Image preprocessing in the spatial domain, local neighborhood operations

Václav Hlaváč

Czech Technical University in Prague
Center for Machine Perception (bridging groups of the)
Czech Institute of Informatics, Robotics and Cybernetics and
Faculty of Electrical Engineering, Department of Cybernetics

http://people.ciirc.cvut.cz/hlavac, hlavac@ciirc.cvut.cz

Thanks to Tomáš Svoboda for several slides.

Outline of the lecture:
◆ Noise filtration (no edge detection here).
◆ Noise and its statistical nature.
◆ Space invariant filters.
◆ Discrete 2D convolution.
◆ Separable filters.
◆ Nonlinear noise filtration.
Image preprocessing, the introduction

The input is an image, the output is an image too.

The image is not interpreted. The image content is not taken into account.

---

The aim

- To suppress the distortion (e.g., correction of the geometric distortion caused by the Earth spherical shape of an satellite image).
- Contrast enhancement (useful only if the human looks at the image).
- Noise suppression.
- Enhancement of some image features needed for further image processing, e.g., edge finding.
Local image preprocessing operations

The specific knowledge about the image and distortions (i.e. semantics) in not used in preprocessing.

**Taxonomy from the usage point of view:**

1. Smoothing (noise filtration).
2. Edge detection (gradient operators, image sharpening).

**Taxonomy according to the character of the mathematical description:**

1. Linear.
2. Nonlinear.
Let us consider almost the simplest image statistical model.

Assume that each image pixel is contaminated by the additive noise:

- which is statistically independent of the image function,
- has a zero mean $\mu$,
- and has a standard deviation $\sigma$.

Let have $i$ realizations of the image, $i = 1, \ldots, n$. The estimate of the correct value is

$$
\frac{g_1 + \ldots + g_n}{n} + \frac{\nu_1 + \ldots + \nu_n}{n}.
$$

The outcome is a random variable with $\mu' = 0$ and $\sigma' = \sigma/\sqrt{n}$.

The thought above is anchored in the probability theory in its powerful Central limit theorem.
The Central limit theorem describes the probability characteristics of the ‘population of the means’, which has been created from the means of an infinite number of random population samples of size $N$, all of them drawn from a given ‘parent population’. The Central limit theorem predicts characteristics regardless of the distribution of the parent population.

1. The mean of the population of means (i.e., the means of many times randomly drawn samples of size $N$ from the parent population) is always equal to the mean of the parent population.

2. The standard deviation of the population of means is always equal to the standard deviation of the parent population divided by the square root of the sample size $N$.

3. The distribution of sample means will increasingly approximate a normal (Gaussian) distribution as the size $N$ of samples increases.
Central limit theorem (2)

- A consequence of the Central limit theorem is that if we average measurements of a particular quantity, the distribution of our average tends toward a normal (Gaussian) one.

- In addition, if a measured variable is actually a combination of several other uncorrelated variables, all of them ‘contaminated’ with a random error of any distribution, our measurements tend to be contaminated with a random error that is normally distributed as the number of these variables increases.

- Thus, the Central limit theorem explains the ubiquity of the bell-shaped ‘Normal distribution’ in the measurements domain.
Central limit theorem (3), the application view

- It is important for applications that there is no need to generate a big amount of population samples. It suffices to obtain one big enough population sample. The Central limit theorem teaches us what is the distribution of population means without the need to generate these population samples.

- What can be considered a big enough population sample? It is application dependent. Trespassing the lower bound of 30-50 random observation is not allowed by statisticians. Recall samples with about 1000 observations serving to estimate outcomes of elections.

- The confidence interval in statistics indicates the reliability of the estimate. It gives the degree of uncertainty of a population parameter. We have talked about sample mean only so far. See a statistics textbook for details.
Smoothing from several images without blurring

**Assumptions:** \( n \) images of the same unchanged scene, in which it can be assumed that noise is independent to the image.

**The correct intensity value:** \( f(i, j) \) is estimated from a random population given by all pixels at the same position in all input images \( g_k(i, j) \), e.g., by ordinary averaging,

\[
f(i, j) = \frac{1}{n} \sum_{k=1}^{n} g_k(i, j).
\]

**Example:** Suppression of thermal noise for cameras utilized for precise measurements. The correct value is usually estimated from 50 images at least.
No other choice, but to resort to the data redundancy in the image.

Neighboring pixels have the same or similar intensity value.

The intensity value can be corrected based on the analysis of intensities in the neighborhood. A single typical sample or the combination of several intensity values in the neighborhood is taken.

The trouble with blurring on the steep intensity transitions occurs..
General local filtration

- The correct (new) intensity value is estimated in the small neighborhood of the current pixel.

- Imagine that the image is systematically traversed line by line, e.g. from top left. The small neighborhood $\mathcal{O}$ around the representative pixel is analyzed. This neighborhood is often a small rectangle, called also a window.

- The result of the analysis is written to the output image at the same position as the representative pixel has in the input image. The method is (sometimes) called Moving window transform.

- In general, the filter properties can vary with the position of the representative point.
Operators independent to the shift, called also the space invariant filters

- It is a special case of the local filtration.
- The properties of the filter remain the same in all positions of the representative point in the image.
- These filters have a counterpart in the frequency domain filtration, e.g. in the Fourier spectrum.
Local linear preprocessing

- The current value is calculated as a linear combination of the image function values in the local neighborhood.

- Reminder: the operator linearity, i.e. two properties: additivity and homogeneity.

- The linearity assumption is violated for real images
  - There is a problem of limit pixel values (intensities), as intensities lie i.e. typically in the interval \( \langle 0, 255 \rangle \).
    The pixel value cannot be multiplied by an arbitrary non-zero number because the result would lie outside the limits. Similarly for addition of two pixel values.
  - Troubles at the image borders.
A discrete 2D convolution

- The contribution of individual pixels in the neighborhood $\mathcal{O}$ is weighted by coefficients $h$ in the linear combination

$$g(x, y) = \sum_{(m,n) \in \mathcal{O}} h(x - m, y - n) f(m, n).$$

- The convolution kernel $h$, also convolution mask.

- A rectangular neighborhood $\mathcal{O}$ with odd number of rows and columns is often used in order to have a representative point in the middle of the mask.
Convolution vs. cross-correlation

- **Convolution ∗** (related to the impulse response)

\[ g(x, y) = (h \ast f)(x, y) = \sum_{(m,n) \in O} h(x - m, y - n) \cdot f(m, n). \]

- **Cross-correlation ∗** (a measure of similarity between two signals/images)

\[ g(x, y) = (h \star f)(x, y) = \sum_{(m,n) \in O} h(x + m, y + n) \cdot f(m, n). \]

- Convolution by cross-correlation:
  - Flip filter \( h \) in both dimension (bottom to top, left to right).
  - Perform cross-correlation.

- The result of convolution and correlation is the same for symmetric filters.
Convolution vs. cross-correlation, 1D illustration

Convolution

Cross-correlation

Autocorrelation

Courtesy: commons.wikimedia.org
Ordinary averaging

Averaging in the $3 \times 3$ neighborhood

$$h = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$ 

Modifications, which stress importance of pixels close to the center of the mask

$$h = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad h = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$
Ordinary averaging, example 1

- Original $256 \times 256$
- Added artif. noise
- Averaging $3 \times 3$
Ordinary averaging, example 2

Original $256 \times 256$

Added artif. noise

Averaging $7 \times 7$
Separable filters

Filter is separable if $h(x, y) = h_1(x) \cdot h_2(y) = \delta * h_1(x, \cdot) * h_2(\cdot, y)$, where $\delta$ is a Dirac function.

Separable filters can be implemented via successive 1D convolutions (associativity of convolution). Let $M, N$ be size of the image, $m, n$ be size of the rectangular convolution mask $h$. Separability reduces the computational complexity of the convolution from $O(MNmn)$ to $O(MN(m + n))$.

Example: a binomical 2D filter of the size $5 \times 5$. The filter element is constituted as the edition of two preceding elements in Pascal triangle.

$$
\begin{bmatrix}
1 & 4 & 6 & 4 & 1 \\
4 & 16 & 24 & 16 & 4 \\
6 & 24 & 36 & 24 & 6 \\
4 & 16 & 24 & 16 & 4 \\
1 & 4 & 6 & 4 & 1 \\
\end{bmatrix}
= 
\begin{bmatrix}
h_1 \\
h_2 \\
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
4 \\
6 \\
4 \\
1 \\
\end{bmatrix}
\begin{bmatrix}
1 & 4 & 6 & 4 & 1 \\
\end{bmatrix}.
$$
Separability $\Rightarrow$ savings in calculations

The size of the convolution mask is $2N + 1$.

$$g(x, y) = \sum_{m=-N}^{N} \sum_{n=-N}^{N} h(x - m, y - n) f(m, n)$$

$$= \sum_{m=-N}^{N} \sum_{n=-N}^{N} h(m, n) f(x + m, y + n)$$

$$= \sum_{m=-N}^{N} h_1(m) \sum_{n=-N}^{N} h_2(n) f(x + m, y + n)$$
Separability $\Rightarrow$ savings in calculations 2

- Our filter of the size $5 \times 5$ needs 25 multiplications and 24 additions in each pixel.

- If a separable filter is used, only 10 multiplications and 8 additions.

The saving would be more dramatic in the case of the convolution-based filtration in the 3D image, e.g. from a tomograph. For the convolution mask of the size $9 \times 9 \times 9$, there is a need for 729 multiplications and 728 additions.

- 27 multiplications and 24 additions suffice per voxel.
Each filter with the rank 1 is separable.

The singular decomposition (SVD).

\[
[u, s, v] = \text{svd}(A);
\]

\[
s = \text{diag}(s);
\]

\[
tol = \text{length}(A) \times \max(s) \times \text{eps};
\]

\[
\text{rank} = \text{sum}(s > tol);
\]

if (rank == 1)

\[
hcol = u(:,1) \times \sqrt{s(1)};
\]

\[
hrow = \text{conj}(v(:,1)) \times \sqrt{s(1)};
\]

\[
y = \text{conv2}(hcol, hrow, x, \text{shape});
\]

else

\[
\text{%Nonseparable stencil}
\]
end
Nonlinear smoothing

The **aim**: to reduce the blur of edges while smoothing.

**1. principle**: find such a subset of the current pixel neighborhood, in which the intensity does not change much.

**2. principle – robust statistics**: The mean value is a bad estimate if outliers are present.
Rotating mask method

The homogeneous part of the $5 \times 5$ is sought by a rotating mask of the size $3 \times 3$.

There are nine possible positions of the $3 \times 3$. One is in the middle and eight are illustrated in the image.

The mask with minimal dispersion of intensity is selected for actual calculation.
Median filtration

- Median = one of quantiles of the population, 2-quantile. 
  Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable (We know it from the histogram equalization). $q$-quantile divides ordered data into $q$ equal-sized data subsets.

- Let $x$ be a random variable. Median $M$ is the value $x$ for which the probability of the phenomenon $x < M$ equals to one half.

- Calculating the median for a discrete image function is simple. It suffices to order intensity values in a local neighborhood. The median is given by the element which is in the middle of the neighborhood.

- The series with the odd number of elements are often used in order to determine uniquely the position in the middle of the series. $3 \times 3$, $5 \times 5$, atd.

- The calculation can be sped up because it can be shown that only partially ordered series suffices for determining the median.
Median, the example

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>98</td>
<td>102</td>
</tr>
<tr>
<td>99</td>
<td>105</td>
<td>101</td>
</tr>
<tr>
<td>95</td>
<td>100</td>
<td>255</td>
</tr>
</tbody>
</table>

- Mean = 117.2
- Median: 95 98 99 100 101 102 105 255
- The median calculation is robust because it copes with up to 50% outliers.
Median example, Prague castle

- Median filtration is sometimes used iteratively.
- The main disadvantage of the median filtration in a rectangular neighborhood is that it violets thin lines and sharp corners in the image.