On-line Corrections for Factors That Affect Uniformity and Linearity

Guy H. Simmons

Veterans Administration Hospital, Lexington, Kentucky

This is the second in a series of four continuing education articles on scintillation camera technology. After reading this article, the reader should be able to: 1) discuss the causes for camera nonuniformities; 2) list the methods for making uniformity and linearity corrections; and 3) describe the relative merits and limitations of various correction techniques.

THE NEED FOR ON-LINE CORRECTIONS

As scintillation camera technology improved in the early 1970s, it became apparent that a significant limitation on performance was caused by the inability to obtain a spatially constant response (within the limits of Poisson statistics) over the useful detector area from a spatially constant photon flux incident on the detector. This limitation became more significant as attempts were made to design cameras with better spatial resolution. The only line of defense against this nonuniformity problem (aside from careful quality control in the manufacturing process) was to keep the photomultiplier (PM) tubes tuned as precisely as possible in the field. Any residual nonuniform response was accepted as unavoidable, and uniformity became perhaps the most important criterion in the selection of instruments.

With the increasing use of scintillation cameras for single-photon emission computed tomography (SPECT), corrections for energy response, nonlinearity, and PM tube gain shifts have become extremely important. Nonuniformities in the detector response are magnified when data from many angles are back-projected to compute tomographic images. Subtle effects that were tolerable in planar imaging have suddenly become intolerable with SPECT. A camera that does not have on-line corrections of the type described in this paper probably cannot be used successfully for SPECT, if for no other reason than the small shifts in PM tube gain that can occur when the tubes are rotated within the earth's magnetic field. One would expect that this problem could be avoided by adequate magnetic shielding, but in some models a small residual effect persists, and satisfactory operation can only be achieved by frequent on-line adjustments.

EARLY CORRECTION METHODS

Matrix Multiplication

The availability of digital computers affords a means to correct for nonuniform response by storing a "flood" image from which a matrix of correction factors can be calculated. The correction factors are applied to image data as a post-processing procedure. A typical computer uniformity correction scheme works as follows:

1. The operator stores a flood image of 1.5 to 2.0 million counts using either a point source with no collimator or a sheet source with a collimator. If the data are to be used for correcting SPECT images prior to reconstruction, 40–60 million counts must be collected, and a collimator must be used. In fact, the flood image must be acquired using the same collimator that was used to acquire the images to be corrected since collimator nonuniformities may create artifacts in SPECT reconstructed images.

2. The operator views the flood image and sets upper and lower thresholds so that no pixels with counts outside the thresholds will be involved in the average pixel count computation. This is to avoid influencing the correction factors for the central field of view by high and low values that result from edge effects.

3. A computer program calculates an average pixel count from those pixels with counts between the threshold limits.

4. A matrix (64×64) of correction factors is computed and stored in which the correction factor for a particular pixel is equal to the ratio of the average pixel count to the individual pixel count. Expressed mathematically,

\[ CF_{ij} = \frac{\bar{P}}{P_{ij}}, \left\{ \begin{array}{l} i = 1, 64 \\ j = 1, 64 \end{array} \right\}, \]

where \( CF_{ij} \) is the calculated correction factor for the element \( i,j \) (pixel), \( \bar{P} \) is the average pixel count, and \( P_{ij} \) is the \( i,j \) pixel count in the stored flood image.

5. Planar images (including those used for SPECT reconstruction) are flood-corrected by applying the correction factors to all pixels in the stored image:

\[ FCI_{ij} = (RDI_{ij}) (CF_{ij}), \left\{ \begin{array}{l} i = 1, 64 \\ j = 1, 64 \end{array} \right\}, \]

where \( FCI_{ij} \) is the \( i,j \) pixel count in the flood-corrected image, \( RDI_{ij} \) is the \( i,j \) pixel count in the raw data image, and \( CF_{ij} \) is the corresponding correction factor.

The method just described is generally known as the "matrix multiplication" or "renormalization" method of uniformity correction.

For reprints contact: Guy H. Simmons, PhD, Nuclear Medicine, Veterans Administration Hospital, Cooper Drive, Lexington, KY 40511.
Whether or not to use a collimator for acquiring flood data to correct planar images has always been a controversy, as well as the question of using scattering material. Some argue that since the count density achieved from 1.5 million to 2.0 million counts is not sufficient to show subtle collimator effects, it is an intrinsic correction that is being made, and no collimator or scattering material is necessary. Furthermore, if one attempts to include collimator effects it would be necessary to store a separate flood for each collimator. Others argue that scattering in both the collimator and the patient are factors in the distribution of the nonuniformities, and therefore collimators and scattering material should be used in acquiring the flood data.

Perhaps the most serious drawback to the post-processing computer (matrix multiplication) method is the inability to apply it to non-digitized (analog) images. This is a drawback, first of all, because not all cameras are interfaced to computers, and secondly, because it is neither feasible nor desirable to digitize all nuclear images. Unless one needs to do computer processing, digitization of static images generally degrades image quality unless large matrices are used. The need therefore existed to develop on-line methods of correcting for non-uniform response.

**Early On-Line Correction Methods**

The first methods for on-line corrections were implemented without full understanding of the causes of the observed count density variations. The methods were developed with the assumption that the nonuniformities were caused by variations in point sensitivity over the detector area. Point sensitivity is defined as the total crystal count rate (the count rate integrated over the entire crystal area) from a collimated point source located at a certain position. As will be seen in a subsequent section, methods based on this assumption treated the symptoms rather than the cause of the problem, but they led to the development of more sophisticated methods.

**Count Skimming.** One correction method based on the assumption of point sensitivity variation is called count skimming. To implement the method, one must store a flood image in a microprocessor memory in the same manner as that described earlier for the external computer method, except in this instance the microprocessor is an integral part of the camera circuitry. A 64×64 matrix is typically used, and the data acquisition continues until 1,000 counts are accumulated in each of the central pixels. The minimum pixel count in the central field of view of the stored image is determined, and correction factors are calculated that are the ratios of the minimum to the individual pixel counts:

$$ CF_{ij} = \frac{P_{\text{min}}}{P_{ij}}, \quad \left\{ i = 1 \text{ to } 64, \quad j = 1 \text{ to } 64 \right\} $$

where $CF_{ij}$ is the correction factor for the $ij$ pixel, $P_{\text{min}}$ is the minimum pixel count, and $P_{ij}$ is the pixel count for the $ij$ pixel. During acquisition, a percentage of the events detected within a particular pixel area are randomly eliminated and do not unblank the CRT (i.e., no Z gate signal is produced). These discarded signals therefore do not contribute to the analog image. The events to be eliminated are chosen by generating random numbers. For example, suppose a pixel has a correction factor of 0.90. Each time an event is detected as occurring within that pixel, a random number between, say, 0 and 1000 is generated. If the random number falls between 0 and 900, the event is recorded; otherwise, it is eliminated. By this method, the elimination of events is randomized in time, thus preserving the temporal distribution of dynamic data. The fraction of the counts eliminated ("skimmed") is equal to one minus the correction factor for a particular pixel. It is seen that a camera with significant count density variations will discard a significant fraction of the detected events using this method. The count loss should not exceed 15% in a properly tuned camera, but even that represents a sizable increase in imaging time for some studies. Figure 1 is taken from Reference 1 by permission and illustrates, with a numerical example, the end result of the count skimming process.

**Count Addition.** Rather than renormalizing an image to its minimum pixel count as in count skimming, one can just as easily use the maximum pixel count. The operation is then called count addition, and the matrix of correction factors contains values greater than or equal to one. During acquisition

![COUNT SKIMMING](image)

**FIG. 1.** Numerical example illustrating the count skimming method of uniformity correction. Each pixel count is multiplied by a factor to make it equal to the minimum pixel count in the field.
of clinical data, extra $Z$ gate signals are randomly generated to increase the pixel count in those pixels with counts less than the maximum in the flood image (i.e., those pixels for which the correction factor is $>1$). If the correction factor is 1.10, for example, for every 10 counts recorded in that pixel, an average of one additional count will be added. The added counts are randomized in time by using a random number generator as described for count skimming.

An important consideration in both methods, count skimming and count addition, is the fact that the statistical accuracy of the data is determined by the raw data pixel counts and not the corrected images. This is true because the correction factors are scaling factors which do not alter the fractional variances in the data. Hence an image produced using the count addition method has no more statistical validity than one recorded without it, even though the pixel counts are higher. An important quality control procedure for cameras that use these methods is to periodically monitor the difference in the recorded count rates with and without the correction circuitry in operation. Another important factor is that some cameras may display the uncorrected counts on the scaler, thereby causing the scaler count to be different from the number of counts in the recorded image. A third consideration is whether or not the correct number of $Z$ signals appears at the output that is used as input to an external digital computer interface. If not, additional circuitry may be required to store on-line corrected images in an external computer.

**Variation of Unblank Time on Display CRT.** An alternative method to count skimming or count addition was introduced in 1977 (2). In this method (called “Autocomp”), the intensity of the displayed dot is modulated by varying the time of the unblank signal on the CRT. A reference or flood image is acquired and stored in a micro-processor memory and a matrix of correction factors is calculated in a manner similar to the count skimming and count addition methods. However, the correction factors in this case determine the relative time per dot that the CRT electron beam is allowed to strike the screen (unblank time). Areas seen as hot spots in the flood image display dots with a shorter unblank time than those with lower count rates, thereby producing a lower intensity dot. Proponents of this method point to the unalteration of the number of recorded counts as its main virtue. A serious drawback is the fact that the corrected data cannot be transmitted to an external computer, and an optional device to perform either the skimming or addition operation must be provided if external digital storage is desired. This method has now been abandoned in favor of those described in the next section.

**STATE-OF-THE-ART CORRECTION METHODS**

**Energy Correction**

A number of investigators performed experiments in the 1970s that showed the major causes of count density variations to be factors other than variations in point sensitivity (3–6). They showed that if the point sensitivity is measured at a large number of locations over the useful crystal area, and if the energy (pulse height analyzer) window is readjusted before each measurement, the observed count variations for most cameras are very close to the limits predicted by Poisson statistics. They concluded that the major cause of the count density variation were the variations in the relative positions of the pulse height spectra within the analog energy window and spatial distortions (errors in positioning) that are the result of several factors, including photocathode sensitivity, phototube gain, preamp gain, crystal light conversion efficiency, light pipe transmission efficiency, and the integrity of the optical couplings between crystal and light pipe and between light pipes (in those cameras that use light pipes). The energy-response variability is illustrated in figure 2. If the point sensitivity experiment is performed without readjusting the energy window, one does observe significant variability in the integrated count rates.

In 1977, a correction method was developed in which the position of the energy window was electronically adjusted to compensate for the local variations in the position of the photopeak (7). A block diagram is shown in figure 3. The circuitry is calibrated periodically by the user. A microprocessor is pro-

![FIG. 2. Hypothetical energy spectra recorded at two different spatial locations within a large area scintillation detector viewed by an array of phototubes show different peak locations (different size Z pulses for the same energy absorbed). By measuring the energy offset ($\Delta E$) for each pixel, one can store a matrix of fractional correction factors ($\Delta E/E$) from which an array of Z-signal correction factors ($\Delta Z$) can be calculated. The result is a superposition of all energy peaks from a given energy photon.](image-url)
programmed to record a pulse height spectrum for each pixel in a 64 × 64 matrix that is within the detector area. The position of the photopeak is determined for each pixel spectrum, and a correction factor is stored in the microprocessor memory that determines the position of the window for each event, depending on the pixel in which the event was detected. If the correction memory is loaded from scratch, ~ 60 million counts are required to analyze the required number of spectra. Since the count rate must not exceed 30,000-35,000 counts/sec (in order to avoid baseline shifts due to high count rates), the calibration takes 30-35 min. If one merely adds to and updates previously stored spectral data, the process requires 5-10 min. During acquisition of clinical data each x and y positioning signal is sampled and held until the microprocessor enters a lookup table and reads the position of the energy peak for the pixel that corresponds to the event location. A window centered about that peak location is established. If the detected signal is within the window, the image CRT is unblanked, otherwise it is not. Z signals with the same pulse height may be included in the image in one pixel and not in another. The energy correction on this camera enables one to use an asymmetric window to eliminate as much small angle scatter as possible.

In this type of microprocessor, a count skimming operation may also be applied to correct for any residual variations in count density that still exist subsequent to the energy correction. Both the energy correction and the count skimming are optional. Recently, this manufacturer* has added a linearity correction of the type described in the next section.

A similar energy correction scheme (8) has been employed by another manufacturer* with one significant difference; it alters the pulse height spectrum as opposed to moving the window. Incremental adjustments to the Z signals are made prior to pulse height analysis, hence the signal that arrives at the input to the stationary PHA window is Z + ΔZ, where Z is the original energy signal, and ΔZ is the increment added by the energy correction circuitry. ΔZ is actually the product f × Z, where f is a fractional energy correction factor (ΔE/E) computed for each pixel from the recorded spectra. An advantage of this method is its applicability to multiple energy windows. Since the fractional energy correction (f) should be independent of energy for a given pixel, one can compute a correction factor for every energy pulse that is within the linear range of the system electronics. The fact that the correction is made prior to pulse height analysis means multiple energy peaks are corrected. In one camera† a 128 × 128 response map is determined and stored in a microprocessor memory at the factory and cannot be updated by the user. Service personnel can, however, update it in the field.

**Linearity Correction**

Energy correction along with linearity correction represented the first comprehensive approach to correcting the root causes of the count density variations in scintillation cameras. The point sensitivity experiments cited earlier showed that the basic problem was due not to variations in photon detection efficiency but rather to mispositioning of events. Indeed, if one images a small disk source at many locations over the detector area of a camera with significant count density variations and measures the source diameter from the images, the apparent diameter will vary with position. This phenomenon suggests that there may be areas of local mispositioning that are not random (i.e., in preferential directions). In fact, if one designs cameras to achieve the best resolutions attainable, by using thin crystals and thin discrete-section light pipes (or no light pipe at all), count compression in the vicinities of individual phototubes is a well-known, predictable phenomenon. If these local nonlinearities could be mapped, a set of correction factors could be generated that would reposition the detected events to the true x, y locations of the original photon.

---

*Manufacturer's name removed for privacy.
†Manufacturer's name removed for privacy.
interactions within the limits of the intrinsic resolution of the camera. Figure 4 shows a diagrammatic representation of exaggerated nonlinearities in one-dimension and the determination of the correction factors. In this camera the nonuniformities are mapped at the factory for each detector assembly by imaging a rectangular grid source of $^{99m}$Tc. Correction factors are generated and stored in ROM chips in the form of incremental $\Delta x$ and $\Delta y$ signals that are added to the original analog $x$ and $y$ signals generated by the detector. A matrix of $64 \times 64$ correction factors are computed and stored initially. Then the numbers are bilinearly interpolated to produce a submatrix of $64 \times 64$ factors for each original factor. The result is a $4096 \times 4096$ array of correction factors ($\Delta x$ and $\Delta y$). Once the increments have been added to the digitized raw data signals, the analog signals that are applied to the deflection plates of the image CRT, and also output to an external computer, are the $x + \Delta x$ and $y + \Delta y$ corrected signals.

One inherently attractive feature of linearity correction by this method is that the number of detected events is unaltered; instead misplaced events are repositioned. The fact that the correction factors cannot be reestablished or updated by the user is perceived by some as a disadvantage. The factors that contribute to the energy and linearity errors may be inherent to a given detector assembly and may not change significantly with time. The manufacturer argues that as long as no disruption of the detector assembly occurs, the original factory generated correction factors should be valid. Clearly one must take great care to insure that the tuning of the phototube array be nearly optimal at all times for this camera to operate properly. Significant drift in the gains of individual phototubes renders the correction factors inaccurate for at least certain pixels.

Some manufacturers incorporate automatic phototube monitoring and gain adjustment, an important feature for cameras that do on-line linearity and energy corrections, because drifts in phototube gains subsequent to the establishment of the correction factors render the corrections invalid. These methods are discussed in the next section.

Phototubes may be replaced in the field in these cameras without reestablishing the correction factors. This can be done by first choosing phototubes that are matched to within very close tolerances in terms of photocathode efficiency and wavelength sensitivity for a particular detector assembly. Then, if individual phototubes require replacement, the replacement tubes are chosen to meet the same close tolerances as the originals. So long as no other detector characteristics are disturbed, according to the manufacturer, the original correction factors should remain valid. Figure 5 shows the effect of energy and linearity corrections on flood and phantom images. The left-hand images are uncorrected, and the ones in the right-hand column are corrected. The nonuniform pattern in the uncorrected flood image is not caused by inadequate PM tube tuning or improper energy calibration. The pattern is typical of cameras that are designed to yield high spatial resolution (i.e., in the trade-off between resolution and uniformity/linearity, the former is emphasized at the expense of the latter). The linearity corrections reposition the events to their proper locations.

Most major camera manufacturers have adopted the energy and linearity correction approach with automatic phototube gain adjustment. In some products, the energy and linearity correction factors are field updatable by the user and/or company service personnel. Figure 6 shows a simplified diagram that illustrates the principle of the energy and linearity correction method. The diagram does not apply to a specific manufacturer's product.

**Automatic PM Tube Adjustment**

Optimum performance of systems that do on-line energy and linearity corrections can be realized only if the PM tubes are properly tuned at all times. One camera with automatic monitoring and adjustment works as follows (Engdahl JC, per-
sonal communication, 1984): light emitting diodes (LEDs) are potted into the PM tubes, one per tube. The LEDs are pulsed on for 1-2 μsec hundreds of times per second. Light from the LEDs travels up the sides of the PM tubes, and is incident on the photocathode of the tube containing the LED as well as its six nearest neighbors. The light from a given LED is distributed approximately equally among the seven PM tubes. Since each tube has an LED, the response of a single tube to the light pulse from all the LEDs is the sum of inputs from many LEDs, but mostly from a group of seven (its own plus that of its six nearest neighbors). The LEDs are temperature compensated to minimize the variation in light output with temperature. Except for the temperature dependence, LEDs exhibit approximately ten times better stability over their expected lifetime than do PM tubes. The result is a constant light input to each PM tube per LED pulse over a time period that is long compared to the life of the PM tube. The number of photoelectrons generated within a PM tube from the LED light pulse is much greater than that resulting from a scintillation event in the crystal. The output charge is integrated over a few LED pulses, then converted to a voltage level which is compared to a reference voltage. If a given PM tube output differs from the reference voltage by more than a predetermined amount, an automatic gain adjustment is made for that tube. The gains are monitored several hundred times each second with actual gain adjustments made every 100 msec. If a tube cannot be brought into adjustment by the automatic feedback circuit, a warning light signals system failure.

The following are perceived by the manufacturer as advantages of the Autotune method:

1. The large number of photoelectrons produced in each PM tube by the LED light output reduces the statistical noise in the signal used to monitor PM tube gain compared to that associated with scintillation events from gamma rays.

2. The time between monitoring pulses (100 msec) is small compared to the time over which external influences are expected to cause gain shifts (e.g., rotating the camera in an external magnetic field).

FIG. 6. Simplified block diagram of the energy and linearity correction method of removing nonuniformity and spatial distortion. The diagram is for illustrative purposes only and does not apply to a specific manufacturer's product.
3. No optical coupling to the crystal or PM tubes is required as is the case in methods that use external light sources with fiber optics coupling.

Another variation uses a single light source that is coupled through the crystal to the PM tubes (9). The calibration procedure is computer controlled. The light source is switched on, stabilized, and the output of each PM tube is compared to a reference signal. The PM tube gains are automatically adjusted as required. The entire calibration process, including energy, linearity, and PM gain adjustment, takes 15 sec on the current model. The procedure is initiated by the operator as needed. An advantage claimed by the manufacturer is that the results of the calibrations are output to a magnetic disk and are retrievable for use as diagnostic aids by service personnel.

A further variation on automatic gain adjustment is a method (Burland K, personal communication, 1985; Woronowicz EM, personal communication, 1987) that monitors the count rate in two energy windows (with widths of 11 keV and 4 keV) positioned on the high side of the photopeak produced by an isotopic source. The windows are so situated in order to minimize the number of scatter pulses detected within the windows. The situation is illustrated in figure 7. Pulses from an individual PM tube are fed into the windows. When the count rates through the two windows are in a predefined ratio, the tube is properly tuned. The ratio is energy dependent, being ~1.0 for $^{99m}$Tc. An advantage of this method is that no light source is required inside the detector assembly. A disadvantage, compared to the LED method previously described, is the relatively high statistical noise in the count rate signals generated by the PM tubes in response to the photon flux emanating from the patient. In order to minimize the effect of the noise, a particular PM tube high voltage is readjusted only when the sum of the counts through the two windows in the vicinity of that tube exceeds 16K. Periodically the user acquires flood data which is used to iteratively apply the gain correction procedure and recalculate new energy correction maps. The advantage of this method cited by the manufacturer is the elimination of the effects of temperature-sensitive LEDs which present a potential problem if not properly compensated. Information on the status of the PM tubes is available to the user.

Another method, called "P correction" (Woronowicz EM, personal communication 1984), adjusts the amplification of each photomultiplier-amplifier combination in order to correct for drift due to tube aging and other causes of long-term instabilities. The correction factors are originally generated at the factory but can be reestablished at any time in the field. The recalibration procedure is automatic and the microprocessor is controlled. The collimator is exchanged for a calibration plate with one hole for each PM tube. A $^{99m}$Tc point source is used. The photopeak location is determined and compared with the initial calibration location for each PM tube. Correction factors are generated from the differences. The procedure takes ~30 min. According to the manufacturers, the accuracy of the corrected peak positions is within 1%.

CONCLUSION

In summary, virtually all scintillation cameras manufactured today have some type of on-line corrections for spatial response variations. State-of-the-art cameras no longer rely on symptom-related methods such as matrix multiplication, count skimming, and count addition. These methods have fallen in favor of energy, linearity, and automatic PM tube adjustment that correct the root causes of the variations. Accurate on-line corrections are essential in order to do acceptable SPECT imaging.

EDITOR'S NOTE

This article has been adapted from a chapter of the forthcoming SNM book, The Scintillation Camera, by permission of the editor and publisher.

NOTES

3. Muehllehner G. Colsher JG. Stoub EW. Correction for field nonuniform-


