

1 **“DISTORTION-FREE IMAGE CONTRAST ENHANCEMENT”**

2

3 FIELD OF THE INVENTION

4 The present invention relates to methods for improving the contrast
5 between neighboring data in digital X-ray images. More particularly, the method
6 determines the actual contrast between neighboring digital image data and
7 stretches each point's intensity to the image's dynamic range without distorting
8 the image.

9

10 BACKGROUND OF THE INVENTION

11 An X-ray image is typically with standard X-ray machines using film
12 photography. In these cases the resulting X-ray image is turned into a computer
13 file by the use of digital scanning technology. More recently, there are X-ray
14 machines that use a bank of light-sensitive sensors for directly capturing a digital
15 version of the X-ray image. The X-ray image is used as a medical diagnostic
16 tool. While there are other related imagery processes that are superior, in
17 particular CAT scans and MRI-s, X-ray images are still widely used and
18 comprise the majority of such images and this is very likely to continue because
19 they are comparatively inexpensive. The current invention improves the
20 usefulness of X-ray images to doctors.

21 Due, in part, to practical limits involved with X-ray imaging, it is
22 difficult to provide an image which both defines variations in density within
23 adjacent soft tissues like lung and variations within adjacent dense tissues like
24 bone. Variations are demonstrated by changes in intensity. Each digital image
25 is associated with a dynamic range. Once developed (for film) or digitally

1 rendered as a positive image, bright intensity (usually depicted as white) areas of
2 dense matter typically occupy the high end of the dynamic range and low
3 intensity (black) occupy the lower end of the dynamic range. While the imaging
4 methods may indeed capture subtle variations, the intensity between such
5 variations are not readily detectable by the human eye. This situation is further
6 worsened when film images, traditionally transilluminated in a light box are
7 converted to digital images and displayed on a digital display. In other words,
8 variation within the black areas and variations within the white areas are not
9 easily distinguished.

10 Accordingly, methods are known for improving the contrast in
11 digital X-ray images, the most well-known of which is contrast stretching.
12 Various methods for means of accomplishing contrast stretching are the subject
13 of several issued patents.

14 For instance in US patent 5,357,549 to Maack et al. (Maack), a
15 technique is provided for stretching image intensity in only a particular area of
16 interest – such as the lung area of a chest X-ray. Maack refers to this as
17 dynamic range compression. Maack locates low frequency components,
18 determines equalization factors and applies them to the image for compressing
19 low frequency components; thus leaving the remainder of the dynamic range
20 available for higher frequency areas of the image intensities. This approach is
21 unable to enhance more than one image intensity area which has been selected
22 and is of immediate interest to the diagnostician, with loss of data the other
23 areas.

1 US patent 5,835,618 to Fang improves on Maack using a method
2 of dynamic range remapping for enhancing the image in both dark and bright
3 intensity areas. This remapping or correction technique amounts to smoothing
4 the data (such as through a low-pass filter), determining the data mean,
5 adjusting the smoothed data to the mean, and then applying that smoothed,
6 adjusted data to the original data. Two curves of adjusted data are provided,
7 each offset from the mean, one of which is offset upwardly by a constant ($\Delta 1$)
8 and one downwardly by a constant ($\Delta 2$) for establishing upper and lower
9 thresholds and separating the data into two distinct populations. Constants $\Delta 1$,
10 $\Delta 2$ define a range. Then, separate additive or multiplicative algorithms are
11 applied firstly to the original data within the range, and secondly to data outside
12 the range. For instance, in the additive option, the original data within the range
13 is adjusted by the difference in the data's original and mean intensity, this
14 difference being scaled by a user-defined control parameter between 0 and 1.
15 Then the adjusted data is scaled to the full dynamic range of the image. Data
16 outside the range is adjusted using a different algorithm.

17 Unfortunately, the method of Fang treats data within and without
18 the range differently with the result that artificial details and other distortions
19 result, such as the creation of a discontinuity at the range boundary. Data just
20 inside the range data just outside the range can result in very different values,
21 distorting stronger signals. The adjusting algorithms produce these distortions
22 whenever the smoothed image deviates from the mean of the entire image, the
23 magnitude of the deviation affecting the magnitude of the distortion. Fang
24 recognizes that the user can manipulate the degree of dynamic compression.

1 However, to minimize distortion, the user must manipulate each of: the
2 attenuation of the correction, the upward offset, and the downward offset. For
3 instance, the larger the chosen range, then the more the distortion is minimized
4 but also the more the subtle details are lost. A smaller range can enhance
5 weaker signals, however, the stronger signals become badly distorted. That is to
6 say, the process requires time and effort and experience on the part of the user
7 to manage three parameters to try to both minimize the image distortion while
8 maximizing the image enhancement.

9 Each of the above methods of image enhancement result in a loss
10 or distortion of the original data. Loss and distortion present artifacts, which can
11 seriously compromise a radiologist's or other diagnostician's interpretation of the
12 image. Ideally, if intensity variations do exist between neighboring data in an
13 image, the contrast between them should be maximally enhanced for detection
14 by the diagnostician without the introduction of artifacts, and regardless of
15 whether the intensity variations are in the light areas or the dark areas of the
16 image.

17 An optimal approach should use control parameters which are
18 independent in their nature. Interrelated variables require the user to make
19 compromises, sacrificing one result in part so as to achieve part of another
20 result. Unfortunately, such decisions require a user to gain expertise in the
21 background to the technique before it could be properly implemented. Further,
22 such control parameters need to be robust so that small changes in a given
23 parameter result in manageable changes, do not cause wild results, and even
24 poor choices should give "livable" results.

SUMMARY OF THE INVENTION

The current invention is a process of maximizing the detail between neighboring points of an image without distorting the image and without adding artificial details. Locally, all contrasts in the image are corrected in the same manner.

In the preferred embodiment, the process is applied to images where variation between neighboring points is digitally significant but the contrast is too low for the human eye to discern. Such cases include X-ray images. There is more information available in an X-ray image than one might guess. The current invention enhances this information in order to make a more revealing, processed X-ray image. The achievements of the current invention are:

- improved contrast - for aiding visual interpretation by doctors and other diagnosticians;
- good correlation of the processed image and the input image – e.g. the processed image of a chest image still resembles the original chest image;
- maximized detail - every part of the X-ray image, is maximized for maximum possible visibility, contrast being improved between subtle variations within dark regions and within light regions;
- automation of the image enhancement - thus it does not require the end user to have prior image enhancement experience; and

- 1 • avoiding distortions and artificial details - thus avoiding addition
2 of false information to the already challenging task of image
3 interpretation.

4 In a broad aspect of the invention, an image is enhanced through
5 the local enhancement of the contrast between neighboring point intensities by
6 fitting a first low frequency upper curve or surface to the local maximums and
7 fitting a second independent, and low frequency lower curve or surface to the
8 local minimums, the space or volume between forming a fairway. The raw
9 image data resides within the fairway. A local range between the fairway local
10 maximum intensity and fairway local minimum intensity is extracted for each
11 point. A local scaling factor is determined as the ratio between the local range
12 and the dynamic range for the image. Each point is then scaled by its local
13 scaling factor so as to maximize the variation in intensity between it and its
14 neighboring points.

15 Preferably the low frequency curves are determined using a two
16 dimensional moving average, accommodating variations in point intensities both
17 between neighboring columns and neighboring rows.

18 Preferably even outlier points falling outside the fairway are
19 enhanced, not by a distorting truncation process but, through a histogram
20 correction which rescales the enhanced image based upon a determination of
21 intensity range of the outliers. More specifically, the locally scaled outliers have
22 intensities outside the image's dynamic range and thus are temporarily stored or
23 preserved at a higher precision.

1 Next, all of the data stored as high precision intensities, including
2 outliers, are placed in a histogram having a predetermined or expanded range
3 greater than the image's dynamic range and large enough to capture
4 substantially all of the outliers. A histogram count is made and a predetermined
5 trim rate is applied to the histograms low end and top end. Because the fairway
6 trends both large variations and small variations in intensity between neighboring
7 points, outliers can appear outside the fairway at almost any local position in the
8 image. As a result, the most deviant of the outliers, selected in this manner, are
9 usually widely dispersed and thus only minimally affect the image enhancement
10 when trimmed. The trimmed image points have a trimmed range which defines
11 a new range having a minimum intensity and a new maximum intensity. All the
12 points are scaled a second time, this time at a scaling factor determined as the
13 ratio of the trimmed range and the image's dynamic range.

14 Preferably the low frequency curves are determined using a two
15 dimensional moving average and, more preferably, using filter box mean
16 determination which markedly reduces the number of calculations required by
17 taking the preceding filter box sums and merely subtracting the lagging point
18 intensities and adding the leading point intensities to obtain a new filter box sum.
19 Merely normalizing each box sum, by dividing by the sum by the number of
20 points in the box, completes the moving average.

21

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is an X-ray of a chest in which many details are immediately visible, but other detail appears lost in the seemingly uniform light areas of the pelvis and vertebral column, and the uniformly dark vascular areas of the lung;

Figure 2 is a digitally enhanced image of the X-ray of Fig. 1, after application of the preferred method of the present invention;

Figure 3 illustrates a fanciful representation of the mathematical selection of one sample column of data from the image of the X-ray of Fig. 1 for processing using the method of the present invention. The sample column is shown extracted from the image of Fig. 1 and set aside, the extracted column being shown with an exaggerated width for ease of viewing;

Figures 4a and 4b are related graphs which illustrate the variation in data intensity within the sample column of data of Fig. 3. In Fig. 4a, the image's bottom to top corresponds to the graph's left to right. Fig. 4b is a copy of the corresponding sample column, again exaggerated in width for viewing;

Figure 5a and 5b are graphs of a simple prior art form of contrast stretching, applied to the data of Fig. 4a. Fig. 5a illustrates the data of Fig. 4a with a smoothed low frequency mean plotted thereon and Fig. 5b which illustrates rescaled output data based upon local data variations from the mean which have then been being stretched to the full dynamic range;

Figures 6a – 6c illustrate related graphs, in which the process of the present invention was applied to the data of Fig. 4a. Fig. 6a illustrates the

1 upper and lower bounding curves of the fairway. Fig. 6b illustrates the
2 corresponding output data. Fig. 6c is a copy of the corresponding image
3 column, exaggerated in width for viewing;

4 Figures 7a and 7b are graphs of the raw input data of Fig. 4 having
5 scaling fairways plotted thereon; the fairway of Fig. 7a being an ideal, freehand
6 sketch; and the fairway of Fig. 7b being calculated using a preferred moving
7 average process of the present invention;

8 Figure 7c is a fanciful and partial representation of 2D bounding
9 surfaces with the raw data residing therebetween;

10 Figures 8a and 8b are isolated images which result from plotting
11 the low frequency upper and lower bounding curves or smoothed trending
12 functions respectively;

13 Figure 9a and 9b are graphs which illustrate and compare the raw
14 input data and a first fairway correction scaled data stored at a precision greater
15 than the dynamic range;

16 Figures 10a and 10b are graphs which illustrate the initial steps in
17 determining the upper trending function. Fig. 10a illustrates a low frequency
18 mean superimposed on the raw input data and Fig. 10b illustrates the residual
19 points at or greater than the mean;

20 Figures 11a and 11b illustrate the iterative convergence
21 determination from the data of Fig. 10b wherein a new mean is iteratively
22 determined for the residual data remaining after applying the previous means so
23 as to reduce the data population to a new and smaller number of residual points;

1 Figure 12 is a graph illustrating data mirroring at the image edges;

2 Figures 13a and 13f illustrate pseudo-code and variable definitions
3 respectively for applying an efficient moving average technique to the image
4 data;

5 Figures 13b – 13e are fanciful illustrations of the moving average
6 technique for improving the efficiency of the summations, Fig. 13b for one 7-
7 point row of column summations incremented one column; Fig. 13c using a 7x7
8 box column by column of row summations; and Fig. 13d and 13e using a 2D box
9 with a row by row of column summations, and then incremented one column;

10 Figures 14a – 14c illustrate pseudo-code and Figure 14d illustrate
11 variable definitions for demonstrating one embodiment of the iterative method for
12 applying the moving average data of Fig. 13a and determining the upper and
13 lower bounding curves;

14 Figure 15a illustrates pseudo-code for an improved convergence
15 technique to replace the technique of Fig. 14b;

16 Figure 15b is a graph which illustrates a faster convergence,
17 iterative determination of the upper bounding curve for the data of Fig. 4a and
18 according to the pseudo code of Fig. 15a;

19 Figure 16 is a graph which illustrates the resulting upper and lower
20 bounding curves of a fairway obtained using the faster convergence technique of
21 Fig. 15a;

22 Figures 17a and 17b illustrate related graphs, more particularly,
23 Fig. 17a illustrates the upper and lower bounding curves of the fairway according

1 to Fig. 16, and Fig. 17b illustrates the corresponding output data after a second
2 histogram correction step;

3 Figures 18a and 18b illustrate pseudo-code and Figure 18c
4 illustrates variable definitions for applying the upper and lower bounding curves
5 determined according to Figs. 13a, 14a-14d and application of the first fairway
6 and second histogram corrections for outputting corrected image data such as
7 the image of Fig. 2;

8 Figures 19, 20a, and 20b illustrate output images which result from
9 variation of the extent of noise suppression. Specifically,

10 Fig. 19 applies minimum bound separation (surf_mdff) of 1.5%,

11 Figs. 20a and 20b apply a minimum separation of the bounding
12 curves of 3% and are identical images to Fig. 2, reproduced for comparative
13 convenience, and

14 Fig. 21 applies minimum bounding curve separation of 12%; and

15 Figures 22 – 25 illustrate output images which result from a
16 variation of the extent of filtering (imag_pcmt), as it applies to the demonstrated
17 moving average technique. Specifically, Fig. 22 applies a filter of 5% of the
18 image size and again is an identical image to Fig. 2, reproduced for comparative
19 convenience, Fig. 23 applies a filter at 7%, Fig. 24 applies a filter at 10%, and
20 Fig. 25 applies a filter at 20% of the image size.

21

1 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

2 An Image's Characteristics

3 Every digital device has a dynamic range which is a measure of it's
4 ability to record relative energy fluctuations. Here, in the case of an X-ray, the
5 intensity of fluorescence directly emitted from a phosphor plate and captured by
6 a charge-coupled (CCD) device or the transillumination through film in a
7 scanner. This dynamic range is usually set by the recording or scanning
8 resolution and the digital storage means (data file). A typical digital file uses 8 bit
9 storage which gives a dynamic range of 0 - 255. Also known are files using 10
10 bit storage which gives a dynamic range of 0 - 1024, 12 bit storage which gives a
11 range of 4096, and 16 bit storage which gives a range of 0 - 65,535. For the
12 purposes of this application, Applicant illustrates the process of the present
13 invention using an 8 bit storage (0 - 255), the process being independent of the
14 actual image dynamic range, and not intending that the disclosed range being
15 limiting.

16 It is an objective of this invention to apply an image enhancement
17 technique for maximizing local contrast and yet avoiding distortion associated
18 with exceeding the image's dynamic range. The present invention recognizes
19 that one part of maximizing local contrast is to stretch the local range to the
20 dynamic range. Relative contrasts, from point to nearby point, are preserved at
21 the local level - each significant sized sub-section of the processed image must
22 make use of most of the dynamic range of the technology being used. Another
23 part of the enhancement technique is to determine the local ranges to be
24 stretched. Another part is to minimize or prevent distortions associated with

1 clipping at the extremes of the dynamic range. Clipping occurs where contrast
2 stretching scales an intensity value to a theoretical value beyond the minimum or
3 maximum. For example, for a dynamic range of 0-255, scaled output values in
4 the hypothetical range of 256 - 300 can only be recorded as 255 and thus details
5 within such areas are lost.

6 Having reference to Fig. 1, an image 10 is provided for image
7 enhancement. The present invention was applied to generate the output image
8 30 of Fig. 2.

9

10 Sample Image

11 Image 10 comprises a data array of about 878 by 1007 pixels,
12 each pixel having being stored with a sample dynamic resolution of 0 - 255
13 levels of intensity. The majority of this disclosure illustrates application of the
14 present invention as it applies in a one dimensional manner to only one sample
15 of the 878 columns of data. As is also described herein, the process actually is
16 performed in two dimensions using nested operations for analyzing both rows
17 and columns.

18 This one dimensional, mid-processing columnar example is
19 introduced as follows.

20 Referring to Fig. 3, a column of data 11, selected from about 2/3 of
21 the way across the image, from the left, is shown as physically being extracted
22 from the image 10 (illustrative purposes only).

1 Turning to Fig. 4a, the raw data of the sample column 11 is plotted
2 as high frequency curve 14. The left-hand side of the graph corresponds to the
3 bottom of the X-ray image 10. The horizontal axis corresponds to an index of
4 the number of dots or points counted from the bottom of the image 10 (0-1007).
5 The vertical axis is the relative intensity of each point in terms of a fraction of the
6 dynamic range. For this image 10, the value 1.0 means as white as possible
7 (the top of the dynamic range limit of 255 for this particular file) and 0 is as black
8 as possible (the lower dynamic range limit). While substantially the entire
9 dynamic range (0 to 1) of black to white was used in this image, this particular
10 column or subset of data only uses only about 75% of it. For example, the
11 relative intensity of the data at the spine of image 10 is nearly 1.0.

12

13 Example of Poor Utilizations of Dynamic Range

14 In Fig. 4b, the sample column data 11 is illustrated shown
15 corresponding to the orientation of the raw data curve 14.

16 Having reference to the results of one prior art technique, in Fig. 5a
17 and 5b, it is known in the art to preserve local contrast by removing the low
18 frequency components from raw data 11 of Fig. 4a and stretching the resulting
19 range to the full dynamic range. In Fig. 5a, a calculated and smoothed low
20 frequency curve or trend 15 is superimposed on the real data 14 from which it
21 was calculated using known techniques.

22 In this prior art technique shown in Fig. 5b, an output curve 16 is
23 produced as a result of subtracting the trend 15 from the raw data 14 and then
24 the minimum and maximum for the data are scaled to the dynamic range. This

1 simple approach preserves the local detail, however, the residual data graph of
2 Fig. 5b is inadequate in terms of how the dynamic range is used locally. Note
3 that the data at 17, having a very small peak-to-trough range at points 330 to
4 350, data at 18 at 580 to 620 and the data at 19 that spans from 780 to 850
5 would ideally be corrected to a similar use of the dynamic range. Note that the
6 range of data at 18 uses about only about one fifth of the dynamic range
7 whereas the data at 17, uses about half, and the data at 19, uses nearly all of
8 the dynamic range. This method treats all sets of data 14 the same and thus
9 does not maximize the visibility of the detail within an X-ray image. Referring to
10 prior art Fig. 5b, one can see that due to the large range intensity data at 19,
11 causes the small intensity range for data at 18 to only be stretched to about 25%
12 of the dynamic range.

13

14 Improved Utilization

15 To optimally stretch all areas of the image 10, even those originally
16 utilizing both a small portion of the dynamic range and those utilizing a large
17 portion, the local contrasts must be handled independently. One way to do this
18 is to identify and choose a sub-section of neighboring data and scale it
19 substantially independent of the remainder of the data 14.

20 Having reference to Fig. 7a, three curves are illustrated, the input
21 or raw data 14 of Fig. 4a and also, sketched freehand, an upper curve 20 and a
22 lower curve 21. These curves 20,21 represent one concept of providing an ideal
23 bounding of the raw data 14. These bounding curves 20,21 have a low
24 frequency and further have the characteristic that they graze the raw data's local

1 extreme values; the upper bound grazing local maximums ($20_i, 20_{i+1} \dots 20_n$) and
2 the lower bound grazing the local minimums ($21_i, 21_{i+1} \dots 21_n$).

3 The low frequency aspect of the two curves 20,21 ensures that the
4 local relative intensities or local contrast are preserved, just as in the case of the
5 single trend curve 15 of Fig. 5a. However, unlike prior art Fig. 5b and the prior
6 art system of US 5,835,618 to Fang, the bounding curves 20,21 are
7 independently responsive to the actual data 14 and not on a mean value 15.

8 Together, the upper and lower bounding curves 20,21 form a
9 bounding fairway 22 which can be analogous to a defined data range which will
10 be ultimately stretched to the dynamic range.

11 Now it can be seen that the intensity range data 18 at 580 – 620
12 and data 19 at 780 - 850 are both directly scalable to about 100% of the dynamic
13 range, in contradistinction to the mere 25% of prior art Fig. 5b.

14 The challenge for the present invention is to determine where
15 these upper and lower bounding curves 20,21 are to be placed and how, and to
16 overcome complicating limitations imposed by computer processing systems.
17 The remainder of this specification discloses methods for determining these
18 bounding curves 20,21.

19

1 The Ideal Case

2 A mathematical correction, as introduced in Equation 1, (also a R,
3 Fig. 18a) applies the bounding curves 20,21 to adjust each raw input data point,
4 one at a time. An input data point is a pixel or point having a known intensity, for
5 the selected column 11 from Fig. 3, selected from the sample images 1007
6 points or rows in each column 11.

$$7 \quad \text{Output Data} = \frac{(\text{Image Point} - \text{Lower Bound})}{(\text{Upper Bound} - \text{Lower Bound})} \text{Dynamic Range} \quad \text{Eqn. 1}$$

8 where:

- 9 • Image Point - is the value of the input data at a given point;
- 10 • Lower Bound - is the value of the minimum curve at the very
- 11 same point;
- 12 • Upper Bound - is the value of the maximum curve at the very
- 13 same point;
- 14 • Dynamic Range Limit - is the Dynamic Range Limit for the file
- 15 or system (here 256); and
- 16 • Output Data - is the corrected data value.

17 The success of this first correction depends upon the method of
18 determining the upper and lower bounding curves 20,21, determined from the
19 raw input image data 14 and two predetermined variables or parameters.

20 If one can achieve the ideal freehand drawn case, no data points
21 fall outside the bounding curves 20,21 and Eqn. 1 is a full solution, all local
22 contrasts being preserved and scaled substantially independently of other areas
23 of the image. Unfortunately, limitations of numerical methods and the law of
24 diminishing returns dictates that the bounding curves will not necessary
25 encompass all the points.

1 Having reference to Fig. 7b, calculated curves 20,21 have a
2 similarity to the previous hand-eye estimated best upper and lower bounding
3 curves. Unlike the ideal case however, there is a difference; namely that there
4 are some outlier points 23, throughout the graph, which fall outside of the
5 calculated fairway 22; in other words points in the raw data 14, and image 10
6 overall, still end up being clipped somewhat, although using the preferred
7 additional techniques disclosed herein, the impact is rendered substantially
8 insignificant.

9 For the best results, a two-step correction is implemented to handle
10 the practicality of these calculated fairways 22. Additional techniques, including
11 use of a histogram, are used to handle the outliers and avoid loss and distortion
12 of the image information contained therein.

13 For example, after the two-step correction, the entire length of the
14 upper curve 20, for the sample column of Fig. 3,4a, produces a failure rate of
15 clipped data of only about 0.2% of the population of the column's data 14. Any
16 clipped data turns out to be the least important of the points as they are spread
17 throughout the corrected image 30 with the result that no neighboring group of
18 points of the original image 10 is distorted. So, even though the bounding
19 curves 20,21 themselves may be associated with a failure rate of about 10% at
20 the first step, the second and final correction results in the much lower failure
21 rate of 0.2%.

22

1 The First Fairway Correction

2 If the outlier points 23, which fall outside of a calculated fairway 22
3 (e.g. see Fig. 9a at about point 330 and 350), are immediately constrained to the
4 dynamic range, the image data integrity would suffer. Such constraints occur
5 when using computer processing systems utilizing variables declared as having
6 the same limited precision as that used for the input image 10. For instance, an
7 outlier point falling above the upper bound could hypothetically be
8 mathematically scaled to 280. Based on a dynamic range of 0 – 255, this outlier
9 would be clipped to 255 and the relative intensity data would be lost.
10 Accordingly, using digital storage means, to avoid clipping outliers at the fairway
11 correction, a temporary data precision can be used which is higher than the
12 dynamic range and the data is maintained therein until an additional and second
13 correction is applied.

14 Eqn. 1 is applied to the raw data of image 10 and the image results
15 are temporarily stored using a variable type having more precision than the
16 image dynamic range limit. For example, for a dynamic range limit of 0-255 (8
17 bit), 1023 (10 bit), or 4095 (12 bit), as are currently common values for stored
18 image data, a standard 2 byte (16 bit) signed integer variable would suffice. The
19 2 byte integer allows for a larger range, to approximately +/- 32,000 and includes
20 negative numbers as well. Accordingly, a 2 byte signed integer can holds the
21 outliers without incurring any distortion effects.

22 Having reference to Fig. 9b, applying Equation 1 to the sample
23 column data 14 of Fig. 3,4a, and using the dynamic range limit of 0-255 as an
24 example, the calculated temporary data 24 ranges almost from -50 to 350. The

1 resulting data ranges would produce distortions for outlier points lower than 0
2 and greater than 255 except that they are being temporarily stored in the higher
3 precision 2 byte integer variables.

4 A second histogram correction is applied to appropriately scale the
5 resulting data 24 of this fairway correction, Fig. 9b, to the actual dynamic range.
6

7 The Second Histogram Correction

8 The temporary high precision image data of Fig. 9b is, for the most
9 part, forced back into the dynamic range of the image. Various means can be
10 employed.

11 One approach is employs a forced failure rate. Simply, a
12 histogram of the temporary data 24 is calculated and a very small level of outlier
13 points 23 are sacrificed for the good of the majority of the data 14. A reasonable
14 first pass is to clip or trim about 1%; that is, the bottom 1% and the top 1% of the
15 data is trimmed and the remaining points are scaled back to the image's
16 dynamic range.

17 Accordingly, as disclosed in greater detail below and in Fig. 18a,
18 the second correction uses a histogram to assembly the fairway correction
19 temporary data 24 of Fig. 9b and trims the lowest and highest 1% of the data.
20 The remaining trimmed range of point values in the higher precision temporary
21 variables are then scaled back again to the lower precision of the dynamic range
22 of 0 to 255. The resulting output data 34 is shown in Fig. 6b with the entire
23 output data for all points being represented in the output image 30 of Fig. 2. The

1 effect on the output data 34 for the sample column 11 is illustrated in Fig. 6b,
2 wherein contrast is significantly increased, and to the same dynamic range,
3 regardless of whether the points were originally neighboring in a substantially
4 low intensity (dark) region of the image 10, or substantially high intensity (light)
5 region. Compare column 31 of Fig. 6b, representing the enhanced data 34, with
6 the original column 11 and data 14 of Fig. 4b.

7 Notice in Fig. 6a, that the fairway's upper and lower bounding
8 curves 20,21 define local dynamic ranges which are all nearly equal, even
9 though they encompass each of the medium range of intensity data 17 present
10 at points 330 to 350, the small range of data 18 at 580 to 620 and one of the
11 largest local ranges of data 19 at 780 to 850. In application, the trim rate is
12 found to be far less than 1%.

13 This process of forcibly trimming only a few of the highest and
14 lowest values, stored in the temporary high precision data, means that only a few
15 sparsely spaced points in the column of data points are affected at all.
16 Accordingly, there is not much sacrificed in terms of image quality. Simply, a few
17 isolated extremely bright (high intensity) points, which theoretically should have
18 been even brighter still, were trimmed. However, it is false economy to save
19 those few points, at a cost of an overall lower contrast for the entire image.

20

1 Determining the Upper and Lower Bounding Curves

2 For the first fairway correction, some criteria are applied to
3 determine the extreme upper and lower bounding curves 20,21. The criteria
4 define the rules which, in one sense, attempt to best duplicate the freehand
5 curves of Fig. 7a. The criteria include that:

- 6 • the maximum and minimum curves forming the upper and lower
7 bounding curves 20,21 are made of low frequency components;
- 8 • the frequency of each curve 20,21 is low enough (smooth
9 enough) to preserve the local detail and yet is high enough to
10 separate sub-areas of interest (for example, so as to preserve
11 each of the contrast within a rib of Fig. 3, contrast within the
12 lung between the ribs, and contrast between a rib and the lung);
- 13 • the maximum and minimum curves 20,21 graze the local
14 extreme data points $20_i - 20_n$, $21_i - 21_n$; and
- 15 • the width of the calculated fairway 22 is similar to the one of
16 Fig. 7a, yet it is understood that some of the points 23 will end
17 up outside of the fairway 22.

18 In the simplest terms, and using one numerical method or another,
19 two bounding curves 20,21 are fit to the raw input data. One upper curve 20
20 having a low order or low frequency is fit, as best as possible, to the maximum
21 intensities for the data 14. A second curve 21, also having a low order or low
22 frequency is fit, as best as possible, to the minimum intensities for the data 14.
23 Ignoring any outliers 23, each point which is between the two bounding curves

1 20,21 can be scaled to the image's dynamic range. Because it is inevitable to
2 have outliers 23 with any numerical method used to determine the upper and
3 lower bounding curves 20,21, the above described second histogram technique
4 is provided to minimize their effects on the image enhancement.

5 One method of determining the upper and lower bounding curves
6 20,21, is to apply a moving average. This simply means that a trend 15 is
7 produced which is made up of a series of points, each of which is an average of
8 its surrounding or neighboring data. A one dimensional (1D) analysis would
9 simply be to sum the intensity data 14 contained within a predetermined interval,
10 and then repeat that for each point up and down the sample column 11. A 2D
11 analysis averages all of the intensity data within an area surrounding each point
12 as it is analyzed. For the purposes of this disclosure, the concept of an area as
13 a rectangular box is implemented, with the average or mean being applied at its
14 center. This box or filter box is dimensioned as a specified number of rows
15 either side of the point (as would be the case in a solely 1D analysis), and a
16 number of columns either side of the point.

17 In one embodiment it is convenient to apply a moving average to
18 the image of Fig. 1 by selecting each of the image's individual columns 11 one-
19 by-one. Other approaches, such as a row-by-row analysis can be employed,
20 affecting only the order of calculation.

21 Upper and lower bounding curves 20,21 are determined for the
22 selected column 11. Use of a 2D analysis ensures that, not only is continuity
23 maintained between adjacent points in a column 11 but, also continuity between
24 adjacent columns 11,11,11 ... is maintained. In fact, using a 2D analysis, the

1 example column-by-column analysis provides exactly the same effect as a row-
2 by-row analysis.

3 Each column 11 is made up of a plurality of neighboring points, one
4 point in each column-intersecting row, each of which is associated with an
5 intensity value (within the dynamic range). A first moving average is performed
6 so as to place a low frequency or smoothed curve 15 through the mean of the
7 raw intensity data. This mean curve divides the raw data 14 into a residual
8 upper population 40 of points; those having an intensity greater than the first
9 mean curve 15. It also divides the raw data 14 into a residual lower population
10 41 of points; those having an intensity less than the first mean curve 15.

11 Then an iterative process is performed, again using a moving
12 average technique, to calculate successive and new mean curves for each of the
13 residual upper and residual lower population of points 40,41. As the populations
14 40,41 become smaller each iteration, each successive mean curve lies shifted
15 further and further towards the maximal and minimal intensities 20_i - 20_n and 21_i -
16 21_n of the raw data 14 respectively.

17 It is understood that other curve fitting techniques could be used,
18 and the example moving average technique can be further optimized. Some of
19 these options are disclosed below.

20 In greater detail, as applied to the determination of the upper
21 bound only, and having reference to Figs. 10a – 12b, a moving average
22 technique is applied to the raw data of Fig. 4a. Turning to Fig. 10a, a first mean
23 low frequency curve 15 is calculated and superimposed on the raw data 14 of
24 Fig. 4a. The result is a partial low frequency curve 24_i ; partial because only the

1 residual population 40 of points greater than the low frequency curve 24_i have
2 been plotted.

3 In this embodiment, determination of the upper bounding curve 20
4 and the lower bounding curve 21 are performed in the same manner, iteratively
5 determining successive mean curves and reducing the respective data
6 populations 40,41. The description for the upper bounding curve 20 applies
7 equally to the lower bounding curve 21.

8 For determining the upper bounding curve 20, the first step is to
9 replace input data that is lower than the calculated mean low frequency
10 components data. Having reference to Fig. 10b, the resulting data 24_i is
11 depicted, and for clarity, the lower points (used for determining the lower
12 bounding curve 21) are shown omitted from the graph. In other words, the first
13 step is to identify the residual upper population 40_i and create an artificial image
14 that is composed, on a point by point basis, of the greater of the input data and
15 the previously calculated low frequency component's data, in this case the first
16 mean curve 15. The resulting population 40_i of artificial data has detail at the
17 local peaks and has only low frequency components at the local valleys.

18 From this residual population 40_i of artificial data, a successive and
19 new mean curve 24_{i+1} is calculated. The new low frequency components of this
20 residual population, being composed of the greater values of the real data local
21 peaks and the valleys of the low frequency mean curve 24_{i+1}, on a point by point
22 basis, are calculated and a second residual population 40_{i+1} is again created
23 being composed of the maximum of the residual population and its low
24 frequency components.

1 As shown in Fig. 11a, this process is repeated in an iteration loop
2 from i to n iterations, each successive mean curve retiring to become a previous
3 mean curve 24_i as each successive mean curve 24_{i+1} is calculated.

4 The same iterative approach can be used to find the lower
5 bounding curve 21. The only difference is that the first and subsequent residual
6 lower population 41_i to 40_n (artificial images) are calculated as the lesser of the
7 input data and successive low frequency components of successive curves 26_i
8 to 26_n . Finding both the upper and lower bounding curves in this way produces
9 an adequate calculated fairway 22 as shown in Fig. 7b.

10 In one dimension, for the sample column 11 and as shown in Fig.
11 7b, the upper and lower bounding curves 20,21 sandwich the bulk of the raw
12 data 14. Referring also to Fig. 7c, with respect to the 2D image, where the
13 calculations really take place, the fairway 22 of the columnar data 14 can be
14 extended in the second dimension (adjacent columns) for each successive
15 analyzed set of raw data 14, and is analogous as the volume between two
16 smooth surfaces made up of each adjacent upper bounding curve 20,20,20 ...
17 (20+) and each adjacent lower bounding curve 21,21,21 ... (21+). Image
18 examples of the upper and lower bounding surfaces a20+,21+ re Figs. 8a and
19 8b, respectively.

20 The nomenclature, bounding curves 20,21 and surfaces 20+,21+,
21 are used interchangeably herein and generally reflect the context of the analysis
22 process, such as whether the image is analyzed on a line-by-line 1D curve basis
23 or in 2D simultaneously defining surfaces for efficiency purposes.

1 In this 2D context, the criteria can be restated in a slightly modified
2 form, being: the maximum and minimum bounding surfaces 20_+ , 21_+ are made
3 of low frequency components; the maximum and minimum surfaces graze the
4 local extreme data points (20_i - 20_n or 21_i - 21_n); and the fairway 22 separating the
5 surfaces is close to the extreme ranges of the data with only a small number of
6 outliers 23 ending up outside of the volume between the surfaces.

7 Having reference to Fig. 11b, the final iteration 40_n ends up with a
8 maximum bounding curve 20 that satisfies each of the criteria above.

9

10 Determining the Order of a Low Frequency Curve

11 As stated, one conventional approach to determining low frequency
12 components of a data set is to use a moving average technique or filter, simply
13 stated to: consider a subset of the data (the image); calculate the average of this
14 subset (average intensity); and place this value in a second image data array at
15 the center of the subset (mean intensity).

16 The most straightforward two dimensional (2D) implementation is
17 to use a filter box having an odd number of points in each of the row and column
18 dimension, making the mean point easy to determine and index in various data
19 arrays. For convenience, a rectangular filter box is disclosed; the average of the
20 points within the box being calculated and the result being stored at the point
21 located at the box's center. The center of the moving average filter box is moved
22 over one point and a new average is then calculated for the next output point.
23 Efficient algorithms ensure that repeated calculations are minimized and in this
24 sense there may never be a physical manifestation of the box. This process is

1 performed for all points in the image. Other shapes, such as circles, can be
2 used and weighted averaging schemes can also be used as well.

3 The size of the box determines whether the order or frequency is
4 low or high. For instance, for the 1007 data points of the example column of raw
5 input data, a filter box of 503 points would hypothetically result in two mean
6 points or a linear curve of zero frequency, blending all of the data of image 10
7 and obscuring any contrast variations. If the box is merely one or three points
8 wide, the mean would be virtually identical to the raw data, resulting in a high
9 frequency curve which would not preserve the local data at all. The choice of the
10 size of the filter is discussed in greater detail below.

11 A moving average filter analysis has several practical weaknesses,
12 including: one being that there is no data past the edges and thus artificial
13 techniques are required to calculate the low frequency components near the
14 edges; and another being that there is a high computational inefficiency, due to
15 many redundant additions.

16 Having reference to Figs. 12 and 14c, the way that image edges
17 are dealt with in the current invention is to mirror the data at the end points or
18 image borders. A temporary data array (imag) is provided which is larger than
19 the dimensions of the image data. The image data is stored in the array with
20 sufficient room available at the beginning and at the end of the image data for
21 the mirrored data. The amount of points mirrored is dependent upon the size of
22 the moving average filter box. At least as many points as $\frac{1}{2}$ of the filter box size
23 (dimensions of col_pad and row_pad) must be mirrored to place the filter box at
24 the image edge and still calculate a mean and store it at the image's edge.

1 More particularly, and having reference to Fig. 12, the original input
2 data is the same as previous graphs except that there is a larger range of point
3 numbers; for the 878x1007 point image of Fig. 1, going into the negative
4 numbers lower than point 1 and beyond the image edge of 1007. The mirrored
5 data 44 is shown prior to point 1 and after point 1007 as denoted by solid vertical
6 lines 51. The extra artificial data is exactly a point-by-point mirror image of the
7 real data 14, which ranges from points 1 to 1007 on this graph.

8 This mirroring is one practical way to handle the image's edges
9 because the resulting artificial mirrored data 44 has similar characteristics to the
10 real data 14 including: the point to point contrasts between real data-and-artificial
11 data or artificial data-and-artificial data are the same as that found within the
12 actual input raw data; and basic statistical qualities of the input data and the
13 artificial data are similar.

14 Moving Average Filter

15 As set forth in the sample Visual Basic (Microsoft Corporation)
16 code of Figs. 13a and 13b, the normally inefficient moving average can be made
17 more efficient through the repeated use of values calculated once and applied
18 many times. This determination is also referred to in the pseudo code of Fig.
19 14a as subroutine effc_2dma when used to calculate the upper and lower
20 bounding curves 20,21. Mirrored data 44 is generated all around the edges of
21 the image 10, by a sufficient number of rows and columns to satisfy at least $\frac{1}{2}$ of
22 the anticipated dimensions of the filter box.

23 Once the image has its edges mirrored, then moving average
24 technique can be applied to find the low frequency bounding curves 20,21.

1 Applied mechanically, a moving average technique is computationally inefficient
2 as there are many repeated calculations (additions). This is of particular concern
3 as the moving average technique is to be used in an iterative process.

4 Turning to the pseudo code of Fig. 13a, at A, the overall image
5 dimensions are initialized as rmin and rmax, including the mirrored data of
6 dimension rows-pad. At B the filter box column dimensions are initialized as
7 c2bg and c3en. A rectangular moving average filter box with uniform weighting
8 (of 1.0) is applied.

9 Note that in the least efficient application of the moving filter, the
10 intensity values for a subset of a line (1D) or a box of points (2D), the filter box, is
11 summed. In such an analysis, the surrounding filter box of points is incremented
12 one point, each point is summed and a mean is calculated. The number of
13 operations is each to the number of point addition and a division to get the mean.

14 Efficiency is improved by acknowledging that the bulk of the filter
15 box averaging calculations are repeated.

16 Having reference to Fig. 13b, in a known and basic concept, in 1D,
17 is that for a row of image points I I I I I I ... , a filter box having a demonstrated
18 dimension of 1 row by seven column points BBBBBBB are summed (7 additions)
19 and are divided by 7 to determine the mean, as indicated in a highlighted square
20 at the center of the BBBBBBB. When the filter is incremented one column, the
21 value of the point in the leftmost column is subtracted from the sum and the
22 value of the point in the rightmost column is added, all of which is again divided
23 by 7. The result is only 2 arithmetic operations and one division. For larger filter
24 boxes, this results in large computational savings.

1 Applicant has improved this known approach to apply in work in 2D
2 and can apply it to an entire image 10 while avoiding repeat calculations in these
3 very large arrays of points.

4 It can be demonstrated that identical analyses result from either a
5 row and column approach or a column and row approach. In Fig. 13c, a plurality
6 of filter subset points in adjacent rows (seven high, three below center and three
7 above) are summed and then the resulting row sums are summed across a
8 plurality of filter columns (7 wide). Resulting in the same calculated mean, in
9 Fig. 13d a plurality of filter subset points in adjacent columns (seven wide, three
10 left of center and three right) are summed and then the resulting column sums
11 are summed across a plurality of filter rows (7 high).

12 Fig. 13e illustrates a 2D application of a row-by-row traverse of the
13 efficient calculation approach as described in Fig. 13b for one dimension.

14 The approach illustrated in Figs. 13d, 13e are only visually
15 illustrative of a moving average technique applied in 2D. Actually, as can be
16 seen in the code in Fig. 13a at C and D, the analysis is actually staged into two
17 iterative steps and a normalization step.

18 In a first iterative step, at C, one dimensional filter box column
19 sums are calculated for the plurality of points within the filter box dimension, and
20 are stored in the array (stot). For the entire image, each successive filter box
21 column sum is calculated using the efficient 1D approach. The array (stot) is
22 indexed by the center point for the filter box column sum.

23 In the second iterative step, at D, for the entire image, a point by
24 point moving average analysis is performed, effectively determining every

1 moving average curve in the image simultaneously. For each and every point,
2 as shown in Fig. 13e for a snap shot of one point, the filter box column sums
3 (stot) below the point, at, and above the point are summed – pulling the column
4 sums from the stot array. The division by the number of filter box points or
5 normalization step is not yet done. For each successive point, the same efficient
6 concept is applied by subtracting each previous filter box column sum stot and
7 adding each successive sum as the calculation loop traverses, and whether on a
8 column-by-column or row-by-row basis.

9 Finally, every point in the image is normalized by the number of
10 points (npts) in the filter box.

11 Referring more particularly to the code in Fig. 13a, at C, in a first
12 nested loop, for irow and ico2, image data for neighboring columns is summed
13 (the filter box column sums or partial sums) for the filter box of the first sample
14 column 11. The first nested loop (ico2) calculates the partial sums for points in
15 columns either side of the sample column direction (the width of the filter box),
16 and at a particular row position for the first column of the input data. The second
17 nested loop (irow) then increments the partial sum routine, row-by-row and up
18 the column between the image limits.

19 Next in C, the second of the nested loops, for icol and irow,
20 calculates the partial sums for each successive column. Here efficiencies are
21 improved as described above; as one moves from the first column to the last
22 column, and re-calculates the partial sums, the new sum is merely the last sum,
23 less a trailing point and plus a leading point. More specifically, in the innermost
24 loop of this second nested loop, as the central point only moves one position, the

1 partial sums for the next incremental filter box are the same as dropping the
2 previous first element (ico5) and adding the current last element (ico4). In this
3 way each partial sum is only done the one time, and not each time the box is
4 moved. The total number of operations for the partial sums are very nearly
5 equal to one addition and one subtraction for each point in the image.

6 At D the filter box sums are performed. Note that a calculation
7 consuming separate step is not required to obtain subtotals for the filter box row
8 dimensions as they were done at C and indexed in an array (stot). The first
9 nested loop at D calculates the filter box by combining the partial sums from C
10 and the sums for neighboring points in the filter box for rows straddling the first
11 row.

12 The second nested loop at D then successively calculates the full
13 filter box sums for the rest of the columns and for the rest of the rows. Again,
14 this second step is again done as efficiently as possible. As the central point of
15 the filter box only moves one position, the new sum is the same as dropping the
16 previous first partial sum (iro5) and adding the last current partial sum (iro4).

17 Accordingly, now between C and D, the total number of operations
18 for the full 2D sums is very nearly equal to only two additions and two subtraction
19 for each point in the image 10.

20 Finally, at E, as we want to know the average of the summed data
21 and not just the sum itself, we must divide the sum of the intensity by the number
22 of points summed (npts) which is the area of the box.

23 Using the above technique, one does not need to physically
24 maintain an intermediate filter box array of intensities, nor the calculations or

1 operations to produce them. The total operations required for the moving
2 average is nearly equal to only 4 additions/subtractions and 1 division per point
3 in the image - regardless of how many points are averaged. Accordingly, in the
4 case of a moving average rectangular box of 81 by 101 points (about 10% of the
5 image dimensions), which is not unreasonable for this sort of application, the
6 normal inefficient moving average technique would have required 8,181
7 additions and 1 division per point in the image – again versus the respective 4
8 and 1 now possible.

9 The problem of the computational inefficiency of the 2D moving
10 average filter is thus solved - by the introduction of an extra holding image data
11 array (stot) for the subtotals.

12 This technique is repeated for each and every other column in the
13 image.

14

15 Iterative Technique

16 As set forth in the sample code of Figs. 14a – 14c and definitions
17 set forth in Fig. 14d, the moving average filter box dimensions are set and the
18 moving average mean curves 24_i-24_n and 26_i-26_n are calculated and are iterated
19 upwardly and downwardly to form the extreme upper and lower bounding curves
20 $20,21$, extreme curves or extreme surfaces $20+,21+$.

21 The filter box dimensions can use predetermined default values.
22 As is shown later, enhancement of the image 10 is rather insensitive to even

1 large variations of the box size. For a certain type of image, a default value
2 produces acceptable results.

3 Having reference to Fig. 14a at F, one convenient way to reference
4 the filter box size is to set it as a fraction of the dimensions of the image 10. Box
5 parameters rows_pad and cols_pad are calculated from the number of image
6 rows and columns; simply as a percentage of the dimension divided by two so as
7 to provide the $\frac{1}{2}$ box dimension, either side of the box's center. Note that this is
8 also how many data points 44 at the end points that need to be mirrored.

9 At G, array limits are defined to hold various temporary versions of
10 the image data. For convenience, the variables cmin, cmax, rmin and rmax
11 were chosen so that the index values referring to the same data point in the
12 original sized image 10 or mirrored sized data arrays are the same. It is
13 understood that the storage, reading and writing of the original image and how
14 columnar, rows and intensity data are extracted from data files or image arrays,
15 is system and programming language specific, the means of which is known to
16 those skilled in the art and is not detailed herein.

17 As defined in Fig. 14a, image_wrk1 contains the raw data,
18 image_wrk2 contains the current upper bound data, image_wrk3 contains the
19 current lower bound data, and image_wrk4 holds the filter box sums for efficient
20 processing. These arrays are initialized to the mirrored dimensions.

21 At H, image_wrk2 is initialized to hold mirrored raw data and then
22 routines data_mirr and effc_2dma calculate the moving average of the mirrored
23 data and place it in image_wrk2, which is then copied to image_wrk3. In other
24 words, at this first stage, the current approximations of the upper and lower

1 extreme surfaces 20+,21+ are, at this early point, are identical and simply equal
2 to the moving average mean curve 15. The subroutines are only called once for
3 the sake of efficiency.

4 At I, a tolerable failure rate, *frat*, for outliers 23 is set to an arbitrary
5 value of 5% (for either above or below). The nature of the histogram correction
6 is such that as long as the failure rate, during the iteration, is not huge, no
7 problems are experienced. There are also upper surface minimum and lower
8 surface maximum parameters, *surf_mdff*, which handle noise. This parameter is
9 described in more detail below.

10 At J the iteration is initiated.

11 At J1, the first step of the iteration, the upper surface iteration array
12 *imag_wrk2* is loaded up with the greater of the residual population of data 40 and
13 the last iteration 50_i of the upper surface 20+. The lower surface iteration array
14 *imag_wrk3* is loaded up with the lesser of the residual population 41 of data and
15 the previous iteration 51_i of the lower surface 21. It is this first step that forces
16 the data in the two arrays to be the upper and lower surfaces 20+,21+.

17 At J2, the second step of the iteration, the data in the two arrays
18 are mirrored.

19 At J3, the third step of the iteration, the low frequency components
20 50,51 of both residual upper and lower populations 40,41 of artificial data sets
21 are found. The two latest approximations of the extreme upper and lower curves
22 20 are output from the subroutine *effc_2dma* and are stored in the arrays
23 *imag_wrk2* and *imag_wrk3*.

1 At J4, the fourth step, the minimum and maximum surfaces
2 20+,21+ are required to have at least a specified minimum separation surf_mdff.
3 Whenever the surfaces are separated by less than the minimum, one point by
4 point basis, the surfaces are then readjusted so that they are separated by the
5 minimum; otherwise insignificant variations, possibly even noise, end up being
6 scaled to the full dynamic range. This separation is done symmetrically, half to
7 each surface. The four conditional "If" statements at the end of the nested loop
8 readjust the surfaces in the unlikely, but possible, event the previous adjustment
9 pushes one outside of the dynamic range. After this step, the two surfaces will
10 be separated by at least the minimum value everywhere.

11 This fourth separation step J4 also aids in convergence of the
12 iteration as it is less likely that data local to this correction will be found outside
13 the adjusted volume between the surfaces 20+,21+.

14 At J5, the fifth step, the number of outliers 23 outside of the volume
15 between the two surfaces 20+,21+ establishes a failure rate. Thus, the process
16 is aware of the degree that the current attempt at the extreme surfaces has failed
17 to contain the data.

18 At J6, the sixth and final step of the iteration, a determination is
19 made whether the iteration should continue or if the current upper and lower
20 bounding curves 50_n,51_n are sufficient. Simply, the iteration will end if the
21 iteration flag (itr_f) is set to no, otherwise the loop repeats the do loop of J1. The
22 itr_f flag will only be set to no, signaling convergence if one of two conditions are
23 met: if the failure rate on both the positive and negative sides is less than the
24 acceptable failure rate; or if the number of iterations is too high. In this case, an

1 arbitrary number of 10 is used. It is important to point out, again, that the failure
2 rate here is not that important due to the secondary correction. Still, a balance is
3 found between the quality of the curve fit and the amount of effort expended
4 without iterating forever.

5 Once the loop is complete, the extreme upper and lower bounding
6 surfaces $50_{n,20+}$ and $51_{n,21+}$ are known and can be output themselves as
7 image files (as illustrated in Figs. 8a,8b), or as other computer files, for quality
8 control purposes. Process for outputting such images for display or storage is
9 not disclosed herein as it is known by those anyone skilled in the art and is
10 dependent upon such variables as the particular image format and display
11 hardware used. Now that the upper and lower bounding curves are established
12 the first fairway and second histogram corrections can be performed to enhance
13 the image of Fig. 1 and produce the image of Fig. 2.

14

15 Improved Convergence

16 As can be seen from multiple curves of Figure 11a, the iterative
17 solution converges somewhat slowly. There is a computational incentive to get a
18 faster convergence - more movement per one iteration pass.

19 In a modified approach, detailed in Fig. 15a, the data analyzed for
20 the next iterative mean curve 60 is not simply the residual data populations
21 40,41, as in Figure 11b, but instead comprises an exaggeration of the data. The
22 difference between the extreme points of residual data 40,41 and the curve
23 50,51 of the current iteration is multiplied by the number of iteration passes. This
24 makes for a quicker convergence as each pass gives more and more correction

1 to the repeated failure points. There are checks to ensure that no extra
2 correction is beyond 20% of the dynamic range and that no value is replaced
3 with a value outside of the dynamic range.

4 Having reference to Figs. 15a - 17, when applied to the raw data of
5 Fig. 4a, this modified approach took only 4 iterations 60_1-60_4 to converge (Fig.
6 15b) in contrast to the previous 10 iterations $24_i - 24_n$ which, regardless of its
7 higher number of iterations (ten), still resulted a failure rate greater than 5%.
8 The fairway 22 resulting from the enhanced convergence technique, as shown in
9 Fig. 16, is not quite as smooth as in the previous 10 iteration case, as is the
10 natural consequence of fewer iterations, but it is acceptable. As shown in Fig.
11 17, the subsequent second histogram correction result 34 is very similar to
12 before, as compared to Figs. 6a and 6b. It is not visually apparent which
13 convergence gives the better result but it is clear that both give a very
14 comparable effect where most of the dynamic range is used locally.

15 In an alternate, non-iterative statistical technique for finding the
16 extreme surfaces, a moving average of the raw input data, is once again
17 calculated initially. Then, the difference of the raw data 14 and the moving
18 average mean curve 15 is calculated. Finally, the standard deviation of the
19 residual data of Fig. 5b, is calculated using the same size filter box of original
20 data as the moving average. Using the same essential logic as the efficient 2D
21 moving average filter of Fig. 13a, but modified to suit a moving standard
22 deviation, a local measure of the standard deviation of every point could
23 efficiently be determined. The extreme surfaces would then be calculated as the
24 moving average $\pm$ (a constant * moving standard deviation) for each point.

1 This constant is set to a number that would typically yield something like a 1%
2 failure rate. Empirically, a suitable constant appears to have a value near 2.0.
3 Note that in the true Gaussian data, or perfect bell curve, that about 98% of the
4 data falls within 3 standard deviations of the mean.

5

6 The Second Histogram Correction

7 Having reference to the coding in Fig. 18a and the definitions in
8 Fig. 18b, separation of the data correction into 2 steps, as discussed above,
9 allows for flexibility in the estimating of the extreme surfaces. Approximate
10 curves 20,21 for the surfaces 20+,21+ are acceptable. Even a high percentage
11 of the data 23 outside of the extreme surfaces at the first step does not result in
12 serious data clipping.

13 This is an important distinction from the prior art techniques which
14 apply conditional statements so as to handle the image data differently if they
15 are above a threshold or below; in a sense creating two separate data
16 populations and two different corrections.

17 In the present invention, while the first fairway correction and
18 makes a preliminary assessment of scaling ranges which identifies and ensure
19 that even outliers 23 are not lost but continue to be preserved and maintained in
20 the main data population 14 throughout the use of the histogram and its
21 expanded scaling range. Finally once as many data points have been
22 enveloped as is possible, only the most deviant of the outliers are finally
23 trimmed.

1 At R, the fairway correction is illustrated in pseudo-code form,
2 wherein the raw data `imag_wrk1` less the lower curve `imag_wrk3` is scaled by the
3 fairway's range (`imag_wrk2-imag_wrk3`) to the dynamic range (`dr1m`), the
4 corrected data being stored in `imag_wrk4`.

5 The second histogram correction allows for control on exactly how
6 much of the data is to be clipped or trimmed. Also, as the data from the first
7 fairway correction is already residual in its nature, only a small amount of data to
8 be trimmed will also be sparsely distributed throughout the image resulting in no
9 noticeable distortion in the image at all. The fairway correction by the application
10 of Equation 1 (at R) retains the outlier information by temporarily storing them in
11 a higher precision variable than is needed for the dynamic range. This fairway
12 correction also allows for negative numbers.

13 Data below the lower extreme curve 21 is calculated as negative
14 and values above the upper extreme curve 20 are calculated as numbers above
15 the dynamic range of the input image file/data. The histogram correction is,
16 then, performed on data that has very similar statistical qualities of what will be
17 the output data, seen by comparing Figs. 9b and 6b. A histogram of the first
18 fairway corrected data is made using a broader range of events than is likely to
19 occur. Empirically, it has been found that the range of values calculated by the
20 fairway correction are unlikely to exceed three times the dynamic range and
21 these limits are arbitrarily used to set the histogram dimensions.

22 The histogram count is then made by looping through the first
23 fairway corrected data and counting the number of occurrences of each intensity
24 value.

1 Referring to Fig. 18a again, at S, the histogram is counted normally
2 where in the fairway correction data lies within the expected range of -drlm to
3 drl2 (three times the output dynamic range). While it is unlikely that any fairway
4 correction data values would be outside this range, these possibilities are dealt
5 with nonetheless. Merely extending the range beyond $3 \cdot \text{drlm}$ is no guarantee.

6 Accordingly, at T, in the unlikely event that a data point lies outside
7 of the expected range, the histogram count is incremented at the extremes of the
8 histogram. It is expected that the histogram counts at the extreme will be zero in
9 virtually all cases and very close to zero in even all but exceptional cases. The
10 Select Case coding jumps out of the counting routine at the first logically correct
11 Case statement. A running total of the histogram is made until the value reaches
12 the acceptable trim (failure) rate, set at step U to 1% of the image data.

13 The trimming points are found by using a running total. As shown
14 at V, for the lower trim point, the loop starts at the low end of the histogram and
15 works up until the running total is above the acceptable trim rate. The position of
16 the trim point is then readjusted down by one value to bring the running total to
17 just below the acceptable rate. This value is constrained to $(-\text{drlm})$ or greater,
18 regardless of the remote likelihood of such a situation.

19 At W, the same approach is used for locating the maximum or
20 upper trim point, except that the loop starts at the top of the histogram and works
21 down.

22 Once the values of the upper and lower trim points are known, then
23 a simple histogram correction of a straight line adjustment from the acceptable
24 trim limits to the image dynamic range limits is made.

1 In the simplest situation shown in one embodiment X1, the first
2 order corrected data is scaled from the trim limits to the dynamic range. Finally,
3 any data point outside of the dynamic range is trimmed to the full range of the
4 dynamic range limits. This is the full correction of the current invention.

5 Each of the new scaled intensity data for each sampled column
6 can be reassembled and the output image is virtually indistinguishable from that
7 in Fig. 2.

8 While the above, at X1, is a perfectly adequate solution in itself, it
9 has been found empirically that the previously determined trim points are
10 preferably be set to a range different from one exactly set to the dynamic range
11 limits, as human vision is less sensitive in the darker regions. The following
12 modification sets the trim point to 20% and 95% of the dynamic range which
13 leads to the favorable result found in image Fig. 2. It also leads to increased
14 data integrity. Fig. 2 is the result of only about 0.2% or about 1800 of some
15 880000 points which ended up being trimmed. Normally, code for counting
16 trimmed data would not be included for improved computational efficiency.

17 At this point, the completely first and second corrected image
18 resides in array `imag_wrk4` and can be output to media, such as a magnetic
19 disk, as a standard image file or any other kind of computer file. Also, if
20 incorporated within a software package with a GUI interface, the processed data
21 can be displayed from electronic (RAM) memory to a monitor or other display
22 device.

23

1 Image Correction Parameters

2 The present invention utilizes the input image and two numeric
3 parameters. These parameters have been previously referred to in the sample
4 code and they are:

- 5 • surf_mdff which holds the minimum allowable difference
6 between surfaces as a Signal to Noise S/N management issue
7 (Fig. 14b, J4); and
- 8 • imag_pcmt which holds the percentage of image size (in both
9 dimensions) to be used for the moving average process (Fig.
10 14a, F).

11

12 If the parameter surf_mdff is set to a value 8, that means that the
13 two extreme surfaces 20+,21+will be constrained so that they will never be less
14 than 8 units of intensity apart. If surf_mdff is expressed in terms of percent of
15 dynamic range, then 3% of 0-255 becomes 8.

16 If the parameter imag_pcmt is set to 10% and the image is 2000 by
17 1500 dots in size, then the moving average filter box dimensions will be set to
18 200 by 150 dots, all of which are averaged to determine the value for a single
19 low frequency point. Note that this is 30,000 additions per point - thus the need
20 for an efficient algorithm. More exactly, the filter box will be set to 151 by 201
21 points.

22 While these two parameters are, in themselves, straightforward
23 enough in concept, Applicant recalls that it was a desired criteria that the user
24 not have to deal with such tasks.

25 Each of two parameters is independent of the other. The
26 parameter surf_mdff manages noise issues. In this sense, noise is considered

1 to be the more or less random fluctuations of the numeric intensity value at each
2 point. Indeed, surf_mdff is a tolerance with respect to the intensity of any given
3 dot expressed in terms of percent of the dynamic range. The parameter
4 imag_pcmt, on the other hand, represents the area of points to be considered at
5 one instant (the filter box). Expressed in Cartesian coordinates, imag_pcmt is a
6 control for both the x and y axis (the domain) and surf_mdff is a control of the z
7 axis (the range).

8 Default values can be provided for images which have a known
9 and historical noise component, and thus is taken completely from the hands of
10 the user. Otherwise, general guidelines can assist the non-technical user in
11 specifying the parameter from higher level choices, such as whether the noise in
12 the image is clean, noticeable, or serious to which values for surf_mdff could be
13 specified for the user at 3%, 6% (a typical default), or 12% of the dynamic range
14 limit.

15 The parameters numeric value is assigned internally in the
16 program by a small lookup table of 3 elements and a multiplication or two. This
17 is a particularly useful approach where the present invention is coded into a GUI
18 and the process is done interactively. The user could estimate the noise level at
19 a glance and then choose from a drop down menu with the above choices. The
20 robustness of the parameter is such that this limited choice of three choices
21 should be perfectly adequate. Note that the details of this scheme are not
22 intending to be limiting but to point the way to a practical implementation.

23 Regarding the box size parameter, imag_pcmt, its selection and
24 magnitude is primarily an issue of the detail within the image. The more detail in

1 the image, the smaller the moving average filter box must be so that the full
2 dynamic range will be used over smaller areas. As the parameter is also robust,
3 having a low gain, a similar scheme can be used for it. In the context of X-ray
4 images, the following values were determined empirically; for a chest X-ray –
5 5%, for a hand - 10% (typical default), or for a leg bone at level of detail at 20%
6 of the image size.

7 The user can estimate the parameter by simply knowing what the
8 subject was. The robustness of the parameter is such that this limited choice of
9 three should be perfectly adequate.

10 It is also the case that if the process was not performed in a GUI
11 this straightforward form of parameter estimation will be particularly useful as so
12 little expertise is needed and the "batch processing" style of input can also be
13 made to be easy.

14 At one extreme, a finer level of choice can be implemented as an
15 option for technically adept or curious and a coarser level of options available for
16 every user. The default evaluations of the user input parameters are also found
17 to produce acceptable results without the user actually being consulted at all.

18 In any case, it is easy to implement the variables in a way that is
19 humanly understandable so that the user is not required to increase their
20 knowledge of the physics and mathematics of the process.

21 Having reference to Figs. 19 - 21, images are presented which are
22 the result of processing the same input image (Fig. 1) with different choices of a
23 value for surf_mdff. Fig. 19 used a surf_mdff value of 1.5%, Figs. 20a and 20b

1 are identical at 3% and Fig. 21 used 12%. Note that Figs. 2, 20a and 20b are
2 identical and have been duplicated for comparative purposes only.

3 The robust nature of the parameter is clearly demonstrated by
4 noticing that it takes a doubling of the parameter to result in a significant effect.
5 The images of Figs. 19, 20a at 1.5 and 3% respectively are virtually
6 indistinguishable. Likewise, Figs. 20b and 21 are virtually indistinguishable from
7 each other. Applicant is aware that there are limits as to how far the parameter
8 can be varied.

9 An increasing surf_mdff value better controls random noise
10 amplification but this also limits the amplification of more subtle effects. In the
11 case of the current example image, subjective analysis finds the Fig. 2 (same as
12 Figs. 20a,20b) to be the superior correction – based upon random noise being
13 hardly noticeable yet the subtle aspects are still very clear. When the surf_mdff
14 value is smaller the noise amplification becomes less desirable and when the
15 surf_mdff value is higher the subtler aspects experience a reduced amplification
16 as in the case in Fig. 21.

17 Yet it is also the case that all three images represent a dramatic
18 improvement over Fig. 1 and, thus, any reasonable choice of the surf_mdff value
19 would have yielded a fine result.

20 Having reference to Figs. 22 – 25, images are presented which are
21 the result of processing the same input image (Fig. 1) with different parameter
22 choices for imag_pcmt. Parameter surf_mdff values remain constant but the
23 imag_pcmt value varies. Once again Fig. 2 is used as the reference and has a
24 value of 5% of the image size. For comparative purposes, Fig. 2 has been

1 duplicated as Fig. 22. Fig. 23 uses a filter box at 7% and Fig. 24 at 10%, and
2 Fig. 25 at 20% of image size.

3 The robust nature of the second, box-size parameter is again
4 clearly demonstrated by the subjective observation that any of Figs. 22 (Fig. 2),
5 23, or Fig. 24 are acceptable results. Finally at 20% for Fig. 25, the result pales
6 compared to the others. However, Fig. 25 it is still a good improvement over Fig.
7 1 and one must keep in mind that to use the value of 20% with this X-ray would
8 represent the worst case of a user possible selections, if so permitted. It is
9 anticipated that diagnosticians would not confuse the differences in detail
10 between a chest and leg image.

11 The applicant-selected box-size parameter setting of 5% (Fig. 2,
12 22) and the suggested default setting of 10% (Fig. 24), both gave an excellent
13 result as did the intermediate value of 7% (Fig. 23).

1 THE EMBODIMENTS OF THE INVENTION FOR WHICH AN
2 EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS
3 FOLLOWS:
4

5 1. A method for maximizing contrast between neighboring points
6 of a data population selected from a digital image, each image point being
7 defined by a position and an intensity, the image having an intensity dynamic
8 range, the method comprising the steps of:

9 (a) determining a first low frequency trending function which is
10 curve fit to the data population's maximum intensities;

11 (b) determining a second low frequency trending function,
12 independent from the first trending function, and which is curve fit to the data
13 population's minimum intensities;

14 (c) establishing a maximum and minimum fairway for the data
15 population bounded by the first and second trending functions; and

16 for each point in the data population,

17 (d) extracting a range of the local minimum and maximum intensity
18 from the fairway;

19 (e) determining a local scaling factor as the ratio between the
20 dynamic range for the image and the extracted local range; and

21 (f) scaling the point by the local scaling factor.

1 2. The method of claim 1 wherein the point is scaled using the
2 relationship of $\frac{(\text{Image Point} - \text{Local Minimum})}{(\text{Local Maximum} - \text{Local Minimum})} \text{Dynamic Range}$.

3

4 3. The method of claim 1 wherein one or more outlier points for
5 the data population exist, which have intensities outside the fairway, further
6 comprising the steps of:

7 (a) scaling each outlier point by the local scaling factor with the
8 result that the scaled intensity is outside the image's dynamic range;

9 (b) storing the outlier point's scaled intensity at a precision greater
10 than the dynamic range to prevent loss of intensity information;

11 (c) forming an intensity histogram of the entire data population, the
12 histogram establishing a predetermined lower intensity and a predetermined
13 upper intensity for forming a range which is greater than the dynamic range and
14 which encompasses the intensities of substantially all the outliers;

15 (d) trimming the intensity histogram of points which have an
16 intensity which is below the predetermined lower intensity, and above the
17 predetermined upper intensity, for establishing a trimmed data population having
18 a trimmed range between minimum and maximum trimmed intensities; and

19 (e) scaling each point of the trimmed data population by the ratio of
20 the trimmed range to the image's dynamic range.

21

1 4. The method of claim 2 wherein the data population is adjusted
2 to the greater intensities suited to human vision, further comprising the steps of:

3 (a) scaling each point of the trimmed data population to an output
4 range less than the image's dynamic range; and

5 (b) offsetting each scaled point by an incremental intensity value
6 which is less than the difference between the dynamic range and the output
7 range so that the scaled points reside in a higher range of intensities which is still
8 within the dynamic range.

9

10 5. The method of claim 1 wherein determination of each of the first
11 and second trending functions comprises the steps of:

12 (a) determining a first function representing the entire data
13 population;

14 (b) iteratively determining a successive upper function for a
15 residual subset of points of the data population which have intensities greater
16 than the greater of the first function or a previous successive upper function,
17 converging upwardly until fewer than a predetermined number of points are
18 greater than the successive upper function, the converged successive upper
19 function forming the first trending function; and

20 (c) iteratively determining a successive lower function for a residual
21 subset of points of the data population which have intensities lower than the
22 lower of the first function or a previous successive lower function, until fewer
23 than a predetermined residual number of points are lower than the successive
24 lower function, the converged successive lower function forming the second
25 trending function.

1 6. The method of claim 5 wherein the convergence of one or both
2 of the upper and lower functions is improved by:

3 (a) determining the differences between the residual subset of
4 points and each successive function in the iteration;

5 (b) amplifying the differences and adding it to a subset of points to
6 form an exaggerated subset of points; and

7 (c) applying the iterative determination of successive functions to
8 the exaggerated subset of points.

9

10 7. The method of claim 6 wherein the differences are amplified by
11 incrementing a counter each iteration and multiplying the differences by the
12 counter.

13

14 8. The method of claim 5 wherein one or more of the functions
15 applied to determine each of the first and second trending functions is a moving
16 average.

17

18 9. The method of claim 8 wherein the function applied for each of
19 the first, successive upper and successive lower curves is a moving average.

20

1 10. The method of claim 8 wherein the moving average has a filter
2 box of predetermined size of a column dimension and a row dimension and
3 which is optimized in two dimensions for the image by:

4 (a) extending the image at the edges of the image by a number of
5 points complementary to the filter box column and row dimensions;

6 (b) determining a first sum of the intensities of a one dimensional
7 subset of points, for the size of the filter box, and about each image point along a
8 first column or row dimension;

9 (c) storing the subset sums, indexed to each point in the image;

10 (d) determining a second sum of the subset sums, for the size of
11 the filter box, and about each image point along a row or column dimension; and

12 (e) normalizing the second sum by dividing by the number of points
13 in the filter box.

14

15 11. The method of claim 10 wherein the image has specified
16 dimensions and the filter box dimensions are set to a percentage of the image
17 dimensions for adjusting the frequency of one or both of the first and second
18 trending functions.

19

20 12. The method of claim 11 wherein the filter box dimensions are
21 between about 5 and 20% of the image dimensions.

22

23 13. The method of claim 12 wherein the image is an X-ray of a
24 chest and the filter box dimensions are about 5%.

25

1 14. The method of claim 10 wherein the local range of the fairway
2 image is constrained to be no narrower than a percentage of the image dynamic
3 range for suppressing noise.

4

5 15. The method of claim 14 wherein the minimum range of the
6 fairway image is constrained between about 3 and 12%.

7

8 16. The method of claim 15 wherein the minimum range of the
9 fairway image is constrained to about 6%.

10

11 17. The method of claim 10 wherein the image is extended at the
12 image edges by mirroring of the image data at the edges.

13

14 18. The method of claim 10 wherein

15 (a) the image has specified dimensions and the filter box
16 dimensions are set to a percentage of the image dimensions for adjusting the
17 frequency of one or both of the first and second trending functions;

18 (b) the local range of the fairway image is constrained to be no
19 narrower than a percentage of the image dynamic range for suppressing noise;
20 and

21 (c) the frequency and noise suppression percentages are
22 adjustable by a diagnostician.

23

1 19. A system for maximizing contrast between neighboring points of
2 a data population selected from a digital image, each image point being defined
3 by a position and an intensity, the image having an intensity dynamic range, the
4 system comprising:

5 (a) means for determining a first low frequency trending function
6 which is curve fit to the data population's maximum intensities;

7 (b) means for determining a second low frequency trending
8 function, independent from the first trending function, and which is curve fit to the
9 data population's minimum intensities; and

10 (c) means for establishing a maximum and minimum fairway for the
11 data population bounded by the first and second trending functions so that for
12 each point in the data population, a range of the local minimum and maximum
13 intensity from the fairway can be extracted, a local scaling factor can be
14 determined as the ratio between the dynamic range for the image and the
15 extracted local range, and each point can be scaled by local scaling factor.