IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 12, 2018, accepted December 26, 2018, date of publication January 10, 2019, date of current version February 6, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2891693

Recent Advances in 3D Data Acquisition and Processing by Time-of-Flight Camera

YU HE[1] AND SHENGYONG CHEN[1,2], (Senior Member, IEEE)
[1]College of Computer Science and Technology, Zhejiang University of Technology, Zhejiang 310023, China
[2]College of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China

Corresponding author: Shengyong Chen (sy@ieee.org)

This work was supported in part by the National Key R&D Program of China under Grant SQ2018YFB130084, and in part by the National Natural Science Foundation of China under Grant U1509207.

ABSTRACT Three-dimensional (3D) data acquisition and real-time processing is a critical issue in an artificial vision system. The developing time-of-flight (TOF) camera as a real-time vision sensor for obtaining depth images has now received wide attention, due to its great potential in many areas, such as 3D perception, computer vision, robot navigation, human–machine interaction, augmented reality, and so on. This paper survey advances in TOF imaging technology mainly from the last decade. We focus only on recent progress of overcoming limitations such as systematic errors, object boundary ambiguity, multipath error, phase wrapping, and motion blur, and address the theoretical principles and future research trends as well.

INDEX TERMS Vision sensor, time of flight, depth image, sensing device, computer vision, 3D vision.

I. INTRODUCTION

New computer vision system mostly needs to solve big data processing problems in real-time, especially when the system works in a dynamic environment and billion bytes of three-dimensional (3D) spatial data generated in every minute. Time of flight (TOF) is a novel method for 3D imaging, which shares the similar principle with 3D laser sensor. The major merit is obtaining the depth information of the whole scene simultaneously, instead of point wise scanning, which is suitable for dynamic scene. TOF cameras have a similar imaging process compared to usual cameras. For example, both of them are consist of light source, optical components, controlling circuit as well as the processing circuit and functional units. However, the key different component is the TOF chip, which implements active light detection, that is, placing a front-end lens before it to collect light, every pixel of TOF chip records the phase shift between the incident light and the reflected light. According to the phase change between the incident and reflected signal, the distance can be measured. Moreover, two more shutters are integrated into TOF chip for sampling the reflected light at different time point [1].

TOF camera produces a depth image, in which every pixel encodes the distance between itself and the corresponding point in the scene [2]. This technology has been applied in many fields for research and engineering solutions. Some practical applications of this sensing modality include robot navigation [3], [4], collision and obstacle detection for robot-assisted surgery [5], 3D reconstruction [6], measurement of structural deformation [7], [8], simultaneous localization and mapping (SLAM) [9], [10], human-computer interaction [11], 3D television (3DTV) [12], plant phenotype [13], [14], debris monitoring [15], etc.

The TOF measurement principle is to calculate the phase delay of the infrared light (IR) reflected from object surface. Of course, there are many other ways of 3D measurement. For instance, Kinect also projects an IR structured pattern onto object surfaces and determine the distance by visual triangulation. This kind of devices shares many applications with TOF cameras [16]–[21]. However, the attractions of TOF camera include its low cost, good accuracy, reliability, single-shot and video-rate depth data collection, and compact size of its hardware.

TOF camera is a 3D vision sensor which modulates its signal of light-emitting diodes (LEDs) and detects the phase delay of the reflected signal with a CMOS/CCD imaging chip at each pixel. The camera can also obtain the amplitude image of the scene. The range of camera can be calculated by the equation $S = c/(2f)$, where S is the depth, f is the modulation frequency and c is the light speed. A 3D point cloud can be derived from the collocated range and the reflected signal amplitude images [22].

TOF camera has a unique sensing architecture, and the raw depth data contains both systematic and nonsystematic bias,

2169-3536 © 2019 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

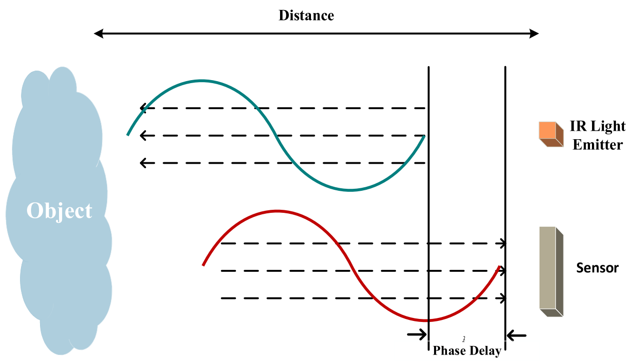


FIGURE 1. TOF imaging principle: measuring the phase difference between the emitted and detected IR signals TOF imaging.

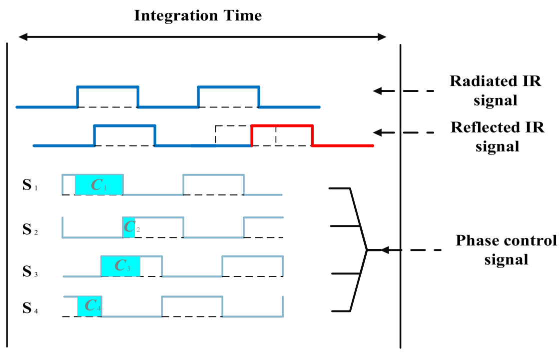


FIGURE 2. Depth is calculated by the phase difference between the emitted and detected IR signals.

which need further process for robust depth information [23]. Moreover, TOF camera suffers various deficiencies in practice, such as low spatial resolution, low depth precision, bias caused by geometric, radiometric and illumination variation. The measurements would be ambiguous when the measured scene is beyond a certain range. The maximum range without phase wrapping is determined by the signal frequency. Motion blur is another critical problem, which is caused by either object or camera motion. Because of the TOF sensing architecture, the motion blur of depth images shows some special characteristics [24].

To deal with the abovementioned challenges, many novel methods have been proposed, which may be categorized into different groups in terms of input depth data. Section II briefly addresses principles, advantages and limitations of these methods. Section III introduces the most widely used TOF cameras and their applications. Section IV presents the deficiencies of TOF cameras, including both systematic and non-systematic errors. Section V gives some usual methods to correct errors. Finally, further research trends and some conclusions are drawn separately in Section VI and Section VII.

II. TOF IMAGING PRINCIPLES

The principle of TOF imaging is illustrated in Fig. 1 [25]. An IR light is emitted from an LED to the object in the scene, and it is reflected by the surface and detected by the TOF sensor. The distance from the sensor to the object can be determined according to the phase difference or time delay between the emitted and reflected IR lights. The phase change is calculated by the relation of the four control signals and the electric charge values. Each phase control signal has a phase delay of 90 degrees, as shown in Fig. 2. The four signals find the collection of electrons from the reflected IR and estimate the phase difference φ as

$$\varphi = \arctan\left(\frac{C_3 - C_4}{C_1 - C_2}\right), \quad (1)$$

$$D = \frac{c}{2f}\frac{\varphi}{2\pi} \quad (2)$$

$$d_{\max} = \frac{c}{2f} \quad (3)$$

where C_1 to C_4 in (1) represent the electric charge amount of four control signals [23], [26], [27]. Then the distance D is determined by (2), where c is the speed and f is the frequency of the light signal. d_{max} constrains the maximum distance of measurement without phase wrapping, which is, of course, determined only by the frequency f. The phase wrapping will be further discussed in Section IV.

III. DEVICES AND APPLICATIONS

A. PROPERTIES AND ADVANTAGES

TOF cameras have been found with many interesting properties which differ from other technologies in obtaining depth images, e.g. (1) video-rate image acquisition, (2) compact and fixed structure, (3) illumination adaptation, (4) self-registration of dense depth data and color image, (5) small and light weight [23]. Compared to conventional cameras, TOF camera exhibits many advantages, including:

- Achieves richer location relationship between objects with depth information [1], [27].
- Depth information also can be competent to traditional applications like image segmentation, tags, recognition, tracking, etc. [28], [29].
- Through further processing, depth information can be used for 3D-reconstruction and other homologous applications [3], [30]–[33]
- Able to quickly applied in target recognition and tracking [18], [34]–[36]
- Costs of main accessories are relatively cheap, including CCD and common LED, and popularizing the production and utilizing the products are in all probability [23]
- With the aid of the characteristics of CMOS, can get a large amount of data and information, the judgment for complex object is very effective [37]
- Without scanning equipment supporting work [25]

B. TYPICAL PRODUCTS

At present, the mainstream TOF camera manufactures include PMD, MESA, Optrima, and Microsoft [38]. MESA is TOF camera manufacturer who is currently the largest provider in the field of scientific research. The main feature is its compactness. PMD products are able to detect multiple range, which can be used both in indoor and outdoor.

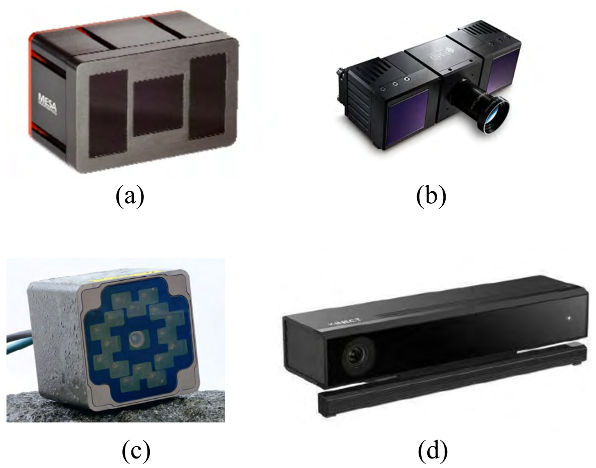


FIGURE 3. Typical TOF cameras, (a) SR4500 (176 × 144), (b) CamCube 3.0 (200 × 200), (c) E-SERIES 70 (160 × 120), (d) Kinect V2 (512 × 424).

Optrima's and Microsoft's products are mainly designed for family and entertainment applications, and therefore their prices are relatively low. Here Four latest professional TOF cameras from MESA, PMD, Fotonic and Microsoft are shown in Fig. 3 and introduced as follows.

1) MESA SR4500.

SR4500 TOF cameras produce depth maps and amplitude images at the resolution 176 × 144. Every pixel is quantified to a 16-bit floating-point word and each pitch represents 40μm size. The amplitude image reflects the detected IR light and forms the depth-map, which provides three dimensional coordinates corresponding to the image pixels. The maximum frame rate is 30 fps and the field of view (FOV) is 44 × 35 degrees. The operating range of SR4500 is up to 9.0 m depending on the modulation frequency.

2) PMD CAMCUBE 3.0.

This type of TOF camera is a state of the art depth camera with a high resolution, high frame rate, superior ambient light suppression and a flexible and modular design. The resolution is 200 × 200 pixels and the frame rate is 40 fps, while the work range with standard settings is 0.3-7.0 m. The FOV of this device is 40 × 40 degrees.

3) FOTONIC E-SERIES 70.

The greatest benefits with the E-SERIES 70 are the very low motion artifacts and high frame rate. These features make it effective for using in dynamic environment, e.g. tracking of moving objects. The maximum frame rate is 58 fps and the pixel array size is 160 × 120. Since the modulation frequency is 15 MHZ, the measurement range is 0.15-10 m. The FOV of E-SERIES 70 is 70×53 degrees.

4) KINECT V2.

Kinect V2 is the new type depth sensor from Microsoft which can produce color image at the resolution 1920 × 1080 and depth image at resolution 512 × 424. The frame rate of color and depth sensor are equal to 30 fps. The range of detection is 0.5-4.5 m and the FOV come up to 70 × 60 degrees.

C. APPLICATIONS

The properties and advantages of TOF cameras make them wide applications in practice, e.g.

1) LOGISTICS INDUSTRY

In the process of logistics, TOF camera can obtain the volume of the packages quickly and track their locations, optimize the packing and shipping [8], [39]–[41].

2) SECURITY AND MONITORING

(1) In some public venues, the security department will count people to ensure the number of people is less than limit [42], [43]. (2) By counting the stream of people or the complicated traffic system, we can complete the statistical analysis design of the security system [44]–[47]. (3) Object detection helps us to monitor in sensitive areas [8], [14], [48]. (4) Machine vision: industry locating, guidance and the volume forecast [35].

3) 3D RECONSTRUCTION

According to the depth image collected by the camera, we can build 3D maps of indoor and outdoor scene and reconstruct objects in the scene [14], [30], [49]–[54]. Even some special environment of 3D reconstruction, such as underwater environment [32], [55], 56].

4) ROBOT

TOF camera provides good obstacle avoidance information for automatic driving [57]–[60]. In industrial production, cameras guide the robots on the installation, quality control and raw material selection [61], [62].

5) MEDICAL AND BIOLOGICAL

In the field of biology and medicine, TOF camera can be applied to many fields, e.g. foot orthopaedics modeling, patient activities/sate monitoring, surgery assistance and 3D facial recognition [34], [63]–[65].

6) INTERACTIVE ENTERTAINMENT

The application in interactive entertainment includes posture detection, expression recognition and human-computer interaction [28], [34], [63], [66]–[71].

IV. EXISTING LIMITATIONS

Although TOF camera takes many advantages, its special sensing architecture still causes a series of problems in applications. A raw depth image taken by TOF camera is still at low spatial resolution. Most of cameras can only get 20-40 pixels in a frame. Both systematic and nonsystematic biases are existing in the resulted data and depth precision is still limited. Errors can be caused by many geometric and radiometric variations. The accuracy is also affected by the limited power of emitted IR. The amplitude of the detected IR value often varies in terms of color and natural material of the object. Depth ambiguities caused by scenic structure and motion blur caused by object and camera movement need to be solved. What's more, calibration is another critical problem of TOF camera. In this section, we especially concern the depth noises and error sources in applications.

The depth image taken by TOF camera suffers from some systematic errors, including integration time (IT) error, amplitude ambiguity, temperature error, depth distortion error and built-in pixel error [1], [23], [72], [73]. On the other hand, there exist many nonsystematic errors in applications, including light scattering error, multipath error, object boundary ambiguity, multipath error, phase wrapping, and motion blur.

A. INTEGRATION TIME

As shown in Fig. 4, the longer IT offers the higher signal-to-noise ratio (SNR) [74]. The IT is related to the frame rate as described in [1], and affects the range of depths and the precision of TOF cameras. This source of IT error is rarely mentioned in majority of existing works, and it is unclear whether this error is explicitly taken into account. Some cameras use an auto mode for the IT. It can be a good feature sometime for non-professional users, but it also makes the calibration not applicable.

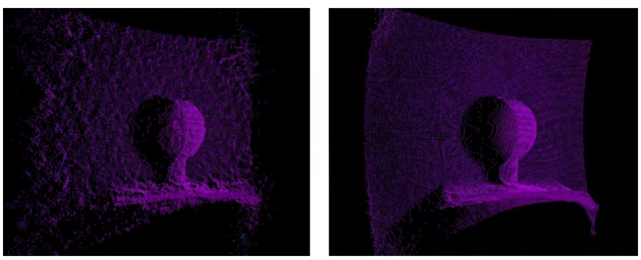


FIGURE 4. Integration time error.

B. AMPLITUDE AMBIGUITY

Several reasons may cause amplitude ambiguity errors, e.g. non-uniform LED radiation, non-uniform scenic illumination due to objects at varying distances, and non-uniform reflection property of the object surfaces. As shown in Fig. 5, the 3D points of same depth have different IR amplitudes of the reflected signal depending on the object color [1], [23].

C. TEMPERATURE DRIFT

The working temperature in a TOF camera affects its depth processing and causes reference drift or systematic errors. Some cameras have an internal fan installed in the device to keep stable temperature. Otherwise, the calculated values will have a drift in the whole 3D image. Impact of internal and external temperature on distance measurement can be found in [75], for understanding the response of semiconductor materials according to temperature changes.

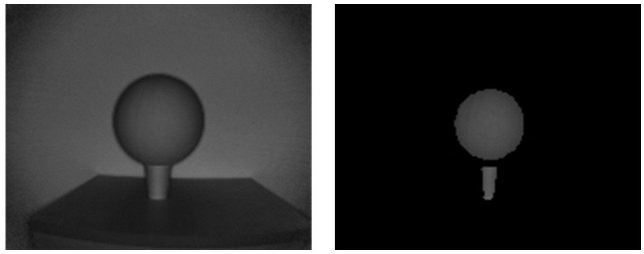


FIGURE 5. Amplitude ambiguity.

D. DEPTH DISTORTION

This type of systematic error appears in TOF cameras because the emitted IR cannot be generated perfectly due to irregularities in the modulation process [76]. Such errors produce a depth offset depending on the distance at each point on the surface. Those error sometimes appears as wiggling or circular error.

E. ELEMENT VARIATION IN SENSOR ARRAY

There are small built-in variations at sensor elements and causes error at pixels. Due to the variation of material properties in each CMOS/CCD element, the depth measured in two adjacent pixels might be different even though they correspond to the same distance in the real scene. Another is the latency related error caused by the time delay of capacitor charge in signal correlation. The built-in element variation causes pixel-dependent errors. These errors are usually small and may be neglected, but we still need to consider the error produced by rotation of the image [23].

F. LIGHT SCATTERING

There are artifacts in the depth image which caused by light scattering [36], [77]. As shown and marked in Fig. 6, due to the low sensitivity of the TOF sensor, IR saturation in a place causes depth distortion in other parts in the depth image.

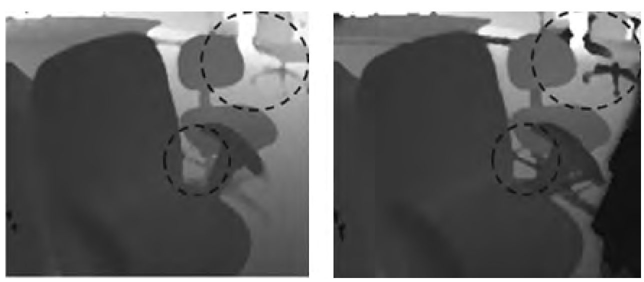


FIGURE 6. Light scattering.

G. BOUNDARY AMBIGUITY

Object boundary ambiguity is a serious problem for reconstructing 3D scenes [78]. The pixels near the boundaries goes to background and foreground simultaneously [79], which results in some distortion in the obtained 3D image, as the example shown in Fig. 7.

FIGURE 7. **Object boundary ambiguity.**

H. MULTIPATH DISTURBANCE

Due to light reflection, multiple reflected IR signals might be superposed for depth calculation in a sensor pixel [80]. This multipath disturbance is serious in some complex places, e.g. the concave corners (Fig. 8).

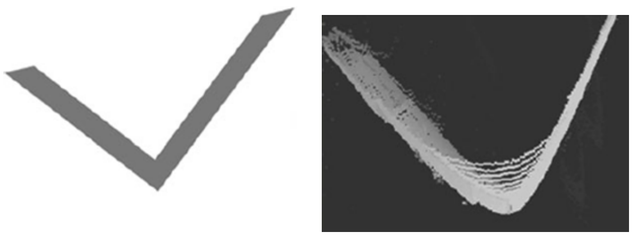


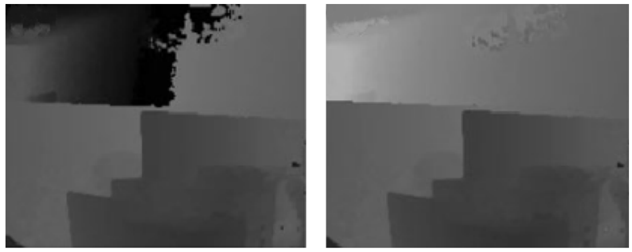
FIGURE 8. **Multipath error around the concave corner.**

I. PHASE WRAPPING

The maximum detecting range of TOF camera is determined by the signal periodicity, or i.e. the signal frequency. Beyond the maximum range, phase wrapping will occur and the measured values are confounded. According to the principle of TOF camera, the detected IR light is gated using its internal reference signals. The function for distance measurement is arctangent of phase φ in the detected signal. Due to the period of 2π, the value has ambiguity at phase $\varphi + 2\pi$, for all $n \geq 0$. Therefore, a modulation frequency f corresponds to a maximum range $d_{\max}$ determined by (3). For a position beyond $d_{\max}$, the actual distance might be $d + nd_{\max}$. This phase wrapping problem requires the algorithm to determine the unknown n, or called phase unwrapping. For example, Fig. 9a shows a typical wrapped TOF depth map obtained by SR4000. The unwrapped version of this map is shown as Fig. 9b.

J. MOTION BLUR

Cameras often produce motion blur in the captured images due to either camera motion or object motion. The blur is a regional or global error and causes image degradation. The corresponding deblurring methods are not yet well explored in practice for TOF cameras. For live 3D data acquisition,



(a) (b)

FIGURE 9. **Phase wrapping.**

a TOF camera has also to deal with this issue. However, the motion blur from TOF images is very different from other color cameras due to the special sensing principle. One special characteristic is that the TOF motion blur often shows overshoot or undershoot, which can be found in the regions between foreground and background transition. Accordingly, the blur results in higher or lower depth value calculated than other depth values near foreground and background.

To study the motion blur in a TOF image, we have to explain the IT. Since the depth value is obtained by measuring the phase delay between the emitted and detected IR signals, the IT has to be sufficient for collecting electric charge to find the phase delay. During the integration period any camera or object motion will cause imaging blur. If the process of collecting electric charge C_1 to C_4 to calculate depth (1) occurs $n$cycles during the IT, the calculation can be repeated n times to increase the SNR.

$$\varphi = \arctan\left(\frac{nC_3 - nC_4}{nC_1 - nC_2}\right) \tag{4}$$

where C1 to C4 are corresponding electric charges of the four control signals S1 to S4 in Fig. 2. The depth calculation (4) assumes that the IR signal comes from a specific 3D point in the scene during the IT. If there is any motion in the period, the resulted depth will be corrupted. As shown in Fig. 10, the red point is a same pixel in the TOF camera. Because of the object motion, the red point comes from different places in the scene during the IT. Thus, the motion causes the false depth calculated from the points around the moving area.

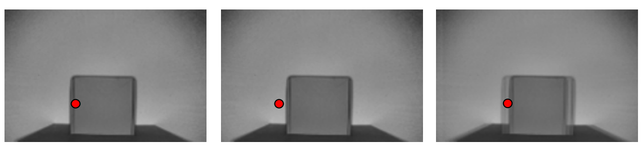


FIGURE 10. **TOF depth motion blur.**

V. METHODS FOR DATA CORRECTION

The enhancement of TOF data is an important issue in the practical applications [73]. In order to deal with the encountered challenges and to reduce the errors, researchers have proposed some useful methods to solve the urgent issues in the TOF camera applications, mostly on system error reductions and nonsystem error reductions.

A. SYSTEMATIC ERROR REDUCTIONS

There are many systematic errors described in Section IV. Here we summarize some available methods attempted to reduction to these errors. Table 1 lists a summation of them which are based on single camera data [23]. There are three main error sources of the systematic error, i.e. built-in pixel, depth distortion, and time integration errors.

TABLE 1. Approaches to systematic error reduction.

Tasks	Method	Representative
Integration time error	Unique integration time range	Fuchs et al. [85]
Built in pixel error	Pan and tilt coef	
Depth distortion	B-splines	
Integration time error	Look up table	Kahlmann et al.[75]
Depth distortion	Look up table	
Built in pixel error	Fixed pattern	
Integration time error	Constant integration time	Lidner et al. [81]
Depth distortion	B-splines	
Built in pixel error	Fixed pattern	
Integration time error	Look up table	Radmer et al. [86]
Depth distortion	B-splines	
Integration time error	Unique integration time range	Kim et al. [87]
Depth distortion	6 degrees polynomial	
Depth distortion	3 degrees polynomial	Schiller et al. [88]
Built in pixel error	Pan and tilt coef	

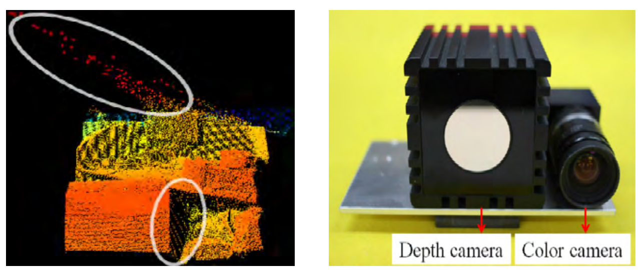
Several recent contributions are supplemented mainly for dealing with the two error sources of the systematic error. The first is depth distortion, which appears when the emitted IR light cannot be practically generated as planned because of irregularities in modulation. Two approaches are addressed for this error. One is comparing depth measurements with a reference ground truth [81], [82], and the other is estimating the error from optimization of multiple measurements [83]. Hussmann *et al.* [84] presented a modulation method based on sine waves for minimizing the wiggling error. A noise distribution model is also derived which predicts the performance of the modulation method in real time for depth images. Applications in 3D reconstruction and modeling should be suited by this type of approach. Fuchs and Hirzinger [85] and Kahlmann *et al.* [75] also studied the systematic error sources and the overall error is reduced to below 3mm. More intensive investigation has been done by González-Ortega *et al.* [23], who had summarized some classic approaches to reduction of the typical systematic errors.

Another systematic error is amplitude ambiguity error, which occurs due to low or overexposed reflected amplitudes. According to the different causes, there are three categories of solutions. A threshold in the amplitude filter can cut the low amplitude errors [83], and the over exposition error can be corrected by accessing the raw measure time of the camera . The third cause of amplitude error is non-uniform surface reflectivity. Generally, a calibration process can be used to analyze different reflective object surfaces [81]. As a novel 3D system, a multiple camera system (MCS) such as combination of TOF cameras with color cameras, has been applied to detect amplitude errors [89].

B. NONSYSTEMATIC ERROR REDUCTION

1) DEPTH DATA DENOISING

The captured depth data from a TOF camera is often starkly contaminated by noise [90], [91]. Multiple light reception and light scattering are the two occurrences of the nonsystematic errors, which make the raw TOF image noisy and being unpredictable. Multiple light reception is mainly caused by object boundaries, e.g. with depth jumps and object concavities (Fig.7 and Fig. 8). Several methods have been proposed to identify and correct the jump edge errors [92]–[95]. Pathak *et al.* [96] used Gaussian analysis in correction of multi-modal measurements. However, its computation cost is very high, this method has to integrate over 100 images for processing every frame, which is difficult to implement for real-time applications. Reynolds et al. used Random Forest Regress to measure pixel confidence and detect flying points according to real world data (Fig. 11a) [78]. Ghorpade *et al.* [79] applied a ''Line-of-Sight'' based edge filter to remove the jump edges. The method has good performance and the computation cost for range image filtering is lower than existing methods. In a different way, Li *et al.* [97] proposed a denoising method for TOF depth images in a weighted least squares framework. The algorithm can well preserve surface edges and improve the Peak Signal-to-Noise Ratio (PSNR) of the denoised images by 0.5–2.6 dB at the same time.



(a) (b)

FIGURE 11. Depth image denoising and resolution improvement. (a) Flying pixels in the white coil [78], (b) RGB-D camera system for upsampling [98].

In addition to edge noises, the accuracy of a TOF camera can also significantly affected by multipath interference when scanning the places with object concavities. by This error may have several centimeters in such places. Fuchs [80] propose a multipath model which can estimate and correct the interference. Such a multipath model was further improved in the recent literatures. Yong *et al.* [99] proposed a denoising algorithm for TOF depth data by a parametric noise model. The results show that the algorithm can yield good denoising effect and preserve edge details as well.

Light scattering effect is another problem affecting the measurement accuracy, due to repeated light reflexions between the lens and the sensor surface in the TOF camera [8]. This effect can be almost minimized following two approaches. The first is applying a filter by combination of amplitude and intensity values to reduce affected scattering pixels [100]. The second is to apply blind deconvolution based on a mathematical model for compensation [101]. Of course, instead of dealing with the scattering effect, people are also seeking for new sensor materials of lower reflectivity to make scattering negligible on the sensor [102].

In practice, MCS can also be used for denoising and resolution improvement [79], [98], [103]–[109]. Oprisescu *et al.* [110] proposed some methods attempting to correct the imaging error of one image based on the other one. The amplitude image is firstly enhanced by using the distance information, and then an algorithm of amplitude-based distance modification corrects some errors of distance estimation for far-distance pixels, rather than treating the amplitude of each TOF sample as a measure of confidence.

2) RESOLUTION IMPROVEMENT

The current TOF sensor limits its image resolution. The low resolution (LR) noisy depth image is big problem faced in the applications [79], [111]–[113]. To improve the depth data resolution [114], [115], Gandhi *et al.* [116] proposed a TOF-stereo fusion method to deal with the LR range data and obtain a dense and accurate depth map. Garicia *et al.* [117] proposed a unified multi-lateral filter which can increase the image resolution in real-time. At present upsampling is an effective method to improve the resolution [118], due to the constraints in upsampling models, the high-resolution depth image obtained in this way suffers from either texture copy artifacts or depth discontinuity blur. An optimization framework proposed in [119] can tackle this problem well. Lately, a deblurring and super-resolution method for blurred TOF images is proposed in [120], which analyzes the image formation model and directly works with raw measurements from the sensor. The reported results outperform most existing methods on both synthetic and real datasets.

Park *et al.* [98] gave a framework to upsample a LR depth map using an auxiliary RGB image with high resolution (HR). They proposed to use registered and potentially HR RGB images as references to enhance the resolution of range images. The number of referenced color images is not restricted [121]. Yeo *et al.* [107] analyzed another framework for upsampling the depth resolution, where the RGB-D camera system is shown in Fig. 11b.

Some other frameworks of multiple camera methods in depth image improvement are also investigated in the literature. Three typical frameworks are shown in Fig. 12. Particularly, Galna *et al.* [22] proposed a method using TOF

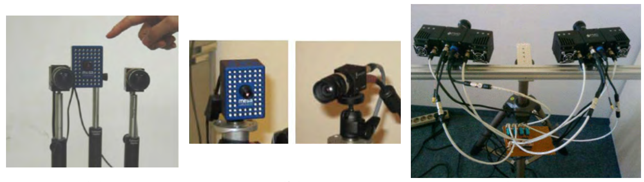

(a) (b) (c)

FIGURE 12. Multiple camera systems. (a) TOF+stereo cameras [123], (b) TOF+Video cameras [125], (c) TOF pair cameras [22].

stereo for depth data acquisition. They combine two cameras in different frequency to obtain depth images, and then improve the accuracy of the depth image with an optimization function. This method can successfully avoid multi-camera interference and improve the TOF data effectively. A pair of TOF cameras is also attempted in [22] (Fig. 12c). Many people have tried to combine a TOF camera with a stereo of color cameras [122]–[124]. Evangelidis *et al.* [124] combined the LR depth map with the HR stereo images. The reconstructed stereo data and depth map are fused according to textural and geometrical likelihoods. This method yields an efficient algorithm for selective growing of correct disparities and runs at 3 fps on a standard personal computer. As another attempt to combining a TOF camera and a video camera Kim *et al.* [87], [125], generated and served 3D video represented by ''video+depth'', where the noisy depth maps are enhanced by performing several steps, e.g. bilateral filtering, outer-boundary refinement, and motion estimation. Finally, it generates high-quality 3D ''video+depth'' in MPEG-4 multimedia.

3) PHASE UNWRAPPING

In the TOF measuring principle formulated in (2), the distance is proportional to the phase difference, but it is restricted by the light modulation wavelength. Distance ambiguity occurs when the measurement distance is larger than the sensor's range $d_{\max}$ (3), which is termed phase wrapping. There are many methods proposed in the last decade for phase unwrapping [126]–[131]. They are categorized into two types, i.e. single depth map-based and multiple depth map-based.

Single map phase unwrapping methods can deal with dynamic environment where the cameras or objects are moving [132]–[134]. For practical applications [132], [133], the depth discontinuity on object boundary is an important cue for relative wrapping estimation. Droeschel *et al.* [127] applied several modulation frequencies to identify wrapping and correct the measured data. In a different way, a generalized approximate message passing (GAMP) framework is used to incorporate both accurate probabilistic modeling for the measurement process and underlying depth map sparsity to accurately extend the unambiguous depth range [129]. Lee [26] proposed ''loopy belief propagation'' for wrapping detection and inference, which is also based on a single map. Some TOF cameras, e.g. SR4500, can take both amplitude image and depth map. The intensity of the amplitude image is inversely proportional to the squared distance. The amplitude values are important in detecting the wrapped regions [134].

Comparing to the single map phase unwrapping methods, multiple map methods are more complicated. These methods use two or more depth maps, which are often taken at different frequencies, and determine the number of wrappings by examining the depth differences for each pixel [127], [135]. It takes the advantage of dealing with occlusion and boundary transition, but it brings another problem for moving cameras and objects because there may be some time difference when switching among different signal frequencies. Considering this problem, a hardware solution will be useful. Actually, it has been reported to make a single shot with two modulation frequencies [136]. It can effectively eliminate the temporal difference in TOF imaging. However, commercial products using such a hardware technique are not yet available probably because of system complexity and cost consideration.

There are some other approaches available in the literature for phase unwrapping, which also utilize a pair of depth maps simultaneously. Among them, Markov random field (MRF) is often applied in the technology. For example, Choi and Lee [137] applied iterative MRF optimization for solving the problem caused by the different viewpoints. Kirmani *et al.* [128] proposed a framework for phase unwrapping in homodyne TOF cameras. As mentioned in [135], the consistency constraint is important in phase unwrapping. Jeong *et al.* [130] described phase unwrapping using single modulated light source and multi photo gate frequencies for TOF camera. To protect human eyes, the illuminating power of TOF cameras is restricted. Consequently, the wrapping points are abundant. In order to achieve a robust estimation against noise, Droeschel *et al.* [127] used an auxiliary depth map of another modulation frequency and incorporates the constraint of depth consistency.

4) MOTION DEBLURRING

As described in Section 4, motion blur in TOF cameras is very different from that in CCD cameras. The motion blur in a CCD camera appears color transition gradually from the foreground to the background in the image [138, 139], but the motion blur in a TOF depth map looks ''overshoot'' or ''undershoot'' near depth jumps. Due to this difference, the existing deblurring algorithms of color images are inapplicable to depth images. A long IT helps to get high SNR depth data, but a shot IT helps to suppress the motion blur. Sun *et al.* [74] found a scheme that can take advantages of both short and long IT and effectively reduce motion artifacts.

In TOF images, there are two types of motion blur artifacts depending on whether due to lateral or axial motion. Some methods are available in solving this problem. In [140], combination of a PMD camera with a conventional color camera is proposed to detect ''lateral motion artifacts'' by an edge detector in the 2D image. Then they filter the image by weighted average of neighbor pixels and perform a

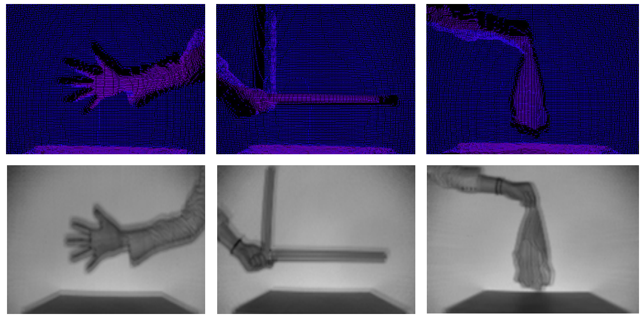


FIGURE 13. Motion blur artifacts of different object. (a) Rigid, (b) Multiple, (c) Deformable.

2-phase depth computation algorithm after sampling analysis of images. In industrial applications, Hussmann *et al.* [141] specifically introduced a method of blur detection for working on a conveyor belt, while this can only deal with one directional motion. Castaneda *et al.* [24] proposed a detection and deblurring method for depth motion blur. These methods are intended to reduce the artifacts to some extent rather than completely eliminate the motion blur. Chang [142], [143] proposed methods for relatively complete systematic deblurring.

As reported in some contributions, motion blur artifacts can be caused by different object motions, e.g. multiple body motion, rigid body motion, and deformable body motion (Fig. 13). Different motion blurs require corresponding methods for deblurring respectively. A notable issue is that motion blur occurs not only just on the object boundaries, but also inside the objects. Furthermore, depth differences inside an object during the integration time can also cause motion blur. A straightforward but ''effective and fast'' method is proposed in [142], which is suitable for realizing hardware with no additional processing time and memory [144], [145].

The representative nonsystematic error reduction and performance improvement approaches in recent years are summarized in Table II.

VI. FUTURE TRENDS

A. MULTIPLE CAMERA SYSTEM

Given the problems of these existing approaches, there are potential trends on TOF cameras. One is the multiple camera system. Although a TOF camera shows the great advantage of real-time depth data acquisition which is rather useful in practical applications, the current TOF imaging products still have two fatal limitations, i.e. resolution and accuracy, as described in the previous sections. A TOF MCS is illustrated in Fig. 14, which is currently a strategy to overcome these shortcomings caused by noise and surface scattering [147], [169]. The MCS uses two or more TOF cameras and combines the acquired depth data sets to improve the accuracy and resolution [14], [31]- [33], [170]. Some 3D object scanning approaches are exploited based on TOF MCS. However, in applications, the multiple camera system often requires

TABLE 2. nonsystematic error reduction and performance improvement approaches.

Tasks	Method	Representative
Multipath effect reduction	Bayesian computational shape model	Adam et al. [31]-2016
Motion deblurring	The Doppler effect of objects in motion	Heide et al.[146]-2015
Motion deblurring	Multi-camera time-of-flight systems	Shrestha et al. [147]-2016
Multipath effect reduction	Estimator based on covariance models	Heijden et al. [148]-2017
Motion deblurring Distortion reduction	Energy-efficient epipolar imaging	Achar et al.[149]-2017
Defocus blur elimination Depth range extension	Focal sweep-based image acquisition	Honnungar et al. [150]-2016
Motion deblurring Distortion reduction Image quality improvement	Pulse based time-of-flight	Sarbolandi et al. [151]-2018
Denoising	Automatically determining the best integration time	Hoegg et al.[152]-2016
Depth range extension	Taking advantage of camera flags or confidence matrix	Kazmi et al.[153]-2014
Phase unwrapping	Thresholding	McClure et al. [133]-2010
Phase unwrapping	Markov random field	Droeschel et al. [132]-2010
Phase unwrapping	Stereo time-of-flight and Markov random field	Choi and Lee [136]-2012
Denoising	Treating phase-delay and amplitude as components of a complex-valued variable	Georgiev et al. [154]-2013
Motion deblurring	Utilizing the relations between different phase offsets from multiple time slots	Lee et al.[155]-2014
Denoising Image quality improvement	Time-of-flight and color sensor fusion	Plank et al. [156]-2016
Denoising Image quality improvement	Weighted error energy minimization	Schwarz et al.[157]-2014
Denoising Image quality improvement	Iteratively reweighted least squares minimization	Choi et al. [158]-2014
Phase unwrapping Motion deblurring	Probabilistic framework	Crabb et al.[159]-2015
Denoising	Finite impulse response filters	Georgiev et al. [160]-2015
Image quality improvement	Calibration	Xie et al.[161]-2014
Precision improvement	Adopting a professional time-to-digital conversion (TDC) chip	Liang et al.[162]-2013
Precision improvement	Employing continuous wavelet transform (CWT)-based local modulus maxima (LMMs) to detect singularities of emitting pulses and receiving echoes	Xiao et al.[163]-2016
Multipath effect reduction	Least squares method	Hofbauer et al.[164]-2014
Phase unwrapping Precision improvement	Time-of-flight and phase shifting (PS)	Zhang et al.[165]-2015
Precision improvement	Multichannel time-delay estimation with linear fitting correction	Li et al.[166]-2013
Denoising	Joint video and sparse 3D transform-domain collaborative filtering	Hach et al.[167]-2015
Precision improvement	Fitting the rising edge of echo envelope	Lai et al.[168]-2014

to capture the same scene by several cameras at different viewpoint. If they work simultaneously at the same frequency, they would interfere with each other. This problem limits the application of multiple camera systems and needs to be solved by some strategies [171], [172].

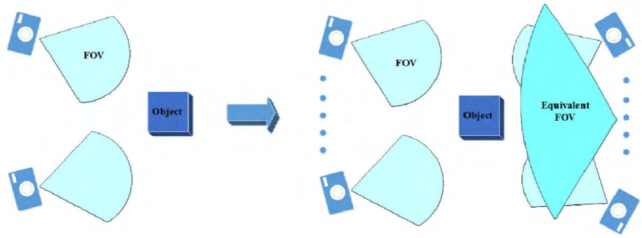


FIGURE 14. Multiple TOF camera system.

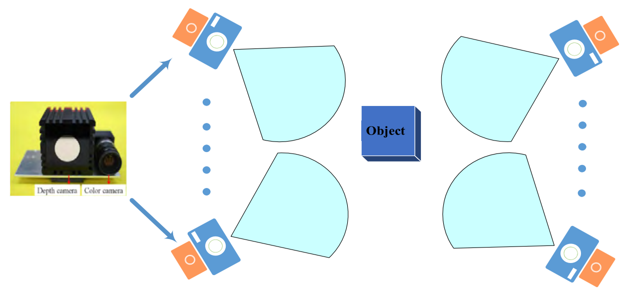


FIGURE 15. RGB-D camera system.

FIGURE 16. People detection and tracking results on office dataset (first row) and mobile camera dataset (second row) [166].

B. RGB-D SYSTEM

Another MCS is to assemble the TOF sensor with one or two color cameras [173], as shown in Fig. 15, which is called RGB-D system. Such a camera system has good performance in many applications, e.g. target detection and tracking [174]–[176] (Fig. 16), human activity analysis [58], [177]–[180] (Fig. 17), object recognition [174], [181]–[188] (Fig. 18), SLAM [189]–[191], hand gesture analysis [192]–[194], and 3D hand pose detection (Fig. 19).

The calibration is a prerequisite in practice for using a measurement system, either stereo sensors or RGB-D MCS, because the color image and depth data have to be corresponded to the world coordinates. A single TOF sensor can usually be calibrated by traditional mathematical methods [195]. However, when calibrating an RGB-D system,

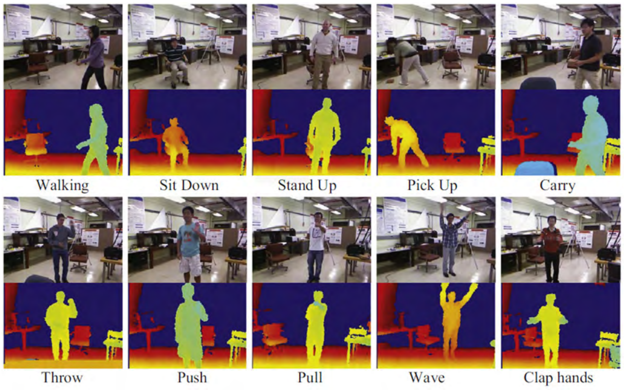


FIGURE 17. Human action recognition from RGB-D data [168].

FIGURE 18. Object recognition in cluttered environment [193].

FIGURE 19. RGB-D system for hand pose detection [147].

the traditional methods do not work well due to the weaknesses of the TOF sensor, such as blurry amplitude image and low resolution. There are some latest calibration methods on various datasets [196]–[200]. However, a better process of calibrating an RGB-D system still desired to overcome the above-mentioned problem.

C. REAL-TIME DATA PROCESSING IN DYNAMIC SCENE

TOF cameras as a developing type of 3D vision sensor are attracting more and more attentions for autonomous mobile robotics [42], [57], [201]. The real-time 3D data can help robots to accomplish autonomous path-planning [57], [202]. Therefore, the vision system now becomes indispensable for the robots to see the environment and avoid possible obstacles, e.g. as the scenario shown in Fig. 20, where the two robots are equipped with a TOF camera. However, some problems are still urgent to be solved in such TOF cameras system, like accurate camera calibration in motion [5].

FIGURE 20. Indoor test with static obstacles [63].

FIGURE 21. Object detection [209].

In order to ensure the correct driving path of a collision free for a mobile robot, we need an effective and robust calibration procedure to obtain the correspondence between the real world and the image [42], [203]. This calibration procedure is required for any mobile robot systems. Extrinsic camera parameters estimation is another task for the system. The reconstruction of the image to real world projection depends strongly on robust estimated camera parameters.

D. INTEGRATION WITH COMPUTER VISION

TOF camera as a vision sensor has to be integrated with vision algorithms for practical applications. Most of contributions in computer vision are dealing with the problems of feature analysis, target detection, recognition, tracking, modeling, and activity analysis. For example of target detection, the algorithm needs to find an area, e.g. showing as a bounding box, to indicate the existence of the target at that place [204]. However, we want to understand the image more meticulous, not just about what is visible but also about what is not visible. After recognizing an object, a better sense is to make clear the exact distance from the observer and the appearance from other views. The TOF image provides a richer representation, and thus computer vision algorithms can take advantages of more cues for practical tasks like object detection, categorization, pose estimation (Fig. 21), and 3D scene labeling (Fig. 22). More corresponding works about these issues can be found in the contributions by [194] and [205]–[208].

Another example of computer vision problem is semantic segmentation for image understanding. A typical attempt of using TOF as a vision sensor for semantic segmentation and

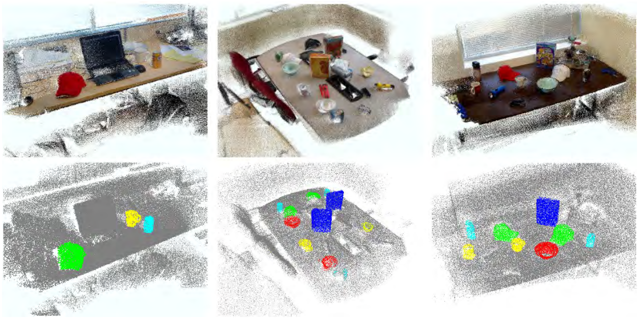


FIGURE 22. Labeling of 3D complex scenes with RGB-D dataset [217].

scene labeling can be found in [210]. Another attempt of adopting a MRF model and using feature descriptor on superpixels is shown in [211]. A recent work in [174] might be inspired by [212], where they chose novel feature descriptors for indoor scene understanding in RGB-D images.

VII. CONCLUSION

This paper summarizes recent advances of TOF data acquisition and processing, including fundamental sensing principles, applications, and current limitations. Typical TOF camera applications include 3D reconstruction, computer vision, medical and biological application, robot navigation, etc. With the mature of TOF camera technology, the performance of TOF camera has been improved obviously. Higher resolution and accuracy depth image are obtained in these years, many systematic and nonsystematic errors have been decreased, and a rich number of projects have been conducted to broaden the application range of TOF cameras. Nevertheless, as summarized in this survey, some challenges still remain. First, the resolution of TOF sensors is still low compared to other vision sensors such as color cameras and laser scanners. Second, superfluous noises still exist though some methods have already led to better SNR. Third, some challenging issues such as phase wrapping still want to be solved better. Although there are some available algorithms to deal with these challenges, more sophisticated approaches are desired in order to increase the unambiguous range. Other concerns are addressed, including motion blur, the relationship between the integration time and SNR, environment light noise and high reflectivity surfaces in the scene.

REFERENCES

[1] M. Hansard, S. Lee, O. Choi, and R. P. Horaud, *Time-of-Flight Cameras: Principles, Methods and Applications*. Berlin, Germany: Springer, 2012, pp. 1–4.

[2] P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo, *Operating Principles of Time-of-Flight Depth Cameras*. Berlin, Germnay: Springer, 2016, pp. 81–113.

[3] A. M. Pinto, P. Costa, A. P. Moreira, L. F. Rocha, G. Veiga, and E. Moreira, "Evaluation of depth sensors for robotic applications," in *Proc. IEEE Int. Conf. Auto. Robot Syst. Competitions*, Vila Real, Portugal, Apr. 2015, pp. 139–143.

[4] M. Kraft, M. Nowicki, R. Penne, A. Schmidt, and P. Skrzypczyński, "Efficient RGB-D data processing for feature-based self-localization of mobile robots," *Int. J. Appl. Math. Comput. Sci.*, vol. 26, no. 1, pp. 63–79, 2016.

[5] H. Yu, J. Zhu, Y. Wang, W. Jia, M. Sun, and Y. Tang, "Obstacle classification and 3D measurement in unstructured environments based on ToF cameras," *Sensors*, vol. 14, no. 6, pp. 10753–10782, 2014.

[6] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments," in *Experimental Robotics*. Berlin, Germany: Springer, 2014, pp. 477–491.

[7] S. Jo, H. Jo, H. M. Cho, and E. Kim, "Pose estimation and 3D environment reconstruction using less reliable depth data," in *Proc. IEEE Int. Conf. Adv. Intell. Mechatronics*, Busan, South Korea, Jul. 2015, pp. 359–364.

[8] D. Liang, Y. Liu, W. Yan, C. Yan, and Q. Dai, "Accurate 3D reconstruction using a multi-phase ToF camera," *Proc. SPIE*, vol. 9273, pp. 92733J-1–92733J-7, Nov. 2014.

[9] D. D. Lichti *et al.*, "Structural deflection measurement with a range camera," *J. Surveying Eng.*, vol. 138, no. 2, pp. 66–76, 2012.

[10] S. Lee, J. Kim, H. Lim, and S. Ahn, "Surface reflectance estimation and segmentation from single depth image of ToF camera," *Signal Process. Image Commun.*, vol. 47, pp. 452–462, Sep. 2016.

[11] S. Hochdorfer and C. Schlegel, "6 DoF SLAM using a ToF camera: The challenge of a continuously growing number of landmarks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Taipei, Taiwan, Oct. 2010, pp. 3981–3986.

[12] D. T. Conway and J. L. Junkins, "Fusion of depth and color images for dense simultaneous localization and mapping," *J. Image Graph.*, vol. 2, no. 1, pp. 64–69, 2014.

[13] H. Lahamy and D. D. Lichti, "Towards real-time and rotation-invariant American sign language alphabet recognition using a range camera," *Sensors*, vol. 12, no. 11, pp. 14416–14441, 2012.

[14] C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, Eds., *3D-TV System with Depth-Image-Based Rendering: Architectures, Techniques and Challenges*. Berlin, Germany: Springer, 2012.

[15] L. Busemeyer *et al.*, "BreedVision—A multi-sensor platform for non-destructive field-based phenotyping in plant breeding," *Sensors*, vol. 13, no. 3, pp. 2830–2847, 2013.

[16] T. T. Nguyen, D. C. Slaughter, N. Max, J. N. Maloof, and N. Sinha, "Structured light-based 3D reconstruction system for plants," *Sensors*, vol. 15, no. 8, pp. 18587–18612, 2015.

[17] W. Koch *et al.*, "A multimodal sensor system for runway debris detection," *Int. J. Microw. Wireless Technol.*, vol. 4, no. 2, pp. 155–162, 2012.

[18] P. Poli, S. B. Healy, and D. P. Dee, "Pose estimation with a Kinect for ergonomic studies: Evaluation of the accuracy using a virtual mannequin," *Sensors*, vol. 15, no. 1, pp. 1785–1803, 2015.

[19] E. E. Stone and M. Skubic, "Fall detection in homes of older adults using the Microsoft Kinect," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 290–301, Jan. 2015.

[20] C. Wang, Z. Liu, and S.-C. Chan, "Superpixel-based hand gesture recognition with Kinect depth camera," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 29–39, Jan. 2015.

[21] R. A. Clark *et al.*, "Instrumenting gait assessment using the Kinect in people living with stroke: reliability and association with balance tests," *J. Neuroeng. Rehabil.*, vol. 12, pp. 1–9, Feb. 2015.

[22] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease," *Gait Posture*, vol. 39, no. 4, pp. 1062–1068, 2014.

[23] D. González-Ortega, F. J. Díaz-Pernas, M. Martínez-Zarzuela, and M. Antón-Rodríguez, "A Kinect-based system for cognitive rehabilitation exercises monitoring," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 620–631, 2014.

[24] V. Castañeda, D. Mateus, and N. Navab, "Stereo time-of-flight," in *Proc. IEEE Int. Conf. Comput. Vis.*, Barcelona, Spain, Nov. 2011, pp. 1684–1691.

[25] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors J.*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.

[26] S. Lee, "Time-of-flight depth camera motion blur detection and deblurring," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 663–666, Jun. 2014.

[27] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor—System description, issues and solutions," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshop*, Washington, DC, USA, Jun./Jul. 2004, p. 35.

[28] B. Kang, S.-J. Kim, S. Lee, K. Lee, J. D. K. Kim, and C.-Y. Kim, "Harmonic distortion free distance estimation in ToF camera," in *Proc. SPIE*, vol. 7864, pp. 357–366, Jan. 2011.

[29] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight cameras in computer graphics," *Comput. Graph. Forum*, vol. 29, no. 1, pp. 141–159, 2010.
[30] M. Van den Bergh and L. Van Gool, "Combining RGB and ToF cameras for real-time 3D hand gesture interaction," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Kona, HI, USA, Jan. 2011, pp. 66–72.
[31] A. Adam, C. Dann, O. Yair, S. Mazor, and S. Nowozin, "Bayesian time-of-flight for realtime shape, illumination and albedo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 5, pp. 851–864, May 2017.
[32] D.-W. Shin and Y.-S. Ho, "Implementation of 3D object reconstruction using multiple Kinect cameras," in *Proc. Electron. Imag., 3D Image Process., Meas. Appl. (3DIPM)*, San Francisco, CA, USA, Feb. 2016, pp. 3DIPM-408.1–3DIPM-408.7.
[33] M. Massot-Campos and G. Oliver-Codina, "Optical sensors and methods for underwater 3D reconstruction," *Sensors*, vol. 15, no. 12, pp. 31525–31557, 2015.
[34] Y. Inoue, S. Nishide, and F. Ren, "Facial expression recognition adaptive to face pose using RGB-D camera," in *Proc. Int. Conf. Ind., Eng. Appl. Appl. Intell. Syst.*, Morioka, Japan, Aug. 2016, pp. 422–427.
[35] D. Dardari, P. Closas, and D. M. Djuric, "Indoor tracking: Theory, methods, and technologies," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1263–1278, Apr. 2015.
[36] R. A. Newcombe *et al.*, "KinectFusion: Real-time dense surface mapping and tracking," in *Proc. IEEE Int. Symp. Mixed Augmented Reality*, Basel, Switzerland, Oct. 2011, pp. 127–136.
[37] F. Chiabrando, R. Chiabrando, 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. 10080–10096, 2009.
[38] P. Fürsattel *et al.*, "A comparative error analysis of current time-of-flight sensors," *IEEE Trans. Comput. Imag.*, vol. 2, no. 1, pp. 27–41, Jan. 2016.
[39] T. K. Kohoutek, D. Droeschel, R. Mautz, and S. Behnke, "Indoor positioning and navigation using time-of-flight cameras," in *TOF Range-Imaging Cameras*. Berlin, Germany: Springer, 2013, pp. 165–176.
[40] K. Tanaka, Y. Mukaigawa, H. Kubo, Y. Matsushita, and Y. Yagi, "Recovering transparent shape from time-of-flight distortion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 4387–4395.
[41] K. Hedenberg and B. Åstrand, "3D sensors on driverless trucks for detection of overhanging objects in the pathway," in *Autonomous Industrial Vehicles: From the Laboratory to the Factory Floor*. West Conshohocken, PA, USA: ASTM International, 2016, pp. 41–56.
[42] A. Gavriilidis, C. Stahlschmidt, J. Velten, and A. Kummert, "Automatic extrinsic camera parameters estimation for a mobile ToF camera application," in *Proc. IEEE Int. Symp. Intell. Control*, Sydney, NSW, Australia, Sep. 2015, pp. 41–46.
[43] S. J. Lee, D. D. Nguyen, and J. W. Jeon, "Design and Implementation of Depth Image Based Real-Time Human Detection," *J. Semicond. Technol. Sci.*, vol. 14, no. 4, pp. 212–226, 2014.
[44] C. A. Luna, C. Losada-Gutierrez, D. Fuentes-Jimenez, A. Fernandez-Rincon, M. Mazo, and J. Macias-Guarasa, "Robust people detection using depth information from an overhead time-of-flight camera," *Expert Syst. Appl.*, vol. 71, pp. 240–256, Apr. 2017.
[45] C. Stahlschmidt, A. Gavriilidis, J. Velten, and A. Kummert, "Applications for a people detection and tracking algorithm using a time-of-flight camera," *Multimedia Tools Appl.*, vol. 75, no. 17, pp. 10769–10786, 2016.
[46] P. Nicolai, J. Raczkowsky, and H. Wörn, "Model-free (human) tracking based on ground truth with time delay: A 3D camera based approach for minimizing tracking latency and increasing tracking quality," in *Proc. 12th Int. Conf. Inform. Control, Autom. Robot.*, Colmar, France, Jul. 2015, pp. 247–266.
[47] Q. Wei, J. Shan, H. Cheng, Z. Yu, B. Lijuan, and Z. Haimei, "A method of 3D human-motion capture and reconstruction based on depth information," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Harbin, China, Aug. 2016, pp. 187–192.
[48] F. Martino, C. Patruno, N. Mosca, and E. Stella, "Material recognition by feature classification using time-of-flight camera," *J. Electron. Imag.*, vol. 25, no. 6, p. 061412, 2016.
[49] W. Liu, X. Chen, J. Yang, and Q. Wu, "Robust color guided depth map restoration," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 315–327, Jan. 2017.
[50] L. Chen, Y. He, J. Chen, Q. Li, and Q. Zou, "Transforming a 3-D LiDAR point cloud into a 2-D dense depth map through a parameter self-adaptive framework," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 165–176, Jan. 2017.
[51] S. Mandal, A. Bhavsar, and A. K. Sao, "Depth map restoration from undersampled data," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 119–134, Jan. 2017.
[52] L. Fang *et al.*, "Comparative evaluation of time-of-flight depth-imaging sensors for mapping and SLAM applications," in *Proc. Austral. Conf. Robot. Autom.*, Brisbane, QLD, Australia, Dec. 2016, pp. 1–7.
[53] M. Li *et al.*, "Fast and robust mapping with low-cost Kinect V2 for photovoltaic panel cleaning robot," in *Proc. Int. Conf. Adv. Robot. Mechatronics*, Macau, China, Aug. 2016, pp. 95–100.
[54] P. Zanuttigh, G. Marin, C. Dal Mutto, F. Dominio, L. Minto, and G. M. Cortelazzo, "3D scene reconstruction from depth camera data," in *Time-of-Flight and Structured Light Depth Cameras*. Berlin, Germany: Springer, 2016, pp. 231–251.
[55] W.-C. Wang, K.-R. Lin, C. L. Tsui, D. Schipf, and J. Leang, "Underwater camera with depth measurement," *Proc. SPIE*, vol. 9805, p. 980529, Mar. 2016.
[56] S. T. Digumarti, G. Chaurasia, A. Taneja, R. Siegwart, A. Thomas, and P. Beardsley, "Underwater 3D capture using a low-cost commercial depth camera," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, Lake Placid, NY, USA, Mar. 2016, pp. 1–9.
[57] S. L. X. Francis, S. G. Anavatti, M. Garratt, and H. Shim, "A ToF-camera as a 3D vision sensor for autonomous mobile robotics," *Int. J. Adv. Robotic Syst.*, vol. 12, no. 11, Nov. 2015.
[58] C. H. Jang, C. S. Kim, K. Jo, and M. Sunwoo, "Design factor optimization of 3D flash lidar sensor based on geometrical model for automated vehicle and advanced driver assistance system applications," *Int. J. Autom. Technol.*, vol. 18, no. 1, pp. 147–156, 2017.
[59] K. Klionovska and H. Benninghoff, "Visual navigation for rendezvous and docking using PMD camera," in *Proc. Int. Conf. Geograph. Inf. Syst. Theory, Appl. Manage.*, Rome, Italy, Apr. 2016, pp. 3–7.
[60] I. Eichhardt, Z. Jankó, and D. Chetverikov, "Novel methods for image-guided ToF depth upsampling," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Budapest, Hungary, Oct. 2016, pp. 2073–2078.
[61] S. Fuchs, S. Haddadin, M. Keller, S. Parusel, A. Kolb, and M. Suppa, "Cooperative bin-picking with Time-of-Flight camera and impedance controlled DLR lightweight robot III," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Taipei, Taiwan, Oct. 2010, pp. 4862–4867.
[62] Y. Tian, H. Liu, and T. Furukawa, "Reliable infrastructural urban traffic monitoring via Lidar and camera fusion," *SAE Int. J. Passenger Cars-Electron. Elect. Syst.*, vol. 10, no. 1, pp. 173–180, 2017.
[63] J. Wetzl, O. Taubmann, S. Haase, T. Köhler, M. Kraus, and J. Hornegger, "GPU-accelerated time-of-flight super-resolution for image-guided surgery," in *Bildverarbeitung für die Medizin*. Berlin, Germany: Springer, 2013, pp. 21–26.
[64] T. Beyl, P. Nicolai, M. D. Comparetti, J. Raczkowsky, E. De Momi, and H. Wörn, "Time-of-flight-assisted Kinect camera-based people detection for intuitive human robot cooperation in the surgical operating room," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 11, no. 7, pp. 1329–1345, 2015.
[65] M. Gilles *et al.*, "Patient positioning in radiotherapy based on surface imaging using time of flight cameras," *Med. Phys.*, vol. 43, no. 8, pp. 4833–4841, 2016.
[66] M. Karbasi, Z. Bhatti, P. Nooralishahi, A. Shah, and S. M. R. Mazloomnezhad, "Real-time hands detection in depth image by using distance with Kinect camera," *Int. J. Internet Things*, vol. 4, no. 1A, pp. 1–6, 2015.
[67] A. Czyżewski, B. Kostek, M. Szykulski, and T. E. Ciszewski, "Building knowledge for the purpose of lip speech identification," in *Multimedia and Network Information Systems*. Berlin, Germany: Springer, 2017, pp. 3–14.
[68] M. B. Holte, "3D scanning of clothing using a RGB-D sensor with application in a virtual dressing room," in *Advances in Applied Digital Human Modeling and Simulation*. Berlin, Germany: Springer, 2017, pp. 143–153.
[69] K. Hoshino, "Hand gesture interface for entertainment games," in *Handbook of Digital Games and Entertainment Technologies*. 2017, pp. 293–312.
[70] J. Yu and J. Park, "Real-time facial tracking in virtual reality," in *Proc. SIGGRAPH ASIA VR Showcase*, Macao, China, Dec. 2016, Art. no. 4.
[71] Y. Lai, C. Wang, Y. Li, S. S. Ge, and D. Huang, "3D pointing gesture recognition for human-robot interaction," in *Proc. Chin. Control Decis. Conf.*, Yinchuan, China, May 2016, pp. 4959–4964.
[72] F. M. M. da Silva and J. A. S. Centeno, "Systematic depth error modeling in range measurements on PMD CamCube 3.0," *Boletim de Ciências Geodésicas*, vol. 21, no. 1, pp. 126–148, 2015.

[73] Y. He, B. Liang, Y. Zou, J. He, and J. Yang, "Depth errors analysis and correction for time-of-flight (ToF) cameras," *Sensors*, vol. 17, no. 1, p. 92, 2017.

[74] K. K. Sun, O. Choi, B. Kang, J. D. Kim, and C.-Y. Kim, "Time-of-flight depth image enhancement using variable integration time," *Proc. SPIE*, vol. 8650, p. 525, Mar. 2013.

[75] T. Kahlmann, F. Remondino, and H. Ingensand, "Calibration for increased accuracy of the range imaging camera SwissRanger," *Image Eng. Vis. Metrol.*, vol. 36, pp. 136–141, Sep. 2006.

[76] S. Hussmann and F. Knoll, "Modulation method for minimizing the depth distortion offset of lock-in TOF cameras," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Minneapolis, MN, USA, May 2013, pp. 551–556.

[77] S. Lee, "Depth camera image processing and applications," in *Proc. IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 545–548.

[78] M. Reynolds, J. Doboš, L. Peel, T. Weyrich, and G. J. Brostow, "Capturing time-of-flight data with confidence," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Colorado Springs, CO, USA, Jun. 2011, pp. 945–952.

[79] V. K. Ghorpade, P. Checchin, and L. Trassoudaine, "Line-of-sight-based ToF camera's range image filtering for precise 3D scene reconstruction," in *Proc. Eur. Conf. Mobile Robots*, Lincoln, U.K., Sep. 2015, pp. 1–6.

[80] S. Fuchs, "Multipath interference compensation in time-of-flight camera images," in *Proc. Int. Conf. Pattern Recognit.*, Istanbul, Turkey, Aug. 2010, pp. 3583–3586.

[81] M. Lindner and A. Kolb, "Lateral and depth calibration of PMD-distance sensors," in *Advances in Visual Computing* (Lecture Notes in Computer Science), vol. 4292. Berlin, Germany: Springer, 2006, pp. 524–533.

[82] M. Lindner, A. Kolb, and T. Ringbeck, "New insights into the calibration of ToF-sensors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Anchorage, AK, USA, Jun. 2008, pp. 1–5.

[83] S. Fuchs and S. May, "Calibration and registration for precise surface reconstruction with time-of-flight cameras," *Int. J. Intell. Syst. Technol. Appl.*, vol. 5, nos. 3–4, pp. 274–284, 2008.

[84] S. Hussmann, F. Knoll, and T. Edeler, "Modulation method including noise model for minimizing the wiggling error of TOF cameras," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 5, pp. 1127–1136, May 2014.

[85] S. Fuchs and G. Hirzinger, "Extrinsic and depth calibration of ToF-cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA, Jun. 2008, pp. 1–6.

[86] J. Radmer, P. M. Fuste, H. Schmidt, and J. Kruger, "Incident light related distance error study and calibration of the PMD-range imaging camera," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Anchorage, AK, USA, Jun. 2008, pp. 1–6.

[87] Y. M. Kim, D. Chan, C. Theobalt, and S. Thrun, "Design and calibration of a multi-view TOF sensor fusion system," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA, Jun. 2008, pp. 1–7.

[88] I. Schiller, C. Beder, and R. Koch, "Calibration of a PMD-camera using a planar calibration pattern together with a multi-camera setup," in *Proc. ISPRS Conf.*, Beijing, China, Jul. 2008, pp. 1–6.

[89] M. Lindner and A. Kolb, "Calibration of the intensity-related distance error of the PMD ToF-camera," *Proc. SPIE*, vol. 6764, p. 67640W, Sep. 2007.

[90] D. D. Lichti, "Self-calibration of a 3D range camera," in *Proc. ISPRS*, Beijing, China, Jul. 2008, pp. 1–6.

[91] F. Chiabrando, R. Chiabrando, 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. 10080–10096, 2009.

[92] S. May *et al.*, "Three-dimensional mapping with time-of-flight cameras," *J. Field Robot.*, vol. 26, nos. 11–12, pp. 934–965, Nov. 2009.

[93] S. Nobuhara, T. Kashino, T. Matsuyama, K. Takeuchi, and K. Fujii, "A single-shot multi-path interference resolution for mirror-based full 3D shape measurement with a correlation-based ToF camera," in *Proc. Int. Conf. 3D Vis.*, Stanford, CA, USA, Oct. 2016, pp. 342–350.

[94] K. Son, M.-Y. Liu, and Y. Taguchi, "Learning to remove multipath distortions in Time-of-Flight range images for a robotic arm setup," in *Proc. IEEE Int. Conf. Robot. Autom.*, Stockholm, Sweden, May 2016, pp. 3390–3397.

[95] V. N. Xuan, K. Hartmann, W. Weihs, and O. Loffeld, "Multi-target super-resolution using compressive sensing arguments for multipath interference recovery," in *Proc. Int. Workshop Compressed Sens. Theory Appl. Radar, Sonar Remote Sens.*, Aachen, Germany, Sep. 2016, pp. 148–152.

[96] K. Pathak, A. Birk, and J. Poppinga, "Sub-pixel depth accuracy with a time of flight sensor using multimodal Gaussian analysis," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nice, France, Sep. 2008, pp. 3519–3524.

[97] Y. Li, J. Li, L. Wang, J. Zhang, D. Li, and M. Zhang, "A weighted least squares algorithm for time-of-flight depth image denoising," *Opt.-Int. J. Light Electron Opt.*, vol. 125, no. 13, pp. 3283–3286, Jul. 2014.

[98] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3D-TOF cameras," in *Proc. IEEE Int. Conf. Comput. Vis.*, Barcelona, Spain, Nov. 2011, pp. 1623–1630.

[99] Y. S. Kim *et al.*, "Parametric model-based noise reduction for ToF depth sensors," *Proc. SPIE*, vol. 8290, pp. 82900A-1–82900A-8, Apr. 2012.

[100] 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.*, Beijing, China, Oct. 2006, pp. 790–795.

[101] J. Mure-Dubois and H. Hügli, "Optimized scattering compensation for time-of-flight camera," *Proc. SPIE*, vol. 6762, p. 67620H, Oct. 2007.

[102] 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.

[103] B. Bartczak and R. Koch, "Dense depth maps from low resolution time-of-flight depth and high resolution color views," in *Proc. Int. Symp. Adv. Vis. Comput.*, Las Vegas, NV, USA, Nov./Dec. 2009, pp. 228–239.

[104] D. Chan, H. Buisman, C. Theobalt, and S. Thrun, "A noise-aware filter for real-time depth upsampling," in *Proc. ECCV Workshop Multi-Camera Multi-Modal Sensor Fusion Algorithms Appl.*, Marseille, France, Oct. 2008, pp. 1–12.

[105] J. Dolson, J. Baek, C. Plagemann, and S. Thrun, "Upsampling range data in dynamic environments," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 1–8.

[106] B. Huhle, T. Schairer, P. Jenke, and W. Strasser, "Robust non-local denoising of colored depth data," in *Proc. IEEE CVPR Workshop Time Flight Camera Based Comput. Vis.*, Anchorage, AK, USA, Jun. 2008, pp. 1–7.

[107] D. Yeo, E. ul Haq, J. Kim, and M. W. Baig, "Adaptive bilateral filtering for noise removal in depth upsampling," in *Proc. Soc Design Conf.*, Las Vegas, NV, USA, Sep. 2010, pp. 36–39.

[108] T. Köhler *et al.*, "ToF meets RGB: Novel multi-sensor super-resolution for hybrid 3-D endoscopy," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, Nagoya, Japan, Sep. 2013, pp. 139–146.

[109] T. Köhler *et al.*, "Multi-sensor super-resolution for hybrid range imaging with application to 3-D endoscopy and open surgery," *Med. Image Anal.*, vol. 24, no. 1, pp. 220–234, 2015.

[110] S. Oprisescu, D. Falie, M. Ciuc, and V. Buzuloiu, "Measurements with ToF cameras and their necessary corrections," in *Proc. Int. Symp. Signals, Circuits Syst.*, Iasi, Romania, Jul. 2007, pp. 1–4.

[111] D. Anderson, H. Herman, and A. Kelly, "Experimental characterization of commercial flash ladar devices," in *Proc. Int. Conf. Sens. Technol.*, Palmerston North, New Zealand, Nov. 2005, pp. 17–23.

[112] K. M. Kyung, K. Bae, S. H. Cho, and T.-C. Kim, "Gesture-dependent depth data extraction for low resolution time-of-flight camera," in *Proc. IEEE Int. Conf. Consum. Electron.-Berlin*, Berlin, Germany, Sep. 2012, pp. 183–184.

[113] H. Lietz, M. M. Hassan, and J. Eberhardt, "Optical lens-shift design for increasing spatial resolution of 3D ToF cameras," *Adv. Opt. Technol.*, vol. 6, no. 1, pp. 39–46, 2017.

[114] C. S. Balure and M. R. Kini, "Depth image super-resolution: A review and wavelet perspective," in *Proc. Int. Conf. Comput. Vis. Image Process.*, Uttar Pradesh, India, Sep. 2017, pp. 543–555.

[115] J. Lei, L. Li, H. Yue, F. Wu, N. Ling, and C. Hou, "Depth map super-resolution considering view synthesis quality," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1732–1745, Apr. 2017.

[116] V. Gandhi, J. Čech, and R. Horaud, "High-resolution depth maps based on TOF-stereo fusion," in *Proc. IEEE Int. Conf. Robot. Autom.*, Saint Paul, MN, USA, May 2012, pp. 4742–4749.

[117] F. Garcia, D. Aouada, B. Mirbach, T. Solignac, and B. Ottersten, "Unified multi-lateral filter for real-time depth map enhancement," *Image Vis. Comput.*, vol. 41, pp. 26–41, Sep. 2015.

[118] C. Jung, S. Yu, and J. Kim, "Intensity-guided edge-preserving depth upsampling through weighted L_0 gradient minimization," *J. Vis. Commun. Image Represent.*, vol. 42, pp. 132–144, Jan. 2017.

[119] W. Liu, Y. Li, X. Chen, J. Yang, Q. Wu, and J. Yu. (2015). "Robust high quality image guided depth upsampling." [Online]. Available: https://arxiv.org/abs/1506.05187

[120] L. Xiao *et al.*, "Defocus deblurring and superresolution for time-of-flight depth cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 2376–2384.

[121] Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Minneapolis, MN, USA, Jun. 2007, pp. 1–8.

[122] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of time-of-flight depth and stereo for high accuracy depth maps," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Anchorage, AK, USA, Jun. 2008, pp. 1–8.

[123] C. D. Mutto, P. Zanuttigh, and G. M. Cortelazzo, "Probabilistic ToF and stereo data fusion based on mixed pixels measurement models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 11, pp. 2260–2272, Nov. 2015.

[124] G. D. Evangelidis, M. Hansard, and R. Horaud, "Fusion of range and stereo data for high-resolution scene-modeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 11, pp. 2178–2192, Nov. 2015.

[125] S.-Y. Kim, J.-H. Cho, A. Koschan, and M. A. Abidi, "3d video generation and service based on a TOF depth sensor in MPEG-4 multimedia framework," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1730–1738, Aug. 2010.

[126] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, no. 4, pp. 713–720, Jul./Aug. 1988.

[127] D. Droeschel, D. Holz, and S. Behnke, "Multi-frequency phase unwrapping for time-of-flight cameras," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Taipei, Taiwan, Oct. 2010, pp. 1463–1469.

[128] A. Kirmani, A. Benedetti, and P. A. Chou, "SPUMIC: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods," in *Proc. IEEE Int. Conf. Multimedia Expo*, San Jose, CA, USA, Jul. 2013, pp. 1–6.

[129] J. Mei, "Algorithms for 3D time-of-flight imaging," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2013.

[130] S. Y. Jeong, K. M. Kyung, T.-C. Kim, K. Bae, and S. H. Cho, "Time of flight camera device and method of driving," U.S. Patent 20 150 204 970, Jul. 2015.

[131] W. H. Wong, N. A. Mullani, E. A. Philippe, R. Hartz, and K. L. Gould, "Image improvement and design optimization of the time-of-fight PET," *J. Nucl. Med.*, vol. 24, no. 1, pp. 52–60, Feb. 1983.

[132] O. Choi *et al.*, "Range unfolding for time-of-flight depth cameras," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, Sep. 2010, pp. 4189–4192.

[133] D. Droeschel, D. Holz, and S. Behnke, "Probabilistic phase unwrapping for time-of-flight cameras," in *Proc. Int. Symp. Robot. German Conf. Robot.*, Munich, Germany, Jun. 2010, pp. 1–7.

[134] S. H. McClure, M. J. Cree, A. A. Dorrington, and A. D. Payne, "Resolving depth-measurement ambiguity with commercially available range imaging cameras," *Proc. SPIE*, vol. 7538, p. 75380K, Jan. 2010.

[135] D. Falie and V. Buzuloiu, "Wide range time of flight camera for outdoor surveillance," in *Proc. Microw., Radar Remote Sens. Symp.*, Kiev, Ukraine, Sep. 2008, pp. 79–82.

[136] A. D. Payne, A. P. Jongenelen, A. A. Dorrington, M. J. Cree, and D. A. Carnegie, "Multiple frequency range imaging to remove measurement ambiguity," in *Proc. 9th Conf. Opt. 3-D Meas. Techn.*, Vienna, Austria, Jul. 2009, pp. 139–148.

[137] O. Choi and S. Lee, "Wide range stereo time-of-flight camera," in *Proc. IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 557–560.

[138] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce, "Non-uniform deblurring for shaken images," *Int. J. Comput. Vis.*, vol. 98, no. 2, pp. 168–186, 2012.

[139] L. Zhang, A. Deshpande, and X. Chen, "Denoising vs. deblurring: HDR imaging techniques using moving cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 522–529.

[140] O. Lottner, A. Sluiter, K. Hartmann, and W. Weihs, "Movement artefacts in range images of time-of-flight cameras," in *Proc. Int. Symp. Signals, Circuits Syst.*, Iasi, Romania, Jul. 2007, pp. 1–4.

[141] S. Hussmann, A. Hermanski, and T. Edeler, "Real-time motion artifact suppression in TOF camera systems," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1682–1690, May 2011.

[142] S. Lee, B. Kang, J. D. K. Kim, and C. Y. Kim, "Motion blur-free time-of-flight range sensor," *Proc. SPIE*, vol. 8298, pp. 105–118, Feb. 2012.

[143] Y. K. Chang, "ToF depth image motion blur detection using 3D blur shape models," *Proc. SPIE*, vol. 8296, pp. 829615-1–829615-6, Feb. 2012.

[144] M. Li, "Three classes of fractional oscillators," *Symmetry*, vol. 10, no. 2, 2018, Art. no. 40.

[145] M. Li, "Fractal time series—A tutorial review," *Math. Problems Eng.*, vol. 2010, Oct. 2010, Art. no. 157264.

[146] F. Heide, W. Heidrich, M. Hullin, and G. Wetzstein, "Doppler time-of-flight imaging," *ACM Trans. Graph.*, vol. 34, no. 4, 2015, Art. no. 36.

[147] S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein, "Computational imaging with multi-camera time-of-flight systems," *ACM Trans. Graph.*, vol. 35, no. 4, 2016, Art. no. 33.

[148] F. van der Heijden, G. Túquerres, and P. Regtien, "Time-of-flight estimation based on covariance models," *Meas. Sci. Technol.*, vol. 14, no. 8, pp. 1295–1304, 2017.

[149] S. Achar, J. R. Bartels, W. L. Whittaker, K. N. Kutulakos, and S. G. Narasimhan, "Epipolar time-of-flight imaging," *ACM Trans. Graph.*, vol. 36, no. 4, 2017, Art. no. 37.

[150] S. Honnungar, J. Holloway, A. K. Pediredla, A. Veeraraghavan, and K. Mitra, "Focal-sweep for large aperture time-of-flight cameras," in *Proc. IEEE Int. Conf. Image Process.*, Phoenix, AZ, USA, Sep. 2016, pp. 953–957.

[151] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse based time-of-flight range sensing," *Sensors*, vol. 18, no. 6, p. 1679, 2018.

[152] T. Hoegg, C. Baiz, and A. Kolb, "Online improvement of time-of-flight camera accuracy by automatic integration time adaption," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Abu Dhabi, United Arab Emirates, Dec. 2016, pp. 613–618.

[153] W. Kazmi, S. Foix, G. Alenyà, and H. J. Andersen, "Indoor and outdoor depth imaging of leaves with time-of-flight and stereo vision sensors: Analysis and comparison," *ISPRS J. Photogram. Remote Sens.*, vol. 88, pp. 128–146, Feb. 2014.

[154] M. Georgiev, A. Gotchev, and M. Hannuksela, "De-noising of distance maps sensed by time-of-flight devices in poor sensing environment," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Vancouver, BC, Canada, May 2013, pp. 1533–1537.

[155] S. Lee, "Time-of-flight depth camera motion blur detection and deblurring," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 663–666, Jun. 2014.

[156] H. Plank, G. Holweg, T. Herndl, and N. Druml, "High performance Time-of-Flight and color sensor fusion with image-guided depth super resolution," in *Proc. Conf. Design, Autom. Test Eur.*, Dresden, Germany, Mar. 2016, pp. 1213–1218.

[157] S. Schwarz, M. Sjöström, and R. Olsson, "Time-of-flight sensor fusion with depth measurement reliability weighting," in *Proc. 3DTV-Conf., True Vis.-Capture, Transmiss. Display 3D Video*, Budapest, Hungary, Jul. 2014, pp. 1–4.

[158] O. Choi and B. Kang, "Denoising of time-of-flight depth data via iteratively reweighted least squares minimization," in *Proc. IEEE Int. Conf. Image Process.*, Melbourne, VIC, Australia, Sep. 2014, pp. 1075–1079.

[159] R. Crabb and R. Manduchi, "Probabilistic phase unwrapping for single-frequency time-of-flight range cameras," in *Proc. Int. Conf. 3D Vis.*, Tokyo, Japan, Dec. 2015, pp. 577–584.

[160] M. Georgiev, R. Bregovic, and A. Gotchev, "Fixed-pattern noise suppression in low-sensing environment of Time-of-Flight devices," in *Proc. Sensors Appl. Symp.*, Zadar, Croatia, Apr. 2015, pp. 1–6.

[161] M. Xie and J. R. Cooperstock, "Time-of-flight camera calibration for improved 3D reconstruction of indoor scenes," in *Proc. Int. Symp. Comput. Intell. Design*, Hangzhou, China, Dec. 2014, pp. 478–481.

[162] K. Liang, H. Liu, and H. Ju, "Accurate ranging method of pulse laser time-of-flight based on the principle of self-triggering," in *Proc. Control Conf.*, Xi'an, China, Jul. 2013, pp. 5583–5587.

[163] J. Xiao, M. Lopez, X. Hu, J. Xiao, and F. Yan, "A continuous wavelet transform-based modulus maxima approach for the walk error compensation of pulsed time-of-flight laser rangefinders," *Opt.-Int. J. Light Electron Opt.*, vol. 127, no. 4, pp. 1980–1987, Apr. 2016.

[164] M. Hofbauer, J. Seiter, M. Davidovic, and H. Zimmermann, "A processing approach for a correlating time-of-flight range sensor based on a least squares method," in *Proc. IEEE Sensors Appl. Symp.*, Queenstown, New Zealand, Feb. 2014, pp. 355–359.

[165] Y. Zhang, Z. Xiong, and F. Wu, "Fusion of time-of-flight and phase shifting for high-resolution and low-latency depth sensing," in *proc. Int. Conf. Multimedia Expo*, Turin, Italy, Jun./Jul. 2015, pp. 1–6.

[166] C. Li, Q. Chen, G. Gu, and W. Qian, "Laser time-of-flight measurement based on time-delay estimation and fitting correction," *Opt. Eng.*, vol. 52, p. 076105, Jul. 2013.

[167] T. Hach, T. Seybold, and H. Böttcher, "Joint video and sparse 3D transform-domain collaborative filtering for time-of-flight depth maps," in *Proc. IEEE Int. Symp. Multimedia*, Miami, FL, USA, Dec. 2015, pp. 272–277.
[168] G. Lai, W. Shi, Q. Xiong, and X. Shen, "A measurement method based on fitting the rising edge of echo envelope for ultrasonic time-of-flight," *Chin. J. Sens. Actuators*, vol. 27, pp. 922–927, Jul. 2014.
[169] M. Lindner, I. Schiller, A. Kolb, and R. Koch, "Time-of-flight sensor calibration for accurate range sensing," *Comput. Vis. Image Understand.*, vol. 114, no. 12, pp. 1318–1328, 2010.
[170] A. Corti, S. Giancola, G. Mainetti, and R. Sala, "A metrological characterization of the Kinect V2 time-of-flight camera," *Robot. Auto. Syst.*, vol. 75, pp. 584–594, Jan. 2016.
[171] L. Li, S. Xiang, Y. Yang, and L. Yu, "Multi-camera interference cancellation of time-of-flight (TOF) cameras," in *Proc. IEEE Int. Conf. Image Process.*, Quebec City, QC, Canada, Sep. 2015, pp. 556–560.
[172] A. Kadambi, J. Schiel, and R. Raskar, "Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 893–902.
[173] B. Huhle, T. Schairer, P. Jenke, and W. Straßer, "Fusion of range and color images for denoising and resolution enhancement with a non-local filter," *Comput. Vis. Image Understand.*, vol. 114, no. 12, pp. 1336–1345, Dec. 2010.
[174] S. Gupta, P. Arbeláez, R. Girshick, and J. Malik, "Indoor scene understanding with RGB-D images: Bottom-up segmentation, object detection and semantic segmentation," *Int. J. Comput. Vis.*, vol. 112, no. 2, pp. 133–149, 2015.
[175] J. Liu, Y. Liu, G. Zhang, P. Zhu, and Y. Q. Chen, "Detecting and tracking people in real time with RGB-D camera," *Pattern Recognit. Lett.*, vol. 53, pp. 16–23, Feb. 2015.
[176] M. Shin, S. Zeigler, A. Gomez, and C. Reid, "Head mounted device for gesture recognition using acoustic time of flight sensor and 2D camera," *J. Acoust. Soc. Amer.*, vol. 140, no. 4, p. 3422, 2016.
[177] A.-A. Liu, W.-Z. Nie, Y.-T. Su, L. Ma, T. Hao, and Z.-X. Yang, "Coupled hidden conditional random fields for RGB-D human action recognition," *Signal Process.*, vol. 112, pp. 74–82, Jul. 2015.
[178] S. Gupta, J. S. Supancic, III, M. Khademi, J. M. M. Montiel, and D. Ramanan, "3D hand pose detection in egocentric RGB-D images," in *Proc. Eur. Conf. Comput. Vis.*, Zürich, Switzerland, Sep. 2015, pp. 1–14.
[179] A. Shahroudy, J. Liu, T.-T. Ng, and G. Wang, "NTU RGB+D: A large scale dataset for 3D human activity analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 1–10.
[180] D. Gürkaynak and H. Yalçın, "Recognition and classification of human activity from RGB-D videos," in *Proc. Signal Process. Commun. Appl. Conf.*, Malatya, Turkey, May 2015, pp. 1745–1748.
[181] S. Song, S. P. Lichtenberg, and J. Xiao, "SUN RGB-D: A RGB-D scene understanding benchmark suite," in *Proc. CVPR*, Jun. 2015, pp. 567–576.
[182] S. Gupta, P. Arbeláez, R. Girshick, and J. Malik. (2015). "Inferring 3D object pose in RGB-D images." [Online]. Available: https://arxiv.org/abs/1502.04652
[183] G. Toscana, S. Rosa, and B. Bona, "Fast graph-based object segmentation for RGB-D images," in *Proc. SAI Intell. Syst. Conf. (IntelliSys)*, in Lecture Notes in Networks and Systems, vol. 16. Cham, Switzerland: Springer, 2018, pp. 42–58.
[184] K. Wu, R. Ranasinghe, and G. Dissanayake, "A fast pipeline for textured object recognition in clutter using an RGB-D sensor," in *Proc. Int. Conf. Control Autom. Robot. Vis.*, Singapore, Dec. 2015, pp. 1650–1655.
[185] A. Perez-Yus, D. Gutierrez-Gomez, G. Lopez-Nicolas, and J. J. Guerrero, "Stairs detection with odometry-aided traversal from a wearable RGB-D camera," *Comput. Vis. Image Understand.*, vol. 154, pp. 192–205, Jan. 2017.
[186] W. Pan, F. Zhu, and Y. Hao, "Combining depth and gray images for fast 3D object recognition," *Proc. SPIE*, vol. 10155, p. 101553C, May 2016.
[187] T. T. Nguyen, K. Vandevoorde, N. Wouters, E. Kayacan, J. G. De Baerdemaeker, and W. Saeys, "Detection of red and bicoloured apples on tree with an RGB-D camera," *Biosyst. Eng.*, vol. 146, pp. 33–44, Jun. 2016.
[188] O. Wasenmüller, M. D. Ansari, and D. Stricker, "DNA-SLAM: Dense noise aware SLAM for ToF RGB-D cameras," in *Proc. Asian Conf. Comput. Vis. Workshop*, Taipei, Taiwan, Nov. 2016, pp. 613–629.
[189] H. Zhang, Y. Liu, and J. Tan, "Loop closing detection in RGB-D SLAM combining appearance and geometric constraints," *Sensors*, vol. 15, no. 6, pp. 14639–14660, 2015.
[190] K. Di, Q. Zhao, W. Wan, Y. Wang, and Y. Gao, "RGB-D SLAM based on extended bundle adjustment with 2D and 3D information," *Sensors*, vol. 16, no. 8, p. 1285, 2016.
[191] D. Gutierrez-Gomez, W. Mayol-Cuevas, and J. J. Guerrero, "Dense RGB-D visual odometry using inverse depth," *Robot. Auto. Syst.*, vol. 75, pp. 571–583, Jan. 2016.
[192] S.-Z. Li, B. Yu, W. Wu, S.-Z. Su, and R.-R. Ji, "Feature learning based on SAE–PCA network for human gesture recognition in RGBD images," *Neurocomputing*, vol. 151, pp. 565–573, Mar. 2015.
[193] X. Chen and M. Koskela, "Using appearance-based hand features for dynamic RGB-D gesture recognition," in *Proc. Int. Conf. Pattern Recognit.*, Stockholm, Sweden, Aug. 2014, pp. 411–416.
[194] J. Molina, J. A. Pajuelo, and J. M. Martínez, "Real-time motion-based hand gestures recognition from time-of-flight video," *J. Signal Process. Syst.*, vol. 86, no. 1, pp. 17–25, 2017.
[195] M. Hansard, R. Horaud, M. Amat, and G. Evangelidis, "Automatic detection of calibration grids in time-of-flight images," *Comput. Vis. Image Understand.*, vol. 121, pp. 108–118, Apr. 2014.
[196] J. Jung, J. Y. Lee, Y. Jeong, and I. S. Kweon, "Time-of-flight sensor calibration for a color and depth camera pair," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 7, pp. 1501–1513, Jul. 2015.
[197] M. Hansard, G. Evangelidis, Q. Pelorson, and R. Horaud, "Cross-calibration of time-of-flight and colour cameras," *Comput. Vis. Image Understand.*, vol. 134, pp. 105–115, May 2015.
[198] B. Karan, "Kalibracija RGB-D senzora tipa Kinekt za roboticarske primene," *FME Trans.*, vol. 47, no. 1, pp. 47–54, 2015.
[199] F. Basso, E. Menegatti, and A. Pretto, "Robust intrinsic and extrinsic calibration of RGB-D cameras," *IEEE Trans. Robot.*, vol. 34, no. 5, pp. 315–1332, Oct. 2017.
[200] V. Villena-Martínez *et al.*, "A quantitative comparison of calibration methods for RGB-D sensors using different technologies," *Sensors*, vol. 17, no. 2, p. 243, 2017.
[201] D. R. dos Santos, M. A. Basso, K. Khoshelham, E. de Oliveira, N. L. Pavan, and G. Vosselman, "Mapping indoor spaces by adaptive coarse-to-fine registration of RGB-D data," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 262–266, Feb. 2016.
[202] I. Villaverde and M. Graña, "Neuro-evolutionary mobile robot egomotion estimation with a 3D ToF camera," *Neural Comput. Appl.*, vol. 20, no. 3, pp. 345–354, 2011.
[203] M. Liu, X. Zhang, Y. Zhang, and S. Lyu, "Calibration algorithm of mobile robot vision camera," *Int. J. Precis. Eng. Manuf.*, vol. 17, no. 1, pp. 51–57, 2016.
[204] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.
[205] S. Gupta, R. Girshick, P. Arbeláez, and J. Malik, "Learning rich features from RGB-D images for object detection and segmentation," in *Proc. Eur. Conf. Comput. Vis.*, Zürich, Switzerland, Sep. 2014, pp. 345–360.
[206] D. Lin, S. Fidler, and R. Urtasun, "Holistic scene understanding for 3D object detection with RGBD cameras," in *Proc. Int. Conf. Comput. Vis.*, Sydney, NSW, Australia, Dec. 2013, pp. 1417–1424.
[207] S. Song and J. Xiao, "Sliding shapes for 3D object detection in depth images," in *Proc. Eur. Conf. Comput. Vis.*, Zürich, Switzerland, Sep. 2014, pp. 634–651.
[208] B.-S. Kim, S. Xu, and S. Savarese, "Accurate localization of 3D objects from RGB-D data using segmentation hypotheses," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Portland, OR, USA, Jun. 2013, pp. 3182–3189.
[209] S. Gupta, P. Arbeláez, R. Girshick, and J. Malik, "Aligning 3D models to RGB-D images of cluttered scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 4731–4740.
[210] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis.*, Florence, Italy, Oct. 2012, pp. 746–760.
[211] D. Fox, L. Bo, and D. Fox, "RGB-(D) scene labeling: Features and algorithms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Providence, RI, USA, Jun. 2012, pp. 2759–2766.
[212] S. Tang *et al.*, "Histogram of oriented normal vectors for object recognition with a depth sensor," in *Proc. Asian Conf. Comput. Vis.*, Daejeon, South Korea, Nov. 2012, pp. 525–538.
[213] K. Lai, L. Bo, X. Ren, and D. Fox, "Detection-based object labeling in 3D scenes," in *Proc. IEEE Int. Conf. Robot. Autom.*, Saint Paul, MN, USA, May 2012, pp. 1330–1337.

YU HE received the bachelor's and master's degrees from the Zhejiang University of Technology, in 2012 and 2015, respectively, where he is currently pursuing the Ph.D. degree in computer vision. His research interests include 3D vision, image processing, and robotics.

SHENGYONG CHEN (M'01–SM'11) received the Ph.D. degree in computer vision from the City University of Hong Kong, Hong Kong, in 2003. He was with the University of Hamburg, from 2006 to 2007. He is currently a Professor with the Tianjin University of Technology, China, and also with the Zhejiang University of Technology, China. He is a Fellow of IET and a Senior Member of IEEE and CCF. He has published over 100 scientific papers in international journals. His research interests include computer vision, robotics, and image analysis. He received a Fellowship from the Alexander von Humboldt Foundation of Germany. He received the National Outstanding Youth Foundation Award of China, in 2013.