

知能制御システム学

画像処理の基礎 (2)
— OpenCV による基本的な例 —

東北大学 大学院情報科学研究科

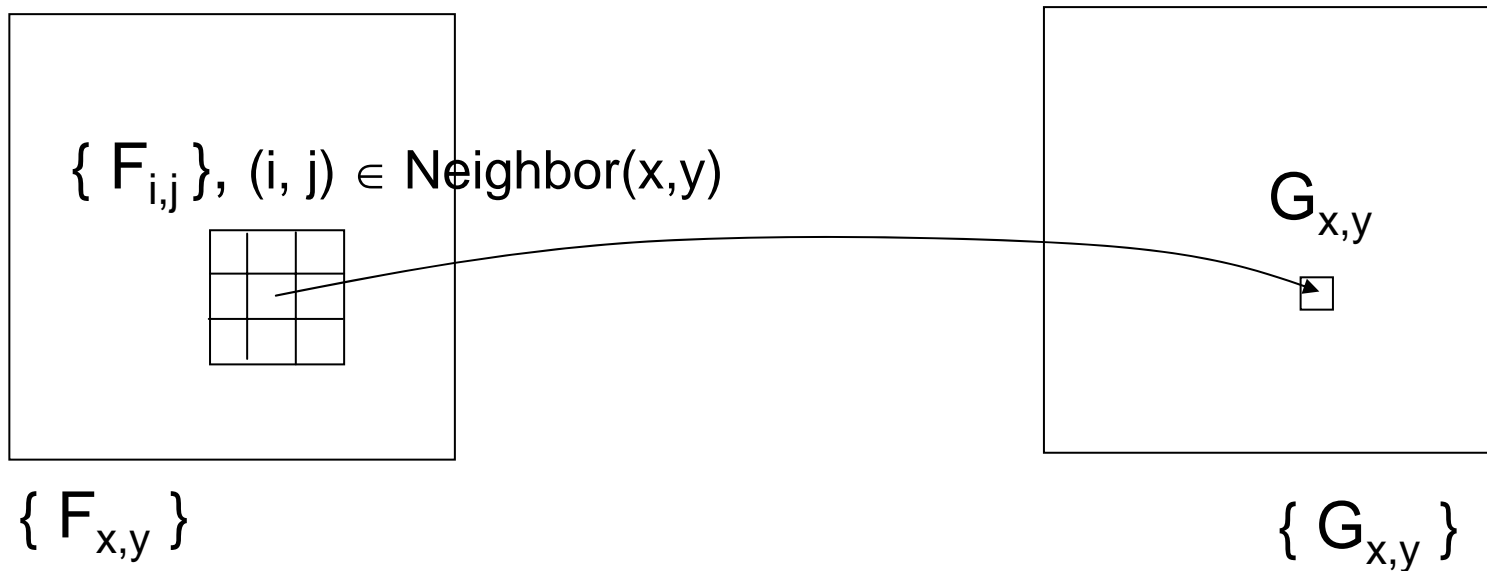
鏡 慎吾

swk(at)ic.is.tohoku.ac.jp

2009.06.30

局所処理の例 — 空間フィルタリング

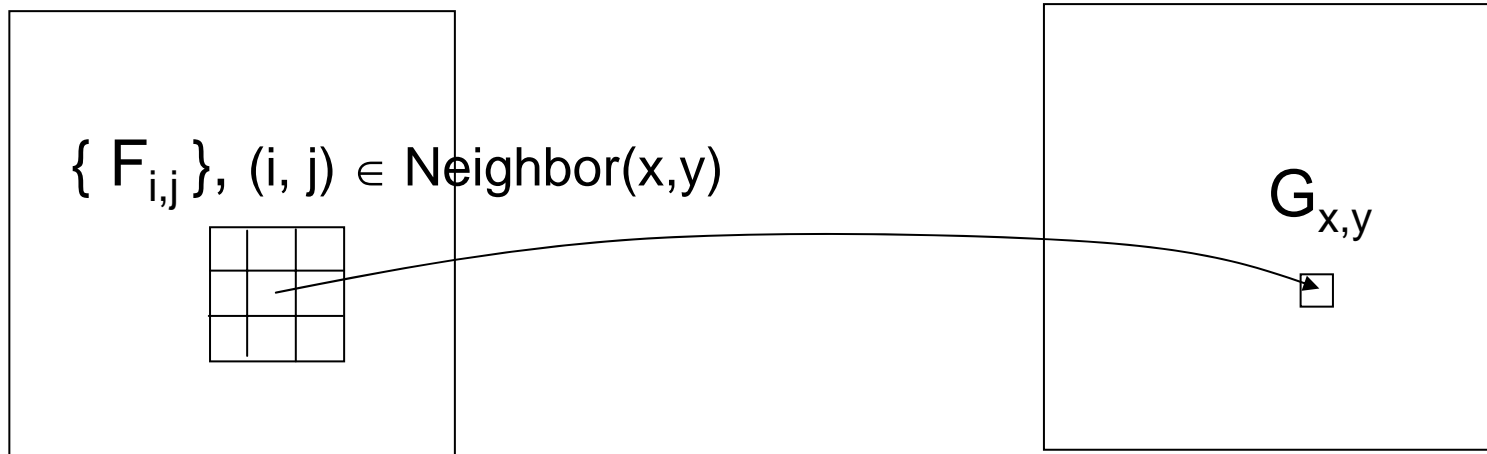
注目点の近傍(典型的には3x3画素, 5x5画素, ... など)の画素値から, 出力 $G_{x,y}$ を定める



典型例: 平滑化 (smoothing), エッジ検出 (edge detection)

典型かつ重要な例: 平滑化 (smoothing)

- ノイズを除去したいときや, 微細な構造を無視したいときに適用
- 近傍の画素値集合の代表値 (平均など) を出力とする



- 例えば周囲の3x3画素 $\{F_{i,j}\}$ の平均を出力 $G_{x,y}$ とする (例1)
- あるいは, 中心画素から遠いほど小さくなるような重みをつけた平均を出力とする (例2)

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

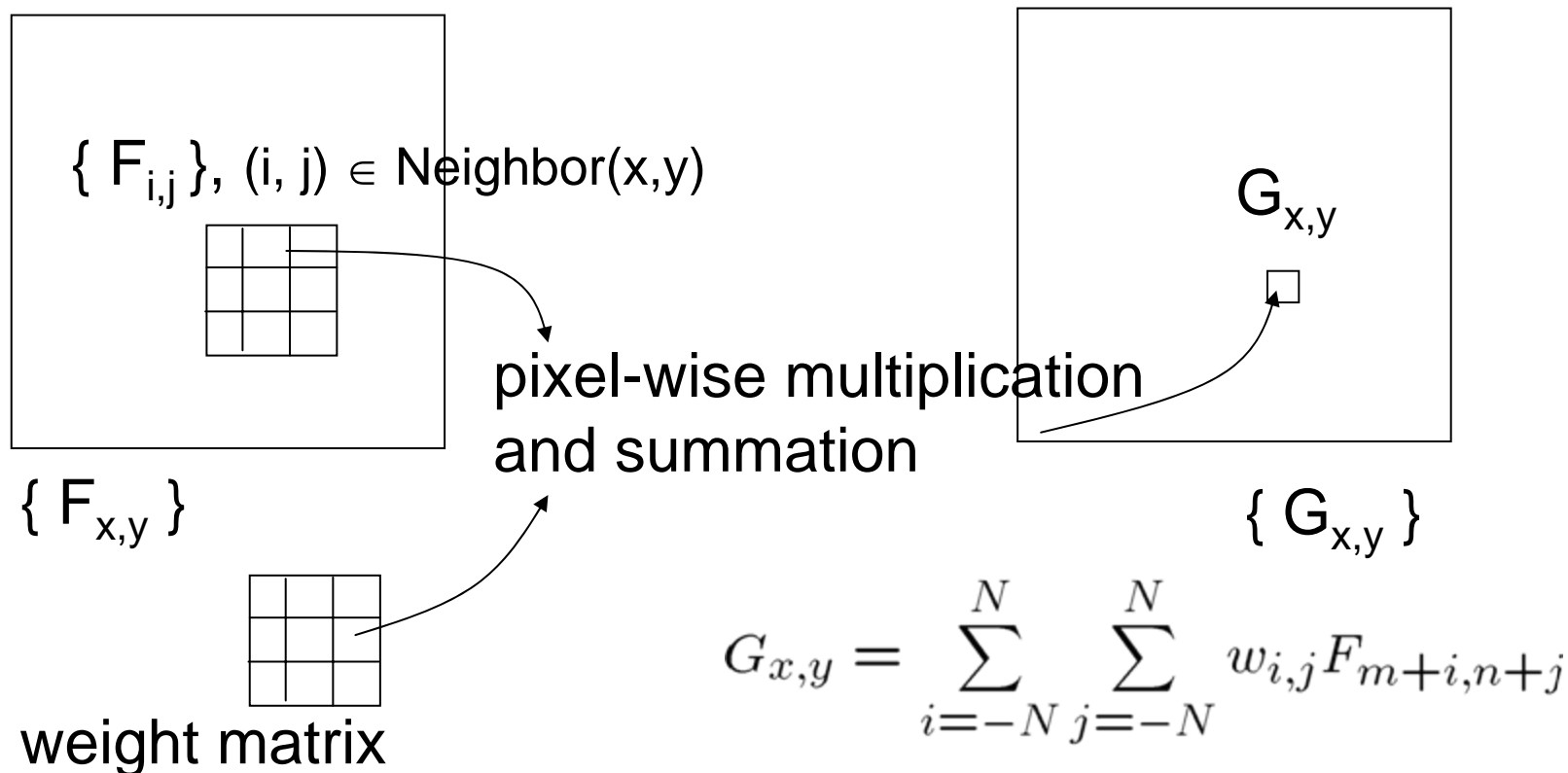
(例1)

1/16	1/8	1/16
1/8	1/4	1/8
1/16	1/8	1/16

(例2)

線形空間フィルタリング

- 前ページの例のように、空間フィルタのうち線形でシフト不変なもの線形空間フィルタ、あるいは単に線形フィルタと呼ぶ
- 加重マトリックス(マスク, カーネル, フィルタ係数, あるいは単にフィルタとも呼ばれる)のたたみこみとして表すことができる



3x3平滑化の例

1	1	1
1	1	1
1	1	1

1	1	1
1	2	1
1	1	1

0	1	0
1	4	1
0	1	0

3x3線形フィルタのコード例

```
double w[3 * 3] = { 0.0, 1.0, 0.0,  
                   1.0, 4.0, 1.0,  
                   0.0, 1.0, 0.0 };  
double scaling = 1.0;  
  
for (j = 0; j < img->height; j++) {  
    for (i = 0; i < img->width; i++) {  
        double sum = 0.0;  
        for (n = 0; n < 3; n++) {  
            for (m = 0; m < 3; m++) {  
                sum += pval(img, i - 1 + m, j - 1 + n)  
                    * w[m + 3 * n];  
            }  
        }  
        PIXVAL(result, i, j) = clipvalue(scaling * sum);  
    }  
}
```

sample program:
•filter3x3.cpp

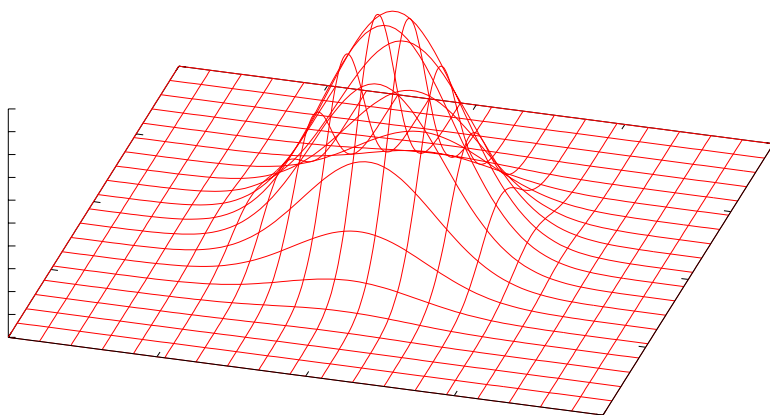
※ OpenCVでは実際には, cvFilter2D() を使うとよい. あるいは平滑化であれば cvSmooth() を使う方がよい場合もある. 後述するエッジ検出等の例も cvSobel(), cvLaplace などを使い分けるとよい

- 画像の読み出しの際に, 座標範囲を越えてしまわないように注意する (サンプルでは pval() の仕事)
- 画素値の範囲を越えてしまわないように注意する (サンプルでは clipvalue() の仕事)

ガウシアンによる平滑化

- 平滑化を実現するような加重マトリックスの中で、最も広く用いられているのが2次元ガウス関数(ガウシアン)である.

$$g_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{x^2 + y^2}{2\sigma^2}\right\}$$



- 適当に離散化して用いる
- 場合によって、加重値を適当に整数にまるめて用いる
- σ によって平滑化の強さを表現できる(大きい σ を使う場合にはマトリックスのサイズも大きくしないと正確に表現できない)

なぜガウシアンなのか? (1)

Marr and Hildreth (1980) の主張

- 理想的な平滑化フィルタは
 - 周波数領域でも滑らかでかつ局在しているべきである(平滑化とはそもそも帯域制限することである)
 - 空間領域で滑らかでかつ局在しているべきである(空間的な明るさの変化は, 照明変化や物体の境界などのように局所的なモノによって発生する)
- 周波数領域で局在していることと, 空間領域で局在していることは相反する要求である.
- この相反する要求を「両領域での分布の標準偏差の積が最小になる」という意味で最大限に満たすのがガウシアンである

なぜガウシアンなのか? (2)

スケールスペース:

- 画像 $I(x, y)$ をある量 t だけ平滑化した $I(x, y; t)$ を考えて、この3次元空間で画像を理解しようという考え方
- この考え方のもとで、適当な条件を仮定すると、それらをすべて満たすような加重マトリックスはガウシアンに一意に定まる

例) Florack et al. (1992) の条件:

- 線形かつシフト不変
- スケール不変 (適当に正規化した無次元量によって平滑化の処理を表現できる)
- 平滑化という操作が「半群」の構造を持つ (ある画像をある量 t_1 だけ平滑化してからさらに t_2 だけ平滑化した結果は、元の画像を $t_1 + t_2$ だけ平滑化したのと等価になる)
- 等方的である

エッジ検出 (edge detection) の例

- 空間微分 (実際には差分) を計算する

-1	0	1
-1	0	1
-1	0	1

単純な水平方向
1次微分

-1	-1	-1
0	0	0
1	1	1

単純な垂直方向
1次微分

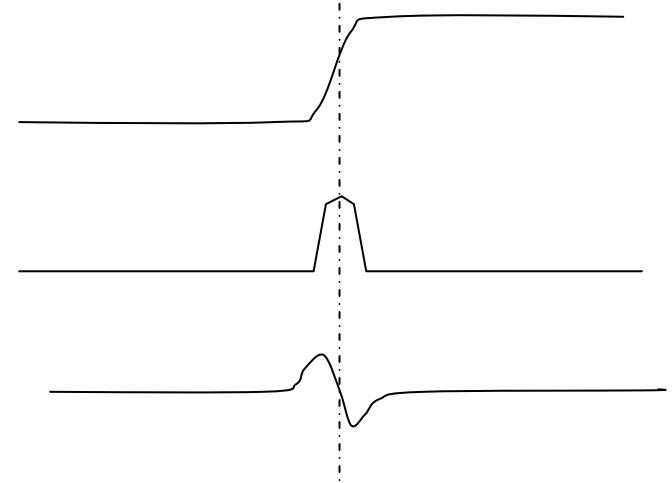
-1	0	1
-2	0	2
-1	0	1

注目画素の近くに重み
(Sobel フィルタ)

-1	-2	-1
0	0	0
1	2	1

2階微分によるエッジ検出

- あるいは2階微分を計算してそのゼロ交差を求める
- ラプラシアン $\partial^2/\partial x^2 + \partial^2/\partial y^2$ は最低次元の等方性微分演算子であり、方向によらずにエッジを得ることができる。



- 1次元の2階微分(差分) $f_{i+1} - 2f_i + f_{i-1}$ の縦横方向を足し合わせた 3×3 マトリックスによってラプラシアンを表現できる

0	1	0
1	-4	1
0	1	0

あるいは周辺
画素も考慮して

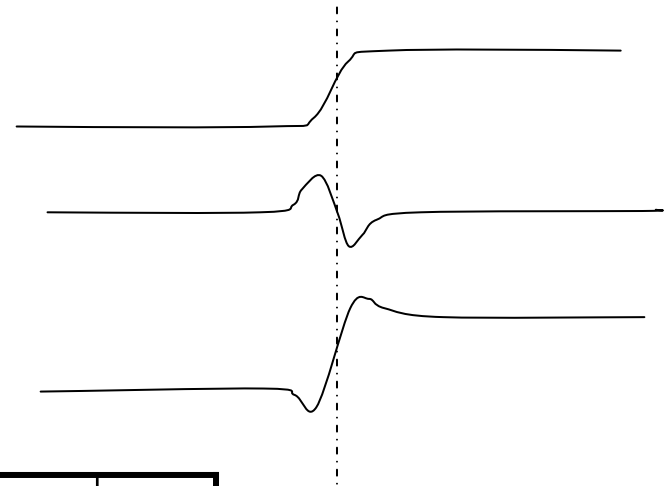
1	1	1
1	-8	1
1	1	1

LoG フィルタと DoG フィルタ

- ガウシアンフィルタによって段階的に平滑化した画像に対してラプラシアンフィルタを適用することで、異なるスケールでの画像特徴を抽出することができる。Laplacian of Gaussian (LoG) フィルタと呼ばれる。
- LoG は、平滑化の量がわずかに異なる2つのガウシアンフィルタ画像の差分として近似できる。Difference of Gaussian (DoG) と呼ばれる。スケールスペースの画像群がそもそも必要な場合は、こちらの方が効率がよい。
$$\text{DoG}(\cdot, t) = G(\cdot, t + \delta t) - G(\cdot, t)$$
- 脊椎動物の視覚野の反応が DoG でよく近似できることが報告されている。

先鋭化 (sharpening) の例

ラプラシアン画像を元画像から引く
ことで高域強調の効果を得る



$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

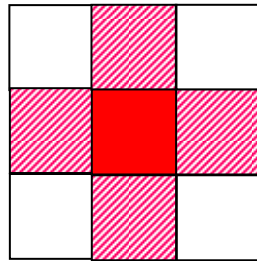
2値画像処理

2値化された後の画像 (binary image) の処理は、それだけで一分野をなすほど独特に発展している。

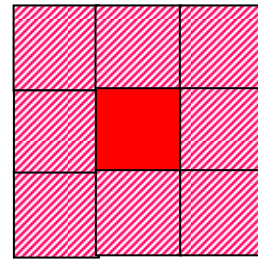
- 実用上重要であった
- 幾何学的に明確な議論がしやすく、体系的な議論が進んだ
- 画素に離散化されているため、従来の連続な図形の幾何学で考えられてきた「連結」や「距離」の概念を考え直す必要がある → デジタル幾何学

連結性

近傍 (neighbor): ある画素の近くにある画素の集合. いろいろな定義が可能.



4-neighbor



8-neighbor

互いに n -近傍の関係にある 2 つの画素は「 n -隣接している」 (n -adjacent) と呼ばれる.

同じ画素値を持つ 2 つの画素 a, b に対して画素の系列 $p_0 (= a), p_1, p_2, \dots, p_{n-1}, p_n (= b)$ が存在し, p_i はすべて同じ画素値を持ち, p_i と p_{i-1} が n -隣接するとき, 画素 a と b は「 n -連結している」 (n -neighbor connected) という

数学的モルフォロジ

dilation: 近傍の誰かが 1 だったら自分も 1 になる

$$G_{i,j} = F_{i,j} \mid F_{i-1,j} \mid F_{i+1,j} \mid F_{i,j-1} \mid F_{i,j+1} \quad (4\text{-近傍の場合})$$

erosion: 近傍の誰かが 0 だったら自分も 0 になる

$$G_{i,j} = F_{i,j} \& F_{i-1,j} \& F_{i+1,j} \& F_{i,j-1} \& F_{i,j+1} \quad (4\text{-近傍の場合})$$

opening: erosion + dilation

closing: dilation + erosion

sample program:

- morph.cpp

※ OpenCVでは実際には, cvErode(), cvDilate() などを使うとよい

画像から1次元データへの変換

projection (射影)

```
for (i = 0; i < img->width; i++) { hx[i] = 0; }
for (j = 0; j < img->height; j++) { hy[j] = 0; }

for (j = 0; j < img->height; j++) {
    for (i = 0; i < img->width; i++) {
        if (PIXVAL(img, i, j) > 128) {
            hx[i]++;
            hy[j]++;
        }
    }
}
```

sample program:

- xyproj.cpp

画像からスカラー値への変換

モーメント特徴

$$m_{p,q} = \sum_i \sum_j i^p j^q F_{i,j}$$

$$m_{0,0} = \sum_i \sum_j F_{i,j} \quad \text{0次モーメント: 2値画像の場合, 面積}$$

$$m_{1,0} = \sum_i \sum_j i F_{i,j} \quad \text{x方向の1次モーメント}$$

$$m_{0,1} = \sum_i \sum_j j F_{i,j} \quad \text{y方向の1次モーメント}$$

0~1次モーメントまでの情報から, 面積と重心が分かる.

$$(g_x, g_y) = (m_{1,0}/m_{0,0}, m_{0,1}/m_{0,0})$$

さらに高次のモーメントから, 形状に関するより詳細な情報が得られる.

sample program:
•moment.cpp

※ OpenCV では cvMoments()
が便利な場合もある.

References

- [1] 田村: コンピュータ画像処理, オーム社, 2002.
- [2] T. Lindeberg: Scale-space: A framework for handling image structures at multiple scales, Proc. CERN School of Computing, 1996.
- [3] T. Lindeberg: Scale-space theory: A basic tool for analysing structures at different scales, J. Applied Statistics, vol.21, no.2, pp.225-270, 1994.
- [4] D. Marr: ビジョン — 視覚の計算理論と脳内表現 —, 産業図書, 1987.
- [5] D. Marr and E. Hildreth: Theory of Edge Detection, Proc. Roy. Soc. London, Series B, Biological Sciences, vol.207, no.1167, pp.187-217, 1980.
- [6] L. M. J. Florack, B. M. T. Romeny, J. J. Koenderink and M. A. Viergever: Scale and the Differential Structure of Images, Image and Vision Computing, vol.10, no.6, pp.376-388, 1992.