COMPUTER GRAPHICS AND IMAGE PROCESSING (1975) 4, (209–224)

Histogram Modification Techniques*

ROBERT A. HUMMEL

University of Maryland, College Park, Maryland 20742

Communicated by A. Rosenfeld

Received November 15, 1974

The theory of histogram modification of continuous real-valued pictures is developed. It is shown that the transformation of gray levels taking a picture's histogram to a desired histogram is unique under the constraint that the transformation be monotonic increasing. Algorithms for implementing this solution on digital pictures are discussed. A gray-level transformation is useful for increasing visual contrast, but may destroy some of the information content. It is shown that solutions to the problem of minimizing the sum of the information loss and the histogram discrepancy are solutions to certain differential equations, which can be solved numerically.

1. INTRODUCTION

Every random variable f has an associated distribution function $P_f(z)$ which specifies the probability that the value of the random variable will be less than or equal to z. By a transformation we mean a one-to-one function T mapping the real numbers (or the range of f) into the real numbers. The composition $T \circ f$ is a new random variable (obtained by applying the transformation T to each value of f) having distribution function $P_{T\circ f}(z)$. In this paper, we are concerned with the problem of finding a transformation T such that the composition $T \circ f$ has some specified distribution function $P_g(z)$. Thus we are given f and P_g and desire a T such that $P_{T\circ f}(z)$ is in some sense close to $P_g(z)$. Later, we will add additional constraints to T so that the solution is a compromise between (1) the degree of match between $P_{T\circ f}(z)$ and $P_g(z)$ and (2) some measure of the cost (in terms of information loss) of transforming f by T.

Although the results in this paper are valid for a general random variable f, most of our examples for motivation and implementation will regard $f = f(x,y)$ as a gray-level-valued continuous picture function. The stated problem is then equivalent to the problem of modifying a picture's histogram by applying a uniform gray-level transformation to the picture. This type of transformation can be useful in picture processing as a method of contrast enhancement or as a preprocessing step for normalizing a set of pictures. This normalization technique has been shown to be useful in texture analysis [1], and in creating photomosaics [2].

Treatments of problems related to histogram modification fall into three categories: (i) transforming a continuous random variable to a continuous random variable; (ii) transforming a continuous random variable to a discrete (quantized)

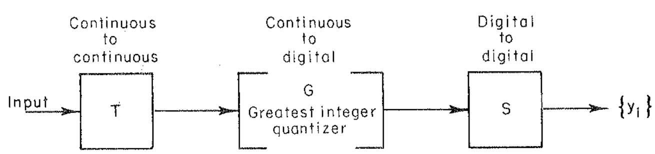
* Supported by U. S. Air Force Office of Scientific Research under Contract F-44620-72C-0062.

random variable; and (iii) transforming a discrete random variable to a discrete random variable. This paper is primarily concerned with the first case. The other two cases are regarded as approximations to the continuous case; thus the results obtained here are applicable to those cases by using standard approximation methods.

2. HISTORY

The continuous-to-continuous case belongs to the theory of distribution transformations in elementary statistics [3]. For example, it is well known that if we regard a distribution function P_f as a mapping from the reals into the reals, and take the composition of this mapping with f, we get a random variable with uniform density. One often wants to solve for the probability distribution when a known function is applied to one or more random variables. In histogram modification, we treat the opposite problem where we solve for the function given the desired final distribution.

Most of the work in the continuous-to-discrete case comes under the heading of optimal quantization. In this field, one is given a continuous picture with a density function $p(x)$, and one seeks a partition of the domain of $p(x)$ into intervals $[x_i,x_{i+1})$ such that the quantizer Q which assigns each interval to an output level y_i minimizes some measure of the quantization error. Equivalently, one seeks transformations T and S such that the composition shown below yields an optimal quantizer. In this scheme, T determines the x_i's, and S determines the y_i's.



It is easy to show that the quantizer which produces a uniform output density (a "flat histogram") maximizes the expected information content, i.e., entropy. It turns out that this is very close to optimal quantization.

In 1964, Roe showed that the quantizer that minimizes the distortion

$$D_r(Q) = \sum_{k=0}^{N-1} \int_{x_k}^{x_{k+1}} |x - y_k|^r p(x)\,dx$$

in the limit as $N \to \infty$ satisfies the approximation

$$\int_0^{x_n} [p(x)]^{1/(r+1)}dx = C_1 n + C_2,$$

where r is a positive constant [4]. His work therefore yields an approximate method for solving the optimal quantization problem as given by Max [5], and rediscovers approximations to the total quantization error given by Panter and Dite [6]. In 1966, Algazi [7] used Roe's results to solve for the optimal transformation T in the scheme shown above:

$$T(z) = (N-1) \int_0^3 [p(x)]^{1/(r+1)}dx \Big/ \int_0^X [p(x)]^{1/(r+1)}dx,$$

where $p(x)$ has support contained in $[0,X]$. In the limit as $r \to 0$, this is essentially the histogram flattening transformation given in this paper. For general $r > 0$, this transformation will broaden histograms by spacing quantization levels close together in those regions where $p(x)$ is large. Most authors specialize to the case $r = 2$, however, rather than examine the limit $r \to 0$.

Elias [8] suggested a different distortion measure,

$$M_r(Q) = \left\{ \sum_{k=0}^{N-1} (x_{k+1} - x_k)^r \int_{x_k}^{x_{k+1}} p(x)\,dx \right\}^{1/r},$$

which in some sense measures the ambiguity in a quantized picture obtained by application of the quantizer Q. Note that this measure no longer depends on the y_i's, and thus depends only on the choice of the transformation T. Interestingly enough, one can show by methods similar to those used by Max that when $r = 1$ this measure yields histogram flattening (sometimes called the equal probability quantizer) in the limit as $K \to \infty$ for continuous histograms. Elias also shows that the equal probability quantizer is asymptotically optimal as $K \to \infty$ for the limit $r \to 0$.

If one could characterize the desired output histogram in terms of the initial distribution, optimal quantization would become equivalent to the problem of histogram modification. This viewpoint, however, is generally less convenient, although the relationship between error minimization and entropy maximization deserves investigation.

Discrete-to-discrete transformations are useful for computer processing of digitized pictures. Most gray-level manipulation studies for contrast enhancement have involved photometric corrections and simple linear rescaling, as at JPL [9], or logarithmic rescaling (density coding) as described by Stockham [10]. "Maximum discriminability" versions of Mariner pictures were produced at JPL by logarithmic stretching about the mean gray value, which produces generally broader and flatter histograms [11]. For the results of a cyclic linear rescaling technique which divides the gray scale range into several equal length intervals, and maps each interval linearly onto the entire available gray scale range, see Selzer [12].

The concept of histogram modification, as defined here, has received relatively little attention. In 1971, Hall *et al.* [13] mentioned distribution linearization (histogram flattening) as an effective contrast enhancement technique, and Haralick, Shanmugan, and Dinstein [14] describe an algorithm for obtaining the equal probability quantizer. Both of these references regard the continuous-to-continuous case as being approximated by the discrete-to-discrete case.

Documentation for a gray-level transformation program designed for histogram modification of digitized pictures is provided by Troy [7,15]. Her program regards histogram modification as an iteration process on constant gray-level subsets of the picture, which is inefficient for sequential processing computers. The program HISTRN written for this report performs the same function as the program GRATRN described in [15], but is considerably more efficient because it defines the transformation prior to operation on the input picture, and is based on the discrete approximations to the continuous case theory developed in this paper. This increased efficiency makes histogram modification by HISTRN an inexpensive and powerful image enhancement technique.

3. PROBLEM FORMULATION

Before developing the theory of histogram modification, we need to make certain assumptions in order to accurately characterize the histogram function. Let $f(x,y)$ be a continuous real-valued function defined on a closed rectangle S of the x–y plane. The cumulative probability distribution is defined by

$$P_f(z) = \mu\{(x,y) \in S \mid f(x,y) \leqslant z\},$$

where μ is the Lebesque measure normalized so that $\mu\{S\} = 1$. We note that f is measurable since it is continuous, and has bounded range since S is compact. We will assume that the range of f is contained in the interval [0,a]. The continuous analog of the picture histogram is the probability density function given by

$$p_f(z) = (d/dz)(P_f(z))$$

at all points where $P_f(z)$ is differentiable. Of course, $P_f(z)$ will be discontinuous at points where

$$\mu\{(x,y) \in S \mid f(x,y) = z\} > 0,$$

but we know that since P_f is monotonic, it is differentiable almost everywhere. Nonetheless, because of the possibility of these jumps, $p_f(z)$ does not satisfy the highly desirable property $\int_0^a p_f(z)\,dz = 1$. One could avoid these difficulties by using the Stieltjes integral whenever integration involving a histogram factor is needed. That is, instead of $\int q(z)p_f(z)\,dz$, one can use $\int q(z)\,dP_f(z)$, where $q(z)$ is any continuous function. In this report, we will assume that all picture cumulative probability functions are absolutely continuous, so that

$$P_f(z) = \int_0^z p_f(z')\,dz'.$$

This is a reasonable assumption since, in the quantization process, all gray levels in the intervals $[z_i,z_{i+1}]$ will be identified. Thus we may adjust f so that gray levels corresponding to jumps in the probability function will be broken up and assigned to a range of gray levels within the quantization interval.

The general problem of histogram modification (for continuous pictures) may be stated as follows. We seek a gray-level transformation T which maps the interval $[0,a]$ into itself such that the output picture obtained by applying T to the value of each point of the input picture f has some specified histogram $p_g(z)$. Thus we are given f and p_g, and desire a T such that

$$p_{T\circ f}(z) = P_g(z), \qquad \text{for all } z \in [0,a], \tag{1}$$

or equivalently,

$$P_{T\circ f}(z) = P_g(z), \qquad \text{for all } z \in [0,a].$$

Furthermore, since we wish T to preserve an ordering relationship on gray levels, we require that T be either monotonic increasing or monotonic decreasing.

Of course, there is no a priori reason to believe that such a T exists. However, under suitable conditions on f, p_g, and T, we will soon prove an existence and uniqueness theorem for solutions to (1). More generally, we can seek a transformation which minimizes the difference between $p_{T\circ f}(z)$ and $p_g(z)$. Our measure of difference will be

$$\int_0^a [p_{T\circ f}(z) - p_g(z)]^2 dz. \tag{2}$$

This suggests the following refinement to the problem of histogram modification. We add to (2) a cost function $C_f(T)$ measuring, in some way, the loss of information in transforming from f to $T \circ f$. Thus we wish to minimize

$$c_1 C_f(T) + c_2 \int_0^a [p_{T\circ f}(z) - p_g(z)]^2 dz \tag{3}$$

for T, where c_1 and c_2 are weighting constants.

In the next section, we treat the case $c_1 = 0$, that is, the problem of minimizing (2). Later, we will suggest two models for $C_f(T)$ and demonstrate techniques for minimizing (3).

4. MINIMIZING THE HISTOGRAM ERROR

We first prove the theorem which solves the histogram modification problem.

THEOREM. Let $f(x,y)$ be a continuous real-valued random variable with range $[0,a]$ and having an absolutely continuous probability distribution $P_f(z)$. Let $p_f(z)$ be the corresponding probability density function, and suppose that p_g is given satisfying

(i) p_g: $[0,a] \to (0,\infty)$,
(ii) $\int_0^a p_g(z)\,dz = 1$.

Denote by $P_g(z)$ the function given by

$$P_g(z) = \int_0^z p_g(z')\,dz'.$$

Then there exist unique monotonic transformations taking f to a random variable having a density function equal to $p_g(z)$. These are given by

$$T(z) = P_g^{-1}(P_f(z)), \tag{4a}$$

$$T(z) = P_g^{-1}(1 - P_f(z)) \tag{4b}$$

for the monotonic increasing and decreasing transformations, respectively.

PROOF. Since P_g is absolutely continuous, we only consider transformations which take the picture f into pictures having absolutely continuous probability functions. For any such T, we have

$$\begin{aligned} P_f(z) &= \mu\{(x,y) \in S \mid f(x,y) \in [0,z]\} \\ &= \mu\{(z,y) \in S \mid T \circ f(x,y) \in T\{[0,z]\}\} \\ &= \int_{T\{[0,z]\}} p_{T\circ f}(z)\,dz, \end{aligned} \tag{5}$$

where the integral is taken in the Lebesgue sense over the set $T\{[0,z]\}$. If T is continuous, two cases arise:

$$T\{[0,z]\} = \begin{cases} [T(0),T(z)] & \text{for } T \text{ monotonic increasing,} \\ [T(z),T(0)] & \text{for } T \text{ monotonic decreasing.} \end{cases}$$

Treating the two cases separately, we have for T increasing,

$$P_f(z) = \int_{T(0)}^{T(z)} p_{T\circ f}(z')\,dz' = \text{const}_1 + P_{T\circ f}(T(z)), \tag{6a}$$

and for T decreasing,

$$P_f(z) = \int_{T(z)}^{T(0)} p_{T\circ f}(z')\,dz' = \text{const}_2 - P_{T\circ f}(T(z)). \tag{6b}$$

Using the facts that $P_f(0) = 0$, $P_f(a) = 1$, and $0 \leq P_{T\circ f}(T(z)) \leq 1$ for all z, one can show that $\text{const}_1 = 0$ and $\text{const}_2 = 1$. Thus

$$P_f(z) = P_{T\circ f}(T(z)) \tag{7a}$$

and

$$P_f(z) = 1 - P_{T\circ f}(T(z)). \tag{7b}$$

Equations (7a) and (7b) are valid even if T has points of discontinuity, since $p_{T\circ f}(z)$ will be zero on the intervals in the complement of the range of T corresponding to the jumps of the function T.

To prove the theorem, we must show two things: first, that solutions (4a) and (4b) solve the histogram modification problem (1); second, that any suitable transformation satisfying (1) will satisfy either (4a) or (4b), thus establishing uniqueness.

Note that P_g is strictly increasing on the interval $[0,a]$ (since $p_g(z) > 0$), and thus has a monotonic inverse. Thus the transformation (4a) is monotonic increasing, while (4b) is decreasing. Substituting (7a) into (4a), we obtain $P_{T\circ f}(z) = P_g(z)$ for all z in the range of T. Since both P_f and P_g^{-1} are continuous, T as given by (4a) maps onto the interval $[0,a]$. Thus (4a) satisfies Eq. (1). Similarly, substituting (7b) into (4b) yields the same result.

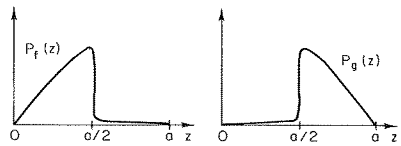
Conversely, if T is a monotonic transformation satisfying Eq. (1), then we can substitute $P_{T\circ f}$ by P_g in the proper form of Eq. (7) (depending on whether T is increasing or decreasing), and solve for T to obtain the respective solution (4a) or (4b). ■

As a corollary to the theorem, we may differentiate Eqs. (7a) and (7b) to get the differential equation form

$$p_f(z) = p_{T\circ f}(T(z)) \cdot |T'(z)|, \tag{8}$$

a formula which will be useful later. The theorem holds even if we allow $p_g(z) = 0$ for some z in the interval $[0,a]$. In this case, however, we must interpret P_g^{-1} as a choice function in order that T be well-defined.

Solution (4b) applied to a picture f is the same as solution (4a) for the "complement" of f defined by $\bar{f}(x,y) = a - f(x,y)$. Note, however, that solution (4b) is not simply the complement of the picture obtained from solution (4a) if the distribution P_g is not linear. In some circumstances, solution (4b) may be preferable to (4a). For example, if one desires the solution T which yields the smaller value of $\int\int_S (T \circ f - f)^2 dxdy$, then solution (4b) will be the appropriate one in certain cases, such as when p_f and p_g are as shown below.



If, however, one desires a transformation which preserves the relationship "is lighter than" on gray levels, then solution (4a) will always be the appropriate one.

5. IMPLEMENTATION

For application to the discrete case, we must approximate the continuous transformation T_c by a mapping T defined on the quantization levels $\{y_0, \ldots, y_{N-1}\}$. Each quantization level y_i represents the gray-level interval $[x_i, x_{i+1})$; thus the discrete histogram of the quantized version, $\hat{f}$, of the picture f is given by

$$p_{\hat{f}}(y_j) = \int_{x_j}^{x_{j+1}} p_f(z)\, dz.$$

Usually, however, we have available only the quantized picture $\hat{f}$ and its histogram $p_{\hat{f}}$, and can only approximate the original density function by assuming it to be constant on the quantization intervals. Thus

$$p_f(y) = p_{\hat{f}}(y_j)/(x_{j+1} - x_j) \text{ for } y \in [x_{,j}, x_{j+1}). \tag{9}$$

This gives the approximate distribution function

$$P_f(y_j) = \frac{y_j - x_j}{x_{j+1} - x_j} p_{\hat{f}}(y_j) + \sum_{i=0}^{j-1} p_{\hat{f}}(y_i).$$

Using a similar model for the desired output distribution P_g, we obtain $T_c(y_j) = P_g^{-1}(P_f(y_j))$ from Eq. (4a). The desired discrete mapping may be obtained from T_c by quantizing the images of the levels $y_1, \ldots, y_N$, or, equivalently, by interpreting P_g^{-1} as a step function which assigns all the values between $P_g(x_j)$ and $P_g(x_{j+1})$ to the level y_j.

Often the actual values of the quantization levels and their corresponding endpoints are not available, and only a digitized version of the quantized picture is supplied. Thus each quantization level y_j is represented by its index j, and the domain of $\hat{p}_{\hat{f}}$ is the set of integers $0,1, \ldots, N-1$. The values $(y_j - x_j)/(x_{j+1} - x_j)$ may be approximated by using Max's solution [5] and the approximation (9):

$$y_j = \frac{\int_{x_j}^{x_{j+1}} y p_f(y)\, dy}{\int_{x_j}^{x_{j+1}} p_f(y)\, dy} = \frac{\int_{x_j}^{x_{j+1}} y dy}{\int_{x_j}^{x_{j+1}} dy} = \frac{x_j + x_{j+1}}{2},$$

so $\quad (y_j - x_j/x_{j+1} - x_j) = 1/2, \qquad j = 0,1, \ldots, N-1.$

Thus the discrete histogram modification transformation for quantized pictures is given by

$$T(k) = P_g^{-1}\Big(\tfrac{1}{2}\,\hat{p}_f(k) + \sum_{m=0}^{k-1} \hat{p}_f(m)\Big), \tag{10}$$

where P_g^{-1} is a step function which assigns all the values between $P_g(k)$ $(= \Sigma^{k-1}{}_{m=0}\,\hat{p}_{\hat{g}}(m))$ and $P_g(k+1)$ to the value k. For the particular case of histogram flattening, Eq. (10) becomes

$$T(k) = \Big[(N-1)\cdot\Big(\tfrac{1}{2}\,\hat{p}_{\hat{f}}(k) + \sum_{m=0}^{k-1} \hat{p}_{\hat{f}}(m)\Big)\Big]. \tag{11}$$

Here $[\cdot]$ denotes the greatest integer function.

Equations (10) and (11) may be regarded as the standard bin-filling technique of [13–15]. However, this viewpoint tends to incorrectly identify histogram modification as an iterative process involving repeated passes through the input picture. In our approach, on the other hand, once the histogram of the input picture is determined, Eq. (10) or (11) may be used to define the transformation on the entire domain of gray levels, which, in turn, may be applied to the entire input picture during a single pass.

Equation (11) is particularly easy to implement, but prone to give only very crude approximations to flat histograms. The problem is inherent in the quantization of pictures to a finite number of gray levels. As long as we insist on assigning the entire set of points having gray level z_j to the single level $T(z_j)$, we can never split up a constant gray-level set, and can only merge such sets together. Thus the net information will at best remain the same, and the improvement, if any, will be purely cosmetic.

In order to include information about the picture objects, and to improve the contrast enhancement effects, the transformation T must be made context sensitive. The choice of context will depend upon the desired effect. In general, then, T will become a function of two or more variables, where one variable corresponds to the gray level, and the others to local context values. In this way, each set $\{f(x,y) = z_j\}$ is broken up into disjoint subsets, based on the range of the expected context values. If we order the entire collection of subsets according to some rule depending on the gray-level and context values, then the net effect is to requantize f to more gray levels, thereby more closely approximating the continuous case. For example, suppose we compute a requantized picture

$$h(x,y) = f(x,y) + \alpha(w_f(x,y) - f(x,y)),$$

where α is a constant, and $w_f(x,y)$ is a local average function of the picture f (e.g., the average of the four or eight neighbors). We may then transform the requantized picture h by using Eq. (10) or (11), where the new histogram $p_h(z)$ is used rather than $p_f(z)$, and the quantization levels z_i have become the range of possible values of h. We will write

$$T(z,w) = T_0(z + \alpha(w - z)),$$

where T_0 is the previously described single-variable transformation applied to the picture h.

a. b. c.

FIG. 1. (a) Original terrain picture. (b) Result after executing FLATRN, CWT = 7. (c) Result after executing FLATRN with Fig. 1b as input, CWT = 15.

A program which implements this last technique for digital pictures has been written. This histogram flattening routine, known as FLATRN, first calculates a requantized version of the input picture by adding a large multiple (~ 56) of each point to its eight neighbors, and simultaneously prepares a histogram of the result. Equation (10) is then used to define a histogram flattening transformation on the requantized picture. The output picture is produced by applying this transformation (see Figs. 1 and 2).

It is suggested that other methods of adding context information in the definition of T might include (a) weighting the local average gray level inversely to the texture graininess t in a medium size neighborhood of the point in order to preserve edge and fine texture detail:

$$T(z,w,t) = T_0(z + (\alpha/t)(w - z)),$$

or (b), darkening a picture point if the average, u, over a large square centered at the point is lighter than the average gray level u_0 over the whole picture:

$$T(z,w,t,u) = T_0(z + \alpha/t)(w - z) - \beta(u - u_0)).$$

6. MINIMIZING THE INFORMATION LOSS

We now turn to the problem of minimizing the sum of the cost of transformation and the error (see Eq. (3)). We wish to transform f so that its histogram is closer to $p_g(z)$, but we shall balance the histogram error against some measure of the loss of information in transforming from f to $T \circ f$. Our treatment will demonstrate the advantage of the "continuous case" approach to histogram modification problems.

As a motivating example, suppose we are given two pictures taken under different lighting conditions, where each picture contains a subregion that corresponds to the same object or material. If we want to combine or compare the two pictures, we would like the histograms on the two subregions to match. One approach would be to transform both pictures so that the respective subregions have some suitably chosen standard histogram. If the subregions are small in relation to the pictures, however, these transformations may cause a substantial loss of information.

Let S_1 be the subregion of S, and let f_1 be the picture f restricted to S_1. We want to transform f so that the histogram $p_{T\circ f_1}(z)$ is close to $p_{g_1}(z)$. If $C_f(T)$ is the cost of transforming f by T, we wish to find a T that minimizes

$$c_1 C_f(T) + c_2 \int_0^a (p_{T\circ f_1}(z) - p_{g_1}(z))^2 dz. \qquad (3')$$

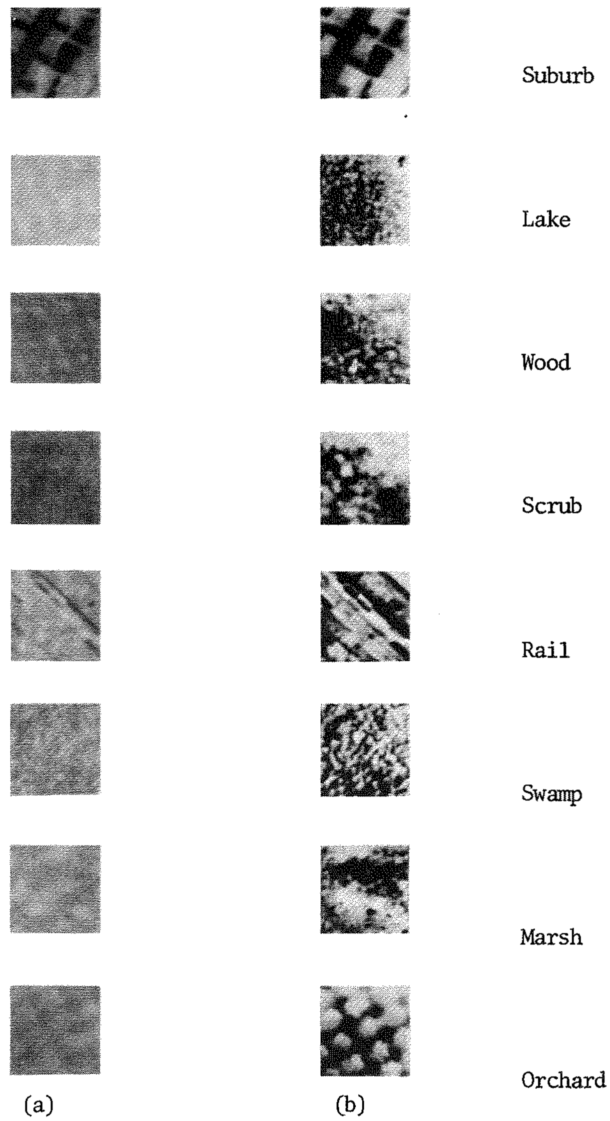


FIG. 2. (a) Original terrain pictures. (b) Results after executing FLATRN.

Now, $C_f(T)$ must be additive over the set of partitions of f, and the cost of mapping z to $T(z)$ will be some function of z, $T(z)$, and perhaps some derivatives of T. Thus $C_f(T)$ can be given by an expression of the form

$$\int_0^a C(z,T(z),T'(z))p_f(z)\,dz. \tag{12}$$

This integral contains the weighting factor $p_f(z)$ since this corresponds to the density of points having gray level z. In the remainder of this paper, we will assume that all the density functions are differentiable.

A simple model for the function $C_f(T)$ is the mean square difference between f and $T \circ f$. Thus in (12) we have $C(z,T(z)) = (z - T(z))^2$. A more sophisticated approach is to take $C(z,T'(z)) = 1/T'(z)$, thus weighting against compression of gray levels. Other approaches might use some information measure to model the loss of content in going from f to $T \circ f$.

To solve Eq. (3′) among the class of monotonic increasing transformations on $[0,a]$ fixing endpoints, we replace T by $T + \epsilon T_1$, where T_1 is a differentiable function on $[0,a]$ vanishing at the endpoints. We then differentiate with respect to ϵ, and set the result equal to 0. For the second integral in (3′), we use Eq. (8) to replace $p_{T\circ f_1}(z)$ by $p_{f_1}(T^{-1}(z))/T'(T^{-1}(z))$ and substitute $x = T^{-1}(z)$ before differentiating:

$$c_2 \int_0^a (p_{T\circ f_1}(z) - p_{g_1}(z))^2 dz = c_2 \int_0^a \left(\frac{p_{f_1}(T^{-1}(z))}{T'(T^{-1}(z))} - p_{g_1}(z)\right)^2 dz$$

$$= c_2 \int_0^a \left[\frac{p_{f_1}(x)}{T'(x)} - p_{g_1}(T(x))\right]^2 T'(x)\,dx.$$

We differentiate

$$c_2 \int_0^a \left[\frac{p_{f_1}(x)}{T'(x) + \epsilon T_1'(x)} - p_{g_1}(T(x) + \epsilon T_1(x))\right]^2 (T'(x) + \epsilon T_1'(x))\,dx$$

with respect to ϵ at $\epsilon = 0$ to obtain

$$c_2 \int_0^a 2\left[\frac{p_{f_1}(x)}{T'(x)} - p_{g_1}(T(x))\right]\left(-\frac{p_{f_1}(x)T_1'(x)}{(T'(x))^2} - p_{g_1}'(T(x))T_1(x)\right)T'(x)\,dx$$

$$+ c_2 \int_0^a \left[\frac{p_{f_1}(x)}{T'(x)} - p_{g_1}(T(x))\right]^2 T_1'(x)\,dx.$$

After factors of $T_1'(x)$ are collected, this becomes

$$2c_2 \int_0^a \left[p_{g_1}(T(x)) - \frac{p_{f_1}(x)}{T'(x)}\right](p_{g_1}'(T(x))T'(x))T_1(x)\,dx$$

$$+ c_2 \int_0^a \left[(p_{g_1}(T(x)))^2 - \left(\frac{p_{f_1}(x)}{T'(x)}\right)^2\right] T_1'(x)\,dx.$$

We may replace the expression $p_{f_1}(x)/T'(x)$ by $p_{T\circ f_1}(T(x))$, and integrate the second integral by parts, using the endpoint conditions $T_1(0) = 0$ and $T_1(a) = 0$ to obtain

$$2c_2 \int_0^a [p_{g_1}(T(x)) - p_{T\circ f_1}(T(x))](p_{g_1}'(T(x))T'(x))T_1(x)\,dx$$

$$- 2c_2 \int_0^a [p_{g_1}(T(x))p_{g_1}'(T(x))T'(x) - p_{T\circ f_1}(T(x))p_{T\circ f_1}'(T(x))T'(x)]T_1(x)\,dx.$$

After collecting terms, and using Eq. (8) once again, we have

$$2c_2 \int_0^a p_{f_1}(x)[p_{T\circ f_1}'(T(x)) - p_{g_1}'(T(x))]T_1(x)\,dx. \tag{13}$$

The first term in (3′) is treated in the same way, but, of course, must be

```
  Z      P(Z)
 --     -----
-2.94   .0116:*
-2.82   .0080:*
-2.70   .0109:*
-2.58   .0148:**
-2.46   .0199:**
-2.34   .0263:***
-2.22   .0345:*****
-2.10   .0445:******
-1.98   .0567:********
-1.86   .0713:**********
-1.74   .0884:*************
-1.62   .1080:****************
-1.50   .1301:*******************
-1.38   .1545:***********************
-1.26   .1809:***************************
-1.14   .2088:*******************************
-1.02   .2376:***********************************
 -.90   .2665:***************************************
 -.78   .2947:********************************************
 -.66   .3212:************************************************
 -.54   .3451:***************************************************
 -.42   .3655:******************************************************
 -.30   .3816:*********************************************************
 -.18   .3928:**********************************************************
 -.06   .3984:***********************************************************
  .06   .3984:***********************************************************
  .18   .3928:**********************************************************
  .30   .3816:*********************************************************
  .42   .3655:******************************************************
  .54   .3451:***************************************************
  .66   .3212:************************************************
  .78   .2947:********************************************
  .90   .2665:***************************************
 1.02   .2376:***********************************
 1.14   .2088:*******************************
 1.26   .1809:***************************
 1.38   .1545:***********************
 1.50   .1301:*******************
 1.62   .1080:****************
 1.74   .0884:*************
 1.86   .0713:**********
 1.98   .0567:********
 2.10   .0445:******
 2.22   .0345:*****
 2.34   .0263:***
 2.46   .0199:**
 2.58   .0148:**
 2.70   .0109:*
 2.82   .0080:*
 2.94   .0058:
```

FIG. 3a. Initial histogram.

handled separately for each function C. If $C(z,T(z)) = (z - T(z))^2$, after differentiating, we obtain

$$2c_1 \int_0^a p_f(z)\,(T(z) - z)\,T_1(z)\,dz. \tag{14a}$$

If $C(z,T'(z)) = 1/T'(z)$, then the calculation is

```
  Z      P(Z)
 --      -----
-2.94   .0021:
-2.82   .0025:
-2.70   .0052:
-2.58   .0108:*
-2.46   .0184:**
-2.34   .0298:****
-2.22   .0422:******
-2.10   .0540:********
-1.98   .0648:*********
-1.86   .0756:***********
-1.74   .0893:*************
-1.62   .1116:****************
-1.50   .2390:***********************************
-1.38   .2039:******************************
-1.26   .2343:***********************************
-1.14   .2642:***************************************
-1.02   .2868:******************************************
 -.90   .2052:******************************
 -.78   .2864:******************************************
 -.66   .3010:*********************************************
 -.54   .3120:**********************************************
 -.42   .3208:************************************************
 -.30   .3274:*************************************************
 -.18   .3317:*************************************************
 -.06   .3339:**************************************************
  .06   .3339:**************************************************
  .18   .3317:*************************************************
  .30   .3273:*************************************************
  .42   .3208:************************************************
  .54   .3120:**********************************************
  .66   .3010:*********************************************
  .78   .2864:******************************************
  .90   .2052:******************************
 1.02   .2868:******************************************
 1.14   .2642:***************************************
 1.26   .2343:***********************************
 1.38   .2039:******************************
 1.50   .2390:***********************************
 1.62   .1116:****************
 1.74   .0893:*************
 1.86   .0756:***********
 1.98   .0648:*********
 2.10   .0540:********
 2.22   .0422:******
 2.34   .0298:****
 2.46   .0184:**
 2.58   .0108:*
 2.70   .0077:*
 2.82   .0076:*
 2.94   .0098:*
```

FIG. 3b. Smoothed output histogram obtained by solving Eq. 15.

$$\frac{d}{d\epsilon} c_1 \int_0^a \frac{p_f(z)}{T'(z) + \epsilon T_1'(z)}\, dz \Big|_{\epsilon=0} = -c_1 \int_0^a \frac{T_1'(z)}{(T'(z))^2}\, p_f(z)\, dz$$

$$= c_1 \int_0^a \frac{d}{dz}\left(\frac{p_f(z)}{(T'(z))^2}\right) T_1(z)\, dz. \quad (14b)$$

For the T which minimizes (3′), the sum of (13) and (14a) or (14b) must be 0.

```
  Z     P(Z)
  --    -----
-3.35  .0015:
-3.16  .0027:
-2.97  .0049:
-2.77  .0085:*
-2.58  .0142:**
-2.39  .0228:***
-2.20  .0354:*****
-2.01  .0530:*******
-1.82  .0764:***********
-1.63  .1063:***************
-1.44  .1424:*********************
-1.24  .1840:***************************
-1.05  .2293:**********************************
 -.86  .2753:*****************************************
 -.67  .3188:***********************************************
 -.48  .3558:*****************************************************
 -.29  .3828:*********************************************************
 -.10  .3971:***********************************************************
  .10  .3971:***********************************************************
  .29  .3828:*********************************************************
  .48  .3558:*****************************************************
  .67  .3188:***********************************************
  .86  .2753:*****************************************
 1.05  .2293:**********************************
 1.24  .1840:***************************
 1.44  .1424:*********************
 1.63  .1063:***************
 1.82  .0764:***********
 2.01  .0530:*******
 2.20  .0354:*****
 2.39  .0228:***
 2.58  .0142:**
 2.77  .0085:*
 2.97  .0049:
 3.16  .0027:
 3.35  .0015:
```

FIG. 4a. Initial normal density function.

Thus

$$\int_0^a [2c_1 p_f(z)(T(z)-z) + 2c_2 p_{f_1}(z)(p'_{T\circ f_1}(T(z)) - p'_{g_1}(T(z)))]T_1(z)\,dz = 0,$$

$$\int_0^a \left[c_1 \frac{d}{dz}\left(\frac{p_f(z)}{(T'(z))^2}\right) + 2c_2 p_{f_1}(z)(p'_{T\circ f_1}(T(z)) - p'_{g_1}(T(z)))\right]T_1(z)\,dz = 0$$

for (14a) and (14b), respectively. In either case, equality holds for all continuous T_1 vanishing at endpoints, which is sufficient to imply that the bracketed expressions are identically 0 on the range $[0,a]$. This yields the differential equations

$$T(z) = z + (c_2/c_1)(p_{f_1}(z)/p_f(z))\ (p'_{g_1}(T(z)) - p'_{T\circ f_1}(T(z))) \qquad (15a)$$

and

$$\frac{p'_f(z)}{p_f(z)}\,T'(z) - 2T''(z) = (2c_2/c_1)(T'(z))^3 \cdot (p'_{g_1}(T(z)) - p'_{T\circ f_1}(T(z))), \qquad (15b)$$

respectively. Both equations may be solved numerically on a digital computer. For reasonably typical functions $p_f(z)$, $p_{f_1}(z)$, and $p_{g_1}(z)$, the desired T will not

```
  Z     P(Z)
 --     -----
-3.35  .0011:
-3.16  .0181:**
-2.97  .0343:*****
-2.77  .0527:*******
-2.58  .0722:**********
-2.39  .0926:*************
-2.20  .1124:****************
-2.01  .1315:*******************
-1.82  .1515:**********************
-1.63  .1679:*************************
-1.44  .1844:***************************
-1.24  .1990:*****************************
-1.05  .2116:*******************************
 -.86  .2225:*********************************
 -.67  .2312:**********************************
 -.48  .2379:***********************************
 -.29  .2424:************************************
 -.10  .2448:************************************
  .10  .2448:************************************
  .29  .2424:************************************
  .48  .2379:***********************************
  .67  .2312:**********************************
  .86  .2225:*********************************
 1.05  .2116:*******************************
 1.24  .1990:*****************************
 1.44  .1844:***************************
 1.63  .1679:*************************
 1.82  .1515:**********************
 2.01  .1315:*******************
 2.20  .1124:****************
 2.39  .0926:*************
 2.58  .0722:**********
 2.77  .0527:*******
 2.97  .0343:*****
 3.16  .0181:**
 3.35  .0011:
```

FIG. 4b. Max's quantizer with $N = 36$. The output levels have been transformed to be spaced equidistantly on the interval -3.35, 3.35.

vary too drastically from the identity transformation. The solutions obtained from Eqs. (15a) and (15b) may not be monotonic increasing, even through that assumption was used in the derivation. Thus the constants c_1 and c_2 may have to be adjusted in order to assure that the solution T will be monotonic.

Suppose that the desired histogram $p_g(z)$ is uniform, and that $p_{f_1}(z) = p_f(z)$ for all z. Then when we replace z by $T^{-1}(z)$, Eq. (15a) becomes

$$T^{-1}(z) - \frac{d}{dz}\left(\frac{c_2}{c_1}\,\frac{p_{f_1}(T^{-1}(z))}{T'(T^{-1}(z))}\right) = z,$$

or, by letting $S = T^{-1}$, we have

$$S(z) - \frac{d}{dz}\left(\frac{c_2}{c_1}\,p_{f_1}(S(z))S'(z)\right) = z. \qquad (15c)$$

This diffusion equation may be solved using a Newton's approximation method for matrices in conjunction with a linear equation solver (for tridiagonal matrices) [16].

Figure 3 shows the result of applying this solution to an adjusted normal density function (Fig. 3a)

$$p_f(z) = [1/(2\pi)^{1/2}]e^{-z^2/2} + (s/a),$$

where

$$s = \int_{-\infty}^{a} [1/(2\pi)^{1/2}]e^{-t^2/2}dt.$$

Here the gray-level range is assumed to be $[-a,a]$, and the weighting constants satisfy $c_2/c_1 = 1/2$. The resulting output distribution has been smoothed slightly, and is shown in Fig. 3b. Since the solution for T behaves rather critically in the regions around $z = \pm 1$ (where $p_{f_1}''(z) = 0$), this smoothing procedure permits a larger value for c_2/c_1 than would be allowed by the monotonicity restriction, and ensures a smooth, near-optimal output distribution.

Figure 3 may be compared with Fig. 4, where we show the initial and output density functions obtained from the minimal distortion quantizer as given by Max [5].

REFERENCES

1. A. Rosenfeld and E. B. Troy, Visual texture analysis, in *Conference Record of the Symposium on Feature Extraction and Selection in Pattern Recognition,* IEEE Publ. 70C 51-C, Argonne, Ill., October 1970, pp. 115–124. See also Univ. of Md. Tech. Rept. 70-116.
2. D. L. Milgram, Computer methods for creating photomosaics, Univ. of Md. Tech. Rept. 313, July 1974.
3. R. V. Hogg and A. T. Craig. *Introduction to Mathematical Statistics,* 3rd ed., Macmillan, New York, 1970.
4. Glenn M. Roe, Quantizing for minimum distortion, *IEEE Trans. Information Theory* **IT-10,** 1964, 384.
5. Joel Max, Quantizing for minimum distortion, *IRE Trans.* **IT-6,** 1960, 7.
6. P. F. Panter and W. Dite. Quantization distortion in pulse-count modulation with nonuniform spacing of levels, *Proc. IRE* **39,** 1951, 44.
7. V. R. Algazi, Useful approximations to optimum quantization, *IEEE Trans. Communication Technology* **COM-14,** 1966, 297.
8. Peter Elias, Bounds on performance of optimum quantizers, *IEEE Trans. Information Theory* **IT-16,** 1970, 172.
9. W. B. Green, J. A. Cutts, E. D. Jahelka, R. A. Johansen, M. J. Sander, J. B. Seidman, A. T. Young, and L. A. Saderblom, Mariner 9: Image processing and products, *Icarus* **18,** 1973, 75.
10. T. G. Stockham, Image processing in the context of a visual model, *Proc. IEEE* **20,** 1972, 828.
11. J. A. Dunne, W. D. Stromberg, R. M. Ruiz, S. A. Collins, and T. E. Thorpe, Maximum discrimiability versions of the near-encounter Mariner pictures, *J. Geophys. Res.* **76,** 1971, 438.
12. R. H. Selzer, Improving Biomedical image quality with computers, Jet Propulsion Laboratory Tech. Rept. 32-1336, October 1968.
13. E. L. Hall, R. P. Kruger, S. J. Dwyer, D. L. Hall, R. W. McLaren, and G. S. Lodwick, A survey of preprocessing and feature extraction techniques, *IEEE Trans. Computers* **C-20,** 1971, 1032.
14. R. M. Haralick, K. Shanmugan, and I. Dinstein, Textural features for image classification, *IEEE Trans. Systems, Man, Cybernetics* **SMC-3,** 1973, 610.
15. E. B. Troy, E. S. Deutsch, and A. Rosenfeld, Gray-level manipulation experiments for textural analysis, *IEEE Trans. Systems, Man, Cybernetics* **SMC-3,** 1973, 91.
16. J. M. Ortega, *Numerical Analysis: A Second Course,* Academic Press, New York, 1972.