Histogram Equalization

Adaptive Histogram Equalization

Contrast Limited Adaptive Histogram Equalization

Histogram Smoothing

Otsu’s method

Gaussian Mixture Model
Histogram Equalization

- computer vision method that adjusts the contrast of the image
- criterion is applied on the density of the brightness function
- ordering is maintained

\[ T^* = \arg\min_T (|c_1(T(k)) - c_0(k)|) \] (1)

- where \( c_0 \) is the desired cumulative histogram

Obrázek: Histogram equalization
Transformation of random variables

- is used to compute $T^*$, utilizes cumulative density of the histogram
- example: mapping a dice to $\{1, 2, 2, 2, 3, 3\}$
  - $p_{dice} = 1/6$  $F_{dice} = \{1/6, 1/3, 1/2, 2/3, 5/6, 1\}$
  - $p_{map}(x) = \{1/6, 1/2, 1/3\}$  $F_{map} = \{1/6, 2/3, 1\}$
- a mapping between $F_{dice}$ and $F_{map}$

Obrázek: Random Variables transformation
Transformation of random variables

- the transformation has to be monotonic, the ordering has to be maintained
- the mapping $F_{map} \mapsto F_{dice}$ is harder to achieve
- one brightness cannot be divided by the transform
- only translation (mind the ordering!) and merging is possible
- our dice problem results in mapping \{1 \mapsto 1, \ 2 \mapsto 4, \ 3 \mapsto 6\}
- we have made use of the whole contrast
Examples

Obrázek: Input histogram and cumulative relative histogram.

Obrázek: Equalized histogram and cumulative relative equalized histogram.
Classic equalization fails

- because the image is handled as a whole, the damage can be seen on the equalized image

Obrázek: Histogram equalization

- but it also influences the rest of the image
Adaptive Histogram Equalization

- used for images with non-uniform lighting
- the equalization is computed piece-wise

Obrázek: Adaptive Histogram equalization

- problems on the edges of the image and salt & pepper noise
- the size of the window affects the result
Contrast Limited Adaptive Histogram Equalization

- method that solves the standard equalization problems
- has a parameter of contrast limitation
- it says that no brightness can have a certain count (based on the image size)
- if a brightness exceeds this level, the value is clipped and the remainder is spread across the other brightnesses
- the method does not operate on the pixels directly, but modifies the histogram first and then finds the transform

Obrázek: Contrast Limited Adaptive Histogram equalization
CLAHE examples

Obrázek: Original image

Obrázek: After Contrast Limited Adaptive Histogram equalization
Classic examples

Obrázek: After Histogram equalization

Obrázek: After Adaptive Histogram equalization
Histogram Smoothing

- is used when finding a threshold automatically
- the threshold lies between peaks of a bimodal histogram
- due to the presence of noise we cannot find only "true" peaks
- peak is a local maximum

Obrázek: Input image and its histogram.
Conventional detection of local maxima

1. If $h'(x) = 0$, then $x$ is an extreme.
2. If $h''(x) < 0$, then $x$ is a local maximum.
3. If $h''(x) > 0$, then $x$ is a local minimum.

- we usually do not have a function available
- we use approximations

\[ h'(x) \approx h(x) - h(x - 1) = \Delta h(x), \quad (2) \]

\[ h''(x) \approx h(x) - 2h(x - 1) + h(x - 2), \quad (3) \]
Peaks and noise

- ideal peak - \{10, 20, 30, 100, 30, 20, 10\}
- noisy (real) peak - \{10, 20, 30, 20, 50, 10, 100, 80, 60\}

Obrázek: Noisy peaks.
Smoothing as a convolution

- convolution can be used for the purpose of smoothing
- the choice of the type and size of the convolution kernel will affect the result

\[(f \ast g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{\infty} f(t-\tau)g(\tau)d\tau = (g \ast f)(t)\]  \(4\)

\[(f \ast g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m]\]  \(5\)
Convolution with different kernels

Obrázek: Kernel is a constant

Obrázek: Kernel is a triangle
Non-maximum suppression

- easy but powerful tool for the local maxima detection
- uses a local window, the center point is a local maximum if it is the global maximum in the window

Obrázek: Non-maximum suppression
Non-maximum suppression Example

Obrázek: Non-maximum suppression
Otsu’s method

▶ used for image segmentation
▶ finds an optimal threshold - a bimodal histogram is desirable
▶ two classes - $C_0 \in \{1, 2, 3, \ldots, k\}$, $C_1 \in \{k + 1, \ldots, L\}$

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=1}^{L} p_i = 1.$$ \hspace{1cm} (6)

$$\omega_0 = Pr(C_0) = \sum_{i=1}^{k} p_i = \omega(k)$$ \hspace{1cm} (7)

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^{L} p_i = 1 - \omega(k)$$ \hspace{1cm} (8)
definitions of the means of the two classes

\[
\mu_0 = \sum_{i=1}^{k} i \Pr(i|C_0) = \sum_{i=1}^{k} i \frac{p_i}{\omega_0} = \sum_{i=1}^{k} \frac{ip_i}{\omega_0} = \frac{\mu(k)}{\omega(k)} \tag{9}
\]

\[
\mu_1 = \sum_{i=k+1}^{L} i \Pr(i|C_1) = \sum_{i=k+1}^{L} i \frac{p_i}{\omega_1} = \sum_{i=k+1}^{L} \frac{ip_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \tag{10}
\]

and the total mean (of the brightness)

\[
\mu_T = \mu(L) = \sum_{i=1}^{L} ip_i \tag{11}
\]
definitions of the variances of the two classes

\[
\sigma_0^2 = \sum_{i=1}^{k} (i - \mu_0)^2 \Pr(i|C_0) = \sum_{i=1}^{k} (i - \mu_0)^2 p_i/\omega_0 \quad (12)
\]

\[
\sigma_1^2 = \sum_{i=k+1}^{L} (i - \mu_1)^2 \Pr(i|C_1) = \sum_{i=k+1}^{L} (i - \mu_1)^2 p_i/\omega_1 \quad (13)
\]

we can proof that the later holds

\[
\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \quad \omega_0 + \omega_1 = 1. \quad (14)
\]
we have to find a criterion to optimize - criteria of discriminative analysis

\[ \lambda = \frac{\sigma_B^2}{\sigma_w^2}, \kappa = \frac{\sigma_T^2}{\sigma_w^2}, \eta = \frac{\sigma_B^2}{\sigma_T^2}, \]  

(15)

\[ \sigma_w^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \]  

(16)

\[ \sigma_B^2 = \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \]  

(17)

\[ \sigma_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 p_i \]  

(18)

the criteria are dependent (because \( \sigma_w^2 + \sigma_B^2 = \sigma_T^2 \)), so we can choose only one to optimize
we choose \( \eta \) because it’s the easiest to compute

the optimal threshold \( k^* \) is computed by maximizing \( \eta \) or equally by maximizing \( \sigma^2_B \)

\[
\sigma^2_B = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]}. \tag{19}
\]

\[
k^* = \arg\max_{1 \leq k < L} \sigma^2_B(k). \tag{20}
\]
Gaussian Mixture Model

- is used to model probability density
- learned via EM

\[
gmm = \sum_{i=1}^{N} \alpha_i \mathcal{N}_i (\mu_i; C_i) \quad (21)
\]

\[
\mathcal{N}_i = \frac{1}{\sqrt{(2\pi)^D |C_i|}} \exp \left( -\frac{1}{2} (x - \mu_i)^T C_i^{-1} (x - \mu_i) \right) \quad (22)
\]