

# Nén dữ liệu đa phương tiện

Nguyễn Đình Hóa

[dinhhoa@gmail.com](mailto:dinhhoa@gmail.com) 0942807711

# Tóm tắt nội dung Bài 3

---

- ▶ Các cấu trúc lưu trữ đa chiều
  - ▶ Cây k-d (k-d trees)
  - ▶ Cây tứ phân (Point Quadrees)
  - ▶ Cây tứ phân MX (MX Quadrees)
  - ▶ Cây tứ phân vùng điểm PR (Point Region Quadrees)
  - ▶ Cây R (R-trees)

# Nén dữ liệu đa phương tiện

---

- ▶ Nén dữ liệu văn bản
- ▶ Nén dữ liệu âm thanh
- ▶ Nén dữ liệu hình ảnh
- ▶ Nén dữ liệu video

# Nén dữ liệu văn bản

---

## ▶ Mã hóa Huffman

- ▶ Dúng ít bits để mã hóa các ký tự thường xuyên xuất hiện,
- ▶ Dùng nhiều bit để mã hóa các ký tự ít xuất hiện,
- ▶ Cần có codebook cho mỗi văn bản được mã hóa.

<i>Symbol</i>	<i>Probability</i>	<i>Code</i>
<i>e</i>	0.8	1
<i>t</i>	0.16	01
<i>x</i>	0.02	001
<i>z</i>	0.02	000

# Nguyên tắc tạo cây Huffman

---

- ▶ Sắp xếp các ký tự theo thứ tự tăng dần về tần suất xuất hiện
- ▶ Gộp 2 ký tự với xác suất nhỏ nhất (các nút lá) vào thành một nút gốc với xác suất bằng tổng của chúng
- ▶ Sắp xếp lại danh sách gồm các ký tự còn lại và nút gốc mới tạo ra.
- ▶ Chu trình lặp lại đến khi hết ký tự. Nút gốc cuối cùng là gốc của cây Huffman.
- ▶ Đi ngược từ gốc tới các lá (các ký tự) và đánh ký hiệu các nhánh là '0' với nhánh trái và '1' với nhánh phải.
- ▶ Mã Huffman của từng ký tự sẽ là chuỗi mã tổng hợp từ gốc tới ký tự đó.

# Nén dữ liệu văn bản

---

## ► Mã hóa Run-length

- Thay thế một chuỗi ký tự trùng lặp bằng một ký tự cùng với số lần trùng lặp.

Sc	X	C
----	---	---

- Cần thiết lập ngưỡng số lần lặp.
- Phát sinh ký tự mới.
- VD: “MMMMannnnnnnnn” (13 ký tự) -> “@M4a@n8” (7 ký tự)

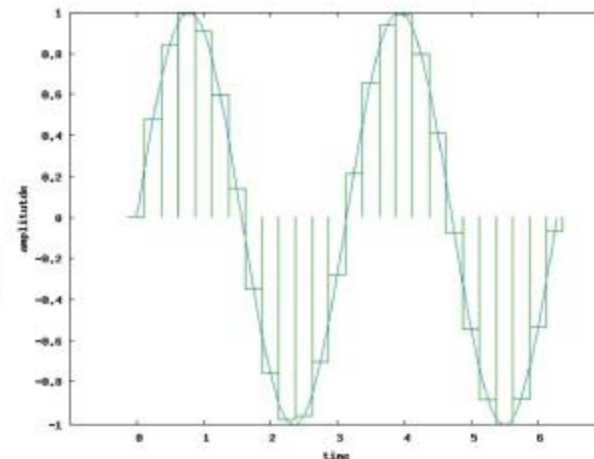
# Nén dữ liệu văn bản

---

- ▶ Mã hóa LZW (Lempel-Ziv-Welch): do Ziv và Lempel giới thiệu năm 1977 và 1978, Welch cải tiến năm 1984.
  - ▶ Sử dụng từ điển của các cụm ký tự. Lưu trữ mã của các cụm ký tự (mã của các từ) thay cho mã của từng ký tự.
  - ▶ VD: văn bản có 10,000 ký tự bao gồm 2000 từ, trong đó có 500 từ khác nhau: ta mã hóa 500 từ thay cho mã hóa 10,000 ký tự.

# Nén dữ liệu âm thanh

- ▶ Tín hiệu âm thanh là liên tục cả về thời gian lẫn biên độ.
- ▶ Chuyển đổi sóng âm thành tín hiệu điện
- ▶ Tín hiệu điện được lấy mẫu, lượng tử hóa, mã hóa.
  - ▶ Cần xác định tần số lấy mẫu, mức lượng tử,...



Sample values

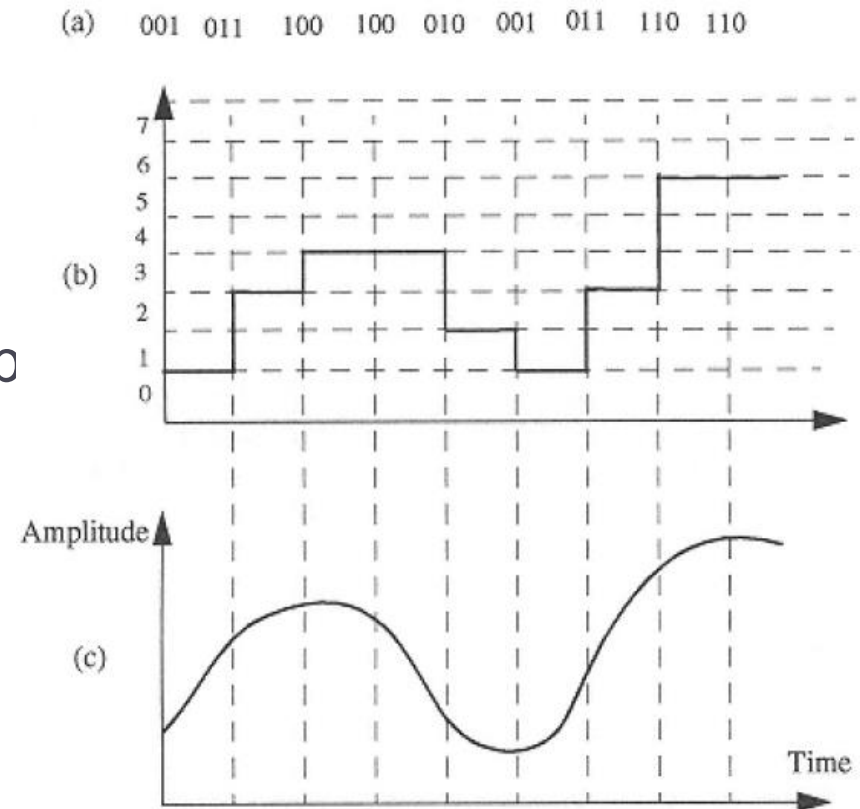
0.0000000000000000
0.479425538604203
0.841470984807897
0.997494986604054
0.909297426825682
0.598472144103956
0.141120008059867
-0.35078322768962
-0.756802495307928
-0.977530117665097
-0.958924274663138

...



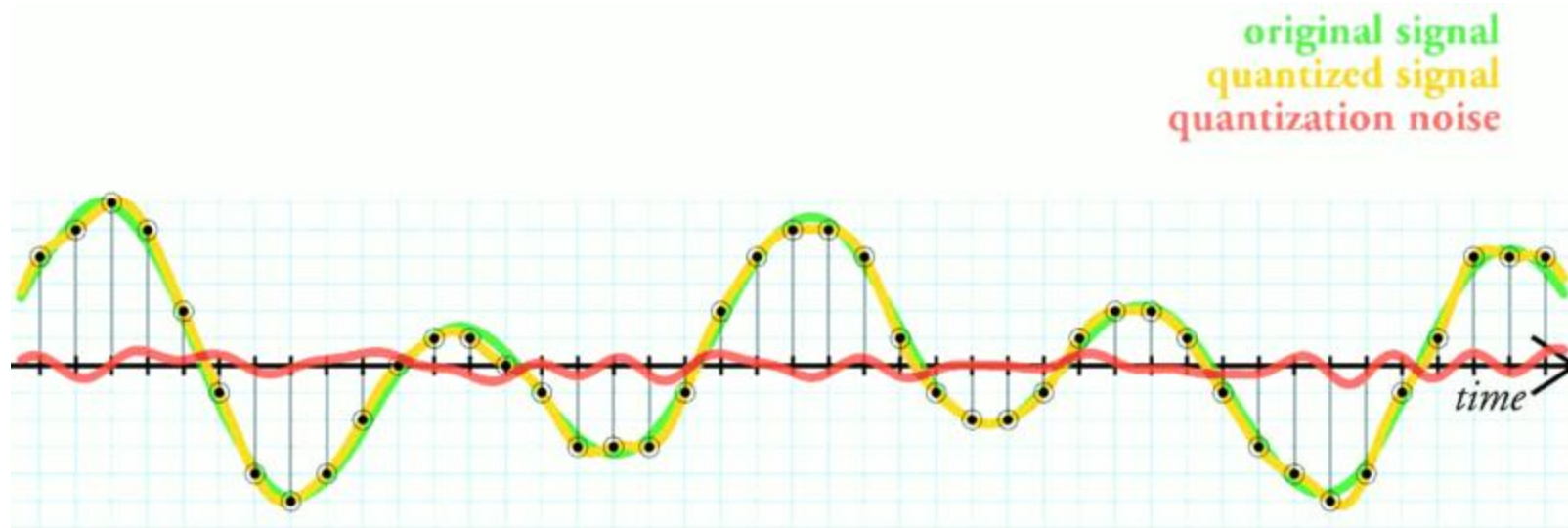
# Nén dữ liệu âm thanh

- ▶ Nhược điểm của lượng tử hóa tuyến tính:
  - ▶ Mất cân đối về tỉ lệ tín hiệu trên nhiễu (SNR) giữa vùng biên độ cao và vùng biên độ thấp
  - ▶ Tai người nhạy cảm hơn với âm thanh ở vùng biên độ thấp



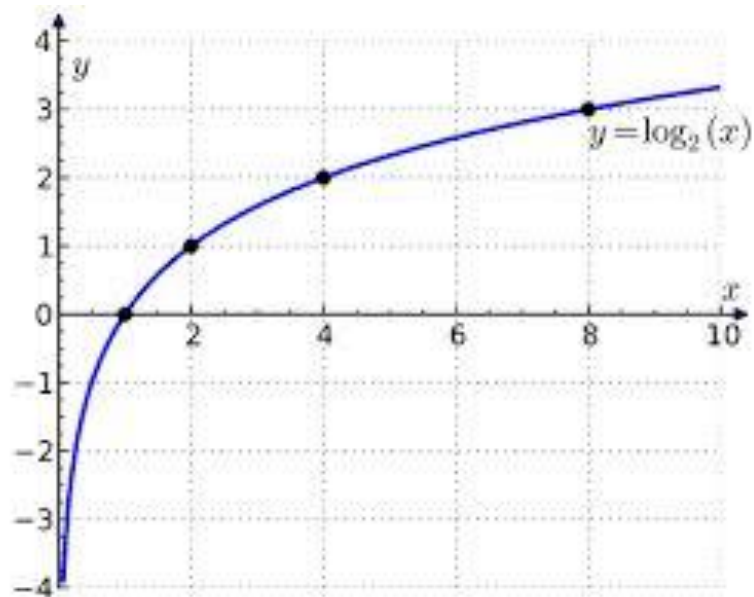
# Nén dữ liệu âm thanh

---



# Nén dữ liệu âm thanh

- ▶ Lượng tử hóa phi tuyến (nonlinear quantization):
  - ▶ Kích thước bước lượng tử biến đổi theo tỉ lệ logarit với biên độ âm thanh
  - ▶ Kích thước bước lượng tử nhỏ khi biên độ âm thanh nhỏ, và ngược lại.



# Nén dữ liệu âm thanh

---

- ▶ Mã hóa dự đoán (predictive coding)
  - ▶ Dự đoán giá trị biên độ của tín hiệu
  - ▶ Tính khoảng sai số giữa giá trị dự đoán và giá trị thực
  - ▶ Khoảng sai số này được lượng tử hóa và mã hóa
    - ▶ Differential Pulse-Coded Modulation (DPCM)
  - ▶ Mã có độ dài cố định được sử dụng trong DPCM

# Nén dữ liệu âm thanh

---

- ▶ DPCM có hiệu quả khi
  - ▶ Các giá trị biên độ liên kề của tín hiệu tương quan với nhau
  - ▶ Khoảng sai số nhỏ hơn khoảng giá trị thực của biên độ tín hiệu (cần dùng ít bit hơn để biểu diễn sai số)

# Nén dữ liệu âm thanh

---

## ▶ Adaptive DPCM

- ▶ Áp dụng đối với tín hiệu có biên độ biến đổi quá nhanh hoặc quá chậm.
- ▶ Nếu biên độ tín hiệu biến đổi nhanh: sử dụng bước lượng tử lớn
- ▶ Nếu biên độ tín hiệu biến đổi chậm: sử dụng bước lượng tử nhỏ.

# Nén dữ liệu âm thanh

- Khuyến nghị của hiệp hội viễn thông Quốc tế (ITU-TS) cho mã hóa tín hiệu thoại (băng thông 3,4kHz hoặc 7kHz)

<i>Recommendation</i>	<i>Compression technique</i>	<i>Speech bandwidth (kHz)</i>	<i>Sampling rate (kHz)</i>	<i>Compressed bit rate (kbps)</i>
G.711	PCM (no compression)	3.4	8	64
G.721	ADPCM	3.4	8	32
G.722	Subband ADPCM	7	16 (14 bits per sample)	48,56,64
G.723	ADPCM	3.4	8	24
G.728	Vector quantization	3.4	8	16

# Nén dữ liệu âm thanh

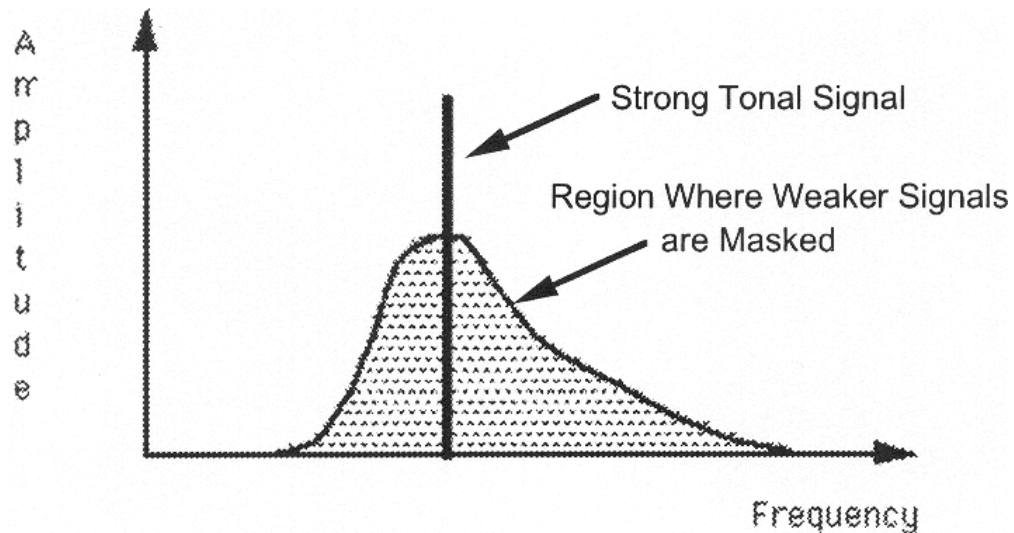
---

- ▶ MPEG-Audio: sử dụng thuộc tính che mặt (masking)
  - ▶ Mã hóa tín hiệu thoại trong toàn dải nghe thấy (20Hz - 20kHz)
  - ▶ Thuộc tính che mặt: một âm thanh có thể che khuất (làm ẩn đi) một tín hiệu âm thanh khác, hoặc nó có thể làm biến đổi cường độ thật của một âm thanh khác.
    - ▶ Nếu một âm thanh bị che khuất, nó có thể được loại bỏ mà không làm ảnh hưởng đến chất lượng tín hiệu
  - ▶ Tần số lấy mẫu: 32, 44.1, 48 kHz.



# Nén dữ liệu âm thanh

- ▶ Ngưỡng che khuất âm thanh
  - ▶ Tùy thuộc vào tần số âm thanh
  - ▶ Tùy thuộc vào năng lượng sóng âm trong một dải nhất định xung quanh tần số chọn sẵn.



# Nén dữ liệu âm thanh

---

## ▶ Các bước cơ bản

- ▶ Chuyển đổi tín hiệu sang miền tần số
- ▶ Phân chia tín hiệu sang 2 dạng: nghe được, không nghe được
- ▶ Xác định ngưỡng tối thiểu che khuất cho từng dải tần số
- ▶ Tính toán tỉ lệ tín hiệu được chọn với tín hiệu che khuất.

# Nén dữ liệu âm thanh

---

- ▶ MPEG-audio: cho phép mã hóa âm thanh ở 3 mức
  - ▶ MPEG-audio layer 1
  - ▶ MPEG-audio layer 2
  - ▶ MPEG-audio layer 3

# Nén dữ liệu âm thanh

---

- ▶ **MPEG-audio layer 1:**
  - ▶ Độ phức tạp ít nhất
  - ▶ Cho phép tốc độ lớn hơn 128kb/s trên mỗi kênh
  - ▶ Mỗi khung tín hiệu bao gồm: header, từ khóa kiểm tra CRC (optional), dữ liệu phụ (optional).
  - ▶ VD: Philips Digital Compact Cassette.

# Nén dữ liệu âm thanh

---

- ▶ MPEG-audio layer 2
  - ▶ Độ phức tạp vừa phải
  - ▶ Tốc độ 128kb/s trên mỗi kênh
  - ▶ Dùng trong phát thanh số (DAB)
  - ▶ Đồng bộ âm thanh và video trong CD-ROM
  - ▶ Các khung tín hiệu chứa 1125 mẫu trên mỗi kênh truyền

# Nén dữ liệu âm thanh

---

- ▶ MPEG-audio layer 3:
  - ▶ Phức tạp nhất
  - ▶ Chất lượng âm thanh tốt nhất
  - ▶ Tốc độ truyền: 64kb/s trên mỗi kênh
  - ▶ Thích hợp cho truyền âm thanh trên mạng ISDN

# Nén dữ liệu âm thanh

---

- ▶ Sự phát triển của MPEG-audio:
  - ▶ MPEG-1: Video CD and MP3.
  - ▶ MPEG-2: Digital Television set top boxes and DVD
  - ▶ MPEG-4: Fixed and mobile web
  - ▶ MPEG-7: description and search of audio and visual content
  - ▶ MPEG-21: Multimedia Framework

# Nén dữ liệu hình ảnh

- ▶ Một bức ảnh số được đặc trưng bởi nhiều thuộc tính
    - ▶ Kích thước ảnh:  $x * y$
    - ▶ Độ phân giải màu (số bit mã hóa màu):  $b$
- Dung lượng ảnh =  $x * y * b$ .
- ▶ Nén ảnh = giảm dung lượng ảnh



(a) JPEG low compression



(b) JPEG high compression



# Nén dữ liệu hình ảnh

---

- ▶ Những cơ sở được áp dụng khi nén hình ảnh
  - ▶ Mắt người khó phân biệt sự khác biệt giữa các điểm ảnh nằm sát nhau
  - ▶ Các điểm ảnh nằm trên các dòng liền kề nhau thường giống nhau (sự dư thừa về không gian)
  - ▶ Sự dư thừa không gian có thể được loại bỏ nhờ kỹ thuật mã hóa dự đoán, mã hóa chuyển đổi
  - ▶ Thông tin của ảnh có thể được lược bớt mà không làm mất đi giá trị truyền tin của ảnh
  - ▶ Các đặc tính của ảnh không nhạy cảm đối với mắt người có thể được loại bỏ

# Nén dữ liệu hình ảnh

---

- ▶ Lấy mẫu theo không gian (spatial subsampling)
  - ▶ Tại bộ mã hóa, các điểm ảnh được lựa chọn cách nhau để truyền đi
  - ▶ Tại bộ giải mã, toàn bộ các điểm ảnh được khôi phục bằng phương pháp nội suy để tạo nên ảnh có độ phân giải thấp hơn
  - ▶ Với các điểm ảnh có cả 2 thành phần chói và màu thì thành phần màu có thể lấy mẫu với tần suất cao hơn, và mức lượng tử thưa hơn bởi vì mắt người kém nhạy cảm với thành phần màu.
  - ▶ Tín hiệu ảnh lấy mẫu theo không gian có thể tiếp tục được nén bằng các kỹ thuật khác.

# Nén dữ liệu hình ảnh

---

- ▶ Mã hóa dự đoán (predictive coding)
  - ▶ Cơ bản giống như nén tín hiệu âm thanh
    - ▶ Giá trị màu của các phần tử ảnh liên kề nhau là tương đồng
    - ▶ Giá trị màu của một điểm có thể dùng để dự đoán màu cho các điểm lân cận
  - ▶ Dự đoán một chiều: sử dụng sự tương đồng giữa các điểm ảnh trên cùng một dòng quét
  - ▶ Dự đoán đa chiều: sử dụng thêm sự tương đồng giữa các điểm ảnh giữa các dòng khác nhau cũng như giữa các khung hình khác nhau.

# Nén dữ liệu hình ảnh

---

- ▶ Mã hóa chuyển đổi (transform coding)
  - ▶ Chuyển đổi sự tương quan giữa các điểm ảnh thành các hệ số độc lập,
  - ▶ Các hệ số này sẽ được truyền đi, giảm sự dư thừa thông tin trong ảnh.

# Nén dữ liệu hình ảnh

---

- ▶ Cách thức của mã hóa chuyển đổi
  - ▶ Ảnh được chia làm các vùng nhỏ (hình chữ nhật) (8x8)
  - ▶ Từng vùng ảnh sẽ được chuyển đổi
    - ▶ Karhunen-Loeve transform (KLT)
    - ▶ Discrete cosine transform (DCT)
    - ▶ Walsh-Hadamard transform (WHT)
    - ▶ Discrete Fourier transform (DFT)
  - ▶ Lựa chọn, lượng tử hóa các hệ số đã được chuyển đổi để truyền đi hoặc lưu trữ.
  - ▶ Mã hóa Huffman hoặc mã hóa run-length có thể được tiếp tục sử dụng.

# Nén dữ liệu hình ảnh

---

- ▶ Lượng tử hóa vector (vector quantization)
- ▶ Mã hóa Fractal
- ▶ Nén dựa trên Wavelet
- ▶ Các hệ thống trên thực tế: kết hợp các kỹ thuật mã hóa nói trên theo nhiều cách khác nhau
  - ▶ Lấy mẫu theo không gian và thời gian
  - ▶ Mã hóa DPCM
  - ▶ Chuyển đổi DCT trên không gian 2 chiều
  - ▶ Mã hóa Huffman
  - ▶ Mã hóa Run-length.

# Nén dữ liệu hình ảnh

---

- ▶ Tiêu chuẩn nén ảnh tĩnh JPEG (Joint Photographic Experts Group)
  - ▶ Là tiêu chuẩn phổ biến từ những năm 1990's
  - ▶ Bao gồm nhiều mức mã hóa
    - ▶ Mã hóa dựa trên DCT liên tiếp (có suy hao): là phương pháp cơ bản, được sử dụng nhiều nhất
    - ▶ Mã hóa dựa trên DCT mở rộng (có suy hao)
    - ▶ Mã hóa không suy hao
    - ▶ Mã hóa phân cấp: ảnh được mã hóa ở nhiều độ phân giải.

# Nén dữ liệu hình ảnh

---

- ▶ Mã hóa dựa trên DCT liên tiếp
  - ▶ Ảnh được chia thành các ô nhỏ (8x8 điểm ảnh)
  - ▶ Chuyển đổi DCT hai chiều cho từng ô, nhận được 8x8 hệ số DCT tại mỗi ô.
  - ▶ Các hệ số được lượng tử hóa bởi bảng lượng tử hóa được xác định trước.
  - ▶ Các hệ số đã lượng tử hóa được quét theo đường zizzag để được một mảng số một chiều
  - ▶ Hệ số đầu tiên được gọi là hệ số một chiều (DC), các hệ số còn lại là các hệ số hai chiều (AC).
  - ▶ Mã hóa Huffman / Run-length được áp dụng trên dãy số này để nén tin hiệu.



# Nén dữ liệu hình ảnh

---

- ▶ JBIG (Joint Bi-level Image Experts Group)
- ▶ JPEG-2000

# Nén dữ liệu video

---

- ▶ Video là chuỗi các hình ảnh
  - ▶ Thông tin dư thừa theo không gian
  - ▶ Thông tin dư thừa theo thời gian
- ▶ Nén video
  - ▶ Nén thông tin theo không gian (giống nén ảnh)
  - ▶ Nén thông tin theo thời gian.

# Nén dữ liệu video

---

- ▶ Ước lượng và bù chuyển động
  - ▶ Chuyển động: điểm ảnh của hình trước thay đổi vị trí tại hình sau.
  - ▶ Mỗi hình ảnh được chia nhỏ thành các khối ảnh
  - ▶ Hai khối ảnh giống nhau tại 2 ảnh khác nhau sẽ được xác định và biểu diễn bởi một vector chuyển động.
  - ▶ Các vector chuyển động cùng với các khối điểm ảnh khác nhau sẽ được mã hóa và truyền đi / lưu trữ.

# Nén dữ liệu video

---

- ▶ MPEG (Moving Picture Experts Group)
  - ▶ MPEG-1: mã hóa video chất lượng VHS (360x280 điểm ảnh, 30 hình/giây) với tốc độ 1.5Mbps
  - ▶ MPEG-2: mã hóa video với tiêu chuẩn truyền hình số CCIR 601 (720x480 điểm, 30 hình/giây) với tốc độ 2 - 10 Mbps.
  - ▶ MPEG-3
  - ▶ MPEG-4
  - ▶ MPEG-7