The design of electro-optical sensors

Jean-Louis MEYZONNETTE

08.03.2023

Table des matières

I.Presentation

Module:

Semiconductor sensors and applications

Author(s):

Jean-Louis MEYZONNETTE - Institut d'Optique – Graduate School

Abstract:

This lesson aims at giving basic elements of optronic captors conception. First, we remind the main applications of optronics ; the steps to follow to conceive such systems and the big family of captors. Then we describe the most useful notions of radiometry and propagation to calculate the optical signal received by the detector. Finally, we end the calculation of the noise signal report, and the evaluation of such captors, under the form of an energetic assessment.

Keywords:

Electro-optical sensors, Bases in optical radiometry, Signal flux in an electro-optical sensor, Performance evaluation,

Prerequisites:

Licence ou première année d'école d'ingénieur

Learning outcomes:

Comprendre le comportement des capteurs optiques à traitement électronique et acquérir les bases de leur conception.

Course overview:

- Introduction
- Electro-optical sensors
- Bases of optical radiometry
- The signal in an electro-optical sensor
- Signal to noise ratio and performance of an electro-optical sensor

Conception & production :

Le Mans Université

License:

1 - http://www.gnu.org/licenses/fdl.txt

II.Cours

This course is aimed at giving some of the bases that are most often used in the design of electro-optical instruments, i.e. sensors assembled from optical and electrical components. The first part of the course deals with their main applications, their main design steps, and their classification. Because the input of an electro-optical sensor is optical, the second part goes over some basic radiometric parameters, corresponding relationships, and ends up by characterizing some typical radiations. The third part is concerned with the propagation of light and basic properties of optical systems; it also evaluates the optical input, or flux, that is incident upon the detector of such sensors, inside which it is converted into the electrical output. Finally, the last part shows how the performance of an EO sensor is based upon the value of its « signal to noise ratio », and often (in domains such as defense, space or telecommunications) expressed in terms of « maximum operating range ».

1. Electro-optical sensors

1.1. Domains of application

The association of **«electronics»** and **«optics»** leads to a large number of terms :

for example, « **electro-optic** » materials are materials with properties (such as transmittance, refractive index, birefringence,...) that may be modified under application of some electrical field.

"Optoelectronics" covers the domain of components that transform photons into electrons, or vice-versa (detectors, lasers, light emitting diodes, laser diodes,...), or those that carry light (optical fibers), amplify it, or modulate it.

As for the term **«electro-optical»**, it is not so much related to components, but rather to the sensors, devices, instuments, equipments or systems which are made up of optical, optoelectronic, electro-optic and electronic components. These sensors are utilized in quite many domains, amongst which :

- **Defense** : electro-optical equipments are more covert than radars, to which they are often associated, because many of them are completely passive (they emit no radiation) and the ones that emit, radiate much narrower beams than radars ; on the other hand, a traditional strong point of optics is its angular precision, which allows EO sensors to greatly improve the quality of observation on the battlefield, and to be of prime interest in target recognition and identification, or in terminal missile guidance. EO countermeasure sensors ar also on the rise in order to neutralize these highly efficient equipments.
- **Space** : electro-optical sensors intervene in space applications as satellite navigation aids (horizon and star sensors), in earth and space observation, in vegetation monitoring, in astronomy,...There also exist research programs in space telecommunications by means of lasers, that could greatly optimize long range transmissions, for example between geostationary and low orbit satellites.
- **Telecommunications** : there are many reasons that explain the increasing use of optical fibers in telecommunications : low losses, high bandwidths, low weight and volume, possible exposure to high voltages, absence of electrical flashes, of ground loops, very little influence of optical or electrical perturbations, ... Very high rate fiber optics terminals are now available, at relatively low cost.
- **Industry** : More and more laser equipments can be found in industrial fabrication processes, such as cutting, soldering, marking. Quite many electro-optical sensors are utilized in industrial controls, in order to measure, among other parameters :

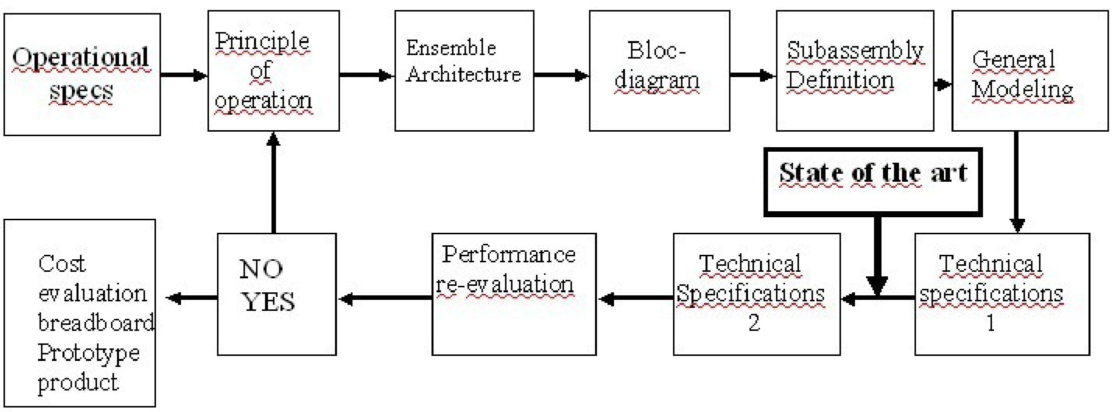
temperatures, flow rates (fluids), displacements, velocity, acceleration, pressure (fluids, acoustic), humidity, stress, forces, charges, gas, pollution, level (liquids), vibrations, shocks, couples, surface shapes...These contactless sensors do not perturb the parameters to be observed, and they may operate in the presence of agressions, either chemical (corrosion), mechanical (vibrations, pressure, shocks), thermal, electrical, magnetic, or radiative (nuclear environment), because of the possibility of locating the sensor far from the measurement area.. Their usage is growing in robotics (shape, object, defect recognition by means of image processing), industrial and environnemental surveillance, chemical analysis, metrology,...They are invading the transportation and automotive industries, particularly as driving aids and for anticollision purposes.

- **Mass Media** : Images are taking more and more room in everyday life, and EO sensors are accompanying this explosion, as well at the image production level as in image display : digital still or movie cameras, projection screens, displays, minicameras for portable phones, compact discs readers, ...
- **Scientific research** : Quite a number of electro-optical sensors are being integrated in scientific projects of national and international importance, such as those of several Atomic Energy Commissions (MegaJoule Laser project of CEA in France, National Ignition Facility in the United States), in astronomy (large telescopes), as well as in numerous research laboratories, either in industry or at Universities.
- **Medical** : Medicine is utilizing more and more electro-optical sensors, be it for diagnosis (thermography, laser imagery) or for medical care (laser surgery).

1.2. Main phases in the design of an electro-optical sensor

a) Introduction

In order to design an electro-optical sensor properly (or any other kind of sensor: capacitive, inductive, with stress gauges, piezoelectric, Hall effect, micro-wave, ultrasonic,...), it is advisable to know the state of the art, i.e. most of the sensors that already exist in the domain of interest, in order to evaluate the possible sources of difficulties, particularly when one wants to introduce a product on a new market : one has to get used to the procedures and norms before introducing a device that may be more complex than usual, which generally means additionnal training and maintenance costs. It is hence necessary for the designer to optimize each phase of its design : operational specifications, principle of operation and architecture, modelling, technical specifications of components and subassemblies, evaluation and test benches, breadboards and prototypes (see design sequences, figure 1 below).



b) Analysing or setting operational specifications

The "operational specifications" of a sensor are a list of its functions and constraints. They are generally defined by the customer, but they may also be established by the designer himself : these specifications must be as precise as possible, about all the missions that are expected from the device, its performances, the environmental conditions of use, hierarchy or priority between functions, when it is a multifunction sensor. Depending upon the sensor complexity, operational specifications may add up to a few pages or correspond to thick documents. Most often, the principle of operation is not mentionned here, neither are its technical specifications, but rather the results to obtain.

c) Choosing an architecture

Starting from the analysis of the sensor operational specifications, the designer will have to define its principle of operation, i.e. the physical (here, optical) phenomenon which the whole system will rely upon to reach its goals. He must lay down the main parameters and their influence on the expected performance, the characteristics of the optical and electrical signals, the signal processing techniques, and then define the sensor subassemblies. In case several principles of operation and architectures are possible in order to solve a given problem, it is necessary to pursue the analyses far enough (i.e. by taking into account the existing technologies) for the advantages and drawbacks of each solution to appear.

At this level, the architecture of the proposed sensor is converted into a « **block-diagram** », that describes its « **electro-optical chain** ».

d) Modelling / performance simulation/ optimization

After defining the architecture of the sensor, the designer starts simulating its performance by means of a « black box » model in which each subassembly is represented by its inputs, issued from the preceding black box, and by its outputs, delivered to the following one, and by its « transfer function ».

Modelling of an electro-optical system is usually separated into two parts : the first one, that goes from the source to the detector (or up to its its preamplifier), must make sure that the electrical output from the detector is strong enough to be processed with success. For that purpose, the model defines the radiometric budget of the sensor, i.e. it evaluates the optical signal, or flux, that is incident upon the detector. This budget ends up by the computation of the signal to noise ratio at the output of the detector, inside the electronic bandwidth of the sensor : this parameter is fundamental in the performance of the sensor and if its value is found to be too low, then it is worthless to start modelling the whole sensor since actual electronic circuits will further degrade this SNR, and no processing technique, whatever its quality, will be able to restaure the situation.

Modelling procedures and corresponding softwares must be specific to each part of the sensor. One can find numerous computer aided design softwares dedicated to optics (emission and propagation of light, optical properties of surfaces and media, optical design, detectors), to mechanics and electronics (signal modulation, demodulation and filtering, analog and digital signal processing, image processing).

These softwares help eliminate the boresome side of computations and lead to finer and finer results because they allow the designer to test and analyze rather large numbers of configurations. However, the designer must not forget to be critical of the results thus obtained, because the validity of such programs rests upon the soundness of the model he himself has chosen.

e) Technical specifications

The results from simulation are then used to define the techinal specifications of all the components and subassemblies in the sensor, taking into account the **state of the art**. This phase will profit from the experience of the designer with previous designs, and is optimized by consulting the literature (technical journals, internet), taking courses in continuing education programs, going to conferences,...

f) Control and Testing

One must not wait till the instrument has been designed before one starts to think about ways of testing and evaluating it : the designs of both the sensor itself and of its test procedures and benches must be carried out in parallel, because the test definition always brings some help into the design of the device, at all levels (components, subassemblies, system). In case one has to design complex electro-optical sensors, or if there is a small number of items to fabricate, as is the case in defense and space applications, test and evaluation equipments generally represent a non negligible part of the project : a wrong definition or an underestimation of the costs for example may be fatal to the projector even to the company.

To conclude about this design phase, a « golden rule » , to be followed as strictly as possible when setting up breadboards and prototypes, is that every component and subassembly must be controlled and tested before assembly, if one wants to avoid costly dissassembly and reassembly efforts in case of unexplained failures at the system's level.

g) Remarks / Traps to avoid

There may not be possible to put the chosen principle into operation because the necessary technology is not available yet. In that case, one has to evaluate the chances for such a technology to get through and how long that will take, by closely following the progress being achieved in the research labs involved in the corresponding domain.

As was mentionned before, it is of the upmost importance to make operational specifications as precise as possible. The designer must insist on clearing up any ambiguity, especially on which goals are vital and which ones are not, if the sensor is multi-functional.

By definition, the design of an electro-optical sensor is a multidisciplinary job, and it must be well balanced in the evaluation and study of the various contributing elements, without neglecting or underestimating any one of them : any weakness concerning one component or another (source, detector, propagation medium, optical instrument, mechanical parts, servo-controls, filtering, analog and digital signal processing, image processing, display, software,...) will affect the global performance of the sensor.

1.3. Main families of electro-optical sensors

a) Introduction

Depending upon the application, electro-optical sensors greatly differ from each other. However, they may be divided into a few families, by means of the following criteria : operating principle, spectral domain, type of output, and detection mode.

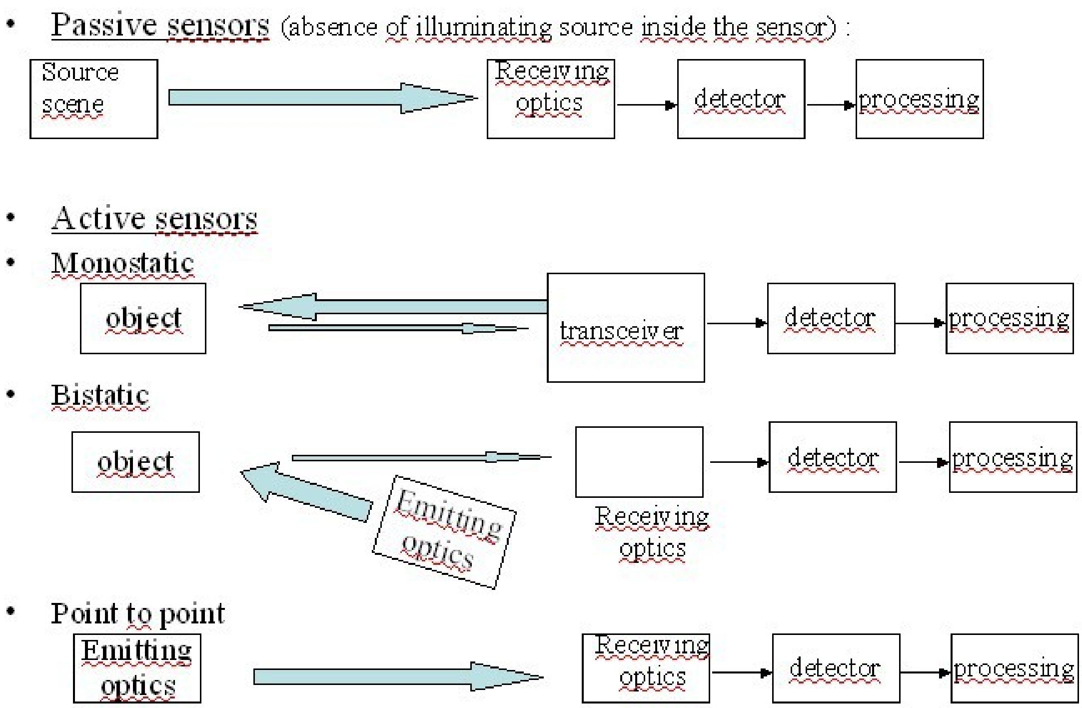
b) Classification by operating principle : passive and active sensors

This first criterion separates EO sensors into two families, i.e. those having their own light source (**active sensors**) and those that don't (**passive sensors**). Passive sensors, which are more widespread than the active ones, observe and detect radiations, without emitting any. In this category, one will find most of the cameras.

Active electro-optical sensors possess their own light source ; some of them use it in order to illuminate some object and collect some of the light reflected from it. These active systems are said to be « **monostatic** », if the light source is situated at the same place or very close to the collector. They are « **bistatic** » whenever the transmitter and the receiver are well separated in space from each other.

There exists a third type of active sensors, corresponding to the configuration where the transmitter is emitting light directly towards the receiver, either in free space (atmosphere, vacuum) or through a guided medium (optical fibers). This type of sensor is said to be a

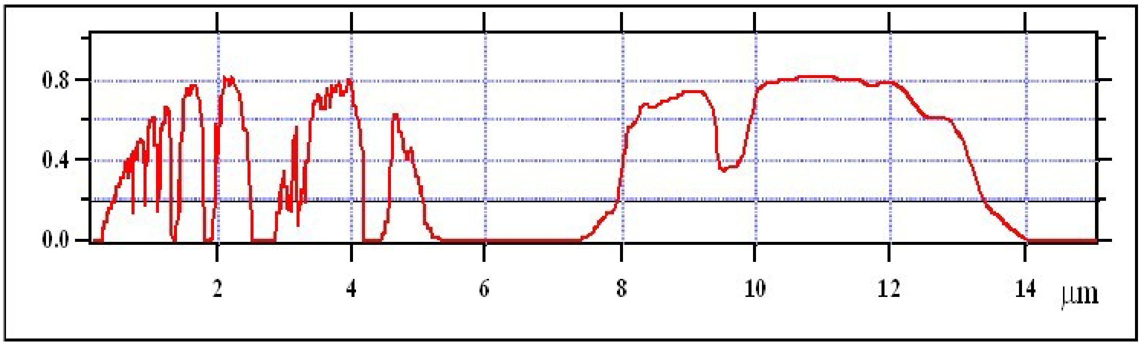
« **point to point** » or « **direct transmission** » sensor. Figure 2 below summarizes these main typical configurations.



c) Classification by spectral domain : UV, visible, IR sensors

Situated between X rays and microwaves, the domain of optics deals with radiations of wavelengths approximately from $0.02\mu m$ and $500\mu m$, and is divided into three parts : Ultraviolet ($0,02$ to $0,4\mu m$), Visible ($0,4$ to $0,7\mu m$), and Infrared ($0,7$ à $500\mu m$).

If a sensor is supposed to operate in vacuum, there will not be any constraint emanating from the propagating medium concerning the best wavelength or spectral domain of operation. But if the radiation to be detected propagates through some medium before reaching the sensor, then that medium will impose some optimal spectal domain, or « window » especially if the propagation path is lengthy. The most usual propagating medium is the atmosphere, which sets its limits between about 0.2 et $15\mu m$ (a typical curve of spectral transmittance of the atmosphere is given, figure 3). Of course the optimal band is not the same for other media : blue-green for water, near IR ($1,3\mu m$, $1,5\mu m$) for optical fibers.



Electro-optical sensors operating in the same spectral domain will share some similarities even if they functions are quite different. This resemblance may concern for example the type of radiation they are dealing with, its origin, its propagation, or its detection .

d) Classification by output information : image forming sensors, flux collectors

Informations delivered by electro-optical sensors are essentially of two types : either images or radiation levels (flux).

An **image forming sensor** delivers the spatial distribution of a given radiation (light source, object) by analyzing it along different directions or positions in space («picture elements» or «pixels») by means of its **radiance** (this parameter will be defined, § 2.1). One of the basic performances of an image forming sensor is its resolution, or number of pixels. This family of sensors corresponds, for example, to movie or still cameras, ...

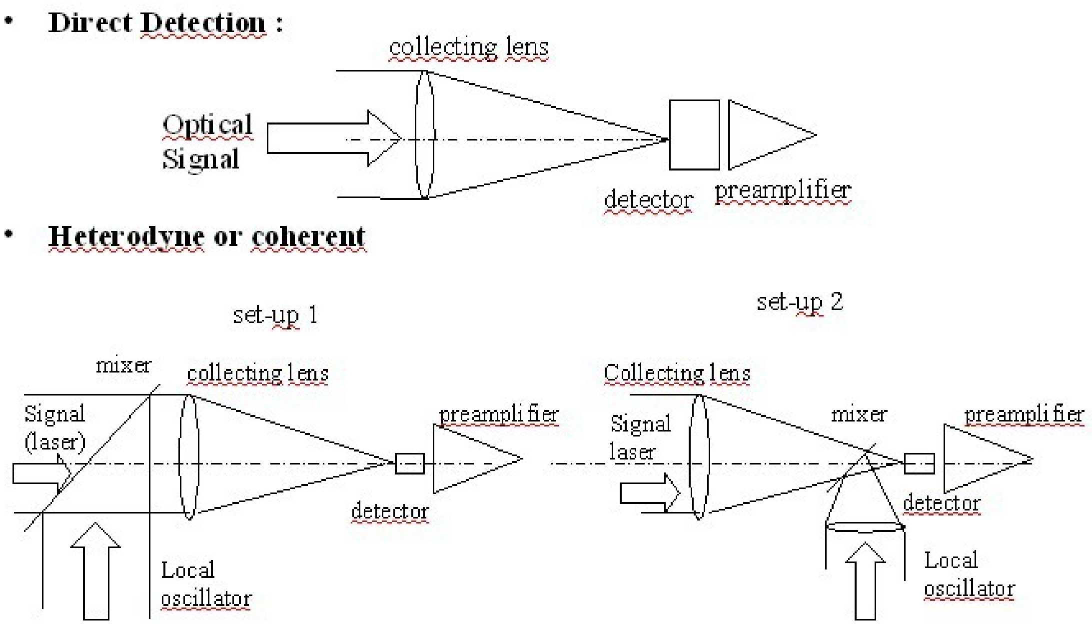
On the other hand, **flux collectors** are used in order to evaluate the level of a given radiation, to monitor it variations in time, but they are not supposed to deliver its shape, as image forming equipments do. Sensors such as detection systems, telecommunication systems, a large percentage of fiber optic sensors, metrologic equipments such as powermeters, ...

e) Classification by Mode of detection : direct, or heterodyne

In optics, there exists two different modes of detection : direct (or incoherent) and heterodyne (or coherent). They differ from each other in the following ways : they are adapted to different types of radiations, their set-ups are different as well as the characteristics of the output signal and of the associated signal processing techniques.

Coherent or heterodyne detection is the most widely used detection mode in the radar and radio domains ; it rests upon the mixing of the (monochromatic) radiation to be detected with a reference wave, generated by the sensor itself and called local oscillator. This mode of detection is not very much used in optics, because, as indicated by its name, it works only on specific radiations : they must be spatially and spectrally coherent with the local oscillator, i.e well aligned and frequency stabilized with respect to it. It means that the experimental set-up must be an interferometer, much more difficult to align and maintain so in optics than in radar, because of the scaling down in optical wavelengths with respect to radar or radio wavelengths.

For these reasons, the design of coherent electro-optical sensors is beyond the scope of this course, which is limited to direct detection sensors. The only reference to heterodyne sensors will be their theoretical set-up on figure 4.



2. Bases of optical radiometry

2.1. Bases of optical radiometry

Introduction

The design and the performance of an electro-optical sensor rest upon a correct evaluation of the electrical output of its detector, and hence it is right here, at the detector outset, that the designer must validate the operating principle of the sensor to be. That is why this part is dedicated to some bases of radiometry, which will be used in part C for computing the optical signal incident upon the sensitive area of the detector, because this optical signal is at the origin of the electrical signal, object of part D.

Définition : Basic radiometric quantities

Most electro-optical sensors respond to optical radiations of wavelengths typically comprised between 10^{-7} and $10^{-5}m$ (or photons of energies between 10^{-18} to $10^{-20}J$), , by converting them into electrical signals, by means of its detector. At each instant, the detector output is proportional to the radiation instantaneous rate, or **flux** Φ, that is incident upon its sensitive area. Consequently, the flux arriving at the detector is the primary parameter to be optimized by the designer of the sensor.

In optics, there exists two types of detectors : quantum and thermal. Quantum detectors are sensitive to the photon rate of arrival of the incoming radiation, and thermal detectors to its energy rate, or power. For that reason, one may express a flux either in terms of either the **photonic flux**Φ_p) (number of photons per second) or the **optical power** (or **radiant flux**, or **energetic flux**) Φ_e.

An ensemble of radiometric parameters have been defined in order to characterize optical radiations from the following points of view : geometry (size and position of sources, angular distribution), spectrum (radiant flux distribution with respect to wavelength), and variation with respect to time. These parameters are rapidly described below, starting with the ones applicable to non monochromatic radiations, i.e. spectrally extended.

In order to define a spectrally extended radiation at a given point, direction and wavelength, the recommended quantity is its **spectral radiance**$dL/d\lambda$, also called **radiance spectral density**.

Définition : densité spectrale de luminance

At the point of interest, spectral radiance is the density of flux that is being radiated towards the said direction, per unit area (centered on the point and normal to that direction), per unit solid angle, and per unit spectral bandwidth. Depending upon the fact that one may be interested either in radiant flux (W) or in photonic flux (s^{-1}), spectral radiance will be expressed in $Wm^{-2}\,sr^{-1}\mu m^{-1}$ or in $s^{-1}m^{-2}\,sr^{-1}\mu m^{-1}$.

It is not always possible to define the spectral radiance of a source, nor is it necessary to know it in all cases : for example, each time the area of a radiating source is not known, it is impossible to evaluate the radiance of that source. In many applications, it may be sufficient to know the angular diagram of the emitter : if so, one will specify its **spectral intensity** $dI/d\lambda$:

Définition : intensité spectrique

For each direction in space starting from the emitter, the flux density per unit solid angle and per unit wavelength which is radiated along that direction. Units are $Wm^{-2}\,sr^{-1}\mu m^{-1}$ for **spectral radiant intensity** and $s^{-1}m^{-2}\,sr^{-1}\mu m^{-1}$ for **spectral photonic intensity**.

To summarize, one will specify spectrally wide radiations by their spectral radiance whenever they are originating from an extended source (and if their emitting area is known), and by their spectral intensity whenever they come from a « **quasi-point source** », or from a source of unknown emitting area.

Définition : éclairement spectrique

If it is desired to characterize a spectrally wide radiation at a given planar surface (surface of an object or of a detector), without taking into account its angular properties (angular properties are included in radiance as well as in intensity), one will have to use another parameter: it is the **spectral irradiance** $dE/d\lambda$, of that plane, which is the spectral flux density per unit area, in $Wm^{-2}\mu m^{-1}$ or in $s^{-1}m^{-2}\mu m^{-1}$.

Now, if the incident radiation is **quasi-monochromatic**, i.e. if it occupies a very narrow spectral domain (typically less than a few percent of its central wavelength), it may be useless to specify its spectral properties : in many cases of quasi monochromatic sources (particularly in laser sensors where the spectral domain is the narrowest), one will specify the values of the previous parameters integrated inside the bandwidth of interest. The radiation will be then characterized either by its **radiant flux** or by its **photonic flux** (in W or s^{-1}), by its **radiance** or by its **photonic radiance** (in $Wm^{-2}sr^{-1}$ or $s^{-1}m^{-2}sr^{-1}$), by its **radiant intensity** or **photonic intensity** (Wsr^{-1} or $s^{-1}sr^{-1}$), or by its **irradiance** or **photonic irradiance** (Wm^{-2}, ou $s^{-1}m^{-2}$), at the wavelength of interest, without any more details about its exact spectral distribution.

In order to convert radiant quantities into photonic quantities, or vice versa, one will notice that, for a given wavelength, or inside a narrow spectral domain, any radiant quantity (either spectral or integrated) is the product of the corresponding photonic quantity by the photon energy at that wavelength (since it is the same for all photons). For example, radiant and photonic fluxes at a given wavelength λ are related to each other by the following equations :

$$\Phi_e(\lambda) = \Phi_p(\lambda)h\nu = \Phi_p(\lambda)h\frac{c}{\lambda}$$

and

$$\left(\frac{d\Phi_e}{d\lambda}\right) = \left(\frac{d\Phi_p}{d\lambda}\right)h\frac{c}{\lambda}$$

It is recalled that all the above mentionned quantities are instantaneous ones, i.e. they express the spatial, angular and spectral properties of a radiation at each instant. The rate of change of these parameters with respect to time must also be taken into consideration, because it impacts upon the design of the electronic part of the sensor (bandpass, noise) : that will be the subject of part D.

In some applications, such as the detection of pulsed signals (laser pulses) or in the case of CCD image forming systems, the detected radiation is integrated during some **integration time**. Then it may be simpler to specify these time-integrated quantities, if one does not need any information about instantaneous values (it may be neccessary to care about instantaneous values whenever there is some risk of component saturation or degradation by high level signals as may be true with lasers). In case integrated quantities are sufficient, integrated radiant flux will lead to **optical energy** (in J), and a photonic flux being integrated over some time will result in a **number of photons** ; time integrated irradiance is called **fluence** (in Jm^{-2} , or in number of photons per m^2).

2.2. Relationships between radiometric quantities

a) Introduction

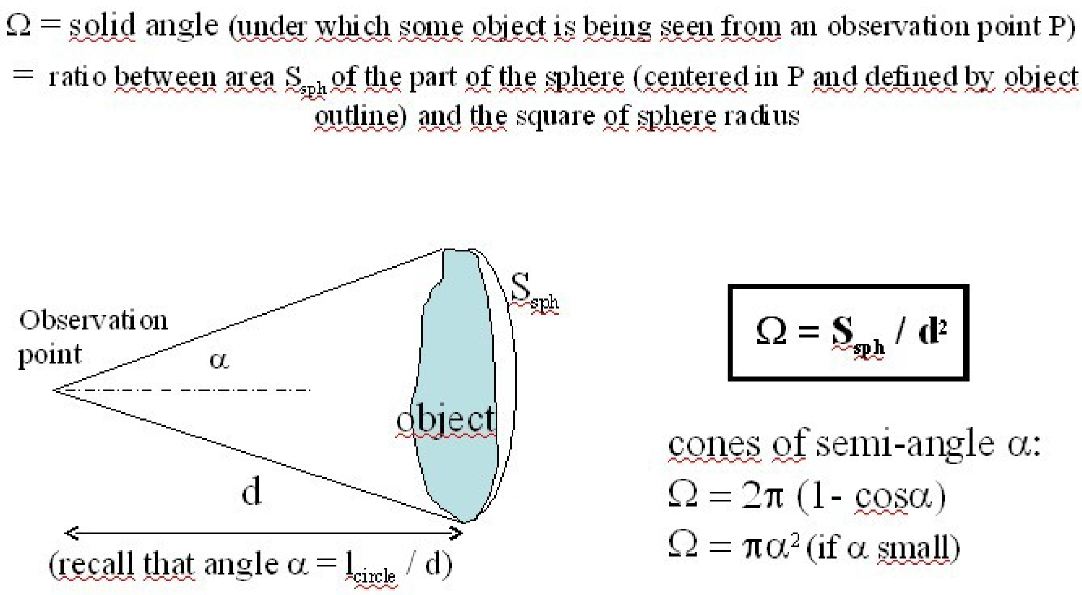
Some of the basic relationships are introduced below, that may help transfer from one radiometric quantity to another, in case the radiation is propagating through an homogeneous medium, i.e. some medium of constant refractive index. These relationships are applicable to spectral quantities in case of spectrally wide radiations, or to the spectrally integrated quantities if the radiation is monochromatic. They are valid wavelength by wavelength but they can also be applied to spectrally integrated quantities of non-monochromatic radiations, if the geometry of the radiated beam is independent of wavelength.

In order to keep mathematic expressions as simple as possible, no mention will be made of the wavelength, nor or of the case under consideration (spectral or spectrally integrated quantities). Furthermore, the propagating medium will be supposed to be perfect (vacuum), the influence of an material medium being treated in Part C.1.

b) Flux and intensity (quasi point sources)

As was mentionned in the previous paragraph, the radiometric quantity that is well fitted to describe radiation from a point source or a source of unknown area is its intensity I, starting from the definition of intensity, the flux Φ (Ω) radiated by a source of intensity I inside a solid angle Ω, (over which the intensity is supposed to be constant) is (figure 5) :

$\Phi(\Omega) = I\Omega$



If a point source is illuminating a planar surface of area $S_{receiver}$ at a distance d with an incidence angle θ' , the solid angle Ωunder which the surface is being seen from the source is given by :

$$\Omega = \frac{S_{apparent,receiver}}{d^2} = \frac{S_{receiver}\cos\theta'}{d^2}$$

and this surface receives the following flux from the source :

$$\Phi_{received} = I_{source}\frac{S_{app,rec}}{d^2} = I_{source}\frac{S_{rec}\cos\theta'}{d^2}$$

The resulting irradiance of the surface is proportional to the source intensity, to the cosine of the incidence angle (or obliquity factor) and inversely proportional to the square of the distance between the source and the receiving surface (« **Bouguer's law** ») :

$$E(d,\theta') = I_{source}\frac{\cos\theta'}{d^2}$$

At a given distance from the source, the plane of maximum irradiance is the one that is perpendicular to the incoming rays (« **normal incidence** »).

c) Flux and radiance (case of a pencil of light, or quasi collimated beam)

If the source cannot be considered as a point source, we have seen that its radiation at each point and towards each direction in space is best described by its radiance L Starting from the definition of radiance given above, one will write that the flux being radiated by the source through a small diaphragm of area S inside a small solid angle Ω centered along an axis normal to that diaphragm is given by :

$\Phi = LS\Omega$

The above configuration, geometrically defined by a (small) diaphragm and a small solid angle directed along its normal, corresponds to an elementary propagation channel, called « pencil of light ». If the diaphragm (or the surface of the emitter in case the emitter is acting as the diaphragm), is not normal to the pencil axis, let θ be the angle between the pencil axis and its perpendicular : then, the projected, or apparent area of the diaphragm along the axix of the pencil is $S_{app}(\theta) = S\cos\theta$ and the flux being radiated along that pencil of light is :

$\Phi = LS\Omega\cos\theta$

A pencil of light, or an elementary propagation channel defined by a diaphragm (of area S) and some small solid angle along an axis at angle θ with respect to the diaphragm normal, is specified by its **geometrical extent**$G = S\Omega\cos\theta$ in $m^2 sr$.

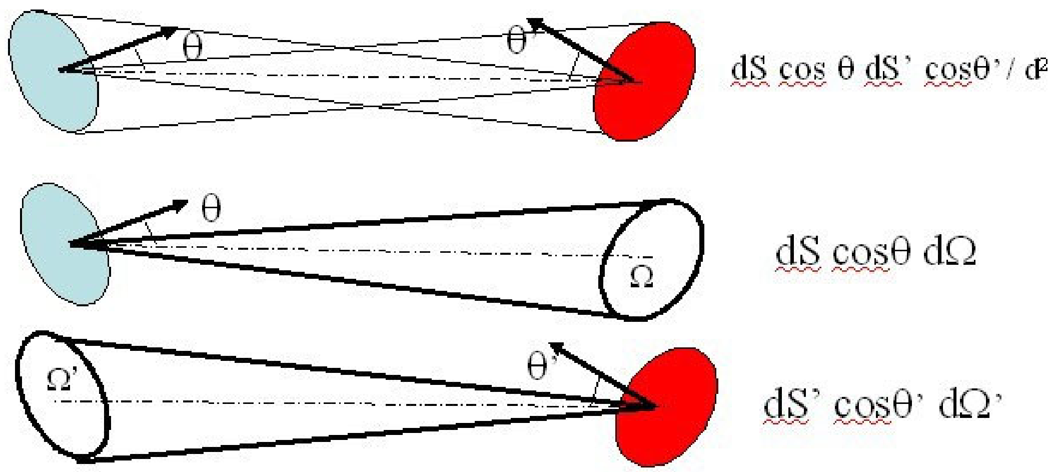
Remarque

Hence, a pencil of light propagates an amount of flux that is the product of its geometrical extent by its radiance.

One may also construct a pencil of light by way of two small diaphragms, of areas respectively equal to S and S', separated from each other by a distance d (such that S and $S' << d^2$) and with their normals at angles θ et θ' with the axis joigning the centers of the diaphragms. In any case, the geometrical extent of a pencil may be expressed by either one of the three following formulas, which are equivalent (figure 6) :

$$G_{pencil} = S\Omega\cos\theta = S'\Omega'\cos\theta' = \frac{S\cos\theta\, S'\cos\theta'}{d^2}$$

where Ω' is the solid angle under which the second diaphragm observes the first one. For this expression to apply, each point of Smust illuminate the whole area of S' .



One can deduce from above that the irradiance of an elementary area being illuminated under an incidence angle θ' by a pencil of light of radiance L inside a (small) solid angle Ω' is the following :

$$E(\theta', \Omega') = \frac{\Phi}{S'} = L\Omega' \cos\theta'$$

d) Flux and radiance (radiation propagating along an extended beam of constant radiance)

What happens when one start increasing the divergence of a pencil of light, without modifying the emitting area nor its radiance? The pencil of light, which is narrow to start with, is said to be transformed into a finite or extended **beam of light**.

The flux being radiated through an extended beam of light is obtained by adding all the fluxes that propagate along each of the elementary pencils, or channels, making up the finite beam. The geometrical extent of the beam is also obtained by adding those of its elementary pencils. Let us consider a beam of light as being made up of elementary pencils exiting from the same diaphragm with all with the same divergence ; one can easily realize that, even though they have the same initial diaphragm and the same divergence, all these pencils don't have the same geometrical extent : indeed, the geometrical extent of a given pencil is decreasing as the pencil get further and further away from the normal to the emitting surface, because the apparent area of the emitter is $S\cos\theta$. One may conclude from this elementary observation that the geometrical extent of an extended beam is no more equal to the product of its emitting area by its solid angle, as is the case for its central pencil, centered along its axis.

If the radiance is constant all over the geometry of the beam, the flux that propagates inside the beam is given by the following expression, where G, the geometrical extent of the beam, must be specifically computed for the configuration at hand :

$$\Phi_{beam} = \sum L(\theta) G_{pencil} = L_{beam} \sum G_{pencil} = L_{beam} G_{beam}$$

Values of geometrical extents for beams of finite size and divergence have been computed for most of simple shapes : below, the result is given concerning a beam made up of a plane emitter that radiates light inside a circular cone of semi angle α_M centered along its normal (figure 7) :

$$G_{em} = \pi S_{source} \sin^2 \alpha_M$$

If a planar source is radiating over a half-space, i.e. over $2\pi sr$, the geometrical extent of its beam is:

$$G_{em,half-space} = \pi S_{source}$$

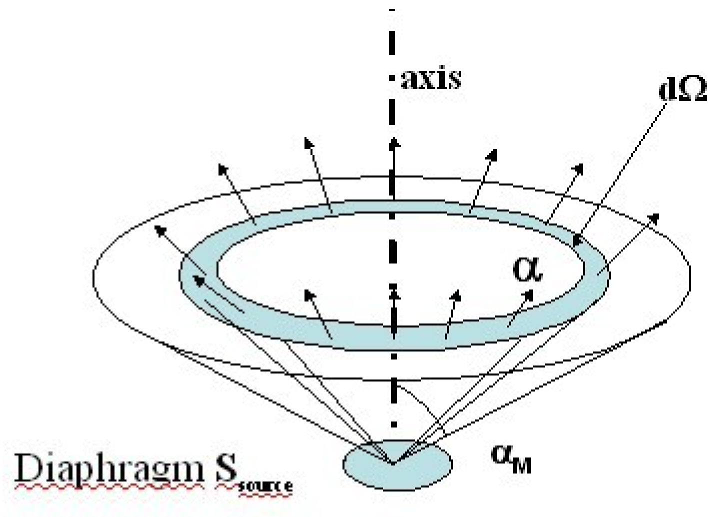
In a similar manner, the geometrical extent of a beam that is being focused down onto a planar surface (such as that of a detector) of area S_{det}, inside a cone of half angle α'_M centered along the axis of the surface, is the following :

$$G_{rec} = \pi S_{rec} \sin^2 \alpha'_M$$

If the incident beam is hemispheric $(\Omega' = 2\pi sr)$, its geometrical extent is then : $G_{rec} = \pi S_{rec}$

$$G = S_{source} \int_0^{\alpha_M} \cos\alpha \, d\Omega \quad \text{with } d\Omega = 2\,\pi \sin\alpha \, d\alpha$$

$$= \pi \, S_{source} \sin^2 \alpha_M$$



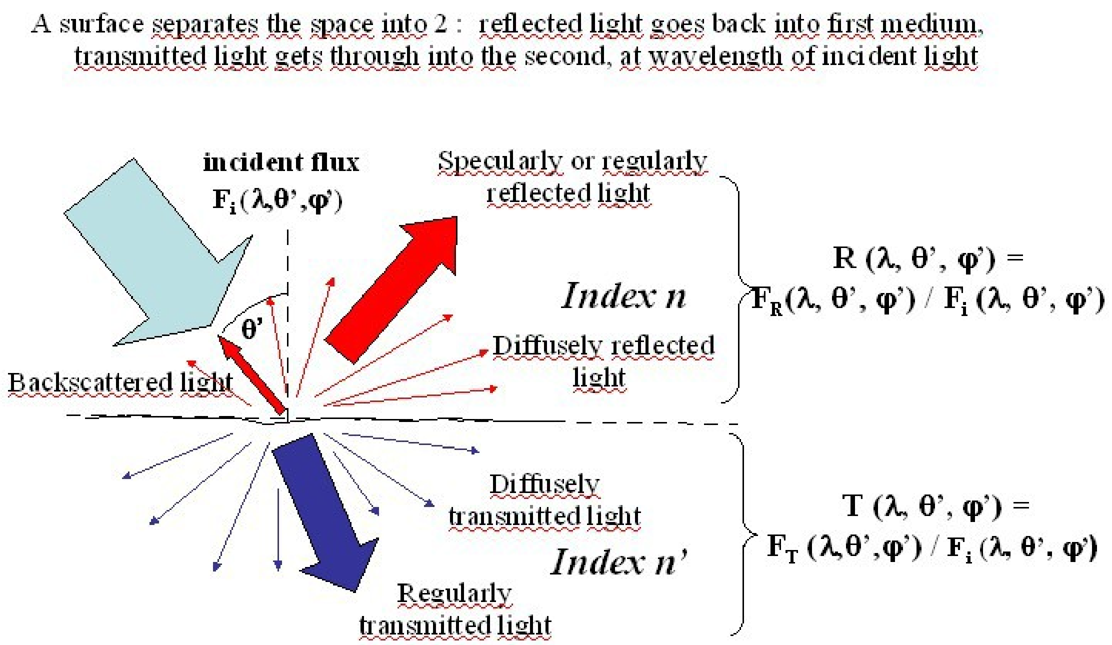
If the radiance is not constant inside the beam, one can no longer write $\Phi_{beam} = G_{beam} L_{beam}$. Then, there is not so much interest in defining the geometrical extent of the beam because this parameter does not intervene as such into the evaluation of the flux being radiated along that beam.

2.3. Some usual examples of radiations

a) Reflection of light from an interface

When the interface between two media is being illuminated, interaction between the photons and matter gives rise to three phenomena : reflection, transmission and absorption (figure 8), that depend essentially upon the following parameters of incoming light : angle of incidence, wavelength and polarization. Considering the input light is incident from a direction (\theta ',φ') and at a wavelength λ, the interface is specified by the following **spectral** and **directional** properties :

- **reflectance**, $R(\lambda, \theta', \varphi')$, or the percentage of incident light being sent back into the initial medium (at the same wavelength),
- **transmittance**, $T(\lambda, \theta', \varphi')$,or the percentage of incident light going through the interface at the same wavelength,
- **absorptance**, $A(\lambda, \theta', \varphi')$, or percentage of incident light being « digested » by the interface and then, as will be seen further down, converted into a spectrally wide radiation, called thermal radiation that depends upon the temperature of the interface.



In most applications, however, it is necessary to know the angular distribution of the light that is being reflected by the surface of an object when it is illuminated, for instance by the sensor itself. That kind of information is not available from the spectral and directional reflectance of the surface, defined above. To be convinved of that it is easy to realize that a wall and a polished mirror may have identical spectral reflectance, but will behave quite differently under some collimated and monochromatic illumination, because the first one (wall) is a diffuse surface and the second (mirror) a specular one.

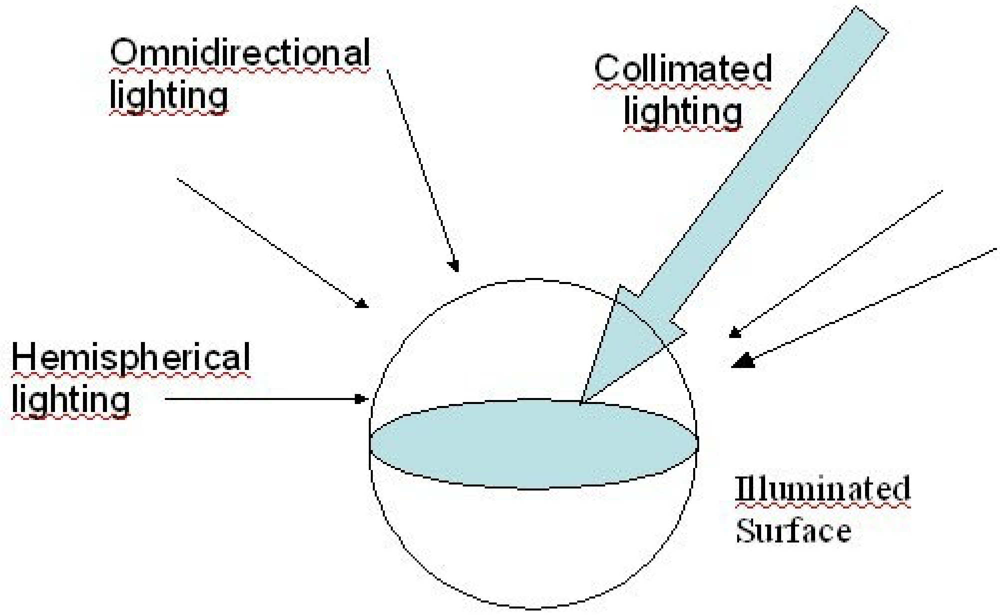
The adequate parameter to use in order to specify the angular distribution of the reflected light from an interface under any kind of illumination is its spectral « Bidirectional Reflectance Distribution Function » or BRDF $(\lambda, \theta, \varphi, \theta', \varphi')$: BRDF (or BTDF for transmitted light) is the ratio between the radiance $L_R(\lambda, \theta, \varphi)$ of the reflected (or transmitted) radiation towards a direction (θ, φ) and its irradiance $E(\lambda, \theta', \varphi')$, when the surface is being illuminated at wavelength λ, and from a direction θ', φ' :

$$BRDF(\lambda, \theta, \varphi, \theta', \varphi') = \frac{L_R(\lambda, \theta, \varphi)}{E(\lambda, \theta', \varphi')}$$

The angular diagram of the light that is reflected from a surface under collimated illumination varies from a surface to another. For that reason, surfaces are classified into different families, the two extremes being diffuse and specular (similar to mirrors) : a perfectly diffuse surface (also called **lambertian surface**) is such that the radiance of reflected light is constant : for a given direction of illumination, its BRDF is constant and equal to $1/\pi$ times the value of its reflectance for that incidence angle.

The surfaces of most natural objects are diffuse, except for example those of calm waters (ponds, lakes). Optical diffuseness of a surface is strongly tied to its **optical roughness**, i.e. the fluctuations of its relief with respect to the incident wavelength. The angular behavior of light reflected by a diffuse surface usually depends upon the angle of incidence and wavelength: Most surfaces tend to become less and less diffuse (hence, more and more specular) with increasing wavelength and incidence angle, i.e. when one goes from the UV to visible, IR and radar), and from normal to grazing incidence. The behavior of specular surfaces will be examined later on along with that of optical surfaces.

If one must design an electro-optical sensor that is intended to observe objects under some kind of (natural or artifical) illumination, one must know the exact conditions of illumination (figure 9), since the surfaces of objects react differently with respect to wavelength and to incidence angle. Hence, for a given value of irradiance, a given surface will appear quite differently if it is illuminated from a single direction (collimated lighting, for example by means of a laser, a flashlight, or the sun), from a set of different directions, or from all the directions inside the half-sphere seen by the surface (which is the case, in the visible, for a horizontal surface under blue sky).



b) Thermal radiation

At temperatures different from 0, any kind of body emits thermal radiation, due to the agitation of its electrons. Kirchhoff's law, below, shows that the spectral radiance of thermal radiation from a surface or a body X at temperature T, towards some direction (θ, φ) in space depends upon its spectral absorptance (introduced in the previous paragraph) for that direction and on its temperature :

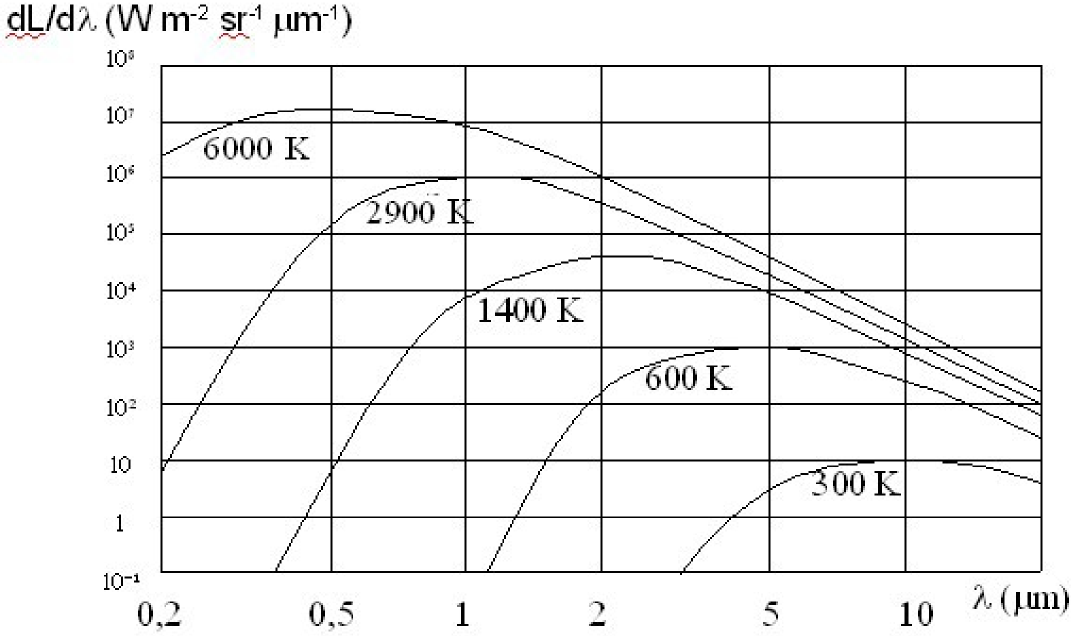
$$\left(\frac{dL}{d\lambda}\right)_{X,T}(\theta,\varphi) = A(\lambda,\theta,\varphi)\left(\frac{dL}{d\lambda}\right)_{BB,T}$$

Since thermal radiation of a source is proportional to its spectral absorptance, this parameter has been in the past, and still is, called **spectral directional emissivity**$\varepsilon(\lambda,\theta,\varphi)$. of the material. The second term which appears on the right hand side of the above equation is the spectral radiance that the material $\lambda$and direction of incidence (θ, φ). The perfect absorber, or **blackbody**, has been the subject of numerous studies and played a fundamental role in the discovery of quantum mechanics. It can be shown that blackbody radiation is lambertian and unpolarized : its spectral radiance is uniform, and its value given (figure 10) by the following relationship (**Planck's law**) :

$$\left(\frac{dL}{d\lambda}\right)_{CN,T} = \frac{2hc^2\lambda^{-5}}{e^{\frac{hc}{\lambda k_B T}} - 1}$$

where h is Planck's constant ($h = 6,62 \times 10^{-34} Js$),
c the speed of light in vacuum ($c \approx 3 \times 10^8 ms^{-1}$)

and k_B Boltzmann's constant ($k_B = 1,38 \times 10^{-23} JK^{-1}$)



c) Global radiation of an object, under omnidirectional illumination

In general, the total radiation from an object originates from reflection of ambient light and from thermal radiation. If some surface element, of temperature T, is receiving some spectral irradiance $dE/d\lambda$ from direction (θ', φ'), its spectral radiance toward any direction (θ, φ) is :

$$\frac{dL}{d\lambda}(\theta,\varphi)_{x,r} = A(\lambda,\theta,\varphi)\Big(\frac{dL}{d\lambda}\Big)_{CN,T} + BRDF(\lambda,\theta,\varphi,\theta',\varphi')\frac{dE}{d\lambda}(\theta',\varphi')$$

If the surface is diffuse, and the illumination hemispherical (that is the case for numerous natural scenes), any source of light that is situated in the half-sphere in front of it will contribute to its radiance towards any direction inside a half-space, by means of the BRDF of the surface. On the contrary, if the surface is strictly specular, each direction of illumination will contribute by reflection only towards its symetric with respect to the normal, as ruled by the law of reflection, other directions of illumination being reflected elsewhere.

In case of natural scenes at ambient temperatures, the main origin of radiation strongly depends upon the spectral bandwidth of the observing sensor :

- In the visible and in the very near infrared, reflection of ambient light is the main contribution, because in that spectral domain, the spectral irradiance of the scene, due to the sun, is quite important and because reflectances of most objects are quite high, while BB thermal radiation at room temperature is infinitesimal.
- In the long wave infrared band ($8 - 12\mu m$), thermal spectral radiance is at its maximum, while spectral irradiance from the sun is rather weak, as well as the reflectance of a vast majority of objects (most objects behave almost like blackbodies in that spectral band, except polished metals).
- In the midwave infrared band ($3 - 5\mu m$), radiation from natural objects originates from both thermal emission and from reflection of ambient light during daytime and from thermal emission at night.

3. The signal in an electro-optical sensor

3.1. Propagation of radiations

a) Scattering and absorption

In vacuum, radiations are not attenuated at all : their basic properties, such as spectral radiance or intensity, keep constant all along the path of propagation. By the contrary, in matter (atmosphere, water, optical fiber,...), the spectral, spatial and time properties of radiations degrade along the propagation, because of two main interactions with the medium : **scattering** and **absorption**.

Scattering originates from collisions between photons and constituents of the medium : large particles for **Mie scattering**, small particles and molecules for **Rayleigh scattering**; because of these collisions, a fraction of the incident photons is being directed in space towards directions other than the initial one ; consequently, the number of photons that go on propagating along the original path is reduced. Absorption originates from the fact that inside matter, a fraction of the energy being radiated is transferred into the medium itself at some specific frequencies, corresponding to vibration, rotation natural frequencies of electrons, atoms and molecules of the medium.

There are two parameters that are well suited to characterize the influence of a given medium upon the propagation of light at each point along the light path (figure 11) : these are its **spectral, linear scattering and absorption coefficients**, $\beta(\lambda, z)$and $\alpha(\lambda, z)$, and they represent the percentages of light being respectively scattered and absorbed per unit length of the propagating medium, at the wavelength of interest.

If the refractive index of the medium is not constant, i.e. if it varies from one point to another, and/or with respect to wavelength, the velocity of light is being modified and so are the time and space characteristics of a light beam propagating in such a medium.

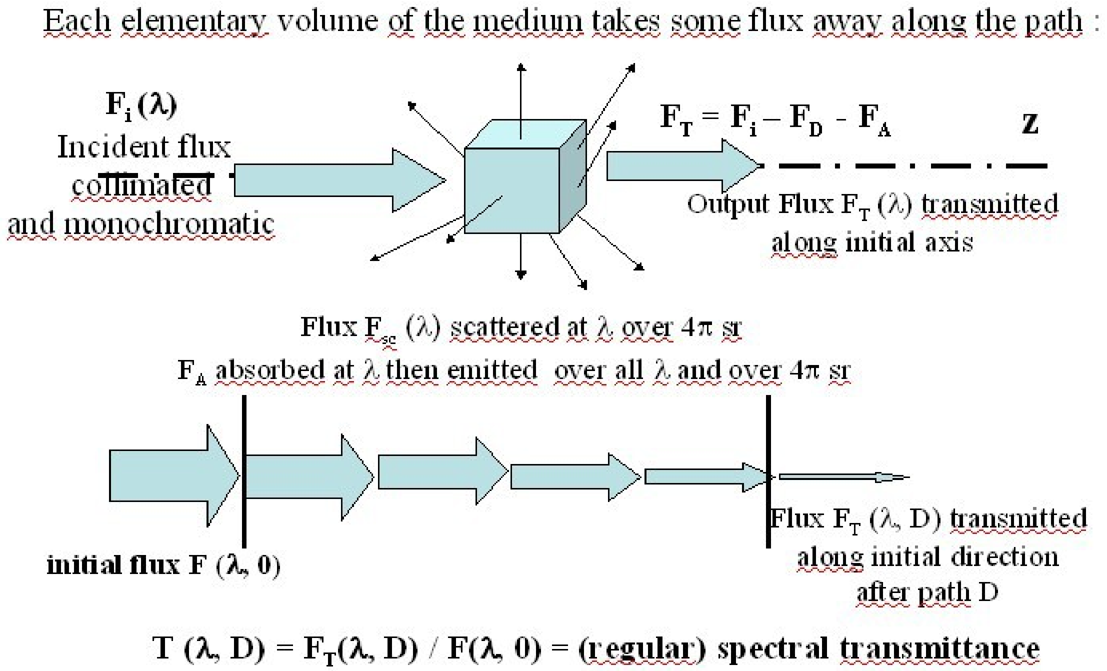
Exemple

For example, in a dispersive medium, refractive index changes with respect to wavelength, and hence the speed of light inside the medium depends upon wavelength, so that short signals (laser pulses, or modulated signals at very high frequencies) are getting distorted along the way.

On the other hand, chaotic fluctuations of refractive indices make light rays depart from the rules of geometrical optics, which modifies the geometric shapes of objects as seen through « **turbulent media** » (well known example of mirages, created by atmospheric turbulence). These phenomena are outside the scope of this course and are only mentionned here for the sake of information, without more detail.

b) Spectral transmittance of a propagating medium

The percentage of the flux that propagates along its original direction at some wavelength λ (figure 11), after traversing a distance D inside a given medium (figure 11) is the **spectral transmittance** of the medium at that wavelength and over that distance.



A medium is said to be « inhomogeneous », if its absorption and scattering coefficients, α and β vary from point to point. One must then write :

$$T_m(\lambda, D) = e^{-\int_0^D [\alpha(\lambda,z)+\beta(\lambda,z)]dz}$$

If the medium is homogeneous, that means that $\alpha$and β sare constant along the path at each wavelength, and the spectral transmittance of the medium over some range D is equal to :

$$T_m(\lambda, D) = e^{-[\alpha(\lambda)+\beta(\lambda)D]}$$

The global attenuation of a propagating medium results from both effects, scattering and absorption : it is specified by the spectral attenuation (or extinction) coefficient $\gamma(\lambda, z)$ of the medium, as :

$$\gamma(\lambda, z) = \alpha(\lambda, z) + \beta(\lambda, z)$$

The spectral transmittance of the medium over some given path length D is therefore :

$$T_m(\lambda, D) = e^{-\int_0^D \gamma(\lambda,z)dz} \text{ (for inhomogeneous media)}$$

$$T_m(\lambda , D) = e^{-\gamma(\lambda)D} \text{ (for homogeneous media)}$$

c) Initial and apparent radiances

The first negative effect of scattering and absorption by a medium upon the performance of an electro-optical sensor is the attenuation of radiations as they propagate along. The second one is that, besides attenuating the radiations of interest for the sensor, a scattering and absorbing medium radiates **stray light** (or undesirable light), either by thermal emission (which is the case of the atmosphere in the infrared) or by scattering of ambient light (that is the origin of the « blue sky » in the visible), or both.

Let us consider, for instance, the case of an image forming sensor observing an extended scenery over a large spectrum : the apparent spectral radiance of an object, i.e. the spectral radiance at the sensor, is the sum of its initial radiance, attenuated by the spectral transmittance $T_m(\lambda, D)$ of the medium over the distance of observation, and of the stray light radiance $L_{app,m}(\lambda, D)$ of the medium over the observation distance:

$$\frac{dL_{app,obj}}{d\lambda}(D) = T_m(\lambda, D)\frac{dL_{ini,obj}}{d\lambda} + \frac{dL_{app,m}}{d\lambda}(D)$$

3.2. Radiometric properties of an optical instrument

a) Introduction

In most electro-optical devices, the optical sensor does not consist in a detector alone, but is made up of some (collecting) optics associated with a detector : this association makes possible to select a particular zone in space and to eliminate the rest, whereas the use of a detector alone does not allow to discriminate the incident the rays with respect to their incidence angle inside the entire half-space in front of the detector. The association of a lens with a detector also allows to optimize the flux collecting capability of the sensor because the area of a lens may be made much larger than that of a detector.

b) Field and aperture stops of an electro-optical sensor

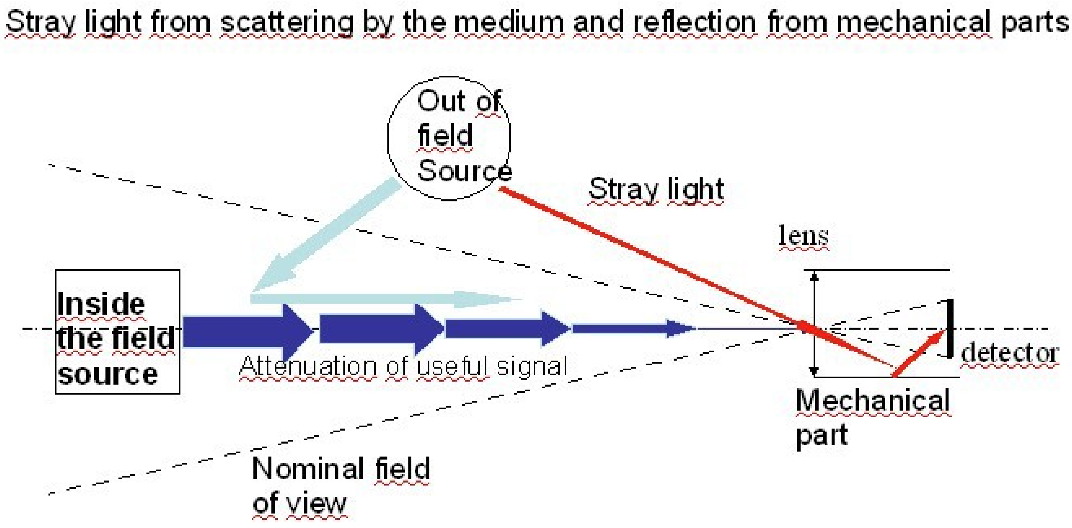
The association « lens + detector » operates a double selection among incident rays. Firstly, it is clear that the photons from the source that do not enter the lens will never reach the detector. The optical diaphragm that is responsible for this first selection is called the **aperture stop**. It is conventionally defined for the beam that propagates along the axis of the sensor (on axis point source). The **entrance pupil**, that defines the size of the beam entering the lens, is the image of the aperture stop in the object space of the optics. The **exit pupil**, or image of the aperture stop in the image space, defines the size of the beam at the outset of the lens.

In most applications, the second selection among the rays is done by the detector itself : among all the photons having gone through the lens, only a fraction succeeds in reaching the sensitive area of the detector : these are the ones originating from the part of the object space which is called the **object field of view of the sensor**. It is why, in this configuration, the sensitive area of the detector is called the **field stop of the sensor**.

If the object, or the source, is at a finite distance from the sensor, the field of view is usually specified or measured in terms of linear dimensions at the object plane. For example, the field of view of a microscope will be said to be $1mm$ in diameter if that is the size of the area being observed. If the object (or the source) is far away, the field of view will be preferably specified, or measured, in terms of angular dimensions, usually along two axes, horizontal and vertical ; if that is the case, one will say, for instance, that the field of view of a camera is $9°$ by $16°$. The half angle θ of the field of view along either of these two axes is given by the following relationship, where a is half the detector size along the corresponding axis, and f' the focal length of the lens:

$$\tan\theta = \frac{a}{2f'}$$

In some cases, there may happen that the detector also collects some light from regions that are theoretically outside the field of view of the sensor, because of scattering by particles along the path, or from reflections by mechanical structures or by optical surfaces inside the sensor. The resulting light must be taken into account when evaluating the incident flux upon the detector (Figure 12), not as part of the useful signal, but as stray light.



c) Geometrical extent of an electro-optical sensor

The shape of a beam, which is defined by the field and aperture stops of the sensor, leads to a basic quantity called the geometrical extent of the sensor. In order to seize what it means, let us consider an E-O sensor that is observing a far away, on-axis, source inside a small solid angle : for this purpose, it is made up of a small detector, of sensitive area S_{det}, at the back focus of a lens, of focal length f. In the object space, the geometrical extent of the sensor may be defined, as that of the pencil of light propagating along the sensor axis inside the solid angle $\Omega_{det} = S_d / f^2$ and passing through the entrance pupil (of area S_{opt}) : in the image space, the geometrical extent is that of the beam passing through the exit pupil and converging onto the detector, so that :

$$G_{sensor} = S_{opt}\Omega_{det} = \frac{S_{opt}S_{det}}{f^2}$$

or, since

$$S_{opt} = \pi\frac{D_e^2}{4}$$

$$G_{sensor} = \pi\frac{S_{det}}{4N^2}$$

where N is the aperture number of the lens ($N = f'/D_e$)

d) Transfer of Radiance and of geometrical extent by an optical instrument

By applying geometrical optics rules, one shows that the input and output geometrical extents, G_1 G_2 of a pencil of light traversing an optical component without being limited by it, are related to each other by :

$$n_1^2 G_1 = n_2^2 G_2$$

where n_1 and n_2 are the refractive indices of the initial and final media.

Whenever the initial and final media of the lens are identical (which happens in a vast majority of EO sensors, the common medium being air), the input and output pencils of light have got the same geometrical extent. One notable exception is underwater imaging systems, where the object medium is water and the image medium is air. Furthermore, if the lens is lossless, input and output fluxes are identical, which means that a perfect optical system that does not limit the geometrical extent of an entering beam delivers an output beam of same extent and radiance as the input beam if the initial and final media are the same. This is called the **radiance conservation theorem by optical systems** (beware : this theorem holds only for perfect systems, without loss and having identical input and output media). If the lens is not perfectly transparent, then the ouput radiance is the product of the input radiance by the transmittance of the lens.

This theorem applies for example to simple components such as mirrors (or specular surfaces) for which, by definition, the final medium is the same as the initial one: the radiance of the reflected beam along the direction of geometrical optics ($\theta = -\theta'$, and $\varphi = \varphi'$) is the product of the input radiance $L_{inc}(\lambda, \theta', \varphi')$ by the reflectance of the mirror :

$$L_R(\lambda, -\theta', \varphi') = R(\lambda, \theta', \varphi') L_{inc}(\lambda, \theta', \varphi')$$

The BRDF (defined in part 3.a) of a mirror or, more generally, of a specular surface, is a Dirac distribution : it is naught in all directions of observation, except along the direction that correspond to geometrical optics, i.e. that is symetrical of the direction of illumination with respect to the normal.

3.3. Evaluation of the flux incident upon the detector of an electro-optical sensor

a) Introduction

As was mentionned above, one of the major tasks in electro-optical system design is evaluating the detector output, and before that, evaluating the flux that is incident on the detector. Computation procedures vary from one application to another, depending upon the goal of the sensor and its configuration (free space propagation, fiber optic sensor, extended or point source, image forming sensor or flux collector,...). The scope of this course is limited to two configurations in free space propagation : image forming sensors and flux collectors.

b) First configuration : image forming sensors operating on extended objects

Image forming sensors are such that they spatially or angularly resolve the light source they are looking at. They are generally comprising a lens and an array of elementary detectors (picture elements or pixels) in the image plane. The field of view of each one of these is much smaller than the object. The number of elementary detectors in the array may vary from a few (coarse image) to several millions (megapixel imagery). The geometrical extent G_{sensor} of the beam incident on each element is then not defined by the whole object, but by the association of the lens with each one of the elemental detectors. Following the relationship of § 2.2, the spectral flux incident upon the sensitive area of each pixel is the product of the geometrical extent of the beam along the direction seen by the pixel by the apparent spectral radiance of the scene along that direction, so that :

$$\frac{d\Phi_{inc,det}}{d\lambda} = G_{sensor} T_{op}(\lambda) \frac{dL_{app}}{d\lambda}$$

and hence, if the object is at infinity :

$$\frac{d\Phi_{inc,det}}{d\lambda} = \frac{\pi T_{op}(\lambda) A_d}{4N^2} \times \frac{dL_{app}}{d\lambda}$$

where N and $T_{op}(\lambda)$ are respectively the aperture number (ratio between the back focal length and the diameter of the entrance pupil) of the lens (in imaging applications, the lens is supposed to be « aplanetic », i.e. devoid of major aberrations) and its spectral transmittance. Ad is the sensitive area of one elemental detector.

From this point, one will compute the electrical output of the detector by means of the procedure which will be described in the next part.

c) Second configuration : flux collectors operating on small sources (quasi point sources)

The main difference between an image forming sensor and a flux collector is the fact that flux collectors do not resolve the source, while image forming sensors do : the field of view of a flux collector is larger than the source, so that it collects light from the entire source. In these conditions, the most adequate parameter for specifying the source is its intensity : the flux that is incident upon the detector is the product of the apparent intensity of the source, I_{app} (evaluated at the sensor entrance pupil) by the transmittance of the lens and by the solid angle under which the entrance pupil is being seen from the source :

$$\frac{d\Phi_{inc,det}}{d\lambda} = T_{op}(\lambda)\frac{dL_{app,source}}{d\lambda}\frac{S_{opt}}{d^2}$$

Since the field of view of a flux collector is larger than the source, that means that the sensor is also observing some part of the surrounding background. The background being an extended object (its solid angle is $(4\pi sr)$ much larger than the field of view of the sensor), a flux collector behaves on the background like an image forming sensor : hence, the background must be specified by its apparent spectral radiance $L_{app,background}$, inside the field of view of the sensor. If the source is very tiny as compared to the sensor field of view, and if the object is far away (« at infinity »), the total spectral flux incident upon the detector is the following :

$$\frac{d\Phi_{inc,det)}}{d\lambda} = T_{op}(\lambda)\left[T_m(\lambda)\left(\frac{dI_0}{d\lambda}\right)\frac{S_{op}}{d^2}\left(\frac{dL_{app,background}}{d\lambda}\right)\frac{\pi A_d}{4N^2}\right]$$

Let us mention a third configuration, which occurs quite often in astronomy, and in which the sensor is an image forming sensor that observes point sources (stars) with an array of elemental detectors. Even though the objects are point sources and their images tiny (the optical instrument is diffraction limited), the flux from a star that enters the lens may not be collected entirely by only one detector, but spread over several of them.

In that case, one must evaluate the percentage of flux from a star which, having gone through the lens, is incident upon each of the corresponding detectors. The specification of the lens answering that problem is its **Point Spread Function** (PSF), which is the relative irradiance distribution in the image plane whenever the source is a point. As for the flux contribution from the background, it remains unchanged with respect to the previous result.

4. Signal to noise ratio and performance of an electro-optical sensor

4.1. Output signal from the detector

a) Total output signal

Two types of detectors may be used in electro-optical sensors : thermal and quantum. The first ones convert incident energy in temperature rise, and the second ones directly into electric charges. Whatever its type, the detector will generally be specified by its spectral sensitivity$R_i(\lambda)$ in AW^{-1} , which is the ratio between its output current and the input flux, with respect to wavelength. In case a quantum detector is being used, one may also specify its spectral quantum efficiency $\eta(\lambda)$, or ratio between the number of output electrons per input photons, with respect to wavelength. Spectral quantum efficiency $\eta(\lambda)$ and sensitivity $R_i(\lambda)$ are related by :

$$R_i(\lambda) = \frac{\eta(\lambda)q}{h\nu} = \frac{\eta(\lambda)q\lambda}{hc}$$

If the incident radiation is spectrally wide, every narrow spectral band, of elemental width $d\lambda$ centered at wavelength λ produces the following elemental output from the detector :

$di_{det} = \eta(\lambda) q d\Phi_p(\lambda) = R_i(\lambda) d\Phi_e(\lambda)$

where $d\Phi_p$ and $d\Phi_e$ are the expressions, in photonic (s^{-1}) and radiant (W),units, of the incident flux inside that band, q the charge of an electron, and h Planck's constant.

As $d\Phi_e = \left(\frac{d\Phi_e}{d\lambda}\right) d\lambda$ and $d\Phi_p = \left(\frac{d\Phi_p}{d\lambda}\right) d\lambda$,there results that, in response to an incoming spectrally wide radiation, the output current from the detector is :

$$i_{det} = \int_0^\infty R_i(\lambda)\left(\frac{d\Phi_e}{d\lambda}\right) d\lambda = \int_0^\infty \eta(\lambda) q \left(\frac{d\Phi_p}{d\lambda}\right) d\lambda$$

If the incident radiation is monochromatic, the output current of the detector is more simply written as :

$i_{det} = R_i(\lambda)\Phi_e(\lambda)$ or $i_{det} = \eta(\lambda) q \Phi_p(\lambda)$

b) The « useful signal » in an electro-optical sensor

Defining the pertinent output signal from the detector of an EO sensor is not always an easy task : in some cases, the detector may be permanently receiving a non negligible flux from the bakground, and hence delivering some permanent signal, even in the absence of the pertinent source. In the visible and in the very near infrared, this background current may be brought down to a minimum level, or **dark current**, simply by putting the sensor in total darkness, because thermal radiation at room temperature is negligible in that spectral band ; by doing so, one can say that the detector output is not far from being representative of the signal to detect. If the same sensor is being operated at high illumination levels (in presence of the sun or of some artificial lamps), then the detector may be exposed to stray light, but these perturbations may be easily be taken into account.

In the infrared, the situation is quite different, because thermal radiation at room temperature is present even if the sensor is operating in complete darkness. This permanent radiation gives rise to a non negligible signal and it is present in temperatures measurements applications such as thermography or thermal imaging. Although applications look very similar at first sight, their « useful or pertinent signals » are defined differently : on one hand, the useful signal in thermography is the value of the detector output, because it is the one that carries information about the temperature of an object, after calibration on standard sources such as blackbodies. In thermal imaging applications, the signal that matters is is not so much the previous output, as its changes from point to point on the object.

At the design stage of an EO sensor, it is very important for the designer to define the nature of the pertinent output signal. What applies to thermal imagery may be applied to a large number of domains, where the interesting information is not the total value of the detector output, but its variation in time or in space : this is true of optical telecommunications (difference between bits 0 and 1), imaging systems (differences in reflectance or in temperature), target detection (target presence or absence), laser pulse detection (rangefinding), surveillance (intrusion detection),...

4.2. Sources of noise in an electro-optical sensor

Shot noise

First of all, the detector of an EO sensor is a noisy electrical source because the photon rate of arrival upon its sensitive area is not constant in time. This fluctuation induces a fluctuation in the electron delivery rate, or **shot noise** in the output. The spectral density of the corresponding variance is given by the **Shottky formula**, below, wher i represents the average output current from the detector :

$$\left(\frac{d\sigma_i^2}{df}\right)_{shot} = 2qi$$

Thermal noise from the load resistor

Besides that, because of the thermal agitation of its electrons, the load resistor of the detector is another source of noise, or thermal noise. Its spectral density with respect to temperature T is given by Johnson formula :

$$\left(\frac{d\sigma_i^2}{df}\right)_{johnson} = \frac{4k_B T}{R_c}$$

in which k_B is Boltzmann's constant, and R_L the detector load resistor

Noise from the signal processor

There are other sources of noise, from the electronics, which must also be taken into account : a large part of it may come from the preamplifier and, generally to a lesser degree, from the next stages (reset noise, digitization noise). In order to appreciate properly their influences, all these origins of noise are by convention evaluated at the same point as the signal itself, i.e. at the immediate outset of the detector. Each electronic noise (at pre-amplifier, amplifier, filters, reset, digitization stages) is brought back to that point by its noise figure F, with respect to the first electronic stage and its spectral density evaluated there, such as th preamplifier noise, below :

$$\left(\frac{d\sigma_i^2}{df}\right)_{PA} = \frac{4k_B T(F-1)}{R_c}$$

Total noise

All noise sources are uncorrelated so that the variance of the total noise current is the sum of the variances of the contributing noise currents :

$$\sigma_i^2 = \sigma_{ishot}^2 + \sigma_{i\,johnson}^2 + \sigma_{iPA}^2$$

hence the following expression for the global variance of the output current at the outset of the detector as defined inside the noise equivalent electronic bandwidth Δf :

$$\sigma_i^2 = \left(2qi_{ave} + \frac{4k_B T F}{R_c}\right)\Delta f$$

$$\sigma_i^2 = (2\,q\,i_{ave} + \frac{4\,k_B\,T\,F}{R_L})\Delta\,f$$

Usually, some sources of noise will emerge as more important than the others : the main origin of noise varies with respect to the spectral domain of the sensor, the mode of detection (direct or heterodyne) and the ambient light environment (darkness, daylight, level of stray light,...). The table of figure 13 lists some of the main origins of noise in some usual cases.

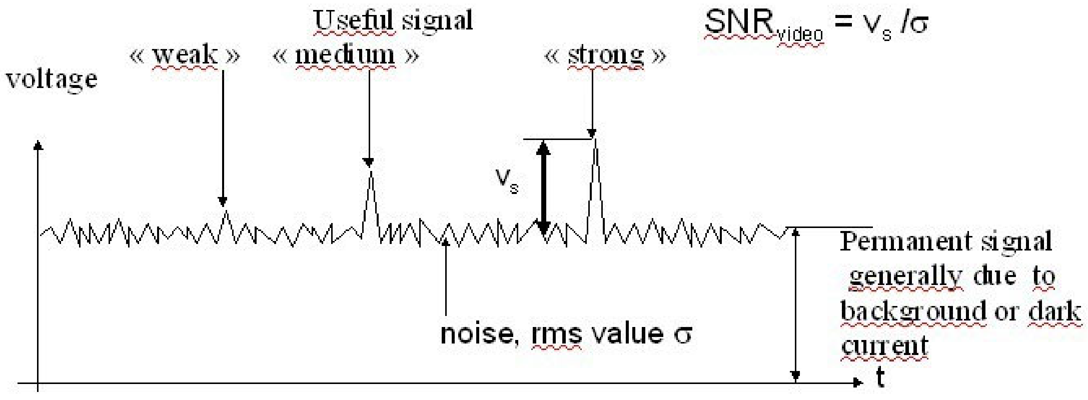
Source of noise	parameter	Variance σ_i^2
background (in IR)	Background current i_{BK}	$2\ q\ i_{BK}\ \Delta f$
signal	Signal current i_S	$2\ q\ i_S\ \Delta f$
darkness	Dark current i_{DK}	$2\ q\ i_{DK}\ \Delta f$
1/f	surface, connexion defects	To be defined
thermal	Temp. of load resistor R_L	$4\ k\ T\ \Delta f\ /\ R_L$
Amplifier	Noise figure F	$4\ k\ (F-1)\ T\ \Delta f\ /R_L$

$$\sigma_i^2 = 2\ q\ (\ i_{BK} + i_S + i_{DK}\ )\ \Delta f + 4\ k\ T\ F\ \Delta f\ /\ R_L$$

4.3. Signal to noise ratio and optimization

a) Definition

In quite many applications, noise adds to the useful signal, while in others, noise is proportional to signal. Whatever the case, the larger the noise fluctuations, the more difficult it is to detect or measure the useful signal. Figure 14, which shows a typical detector output with respect to time, is intended to illustrate the fact that the performance of an electro-optical sensor does not depend on its signal alone but on its **signal to noise ratio**.



Since the sensor output is noisy, it fluctuates above and below its average value by an instantaneous amount, for example $i_b(t)$ if one is concerned with the output current. The corresponding current or voltage variances, σ_i^2 or σ_v^2, inside the electronic bandpass of the sensor generate the following electrical noise power P_n across the load resistor R_L:

$$P_n = R_c\sigma_i^2 = \frac{\sigma_v^2}{R_c}$$

If $i_s$is the instantaneous pertinent output from the detector, the corresponding electric power of the signal is :

$$P_s = R_c i_s^2$$

By definition, the **power signal to noise ratio** of the sensor at that corresponding instant is the ratio between the electrical pertinent signal power and that of the noise, both being evaluated inside the sensor bandpass:

$$SNR_P = \frac{P_s}{P_n} = \frac{i_s^2}{\sigma_i^2} = \frac{v_s^2}{\sigma_v^2}$$

Attention

Another definition of the signal to noise ratio may be used : the video signal to noise ratio is the ratio between signal voltage and rms noise voltage (or ratio between corresponding currents). Its value is the square root of the previous one :

$$SNR_v = \frac{i_s}{\sigma_i} = \frac{v_s}{\sigma_v} = \sqrt{SNR_P}$$

Furthermore, signal to noise ratio is often expressed in decibels (dB), as follows :

$$SNR_{dB} = 10\log_{10}\left(\frac{P_s}{P_n}\right) = 20\log_{10}\left(\frac{v_s}{\sigma_v}\right)$$

b) Optical filtering, before detection

Introduction

Spatial (geometric) and spectral filtering are aimed at minimizing shot noise, due for example to stray light, and at maximizing lens transmittance for the pertinent signal.

Spatial filtering (examples)

- In infrared sensors, one will limit field of view of the (cryogenically cooled) detector to the exit pupil of the associated lens, in order to reduce its sensitivity to stray light radiated by the mechanical parts, main source of noise.
- Whenever possible, one will eliminate stray light from intense sources of light outside the field of view, by means of diaphragms, baffles, or protective screens.
- Early infrared missile guidance sensors were using finely etched rotating graticules aimed at reducing signals from extended sources such as clouds.
- In image forming sensors, the elemental field of view must be matched to the size of the finest details to be detected on the object.

Spectral filtering

The spectral bandwidth of the sensor must be optimized with respect to the signatures of useful and parasitic radiations. Spectral filtering is most efficient in active sensors (particularly laser sensors) since the signal to detect is then monochromatic, while stray light is usually spectrally wide. Under these conditions, interference filtering is recommended : centered upon the laser wavelength it will isolate a narrow bandpass : relative width ($\Delta\lambda/\lambda$) less than or of the order of 1% is rather commonplace.

If the sensor operates in the long wave IR band (8 to $12\mu m$ band), one may also use filters, but then they should be cooled down at the detector temperature, in order to eliminate their out of band thermal radiation (Kirchhoff's law).

With narrowband interference filters, one must not forget the spectral shifts due to changes in incidence angle, temperature, humidity, as well as possible problems of lifetime and when they are in or out of operating conditions.

c) Electronic filtering after detection : matched electronic filtering

In many applications, the variation in time of the expected signal is known. If furthermore, this signal is band limited and if the noise spectrum is white, signal processing techniques such as matched filtering are a good choice. A matched filter is such that its electronic gain is proportional to the conjugate of the Fourier transform of the signal amplitude spectrum : it amplifies components that contribute most to the signal and puts them back into phase. If $s(f)$ is the Fourier transform of the signal $S(t)$, then the gain of the matched filter is :

$$H_{matched}(f) = S(f)e^{-2\pi i f t_0}$$

and the pulse response of the matched filter is : $h(t) = K_S(t_0 - t)$: i.e. proportional to the time reversed signal $s(t)$ The pulse response is also delayed by some amount t_0, which means that the output pulse does not occur at time $t = 0$ but at some time $t = t_0$, (because of the « causality principle »).

Matched filtering is particularly well suited to the detection of known pulses imbedded in noise : if the pulse to detect is rectangular, of duration T, its amplitude spectrum is the following :

$$S(f) = \sin c(Tf) = \frac{\sin(\pi T f)}{\pi T f}$$

The power gain curve $G(f)$ of the matched filter is given by :

$$G(f) = G_0\left[\frac{\sin \pi T f}{\pi T f}\right]^2$$

The bandpass Δf of the filter that would produce the same output noise power with constant gain is called the **equivalent noise bandwidth of the filter**, and its value is the following :

$$\Delta f = \frac{1}{G_0}\int_0^\infty G(f)df = \frac{0.5}{T}$$

d) TDI : time delay and integration

Whenever it is possible to repeat measurements in time, it is advisable to average the results by summing them up in a coherent way, i.e. by synchronizing them. By means of this method, called **Time Delay and Integration** (or **TDI**), the signal that is obtained after summation is equal to :

$$S(t) = \sum s_i(t)$$

where $s_i(t)$ is the value of the signal at instant t. Noise values are considered to be uncorrelated from one experiment to the next, so that the noise variance of the sum is :

$$\sigma_\Sigma^2 = \sum \sigma_i^2 = n\sigma_i^2$$

There results from above that the video signal to noise ratio on the sum of n curves is higher than the average signal to noise ratio, on each individual curve, by $n^{1/2}$.

4.4. Performance and range evaluation of an electro-optical sensor

Introduction

Criteria used for performance simulation and evaluation differ from sensor to sensor, depending upon the application at hand : optical telecommunications, detection, imaging, astronomy, metrology,.... For example, the performance of a detection system is defined in terms of probability of detection and false alarm rate, that of a telecommunication system in terms of Bit Error Rate (BER), that of a measuring instrument by its uncertainty, repeatability,... In all cases performance optimization comes from signal to noise ratio optimization..

Radiometric budget and range equation

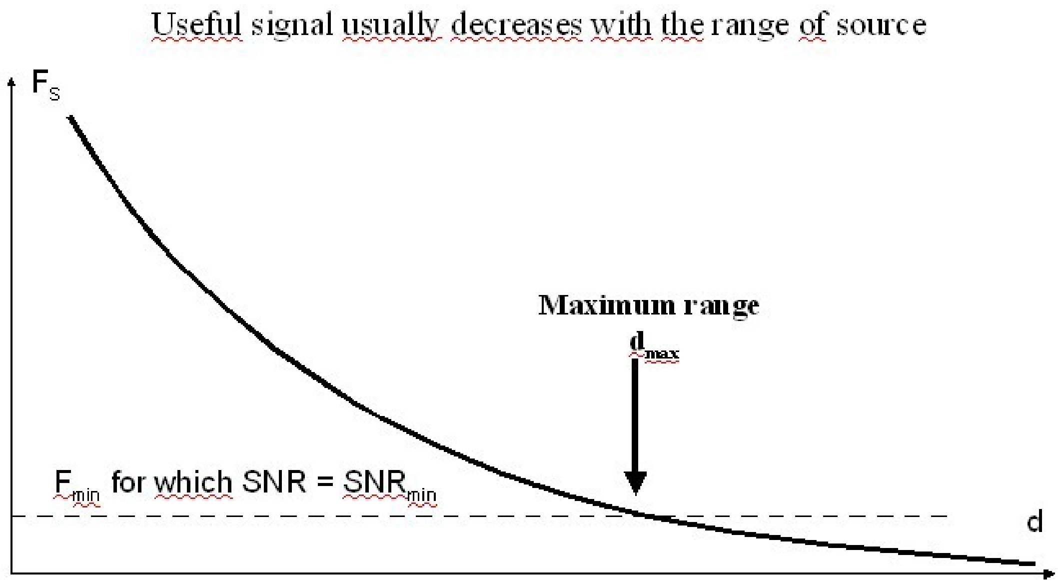
In order for the sensor to satisfy its operational specifications, the designer must make sure that the signal to noise ratio be larger than some threshold value. In order to achieve that result, the useful flux incident upon the detector be larger than some lower limit imposed upon the sensor by a set of criteria such as those mentionned above. By definition, the useful flux level corresponding to a signal to noise of 1 is the "**noise equivalent power**" (NEP) **of the sensor**. If the main source of noise is the detector, the NEP of the sensor is equal to the product of the power spectral density of the detector noise, in $W\,Hz^{-1/2}$ (data issued from the detector manufacturer) by the square root of the noise equivalent bandwidth of the sensor :

$$NEP_{sensor} = \left(\frac{d\,NEP}{\sqrt{df}}\right)_{det} \sqrt{\Delta f_{sensor}}$$

If SNR_{min} is the threshold value of the signal to noise ratio, there follows that the useful flux incident upon the detector must be larger than :

$$F_{signal} > (S/B)_{min} \times NEP_{sensor}$$

One of the main outputs from a sensor model is the **sensor Range equation** : it is the evaluation of the useful signal flux incident upon the detector with respect to the distance between the source and the sensor. Usually, this flux is decreasing with the distance. From the sensor range equation, it is hence possible to deduce the range (called **maximum range**) for which the signal to noise ratio reaches its threshold value : up to that range, the sensor should theoretically perform better than what is asked from it by the operational specifications and, beyond that range, the sensor should perform more poorly than expected (figure 15).



The range equation of an electro-optical sensor applies to only one configuration, and it is no more valid when any one of its parameters is being modified.

Exemple

For example, in defense applications, the maximum range of an electro-optical system (such as a rangefinder, a detection, surveillance or recognition system,...) on a given target varies widely if any one of these parameters is changed : target orientation, atmospheric conditions (day or night environment, meteorological range, climate, aerosols), target and sensor altitudes, optical path (ground to ground, air to air, air to ground, ground to air,...), surrounding background,...

In many cases, the sensor Noise Equivalent Power does not mean much to the user in terms of performance. That is why the sensor NEP is generally converted into its equivalent value in terms of the parameter of interest for the user : that is the case for example in thermal imagery where the sensor performance is very seldom evaluated by the NEP but by its equivalent quantity in temperature difference. This quantity, called **NETD**, or **Noise Equivalent Temperature Difference**, represents the change in temperature, between two blackbodies, that induces a signal to noise ratio of 1 from the thermal sensor.

III.Case study : detection systems

This case study completes the course on the importance of the signal to noise ratio value upon the performance of electro-optical sensors, and its importance in the design of detection systems. Some specific criteria are introduced here, such as the probability of detection and the probability of false alarm : these have been used in the radar domain for quite a while and they are being transposed here, with some simplification, to the case of electro-optical sensors.

Similarly to digital optical telecommunication systems, which are supposed to decode a two level signal (bit 1 and bit 0), detection systems (radar or electro-optical sensors) must decide upon the presence or the absence of a target, or of some given signal. Hence, their output is binary, summarized by the truth table below : the left column represents the actual situation (target is present or not), the middle one shows the expected output from the sensor, and the right ones possible errors :

Sensor output => actual situation	correct	erroneous
Presence of signal (target)	**presence : détection**	absence : non detection
Absence of signal (target)	Absence : surveillance	**presence : false alarm**

The optimization of a detection sensor consists in reducing the probability of occurrence of its erroneous answers (non detection of the target while it is present, and detection of a target while there is none, or false alarm) and in maximizing its probability of detection (probability of detecting a target when present). As we shall see, this is obtained by maximizing the signal to noise ratio (SNR) and making it larger than some threshold level : the better the expected performance, the larger the threshold and SNR values.

1. Statistics of the electrical output signal

The output signal from an electro-optical sensor is supposed to be different when the pertinent signal is present and when there is only noise : one of the most usual procedures for discriminating signal from noise (or bit 1 from bit 0 in digital telecommunications) is « **threshold logic** » in which, by comparison of the output voltage with some preset threshold value V_{th}, a binary decision is made by the sensor about the input signal : presence or absence of target, bit 1 or bit , in the following manner :

- if $v > V_{th}$===> presence of target (or presence of bit 1)
- if $v < V_{th}$ ===> absence of target (or presence of bit 0)

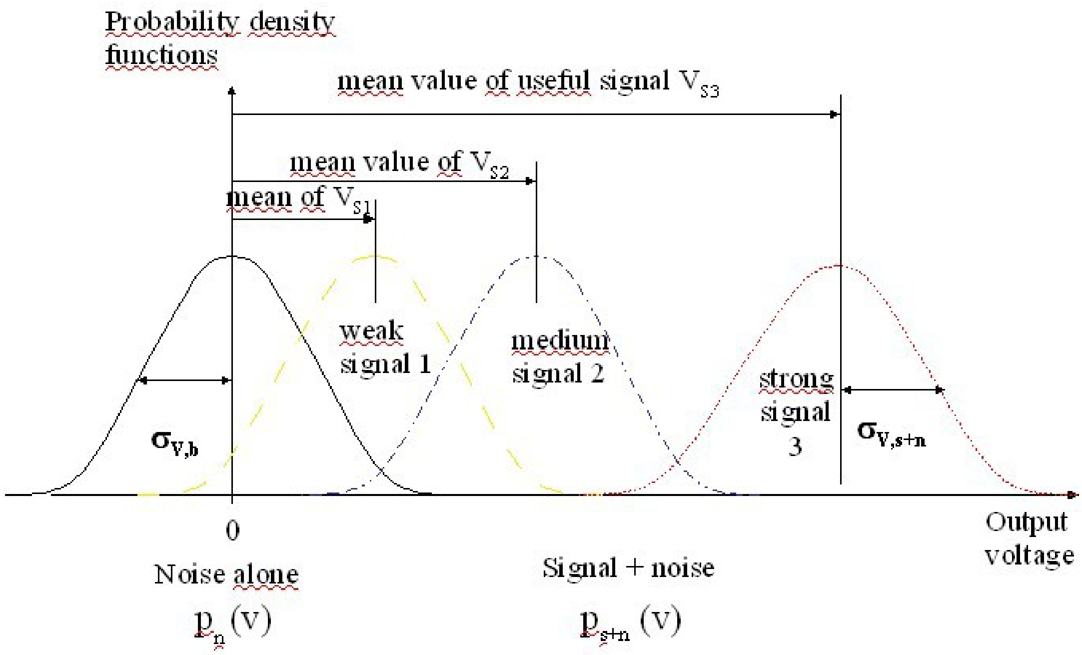
We have seen above that the voltage across the detector load resistor is noisy. Without getting into the details of each case, one must recall that noise levels depend upon such parameters as : mode of detection, spectral domain, scene lighting, type of source, detector,... One must also remember that noise and signal have to be defined, evaluted or measured at the same point (usually at the outset of the detector, and inside the same electrical bandpass, i.e. the electronic bandpass of the sensor.

Among other problems, the designer of an electro-optical sensor is facing the following ones : which values must be assigned to the threshold and to the SNR in order for the sensor to satisfy its specifications, generally expressed as **probability of detection** (this probability of detecting the target when present must be maximized) and **false alarm rate** (the number of false detections per unit time or per usage duration must be minimized). In order to answer those two questions, one characterizes the statistical properties of the output signal by means of its **probability density**. If the output is the instantaneous voltage across the load resistor, the probability density $p(v)$, then represents the voltage distribution function resulting from a large number of samples (either from modelling or from measurements). The probability Pr for an instantaneous voltage sample to be comprised between two values, v and $v + dv$, is given by the following relationship :

$$Pr(v\,; v + dv) = p(v)dv$$

Let us consider that, in the absence of the target or of the expected signal, the output voltage is characterized by the « noise alone probability density function » $p_n(v)$, and let us suppose, for mathematical convenience, that its average value is equal to zero, for example by means of some decoupling capacitance. Now, let us suppose that, in the presence of the target, the output voltage probability density is modified (if not, one can easily understand that there is no use going any further in the design of the sensor) : it becomes the « signal + noise probability density function », $p_{s+b}(v)$, Its average value, corresponds to that of the signal flux that is incident upon the detector, v_s, (v_s is the value that is computed by the procedure described in part D.1. of the course). Let $\sigma_{v+n}$and $\sigma_{v,s+n}$be the rms values of these voltage probability density functions, respectively obtained in the absence and in the presence of the target.

In order to illustrate these parameters, figure 1 (CS) below compares three distinct configurations, where one can observe a typical noise voltage probability density function and probability density functions corresponding to signals at three different levels : weak, medium, and strong. A rapid analysis of this figure shows that, as the separation between the average values the curves (« noise alone » and « signal + noise ») gets larger, decision making by the sensor (presence or absence of target) gets easier. That means that the larger the ratios $v_s/\sigma_{v,n}$and $v_s/\sigma_{v,s+n}$ the better the performance of the sensor.



These two ratios are respectively the « **signal to noise ratio with respect to internal noise of the sensor** » (i.e. noise level in the absence of target) and the signal to « **noise ratio due**

to the signal ». **Their influence upon the performance of an electro-optical detection sensor is fundamental.**

2. False alarm rate and setting of the threshold value

Let us come back to the « **threshold logic** » of a detection sensor, in which the presence or the absence of the target is deduced from the comparison between the output voltage and a preset threshold voltage. According to that logic, the probability of detection of a given target or signal is the probability that the instantaneous output voltage, or sample, be larger than the threshold whenever the target (or signal to be detected, such as a laser pulse) is present. Mathematically, the probability of detection is the integral of the « signal + noise » probability density function, $p_{s+n}(v)$ above the threshold value :

$$Pd = \int_{threshold}^{\infty} p_{s+n}(v)dv$$

Similarly, the probability of false alarm, P_{FA}, is the probability for a voltage sample to be larger than the threshold value while the target, or the signal to be detected, is absent (i.e. in the presence of noise alone). As for the probability of detection, the probability of false alarm is the integral of the « noise alone » density probability function, $p_n(v)$, above the threshold :

$$P_{FA} = \int_{seuil}^{\infty} p_n(v)dv$$

Probability density functions of output voltage in the presence of the target (or of the signal to be detected) vary quite a lot from one application to the other : the fluctuations on « useful signals » depend upon many parameters : in some cases, the signal is said to be stationary, which means that from sample to sample the difference comes mainly from the noise of the sensor, which may vary widely according to the mode of detection (direct or heterodyne), the spectral bandwidth of operation. Otherwise, these fluctuations may arise from the source itself, in case it is highly coherent, as is the case for single mode lasers, and if there is presence of speckle. They may arise from the object under illumination (changes in its orientation, objects with diffuse or specular surface,...).

For a given average value of the signal, one can easily understand that the actual performance of an EO sensor will be more or less degraded because of signal fluctuations from sample to sample. Indeed, some of the signal sampled values will be unnecessarily high, while other will be too low for acceptable detection. Among the most often encountered probability density functions of « useful signals », one will mention the following : Gauss, Laplace, Rayleigh, Gamma functions, ...

As for the output fluctuations due to the sensor itself (coming for a large part from shot and Johnson noises), their statistics is essentially gaussian. The performance analysis of this case study is limited to EO sensors in which probability density functions are gaussian, on the signal as well as on the noise. This hypothesis is far from being representative of all cases, but it is the most simple in its results and its bases may be extended to the other configurations, main difference being in the results.

This hypothesis being agreed upon, let us write down the two probability density functions, applicable either on the « signal » or on the « noise alone », i.e. in the absence or in the presence of the target (or of the useful signal) :

$$p_n(v) = \frac{1}{\sqrt{2\pi\sigma_{v,n}^2}} e^{\frac{-v^2}{2\sigma_{v,n}^2}}$$

$$p_{s+n}(v) = \frac{1}{\sqrt{2\pi\sigma^2_{v,s+n}}} e^{\frac{-(v-v_s)^2}{2\sigma^2_{v,s+n}}}$$

In these conditions, the probability of detection and the probability of false alarm are respectiveley given by :

$$P_d = \frac{1}{\sqrt{2\pi\sigma^2_{v,s+n}}} \int_{Vth}^{\infty} exp(\frac{-(v-v_s)^2}{2\sigma^2_{v,s+n}})dv$$

$$P_{FA} = \frac{1}{\sqrt{2\pi\sigma^2_{v,n}}} \int_{Vth}^{\infty} exp(\frac{-v^2}{2\sigma^2_{v,n}})dv$$

One of the first parameters to choose when designing a detection system (be it radar or electro-optical) is its threshold value (for instance, in voltage : V_{th}), which must be computed from the specification on the probability of false alarm and from the noise level of the equipment. From the table below, one will realize that the probability of false alarm is a rapidly decreasing function of the ratio, $v_{th}/\sigma_{v,n}$, between the threshold value and the rms value of the « noise alone ». For example, the probability of false alarm is 10^{-3} when this ratio is equal to 3 ; 10^{-9} for a ratio of 6, 10^{-12} for a ratio of 7,...

v_{th}/σ_v	3	4,25	5,2	6	7	8	9
P_{FA}	10^{-3}	10^{-5}	10^{-7}	10^{-9}	10^{-12}	10^{-15}	10^{-18}

Up to now we have shown that what counts in the choice of the threshold level is the probability of false alarm, i.e. the probability that an output sample be larger than the threshold in the absence of the target. For most detection systems, that parameter is not mentionned at all in the customer's specifications ; from the user's point of view, what is important in the specifications of false alarms is not the probability of false alarm itself but rather what is called the « **false alarm rate** », or **FAR**. FAR is the maximum number of false alarms tolerated by the user (or by the designer) per unit time, or during the operating time of the device ; it represents the number, per second or during the period of operation, of false detections, i.e. of voltage samples larger than the threshold, while no target is present inside the field of view of the equipment. FAR is hence one of the fundamental operating parameters of a detection system.

So the designer of a detection system must convert the operational FAR specification from the customer into the probability of false alarm, which is the mathematical parameter to be introduced into the design model. As mentionned before, the probability of false alarm (PFA) is the probability that an instantaneous sample be above the threshold value, while no target is present. FAR and PFA are related to each other by the fact that the number of false alarms over some time duration is the product of the false alarm rate per individual sample (which is nothing else than the PFA) by the number of samples being made during that time :

$$PFA = \frac{\text{FAR for a giventime duration}}{\text{number of mesurements during that time}}$$

For example, let us consider the case of a sensor operating without interruption, (100% of the time) : the number of samples being obtained during a given period of time is the product of that duration by the sampling rate of the sensor, so that :

$$PFA = \frac{\text{TFA (per second)}}{\text{sampling rate}}$$

Because he generally does not know the sampling rate of the sensor and is not directly interested in that parameter, the user of a detection system will only specify the desired FAR. Because he knows the necessary sampling rate of the sensor, defined from the signal duration and from such considerations as **Shannon sampling theorem** (« sampling frequency should be at least twice as high as the maximum frequency of the signal to be detected »), the designer is able to deduce PFA from FAR.

In quite many applications, the output signal is not relevant 100% of the operating time of the sensor : that typically happens with pulsed laser rangefinders, where target range is deduced from the time of flight of laser pulses ; if the maximum distance of the target , d_{max}, is known, it is recommended not to take the output signal into consideration for times of flight corresponding to longer ranges.For example, let us consider that the maximum target range is $15km$: that means that the flight time of the laser pulses is at most equal to :

$$\tau_{timeofflight} = \frac{2d_{max}}{c} = 10^{-4}s$$

Beyond that duration, it useless to sample the output voltage, and hence it is advisable to stop the counter, which eliminates risks of false alarm from targets thought to be too far away to give rise to some detectable laser echo. If the pulse repetition frequency (PRF) of the rangefinder is $10Hz$, each second of operation of the device leads to only $10^{-3}s$ of effective measurements. This corresponds to $(T_u/T) = 10^{-3}$ as being the percentage of time effectively dedicated to measurements. In the case of such « part-time » sensors, the relationship between FAR and PFA is the following :

$$PFA = \frac{\text{FAR (per second of use)}}{(T_u/T)\ \text{sampling rate}}$$

By discarding the output signal whenever it is not relevant, the designer will be able to set a much higher value to the probability of false alarm. As will be shown in the following paragraph, this procedure, i.e. reducing the number of samples to the minimum, leads to a reduction in the threshold value and hence to an improvement on the probability of detection, while the false alarm rate is being kept compatible with the sensor specifications.

3. Probability of detection and minimal value of SNR

From the point of view of the user, an EO detection sensor is characterized by its false alarm rate and by its probability of detection. We have seen that the first criterion, FAR, is the one that defines the threshold value to be imposed upon the output. In what follows, it is shown how the designer must take into account the second one, i.e. the probability of detection, in order to evaluate the minimum value of the signal to noise ratio.

By setting :

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt$$

then :

$$y = f(x) \rightarrow x = g(y)$$

f and g are said to be reciprocal functions from each other, and there comes :

$$P_{FA} = f(\frac{-v_{th}}{\sigma_{v,n}})$$

and hence :

$$\frac{-v_{th}}{\sigma_{v,n}} = -g(P_{FA})$$

On the other hand, if the gaussian hypothesis is being representative of the « signal + noise » probability density function, the probability of detection is equal to :

$$P_d = f(\frac{v_s - v_{th}}{\sigma_{v,s+n}})$$

and hence :

$$\frac{v_s - v_{th}}{\sigma_{v,s+n}} = g(P_d)$$

For a further simplification, we will consider here the case where the incoming signal is not fluctuating from sample to sample (signal is then said to be « **certain** »or « **stationary** »), and weak enough so that its presence or absence does not modify the noise level significantly. This approach can be justified in a vast majority of thermal infrared systems, for which the permanent background flux is much larger than the flux variation brought by the presence of the target to be detected. Consequently, the average output voltage is almost not modified by the presence of the target. If the noise is essentially due to the background radiation (in which case the sensor is said to be « **BLIP** », *Background Limited Infrared Photodetector*), the noise level is about the same in the absence or in the presence of the signal, which allows to write :

$$\sigma_{v,s+n} \approx \sigma_{v,n}$$

and

$$V_s/\sigma_{vn} = g(P_d) - g(P_{FA})$$

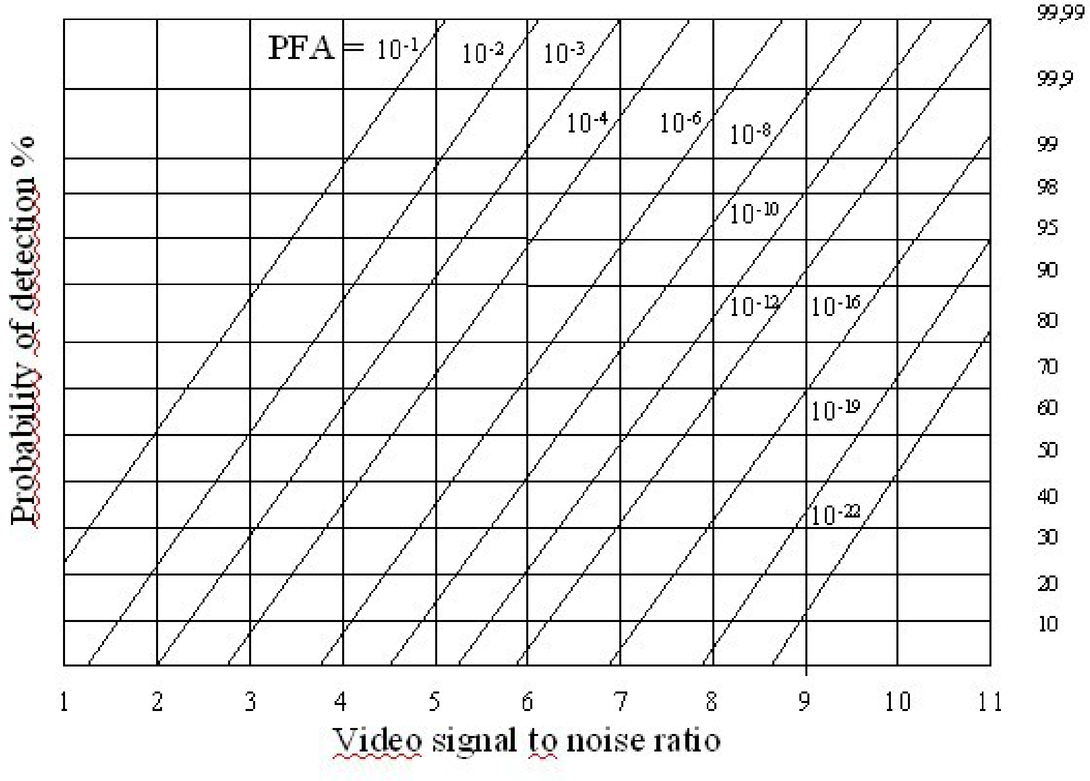
One concludes then that the minimum signal to noise ratio necessary for the sensor to satisfy the probability of detection and probability of false alarm (after conversion from the false alarm rate as explained above) :

$$(SNR)_{min} = g(P_d) - g(P_{FA})$$

The table below makes the correpondence between reciprocal functions y and $g(y)$ for several values of y, in case of gaussian noise :

y	g(y)	y	g(y)	y	g(y)	y	g(y)
10^{-12}	**-7,03**	$2\ 10^{-3}$	**-2,87**	0,6	**0,25**	$1 - 10^{-4}$	**3,71**
10^{-11}	**-6,7**	$5\ 10^{-3}$	**-2,57**	0,7	**0,52**	$1 - 10^{-5}$	**4,26**
10^{-10}	**-6,36**	10^{-2}	**-2,32**	0,8	**0,84**	$1 - 10^{-6}$	**4,75**
10^{-9}	**-5,99**	$2\ 10^{-2}$	**-2,05**	0,9	**1,28**	$1 - 10^{-7}$	**5,19**
10^{-8}	**-5,61**	$5\ 10^{-2}$	**-1,64**	0,95	**1,64**	$1 - 10^{-8}$	**5,61**
10^{-7}	**-5,19**	10^{-1}	**-1,28**	0,98	**2,05**	$1 - 10^{-9}$	**5,99**
10^{-6}	**-4,75**	$2\ 10^{-1}$	**-0,84**	0,99	**2,32**	$1 - 10^{-10}$	**6,36**
10^{-5}	**-4,26**	$3\ 10^{-1}$	**-0,52**	1	**2,57**	$1 - 10^{-11}$	**6,7**
10^{-4}	**-3,71**	$4\ 10^{-1}$	**-0,25**	1	**2,87**	$1- 10^{-12}$	**7,03**
10^{-3}	**-3,09**	$5\ 10^{-1}$	**0**	1	**3,09**		

The abacus that follows (Figure 17) may then be used in order to evaluate the minimum value of signal to noise ratio that is necessary to obtain the couple of specifications requested for P_d and PFA. This abacus is meant to be used only for the detection of stationary signals embedded in noise (the signal must not fluctuate from sample to sample), so that the « noise alone » and « signal + noise » probability density functions have the same rms value.



For example, if the required probability of detection Pd is 90%, and if the probability of false alarm PFA is 10^{-10}, then the minimum value of the video SNR is 8.

This abacus may also be read the other way around, if simulation of the sensor shows that the expected value for its signal to noise ratio is at best SNR_{max}. In these conditions, the optimal performances of the sensor are all aligned along the vertical line of abscissa SNR_{max}. For example, let us suppose that the maximum value that is predicted by the sensor model for the signal to noise ratio is $SNR_{max} = 6$. Then, by consulting the vertical line drawn from that limiting value, one may predict that performances of the sensor will be limited to a set of (P_d, PFA) values such as : $P_d = 70\%$ and $PFA = 10^{-7}$, $P_d = 85\%$ et $PFA = 10^{-6}$, $P_d = 93\%$ and $PFA = 10^{-5}$, $P_d = 98\%$ et $PFA = 10^{-4}$, ...

Remarque

This performance analysis has been limited to the detection of stationary signals or targets as a first approach to the more general problem of electro-optical detection. In actual systems, the incoming signal to be detected is much more fluctuating than the internal noise of the sensor : this is the case, for example, of laser sensors in the presence of speckle. Among the main consequences of signal non stationarity on the design and on the performances of such systems, as compared to the stationary case, one may say the following : for the same average value in SNR, the probability of detection is increased for low SNRs, (because the probability of high signal returns is more favorable than with stationary targets), but unfortunately, at more comfortable SNR values, the probability of detection becomes lower than with stationary signals.

Hence, in order to obtain a decent probability of detection (typically larger than 80%) in presence of a wildly fluctuating signal (detection of laser pulses in the presence of speckle, for example), the designer will be forced to increase by a non negligible factor (which may be of the order of 10) the minimum SNR value by comparison with what would be necessary to detect a stationary signal in the same conditions : that means that, for obtaining comparable performances from two laser sensors, the laser power that is necessary to detect fluctuating returns may have to be ten times larger than the one that would be necessary in case of stationary returns. Signal fluctuations are very costly in terms of sensor performance.

IV.Exercice

1. Knowledge Test

Separate Questions :

Question 1

[Solution n°1 p 43]

What is the radiant intensity, assumed to be constant inside the emitted cone of light, of a laser source of power P and total divergence α ? What is the number of photons emitted per second by the laser, if its wavelength is λ ?

Numerical application : $P = 1W$, $\alpha = 1mrad$, $\lambda = 1m$

Question 2

[Solution n°2 p 43]

What is the irradiance of the solar panel of an earth satellite that is oriented perpendicularly to the solar rays (consider the sun as a $6000K$ Blackbody, with angular diameter $\alpha_s$égal à $30'$) ?

Question 3

[Solution n°3 p 43]

An image forming electro-optical sensor is observing the ambient light being reflected from a diffuse surface of irradiance E and reflectance R, some distance d away. What is the irradiance E' of a (small) detector placed on axis in the image plane of the sensor if the focal length of the lens isf', the entrance pupil diameter D_e and the transmittance T_{op}. Consider that the propagating medium is the vacuum and the the object distance is very large as compared to the focal length.

2. Problem : Laser sensor for illumination and ranging of the moon

Capteur laser pour éclairage et télémétrie de la lune

Installed in several countries, a few pulsed laser systems illuminate a small area of the moon for precise ranging of the moon. We will consider that, for each pulse, the laser beam, of wavelength $\lambda(= 1,5m)$ radiate a peak power $P_c(= 1GW)$, during pulse duration $\tau(= 0,1ns)$, with a total divergence $\alpha(= 10\mu rad)$and that its intensity is constant inside the cone (and naught outside). The ground of the moon, supposed to be perpendicular to the laser beam, is lambertian, and its diffuse reflectance at the laser wavelength is $R(= 0,1)$.

A telescope, situated near the laser transmitter, is aligned onto the laser impact on the moon. Its entrance pupil diameter is $D_e(= 1m)$ and its aperture number is $N(= 5)$

Question 1

[Solution n°4 p 43]

1) What is the number $N_0$of photons being emitted inside each pulse ?

Question 2

[Solution n°5 p 44]

2) What is the peak laser radiant irradiance of the lunar surface ?

Question 3

[*Solution n°6 p 44*]

3) What is the peak radiance $L_{p,moon}$ and what is the peak intensity $L_{p,moon}$ of the lunar area being illuminated by the laser, back towards the earth ?

Question 4

[*Solution n°7 p 44*]

4) What is the number $N_1$of laser photons laser returning from the moon onto the detector of the telescope, if the detector diameter is $D_d(= 100\mu m)$?

Question 5

[*Solution n°8 p 44*]

5) What would this number be, N_2, if the existing detector was replaced by another one, smaller and of diameter $D'_d(= 60\mu m)$?

Question 6

[*Solution n°9 p 44*]

6) In order for the photons to originate from a precisely known and small area of the moon, the laser beam is aligned onto a lunar region where a set of large cube corners (or retro-reflectors) have been deposited. Their collecting area is $S_{RR}(= 1m^2)$ and their BRDF ($= 10^7 sr^{-1}$). What is then the number of photons returning from the cube corners onto the detector at each of the laser echoes?

Conclusion

Vous pouvez consulter les références suivantes :

- [Bases de radiométrie optique]- [Infrared and electro-optical systems handbook] - [La thermographie infrarouge]- [Electro-optical systems performance modeling]- [Thermal imaging systems]- [Introduction to sensor systems]- [The microwave engineering handbook]

Solution des exercices

>Solution n°1 *(exercice p. 40)*

If the intensity is constant inside the laser beam, the emitted power is equal to : $P = I\Omega$ where the solid angle of the beam is $\Omega = \pi\alpha^2/4$

Hence :

$I = 4P/\alpha^2 = 4/10^{-6}\pi = 1.3MW\ sr^{-1}$

The photonic (or photon) laser flux is the ratio between its power and the energy of each photon (since the energy is the same for all the photons from the laser) :

$\Phi p = P/h\nu = \frac{P}{hc}\lambda = 5.10^{18} s^{-1}$ (photons per second)

>Solution n°2 *(exercice p. 40)*

The geometrical extent of the pencil defined by the solar rays that illuminate the panel (of area S) under normal incidence is the following :

$G = S\Omega_{sun}$, where $\Omega_{sun}$is the solid angle under which the sun is being seen from the panel, which then receives the following solar flux :

$\Phi = L_{sun}G = L_{sun}S\Omega_{sun}$

where $L_{sun}$is the sun radiance = $K_3 T_{sun}^4$

There results that the panel irradiance is :

$E = \Phi/S\ = L_{sun}\Omega_{sun} = (\pi\alpha_s^2/4)K_3T_{sun}^4$

numerical application : $K_3 = 1,8\ 10^{-8} Wm^{-2}sr^{-1}K^{-4}$ (Stefan's constant) :

$E = 1400W\ m^{-2}$

>Solution n°3 *(exercice p. 40)*

The radiance of the diffuse object under irradiance E is $L_{objet} = RE/\pi$.

Since the sensor is an image forming system, the object is supposed to be larger than the sensor field of view and the flux incident upon its detector is :$F = T_{op}L_{objet}G_{detector}$.

From the results of the the main course, the geometrical extent of the conical beam that converges onto the the detector is : $G = \pi S_d \sin^2\alpha_M$ where S_d is the detector area.

One may conclude that the detector irradiance E' is given by :

$E'\ = RET_{op}\sin^2\alpha_M = RET_{op}/4N^2$

If the propagating medium is perfectly transparent, the detector irradiance in an image forming sensor depends upon the object irradiance, upon its diffuse and the aperture number of the lens, but it is independent from the distance from the object. Hence the the flux received by the detector is constant whatever the distance.

>Solution n°4 *(exercice p. 40)*

Number of emitted photons :

The number of emitted laser photons emitted during a pulse is the ratio between the pulse and the photon energies, i.e. :

$N_0 = P_p\tau/h\nu = 7.5\ 10^{17}s^{-1}$

>Solution n°5 *(exercice p. 40)*

Peak laser irradiance of the moon surface :

According to Bouguer's law, this irradiance is $E_{p,moon} = I_{p,moon}/d^2$où *d* where d is the distance between the earth and the moon ($380000km$), hence :

$E_{p,moon} = 4P_p/\pi\alpha^2 d^2 \sim 100W\ m^{-2}$

>Solution n°6 *(exercice p. 41)*

Peak radiance and intensity of the moon in return :

Since the moon surface is considered as lambertian, its peak radiance is :

$E_{p,moon} = RE_{p,moon}/\pi = 3W\ m^{-2}sr^{-1}$

and the peak intensity of the backscattered laser beam is:

$I_{p,moon} = RP_p/\pi = 30 \times 10^6 W\ sr^{-1}$

>Solution n°7 *(exercice p. 41)*

Number of photons de returning onto the detector :

the reception field of view of the sensor is :

$\alpha' = D_d/f' = 10^{-4}/5 = 20\mu rad$

Hence it is twice as large as the divergence of the transmitted laser beam, which means that theis sensor is a flux collector with respect to the laser spot on the moon. The laser flux being backscattered into the entrance pupil of the sensor, and hence onto the detector, if one considers the telescope as being perfectly transparent, is the following :

$P_d = I_{moon}S_{op}/d^2$

The number of photons N_1 incident upon the detector is then :

$N_l = P_d\tau/h\nu = 0.1$ photon per return

That means that in the average, there is one photon every ten shots.

>Solution n°8 *(exercice p. 41)*

If the existing detector is being replaced by another one that is smaller, the flux on the detector remains unchanged as long as the field of view of the receiver is larger than the laser spot on the moon. This is the case here, because with a detector diameter of $60mm$, the receiver field of view is $12\mu rad$, till slightly larger than the total divergence of the emitted laser beam. Admitting that there is no misalignment between laser axis and telescope axis, one finds that :

$N_2 = N_1$

>Solution n°9 *(exercice p. 41)*

The on-axis laser intensity of the set of cube corners (or retro-reflectors RR) is the product of its radiance by its area :

$I_{c,RR} = L_{c,RR} \times S_{RR} = BRDF \times E_{c,moon}S_{RR} = 10^7 \times 100 \times 1 = 10^9 W\ sr^{-1}$

It is is hence about 30 times larger than that of the ground of the moon. This advantage is the same in the number of photons returning from the cube corners with respect with those returning from the ground, But the largest advantage of the photons coming from the cubes as compared those coming from the ground is that they all come from the same tiny area ($1m^2$

instead of $10km^2$). So they all come back at the same instant onto the telescope, which is not the case for the other photons. In this application, where the number of “useful”photons est extremely small, signal processing is based upon photon counting techniques implying signal integration over very short periods of time (corresponding to “range windows”) case distance per case distance, and correlation from pulse to pulse.

Bibliographie

[Bases de radiométrie optique] Jean-Louis Meyzonnette, Thierry Lépine, *Bases de radiométrie optique*, Cépaduès-Editions, 2001.

[Electro-optical systems performance modeling] Waldman, Wooton, *Electro-optical systems performance modeling*, Artech, house, 1993.

[Infrared and electro-optical systems handbook] Acetta, Shumaker, *Infrared and electro-optical systems handbook*, SPIE Press, 1993.

[Introduction to sensor systems] Hovanessian, *Introduction to sensor systems*, Artech house, 1988.

[La thermographie infrarouge] G. Gaussorgues, *La thermographie infrarouge*, Lavoisier, 4° éditions, 1999.

[The microwave engineering handbook] Smith, Carpentier, *The microwave engineering handbook*, vol. 3, Chapman and Hall, 1993.

[Thermal imaging systems] Lloyd, *Thermal imaging systems*, Plenum Press, Third printing, 1982.