P. Zanuttigh, G, Marin, C. Dal Mutto, F. Dominio, L. Minto, G. Cortelazzo
Time-of-Flight and Structured Light Depth Cameras: Technology and Applications
Springer 2016

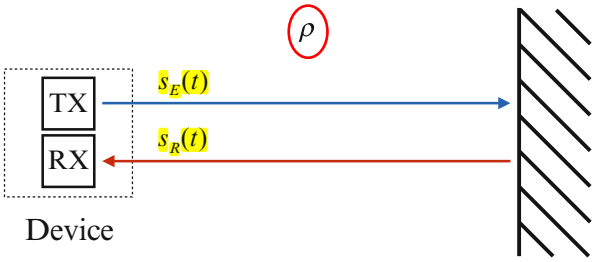

Fig. 1.14 Scheme of principle of ToF measurement

for a distance ρ. It is then echoed or back-reflected by a point P on the scene surface and travels a distance ρ. At time τ it reaches the ToF receiver (or sensor) RX, ideally co-positioned with the transmitter, as signal $s_R(t)$. Since at time τ the path length covered by the radiation is 2ρ, the relationship between ρ and τ is

$$\rho = \frac{c\tau}{2} \tag{1.51}$$

which is the basic expression of a ToF camera's distance measurement.

In stereo or structured light systems, occlusions are inevitable due to the presence of two cameras, or a camera and a projector, in different positions. Additionally, the distance between the camera positions (i.e. the baseline) improves the distance measurement accuracy. This is an intrinsic difference with respect to ToF, in which measurements are essentially occlusion-free, because the ToF measurement scheme assumes the transmitter and receiver are collinear and ideally co-positioned. In common practice such a requirement is enforced by placing them as close together as possible.

In spite of the conceptual simplicity of relationship (1.51), its implementation presents tremendous technological challenges because it involves the speed of light. For example, since

$$c = 3 \times 10^8 \, \frac{[m]}{[s]} = 2 \times 150 \, \frac{[m]}{[ps]} = 2 \times 0.15 \, \frac{[\text{mm}]}{[\text{ps}]} \tag{1.52}$$

it takes 6.67 [ns] to cover a 1 [m] path and distance measurements of nominal resolution of 1 [mm] need time measurement mechanisms with accuracy superior to $6.67 \div 7$ [ps], while a nominal resolution of 1 [cm] needs accuracy superior to 70 [ps]. The accurate measurement of round-trip time τ is the fundamental challenge in ToF systems and can be solved by two approaches: direct methods, addressing either the measurement of time τ by pulsed light or of phase φ in case of continuous wave operation (see Fig. 1.15), and indirect methods deriving τ (or φ as an intermediate step) from time-gated measurements of signal $s_R(t)$ at the receiver. The former methods are considered next while the latter methods will be discussed in detail in Chap. 3.

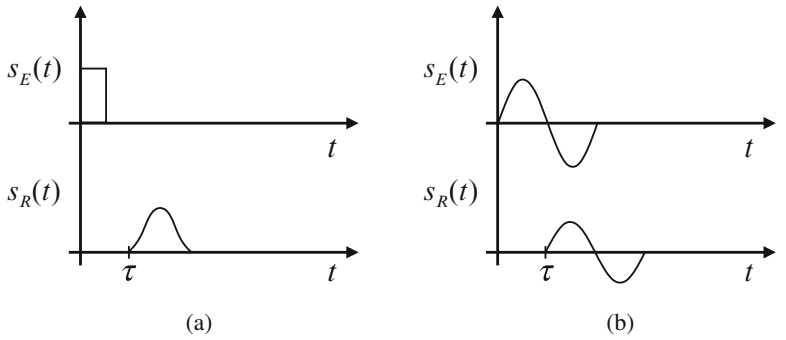

Fig. 1.15 Schematic operation of direct ToF measurement: (**a**) pulse modulation; (**b**) CW (sinusoidal) modulation

1.4.2 Direct ToF Measurement Methods

1.4.2.1 Direct Pulse Modulation

Figure 1.15a shows a square pulse $s_E(t)$ transmitted at time 0 and the received signal $s_R(t)$ arriving at time τ. Pulsed light has several advantages since it allows for a high signal to noise ratio due to the characteristics of pulse signals, which concentrate high energy values in a short time. High signal to noise ratio gives good resilience with respect to background illumination, enabling long distance measurements, reducing average optical power (with eye safety and power consumption advantages) and eliminating the need for high sensibility detectors at the receiver front-end.

The receiver back-end has a measurement device called a stop-watch, typically implemented by a counter with a start coinciding with the transmission of $s_E(t)$ and a stop coinciding with at the detection of $s_R(t)$. The accurate detection of $s_R(t)$ represents the most challenging task of the receiver. Indeed, the actual shape and values of $s_R(t)$ are hardly predictable since they do not only depend on the target distance and material but on the atmospheric attenuation of the signal path, especially for long distances. Therefore, simple fixed thresholding on the receiver optical power or straightforward correlation are not adequate, also because pulses with sharp rise and fall times at the transmitter side are not easy to implement. In addition, pulse signaling requires large bandwidth and high dynamic range at the receiver sensors.

Light-emitting diodes (LEDs) and laser diodes are commonly used for the pulse generation of ToF systems with repetition rates on the order of tens of KHz. Current commercial products for architectural and topographic surveys [1, 4, 8, 10] covering ranges from tens of meters to kilometers use direct ToF measurements based on pulse modulation.

1.4.2.2 Direct CW Modulation

Direct continuous wave (CW) modulation is not as common as direct pulse modulation. However, because CW modulation is commonly used also with indirect ToF measurements as we will discuss in Chap. 3, we introduce its basic principles here for sinusoidal signals.

Figure 1.15b shows a sinusoid $s_E(t) = A_E \sin(2\pi f_c t + \varphi_c)$ transmitted at $t = 0$ and the received attenuated sinusoidal $s_R(t)$ which arrives at time $t = \tau$

$$\begin{aligned} s_R(t) &= A_R \sin\left[2\pi f_c (t+\tau) + \varphi_c\right] \\ &= A_R \sin\left[2\pi f_c t + \varphi_c + \Delta\varphi\right] \\ &= \frac{A_R}{A_E} s_E(t+\tau) \end{aligned} \tag{1.53}$$

where

$$\Delta\varphi = 2\pi f_c \tau \tag{1.54}$$

is the phase difference between $s_E(t)$ and $s_R(t)$.

The round trip transmitter to target point distance ρ from (1.54) and (1.51) can be directly computed from the phase difference as

$$\rho = c\frac{\Delta\varphi}{4\pi f_c}. \tag{1.55}$$

Many methods can be employed to estimate the phase difference between sinusoids as this is a classical telecommunications and control issue (see, for example, the phase-locked loop literature). It is important to note that the measurement of $\Delta\varphi$ only requires knowledge of f_c but not φ_c, hence the transmitter can operate with incoherent sinusoids.

An important characteristic of direct ToF systems which differs from stereo and structured light systems is that the measurement accuracy is distance independent, only depending on the accuracy of the time or phase measurement devices. Such an important characteristic will be retained also by indirect ToF measurement systems as discussed in Chap. 3.

1.4.3 Surface Measurement by Single Point and Matricial ToF Systems

ToF systems made by a single transmitter and receiver, as schematically shown in Fig. 1.14, are typically used in range-finders for point-wise measurements. Such systems can also be mounted on time sequential 2D (or 1D) scanning mechanisms

for measuring distance on surfaces or along lines. It is typical to move a single ToF transmitter and receiver system along a vertical linear support placed on a rotating platform, with motion both in vertical and horizontal directions, as in the case of the scanning systems used for topographic or architectural surveys (e.g., [1, 4, 8, 10]). Since any time sequential scanning mechanism requires time in order to scan different scene points, such systems are intrinsically unsuited to acquire dynamic scenes, i.e., scenes with moving objects. Nevertheless, fast ToF scanning systems have been studied and some products, targeting automotive applications, reached the market [9]. 3D scanning mechanisms are bulky and unsuited to vibrations, as well as expensive, since they require high precision mechanics and electro-optics.

The progress of microelectronics fostered a different surface measurement solution, based on the concept of scanner-less ToF systems [38] and appeared in the literature under various names such as Range IMaging (RIM) cameras, ToF depth cameras, or ToF cameras, which are the denominations used in this book. Unlike systems which acquire scene geometry one point at a time by a point-wise ToF sensor mounted on time sequential scanning mechanisms, ToF cameras estimate the scene geometry in a single shot by a matrix of $N_R \times N_C$ in-pixel ToF sensors where all the pixels independently but simultaneously measure the distance of the scene point in front of them. ToF cameras deliver depth maps, i.e., measurement matrices with entries giving the distance between the matrix sensor and the imaged scene points, at video rates.

Even though a ToF camera may be conceptually interpreted as a matricial organization of many single devices, each made by an emitter and a co-positioned receiver as discussed previously, in practice implementations based on a simple juxtaposition of a multitude of single-point measurement devices are not feasible. Currently, it is not possible to integrate $N_R \times N_C$ emitters and $N_R \times N_C$ receivers in a single chip, especially for high values of N_R and N_C as required by imaging applications. However, each receiver does not require a specific co-positioned emitter; instead, a single emitter may provide an irradiation that is reflected back by the scene and collected by a multitude of receivers close to each other. Once the receivers are separated from emitters, the former can be integrated in a $N_R \times N_C$ matrix of lock-in pixels commonly called *ToF camera sensor* or more simply, *ToF sensor*.

The implementation of lock-in pixels and their integration into matricial configurations are fundamental issues of current ToF systems research and industrial development. Different technological choices lead to different ToF camera types. An in depth review of state of the art ToF technology can be found in [56].

1.4.4 ToF Depth Camera Components

ToF depth cameras lend themselves to a countless variety of different solutions, however, all the current implementations share the same structure shown in Fig. 1.16 made by the following basic components:

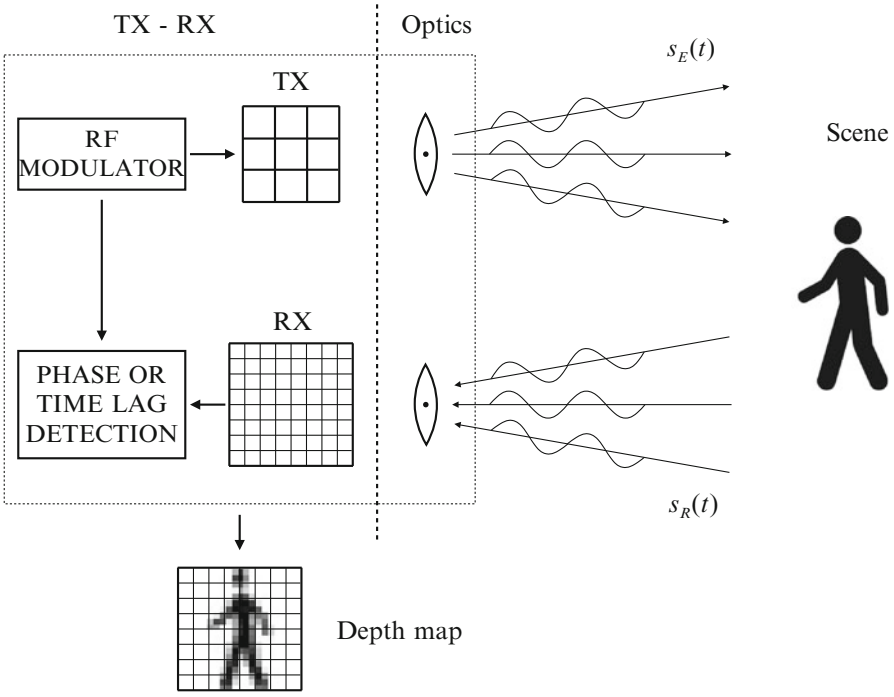

Fig. 1.16 Basic ToF depth camera structure

(a) The transmitter is typically made by an array of laser emitters or LEDs with an optics allowing illumination of the whole scene by the same signal $s_E(t)$.
(b) The receiver can be interpreted as a matrix of single ToF receivers coupled with an imaging system framing a scene frustum. Each pixel receives the signal $s_R(t)$ echoed by a surface point P.
(c) The emitted light is modulated by frequencies of some tens of MHz (in general in the high HF or low VHF bands), e.g., 16, 20, 30, 60, 120 MHz and each pixel of the receiver sensor matrix typically estimates the phase difference $\Delta\varphi$ between $s_E(t)$ and $s_R(t)$. Such quantities are computed by time-gated measurements (indirect measurement approach)

A number of clarifications and comments about the above listed ToF cameras features are presented in the following sections.

1.4.4.1 Modulation Methods for ToF Depth Cameras

The choice of modulation is the best starting point in order to understand the characteristics of current ToF depth cameras. Indeed, the modulation operation determines the basic transmitter and receiver functions and structure. Although in

principle many modulation types suit ToF depth cameras, in practice, all current commercial ToF depth camera products [2, 5, 7] adopt only one type of CW modulation, namely homodyne amplitude modulation with either a sinusoidal or square wave modulating signal $m_E(t)$. This is because current microelectronic technology solutions for homodyne AM are more mature than others for commercial applications. The advantages of AM modulation, besides its effective implementability by current CMOS solutions, are that it uses a single modulation frequency f_m and does not require a large bandwidth. A major disadvantage is that it offers little defense against multipath and other propagation artifacts (see Chap. 5 of [56]) as will be seen in Chap. 3.

Other modulation types than AM could be usefully employed in ToF depth cameras and their implementation is being actively investigated [56]. Other candidate modulation types include pulse modulation and pseudo-noise modulation. The former, as already mentioned, is the preferred choice for single transmitter and receiver ToF systems. Although in principle it would be equally suitable for matrix ToF sensors, in practice its application is limited by the difficulties associated with implementing effective stop-watch at pixel level within matrix arrangements. Current research approaches this issue in various ways (see Chaps. 2 and 3 of [56]). Pseudonoise modulation would be very effective against multipath, as other applications such as indoor radio localization [21] indicate.

CW modulation itself offers alternatives to homodyne AM, such as heterodyne AM or frequency modulation (FM) with chirp signals. Such properties, although reported in ToF measurement literature, are still problematic for matrix ToF sensor electronics.

The remainder of this section considers the basic characteristics of ToF depth camera transmitters and receivers assuming the underlying modulation is Continuous Wave Amplitude Modulation (CWAM).

1.4.4.2 ToF Depth Camera Transmitter Basics

Lasers and LEDs are the typical choice for the light sources at the transmitter since they are inexpensive and can be easily modulated by signals within the high HF or low VHF bands up to some hundreds of MHz. The LED emissions typically used are in the near infrared (NIR) range, with wavelength around $\lambda_c = 850$ [nm], corresponding to

$$f_c = \frac{c}{\lambda_c} = 3 \times 10^8 \frac{[\text{m}]}{[\text{s}]} \frac{1}{850 \times 10^{-9}\,[\text{m}]} \approx 352\ [\text{THz}]. \tag{1.56}$$

The transmitter illuminates the scene by an optical 2D wavefront signal which, for simplicity, can be modeled as

$$s_E(t) = m_E(t)\ \cos(2\pi f_c t + \varphi_c) \tag{1.57}$$

where $s_E(t)$ denotes the emitter NIR signal structured as the product of a carrier with NIR frequency f_c, of some hundreds of THz, and phase φ_c and a modulating signal $m_E(t)$. Signal $m_E(t)$, in turn, incorporates AM modulation of either sinusoidal or square wave type in current products with frequency f_m, of some tens of MHz, and φ_m. In current products there are two levels of AM modulation, as shown in Fig. 1.18 and further explained in the next section. The first is AM modulation at NIR frequencies concerning the optical signal $s_E(t)$ used to deliver the modulating signal $m_E(t)$ at the receiver. The second is AM modulation in the high HF or low VHF bands embedded in $m_E(t)$, which delivers information related to round-trip time τ to the receiver, either in terms of phase or time lag. This will be seen in detail in Chap. 3.

The current ToF camera NIR emitters are either lasers or LEDs. Since they cannot be integrated, they are typically positioned in configurations mimicking the presence of a single emitter co-positioned with the optical center of the ToF camera. The geometry of the emitters' position is motivated by making the sum of all the emitted NIR signals equivalent to a spherical wave emitted by a single emitter, called *simulated emitter*, placed at the center of the emitters constellation. The LED configuration of the Mesa Imaging SR4000, shown in Fig. 1.17, is an effective example of this concept.

The arrangement of the actual emitters, such as the one of Fig. 1.17, is only an approximation of the non-feasible juxtaposition of single ToF sensor devices with emitter and receiver perfectly co-positioned and it introduces a number of artifacts, including a systematic distance measurement offset that is larger for close scene points than for far scene points.

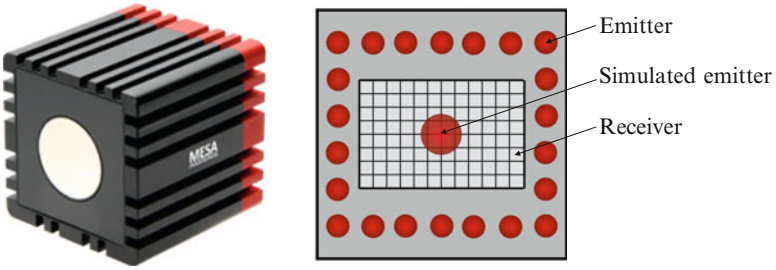

Fig. 1.17 The NIR emitters of the MESA Imaging SR4000: (**a**) the actual depth camera; (**b**) in the scheme the emitters are distributed around the lock-in pixels matrix and mimic a simulated emitter co-positioned with the center of the lock-in pixel matrix

1.4.4.3 ToF Depth Camera Receiver Basics

The heart of ToF camera receivers is a matricial sensor with individual elements, called pixels because of their imaging role, individually and simultaneously capable of independent ToF measurements. In other words each pixel independently computes the delay between the departure of the sent signal $s_E(t)$ and the arrival of the signal $s_R(t)$ back-projected by the scene point P imaged by the pixel. Currently there are three main technological solutions (Chap. 1 of [56]) considered best suited for the realization of such matricial ToF sensors, namely Single-Photon Avalanche Diodes (SPADs) assisted by appropriate processing circuits, standard photo diodes coupled to dedicated circuits and the In-Pixel Photo-Mixing devices. The latter technology includes the lock-in CCD sensor of [38], the photonic mixer device (PMD) [52, 62], and other variations [11, 12]. Chapter 3 will only recall the main characteristics of the In-Pixel Photo-Mixing devices, since so far it is the only one adopted in commercial products [2, 5, 7]. The reader interested in an in-depth treatment of such a technology can refer to [38] and to Chap. 4 of [56]. Current solutions about matricial ToF sensors based on SPADs and standard photo diodes are reported in Chaps. 2 and 3 of [56].

Figure 1.18 offers a system interpretation of the basic functions performed by each pixel of a sensor based on photo-mixing device technology, which are

(a) photoelectric conversion
(b) correlation or fast shutter
(c) signal integration by charge storage on selectable time intervals

The logic of system analysis and circuit effectiveness are opposite. For analysis purposes it is useful to recognize and subdivide the various operations as much as possible. On the contrary, multifunctional components are the typical choice for circuit effectiveness. This chapter and Chap. 3 present ToF depth cameras from a system perspective and the reader must be aware that it does not always coincide with the circuit block description.

Each sensor pixel receives as input the optical NIR signal back-projected by the scene point P imaged by the pixel itself, which can be modeled as

$$s_R(t) = m_R(t)\ \cos(2\pi f_c t + \varphi_c') + n_R(t) \tag{1.58}$$

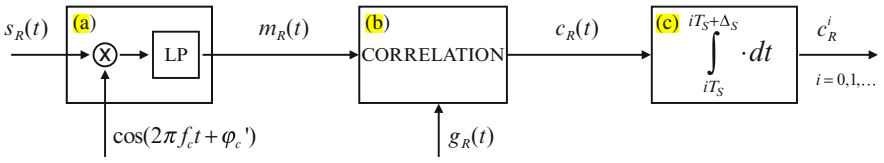

Fig. 1.18 System interpretation of the operation of a single pixel of a sensor based on In-Pixel Photo-Mixing devices technology

where $m_R(t)$ denotes the transformations of the modulating signal $m_E(t)$ actually reaching the receiver, since direct and reflected propagation typically affect some parameters of the transmitted signal $m_E(t)$ (for instance amplitude attenuation is inevitable) and $n_R(t)$ is the background wide-band light noise at the receiver input.[1]

The photoelectric conversion taking place at the pixel in the scheme of Fig. 1.18 is modeled as a standard front-end demodulation stage (a) with a carrier $\cos(2\pi f_c t + \varphi_c')$ at NIR frequency f_c followed by a low pass filter (LP). The input of stage (a) is the optical signal $s_R(t)$ and the output is the baseband electrical signal $m_R(t)$. Stage (a) is, however, only a simple model for the light detection operations taking place at pixel level; details about the actual circuits converting optical signal $s_R(t)$ into electrical signal $m_R(t)$ can be found in [38] and in Chap. 4 of [56]. Stage (b) represents the correlation between baseband signal $m_R(t) + n(t)$ and reference signal $g_R(t)$. Details about the most common types of $g_R(t)$ will be given in Chap. 3. Stage (c) models the charge accumulation process as an integrator operating on time intervals of selectable lengths Δ_S starting at uniformly spaced clock times iT_s, $i = 1, 2, \ldots$ where T_s is the sampling period. Details about indirect ToF measurement methods used to compute phase increment $\Delta\varphi$ or round-trip time delay τ from sampled data will be presented in Chap. 3. Chapter 3 also covers the imaging characteristics of ToF depth cameras and gives an operating paradigm of current commercial ToF depth cameras.

[1] It is worth to notice that the phase φ_c of the carrier at the transmitter side is generally different from the phase φ_c' at the receiver. Both φ_c and φ_c' are usually unknown, especially in the case of a non-coherent process which is the typical practical solution. However, the system does not need to be aware of the values of φ_c and φ_c' and it is inherently robust to the lack of their knowledge.

Chapter 3
Operating Principles of Time-of-Flight Depth Cameras

Time-of-Flight depth cameras (or simply *ToF cameras* in this book) are active sensors capable of acquiring 3D geometry of a framed scene at video rate. Microsoft [6] is currently the major actor in the ToF camera technology arena since it acquired Canesta, a U.S. ToF camera manufacturer, in 2010. This led to the commercialization of the Kinect™ v2 [5]. Commercial products are also available from other manufacturers, such as MESA Imaging [4], PMD Technologies [8] and Sony (which acquired Optrima Softkinetic [9] in 2015). Other companies (e.g., Panasonic [7] and IEE [3]) and research institutions (e.g., CSEM [1] and Fondazione Bruno Kessler [2]) are also actively investigating ToF cameras.

As mentioned in Sect. 1.4.1, this chapter examines the technology adopted in all current commercial ToF camera products. Section 3.1 presents how AM modulation is used within In-Pixel Photo-Mixing devices. Section 3.2 discusses the imaging system supporting ToF sensors. Section 3.3 presents the practical issues driving ToF's performance limits and noise characteristics. Section 3.4 describes how principles introduced in this chapter enter various ToF camera products.

3.1 AM Modulation Within In-Pixel Photo-Mixing Devices

As seen in Sect. 1.4, all the current commercial products adopt homodyne AM modulation with circuitry based on various solutions related to In-Pixel Photo-Mixing devices [17, 23], simply called in-pixel devices in the rest. Figure 3.1 shows a conceptual model of the operation of an homodyne AM transmitter and receiver, which are co-sited in a ToF camera, unlike in typical communication systems. Telecommunication systems convert the signal sent by the transmitter into useful information. In contrast, ToF systems only estimate the round-trip delay of the signal rather than the information encoded inside the signal. The transmitted and received signal from (1.57) and (1.58) reported below for convenience are respectively

P. Zanuttigh et al., *Time-of-Flight and Structured Light Depth Cameras*,
DOI 10.1007/978-3-319-30973-6_3

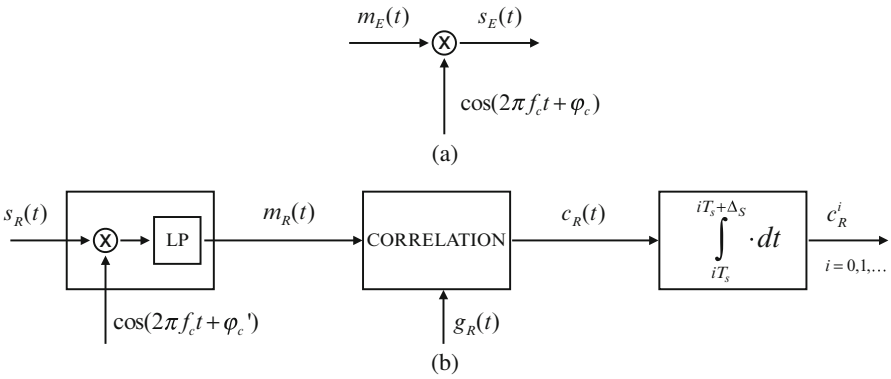

Fig. 3.1 ToF camera: (**a**) transmitter model; (**b**) model of the in-pixel receiver

$$s_E(t) = m_E(t)\cos(2\pi f_c t + \varphi_c) \tag{3.1}$$

and

$$s_R(t) = m_R(t)\cos(2\pi f_c t + \varphi_c') + n_R(t) \tag{3.2}$$

where the carrier frequency f_c corresponds to a NIR wavelength and is in the few hundreds of THz (e.g., $\lambda_c = c/f_c = 860$ [nm] corresponds to $f_c = 348$ [THz]) and $m_E(t)$ is the modulating signal with frequency f_m in the tens of MHz (HF band) and wavelength of a few meters (e.g., $f_m = 16$ [MHz] corresponds to $\lambda_m = 18$ [m]).

It is worth noting that both $s_E(t)$ and $s_R(t)$ are optical signals and that the modulation schemes of Fig. 3.1 are an appropriate description for the operation of the transmitter but not for the photoelectric conversion of the receiver. The schemes of Fig. 3.1 have a conceptual rather than a physical circuit meaning. Indeed, the actual light detection mechanism of the in-pixel devices is such that a baseband voltage signal $m_R(t)$ is generated from the optical input $s_R(t)$, without direct demodulation as in the transmitter side. Therefore the demodulator at the front end of the receiver and the explicit reference to the carrier phase φ_c and φ_c' are only used for modeling simplicity; they hide processes more complex than ones of interest for the purposes of this book.

The electric modulating signal $m_E(t)$ can be either a sine wave of period T_m

$$m_E(t) = A_E[1 + \sin(2\pi f_m t + \varphi_m)] \tag{3.3}$$

with $f_m = 1/T_m$, or a square wave of support $\Delta_m < T_m$ spaced by the modulation period T_m

$$m_E(t) = A_E \sum_{k=0}^{\infty} p(t - kT_m + \varphi_m;\ \Delta_m) \tag{3.4}$$

where

$$p(t;\Delta) = \text{rect}\left(\frac{t-\frac{\Delta}{2}}{\Delta}\right) = \begin{cases} 1 & 0 \le t \le \Delta \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

The pulse $p(t;\Delta)$ in (3.5) is modeled by a rectangle for simplicity, however, this is only a nominal reference signal given the practical difficulty of obtaining sharp rise and fall signals.

At the receiver, after the demodulation of the optical signal $s_R(t)$, the baseband electrical signal $m_R(t)$, of shape similar to that of $m_E(t)$, is correlated with the reference signal $g_R(t)$ with period T_m, obtaining

$$c_R(t) = \int_0^{T_m} m_R(t) g_R(t+t')\, \mathrm{d}t'. \tag{3.6}$$

The signal $c_R(t)$ is sampled according to the "natural sampling" paradigm by the charge accumulator circuit at the back-end of the receiver and can be modeled as a system which at each sampling time iT_s, $i = 0, 1, \ldots$, returns the integration of $c_R(t)$ in the support Δ_S

$$c_R^i = \int_{iT_S}^{iT_S+\Delta_S} c_R(t)\, \mathrm{d}t. \tag{3.7}$$

Clearly for designing ToF camera sensors there is a countless number of combinations for the value of $m_E(t)$, $g_R(t)$ and Δ_S. The two basic situations of sinusoidal and square modulating signal $m_E(t)$ and related choices of $g_R(t)$ and Δ_S will be discussed next.

3.1.1 Sinusoidal Modulation

In the case of sinusoidal modulation, according to the scheme of Fig. 3.1 the ToF camera transmitter modulates the NIR optical carrier by a modulation signal $m_E(t)$ made by a sinusoidal signal of amplitude A_E and frequency f_m, namely

$$m_E(t) = A_E[1 + \sin(2\pi f_m t + \varphi_m)]. \tag{3.8}$$

Signal $m_E(t)$ is reflected back by the scene surface within $s_E(t)$ and travels back towards the receiver ideally co-positioned with the emitter.

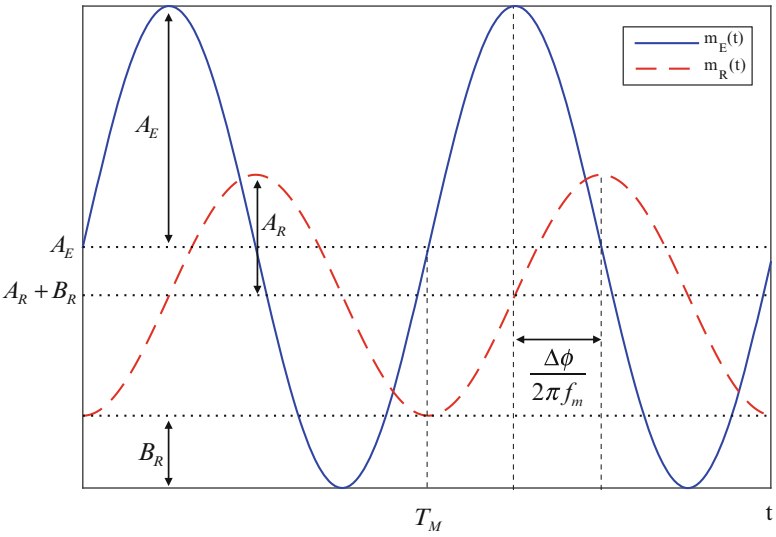

Fig. 3.2 Example of an emitted modulating signal $m_E(t)$ and a received modulating signal $m_R(t)$

The HF/VHF modulating signal reaching the receiver, due to factors such as the energy absorption associated with the reflection, the free-path propagation attenuation (proportional to the square of the distance), and the non-instantaneous propagation of IR optical signals leading to a phase delay $\Delta\varphi$, can be written as

$$\begin{aligned} m_R(t) &= A_R[1 + \sin(2\pi f_m t + \varphi_m + \Delta\varphi)] + B_R \\ &= A_R \sin(2\pi f_m t + \varphi_m + \Delta\varphi) + (A_R + B_R) \end{aligned} \tag{3.9}$$

where A_R is the attenuated amplitude of the received modulating signal and B_R is due to the background light interfering with λ_c and to other artifacts. Figure 3.2 shows an example of emitted and received modulating signal.

For simplicity we will call A_R simply A and $A_R + B_R$ simply $B/2$, obtaining

$$m_R(t) = A \sin(2\pi f_m t + \varphi_m + \Delta\varphi) + \frac{B}{2}. \tag{3.10}$$

Quantity A is called *amplitude*, since it is the amplitude of the useful signal. Quantity B is called *intensity* or *offset*, and it is the sum of the received modulating signal, with a component A_R due to the sinusoidal modulation component at f_m, and an interference component B_R, mostly due to background illumination. It is common to call A and B amplitude and intensity respectively, even though both A and B are signal amplitudes (measured in $[V]$).

If the correlation signal at the receiver is

$$g_R(t) = \frac{2}{T_m}[1 + \cos(2\pi f_m t + \varphi_m)] \tag{3.11}$$

the output of the correlation circuit is

$$
\begin{aligned}
c_R(t) &= \int_0^{T_m} m_R(t')g_R(t'+t)\,\mathrm{d}t' \\
&= \frac{2}{T_m}\int_0^{T_m}\left[A\sin(2\pi f_m t' + \varphi_m + \Delta\varphi) + \frac{B}{2}\right]\left[1+\cos(2\pi f_m(t'+t)+\varphi_m)\right]\,\mathrm{d}t' \\
&= \frac{2}{T_m}\int_0^{T_m} A\sin(2\pi f_m t' + \varphi_m + \Delta\varphi)\,\mathrm{d}t' + \frac{2}{T_m}\int_0^{T_m}\frac{B}{2}\,\mathrm{d}t' + \\
&+ \frac{2}{T_m}\int_0^{T_m} A\sin(2\pi f_m t' + \varphi_m + \Delta\varphi)\cos(2\pi f_m(t'+t)+\varphi_m)\,\mathrm{d}t' + \\
&+ \frac{2}{T_m}\int_0^{T_m}\frac{B}{2}\cos(2\pi f_m(t'+t)+\varphi_m)\,\mathrm{d}t' \\
&= A\sin(\Delta\varphi - 2\pi f_m t) + B.
\end{aligned}
\tag{3.12}
$$

Note that since transmitter and receiver are co-sited, the modulation sinusoidal signal (therefore including its phase φ_m) is directly available at the receiver side.

The unknowns of (3.12) are A, B and $\Delta\varphi$, where A and B are measured in Volts $[V]$ and $\Delta\varphi$ as phase value is a pure number. The most important unknown is $\Delta\varphi$, since as already seen by (1.54) and (1.55), reported below for convenience,

$$
\Delta\varphi = 2\pi f_m \tau = 2\pi f_m \frac{2\rho}{c} \tag{3.13}
$$

it can deliver distance ρ

$$
\rho = \frac{c}{4\pi f_m}\Delta\varphi. \tag{3.14}
$$

Unknowns A and B will be shown later to be important for SNR considerations.

In order to estimate the unknowns A, B and $\Delta\varphi$, $c_R(t)$ must be sampled by an ideal sampler, i.e. with $\Delta_S \to 0$ in (3.7), at least 4 times per modulation period T_m [23], i.e., $T_s = T_m/4$. For instance, if the modulation frequency is 30 [MHz], signal $c_R(t)$ must be sampled at least at 120 [MHz]. Assuming a sampling frequency $F_S = 4f_m$, given the 4 samples per period $c_R^0 = c_R(t=0)$, $c_R^1 = c_R(t=1/F_S)$, $c_R^2 = c_R(t=2/F_S)$ and $c_R^3 = c_R(t=3/F_S)$, the receiver estimates values $\hat{A}$, $\hat{B}$ and $\widehat{\Delta\varphi}$ as

$$(\hat{A}, \hat{B}, \widehat{\Delta\varphi}) = \underset{A,B,\Delta\varphi}{\operatorname{argmin}} \sum_{n=0}^{3} \left\{ c_R^n - \left[A \sin\left(\Delta\varphi - \frac{\pi}{2} n \right) + B \right] \right\}^2 . \tag{3.15}$$

After some algebraic manipulations of (3.15) one obtains

$$\begin{aligned} \hat{A} &= \frac{\sqrt{\left(c_R^0 - c_R^2\right)^2 + \left(c_R^3 - c_R^1\right)^2}}{2} \\ \hat{B} &= \frac{c_R^0 + c_R^1 + c_R^2 + c_R^3}{4} \\ \widehat{\Delta\varphi} &= \operatorname{atan2}\left(c_R^0 - c_R^2, c_R^3 - c_R^1\right). \end{aligned} \tag{3.16}$$

The final distance estimate $\hat{\rho}$ can be obtained from (3.14) as

$$\hat{\rho} = \frac{c}{4\pi f_m} \widehat{\Delta\varphi}. \tag{3.17}$$

If one takes into account that the sampling is not ideal but actually made by a sequence of rectangular pulses of width Δ_S within the standard natural sampling model, the estimates of A, B, and $\widehat{\Delta\varphi}$ in this case become [24]

$$\begin{aligned} \hat{A}' &= \frac{\pi}{T_S \sin\left(\frac{\pi \Delta_S}{T_S}\right)} \hat{A} \\ \hat{B}' &= \frac{\hat{B}}{\Delta_S} \\ \widehat{\Delta\varphi}' &= \widehat{\Delta\varphi} \end{aligned} \tag{3.18}$$

showing that the phase shift $\Delta\varphi$ is independent from the size of the sampling duration Δ_S, that instead affects both the estimate of A and B. A typical value of Δ_S is $\Delta_S = T_m/4 = 1/(4f_m)$.