A Discrete Model for Histogram Shaping

S.M.Tabirca I.Pitt D.Murphy
Department of Computer Science
National University of Ireland, Cork
s.tabirca@cs.ucc.ie

Abstract

The aim of this article is to present a discrete model for histogram shaping. This is an important image transformation with several practical applications. The model that is proposed is based on a generalization of the inferior part function. Finally, an algorithm based on this model is developed.

Key Words: histogram, histogram shaping, discrete random variable.

1 Introduction

Histogram equalization or histogram flattening is one of the most important nonlinear point operations. This transformation aims to distribute uniformly the gray levels of the input image such that the histogram of the output image is flat. Histogram equalization has been studied for many years (see [1], [3], [4]) and many practical applications have been proposed so far. A direct generalization of this transformation is represented by histogram shaping or histogram specification (see [1], [3]). The idea of histogram shaping is to transform the input image into one which has histogram of a specific shape. Obviously, when the output shape is flat, histogram equalization is obtained. Both histogram equalization and histogram shaping have become classical image transformations, therefore it has been quite difficult to find the initial reference source. One of the earliest references about is [2].

The mathematical model of histogram equalization and shaping is based on stochastic approach. Let us consider that the input digital image is $f = (f_{i,j} :\ i = 1, 2, ..., n;\ j = 1, 2, ..., m)$ where

$$1 \le f_{i,j} \le G$$

represents the gray value of pixel (i, j). The probability or frequency of gray level $k \in 1, ..., G$ is defined by

$$p_f(k) = \frac{\#\{(i,j) :\ f(i,j) = k\}}{m \cdot n},\ k = 1, ..., G, \tag{1}$$

where $\#\{(i,j) :\ f(i,j) = k\}$ gives the number of pixel with the gray level equal to k. Based on these probabilities, the digital image f can be considered a discrete random variable

$$p_f = \begin{pmatrix} 1 & 2 & ... & G \\ p_f(1) & p_f(2) & ... & p_f(G) \end{pmatrix} \tag{2}$$

for which $\sum_{k=1}^{G} p_f(k) = 1$. Recall that the cumulative probability distribution of p_f is

$$P_f : \{1,2,...,G\} \to [0,1],\ P_f(k) = \sum_{l=0}^{k} p_f(l). \tag{3}$$

A more productive approach is to consider the digital image as a continuous random variable $p_f : [0,\infty) \to [0,1]$ such that $\int_0^\infty p_f(x)dx = 1$. In this case the cumulative probability distribution is

$$P_f : [0,\infty) \to [0,1),\ P_f(x) = \int_0^x p_f(t)dt.$$

Based on this continuous model the histogram shaping transformation can be defined more easily. Consider that the input digital image f is transformed such that the histogram of the output image g has a shape given by the cumulative distribution $Q : [0,\infty) \to [0,1)$. The equation that gives histogram shaping is [3]

$$g = Q^{-1}(P_f(f)). \tag{4}$$

The main inconvenience arising form Equation(4) is represented by the inverse function Q^{-1}. Firstly, because the calculation of Q^{-1} might not be easier even for simple shapes. Secondly, we cannot define Q^{-1} for the discrete case therefore it would be difficult to apply (4) to a discrete computation. Perhaps, this is the real reason for seeing no discrete models for histogram shaping. In the following we will propose a discrete model for this transformation.

2 The Superior Smarandache f-Part

In order to propose an equation for the discrete case, we have to find a substitute for Q^{-1}. This is given by the Superior Smarandache f-Part , which represents a direct generalization of the classical ceiling function. Smarandache proposed [5] a generalization of the ceiling function as following. Consider that $f : Z \to R$ an increasing function such that $lim_{n\to-\infty} f(n) = -\infty$ and $lim_{n\to\infty} f(n) = \infty$. The Superior Smarandache f-Part associated with f is $f^{\lceil\rceil} : R \to Z$ defined by

$$f^{\lceil\rceil}(x) = k \Leftrightarrow f(k-1) < x \le f(k).$$

Smarandache studied this function in relation to some functions of Number Theory and proposed several conjectures on them [6]. Tabirca also studied the Superior Smarandache f-Part [7] when $f(n) = \sum_{i=0}^{n} i^a$ and proposed equations for $f^{\lceil\rceil}$ when $a = 0,1,2$. Tabirca also applied this function to static parallel loop scheduling [8].

Now, we propose a version of the Superior Smarandache f-Part for our discrete case. Consider that $f : \{1,2,...,G\} \to (0,1]$ is an increasing function such that $f(G) = 1$. We also consider that this function is extended to 0 with $f(0) = 0$. The Superior Smarandache f-Part associated with f is $f^{\lceil\rceil} : (0,1] \to \{1,...,G\}$ defined by

$$f^{\lceil\rceil}(x) = k \Leftrightarrow f(k-1) < x \le f(k),\ \forall x \in [0,1]. \tag{5}$$

This function is also extended in 0 by $f^{\lceil\rceil}(0) = 0$.

Some properties of the function $f^{\lceil\rceil}$ are proposed in the following.

Theorem 1

$$f^{\lceil\rceil}(f(k)) = k,\ \forall k \in \{1,2,...,G\}. \tag{6}$$

Proof The proof is based on the definition of $f^{\square}$ and on the double inequality

$$f(k-1) < f(k) \leq f(k).$$

$\spadesuit$

Theorem 2

$$x \leq f\left(f^{\square}(x)\right) < x + \sup_{k}\left(f(k+1) - f(k)\right), \ \forall x \in (0,1]. \tag{7}$$

Proof

Let us denote $k = f^{\square}(x)$. The definition of $f^{\square}$ provides $f(k-1) < x \leq f(k)$. From this equation, it directly follows that $x \leq f(f^{\square}(x))$.
The second part of Equation (7) comes from the following implication:

$$f(k) < f(k) + x - f(k-1) \Rightarrow$$

$$f\left(f^{\square}(x)\right) < x + \sup_{k}\left(f(k) - f(k-1)\right).$$

$\spadesuit$

Based on these properties, the histogram shaping model of the discrete case is proposed.

3 Histogram Shaping for the Discrete Case

Consider that the input image $f = (f_{i,j} : \ i = 1,2,...,n; \ j = 1,2,...,m)$ is transformed into the output image $g = (g_{i,j} : \ i = 1,2,...,n; \ j = 1,2,...,m)$ such that the histogram of g has a certain shape. Let us presume that the shape of the output histogram is given by the discrete random variable

$$p_h = \begin{pmatrix} 1 & 2 & ... & G \\ p_h(1) & p_h(2) & ... & p_h(G) \end{pmatrix} \tag{8}$$

where $\sum_{k=1}^{G} p_h(k) = 1$.

The general equation of histogram shaping is similar with Equation (4) but $P_h^{\square}$ is used in place of P_h^{-1}. Let consider that the equation of image g is

$$g(i,j) = P_h^{\square}\left(P_f(f(i,j))\right), \ \forall \ (i,j) \in \{1,...,n\} \times \{1,2,...,m\}. \tag{9}$$

We prove that the cumulative probability distribution of g is very close to the cumulative probability distribution of h.

Theorem 3

$$P_g(k) = P_f\left(P_f^{\square}\left(P_h(k)\right)\right), \ \forall k \in \{1,2,...,G\}. \tag{10}$$

Proof The proof is given by the following transformations:

$$P_g(k) = \Pr\left[g(i,j) \leq k\right] = \sum_{l=1}^{k} \Pr\left[P_h^{\square}\left(P_f(f(i,j))\right) = l\right] =$$

$$= \sum_{l=1}^{k} \Pr\left[P_h(l-1) < P_f(f(i,j)) \le P_h(l)\right] =$$

$$= \sum_{l=1}^{k} \Pr\left[P_f^{[]}(P_h(l-1)) < P_f^{[]}(P_f(f(i,j))) \le P_f^{[]}(P_h(l))\right] =$$

$$= \sum_{l=1}^{k} \Pr\left[\{f(i,j) \le P_f^{[]}(P_h(l))\} - \{f(i,j) \le P_f^{[]}(P_h(l-1))\}\right] =$$

$$= \sum_{l=1}^{k} \left(\Pr\left[f(i,j) \le P_f^{[]}(P_h(l))\right] - \Pr\left[f(i,j) \le P_f^{[]}(P_h(l-1))\right]\right) =$$

$$= \sum_{l=1}^{k} \left(P_f\left(P_f^{[]}(P_h(l))\right) - P_f\left(P_f^{[]}(P_h(l-1))\right)\right) =$$

$$= P_f\left(P_f^{[]}(P_h(k))\right) - P_f\left(P_f^{[]}(P_h(0))\right)$$

Since $P_f\left(P_f^{[]}(P_h(0))\right) = 0$ we find that $P_g(k) = P_f\left(P_f^{[]}(P_h(k))\right)$ holds.
♠

From Theorems 2 and 3 the following equation is directly obtained.

$$P_h(k) \le P_g(k) < P_h(k) + \sup_j (P_f(j+1) - P_f(j)), \ \forall k \in \{0, 1, ..., G-1\}. \tag{11}$$

Equation (11) provides an estimation of the gap between the quantities $P_h(k)$ and $P_g(k)$. When $\sup_j (P_f(j+1) - P_f(j))$ is smaller these two quantities are very close. Although Equation (11) does not give a perfect equality we can say that the histogram of the image g has the shape very close to h.

The algorithm based on this model firstly finds the functions P_f, P_h and $P_h^{[]}$. Secondly, Equation (9) is applied to obtain the value $g(i,j)$ for each pair (i,j). A full description of this algorithm is presented below.

```
Inputs:
       n,m - the image sizes.
       f=(f[i,j]: i=1,...,n; j=1,...,m) - the input image.
       p_h=(p_h[i]: i=1,...,G) - the desired shape.
Output:
       g=(g[i,j]: i=1,...,n; j=1,...,m) - the input image.
double P_h(int k){
       double s=0;
       if(k<=0 || k>G) return 0;
       for(int i=1;i<=k;i++)s=s+p_h[i];
       return s;
}
int P_h_Inv(double x){
       int k;
       if(x<=0 || x>1) return 0;
```

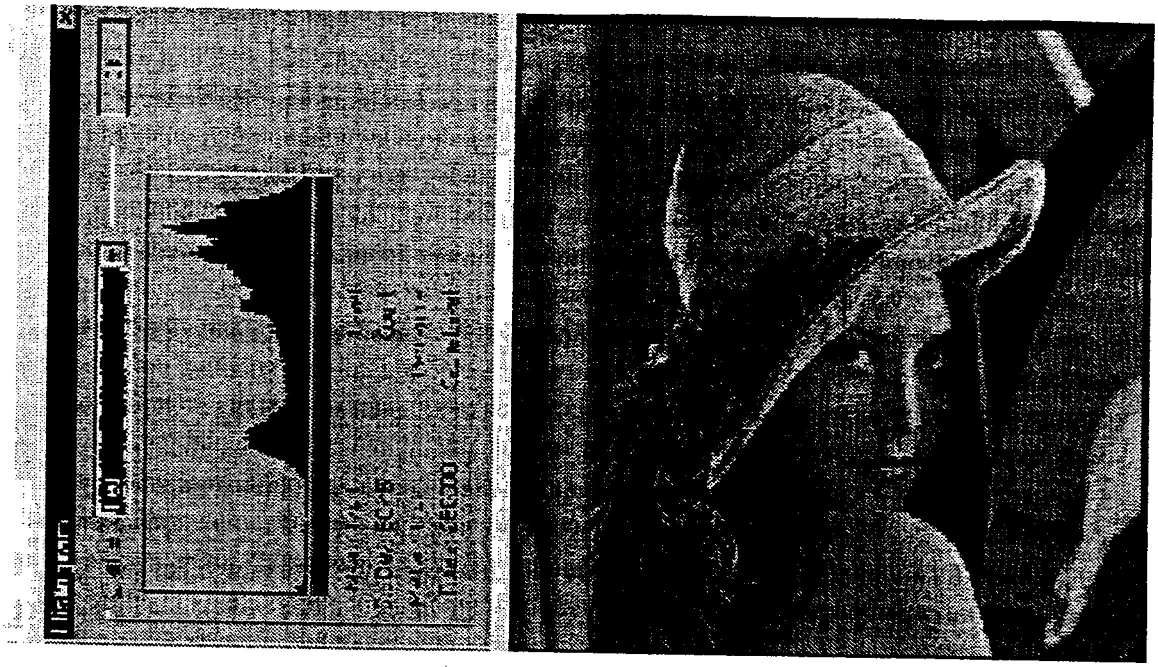

Figure 1: Lenna's picture 1.

```
        for(k=0;P_h(k)<=x;k++);
        return k-1;
}
double P_f(int k){
        double s=0, p_f[G];
        if(k<=0 || k>G) return 0;
        for(int i=1;i<=G;i++)p_f[i]=0;
        for(int i=1;i<=n;i++)for(int j=1;j<=m;j++)p_f[f[i,j]]=p_f[f[i,j]]+1./(n*m);
        for(int i=1;i<=k;i++)s=s+p_h[i];
        return s;
}
int [][] Hist_Shaping(int n, int m, int f[][]){
        int g[n][m];
        for(int i=1;i<=n;i++)for(int j=1;j<=m;j++)
           g[i,j] = P_h_Inv(P_f(f[i,j]));
        return g;
}
```

In order to show that the algorithm performs well we consider an example presented in [1]. Histogram shaping can be used to compare two images of the same scene, which have been taken under different lighting conditions. When the histogram of the first image is shaped to match in the histogram of the second image, the lighting effects might be eliminated.

Consider that we have the images presented Figures 1 and 2. They are two different Lenna's images where the second one has a poor lighting. Each image also contains the histogram for the red channel. The histogram shaping algorithm was applied to transform the second image

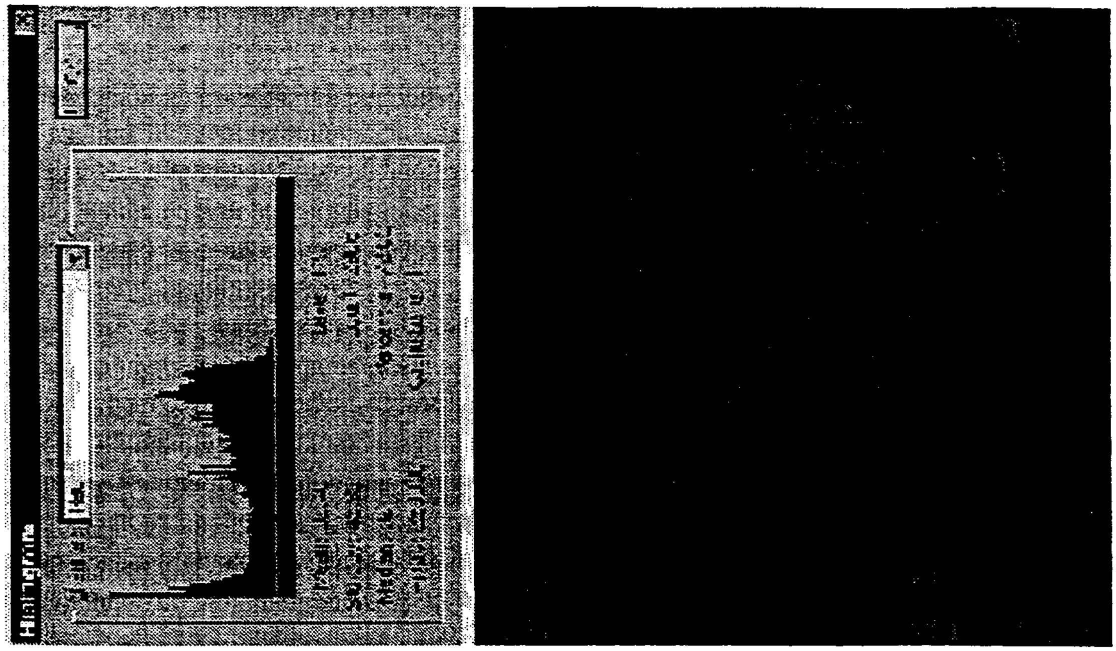

Figure 2: Lenna's picture 2.

according to the histogram of the first image. Figure 3 shows the resulting image which is the same as the first image. Moreover, the histograms of the first and third images are very alike with similar positions for peaks and valleys.

4 Conclusions

This article has introduced a discrete model for the histogram shaping transformation. The model that has been proposed uses the Smarandache ceiling function and is based on the equation $g = P_h^{\square}(P_f(f))$. A example has been also presented in order to prove that the method is viable.

References

[1] Bovik, A. (2000) *Handbook of Image and Video Processing*, Academic Press, London, UK.

[2] Humel, R. (1977) Image Enhancement by Histogram Transformation, *Comput. Graphics Image Process*, 6, pp 184-195.

[3] Jain, A. (1989) *Fundamental of Digital Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA.

[4] Petrou, M. and Bosdogianni, P. (1999) *Image Processing - The Fundamentals*, Wiley, UK.

[5] The Web Page of Smarandache Notions Journal, www.gallup.unm.edu/ smarandache.

[6] Smarandache, F. (1992) *Only Problems ... not Solutions*, Xiquan Publishing House, Phoenix, USA.

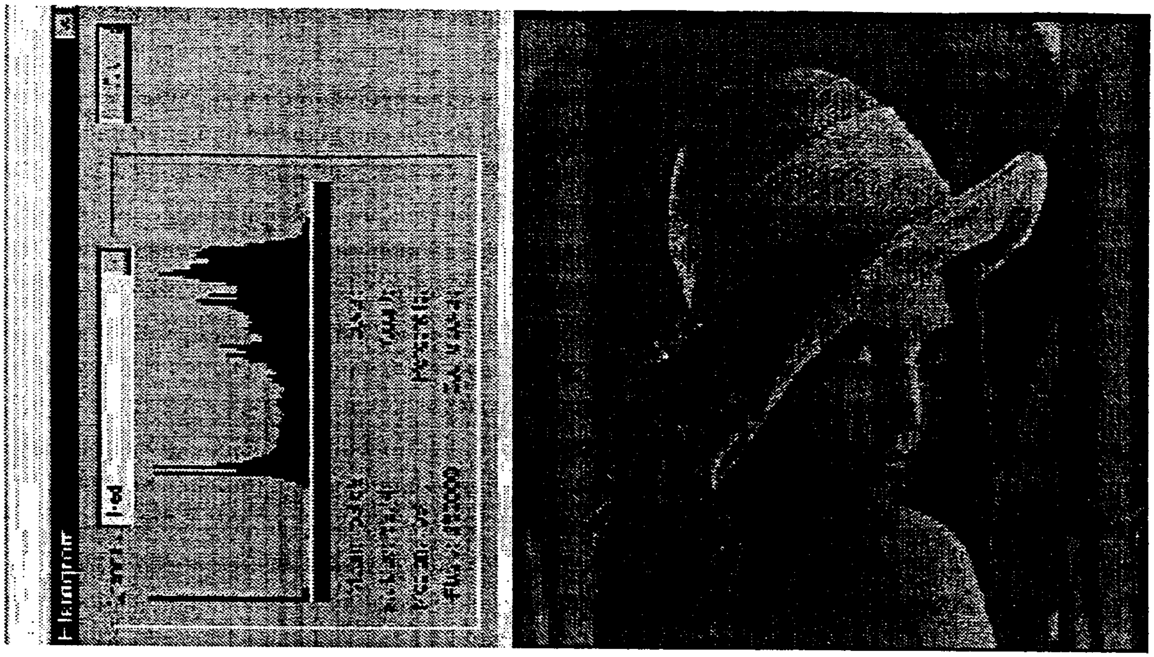

Figure 3: Lenna's Picture after Histogram Shaping.

[7] Tabirca, T. and Tabirca, S. (2001) A Parallel Loop Scheduling Based on the Smarandache f-inferior Part Function, *Smarandache Notions Journal*, 12, pp.28-35.

[8] Tabirca, T. and Tabirca, S. (2001) A New Equation for The Loop Balanced Scheduling Based on the Smarandache f-inferior Part Function, *Proceedings of the Second International Conference on Smarandache Notions Journal* [to appear].