

RADIANCE AND PHOTON NOISE: Imaging in geometrical optics, physical optics, quantum optics and radiology

Harrison H. Barrett^{a,b}, Kyle J. Myers^c and Luca Caucci^b

^aCollege of Optical Sciences, University of Arizona, Tucson AZ 85721

^bCenter for Gamma-Ray Imaging, Dept. of Medical Imaging, Tucson AZ 85724

^cDiv. of Imaging and Applied Mathematics, Center for Devices and Radiological Health,
Food and Drug Administration, Silver Spring, MD 20993

ABSTRACT

A fundamental way of describing a photon-limited imaging system is in terms of a Poisson random process in spatial, angular and wavelength variables. The mean of this random process is the spectral radiance. The principle of conservation of radiance then allows a full characterization of the noise in the image (conditional on viewing a specified object). To elucidate these connections, we first review the definitions and basic properties of radiance as defined in terms of geometrical optics, radiology, physical optics and quantum optics. The propagation and conservation laws for radiance in each of these domains are reviewed. Then we distinguish four categories of imaging detectors that all respond in some way to the incident radiance, including the new category of photon-processing detectors. The relation between the radiance and the statistical properties of the detector output is discussed and related to task-based measures of image quality and the information content of a single detected photon.

Keywords: Radiance, geometrical optics, physical optics, quantum imaging, radiology, point processes, image quality

1. INTRODUCTION

Radiance is a familiar and important concept in optical engineering; it describes the radiant flux in an optical system as a function of 3D spatial position and ray direction, but only in the limit where diffraction can be ignored. If one knows the radiance throughout a system, other familiar radiometric quantities such as irradiance, radiant intensity and flux through some area can readily be computed (see Sec. 2).

Because x rays and gamma rays have very small wavelengths compared to most structures we encounter in imaging, these same definitions and properties of radiance apply with minor modifications in radiological imaging. On the other hand, diffraction often cannot be ignored at longer wavelengths, and for those cases a physical-optics counterpart of radiance, with similar conservation and propagation formulas, is available. There is even a form of radiance for the emerging field of quantum imaging. Details are given in Sec. 3.

Some basic propagation and conservation laws, discussed in Sec. 4, often allow the calculation of radiance for an arbitrary point in a system and an arbitrary ray direction if the radiance function is known just on a single plane. In imaging, these relations lead to simple ways of calculating the radiance in the image when one knows the radiance of the object, even for very complex imaging systems.

Nothing we have said so far about radiance has any obvious relation to the second term in the title, photon noise, but a common problem in optical engineering is the limited number of photons that can be collected and the resulting Poisson noise in any measurement of interest, including an image. As we shall see in detail in Secs. 5 and 6, any measurement we make in optics or radiology is ultimately described by a Poisson point process (PPP), and the complete statistical properties of this PPP are described by a radiance function. Thus the conservation

Further author information: Harrison H. Barrett, Center for Gamma-Ray Imaging, Radiology Research Laboratory, Arizona Health Sciences Center, Tucson, AZ 85724. E-mail: barrett@radiology.arizona.edu

and propagation rules for radiance tell us all we will ever need to know about the image statistics, at least for a single object.

Image statistics for a single object are important in assessment of image quality, but a full discussion of image quality must also account for the randomness of the objects being imaged and in some cases the randomness of the imaging system itself. This line of research is summarized in Sec. 7, where the concept of information content of a photon is introduced.

A useful general reference for this paper is Barrett and Myers, *Foundations of Image Science*,¹ especially Chapter 10, which deals with radiance, and Chapter 11, which deals with Poisson point processes.

2. RADIANCE IN GEOMETRICAL OPTICS AND RADIOLOGY

2.1. Definitions and notation

As noted in the introduction, radiance is a function of spatial position and ray or flux direction. The position of a point in space is specified by a 3D vector \mathbf{R} from some arbitrary origin, and the flux direction is specified by a unit vector $\hat{\mathbf{s}}$. Thus radiance is denoted $L(\mathbf{R}, \hat{\mathbf{s}})$. Because $\hat{\mathbf{s}}$ is a unit vector, it has only two independent components, say s_x and s_y ; the third component is given by $s_z = \pm\sqrt{1 - s_x^2 - s_y^2}$, so that $s_x^2 + s_y^2 + s_z^2 = 1$.

To define the radiance for the given point and flux direction, we construct an arbitrary plane P passing through the point. We denote the normal to this plane by the unit vector $\hat{\mathbf{n}}$. The equation of the plane is $\hat{\mathbf{n}} \cdot \mathbf{R} = \text{constant}$, so the position of the observation point within the plane is determined by the two components of \mathbf{R} perpendicular to $\hat{\mathbf{n}}$. We denote this 2D vector as \mathbf{r} ; if we choose coordinates so that the reference plane is perpendicular to the z axis (i.e. $\hat{\mathbf{n}}$ parallel to $\hat{\mathbf{z}}$), then $\mathbf{R} = (x, y, z)$ and $\mathbf{r} = (x, y)$

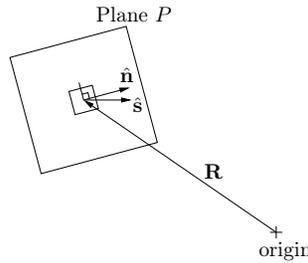


Figure 1. Geometry for specifying radiance on a reference plane. The 3D vector \mathbf{R} specifies a point on the plane; the 3D unit vector $\hat{\mathbf{n}}$ is perpendicular to the plane, and the 3D unit vector $\hat{\mathbf{s}}$ specifies the flux direction.

With this geometry, the radiance is defined as

$$L(\mathbf{R}, \hat{\mathbf{s}}) \equiv \frac{1}{\cos \theta} \frac{\partial^2 \Phi}{\partial A \partial \omega}. \quad (2.1)$$

where Φ is radiant flux (W), $d\omega$ is the differential solid angle of a cone of flux directions from a point on the plane, and dA is the differential area on the reference plane P (see Fig. 1). The $\cos \theta$ in (2.1) is defined as the scalar product between the flux direction $\hat{\mathbf{s}}$ and the unit vector $\hat{\mathbf{n}}$ normal to the reference plane:

$$\cos \theta = \hat{\mathbf{n}} \cdot \hat{\mathbf{s}}. \quad (2.2)$$

The cosine can be associated with either of the partial derivatives in (2.1), leading to two alternative definitions of radiance:

$$L(\mathbf{R}, \hat{\mathbf{s}}) = \frac{\partial^2 \Phi}{\partial A_{proj} \partial \omega} = \frac{\partial^2 \Phi}{\partial A \partial \Omega}, \quad (2.3)$$

where dA_{proj} is the projected differential area and $d\Omega$ is the projected differential solid angle. If we again choose coordinates so that the reference plane is perpendicular to the z axis, these projections are given by

$$dA = dx dy, \quad dA_{proj} = \cos \theta dA = \cos \theta dx dy; \quad (2.4)$$

$$d\omega = \sin \theta d\theta d\phi, \quad d\Omega = \cos \theta d\omega = \cos \theta \sin \theta d\theta d\phi. \quad (2.5)$$

The two alternative definitions in (2.3) show that radiance can be interpreted either as the flux per unit projected area per unit solid angle or as the flux per unit projected solid angle per unit actual area (on the reference plane). In both cases the word *projected* refers to projection onto a plane perpendicular to the flux direction through multiplication with $\cos \theta$. With either interpretation, radiance is invariant as the orientation of the reference plane is changed.

Since radiant flux is measured in watts (W), the SI units of radiance are watts per square meter per steradian, which we indicate as $[L] = W/(m^2 \cdot sr)$.

2.2. Other radiometric quantities derived from radiance

Irradiance, radiant intensity and radiant flux can all be found by appropriate integration if the radiance is known on a reference plane.

Irradiance (power per unit area) incident on the plane P is given by

$$E^P(\mathbf{r}) \equiv \frac{\partial \Phi}{\partial A} = \int_{hemi} L(\mathbf{R}, \hat{\mathbf{s}}) d\Omega = \int_{hemi} L(\mathbf{R}, \hat{\mathbf{s}}) \cos \theta d\omega, \quad [E] = W/m^2. \quad (2.6)$$

The 3D observation point \mathbf{R} is fully specified by the plane P indicated in the superscript and the 2D vector \mathbf{r} in that plane. For $\hat{\mathbf{n}}$ parallel to the z axis, for example, $E^P(\mathbf{r})$ is interpreted as the power per unit area on a plane of constant z , and $\mathbf{r} = (x, y)$. So long as the plane is specified, \mathbf{r} and \mathbf{R} determine the same point. The notation $E^P(\mathbf{r})$ indicates that the irradiance depends on (x, y) and on the location of the plane, z . Moreover, tilting the plane tilts the z axis and changes the irradiance; unlike radiance, the irradiance on a plane does depend on the orientation of the plane.

The two integrals in (2.6) run over hemispheres centered on the normal to the plane, so (2.5) applies, but $\int_{hemi} d\omega = 2\pi$ while $\int_{hemi} d\Omega = \pi$.

Radiant exitance, denoted $M^P(\mathbf{r})$, has the same definition as irradiance, (2.6), but it refers to flux reflected or emitted from plane P rather than incident on it. For a planar Lambertian source or reflector, where the radiance at the plane is independent of direction $\hat{\mathbf{s}}$, it follows from (2.6) that $M = \pi L$.

Radiant intensity, used to express the directional dependence of the flux from some area A on plane P , is given by

$$I^P(\hat{\mathbf{s}}) \equiv \frac{\partial \Phi}{\partial \omega} = \int_A L(\mathbf{R}, \hat{\mathbf{s}}) \cos \theta d^2r, \quad [I] = W/sr. \quad (2.7)$$

The radiant flux (power) through some area A on plane P is given by

$$\Phi^A = \int_A d^2r \int_{hemi} L(\mathbf{R}, \hat{\mathbf{s}}) \cos \theta d\omega d^2r, \quad [\Phi] = W. \quad (2.8)$$

2.3. Spectral radiance

If we wish to specify the spectral dependence of the radiance (per unit wavelength), we define [cf. (2.3)]

$$L_\lambda(\mathbf{R}, \hat{\mathbf{s}}, \lambda) \equiv \frac{\partial^3 \Phi}{\partial \lambda \partial A \partial \Omega}, \quad [L_\lambda] = \frac{W}{m^2 \cdot sr \cdot nm} \quad \text{if } [\lambda] = nm. \quad (2.9)$$

We can also define the spectral radiance in terms of the frequency of the radiation, $\nu = c/\lambda$:

$$L_\nu(\mathbf{R}, \hat{\mathbf{s}}, \nu) \equiv \frac{\partial^3 \Phi}{\partial \nu \partial A \partial \Omega}, \quad [L_\nu] = \frac{W}{m^2 \cdot sr \cdot Hz} \quad \text{where } [\nu] = Hz = s^{-1}, \quad (2.10)$$

or in terms of the photon energy $\mathcal{E} = h\nu$:

$$L_{\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}) \equiv \frac{\partial^3 \Phi}{\partial \mathcal{E} \partial A \partial \Omega}, \quad [L_{\mathcal{E}}] = \frac{W}{m^2 \cdot sr \cdot eV} \quad \text{if } [\mathcal{E}] = eV. \quad (2.11)$$

2.4. Flux and radiance in photon units

So far radiant flux has been in SI units of watts or joules per second, but for monochromatic radiation or for the spectral radiances defined above, we can express the flux in photons per second just by dividing the flux in joules/s by the photon energy in joules, denoted as \mathcal{E}_J . Thus, for example, the spectral radiance per unit wavelength from (2.9) becomes a spectral *photon* radiance, given by

$$L_{p,\lambda}(\mathbf{R}, \hat{\mathbf{s}}, \lambda) \equiv \frac{1}{\mathcal{E}_J} L_\lambda(\mathbf{R}, \hat{\mathbf{s}}, \lambda), \quad [L_{p,\lambda}] = \frac{\text{photons/s}}{m^2 \cdot sr \cdot nm} \quad \text{if } [\lambda] = nm. \quad (2.12)$$

The photon energy in joules is equal to the photon energy in electron volts multiplied by the charge of an electron in coulombs: $\mathcal{E}_J = e\mathcal{E}_{eV}$, $e = 1.6 \times 10^{-19}$.

The word *photon* in this usage has no quantum-mechanical implications; it is just a way of changing the units of flux and spectral radiance.

3. RADIANCE IN PHYSICAL OPTICS AND QUANTUM OPTICS

3.1. Physical optics

Physical optics deals with coherence and diffraction. The basic difficulty in integrating radiometry with coherence theory and scalar diffraction theory is that scalar fields are complex, while radiometric quantities are real and nonnegative.

For quasimonochromatic light of frequency ν (wavelength $\lambda = c/\nu$), Adriaan Walther²⁻⁴ defined the *generalized spectral radiance* (actually a spectral radiance per unit frequency) at a point \mathbf{R} on plane P by

$$L_\nu^P(\mathbf{R}, \hat{\mathbf{s}}, \nu) = \frac{\cos \theta}{\lambda^2} \int_P d^2 r' \langle e(\mathbf{R} + \frac{1}{2}\mathbf{R}') e^*(\mathbf{R} - \frac{1}{2}\mathbf{R}') \rangle \exp(-ik \hat{\mathbf{s}} \cdot \mathbf{R}'), \quad (3.1)$$

where θ is the usual angle between $\hat{\mathbf{s}}$ and the normal to the plane, $k = 2\pi/\lambda = 2\pi\nu/c_m$, $e(\mathbf{R})$ is a scalar electric field (e.g., one component of the electric field vector), the angle brackets denote a statistical average over an ensemble of scalar fields, and \mathbf{R} and \mathbf{R}' are 3D position vectors specifying points in the plane.

The expectation in (3.1) is a familiar quantity from coherence theory, the *mutual intensity*, and (3.1) states that (within constants) the generalized spectral radiance is the 2D Fourier transform of the mutual intensity. Thus Walther's radiance definition makes an important connection between radiometry and coherence.

Another connection, noted by Walther in his 1968 paper, is that the integral in (3.1) is a Wigner Distribution Function (WDF), which originated in quantum mechanics but is now an important tool in *phase-space optics*; see the book by Testorf et al.⁵ and the review article by Alonso.⁶

By construction, the WDF is real, but it can take on negative values, which of course the radiance defined by (2.1) cannot. Nevertheless, when one uses the generalized radiance to compute irradiance, radiant intensity or flux with the integral formulas from Sec. 2, the results are nonnegative.

3.2. Quantum optics

A quantum-mechanical definition of radiance can be given by a straightforward generalization of (3.1)¹:

$$L_\nu^P(\mathbf{R}, \hat{\mathbf{s}}, \nu) = \frac{\cos \theta}{\lambda^2} \int_P d^2 r' \langle \hat{e}^-(\mathbf{R} + \frac{1}{2}\mathbf{R}') \cdot \hat{e}^+(\mathbf{R}' - \frac{1}{2}\mathbf{R}') \rangle \exp(-ik \hat{\mathbf{s}} \cdot \mathbf{R}'), \quad (3.2)$$

where \hat{e}^+ and \hat{e}^- are electric field operators and the angle brackets now denote a quantum-mechanical expectation. The other radiometric quantities can then be computed just as in the classical case.

An example of a quantum imaging systems that can be described radiometrically by (3.2) is shown in Fig. 2. In the optical or infrared realm, the source of entangled photons in this figure can be a process called spontaneous parametric down-conversion (SPDC), but in medical imaging with positron-emission tomography (PET) it can be the two entangled gamma-ray photons produced by electron-positron annihilation.

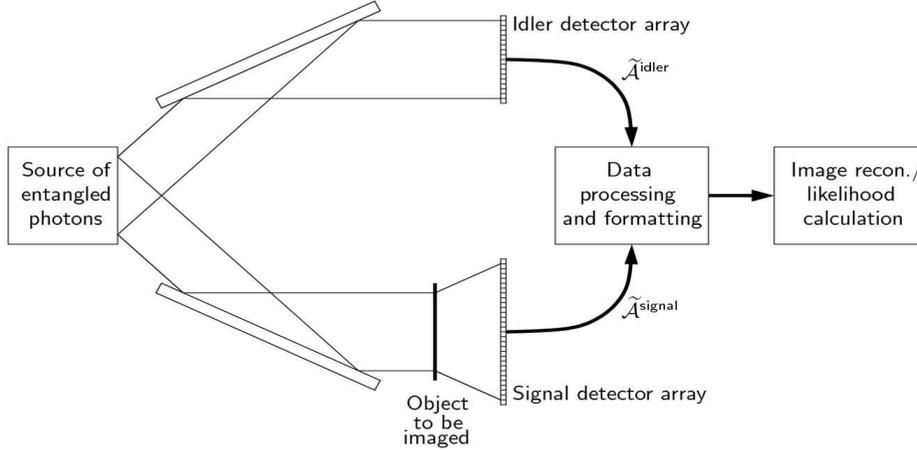


Figure 2. Quantum imaging with entangled photons.

4. TRANSPORT AND CONSERVATION OF RADIANCE

4.1. Boltzmann transport equation

The Boltzmann equation is an equation for the time derivative of the spectral photon radiance $L_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t)$, or $L_{p,\mathcal{E}}$ for short. This derivative has contributions from the physical processes of absorption, emission, scattering and propagation radiation, so we have

$$\frac{dL_{p,\mathcal{E}}}{dt} = \left[\frac{\partial L_{p,\mathcal{E}}}{\partial t} \right]_{abs} + \left[\frac{\partial L_{p,\mathcal{E}}}{\partial t} \right]_{em} + \left[\frac{\partial L_{p,\mathcal{E}}}{\partial t} \right]_{sc} + \left[\frac{\partial L_{p,\mathcal{E}}}{\partial t} \right]_{prop}, \quad (4.1)$$

where the subscripts have the obvious meanings. The four terms in this equation are derived in Sec. 10.3 of Ref. 1, and the resulting transport equation is

$$\frac{dL_{p,\mathcal{E}}}{dt} = -c_m \mu_{tot} L_{p,\mathcal{E}} + c_m \Xi_{p,\mathcal{E}} + \mathcal{K} L_{p,\mathcal{E}} - c_m \hat{\mathbf{s}} \cdot \nabla L_{p,\mathcal{E}}, \quad (4.2)$$

where c_m is the speed of light in the medium, μ_{tot} is the total attenuation coefficient (with contributions from both absorption and scattering processes), \mathcal{K} is an integral operator describing the angular and energy dependence of the scattering, and $\Xi_{p,\mathcal{E}}$ describes the photon source. All of these quantities can depend in general on \mathbf{R} , $\hat{\mathbf{s}}$, \mathcal{E} and t .

For a self-luminous volumetric source, as in fluorescent microscopy or nuclear medicine, the units of the source term are $[\Xi_{p,\mathcal{E}}] = (\text{photons/s})/(\text{m}^3 \cdot \text{sr} \cdot \text{eV})$. If the source emits isotropically, we can write

$$\Xi_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t) = \frac{1}{4\pi} f(\mathbf{R}, \mathcal{E}, t), \quad (4.3)$$

where $f(\mathbf{R}, \mathcal{E}, t)$ is the volumetric source we wish to image. Note that the 4π in (4.3) carries the units of sr .

For planar Lambertian sources (a common model in optics), the source is confined to a plane (say, $z = 0$), and the source term is written as

$$\Xi_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t) = \frac{1}{\pi} f(\mathbf{r}, \mathcal{E}, t) \delta(z). \quad (4.4)$$

Now $f(\mathbf{r}, \mathcal{E}, t)$ is the planar Lambertian source we wish to image. Note that $[\delta(z)] = 1/\text{m}$.

If the source is independent of time, or at least slowly varying, we can derive a steady-state form of the Boltzmann equation by setting $dL_{p,\mathcal{E}}/dt = 0$; the result is

$$c_m \hat{\mathbf{s}} \cdot \nabla L_{p,\mathcal{E}} + c_m \mu_{tot} L_{p,\mathcal{E}} - \mathcal{K} L_{p,\mathcal{E}} = c_m \Xi_{p,\mathcal{E}}, \quad (4.5)$$

If we use either of the two source models, (4.3) or (4.4), then (4.5) can be written abstractly as

$$\mathcal{B}\mathbf{L}_{p,\mathcal{E}} = \mathbf{f}, \quad (4.6)$$

where \mathbf{f} is the object to be imaged and (within constants) \mathcal{B} is the operator defined by the left side of (4.5).

If we assume there are no sources outside some volume V , then \mathcal{B} has a left inverse, which we denote by \mathcal{P} , so that

$$\mathcal{P}\mathbf{f} = \mathbf{L}_{p,\mathcal{E}}. \quad (4.7)$$

If we know the source distribution and the characteristics of the propagation medium and/or imaging system, we can calculate the spectral photon radiance at any point outside V .

4.2. Conservation of radiance along a ray in lossless media

A simple but important example of the Boltzmann formalism is to consider a volumetric, monochromatic, time-independent source $f(\mathbf{R})$ inside a medium with constant speed of light c_m and no absorption, scattering or other radiation sources. For points outside the source region, (4.5) becomes

$$\hat{\mathbf{s}} \cdot \nabla L_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}) = \frac{1}{4\pi} f(\mathbf{R}) \delta(\mathcal{E} - \mathcal{E}_0), \quad (4.8)$$

To solve this equation,¹ we choose a Cartesian coordinate system such that the z axis is parallel to $\hat{\mathbf{s}}$. Then $\hat{\mathbf{s}} \cdot \nabla = \frac{\partial}{\partial z}$, and (4.8) is an ordinary differential equation in z , which integrates to

$$L_{p,\mathcal{E}}(x, y, z, \hat{\mathbf{s}}, \mathcal{E}) = \frac{\delta(\mathcal{E} - \mathcal{E}_0)}{4\pi} \int_{-\infty}^z dz' f(x, y, z'). \quad (4.9)$$

If we define $\ell = z - z'$ and go back to a general vector notation, we can also write this result as

$$L_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}) = \frac{\delta(\mathcal{E} - \mathcal{E}_0)}{4\pi} = \int_0^\infty d\ell f(\mathbf{R} - \hat{\mathbf{s}}\ell, \hat{\mathbf{s}}) \equiv [\mathcal{P}\mathbf{f}](\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}). \quad (4.10)$$

As illustrated in Fig. 3, the interpretation of this equation is that $L_{p,\mathcal{E}}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$ is obtained by integrating the source distribution along a line parallel to $\hat{\mathbf{s}}$ and passing through the point \mathbf{R} . The operator that performs this integration is a special case of the operator \mathcal{P} defined in (4.7), and $[\mathcal{P}\mathbf{f}](\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$ is a spectral photon radiance with the indicated arguments.

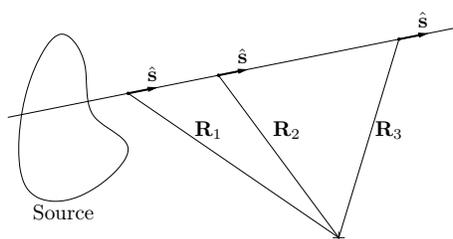


Figure 3. Diagram showing that radiance in a homogeneous medium is a line integral of a volumetric source that does not absorb or scatter its own radiation. The cross indicates the origin of coordinates for the three 3D position vectors.

It follows immediately from (4.10) that

$$L_{p,\mathcal{E}}(\mathbf{R}_1, \hat{\mathbf{s}}) = L_{p,\mathcal{E}}(\mathbf{R}_2, \hat{\mathbf{s}}) = L_{p,\mathcal{E}}(\mathbf{R}_3, \hat{\mathbf{s}}), \quad (4.11)$$

where \mathbf{R}_1 , \mathbf{R}_2 and \mathbf{R}_3 are three different points along the same line through the source, but all three lying outside the source region (see Fig. 3).

Since this expression holds for all photon energies and all units for the flux, it is valid for the total radiance as well as the spectral photon radiance:

$$L(\mathbf{R}_1, \hat{\mathbf{s}}) = L(\mathbf{R}_2, \hat{\mathbf{s}}) = L(\mathbf{R}_3, \hat{\mathbf{s}}). \quad (4.12)$$

Noting that the three radiances in (4.12) have the same angular argument $\hat{\mathbf{s}}$, we can call this line a ray, and (4.12) shows that radiance is constant along a ray in a homogeneous, lossless medium. Alternatively, we can define a ray as the path along which the radiance is constant, a path found by solving the Boltzmann transport equation.

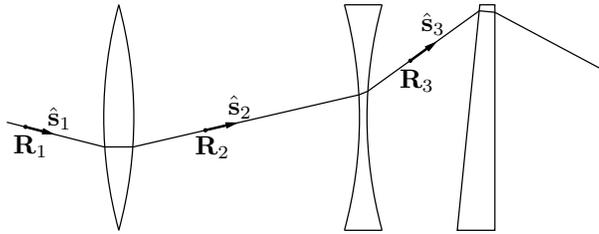


Figure 4. Diagram showing the invariance of radiance along a ray in a lossless optical system, as expressed by (4.13). The three points are vectors from a common origin of coordinates, as in Fig. 3, but the vectors themselves are omitted for simplicity in this figure. Note that all three points are in media of the same refractive index; otherwise L/n^2 would be the conserved quantity.

Moreover, this same conclusion holds if we consider a lossless optical system, i.e., one with no absorption, scattering or reflection losses, and trace any ray through it as in Fig. 4. Again, radiance is constant along the ray, but now the ray direction depends on which portion of the ray we examine, so we have to write

$$L(\mathbf{R}_1, \hat{\mathbf{s}}_1) = L(\mathbf{R}_2, \hat{\mathbf{s}}_2) = L(\mathbf{R}_3, \hat{\mathbf{s}}_3), \quad (4.13)$$

so long as all three points are in media of the same refractive index (e.g., air). If the points are in homogeneous media of different refractive indices, but again there are no losses, it can be shown that L/n^2 is the conserved quantity:

$$\frac{L(\mathbf{R}_1, \hat{\mathbf{s}}_1)}{n_1^2} = \frac{L(\mathbf{R}_2, \hat{\mathbf{s}}_2)}{n_2^2} = \frac{L(\mathbf{R}_3, \hat{\mathbf{s}}_3)}{n_3^2}. \quad (4.14)$$

4.3. Conservation of radiance in a lossless paraxial imaging system

ABCD matrices are an elementary tool for analyzing paraxial optical systems. If we consider two arbitrary planes perpendicular to the optical axis in a system with rotational symmetry, the ray height x and the angle of the ray with respect to the optical axis, α , in the two planes are well known to be related by

$$\begin{bmatrix} x \\ \alpha \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x' \\ \alpha' \end{bmatrix}, \quad (4.15)$$

If the system is paraxial but not rotationally symmetric, then we write

$$\begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{r}' \\ \mathbf{s}'_\perp \end{bmatrix} \equiv \mathbf{M} \begin{bmatrix} \mathbf{r}' \\ \mathbf{s}'_\perp \end{bmatrix}, \quad (4.16)$$

where \mathbf{s}_\perp is the projection of the flux direction $\hat{\mathbf{s}}$ onto the plane (i.e., perpendicular to $\hat{\mathbf{n}}$), and \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} are all 2×2 matrices, so \mathbf{M} is 4×4 . We refer to matrices in this form as **ABCD** matrices rather than $ABCD$ matrices.

The principle of conservation of radiance in a lossless paraxial system is most easily stated if we write the radiance as

$$L(\mathbf{r}, \hat{\mathbf{s}}) = L \left(\begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} \right). \quad (4.17)$$

With this notation, the radiance on an output plane is related to the radiance on a parallel input plane by

$$L_{out} \left(\begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} \right) = L_{in} \left(\mathbf{M} \begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} \right). \quad (4.18)$$

4.4. Transport equations for physical optics

A lesser-known application of **ABCD** matrices is to paraxial diffraction problems. The field propagation from input to output plane is described in the Fresnel approximation by⁷

$$e_{out}(\mathbf{r}) = -\frac{i}{\lambda \sqrt{\det \mathbf{B}}} \int_\infty d^2 r_0 e_{in}(\mathbf{r}_0) \exp \left[\frac{i\pi}{\lambda} (\mathbf{r}_0^t \mathbf{B}^{-1} \mathbf{A} \mathbf{r}_0 + \mathbf{r}^t \mathbf{D} \mathbf{B}^{-1} \mathbf{r} - 2 \mathbf{r}_0^t \mathbf{B}^{-1} \mathbf{r}) \right]. \quad (4.19)$$

Friberg⁸ studied the propagation of Walther's generalized radiance through systems described by **ABCD** matrices. With notation similar to (4.18), he showed that

$$L_\nu^{out} \left(\begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} \right) = L_\nu^{in} \left(\mathbf{M} \begin{bmatrix} \mathbf{r} \\ \mathbf{s}_\perp \end{bmatrix} \right). \quad (4.20)$$

Thus the physical-optics radiance is constant along the ray defined by geometrical optics. The same conclusion was reached by Walther (1978) by a stationary-phase approximation, valid asymptotically in the limit of zero wavelength. Friberg's derivation does not require this limit (though it does use the Fresnel approximation), and it is valid for any state of coherence of the fields.^{1,6}

A fully quantum mechanical version of Fresnel diffraction theory has been given by Shapiro.⁹ Combining this theory with the quantum-mechanical form of the Walther generalized radiance in (3.1) leads to a transport equation for quantum radiance, at least for paraxial propagation.

5. RADIANCE DETECTORS FOR IMAGING SYSTEMS

All radiation detectors are sensitive to the radiance incident on them. In particular, if we exclude so-called thermal detectors, they all respond in some way to the spectral photon radiance defined in Sec. 2, but in the subsections below we distinguish four different classes of detectors, depending on what information they yield about the spectral photon radiance function.

5.1. Integrating detectors

A CCD or CMOS camera produces a pixellated image related to the optically induced electrical charge accumulated at each pixel. A first approximation is that the mean charge (averaged over many images of the same object) is proportional to mean number of incident photons with energy above the bandgap energy during the exposure time. In fact, however, the mean charge produced is influenced also by the spectral, spatial, angular and temporal distribution of the flux. If the spectral photon radiance is known in the detector plane, and if the sensor array responds linearly to the radiation incident on it, the most general expression for the mean output at pixel m has the form¹

$$\bar{g}_m = \int_{det} d^2 r \int_0^\infty d\mathcal{E} \int_{hemi} d\Omega \int_0^\tau dt \eta_m(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}) L_{p,\mathcal{E}}^{det}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t), \quad (5.1)$$

where $\eta_m(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$ specifies the quantum efficiency of the detector, τ is the exposure time and \mathbf{R} is a general 3D vector, but the spatial integral is over the two variables needed to specify points in the detector plane.

Other locations for P may also be chosen. Various optical elements such as lenslets, pinholes, spectral filters, choppers or other optical modulators may be placed in the optical path between P and the detector array. The general form of the mean CCD/CMOS output signal in this case is

$$\bar{g}_m = \int_P d^2 r \int_0^\infty d\mathcal{E} \int_{hemi} d\Omega \int_0^\tau dt d_m(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t) L_{p,\mathcal{E}}^P(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t). \quad (5.2)$$

The function $d_m(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t)$ describes the propagation of radiance through the optical elements between plane P and the detector as well as the response of the detector to the radiance incident on it. One way to summarize (5.2) is to say each mean pixel output is a linear functional of the radiance in plane P , where the optical elements between the reference plane and the detector give us great freedom in choosing the weighting functions $d_m(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t)$. Cell-phone cameras and DSLRs are now available with 50 MP sensors, which means that we can measure 50 million independent linear functionals of the radiance function in an arbitrary plane.

One application of (5.2), of considerable current interest, is the *light-field camera*, pioneered by Ren Ng¹⁰ and illustrated in Fig. 5. The microlens array makes the photosensor array sensitive to the angular dependence of the radiance and not just the irradiance, and that allows the image to be refocused on different planes after it is recorded. Plane P in (5.2) is tangent to the left face of the microlens array in Fig. 5, so the light-field camera essentially captures samples of $L_{p,\mathcal{E}}^P(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}, t)$.

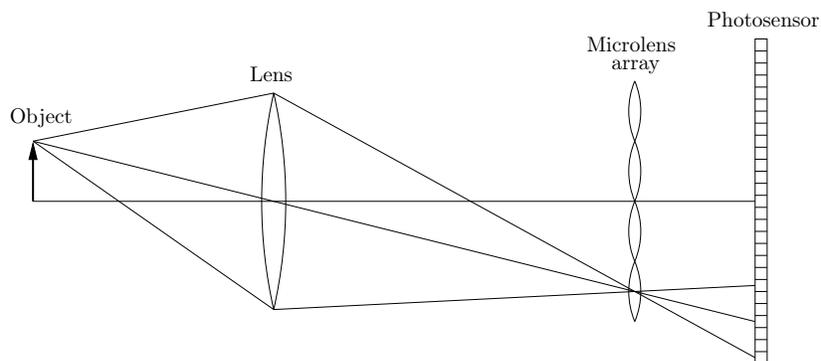


Figure 5. Schematic of a light-field camera (adapted from Ng¹⁰). The object is imaged onto the microlens array, and each microlens images the pupil of the main imaging system onto the photosensor array.

5.2. Pixellated photon-counting detectors

So-called photon counting detectors have circuitry at each pixel to detect the presence of a photoelectric interaction in the pixel and record the number of interactions (“counts”) during some exposure time for later readout. What is counted, therefore, is photoelectric interactions rather than photons; even a classical electromagnetic field induces photoelectric interactions. Nevertheless, we will continue to use common parlance and refer to a photoelectric interaction as absorption of a photon.

As a practical matter, it is easier to distinguish radiation-induced photoelectric interactions from thermally induced electrons (dark current) if some form of optical gain is used. Intensified CCDs, for example, use an optical image intensifier ahead of the CCD sensor so that a single optical photon absorbed in the photocathode of the intensifier produces a burst of light and hence multiple photons absorbed in the CCD element.

Another way of incorporating gain is to consider an array of avalanche photodiodes (APDs) as the image detector. In this case an absorbed photon produces an electron-hole pair, and each charge carrier is accelerated by an applied electric field to energies greater than the bandgap energy, so it can produce additional electron-hole pairs.

For x rays or gamma rays, the gain can be provided by scintillation materials for which each absorbed high-energy photon produces a burst of low-energy optical photons that can be detected by optical sensors.

Perhaps surprisingly, (5.1) and (5.2) are still valid descriptions of the mean output at pixel m for arrays of photon-counting detector. The actual outputs, g_m without the bar, are integers for photon-counting arrays, but the means are still linear functionals of the radiance so long as the event rate is low enough that there is no loss of counts from temporally overlapping events (pile-up) and no overflow of the registers that store the number of counts at each pixel.

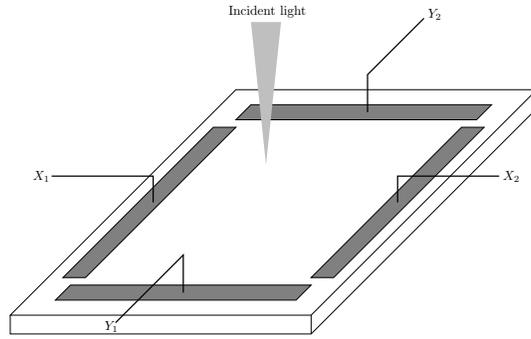


Figure 6. Position-sensitive silicon detector.

5.3. Position-sensitive photon-counting detectors

It is also possible to build photon-counting detectors without pixels. A common configuration is shown in Fig. 6, which shows a large-area silicon photodiode with strip electrodes on the four edges. Light incident in a spot, as shown, creates signals in all four electrodes, and a simple centroid calculation provides a measure of the x - y position of the spot.

The primary application of position-sensitive detectors is to track a beam or localized spot of light, but the addition of an image intensifier makes it possible for them to respond to individual photoelectric interactions. As in Sec. 5.2, a single interaction in the photocathode of the intensifier produces a burst of light and hence measurable signals at each of the four output electrodes.

Thus the x - y coordinates of the individual interactions can be computed and stored in real time, as the events occur; no integration time is needed, and there are no detector pixels. Instead, the data consist of a list of the coordinates, possibly tagged with the time at which the event occurred. This form of data storage is referred to as *list mode*, in contrast to the pixellated or *binned mode* implicit in (5.1) and (5.2). Listmode data storage has received considerable attention in the radiology literature,^{11–13} and it is used routinely in PET.

5.4. Photon-processing detectors

A photon-absorption event in a detector can be characterized by several *attributes*, including the position \mathbf{R} of the interaction, the energy \mathcal{E} deposited in the detector by the event, the direction $\hat{\mathbf{s}}$ of the radiant flux that produced the event and possibly the time t at which the event occurs. For the j^{th} event ($j = 1, \dots, J$), we can organize these attributes into an *attribute vector* with 5–7 components (2 or 3 components for \mathbf{R}_j , depending on the thickness of the detector; one component for energy \mathcal{E}_j ; two components for the unit vector $\hat{\mathbf{s}}$, and possibly time t_j). We denote this attribute vector for the j^{th} event by \mathbf{A}_j .

Of course we cannot directly measure the components of an attribute vector, but we can estimate them from suitable data sets. For example, if the position-sensitive detector of Fig. 6 is operated with an image intensifier, the pulse signals from the four electrodes can be used to estimate x_j and y_j , which are two of the 5–7 components of \mathbf{A}_j . To get sufficient information to estimate additional components, we need to add optical elements such as lenslets or color filters in front of the detector to alter its sensitivity to the spectral photon radiance [see (5.2) and Fig. 5).

The goal of a *photon-processing detector* is estimate the components of each attribute vector as accurately as possible. We denote the estimated attributes with tildes. Thus, if $\mathbf{A}_j = (\mathbf{R}_j, \hat{\mathbf{s}}_j, \mathcal{E}_j)$, then $\tilde{\mathbf{A}}_j = (\tilde{\mathbf{R}}_j, \tilde{\hat{\mathbf{s}}}_j, \tilde{\mathcal{E}}_j)$.

Of course these estimates will not constitute perfect measurements of the true photon attributes, but statistical estimation theory^{1, 14, 15} provides us with the tools to analyze and minimize the bias and variance of the estimates with respect to the true attributes.

An *efficient* estimator is one that is unbiased and has the smallest possible variance, as specified by the Cramèr-Rao lower bound. It is well known from statistical estimation theory that the *maximum-likelihood* (ML)

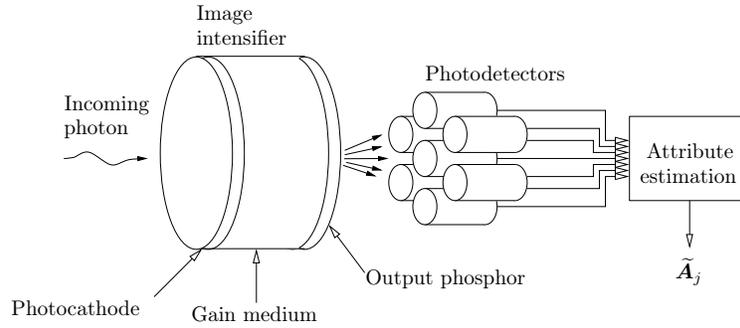


Figure 7. Generic photon-processing detector for optical wavelengths.

estimator is efficient in any problem where an efficient estimator exists, and in any case the ML estimator is asymptotically efficient as more or better data are acquired about each photon-absorption event.

For a review of ML estimation as applied to photon-processing detectors, see Barrett et al.¹⁶

To summarize what we know about how to implement optimum estimation of photon attributes, an ideal photon-processing detector must:

- Use a gain mechanism and multiple sensors (not just a single pixel) to obtain multiple non-redundant measurements for a single absorbed photon.
- Collect data from all sensors that respond to each event, at full precision; this form of data has been called *super listmode*.
- Use the super-listmode sensor data and specifically maximum-likelihood estimation (MLE) to estimate multiple attributes (e.g., 3D position, direction, energy, time) of each absorbed photon.
- Store all ML estimates of attributes in a list, at full precision

Figure 7 shows a conceptual diagram of a photon-processing detector capable of operating in the visible or NIR portions of the spectrum. A practical realization of a photon-processing detector for x rays, gamma rays and charged particles is the intensified Quantum Imaging Detector (iQID) developed by Brian Miller.^{17,18} Computer systems for performing the MLE step in real time are discussed by Hesterman et al.¹⁹

6. PROPERTIES OF IMAGE DATA FROM PHOTON-PROCESSING DETECTORS

So far we have discussed the mean values of data obtained by integrating detectors and photon-counting arrays (see Secs. 5.1 and 5.2), but we have not done the same for photon-processing detectors, and we have not discussed the statistical properties of the data in any case.

6.1. List-mode data and point processes

In practice, the estimated photon attributes will be stored in listmode, where each list entry is a K -dimensional vector, with K being the number of estimated attributes per event; $K = 6$ in our example where 3D position, two components of a unit vector and an energy are estimated.

A mathematically equivalent representation of the same information is a *random point process*, a sum of delta functions defined by

$$u(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}) \equiv \sum_{j=1}^J \delta(\mathbf{R} - \tilde{\mathbf{R}}_j) \delta(\hat{\mathbf{s}} - \tilde{\mathbf{s}}_j) \delta(\mathcal{E} - \tilde{\mathcal{E}}_j). \quad (6.1)$$

For a discussion of angular delta functions, $\delta(\hat{\mathbf{s}} - \tilde{\hat{\mathbf{s}}}_j)$, see Sec. 2.4.7 in Barrett and Myers,¹ and for a thorough treatment of random point processes, see Sec. 11.3.

Because of the independence of the photon-absorption events, the number of events J is a Poisson random variable, and $u(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$ is called a Poisson random process.

The mean of this point process is obtained by averaging over the statistics of each of the J vectors \mathbf{A}_j and over J itself for a given object \mathbf{f} that produced the data. The result is a function of \mathbf{A} which depends on \mathbf{f} , so we can denote it as $\bar{u}(\mathbf{A}|\mathbf{f})$, or more explicitly as $\bar{u}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}|\mathbf{f})$. Technically, this function can be regarded as a vector $\bar{\mathbf{u}}(\mathbf{f})$ in the Hilbert space $\mathbb{L}_2(\mathbb{R}^K)$, which is the space of square-integrable functions of K continuous variables.

6.2. Linear imaging of emissive objects

For self-luminous objects as in astronomy, nuclear medicine or fluorescence microscopy, Caucci and Barrett¹³ have shown that

$$\bar{\mathbf{u}}(\mathbf{f}) = \mathcal{L}\mathbf{f}, \quad (6.2)$$

where \mathcal{L} is a linear operator that maps $\mathbb{L}_2(\mathbb{R}^3)$ to $\mathbb{L}_2(\mathbb{R}^K)$. We emphasize that \mathcal{L} is linear even when the relation between raw data (e.g., sensor signals) and estimated attributes is nonlinear. Linearity depends only on the independence of the individual photon interactions.

For comparison, all linear photon-counting imaging systems that record only the number of counts in each of M bins are described by

$$\bar{\mathbf{g}}(\mathbf{f}) = \mathcal{H}\mathbf{f}, \quad (6.3)$$

where now \mathcal{H} maps $\mathbb{L}_2(\mathbb{R}^3)$ to an M -dimensional Euclidean space \mathbb{E}_M . In the terminology of Barrett and Myers,¹ \mathcal{H} is a continuous-to-discrete (CD) operator but \mathcal{L} is a continuous-to-continuous (CC) operator. The crucial difference between CC and CD operators is that the latter must necessarily have an infinite-dimensional null space; the system maps a vector in an infinite-dimensional Hilbert space to a finite set of numbers, so there is an infinite set of objects \mathbf{f}_{null} that yield no data at all: $\mathcal{H}\mathbf{f}_{null} = 0$. Because real objects almost always have null components for any CD system, the object itself cannot be recovered from the discrete data, and even simple linear functionals such as integrals of the object over voxels cannot be estimated without a possibly infinite bias. A CC operator like \mathcal{L} may have a null space, but if $K \geq 3$ the null space is not demanded by dimensionality considerations, so it is possible that many of the indeterminacy problems inherent in the world of conventional digital imaging may disappear with photon processing rather than mere photon counting.

6.3. Nonlinear imaging of transmissive objects

In many situations in optics and radiology, an image is formed by illuminating a transmissive object with a known radiation source. In these cases, the object is described by the spatial and spectral distributions of its attenuation and scattering coefficients. In terms of the Boltzmann equation (4.5), the object properties are embedded in the attenuation term $c_m \mu_{tot} L_{p,\mathcal{E}}$ and the scattering term $\mathcal{K} L_{p,\mathcal{E}}$ rather than the source term $c_m \Xi_{p,\mathcal{E}}$. The solution of the Boltzmann equation shows that the radiance is exponentially attenuated as it passes through the object, so the radiance on the image detector depends nonlinearly on the object.

We can still use (6.3) to describe the mean image data with a photon-counting or integrating detector, and (6.2) for a photon-processing detector, but now the operators \mathcal{H} and \mathcal{L} are nonlinear, and it is much more difficult to say anything about null functions.

6.4. Probability density functions for photon-processing data

The full probability law for a univariate Poisson random *variable* is determined by its mean. The same thing is true for a Poisson random *process* but now the mean is a function, $\bar{u}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$ and the probability law is a *probability density function* or PDF.

The 6D PDF for the attribute estimates is obtained by normalizing the mean:

$$\text{pr}(\tilde{\mathbf{R}}, \tilde{\hat{\mathbf{s}}}, \tilde{\mathcal{E}}|\mathbf{f}) = \frac{[\mathcal{L}\mathbf{f}](\tilde{\mathbf{R}}, \tilde{\hat{\mathbf{s}}}, \tilde{\mathcal{E}})}{\int_V d^3 R \int_{hemis} d\Omega_s \int_0^\infty d\mathcal{E} [\mathcal{L}\mathbf{f}](\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})}. \quad (6.4)$$

Because the events are statistically independent (conditional on \mathbf{f}), the overall PDF for the entire data set can be expressed as a product of 6D PDFs like those in (6.4).^{11,13}

6.5. Computation of the conditional mean of the Poisson Point Process

According to (6.4), the mean of the Poisson point process tells you everything about the statistics of the process, conditional on a specific object \mathbf{f} , but how do we compute that mean?

The first approach is to recognize that, within constants, the mean $\bar{u}(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}|\mathbf{f})$ is just a blurred version of the spectral photon radiance in plane P , $L_{p,\mathcal{E}}^P(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E}|\mathbf{f})$, with the blurring arising from the inaccuracies in estimating the photon attributes from the sensor data. Ways of accounting mathematically for this estimation blur have been discussed by Caucci and Barrett,¹³ so finding the mean of photon-processing data, and hence the full PDF by (6.4), boils down to computing the spectral photon radiance in plane P .

As we saw in Sec. 4.1, the spectral photon radiance is the solution of the Boltzmann transport equation, and often analytical solutions can be found; a simple example is (4.10). In many other situations $L_{p,\mathcal{E}}^P$ in a plane of interest P can be computed from the radiance in another plane, such as a source plane, by simple invocation of the principle of conservation of radiance along a ray, as discussed in Secs. 4.2 and 4.3. As we know from Sec. 4.4, these conservation and propagation laws apply to physical-optics and quantum-optics versions of radiance also.

For emissive objects, it may be possible to approximate the linear operator \mathcal{L} with a very large matrix and find its elements by simulation of the imaging system. For transmissive objects, \mathcal{L} is nonlinear, but nevertheless the Boltzmann can be solved to get the spectral photon radiance on the detector and from there get the mean data for any of the four detection modes discussed in Sec. 5. For an example of this approach as applied to diffuse optical imaging, see Jha et al.^{20–22}

When analytical solutions are not readily available, we can use Monte-Carlo simulation to estimate the spectral photon radiance. There is even a version of Monte-Carlo simulation, called adjoint Monte Carlo,¹ that allows direct estimation of the spectral photon radiance blurred by the estimation errors.

7. RELATION TO IMAGE QUALITY

7.1. Task-based assessment

In much of the medical-imaging community, and increasingly in other areas of imaging, image quality is defined in terms of the performance of some observer on some clinically or scientifically relevant task. Generically, the task can be either classification or estimation. An estimation (or quantitation) task is to determine a numerical value for one or more parameters associated with the object being imaged. Examples include cardiac ejection fraction or the concentration of a dye or contrast agent in some region of interest. In a classification task, the goal is to assign the object to one of two or more classes. A simple example is tumor detection, where the classes are tumor-present and tumor-absent. The observer for a classification task can be either a human or a mathematical algorithm. The performance of a human observer can be measured by well-established psychophysical methods and analysis of the resulting ROC (receiver operating characteristic) curves. Psychophysical studies are complicated and time-consuming, however, so there is considerable interest in developing mathematical model observers that will predict human performance. Another form of mathematical observer is the ideal observer, which makes optimal use of the available data in such a way as to maximize an ROC-based figure of merit. If we know the performance of the ideal observer, we can compare it to the performance of a human and determine how well the human can extract the information contained in the data. Often the performance of the true ideal observer is difficult to compute, so we use the optimal linear observer, which we call the Hotelling observer. It is optimal among all observers constrained to perform only linear observations on the data. An important variant on the Hotelling observer is the channelized Hotelling observer (CHO), which is the Hotelling observer applied not to an image but to a reduced data set consisting of the sampled outputs of a small number of linear filters, called channels. An important example is spatial-frequency-selective channels such as those known to be operative in the human visual system.

7.2. Object variability

For all of these tasks and observers, task performance is limited by the randomness in the data. So far we have discussed randomness arising from photon noise in different kinds of detectors, but there is also an inevitable statistical variability in the objects being imaged,²³ and in some cases the imaging system itself introduces a form of noise.²⁴

To account for the object variability in a classification task, we have to understand the statistics of the data conditional on a class \mathbf{C} of objects. For any data set \mathbf{g} , including the listmode data acquired by photon-processing detectors, it follows from elementary laws of probability that

$$\text{pr}(\mathbf{g}|\mathbf{C}) = \langle \text{pr}(\mathbf{g}|\mathbf{f}) \rangle_{\mathbf{f} \in \mathbf{C}} , \quad (7.1)$$

where $\text{pr}(\mathbf{g}|\mathbf{f})$ is the PDF of the data given a particular object \mathbf{f} , and the expectation is over all objects drawn from the class.

To visualize this expectation and its impact on classification performance, we can imagine using an accurate simulation program to generate N sample objects from class \mathbf{C} . Denoting the n^{th} sample object as \mathbf{f}_n ($n = 1, \dots, N$), we can approximate $\text{pr}(\mathbf{g}|\mathbf{C})$ by a sample average as

$$\text{pr}(\mathbf{g}|\mathbf{C}) \approx \frac{1}{N} \sum_{n=1}^N \text{pr}(\mathbf{g}|\mathbf{f}_n) . \quad (7.2)$$

For any form of photon noise, which is the subject of this paper, the PDF $\text{pr}(\mathbf{g}|\mathbf{f}_n)$ is fully determined by its conditional mean $\bar{\mathbf{g}}(\mathbf{f}_n)$. For a photon-counting detector with M pixels, therefore, (6.3) and (7.2) show that

$$\text{pr}(\mathbf{g}|\mathbf{C}) \approx \frac{1}{N} \sum_{n=1}^N \text{pr}(\mathbf{g}|\mathcal{H}\mathbf{f}_n) , \quad (7.3)$$

where \mathbf{g} is a vector in an M -D space with components $(g_1, g_2, g_3, \dots, g_M)$. These components are statistically independent conditional on an object \mathbf{f} , so $\text{pr}(\mathbf{g}|\mathcal{H}\mathbf{f}_n)$ is a product of M univariate Poissons, but the average over objects in (7.1) introduces statistical dependence. The accuracy of the approximation in (7.3) increases as the number of sample objects, N , increases.

To display this M -D PDF, we consider two components at a time. For example, we can plot the bivariate PDF $\text{pr}(g_m, g_{m'}|\mathbf{C})$ as in Fig. 8 for different total exposure times τ . The axes on these plots represent g_m and $g_{m'}$ (not m and m'), and the gray level is the PDF. The leftmost panel is for the largest τ , and the rightmost panel is for the smallest τ .

The points shown by \times on all three panels correspond to the values of $\mathcal{H}\mathbf{f}_n$ for the N sample objects used in (7.3). For this illustration, $N = 20$, and the bivariate PDFs for different objects are well separated on the leftmost panel where τ is large so the standard deviation of the counts per pixel is small compared to its mean. For smaller τ , however, the standard deviation of the counts on each pixel increases relative to its mean, and the individual PDFs, $\text{pr}(g_m, g_{m'}|\mathcal{H}\mathbf{f}_n)$ merge into one big blob, which for large N will be a good approximation to the class PDF, $\text{pr}(g_m, g_{m'}|\mathbf{C})$. The overall variance and covariance for the class PDF depends on both the exposure time (mean number of counts per pixel for a given object) and the randomness of the objects themselves.

A similar illustration holds for photon-processing detectors, but there the discrete index m is replaced by continuous indices like $(\mathbf{R}, \hat{\mathbf{s}}, \mathcal{E})$. The exposure time τ controls the total number of recorded events J in the data list for a particular object, and again the overall variance and covariance for the class PDF depend on both the exposure time and the randomness of the objects. For mathematical details, see Caucci.¹³

7.3. Information content of a photon

Consider an imaging problem where the task is to decide if the object that produced a particular image \mathbf{g} was from class \mathbf{C}_1 or class \mathbf{C}_2 . To perform this task, we need some knowledge the two class PDFs, $\text{pr}(\mathbf{g}|\mathbf{C}_1)$ and $\text{pr}(\mathbf{g}|\mathbf{C}_2)$. These densities can again be approximated by expressions like (7.3) if we can accurately simulate objects from the two classes.

A depiction of the data statistics like that in Fig. 8, but for two classes, is given in Fig. 9. It is visually evident that the classification task can be performed with high accuracy at large exposure time (leftmost image in the figure) but not for shorter exposures where the photon noise dominates.

It is beyond the scope of this paper, but it is possible to make these statements much more precise by specifying the observer that will perform the tasks (e.g., human or Hotelling observer) and by adopting some

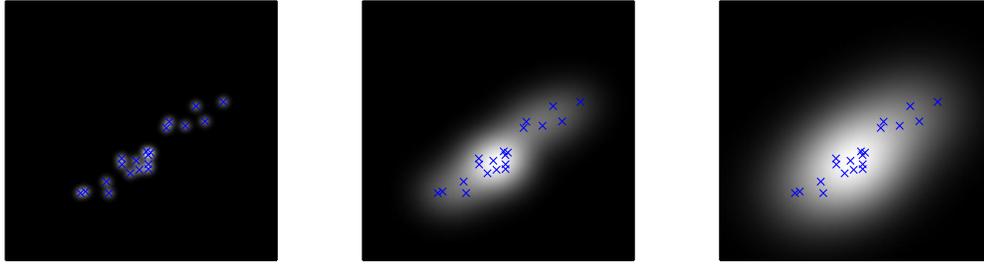


Figure 8. Illustration of the bivariate class PDF $\text{pr}(g_m, g_{m'} | \mathbf{C})$ for a class of objects \mathbf{C} and a photon-counting pixel array. Distributions are shown for three different values of the exposure time τ (increasing right to left). See text for explanation.

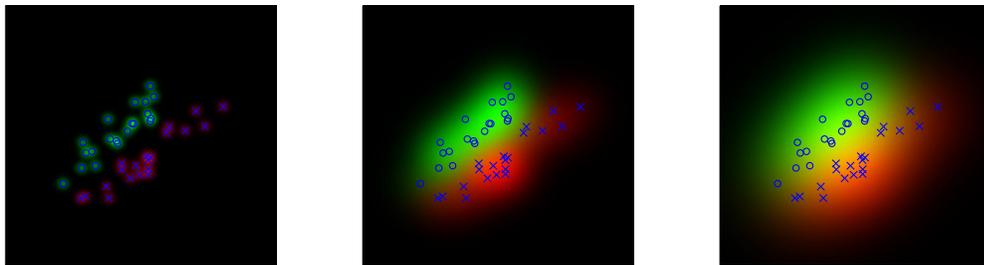


Figure 9. Illustration of the bivariate PDFs, $\text{pr}(g_m, g_{m'} | \mathbf{C}_1)$ and $\text{pr}(g_m, g_{m'} | \mathbf{C}_2)$, for two classes and a photon-counting pixel array. Distributions are shown for three different values of the exposure time τ (increasing right to left). See text for explanation.

measure of the task performance, such as area under an ROC curve. If one then plots this performance metric as a function of the mean number of absorbed photons, then one can define the information content per photon for this particular imaging system, task and observer. For a full mathematical exposition of this approach, with application to imaging with ionizing radiation, the reader can consult a forthcoming tutorial review article in *Physics in Medicine and Biology*.²⁵

8. SUMMARY AND CONCLUSIONS

This paper has considered two topics that at first blush might appear to be unrelated: radiance and photon noise. In one sense the connection is elementary: All radiation detectors respond to electromagnetic radiation through photoelectric interactions, and the statistical properties of the detector output are related to the inherent Poisson randomness of these interactions. Poisson probability laws are determined by their means, and the mean output from a detector is some functional of the radiance incident on the detector. In practical engineering problems we strive to develop highly linear detectors, so the relation between radiance incident on the detector and the mean of some Poisson process is usually also linear.

In the context of imaging, the radiance incident on the detector is some other functional of the radiance incident on or emitted by the object being imaged; this functional is linear for self-luminous or reflecting objects but nonlinear for transmissive objects. In either case, the basic problem in image analysis or task-based assessment of image quality is to understand the full relationship between object properties and the Poisson statistics on the detector output. This program requires clear definitions of radiance in various circumstance; methods for transport of the radiance from object to imaging detector; mathematical expressions for the mean response of

the detector to a given incident radiance; a clear statement of just what constitutes a detector output; an understanding of how the statistics of this output relate to the randomness not only of the photoelectric interactions but also of the objects being image, and finally a rigorous way of relating these statistics to image quality.

This paper has attempted to survey the current state of knowledge for all steps in this chain. We started with the usual elementary definition of radiance in terms of radiant flux, in the limit of geometrical optics, and we related this radiance to other radiometric quantities, such as irradiance and radiant intensity in the usual way. A particularly important radiometric quantity for this discussion turned out to be the spectral photon radiance, defined in (2.12).

Then we explored definitions of radiance valid for physical optics and quantum optics, and we studied the transport and conservation of radiance beginning with the Boltzmann transport equation. In particular, the crucial principle that radiance is constant along a ray for lossless optical systems was developed for both geometrical and physical optics.

We discussed four generic types of radiance detector for imaging systems. In addition to the familiar distinction between integrating and photon-counting detectors, we briefly discussed position-sensitive photon-counting detectors, which are usually used for beam-tracking applications but can be used for single-photon counting and position estimation with the addition of an image intensifier or other optical gain medium. An interesting facet of this class of detectors is that the output is a list of estimated interaction positions rather than an array of pixel values. Finally, we introduced a new class of devices, called photon-processing detectors, where multiple attributes of each photon interaction event can be estimated and stored in an expanded list.

The statistical properties of the listmode data from these last two classes of detectors were analyzed by introducing a Poisson point process, defined as a sum of delta functions in photon-attribute space; the mean of this point process is the spectral photon radiance.

We concluded with a brief summary of task-based assessment of image quality, with an emphasis on photon-counting and photon-processing detectors. This analysis led to a definition of the task-based information content of a single photon.

ACKNOWLEDGMENTS

We have benefited significantly from discussions with Abhinav Jha and Eric Clarkson. This research was supported by the National Institutes of Health under grants R37 EB000803 and P41 EB002035.

REFERENCES

1. Barrett, H. H. and Myers, K. J., [Foundations of Image Science], Wiley, New York (2004).
2. Walther, A., "Radiometry and coherence," *J. Opt. Soc. Am.* 58, 1256–1259 (1968).
3. Walther, A., "Radiometry and coherence," *J. Opt. Soc. Am.* 63, 1622–1623 (1973).
4. Walther, A., "Propagation of the generalized radiance through lenses," *J. Opt. Soc. Am.* 68, 1606–1610 (1978).
5. Testorf, M., Hennelly, B., and Ojeda-Castaneda, J., [Phase-Space Optics: Fundamentals and Applications], McGraw-Hill Professional (2009).
6. Alonso, M. A., "Wigner functions in optics: describing beams as ray bundles and pulses as particle ensembles," *Advances in Optics and Photonics* 3(4), 272–365 (2011).
7. Siegman, A. E., [Lasers], University Science Books, Mill Valley, CA (1986).
8. Friberg, A. T., "Propagation of a generalized radiance in paraxial optical systems," *J. Opt. Soc. Am. A* 30(18), 2443–2446 (1991).
9. Shapiro, J. H. "The Quantum Theory of Optical Communications." *IEEE Journal of Selected Topics in Quantum Electronics* 15(6), 1547–1569 (2009).
10. Ng, R., Levoy, M., Bredif, M., Duval, G., Horowitz, M., and Hanrahan, P., "Light field photography with a hand-held plenoptic camera," Stanford Tech Report CTSR 2005-02 (2005).
11. Barrett, H. H., Parra, L., and White, T. A., "List-mode likelihood," *J. Opt. Soc. Am. A* 14, 2914–2923, (1997).

12. Parra, L., and Barrett, H. H. "List-mode likelihood-EM algorithm and noise estimation demonstrated on 2D-PET," *IEEE Trans. Med. Imag.*, MI-17, 228-235 (1998).
13. Caucci, L. and Barrett, H. H., "Objective assessment of image quality V: Photon-counting detectors and list-mode data." *J. Opt. Soc. Am. A*, 29(6), 1003-1016 (2012).
14. Melsa, J. L. and Cohn, D. L., [Decision and Estimation Theory], McGraw-Hill, New York (1978).
15. Van Trees, H. L., [Detection, Estimation, and Modulation Theory, Vol. 1], Wiley, New York (1968).
16. Barrett, H. H., Hunter, W. C. J., Miller, B. W., Moore, S. K., Chen, Y., and Furenlid, L. R., "Maximum-likelihood methods for processing signals from gamma-ray detectors," plenary paper, *IEEE Trans. Nucl. Sci.* 56, 725-735 (2009).
17. Miller, B. W., Barber, H. B., Barrett, H. H., Liu, Z., Nagarkar, V. V., and Furenlid, L. R., "Progress in BazookaSPECT: high-resolution dynamic scintigraphy with large-area imagers," *Proc. SPIE.* 8508, 85080F (2012).
18. Miller, B. W., Gregory, S. J., Fuller, E. S., Barrett, H. H., Barber, H. B., and Furenlid, L. R., "The iQID camera: An ionizing-radiation quantum imaging detector," *Nuclear Instruments and Methods in Physics Research A*, in press (2014).
19. Hesterman, J. Y., Caucci, L., Kupinski, M. A., Barrett, H. H., and Furenlid, L. R., "Maximum-likelihood estimation with a contracting-grid search algorithm," *IEEE Trans. Nucl. Sci.*, 57(3), 1077-1084 (2010).
20. Jha, A. K., Kupinski, M. A., Masumura, T., Clarkson, E., Maslov, A. A., and Barrett, H. H., "A three-dimensional Neumann-series approach to model light transport in uniform media," *J. Opt. Soc. Am. A*, 29(8), 1741-1757 (2012).
21. Jha, A. K., Kupinski, M. A., Barrett, H. H., Clarkson, E., and Hartman, J. H., "A three-dimensional Neumann-series approach to model light transport in non-uniform media," *J. Opt. Soc. Am. A*, 29(8), 1885-1899 (2012).
22. Jha, A. K., Clarkson, E., and Kupinski, M. A., "An ideal-observer framework to investigate signal detectability in diffuse optical imaging", *Biomed. Opt. Expr.*, 4 (10), 2107-2123 (2013).
23. Barrett, H. H., "Objective assessment of image quality: Effects of quantum noise and object variability," *J. Opt. Soc. Am. A* 7, 1266-1278 (1990).
24. Barrett, H. H., Myers, K. J., Devaney, N., and Dainty, J. C., "Objective assessment of image quality: IV. Application to adaptive optics," *J. Opt. Soc. Am. A* 23, 3080-3105 (2006).
25. Barrett, H. H., Myers, K. J., Hoeschen, C., Kupinski, M. A., and Little, M. P., "Task-based measures of image quality and their relation to radiation dose and patient risk," Invited tutorial review in *Phys. Med. Biol.*, pending (2014).