PART 1

An Overview of Image Processing and Analysis (IPA)
Gray-Tone Images

In this textbook, the term image will have a physical meaning and will refer to a one-, two- or three-dimensional (3D), continuous or discrete (including the digital form) radiometric spatial distribution of light (or another radiation) intensities.

1.1. Intensity images, pixels and gray tones

Radiometric images are spatially defined on pixels (contraction of “picture elements”) with intensity values called gray tones. Such images are often abusively called “black and white” images in the common language (panchromatic images is better suited and sometimes used in relation to the visible light and the human eye) [ALL 10; Original 1st ed., 1890]. In this book, they will be naturally designated as gray-tone images. Color images (e.g. three colors according to the human visual perception), multispectral images (e.g. four or five colors as in satellite imagery) [LEE 05] [PET 10; p. 665] and hyperspectral images (i.e. numerous almost monochromatic channels) [CHA 03b] will not be discussed because they require specific frameworks and approaches, still subject to particular mathematical research works.

The term illumination designates the incident light (or another radiation, such as an electromagnetic or nuclear radiation, e.g. gamma rays and X-rays) [DAI 74, HEN 02, BAR 04, HOR 06].

There exist a lot of imaging modalities, in particular for materials investigation, and in biological and medical imaging, but also in many other scientific, engineering or technical fields, as well for professional and personal purposes (e.g. magnetic resonance imaging (MRI), positron emission tomography (PET), scanning electron microscopy (SEM), transmission electron microscopy (TEM) and ultrasound imaging (US)).
The term **imaging**, although wider than the terms “image processing” and “image analysis”, as it also includes “image acquisition” and “image visualization” aspects (see section I.1), will be employed with a general meaning in this book.

### 1.2. Scene, objects, context, foreground and background

The term **scene** designates generically everything that is observed, i.e. a particular physical environment, and takes different names according to the addressed situations and concerned users: (1) sample (e.g. in metallurgy or histology), (2) raw or manufactured component (e.g. in industrial inspection), (3) body or organ (e.g. in medical imaging), (4) specimen (e.g. in biology or zoology), etc.

An image is an **observation** of a scene, which is most often a partial and incomplete view of it, called **field of view** (FoV). The observed scene, depending on (1) the light (e.g. visible, infrared, ultraviolet or X-rays light) or (2) another radiation (e.g. electronic) constituting the illumination (see section I.1), (3) the nature of its interaction with this illumination (e.g. simple reflection or fluorescence) and (4) the type of collected images (e.g. by transmission or reflection), will be investigated through a field of observation (i.e. FoV) corresponding to a width and length for two-dimensional (2D) imaging, as well as a **depth of field** (DoF) for 3D imaging. It is a major difficulty in imaging.

In a scene, there are **objects** located in a **context**, generally called the “background”. Depending on the addressed situation or the application issue, the context is also called (1) ambient space (e.g. in Geometry or Physics), (2) matrix (e.g. in Geology, or in Materials Sciences and Engineering), (3) medium (e.g. in Biology, or Chemical Sciences and Engineering), etc., while a class of similar objects will be called (1) phase (e.g. in Physics), or (2) population (e.g. in Biology, or Chemical Sciences and Engineering), etc. The term “object” is thus very general.

Therefore, an image ‘ideally’ includes **background pixels**, corresponding to the context, and **foreground pixels**, corresponding to the objects, or more often only to parts of objects. The term ‘ideal’ expresses the fact that in reality some pixels or groups of pixels can be wrongly considered as belonging to the other category. This is another major difficulty in imaging.

### 1.3. Simple intensity image formation process models

The purpose of this section is to present several **image formation process models and laws** that form the basis of the main imaging processes.
1.3.1. The multiplicative image formation process model

The basic nature of an intensity image, denoted by $f$, may be considered as being characterized by two components [OPP 68, HUA 71] [GON 87; section 2.2; 1st ed., 1977]. One component is the amount of light (or another radiation) incident on the scene being observed, while the other is the amount of light (or another radiation) reflected (or transmitted) by the scene. These two components are appropriately called the illumination and reflectance (or transmittance), and are denoted by $i$ and $r$ (or $t$), respectively. These two functions combine as a product to form an intensity image $f$, which is given at spatial location $x$ by [OPP 68, STO 72, PIN 97a]:

$$f(x) := i(x) \cdot r(x),$$  \hspace{1cm} [1.1] \\
or $$f(x) := i(x) \cdot t(x),$$  \hspace{1cm} [1.2]

where $0 < i(x) < +\infty$ and $0 \leq r(x) \leq 1$, and “.” is the standard product operation.

1.3.1.1. Lambert’s reflection cosine law

Lambert’s reflection cosine law [LAM 60] states that the apparent brightness of a Lambertian surface is proportional to the cosine of the angle between the surface normal and the direction of the incident light at spatial location $x$ [PRA 07; p. 55.; 1st ed., 1978] [PED 93] [BRO 08; p. 273]:

$$f_r(x) = i(x) \cos(\theta(x)),$$  \hspace{1cm} [1.3]

where $f_r(x)$ is the intensity of the diffusely reflected light (i.e. surface brightness), $i(x)$ is the intensity of the incoming light and $\theta(x)$ is the angle between the direction of the two vectors (assuming that $f_r(x) = 0$ when the cosine takes on negative values).

The reflected intensity will be the highest if the surface is perpendicular to the direction of the light, and the lowest if the surface runs parallel with the direction of the light.

The reflection coefficient, reflection ratio, denoted $r(x)$ ($0 \leq r(x) \leq 1$), given by [BRO 08; p. 273]:

$$r(x) := \cos(\theta(x)),$$  \hspace{1cm} [1.4]

is called the albedo [LAM 60] (from the Latin term albedo which means “whiteness”), and is the ratio of reflected radiation from the surface to the incident radiation upon it.

In general, the albedo depends on the directional distribution of incident radiation, except for Lambertian surfaces which scatter radiation in all directions according to a cosine function and therefore have an albedo that is independent of the incident radiation distribution.
1.3.1.2. **Bouguer–Beer–Lambert’s attenuation law**

In Optics, Bouguer–Beer–Lambert’s attenuation law [BOU 29, LAM 60, BEE 52] relates the absorption of light (or another radiation) to the properties of the material through which the light (or another radiation) travels.

**Bouguer–Beer–Lambert’s attenuation law** can be expressed at spatial location \( x \) by [HUN 75] [ATK 10; 1st ed., 1978]:

\[
f_t(x) = i(x) \exp (-c_{BBL}z(x)),
\]

where \( f_t(x) \) is the transmitted intensity, \( i(x) \) is the incident intensity, \( c_{BBL} \) is an attenuation coefficient (i.e. a strictly positive real number) that depends on the material and \( z(x) \) is the traveled thickness through the material.

The ratio of intensities is a real-number value, denoted \( t(x) \), called the transmittance ratio \((0 \leq t \leq 1)\):

\[
t(x) = \frac{f_t(x)}{i(x)},
\]

assuming that the incident intensity is not zero.

1.3.1.3. **Hounsfield’s X-ray unit**

In X-ray imaging, **Hounsfield unit (HU)** is defined by [FEE 10]:

\[
u_H := 1000 \frac{c_a - c_w}{c_w},
\]

where \( c_a \) and \( c_w \) are the attenuation coefficients of the material and the (distilled) water (under specific conditions, i.e. standard pressure and temperature), respectively.

1.3.1.4. **Hurter–Drifffield’s photographic recording law**

The Hurter–Drifffield photographic recording law [HUR 90, HUR 98], which was stated in the 1870s for photographic film recording, relates the film density (i.e. the logarithm of opacity) versus the logarithm of the total exposure called the characteristic **Hurter–Drifffield’s curve**. The overall shape of such a curve is a bit like an “S” slanted so that its base (the ‘fog’ region) and top (the saturation region) are horizontal (i.e. a sigmoid curve [SEG 07]), and with a central region which approximates to a straight line. The slope of this ‘straight-line’ portion is called the **HD-gamma** [HUN 75].
Within this portion, **Hurter–Driffield’s photographic recording law** can be expressed at spatial location \( x \) by [HUN 75] [PRA 07; p. 356, 1st ed., 1978] [CAR 00]:

\[
d_{HD}(x) = -c_{HD} \ln \left( \frac{f_i(x)}{i_{ref}} \right),
\]

where \( d_{HD}(x) \) is the optical density, \( f_i(x) \) is the incident intensity, \( i_{ref} \) is the reference intensity value (\( i_{ref} \geq f(x) \)) and \( c_{HD} \) is the **HD-gamma** proportionality constant that depends on the units used.

### 1.3.2. The main human brightness perception laws

Subjective or perceptual **brightness** is an attribute of (human) visual perception in which a scene appears to be reflecting or transmitting light. This is a subjective attribute of a scene being observed.

The specialized literature describing human brightness response to **stimulus intensity** includes many uncorrelated results due to the various viewpoints and focus interests of researchers from different scientific disciplines [XIE 89, KRU 89, KRU 91]. Several human brightness perception laws have been studied and reported, e.g. Weber’s law, Fechner’s law, devVries-Rose’s law, Stevens’s law and Naka–Rushton’s electrophysiological law.

#### 1.3.2.1. Weber’s brightness perception law

The response to light intensity by the human visual system has been known to be nonlinear since the mid-19th Century, when the psychophysicist E.H. Weber [WEB 46] established the now so-called “Weber’s visual law”. He argued that the human visual detection depends on the ratio, rather than the difference, between two incident light intensity values \( f \) and \( f + df \), where \( df \) is the so-called **just noticeable difference** (JND), also called the least perceptible difference [JUD 32], which is the amount of light necessary to add to a visual test field of constant intensity value \( f \) such that it can be discriminated from the reference light field of constant intensity value \( f \)[GOR 89; p. 17] [WAT 91].

**Weber’s brightness perception law** is expressed as [GOR 89; p. 18]:

\[
\frac{df}{f} = c_W,
\]

where \( f \) and \( f + df \) are two just noticeable incident light intensities (i.e. the magnitudes of the physical stimuli), and \( c_W \) is a real-number constant called **Bouguer-Weber’s constant** that has been found to be near 0.025 for retinal rods [COR 65].
1.3.2.2. *Fechner’s brightness perception law*

A few years after Weber, G. Fechner [FEC 60] (Weber’s student) explained the nonlinearity of the human visual perception as follows: in order to produce incremental arithmetic steps in sensation, the light intensity must grow geometrically. He proposed the following relationship between the incident light intensity \( f \) (the so-called *stimulus*) and the brightness \( b_F \) (the so-called *sensation*):

\[
\frac{db_F(x)}{df} = c_F \frac{df}{f}(x),
\]

where \( df \) is the increment of incident light that produces the increment \( db_F \) of visual sensation (brightness), and \( c_F \) is a real-number constant that depends on the units used.

*Fechner’s brightness perception law* can then be expressed as [GOR 89; p. 17]:

\[
b_F(x) = c_F \ln \left( \frac{f(x)}{f_{\text{min}}} \right),
\]

where \( f(x) \) is the incident light intensity (i.e. the magnitude of the physical stimulus), \( b_F(x) \) is the brightness (i.e. the subjective magnitude of the sensation evoked by the stimulus), \( f_{\text{min}} \) is the *absolute threshold* [COR 70; Chapters 2 and 4] [GOR 89; p. 15] of the human visual system, which is known to be very close to the physical complete darkness [PIR 67, ZUI 83], and \( c_F \) is a strictly positive real-number proportionality constant that depends on the units used.

Fechner’s brightness perception law can be equivalently expressed as [PIN 97b]:

\[
b_F(x) = c_F \ln \left( \frac{f(x)}{f_{\text{max}}} \right) + c_F \ln \left( \frac{f_{\text{max}}}{f_{\text{min}}} \right),
\]

where \( f_{\text{max}} \) is the *upper threshold* (or *glare limit*) of the Human Vision [GON 87; p. 39, 1st ed., 1977] [LEV 00].

1.3.2.3. *Stevens’s brightness perception law*

In the 1950s, S.S. Stevens [STE 57b, STE 57a, STE 64] proposed a power law for describing the relationship between the magnitude of a physical stimulus and its perceived intensity or strength.

The general form of *Stevens’s brightness perception law* is [GOR 89; p. 30]:

\[
b_S(x) = c_S f(x)^{e_S},
\]
where \( f(x) \) is the incident light intensity (i.e. the magnitude of the physical stimulus), \( b_S(x) \) is the brightness (i.e. the subjective magnitude of the sensation evoked by the stimulus), \( c_S \) is a strictly positive real-number proportionality constant that depends on the units used and \( e_S \) is an exponent, called Stevens’ exponent, that depends on the type of stimulation that has been found to be near to 0.3 [HAL 77].

1.3.2.4. Other brightness perception laws

Other human brightness perception laws have been reported such as devries-Rose’s square-root law [DEV 43, ROS 48, ROS 73a, ZEE 78] and Naka-Rushton’s electrophysiological law [NAK 66, NOR 74, HOO 79b].

1.3.2.5. Concluding discussion

Considerable debate has occurred in the specialized literature about the form of the response-intensity function, in particular between the Fechnerians and the Stevensians. Indeed, Fechner’s brightness perception law has been largely criticized and rejected by Stevens and his school (e.g. [STE 57b, STE 57a, STE 64]), who claimed a power law instead of a logarithmic law. The Stevens’ brightness perception law has also been criticized on both empirical and theoretical grounds (e.g. [TRE 64, POU 68, GRE 88]). Some authors have tried to relate some of these human visual laws (e.g. [EKM 64, TEG 71, KVA 92]). Others have tried to reconcile the partisans of some of these laws by proposing modified or unified human visual laws (e.g. [MCG 68, GRA 74, MAN 76]). Many researchers even argued that no general law exists (e.g. [HOO 79a]) or can exist (e.g. [WEI 81]). The debate still continues (e.g. [LUC 02]).

1.4. The five main requirements for a relevant imaging approach

Addressed from a scientific and technical viewpoint, and especially mathematically, an image processing and analysis approach needs to satisfy the following five main requirements [HUA 71, STO 72, GRA 81, MAR 82, SCH 86, PIN 97a, PAN 08]:

1) Physical and/or psycho-physical relevance: it is based on a physically relevant image formation process model (e.g. it is known that reflected light or transmitted light images follow multiplicative laws based on the product of an illumination (i.e. an incident radiation) with a reflectance or a transmittance component (i.e. the response of the illuminated scene)) [STO 72], or on a psychophysically relevant image formation process model (e.g. the human visual perception is known to be nonlinear in brightness) (see section 1.3), [GRA 81].

2) Mathematical powerfulness and consistency with the physical nature of the images, that is to say compatible with the physical or psychophysical laws underlying the image formation and combination processes (see section 1.3), [PIN 97a].
3) **Computational tractability**: its operations and operators are computationally implementable with sufficient effectiveness (e.g. optical processings are often preferable to computer processings because of their high speed, but the latter offer more opportunities concerning the transformations made) [MAR 82, SCH 86].

4) **Practical fruitfulness**: it is proved to be practically fruitful in the sense that it allows us to successfully address application issues in real situations (e.g. computer tomography allows the investigation of real objects in three spatial dimensions) [STO 72, GON 87].

5) **Quality assertion**: the quality of resulting images is asserted from a subjective or/and an objective viewpoint [HUA 71, PAN 08].

### 1.5. Additional comments

#### Historical comments and references

The term *albedo* was introduced into Optics by J. Lambert in 1760 [LAM 60].

P. Bouguer [BOU 29] is considered as the father of *Photometry*.

#### Bibliographic notes and additional readings

Examples of overview papers are [RUS 04a, RUS 04b]. The survey papers of [HUA 71] and [HUN 75] are good introductory journal articles to Image Processing containing relatively detailed matter on image formation.

The reading of introductory parts of classical books is recommended (e.g. [ROS 69a] [GON 87; 1st ed., 1977] [PRA 07; 1st ed., 1978] [GON 08; Chapters 1 and 2]). See also [SCH 86].

For *photographic-type imaging processes*, see [DAI 74]; for *medical imaging physics*, see [HEN 02]; for *Machine Vision*, see [HOR 06]; for *Computer Vision*, see [BAL 82, SHA 01, FOR 03]; for *Human Vision models* applied to image processing, see [STO 72, GRA 81, XIE 89, NAD 00].

Concerning *Human Vision*, refer to [COR 70, WAT 91], while to have an understanding of the *theories of visual perception*, see [GOR 89].

#### Further topics and readings

More realistic, but sophisticated reflection models (more or less empirical) have been developed in *Computer Graphics* such as *Phong’s reflection model* [PHO 75],
Cook–Torrance’s reflection model [COO 81] and Oren–Nayar’s reflection model [ORE 94].

Physiological experiments have reported that the human eye can detect a few number of light photons, i.e. of quantum of light [COR 70, LEV 00].

Albedo is very useful as a first approximation. In practice, a bidirectional reflectance distribution function (BRDF) [NIC 65] may be required to accurately characterize the scattering properties of a surface. Lambertian surfaces (see section 1.3.1.1) represent perfectly diffuse (i.e. matte) surfaces by a constant BRDF.