UNIVERSITY OF READING

DEPARTMENT OF MATHEMATICS

SMOOTHED HISTOGRAM MODIFICATION FOR IMAGE PROCESSING

Jaroslav Kautsky, Nancy K. Nichols
and David L.B. Jupp

Numerical Analysis Report 9/81

This paper is also issued as a Report from the School of Mathematical Sciences at the Flinders University of South Australia.

December 1981

ABSTRACT

The technique of constructing a transformation, or *regrading*, of a discrete data set such that the histogram of the transformed data matches a given reference histogram is commonly known as histogram modification. The technique is widely used for image enhancement and normalization. In this paper we show that a method which we have previously defined for producing such a regrading is "best" in the sense that it minimizes the error between the cumulative histogram of the transformed data and that of the given reference function, over-all single-valued, monotone, discrete transformations of the data. We also examine techniques for *smoothed* regrading, which provide a means of balancing the error in matching a given reference histogram against the information lost with respect to a linear transformation. The smoothed regradings are shown to optimize certain cost functionals. Numerical algorithms for generating the smoothed regradings, which are simple and efficient to implement, are described, and practical applications to the processing of LANDSAT image data are discussed.

1. Introduction

Histogram modification is widely used in image processing; for example in image enhancement to improve visual contrast by histogram "equalization," or flattening [1], [4], [5], and in precision processing to calibrate corrections for balancing sensor differences ("destriping") [3]. It may also be used to normalize images for global atmospheric and scene radiance changes before multi-temperal analysis or mosaicing. Simple systematic procedures for mapping image data linearly into a restricted grey scale range for colour display or for transforming "standard colours" between colour monitors may also be provided by histogram modification methods [7], [10].

The technique of histogram modification requires the construction of a transformation such that the histogram of the transformed image data matches a given reference function. In this paper we consider only transformations which are single-valued and monotonic, that is, which preserve equality and order relations between grey levels. We refer to such a transformation as a *regrading*. Regradings are *spatially* independent of the image data and are easily represented by simple look-up tables. Histogram modification by regrading is therefore rapid and efficient to implement.

The regrading obtained by matching a *flat* reference histogram is well-known to have certain optimal information theoretic characteristics [11]. This regrading will be called *equidistributing* and has the property that any two images which differ only by a monotone transformation of their grey levels will become identical after such a regrading. It is optimal in the sense that amongst all transformations into a fewer number of grey levels it retains maximal information.

A linear compression or stretch of the grey scale gives a regrading which matches a linear scaling of the original histogram. Such a transformation will be called a *linear* regrading and has the property that it causes any two images which differ only by a linear transformation to become identical. Although not optimal in the sense of information theory, the linear regrading preserves quantitative relations which are often significant for interpretation.

In an earlier paper [10] we have shown that a range of *smoothed* regradings between these two cases: equidistributing and linear, may be defined. These smoothed regradings provide a means of balancing the gains and losses of information between the two extreme cases. Smoothed regradings which offer a compromise between the linear regradings and the regrading that matches a specified reference histogram may also be determined.

In theory, for *continuous* image data, there always exists a unique monotone continuous transformation which exactly matches an image to a reference histogram [4]. In general, a discrete-to-discrete transformation can only give an approximation to the continuous case. Various choices of rounding from the continuous to the discrete have been used in the literature [2], [3], [4]. In this paper we show that the explicit *weighted* regrading we have previously defined [10] gives the "best" approximation to the continuous-to-continuous transformation in the sense that it minimizes the error between the cumulative histograms (or equivalently, the probability distribution functions) of the transformed images. We show also that *smoothed* weighted regradings are best approximations to the continuous transformations which minimize certain cost functions. The cost function is a measure which balances the error in matching the reference function against the information lost with respect to the linear regrading.

In cases where the image histogram is extremely unevenly distributed, for example, where it has large peaks, even the "best" discrete regrading may give only a poor approximation to the specified reference histogram. The use of smoothing mollifies these difficulties without any extra preprocessing of the data and without loss of speed and efficiency in implementation.

A computer package for the display of regraded images has been developed at CSIRO, Land Use Research Division. This package, documented as PEEK [6], incorporates the algorithms discussed here for producing smoothed regradings. Applications and results are described in section 6 of this paper.

In section 2 we introduce general notation and give the formulation of the problem. In section 3 we examine algorithms for histogram matching, and in section 4 we investigate smoothing techniques. Methods for implementing smoothed regradings are described in section 5. The paper concludes with a discussion of the practical use of these techniques in processing LANDSAT image data.

2. Formulation of the Problem

In this section we introduce the necessary notation and terminology. The problem is essentially to determine a transformation, say τ, of the input data - a sequence of numbers giving the grey levels of individual pixels of the scene, into the output data - a similar sequence of grey levels suitable for display. The input data are assumed to be from a discrete set of *grades*, say $\{1,2,...,n\}$, dependent on the instruments collecting the data; and the output data to be from another discrete set of grades, say $\{1,2,...,m\}$, determined by the characteristics of the output display unit. The transformation τ is such that input level $j \in \{1,2,...,n\}$ becomes output level $\tau(j) \in \{1,2,...,m\}$, and τ thus maps discrete data into discrete data. We make the following definition:

Definition 1: τ is called a *regrading* of n grades into m grades if $\tau \in T_n^m$, the set of transformations such that

(i) $\tau : \{1,2,...,n\} \longmapsto \{1,2,...,m\}$

(ii) τ is single-valued

(iii) τ is monotonic, *i.e.* $i \le j \Rightarrow \tau(i) \le \tau(j)$.

Regradings thus preserve equality and order relations of the input scale.

A regrading from n into m grades is easily represented by a look-up-table, that is, by an n-dimensional vector $\underline{\tau} = (\tau_1,\tau_2,...,\tau_n)$ where $\tau_j = \tau(j)$ and τ_j are integer values satisfying $1 \le \tau_1 \le \tau_2 ... \le \tau_n \le m$. Such a regrading may also be represented by an ordered set of so-called *break-points* x_k, $k = 0,1,\ldots,m$, such that

$$\tau_j = k \quad \forall j \quad \text{such that} \quad x_{k-1} < j \le x_k \quad .$$

(We note that a regrading τ is uniquely representable using *integer* break-points).

For a given scene, we denote the histogram of the input data by the n-dimensional vector $\underline{f} = (f_1,f_2,\ldots,f_n)$, where f_j equals the number of pixels with grey level j in the scene, and $f_j \ge 0$. The histogram of the output data under transformation $\tau \in T_n^m$ is denoted by $\tau \otimes \underline{f}$ and is defined as the m-dimensional vector $\tau \otimes \underline{f} = (\hat{f}_1, \hat{f}_2,\ldots,\hat{f}_m)$ with $\hat{f}_k = \sum_{j \ni \tau_j = k} f_j$. (Equivalently $\hat{f}_k = \sum_{j=x_{k-1}+1}^{x_k} f_j$, where x_k are the *integer* break-points defining τ). Clearly $\hat{f}_k$ equals the number of pixels in the output data with grey level k, and

$$|\tau \otimes \underline{f}| \equiv \sum_{k=1}^{m} \hat{f}_k = \sum_{j=1}^{n} f_j \equiv |\underline{f}| \; .$$

The discrete histogram modification problem is then expressed as follows: given input histogram $\underline{f}$ and reference $\underline{g}$, find $\tau \in T_n^m$ such that $\tau \otimes \underline{f} = \underline{g}$.

Two examples are the equidistributing regrading and the linear regrading described in the introduction. The equidistributing regrading is obtained by matching the input histogram $\underline{f}$ to a flat histogram $\underline{g} = (g_1,g_2,\ldots,g_m)$ where $g_i = c \;\; \forall i$, and c is a positive constant. The linear regrading is obtained by matching $\underline{f}$ to a linear scaling of itself in such a way that the original n grades are distributed as equally as possible into the new m grades.

We observe that, in general, the discrete histogram modification problem cannot be solved exactly. In practice, it may even happen that the reference histogram $\underline{g}$ arises from data defined on q grades where $q \neq m$, and the problem is then not clearly defined.

To make the problem more precise, we use the *cumulative* histogram of the data: given histogram $\underline{f}$ (of dimension n), we define $D_{\underline{f}}$ as the continuous piecewise-linear function on [0,1] such that

$$D_{\underline{f}}(j/n) = \sum_{i=1}^{j} f_i/|\underline{f}| \quad , \quad D_{\underline{f}}(0) = 0 \quad ,$$

$$D'_{\underline{f}}(x) = nf_i/|\underline{f}| \qquad \forall\, x \in ((j-1)/n,\ j/n) \quad .$$

Then $D_{\underline{f}} \in \mathcal{D}$ where $\mathcal{D}$ is the set of functions D such that

(i) $D : [0,1] \longmapsto [0,1]$

(ii) $D(0) = 0 \quad , \quad D(1) = 1$

(iii) D' exists a.e.

(iv) $D^{-1} \in \mathcal{D}$ exists.

(*Remark*: If $D' = 0$ on a subinterval of [0,1], then D^{-1} is not uniquely defined on that subinterval. In order that such functions are included in $\mathcal{D}$, we explicitly define D^{-1} such that

$$D^{-1}(K) = \max\,\{x \mid D(x) = K\} \quad .)$$

The function $D_{\underline{f}}$ may be regarded as a continuous (scaled) representation of the cumulative histogram of the input data, and $d_{\underline{f}} \equiv D'_{\underline{f}}$, its derivative, as a piecewise continuous (scaled) representation of the histogram $\underline{f}$. If we make the assumption that the discrete image data arise from a continuous image where the grey levels in the continuous range [j-1,j] are equally distributed over each pixel having grey level j, then $D_{\underline{f}}$ is just the probability distribution function of the continuous image data, and $d_{\underline{f}}$ is the corresponding probability density function.

The discrete histogram modification problem then becomes:

Problem 1: Given input histogram $\underline{f}$ and reference histogram $\underline{g}$, find $\tau \in T_n^m$ such that $D_{\tau\otimes\underline{f}} = D_{\underline{g}}$.

(We observe that with this definition of the problem $\underline{g}$ may be of any dimension.)

The exact solution to this problem may not exist, however, due to the discrete nature of transformation τ. We therefore must choose a regrading τ which gives the "best" approximation to the solution in some sense. In the next section we give the definition of a weighted regrading of $\underline{f}$ into $\underline{g}$ (as in [10]) and show that it minimizes the difference between $D_{\tau\otimes\underline{f}}$ and $D_{\underline{g}}$ in a certain measure.

3. Algorithms for histogram modification

For continuous image data it may be shown that there always exists a unique monotonic continuous transformation which matches an image histogram to a reference histogram. We have, precisely, the following theorem.

Theorem 1: Given any $D_{\underline{f}}, D_{\underline{g}} \in \mathcal{D}$ then there exists a unique $T \in \mathcal{D}$ (i.e. a unique single-valued, monotonic transformation from [0,1] into [0,1]) such that $D_{\underline{f}}(s) = D_{\underline{g}}(T(s))$, and T is given explicitly by $T = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$.

In other words, given cumulative histograms (or probability distribution functions) $D_{\underline{f}}(s)$, and $D_{\underline{g}}(s')$, there exists a continuous transformation T of the continuous data s into s' such that the cumulative histogram (probability distribution) of the transformed data s' takes given form $D_{\underline{g}}$.

For histograms $\underline{f}$ and $\underline{g}$, if a *discrete* regrading τ exists such that $D_{\tau\otimes\underline{f}}$ exactly equals $D_{\underline{g}}$, then for all $k = 1,2,...,m$, there exists an *integer* $j_k \in \{1,2,...,n\}$ such that $D_{\underline{f}}(jk/n) = D_{\underline{g}}(k/m)$; that is, letting $T = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$ be the exact continuous transform matching $D_{\underline{f}}$ to $D_{\underline{g}}$, then $j_k = nT^{-1}(k/m)$ is an integer belonging to [0,n] and the solution to Problem 1 is given *exactly* by

$$\tau_j = k \quad \forall_j \quad \text{such that} \quad j_{k-1} < j \leq j_k \; .$$

In general, however, $x_k = nT^{-1}(k/m)$ is *not* an integer, and some "rounding" procedure is required to define the discrete transformation.

Procedures used widely in the literature for obtaining a discrete regrading involve "bin-filling" ([2],[3],[4]), and generally work from "left to right". The algorithm defining the transformation in this case is given by

$$\begin{aligned}\tau_j &= \min\{k \mid \sum_{i=1}^{j} f_i \le \sum_{i=1}^{k} g_i\} \\ &\equiv \min\{k \mid D_{\underline{f}}(j/n) \le D_{\underline{g}}(k/m)\} \quad .\end{aligned} \tag{1}$$

Using the definition $T = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$ and the monotonicity properties of $D_{\underline{f}}$, $D_{\underline{g}}$ and T, (1) is equivalent to the algorithm

$$\tau_j = k \quad \forall j \text{ such that } x_{k-1} < j \le x_k \tag{2}$$

where $x_k = nT^{-1}(k/m)$. This algorithm produces the exact regrading matching $D_{\tau\otimes\underline{f}}$ to $D_{\underline{g}}$ when it exists.

The complementary "right-to-left" technique

$$\begin{aligned}\tau_j &= \max\{k \mid \sum_{i=1}^{k-1} g_i \le \sum_{i=1}^{j-1} f_i\} \\ &\equiv \max\{k \mid D_{\underline{g}}((k-1)/m) \le D_{\underline{f}}((j-1)/n)\}\end{aligned} \tag{3}$$

has this same property, and is equivalent to the algorithm

$$\tau_j = k \quad \forall_j \text{ such that } x_{k-1} \le j-1 < x_k \tag{4}$$

where $x_k = nT^{-1}(k/m)$.

Regradings (2) and (4) are representable by integer break-points j_k, $k = 1,2,\ldots,m$, where, in case (2) j_k is the largest integer such that $j_k \le x_k$, and in case (4) j_k is the smallest integer such that $x_k \le j_k$. These algorithms thus consistently round the continuous transform up (or down) to obtain the discrete regrading. Such "one-way" algorithms tend to give poor results, especially when the input histogram is sharply peaked at some points.

In [10] we have given a more symmetric alternative algorithm for determining the discrete regrading matching $D_{\underline{f}}$ to $D_{\underline{g}}$:

Definition 2: The regrading $\tau^* \in T_n^m$ satisfying

$$\tau_j^* = k \quad \forall_j \quad \text{such that} \quad x_{k-1} < j-\tfrac{1}{2} \le x_k \qquad (5)$$

where $x_k = nT^{-1}(k/m)$, $T = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$ is called the *weighted* regrading matching $D_{\underline{f}}$ to $D_{\underline{g}}$.

The weighted regrading τ^* given by (5) is representable by integer break-points j_k, $k = 1,2,\ldots,m$ such that $j_k - \tfrac{1}{2} \le x_k < j_k + \tfrac{1}{2}$, *i.e.* such that j_k is the *closest* integer to $x_k = nT^{-1}(k/m)$. This algorithm also produces the *exact* regrading matching $D_{\underline{f}}$ to $D_{\underline{g}}$ when it exists. Futhermore the weighted regrading τ^* of Definition 2 gives the "best" approximation to the solution of Problem 1 in the following sense:

Theorem 2: The weighted regrading τ^* minimizes the error between $D_{\tau\otimes\underline{f}}$ and $D_{\underline{g}}$ in the discrete l_p-norm $(1 \le p \le \infty)$ over all $\tau \in T_n^m$, that is

$$\| D_{\tau^*\otimes\underline{f}} - D_{\underline{g}} \|_{p,m} = \min_{\tau\in T_n^m} \| D_{\tau\otimes\underline{f}} - D_{\underline{g}} \|_{p,m} \qquad (6)$$

(The discrete l_p-norms of $D_{\underline{f}}$ are given by

$\| D_{\underline{f}} \|_{p,n} = [\sum_{j=1}^{n} D_{\underline{f}}(j/n)^p]^{1/p}$, $1 \le p < \infty$; and $\| D_{\underline{f}} \|_{\infty,n} = \max_j |D_{\underline{f}}(j/n)|$.)

Proof: Let $j_k \in \{1,2,\ldots,n\}$ be the integer closest to $x_k = nT^{-1}(k/m)$, $k = 1,2,\ldots,m$, and let τ^* be the regrading defined by these break-points. Then $D_{\tau^*\otimes\underline{f}}(k/m) = D_{\underline{f}}(j_k/n)$. There are two cases. Suppose first that

$j_k - 1 < x_k \le j_k$; then, since $D_{\underline{f}}$ is monotonic non-decreasing we have

$$D_{\underline{f}}((j_k-1)/n) < D_{\underline{f}}(x_k) \le D_{\underline{f}}(j_k/n)$$

and since $D_{\underline{f}}$ is linear on $[(j_k-1)/n\ ,\ j_k/n]$, we have

$$|D_{\underline{f}}(x_k) - D_{\underline{f}}(j_k/n)| < |D_{\underline{f}}(x_k) - D_{\underline{f}}(j/n)|\ ,\ \forall j \ne j_k\ .$$

Using $T^{-1} = D_{\underline{f}}^{-1} \circ D_{\underline{g}}$ and the definition of x_k we obtain

$$|D_{\underline{g}}(k/m) - D_{\underline{f}}(j_k/n)| = \min_{j\in\{1,2,\ldots,n\}} |D_{\underline{g}}(k/m) - D_{\underline{f}}(j/m)|\ .$$

The same result follows in the case $j_k \le x_k < j_k + 1$. Hence

$$|D_{\underline{g}}(k/m) - D_{\tau^*\otimes \underline{f}}(k/m)| = \min_{\tau\in T_n^m} |D_{\underline{g}}(k/m) - D_{\tau\otimes\underline{f}}(k/m)|\ ,$$

and (6) follows directly by definition.

We observe that in the case $\underline{f}$ is of dimension m, the error between $D_{\tau\otimes\underline{f}}$ and $D_{\underline{g}}$ is also minimized by τ^* in the L_∞ *continuous* norm. This follows because $D_{\underline{g}}$ and $D_{\tau\otimes\underline{f}}$ are piecewise-linear between points k/m, k = 1,2,...,m , and therefore the maximum errors on [0,1] are bounded by the errors at points k/m.

We conclude that amongst all choices of discrete transformation $\tau \in T_n^m$ which approximate the continuous transformation T, the weighted regrading τ^* of Definition 2 minimizes the maximum error between $D_{\tau\otimes\underline{f}}$ and $D_{\underline{f}}\circ T^{-1} \equiv D_{\underline{g}}$. Other properties and the behaviour of the regrading τ^* are given in [10] .

A direct procedure for constructing τ^* is easy to implement using inverse linear interpolation to find the break-points x_k satisfying

$D_{\underline{f}}(x_k/n) = D_{\underline{g}}(k/m)$, where $\underline{f}$, $\underline{g}$ are given. A simple FORTRAN subroutine for determining τ^* is given in [9].

4. Smoothed regradings

The regrading from n into m grades which maximally retains information (in the sense of maximum entropy) is obtained by matching the input histogram to a flat reference histogram, and is called equidistributing. Experience has shown that, although optimal in the information theory sense, images produced by the equidistributing regrading tend to be too sharp, and to reduce many of the contrasts in the image which are significant for interpretation. On the other hand, the regrading which uniformly distributes the old grey levels onto the new grey scale, called the linear regrading, preserves the original shape of the histogram as accurately as possible. In an earlier paper [10], we have described methods for producing a continuous range of *smoothed* regradings between these two cases: linear and equidistributing. The smoothing is regulated by a single parameter which provides a means of precisely balancing the gains and losses in image clarity between the extremes.

We have also defined methods for constructing single-parameter families of smoothed regradings in a range between the linear regrading and the weighted regrading which matches a specified reference histogram [9]. These smoothed regradings minimize cost functions which measure the error in matching the reference histogram against the information lost from the linearly scaled input histogram. This result is most easily demonstrated by first examining continuous transformations on continuous image data. The smoothed regradings are then obtained by algorithm (5) as "best" approximations to the continuous transformations. In this section we consider the continuous case and in the next section we describe the discrete implementation.

Definition 3: Given $D_{\underline{f}}$, $D_{\underline{g}} \in \mathcal{D}$, let F_λ, $G_\lambda \in \mathcal{D}$ be functions continuously dependent on parameter $\lambda \in [0,1]$ such that $F_0 = D_{\underline{f}}$ $G_0 = D_{\underline{g}}$ and $F_1 \equiv G_1$. Then $T_\lambda = G_\lambda^{-1} \circ F_\lambda$ is a *smoothed* transformation of $D_{\underline{f}}$ into $D_{\underline{g}}$.

The transformation T_λ has the property that for $\lambda = 0$, $T_0 = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$, *i.e.* $D_{\underline{f}}(s) = D_{\underline{g}}(T_0(s))$ and T_0 is just the continuous transformation which matches cumulative histogram $D_{\underline{f}}$ to cumulative histogram $D_{\underline{g}}$. For $\lambda = 1$, $T_1 = G_1^{-1} \circ F_1 \equiv I$, *i.e.* T_1 corresponds to a linear map between grey scales. For $\lambda \in (0,1)$, T_λ provides a compromise between the two extremes.

Linear Smoothing

The simplest smoothing depends linearly on the parameter λ. We have the following two cases:

(i) Let $F_\lambda \equiv D_{\underline{f}}$ for all λ, and let $G_\lambda \equiv (1-\lambda)D_{\underline{g}} + \lambda D_{\underline{f}}$. Then we obtain the smoothed transformation

$$\tilde{T}_\lambda \equiv ((1-\lambda)D_{\underline{g}} + \lambda D_{\underline{f}})^{-1} \circ D_{\underline{f}} \quad . \tag{7}$$

(ii) Let $F_\lambda = (1-\lambda)D_{\underline{f}} + \lambda D_{\underline{f}}$, and let $G_\lambda \equiv D_{\underline{g}}$ for all λ. Then we obtain the smoothed transformation

$$\tilde{\tilde{T}}_\lambda \equiv D_{\underline{g}}^{-1} \circ ((1-\lambda)D_{\underline{f}} + \lambda D_{\underline{g}}) \tag{8}$$

Both transformations optimize quadratic cost functionals. For the proof we require the following lemma, which is easily demonstrated by standard least-square arguments:

Lemma Given $D_i \in \mathcal{D}$, $i = 1,2,\ldots,p$, and continuous, positive functions $\sigma_i = \sigma_i(s)$, the functional

$$\tilde{F}(D) = \sum_{i=1}^{p} \int \sigma_i(s)(D(s) - D_i(s))^2 ds$$

is minimized by

$$D^*(s) = \left[\sum_{i=1}^{p} \sigma_i(s) D_i(s)\right] / \left[\sum_{i=1}^{p} \sigma_i(s)\right].$$

We obtain the following theorems:

Theorem 3: Given $D_{\underline{f}}, D_{\underline{g}}$, the transformation $\tilde{T}_\lambda$ defined by (7) minimizes

$$(1-\lambda)\int_0^1 \left[d_{\langle T\underline{f}\rangle} - d_{\underline{g}}\right]^2 \sigma ds + \lambda \int_0^1 \left[d_{\langle T\underline{f}\rangle} - d_{\underline{f}}\right]^2 \sigma ds, \tag{9}$$

where $\sigma = \sigma(s)$ in any continuous, positive function, $d_{\underline{f}} \equiv D'_{\underline{f}}$, $d_{\underline{g}} \equiv D'_{\underline{g}}$, and $d_{\langle T\underline{f}\rangle} \equiv (D_{\underline{f}} \circ T^{-1})'$. (The function $d_{\langle T\underline{f}\rangle}$ may be regarded as the probability density function of the continuously transformed continuous image data defined in section 1.)

Proof: The functional (9) may be written

$$(1-\lambda)\int_0^1 ((D_{\underline{f}} \circ T^{-1})' - D'_{\underline{g}})^2 \sigma ds + \lambda \int_0^1 ((D_{\underline{f}} \circ T^{-1})' - D'_{\underline{f}})^2 \sigma ds. \tag{10}$$

By the Lemma, (10) is minimized by

$$(D_{\underline{f}} \circ T^{-1})' = (1-\lambda) D'_{\underline{f}} + \lambda D'_{\underline{g}}.$$

Integrating and using $D_{\underline{g}}(1) = D_{\underline{f}}(1) = 1$, and then re-arranging, we obtain the minimum of (10) with

$$T = ((1-\lambda) D_{\underline{f}} + \lambda D_{\underline{g}})^{-1} \circ D_{\underline{f}} \equiv \tilde{T}_\lambda.$$

Theorem 4: Given $D_{\underline{f}}, D_{\underline{g}}$, the transformation $\tilde{\tilde{T}}_\lambda$ defined by (8) minimizes

$$(1-\lambda)\int_0^1 (d_{\langle T^{-1}\underline{g}\rangle} - d_{\underline{f}})^2 \sigma ds + \lambda\int_0^1 (d_{\langle T^{-1}\underline{g}\rangle} - d_{\underline{g}})^2 \sigma ds \quad (11)$$

where $\sigma = \sigma(s)$ is any continuous, positive function, and $d_{\underline{f}} \equiv D'_{\underline{f}}$, $d_{\underline{g}} \equiv D'_{\underline{g}}$, $d_{\langle T^{-1}\underline{g}\rangle} \equiv (D_{\underline{g}} \circ T)'$.

Proof: The functional (11) may be written

$$(1-\lambda)\int_0^1 ((D_{\underline{g}} \circ T)' - D'_{\underline{f}})^2 \sigma ds + \lambda\int_0^1 ((D_{\underline{g}} \circ T)' - D'_{\underline{g}})^2 \sigma ds \quad (12)$$

The rest of the proof follows from the Lemma as in Theorem 3.

It may be observed that when $\lambda = 0$, $\tilde{T}_\lambda \equiv \tilde{\tilde{T}}_\lambda = D_{\underline{g}}^{-1} \circ D_{\underline{f}}$, and the minimum value of functions (9) and (11) is zero. Similarly, when $\lambda = 1$, $\tilde{T}_\lambda \equiv \tilde{\tilde{T}}_\lambda = I$, and the functionals (9) and (11) also take minimum value zero. We remark that the cost functional (9) is equivalent to the minimum information loss function of Hummel [4] and that the corresponding "best" discrete approximation to the minimizing continuous transformation is easily computed by the algorithms described in section 5 of this paper.

Generalized Linear Smoothing

A more general form of smoothing can be obtained as a composition of linear smoothings. The resulting transformation is non-linearly dependent upon the parameter λ:

(iii) Let $F_\lambda = (1-\lambda)D_{\underline{f}} + \lambda D_{\underline{y}}$, and let $G_\lambda = (1-\lambda)D_{\underline{g}} + \lambda D_{\underline{y}}$, where $D_{\underline{y}} \in \mathcal{D}$ is an arbitrary function. Then we obtain the smoothed transformation

$$\tilde{\tilde{T}}_\lambda \equiv ((1-\lambda)D_{\underline{g}} + \lambda D_{\underline{y}})^{-1} \circ ((1-\lambda)D_{\underline{f}} + \lambda D_{\underline{y}}). \quad (13)$$

(We note that if S_1 is the transformation matching F_λ to $D_{\underline{y}}$ and S_2 is the transformation matching $D_{\underline{y}}$ to G_λ, then $\tilde{\tilde{\tilde{T}}}_\lambda = S_2 \circ S_1$.)

Using arguments similar to those of Theorems 3 and 4 we can show the following:

Theorem 5: Given $D_{\underline{f}}$, $D_{\underline{g}}$ and $D_{\underline{y}}$, the transformation $\tilde{\tilde{\tilde{T}}}_\lambda$ defined by (13) minimizes

$$(1-\lambda)\int_0^1 \sigma_1(d_{\langle T\underline{f}_\lambda\rangle} - d_{\underline{g}})^2 + \sigma_2(d_{\langle T^{-1}\underline{g}_\lambda\rangle} - d_{\underline{f}})^2 ds$$

$$+ \lambda\int_0^1 \sigma_1(d_{\langle T\underline{f}_\lambda\rangle} - d_{\underline{y}})^2 + \sigma_2(d_{\langle T^{-1}\underline{g}_\lambda\rangle} - d_{\underline{y}})^2 ds,$$

where $\sigma_1 = \sigma_1(s)$, $\sigma_2 = \sigma_2(s)$ are any continuous, positive functions, and $d_{\underline{f}} \equiv D'_{\underline{f}}$, $d_{\underline{g}} = D'_{\underline{g}}$, $d_{\underline{y}} = D'_{\underline{y}}$, $d_{\langle T\underline{f}_\lambda\rangle} \equiv (F_\lambda \circ T^{-1})'$, $d_{\langle T^{-1}\underline{g}_\lambda\rangle} \equiv (G_\lambda \circ T)'$.

We observe that the arbitrariness of the weight functions σ, and σ_1, σ_2 in (9), (11) and (14) means that the transformations $\tilde{T}_\lambda$, $\tilde{\tilde{T}}_\lambda$, $\tilde{\tilde{\tilde{T}}}_\lambda$ defined by (7), (8) and (13) each minimize a large class of functionals. However, despite these optimality properties, practical experience indicates that none of these rather unsophisticated smoothings is as effective as *non-linear* methods based on "padding" techniques. In [10] we describe methods of this type for determining smoothed equidistributing regradings. Generalizations of the discrete non-linear techniques for obtaining smoothed weighted regradings are given in the next section. Optimality properties of the non-linear regradings are unfortunately not easy to determine.

5. Implementation of smoothing

In principle, the discrete implementation of the continuous smoothing procedures described in section 4 is very simple. The input histogram $\underline{f}$ is replaced by a new histogram $\underline{f}_\lambda$ and the reference histogram $\underline{g}$ is replaced by $\underline{g}_\lambda$; then $\tau^*_\lambda \in T^m_n$, the weighted regrading satisfying Definition 2, which matches $D_{\underline{f}_\lambda}$ to $D_{\underline{g}_\lambda}$ is the smoothed weighted regrading. By Theorem 2, τ^*_λ is the best discrete approximation to the smoothed *continuous* transformation $T_\lambda = G^{-1}_\lambda \circ F_\lambda$ of Definition 3, where $F_\lambda \equiv D_{\underline{f}_\lambda}$ and $G_\lambda \equiv D_{\underline{g}_\lambda}$. (It is assumed that $\underline{f}_0 \equiv \underline{f}$, $\underline{g}_0 \equiv \underline{g}$ and that $\underline{f}_1$, $\underline{g}_1$ are such that $D_{\underline{f}_1} = D_{\underline{g}_1}$).

Linear Smoothing

Exact implementation of the linear smoothing (cases (i) and (ii)) of section 4 requires that the input histogram $\underline{f}$ and the reference histogram $\underline{g}$ *both* be of dimension n. Then for case (i) $\underline{f}_\lambda$ and $\underline{g}_\lambda$ are chosen as:

(i) $\underline{f}_\lambda \equiv \underline{f}$,

$\underline{g}_\lambda \equiv (1-\lambda)\underline{g} + \lambda\underline{f}$;

and for case (ii) we take

(ii) $\underline{f}_\lambda = (1-\lambda)\underline{f} + \lambda\underline{g}$,

$\underline{g}_\lambda \equiv \underline{g}$.

Clearly the cumulative histograms $D_{\underline{f}_\lambda}$ and $D_{\underline{g}_\lambda}$ are equal, respectively, to F_λ and G_λ as required by the definitions in section 4, and the smoothed weighted regrading τ^*_λ matching $D_{\underline{f}_\lambda}$ to $D_{\underline{g}_\lambda}$ gives the best approximation in $\mathcal{T}_n^m$ to $\tilde{T}_\lambda$ in case (i) and to $\tilde{\tilde{T}}_\lambda$ in case (ii). These regradings are thus optimal in the sense that they minimize information loss as measured by specific cost functionals. In particular, these regradings minimise a specific balance between the error in matching $D_{\underline{f}}$ to $D_{\underline{g}}$, which gives the weighted regrading, and the error in matching $D_{\underline{f}}$ to $D_{\underline{f}}$, which gives the linear regrading.

Generalized Linear Smoothing

The linear smoothings (i) and (ii) are special cases of the generalized linear smoothing (iii) where $\underline{y}$ is chosen in (i) equal to $\underline{f}$ and in (ii) equal to $\underline{g}$. The more general form (iii) of the continuous smoothing is implemented by choosing

$$\text{(iii)}\quad \underline{f}_\lambda = (1-\lambda)\underline{f} + \lambda\underline{y},$$
$$\underline{g}_\lambda = (1-\lambda)\underline{g} + \lambda\underline{\tilde{y}}.$$

If $\underline{f}$ and $\underline{g}$ are both of dimension n, then $\underline{y} = \underline{\tilde{y}}$ is taken, and $\underline{y}$ may be chosen arbitrarily. In the case where $\underline{g}$ is of dimension $q \neq n$, we take $y_j = c$, $j = 1,2,\ldots,n$, and $\tilde{y}_j = \tilde{c}$, $j = 1,2,\ldots,q$, where c, $\tilde{c}$ are positive constants. In either case $D_{\underline{y}} \equiv D_{\underline{\tilde{y}}}$, and the smoothed weighted regrading τ_λ^* matching $D_{\underline{f}_\lambda}$ to $D_{\underline{g}_\lambda}$ gives the best approximation in $\mathcal{T}_n^m$ to the smoothed continuous transformation $\tilde{\tilde{\tilde{T}}}_\lambda$ defined by (13).

Non-linear Smoothing

Although all three smoothing procedures (i), (ii) and (iii), based on linear combinations of the weight and reference functions, possess certain optimality properties, we have found that non-linear "padding" procedures often produce more effective results in practice. These methods are based on mesh selection techniques used in solving ordinary differential equations (see [8]), and operate by adding artificial "pixels" to pad out the grey levels which have no, or few, entries, thus "smoothing" the input histogram (or reference histogram). In [10] we have described two methods of this type for finding smoothed equidistributing regradings. These methods match certain paddings of the input histogram to flat reference histograms; the padding is regulated by a single parameter $\lambda \in [0,1]$ such that, with no padding ($\lambda = 0$) the equidistributing regrading is obtained and with heavy padding ($\lambda = 1$) the linear regrading results.

Methods for determining smoothed weighted regradings, dependent on a single parameter $\lambda \in [0,1]$, are similarly obtained by padding *both* input and reference histograms, and matching the padded histograms. With $\lambda = 0$, no padding is applied and the weighted regrading of Definition 2 results; with $\lambda = 1$ the padded histograms both become constant and the linear regrading is achieved.

The first of these nonlinear methods uses constant padding. The input histogram $\underline{f}$ and the reference histogram $\underline{g}$ are replaced by $\underline{f}_\lambda$ and $\underline{g}_\lambda$ where

$$\text{(iv)} \quad (\underline{f}_\lambda)_j = \max(f_j,\ c_1(\lambda)), \qquad j = 1,2,\ldots,n,$$

$$(\underline{g}_\lambda)_j = \max(g_j,\ c_2(\lambda)), \qquad j = 1,2,\ldots,q,$$

$$\text{and } c_1(\lambda) = \lambda\ (\lambda \max_j[f_j] + 2(1-\lambda)|\underline{f}|/n),$$

$$c_2(\lambda) = \lambda\ (\lambda \max_j[g_j] + 2(1-\lambda)|\underline{g}|/m).$$

The smoothed weighted regrading is then determined by matching $D_{\underline{f}_\lambda}$ to $D_{\underline{g}_\lambda}$. We note that any choice of the constants c_1, c_2 such that $c_1(0) = 0 = c_2(0)$ and $c_1(1) = \max_j f_j$, $c_2(1) = \max_j g_j$ would give a range of regradings satisfying our requirements. The constants chosen here are designed to give the regradings a meaningful dependence on λ for histograms with different average and maximal values.

The second nonlinear method uses padding by an inverse linear function rather than a constant. The smoothed weighted regrading is obtained by matching $D_{\underline{f}_\lambda}$ to $D_{\underline{g}_\lambda}$ where $\underline{f}_\lambda$ and $\underline{g}_\lambda$ are now given by

$$\text{(v)}\quad (\underline{f}_\lambda)_j = \max_i \left\{ f_i \;/\; (1 + c_1(\lambda) f_i |i - j|) \right\}, \qquad j = 1,2,\ldots,n,$$

$$(\underline{g}_\lambda)_j = \max_i \left\{ g_i \;/\; (1 + c_2(\lambda) g_i |i - j|) \right\}, \qquad j = 1,2,\ldots,q,$$

with $c_1(\lambda) = m \log(1/\lambda) \;/\; |\underline{f}|$,

$c_2(\lambda) = m \log(1/\lambda) \;/\; |\underline{g}|$.

(We note that as $\lambda \to 0$, $\underline{f}_\lambda \to \underline{f}$ and $\underline{g}_\lambda \to \underline{g}$, and we may continuously define $\underline{f}_0 \equiv \underline{f}$ and $\underline{g}_0 \equiv \underline{g}$.)

Both of the nonlinear methods (iv) and (v) give the smoothed equidistributing regradings defined in [10] when the reference histogram $\underline{g}$ is constant.

Properties of the two different padding procedures are illustrated in [10]. The smoothed weighted regradings obtained by these paddings are "best" discrete approximations (in the sense of Theorem 2) to the smoothed *continuous* transformation $T_\lambda = D_{\underline{g}_\lambda}^{-1} \circ D_{\underline{f}_\lambda}$ of $D_{\underline{f}}$ into $D_{\underline{g}}$, where $\underline{f}_\lambda$, $\underline{g}_\lambda$ are defined by (iv) or (v). Both non-linear smoothings clearly provide a compromise between the weighted regrading matching $D_{\underline{f}}$ to $D_{\underline{g}}$ (obtained with $\lambda = 0$) and the linear regrading from n into m grades (obtained with $\lambda = 1$). Unfortunately, it seems difficult to show that these non-linear transformations optimize any meaningful cost functionals. The effect of the nonlinear smoothings on the break-points of the regrading is, however, well-understood in the equidistributing case [10]. Constant padding (method (iv)) limits the ratio of the largest interval between breakpoints to the smallest such interval, while padding by the inverse linear function (method (v)) bounds the ratios of consecutive intervals between break-points; that is, for method (iv) $\max_k(x_k - x_{k-1})/\min_k(x_k - x_{k-1}) \le K_\lambda$, and for method (v) $1/\tilde{K}_\lambda \le (x_{k+1} - x_k)/(x_k - x_{k-1}) \le \tilde{K}_\lambda$, $\forall k = 1,2,\ldots,m-1$, where $x_k = nT_\lambda^{-1}(k/m)$ are the breakpoints defining the smoothed regrading, and K_λ, $\tilde{K}_\lambda$ are positive constants dependent upon λ.

The discrete algorithms (i) - (v) for determining smoothed weighted regradings are all rapid and efficient to implement. Look-up tables for the regradings are produced easily by modifying the input and reference histograms grade by grade and then computing the breakpoints x_k of Definition 2 which satisfy $D_{\underline{f}_\lambda}(x_k/n) = D_{\underline{g}_\lambda}(k/m)$, using inverse linear interpolation. Simple FORTRAN subroutines for these operations are incorporated in the computer package PEEK [6] and are listed in [9]. Applications of the regrading algorithms are discussed in the next section.

6. Applications

In this section we report results obtained by the application of the discrete regrading and smoothing techniques described in sections 3 and 5 to the processing of LANDSAT image data. These techniques are highly flexible and may be used to achieve a wide variety of objectives.

Essentially any discrete regrading of the input data can be determined by the weighted regrading algorithm (5) of Definition 2- including, for example, a simple linear stretch. The linear regrading from n into m grades is obtained by matching any cumulative histogram to itself. The algorithm (5) then reduces to the explicit formula:

$$\tau_j = \underline{[}(j+\tfrac{1}{2})m/n\underline{]} \quad ,$$

where $\underline{[z]}$ is "the largest integer less than z." This regrading has been used on LANDSAT data with good results for mapping input data into a restricted grey scale range for colour display by various media, including different video monitors, and "spray gun" reproduction. It has also been used successfully for transforming "standard colours" between colour monitors.

The regrading technique may also be used for calibrating corrections to balance sensor differences, that is for "destriping". Here the histogram of each sensor in a given band is matched to a reference histogram which may be one of the sensor histograms, or an average of them all (see [3]). Good visual results are achieved by this method over land masses, but difficulties arise when the sensors do not cover comparable areas. It has been found that the use of smoothing in these cases tends to offset the poor effects.

Image enhancement by histogram "flattening" may also be achieved using algorithm (5). The equidistributing regrading is then obtained by matching the input histogram to a constant reference histogram. Experience has shown, however, that significant contrasts are often lost in images produced by this regrading. Introducing smoothing in these cases results in a much more satisfactory over-all image with good contrast and definition of land covers. The smoothing techniques also allow for the identification of special features and for finding the most satisfactory visual images of a scene (or subscene) by continuously varying the free parameter.

Image normalization requires that the histograms of two different scenes, or subscenes, be matched to the same reference histogram, and may also be achieved by the weighted regrading of Definition 2. Normalization, like destriping, involves a balance between the given information contained in the data histograms and a "prior" view that these histograms should be equal. Smoothing, in these cases, provides a simple technique for balancing image colour by continuously varying one or two parameters, and may be used with effect in mosaicing and in multitemporal analysis.

An interactive computer package, known as PEEK [6], for implementing these applications of the regrading techniques has been developed at CSIRO, Division of Land Use Research (DLUR). For a given choice of scene (or subscene) the histograms of the six sensors in each of the four colour bands are accumulated by the system, together with the total histogram of the data in each band, and the true minimal and maximal values of the grey levels in each band are returned. The following options are then available to the user:

(a) choice of "window" - to replace the calculated (default), minimal and maximal grey levels in each band;

(b) "destriping" - with linear smoothing of type (i);

(c) "enhancement" - by equidistribution, with smoothings of type (i),(iv), or (v);

(d) "normalization" - by matching a specified reference, with smoothing of type (i);

(e) colour "display" - with choice of up to three colour bands and output to video terminal or to disc file for subsequent reproduction.

Option (a) fixes an initial linear stretch of the data from the "window" onto the full set of output levels; at any subsequent operation the user may redefine the number of output grades required. For colour display, a linear regrading is applied if the number of output levels requested differs from the number of input grades to that option. Except in option (e), the data is not re-read; all other options apply operations to the accumulated histograms only, and return look-up tables which define the required regrading directly. In options (b),(c) and (d) the user chooses the smoothing parameter to be used, and in option (c) the type of smoothing to be applied. The histograms of the processed data can be printed at any stage, and the raw histogram data can be reset if necessary. The system is highly flexible and adjustable to different colour display media.

This package is currently being used at DLUR for the investigation of forest and coastal regions in New South Wales and in management studies of the Great Barrier Reef. The effects of the operations described here are illustrated for a coastal subscene of the Morgan River basin. In Figures 1 - 4, the histograms for colour bands 4, 5 and 7 are shown on a scale from 0% to 100%, where the aboslute value at 100% is given by MAX. The raw input data for bands 4, 5 and 7 were given on 105, 119 and 48 grades, and the output data were obtained on 16, 16 and 8 grades, respectively. In the Figures, the output for band 4 is shown on the grey scale levels 0 - 15, band 5 is shown on levels 22 - 37 and band 7 on levels 44 - 51.

In Figure 1 the histograms of the raw data (with a linear regrading) are displayed. Figure 2 shows the histograms of the three colour bands after destriping was applied in each band. The destriping visually improved the image, removing almost all of the original striping. It may be observed that the total histograms in each band are more peaked than those of the raw data, as a result of regrading the data of each sensor into the same average form. Figure 4 shows the histograms after enhancement by equidistribution (histogram "flattening"). The data is clearly distributed over all the grades and the MAX value is greatly reduced. In the original image only one or two major features were discernible, but in the enhanced image a wide range of different ground covers were visible; the major features, however, were no longer clearly distinct in the enhanced image. In Figure 3 the results of smoothing by padding procedure (iv) with parameter value 0.2 are shown. The histograms are much flatter than for the raw data, but are not nearly so wide-spread as in the equidistributed case. In the

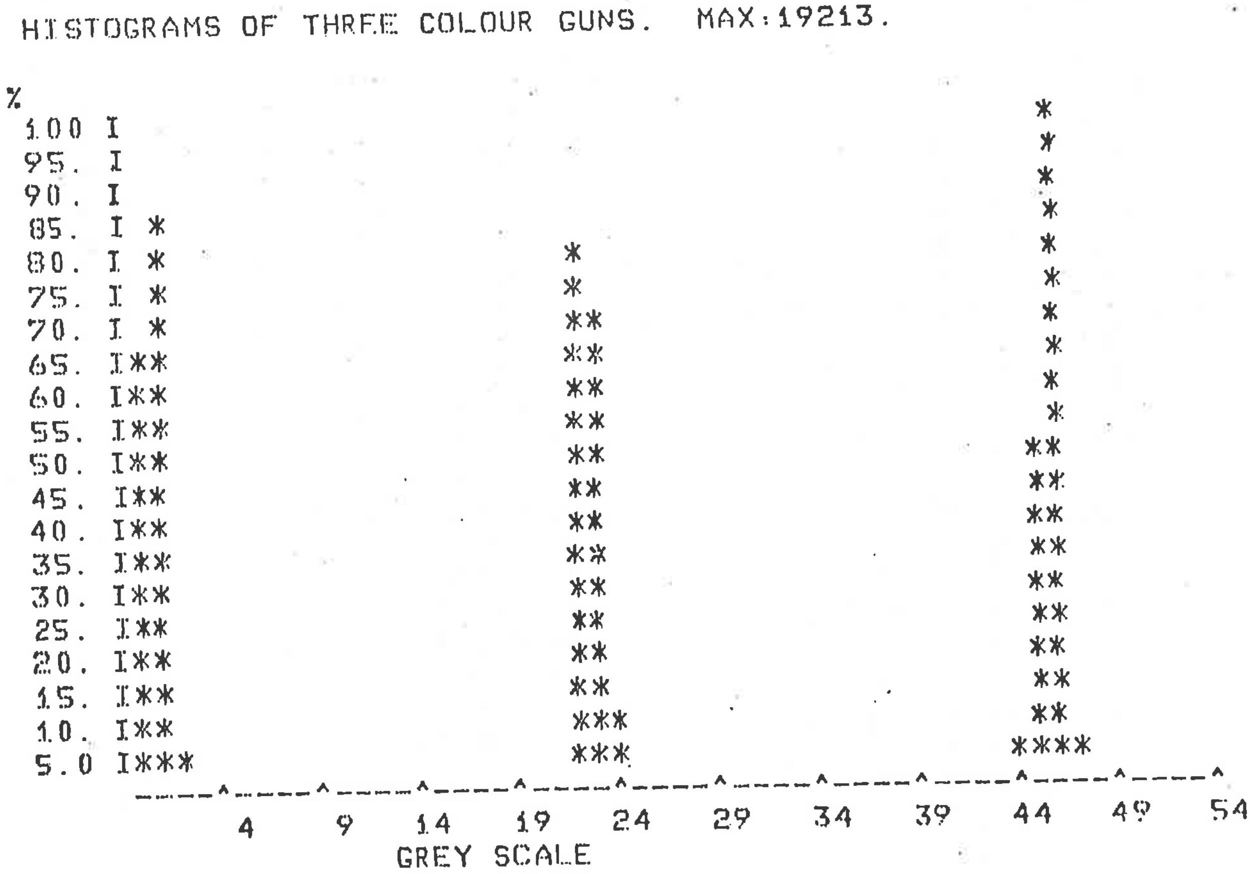

FIGURE 1

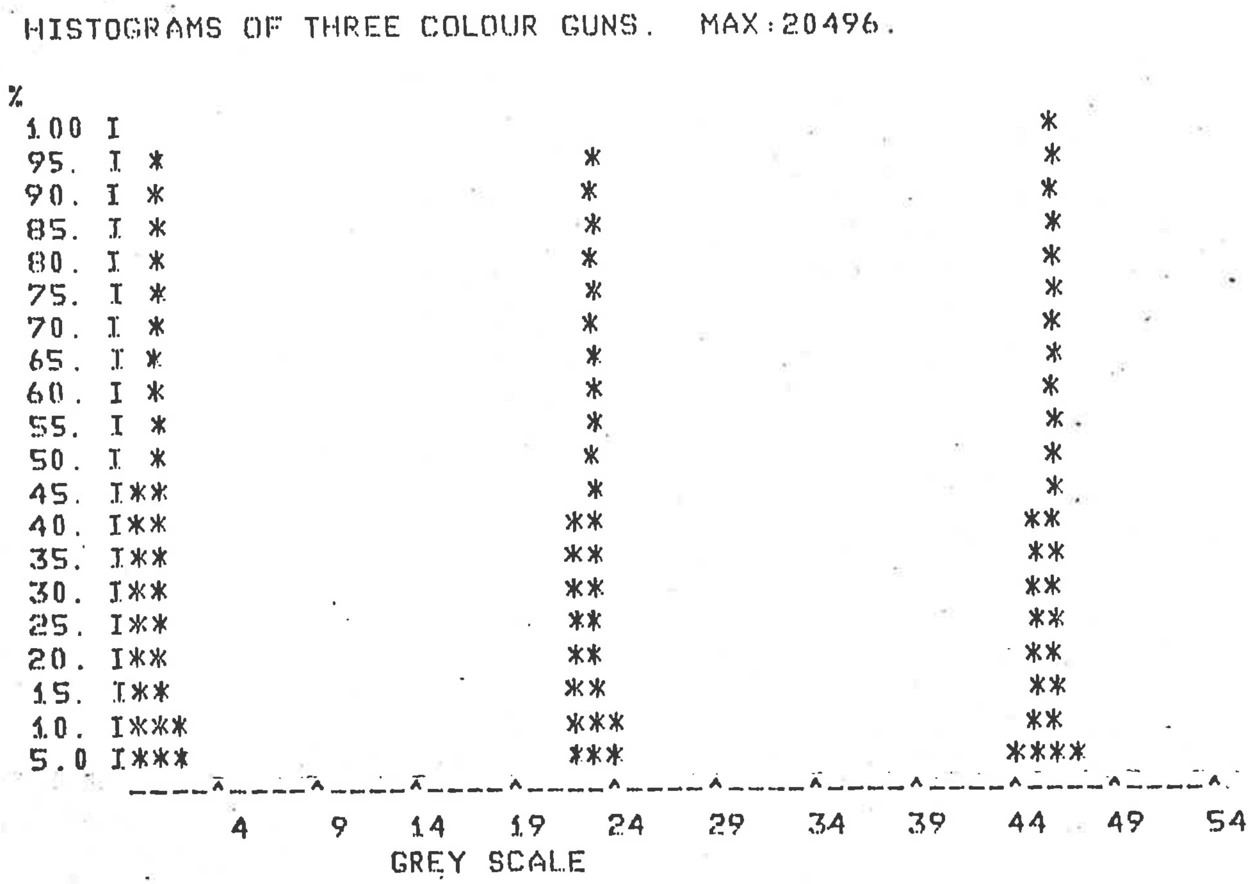

FIGURE 2

HISTOGRAMS OF THREE COLOUR GUNS. MAX: 8647.

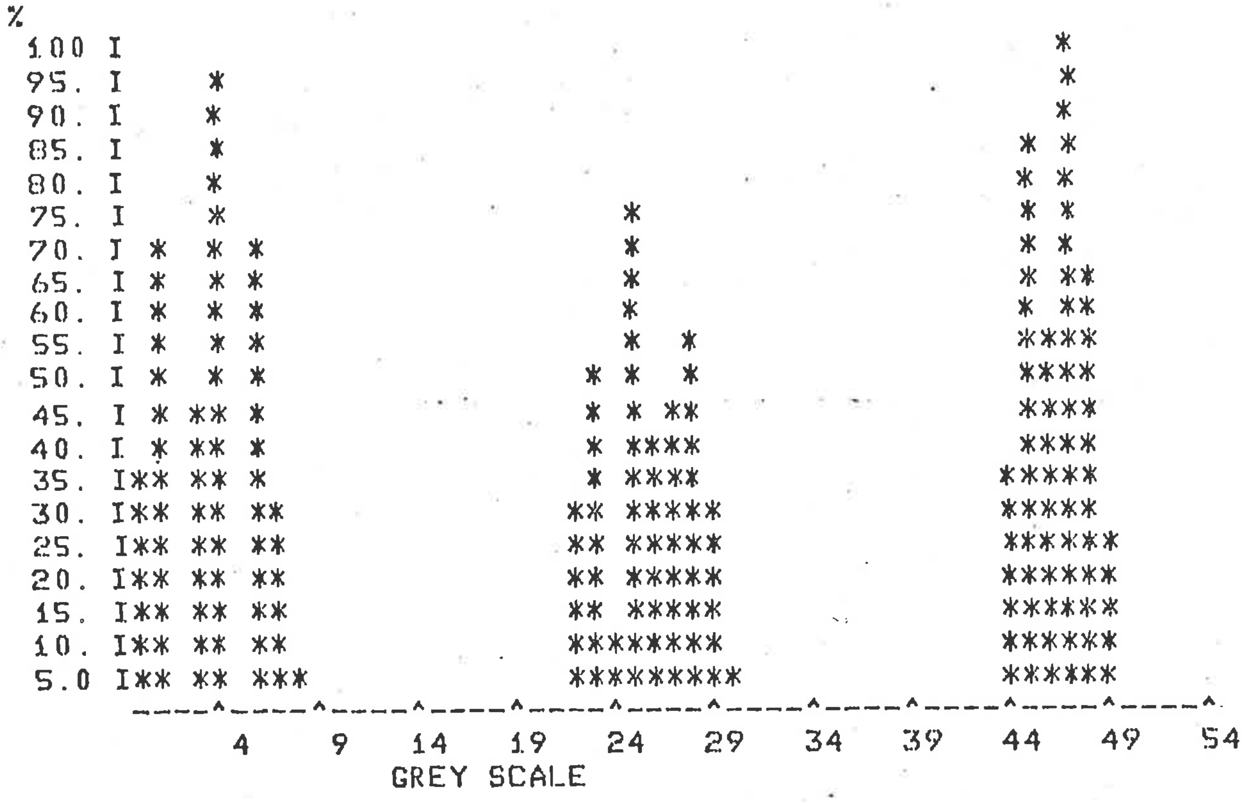

FIGURE 3

HISTOGRAMS OF THREE COLOUR GUNS. MAX: 8496.

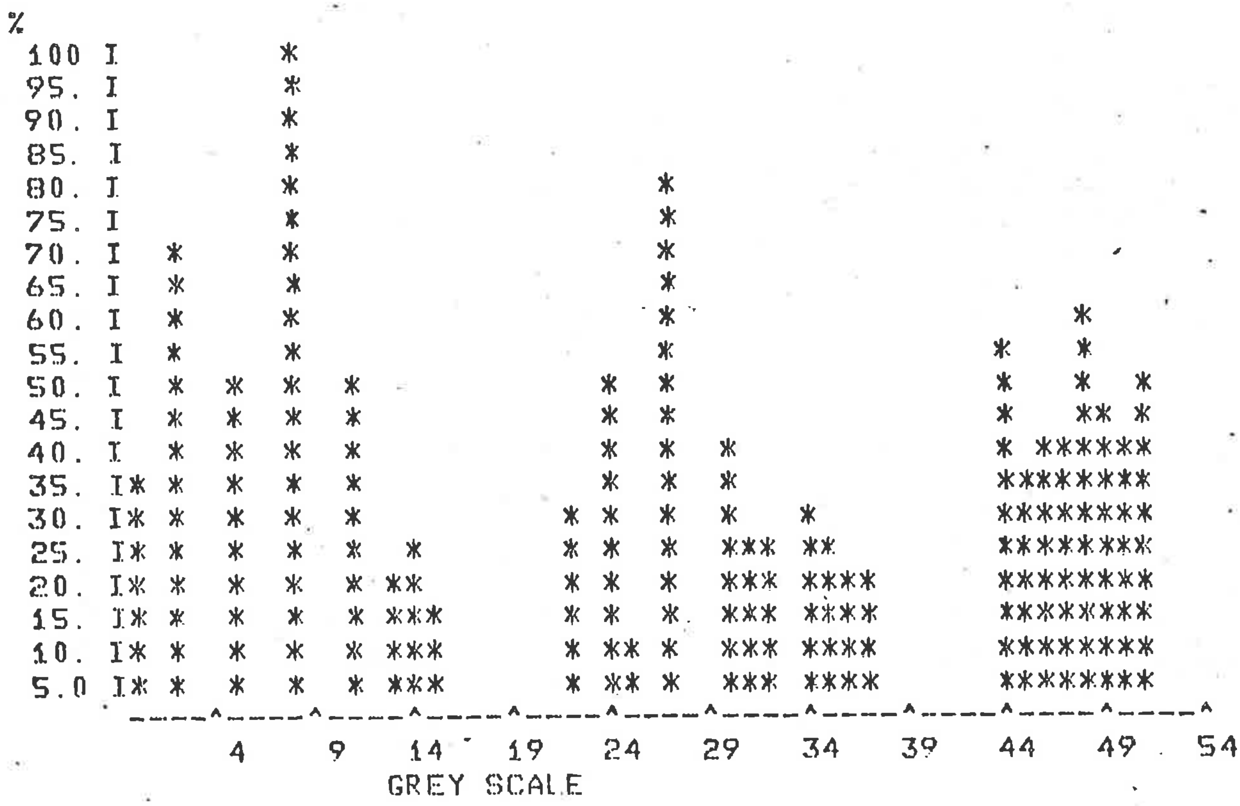

FIGURE 4

smoothed image, the major features appear clearly, but the other interesting ground features are also shown with good contrast. By varying the parameter a sequence of images were produced and the best smoothing for enhancing a particular feature could be chosen. The effects of the smoothing were principally apparent in the composite colour images and not in separate grey scale maps, and, therefore, could not easily be illustrated here in black and white.

The flexibility of the PEEK system for image processing requires choices to be made for a large number of different variables. Experience over a wide range of experiments at DLUR suggests certain strategies for these choices. For destriping, the best results seem to be achieved by regrading from the full number of input grades into an equal number of output levels. It has been observed that enhancement of destriped images by histogram flattening causes striping to reappear. By smoothing the enhancement, the effect of this striping can be minimized and a balance achieved between striping and enhancement of desired features.

It has also been noted that with enhancement by full equidistribution, an implicit "window" is defined by the majority of the data, such that data outside this "window" is reduced to a common grey level. Using a linear stretch with an explicit "window" has the same effect. Smoothed regradings also define an implicit window, but allow minor features of significance to be preserved by varying the smoothing parameter.

Experiments with smoothed equidistribution, using different padding procedures indicates that the non-linear method (iv) of section 5 is the most effective. Distinctive images occur for different choices of the parameter, showing good contrast and exposing different visible features. This characteristic arises from the nature of the padding - which moves the regrading break-points individually in distinctive jumps from the equidistributing to the linear positions (see [9], [10]).

One of the major advantages of the smoothing procedures discussed here is that, given a particular option, the required number of output levels and the smoothing to be applied, a whole range of regraded images can be generated by simply varying a single parameter. As the regradings are represented by look-up tables, which can be loaded into hardware, it is possible to display the continuous range of images directly on a colour monitor (without even re-processing the data). A striking example of the value of smoothing is given by the detection of subtle patterns in Eucalypt forests on the south coast of New South Wales where burning had been used to remove under-growth six months before a LANDSAT overpass in November, 1975. In this area the fire had caused some minor crown damage. Here, equidistribution and severe linear stretching both enhanced low contrast features, but did not show clearly the fire damage. Smoothed regrading, however, enhanced the burn pattern associated with the crown damage and produced a much more satisfactory over-all image, which had good contrast and definition for interpretation of land covers in the whole subset.

We conclude that the regrading algorithm and smoothing procedures described here offer a rich source of enhanced and normalized images, which can be rapidly and conveniently constructed, and provide the means for reducing some of the severe problems encountered in image processing.

References

[1] Haralick, R. M., "Automatic remote sensor image processing," Topics in Applied Physics, 11, 5-63, (1976).

[2] Hogan, A., "A piecewise linear contrast stretch algorithm suitable for batch LANDSAT image processing," Proc. of 2nd Australasian Remote Sensing Conf., LANDSAT 81, Canberra pp. 6.4.1 - 6.4.4, (1981).

[3] Horn, B. K. P., and Woodham, R. J., "Destriping LANDSAT images by histogram modification," Comp. Graph. Im. Proc. 10, 69-83, (1979).

[4] Hummel, R. A., "Histogram modification techniques," Comp. Graph. Im. Proc. 4, 209-224, (1975).

[5] Hummel, R. A., "Image enhancement by histogram modification," Comp. Graph. Im. Proc. 6, 184-195, (1977).

[6] Jupp, D. L. B., Kautsky, J. and Nichols, N. K., "PEEK-2, a computer package for LANDSAT image processing," CSIRO, DLUR Rpt. (1981).

[7] Kautsky, J., Jupp, D. L. B., and Nichols, N. K., "Image Enhancement by Smoothed Histogram Modification," Proc. of 2nd Australasian Remote Sensing Conf., LANDSAT 81, Canberra, pp. 6.6.1-6.6.5, (1981).

[8] Kautsky, J. and Nichols, N. K., "Equidistributing meshes with constraints," SIAM J. Sci. Stat. Comp. 1, 499-511, (1980).

[9] Kautsky, J. and Nichols, N. K., Smooth regrading of discretized data. University of Reading Department of Mathematics Numerical Analysis Rpt. 1/80, (1980).

[10] Kautsky, J. and Nichols, N. K., "Smooth regrading of discretized data," SIAM J. Sci. Stat. Comp. (to appear).

[11] Max, J., "Quantizing for minimum distortion," IRE Trans., IT-6, 7, (1960).