image processing Lecture

Book · January 2020

1 author:

nguyen dinh Cuong
Nha Trang University

LỜI NÓI ĐẦU

Xử lý ảnh là môn học đang được quan tâm và đã trở thành một môn học chuyên ngành của sinh viên ngành Công nghệ Thông tin cũng như những ngành kỹ thuật trong các trường Đại học kỹ thuật. Môn học này có liên quan đến nhiều ngành khác như: hệ thống tin học, lý thuyết thông tin, lý thuyết thống kê, nhận dạng.

Với mong muốn cung cấp tóm lược những kiến thức cơ bản của xử lý ảnh, bài giảng này được tham khảo trên nhiều nguồn tài liệu khác nhau nhằm cung cấp cho sinh viên có được cái nhìn tổng quát về lĩnh vực xử lý ảnh và những ứng dụng của nó trong cuộc sống.

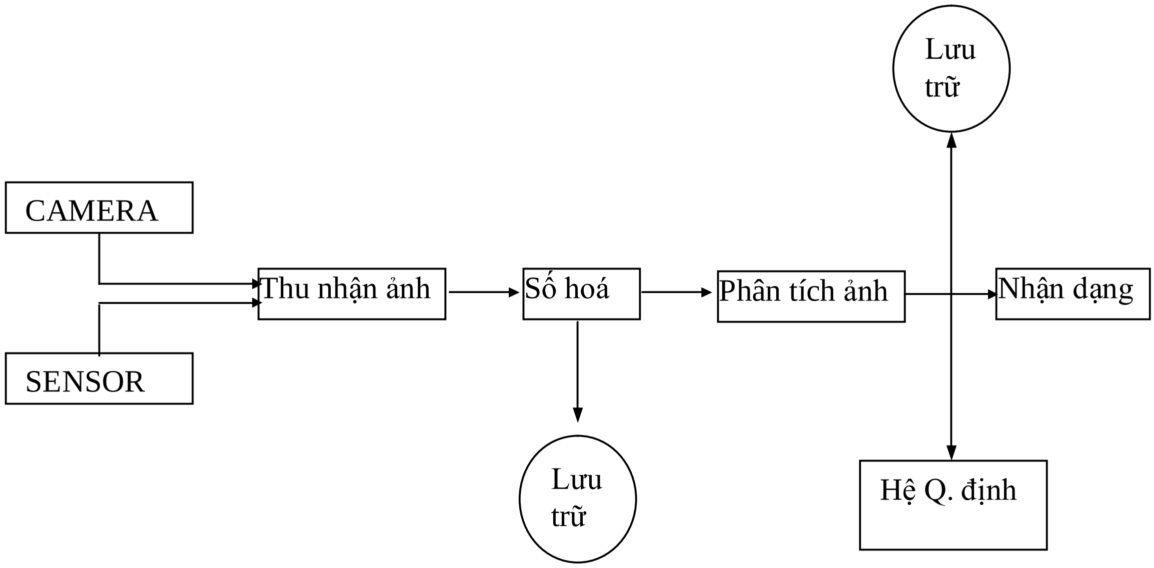
Xin chân thành cám ơn những ý kiến đóng góp quý báu của các đồng nghiệp và các bạn sinh viên đã góp ý cho bài giảng hoàn thiện hơn.

Nha Trang ngày 10 tháng 01 năm 2011

CHƯƠNG I NHẬP MÔN XỬ LÝ ẢNH

1.1 Tổng quan về một hệ thống xử lý ảnh

Xử lý ảnh là một khoa học còn tương đối mới mẻ so với nhiều ngành khoa học khác, nhất là trên quy mô công nghiệp, song trong xử lý ảnh đã bắt đầu xuất hiện những máy tính chuyên dụng. Để hình dung cấu hình của một hệ thống xử lý ảnh chuyên dụng, hay một hệ thống xử lý ảnh dùng trong nghiên cứu, ta có mô hình tổng quát của hệ xử lý ảnh như sau:



Hình 1.1 Các giai đoạn chính trong xử lý ảnh.

Thu nhận ảnh

Ảnh có thể thu nhận qua camera. Thông thường ảnh thu nhận qua camera là tín hiệu tương tự, nhưng cũng có thể là tín hiệu số hoá. Ảnh cũng có thể thu nhận từ vệ tinh qua bộ cảm ứng (sensor), hay ảnh, tranh được quét trên scanner.

Số hoá

Là quá trình biến đổi tín hiệu tương tự sang tín hiệu rời rạc (lấy mẫu) và số hoá bằng lượng hoá, trước khi chuyển sang giai đoạn xử lý, phân tích hay lưu trữ lại.

Phân tích ảnh

Bao gồm nhiều công đoạn nhỏ, trước hết là công việc tăng cường nâng cao chất lượng ảnh. Do những nguyên nhân khác nhau: có thể do thiết bị thu nhận ảnh, do nguồn sáng hay do nhiễu, ảnh có thể bị suy biến. Do vậy cần phải tăng cường và khôi phục lại ảnh để làm nổi bậc một số đặc tính chính của ảnh, hay làm cho ảnh gần giống với trạng thái gốc (trạng thái trước khi ảnh bị biến dạng). Giai đoạn tiếp theo là phát hiện các đặc tính như biên, phân vùng ảnh, trích chọn các đặc tính, ….

Nhận dạng ảnh

Cuối cùng tuỳ theo mục đích của ứng dụng, sẽ là giai đoạn nhận dạng, phân lớp hay các quyết định khác.

1.2 Các vấn đề cơ bản trong xử lý ảnh

1.2.1 Một số khái niệm

✓ **Pixel** (picture element): phần tử ảnh

Ảnh trong thực tế là một ảnh liên tục về không gian và giá trị độ sáng. Để có thể xử lý ảnh bằng máy tính cần thiết phải tiến hành số hoá. Trong quá trình số hoá người ta biến đổi tín hiệu liên tục sang tín hiệu rời rạc thông qua quá trình lấy mẫu(rời rạc hoá về không gian) và lượng hoá thành phần giá trị, về nguyên tắc bằng mắt thường không thể phân biệt được hai điểm kề nhau. Trong quá trình này, người ta sử dụng khái niệm *Picture element* mà ta quen gọi hay viết tắc là *Pixel* (phần tử ảnh). Như vậy ảnh là một tập hợp các pixel.

Như vây, một ảnh là tập hợp các điểm ảnh. Khi được số hoá, nó thường được biểu diễn bởi bảng 2 chiều I(n, p): n dòng và p cột. Ta nói ảnh nxp pixels. Người ta thường kí hiệu I(x,y) để chỉ một pixel. Một pixel có thể lưu trữ trên 1, 4, 8 hay 24 bit.

✓ **Grey leval (mức xám)**

Mức xám là kết quả sự mã hoá tương ứng cường độ một điểm ảnh với một giá trị số, kết quả của quá trình lượng hoá. Cách mã hoá kinh điển thường dùng là 16, 32, 64 mức. Mã 256 là phổ dụng nhất, mỗi pixel được mã hoá bởi 8 bit

1.2.2 Các bài toán cơ bản trong xử lý ảnh

A. Bài toán cải thiện ảnh

✓ **Mục đích**:

Tăng cường những thuộc tính cảm nhận ảnh, để phục vụ cho các bước xử lý tiếp theo.

✓ **Các thao tác:**

- Thay đổi độ tương phản, màu sắc, cường độ sáng dựa trên mô hình cảm nhận.
- Làm nét, làm trơn ảnh.

✓ **Phương pháp:**

- Phương pháp thao tác trên điểm.
- Các thao tác không gian, sử dụng các phép toán lọc, làm nét, làm trơn.

B. Bài toán khôi phục ảnh

✓ **Mục đích**

Khôi phục lại ảnh ban đầu, loại bỏ những biến dạng ra khỏi ảnh ban đầu.

✓ **Các thao tác**

- Lọc nhiễu.
- Giảm độ méo của tín hiệu.

✓ **Phương pháp**

- Lọc thích nghi(wiener, Kalman).
- Khôi phục ảnh từ các hình chiếu.

C. Bài toán phân tích ảnh

✓ **Mục đích**

Tìm ra các đối tượng ảnh, xây dựng các mối quan hệ của đối tượng ảnh dựa vào các đặc trưng cục bộ.

✓ **Các thao tác và phương pháp**

- Tìm biên và tách biên.
- Phân vùng ảnh.
- Phân loại đối tượng: đưa các đối tượng trong ảnh vào những nhóm đã biết trước.

D. Bài toán nén ảnh

✓ **Mục đích**

Làm giảm khối lượng thông tin chứa trong đối tượng, để phục vụ cho các bài toán khác(lưu trữ, truyền ảnh).

✓ **Các thao tác và phương pháp**

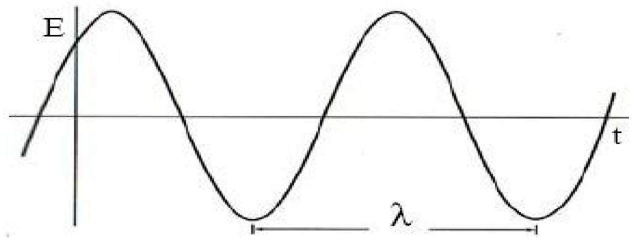
- Nén không mất thông tin: có các phương pháp sau, Huffman, Fano-Shannon, RLE, Zip-Lempel.
- Nén mất thông tin: JPEG (dựa trên phép biến đổi cosin), các phương pháp nén cho định dạng MPEG I, MPEG II.

1.3 Các mô hình màu.

Màu là gì

Có nhiều định nghĩa về màu (không có định nghĩa hình thức)

Từ góc nhìn khoa học: Màu là phân bổ các bước sóng l (red: 700 nm, violet: 400 nm), và tần số f, tốc độ ánh sáng c=lf.



Hình 1.2 Bước sóng màu

Từ góc nhìn về nghệ thuật và cuộc sống: Màu là Hue, Brightness, Saturation của ánh sáng. Sắc, độ sáng và bão hòa của đối tượng

Mô hình màu

Là phương pháp diễn giải các đặc tính và tác động của màu trong ngữ cảnh nhất định.

Không có mô hình màu nào là đầy đủ cho mọi khía cạnh của màu Người ta sử dụng các mô hình màu khác nhau để mô tả các tính chất được nhận biết khác nhau của màu.

Thí dụ

Mô hình màu RGB: ánh sáng Red, Green và Blue ứng dụng cho màn hình, TV.

Mô hình HSV: Nhận thức của con người.

Mô hình CMYK: Máy in.

1.3.1 Màu cơ sở và biểu đồ màu CIE

Năm 1931: CIE (Commision Internationale de l'Éclairage) xây dựng màu cơ sở chuẩn quốc tế:

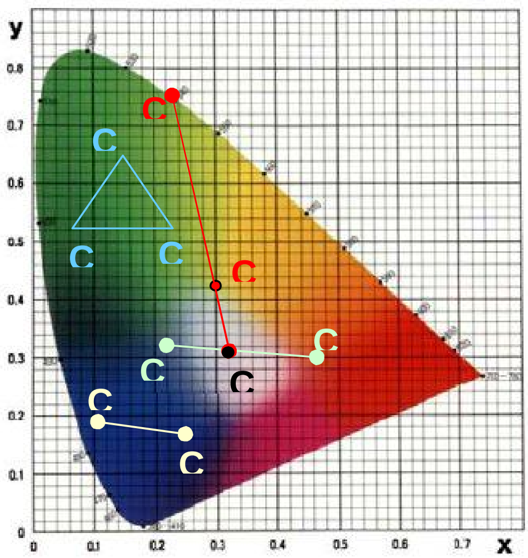
- Cho phép các màu khác được định nghĩa như tổng trọng lượng của ba màu cơ sở.
- Do không tồn tại 3 màu cơ sở chuẩn trong phổ nhìn thấy để tổng hợp màu mới do đó CIE sử dụng các màu tưởng tượng.
- Mỗi màu cơ sở trong CIE được xác định bằng đường cong phân bổ năng lượng.
- Nếu A, B, C là tổng số các màu cơ sở chuẩn cần xác định màu cho trước trong phổ nhìn thấy thì các thành phần của màu sẽ là:

$$x = \frac{A}{A+B+C} \quad y = \frac{B}{A+B+C} \quad z = \frac{C}{A+B+C}$$

- Nhưng x+y+z=1 cho nên chỉ cần 2 giá trị có thể xác định màu mới
- Cho khả năng biểu diễn mọi màu trên biểu đồ 2D ta có biểu đồ CIE

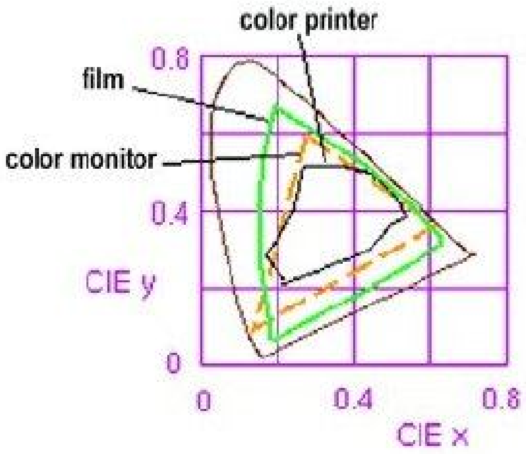
Biểu đồ CIE

- Khi vẽ các giá trị x, y của màu trong phổ nhìn thấy ta có biểu đồ CIE là đường cong hình lưỡi (còn gọi là biểu đồ kết tủa – CIE Chromaticity Diagram)
- Các điểm màu gán nhãn trên đường cong từ violet (400 nm) đến red (700 nm)
- Điểm C tương ứng màu trắng (ánh sáng ban ngày)
- Biểu đồ CIE là phương tiện lượng hóa độ tinh khiết và bước sóng trội:
- Độ tinh khiết của điểm màu C_1: được xác định bằng khoảng cách tương đối của đoạn thẳng nối C với đường cong qua C_1.
- Màu bù: biểu diễn bởi 2 điểm cuối C_3, C_4 của đoạn thẳng đi qua C.
- Gam màu xác định bởi 2 điểm: biểu diễn bởi đoạn thẳng nối hai điểm màu C_5, C_6.
- Gam màu xác định bởi 3 điểm: ba điểm C_7, C_8, C_9 chỉ xác định màu trong tam giác.



Hình 1.3 Vẽ biểu đồ màu CIE

- Ứng dụng biểu đồ CIE để so sánh gam màu các thiết bị ngoại vi. Máy in không thể in mọi màu hiển thị trên màn hình.



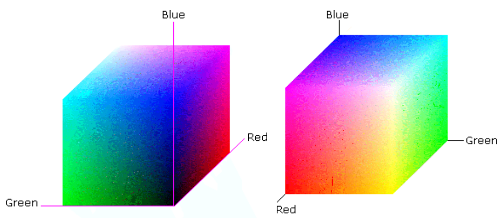
Hình 1.4 Phân bố hiển thị màu

Quan niệm về màu trực giác

- Họa sỹ vẽ tranh màu bằng cách trộn các chất màu với chất màu trắng và chất màu đen để có shade, tint và tone khác nhau:bắt đầu từ màu tinh khiết, bổ sung đen để có bóng (shade) màu. Nếu bổ sung chất màu trắng sẽ có tint khác nhau. Bổ sung cả chất màu trắng và đen sẽ có tone khác nhau.
- Cách biểu diễn này trực giác hơn mô tả màu bằng ba màu cơ sở. Các bộ chương trình đồ họa có cả hai mô hình màu: cho người sử dụng dễ tương tác với màu, các thành phần màu ứng dụng trên các thiết bị.

1.3.2 Mô hình màu RGB

- Mô hình màu RGB được biểu diễn bởi lập phương với các trục R, G, B.
- Gốc biểu diễn màu đen.
- Tọa độ (1, 1, 1) biểu diễn màu trắng.
- Tọa độ trên các cạnh trục biểu diễn các màu cơ sở.
- Các cạnh còn lại biểu diễn màu bù cho mỗi màu cơ sở.



Hình 1.5 Biểu diễn màu RGB

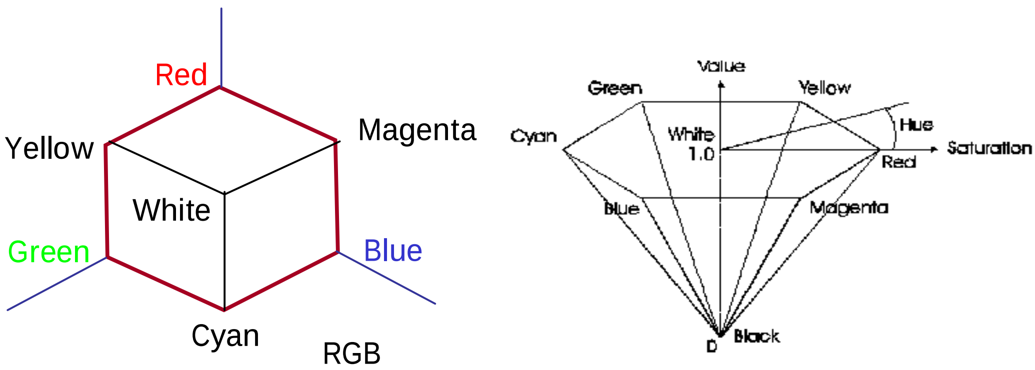
- Biểu đồ RGB thuộc mô hình cộng: Phát sinh màu mới bằng cách cộng cường độ màu cơ sở. Gán giá trị từ 0 đến 1 cho R, G, B. Red+Blue -> Magenta (1, 0, 1), đường chéo từ (0, 0, 0) đến (1, 1, 1) biểu diễn màu xám.

Nhận xét

- Mô hình này không thể biểu diễn mọi màu trong phổ nhìn thấy
- Đủ cho các ứng dụng máy tính
- Màn hình máy tính và TV sử dụng mô hình này
- Được sử dụng rộng rãi nhất
- Đơn giản

1.3.3 Mô hình màu HSV

- Thay vì chọn các phần tử RGB để có màu mong muốn, người ta chọn các tham số màu: Hue, Saturation và Value (HSV)
- Mô hình HSV suy diễn từ mô hình RGB: hãy quan sát hình hộp RGB theo đường chéo từ White đến Black (gốc) ta có hình lục giác, sử dụng làm đỉnh hình nón HSV.

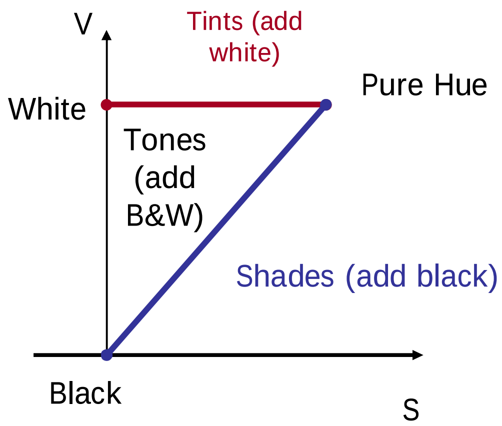


Hình 1.6 Tương quan màu RGB và HSV

- Hue: Bước sóng gốc của ánh sáng. Trong mô hình Hue được biểu diễn bằng góc từ 0^0 đến 360^0nValue: Cường độ hay độ chói ánh sáng.
- Value có giá trị [0, 1], V=0 màu đen. Đỉnh lục giác có cường độ màu cực đại.
- Saturation: Thước đo độ tinh khiết ánh sáng gốc. S trong khoảng [0, 1]. Biểu diễn tỷ lệ độ tinh khiết của màu sẽ chọn với độ tinh khiết cực đại.

Nhận xét

- Mô hình HSV trực giác hơn mô hình RGB. Bắt đầu từ Hue (H cho trước và V=1, S=1). thay đổi S: Bổ sung hay bớt trắng, thay đổi V: Bổ sung hay bớt đen cho đến khi có màu mong muốn.
- Mắt người có thể phân biệt 128 Hues, 130 tints và cực đại 30 shades (Yellow):128 x 130 x 30 = 382 720 màu khác nhau.



Hình 1.7 Biểu diễn HSV.

CHƯƠNG II HỆ THỐNG XỬ LÝ TÍN HIỆU SỐ 2 CHIỀU

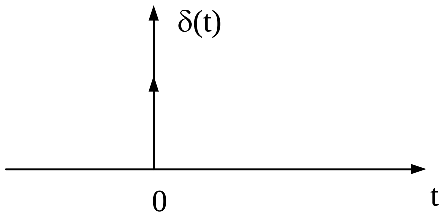
2.1 Một số tín hiệu 2 chiều cơ bản

2.1.1 Xung Dirac và xung đơn vị:

a Hệ thống một chiều

Xung dirac cho tín hiệu một chiều:

$$\delta(t) = \begin{cases} \infty\,; & t = 0 \\ 0\,; & t \neq 0 \end{cases} \qquad (2\text{-}1)$$



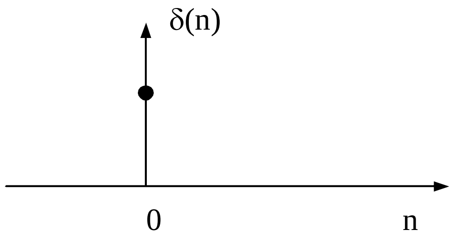
Hình 2.1 xung dirac tín hiệu 1 chiều

Biểu diễn dưới dạng công thức:

$$S(t) = \int_{-\infty}^{\infty} S(\tau)\delta(t-\tau)d\tau \qquad (2\text{-}2)$$

Xung đơn vị, tác động tại thời điểm t=0

$$\delta(n) = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \qquad (2\text{-}3)$$



Hình 2.2 Xung đơn vị, tín hiệu một chiều

Tại thời điểm n_0 tác động là $\delta(n-n_o)$

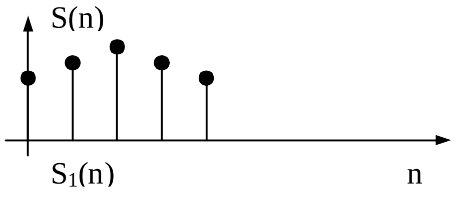
Một tín hiệu S(n), được biểu diễn tổng quát như sau

$$S(n) = \sum_{k=-\infty}^{\infty} S(k)\delta(n-k) \qquad (2\text{-}4)$$

Khai triển công thức trên ta có

$S(n)=S_1(n) + S_2(n)+\ldots+\ldots=S(1)\delta(n)+ S(2)\delta(n-1)+\ldots.+\ldots..$

Ví dụ

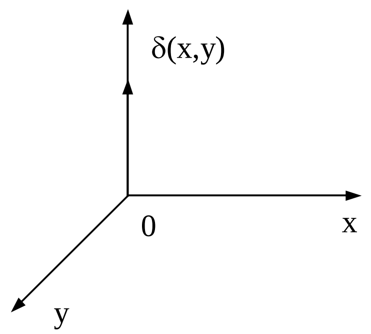


Hình 2.3 Biểu diễn tín hiệu 1 chiều

b Hệ thống hai chiều

Xung Dirac cho tín hiệu 2 chiều

$$\delta(x,y)=\begin{cases}\infty & x=0,y=0\\ 0 & x\neq 0,y\neq 0\end{cases} \qquad (2\text{-}5)$$



Hình 2.4 Xung dirac tín hiệu 2 chiều

Xung đơn vị cho tín hiệu 2 chiều``

$$\delta(m,n)=\begin{cases}1 & m=0,n=0\\ 0 & m\neq 0,n\neq 0\end{cases} \qquad (2\text{-}6)$$

Biểu diễn một tín hiệu 2 chiều

$$S(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(u,v)\delta(x-u,y-v)dudv$$ Dùng cho tín hiệu liên tục (2-7)

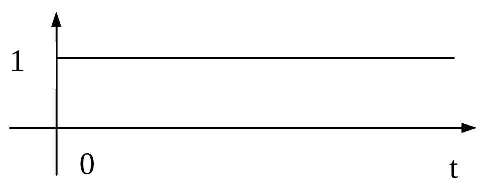
$$S(m,n)=\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}S(k,l)\delta(m-k,n-l)$$ Dùng cho tín hiệu rời rạc (2-8)

2.1.2 Tín hiệu đơn vị và bước nhảy đơn vị

a. Hệ thống một chiều

Tín hiệu đơn vị

$$u(t)=\begin{cases}1 & t\geq 0\\ 0 & t<0\end{cases} \qquad (2\text{-}9)$$



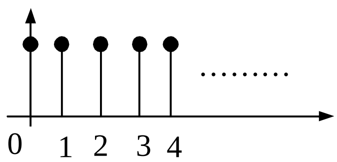
Hình 2.5 tín hiệu đơn vị

Tín hiệu đơn vị là tổng các xung Dirac tính từ -∞ đến thời điểm t.

$$u(t)=\int_{-\infty}^{t}S(\tau)dt \qquad \text{với điều kiện } \begin{cases}t<0 & u(t)=0\\ t\geq 0 & u(t)=1\end{cases} \qquad (2\text{-}10)$$

Bước nhảy đơn vị

$$u(n)=\begin{cases}1 & n\geq 0\\ 0 & n<0\end{cases} \qquad (2\text{-}11)$$



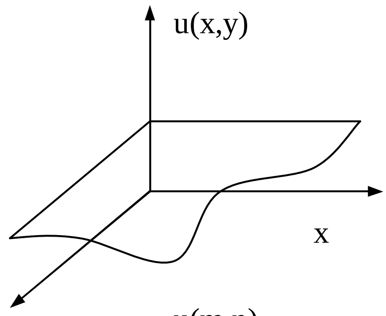
Hình 2.6 Bước nhảy đơn vị

$$u(n)=\sum_{k=-\infty}^{n}\delta(k) \quad \text{Ta cũng có } u(n-v)=\begin{cases}1 & n-k\geq 0\\ 0 & n-k<0\end{cases} \qquad (2\text{-}12)$$

b. Hệ thống 2 chiều

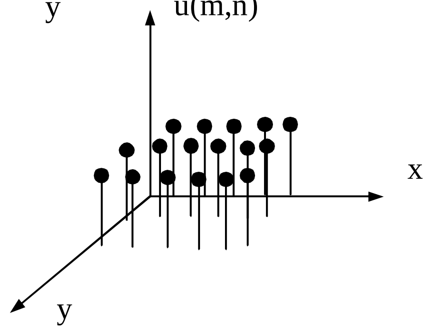
Với tín hiệu liên tục

$$u(x,y)=\begin{cases}1 & x\ge 0, y\ge 0\\ 0 & x<0, y<0\end{cases} \quad (2\text{-}13)$$



Với tín hiệu rời rạc

$$u(m,n)=\begin{cases}1 & m\ge 0, n\ge 0\\ 0 & m<0, n<0\end{cases} \quad (2\text{-}14)$$



Ta có cặp công thức

$$u(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(u,v)dudv \qquad (2\text{-}15)$$

$$u(m,n)=\sum_{k=-\infty}^{m}\sum_{l=-\infty}^{n}\delta(k,l) \qquad (2\text{-}16)$$

2.1.3 Một số tính chất với tín hiệu phân tách được

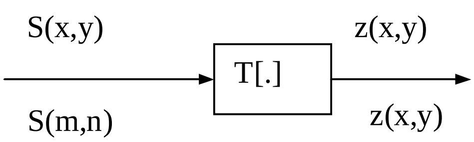
$S(x,y)=S_1(x)S_2(y)$

$\delta(x,y)=\delta(x)\delta(y)$

$\delta(m,n)=\delta(m).\delta(n)$

2.2 Hệ thống xử lý tín hiệu 2 chiều

2.2.1 Toán tử hệ thống



Ta có:

$$z(x,y) = T[S(x,y)]$$
$$z(m,n) = T[S(m,n)]$$

T : gọi là toán tử của hệ thống

Z : đáp ứng của hệ thống

Với một hệ thống tuyến tính ta có các công thức sau

Với $S_1(x,y) \xrightarrow{T} Z_1(x,y) \quad S_2(x,y) \xrightarrow{T} Z_2(x,y)$,

Ta có

$$\begin{cases} S(x,y) = S_1(x,y) + S_2(x,y) \xrightarrow{T} z(x,y) = z_1(x,y) + z_2(x,y) \\ \quad k = const \quad kS(x,y) \xrightarrow{T} kz(x,y) \end{cases} \tag{2-17}$$

Tổng quát ta có:

S(x, y) $\xrightarrow{T[.]}$ Z(x,y)

Với toán tử T[.] tuyến tính ta có quan hệ S(x,y) và z(x,y)

Ta có

$$S(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)\delta(x-u,y-v)dudv$$

$$Z(x,y) = T[S(x,y)] \quad = T[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)\delta(x-u,y-v)dudv] \quad = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)T[\delta(x-u,y-v)]dudv$$

Mặt khác

$T[\delta(x-u,y-v)] = h(x,y;u,v)$: gọi là đáp ứng xung của hệ thống tuyến tính bất biến

Với hệ thống tuyến tính bất biến dịch ta có:

$$h(x,u;0,0) = h(x,y)$$
$$h(x,y,u,v) = h(x-u,y-v)$$
$$T[\delta(x-u,y-v)] = h(x-u,y-v)$$

Ta có công thức tích chập

$$Z(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(u,v)h(x-u,y-v)dudv \tag{2-18}$$

$$Z(x,y) = S(x,y) \otimes h(x,y)$$

Với tín hiệu rời rạc

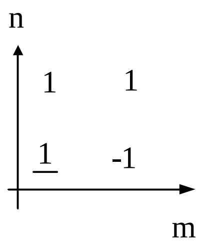
Ta có công thức tổng chập

$$Z(m,n) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} S(k,l)h(m-k,n-l)$$
$$Z(m,n) = S(m,n)h(m,n) \qquad (2\text{-}19)$$

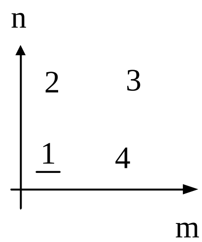
Ví dụ cho hệ thống H(m,n) xác định x(m,n) với S(m,n) cho trước

$x(m,n) = S(m,n) \otimes h(m,n)$

S(m,n)

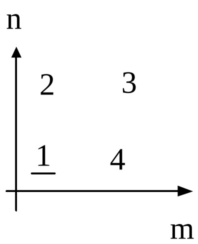


h(m,n)

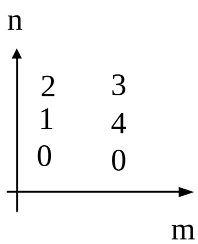


$$x(m,n) = S(m,n) \otimes h(m,n) = \sum_{k=\infty}^{\infty} \sum_{l=-\infty}^{\infty} S(k,l)xh(m-k,n-l)$$
$$= \sum_{k=0}^{1} \sum_{l=0}^{1} S(k,l)xh(m-k,n-l) = \sum_{l=0}^{1} S(0,l)h(m,n-l) + \sum_{l=0}^{1} S(1,l)h(m-1,n-1)$$
$$= S(0,0)h(m,n) + S(0,1)h(m,n-1) + S(1,0)h(m-1,n) + S(1,1)h(m-1,n-1)$$
$$= h(m,n) + h(m,n-1) - h(m-1,n) + h(m-1,n-1)$$

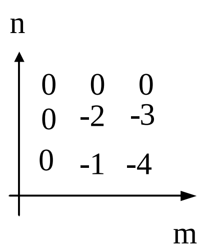
h(m,n)



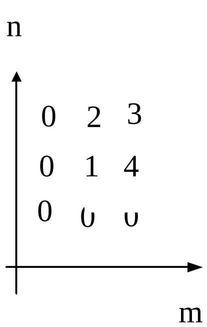
h(m,n-1)



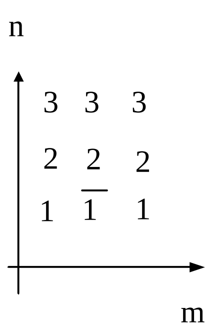
-h(m-1,n)



h(m-1,n-1)



S(m,n)



2.3 Các tính chất của tổng chập

a. Tính giao hoán

$$S(m,n)\otimes g(m,n) = g(m,n)\otimes S(m,n)$$

$$\sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} S(k,l)g(m-k,n-k) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} g(k,l)S(m-k,n-l) \tag{2-20}$$

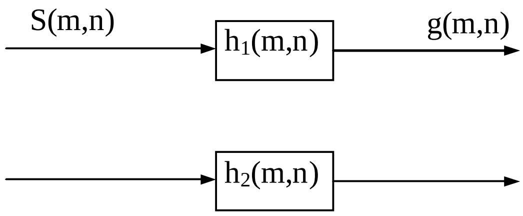
b. Tính kết hợp

Với các $S_1(m,n)$, $S_2(m,n)$ và $S_3(m,n)$

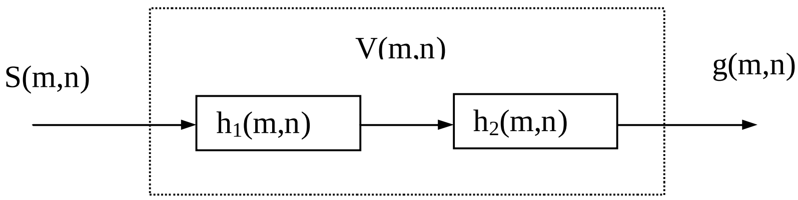
Ta có

$$S_1(m,n)\otimes\left[S_2(m,n)\otimes S_3(m,n)\right]=\left[S_1(m,n)\otimes S_2(m,n)\right]\otimes S_3(m,n)=S_1(m,n)\otimes S_2(m,n)\otimes S_3(m,n) \tag{2-21}$$

Suy ra ta có cách ghép nối 2 hệ thống tuyến tính bất biến như sau



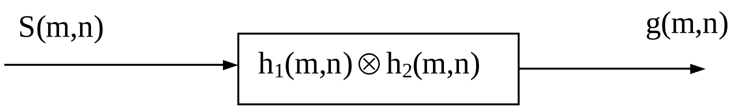
Ghép nối tiếp 2 hệ thống h_1 và h_2 là ghép đầu ra của h_1 nối vào đầu vào của h_2 nhưng tính chất của 2 hệ thống không thay đổi.



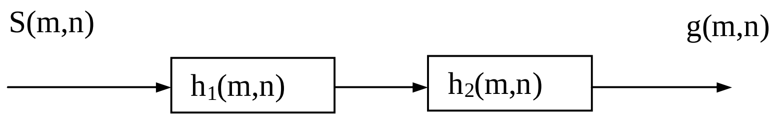
Ta có

$$\begin{aligned}
g(m,n) &= v(m,n)\otimes h_2(m,n)\\
v(m,n) &= S_1(m,n)\otimes h_1(m,n)\\
g(m,n) &= \left[S(m,n)\otimes h_1(m,n)\right]\otimes h_2(m,n) = S(m,n)\otimes\left[h_1(m,n)\otimes h_2(m,n)\right]\\
&= S(m,n)\otimes\left[h_2(m,n)\otimes h_1(m,n)\right] = \left[S(m,n)\otimes h_2(m,n)\right]\otimes h_1(m,n)
\end{aligned}$$

Như vậy hệ thống có thể tính toán như sau:



hoặc

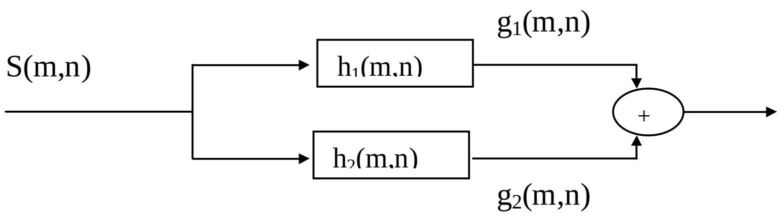


c. Tính chất phân phối với phép cộng

Với các tín hiệu $S_1(m,n)$, $S_2(m,n)$ và $S_3(m,n)$ ta có:

$$S_1(m,n) \otimes [S_2(m,n) + S_3(m,n)] = S_1(m,n) \otimes S_2(m,n) + S_1(m,n) \otimes S_3(m,n) \quad (2\text{-}22)$$

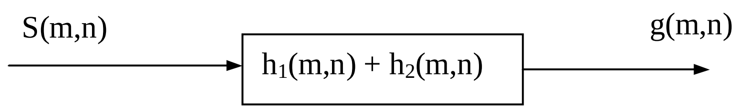
Từ đó ta có ghép nối song song của 2 hệ thống TTBB



Ta có

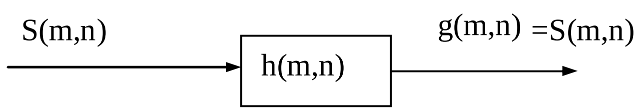
$$\begin{aligned} g(m,n) &= g_1(m,n) + g_2(m,n) \\ &= S(m,n) \otimes h_1(m,n) + S(m,n) h_2(m,n) \\ &= S(m,n) \otimes [h_1(m,n) + h_2(m,n)] \end{aligned}$$

Mô hình có thể tính toán như sau:



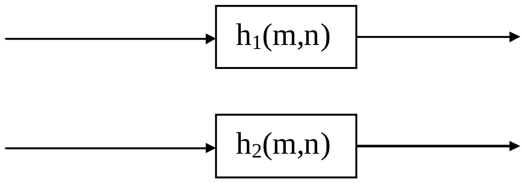
2.4 Một số hệ thống tuyến tính bất biến cơ bản

a. Hệ đồng nhất

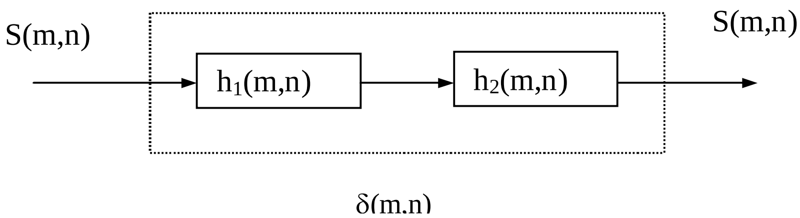


$h(m,n)=\delta(m,n)$

b. Hệ thống đảo

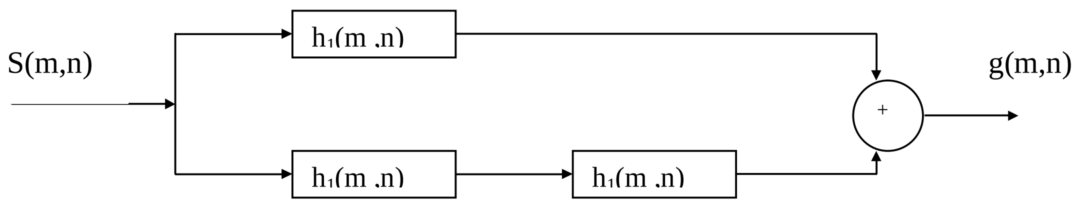


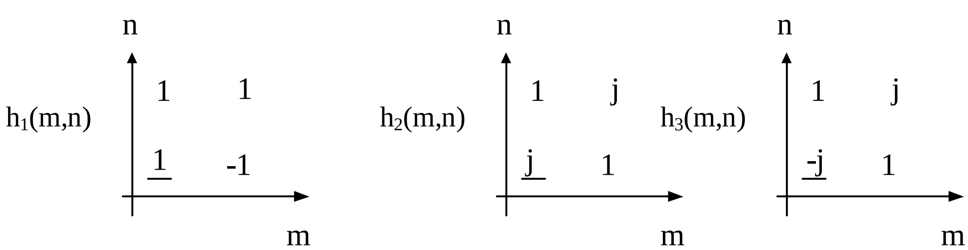
$h_2(m, n)$ gọi là hệ thống đảo của $h_1(m,n)$ khi và chỉ khi ghép nối tiếp 2 hệ thống sẽ trở thành hệ đồng nhất

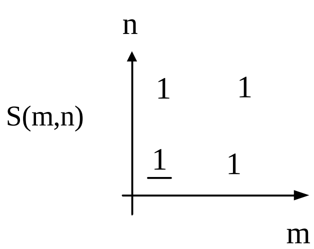


Ví dụ:

Cho một hệ thống xử lý ảnh được thiết kế như hình vẽ, hãy xác định tín hiệu ra g(m,n) ứng với tín hiệu vào S(m,n).







$$g(m,n) = S(m,n) \otimes h(m,n) + S(m,n)\left[h_2(m,n) + h_3(m,n)\right]$$
$$= S(m,n) \otimes \left[h_1(m,n) + h_2(m,n) \otimes h_3(m,n)\right]$$

Tính riêng: $h_2(m,n) \otimes h_3(m,n)$

$$h_2(m,n) \otimes h_3(m,n) = \sum_{k=0}^{1}\sum_{l=0}^{1} h_2(k,l) \cdot h_3(m-k,n-l)$$
$$= \sum_{l=0}^{1} h_2(0,l)h_3(m,n-l) + \sum_{l=0}^{1} h_2(1,l)h_3(m-1,n-l)$$
$$= h_2(0,0)h_3(m,n) + h_2(0,1)h_3(m,n-1) + h_2(1,0)h_3(m-1,n) + h_2(1,1)h_3(m-1.n-1)$$
$$= jh_3(m,n) + h_3(m,n-1) + h_3(m-1,n) + jh_3(m-1,n-1)$$

$$h_3(m,n-1):\ \begin{array}{cc} 1 & j \\ -j & 1 \\ 0 & 0 \end{array} \qquad h_3(m-1,n):\ \begin{array}{ccc} 0 & 1 & j \\ 0 & -j & 1 \end{array} \qquad jh_3(m-1,n-1):\ \begin{array}{ccc} 0 & j & -1 \\ 0 & 1 & j \\ 0 & 0 & 0 \end{array}$$

$$jh_3(m,n):\ \begin{array}{cc} j & -1 \\ 1 & j \end{array} \quad \xrightarrow{h_2 \otimes h_3} \quad H^*(m,n):\ \begin{array}{ccc} 1 & 2j & -1 \\ 0 & 2 & 2j \\ 1 & 0 & 1 \end{array}$$

$h(m,n)=h_1(m,n)+h^*(m,n)$

$$H(m,n):\ \begin{array}{ccc} 1 & 2j & -1 \\ 1 & 3 & 2j \\ 2 & -1 & 1 \end{array}$$

Kết quả cuối cùng của hệ thống ta có:

$S(m,n) \otimes h(m,n) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} S(k,l)h(m-k,n-l)$ Khai triển công thức trên với S(m,n) và H(m,n) ta sẽ thu được tín hiệu ra G(m,n).

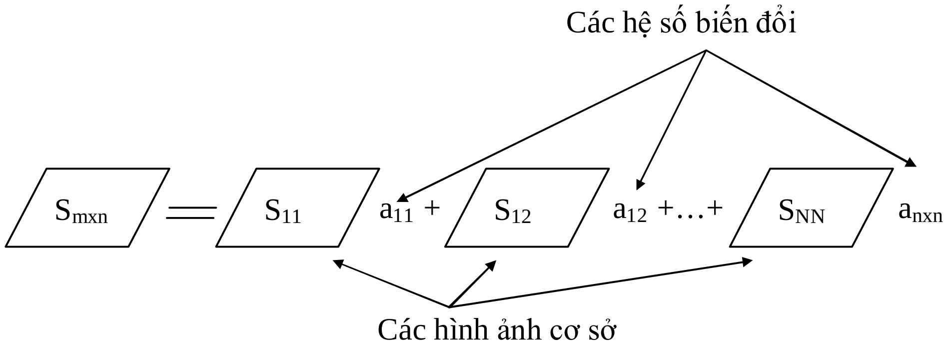
CHƯƠNG III CÁC PHÉP BIẾN ĐỔI ẢNH

3.1 Tổng quan về biến đổi ảnh trong không gian

Các phép biến đổi ảnh là cách tiếp cận thứ hai được áp dụng trong tín hiệu số nói chung và trong xử lý ảnh nói riêng. Phép biến đổi (transform) là thuật ngữ dùng để chỉ việc chuyển đổi sự biểu diễn của một đối tượng từ không gian này sang một không gian khác.

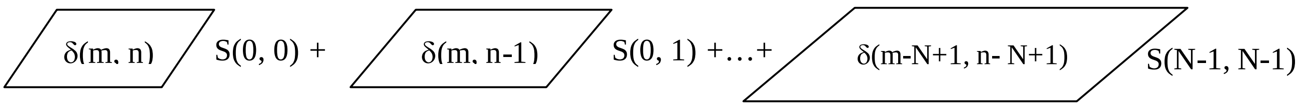
Ví dụ 1 :

Với 1 ảnh S_{mxn} ta có thể biểu diễn



$$S(m,n)= \sum_{k,l=-\infty}^{\infty} \sum S(k,l)\delta(m-k,n-k) \quad 0 \le m, n \le N-1$$

Hay: S(m,n) =



Ví dụ 2:

Biến đổi Furier với tín hiêu 1 chiều N mẫu:

$$x(n) = \sum_{k=0}^{N-1} X(k) e^{j\frac{2k\pi n}{N}}$$

$$X(k) = \frac{1}{N}\sum_{n=0}^{N-1} x(n) e^{-j\frac{2k\pi n}{N}}$$

3.2 Các phép biến đổi đơn vị

3.2.1 Ma trận trực giao và ma trận Unitar

Cho A là một ma trận vuông

A trực giao khi: $A^{-1} = A^T$ hay $A A^T = I$ (3-1)

Trong đó A^{-1} là ma trận đảo của A.

A^T là ma trận chuyển vị của A.

Ma trận A được gọi là ma trận Unitar nếu:

$$A^{-1} = A^{*T} \quad \text{hay} \quad AA^{*T} = I \tag{3-2}$$

Các phần tử của A^* được xác định như sau với aij= x + jy thì a^*_{ij} = x – jy (dạng số phức tổng quát).

Nhận xét :

Nếu các phần tử của ma trận A có giá trị là số thực thì nếu A trực giao $\Leftrightarrow$ A unitar

Ví dụ 1

Xét xem ma trận A sau đây có phải là ma trận Unitar không

$$A = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}$$

Giải :

Ta có

$$A^T = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \qquad AA^T = \frac{1}{2}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = I$$

A trực giao $\Rightarrow$ A Unitar

Ví dụ 2

Kiểm tra tính Unitar của ma trận sau

$$A = \begin{vmatrix} \sqrt{2} & j \\ -j & \sqrt{2} \end{vmatrix}$$

Nhận xét

$$A^T = \begin{vmatrix} \sqrt{2} & -j \\ j & \sqrt{2} \end{vmatrix} \qquad AA^T = \begin{vmatrix} \sqrt{2} & j \\ -j & \sqrt{2} \end{vmatrix}\begin{vmatrix} \sqrt{2} & -j \\ j & \sqrt{2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} = I$$

Tuy nhiên

$$A^* = \begin{vmatrix} \sqrt{2} & -j \\ j & \sqrt{2} \end{vmatrix}, \quad A^{*T} = \begin{vmatrix} \sqrt{2} & j \\ -j & \sqrt{2} \end{vmatrix}, \quad AA^{*T} = \begin{vmatrix} \sqrt{2} & j \\ -j & \sqrt{2} \end{vmatrix}\begin{vmatrix} \sqrt{2} & j \\ -j & \sqrt{2} \end{vmatrix} = \begin{vmatrix} 3 & 2j\sqrt{2} \\ -2j\sqrt{2} & 3 \end{vmatrix} \neq I$$

Vậy A không Unitar

Ví dụ 3

Xét ma trận

$$A = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & j \\ j & 1 \end{vmatrix}, A^T = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & j \\ j & 1 \end{vmatrix}, AA^T = \frac{1}{2}\begin{vmatrix} 1 & j \\ j & 1 \end{vmatrix}\begin{vmatrix} 1 & j \\ j & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 0 & 2j \\ 2j & 0 \end{vmatrix} \neq I$$

Tuy nhiên ta lại có:

$$A^{*T} = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & -j \\ -j & 1 \end{vmatrix}, \ AA^{*T} = \frac{1}{2}\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = I$$

$\Rightarrow$ A là ma trận Unitar

Bài tập tự làm:

Xét tính Unitar của ma trận sau:

$$A = \frac{1}{\sqrt{3}}\begin{vmatrix} 1 & 1 & 1 \\ 1 & \frac{-1}{2} - j\frac{\sqrt{3}}{2} & \frac{-1}{2} + j\frac{\sqrt{3}}{2} \\ 1 & \frac{-1}{2} + j\frac{\sqrt{3}}{2} & \frac{-1}{2} - j\frac{\sqrt{3}}{2} \end{vmatrix}$$

3.2.2 Phép biến đổi Unitar một chiều

Cho $\vec{S}$ = S(n) = (S(0), S(1), S(2),…S(N-1)) và A_{nxn} là ma trận Unitar. Ta có ảnh của $\vec{S}$ qua phép biến đổi Unitar biểu diễn bằng ma trận A :

$$\vec{V} = A\vec{S} \quad \text{hay} \quad v(k) = \sum_{n=0}^{N-1} a_{kn}\, S(n) \tag{3-3}$$

Ví dụ:

$S(n) = (S_1, S_2, S_3)^T$, ma trận unitar $A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$

Ta có $\vec{V} = A\vec{S} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \times \begin{vmatrix} S_1 \\ S_2 \\ S_3 \end{vmatrix} = \begin{vmatrix} a_{11}S_1 + a_{12}S_2 + a_{13}S_3 \\ a_{21}S_1 + a_{22}S_2 + a_{23}S_3 \\ a_{31}S_1 + a_{32}S_2 + a_{33}S_3 \end{vmatrix}$

Phép biến đổi Unitar ngược:

$$\vec{S} = A^{-1}\vec{V} \quad \text{Suy ra:} \quad \vec{S} = A^{*T}\vec{V} \qquad (3\text{-}4)$$

Hay ta có công thức:

$$S(n) = \sum_{k=0}^{N-1} b_{nk}\, v(k) = \sum_{k=0}^{N-1} \overset{*}{a}_{kn}\, v(k) \qquad (3\text{-}5)$$

Trong đó:

$$A^{*T} = \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix}$$

k

n

Như vậy: $\overset{*}{a}_{kn} = b_{nk}$

Kết luận:

Với hình ảnh cơ sở $\overset{*}{a}_k$ là cột k của ma trân A^{*T}, ta tách $\vec{S}$ thành các hình ảnh cơ sở thông qua các hệ số của $\vec{V}$

Các hệ cơ sở

$$\vec{S} = \overset{*\rightarrow}{a_0}\, v(0) + \overset{*\rightarrow}{a_1}\, v(1) + \ldots + \overset{*\rightarrow}{a_{N-1}}\, V(N-1)$$

Các hinh ảnh cơ sở

3.2.3 Phép biến đổi Unitar 2 chiều

Cho ma trận Unitar A_{nxn} , với ảnh S(m, n) ta có công thức biến đổi Unitar của ảnh S như sau:

Cặp biến đổi Unitar 2 chiều:

$V = ASA^T$ (Xác định hệ cơ sở)

$S = A^{*T}VA^*$ (Xác định ảnh cơ sở) (3-6)

Hay $S = \sum_{k,l=0}^{N-1}\sum \overset{*}{A}_{k,l} V(k,l)$, với $\overset{*}{A}_{k,l}$: là hình ảnh cơ sở

$$\overset{*}{A}_{k,l} = \overset{*}{a}_k \overset{*T}{a}_l$$

Trong đó : $\overset{*}{a}_k$ và $\overset{*}{a}_l$ là các cột thứ k và l của A^{*T}

Ví dụ: Cho ma trận Unitar A và ảnh S, hãy xác định các ảnh của S qua phép biến đổi

$$A = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \quad \text{và} \quad S = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

Giải:

Xác định hệ cơ sở:

$$V = ASA^T = \frac{1}{2}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 4 & 6 \\ -2 & -2 \end{vmatrix}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 10 & -2 \\ -4 & 0 \end{vmatrix}$$

$$A^{*T} = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}$$

Xác định các $A^*_{k,l} = a^*_k a^{*T}_l$

Ta có : $a^*_0 = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 \\ -1 \end{vmatrix}$ và $a^*_1 = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 \\ -1 \end{vmatrix}$

$$A^*_{00} = a^*_0 a^{*T}_0 = \frac{1}{2}\begin{vmatrix} 1 \\ 1 \end{vmatrix}\begin{vmatrix} 1 & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}, \quad A^*_{10} = a^*_1 a^{*T}_0 = \frac{1}{2}\begin{vmatrix} 1 \\ -1 \end{vmatrix}\begin{vmatrix} 1 & 1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix}$$

$$A^*_{01} = a^*_0 a^{*T}_1 = \frac{1}{2}\begin{vmatrix} 1 \\ 1 \end{vmatrix}\begin{vmatrix} 1 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix}, \quad A^*_{11} = a^*_1 a^{*T}_1 = \frac{1}{2}\begin{vmatrix} 1 \\ -1 \end{vmatrix}\begin{vmatrix} 1 & -1 \end{vmatrix} = \frac{1}{2}\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}$$

Như vậy S có thể biểu diễn qua các hình ảnh cơ sở như sau:

$$S = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \frac{5}{2}\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} - \frac{1}{2}\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} - \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}$$

Hình ảnh cơ sở

Ví dụ 2:

Cho ma trận Unitar A và ảnh S, hãy xác định V và $A^*_{k,l}$

$$A = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & j \\ j & 1 \end{vmatrix} \quad \text{và} \quad S = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$$

Giải:

$$V= ASA^T = \frac{1}{2}\begin{vmatrix}1 & j\\ j & 1\end{vmatrix}\begin{vmatrix}1 & 2\\ 3 & 4\end{vmatrix}\begin{vmatrix}1 & j\\ j & 1\end{vmatrix} = \frac{1}{2}\begin{vmatrix}1+3j & 2+4j\\ 3+j & 4+2j\end{vmatrix}\begin{vmatrix}1 & j\\ j & 1\end{vmatrix} = \frac{1}{2}\begin{vmatrix}-3+5j & -1+5j\\ 1+5j & 3+5j\end{vmatrix}$$

$$A^{*T}= \frac{1}{\sqrt{2}}\begin{vmatrix}1 & -j\\ -j & 1\end{vmatrix}$$

Tính $A^{*}_{k,l} = a^{*}_{k} a^{*T}_{l}$ với $a^{*}_{0} = \frac{1}{\sqrt{2}}\begin{vmatrix}1\\ -j\end{vmatrix}$ và $a^{*}_{1} = \frac{1}{\sqrt{2}}\begin{vmatrix}-j\\ 1\end{vmatrix}$

$$A^{*}_{00} = a^{*}_{0} a^{*T}_{0} = \frac{1}{2}\begin{vmatrix}1\\ -j\end{vmatrix}\begin{vmatrix}1 & -j\end{vmatrix} = \frac{1}{2}\begin{vmatrix}1 & -j\\ -j & -1\end{vmatrix}$$

$$A^{*}_{01} = a^{*}_{0} a^{*T}_{1} = \frac{1}{2}\begin{vmatrix}1\\ -j\end{vmatrix}\begin{vmatrix}-j & 1\end{vmatrix} = \frac{1}{2}\begin{vmatrix}-j & 1\\ 1 & -j\end{vmatrix}$$

$$A^{*}_{10} = a^{*}_{1} a^{*T}_{0} = \frac{1}{2}\begin{vmatrix}-j\\ 1\end{vmatrix}\begin{vmatrix}1 & -j\end{vmatrix} = \frac{1}{2}\begin{vmatrix}-j & -1\\ 1 & -j\end{vmatrix}$$

$$A^{*}_{11} = a^{*}_{1} a^{*T}_{1} = \frac{1}{2}\begin{vmatrix}-j\\ 1\end{vmatrix}\begin{vmatrix}-j & 1\end{vmatrix} = \frac{1}{2}\begin{vmatrix}-1 & -j\\ -j & 1\end{vmatrix}$$

3.3 Biến đổi Fourier và biến đổi KL

3.3.1 Biến đổi Fourier 1 chiều

Cho f(x) là hàm liên tục với biến thực x. Biến đổi Furier của f(x) là $\Im\{f(x)\}$:

$$\Im\{f(x)\} = F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux}\, dx \qquad (3\text{-}7)$$

Trong đó j= $\sqrt{-1}$

Cho F(u), f(x) có thể nhận được bằng cách biến đổi Furier ngược (IFT):

$$\Im\{F(u)\} = f(x) = \int_{-\infty}^{\infty} F(u)\, e^{j2\pi ux} du \qquad (3\text{-}8)$$

Công thức trên là cặp biến đổi Furier tồn tại nếu f(x) liên tục và có thể tích phân được, và F(u) cũng có thể tích phân được. Trong thực tế các điều kiện trên luôn thoả mãn.

Với f(x) là hàm thực, biến đổi Furier của hàm thực nói chung là số phức:

$$F(u) = R(u) + j\, I(u) \qquad (3\text{-}9)$$

Trong đó R(u) và I(u) là thành phần thực và thành phần ảo của F(u). Ta thường biểu diễn dưới dạng hàm mũ

$$F(u)= |F(u)| e^{j\phi(u)} \qquad (3\text{-}10)$$

Trong đó:

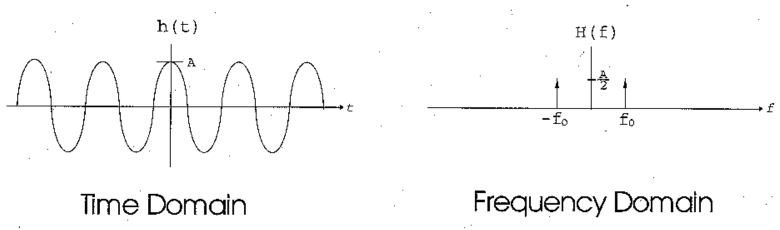
$$|F(u)| = \left[R^2(u) + I^2(u)\right]^{\frac{1}{2}} \text{ và } \phi(u) = \tan^{-1}\left[\frac{I(u)}{R(u)}\right] \qquad (3\text{-}11)$$

Hàm biên độ $|F(u)|$ được gọi là phổ Furier của f(x), và $\phi(u)$ gọi là góc pha. Bình phương của phổ gọi là phổ công suất của f(x).

Biến u thường được gọi là biến tần số(phần biểu diễn hàm mũ) $e^{-j2\pi ux}$, theo công thức Euler:

$$e^{-j2\pi ux} = \cos 2\pi ux - j\sin 2\pi ux \qquad (3\text{-}12)$$

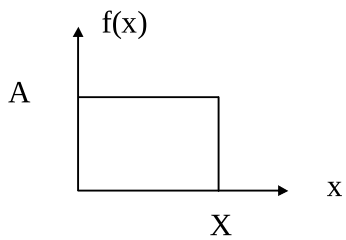
Khai triển tích phân ta có tổng các phần tử rời rạc của các thành phần cosine và sine, mỗi giá trị u xác định tần số của cặp sine- cosine.



Hình 3.1 Biểu diễn tín hiệu miền thời gian và tần số

Ví dụ:

Ta có hàm f(x) như sau:



$$F(u) = \int_{-\infty}^{\infty} f(x)\, e^{-j2\pi ux} dx = \int_0^X A e^{-j2\pi ux} dx = \frac{-A}{j2\pi u}\left[e^{-j2\pi ux}\right]_0^X = \frac{-A}{j2\pi ux}\left[e^{-j2\pi ux} - 1\right]$$

$$= \frac{A}{j2\pi u}\left[e^{j2\pi ux} - e^{-j2\pi ux}\right] e^{-j\pi ux} = \frac{A}{\pi u}\sin(\pi ux)\, e^{-j\pi ux}$$

Đó là một hàm phức, phổ Fourier: $|F(u)| = \left|\frac{A}{\pi u}\right| |\sin(nux)| \left|e^{-j\pi ux}\right| = Ax\left|\frac{\sin(\pi ux)}{(\pi ux)}\right|$

3.3.2 Biến đổi Fourier 2 chiều

Biến đổi Fourier có thể mở rộng cho hàm f(x, y) với 2 biến. Nếu f(x, y) là biến liên tục và tích phân được và F(u, v) cũng tích phân được, thì cặp biến đổi Fourier 2 chiều sẽ là :

$$\Im\ \{f(x,y)\} = F(u,v) = \int\int_{-\infty}^{\infty} f(x,y)\, e^{-j2\pi(ux+vy)}\, dxdy \qquad (3\text{-}13)$$

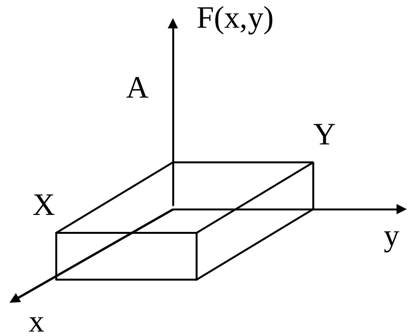
$$\Im^{-1}\ \{f(u,v)\} = f(x,y) = \int\int_{-\infty}^{\infty} F(u,v)\, e^{j2\pi(ux+vy)}\, dudv \qquad (3\text{-}14)$$

Trong đó u, v là biến tần số.

Cũng như biến đổi Fourier 1 chiều, ta có phổ Fourier, pha, phổ công suất cho trường hợp 2 chiều:

$$|F(u,v)| = \left[R^2(u,v) + I^2(u,v)\right]^{\frac{1}{2}} \quad \text{và} \quad \phi(u,v) = \tan^{-1}\left|\frac{I(u,v)}{R(u,v)}\right| \qquad (3\text{-}15)$$

Ví dụ: xác định biến đổi Fourier của hàm trên hình sau:



F(u, v)= $\int\int_{-\infty}^{\infty} f(x,y)\, e^{-j2\pi(ux+vy)}\, dxdy = A\int_0^X e^{-j2\pi ux}\, dx \int_0^Y e^{-j2\pi vy}\, dy = A\left[\frac{e^{-j2\pi ux}}{-j2\pi ux}\right]_0^X \left[\frac{e^{-j2\pi vy}}{-j2\pi vy}\right]_0^Y$

$= \frac{A}{-j2\pi u}\left[e^{-j2\pi uX} - 1\right]\frac{1}{-j2\pi v}\left[e^{-j2\pi Y} - 1\right] = AXY\left[\frac{\sin(\pi uX)\, e^{-j\pi uX}}{\pi uX}\right]\left[\frac{\sin(\pi vY)\, e^{-j\pi vY}}{\pi vY}\right]$

Phổ của nó: $|F(u,v)|^2 = \mathrm{AXY}\left|\frac{\sin(\pi \mathrm{uX})}{(\pi \mathrm{uX})}\right| \left|\frac{\sin(\pi \mathrm{vY})}{(\pi \mathrm{vY})}\right|$

3.3.3 Biến đổi Fourier rời rạc (DFT)

Giả thiết cho hàm liên tục f(x), được rời rạc hoá thành chuỗi:

$\{f\{x_0\},\ f\{x_0+\Delta x\},\ f\{x_0+2\Delta x\},\ f\{x_0+[N-1]\Delta x\}\}$

Trong đó:N- số mẫu, Δx bước rời rạc. Ta dùng biến x vừa là biến liên tục vừa là biến rời rạc.

Ta định nghĩa : $f(x)= f(x_0 + x\Delta x)$

x: - là các giá trị rời rạc 0, 1, 2,…, N-1.

Chuỗi $\{f(0),\ f(1),\ f(2),...f(N-1)\}$ là các mẫu đều bất kì N từ một hàm liên tục. Cặp biến đổi Fourier cho các hàm lấy mẫu:

$$F(u)=\frac{1}{N}\sum_{x=0}^{N-1} f(x)e^{\frac{-j2\pi ux}{N}} \quad \text{với } u= 0, 1, 2, \ldots N-1 \qquad (3\text{-}16)$$

$$\text{Và } f(x) = \sum_{x=0}^{N-1} F(u)e^{\frac{j2\pi ux}{N}} \quad \text{với } x=0, 1, 2,\ldots N-1 \qquad (3\text{-}17)$$

Các giá trị u=0, 1, 2…N-1 trong DFT (u) ứng với các mẫu trong biến đổi liên tục 0, Δu, 2Δu,…(N-1)Δu.

Nói cách khác F(u) biểu diễn F(uΔu): $\Delta u = \dfrac{1}{N\Delta x}$ (3-18)

Trường hợp DFT 2 chiều:

$$F(u, v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(\frac{ux}{M}+\frac{vy}{N})} \qquad (3\text{-}19)$$

$$f(x,y)=\sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)e^{j2\pi(\frac{ux}{M}+\frac{vy}{N})} \qquad (3\text{-}20)$$

với $u=\overline{0, M-1}$, $v=\overline{0, N-1}$ và $x=\overline{0, M-1}$, $y=\overline{0, N-1}$

Tương tự như trong một chiều, hàm rời rạc f(x, y) biểu diễn các mẫu của hàm $f(x_0 + x\Delta x, y_0+ y_0\Delta y)$. Tương tự ta tính F(u, v) quan hệ giữa miền không gian và miền tần số được tính như sau:

$$\Delta u = \frac{1}{M\Delta x} \quad \text{và} \quad \Delta v = \frac{1}{N\Delta x} \qquad (3\text{-}21)$$

Nếu M=N (lấy mẫu vuông):

Ta có : $$F(u,v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(\frac{ux+vy}{N})} \qquad (3\text{-}22)$$

$$f(x,y)=\frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}F(u,v)e^{j2\pi(\frac{ux+vy}{N})} \qquad (3\text{-}23)$$

với x, y=0, 1, 2,…N-1

3.3.4 Một số đặc tính của biến đổi Fourier 2-D

Phạm vi động của phổ Fourier thường cao hơn so với hiển thị thường dùng(trong đó chỉ có phần tử sáng nhất của ảnh được nhìn thấy). Nó làm tăng khả năng nhìn thấy các chi tiết

a. phân biệt trong biến đổi Fourier

Với cặp F(u, v), f(x,y) 2-D ta có thể nhận được theo 2 bước biến đổi Fourier 1-D thuận hoặc biến đổi ngược.

Ví dụ: $F(u,v)=\frac{1}{N}\sum_{v=0}^{N-1}F(x,v)e^{\frac{-j2\pi ux}{N}}$ trong đó $F(x,v)=N\left[\frac{1}{N}\sum_{y=0}^{N-1}f(x,y)e^{\frac{-j2\pi vy}{N}}\right]$

b. Chuyển dịch

Đặc trưng chuyển dich của cặp biến đổi Fourier là

$$f(x,y)e^{\frac{j2\pi(u_0x+v_0y)}{N}} \Leftrightarrow F(u-u_0,v-v_0) \qquad (3\text{-}24)$$

và $f(x-x_0,y-y_0)\Leftrightarrow F(u,v)e^{\frac{-j2\pi(ux_0+vy_0)}{N}}$ (3-25)

Trong đó mũi tên kép biểu diễn quan hệ giữa 1 hàm và biến đổi Fourier của nó(hoặc ngược lại)

c. Chu kì và liên hiệp đối xứng

DFT và biến đổi ngược theo chu kì N ta có:

$$F(u,v)=F(u+N,v)=F(u,v+N)=F(u+n,v+N) \qquad (3\text{-}26)$$

Nếu f(x, y) là thực, thì biến đổi Fourier có liên hiệp đối xứng

$$F(u,v)=F^{*}(-u,-v) \text{ và } |F(u,v)|=|F(-u,-v)| \qquad (3\text{-}27)$$

d. Quay

Nếu sử dụng tọa độ cực: $x=r\cos\theta$; $y=r\sin\theta$; $u=\varpi\cos\phi$; $v=\varpi\sin\theta$ thì f(x,y) và F(u, v) trở thành $f(r,\theta)$ và $F(\varpi,\phi)$. Thay các biểu thức trên vào DFT hoặc FT liên tục, ta có:

$$f(r,\theta+\theta_0)\Leftrightarrow F(\varpi,\phi+\theta_0) \qquad (3\text{-}28)$$

Nói cách khác, quay f(x, y) một góc θ_0 sẽ làm quay F(u, v) cùng một góc. Tương tự, ta quay F(u, v) cũng sẽ làm quay f(x,y) cùng một góc.

e. Phân bố và thang độ

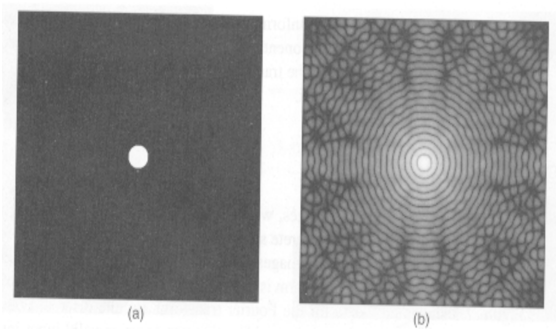
Định nghĩa cặp biến đổi tương tự hoặc rời rạc:

$$\Im\{f_1(x, y) + f_2(x, y)\} = \Im\{f_1(x, y)\} + \Im\{f_2(x, y)\} \tag{3-29}$$

và $\Im\{f_1(x, y).f_2(x, y)\} = \Im\{f_1(x, y)\}.\Im\{f_2(x, y)\}$

Đối với 2 thang a và b ta có:

$$af(x, y) \Leftrightarrow aF(u, v) \text{ và } f(ax, bx) \Leftrightarrow \frac{1}{ab}F\left(\frac{u}{a}, \frac{v}{b}\right) \tag{3-30}$$



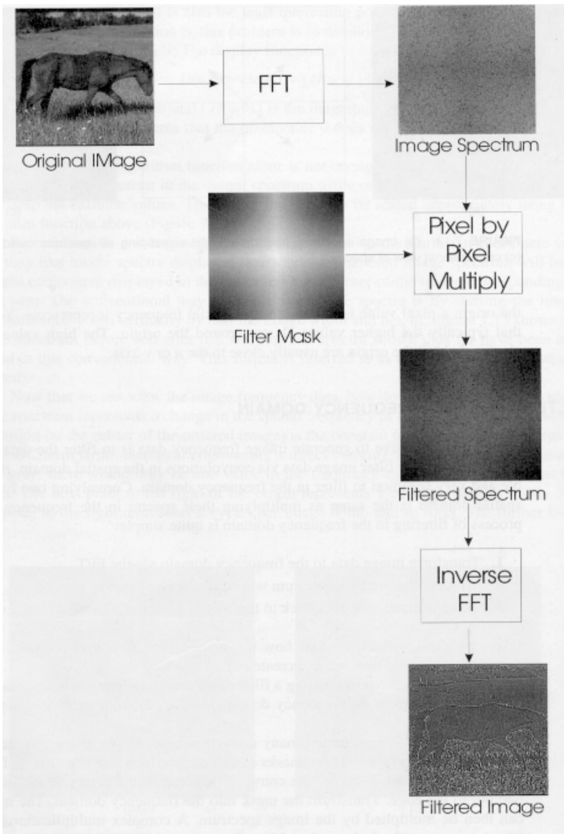
Hình 3.2 Biến đổi Fourier của tín hiệu, a. ảnh gốc, b. kết quả biến đổi Fourier

3.4 Lọc tín hiệu trong miền tần số

Việc xử lý lọc ảnh trong miền tần số được thực hiện theo các bước đơn giản sau:

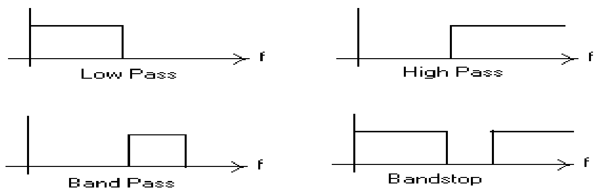
1. Biến đổi dữ liệu ảnh vào miền tần số (xử dụng biến đổi Fourier)
2. Nhân phổ của ảnh với một mặt nạ lọc
3. Biến đổi ngược ảnh về miền không gian.

Chúng ta đã biết cách biến đổi FFT thuận và ngược để xử dụng lọc tín hiệu. Để tạo mặt nạ lọc có 2 cách: biến đổi convoluotion của mặt nạ từ miền không gian vào miền tần số, hoặc tính toán trực tiếp trên miền tần số. Hình cho 3.3 cho thấy ảnh được lọc trên miền tần số.



Hình 3.3 Ảnh được lọc trên miền tần số

Trên miền tần số một số bộ lọc hay dùng: lọc thông thấp, thông cao và thông dải. Hình 3.4 biểu diễn các bộ lọc trên miền tần số.



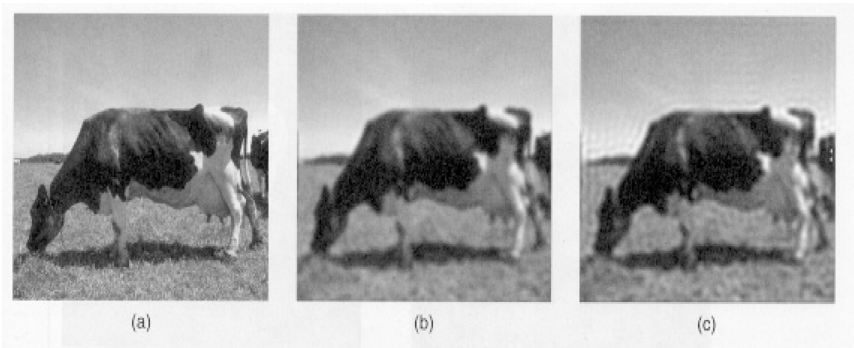
Hình 3.4 Biểu diễn bộ lọc thông thấp, thông cao và thông dải trên miền tần số

Ví dụ:

Ta có bộ lọc làm trơn tín hiệu trên miền tần số là bộ lọc Butterworth. Biểu diễn bộ lọc như sau

$$H(u,v) = \frac{1}{1+\left[\frac{D(u,v)}{D_0}\right]^{2n}} \quad (3\text{-}31)$$

$$D(u,v) = \sqrt{\left(u^2 + v^2\right)} \quad (3\text{-}32)$$



Hình 3.5 (a) ảnh ban đầu, (b) ảnh sau khi lọc thông thấp, (b) ảnh xử dụng bộ lọc Buuterworth

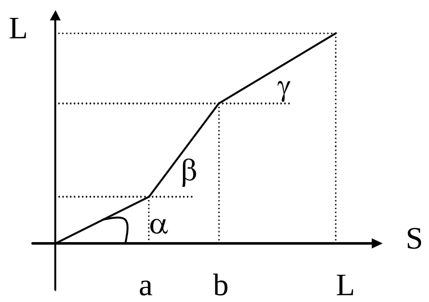
CHƯƠNG IV XỬ LÍ NÂNG CAO CHẤT LƯỢNG ẢNH

4.1 Các phương pháp tác động lên điểm ảnh

4.1.1 Tăng độ tương phản

Trước tiên cần làm rõ khái niệm độ tương phản. Ảnh số là tập hợp các điểm, mà mỗi điểm có giá trị sáng khác nhau. Ở đây, độ sáng để mắt người dễ cảm nhận ảnh song không phải là quyết định. Thực tế chỉ ra rằng hai đối tượng có cùng độ sáng nhưng đặt trên hai nền khác nhau sẽ cho cảm nhận khác nhau. Như vậy, độ tương phản biểu diễn sự thay đổi độ sáng của đối tượng so với nền. Một cách nôm na, độ tương phản là độ nổi của điểm ảnh hay vùng ảnh so với nền. Với khái niệm này, nếu ảnh của ta có độ tương phản kém, ta có thể thay đổi tuỳ theo ý muốn.

Ta có phương pháp dàn trải độ tương phản:



Các hàm tuyến tính được xác định như sau:

$$f = \begin{cases} \alpha S & 0 \le S \le a \\ \beta(S-a) + V_a & a < S \le b \\ \gamma(S-b) + V_b & b < S \le L \end{cases} \tag{4-1}$$

- $\alpha = \beta = \gamma$ ảnh kết quả trùng với ảnh gốc
- $\alpha, \beta, \gamma > 1$ giãn độ tương phản
- $\alpha, \beta, \gamma < 1$ co độ tương phản
- $\alpha, \gamma < 1$ và $\beta > 1$ dàn trải trong đoạn $a < S \le b$

Ví dụ:

$$S(m,n) = \begin{vmatrix} 10 & 20 & 20 & 30 \\ 20 & 22 & 30 & 26 \\ 23 & 24 & 27 & 26 \\ 120 & 160 & 170 & 130 \\ 180 & 190 & 100 & 200 \end{vmatrix}$$

a b

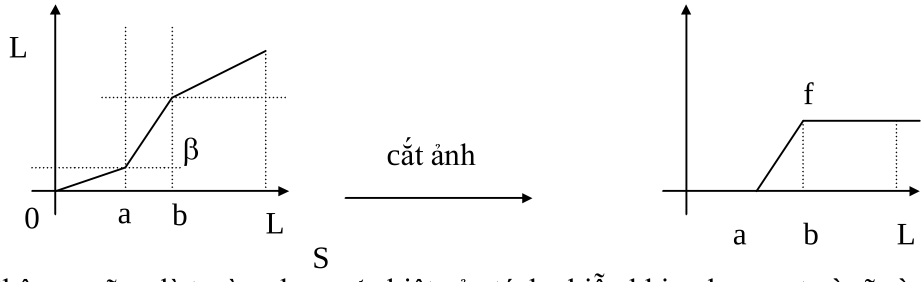
10 | 20 22 23 24 26 27 30 | 100 120 130 160 170 180 190 200

α=0.5 β=8 γ=0.9

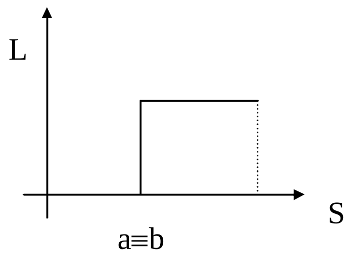
5 85 101 109 117 133 141 165 228…….

4.1.2 Tách nhiễu và phân ngưỡng

Tách nhiễu là trường hợp đặc biệt của giãn độ tương phản khi hệ số góc $\alpha=\beta=0$. Tách nhiễu được ứng dụng để giảm nhiễu khi biết tín hiệu vào nằm trên [a, b].



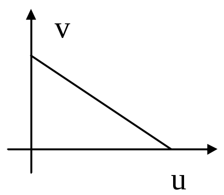
Phân ngưỡng là trường hợp ɑặc biệt của tách nhiễu khi a=b=const và rõ ràng trong trường hợp này, ảnh đầu ra là nhị phân(vì chỉ có 2 mức)



$$a \equiv b$$
$$S<a \Rightarrow S=0$$
$$S \geq a \Rightarrow S=L \qquad (4\text{-}2)$$

4.1.3 Biến đổi âm bản

Biến đổi âm bản nhận được khi dùng phép biến đổi f(u)=255-u . Biến đổi âm bản rất có ích khi hiện các ảnh y học và trong quá trình tạo ảnh âm bản.



4.2 Các phương pháp xử lý dựa trên lược đồ xám (histogram)

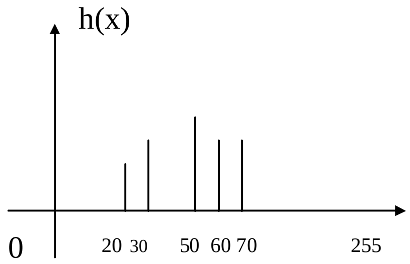
4.2 .1 Phương pháp cân bằng histogram

Histogram: giản đồ xác định tần suất xuất hiện của các giá trị mức xám trong ảnh.

$$S = \begin{vmatrix} 20 & 30 & 50 & 70 \\ 60 & 30 & 50 & 70 \\ 70 & 20 & 60 & 30 \\ 60 & 50 & 50 & 50 \end{vmatrix}$$

Biểu đồ

Mức xám	20	30	50	60	70
Tần suất	2	3	5	3	3



Hình 4.1 Biểu diễn histogram

Xác suất xuất hiện của một điểm ảnh tửrên tấm ảnh:

$$P(X_0) = \frac{h(x_0)}{\sum_{x_i=0}^{L} h(x_i)} \qquad (4\text{-}3)$$

- h(x) : histogram của ảnh S
- 0≤x ≤ L : giá trị các màu
- Ứng với mỗi giá trị x: $P(x) = \frac{h(x)}{\sum_{x_i=0}^{L} h(x_i)}$
- Cân bằng mức xám, bằng cách biến đổi S → V

$$V_S^* = \sum_{x_i=0}^{S} P(x_i) \quad ; \quad V = \left[\frac{V_S^* - V_S^* \min}{1 - V_S^* \min}(L-1) + 0.5 \right] \qquad (4\text{-}4)$$

Áp dụng công thức với ảnh trên ta có:

$p(20) = \frac{2}{16}$, $V^*(20) = \sum_{x_i=0}^{20} p(x_i) = \frac{2}{16}$ $\qquad$ $p(60) = \frac{3}{16}$, $V^*(60) = \frac{13}{16}$

$p(30) = \frac{3}{16}$, $V^*(30) = \sum_{x_i=0}^{30} p(x_i) = \frac{5}{16}$ $\qquad$ $p(70) = \frac{3}{16}$, $V^*(70) = 1$

$p(50) = \frac{5}{16}$, $V^*(50) = \frac{10}{16}$

Từ các giá trị V^*, ta có thể tính ra các giá trị V

Ví dụ 2

Cân bằng histogram của ảnh S

$$S = \begin{vmatrix} 10 & 20 & 30 & 40 & 50 \\ 20 & 40 & 70 & 30 & 30 \\ 40 & 60 & 50 & 50 & 70 \\ 70 & 70 & 60 & 60 & 30 \\ 20 & 10 & 10 & 20 & 30 \end{vmatrix}$$

Xác định tần số mức xám

mức xám	10	20	30	40	50	60	70
tần số	3	4	5	3	3	3	4

$$p(10) = \frac{3}{25} \quad , \quad p(50) = \frac{3}{25}$$

$$p(20) = \frac{4}{25} \quad , \quad p(60) = \frac{3}{25}$$

$$p(30) = \frac{5}{25} \quad , \quad p(70) = \frac{4}{25}$$

$$p(40) = \frac{3}{25}$$

$$V^*(10) = \sum_{x_i=0}^{10} p(x_i) = \frac{3}{25} \quad , V^*(50) = \sum_{x_i=0}^{50} p(x_i)$$

$$V^*(20) = \sum_{x_{i=0}}^{20} p(x_i) = \frac{7}{25} \quad , V^*(60) = \sum_{x_i=0}^{60} p(x_i)$$

$$V^*(30) = \sum_{x_{i=0}}^{30} p(x_i) = \frac{12}{25} \quad , V^*(70) = \sum_{x_i=0}^{70} p(x_i)$$

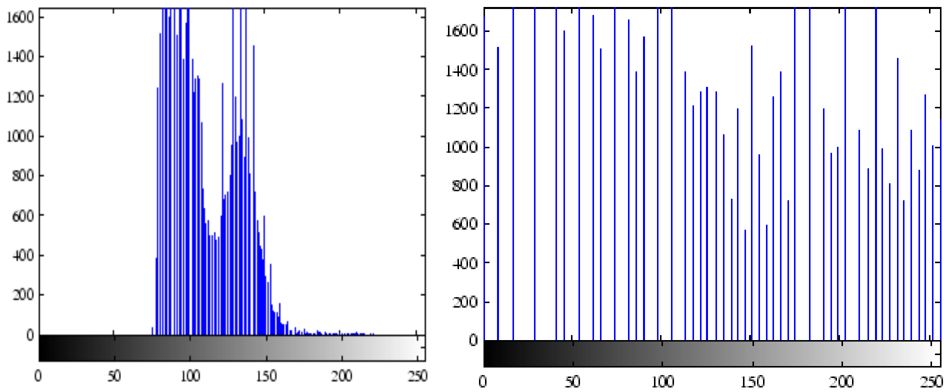
$$V^*(40) = \sum_{x_i=0}^{40} p(x_i) = \frac{15}{25}$$

$$V^*_{\min} = \frac{3}{25}; \quad 1 - V^*_{\min} = \frac{22}{25}$$

Kết quả:

V(10)=1, V(20)=47, V(30)=104, V(40)=139, V(50)=173, V(60)=208, V(70)=255

(a) Ảnh ban đầu (b). Ảnh cân bằng histigram



(c). Lược đồ histogram ảnh ban đầu(d). Lược đồ histogram sau cân bằng

Hình 4.2 Cân bằng histogram

4.2 .2 Kỹ thuật tách ngưỡng tự động

Thuật toán 1

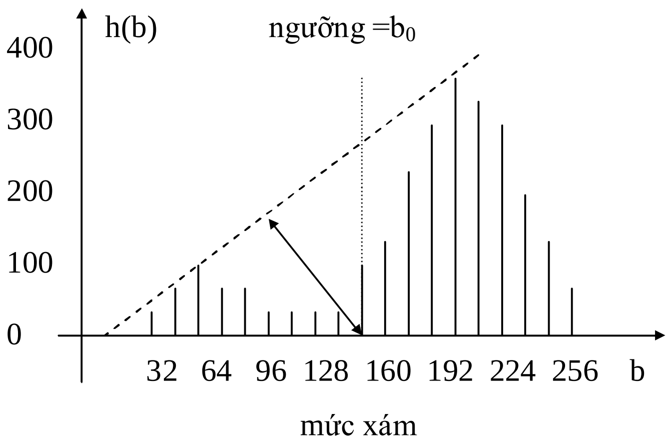
- Gọi t(g) là số điểm ảnh có giá trị ≤g
- G là số cấp sáng được xét
- P là số điểm ảnh được xét
- m(g) giá trị trung bình của các cấp xám ≤g

$$m(g) = \frac{\sum i * h(i)}{t(g)} \quad , i \le g \tag{4-5}$$

- $A(g) = \dfrac{t(g)}{(p - t(g))}$, $B(g) = [m(g) - m(G-1)]^2$ (4-6)
- $f(g) = A(g) * B(g) - 1$ (4-7)
- Tìm θ sao cho $f(\theta) = \max\{f(g)\}$ với g≤G-1 (4-8)

Thuật toán 2

Xác định ngưỡng tự động bằng thuật toán tam giác



Hình 4.3 Xác định ngưỡng

Chúng ta có thể quan sát thấy một đường thẳng đã được dựng bằng cách nối từ giá trị lớn nhất của lược đồ tại độ sáng b_{max} đến giá trị nhỏ nhất của lượt đồ tại độ sáng b_{min}. Với mỗi độ sáng b trong khoảng [b_{min}, b_{max}] chúng ta đi tính khoảng cách **d** từ giá trị lược đồ tại b là h[b] đến đường thẳng đã có. Giá trị b_0 ứng với khoảng cách **d** lớn nhất sẽ được chọn làm giá trị ngưỡng θ.

4.3 Lọc ảnh

4.3.1 Cuộn ảnh với mẫu

- **I**(x, y) là ảnh, **T(m,n)** là mẫu (cửa sổ): T*I được xác định như sau:
- $T * I(x,y) = \sum_{i=0}^{m-1}\sum_{j=0}^{n-1} T(i,j) I(x+i, y+j)$ (4-9)

Ví dụ $I = \begin{vmatrix} 4 & 7 & 2 & 7 & 1 \\ 5 & 7 & 1 & 7 & 1 \\ 6 & 6 & 1 & 8 & 3 \\ 5 & 7 & 5 & 7 & 1 \\ 5 & 7 & 6 & 1 & 2 \end{vmatrix}$ $T = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{vmatrix}$ $T * I = \begin{vmatrix} 23 & 26 & 31 & 19 & 16 \\ 35 & 39 & 46 & 31 & 27 \\ 36 & 43 & 49 & 34 & 27 \\ 36 & 43 & 48 & 34 & 12 \\ 24 & 35 & 33 & 22 & 11 \end{vmatrix}$

4.3.2 Lọc trung vị với ngưỡng θ

Cho I là ảnh, với mỗi điểm p ta lấy cửa sổ W(p), xét các điểm S ∈ W(p): $S_1 < S_2 < \ldots < S_k$

$S_1, S_2, \ldots S_k \in W(p)$.

Xác định trung vị của W(p)

- K chẵn : $TV(p) = \dfrac{S_{\frac{k}{2}} + S_{\frac{k}{2}+1}}{2}$ (4-10)
- K lẻ : $TV(p) = S_{\frac{k}{2}+1}$ (4-11)

Với θ là ngưỡng cho trước khi đó ta xác định giá trị I(x,y) mới như sau:

$$I(x,y) = \begin{cases} I(x,y) & if \ |I(x,y) - TV(p)| \le \theta \\ TV(p), otherwise \end{cases} \quad (4\text{-}12)$$

Ví dụ:

$$S(m,n) = \begin{vmatrix} 9 & 8 & 5 & 5 & 5 & 5 & 8 & 9 \\ 8 & 1 & 4 & 2 & 8 & 5 & 7 & 8 \\ 5 & 4 & 2 & 8 & 5 & 7 & 1 & 5 \\ 5 & 2 & 8 & 5 & 7 & 1 & 4 & 5 \\ 5 & 8 & 5 & 7 & 1 & 4 & 2 & 5 \\ 5 & 5 & 7 & 1 & 4 & 2 & 8 & 5 \\ 9 & 7 & 1 & 4 & 2 & 8 & 5 & 8 \\ 9 & 9 & 5 & 5 & 5 & 5 & 8 & 9 \end{vmatrix}$$

B1: Xác định ma trận trung vị

$$TV(p) = \begin{vmatrix} 5 & 4 & 5 & 5 & 5 & 7 \\ 4 & 4 & 5 & 5 & 5 & 5 \\ 5 & 5 & 5 & 5 & 4 & 4 \\ 5 & 5 & 5 & 4 & 4 & 4 \\ 5 & 5 & 4 & 4 & 4 & 5 \\ 7 & 5 & 4 & 4 & 5 & 8 \end{vmatrix}$$

B2: Xác định độ lệch

$$|\Delta| = \begin{vmatrix} 4 & 0 & 3 & 3 & 0 & 0 \\ 0 & 2 & 3 & 0 & 2 & 4 \\ 3 & 3 & 0 & 2 & 3 & 0 \\ 3 & 0 & 2 & 3 & 0 & 2 \\ 0 & 2 & 3 & 0 & 2 & 3 \\ 0 & 4 & 2 & 2 & 3 & 3 \end{vmatrix}$$

B3: So sánh với ngưỡng θ=3

$$I(x,y)=\begin{vmatrix}5&4&2&8&5&7\\4&2&8&5&7&5\\2&8&5&7&1&4\\8&5&7&1&4&2\\5&7&1&4&2&8\\7&5&4&2&8&5\end{vmatrix}$$

4.3.3 Lọc trung bình với ngưỡng θ

- Cho I(x,y) là ảnh số, với mỗi p trong ảnh ta lấy cửa sổ W(p). Sau đó lấy trung bình và gán cho TB(p) .
- Với θ là ngưỡng cho trước, khi đó ta xác định giá trị I(x,y) mới như sau:

$$I(x,y)=\begin{cases}I(x,y) & if\ |I(x,y)-TB(p)\le\theta| \\ TB(p), otherwise\end{cases} \tag{4-13}$$

Các bước thực hiện lọc trung bình giống như lọc trung vị:

B1: $\forall p \rightarrow TB(p)$

B2: Xác định ma trận độ lệch $|I-TB|$

B3: Lọc với ngưỡng θ ⇒ I(x,y) mới

Bài tập:

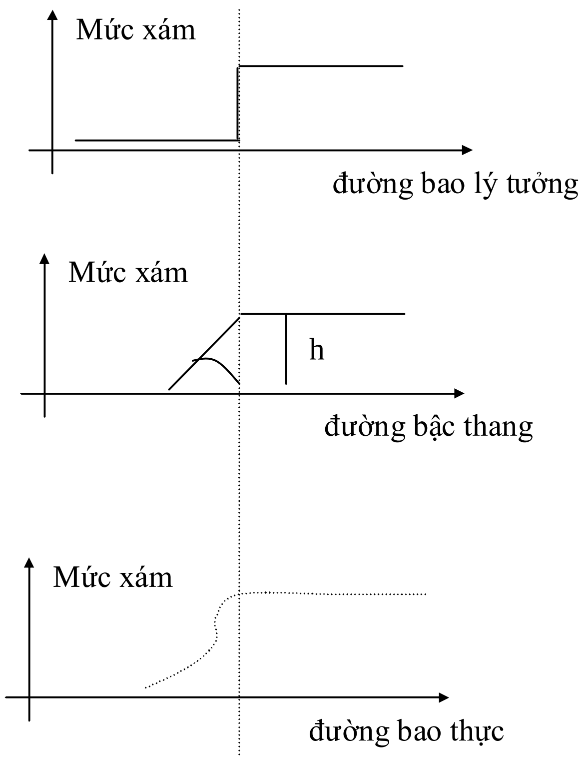
Xác định giá trị của I(x,y) sau khi áp dụng thuật toán lọc trung bình với ngưỡng θ=3.

$$I(x,y)=\begin{vmatrix}1&4&2&8&5&7\\4&2&8&5&7&1\\2&8&5&7&1&4\\8&5&7&1&4&2\\5&7&1&4&2&8\\7&1&4&2&8&5\end{vmatrix}$$

CHƯƠNG V PHÁT HIỆN BIÊN VÀ PHÂN VÙNG ẢNH

5.1 Biên và kỹ thuật phát hiện biên

Nhìn chung về mặt toán học người ta coi điểm biên của ảnh là điểm có sự biến đổi đột ngột về độ xám như chỉ ra trong hình dưới đây:



Hình 5.1 Đường bao của ảnh

Như vậy, phát hiện biên một cách lý tưởng là xác định được tất cả các đường bao trong các đối tượng. Định nghĩa toán học của biên ở trên là cơ sở cho các kỹ thuật phát hiện biên. Điều quan trọng là sự biến thiên giữa các điểm ảnh là nhỏ, trong khi đó biến thiên độ sáng của điểm biên (khi qua biên) lại khá lớn. Xuất phát từ cơ sở này người ta thường sử dụng 2 phương pháp phát hiện biên sau:

- **Phương pháp phát hiện biên trực tiếp:** phương pháp này nhằm làm nổi đường biên dựa vào biến thiên về giá trị độ sáng của điểm ảnh. Kỹ thuật chủ yếu là dùng kỹ thuật đạo hàm. Nếu lấy đạo hàm bậc nhất của ảnh ta có phương pháp gradient, nếu lấy đạo hàm bậc 2 ta có kỹ thuật Laplace.

- **Phương pháp gián tiếp:** nếu bằng cách nào đấy ta phân ảnh thành các vùng thì đường phân ranh giữa các vùng đó chính là biên.

5.1.1 Phương pháp gradient

Phương pháp gradient là phương pháp dò biên cục bộ dựa vào cực đại của đạo hàm. Theo định nghĩa gradient là một vectơ có các thành phần biểu thị tốc độ thay đổi giá trị của điểm ảnh theo hai hướng x và y. Các thành phần của gradient được tính bởi:

$$\frac{\partial f(x,y)}{\partial x} = f_x \approx \frac{f(x+dx,y)-f(x,y)}{dx} \qquad (5\text{-}1)$$

$$\frac{\partial f(x,y)}{\partial y} = f_y \approx \frac{f(x,y+dy)-f(x,y)}{dy} \qquad (5\text{-}2)$$

Với dx là khoảng cách các điểm theo hướng x(khoảng cách tính bằng số điểm) và tương tự với dy. Trên thực tế người ta hay dùng dx=dy=1

Xét trong hệ toạ độ cực ta có:

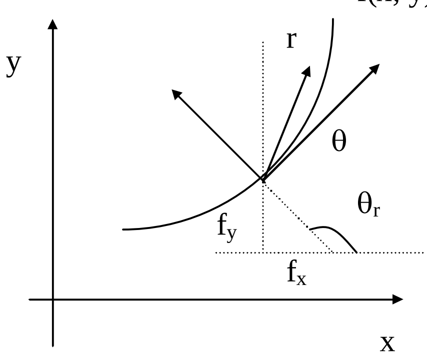
Tính đạo hàm theo hướng bất kì

$$\frac{df}{dr} = \frac{\partial f}{\partial x}\frac{dx}{dr} + \frac{\partial f}{\partial y}\frac{\partial y}{dr} = f_x\cos\theta + f_y\sin\theta \qquad (5\text{-}3)$$

$$\frac{df}{d\theta} = -rf_x\sin\theta + rf_y\cos\theta \qquad (5\text{-}4)$$

$\frac{df}{dr}$ đối với θ đạt cực đại khi

$$\frac{df}{d\theta} = -rf_x\sin\theta + rf_y\cos\theta = 0$$



Tinh góc θ_r:

$$\theta_r = \tan^{-1}\frac{f_y}{f_x} \text{ và } \frac{df}{dr}_{max} = \left(f_x^{\,2} + f_y^{\,2}\right)^{\frac{1}{2}} \qquad (5\text{-}5)$$

Ta gọi θ_r là hướng của biên

Trong kỹ thuật gradient, người ta chia nhỏ thành 2 kỹ thuật(do dùng 2 toán tử khác nhau) : kỹ thuật gradient và kỹ thuật la bàn. Kỹ thuật gradient dùng toán tử gradient lấy đạo hàm theo một hướng; còn kỹ thuật la bàn dùng toán tử la bàn lấy đạo hàm theo 8 hướng: Bắc, Nam, Đông, Tây và Đông Bắc, Tây Bắc, Đông Nam, Tây Nam.

5.1.1.1 Kỹ thuật gradient

Sử dụng một cặp mặt nạ H_1 và H_2 trực giao(theo 2 hướng vuông góc). Gọi g_1 và g_2 là gradient tương ứng theo 2 hướng x và y, thì biên độ của gradient, kí hiệu là g tại điểm(m,n) được tính theo công thức:

$$A_0 = g(m,n) = (g_1^2(m,n) + g_2^2(m,n))^{\frac{1}{2}} \text{ và } \theta_r(m,n) = \tan^{-1}\frac{g_2(m,n)}{g_1(m,n)} \quad (5\text{-}6)$$

Với mỗi điểm ảnh I(x, y) của I, đạo hàm theo x, theo y được kí hiệu tương ứng bởi g_x, g_y

Ta có:

$$\begin{cases} g_x = I(x+1,y) - I(x,y) \\ g_y = I(x,y+1) - I(x,y) \end{cases} \quad (5\text{-}7)$$

Điều này tương đương với nhân chập ảnh với 2 mặt nạ H_1 và H_2

$$H_1 = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} \quad H_2 = \begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}$$

Ta gọi H_1 và H_2 là mặt nạ Robert

Ngoài ra trong kỹ thuật Sobel và Prewitt người ta sử dụng 2 mặt nạ sau:

$$H_1 = \begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix} \quad H_2 = \begin{vmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{vmatrix} \quad \text{Mặt nạ Sobel}$$

Ngang (hướng x)   Dọc (hướng y)

Hay:

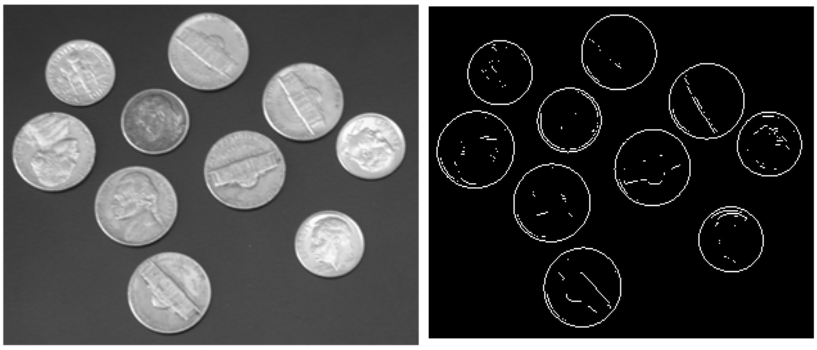
$$H_1 = \begin{vmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{vmatrix} \quad H_2 = \begin{vmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{vmatrix} \text{ Mặt nạ Prewitt}$$

Ngang (hướng x)   Dọc (hướng y)

Gradient được xấp xỉ bởi công thức: $G_x = H_x \otimes I$ và $G_y = H_y \otimes I$ (5-8)

Thường để giảm thời gian tính toán, người ta còn tính gradient theo các chuẩn sau:

$$A_1 = |g_1(m,n) + g_2(m,n)| \text{ hoặc } A_2 = Max(|g_1(x,y)|, |g_2(x,y)|) \quad (5\text{-}9)$$

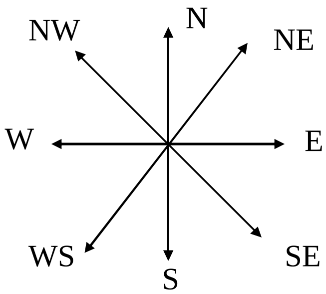


Hình 5.2 Tìm biên dùng mặt nạ sobel

5.1.1.2 Toán tử la bàn

Toán tử la bàn đo gradient theo một số hướng đã chọn. Nếu kí hiệu g_k là gradient la bàn theo hướng $\theta_k=\pi/2+2k\pi$ với k=0,1, 2,…7. Như vậy ta có gradient E theo 8 hướng ngược chiều kim đồng hồ.

Có nhiều toán tử la bàn khác nhau. Nhưng ở đây, trình bày một cách chi tiết toán tử Kish. Toán tử này sử dụng mặt nạ 3x3.



Mô hình 8 hướng

$$H_1=\begin{vmatrix}5 & 5 & 5\\ -3 & 0 & -3\\ -3 & -3 & -3\end{vmatrix}\quad H_2=\begin{vmatrix}-3 & 5 & 5\\ -3 & 0 & 5\\ -3 & -3 & -3\end{vmatrix}$$

$$H_3=\begin{vmatrix}-3 & -3 & 5\\ -3 & 0 & 5\\ -3 & -3 & 5\end{vmatrix}\quad H_4=\begin{vmatrix}-3 & -3 & 5\\ -3 & 0 & 5\\ -3 & -3 & 5\end{vmatrix}$$

$$H_5=\begin{vmatrix}-3 & -3 & -3\\ -3 & 0 & -3\\ 5 & 5 & 5\end{vmatrix}\quad H_6=\begin{vmatrix}-3 & -3 & -3\\ 5 & 0 & -3\\ 5 & 5 & -3\end{vmatrix}\quad H_7=\begin{vmatrix}5 & -3 & -3\\ 5 & 0 & -3\\ 5 & -3 & -3\end{vmatrix}\quad H_8=\begin{vmatrix}5 & 5 & -3\\ 5 & 0 & -3\\ -3 & -3 & -3\end{vmatrix}$$

Trong đó H_1, H_2, H_3, …H_8 tương ứng với 8 hướng: 0^0, 45^0, 90^0, 135^0, 180^0, 225^0, 315^0. Nếu ta kí hiệu A_i, i=1, 2, …8 là gradient thu được theo 8 hướng bởi 8 mặt nạ, biên độ gradient tại (x, y) được tính như sau:

$$A(x,y)=Max\left(\left|g_1(x,y)\right|, i=1,2,....8\right) \qquad (5\text{-}10)$$

Ngoài toán tử Kish, người ta còn sử dụng một số toán tử khác (có thể tham khảo trong tài liệu tham khảo) .

5.1.2 Kỹ thuật Laplace

Các phương pháp đánh giá gradient ở trên làm việc khá tốt khi độ sáng thay đổi rõ nét. Khi mức xám thay đổi chậm, miền chuyển tiếp trải rộng, phương pháp hiệu quả hơn đó là phương pháp sử dụng đạo hàm bậc2, gọi là phương pháp Laplace. Toán tử Laplace được định nghĩa như sau:

$$\nabla^2 f = \frac{\partial^2 f}{dx^2} + \frac{\partial^2 f}{dy^2} \tag{5-11}$$

Toán tử Laplace dùng nhiều kiểu mặt nạ khác nhau để xấp xỉ rời rạc đạo hàm bậc hai. Dưới đây là 3 kiểu mặt nạ hay dùng:

$$H_1 = \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix} \quad H_2 = \begin{vmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{vmatrix} \quad H_3 = \begin{vmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{vmatrix} \tag{5-12}$$

Với mặt nạ H_1, đôi khi người ta dùng phần tử ở tâm có giá trị là 8 thay vì giá trị là 4 như đã chỉ ra. Để dễ hình dung việc xấp xỉ đạo hàm bậc hai trong không gian rời rạc bởi mặt nạ H_1 hay là ý nghĩa của mặt nạ H_1, ta xét chi tiết cách tính đạo hàm bậc 2. Trong không gian rời rạc đạo hàm bậc 2 có thể tính:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= 2f(x,y) - f(x-1,y) - f(x+1,y) \\ \frac{\partial^2 f}{\partial y^2} &= 2f(x,y) - f(x,y-1) - f(x,y+1) \end{aligned} \tag{5-13}$$

Vậy $\nabla^2 f = -f(x-1,y) - f(x,y-1) + 4f(x,y) - f(x,y+1) - f(x+1,y)$ (5-14)

5.1.3 Phương pháp khớp nối lỏng

a. Khái niệm liên thông 4 và liên thông 8

P_3	P_2	P_1
P_4	P	P_0
P_5	P_6	P_7

Với điểm P được bao phủ xung quanh bởi 8 điểm: P_0, P_1, …P_8

Ta có láng giềng 8 của P gồm các điểm: P_0, P_1, P_2, P_3, P_4, P_5, P_6, P_7

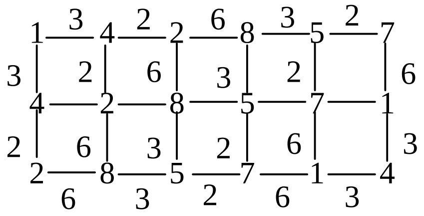
Láng giềng 4 của P gồm các điểm: P_0, P_2, P_4, P_6.

b. Phương pháp khớp nối lỏng

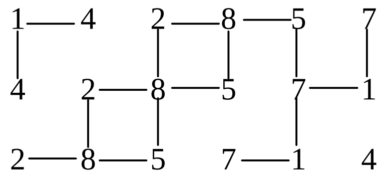
- Xét các điểm p và q là 2 điểm 4 láng giềng.
- I(p), I(q): giá trị mức xám của điểm p và q
- Nếu $|I(p) - I(q)| > \theta$ thì coi như có cặp biên (p, q).

Ví dụ:

Cho ma trận ảnh



chọn θ =3 ta có



Nhận xét:

Nếu θ giảm thì số khớp sẽ tăng và ngược lại.

5.1.4 Làm mảnh đường biên

Ý tưởng chính của quá trình làm mảnh biên là xoá nhiều lần các điểm biên để giảm bề rộng dòng còn một Pixel. Điều này phải được thực hiện không tách rời đối tượng (tách đối tượng thành 2 phần) cũng như không cần xoá các điểm cuối dòng. Nói cách khác chúng ta có thể nhận thấy làm mảnh là xoá các điểm biên có điều kiện. Ta có giải thuật như sau:

Sử dụng quy ước lân cận như hình vẽ, đặt $NP(P_1)$ là số các biến đổi 0(trắng) thành 1(đen) trong chuỗi có thứ tự<P_2, P_3,…P_9, P_2> và $ZN(P_1)$ là số các lân cận khác 0 của P_1. Điểm ảnh P_1 bị xoá(đặt bằng 0) nếu :

$$\begin{cases} 2 \le NZ(P_1) \le 6 \\ NT(P_1) = 1 \\ P_2 \cdot P_4 \cdot P_8 = 0 \text{ hay } NT(P_2) \ne 1 \\ P_2 \cdot P_4 \cdot P_6 = 0 \quad \text{hay } NT(P_4) \ne 1 \end{cases} \qquad (5\text{-}15)$$

P_3	P_2	P_9
P_4	P_1	P_8
P_5	P_6	P_7

Quy ước các lân cận làm mảnh biên

5.2 Kỹ thuật dò biên

5.2.1 Kỹ thuật Freeman(dò biên theo ảnh đen trắng)

Thuật toán

Bước1: Quét ảnh đến khi gặp điểm đen. Gọi nó là pixel 1.

Bước 2: Lặp

Nếu "điểm ảnh hiện thời là đen" rẽ trái

Ngược lại thì rẽ phải.

Dừng khi gặp điểm 1 ban đầu.

Cải tiến thuật toán trên

Thuật toán

Bước1: Quét ảnh đến khi gặp điểm đen. Gọi nó là pixel 1.

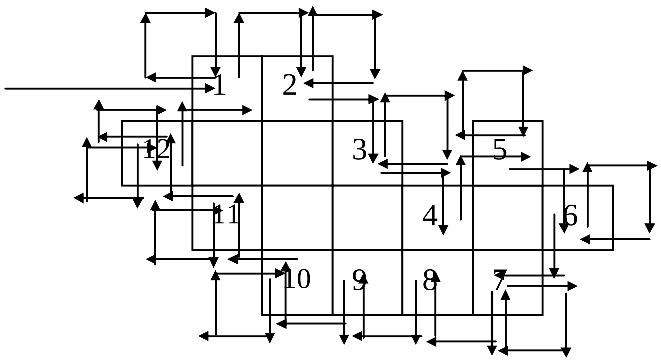
Bước 2: Lặp

Nếu "điểm ảnh hiện thời là đen"

Thì "dò ngược".

Ngược lại "sang phải".

Đến khi gặp pixel 1



Hình 5.3 Dò biên theo thuật toán cải tiến

5.2.2 Dò biên theo cặp nền vùng

Phương pháp

Tìm cặp điểm (n,v), trong đó n và v là điểm 8 láng giềng, n là điểm nền và v là điểm vùng.

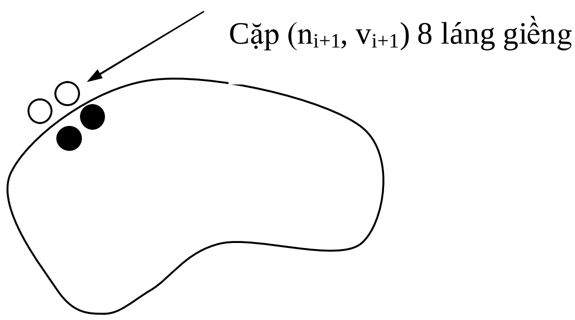
Ban đầu có (n_0, v_0) dựa vào đó ta tìm được (n_1, v_1), qua trình này cứ tiếp tục. Tổng quát nếu có (n_i, v_i) ta sẽ tìm (n_{i+1}, v_{i+1}), sao cho n_i và n_{i+1} là 8 láng giềng , v_i và v_{i+1} là 8 láng giềng.

Quá trình dò biên theo nền vùng là: tìm 1 dãy các điểm (n_0, v_0), (n_1, v_1)…(n_k, v_k) sao cho

n_0, n_1, ….n_k : chu tuyến nền

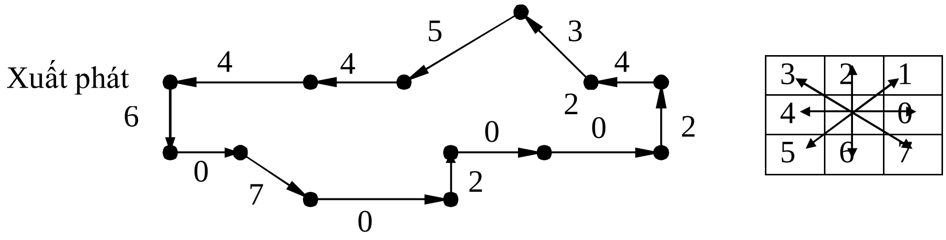
v_0, v_1, ….v_k : chu tuyến vùng

```
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 1 1 1 1 1 1 1 1 0 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 1 1 1 1 1 1 1 1 1 1 0 0
0 0 0 1 1 1 1 1 1 1 1 1 0 0 0
0 0 0 1 1 1 1 1 1 1 1 0 0 0 0
0 0 0 0 1 1 1 1 1 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```



5.2.3 Mã hoá biên (mã hoá Freeman)

Phương pháp mã hoá này cũng là một cách biểu diễn chính xác các điểm đường bao bằng việc sử dụng vị trí tương đối của điểm trên đường bao với điểm trước.



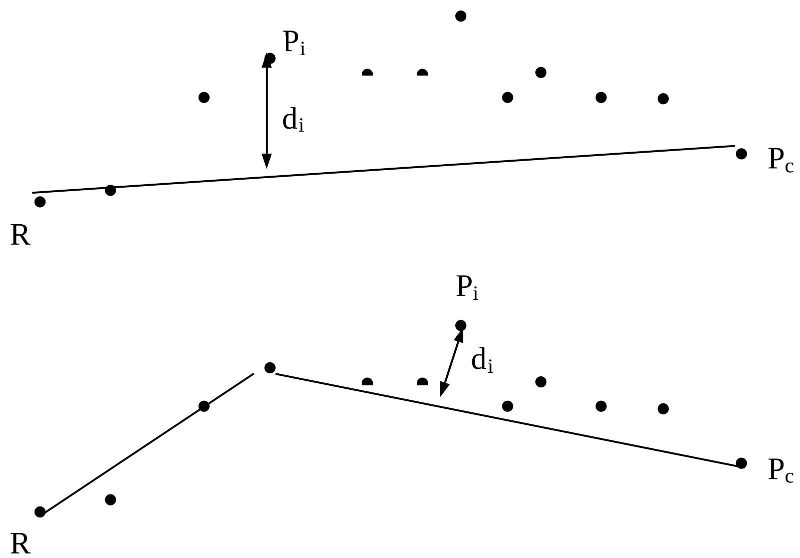
Hình 5.4 8 liên thông và mã hướng tương ứng

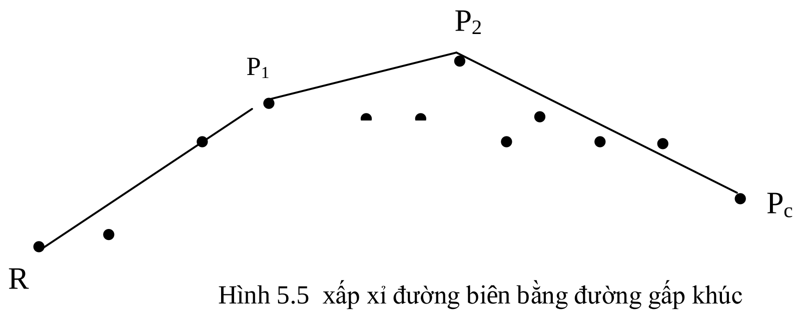
Xích mã hoá đường biên freeman như sau:

Cung=(6 0 7 0 2 0 0 2 4 3 5 4 4)

5.2.4 Xấp xỉ bởi đoạn thẳng

Nối điểm xuất phát R với điểm đang xét P_c bởi một đoạn thẳng. Sau đó tính toạ độ của P_i, một điểm nằm giữa R và P_c sao cho khoảng cách từ P_i đến đoạn thẳng là cực đại. Gọi khoảng cách này là d_1. Nếu d_1 lớn hơn một ngưỡng cho trước (độ chính xác của xấp xỉ) người ta phân đoạn RP_c thành 2 đoạn RP_i và P_iP_c và tiếp tục thực hiện lấy mẫu với từng đoạn cho tới khi đoạn thẳng tìm được là "rất gần" với đường bao.





Hình 5.5 xấp xỉ đường biên bằng đường gấp khúc

5.3 Các phương pháp phân vùng ảnh

Phân vùng là tiến trình chia ảnh thành nhiều vùng, mỗi vùng chứa một đối tượng hay nhóm đối tượng cùng kiểu. Chẳng hạn, một đối tượng có thể là một kí tự trên một trang văn bản hoặc một đoạn thẳng trong một bản vẽ kỹ thuật hoặc một nhóm các đối tượng có thể biểu diễn một từ hay hay đoạn thẳng tiếp xúc nhau. Ta có một số phương pháp phân vùng ảnh như sau:

5.3.1 Thuật toán gán nhãn thành phần liên thông

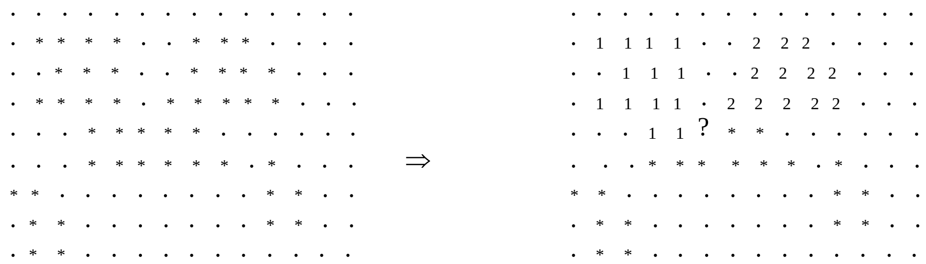
Kỹ thuật này gán cho mỗi thành phần liên thông của ảnh nhị phân một nhãn riêng biệt. Nhãn thường là các số tự nhiên bắt đầu từ một đến tổng số các thành phần liên thông có trong ảnh. Giải thuật quét ảnh từ trái sang phải và từ trên xuống dưới. Trong dòng thứ nhất của các pixel đen, một nhãn duy nhất được gán cho mỗi đường chạy liên tục của pixel đen. Với mỗi pixel đen của các dòng tiếp theo, các pixel lân cận trên dòng trước và pixel bên trái được xem xét. Nếu bất kì pixel lân cận nào được gán nhãn, nhãn tương tự được gán cho pixel đen hiện thời; ngược lại nhãn tiếp theo chưa được sử dụng được chọn. Thủ tục này được tiếp tục cho tới dòng cuối của ảnh.

Lúc kết thúc tiến trình này, một thành phần liên thông có thể chứa các pixel có các nhãn khác nhau vì khi chúng ta xem xét lân cận của pixel đen, chẳng hạn pixel "?" trong hình vẽ. Pixel đối với lân cận trái và những lân cận trong dòng trước có thể được gán nhãn một cách riêng biệt. Một tình huống như vậy phải được xác định và ghi lại. Sau tiến trình quét ảnh, việc gán nhãn được hoàn tất bằng cách thống nhất các mâu thuẫn các nhãn và gán lại các nhãn chưa sử dụng.

Để minh hoạ ta có hình biểu diễn sau :

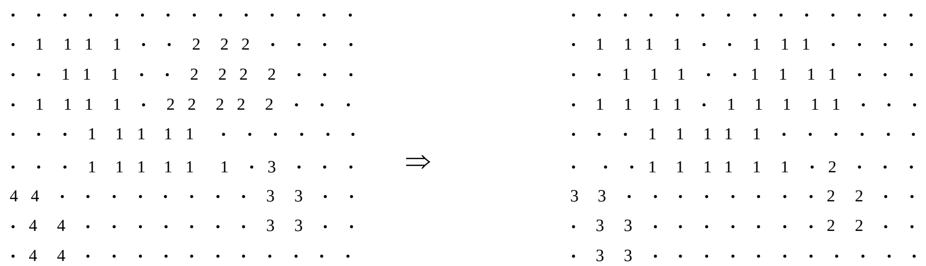
```
. . . . . . . . . ……..
. . . . P P P P . . . .
. . . . L ? . . . . . . .
. . . . . . . .. . . . . ..
```

P: lân cận trước, L lân cân trái



Hình b . Ảnh ban đầu

Hình c . Tiến trình gán nhãn



Hình d . Sau khi quét đầy đủ

Hình e .Kết quả sau cùng

Hình 5.6 Gán nhãn thành phần liên thông

5.3.2 Phân vùng bằng cây tứ phân

Về nguyên tắc, phương pháp này kiểm tra tính hợp thức của tiêu chuẩn một cách tổng thể trên miền lớn của ảnh. Nếu tiêu chuẩn được thỏa, việc phân đoạn coi như kết thúc. Trong trường hợp ngược lại, ta chia miền đang xét thành 4 miền nhỏ hơn. Với mỗi miền nhỏ, ta áp dụng một cách đệ quy phương pháp trên cho đến khi tất cả các miền đều thỏa.

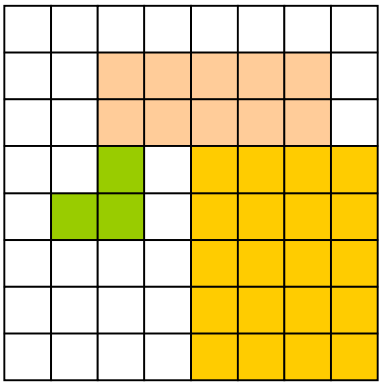
Thuật toán này tạo nên một cây mà mỗi nút cha có 4 nút con ở mọi mức trừ mức ngoài cùng. Vì thế cây này có tên là cây tứ phân. Cây này cho ta hình ảnh rõ nét về cấu trúc phân cấp của các vùng tương ứng với tiêu chuẩn.

Một vùng thỏa chuẩn sẽ tạo nên một *nút lá*, nếu không nó sẽ tạo nên một nút trong và có 4 nút con tương ứng với việc chia làm 4 vùng. Ta cứ tiếp tục như vậy cho đến khi phân xong. Các nút của cây biểu diễn số vùng đã phân.

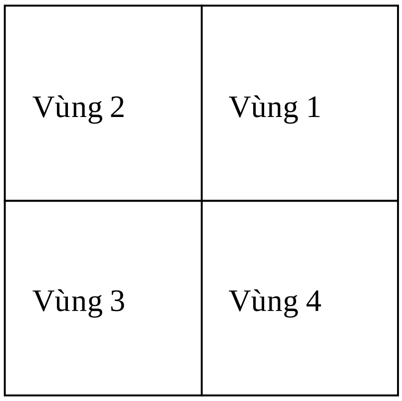
Tiêu chuẩn phân vùng ở đây là màu sắc. Nếu mọi điểm của vùng đều là màu trắng thì sẽ tạo nên nút lá trắng và tương tự như vậy với nút lá đen. Nút màu ghi vùng không thuần nhất và phải tiếp tục chia.

Với ngưỡng θ cho trước, vùng thuần nhất phải thỏa điều kiện

- Độ lệch chuẩn σ < θ
- Hoặc $|Max - Min| < \theta$ với Max, Min lần lượt là giá trị lớn nhất và nhỏ nhất của mức xám trong vùng cần chia.
- Giá trị điểm ảnh trong vùng bằng cách lấy trung bình giá trị của vùng đó



Ảnh gốc



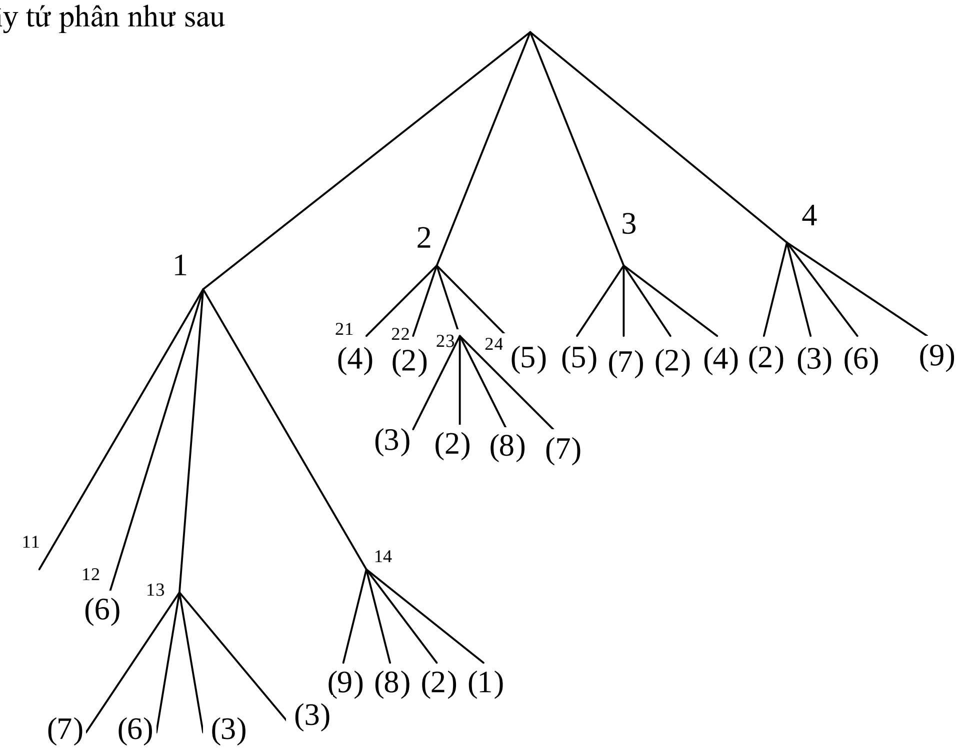
Phân đoạn ở mức 1

Ví dụ:

Cho ảnh S(m, n) , hãy phân vùng theo tiêu chí: ngưỡng θ= 2 và $|Max - Min| < \theta$

$$S(m,n) = \begin{vmatrix} 2 & 2 & 4 & 4 & 6 & 6 & 8 & 8 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 8 \\ 2 & 3 & 5 & 5 & 6 & 7 & 8 & 9 \\ 8 & 7 & 5 & 5 & 3 & 3 & 2 & 1 \\ 7 & 7 & 5 & 5 & 3 & 3 & 2 & 2 \\ 7 & 7 & 5 & 5 & 3 & 3 & 2 & 2 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 9 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 9 \end{vmatrix} \xrightarrow{\text{kết quả}} S(m,n) = \begin{vmatrix} 2 & 2 & 4 & 4 & 6 & 6 & 8 & 8 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 8 \\ 2 & 3 & 5 & 5 & 6 & 7 & 8 & 9 \\ 8 & 7 & 5 & 5 & 3 & 3 & 2 & 1 \\ 7 & 7 & 5 & 5 & 3 & 3 & 2 & 2 \\ 7 & 7 & 5 & 5 & 3 & 3 & 2 & 2 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 9 \\ 2 & 2 & 4 & 4 & 6 & 6 & 8 & 9 \end{vmatrix}$$

Ta có cây tứ phân như sau



Hình 5.7 Cây tứ phân kết quả phân vùng

Nhận xét về cây tứ phân

- Nếu quay ảnh một góc 90^0 thì không cần phải xây dựng lại cây tứ phân mà chỉ đổi vị trí nhánh.
- Ngưỡng càng nhỏ ảnh phân vùng càng chính xác nhưng cây tứ phân lại rất to.

5.3.3 Hợp vùng

Nhận xét trong phương pháp tách vùng bằng cây tứ phân ta thấy vẫn còn có những vùng mà độ lệch mức xám giữa các vùng này nhỏ hơn ngưỡng θ. Nghĩa là đối với các vùng này ta có thể hợp lại thành một vùng có diện tích liên thông lớn hơn.

Như vậy để khắc phục trường hợp này, sau khi phân vùng theo phương pháp cây tứ phân, ta tiến hành duyệt cây theo chiều ngược lại và hợp các vùng có cùng tiêu chuẩn. Với cách này ta thu được một miêu tả cấu trúc của ảnh với các miền liên thông có kích thước tối đa.

Phương pháp hợp vùng được thực hiện như sau:

- Giả sử có 2 vùng ω và ω'

- Ta xác định cặp các điểm 4 láng giềng (p, q) với p $\in \omega$ và q $\in \omega'$
- Xác định KHOP(ω,ω') : là các cặp điểm (p,q) với p $\in \omega$ và q $\in \omega'$, p và q là các cặp 4 láng giềng.
- Xác định $T(p,q) = \begin{cases} 1 & if\ |I(p) - I(q)| \le \theta_1 \\ 0 & otherwise \end{cases}$ (5-16).

Trong đó I(p), I(q) là giá trị mức xám của điểm p và q, θ1 là giá trị ngưỡng cho trước.

- Gọi b(ω) và b(ω') là số điểm biên của 2 vùng ω và ω'
- Xét hàm khả năng hợp 2 vùng : $KNG(\varpi,\varpi') = \dfrac{\sum T(p,q)}{Min(b(\varpi), b(\varpi'))}$ (5-17)
- Nếu $KNG(\varpi,\varpi') \ge \theta^2$ thì có thể hợp 2 vùng ω và ω' thành 1 vùng.

Ví dụ:

Xét khả năng hợp các vùng của ảnh sau

$$S(m,n) = \begin{vmatrix} 1 & 1 & 6 & 6 & 6 & 6 & 6 & 6 \\ 1 & 1 & 1 & 6 & 6 & 6 & 6 & 6 \\ 1 & 1 & 1 & 6 & 8 & 8 & 8 & 6 \\ 2 & 1 & 6 & 6 & 6 & 8 & 8 & 8 \\ 2 & 1 & 2 & 2 & 6 & 6 & 8 & 8 \\ 2 & 2 & 2 & 4 & 4 & 6 & 8 & 8 \\ 2 & 2 & 2 & 2 & 4 & 4 & 4 & 4 \end{vmatrix}$$

Gọi A, B, C, D lần lượt là các vùng chứa giá trị 1, 2, 4, 6, 8

Ta có bảng 1

Vùng	Biên	Với A	Với B	Với C	Với D	Với E
A	8	---	5	0	0	0
B	11	5	---	4	0	0
C	6	0	4	---	3	0
D	20	0	0	3	---	11
E	9	0	0	0	11	---

Xác định hợp vùng Bảng 2

Ghép vùng	A	B	C	D	E
A	---	5/8	0	0	0
B	5/8	---	4/6	0	0
C	0	4/6	---	3/6	0
D	0	0	3/6	---	11/9
E	0	0	0	11/9	---

Kết luận : Có thể hợp được 2 vùng D và E vì 11/9 >θ_2

5.3.4 Biến đổi Hough

Phép biến đổi Hough dùng để xác định một đường nào đó đi qua số điểm ảnh lớn hơn một ngưỡng θ cho trước. Dựa vào biến đổi Hough ta có ta có thể xác định được tập các điểm nằm trên đường thẳng, đường tròn, elip…

Đường thẳng Hough

Giả sử (x', y') là một điểm trong ảnh. Mọi điểm qua (x', y') phải thoả phương trình :

$$y' = mx' + c \tag{5-18}$$

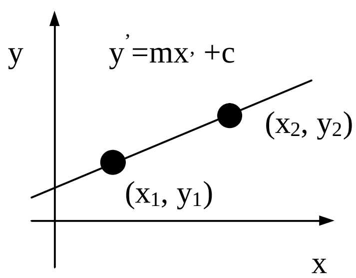
Với m, c là các tham số, viết lại phương trình ta có $c=-mx' + y'$.

Như vậy, mọi đường thẳng qua (x', y') tương ứng với một điểm trong không gian tham số (c, m). Xét 2 điểm (x_1, y_1) và (x_2, y_2) nằm trên cùng một đường thẳng:

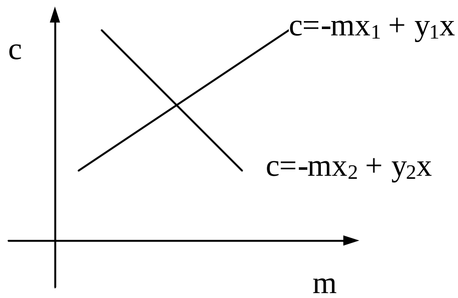
Với mỗi điểm ảnh, mọi đường thẳng qua m đó được biểu diễn bởi một điểm trong(c, m) .

$P_1(x_1, y_1)$: $c=-mx_1 + y_1$

$P_2 (x_2, y_2)$: $c=-mx_2 + y_2$



đường thẳng qua (x_1, y_1) và (x_2, y_2)



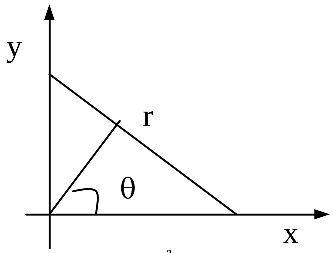
Giao điểm của 2 đường

Nhưng một đường thẳng duy nhất trong không gian (x, y) qua 2 điểm(x_1, y_1) và (x_2, y_2) được biểu diễn bởi giao của 2 đường trong không gian tham số (c, m). Điểm giao cho giá trị của c và m trong phương trình y=mx + c

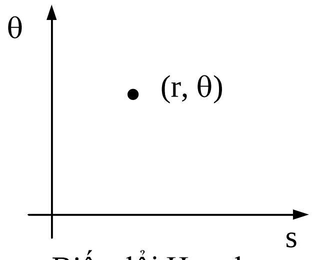
Để áp dụng kỹ thuật này, không gian tham số (c, m) cần phải được lượng hoá và như vậy ta cần dùng một ma trận tham số P(c, m) với:$c_1 \le c \le c_k$ và $m_1 \le m \le m_k$: K là số điểm chia của C và N là số điểm ảnh.

Tuy nhiên, cần xem phương pháp này có áp dụng trong thực tế được không? Điều đó có nghĩa là cần tính độ phức tạp tính toán của phương pháp. Dễ dàng thấy rằng độ phức tạp tính toán là O(KL) với K,L có nghĩa như trên. Độ phức tạp này là chấp nhận được. Song cũng dễ thấy, nếu biểu diễn bởi đường thẳng thì khi biểu diễn các đường đứng thì c có xu hướng tiến ra vô cùng. Một cách khắc phục là dùng hệ toạ độ cực (r, θ)

Trong biến đổi Hough, một đường thẳng trong một mặt phẳng với khoảng cách r và hướng θ có thể biểu diễn bởi: $r = x\cos\theta + y\sin\theta$



đường thẳng



Biến đổi Hough

$$\begin{aligned} r &= \frac{x}{\cos\theta} + (y - x\tan\theta)\sin\theta \\ &= \frac{x}{\cos\theta} + y\sin\theta - x\frac{\sin^2\theta}{\cos\theta} \\ &= x\left(\frac{1-\sin^2\theta}{\cos\theta}\right) + y\sin\theta = x\cos\theta + y\sin\theta \end{aligned}$$

Đường thẳng này có thể coi như một điểm trong mặt phẳng (x,θ). Kỹ thuật này sử dụng tất cả các điểm của đường bao và bản đồ gradient hướng. Biến đổi Hough tạo nên một ánh xạ các điểm của đường thẳng với một điểm trong mặt phẳng.

$$-\sqrt{N_1^2 + N_2^2} \le s \le \sqrt{M_1^2 + M_2^2} \text{ và } -\frac{\pi}{2} \le \theta \le \frac{\pi}{2} \qquad (5\text{-}19)$$

Ngoài đường thẳng Hough được trình bày trên, ta còn có đường tròn Hough và elip Hough có thể tham khảo trong các tài liệu tham khảo đã được giới thiệu trong giáo trình này.

CHƯƠNG VI NÉN ẢNH

Tại sao phải nén ảnh

Tính lượng ảnh để hiển thị theo Video trên máy tính trong 1 tiếng. Nếu 1 màn hình có độ phân giải 1024x768, số lượng màu RGB tương ứng với 3 bytes.Vậy dung lượng cần cho bộ nhớ cho 1 khung hình có số lượng là 1024x768 x3 bytes ≈ 2MBs. Có 24 hình/1s, do đó số lượng các bytes cần dùng là 48MBs/s. Vậy để xem phim trong 1h cần khối lượng nhớ là 180GBS.

Từ kết quả tính toán trên ta thấy cần một khối lượng nhớ rất lớn mới có thể lưu trữ và hiển thị phim ảnh, điều này càng không khả thi khi truyền dữ liệu phim ảnh qua mạng truyền thông. Do đó cần có phương pháp làm giảm kích thước lưu trữ thông tin ảnh. Điều này gọi là nén ảnh.

6.1 . Một số khái niệm

a. Nén dữ liệu (data compression)

Nén dữ liệu là quá trình làm giảm lượng thông tin "dư thừa"trong dữ liệu gốc và do vậy, lượng thông tin thu được sau nén thường nhỏ hơn dữ liệu gốc rất nhiều. Với dữ liệu ảnh, kết quả thường là 10:1. Một số phương pháp còn cho kết quả cao hơn. Theo kết quả nghiên cứu được công bố gần đây tại viện Kỹ thuật George, kỹ thuật nén fractal cho tỷ số nén là 30 trên 1.

b. Tỷ lệ nén (compression rate)

- Tỷ lệ nén=1/r.
- r= kích thước dữ liệu gốc/kích thước dữ liệu thu được sau nén, hiệu suất của nén=(1- Tỷ lệ nén)x %.

6.2 Các loại dư thừa dữ liệu

a. Sự phân bố kí tự

Trong một dãy kí tự, có một số kí tự có tần suất xuất hiện nhiều hơn một số dãy khác. Do vậy, ta có thể mã hoá dữ liệu một cách cô động hơn. Các dãy kí tự có tần suất cao được thay bởi một từ mã nhị phân với số bít nhỏ; ngược lại các dãy có tần suất thấp sẽ được mã hóa bởi từ mã có nhiều bít hơn. Đây chính là bản chất của phương pháp mã hoá Huffman.

b. Sự lặp lại của kí tự

Trong một số tình huống như trong ảnh, 1 ký hiệu(bít "0" hay bít "1") được lặp đi lặp lại một số lần. Kỹ thuật nén dùng trong trường hợp này là thay dãy lặp đó bởi dãy mới gồm 2 thành phần: số lần lặp và kí hiệu dùng để mã. Phương pháp mã hoá kiểu này được gọi là mã hoá loạt dài RLC(Run length Coding).

c. Những mẫu sử dụng tần suất

Có thể có dãy kí hiệu nào đó xuất hiện với tần xuất tương đối cao. Do vậy, có thể mã hoá bởi ít bit hơn. Đây là cơ sở của phương pháp mã hoá kiểu từ điển do Lempel-Ziv đưa ra và có cải tiến vào năm 1977, 1978 và do đó có tên là phương pháp nén LZ77, LZ78. Năm 1984, Terry Welch đã cải tiến hiệu quả hơn và đặt tên là LZW(Lempel-Ziv-Welch)

d. Độ dư thừa vị trí

Do sự phụ thuộc lẫn nhau của dữ liệu, đôi khi biết được kí hiệu(giá trị) xuất hiện tại một vị trí, đồng thời có thể đoán trước sự xuất hiện của các giá trị ở các vị trí khác nhau một cách phù hợp. Chẳng hạn, ảnh biểu diễn trong một lưới 2 chiều, một số điểm ở hàng dọc trong một khối dữ liệu lại xuất hiện trong cùng vị trí ở các hàng khác nhau. Do vậy, thay vì lưu trữ dữ liệu, ta chỉ cần lưu trữ vị trí hàng cột. Phương pháp nén dựa trên sự dư thừa này gọi là phương pháp mã hoá dự đoán.

Cách đánh giá độ dư thừa như trên hoàn toàn mang tính trực quan nhằm biểu thị cái gì đó xuất hiện nhiều lần. Đối với dữ liệu ảnh, ngoài đặc thù chung đó, nó còn có những đặc thù riêng. Thí dụ có ứng dụng không cần toàn bộ dữ liệu thô của ảnh mà chỉ cần các thông tin đặc trưng biểu diễn ảnh như biên ảnh hay vùng đồng nhất. Do vậy, có những phương pháp nén riêng cho ảnh dựa vào biến đổi ảnh hay dựa vào biểu diễn ảnh.

6.3 Phân loại các phương pháp nén

Có nhiều cách phân loại các phương pháp khác nhau. Cách thứ nhất dựa vào nguyên lý nén. Cách này phân phương pháp nén thành hai họ lớn:

- Nén chính xác hay nén không mất mác thông tin

Họ này bao gồm các phương pháp nén mà sau khi giải nén ta thu được chính xác dữ liệu gốc.

- Nén có mất mác thông tin

Họ này bao gồm các phương pháp mà sau khi giải nén ta không thu được dữ liệu như bản gốc. Trong nén ảnh, người ta gọi là các phương pháp "tâm lý thị giác". Các phương pháp này lợi dụng tính chất của mắt người, chấp nhận một số vặn xoắn trong ảnh khi khôi phục lại. Tất nhiên, khi phương pháp này chỉ có hiệu quả khi mà độ vặn xoắn là chấp nhận được bằng mắt thường hay với dung sai nào đó.

Cách phân loại thứ hai dựa vào cách thực hiện nén. Theo cách này người ta cũng phân thành 2 họ:

- Phương pháp không gian

Các phương pháp thuộc họ này thực hiện nén bằng cách tác động trực tiếp lên việc lấy mẫu của ảnh trong miền không gian.

- Phương pháp sử dụng biến đổi

Gồm các phương pháp tác động lên sự biến đổi của ảnh gốc mà không tác động trực tiếp như các họ trên.

Có một cách phân loại thứ 3, dựa vào triết lý của sự mã hoá. Cách này cũng phân các phương pháp nén thành hai họ

- Các phương pháp nén thế hệ thứ nhất: gồm các phương pháp mà mức độ tính toán là đơn giản, thí dụ như việc lấy mẫu, gán từ mã…
- Các phương pháp nén thế hệ thứ hai: dựa vào mức độ bảo hoà của tỷ lệ nén. Trong nội dung trình bày tiếp theo, ta sẽ đi theo cách phân loại này.

6.4 Các phương pháp nén thế hệ thứ nhất

Trong lớp các phương pháp này, ta lần lượt xét các phương pháp: mã hoá loạt dài, mã hoá Huffman, mã hoá LZW .

6.4.1 Phương pháp mã hoá loạt dài

Phương pháp mã hoá loạt dài lúc đầu được phát triển dành cho ảnh số 2 mức: mức đen(1) và mức trắng(0) như các văn bản trên nền trắng, trang in và các bức vẽ kỹ thuật.

Nguyên tắc của phương pháp là phát hiện một loạt các bít lặp lại, thí dụ như một loạt bít 0 nằm giữa hai bít 1, hay ngược lại, một loạt bít một nằm giữa hai bít 0. Phương pháp này chỉ có hiệu quả khi chiều dài của dãy lặp lớn hơn một ngưỡng nào đó. Dãy các bít lặp gọi là một mach(run) . Tiếp theo, thay thế chuỗi đó bởi một chuỗi mới gồm 2 thông tin chiều dài chuỗi và bit lặp(kí tự lặp). Như vậy, chuỗi thay thế sẽ có chiều dài ngắn hơn chuỗi cần thay.

Cần lưu ý rằng, đối với ảnh, chiều dài của chuỗi lặp có thể lớn hơn 255. Nếu ta dùng một byte để mã hoá thì sẽ không đủ. Giải pháp được dùng là tách chuỗi thành 2 chuỗi: một chuỗi có chiều dài 255, chuỗi kia là số bit còn lại.

Phương pháp RLC được sử dụng trong việc mã hoá lưu trữ các ảnh Bitmap theo dạng PCX, BMP.

Phương pháp RLC có thể chia thành 2 phương pháp nhỏ: phương pháp dùng chiều dài từ mã cố định và phương pháp thích nghi như kiểu Huffman. Giả sử các mạch gồm M bít. Để tiện trình bày đặt $M=2^m-1$. Như vậy mạch cũ được thay bởi mạch mới gồm m bít. Với cách thức này, mọi mạch đều được mã hoá bởi từ mã có cùng độ dài. Người ta cũng tính được, với M=15, p=0.9 ta sẽ có m=4 và tỷ số nén là 1.95

Với chiều dài cố định việc cài đặt thuật toán là đơn giản. Tuy nhiên tỷ lệ nén sẽ không tốt bằng dùng chiều dài biến đổi hay gọi là mã RLC thích nghi.

6.4.2 Phương pháp mã hoá Huffman

Nguyên tắc

Phương pháp mã hoá Huffman là phương pháp mã hoá dựa vào mô hình thống kê. Dựa vào dữ liệu gốc, người ta tính tần suất xuất hiện của các kí tự. Việc tính tần suất được thực hiện bằng cách duyệt tuần tự tệp gốc từ đầu đến cuối. Việc sử lý ở đây tính theo bit. Trong phương pháp này, người ta gán cho các kí tự có tần suất cao một từ mã ngắn, các kí tự có tần suất thấp từ mã dài. Nói một cách khác, các kí tự có tần suất càng cao được gãn mã càng ngắn và ngược lại. Rõ ràng với cách thức này, ta đã giảm chiều dài trung bình của từ mã hoá bằng cách dùng chiều dài biến đổi. Tuy nhiên, trong một số tình huống khi tần suất là rất thấp, ta có thể không được lợi một chút nào, thậm chí còn bị thiệt một ít bit.

Thuật toán:

Thuật toán gồm 2 bước chính:

- Giai đoạn tính tần suất của các kí tự trong dữ liệu gốc: duyệt tệp gốc một cách tuần tự từ đầu đến cuối để xây dựng bảng mã. Tiếp theo đó sắp xếp lại bảng mã theo thứ tự giảm dần tần suất.
- Giai đoạn thứ hai: mã hoá, duyệt bảng tần suất từ cuối lên đầu để thực hiện ghép hai phần tử có tần suất thấp thành một phần tử duy nhất. Phần tử này có tần suất bằng tổng hai tần suất thành phần.Tiến hành cập nhật lại bảng và đương nhiên loại bỏ 2 phần tử đã xét. Quá trình lặp lại cho đến khi bảng chỉ có một phần tử.Quá trình này gọi là quá trình tạo cây mã Huffman vì việc tập hợp được tiến hành nhờ vào một cây nhị phân với hai nhánh. Phần tử có tần suất thấp ở bên phải, phần tử kia ở bên trái.

Với cách tạo cây này, tất cả các bít dữ liệu, ký tự là nút lá; các nút trong là nút tổng hợp. Sau khi cây đã tạo xong, người ta tiến hành gán nhãn mã cho các nút lá. việc mã hoá rất đơn giản: mỗi lần xuống bên phải ta thêm 1 bit “1” vào từ mã; mỗi lần xuống bên trái ta thêm 1 bit “0”. Tất nhiên có thể làm ngược lại, chỉ có giá trị mã thay đổi còn tổng chiều dài là không đổi. Cũng chính do lý do này mà cây có tên gọi là Huffman.

- Quá trình giải nén được thực hiện ngược lại khá đơn giản. Người ta cũng phải dựa vào bảng mã tạo ra trong giai đoạn nén(bảng này được giữ lại trong cấu trúc đầu của tệp nén cùng với dữ liệu nén). Thí dụ với một tệp dữ liệu mà tần suất các ký tự cho bởi:

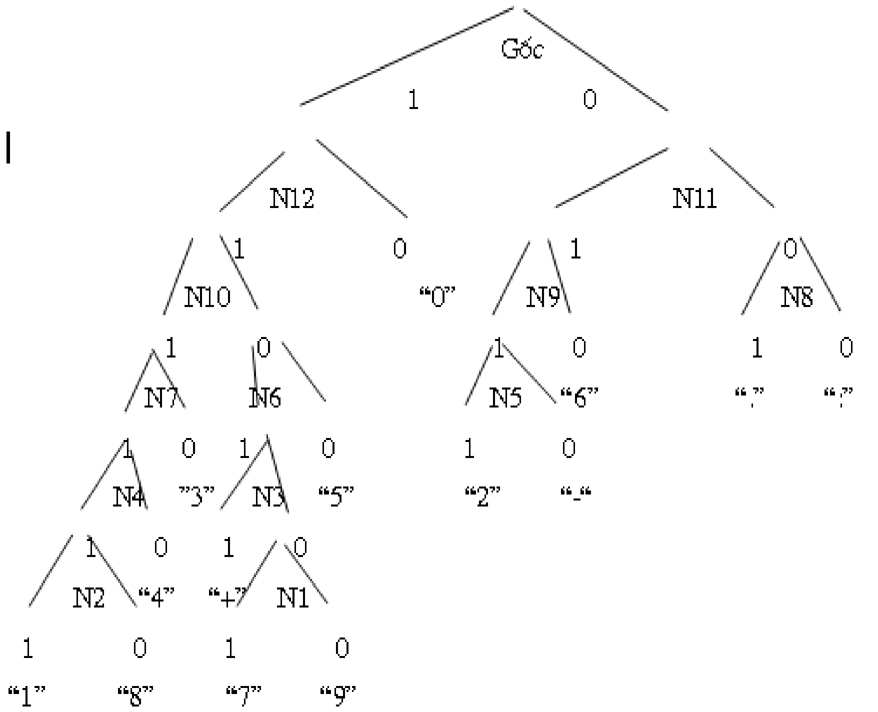
Kí tự	Tần suất	Kí tự	Tần suất	Xác suất
“1”	153	“0”	1532	0,2770
“2”	323	“0”	602	0,1088
“3”	412	“6”	536	0,0969
“4”	226	“.”	535	0,0967
“5”	385	“+”	112	0,0746
“6”	602	“3”	385	0,0696
“7”	92	“5”	323	0,0585
“8”	112	“2”	315	0,0569
“9”	87	“-“	226	0,0409
“0”	1532	“4”	220	0,0396
“.”	536	“1”	152	0,0275
“+”	220	“8”	112	0,0203
“-“	315	“7”	92	0,0167
“;”	535	“9”	87	0,158

Bảng tần suất **Bảng tần suất sắp theo thứ tự giảm dần**

Hình 6.1 Bảng tần suất mã hóa Huffman

Lưu ý rằng trong mã hoá Huffman, mã của ký tự là duy nhất và không mã nào là phần bắt đầu của mã khác. Vì vậy, khi đọc tệp nén từng bit từ đầu đến cuối ta có thể duyệt cây mã cho đến một nút lá, tức là kí tự đã được giải nén.

Cây mã Huffman tương ứng



Hình 6.2 Cây mã hóa Huffman

6.4.3 Phương pháp LZW

Khái niệm nén từ điển được Jacob Lempel và Abraham Ziv đưa ra lần đầu tiên vào năm 1977, sau đó phát triển thành một họ giải thuật LZ. Năm 1984, Terry Welch đã cải tiến giải thuật LZ thành một giải thuật mới hiệu quả hơn và đặt tên là LZW. Phương pháp nén từ điển dựa trên việc xây dựng từ điển lưu các chuỗi kí tự có tần suất lặp lại cao và thay thế bằng từ mã tương ứng mỗi khi gặp lại chúng. Giải thuật LZW hay hơn các giải thuật trước nó ở kĩ thuật tổ chức từ điển cho phép nâng cao tỉ lệ nén

Giải thuật nén LZW được sử dụng cho tất cả các file nhị phân. Nó thường được sử dụng 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. Mức độ hiệu quả của LZW không phụ thuộc vào số bit màu của ảnh

- Phương pháp:

Giải thuật nén LZW xây dựng một từ điển lưu trữ các mẫu có tần suất xuất hiện cao trong ảnh. Từ điển là tập hợp những cặp từ vựng và nghĩa của nó. Trong đó, từ vựng sẽ là các từ mã được sắp xếp theo thứ tự nhất định. *Nghĩa* là một chuỗi con trong dữ liệu ảnh. 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 suất 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 từ mã. Như vậy độ dài lớn nhất của từ mã là 12 bit($4096=2^{12}$) . Cấu trúc từ điển như sau

0
1

255
256
257
chuỗi
chuỗi

---
chuỗi

Hình 6.3 Bảng cấu trúc từ điển

- 256 từ mã đầu tiên theo thứ tự từ 0…255 chứa các số nguyên từ 0…255. Đây là mã của 256 kí tự cơ bản trong bảng mã ASCII
- Từ mã thứ 256 chứa một mã đặc biệt là "mã xoá" (CC- Clear Code) mục đích việc dùng mã xóa nhằm khắc phục tình trạng số mẫu lặp trong ảnh lớn hơn 4096. Khi đó một ảnh được quan niệm là nhiều mảnh ảnh, và từ điển là một bộ từ điển gồm nhiều từ điển con. Cứ hết một mảnh ảnh người ta lại gửi một mã xoá để báo hiệu kết thúc mảnh ảnh cũ, bắt đầu mảnh ảnh mới đồng thời khởi tạo từ điển cho mảnh ảnh mới. Mã hoá có giá trị là 256.

- Từ mã thứ 257 chứa mã kết thúc thông tin (EOI- End of Information). Mã này có giá trị là 257. Như chúng ta đã biết, một file ảnh GIF có thể chứa nhiều ảnh. Mỗi một ảnh sẽ được mã hoá riêng. Chương trình giải mã sẽ lặp đi lặp lại thao tác giải mã từng ảnh cho đến khi gặp mã kết thúc thông tin thì dừng lại.

- Các từ mã còn lại(từ 258 đến 4096) chứa các mẫu thường lặp lại trong ảnh. 512 phần tử đầu tiên của từ điển biểu diễn bằng 9 bít. Các từ mã từ 512 đến 1023 biểu diễn bởi 10 bít, 1024 đến 2047 biểu diễn bởi 11 bít và 2048 đến 4095 biểu diễn bởi 12 bit.
- Cho chuỗi đầu vào là "ABCBCABCABCD" (mã ASCII của A là 65, B là 66, C là 67)

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

Bảng 1. Bảng mã hóa kí tự

Chuỗi đầu ra sẽ là:

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

Đầu vào có kích thước 12x8 =96 bit. Đầu ra có kích thước:4x8+3x9=59 bít.

Tỉ lệ nén:96:59 =1.63

Ngoài các phương pháp nén được trình bày trên đây, còn có một số chuẩn nén MPEG, JPEG dựa trên cơ sở toán học của phép biến đổi cosin, có thể tham khảo trong các tài liệu tham khảo được giới thiệu.

CHƯƠNG VII NHẬN DẠNG ẢNH

Nhận dạng là giai đoạn cuối cùng của hệ thống xử lý ảnh. Nhận dạng ảnh dựa trên nền tảng lý thuyết nhận dạng. Trong lý thuyết nhận dạng nói chung và nhận dạng ảnh nói riêng có 3 cách tiếp cận khác nhau: nhận dạng dựa vào phân hoạch không gian, nhận dạng cấu trúc, nhận dạng dựa vào kỹ thuật mạng nơron.

Hai cách tiếp cận đầu là kỹ thuật kinh điển. Các đối tượng ảnh quan sát và thu nhận được phải trải qua giai đoạn tiền xử lý nhằm tăng cường chất lượng, làm nổi các chi tiết, tiếp theo là trích chọn và biểu diễn các đặc trưng, và cuối cùng mới qua giai đoạn nhận dạng. Cách tiếp cận thứ 3 hoàn toàn khác, nó dựa vào cơ chế đoán nhận, lưu trữ và phân biệt đối tượng mô phỏng theo hoạt động của hệ thần kinh con người. Do cơ chế đặc biệt, các đối tượng thu nhận bởi thị giác người không cần qua giai đoạn cải thiện mà chuyển ngay sang giai đoạn tổng, đối sánh các mẫu đã lưu trữ để nhận dạng. Đây là cách tiếp cận có nhiều hứa hẹn. Các cách tiếp cận trên sẽ được trình bày chi tiết ở các phần dưới đây.

7.1 Tổng quan về nhận dạng

Nhận dạng là quá trình phân loại các đối tượng được biểu diễn theo một mô hình nào đó và gán cho chúng vào một lớp(gán cho đối tượng một tên gọi) dựa vào những quy luật và các mẫu chuẩn. Quá trình nhận dạng dựa vào các mẫu học biết trước gọi là nhận dạng có thày hay học có thày; trong trường hợp ngược lại gọi là học không có thày.

7.1.1 Không gian biểu diễn đối tượng, không gian diễn dịch

Không gian biểu diễn đối tượng

Các đối tượng khi quan sát hay thu thập được, thường được biểu diễn bởi tập các đặc trưng hay đặc tính. Như trong trường hợp xử lý ảnh, ảnh sau khi được tăng cường để nâng cao chất lượng, phân vùng và trích chọn đặc tính như đã trình bày. Người ta thường phân các đặc trưng này theo các loại: đặc trưng tôpô, đặc trưng hình học, và đặc trưng chức năng. Việc biểu diễn ảnh theo đặc trưng nào phụ thuộc vào ứng dụng tiếp theo.

Ở đây ta đưa ra một cách nhìn hình thức về biểu diễn các đối tượng. Giả sử đối tượng X(ảnh, chữ viết, dấu vân tay, …) được biểu diễn bởi n thành phần(n đặc trưng): X=(x_1, x_2,…x_n) ; mỗi x_i biểu diễn một đặc tính. Không gian biểu diễn đối tượng thường được gọi tắc là không gian đối tượng χ được định nghĩa: $\chi = \left(X_1, X_2, ..., X_m\right)$, trong mỗi X_i biểu diễn một đối tượng. Không gian này có thể là vô hạn. Để tiện xem xét chúng ta chỉ xét tập χ hữu hạn.

Không gian diễn dịch

Không gian diễn dịch là tập các tên gọi của đối tượng. Kết quả quá trình nhận dạng ta xác định được tên gọi cho các đối tượng trong tập không gian đối tượng hay nói là đã nhận dạng được đối tượng. Một cách hình thức gọi Ω là tập tên đối tượng:

$\Omega = \{w_1, w_2, ...w_k\}$ với w_i , i=1, 2,…k là tên các đối tượng.

Quá trình nhận dạng một đối tượng f là một ánh xạ f: $\chi \rightarrow \Omega$ với f là tập các quy luật để định một phần tử trong χ ứng với một phần tử trong Ω. Nếu tập các quy luật và tập tên các đối tượng là biết trước người ta gọi là nhận dạng có thày. Trường hợp thứ hai là nhận dạng không có thày, trong trường hợp này nhận dạng sẽ khó khăn hơn.

7.1.2 Bản chất của quá trình nhận dạng

Quá trình nhận dạng gồm 3 giai đoạn:

- Lựa chọn mô hình biểu diễn đối tượng.
- Lựa chọn luật ra quyết định(phương pháp nhận dạng) và suy diễn quá trình học.
- Học nhận dạng

Khi mô hình biểu diễn đối tượng đã được xác định, có thể là định lượng(mô hình tham số) hay định tính(mô hình cấu trúc), quá trình nhận dạng chuyển sang giai đoạn học. Học là giai đoạn rất quan trọng. Thao tác học nhằm cải thiện, cải thiện việc phân hoạch tập đối tượng thành lớp.

Việc nhận dạng chính là tìm quy luật, thuật toán để có thể gán đối tượng vào một lớp hay nói cách khác cho đối tượng một tên.

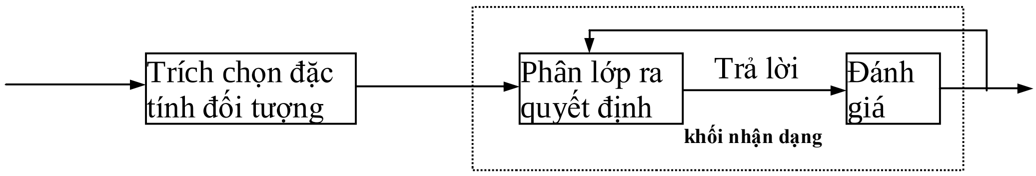
Học có thày

Kỹ thuât phân loại nhờ vào kiến thức biết trước gọi là học có thày. Đặc điểm cơ bản của phương pháp này là người ta có một thư viện các mẫu chuẩn. Mẫu cần nhận dạng sẽ được đem sánh cùng với các mẫu chuẩn để xem nó thuộc loại nào. Thí dụ như trong ảnh viễn thám, người ta muốn phân biệt một cánh đồng lúa, một cánh rừng hay một vùng đất hoang mà đã có các miêu tả về đối tượng đó. Vấn đề chủ yếu là thiết kế một hệ thống để có thể đối sánh trong ảnh mẫu với chuẩn và quyết định gán cho chúng vào một lớp. Việc đối sánh nhờ vào thủ tục ra quyết định dựa trên một công cụ gọi là *hàm phân lớp* hay *hàm ra quyết định.*

Học không có thày

Kỹ thuật này phải tự định ra các lớp khác nhau và xác định các tham số đặc trưng cho từng lớp. Học không có thày đương nhiên là khó khăn hơn. Một mặt, do số lớp không được biết trước, mặt khác những đặc trưng của các lớp cũng không được biết trước. Kỹ thuật này nhằm tiến hành một cách gộp nhóm có thể và chọn lựa cách tốt nhất. Bắt đầu từ tập dữ liệu, nhiều thủ tục xử lý khác nhau nhằm phân lớp và nâng cấp dần để đạt được một phương án phân loại

Sơ đồ tổng quát của hệ thống nhận dạng:



Hình 7.1 Sơ đồ tổng quát hệ thống nhận dạng

7.2 Một số thuật toán nhận dạng

7.2.1 Nhận dạng trên phân hoạch không gian

Trong kỹ thuật này đối tượng nhận dạng là đối tượng định lượng. Mỗi đối tượng được biểu diễn bởi một véctơ nhiều chiều. Trước tiên, ta xét một số khái niệm như: phân hoạch không gian, hàm phân biệt sau đó sẽ đi vào cụ thể.

a. phân hoạch không gian

Giả sử trong không gian đối tượng χ được định nghĩa: $\chi = \{X_i, i=1, 2,\ldots m\}$. X_i là một véctơ. Người ta nói P là một phân hoạch của không gian χ thành các lớp Ci, Ci $\subset \chi$ nếu:

$$C_i \cap C_j = \emptyset; i \neq j; \quad C_i \cup C_j = \chi \tag{7-1}$$

Nói chung đây là trường hợp lý tưởng: tập χ tách được hoàn toàn. Trong thực tế, thường gặp không gian biểu diễn tách được từng phần. Như vậy phân loại là dựa vào ánh xạ f: $\chi \rightarrow$P. Công cụ xây dựng ánh xạ là hàm phân biệt.

b. Hàm phân lớp hay hàm ra quyết định

Để phân đối tượng vào các lớp, ta phải xác định số lớp và ranh giới giữa các lớp đó. Hàm phân lớp hay phân biệt là một công cụ rất quan trọng. Gọi g_i là lớp các hàm phân lớp. Lớp này được định nghĩa như sau:

nếu $\forall$ i≠k, $g_k(X) > g_i(X)$ thì ta quyết định X thuộc lớp k.

Như vậy để phân biệt k lớp ta cần có k-1 hàm phân biệt. Hàm phân biệt g của một lớp nào đó thường dùng là hàm tuyến tính, có nghĩa là:

$$g(X)=W_o + W_1X_1 + W_2X_2+\ldots+W_kX_k \text{ trong đó} \tag{7-2}$$

W_1 là các trọng số gán cho các thành phần X_i.

W_0 là trọng số để viết cho gọn

Trong trường hợp g là tuyến tính, người ta nói việc phân lớp là tuyến tính hay siêu phẳng. Các hàm phân biệt thường được xây dựng dựa vào khoảng cách hay dựa vào xác suất có điều kiện.

Lẽ tự nhiên, khoảng cách là một công cụ rất tốt để xác định xem đối tượng có "gần nhau" hay không. nếu khoảng cách nhỏ hơn một ngưỡng τ nào đấy ta coi hai đối tượng là giống nhau và gộp chúng vào một lớp. Ngược lại nếu khoảng cách lớn hơn ngưỡng, có nghĩa là chúng khác nhau và ta tách thành hai lớp.

Trong một số trường hợp người ta dựa vào xác suất có điều kiện để phân lớp cho đối tượng. Lý thuyết xác suất có điều kiện được Bayes nghiên cứu khá kỹ và chúng ta có thể áp dụng lý thuyết này để phân biệt đối tượng.

Gọi $P(X/C_i)$ là xác suất để có X biết rằng có suất hiện lớp C_i

$P(C_i/X)$ là xác suất có điều kiện để X thuộc C_i.

Với X là đối tượng nhận dạng, C_i là lớp đối tượng.

Quá trình học cho phép ta xác định $P(X/C_i)$ và nhờ công thức Bayes về xác suất có điều kiện áp dụng trong điều kiện nhiều biến, chúng ta sẽ tính được P(C/X) theo công thức:

$$P(C_i, X) = \frac{P(X/C_i)P(C_i)}{\sum_{i=1}^{n} P(C/X_i)P(C_i)} = \frac{P(X/C_i)P(C_i)}{P(X)} \tag{7-3}$$

Nếu $P(C_i/X) > P(C_k/X)$ với $\forall$i, k thì $C_i \in$ X. Tuỳ theo các phương pháp nhận dạng khác nhau, hàm phân biệt sẽ có các dạng khác nhau.

7.2.2 Nhận dạng thống kê

Nếu các đối tượng nhận dạng tuân theo phân bố Guass, mà hàm mật độ xác suất cho bởi:

$$f(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-m)^2}{2\pi\sigma^2}} \quad (7\text{-}4)$$

Người ta dùng phương pháp quyết định dựa vào lý thuyết Bayes có tên là phương pháp thống kê.

Quy tắc Bayes:

Cho không gian đối tượng χ ={X_1, l=1,2,…L} với X_i={x_1, x_2,…x_p}

Cho không gian diễn dịch Ω={C_1, C_2,…C_r}, r là số lớp

Quy tắc bayes phát biểu như sau:

$\varepsilon : \chi \to \Omega$ sao cho X ∊ C_k nếu P(C_k/X) > P(C_l/X) ∀l<>k, l=1,2,…r

Trường hợp lý tưởng là nhận dạng luôn đúng, có nghĩa là không có sai số. Thực tế luôn tồn tại sai số ε trong quá trình nhận dạng. Vấn đề ở đây là xây dựng quy tắc nhận dạng với sai số ε là nhỏ nhất.

Phương pháp ra quyết định với ε tối thiểu

Ta xác định X∊ C_k nhờ xác suất P(C_k/X) . Vậy nếu có sai số, sai số sẽ được tính bởi 1-P(C_k/X). Để đánh giá sai số trung bình, người ta xây dựng một ma trận L(r, r) giả thiết có n lớp.

Ma trận L được định nghĩa như sau

$$L_{k,j} = \begin{cases} l_{k,j} > 0 \; neu \; k <> j \\ l_{k,j} \le 0 \, , neu \; k = j \end{cases} \quad (7\text{-}5)$$

Như vậy sai số trung bình của sự phân lớp sẽ là

$$r_k(X) = \sum_{j=1}^{r} l_{k,j} P(C_j / X) \quad (7\text{-}6)$$

Để sai số là nhỏ nhất ta cần có r_k là min trong công thức sau

$$r_k(X) = \sum_{j=1}^{r} l_{k,j} P(X / C_j) P(C_j) \quad (7\text{-}7)$$

Vậy quy tắc quyết định dựa trên Bayes có tính sai số quyết định được phát biểu như sau:

$X \in C_k$ nếu P_k<P_p với k<>p, p=1, 2, ,…r.

Với ρ_k là r_k(X), trường hợp đặc biệt với 2 lớp C_1 và C_2 ta dễ dạng có:

$$X \in C_1 : P(X / C_1) > \frac{l_{12} - l_{22}}{l_{11} - l_{21}} \frac{P(C_2)}{P(C_1)} P(X / C_2) \quad (7\text{-}8)$$

Giả sử thêm rằng xác suất phân bố đều ($P(C_1)=P(C_2)$ sai số là như nhau ta có:

$$X \in C_1 : P(X / C_1) > P(X / C_2)$$

7.2.3 Một số thuật toán nhận dạng tiêu biểu trong tự học

a. Thuật toán dựa vào khoảng cách lớn nhất

Nguyên tắc

Cho một tập gồm m đối tượng. Ta xác định khoảng cách giữa các đối tượng và khoảng cách lớn nhất ứng với phần tử xa nhất tạo nên lớp mới. Sự phân lớp được hình thành dần dần dựa vào việc xác định khoảng cách giữa các đối tượng và các lớp.

Thuật toán

Bước 1: Chọn hạt nhân ban đầu: giả sử $X_1 \in C_1$ gọi là lớp g_1. Gọi Z_1 là phần tử trung tâm của g_1.

Tính tất cả các khoảng cách $D_{j1}= D(X_j, Z_1)$ với j=1,2,….m.

Tìm $D_{k1}=\max_j D_{jl}$, X_k là phần tử xa nhất của nhóm g_1. Như vậy X_k là phần tử trung tâm của lớp mới g_2, kí hiệu X_2

Tính $d_1=D_{12}= D(Z_1, Z_2)$

Bước 2: tính khoảng cách D_{j1}, D_{j2}.

$D_{j1}= D(X_j,Z_1)$, $D_{j2}= D(X_j, Z_2)$. Đặt $D_k^{(2)} = \max_j D_j$

Nguyên tắc chọn:

Nếu $D_k^{(2)} < \theta d_1$ kết thúc thuật toán. Phân lớp xong.

Nếu không, sẽ tạo nên nhóm thứ 3. Gọi X_3 là phần tử trung tâm của g_3, kí hiệu là Z_3.

Tính $d_3=(D_{12} + D_{13} + D_{23})/3$ với θ là ngưỡng cho trước và $D_{13}= D(Z_1\ Z_3)$, $D_{23}=D(Z_2,Z_3)$.

Quá trình cứ lặp lại cho đến khi phân lớp xong. Kết quả ta thu được các lớp với các đại diện $Z_1, Z_2,\ldots Z_m$.

b Thuật toán k trung bình

Nguyên tắc

Khác với thuật toán trên, ta xét K phần tử đầu tiên trong không gian đối tượng. Hay nói một cách khác ta cố định K lớp. Hàm để đánh giá là hàm khoảng cách Euclide:

$$J_k = \sum_{X \in gk} D(X, Z_k) = \sum_{j=1}^{k} D_2(X_j, Z_k) \quad (7\text{-}9)$$

J_k là hàm chỉ tiêu với C_k. Việc phân vùng cho k hạt nhân đầu tiên được tiến hành theo nguyên tắc khoảng cách cực tiểu. Ở đây ta dùng phương pháp đạo hàm để tính cực tiểu.

Xét $\frac{\partial J_{jk}}{\partial Z_k} = 0$ với Z_k là biến. Ta dễ dàng có min khi:

$$\sum_{i=1}^{N} (X_i - Z_k) = 0 \Rightarrow Z_k = \frac{1}{N} \sum_{j=1}^{Nc} X_j \quad (7\text{-}10)$$

Công thức này là giá trị trung bình của lớp C_k và điều này lý giải cho tên của phương pháp.

Thuật toán

- Chọn N_c phần tử(giả thiết có N_c lớp) của tập T. Gọi các phần tử trung tâm của các lớp đó là: X_1, X_2,…X_{nc} và kí hiệu là Z_1, Z_2,…Z_{Nc}.
- Thực hiện phân lớp

  $X \in C_k$ nếu $D(X, Z_k) = \text{Min } D(X, Z_j)^{(1)}$, j=1, 2, …$N_c$ là lần lặp thứ nhất
- Tính tất cả các $Z_k$theo công thức đã trình bày trên.
- Tiếp tục như vậy cho đến bước q.

  $X \in G_k(q-1)$ nếu $D(X,Z_k^{(q-1)}) = \min_1 D(X, Z_1^{(q-1)})$

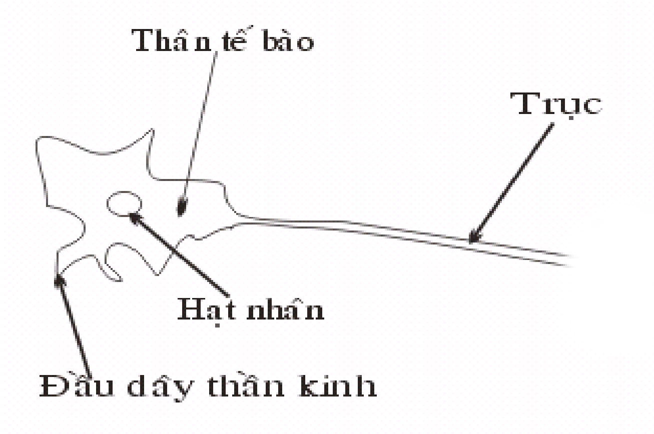
Nếu $Z_k^{(q-1)} = Z_k^{(q)}$ thuật toán kết thúc, nếu không ta tiếp tục thực hiện phân lớp

7.3 Mạng nơron nhân tạo và nhận dạng theo mạng nơron

Phần này giới thiệu một cách tổng quát về mạng nơron nhân tạo và phương pháp nhận dạng dựa trên mạng nơron nhân tạo. Để nắm thêm chi tiết, có thể tham khảo trong một số tài liệu tham khảo đã được giới thiệu.

7.3.1 Bộ não và nơron sinh học

Các nhà nghiên cứu sinh học về bộ não cho ta thấy rằng các nơron(tế bào thần kinh) là đơn vị cơ sở đảm nhiệm những chức năng xử lý nhất định trong hệ thần kinh, bao gồm não, tuỷ sống và các dây thần kinh. Mỗi nơron có phần than với nhân bên trong(gọi là soma), một đầu thần kinh ra(gọi là sợi trục ason) và một hệ thống dạng cây các dây thần kinh vào(gọi là dendrite). Các dây thần kinh vào tạo thành lưới dày đặt xung quanh thân tế bào, chiếm diện tích 0.25 mm^2, còn dây thần kinh ra tạo thành trục dài có thể từ 1cm đến hàng mét. Đường kính của thân tế bào thường là 10^{-4}m. Trục dây thần kinh ra cũng có thể phân nhánh theo dạng cây để nối với các dây thần kinh vào hoặc trực tiếp với nhân tế bào các nơron khác thông qua các khớp nối. Thông thường mỗi nơ ron có thể có vài chục đến vài trăm ngàn khớp nối với các nơron khác. Người ta ước lượng rằng lưới các dây thần kinh ra cùng với các khớp nối bao phủ diện diện tích 90% bề mặt nơron.



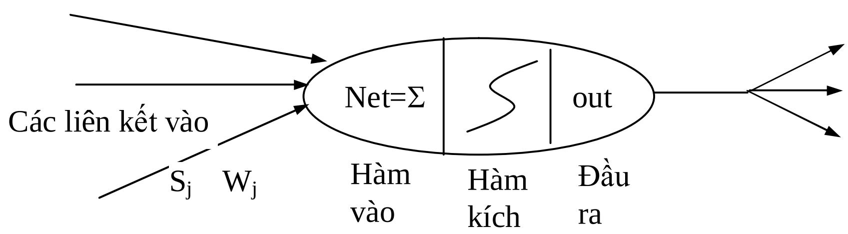
Hình 7.2 vẽ mạng nơron sinh học

7.3.2 Mô hình mạng nơron nhân tạo

Mạng nơron nhân tạo(Artifical Neural Network) gọi tắt là ANN bao gồm các nút (đơn vị xử lý, nơron) được nối với nhau bởi các liên kết nơ ron. Mỗi liên kết kèm theo một trọng số nào đó, đặc trưng cho tính kích hoạt /ức chế giữa các nơron. Có thể xem trọng số là phương tiện để lưu giữ thông tin dài hạn trong mạng nơron và nhiệm vụ của quá trình huấn luyện(học) mạng là cập nhật các trọng số khi có thêm các thông tin về các mẫu học, hay nói một cách khác, các trọng số được điều chỉnh sao cho dáng điệu vào ra của nó mô phỏng hoàn toàn phù hợp môi trường đang xem xét.

Trong mạng, một số nơron được nối với môi trường bên ngoài như các đầu ra, đầu vào.

a mô hình nơron nhân tạo:



Hình 7.3 Mô hình nơron nhân tạo

Mỗi nơron được nối với các nơron khác và nhận được các tín hiệu s_j từ chúng với các trọng số w_j. Tổng các thông tin vào có trọng số là:

$$Net = \Sigma\ w_j s_j \tag{7-11}$$

Người ta gọi đây là thành phần tuyến tính của nơron. Hàm kích hoạt g(còn gọi là hàm chuyển) đóng vai trò biến đổi từ Net sang tín hiệu đầu ra:

$$Out = g(Net) \tag{7-12}$$

Đây là thành phần phi tuyến của nơron. Có 3 dạng hàm kích hoạt thường được dùng trong thực tế.

Hàm dạng bước $step(x) = \begin{cases} 1 & ,x \geq 0 \\ 0 & x < 0 \end{cases}$ hoặc $step(x) = \begin{cases} 1 & ,x \geq \theta \\ 0 & x < \theta \end{cases}$ (7-13)

Hàm dấu $sign(x) = \begin{cases} 1, & x \geq 0 \\ -1 & x < 0 \end{cases}$ hoặc $sign(x) = \begin{cases} 1 & x \geq \theta \\ 0 & x < \theta \end{cases}$ (7-14)

Hàm sigmoid $Sigmoid(x) = \dfrac{1}{1 + e^{-\alpha(1+\theta)}}$ (7-15)

Ở đây ngưỡng θ đóng vai trò làm tăng tính thích nghi và khả năng tính toán của mạng nơron. Sử dụng kí pháp vectơ $S=(s_1, s_2, \ldots s_n)$ vectơ tín hiệu vào $W=(w_1, w_2,..w_n)$ vectơ trọng số ta có

$$Out=g(Net)\ ,\ Net=SW \tag{7-16}$$

Trường hợp xét theo ngưỡng θ ta dùng biểu diễn vectơ mới $S^{'}=(s_1, s_2, \ldots s_n, \theta)$ $W^{'}=(w_1, w_2,\ldots w_n,-1)$

b mạng nơron

Mạng nơron là hệ thống bao gồm nhiều phần tử xử lý đơn giản(nơron) hoạt động song song. Tính năng của hệ thống này phụ thuộc vào cấu trúc của hệ, các trọng số liên kết nơron và quá trình tính toán tại các nơron đơn lẻ. Mạng nơron có thể học từ dữ liệu mẫu và tổng quát hoá dựa trên dữ liệu mẫu học. Trong mạng nơron, các nơron đón nhận tín hiệu vào gọi là nơron vào và các nơron đưa thông tin ra gọi là nơron ra.

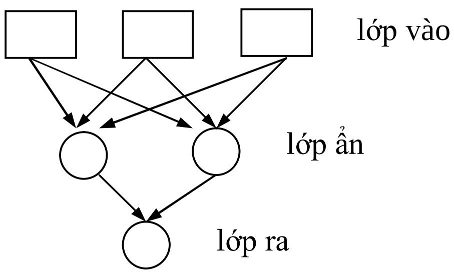
7.3.3 Phân loại các mạng

Theo kiểu liên kết nơron:

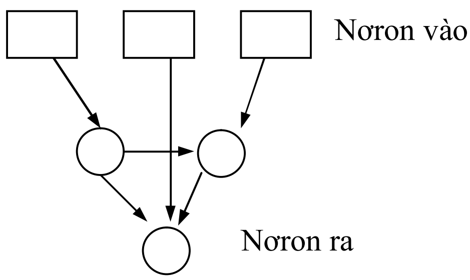
Ta có mạng nơron truyền thẳng, và mạng hồi quy. Trong mạng nơron truyền thẳng, các liên kết nơron đi theo một hướng nhất định, không tạo thành đồ thị có chu trình, với các đỉnh là các nơron, các cung là các liên kết giữa chúng. Ngược lại, các mạng quy hồi cho phép các liên kết nơron tạo thành chu trình. Vì các thông tin ra của các nơron được truyền lại cho các nơron đã góp phần kích hoạt chúng, nên mạng hồi quy còn có khả năng lưu giữ trạng thái trong của nó dưới dạng các ngưỡng kích hoạt ngoài các trọng số liên kết nơron.

Theo số lớp:

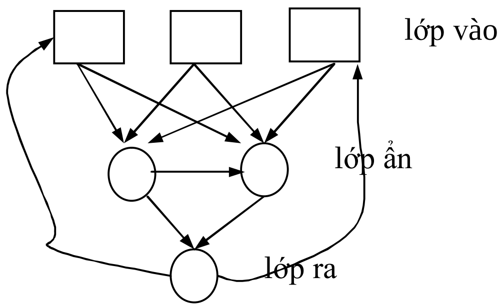
Các nơron có thể tổ chức lại thành các lớp sao cho mỗi nơron của lớp này chỉ được nối với các nơron ở lớp tiếp theo, không cho phép các liên kết giữa các nơron trong cùng một lớp, hoặc từ nơron lớp dưới lên nơron lớp trên. Ở đây cũng không cho phép các liên kết nơron nhảy qua một lớp.



Mạng nơron nhiều lớp



Mạng nơron truyền thẳng



Mạng nơron hồi quy

Dễ dạng nhận thấy các nơron trong cùng một lớp nhận được tín hiệu từ lớp trên cùng một lúc, do vậy về nguyên tắc chúng có thể xử lý song song. Thông thường, lớp nơron vào chỉ chịu trách nhiệm truyền đưa tín hiệu vào, không thực hiện một tính toán nào nên khi tính số lớp của mạng, người ta không tính lớp vào.

7.3.4 Hai cách nhìn về mạng nơron

- Mạng nơron như một công cụ tính toán:

Giả sử mạng nơron NN có m nơron vào và n nơron ra, khi đó với mỗi vectơ các tín hiệu vào $X=(x_1, x_2,\ldots x_m)$ sau quá trình tính toán tại các nơron ẩn, ta nhận được kết quả ra $Y=y_1, y_2,\ldots y_n$. Theo nghĩa nào đó, mạng nơron làm việc với tư cách là một bảng tra, mà không cần biết phụ thuộc hàm tường minh giữa Y và X. Khi đó ta viết:

$$Y=Tinh(X, NN)$$

Cần lưu ý thêm rằng các nơron trên cùng một lớp có thể tính toán đồng thời, do vậy độ phức tạp tính toán nói chung sẽ phụ thuộc vào số lớp của mạng.

Các thông số cấu trúc mạng nơron bao gồm:

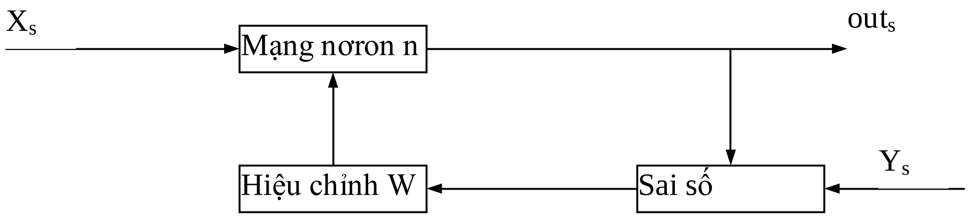
- Số tín hiệu vào, số tín hiệu ra.
- Số lớp nơron.
- Số nơron trên mỗi lớp ẩn.
- Số lượng liên kết của mỗi nơron(liên kết đầy đủ, liên kết bộ phận và liên kết ngẫu nhiên).
- Các trọng số liên kết nơron.

- Mạng nơron như một hệ thống thích nghi:

Có thể học(huấn luyện) để tinh chỉnh các trọng số liên kết cũng như cấu trúc của mình sao cho phù hợp với các mẫu học. Người ta phân biệt 3 loại kỹ thuật học: (i) học quan sát(hay còn gọi là học có thày), (ii) học không có giám sát hay còn gọi là học không thày và (iii) học tăng cường. Trong học giám sát, mạng được cung cấp một tập mẫu học $\{(Xs,Y_s)\}$ theo nghĩa X_s là các tín hiệu vào, thì kết quả ra của hệ phải là Y_s với kết quả tính toán out_s. Sai số này dùng để hiệu chỉnh các trọng số liên kết trong mạng. Quá trình cứ tiếp tục cho đến khi thoả mãn một tiêu chuẩn nào đó. Có 2 cách sử dụng tập mẫu học: hoặc dùng các mẫu lần lượt, hết mẫu này đến mẫu khác, hoặc sử dụng đồng thời tất cả các mẫu cùng lúc. Các mạng với các cơ chế học không giám sát được gọi là mạng tự tổ chức. Các kỹ thuật học trong mạng nơron có thể nhằm vào hiệu chỉnh các trọng số liên kết(gọi là học tham số) hoặc điều chỉnh, sửa đổi cấu trúc của mạng bao gồm số lớp, số nơron, kiểu và trọng số các liên kết(gọi là học cấu trúc). Cả hai mục đích học này có thể học đồng thời hoặc cách biệt.

Học tham số: Giả sử có k nơron trong mạng và mỗi nơron có đúng 1 liên kết vào với các nơron khác. Khi đó, ma trận trọng số liên kết W sẽ có kích thước kxl. Các thủ tục học tham số nhằm mục đích tìm kiếm ma trận W sao cho:

Y_s= Tính(X_s, W) đối với mọi mẫu học S=(X_s,Y_s)



Hình 7.4 Học tham số có giám sát

Học cấu trúc: Với học tham số ta giả định rằng mạng có một cấu trúc cố định. Việc học cấu trúc của mạng truyền thẳng gắn với yêu cầu tìm ra số lớp của mạng L và số nơron trên mỗi lớp n_i. Tuy nhiên, với mạng hồi quy còn phải xác định thêm các tham số ngưỡng θ của các nơron trong mạng. Một cách tổng quát phải xác định bộ tham số P=(L, n_1,…n_l, θ_1,…θ_k) .

Ở đây k=Σ n_j sao cho, Y_s = Tính(X_s, P) đối với mọi mẫu học s=(X_s, Y_s)

Về thực chất, việc điều chỉnh các vectơ tham số W đều quy về bài toán tìm kiếm tối ưu trong không gian tham số. Do vậy, có thể áp dụng các cơ chế tìm kiếm kinh điển theo gradient hay giải thuật di truyền trong lập trình tiến hoá.

7.3.5 Mạng tính toán và biểu diễn phụ thuộc dữ liệu của mạng nơron

Mạng nơron truyền thẳng chỉ đơn thuần tính toán các dữ liệu ra dựa trên các tín hiệu vào và các trọng số liên kết nơron đã xác định sẵn trong mạng. Do đó chúng không có trạng thái bên trong nào khác ngoài các vectơ trọng số W. Đối với mạng hồi quy, trạng thái trong của mạng được lưu giữ tại các ngưỡng của các nơron. Điều này có nghĩa là quá trình tính toán trên mạng truyền thẳng có lớp lang hơn mạng hồi quy. Nói chung các mạng hồi quy có thể không ổn định, thậm chí rối loạn theo nghĩa, khi cho vectơ giá trị đầu vào X nào đó, mạng cần phải tính toán rất lâu, thậm chí có thể bị lập vô hạn trước khi đưa ra được kết quả mong muốn. Quá trình học của mạng quy hồi cũng phức tạp hơn rất nhiều. Tuy vậy, các mạng quy hồi có thể cho phép mô phỏng các hệ thống tương đối phức tạp trong thực tế.

7.3.6 Xác định cấu trúc của mạng tối ưu

Như đã nói ở trên, lựa chọn sai cấu trúc của mạng có thể dẫn đến hoạt động của mạng trở nên kém hiệu quả. Nếu ta chọn mạng quá nhỏ có thể chúng không biểu diễn được sự phụ thuộc dữ liệu mong muốn. Nếu chọn mạng quá lớn để có thể nhớ được tất cả các mẫu học dưới dạng bảng tra, nhưng hoàn toàn không thể tổng quát hoá được cho tín hiệu vào chưa biết trước. Nói cách khác, cũng giống như trong các mô hình thống kê, các mạng nơron có thể đưa đến tình trạng quá thừa tham số.

Bài toán xác định cấu trúc mạng tốt có thể xem như bài toán tìm kiếm trong không gian tham số. Một cách làm là sử dụng giải thuật di truyền. Tuy vậy, không gian tham số có thể rất lớn và để xác định một trạng thái W(hoặc P) trong không gian đòi hỏi phải huấn luyện mạng, do vậy rất tốn thời gian. Có thể áp dụng tư tưởng tìm kiếm leo đồi nhằm sửa đổi một cách có lựa chọn, mang tính địa phương cấu trúc mạng hiện có. Có hai cách làm:

- Hoặc bắt đầu một mạng lớn, sau đó giảm nhỏ dần xuống
- Hoặc bắt đầu một mạng nhỏ, sau đó tăng dần lên.

Một kỹ thuật khác có thể áp dụng gọi là "tổn thương tối ưu" nhằm loại bỏ một số liên kết trọng số trong mạng dựa trên cách tiếp cận lý thuyết thông tin. Đơn giản nhất là cách liên kết có trọng số bằng 0. Quá trình cư tiếp tục như vậy, thực nghiệm chỉ ra rằng, kỹ thuật này có thể loại trừ tới ¾ các liên kết, do đó nâng cao đáng kể hiệu quả của mạng.

Ngoài việc loại trừ các liên kết nơron thừa, người ta còn vứt bỏ những nơron không đóng góp nhiều trong quá trình thực hiện của mạng.

Giải thuật "lợp ngói" là một biến thể của một kỹ thuật tăng trưởng mạng suất phát từ cấu hình ban đầu tương đối nhỏ. Ý tưởng ở đây là xác định một cấu hình mạng cho phép tính đúng các mẫu học đã biết. Sau đó, mỗi khi thêm dần mẫu học mới, mạng được phép thêm một số nơron cho phép đoán đúng kết quả học hiện tại và quá trình cư tiếp như vậy.

Tài liệu tham khảo

[1] “Nhận dạng các phương pháp và ứng dụng”, Hoàng Kiếm, Nguyễn Ngọc Kỷ và các tác giả, Nhà xuất bản thống kê 7/1992.

[2] ”Nhập môn xử lý ảnh số”, Lương Mạnh Bá, Nguyễn Thanh Thủy, nhà xuất bản khoa học kỹ thuật, năm 1995.

[3] “Giáo trình xử lý ảnh”, Võ Đức Khánh, Đại học quốc gia Thành phố Hồ Chí Minh, năm 2006

[4] “Xử lý ảnh số và Video số”, Nguyễn Kim Sách, nhà xuất bản khoa học kỹ thuật, năm 2000.