Fundamentals of Image Processing (part I)

Alexandre Xavier Falcão

Institute of Computing - UNICAMP

afalcao@ic.unicamp.br

Fundamentals of Image Processing

Image processing [1] is the area of Computer Science that transforms images from one representation to another for

- quality enhancement,
- geometric transformations,
- reconstruction from projections,
- coding (compression), decoding,
- representation, description,
- segmentation, ...

Most image transformations may be divided into three categories.

Fundamentals of Image Processing

Most image transformations may be divided into three categories.

- point-based transformations: the value of a pixel p in the ouput image depends only on the value of p in the input image(s).
- adjacency-based transformations: the value of a pixel p in the ouput image depends on the values of its adjacent pixels in the input image.
- connectivity-based transformations: the value of a pixel p in the ouput image depends on the values of a sequence of adjacent pixels in the input image(s) with terminus at p.

Fundamentals of Image Processing

Most image transformations may be divided into three categories.

- point-based transformations: the value of a pixel p in the ouput image depends only on the value of p in the input image(s).
- adjacency-based transformations: the value of a pixel p in the ouput image depends on the values of its adjacent pixels in the input image.
- connectivity-based transformations: the value of a pixel p in the ouput image depends on the values of a sequence of adjacent pixels in the input image(s) with terminus at p.

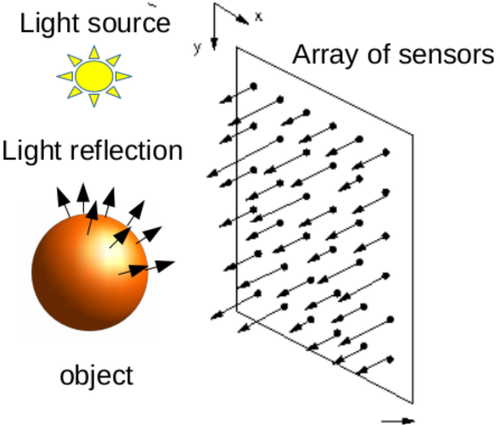
The initial lectures will provide concepts and examples from the above categories that are important for this course.

Agenda

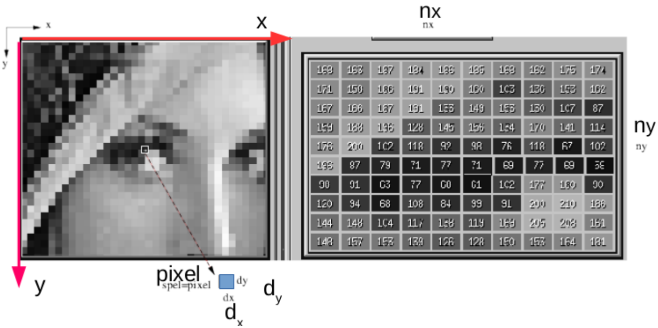
- Digital Images: definition and representation.
- Point-based transformations.
- Adjacency relations and simple morphological operations.

Gray-scale image

A gray-scale image may be acquired from an array of sensors that measure light reflection on visible object points (sampling), transforming those measures into integer numbers (quantization).



A gray-scale image may be acquired from an array of sensors that measure light reflection on visible object points (sampling), transforming those measures into integer numbers (quantization).



Sampling and quantization

- Sampling: A continuous function $f(x, y)$ of light reflection is measured on the 2D array of sensors at $p = (x, y)$ positions (pixels) separated by (d_x, d_y) (spatial resolution) to form a discrete function $I(x, y)$.

$$
\begin{aligned}
I(x, y) &= f(x, y) \sum_{j=0}^{n_y - 1} \sum_{i=0}^{n_x - 1} \delta(x - id_x, y - jd_y) \\
\delta(x - id_x, y - jd_y) &= \begin{cases} 1 & \text{if } x = id_x \text{ and } y = jd_y, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

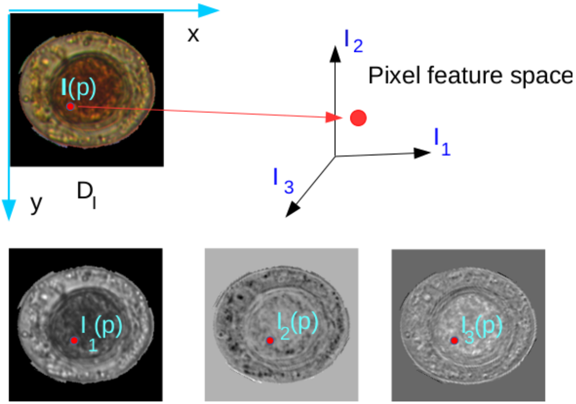
- Quantization: The values $I(x, y)$ are quantitized into integer values $I(p) \in [0, 2^b - 1]$ of b bits (image depth).

Digital image

- Images may be acquired using multiple filters of distinct wavelengths, which creates a grayscale image for each channel (band).
- By image processing, the image values can also be transformed into real values.
- Formally, a 2D multi-band image $\hat{I}$ is a pair $(D_I, \mathbf{I})$ in which $\mathbf{I}(p) \in \Re^m$ assigns m scalar values to each pixel $p \in D_I \subset \mathcal{Z}^2$.
- The image domain D_I may also have more than 2 dimensions, but this is out of the scope of this course.

Digital image

For $m = 3$ bands, each pixel p is represented by a point $\mathbf{I}(p) = (I_1(p), I_2(p), I_3(p)) \in \Re^3$.

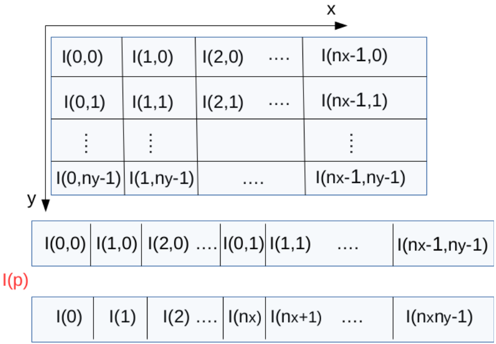


A color image with its three color components (gray-scale images).

Representation

A gray-scale image (or each band of a color image) may be stored in a 2D array $I(x_p, y_p)$ or a vector $I(p)$,

$$\begin{aligned} p &= x_p + y_p n_x, \\ x_p &= p \% n_x, \\ y_p &= \lfloor p / n_x \rfloor. \end{aligned}$$



Point-based operations

Examples of point-based operations are mathematical operations between images.

- Let $\hat{I} = (D_I, I)$ and $\hat{J} = (D_I, J)$ be two gray-scale images with the same domain and $\odot$ be any logical/arithmetical operation.

- The resulting image $\hat{K} = (D_I, K) = \hat{I} \odot \hat{J}$ is obtained by applying the operation pixel by pixel.

  - Adition and subtraction: $K(p) = I(p) + J(p)$ and $K(p) = I(p) - J(p)$.

  - Union and intersection: $K(p) = \max\{I(p), J(p)\}$ and $K(p) = \min\{I(p), J(p)\}$.

Point-based operations

Mathematical operations might also involve $\hat{I}$ and a scalar or a function to create $\hat{K}$.

- Square root $\hat{K} = \sqrt{\hat{I}}$: $K(p) = \sqrt{I(p)}$.
- Logarithm $\hat{K} = \log \hat{I}$: $K(p) = \log I(p)$.
- Module of the difference $\hat{K} = |\hat{I} - \hat{J}|$: $K(p) = |I(p) - J(p)|$.

Point-based operations

- Point-based operations may involve color space transformations for image representation.
- From RGB to YCbCr color space, an RGB image $\hat{I}$ can be transformed into a YCbCr image $\hat{K}$ pixel by pixel as follows.

$$\begin{bmatrix} K_1(p) \\ K_2(p) \\ K_3(p) \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} I_1(p) \\ I_2(p) \\ I_3(p) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{2^b}{2} \\ \frac{2^b}{2} \end{bmatrix}$$

where b is the depth of image $\hat{I}$.

Adjacency relation

An adjacency relation $\mathcal{A} \subseteq D_I \times D_I$ may be defined in the image domain and/or feature space as a binary relation.

- $\mathcal{A}_1$: $\{(p, q) \in D_I \times D_I \mid \|q - p\| \leq \alpha_1\}$,
- $\mathcal{A}_2$: $\{(p, q) \in D_I \times D_I \mid \|\mathbf{I}(q) - \mathbf{I}(p)\| \leq \alpha_2\}$,
- $\mathcal{A}_3$: $\{(p, q) \in D_I \times D_I \mid \|q - p\| \leq \alpha_1 \text{ and } \|\mathbf{I}(q) - \mathbf{I}(p)\| \leq \alpha_2\}$,

$\alpha_1, \alpha_2 \in \Re^+$, such that $\mathcal{A}(p)$ is the set of pixels q adjacent to p.

An adjacency relation $\mathcal{A} \subseteq D_I \times D_I$ may be defined in the image domain and/or feature space as a binary relation.

- $\mathcal{A}_1$: $\{(p,q) \in D_I \times D_I \mid \|q - p\| \leq \alpha_1\}$,
- $\mathcal{A}_2$: $\{(p,q) \in D_I \times D_I \mid \|\mathbf{I}(q) - \mathbf{I}(p)\| \leq \alpha_2\}$,
- $\mathcal{A}_3$: $\{(p,q) \in D_I \times D_I \mid \|q - p\| \leq \alpha_1$ and $\|\mathbf{I}(q) - \mathbf{I}(p)\| \leq \alpha_2\}$,

$\alpha_1, \alpha_2 \in \Re^+$, such that $\mathcal{A}(p)$ is the set of pixels q adjacent to p.

For the image on the right, what is the adjacency set of $p = (2,3)$ for $\mathcal{A}_1, \mathcal{A}_2$, and $\mathcal{A}_3$, when $\alpha_1 = \sqrt{5}$ and $\alpha_2 = 0$?

y \ x	0	1	2	3	4	5
0	0	1	1	0	3	2
1	2	1	2	3	2	1
2	1	0	2	3	2	0
3	2	1	2	0	1	1
4	2	2	1	2	1	0
5	1	0	2	1	0	1

Adjacency relation

- It is said translation-invariant when $\mathcal{A}$: $\{(p, q_k) \in D_I \times D_I \mid (x_{q_k}, y_{q_k}) - (x_p, y_p) = (dx_k, dy_k), k = 1, 2, \ldots, K\}$, where $\{(dx_k, dy_k)\}$ is a set of K displacements.

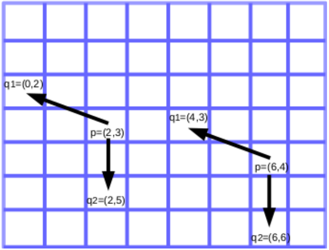
Adjacency relation

- It is said translation-invariant when $\mathcal{A}$: $\{(p, q_k) \in D_I \times D_I \mid (x_{q_k}, y_{q_k}) - (x_p, y_p) = (dx_k, dy_k), k = 1, 2, \ldots, K\}$, where $\{(dx_k, dy_k)\}$ is a set of K displacements.
- One can store the displacements and generate the set $\mathcal{A}(p) = \{q_k\}$, $q_k = (x_{q_k}, y_{q_k}) = (x_p + dx_k, y_p + dy_k)$, $k = 1, 2, \ldots, K$, for any $p \in D_I$.

- It is said translation-invariant when $\mathcal{A}$: $\{(p, q_k) \in D_I \times D_I \mid (x_{q_k}, y_{q_k}) - (x_p, y_p) = (dx_k, dy_k), k = 1, 2, \ldots, K\}$, where $\{(dx_k, dy_k)\}$ is a set of K displacements.

- One can store the displacements and generate the set $\mathcal{A}(p) = \{q_k\}$, $q_k = (x_{q_k}, y_{q_k}) = (x_p + dx_k, y_p + dy_k)$, $k = 1, 2, \ldots, K$, for any $p \in D_I$.

For fixed displacements $\{(-2, -1), (0, 2)\}$, examples of sets $\mathcal{A}(p) = \{q_1, q_2\}$ are



Adjacency relation

Rectangular adjacency relations $\mathcal{A}$ are very popular.

$$\mathcal{A}(p) = \{q \in D_I | \quad |x_q - x_p| \leq \frac{w}{2} \quad and \quad |y_q - y_p| \leq \frac{h}{2}\}.$$

y \ x	0	1	2	3	4	5
0	0	1	1	0	3	2
1	2	1	2	3	2	1
2	1	0	2	3	2	0
3	2	1	2	0	1	1
4	2	2	1	2	1	0
5	1	0	2	1	0	1

Which are the neighbors of $p = (2, 3)$ when $w = 3$ and $h = 5$?

Simple morphological operations

- A dilation $\hat{J} = \hat{I} \oplus \mathcal{A}$ is obtained pixel by pixel as

$$J(p) = \max_{q \in \mathcal{A}(p)} \{I(q)\}.$$

Simple morphological operations

- A dilation $\hat{J} = \hat{I} \oplus \mathcal{A}$ is obtained pixel by pixel as

$$J(p) = \max_{q \in \mathcal{A}(p)} \{I(q)\}.$$

- It reduces dark regions, increases bright ones, and is also known as max-pooling.

Simple morphological operations

- A dilation $\hat{J} = \hat{I} \oplus \mathcal{A}$ is obtained pixel by pixel as

$$J(p) \quad = \quad \max_{q \in \mathcal{A}(p)} \{I(q)\}.$$

- It reduces dark regions, increases bright ones, and is also known as max-pooling.

- Compute $\hat{J} = \hat{I} \oplus \mathcal{A}$ when $\mathcal{A}$ is a 4-neighborhood relation.

I

-2	3	7	8
-8	-2	6	-9
8	1	2	7
7	9	-8	0

J

3	7	8	8
8	6	7	8
8	9	7	7
9	9	9	7

Dilation algorithm

Input: $\hat{I} = (D_I, I)$ and $\mathcal{A}$.
Output: $\hat{J} = (D_I, J)$.

1. For each $p \in D_I$, do
2. $\quad J(p) \leftarrow I(p)$.
3. $\quad$ For every $q \in \mathcal{A}(p) \setminus \{p\}$, such that $q \in D_I$, do
4. $\quad\quad$ If $I(q) > J(p)$, then
5. $\quad\quad\quad J(p) \leftarrow I(q)$.
6. Return $\hat{J}$.

Dilation algorithm

Input: $\hat{I} = (D_I, I)$ and $\mathcal{A}$.
Output: $\hat{J} = (D_I, J)$.

1. For each $p \in D_I$, do
2. $\quad J(p) \leftarrow I(p)$.
3. $\quad$ For every $q \in \mathcal{A}(p) \setminus \{p\}$, such that $q \in D_I$, do
4. $\quad\quad$ If $I(q) > J(p)$, then
5. $\quad\quad\quad J(p) \leftarrow I(q)$.
6. Return $\hat{J}$.

When combined with its dual operation, named erosion, they can create several types of morphological filters.

Examples

Let's see MathMorphology.ipynb in notebooks.tar.gz

[1] Rafael C. Gonzalez and Richard E. Woods.
Digital Image Processing (4th Edition).
Pearson, 2018.
www.imageprocessingplace.com.