Computer Vision I

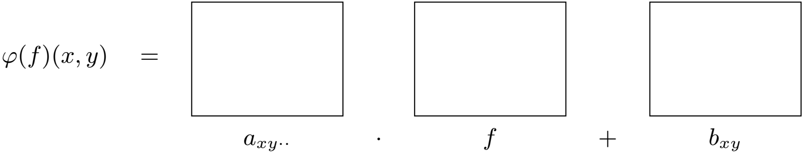
Bjoern Andres, Holger Heidrich

Machine Learning for Computer Vision
TU Dresden

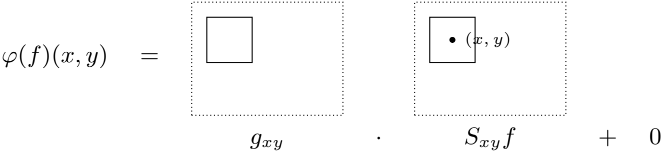
Winter Term 2022/2023

Lemma 1. An operator $\varphi\colon \mathbb{R}^{[n_0]\times[n_1]} \to \mathbb{R}^{[n_0]\times[n_1]}$ is **linear** if and only if there exist $a\colon ([n_0]\times[n_1])^2 \to \mathbb{R}$ and $b\colon [n_0]\times[n_1] \to \mathbb{R}$ such that for any (image) $f \in \mathbb{R}^{[n_0]\times[n_1]}$ and any (pixel) $(x,y) \in [n_0]\times[n_1]$ we have

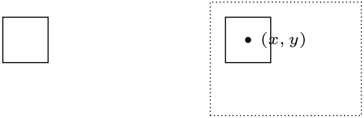
$$\varphi(f)(x,y) = \sum_{j=0}^{n_0-1} \sum_{k=0}^{n_1-1} a_{xyjk} f(j,k) + b_{xy} \quad . \tag{1}$$



More restrictive than such an operator with $(n_0 n_1)^2 + (n_0 n_1)$ coefficients is:



Even more restrictive is the typical setting in which we are given $m_0, m_1 \in \mathbb{N}$ and $g\colon [m_0] \times [m_1] \to \mathbb{R}$ and



$$\varphi(f)(x,y) = g \cdot S_{xy} f + 0$$

$$= \sum_{j=0}^{m_0-1} \sum_{k=0}^{m_1-1} g(j,k) f\left(x + j - \left\lfloor \tfrac{m_0-1}{2} \right\rfloor, y + k - \left\lfloor \tfrac{m_1-1}{2} \right\rfloor\right)$$

Remark 1.

1. f needs to be extended in order for $\varphi(f)$ to be well-defined.
2. g defines the linear operator $\varphi =: \varphi_g$ uniquely.
3. g is itself a digital image.
4. The application of operators φ_g to images f defines a binary operation $f \otimes g := \varphi_g(f)$.

Definition 1. For the set $\mathbb{R}^{\mathbb{Z}}$ of all functions from $\mathbb{Z}$ to $\mathbb{R}$, **convolution** is the operation $*\colon \mathbb{R}^{\mathbb{Z}} \times \mathbb{R}^{\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}}$ such that for any $f, g\colon \mathbb{Z} \to \mathbb{R}$ and any $t \in \mathbb{Z}$:

$$(f * g)(t) = \sum_{s=-\infty}^{\infty} f(t+s)\, g(-s) \ . \tag{2}$$

For the set $\mathbb{R}^{\mathbb{Z}\times\mathbb{Z}}$ of all functions from $\mathbb{Z} \times \mathbb{Z}$ to $\mathbb{R}$, **convolution** is the operation $*\colon \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}} \times \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}} \to \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}}$ such that for any $f, g\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ and any $(x, y) \in \mathbb{Z} \times \mathbb{Z}$:

$$(f * g)(x, y) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(x+j, y+k)\, g(-j, -k) \ . \tag{3}$$

Lemma 2. For any $f, g, h \in \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}}$ and any $\alpha \in \mathbb{R}$, we have:

$$f * g = g * f \qquad \text{(commutativity)} \qquad (4)$$

$$f * (g * h) = (f * g) * h \qquad \text{(associativity)} \qquad (5)$$

$$f * (g + h) = (f * g) + (f * h) \qquad \text{(distributivity)} \qquad (6)$$

$$\alpha(f * g) = (\alpha f) * g \qquad \text{(associativity with } \cdot\text{)} \qquad (7)$$

Definition 2. For any $C \neq \emptyset$, the operator $X : \bigcup_{n_0, n_1 \in \mathbb{N}} C^{[n_0] \times [n_1]} \to C^{\mathbb{Z} \times \mathbb{Z}}$ such that for any $n_0, n_1 \in \mathbb{N}$, any $f \colon [n_0] \times [n_1] \to C$ and any $(x, y) \in \mathbb{Z}^2$ we have

$$X(f)(x,y) = \begin{cases} f(x,y) & \text{if } (x,y) \in [n_0] \times [n_1] \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

is called the **infinite** 0-**extension** of digital images.

Definition 3. For any $C \neq \emptyset$ and any $n_0, n_1 \in \mathbb{N}$, the map $R_{n_0,n_1} \colon C^{\mathbb{Z} \times \mathbb{Z}} \to C^{[n_0] \times [n_1]}$ such that for any $f \colon \mathbb{Z} \times \mathbb{Z} \to C$ and any $(x, y) \in [n_0] \times [n_1]$, we have $R_n(f)(x, y) = f(x, y)$ is called the (n_0, n_1)-**restriction** of infinite digital images.

Definition 4. For any $j, k \in \mathbb{Z}$, the operator $S_{jk} \colon C^{\mathbb{Z}\times\mathbb{Z}} \to C^{\mathbb{Z}\times\mathbb{Z}}$ such that for any $x, y \in \mathbb{Z}$, we have $S_{jk}(f)(x, y) = f(x + j, y + k)$ is called the (x, y)-**shift** of infinite digital images.

Definition 5. The operator $L \colon C^{\mathbb{Z}\times\mathbb{Z}} \to C^{\mathbb{Z}\times\mathbb{Z}}$ such that for any $x, y \in \mathbb{Z}$, we have $L(f)(x, y) = f(-x, -y)$ is called the **reflection** of infinite digital images.

Definition 6. For any $n_0, n_1, m_0, m_1 \in \mathbb{N}$, any $f \in C^{[n_0]\times[n_1]}$, any $g \in C^{[m_0]\times[m_1]}$, $d_0 = -\left\lfloor \frac{m_0-1}{2} \right\rfloor$ and $d_1 = -\left\lfloor \frac{m_1-1}{2} \right\rfloor$, the **convolution** of f and g is defined as

$$f * g := R_{n_0 n_1}(X(f) * S_{d_0 d_1}(X(g))) \tag{9}$$

Lemma 3. For any $n_0, n_1, m_0, m_1 \in \mathbb{N}$, any $f \in C^{[n_0]\times[n_1]}$ and any $g \in C^{[m_0]\times[m_1]}$:

$$f \otimes g = f * L(g) \tag{10}$$

Definition 7. For any $\sigma \in \mathbb{R}^+$ and any $m \in \mathbb{N}_0$ (typically: $m \geq 3\sigma$), for the function

$$w\colon \quad \mathbb{R} \to \mathbb{R}\colon \quad t \mapsto e^{-\frac{t^2}{2\sigma^2}} \tag{11}$$

and the number

$$N := \sum_{j=-m}^{m} w(j) \ , \tag{12}$$

the functions

$$g_0\colon \quad [2m+1] \times [1] \to \mathbb{R}\colon \quad (x, 0) \mapsto \frac{w(j-m)}{N} \tag{13}$$

$$g_1\colon \quad [1] \times [2m+1] \to \mathbb{R}\colon \quad (0, y) \mapsto \frac{w(j-m)}{N} \tag{14}$$

are called **Gaussian averaging filters**.

f

$f * g_0 * g_1$

$\sigma = 3.0$
$m = 9$

$f * g_0 * g_1$

$\sigma = 10.0$
$m = 30$

Linear operators

f

f

$2f - (f * g_0 * g_1)$

$\sigma = 1.0$
$m = 3$

Definition 8. The **discrete derivatives** of an infinite digital image $f\colon \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$ are defined as

$$\partial_0 f := g * d_0 \tag{15}$$

$$\partial_1 f := g * d_1 \tag{16}$$

with

$$d_0 = \frac{1}{2}(1, 0, -1) \tag{17}$$

$$d_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \tag{18}$$

The **discrete gradient** is defined as

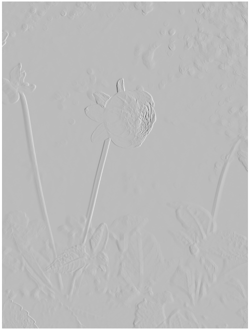
$$\nabla f = \begin{pmatrix} \partial_0 f \\ \partial_1 f \end{pmatrix} \ , \tag{19}$$

and $|\nabla f| = \sqrt{(\partial_0 f)^2 + (\partial_1 f)^2}$ is commonly referred to as its **magnitude**.

Linear operators

f

$f * d_0$



$f * d_1$

Linear operators

f

$\sqrt{(f * d_0)^2 + (f * d_1)^2}$