nm/ CONCISE COMMUNICATION

A MEASURE FOR RADIOISOTOPE SCAN IMAGE QUALITY

Henri G. Vetter
New England Medical Center Hospitals, Boston, Massachusetts

and Stephen M. Pizer
Massachusetts General Hospital, Boston, Massachusetts
and University of North Carolina, Chapel Hill, North Carolina

A perennial difficulty in the field of radioisotope scanning* is the fact that the requirements for resolution and number of counts are specified independently although everyone knows that both contribute to the ultimate picture quality. This difficulty results from the fact that the image degradation due to blurring and that due to statistics are not directly comparable. Blurring introduces positional uncertainty whereas statistical fluctuations introduce uncertainty in brightness, that is to say, dot population density.

We assume that ultimately positional uncertainty of picture features is what we mean by quality (or rather the lack of quality). However, brightness uncertainties, far from being irrelevant, contribute to the overall positional uncertainties since it is only through some systematic brightness contrast that features are revealed to the observer.

"Brightness" can only be measured by evaluating the dot population over a sampling area of some significant extent. The statistical uncertainty of the brightness given by N counts can be decreased if N is made larger, but this can be done only by increasing the sampling area, and thus the positional uncertainty related to that reading. What we are after, then, is a formula giving the amount of positional uncertainty perceived by a human observer as a result of the brightness uncertainties and brightness contrast present in the picture. Combining this positional uncertainty due to statistics with the positional uncertainty due to blurring, one can then create a measure of overall positional uncertainty which presumably will reflect overall subjective quality of the scan. The existence of such a measure will permit us to optimally make the resolution sensitivity trade-off embodied in any instrument.

* In this paper the word "scanning" will refer to any radioisotope activity mapping process. The word "scan" will refer to the raw output of that device, consisting of a set of dots, each displayed as point of fixed density.

INDEX FOR POSITIONAL UNCERTAINTY DUE TO STATISTICS*

We have said that brightness uncertainty due to counting statistics is reflected as positional uncertainty through the size of the sampling area necessary to detect significant brightness changes (contrast) representing basic positional features, e.g. boundary elements. As an index for this positional uncertainty, we will take the diameter d_s of this sampling area. To determine the size of this sampling area a model must be postulated which embodies specific perceptual assumptions.

Three such assumptions were considered: (A) the eye counts the dots present on the sampling area (counting models); (B) the eye appraises the average distance between dots (distance models); (C) the eye appraises the average size of empty areas between dots (area models).

For each case the procedure involves reducing the sampling area until we reach a specified upper limit for the probability that an observed difference between the value of a particular index (e.g., number of dots) evaluated on either side of a presumed boundary might result by chance from a uniform random distribution. Making various assumptions specifying this composite uniform distribution leads to variations on each model identified by the subscripts 1, 2, and 3 in Table 1.

A summary of the model measures investigated is given in Table 1. We see that all the measures obtained are of the form $d_s = k \cdot f(\alpha) \cdot g(\lambda_1)$, where $g(\lambda_1)$ has in all cases the form $g(\lambda_1) = \lambda_1^{-1/2}$. Thus we can discuss the models on the basis of the α dependence only (Fig. 1).

Received Aug. 27, 1971; revision accepted March 9, 1971.

For reprints contact: Henri G. Vetter, Dept. of Therapeutic Radiology, New England Medical Center, 171 Harrison Ave., Boston, Mass. 02111.

* A more detailed account of the theory and experiments described in this and the following section has been reported elsewhere (*1*). Reprints can be obtained from the authors.

We note the following points: (A) the curve corresponding to the counting model is qualitatively different from all the other curves, and (B) the area model is distinguished from the other two models by its sensitivity to the statistical assumptions, behavior which brings this model into question because we consider the essence of the model to be more in its physical basis than in its statistical assumptions.

To evaluate the various models, sets of quantum-limited images (scans) were produced by computer simulation (2). Two-tone patterns are used so that the contrast, and thus the size of the sampling area, can be assumed constant for a given picture. Each set represents a different pattern, and within each set, the pattern was generated at various contrast levels and with various dot populations (examples in Fig. 2). All these pictures were simulated under the assumption of perfect geometric imaging.

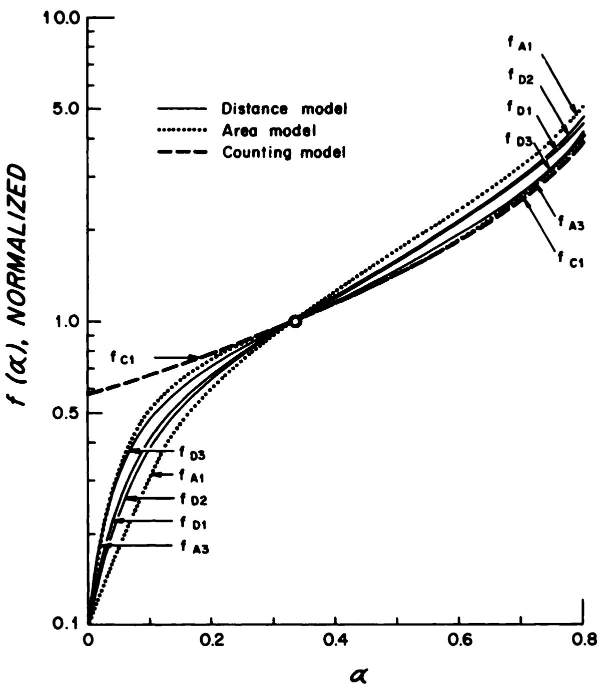

FIG. 1. Contrast dependence for various models (common point results from arbitrary normalization).

These sets, consisting of about 24 pictures each, with contrasts varying between 10 to 1 and 3 to 2, and dot populations between 1,000 and 10,000 were submitted to a group of 15 human observers who ranked each set in order of pictorial quality. These human rankings were then compared to model rankings based on the calculated index d_s.

In comparison with the human rankings, the distance models scored consistently better than the area models and very significantly better than the counting models. The analysis of these differences required the development of (A) measures for the agreement between a model ranking and the "human consensus," and (B) criteria for the significance of changes in the value of those measures from one model to the next. This matter was handled at length in another paper (*1*).

The distance model selected gives rankings which fit the "human consensus" as well as individual human rankings do on the average. In this model d_s is given as a function of contrast and dot density by the following equation

$$d_s = k\frac{\alpha^{1/2}(1+\alpha^{1/2})}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)} \qquad (1)$$

where α is the contrast defined as the ratio of the population densities (λ_2/λ_1) across the boundary to

TABLE 1. SUMMARY OF MODEL MEASURES

Statistical assumptions		Model type		
Subscript	Description	Counting (subscript C)	Distance (subscript D)	Area (subscript A)
1	σ of composite uniform distribution derived from uniform redistribution of dots	$d_{C1} = \frac{K(1+\alpha)^{1/2}}{\lambda_1^{1/2}(1-\alpha)}$	$d_{D1} = \frac{K\alpha^{1/2}(1+\alpha^{1/2})}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)}$	$d_{A1} = \frac{K\alpha}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$
2	σ of composite uniform distribution derived from index evaluated over whole sample area	$d_{C2} = d_{C1}$	$d_{D2} = \frac{K\alpha^{1/2}(1+\alpha^{1/2})^2}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$	$d_{A2} = d_{A1}$
3	σ of composite uniform distribution perceived directly from sample	$d_{C3} = d_{C1}$	$d_{D3} = \frac{K\alpha^{1/2}(1+\alpha^{1/2})\left(\frac{8}{\pi}-\frac{(1+\alpha^{1/2})^2}{1+\alpha}\right)^{1/2}}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)}$	$d_{A3} = \frac{K\alpha^{1/2}(1+\alpha^{3})^{1/2}}{\lambda_1^{1/2}(1-\alpha)(1+\alpha)^{3/2}}$

d = effective diameter of sampling area
λ_1 = dot density in denser region
α = contrast

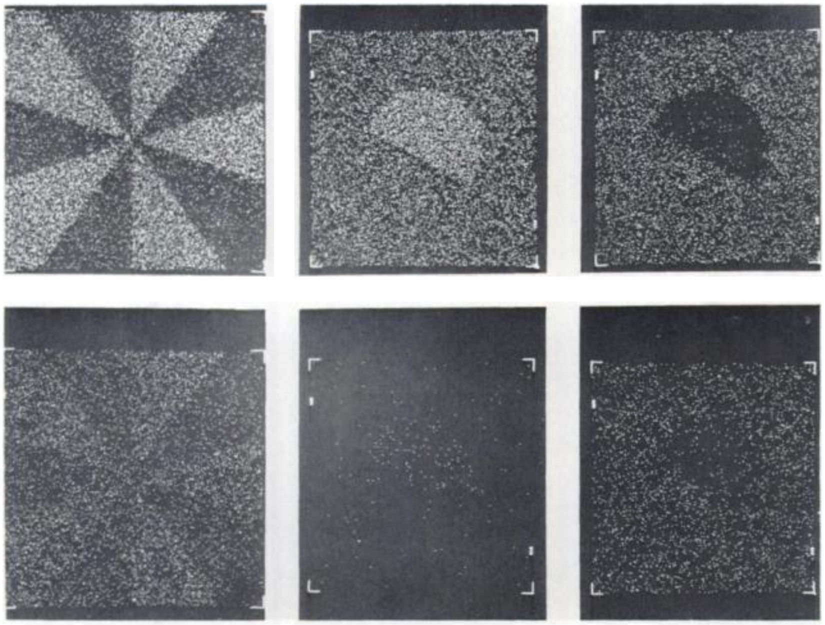

FIG. 2. Examples of scans used in this study.

be perceived, λ_1 is the dot population density in the denser area, and k is a constant scale factor which includes a human constant related to a confidence threshold level.

The superiority of the distance model suggests that the visual system approaches scan interpretation as a texture problem rather than blurring the dots and averaging luminous intensities, to treat it as an intensity contrast problem. This psychological result is itself interesting and deserves far more investigation, but this is not the object of this paper.

INDEX OF OVERALL POSITIONAL UNCERTAINTY

The quantity d_s, or the measure resulting from the statistical model, must be combined with a measure of geometric degradation, d_g, to produce an overall measure of degradation for a scan which is consistent with human evaluation of the scan. The geometric degradation encountered with any nonideal imaging device is fully characterized by the device's point-spread function. However, in the statistical context of scans the geometric degradation can be adequately characterized by a single index, the geometric resolution distance, d_g, defined as the diameter of the circle within which one half the flux of the two-dimensional spread function is contained. This measure has been shown to be reasonably independent of the shape of the spread function and to be in agreement with subjective human quality ratings (*3*).

The sampling area used by the eye for its statistical analysis also involves a degrading spread function, of unspecified shape, but characterized by d_s*. Since spread-function shapes are not crucial in this argument, we will estimate the overall effect by the formula:

$$d^2 = d_s^2 + d_g^2 \qquad (2)$$

where d is the resulting overall resolution distance. This formula is rigorously valid in the case of Gaussian functions.

Note that d_s has been determined only up to a scale factor, k, because of the relative nature of the ranking procedure used. A specific numerical value for k is needed, however, if Eq. 2 is to be used. The selection of k and the evaluation of Eq. 2 were achieved as follows. The observers were presented with 23 pictures with varying amounts of both statistical and geometric degradations and asked to match each picture in terms of overall quality with one of another set of pictures, containing statistical degradation only.

The value k = 8 was chosen as that value which produced matchings in closest agreement with human matchings. For k = 8 the matches predicted by the model do not differ more from the human consensus than an individual does on the average. With this value, d is given by

$$d^2 = d_g^2 + \left[\frac{8\alpha^{1/2}(1+\alpha^{1/2})}{\lambda_1^{1/2}(1+\alpha)(1-\alpha)}\right]^2. \qquad (3)$$

APPLICATIONS

Optimization of scanning parameters in clinical situations. If d is a valid measurement of overall degradation, the best compromise for a given clinical situation is that which minimizes d. This cannot be done freely but must satisfy physical constraints set

* Note, however, that in contrast with d_g, d_s varies with local scan characteristics.

by any practical instrument design and the nature of the source. In radioisotope scanning the constraint most frequently encountered is of the form $\lambda = C_1 d_g^2$, i.e., counting rate proportional to the square of the resolution distance, where the resolution distance is controllable by the selection of the collimator. Since d_s is proportional to $1/\lambda^{1/2}$, the constraint becomes $d_s = C_2/d_g$ and minimizing d under this constraint gives $d_s = d_g$. This condition has been tested on various sets of computer-generated radioisotope scans where each set embodied a different constraining constant C_2, and it leads to production of the "best picture" in good agreement with human selection. This condition can be used as a guideline to adjust or select those parameters which are left under the control of the clinician in the use of a specific instrument: duration of exposure, selection of a collimator, selection between two possible radionuclides where one produces higher differential uptakes (i.e., contrast), and the other higher counting rates.

Table 2, which gives the value of d_s for typical count densities and target contrasts, is provided for this purpose.

When dealing with an instrument with a constraint of the form $\lambda = C\, d_g^n$, the best balance occurs for $d_s = \left(\frac{2}{n}\right)^{1/2} d_g$, rather than $d_s = d_g$, but the Table is still valid.

Scan processing. The model permits the measurement of the overall degradation in a scan and the decomposition of this degradation into geometric and statistical components. The relative values of these components determine the nature of the processing methods to be applied. For example, if d_s is significantly less than d_g, the picture is called geometrically limited and processing should focus on retrieving geometric degradation. If the inverse is true, the picture is called statistically limited and processing should focus on cancelling statistical fluctuations by smoothing.

In either case, the improvement in one component is achieved at the expense of further degradation in the other, but some overall improvement can be achieved for a human observer to the extent that the overall degradation measure may be reduced. In this sense, the study gives a framework in which to consider processing methods and their design (*4*).

TABLE 2. VALUES OF d_s IN CM COMPUTED FOR VARIOUS DOT DENSITIES (λ) AND CONTRASTS (α)

Contrast	α:	0.1	0.2	0.333	0.5	0.667	0.800
	Count-ratio:	10:1	5:1	3:1	2:1	3:2	5:4
Count Density (Count/cm² in denser area.)	4,096	0.05	0.08	0.13	0.20	0.33	0.59
	2,048	0.07	0.12	0.18	0.28	0.47	0.83
	1,024	0.11	0.17	0.26	0.40	0.67	1.18
	512	0.15	0.24	0.36	0.57	0.95	1.66
	256	0.21	0.34	0.51	0.80	1.33	2.35
	128	0.30	0.48	0.73	1.14	1.89	3.33
	64	0.42	0.67	1.02	1.61	2.67	4.70
	32	0.59	0.95	1.45	2.27	3.77	6.65
	16	0.84	1.35	2.05	3.22	5.34	9.41
	8	1.19	1.90	2.89	4.55	7.55	13.30
	4	1.68	2.70	4.09	6.43	10.70	18.80

SUMMARY

A measure of scan quality has been developed by making quantitatively comparable the effects of statistics and blurring. Strictly speaking, the model has been shown to be useful for some two-tone patterns only, but because locally scans approximately reduce to such patterns, we can expect the measure to be generally applicable. This measure is useful in the specification of clinical scanning parameters and also in providing a framework in which to consider the design of scan processing methods.

ACKNOWLEDGMENTS

The authors are indebted to Gordon L. Brownell and Saul Aronow of the Physics Research Laboratory of the Massachusetts General Hospital for stimulating discussions and to the scientific personnel of the laboratory for their cooperation in ranking endless sets of pictures. This work was partially supported by U. S. Public Health Services Grant GM-889. The work reported was done while both authors were at the Physics Research Laboratory, Massachusetts General Hospital.

REFERENCES

1. VETTER HG, PIZER SM: Perception of quantum limited images. *Photogrammetric Engineering* 36: 1179–1188, 1970

2. PIZER SM: Simulation of radioisotope scans by computer. *Comm ACM 9* 5: 358, 1966

3. VETTER HG: Characterization of geometric imaging properties of gamma-ray detectors. *Int J Appl Radiat* 18: 231, 1967

4. PIZER SM, VETTER HG: Perception and processing of medical radioisotope scans. In *Pictorial Pattern Recognition*, Washington, DC, Thompson Book Co, 1968, pp 147–156