

# Lock-in Time-of-Flight (ToF) Cameras: A Survey

Sergi Foix, Guillem Alenyà and Carme Torras

**Abstract**—This paper reviews the state-of-the art in the field of lock-in ToF cameras, their advantages, their limitations, the existing calibration methods, and the way they are being used, sometimes in combination with other sensors. Even though lock-in ToF cameras provide neither higher resolution nor larger ambiguity-free range compared to other range map estimation systems, advantages such as registered depth and intensity data at a high frame rate, compact design, low weight and reduced power consumption have motivated their increasing usage in several research areas, such as computer graphics, machine vision and robotics.

**Index Terms**—Lock-in, time-of-flight, calibration.

## I. INTRODUCTION

**T**OF camera is a relatively new type of sensor that delivers 3-dimensional imaging at a high frame rate, simultaneously providing intensity data and range information for every pixel. Despite the number of pixels in the images is still small (i.e.  $176 \times 144$  in Swissranger SR3000 and SR4000 cameras, and  $204 \times 204$  in PMD CamCube camera) and noise in the depth values can not yet be completely removed after calibration, ToF imaging is rapidly showing a great potential in numerous scientific domains.

Due to continuous progress in microelectronics, micro optics and micro technology, the development of ToF cameras has been possible over the last decade. They outperform past technologies at the still difficult and slow task of depth-intensity image matching. Further efforts are being devoted to the optimisation of the cameras themselves. More compact and lighter cameras with better signal-to-noise ratio are being developed, and work continues in order to improve present-day products. New camera models have recently appeared, such as PMD CamCube and Swissranger 4K, and impressive results are expected once researchers start to work extensively with these new models.

Depth-intensity pixel-associated images at a high frame rate without need of mobile components, combined with other technical advantages such as robustness to illumination changes and low weight, make it foreseeable that ToF cameras will replace previous solutions, or alternatively complement other technologies, in many areas of application.

Thus, this paper tries to give a comprehensive overview of the state-of-the-art for the off-the-shelf, most widely used ToF cameras, mainly those relying on demodulation lock-in pixels, describing not only their principles and advantages,

This work has been partially supported by the Spanish Ministry of Science and Innovation under project DPI2008-06022, the MIPRCV Consolider Ingenio 2010 project, and the EU PACO PLUS project FP6-2004-IST-4-27657. S. Foix and G. Alenyà are supported by PhD and postdoctoral fellowships, respectively, from CSIC's JAE program.

The authors are with the Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Llorens i Artigas 4-6, 08028 Barcelona, Spain (e-mails: sfoix, galenya, torras@iri.upc.edu).

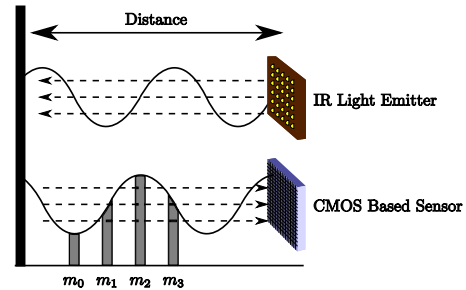


Fig. 1. Distance measurement using the phase offset

but also their current limitations and the research that is in progress. The survey is structured as follows. Section II explains the underlying principle of lock-in ToF cameras. Section III discusses their advantages in comparison with alternative systems. Systematic and non-systematic errors are classified in Sec. IV, where some methods to compensate them are also presented. Section V gives an overview of the current intrinsic and extrinsic calibration methodologies, useful e.g. for sensor fusion. Section VI discusses the main ToF advantages that are being exploited in applications. Finally, conclusions and some unresolved challenges are drawn in Sec. VII.

## II. TOF CAMERA PRINCIPLE

Depth measurements are based on the well-known time-of-flight principle. Time-of-flight can be measured by using either pulsed or continuous-wave (CW) modulation. Although there are ToF cameras based on both technologies, this article will focus on those based on CW modulation, and more precisely on those that use demodulation lock-in pixels [1], no matter whether the demodulation is digital or analog. Lock-in ToF cameras are surveyed because they have been commercially available for more than half a decade and have been extensively used in multiple applications [2], while applications using pulsed-based ToF cameras are still scarce.

Whereas sensors based on discrete pulsed modulation measure the time of a light pulse trip to calculate depth, sensors based on lock-in measure phase differences between emitted and received signals (see Fig. 1). A near-infrared light (NIR), via light-emitting diodes (LED), is emitted by the system and then reflected back to the sensor. Many authors [9]–[12] provide formulations for sinusoidal signals, although other periodic functions can be used. Every pixel on the sensor samples the amount of light reflected by the scene four times at equal intervals for every period ( $m_0$ ,  $m_1$ ,  $m_2$  and  $m_3$  in Fig. 1), which allows for the parallel measurement of its phase



Fig. 2. Current commercial lock-in ToF cameras. (a-b) Mesa Imaging AG<sup>©</sup> [3]. (c) Ifm electronic<sup>©</sup> [4]. (d) CanestaVision<sup>TM</sup> [5]. (e-f) PMD[Vision]<sup>®</sup> [6]. Particularities of each solution include the use by CanestaVision<sup>TM</sup> of square modulated waves [7], the use of a smart pixel - photonic mixer device (PMD) for simultaneous wave sensing and mixing by PMD[Vision]<sup>®</sup> [8], and the addition by Mesa Imaging AG<sup>©</sup> of a coded binary sequence (CBS) modulation for multi-camera operation on SR4000 new models.

$$\varphi = \arctan\left(\frac{m_3 - m_1}{m_0 - m_2}\right), \quad (1)$$

its offset

$$B = \frac{m_0 + m_1 + m_2 + m_3}{4}, \quad (2)$$

and amplitude

$$A = \frac{\sqrt{[m_3 - m_1]^2 + [m_0 - m_2]^2}}{2}. \quad (3)$$

This phase demodulation technique is commonly known as “four-bucket” sampling and it permits to calculate easily the target depth

$$D = L \frac{\varphi}{2\pi}, \quad (4)$$

and the intensity ( $B$ ), whose amplitude ( $A$ ) helps to predict the quality of the measurements. The modulation frequency ( $f_m$ ) of the emitted light determines the *ambiguity-free distance range* of the sensor

$$L = \frac{c}{2f_m}, \quad (5)$$

where  $c$  is the speed of light in vacuum.

Although current off-the-shelf lock-in ToF cameras are based on analog phase demodulation, such as the ones shown in Fig. 2, new prototypes based on digital phase demodulation using single-photon synchronous detection (SPSD) are emerging and claiming better performance [13], [14]. SPSD prototypes use single-photon avalanche diodes (SPADs) as digital single-photon detectors instead of CCD/CMOS photogates used by lock-in pixels. Due to its digital nature, typical analog accumulating diffusion used by previous approaches is simply replaced by a digital counter. Since SPSD does not use any analog processing or analog-to-digital conversion, it is considered virtually noise-free at signal detection and demodulation. Digital and analog approaches share the same mathematical representation shown previously.

From now on and for the sake of simplicity, we will refer to lock-in ToF cameras as just ToF cameras.

### III. DEPTH COMPUTATION AND TOF CAMERAS

Compared to other technologies to obtain scene depth, ToF cameras exhibit some interesting properties:

- 1) Registered dense depth and intensity images
- 2) Complete image acquisition at a high frame rate
- 3) Small, low weight and compact design

4) No mobile parts needed

5) Auto-illumination

Traditionally, depth computation has been carried out by camera and laser-based systems (see [15] for a complete review on laser and other light emitting devices). The following subsections discuss their main disadvantages as compared to ToF cameras.

#### A. Camera-based Systems

In this group we can place methods such as depth-from-focus/defocus/blur, depth-from-motion, depth-from-shape, stereo and structured light triangulation methods [16]. Depth-from-focus, depth-from-motion and depth-from-shape methods are based on focus variation, motion estimation, and shape change determination, respectively. Generally, they produce ambiguities and singularities, and often require using multiple images and solving a correspondence problem, which implies additional temporal, spatial and computational costs. Conversely, depth information obtained with ToF cameras is generally more precise, and it is obtained using only one image.

Triangulation methods can be divided into passive (stereo vision) and active (such as projected structured light methods). Table I shows the main differences between ToF cameras and common stereo and structured light methods.

##### 1) Passive triangulation methods:

Passive triangulation methods require two cameras separated by a baseline that determines a limited working depth range (the higher the needed depth resolution, the larger the needed base). These algorithms have to solve the so-called correspondence problem: determining what pairs of points in the two images are projections of the same 3D point. This is a computationally expensive and complex problem, as stereo vision systems are unable to match corresponding points in homogeneous regions [17]. In contrast, ToF cameras naturally deliver depth and simultaneous intensity data avoiding the correspondence problem, and do not require a baseline in order to operate. In addition, the ambiguity-free range of ToF cameras (usually from 30cm to 7m) can easily be extended by varying the modulation frequency<sup>1</sup>, while that of stereo systems is limited and usually requires changing the baseline, controlled camera motions, or zooming techniques.

<sup>1</sup>In this case, however, some internal parameters would change, making camera recalibration necessary.

TABLE I  
TOF CAMERA VS. TRIANGULATION METHODS.

Differences	ToF cameras	Stereo vision	Structured light
Correspondence problem	No	Yes	Yes
Extrinsic calibration	No, when used alone	Yes	Yes
Auto illumination	Yes	No	Yes
Untextured surfaces	Good performance	Bad performance	Good performance
Depth range	0.3 ÷ 7.5 m.	Base-line dependent	Light-power dependent
Image resolution	Up to 204x204	High resolution. Camera dependent	
Frame rate	Up to 25 fps.	Typically 25 fps. Camera dependent	

## 2) Active triangulation methods:

Contrarily to the preceding methods, active triangulation ones require only one camera together with a structured light emitter that projects one line or a complete set of patterns. Disadvantages here, in comparison with ToF cameras, include partial occlusions that involve missing depth measurements, a need of highly powered and focused light, occasional scanning of the light through the scene which results in low frame rates, and a very controlled light environment that leads to a big restriction in domestic or outdoor robotics applications. Recent approaches [18] solve the partial occlusions problem and the low frame rate by projecting the structured light along the optical path of the camera, and using pattern defocus as a depth estimation technique.

## B. Laser-based Systems

Laser-based systems provide very precise sliced 3D measurements. Albeit they have been successfully applied to solve Simultaneous Localization and Mapping (SLAM) problems [19], difficulties in collision avoidance have been reported due to their 3D reduced field of view [11]. The common solution has been mounting the sensor on a pan-and-tilt unit. This implies row by row sampling, and makes this solution inappropriate for real-time, dynamic scenes, as opposed to ToF cameras. Although high depth range, accuracy and reliability are advantageous in these systems, they are voluminous, heavy, increase the power consumption, and add additional moving parts. ToF cameras, on the contrary, are compact and portable, they do not require the control of mechanical moving parts, thus reducing power consumption, and they do not need row by row sampling, thus reducing image acquisition time.

In sum, ToF cameras have evolved rapidly during the last two decades and, despite their low resolution and low ambiguity-free range, they are already showing great potential in many applications where not very precise but fast 3D image range data acquisition is needed, such as obstacle avoidance [11], [20], pose estimation [21], [22], coarse 3D object reconstruction [23], [24], human body parts recognition and tracking [25]–[27] among others (see [2] for a detailed

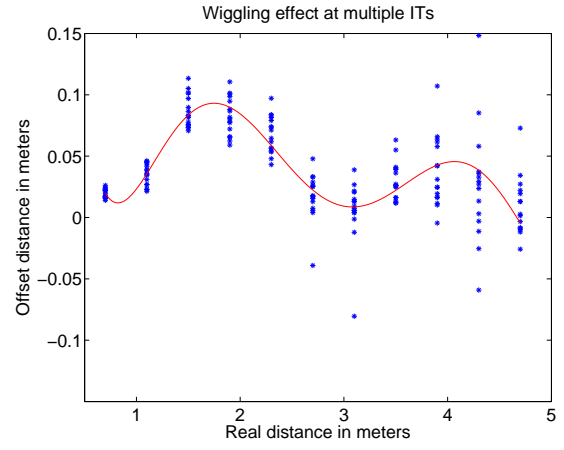


Fig. 3. Depth distortion offset (Wiggling effect). (Blue dots) Measurements captured with a SR3100 ToF camera at multiple integration times (2ms - 32 ms). (Red line) 6 degrees polynomial approximated function.

application review). Although ToF cameras can not be considered yet as a mature sensor compared to other camera-based measuring techniques and other depth sensors, a very promising future can be foreseen.

## IV. DEPTH MEASUREMENT ERRORS AND COMPENSATION

ToF cameras are evolving and a lot of work is devoted to understanding the sources of errors and to minimizing them [28]–[30], as well as to model their effect for camera simulation [31]. In this section we will present a classification and characterisation of the different errors as well as the currently available compensation methods and the quantitative error reduction attained.

Depth measurements with ToF cameras face the appearance of both systematic and non-systematic errors. Generally, systematic errors can be managed by calibration and non-systematic ones by filtering.

### A. Systematic Errors

Five types of systematic errors have been identified:

**Depth distortion** appears as a consequence of the fact that the emitted infrared light can not be generated in practice as theoretically planned (generally sinusoidal) due to irregularities in the modulation process. This type of error produces an offset that depends only on the measured depth for each pixel. Usually, the error plotted against the distance follows a sinusoidal shape<sup>2</sup> (see Fig. 3). This error is sometimes referred to as *wiggling* or *circular error*.

This type of error depends on the measured depth distance, and it can be addressed by comparing camera depth measurements with a reference ground truth distance, or by means of an optimisation process that models the error from multiple relative measurements. While the first approach has the disadvantage of needing an additional sensor in order to acquire the reference distance, i.e. high accuracy track line as

<sup>2</sup>This has been explained by means of perturbations on the measured signal phase caused by wrapping of odd harmonics contained in the emitted reference signal [32].

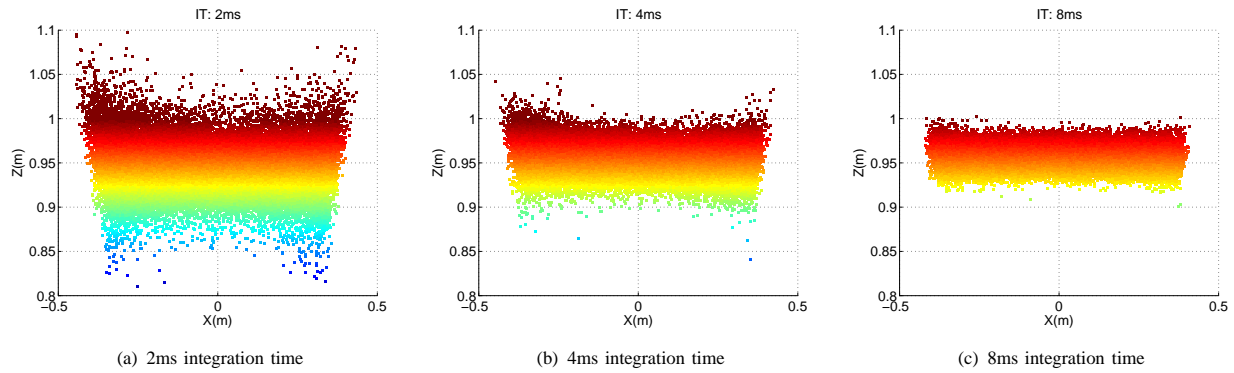


Fig. 4. Depth-colored 3D point cloud view of a white wall at a constant distance of 1 meter. Each figure shows the X/Z view at different integration times (a) 2ms (b) 4ms (c) 8ms. A systematic depth offset can be observed dependent on the integration time. Amplitude-related errors also appear on the boundary edges in Fig. (a) and (b) due to low amplitudes.

in [33], [34] or a calibrated color camera as in [35]–[37], the second approach has the disadvantage of being only suitable in a limited operating range [38], [39]. Applications in robot navigation, localization and mapping should be better suited by the first approach in order to ensure the most reliable acquisition depth range, while for applications such as object modelling the second one will be more adequate.

There are several approaches to encode the error data. A Look-up Table (LUT) has been proposed [33] that stores the depth errors depending on the measured depth distance using only one central pixel. The representation of depth errors has a sinusoidal shape, so a B-Spline can be used to store these values in a more compact form [34], [39]. Alternatively, a polynomial function has been also used, although, on the contrary of B-Splines, undesirable border effects can appear outside the interpolation range. The degree of the polynomial that models the depth error has to be chosen depending on the required measurement depth range. This detail has been left undetermined in some works [38], [40]. In the general case a 6-degree function is adequate [41]. For small ranges (1-2 meters) only a portion of the function has to be represented and a simple 3-degree polynomial function suffices [37]. In general this is a time-consuming process as several distances have to be measured.

In a different way, Lindner *et al.* [36] present a new demodulation algorithm applicable to the PMD camera. They use the fact that the modulated signal is composed of a sinusoidal with a rectangular reference signal. The combination of both does not provide more accurate depth images, but can be used to better determine the *depth distortion* errors.

**Integration-time-related error.** Integration time (IT) can be selected by the user. It has been observed that for the same scene different IT cause different depth values in the entire scene (see Fig. 4). The main reason for this effect is still a subject of investigation.

IT affects the range of depths that the camera is sensing with more precision. This has the effect of changing the former calibration solutions. A lot of works do not mention this source of error and usually it is not reported whether it is explicitly taken into account or not. We note that some cameras have an auto mode for the IT. Although it may seem as a good feature, its use makes the calibration methods hard to apply.

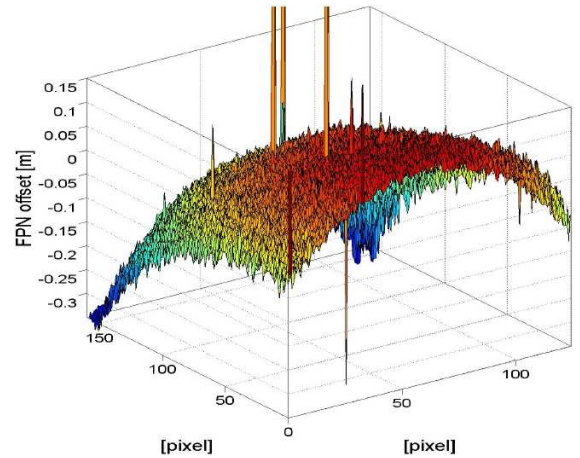


Fig. 5. Depth-colored Fixed Pattern Noise (FPN) offset per pixel. Figure extracted from Kahlmann *et al.* [33].

There are two main strategies to solve this problem. The first one is to choose one integration time value, perform the calibration for the rest of the errors with this value, and never change it [34], [39], [41], [42]. This is possible when the range of depths is small.

For the second one the idea is to repeat the *depth distortion* calibration process for different integration times [33], [35], [43] and then apply the corresponding correction values taking into account the current IT.

**Built-in pixel-related errors** arise from two main sources. On the one hand, errors due to different material properties in CMOS-gates. This produces a constant pixel-related distance offset, leading to different depths measured in two neighbour pixels corresponding to the same real depth. On the other hand, there are latency-related offset errors due to the capacitor charge time delay during the signal correlation process. This can be observed as a rotation of the image plane, i.e. a perpendicular flat surface is viewed with a wrong orientation.

Such errors are related to the position of the pixel in the sensor array. A common representation of this error is a Fixed Pattern Noise (FPN) table (see Fig. 5) that is obtained by comparing the computed depths with a reference distance [33]. However, with this procedure the contribution of *amplitude-*

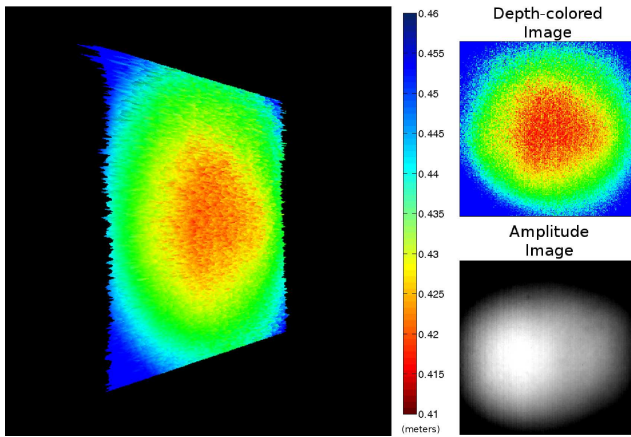


Fig. 6. Depth-colored amplitude-related errors. Depth image of a flat wall at 0.43 meters. Depth overestimation can be observed due to low illumination (borders of the image).

related errors cannot be separated and FPN accounts for both error sources.

Neighbouring pixel errors are small, and can be considered negligible. In that case, only the error from the rotation of the image plane has to be modeled. A compact representation is a function depending on the row and column position of the pixel [39]. Sometimes the parameters of this function are specified inside the polynomials that define the *Depth distortion error* and they are solved jointly in the same minimisation process [37], [38]. We note that the Swissranger camera manufacturer provides such a FPN matrix in the calibration file [3]. However, some authors prefer to recalibrate for this error effects when using this camera [44].

**Amplitude-related errors** occur due to low or overexposed reflected amplitudes. Depth accuracy is highly related to the amount of incident light as it can be deduced from (1) and (4). The higher the reflected amplitudes, the higher the depth accuracy. Low amplitude appears more often in the border of the image as the emitted light power is lower than in the center, leading to overestimating depth (see Fig. 6). Contrarily, when the object is too close to the camera or integration time has been chosen too high, saturation can appear and depth measurements will not be valid.

This type of error arises due to three main causes. First, systematic non-uniform NIR LEDs illumination causes depth misreadings at pixels distant from the image center. A second cause is low illumination for scenes with objects at different distances. And third, differences in object reflectivities cause different depth measurements for pixels at the same constant distance. Non-specular materials retain energy and modify consequently the reflected light phase, depending on their refraction indices.

Low amplitude errors can be avoided easily by filtering pixels with lower amplitude than a threshold [30], [38], but this solution may discard a large region of the image. Additionally, the threshold may need to vary when moving. An earlier solution was to increase the overall depth accuracy in scenes with nearby and distant objects by combining depth measurements from two range images with different exposure settings [7].

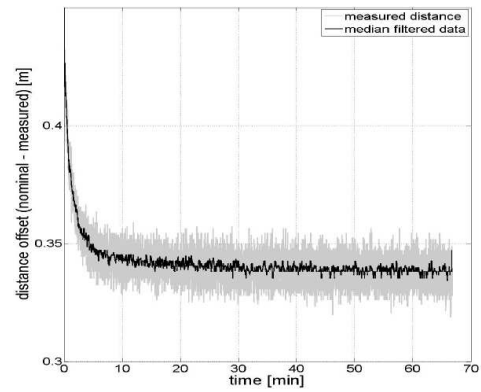


Fig. 7. Temperature-related error. Figure extracted from Kahlmann *et al.* [33].

Nevertheless, the preferred solution to this problem is still filtering.

The second error source, over-exposition, can be detected if the raw time measures of the camera can be accessed [45]. This is not possible in Swissranger cameras. However, the new SR4k camera provides a confidence value that can be used for this purpose.

The third amplitude-related error cause, different object reflectivities, is quite difficult to handle. A common solution is to reproduce the *Built-in pixel-related errors* and *Amplitude-related errors* calibration methods for different reflective surfaces [33] and store all the median values and use them as a look-up table depending on intensity values. As the amplitude plays an important role, the combination of the ToF camera with a color camera has been also suggested [35] to better measure intensity.

In fact, depth and amplitude measurements are highly correlated. Guomundsson *et al.* [29] propose to improve depth ones by simply subtracting the standardised amplitude inverse ( $1/A$ ), where standardised means taking away the mean and dividing by the standard deviation. Taking into account the same correlation principle, Oprisescu *et al.* [42] provide two methods to correct inaccuracies of depth and amplitude by using information based on the other. This approach is continued by Falie *et al.* [46], who provide a noise model for phenomena analysis [47] that predicts distance error at a pixel as a function of the amplitude at that pixel and the distance itself.

**Temperature-related errors** happen because internal camera temperature affects depth processing, explaining why some cameras include an internal fan. Depth values suffer from a drift in the whole image until the temperature of the camera is stabilised.

Impact of internal and external temperature on distance measurements is studied in [33], [45] as a result of the high response of the semiconductor materials to changes in temperature. A SwissRanger camera SR-2 showed an overestimation in measured distances when the sensor started working, and when operating at higher temperatures (see Fig. 7). The next generation of the camera tried to palliate this problem by incorporating a fan to stabilize the temperature. The general strategy to palliate temperature depth errors is to switch

on the camera and let it take a stable working temperature before calibrating it and using it. While some past approaches recommended to wait around 4 minutes for the SR-3000 [48], new studies with new camera models (SR-4000) recommend to wait up to 40 minutes [49]. New models did not get worse but more accurate ( $\pm 1\text{cm}$ ), and higher waiting time is considered necessary to ensure stabilization.

### B. Quantitative Error Analysis

The preceding section has described how several authors have applied different calibration methods in order to reduce each systematic error. In order to better understand the amount of improvement achieved in each work, Table II summarizes the error reductions attained by the main compensation approaches found in literature. Although a comparison between the different methods is a difficult task due to the variety of cameras being used, some conclusions can be drawn. Special attention has to be paid to the results obtained by Fuchs *et al.* [39] and Kahlmann *et al.* [33], since they managed to reduce the overall standard error to less than 3 mm. The reason why these two approaches achieve such a good performance is because they reduce the three main error sources: depth distortion, built-in pixel and integration-time-related errors. Rapp [45] quantified the proportion of reduction attributable to each of these systematic errors. The temperature-related error was not considered in his work and amplitude was just used for pixel validation purposes. Three different ToF cameras (Effector O3D, PMD 19k and SR-3000) were used in his experiments, all of them leading to similar error reduction results. Approximately 40 % of the overall error reduction was found to be attributable to the compensation of depth distortion, 33.3 % to the mitigation of the integration-time-related error, and only 6.6 % to correction of the built-in pixel-related error. Further evidence for these results can be observed in Table II, where the worst overall error reduction is obtained by a method that does not treat the integration-time error, and deals only partially with the built-in pixel-related error.

### C. Non-systematic Errors

Four non-systematic errors can also be identified in depth measurements with ToF cameras, the occurrence of the last three being unpredictable.

**Signal-to-noise ratio** distortion appears in scenes not uniformly illuminated. Low illuminated areas are more susceptible to noise than high illuminated ones. This type of error is highly dependent on the amplitude, the IT parametrisation and the depth uniformity of the scene. Non-uniform depth over the scene can lead to low-amplitude areas (far objects) that will be highly affected by noise.

Signal-to-noise ratio can be improved by several means. Low-amplitude filtering can be easily used and corrupted readings can be simply removed [40] or a more sophisticated procedure can actively decide the optimal IT depending on the desired areas [38]. Other approaches try to minimise noise effects by computing the average of those readings

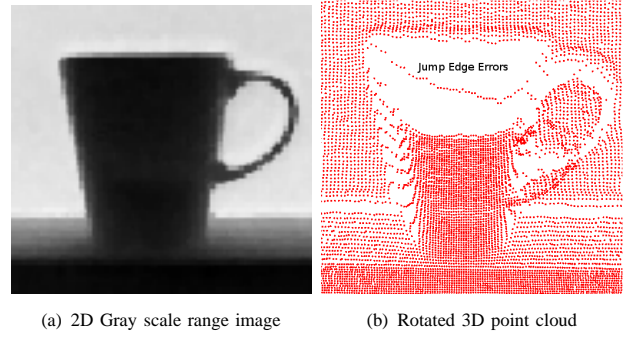


Fig. 8. (a) 2D Gray scale range image of a mug. (b) Rotated 3D point cloud view. False depth readings appear at the edges between foreground and background objects due to the integration of the reflected light of both surfaces in the corresponding pixels.

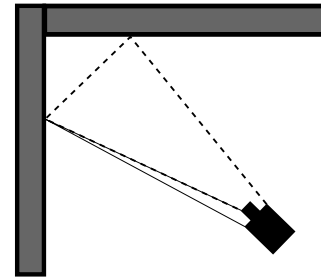


Fig. 9. Multiple light reception due to concavities in the scene.

and surpassing a certain accuracy threshold based on pixels variance [29], [45], [47].

**Multiple light reception** errors appear due to the interference of multiple light reflections captured at each sensor's pixel. These multiple light reflections depend on the low lateral sensor resolution and the geometric shape of the objects in the scene.

Multiple light reception errors are mainly due to the presence of surface edges (jump edges) and object concavities (see Fig. 8 and Fig. 9, respectively). On the one hand, jump edge errors are generally removed by comparing the angle of incidence of neighboring pixels [28], [38], [50]. On the other hand, it is still an open question how to deal with multiple reflections originated by concavities [29].

**Light scattering** effect arises due to multiple light reflexions between the camera lens and its sensor (see Fig. 10). This effect produces a depth underestimation over the affected pixels, because of the energy gain produced by its neighbouring pixel reflections [50]. Errors due to light scattering are only relevant when nearby objects are present in the scene. The closer an object, the higher the interference [51].

Light scattering effects have been minimised following two approaches. Firstly, [53] suggested selecting an optimal IT in order to minimise saturation problems and remove scattering-affected pixels using a filter based on the combination of amplitude and intensity values. And secondly, a compensation method based on blind deconvolution was proposed based on

TABLE II  
ERROR REDUCTIONS ATTAINED BY DIFFERENT COMPENSATION APPROACHES FROM LITERATURE

Article	Camera	Compensation method used				Overall remaining error	
		Depth Distortion	Amplitude	Built-in pixel	Integration Time	Mean	Std. Dev.
Fuchs <i>et al.</i> [39]	O3D100	B-Splines/multiple range amplitudes	-	Pan and tilt coef.	Unique IT/range	$\pm 1.2 \text{ mm.}$	$\pm 5.7 \text{ mm.}$
Kahlmann <i>et al.</i> [33]	SR-2	Look-up table	-	Fixed pattern noise	Look-up table	$\pm 1.0 \text{ mm.}$	$\pm 10.0 \text{ mm.}$
Lidner <i>et al.</i> [34]	PMD (64x48)	B-Spline	-	Fixed pattern noise	Constant IT	$\pm 10.0 \text{ mm.}$	$\pm 3.0 \text{ mm.}$
Radmer <i>et al.</i> [43]	PMD	B-Spline	Look-up table	-	Look-up table	$\pm 10.0 \text{ mm.}$	$\pm 25.0 \text{ mm.}$
Kim <i>et al.</i> [41]	SR3000	6-degree polynomial	Radial pattern	-	Unique IT/range	$\pm 13.6 \text{ mm.}$	$\pm 8.8 \text{ mm.}$
Schiller <i>et al.</i> [37]	PMD (64x48)	3-degree polynomial	-	Pan and tilt coef.	-	$\pm 50.0 \text{ mm.}$	$\pm 100.0 \text{ mm.}$

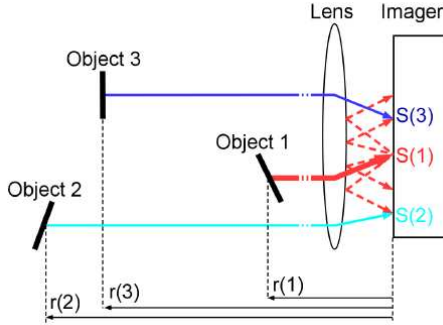
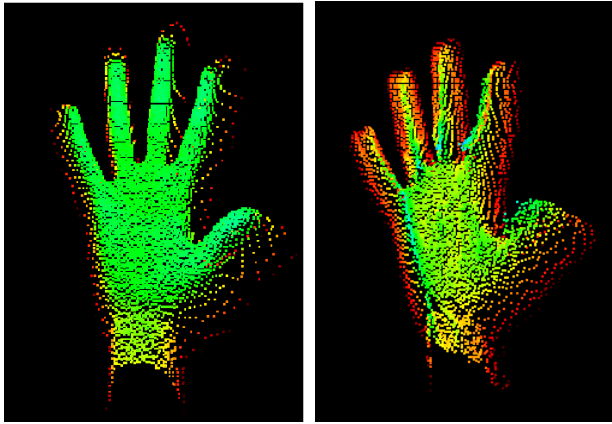


Fig. 10. Light scattering. Figure extracted from Mure-Dubois, J. and Hügli, H. [52].



(a) Depth colored point cloud of a static hand (b) Depth colored point cloud of a hand moving to the left

Fig. 11. Motion blurring appears due to a quick movement of the camera or the objects in the scene during the integration time. Its effect can be seen by comparing subfigures (a) and (b).

a mathematical model [52]. Because empiric parametrisation was still needed, further research must be carried out to optimally mitigate its effect. Instead of trying to detect and decrease the scattering effect, some researchers point out that new sensor materials with lower reflectivity will arise in the future that will make scattering negligible [50].

**Motion blurring**, present when traditional cameras are used in dynamic environments, appears also with ToF cameras. This is due to the physical motion of the objects or the camera during the integration time used for sampling (see Fig. 11).

Motion blurring errors can be classified in two different

types of artifacts depending on whether their appearance is due to lateral or axial motion. In [54] a combination of a conventional 2D image sensor and a PMD camera is used in order to detect lateral motion artifacts by means of a classical 2D image edge detector. Instead of discarding the corrupted data, the authors present two possible correction approaches. On the one hand, an average of positionally weighted neighbouring pixels is recommended, and on the other hand, after a phase sampling analysis of images, 2 phase depth computation can be used instead of the common 4 phase algorithm. Another approach, this time for solving both lateral and axial motion blurring consecutively, is presented in [55]. Lateral motion artifacts are identified first, by estimating optic flow from some pre-processed phase-sampled images, and afterwards, axial motion artifacts are removed using both an axial motion estimation approach and a theoretical model for axial motion deviation errors.

## V. INTRINSIC AND EXTRINSIC CALIBRATION

While the error compensation procedures described in the preceding section are specific of ToF cameras (with the exception of those dealing with motion blurring), the calibration techniques explained next are mostly similar to those used for traditional cameras that provide depth through stereo or structured light.

### A. Intrinsic

3D data are not delivered in the same manner by all different camera models, i.e. SR3 delivers cartesian 3D points, while PMD delivers the absolute distance to the center of projection of the optical system. The pinhole model and the intrinsic calibration parameters [40], [45] are needed to compute cartesian 3D points from depth points. This is mandatory for PMD cameras. For SwissRanger cameras, some authors recompute depth maps with the obtained focal depth to improve depth precision. As shown below, intrinsic parameters are also useful when ToF camera images have to be combined with other sensors.

Intrinsic camera parameters have often been obtained by applying classical calibration procedures based on intensity images and calibration patterns [34], [38], [56]. Alternatively, using an array of infrared LEDs has been also proposed to improve the localization of the calibration pattern [33]. However, the characteristic low resolution of these cameras leads

to a poor localization of the calibration pattern characteristics and obtained calibration parameters are usually erroneous.

In sum, intrinsic TOF camera calibration has to be improved by using also the depth information provided by the camera [40], and it is usually included in the extrinsic calibration methods explained below.

### B. Extrinsic

The extrinsic parameters encode the coordinate system transformation from 3D world coordinates to 3D camera coordinates. They are useful in two scenarios: first when a ToF camera has to be referenced to an external device, i.e. a robot arm; second, when ToF images have to be combined with other sensor data.

For the first scenario, let us imagine a camera mounted on the end-effector of a robotic arm. As will be presented in Sec. VI, this is a common configuration e.g. for object modelling [24]. The transformation between the robot coordinate system and the camera coordinate system is the so-called hand-eye calibration. To compute it, the image measures have to be used. Recently, a calibration algorithm that integrates depth calibration and hand-eye calibration has been proposed [38].

For the second scenario, lasers and one or various color cameras are rigidly mounted with one or various ToF cameras. Although extracting precise extrinsic parameters is highly recommended, some simplifications can be performed when sensors are mounted in particular configurations [57]–[59], and, depending on the application, an inaccurate calibration is sometimes sufficient [60]. Combination of a ToF camera and a 2D laser scanner is common, and extrinsic parameters can be obtained with a specialised pattern [61].

One of the most used sensor systems is to combine ToF cameras with color cameras. Traditionally the extrinsic calibration has been addressed by considering the intensity image of the ToF camera and using classical stereo calibration algorithms [56], [62]–[64]. Unfortunately, due to the low resolution of the sensor, this approach suffers from the same problems as the ones presented for intrinsic calibration.

Hence the idea is to take advantage of depth information when calibrating, either coming from the ToF camera itself, when used together with a single intensity camera, or also derived from triangulation when used in combination with a stereo rig or structured light. The aim in the monocular setting is to backproject the points using the available depth data to refine the calibration, while in the latter case, 3D-3D correspondences can be used to estimate the calibration between a ToF camera and a stereo rig. This has been applied for a small depth range (only 400mm) [65].

However, stereo is not strictly required. Once a color camera has been calibrated with a known pattern, reconstruction of the calibration poses is possible, and this can be used to find better extrinsic parameters [37]. A software to calibrate one or multiple color cameras with a TOF camera using this principle is available [66]. This algorithm also includes a depth calibration model that represents the depth deviation as a polynomial function, similar to [38].

Multiple ToF cameras can be used to observe the same scene. Different modulation frequencies for each ToF camera should be used to avoid interference problems between them [41]. Obviously, the resulting images can be also registered with color cameras.

## VI. EXPLOITATION OF TOF CAMERA ADVANTAGES

The distinctive characteristics of ToF cameras have proved to provide important advantages in several fields. After reviewing their main applications to date, fully discussed in Foix *et al.* [2], our conclusion is that the most exploited feature of ToF cameras is their capability of delivering complete scene depth maps at high frame rate without the need of moving parts. Moreover, foreground/background segmentation methods based on depth information are quite straightforward, so ToF images are used in many applications requiring them. A good characteristic is that geometric invariants as well as metric constraints can be naturally used with ToF depth images. ToF cameras are also used satisfactorily in human environments because they are eye-safe and permit avoiding physical contact and dedicated markers or hardware.

The depth-intensity image pair is also often used, exploiting the fact that both images are delivered already registered. In applications where the reduced resolution of a ToF camera is critical, it is complemented with other sensors, usually color cameras. Once the extrinsic parameters of the coordinate transformation between a color camera and a ToF camera have been obtained, data fusion is possible. The easy part is to find the correspondences between them and put color to the depth image, but more can be done. Due to the difference in resolution, between each pair of neighbouring points in the ToF image there are several points on the color image. As a consequence, these points can be interpolated to obtain a dense depth map [56].

Some of the reviewed works do not apply any calibration method to rectify the depth images. We believe that this explains several of the errors and inaccuracies reported in some experiments, and that with proper calibration better results can be obtained. We note that ToF technology is evolving and depth correction methods for this type of sensor are still subject to investigation.

Albeit ToF cameras are increasingly being used in more applications everyday due to their distinctive features, previous technologies such as stereo vision, structured light or lidar systems are still leading the 3D depth acquisition field. Mainly because of their higher precision and higher acquisition range.

## VII. CONCLUSIONS

Over the last years, the performance of ToF cameras has improved significantly; errors have been minimised and higher resolution and frame rates are being obtained. Although ToF cameras cannot yet attain the depth accuracy offered by other types of sensors such as laser scanners, structured light or stereo vision systems, plenty of research demonstrates that their distinctive features make this type of sensors a suitable solution or alternative in many applications.

Advantages of this type of sensors are multiple, as demonstrated in the previous sections: they are compact and portable, easing movement; they make data extraction simpler and quicker, reducing power consumption and computational time; and they offer a combination of images that show great potential in the development of data feature extraction, registration, reconstruction, planning and optimisation algorithms, among other positive characteristics. Thus, ToF cameras prove to be especially adequate for real-time applications and, in particular, for automatic acquisition of 3D models requiring sensor movement and on-line mathematical calculation.

Finally, some broad challenges need to be mentioned. First, resolution is still generally low for ToF cameras, despite some efforts have already led to better resolutions as explained above. Second, short integration times contribute to obtain a strong noise ratio, and high integration times can result in pixel saturation [67]. Although some algorithms dealing with this problem have already been proposed, more research is needed in this direction. Third, an important issue for ToF cameras is the wrapping effect, a consequence of the periodicity of the modulated signal. Distances to objects that differ  $360^\circ$  in phase are indistinguishable. Use of multiple modulated frequencies can be a solution here, or lowering the modulation frequency since it would increase the unambiguous metric range.

Other concerns include ambient light noise, motion artifacts and high-reflectivity surfaces in the scene. Ambient light may contain unwanted light of the same wavelength as that of the ToF light source which may cause false measurements in the sensor. Frequency-based filters can be used in order to minimise this effect. Motion artifacts are errors caused by receiving light from different depths at the same time due to object motion in the scene. This type of errors are mostly observed around the edges of the moving object and can be attenuated by either increasing the frame rate, or by correction using motion estimation. Finally, errors due to the coexistence of low-reflective and high-reflective objects (mirroring effect) can be addressed by combining multiple exposure settings.

## REFERENCES

- [1] T. Spirig, P. Seitz, O. Vietze, and F. Heitger, "The lock-in CCD - two-dimensional synchronous detection of light," *IEEE J. Quantum Electron.*, vol. 31, no. 9, pp. 1705–1708, Sept. 1995.
- [2] S. Foix, G. Alenyà, and C. Torras, "Exploitation of Time-of-Flight (ToF) cameras," IRI, UPC, Tech. Rep. IRI-DT-10-07, 2010.
- [3] "SR-Cameras, <http://www.mesa-imaging.ch>," MESA Imaging AG, 2009.
- [4] "3D vision sensors, <http://www.ifm.com>," Ifm electronic gmbh, 2009.
- [5] "Camera Modules, <http://www.canesta.com>," CanestaVision<sup>TM</sup>, 2009.
- [6] "PMD-Cameras, <http://www.pmdtec.com>," PMDTechnologies GmbH, 2009.
- [7] S. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor - system description, issues and solutions," in *Proc. IEEE CVPR Workshops*, Washington, D. C., June 2004, pp. 35–35.
- [8] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, and T. Ringbeck, "Smart pixel - photonic mixer device (PMD) / New System Concept of a 3D-imaging-on-a-chip," in *Proc. 5th Int. Conf. Mechatronics and Machine Vision in Practice*, Nanjing, Sep. 1998, pp. 259–264.
- [9] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE J. Quantum Electron.*, vol. 37, no. 3, pp. 390–397, Mar. 2001.
- [10] T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger<sup>TM</sup>)," in *Proc. of SPIE*, vol. 5249, no. 1, St. Etienne, Feb. 2004, pp. 534–545.
- [11] J. Weingarten, G. Gruener, and R. Siegwart, "A state-of-the-art 3D sensor for robot navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 3, Sendai, Sep. 2004, pp. 2155–2160.
- [12] A. Kolb, E. Barth, and R. Koch, "ToF-sensors: New dimensions for realism and interactivity," in *Proc. IEEE CVPR Workshops*, vol. 1-3, Anchorage, June 2008, pp. 1518–1523.
- [13] C. Niclass, C. Favi, T. Kluter, F. Monnier, and E. Charbon, "Single-photon synchronous detection," *IEEE J. Solid-State Circuits*, vol. 44, no. 7, pp. 1977–1989, 2009.
- [14] D. Stoppa, L. Pancheri, M. Scandiuozzo, L. Gonzo, G.-F. D. Betta, and A. Simoni, "A CMOS 3-D imager based on single photon avalanche diode," *IEEE Trans. Circuits Syst.*, vol. 54, no. 1, pp. 4–12, 2007.
- [15] F. Blais, "Review of 20 years of range sensor development," *J. Electronic Imag.*, vol. 13, no. 1, pp. 231–243, 2004.
- [16] E. Stoykova, A. Alatan, P. Benzie, N. Grammalidis, S. Malassiotis, J. Ostermann, S. Piek, V. Sainov, C. Theobalt, T. Thevar, and X. Zabulis, "3-D time-varying scene capture technologies: a survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1568–1586, Nov. 2007.
- [17] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge: Cambridge University Press, 2004.
- [18] F. Moreno-Noguer, P. N. Belhumeur, and S. K. Nayar, "Active refocusing of images and videos," *ACM T. Graphics*, vol. 26, no. 3, July 2007.
- [19] J. Andrade-Cetto and A. Sanfeliu, *Environment Learning for Indoor Mobile Robots. A Stochastic State Estimation Approach to Simultaneous Localization and Map Building*, ser. Springer Tracts in Advanced Robotics. Springer, 2006, vol. 23.
- [20] A. Prusak, O. Melnychuk, H. Roth, I. Schiller, and R. Koch, "Pose estimation and map building with a time-of-flight camera for robot navigation," *Int. J. Int. Syst. Tech. App.*, vol. 5, no. 3-4, pp. 355–364, 2008.
- [21] S. May, D. Droschel, D. Holz, C. Wiesen, and S. Fuchs, "3D pose estimation and mapping with time-of-flight cameras," in *Proc. IEEE/RSJ IROS Workshop on 3D-Mapping*, Nice, Sep. 2008.
- [22] S. Hussmann and T. Liepert, "Robot vision system based on a 3D-ToF camera," in *Proc. 24th IEEE Instrum. and Meas. Tech. Conf.*, vol. 1-5, Warsaw, May 2007, pp. 1405–1409.
- [23] B. Dellen, G. Alenyà, S. Foix, and C. Torras, "3D object reconstruction from Swissranger sensors data using a spring-mass model," in *Proc. 4th Int. Conf. Comput. Vision Theory and Applications*, vol. 2, Lisbon, Feb. 2009, pp. 368–372.
- [24] S. Foix, G. Alenyà, J. Andrade-Cetto, and C. Torras, "Object modeling using a ToF camera under an uncertainty reduction approach," in *Proc. IEEE Int. Conf. Robot. Automat.*, Anchorage, May 2010, pp. 1306–1312.
- [25] X. Liu and K. Fujimura, "Hand gesture recognition using depth data," in *Proc. 6th IEEE Int. Conf. Automatic Face Gesture Recog.*, Seoul, May 2004, pp. 529–534.
- [26] S. B. Gokturk and C. Tomasi, "3D head tracking based on recognition and interpolation using a time-of-flight depth sensor," in *Proc. 18th IEEE Conf. Comput. Vision Pattern Recog.*, vol. 2, Washington DC, June-July 2004, pp. 211–217.
- [27] K. Nanda, H.; Fujimura, "Visual tracking using depth data," in *Proc. IEEE CVPR Workshops*, vol. 3, Washington, D. C., June 2004, pp. 37–37.
- [28] W. Karel, P. Dorninger, and N. Pfeifer, "In situ determination of range camera quality parameters by segmentation," in *Proc. 8th Int. Conf. on Opt. 3D Meas. Tech.*, Zurich, July 2007, pp. 109 – 116.
- [29] S. A. Guomundsson, H. Aanaes, and R. Larsen, "Environmental effects on measurement uncertainties of time-of-flight cameras," in *Proc. Int. Sym. Signals, Circuits and Systems*, vol. 1-2, Lasi, July 2007, pp. 113–116.
- [30] C. A. Weyer, K. H. Bae, K. Lim, and D. D. Lichti, "Extensive metric performance evaluation of a 3D range camera," in *Proc. ISPRS Conf.*, vol. 37, Beijing, Jul. 2008, pp. 939–944.
- [31] M. Keller and A. Kolb, "Real-time simulation of time-of-flight sensors," *Sim. Mod. Pract. Theory*, vol. 17, pp. 967–978, May. 2009.
- [32] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. dissertation, Univ. Siegen, Germany, 2000.
- [33] T. Kahlmann, F. Remondino, and H. Ingsand, "Calibration for increased accuracy of the range imaging camera Swissranger<sup>TM</sup>," in *ISPRS Commission V Symposium*, Dresden, Sep. 2006, pp. 136–141.
- [34] M. Lindner and A. Kolb, "Lateral and depth calibration of PMD-distance Sensors," in *Proc. 2nd Int. Sym. Visual Computing*, vol. 4292, Lake Tahoe, Nov. 2006, pp. 524–533.
- [35] —, "Calibration of the intensity-related distance error of the PMD ToF-camera," in *Proc. SPIE*, vol. 6764, no. 67640W, Boston, Sept. 2007.

- [36] M. Lindner, A. Kolb, and T. Ringbeck, "New insights into the calibration of ToF-sensors," in *Proc. 22nd IEEE Conf. Comput. Vision Pattern Recog.*, vol. 1-3, Anchorage, June 2008, pp. 1603–1607.
- [37] I. Schiller, C. Beder, and R. Koch, "Calibration of a PMD camera using a planar calibration object together with a multi-camera setup," in *Proc. ISPRS Conf.*, vol. 37, Part B3a, Beijing, Jul. 2008, pp. 297–302.
- [38] S. Fuchs and S. May, "Calibration and registration for precise surface reconstruction with time of flight cameras," *Int. J. Int. Syst. Tech. App.*, vol. 5, no. 3-4, pp. 274–284, 2008.
- [39] S. Fuchs and G. Hirzinger, "Extrinsic and depth calibration of ToF-cameras," in *Proc. 22nd IEEE Conf. Comput. Vision Pattern Recog.*, vol. 1-12, Anchorage, June 2008, pp. 3777–3782.
- [40] M. Wiedemann, M. Sauer, F. Driewer, and K. Schilling, "Analysis and characterization of the PMD camera for application in mobile robotics," in *Proc. 17th IFAC World Congress on Aut. Control*, Seoul, July 2008, pp. 13 689–13 694.
- [41] Y. M. Kim, D. Chan, C. Theobalt, and S. Thrun, "Design and calibration of a multi-view ToF sensor fusion system," in *Proc. IEEE CVPR Workshops*, vol. 1-3, Anchorage, June 2008, pp. 1524–1530.
- [42] S. Oprisescu, D. Falie, M. Ciuc, and V. Buzuloiu, "Measurements with ToF cameras and their necessary corrections," in *Proc. Int. Sym. Signals, Circuits and Systems*, vol. 1-2, Lasi, July 2007, pp. 221–224.
- [43] J. Radmer, P. Fuste, H. Schmidt, and J. Kruger, "Incident light related distance error study and calibration of the PMD-range imaging camera," in *Proc. IEEE CVPR Workshops*, vol. 1-3, Anchorage, June 2008, pp. 1579–1584.
- [44] S. May, S. Fuchs, D. Droschel, D. Holz, and A. Nuechter, "Robust 3D-mapping with time-of-flight cameras," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Saint Louis, Oct. 2009, pp. 1673–1678.
- [45] H. Rapp, "Experimental and theoretical investigation of correlating ToF-camera systems," Master's thesis, University of Heidelberg, Sept. 2007.
- [46] D. Falie and V. Buzuloiu, "Distance errors correction for the time-of-flight (ToF) cameras," in *Proc. IEEE Int. Workshop Imag. Syst. Tech.*, Chania, Sep. 2008, pp. 123–126.
- [47] —, "Noise characteristics of 3D time-of-flight cameras," in *Proc. Int. Sym. Signals, Circuits and Systems*, vol. 1-2, Lasi, July 2007, pp. 229–232.
- [48] O. Steiger, J. Felder, and S. Weiss, "Calibration of time-of-flight range imaging cameras," in *Proc. IEEE Int. Conf. Image Process.*, San Diego, Oct. 2008, pp. 1968–1971.
- [49] F. Chiabrand, R. Chiabrand, D. Piatti, and F. Rinaudo, "Sensors for 3D imaging: metric evaluation and calibration of a CCD/CMOS time-of-flight camera," *Sensors*, vol. 9, no. 12, pp. 10 080–10 096, 2009.
- [50] T. Kahlmann and H. Ingensand, "Calibration and development for increased accuracy of 3D range imaging cameras," *J. Appl. Geodesy*, vol. 2, no. 1, pp. 1–11, 2008.
- [51] W. Karel, "Integrated range camera calibration using image sequences from hand-held operation," in *Proc. ISPRS Conf.*, vol. 37, Beijing, Jul. 2008, pp. 945–952.
- [52] J. Mure-Dubois and H. Hügli, "Real-time scattering compensation for time-of-flight camera," in *Proc. 5th Int. Conf. Comput. Vision Systems*, Bielefeld, March 2007.
- [53] S. May, B. Werner, H. Surmann, and K. Pervolz, "3D time-of-flight cameras for mobile robotics," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 1-12, Beijing, Oct. 2006, pp. 790–795.
- [54] O. Lottner, A. Sluiter, K. Hartmann, and W. Weihs, "Movement artefacts in range images of time-of-flight cameras," in *Proc. Int. Sym. Signals, Circuits and Systems*, vol. 1-2, Lasi, July 2007, pp. 117–120.
- [55] M. Lindner and A. Kolb, "Compensation of motion artifacts for time-of-flight cameras," in *Proc. Dynamic 3D Vision Workshop*, vol. 5742, Jena, Sep. 2009, pp. 16–27.
- [56] M. Lindner, A. Kolb, and K. Hartmann, "Data-fusion of PMD-based distance-information and high-resolution RGB-images," in *Proc. Int. Sym. Signals, Circuits and Systems*, vol. 1-2, Lasi, July 2007, pp. 121–124.
- [57] K. D. Kuhnert and M. Stommel, "Fusion of stereo-camera and PMD-camera data for real-time suited precise 3D environment reconstruction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 1-12, Beijing, Oct. 2006, pp. 4780–4785.
- [58] F. Wallhoff, M. Russ, G. Rigoll, J. Gobel, and H. Diehl, "Improved image segmentation using photonic mixer devices," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1-7, San Antonio, Sep. 2007, pp. 2849–2852.
- [59] S. A. Guomundsson, R. Larsen, H. Aanæs, M. Pardas, and J. R. Casas, "ToF imaging in smart room environments towards improved people tracking," in *Proc. IEEE CVPR Workshops*, vol. 1-3, Anchorage, June 2008, pp. 1486–1491.
- [60] U. Hahne and M. Alexa, "Combining time-of-flight depth and stereo images without accurate extrinsic calibration," *Int. J. Int. Syst. Tech. App.*, vol. 5, no. 3-4, pp. 325–333, 2008.
- [61] F. Yuan, A. Swadzba, R. Philippsen, O. Engin, M. Hanheide, , and S. Wachsmuth, "Laser-based navigation enhanced with 3D time of flight data," in *Proc. IEEE Int. Conf. Robot. Automat.*, Kobe, May 2009, pp. 2844–2850.
- [62] J. Fischer, B. Huhle, and A. Schilling, "Using time-of-flight range data for occlusion handling in augmented reality," in *Proc. Eurographics Sym. Virtual Environments*, Sep. 2007, pp. 109–116.
- [63] S. A. Guomundsson, H. Aanæs, and R. Larsen, "Fusion of stereo vision and time-of-flight imaging for improved 3D estimation," *Int. J. Int. Syst. Tech. App.*, vol. 5, no. 3-4, pp. 425–433, 2008.
- [64] T. Grundmann, Z. Xue, J. Kuehnle, R. Eidenberger, S. Ruehl, A. Verl, R. D. Zoellner, J. M. Zoellner, and R. Dillmann, "Integration of 6D object localization and obstacle detection for collision free robotic manipulation," in *Proc. IEEE/SICE Int. Sym. System Integration*, Nagoya, Dec. 2008, pp. 66–71.
- [65] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of time-of-flight depth and stereo for high accuracy depth maps," in *Proc. 22nd IEEE Conf. Comput. Vision Pattern Recog.*, vol. 1-12, Anchorage, June 2008, pp. 3262–3269.
- [66] "http://mip.informatik.uni-kiel.de," 2009.
- [67] P. Einramhof, S. Olu, and M. Vincze, "Experimental evaluation of state of the art 3D-sensors for mobile robot navigation," in *Proc. in the 31st Austrian Association for Pattern Recog. Workshop (OAGM)*, Krumbach, May 2007.



**Sergi Foix** is a JAE predoc fellow at the Spanish Scientific Research Council (CSIC). He received M.Sc. degrees in Intelligent Systems and Automatic Control and Robotics from the University of Sunderland and the Technical University of Catalonia (UPC), respectively. His fields of interest are next-best-view planning for active vision, 3D object modelling and pose estimation for robot manipulation, and 3D feature extraction.



**Guillem Alenyà** is a JAE postdoc fellow at the Spanish Scientific Research Council (CSIC). He received Ph.D. degree in 2007 from the Technical University of Catalonia (UPC) with a work on mobile robot navigation using active contours. He participates in numerous scientific and technological projects involving image understanding and robot localization. Dr. Alenyà was a Marie Curie fellow of the European Commission in the period 2002-2004. His areas of interest include robot active vision, egomotion estimation, and planning for active modelling.



**Carme Torras** (M'07) is Research Professor at the Spanish Scientific Research Council (CSIC). She received M.Sc. degrees in Mathematics and Computer Science from the Universitat de Barcelona and the University of Massachusetts, respectively, and a Ph.D. degree in Computer Science from the Technical University of Catalonia (UPC). Prof. Torras has published five books and about two hundred papers in the areas of robotics, computer vision, and neurocomputing. She has been local project leader of several European projects, such as "Planning RObot Motion" (PROMotion), "Robot Control based on Neural Network Systems" (CONNY), "Self-organization and Analogical Modelling using Subsymbolic Computing" (SUBSYM), "Behavioural Learning: Sensing and Acting" (B-LEARN), the 6th framework IP project "Perception, Action and COgnition through Learning of Object-Action Complexes" (PACO-PLUS), and the ongoing 7th framework STREP projects "GARdeNing with a Cognitive System" (GARNICS) and "Intelligent observation and execution of Actions and manipulations" (IntellAct).