

Histogram Specification

CHEN Siyu

April 6, 2022

Abstract

This article describes Histogram Specification, whose purpose is to obtain a function $f(p)$ to map the given data into a specified histogram.

This is useful when raw images directly read from sensors are not ideal in brightness or contrast.

I Problem definition

Given a series of data $P = \{p_1, p_2, \dots, p_N\}, p \in \mathbb{N}$, its histogram can be expressed as

$$h_P(p) = n_p, \quad (1)$$

in which, n_p denotes the number of element that values p in sequence P .

Given a desired target histogram $h_Q(q) = n_q$, the objective is to find a monotonic mapping function $f(p) = q$ such that:

$$Q = P' = \{f(p_1), f(p_2), \dots, f(p_N)\} = \{q_1, q_2, \dots, q_N\}, \quad (2)$$

$$h_P(p) = h_Q(f(p)). \quad (3)$$

2 Derivations

The former problem definition is in discrete form. In some extreme but not rare cases, optimal solution may not exist for discrete histogram specification. For example:

$$P = \{1, 1, 1, 1, 1\}, \quad (4)$$

$$Q = \{1, 2, 3, 4, 5\}. \quad (5)$$

In the above scenario, a desired $f(p)$ does not exist.

To simplify the problem, we assume continuous distribution on set P and Q such that $P, Q \in [0, 1]^N$. In other word, $h_P(x)$ and $h_Q(x)$ are both continuous.

According to the definition of the problem, several constrains hold as listed below:

$$h_P(x) = h_Q(f(x)), \quad (6)$$

$$\frac{\partial f(x)}{\partial x} \geq 0, \quad (7)$$

in which $x \in [0, 1]$.

According to Eq. 6, if the distribution density is identical everywhere, then the cumulative distribution shall also be equal-valued.

$$\int_0^x h_P(z) \, dz = \int_0^{f(x)} h_Q(z) \, dz. \quad (8)$$

The above equation can be transformed into the following form:

$$\int_0^x h_P(z) \, dz = \int_0^{f(x)} h_Q(z) \, dz, \quad (9)$$

$$H_P(x) = H_Q(f(x)), \quad (10)$$

where the original functions $H_P(x), H_Q(x)$ are the cumulative intensity functions of $h_P(x), h_Q(x)$. To obtain $f(x)$, the inverse function of H_Q can be applied to both side:

$$H_Q^{-1}(H_P(f(x))) = H_Q^{-1}(H_Q(f(x))) \quad (11)$$

$$H_Q^{-1}(H_P(x)) = f(x), \quad (12)$$

which is the desired optimal solution.

3 Histogram Equalization

A mapping function $f(x)$ that equalizes the histogram can be computed given that:

$$h_Q(x) = C, \quad (13)$$

where C is a constant value.

Under continuous distribution assumption, we have:

$$H_Q(x) = x, \tag{I4}$$

$$H_Q^{-1}(x) = x, \tag{I5}$$

$$f(x) = H_P(x). \tag{I6}$$