	...
255	255
256	256
257	257
258	Chuỗi
259	Chuỗi
...	...
...	...
4095	Chuỗi

(Clear Code)

(End Of Information)

Mã LZW

- Từ điển được xây dựng đồng thời với quá trình đọc dữ liệu. Sự có mặt của một chuỗi con trong từ điển khẳng định rằng chuỗi đó đã từng xuất hiện trong phần dữ liệu đã đọc.
- Thuật toán liên tục “tra cứu ” và cập nhật từ điển sau mỗi lần đọc một kí tự ở dữ liệu đầu vào.
- Do kích thước bộ nhớ không phải vô hạn và để đảm bảo tốc độ tìm kiếm, từ điển chỉ giới hạn 4096 ở phần tử dùng để lưu lớn nhất là 4096 giá trị của các từ mã. Như vậy độ dài lớn nhất của mã là 12 bít($4096 = 2^{12}$).

Ví dụ

Cho ma trận ảnh 4x4 8bit

39	39	126	126
39	39	126	126
39	39	126	126
39	39	126	126

LEMPER ZIV – WENCH (LZW)

Ví dụ: bảng mã hóa và xây dựng tự điển

Currently Recognized Sequence	Pixel Being Processed	Encoded Output	Dictionary Location (Code Word)	Dictionary Entry
	39			
39	39	39	256	39-39
39	126	39	257	39-126
126	126	126	258	126-126
126	39	126	259	126-39
39	39			
39-39	126	256	260	39-39-126
126	126			
126-126	39	258	261	126-126-39
39	39			
39-39	126			
39-39-126	126	260	262	39-39-126-126
126	39			
126-39	39	259	263	126-39-39
39	126			
39-126	126	257	264	39-126-126
126		126		

Kết quả: input: $16 \times 8 = 128 \text{bit}$
output : $5 \times 8 + 5 \times 9 = 85 \text{bit}$
tỷ số nén $128/85 = 1,5$

Mã LZW

- Ví dụ cơ chế nén LZW

Cho chuỗi ban đầu là
“**ABCBCABCABCD**”

(Mã ASCII của A là 65, B là 66,
C là 67).

Từ điển ban đầu đã gồm 256 kí tự cơ bản.

Chuỗi đầu ra sẽ là:

65 - 66 - 67 - 259 - 258 - 67 - 262

Đầu vào có kích thước : $12 \times 8 = 96$
bits.

Đầu ra có kích thước là: $4 \times 8 + 3 \times 9 = 59$
bits Tỷ số nén là: $96:59 \cong 1,63$.

Đầu vào	Đầu ra	Thực hiện
A (65)		A đã có trong từ điển \Rightarrow Đọc tiếp
B (66)	65	Thêm vào từ điển mã 258 đại diện cho mã AB
C (67)	66	Thêm vào từ điển mã 259 đại diện cho mã BC
B	67	Thêm vào từ điển mã 260 đại diện cho mã CB
C		BC đã có trong từ điển \Rightarrow Đọc tiếp
A	259	Thêm vào từ điển mã 261 đại diện cho mã BCA
B		AB đã có trong từ điển \Rightarrow Đọc tiếp
C	258	Thêm vào từ điển mã 262 đại diện cho mã ABC
A	67	Thêm vào từ điển mã 263 đại diện cho mã CA
B		AB đã có trong từ điển \Rightarrow Đọc tiếp
C		ABC đã có trong từ điển \Rightarrow Đọc tiếp
D	262	Thêm vào từ điển mã 263 đại diện cho mã ABCD

Bài tập

- Dùng mã hóa LZ để mã hóa chuỗi ACCBCABCABACD tính tỉ lệ nén và hiệu suất nén
- Tỉ lệ nén $\frac{1}{c_r} \times \%$
- Hiệu suất nén $(1 - \frac{1}{c_r}) \times \%$

Biến đổi Cosin và chuẩn JPEG

$$2-D DCT = X_{k,l} = \frac{c(k)c(l)}{4} \sum_{m=0}^7 \sum_{n=0}^7 x_{m,n} \cos\left(\frac{(2m+1)k\pi}{16}\right) \cos\left(\frac{(2n+1)l\pi}{16}\right)$$

Trong đó $k, l = 0, 1, \dots, 7$

$$2-D IDCT = x_{m,n} = \sum_{k=0}^7 \sum_{l=0}^7 \frac{c(k)c(l)}{4} X_{k,l} \cos\left(\frac{(2m+1)k\pi}{16}\right) \cos\left(\frac{(2n+1)l\pi}{16}\right)$$

Trong đó $m, n = 0, 1, \dots, 7$

$$\text{và } c(k), c(l) = \begin{cases} 1/\sqrt{2}, & k \text{ \& } l = 0 \\ 1, & k^2 + l^2 \neq 0 \end{cases}$$

Thuật toán để tính 2-D DCT và IDCT là: thực hiện phép biến đổi 1-D lần lượt cho hàng rồi đến cột của ma trận.

Trong đó

$$X_{0,0} = \frac{1}{8} \sum_{m=0}^7 \sum_{n=0}^7 x(m, n)$$

Biến đổi Cosin và chuẩn JPEG

$$X_{(k,l)} = \frac{2c(k)c(l)}{N} \sum_{m=0}^7 \sum_{n=0}^7 x_{m,n} \cos\left(\frac{(2m+1)k\pi}{2N}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right)$$

Chuẩn JPEG

-

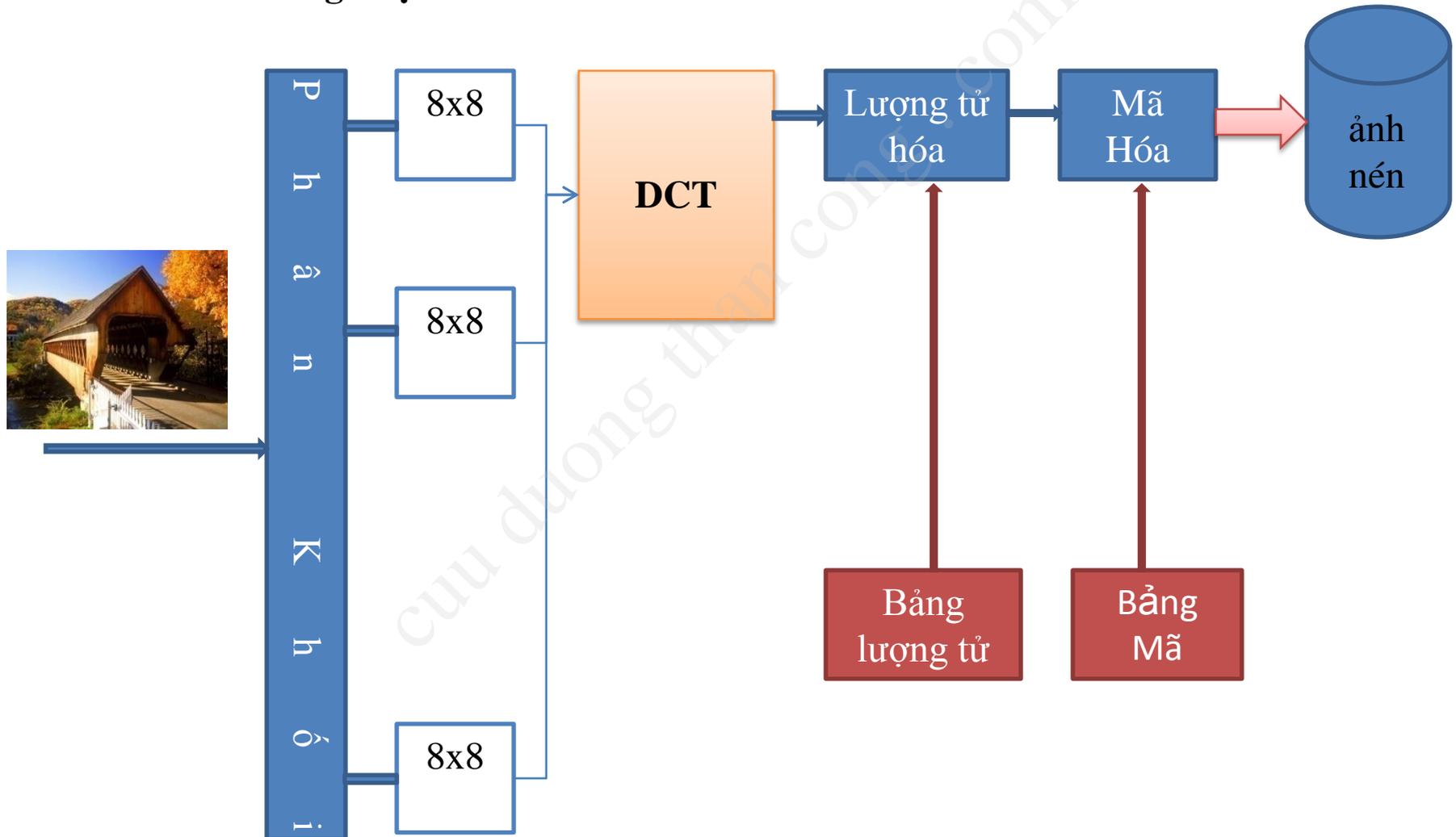
V.V...

-

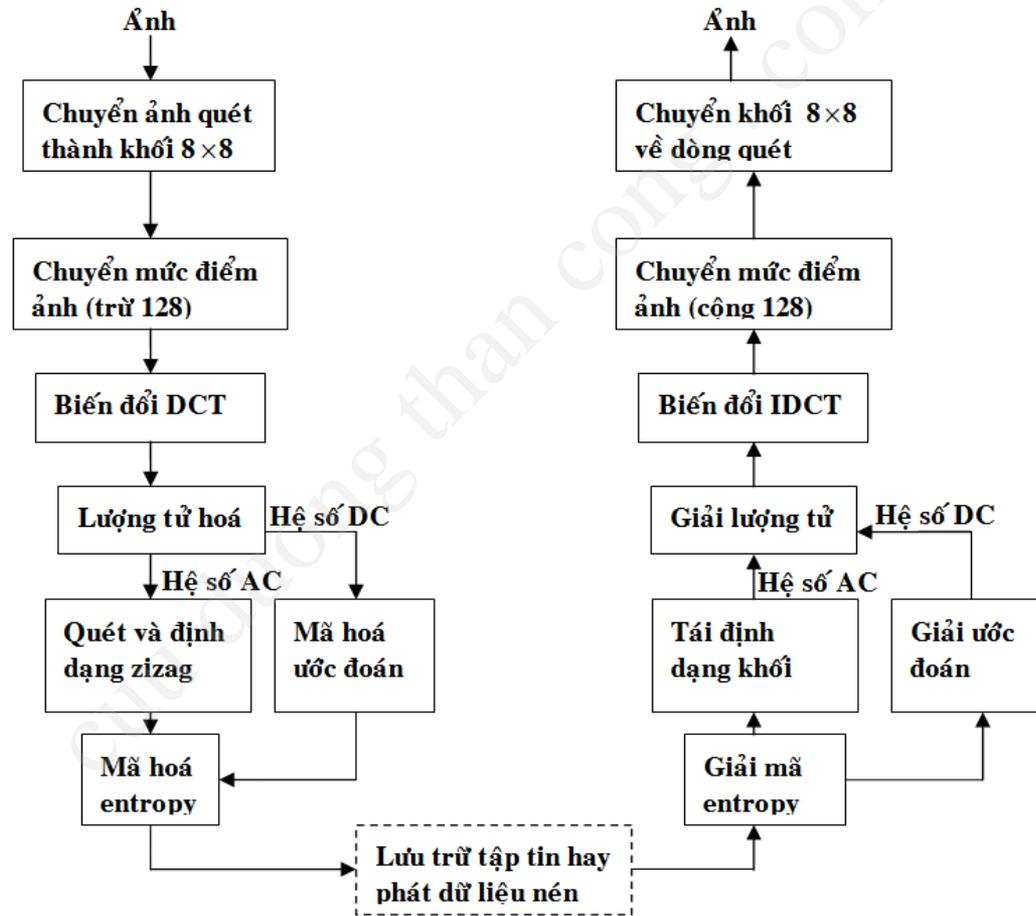
n MPEG.

Chuẩn JPEG

Các công đoạn Nén ảnh JPEG



Chuẩn JPEG



Chuyển ảnh thành các MB

- nh $8\times$

do sau:

$8\times$

- 8.

- c block tăng.

Chuyển mức điểm ảnh (trừ 128)

Ví Dụ:

$$\begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 55 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix} \xrightarrow{-128} \begin{bmatrix} -76 & -73 & -67 & -62 & -58 & -67 & -64 & -55 \\ -65 & -69 & -73 & -38 & -19 & -43 & -59 & -56 \\ -66 & -69 & -60 & -15 & 16 & -24 & -62 & -55 \\ -65 & -70 & -57 & -6 & 26 & -22 & -58 & -59 \\ -61 & -67 & -60 & -24 & -2 & -40 & -60 & -58 \\ -49 & -63 & -68 & -58 & -51 & -60 & -70 & -53 \\ -43 & -57 & -64 & -69 & -73 & -67 & -63 & -45 \\ -41 & -49 & -59 & -60 & -63 & -52 & -50 & -34 \end{bmatrix}$$

Biến đổi DCT

Ví Dụ:

Biến đổi DCT và làm tròn các hệ số



$$\begin{bmatrix} -415 & -30 & -61 & 27 & 56 & -20 & -2 & 0 \\ 4 & -22 & -61 & 10 & 13 & -7 & -9 & 5 \\ -47 & 7 & 77 & -25 & -29 & 10 & 5 & -6 \\ -49 & 12 & 34 & -15 & -10 & 6 & 2 & 2 \\ 12 & -7 & -13 & -4 & -2 & 2 & -3 & 3 \\ -8 & 3 & 2 & -6 & -2 & 1 & 4 & 2 \\ -1 & 0 & 0 & -2 & -1 & -3 & 4 & -1 \\ 0 & 0 & -1 & -4 & -1 & 0 & 1 & 2 \end{bmatrix}$$

a

Ví dụ:

Sử dụng Ma trận lượng tử hóa (Q)

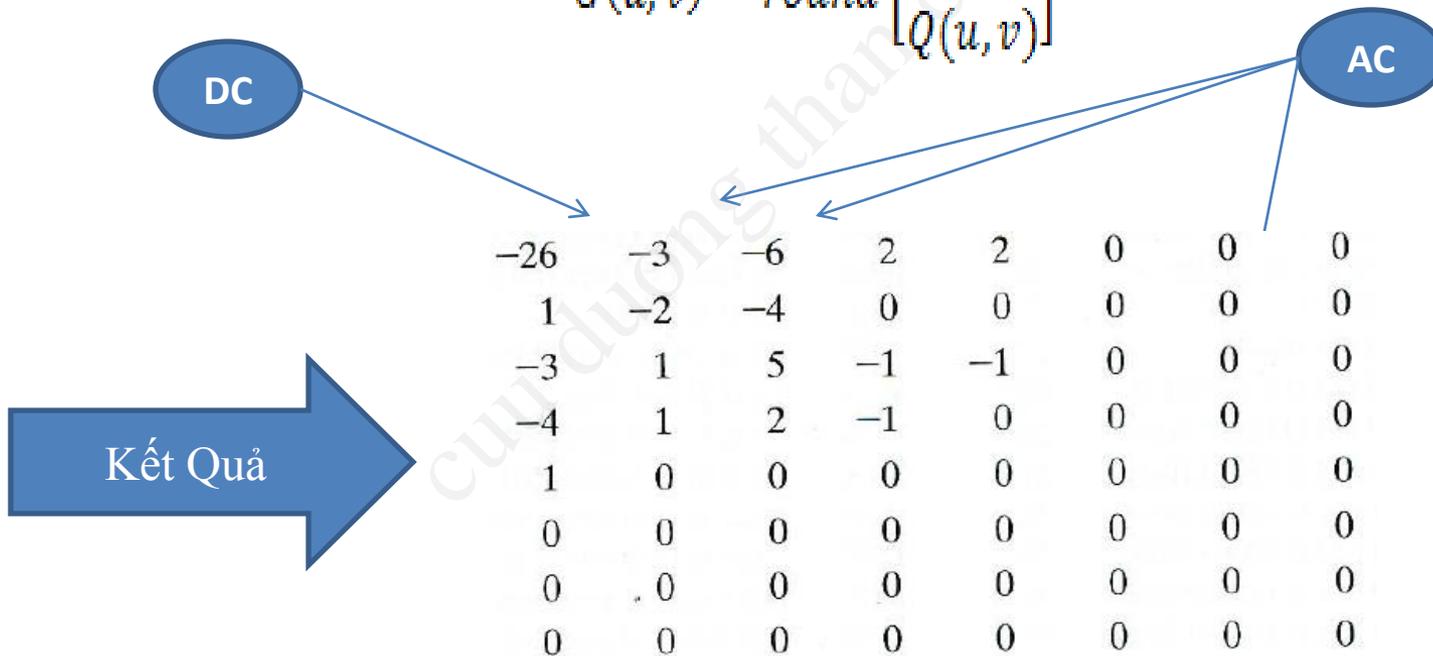
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

a

Ví Dụ (tt):

Chia các phần tử của ma trận DCT với các phần tử tương ứng của ma trận lượng tử hóa ở trên theo công thức:

$$G(u, v) = \text{round} \left[\frac{T(u, v)}{Q(u, v)} \right]$$



$$G(0, 0) = \text{round} \left[\frac{T(0, 0)}{Q(0, 0)} \right] = \text{round} \left[\frac{-415}{16} \right] = -26$$

Mã hóa AC

Sử Dụng Mã hóa Huffman kết thúc sớm chuỗi khi các ký tự còn lại là 0

Ký tự đặc biệt là EOB



Kết Quả

[-26 -3 1 -3 -2 -6 2 -4 1 -4 1 1 5 0 2 0 0 -1 2 0 0 0 0 0 -1 -1 EOB]

Giả sử nếu DC của khối trước là -17

Mã hóa nhị phân

1010110 0100 001 0100 0101 100001 0110 100011 001 100011 001
001 100101 11100110 110110 0110 11110100 000 1010

Ví dụ

qdct =

2	5	0	-2	0	-1	0	0
9	1	-1	2	0	1	0	0
14	1	-1	0	-1	0	0	0
3	-1	-1	-1	0	0	0	0
2	-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

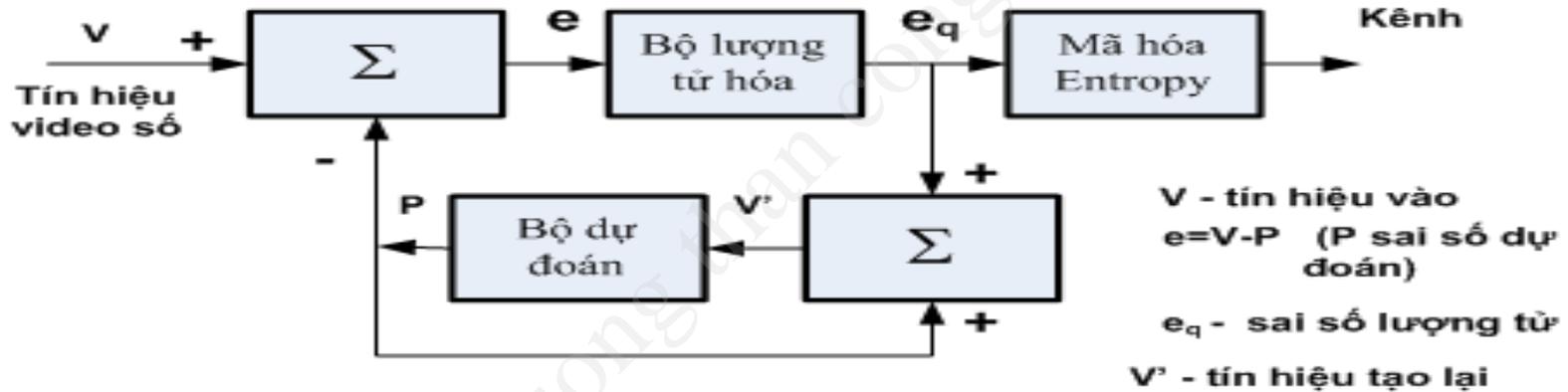
Run-length symbol representation:

{2,(0,5),(0,9),(0,14),(0,1),(1,-2),(0,-1),(0,1),(0,3),(0,2),(0,-1),(0,-1),(0,2),(1,-1),(2,-1), (0,-1), (4,-1),(0,-1),(0,1),EOB}

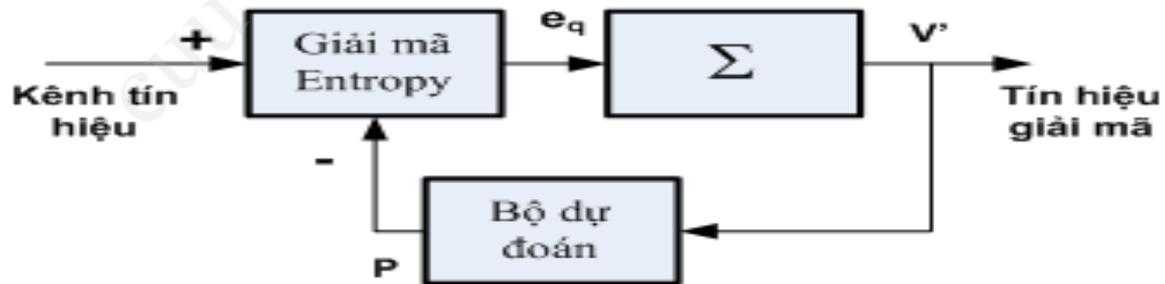
EOB: End of block, one of the symbol that is assigned a short Huffman codeword

Mã hóa DC

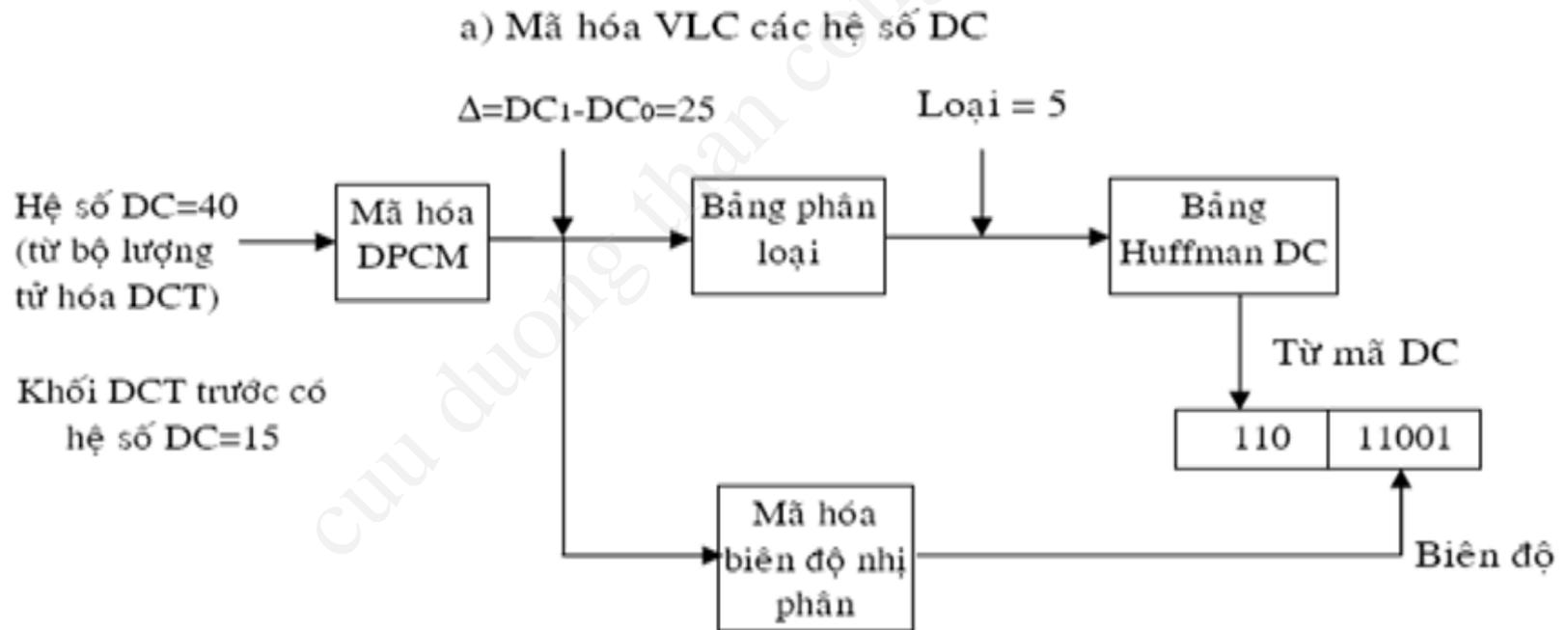
Mã hóa DPCM



Giải mã DPCM



Mã hóa DC



Mã hóa DC

Example:

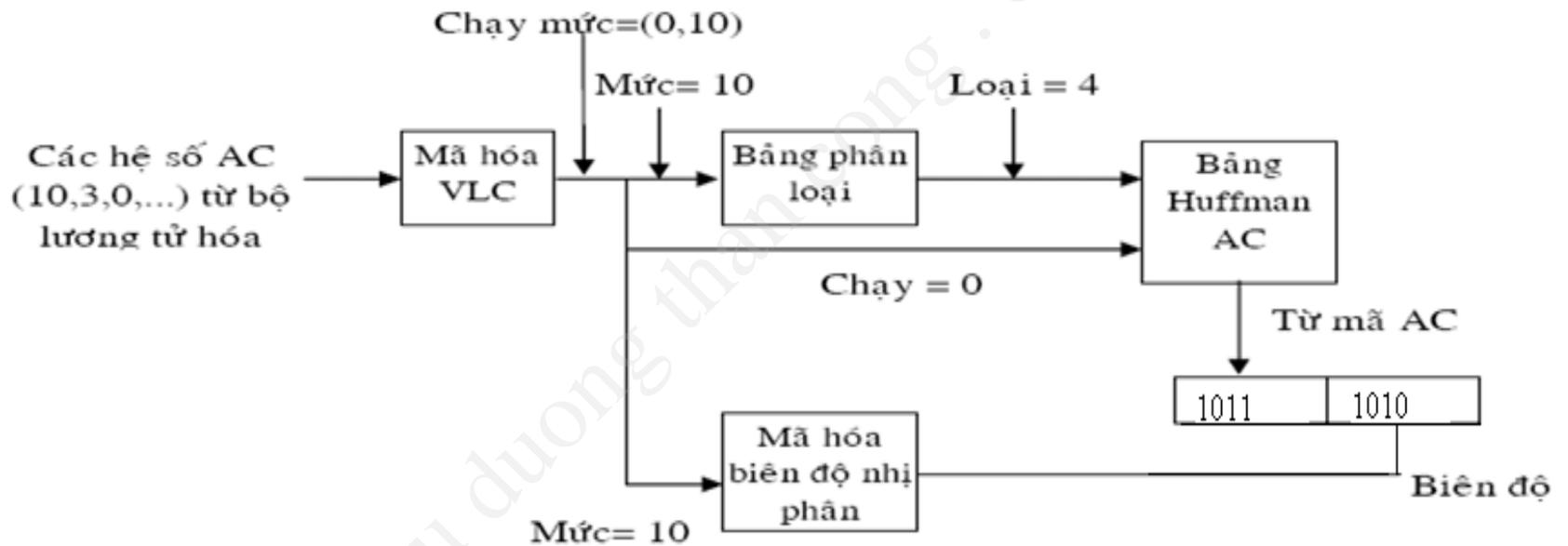
- Current quantized DC index: 2
- Previous block DC index: 4
- Prediction error: -2
- The prediction error is coded in two parts:
 - Which category it belongs to (Table of JPEG Coefficient Coding Categories), and code using a Huffman code (JPEG Default DC Code)
 - DC= -2 is in category “2”, with a codeword “100”
 - Which position it is in that category, using a fixed length code, length=category number
 - “-2” is the number 1 (starting from 0) in category 2, with a fixed length code of “01”
 - Mã tổng thể 10001

Bảng mã DC

Range	DC Difference Category	AC Category
0	0	N/A
-1, 1	1	1
-3, -2, 2, 3	2	2
-7, ..., -4, 4, ..., 7	3	3
-15, ..., -8, 8, ..., 15	4	4
-31, ..., -16, 16, ..., 31	5	5
-63, ..., -32, 32, ..., 63	6	6
-127, ..., -64, 64, ..., 127	7	7
-255, ..., -128, 128, ..., 255	8	8
-511, ..., -256, 256, ..., 511	9	9
-1023, ..., -512, 512, ..., 1023	A	A
-2047, ..., -1024, 1024, ..., 2047	B	B
-4095, ..., -2048, 2048, ..., 4095	C	C
-8191, ..., -4096, 4096, ..., 8191	D	D
-16383, ..., -8192, 8192, ..., 16383	E	E
-32767, ..., -16384, 16384, ..., 32767	F	N/A

Category	Base Code	Length	Category	Base Code	Length
0	010	3	6	1110	10
1	011	4	7	11110	12
2	100	5	8	111110	14
3	00	5	9	1111110	16
4	101	7	A	11111110	18
5	110	8	B	111111110	20

Mã hóa AC



Mã hóa AC

- Example:
 - First symbol (0,5)
 - The value '5' is represented in two parts:
 - Which category it belongs to (Table of JPEG Coefficient Coding Categories), and code the "(runlength,category)" using a Huffman code (JPEG Default AC Code)
 - AC=5 is in category "3",
 - Symbol (0,3) has codeword "100"
 - Which position it is in that category, using a fixed length code, length=category number
 - "5" is the number 5 (starting from 0) in category 3, with a fixed length code of "101".
 - The overall codeword for (0,5) is "100101"
 - Second symbol (0,9)
 - '9' in category '4', (0,4) has codeword '1011', '9' is number 9 in category 4 with codeword '1001' -> overall codeword for (0,9) is '10111001'
 - ETC

Bảng mã AC

Range	DC Difference Category	AC Category
0	0	N/A
-1, 1	1	1
-3, -2, 2, 3	2	2
-7, ..., -4, 4, ..., 7	3	3
-15, ..., -8, 8, ..., 15	4	4
-31, ..., -16, 16, ..., 31	5	5
-63, ..., -32, 32, ..., 63	6	6
-127, ..., -64, 64, ..., 127	7	7
-255, ..., -128, 128, ..., 255	8	8
-511, ..., -256, 256, ..., 511	9	9
-1023, ..., -512, 512, ..., 1023	A	A
-2047, ..., -1024, 1024, ..., 2047	B	B
-4095, ..., -2048, 2048, ..., 4095	C	C
-8191, ..., -4096, 4096, ..., 8191	D	D
-16383, ..., -8192, 8192, ..., 16383	E	E
-32767, ..., -16384, 16384, ..., 32767	F	N/A

Mã hóa AC

Run/ Category	Base Code	Length	Run/ Category	Base Code	Length
0/0	1010 (= EOB)	4			
0/1	00	3	8/1	11111010	9
0/2	01	4	8/2	111111111000000	17
0/3	100	6	8/3	1111111110110111	19
0/4	1011	8	8/4	1111111110111000	20
0/5	11010	10	8/5	1111111110111001	21
0/6	111000	12	8/6	1111111110111010	22
0/7	1111000	14	8/7	1111111110111011	23
0/8	111110110	18	8/8	1111111110111100	24
0/9	1111111110000010	25	8/9	1111111110111101	25
0/A	1111111110000011	26	8/A	1111111110111110	26
1/1	1100	5	9/1	111111000	10
1/2	111001	8	9/2	1111111110111111	18
1/3	1111001	10	9/3	1111111111000000	19
1/4	111110110	13	9/4	1111111111000001	20
1/5	11111110110	16	9/5	1111111111000010	21
1/6	1111111110000100	22	9/6	1111111111000011	22
1/7	1111111110000101	23	9/7	1111111111000100	23
1/8	1111111110000110	24	9/8	1111111111000101	24
1/9	1111111110000111	25	9/9	1111111111000110	25
1/A	1111111110001000	26	9/A	1111111111000111	26
2/1	11011	6	A/1	111111001	10
2/2	11111000	10	A/2	1111111111001000	18
2/3	1111110111	13	A/3	1111111111001001	19
2/4	1111111110001001	20	A/4	1111111111001010	20
2/5	1111111110001010	21	A/5	1111111111001011	21
2/6	1111111110001011	22	A/6	1111111111001100	22
2/7	1111111110001100	23	A/7	1111111111001101	23

Ví dụ: DC coefficient of the previous block is 60

Cho ma trận ảnh lượng tử:

$$QDCTindex = \begin{bmatrix} 85 & -5 & -2 & 2 \\ 1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

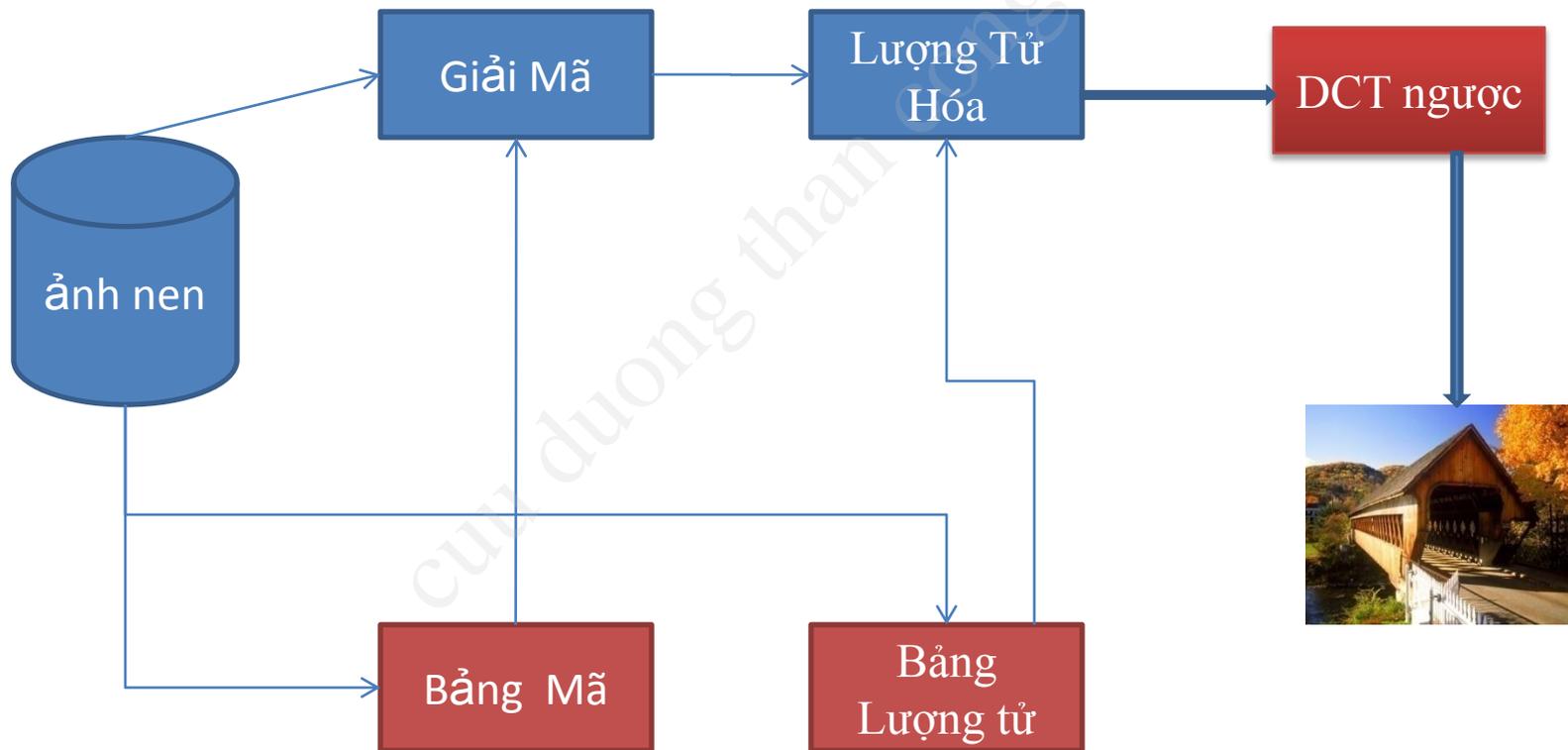
Biết rằng hệ số DC của khối trước là 60. Hãy mã hóa ma trận ảnh thành mã nhị phân. Tính tỉ số nén

Kết quả

- “11011001, 100010, 001, 1111100001, 0110, 0110, 000,1010”.

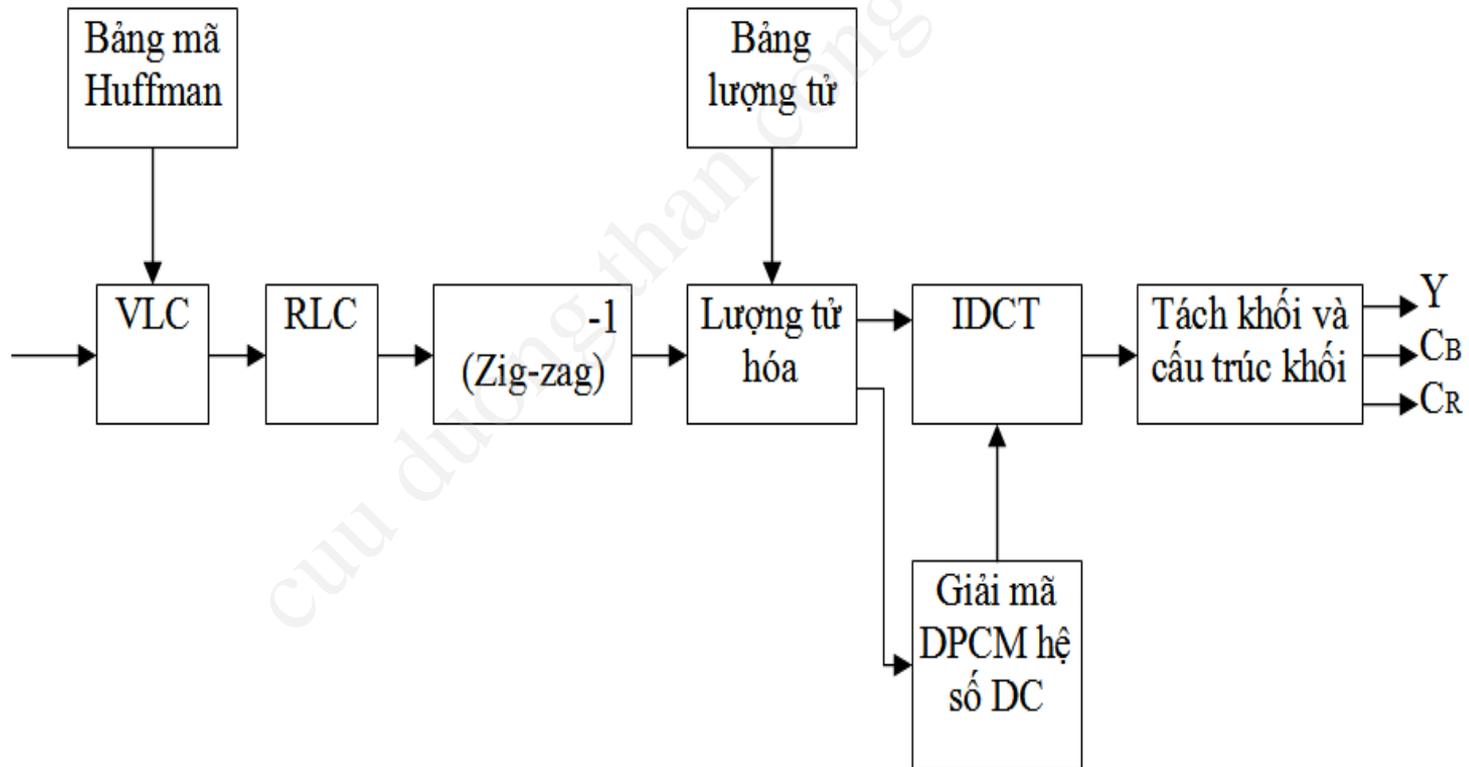
Khôi phục ảnh JPEG

Các công đoạn khôi phục ảnh:



Khôi phục ảnh JPEG

$$R(u,v) = Fq(u,v)Q(u,v)$$



i 8x8

0	u → 7							
v ↓	40	10	-2	2	-1	0	0	0
7	3	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	0
	0	0	0	0	0	0	0	0

a) Các hệ số DCT lượng tử hóa và chuẩn hóa $F_q(u,v)$

Biến đổi IDCT

0	u → 7							
v ↓	16	11	10	16	24	40	51	61
7	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

b) Bảng trọng số $Q(u,v)$

0	i → 7							
k ↓	98	97	90	81	80	79	67	50
7	97	96	89	80	79	78	66	49
	95	94	87	78	77	77	64	47
	93	92	85	75	75	74	62	45
	90	89	82	74	72	72	59	43
	88	87	80	71	70	69	57	40
	86	85	78	70	68	68	55	39
	85	79	77	69	67	67	54	38

d) Khối 8x8 các điểm chói $f^*(j,k)$

0	u → 7							
v ↓	591	110	-20	32	-24	0	0	0
7	36	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	0
	0	0	0	0	0	0	0	0

c) Khối các hệ số DCT tương ứng sau khi giải lượng tử

c (IDCT)

c:

$$f^*(j,k) = \sum_{u=0}^7 \sum_{v=0}^7 \frac{C_u C_v}{4} F(u,v) \cos \frac{(2j+1)u\pi}{16} \cos \frac{(2k+1)v\pi}{16}$$

c tính như sau:

$$e(j,k) = f(j,k) - f^*(j,k)$$

Biến đổi cosin và chuẩn JPEG

Ảnh Gốc



Ảnh nén và giải nén



Biến đổi cosin và chuẩn JPEG

Ảnh sau khi nén và giải nén bằng phép biến đổi Cosin cho chất lượng không tốt như cũ.

Khắc phục bằng việc làm trơn ảnh sau khi giải nén .

Chuẩn JPEG

Chuẩn JPEG cho nhiều tùy chọn khác nhau để nén ảnh, tùy vào những tùy chọn mà cho hệ số nén và chất lượng ảnh khác nhau



31KB



46KB



60KB



100KB

Ảnh gốc có kích thước $400 \times 300 = 120\text{KB}$

Bài tập

Cho ma trận ảnh S là 2×2

$$S = \begin{pmatrix} 9 & 1 \\ 1 & 9 \end{pmatrix}$$

Tính các hệ số DCT, và viết chuỗi nhị phân sau khi sử dụng phương pháp nén JPEG, tìm ảnh khôi phục sau khi giải nén, biết rằng ma trận lượng tử

$$Q = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}$$

Lời giải

- Áp dụng công thức DCT để tính hệ số DC và AC

$$2^{-DCT} = X_{(k,l)} = \frac{2}{N} c(k)c(l) \sum_{m=0}^7 \sum_{n=0}^7 x_{(m,n)} \cos\left(\frac{(2m+1)k\pi}{2N}\right) \cos\left(\frac{(2n+1)l\pi}{2N}\right) \quad \text{Trong đó } k,l=0,1,\dots,7$$

$$c(k), c(l) = \begin{cases} \frac{1}{\sqrt{2}} & , k \& l = 0 \\ 1 & , k^2 + l^2 \neq 0 \end{cases}$$

Hệ số DCT



$$T_{0,0} = (S, U_{0,0}) = (9 + 1 + 1 + 9) / 2 = 10, T_{0,1} = (S, U_{0,1}) = (9 - 1 + 1 - 9) / 2 = 0,$$

$$T_{1,0} = (S, U_{1,0}) = (9 + 1 - 1 - 9) / 2 = 0, T_{1,1} = (S, U_{1,1}) = (9 - 1 - 1 + 9) / 2 = 8.$$

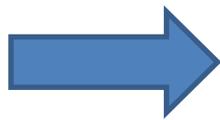
Ảnh khôi phục sau khi giải nén



$$\hat{S} = \hat{T}_{0,0} U_{0,0} + \hat{T}_{0,1} U_{0,1} + \hat{T}_{1,0} U_{1,0} + \hat{T}_{1,1} U_{1,1} = \frac{9}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{10}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 9.5 & -0.5 \\ -0.5 & 9.5 \end{bmatrix}$$

TỔNG QUAN JPEG 2000

- Ảnh số hóa ngày càng phổ biến và yêu cầu chất lượng ngày càng cao, ngoài ra còn có các nhu cầu về xử lý ảnh kéo theo
- Nén hình ảnh không chỉ làm giảm dung lượng mà còn phải cho phép tách ghép để sắp xếp xử lý và đáp ứng các nhu cầu trên các thiết bị cụ thể
- Yêu cầu về hiệu suất nén với tỉ số nén cao



SỰ PHÁT TRIỂN CỦA JPEG 2000 LÀ TẤT YẾU

TỔNG QUAN JPEG 2000



JPEG2000

JPEG-2000 có tỉ lệ xuyên âm thấp hơn hẳn công nghệ JPEG truyền thống



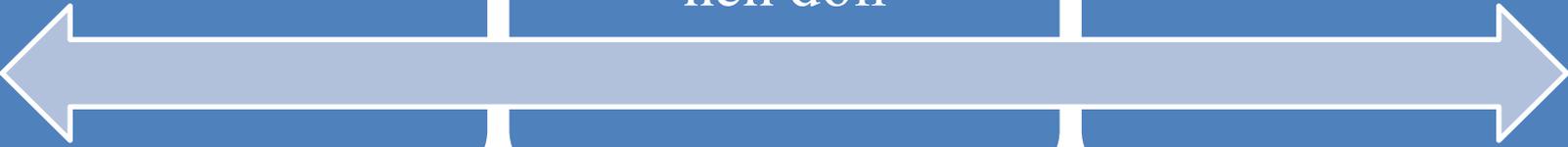
JPEG2000

Cho phép tách các phân giải khác nhau, các miền quan tâm và đưa ra được một dòng bit nén đơn



JPEG2000

JPEG-2000 sử dụng kỹ thuật mã hóa dạng sóng rời rạc DWT-Discrete Wavelet Transform



JPEG2000

ƯU ĐIỂM CỦA JPEG2000 SO VỚI JPEG

JPEG2000

Sử dụng được với truyền dẫn và hiển thị lũy tiến về chất lượng, độ phân giải, các thành phần màu có tính định vị không

JPEG2000

Sử dụng một cơ chế nén ảnh cho cả 2 dạng thớ nén (có tổn thất và không tổn thất)

JPEG2000

JPEG2000 có thể đưa ra tỷ lệ nén cao hơn nhiều so với JPEG. Tỷ lệ nén có thể lên tới 200:1

JPEG2000

Có khả năng mã hóa ảnh với tỉ lệ nén theo từng vùng khác nhau (ROI)

Chuẩn JPEG2000 và nén ảnh màu

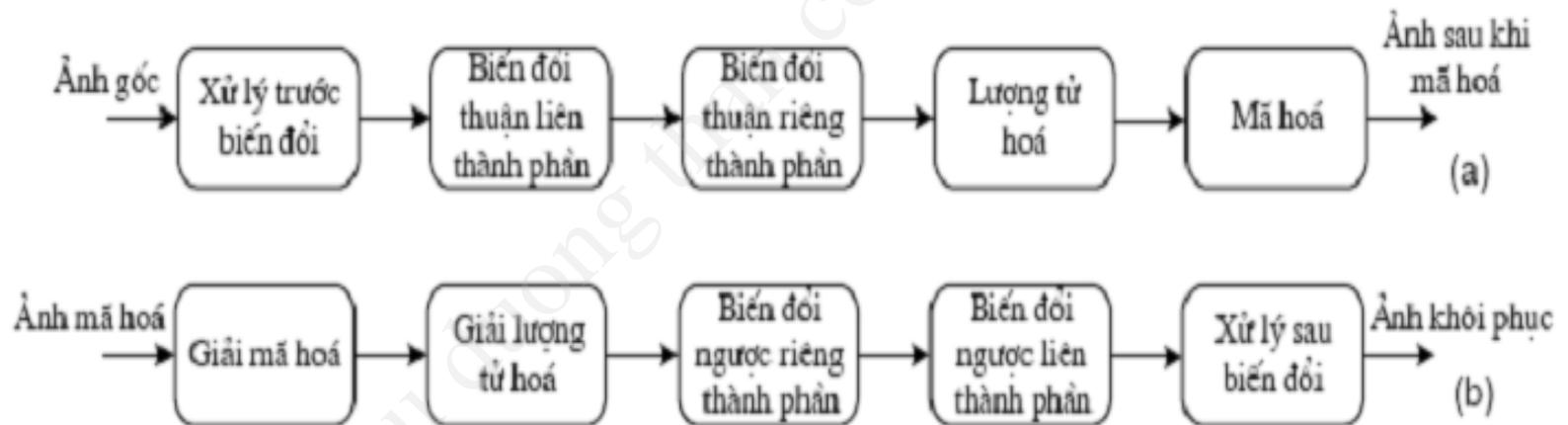
- *JPEG2K sử dụng phép biến đổi wavelet và các phương pháp đặc biệt để có được ảnh nén tối ưu nhất*
- Cho chất lượng tốt nhất khi sử dụng nén ảnh tĩnh
- Sử dụng được với truyền dẫn và hiển thị lũy tiến về chất lượng, độ phân giải.
- Truy cập và giải nén tại mọi thời điểm trong khi nhận dữ liệu
- Giải nén từng vùng ảnh mà không giải nén toàn bộ ảnh

CHUẨN JPEG2K

- Có khả năng mã hoá ảnh với tỷ lệ nén theo từng vùng khác nhau
- Nén một lần nhưng có thể giải nén với nhiều cấp chất lượng tùy theo yêu cầu của người sử dụng

Hiện tại, *ISO* và uỷ ban *JPEG* đã đưa ra khuyến nghị thay thế *JPEG* bằng *JPEG2000*.

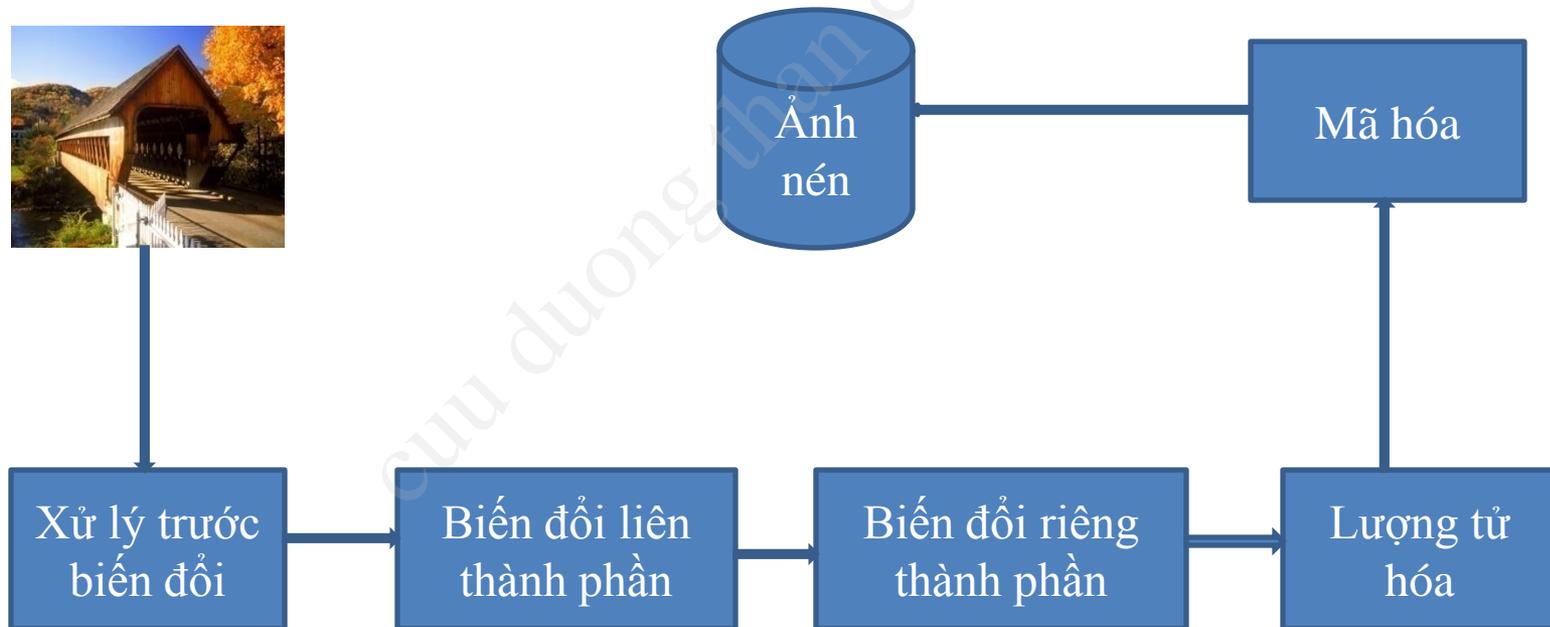
Các bước nén ảnh trong JPEG 2000



Sơ đồ nén và giải nén trong jpeg 2000

CHUẨN JPEG2000 VÀ NÉN ẢNH MÀU

- Các bước thực hiện nén ảnh theo chuẩn JPEG2K



Các bước nén ảnh trong JPEG 2000

BUỐC 1: Xử lý trước biến đổi

Xử lý trước biến đổi chính là giai đoạn đảm bảo dữ liệu đưa vào nén ảnh có dạng đối xứng qua 0

Các bước nén ảnh trong JPEG 2000

Bước 2: biến đổi liên thành phần

Loại bỏ tính tương quan giữa các thành phần của ảnh

Chuyển từ RGB sang YCrBr sử dụng ICT và RCT

Làm tăng hiệu quả nén

Các bước nén ảnh trong JPEG 2000

BƯỚC 3: Biến đổi riêng thành phần (biến đổi Wavelet)

Sử dụng biến đổi wavelet.
Chia tín hiệu thành các băng con và mỗi băng sẽ được mã hóa riêng rẽ

Thường sử dụng hàm wavelet Daubechies

Các bước nén ảnh trong JPEG 2000

BUỚC 4 : Lượng tử hóa

Cho phép đạt tỉ lệ nén cao hơn bằng cách thể hiện các giá trị biến đổi với độ chính xác tương ứng cần thiết với mức chi tiết của ảnh cần nén.

Các hệ số biến đổi sẽ được lượng tử hoá theo phép lượng tử hoá vô hướng

Các bước nén ảnh trong JPEG 2000

Bước 5 : Mã hóa

JPEG-2000 cho phép sử dụng nhiều phương pháp mã hóa khác nhau

Trong thực tế sử dụng 2 phương pháp phổ biến là SPIHT và EZW

Các bước nén ảnh trong JPEG 2000

SPIHT

Là viết tắt của Set
**Partitioning in
Hierarchical Trees**

-phương pháp mã
hoá phân cấp theo
phân vùng

SPIHT được thiết kế
tối ưu cho truyền
dẫn luỹ tiến

Sử dụng kỹ thuật embedded
coding

- Có nghĩa là: một ảnh sau nén
với kích cỡ (lưu trữ) lớn (tỷ lệ
nén thấp) sẽ chứa chính dữ liệu
sau nén của ảnh có kích cỡ (lưu
trữ) nhỏ (tỷ lệ nén cao)

Các bước nén ảnh trong JPEG 2000

EZW

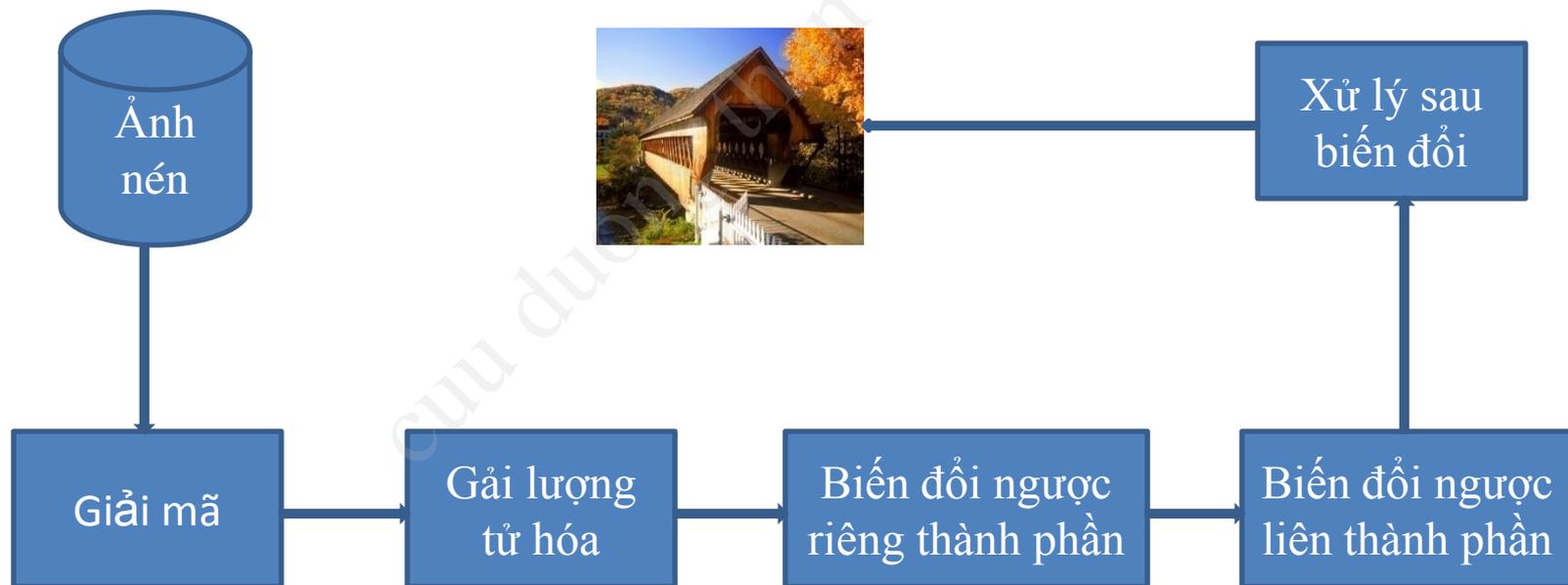
Viết tắt của
**Embedded
Zerotree Wavelet
Encoder**

Cũng dựa trên
phương pháp mã
hóa lũy tiến

Phương pháp này
chủ yếu dựa trên
khái niệm về cây
zero (zerotree).

CHUẨN JPEG2000 VÀ NÉN ẢNH MÀU

- Giải nén: Làm ngược lại các bước trên



Các chuẩn nén ảnh hiện nay

Các định dạng ảnh và tỉ lệ nén:

1600x1200	5760KB
BMP	5626KB
TGA	5626KB
PGM	5626KB
PBM	5626KB
TIF	5565KB
PNP	4405KB
XPM	3759KB
GIF	1802KB
JPEG	754KB

Ưu điểm của chuẩn nén JPEG2k

Một tính năng ưu việt nữa của *JPEG2000* là tính năng mã hoá ảnh quan trọng theo vùng (*ROI - Region of Interest*) mà chúng ta đã đề cập trong phần mã hoá ảnh theo *JPEG2000*.

Tính năng ưu việt thứ 3 của *JPEG2000* so với *JPEG* là chuẩn nén ảnh này có thể hiển thị được các ảnh với độ phân giải và kích thước khác nhau từ cùng một ảnh nén. Với *JPEG* thì điều này là không thể thực hiện được. Sở dĩ có điều này là do *JPEG2000* sử dụng kỹ thuật phân giải ảnh và mã hoá đính kèm

JPEG và JPEG2k

Bit per pixel	0.125	0.50	2.00
Ảnh 1 theo <i>JPEG</i>	24.42	31.17	35.15
Ảnh 1 theo <i>JPEG2000</i>	28.12	32.95	37.35
Ảnh 2 theo <i>JPEG</i>	22.6	28.92	35.99
Ảnh 2 theo <i>JPEG2000</i>	24.85	31.13	38.80

JPEG và JPEG2k



JPEG



JPEG2000



JPEG



JPEG2000

Bài tập

RGB <-> YCbCr

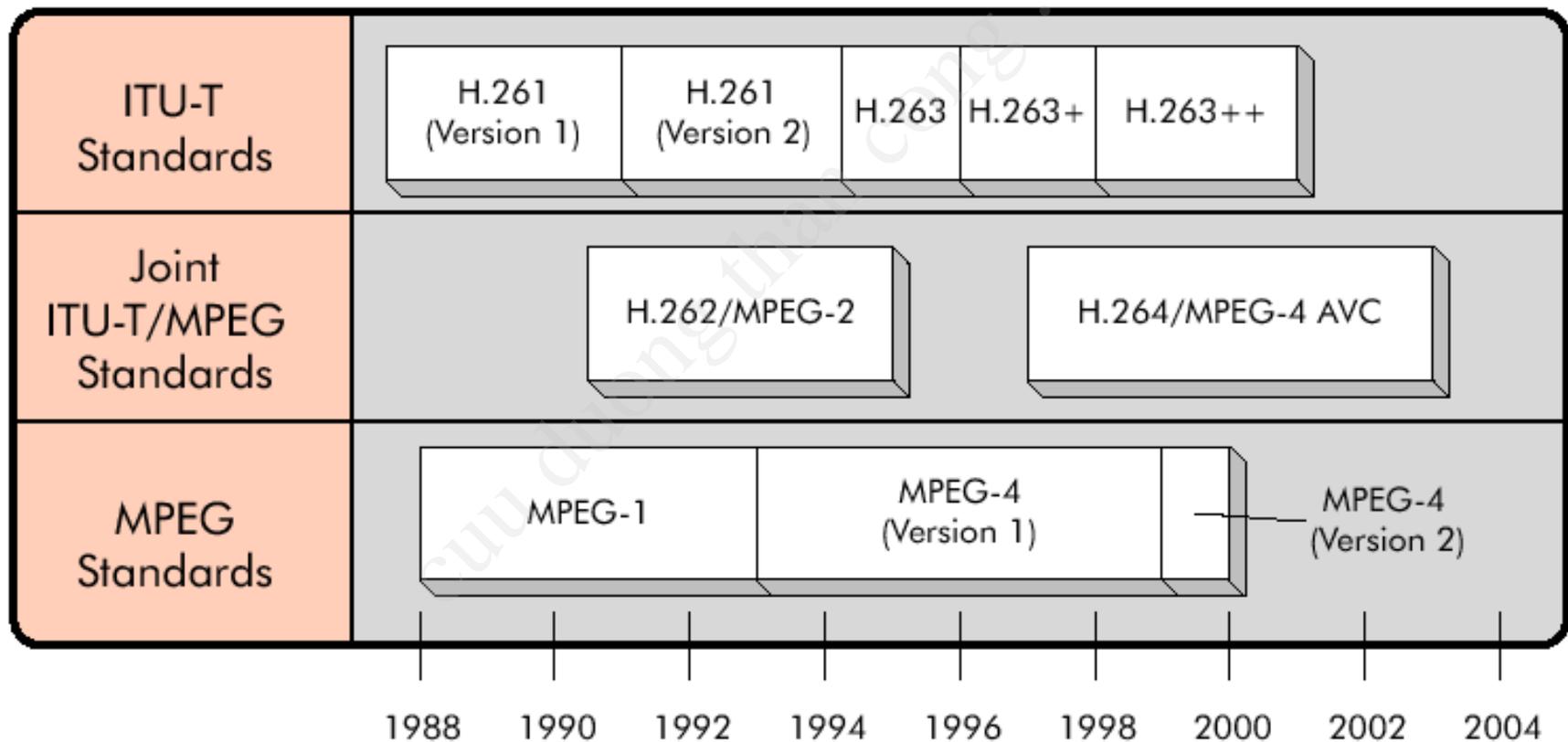
$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & -0.001 & 1.402 \\ 1.000 & -0.344 & -0.714 \\ 1.000 & 1.772 & 0.001 \end{bmatrix} \begin{bmatrix} Y \\ C_b - 128 \\ C_r - 128 \end{bmatrix}$$

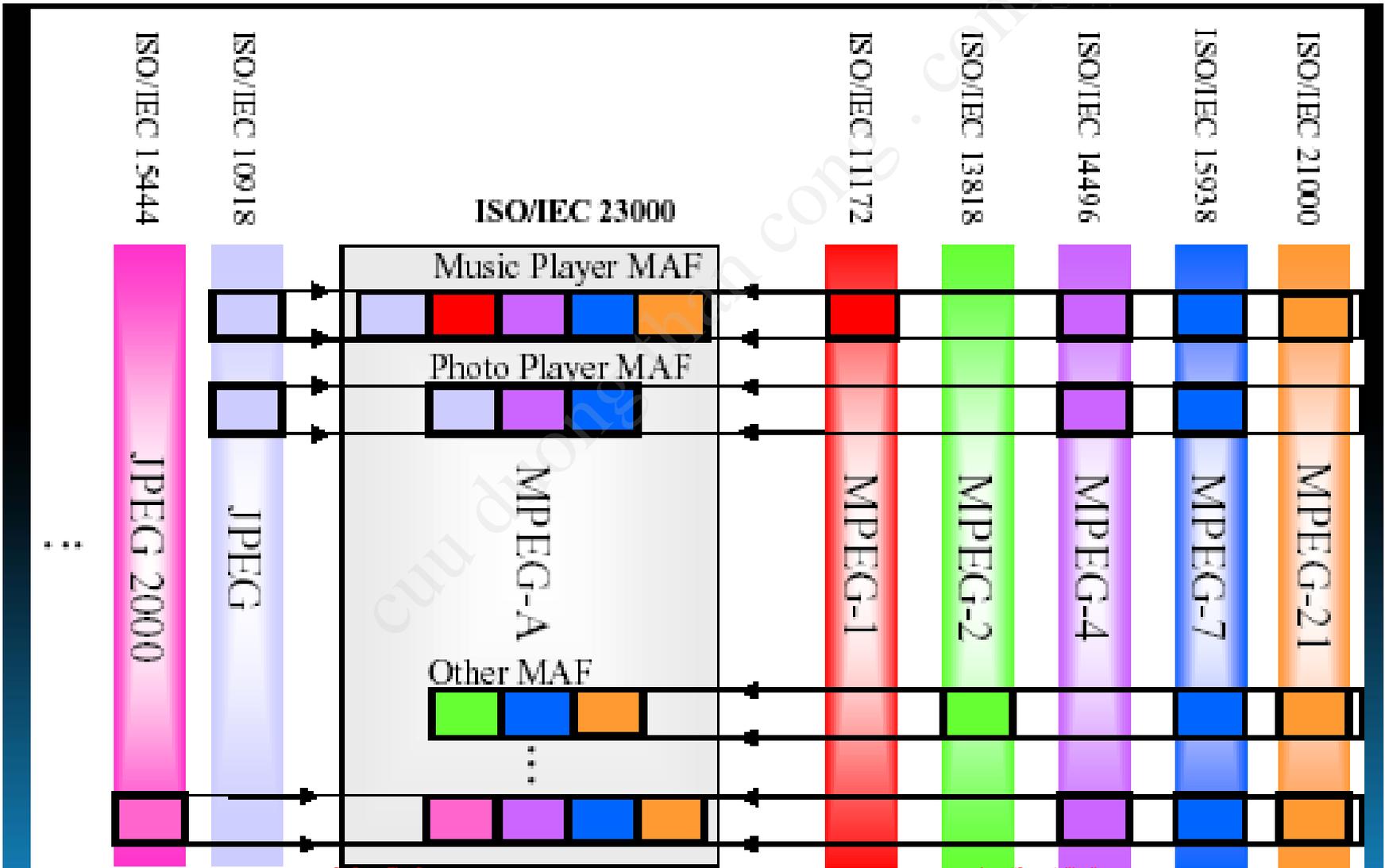
Bài tập

- What is the (Y,Cb,Cr) value for (R,G,B)=(255,128,64)?
- What is the (R,G,B) value for (Y,Cb,Cr)=(128,25,43)?

Các chuẩn nén video



MPEG



MPEG

- MPEG-1 ~ *1-1.5Mbps (early 90s)*
 - For compression of 320x240 full-motion video at rates around 1.15Mb/s
 - Applications: video storage (VCD)
- MPEG-2 ~ *2-80Mbps (mid 90s)*
 - For higher resolutions
 - Support interlaced video formats and a number of features for HDTV
 - Address scalable video coding
 - Also used in DVD
- MPEG-4 ~ *9-40kbps (later 90s)*
 - For very low bit rate video and audio coding
 - Applications: interactive multimedia and video telephony
- MPEG-21 ~ *ongoing*

MPEG

-

nh.

-

i DCT

-

ng.

-

ng $Y C_r C_b$.

ng

cuu duong than cong . com

macroblock (MB)

nh.

MPEG-1 Picture Types and Group-of-Pictures

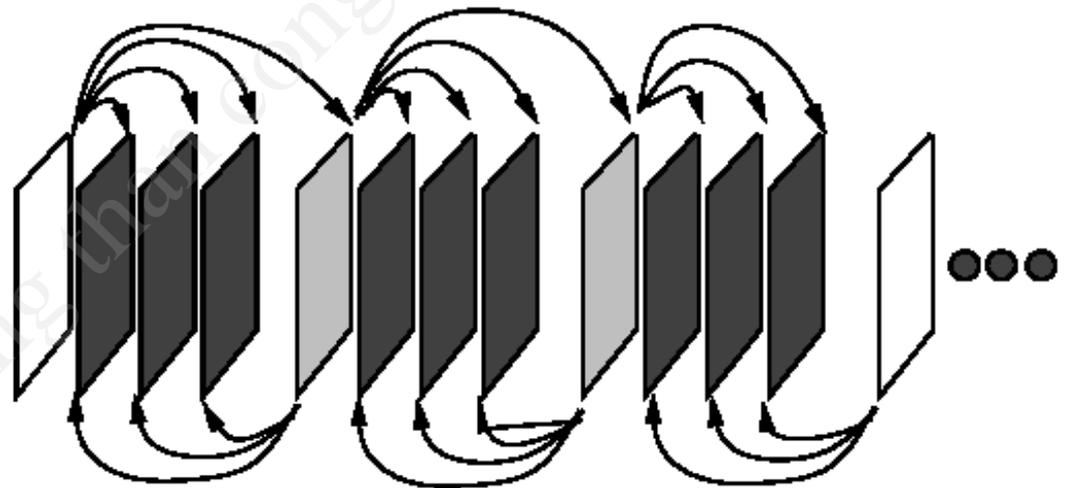
- A Group-of-Picture (GOP) contains 3 types of frames (I/P/B)

- Frame order

$I_1 \text{ BBB } P_1 \text{ BBB } P_2 \text{ BBB } I_2 \dots$

- Coding order

$I_1 P_1 \text{ BBB } P_2 \text{ BBB } I_2 \text{ BBB } \dots$



Intra-frame



Forward predicted frame

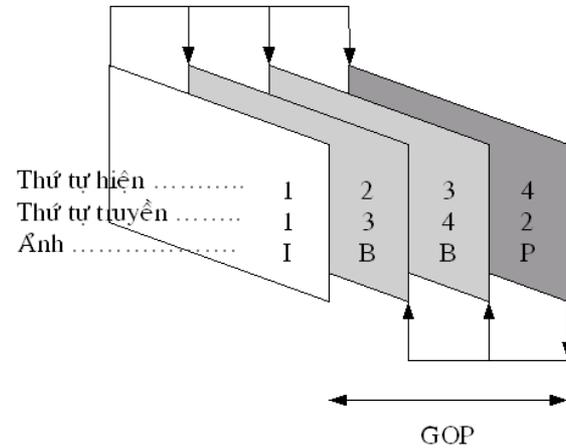
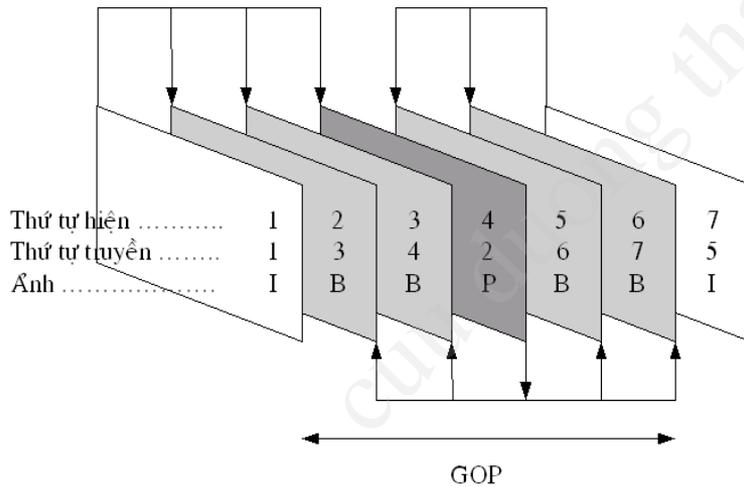


Bidirectionally predicted frame

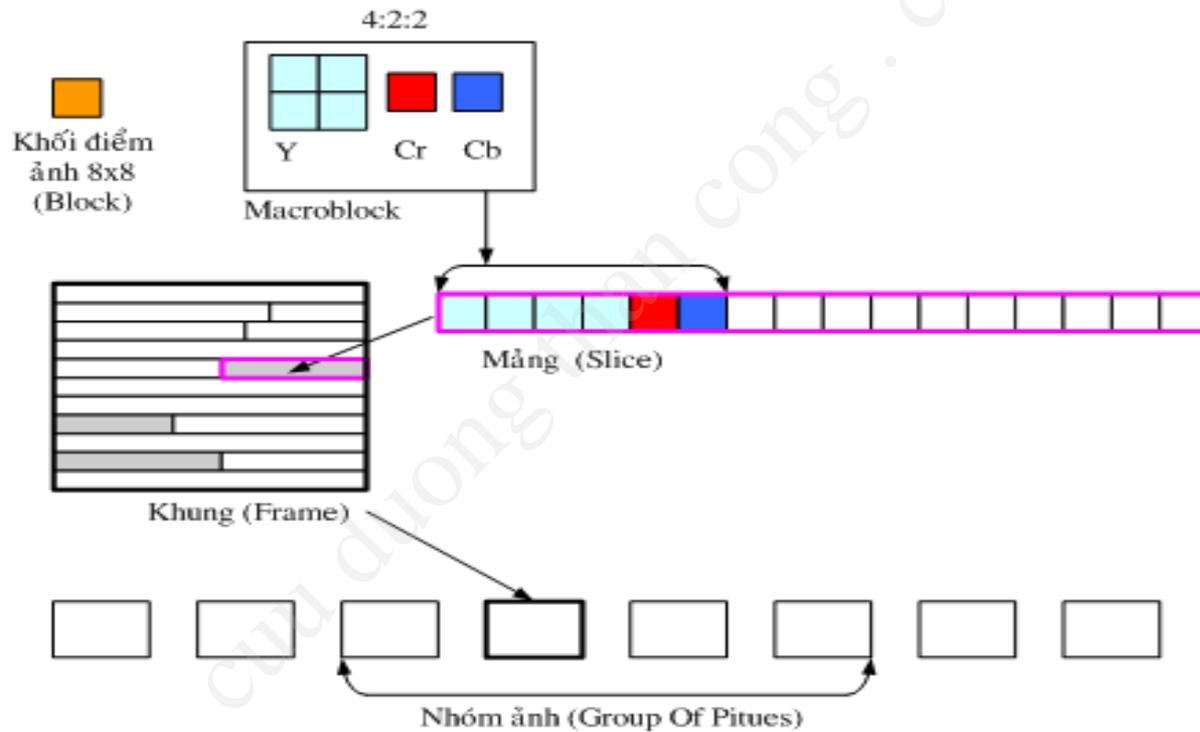
(Priority: Intra > Forward > Bidirectional)

Motion Compensation

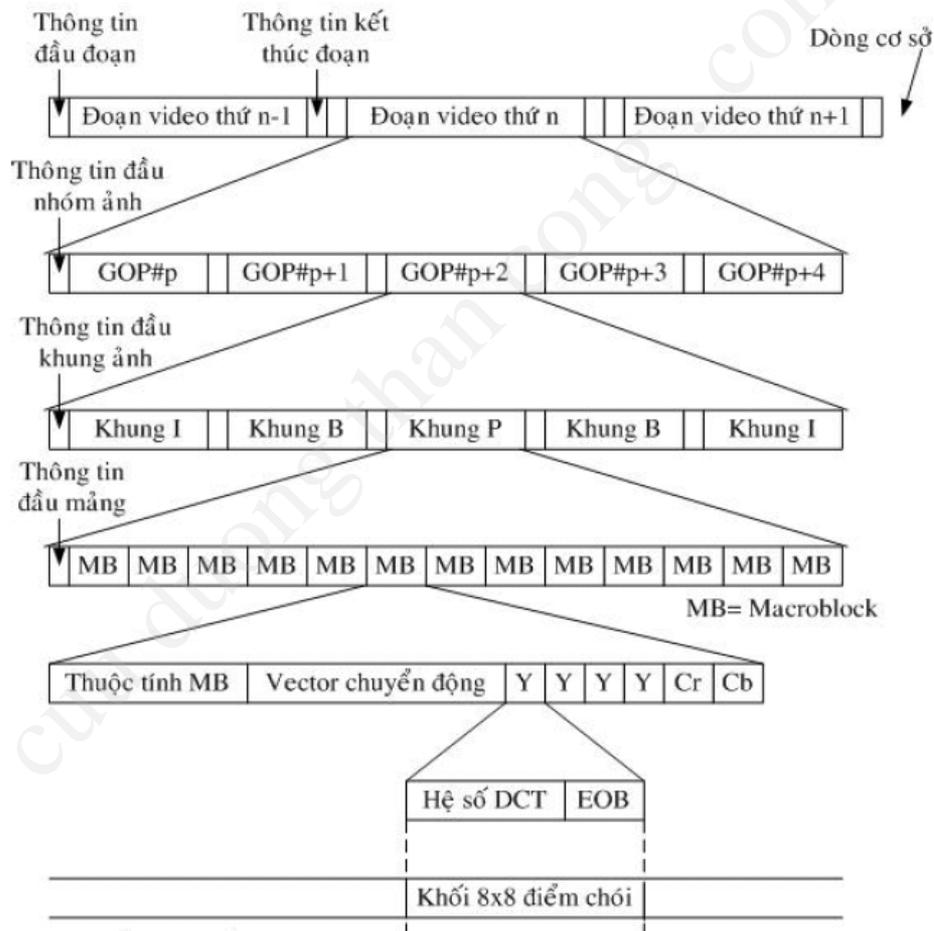
- i I (Intra-picture)
- i P (Predicted-picture)
- i B (Bi-directional predicted picture)



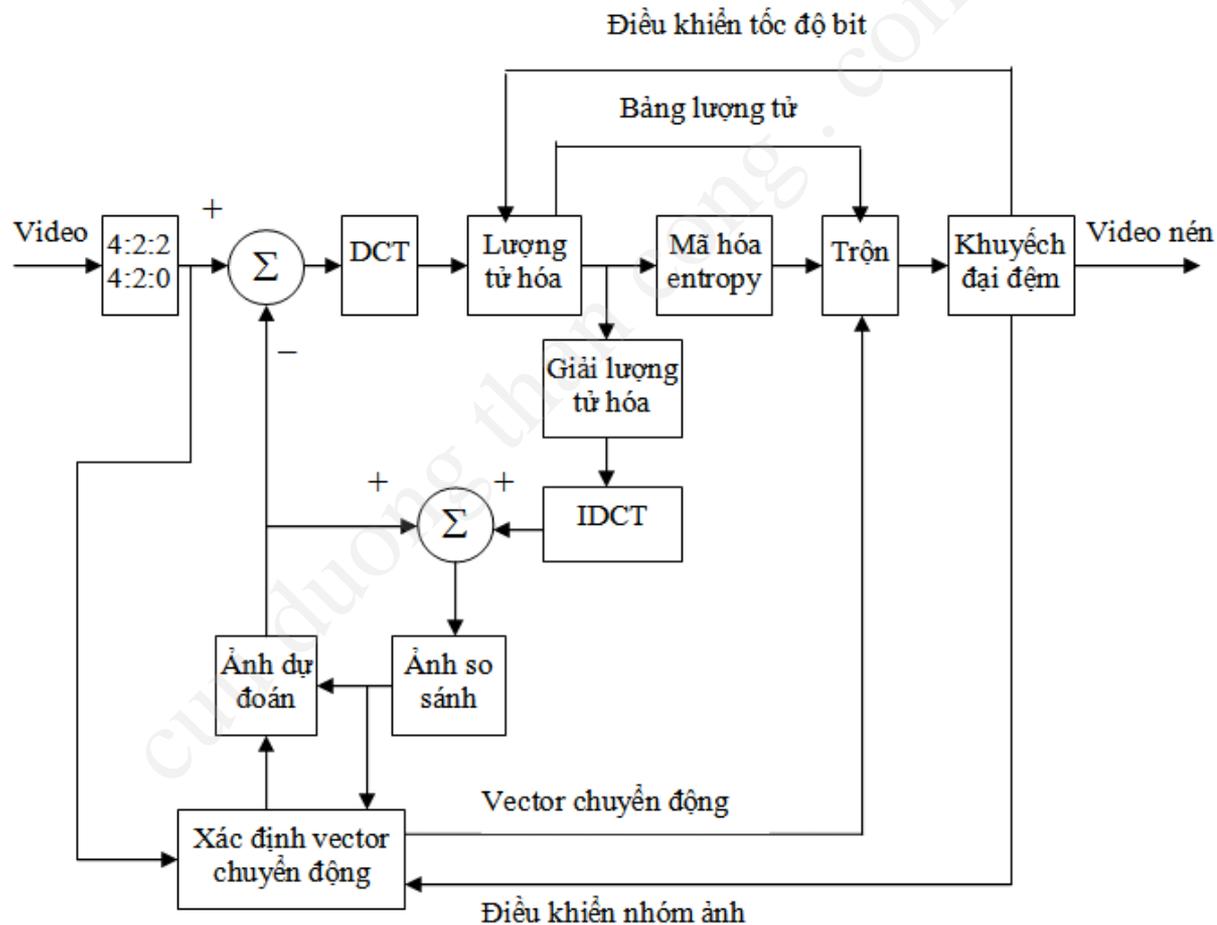
ng bit MPEG



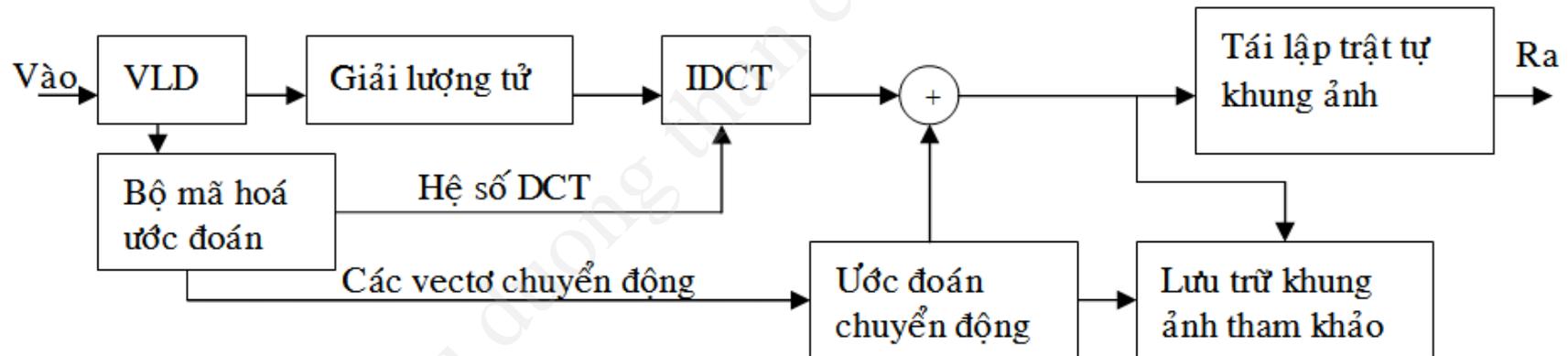
Kiến trúc dòng dữ liệu MPEG



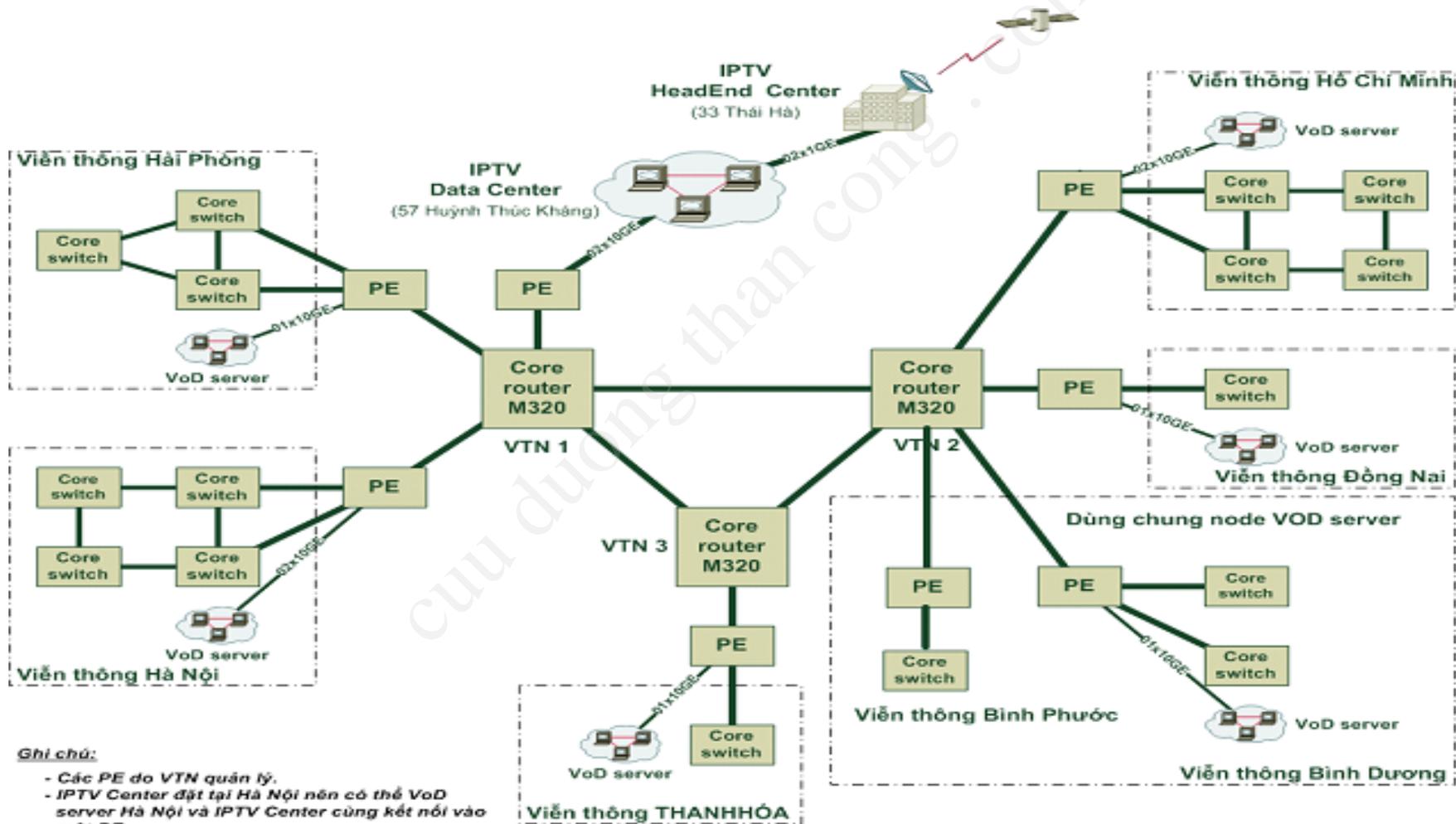
Bộ mã hóa MPEG tiêu biểu



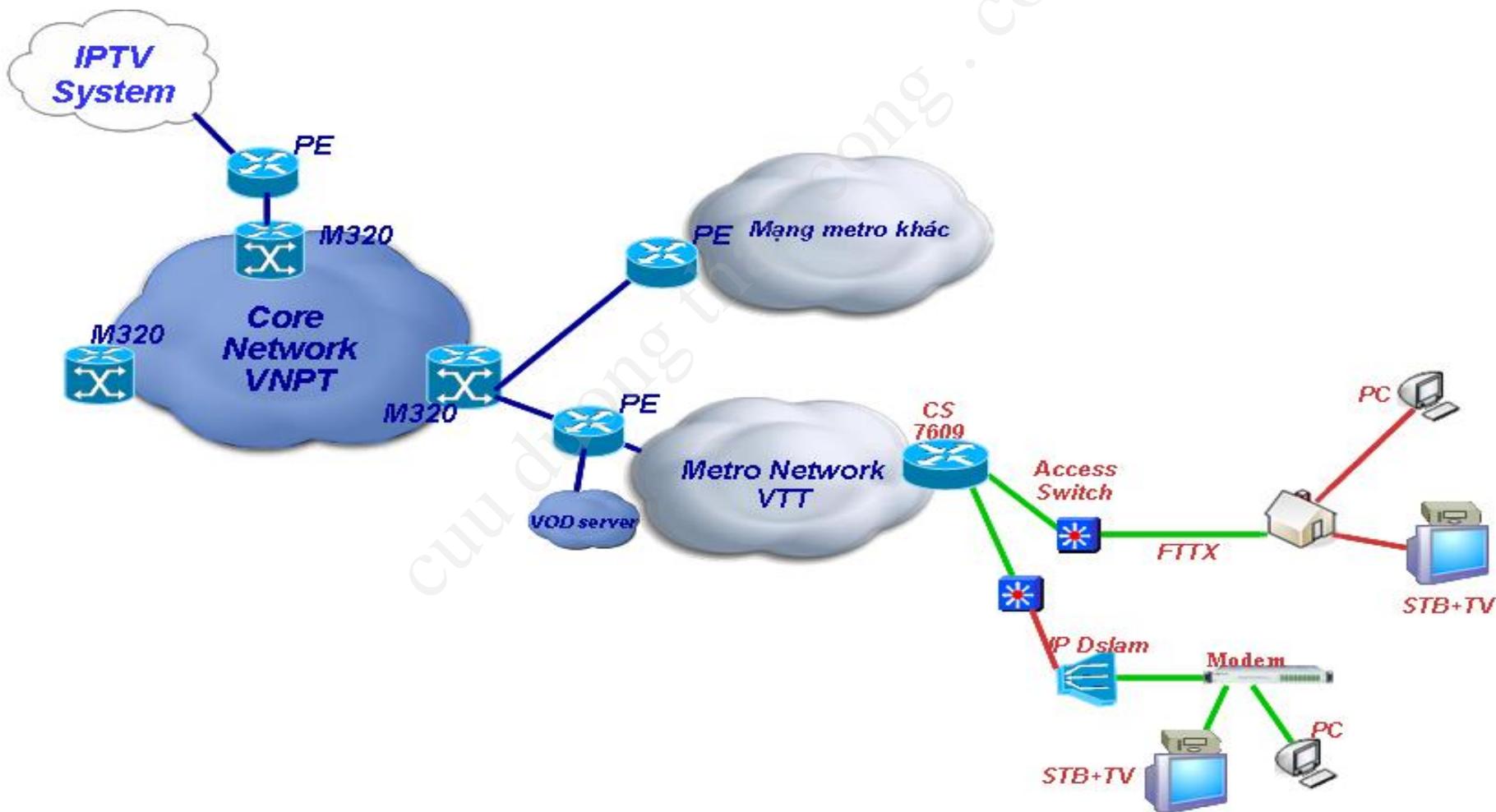
Bộ mã hóa MPEG tiêu biểu



Ứng dụng MPEG trong truyền thông đa phương tiện



Ứng dụng MPEG trong truyền thông đa phương tiện



Ứng dụng MPEG trong truyền thông đa phương tiện

