

Xử lý ảnh - Spatial Convolution

Đỗ Thanh Hà

Bộ môn Tin học
Khoa Toán - Cơ - Tin học
Trường Đại học Khoa học Tự nhiên

Lọc không gian (*Spatial Filtering*)

Gọi **I** và **J** là các ảnh sao cho $\mathbf{J} = T[\mathbf{I}]$. Trong đó $T[.]$ là một phép biến đổi sao cho:

$$\mathbf{J}(r, c) = T[\mathbf{I}](r, c) = f(\{\mathbf{I}(\rho, \chi) | \rho \in \{r-s, \dots, r, \dots, r+s\}, \chi \in \{c-d, \dots, c, \dots, c+d\}\})$$

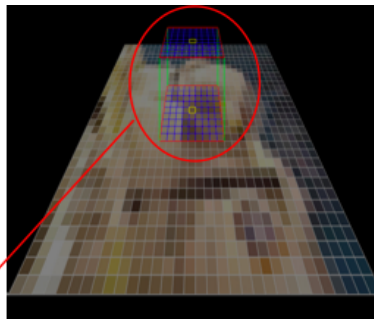
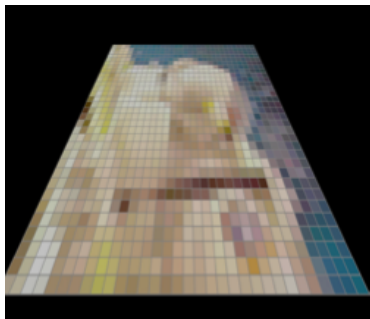
Hay giá trị của ảnh biến đổi **J** tại vị trí pixel (r, c) là một hàm các giá trị của các pixels của ảnh **I** mà nằm trong hình chữ nhật kích thước $2s + 1 \times 2d + 1$ và tâm của hình chữ nhật tại vị trí (r, c) .

Cửa sổ dịch chuyển (*Moving Windows*)

- Giá trị $J(r, c) = T[I](r, c)$ là một hàm của các pixel trong hình chữ nhật hàng xóm có tâm tại vị trí (r, c)
- Tại mỗi vị trí pixel thì có các hàng xóm khác nhau. Trong trường hợp kích thước của các hàng xóm là giống nhau tại mỗi vị trí, thì phép biến đổi T được gọi là *phép biến đổi cửa sổ dịch chuyển*

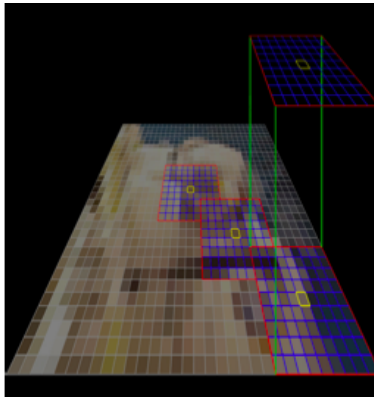
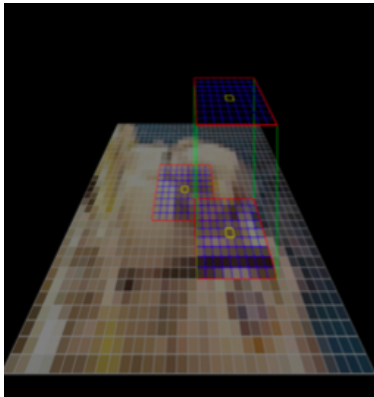
Cửa sổ dịch chuyển - minh hoạ

Hàng xóm được định nghĩa bởi một ma trận có trọng số



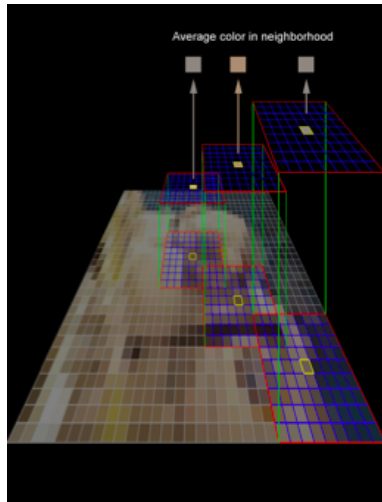
Cửa sổ dịch chuyển - minh hoạ

Hàng xóm tại các vị trí pixel khác



Biến đổi cửa sổ dịch chuyển tuyến tính

Đầu ra (*output*) của phép biến đổi tại mỗi pixel là giá trị trung bình (có trọng số) của các pixel trong hàng xóm



Tích chập (*convolution*): Biểu diễn toán học

Nếu phép biến đổi của số dịch chuyển là tuyến tính thì nó là một tích chập:

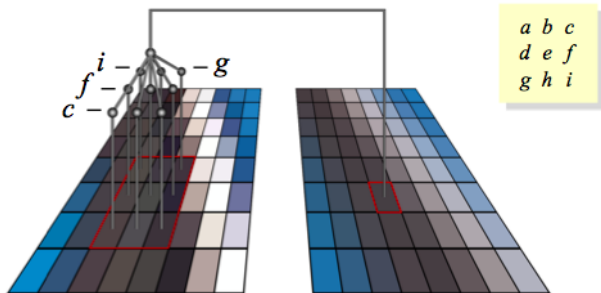
$$\mathbf{J}(r, c) = [\mathbf{I} * \mathbf{H}](r, c) = \sum_{\rho=-s}^s \sum_{\chi=-d}^d \mathbf{I}(r-\rho, c-\chi) \mathbf{H}(\rho, \chi)$$

với \mathbf{I} là ảnh số $\mathbf{I} : Z \times Z \rightarrow Z$

Mặt nạ Convolution (Ma trận trọng số)

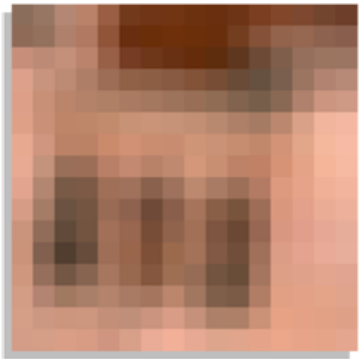
- **H** trong phương trình là một ma trận chữ nhật các số
- Ma trận là cửa sổ dịch chuyển
- Pixel (r, c) trong ảnh đầu ra là tổng có trọng số các pixel từ ảnh ban đầu trong hàng xóm của (r, c) theo vết bởi ma trận
- Mỗi pixel trong hàng xóm tâm tại (r, c) được nhân với giá trị ma trận tương ứng - sau khi ma trận được quay một góc 180 độ
- Tổng các tích này là giá trị của pixel (r, c) của ảnh đầu ra

Convolution bởi cửa sổ dịch chuyển



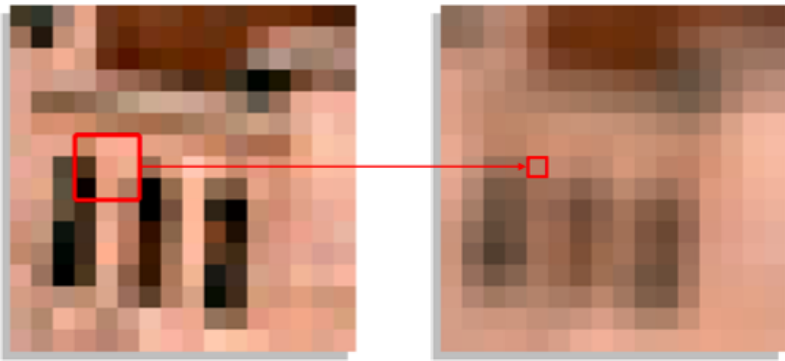
Biến đổi cửa sổ dịch chuyển - Ví dụ

Ảnh ban đầu và ảnh đầu ra sử dụng cửa sổ $\mathbf{H} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

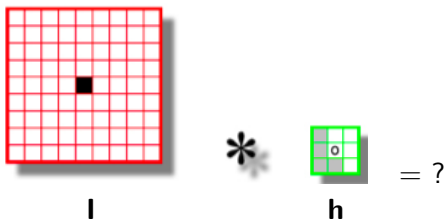


Biến đổi cửa sổ dịch chuyển - Ví dụ

Ảnh ban đầu và ảnh đầu ra sử dụng cửa sổ $\mathbf{H} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$



Tích chập bằng cách quay và dịch chuyển ma trận trọng số



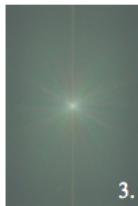
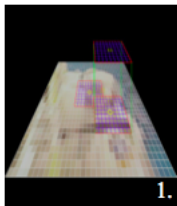
Ma trận trọng số đối xứng

| | | | | |
|----------|----------|----------|----------|----------|
| <i>f</i> | <i>e</i> | <i>d</i> | <i>e</i> | <i>f</i> |
| <i>e</i> | <i>c</i> | <i>b</i> | <i>c</i> | <i>e</i> |
| <i>d</i> | <i>b</i> | <i>a</i> | <i>b</i> | <i>d</i> |
| <i>e</i> | <i>c</i> | <i>b</i> | <i>c</i> | <i>e</i> |
| <i>f</i> | <i>e</i> | <i>d</i> | <i>e</i> | <i>f</i> |

Nếu **H** là ma trận trọng số đối xứng thì không cần thiết phải quay 180 độ

Có 3 cách để tính tích chập

- Sử dụng biến đổi cửa sổ dịch chuyển (vừa giới thiệu)
- Sử dụng phương pháp Cộng-nhân-dịch chuyển
- Sử dụng biến đổi Fourier

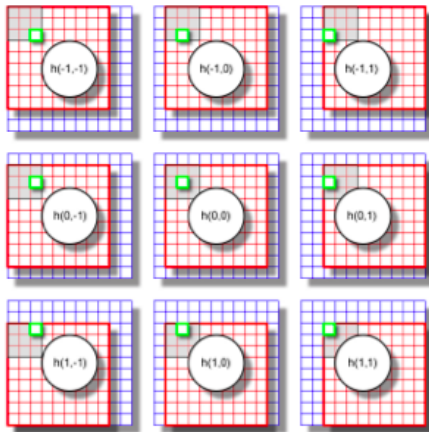


Phương pháp Cộng - Nhân - Dịch chuyển² (*Shift-Multiply-Add*)

- Với mỗi thành phần thứ i trong mặt nạ \mathbf{H} , ảnh \mathbf{I} được sao chép một lần sang ảnh \mathbf{P} gồm toàn số 0, thu được ảnh \mathbf{P}_i
- Vị trí bắt đầu sao chép phụ thuộc vào vị trí tương ứng của thành phần trong mặt nạ
- Mỗi ảnh \mathbf{P}_i được nhân với giá trị của thành phần mặt nạ liên kết với nó
- Cộng các ảnh \mathbf{P}_i lại thu được ảnh \mathbf{P}^*
- Điều chỉnh (cắt) ảnh \mathbf{P}^* để thu được ảnh \mathbf{J} có kích thước bằng kích thước của ảnh ban đầu \mathbf{I}

Minh họa phương pháp Cộng - Nhân - Dịch chuyển

Ứng với mỗi thành phần $\mathbf{H}(r_h, c_h)$ trong ma trận trọng số \mathbf{H} , ảnh \mathbf{I} sẽ được sao chép sang ảnh toàn số 0, \mathbf{P} , bắt đầu tại vị trí (r_h, c_h) . Mỗi \mathbf{P} được nhân với trọng số tương ứng $\mathbf{H}(r_h, c_h)$. Sau đó, tất cả các ảnh \mathbf{P} được cộng lại và chia cho số thành phần trong \mathbf{H} . Ảnh thu được sẽ được cắt sao cho có kích thước bằng kích thước ảnh \mathbf{I} ban đầu



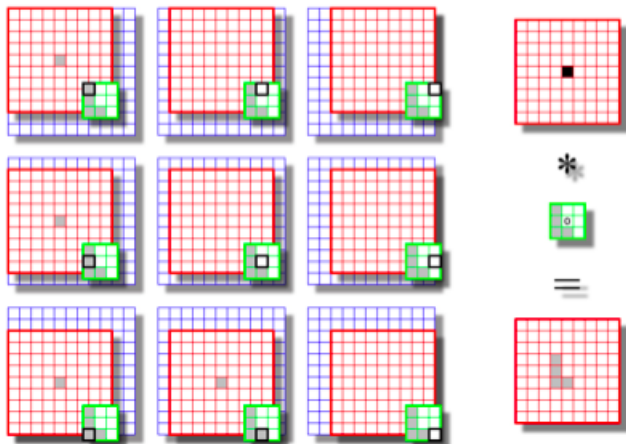
I: original image
P: padded image

| | | |
|------------|-----------|-----------|
| $h(-1,-1)$ | $h(-1,0)$ | $h(-1,1)$ |
| $h(0,-1)$ | $h(0,0)$ | $h(0,1)$ |
| $h(1,-1)$ | $h(1,0)$ | $h(1,1)$ |

weight matrix

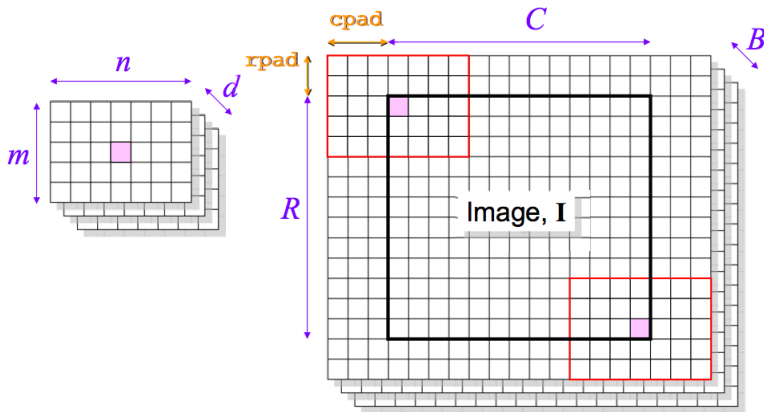
Dóng hàng pixel để cộng

Phương pháp Cộng - Nhân - Dịch chuyển: Ví dụ



Kích thước ảnh **P**

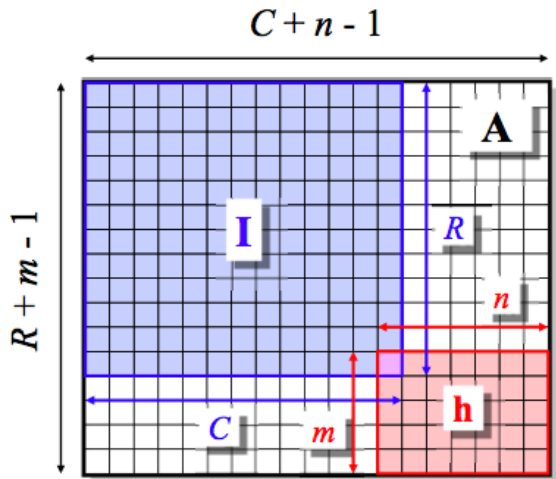
Ma trận trọng số **H** có kích thước là $m \times n$



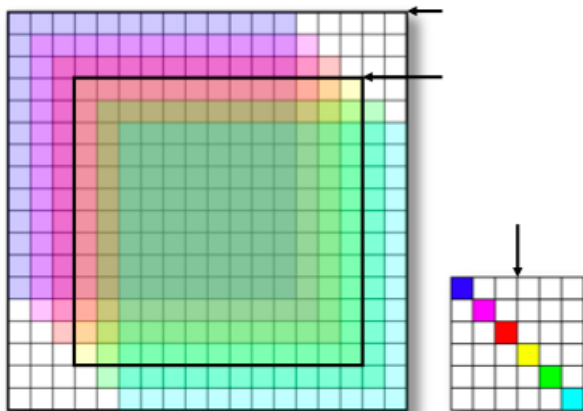
$$cpad = \text{floor}(n/2); rpad = \text{floor}(m/2);$$

Kích thước ảnh P

Để sử dụng
Phương pháp
Cộng - Nhân
- Dịch
chuyển, tạo
ma trận tích
luỹ **A** có kích
thước là
 $R + m - 1 \times$
 $C + n - 1$
(ảnh **I** có
kích thước là
 $R \times C$)

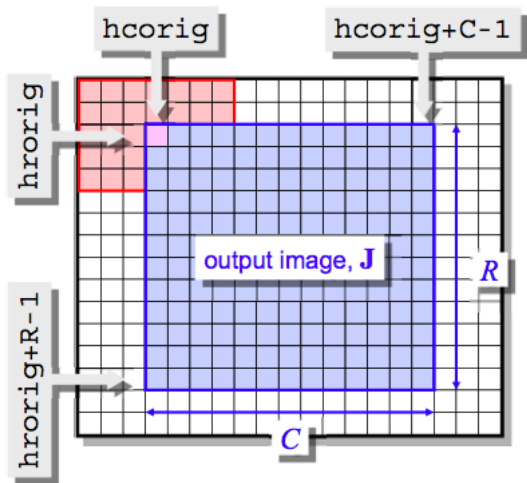


Tích chập bằng phương pháp Cộng - Nhân - Dịch chuyển

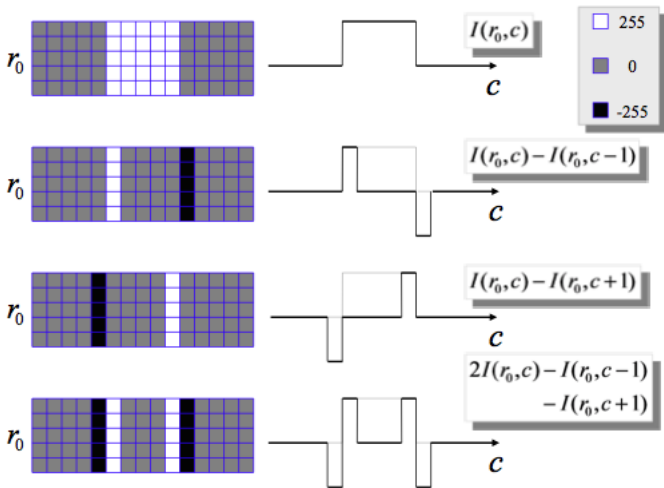


Tích chập bằng phương pháp Cộng - Nhân - Dịch chuyển

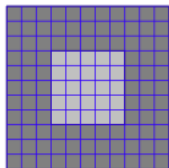
Bước cuối cùng, điều chỉnh (cắt) ảnh tích lũy cuối cùng thu được bắt đầu tại vị trí (h_{orig} , h_{orig}) và kết thúc tại vị trí ($h_{orig} + R - 1$, $h_{orig} + C - 1$) với $h_{orig} = cpad + 1$, $h_{orig} = rpad + 1$



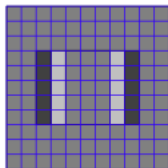
Phát hiện cạnh



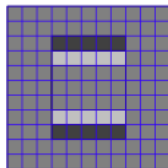
Phát hiện cạnh



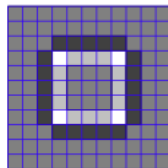
$I(r,c)$



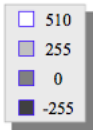
$$2I(r,c) - I(r,c-1) - I(r,c+1)$$



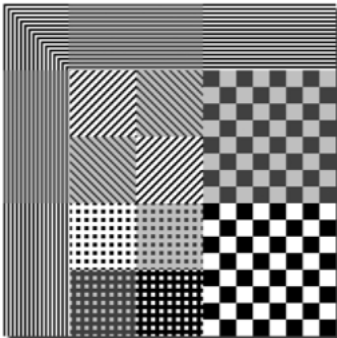
$$2I(r,c) - I(r-1,c) - I(r+1,c)$$



$$4I(r,c) - I(r-1,c) - I(r+1,c) - I(r,c-1) - I(r,c+1)$$

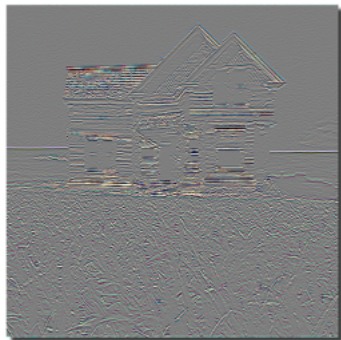
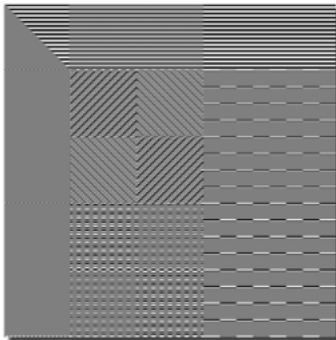


Tích chập: Ví dụ



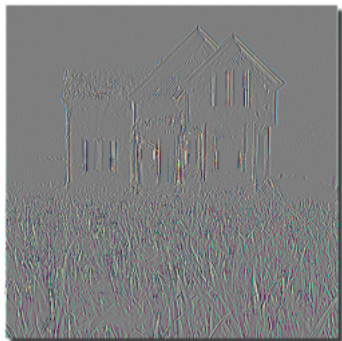
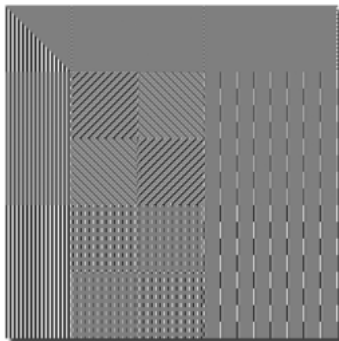
Tích chập: Sự khác nhau theo chiều dọc

$$\mathbf{h} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$



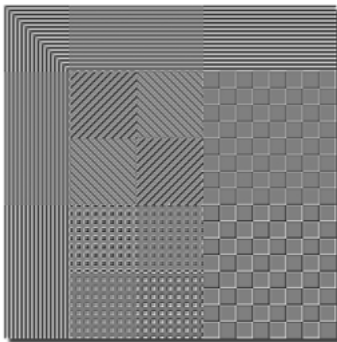
Tích chập: Sự khác nhau theo chiều ngang

$$\mathbf{h} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$



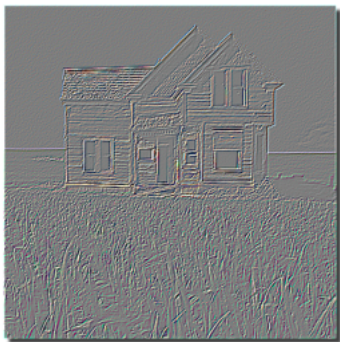
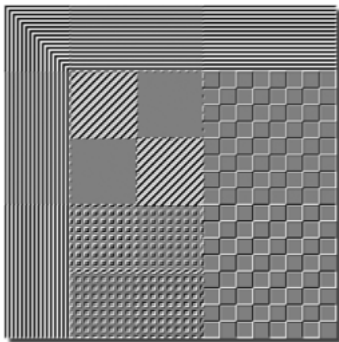
Tích chập: Sự khác nhau theo chiều ngang + chiều dọc

$$\mathbf{h} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$



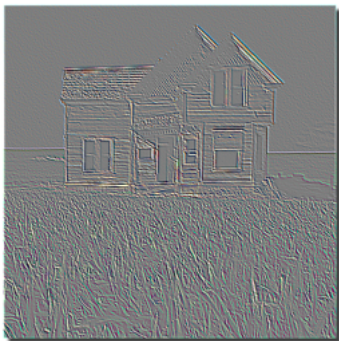
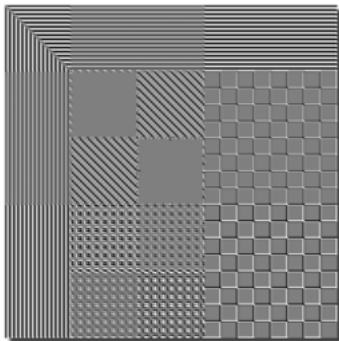
Tích chập: Sự khác nhau theo đường chéo

$$\mathbf{h} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$



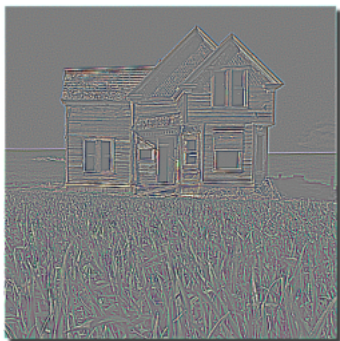
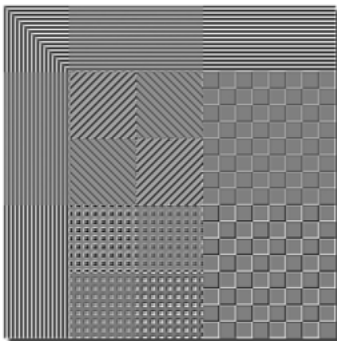
Tích chập: Sự khác nhau theo đường chéo

$$\mathbf{h} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$



Tích chập: Sự khác nhau theo đường chéo

$$\mathbf{h} = \begin{bmatrix} -1 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$



Tích chập: Sự khác nhau H + V + D

$$\mathbf{h} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

