Technical Report

IRI-TR-10-07

Exploitation of Time-of-Flight (ToF) Cameras

IRI Technical Report

Sergi Foix
Guillem Alenyà
Carme Torras

Institut de Robòtica i Informàtica Industrial

Abstract

This technical report reviews the state-of-the art in the field of ToF cameras, their advantages, their limitations, and their present-day applications sometimes in combination with other sensors. Even though ToF cameras provide neither higher resolution nor larger ambiguity-free range compared to other range map estimation systems, advantages such as registered depth and intensity data at a high frame rate, compact design, low weight and reduced power consumption have motivated their use in numerous areas of research. In robotics, these areas range from mobile robot navigation and map building to vision-based human motion capture and gesture recognition, showing particularly a great potential in object modeling and recognition.

Institut de Robòtica i Informàtica Industrial (IRI)
Consejo Superior de Investigaciones Científicas (CSIC)
Universitat Politècnica de Catalunya (UPC)
Llorens i Artigas 4-6, 08028, Barcelona, Spain

Tel (fax): +34 93 401 5750 (5751)
`http://www.iri.upc.edu`

Corresponding author:

S Foix
tel: +34 93 401 5791
`sfoix@iri.upc.edu`
`http://www.iri.upc.edu/staff/sfoix`

1 Introduction

The distinctive characteristics of ToF cameras have proved to give important advantages in several fields. We can classify the wide range of applications where ToF sensors are used by considering their scenario of application, yielding *scene-related tasks*, *object-related tasks* and *applications involving humans*. *Scene-related tasks* generally involve mobile robots and large displacements, while *object-related tasks* involve instead robotic arms or humanoid-like robots, and small depths. To give a comprehensive overview we also include *applications involving humans*, since a lot of work has been done in face, hand, and body recognition with applications to man-machine interfaces.

The first works in these areas are comparisons between ToF and other technologies. Then new works appear where these technologies are gradually complemented, and sometimes substituted, by ToF sensors.

A table is provided for each section to summarise and give a comprehensive view of the section contents.

Our conclusion is that the most exploited feature of ToF cameras is their capability of delivering complete scene depth maps at high frame rate without the need of moving parts. Moreover, foreground/background segmentation methods based on depth information are quite straightforward, so ToF images are used in many applications requiring them. A good characteristic is that geometric invariants as well as metric constraints can be naturally used with the ToF depth images.

The depth-intensity image pair is also often used, exploiting the fact that both images are delivered already registered. In applications where the reduced resolution of a ToF camera is critical, it is complemented with other sensors, usually color cameras. ToF cameras are used in human environments because they are eye-safe and permit avoiding physical contact and dedicated markers or hardware.

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

2 Exploitation of ToF Camera Advantages

2.1 Scene-related tasks

This kind of applications deal with tasks involving scenes that contain objects like furniture and walls. Observe that the expected range of distances to these objects is relatively wide. A usual framework in these applications is to install the camera on a mobile robot and use it for robot navigation and mapping. As it will be seen, one of the areas where ToF sensors are adequate is in obstacle avoidance, because the detection region is not only horizontal (like in laser scanners) but also vertical, allowing to detect obstacles with complex shapes. Clearly, *the most appreciated characteristic of ToF sensors here is the high frame rate* (see Table 1). Some applications also benefit from the metric information obtained with depth images.

Comparison. The first works were devoted to the comparison of ToF with other sensors, mainly laser scanners. Thanks to the larger vertical field of view of ToF cameras, difficult obstacles (like tables) are better detected by a ToF camera than by a 2D laser scanner. For example, Weingarten *et al.* [67] demonstrated this in the context of an obstacle avoidance algorithm.

To obtain a comparable detection area, a 3D scanner can be built from a pivoted 2D laser scanner. May *et al.* [44, 42] compared the performance of their robot navigation algorithm with such sensor and with a ToF camera. One of the main difficulties they encountered is the

Table 1: Scene-related tasks

Article	**Topic**	**Advantages**	**Type of Sensor**
Weingarten *et al.* [67]	Obstacle avoidance in static env.	3D at high rate	SR2 (depth)
May *et al.* [44, 42]	3D mapping	3D at high rate/No required Pan-Tilt	SR2 (depth)
May *et al.* [43]	Pose estimation/3D mapping	Registered depth-intensity	SR3 (depth + intensity)
Hedge and Ye [28]	Planar feature 3D mapping	3D at high rate/No required Pan-Tilt	SR3
Ohno *et al.* [47]	3D mapping	3D at high rate	SR2
Stipes *et al.* [59]	3D mapping / Point selection	Registered depth-intensity	SR3
May *et al.* [41]	3D mapping/SLAM	3D at high rate	SR3
Gemeiner *et al.* [18]	Corner filtering	Registered depth-intensity	SR3 (depth + intensity)
Thielemann *et al.* [64]	Navigation in pipelines	3D allow geometric primitives search	SR3
Sheh *et al.* [56]	Navigation in hard env.	3D at high rate	SR3 + inertial
Swadzba *et al.* [62]	3D mapping in dynamic env.	3D at high rate/Registered depth-intensity	SR3 (depth + intensity)
Acharya *et al.* [2] Gallo *et al.* [17]	Safe car parking	Improved depth range/3D at high rate	Canesta
Gortuk *et al.* [20]	Object classification (airbag app.)	light/texture/shadow independence	Canesta
Yuan *et al.* [68]	Navigation and obst. avoidance	Increased detection zone	SR3 + laser
Kuhnert and Stommel *et al.* [37]	3D reconstruction	Easy color registration	PMD + stereo
Netramai *et al.* [46]	Motion estimation	3D at high rate	PMD + stereo
Huhle *et al.* [31]	3D mapping	Easy registration of depth and color	PMD + color camera
Prusak et al. [52]	Obst. avoidance/Map building	Absolute scale/better pose estimation	PMD + spherical camera
Swadzba *et al.* [63]	3D mapping/Map optimisation	3D at high rate	SR3
Vaskevicius *et al.* [66] Poppinga [50]	Localization/Map optimisation	Neighbourhood relation of pixels No color restrictions	SR3

accumulated error in the map created with the ToF camera, leading to failures when closing loops, for instance. Compared to pivoted laser scanners, accumulated errors usually occur more often with ToF cameras due to their smaller field of view. As we will see in the next section, this problem is also present in modeling objects tasks.

Only ToF. ToF sensors have been used successfully as the unique sensor in some mobile robotic applications, despite their characteristic limited resolution. For mapping purposes, *ToF sensors are very interesting because they allow to extract geometric features.* Most of the reviewed applications extract planar regions using both intensity and depth images. In [43], May *et al.* explored different methods to improve pose estimation. They propose additionally a final refinement step that involves the alignment of corresponding surface normals leading to improved 3D scene maps computed at frame rate. The normal of the extracted planes is also used by Hedge and Ye [28] to detect badly conditioned plane detection, as horizontal planes in a staircase. Also Pathak *et al.* [48] have reported the use of ToF to extract planes for 3D mapping.

Alternatively, the acquired crude point clouds can be processed by a variant of the Iterative Closest Point (ICP) algorithm to find the relation between two point clouds. For example, a real time 3D map construction algorithm is proposed by Ohno *et al.* [47] in the context of a snake-like rescue robot operating in complex environments, like rubble in disaster-like scenarios. Here, a modification of the classical ICP algorithm is proposed to cope with ToF noisy readings and to speed up the process.

Another adaptation of an ICP-like algorithm for ToF images is presented by Stipes *et al.* [59], where both the depth and the intensity images are used. They present a probabilistic point sampling process to obtain significant points used in the registration process.

ICP assumes that both point clouds overlap, so wrong depth points can skew the result. May *et al.* [41] presented an ICP variant to take this explicitly into account. They propose a mapping algorithm using a Simultaneous Localization and Mapping (SLAM) technique to reduce the reconstruction error that is specially useful when a zone of the scenario is revisited, i. e., when closing a loop.

Also with potential applications to SLAM, Gemeiner *et al.* [18] proposed a corner filtering scheme combining both the intensity and depth image of a ToF camera.

Complex environments are a good test field for ToF sensors, as they are capable of naturally recovering their geometry. In the context of pipeline inspection, Thielemann *et al.* [64] have proposed to use a ToF camera to detect the different junctions based not on appearance but on geometric properties. Here the self-illumination mechanism of ToF sensors is appreciated. Furthermore, Sheh *et al.* [56] have proposed a ToF based navigation system for a random stepfield terrain[1]. They use the depth information to color an array of pixels and then perform some classical edge detection algorithms in this array, that is called *heightfield.* The heading and attitude compensation of the image is performed using an inertial unit.

ToF sensors have proved to be also applicable in dynamic environment mapping thanks to their characteristic high frame rate. Swadzba *et al.* [62] present a scene reconstruction algorithm that discards dynamic objects, like pedestrians, using a static camera in the difficult case of short sequences (2-3 sec.). Motion is recovered via optical flow in the intensity images, and then transferred to the depth image to compute a 3D velocity vector.

ToF cameras have been employed also in the automotive field to assist in parking operations. In [2] Acharya *et al.* describe the system design of a ToF camera for backup obstacle detection. In [17] the same group presents an application of a similar camera for the detection of curves and ramps also in parking settings. A modified Ransac algorithm, that uses only the best inliers, is used to find the best fitting of the planar patches that model the environment. ToF has been used also to control the deployment of the airbag system depending on the nature of the occupant in a car [20]: adult, child, child seat or objects.

[1]Stepfield terrains are the NIST proposal to generate repeatable terrain for evaluating robot mobility.

Fusion with other sensors. Some other authors have started recently to fuse ToF cameras with other sensors, i.e. laser scanners, and different types of color cameras. A simple approach is to integrate ToF into existing algorithms. For example, Yuan *et al.* [68] propose a fusion process to integrate 3D data in the domain of laser data by projecting ToF point clouds onto the laser plane. This is applicable when considering a simple shaped robot, i.e. one that can be approximated by a cylinder, and it is a minimal update to their previous laser-scanner-based algorithm. Nevertheless, the resulting algorithm can cope with new kinds of obstacles in a simple way. Note that this is not a pure 3D approach and it is not using the potentiality of having full 3D information at a high frame rate.

Fusion of color and depth information in scene tasks seems to have a great potential. In a preliminary work, Kuhnert and Stommel [37] present a revision of their 3D environment reconstruction algorithm combining information from a stereo system and a ToF sensor. Later, Netramai *et al.* [46] compared the performance of a motion estimation algorithm using both ToF and depth from stereo. They also presented an oversimplified fusion algorithm that relies on the optical calibration of both sensors to solve the correspondence problem. These works propose fusion paradigms combining the results produced in two almost independent processes.

Contrarily, Huhle *et al.* [31] present a color-ICP algorithm useful for scene-based image registration, showing that introducing color information from a classical camera in the beginning of the process effectively increases the registration quality.

Depth information allows to identify in a robust manner not only obstacles but also holes and depressions. Prusak *et al.* [52] proposed a join approach to pose estimation, map building, robot navigation and collision avoidance. The authors use a PMD camera combined with a high-resolution spherical camera in order to exploit both the wide FOV of the latter for feature tracking and pose estimation, and the absolute scale of the former. The authors relied on a previous work on integration of 2D and 3D sensors [51, 60], showing how restrictions of standard Structure-from-Motion (SfM) approaches (mainly scale ambiguity and the need for lateral movement) could be overcome by using a 3D range camera. The approach produced 3D maps in real-time, up to 3 frames per second (fps), with an ICP-like algorithm and an incremental mapping approach.

Noisy data enhancement. Swadzba *et al.* [63] propose a new algorithm to cluster redundant points using a virtual plane, which apparently gives better results in planar regions and reduces noise, improving registration results. Furthermore, a group at Jacobs University [66, 50] has proposed to identify surfaces using a region growing approach that allows the poligonization of the resulting regions in an incremental manner. The nature of the information delivered by ToF cameras, specially the neighbourhood relation of the different points, is explicitly exploited and also their noisy nature is taken into account. Moreover, some comparisons with results from stereo rigs are reported.

Finally, Huhle *et al.* [30] propose an alternative representation of the map by means of the Normal Distribution Transform, which efficiently compresses the scan data reducing memory requirements. This representation seems to be well suited also for the typical noisy ToF depth images.

2.2 Object-related tasks

ToF cameras have also been successfully used for object and small surface reconstruction, where the range of distances is small. It is expected that some oversaturation problems occur when acquiring depth images. Contrarily, as the range of depths is short, some calibration processes can be simplified. In general the scenario for these applications involves a robotic manipulator or a human-like robot with the task of modeling the object shape.

As before, ICP-like techniques are the preferred solution to reconstruct the surfaces. A

Table 2: Object-related tasks

Reference	Topic	Advantages	Type of Sensor
Ghobadi *et al.* [19]	Dynamic object detection and classification	Color and light independence	PMD
Hussmann and Liepert [32]	Object pose	Easy object/background segmentation	PMD
Guomundsson *et al.* [24]	Known object pose estimation	Light independent / Absolute scale	SR3
Beder *et al.* [3]	Surface reconstruction using patchlets	ToF easily combines with stereo	PMD
Fuchs and May [14]	Precise surface reconstruction	3D at high rate	SR3/O3D100 (Depth)
Dellen *et al.* [9] Foix *et al.* [12]	3D object reconstruction	3D at high rate	SR3 (Depth)
Kuehnle *et al.* [36]	Object recognition for grasping	3D allow geometric primitives search	SR3
Grundmann *et al.* [22]	Collision free object manipulation	3D at high rate	SR3 + stereo
Reiser and Kubacki [53]	Position based visual servoing	3D is simply obtained / No model needed	SR3 (Depth)
Gachter *et al.* [16] Shin *et al.* [57]	Object part detection for classification	3D at high rate	SR3 SR2
Marton *et al.* [40]	Object categorisation	ToF easily combines with stereo	SR4 + color
Saxena *et al.* [54]	Grasping unknown objects	3D at high rate	SR3 + stereo
Zhu *et al.* [69]	Short range depth maps	ToF easily combines with stereo	SR3 + stereo
Lindner *et al.* [38]	Object segmentation for recognition	Easy color registration	PMD + color camera
Fischer *et al.* [11]	Occlusion handling in virtual objects	3D at high rate	PMD + color camera

common technique to identify objects is Support Vector Machines (SVM), which performs adequately when considering the noisy point models obtained with one ToF image or when merging different ToF views.

A comprehensive summary is given in Table 2, where we can observe that the high frame rate of ToF sensors is a key advantage, but also the natural combination with color cameras and stereo rigs. Here, contrary to Table 3 in the next section, the intensity image provided by the ToF camera is not much used, preferring the combination with high resolution conventional cameras.

Comparison. A classical solution in this area is the use of calibrated stereo rigs. Therefore, first works were devoted to their comparison with ToF sensors, showing the potential of the latter when poorly textured objects are considered, and when background-foreground segmentation is difficult. In planar and untextured object surfaces, where stereo techniques clearly fail, Ghobadi *et al.* [19] compared the results of a dynamic object detection algorithm based on SVM using stereo and ToF depth images. In the same manner, Hussmann and Liepert [32] also compared ToF and stereo vision for object pose computation. The key difference favourable to ToF camera is its ability to effectively segment the object and the background, even if their color or texture is exactly the same (i.e. a white object in a white table). They also propose a simple method to obtain object pose from a depth image.

Another comparison is presented by Guomundsson *et al.* [24]. They classify and estimate the pose of some simple geometric objects using a Local Linear Embedding (LLE) algorithm, and compare the results of using the intensity image and the depth image. Their analysis shows that range data adds robustness to the model, simplifies some preprocessing steps, and in general the generated models capture better the nature of the object. Stereo and ToF have also been compared by Beder *et al.* [3] in the framework of surface patchlet identification and pose estimation. In their setup, using a highly textured surface for stereo experiments, ToF slightly outperforms stereo in terms of depth and normal direction to the patchlet. Thus, ToF can be used to benchmarking stereo surface reconstruction algorithms.

ToF for surface reconstruction. To obtain 3D object surfaces, multiple 3D images need to be taken and the resulting 3D point clouds should be combined. The setups for these object modeling algorithms usually include a ToF camera mounted on the end-effector of a robotic arm. *Point cloud registration is more critical in object modeling than in scene modeling.* Even if the hand-eye system is precisely calibrated, the displacement given by the robot is usually not enough and the transformation between different point clouds has to be calculated. The application of ICP in two consecutive views naturally accumulates errors and consequently more precise algorithms need to be used.

To obtain precise object models, Fuchs and May [14] perform a circular trajectory around the object to acquire equally spaced images, and use a *simultaneous matching* algorithm [61] instead of classical ICP to distribute the errors in all the estimated displacements. Their work also includes a comparison of two different ToF cameras. Alternatively, Dellen *et al.* [9] propose a fine registration algorithm based on an ICP algorithm using invariant geometric features. The resulting model (Fig. 1(d)) is obtained after reducing noise and outliers by treating the coarse registered point cloud (Fig. 1(b)) as a system of interacting masses connected via elastic forces. More recently, Foix *et al.* [12] propose a method to compute the covariance of the point clouds registration process (ICP), and apply an iterative view-based aggregation method to build object models under noisy conditions. Their method does not need accurate hand-eye calibration since it uses globally consitent probilistic data fusion by means of a view-based information-form SLAM algorithm [65], and can be executed in real time taking fully advantage of ToF camera high frame rate.

ToF for object manipulation. Object recognition and object pose estimation algorithms are usually related to robotic manipulation frameworks: objects have to be identified or categorised with the aim of finding and extracting some characteristics that permit interacting with them. This is usually a challenging task as ToF depth images are noisy, and low sensor resolution leads to only few depth points per object.

Kuehnle *et al.* [36] explore the use a ToF camera to recognise and locate 3D objects in the framework of the robotic manipulation system DESIRE. Objects are modelled with geometric primitives. Although they use depth images rectified up to some level, their system is not reliable enough. In a subsequent work [22] they use the ToF camera to detect unknown objects and classify them as obstacles, and use a stereo camera system to identify known objects using SIFT features. As it is widely known, this second approach requires textured objects while their first approach does not. In the same project, Reiser and Kubacki [53] have proposed a method to actively orientate the camera using a visual servoing approach to control a pan-and-tilt unit. They proved that position-based visual servoing is straightforward by using a ToF camera, because of its ability to deliver 3D images at high rate.

In a different way, Gächter *et al.* [16] propose to detect and classify objects by identifying their different parts. For example, chairs are modelled by finding their legs, which in turn are modelled with vertical bounding boxes. The tracking of the different parts in the image sequence is performed using an extended particle filter, and the recognition algorithm is based on a SVM, that proves again to be useful in typical noisy ToF images. Later, Shin *et al.* [57] used this incremental part detector to propose a classification algorithm based on a geometric grammar. However, they use a simulated environment because the classification in real scenarios does not seem to be reliable enough.

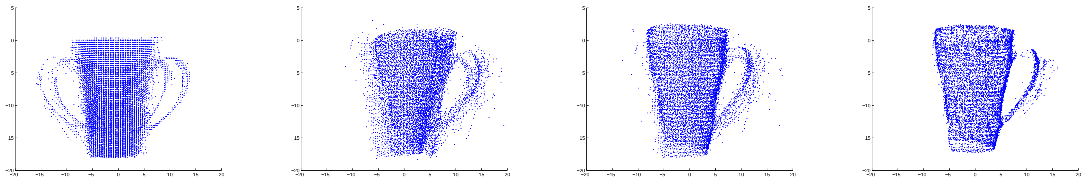
Depth information is very useful in cluttered environments to detect and grasp unknown objects: the 3D region of interest can be extracted easily, and some object segmentation algorithms can be developed combining cues from both a ToF sensor and a color camera. Using such a combined sensor, Marton *et al.* [40] proposed a probabilistic categorisation algorithm for kitchen objects. This work uses a new SR4000 camera. This sensor assigns a confidence value to each depth reading that allows to infer if the object material is producing bad sensor readings.

Thanks to the depth information, some grasping properties can be easier evaluated, i.e. form- and force-closure, sufficient contact with the object, distance to obstacles, and distance between the center of the object and the contact point. Saxena *et al.* [54] used this advantage to propose a learning grasp strategy that identifies good grasp points using partial shape information of unknown objects. The contribution of the depth information allows to update an already presented method using a color camera, with the advantage of having depths even in textureless portions of the objects.

Fusion algorithms. In fact, ToF and stereo systems naturally complement one another. As has been argued before, ToF performs correctly in poorly textured surfaces and object segmentation becomes easy even in poorly contrasted situations. Contrarily, it has difficulties precisely in textured surfaces and in short distances, where stereo outperforms it. This fact has been exploited in several works. For example, Zhu *et al.* [69] propose a probabilistic framework to fuse depth maps from stereo and the ToF sensor. They use a depth calibration method to improve the ToF image, which is useful in small depth ranges (from 1m to 1.4m).

Another fusion framework is proposed by Lindner *et al.* [38] using calibration and scaling algorithms. They obtain a dense colored depth map using the geometrical point correspondence between the ToF and color cameras by assigning a color to the ToF depth points, and interpolating the depth of the rest of the color camera pixels. A way to detect areas not seen by the color camera is also provided, as well as some techniques to enhance edges and detect invalid pixels.

Finally, in the context of augmented reality, Fischer *et al.* [11] combine a ToF camera and a standard color camera to handle virtual object occlusions caused by real objects in the scene. Fast 3D information is highly valuable, but also the independence on lightning conditions, object texture and color. They do not use any depth calibration or noise outlier removal algorithm, and consequently the negative effect of noise is clearly visible in their results.



(a) Raw data (b) Coarse registration (c) Fine registration (ICP) (d) Spring-mass model result

Figure 1: Modeling process of Swissranger data taken from a mug. (a) Raw data with superimposed unmerged views. (b) Coarse registration. (c) Fine registration (ICP). (d) Final result using a spring-mass model.

2.3 Tasks involving humans

One of the areas where the use of ToF cameras is most active is in human activity recognition and man-machine interaction. A recent survey on ToF sensors with special attention to 3D graphics and realism has been recently presented [35]. Here we like to concentrate on technologies appropriate for man-machine interaction. One important characteristic of ToF cameras appreciated in this area is their being a non-invasive technology, contrary to the widely extended use of special gloves, artificial marks, special skin color or special attached devices. ToF camera also offers the advantaged that no special background is needed.

In contrast to the preceding section, here we observe the use of many different camera prototypes (Table 3). Again, the ToF high frame rate is highly appreciated, as most of the applications we review involve tracking (see Table 3). We observe also that most of the methods rely on depth but also on appearance. Hence, the intensity image delivered by the ToF sensor is sometimes used. Alternatively, to obtain higher resolution, depth is combined with color cameras and stereo rigs.

People tracking. ToF sensors have been used to perform people tracking, with applications, for instance, to common path detection and activity understanding. One common way is to place the sensor in a zenithal configuration and fix the attention in the person head. Following this idea, a single person tracking algorithm is presented by Gokturk *et al.* [21] which uses depth signatures combined with a clustering algorithm to identify the target, useful even when partial occlusions or partial out of image situations occur. This algorithm is possible because ToF sensors deliver a depth image from which it is possible to infer geometry and 3D location.

Along the same line, a multiple people tracking algorithm has been proposed by Bevilacqua *et al.* [4] as an update of a stereo based algorithm. Changing illumination conditions are specifically tested and it is proved that ToF camera performs adequately also in this situation.

Alternatively, Guomundsson *et al.* [23] use a ToF camera in a smart room environment to enhance their foreground/background segmentation algorithm, based on a Shape from Silhouettes (SfS) method, with the objective of segmenting people. Here different cameras are used, placed in zenithal position, but also placed elevated and not pointing vertical.

With the camera also elevated, Kahlmann *et al.* [33] presented a person tracking algorithm based on a particle filter. The segmentation is performed in the depth image, as with their algorithms an intensity image doesn't offer enough invariant characteristics. Using the depth of the segmented points a reduced range histogram is created, which is used for the similarity

Table 3: Tasks involving humans

Article	Topic	Advantages	Type of Sensor
Gokturk *et al.* [21]	Head tracking	Easy background segmentation Easy geometrical data	Canesta
Bevilacqua *et al.* [4]	Multiple person tracking	Light independent	Canesta
Guomundsson *et al.* [23]	People in smart room	Easy background segmentation/Light independent	SR3+stereo
Kahlmann *et al.* [33]	People tracking	3D at high rate	SR3
Haker *et al.* [25]	Human nose detector	3D at high rate/Registered depth-intensity	SR3(depth + intensity)
Haker *et al.* [26]	Face detection	3D at high rate/Registered depth-intensity	SR3 (depth + intensity)
Hansen *et al.* [27]	Face detection	3D at high rate/Registered depth-intensity	SR3 (depth + intensity)
Nanda *et al.* [45]	Hand tracking	Easy segmentation in cluttered env.	prototype
Liu and Fujimura [39] Fujimura and Liu [15]	Hand gesture recognition	Easy segmentation Robustness to illumination changes	prototype
Breuer *et al.* [6]	7 dof hand motion tracking	Easy background segmentation Skin color and light independent	SR2 (depth)
Soutschek *et al.* [58]	Gesture-based user interface	Non invasive/metric information	SR3 (depth + intensity)
Penne *et al.* [49]	Human respiratory motion analysis	3D at high rate/Multiple detection areas	SR3
Bianchi *et al.* [5]	video editing	Easy background segmentation Easy color registration	SR3+color camera
Frick *et al.* [13]	3DTV	Easy background segmentation	SR3
Cho *et al.* [7]	Virtual 3D actor generation	3D at high rate/Easy background segmentation	HDTV + color camera
Knoop *et al.* [34]	3D human body tracking	3D at high rate/Non invasive	SR2 + stereo
Holte *et al.* [29]	Body gesture recognition	Registered depth-intensity	SR2 (depth + intensity)
Dariush *et al.* [8]	Human-robot motion transfer	3D at high rate	SR3

measurement. Thus, obtaining 3D information at high rate is crucial for this algorithm.

Human face detection. Some work has been carried out in the area of human face detection and tracking. Here, one of the most interesting characteristic of ToF sensors is that the *light that is emitted is eye-safe* and thus can be directly pointed to the person face. Haker *et al.* [25] presented a face detector algorithm based on the detection of the human nose. The key point of the algorithm is the combination of intensity and depth based detectors, that can be easily performed thanks to the already registered images provided by the ToF sensor. Later, they show that some scale-invariants can be formulated by expressing the depth image in the frequency domain [26]. They propose a new face detector combining both mentioned approaches that also benefits from the already registered intensity and depth images.

Boost-based algorithms are one of the most popular face detection methods. Hansen *et al.* [27] demonstrate that boost classifiers can be used in the low-resolution intensity images typically delivered by ToF sensors. Moreover, they propose to use the depth information as an additional cue to resolve ambiguities, exploiting again the advantage of ToF sensors of having intensity and depth images registered by construction.

Hand tracking. Another area of interest is hand tracking for gesture recognition. Some initial works on hand tracking [45], hand gesture [39] and sign recognition [15] use the ToF sensor to segment the hand from the background, and then use the intensity image to recover the hand configuration based on an appearance algorithm. In these works a prototype of a ToF camera was used, which later has evolved to ZCam [1].

Breuer *et al.* [6] present a system that recognises 7dof hand movements, including translation, rotation and scaling. A first crude estimation of hand position is performed using PCA, and then a fine matching involving a model of the hand is performed. ToF has the advantage that no special background or skin color is needed for segmentation, besides its robustness to illumination changes. This system does not include the detection of finger motion, and thus cannot identify gestures or signs.

An algorithm to identify gestures has been recently presented by Soutschek *et al.* [58] using a ToF camera system to build a gesture-based user-interface for 3D medical data exploration. The goal of this application is to preserve the sterility of surgeons by eliminating physical interaction with the system. ToF is used because, as stated before, hand segmentation from 3D data is easy and independent of appearance. However, this application requires the computation not only of the distinctive sign but also the hand translation and rotation. This can be accomplished with the ToF camera in a very natural way at a high frame rate.

Body gesture. ToF cameras have proved useful also to detect and track not only the head or hands but also the full human body. In [34], Knoop *et al.* present a human body tracking system based on an articulated 3D body modelled using cylinders. The use of 3D images is a key difference with respect to other tracking methods, and the high frame rate exhibited by the ToF camera is crucial. The authors propose also a fusion framework and use a stereo camera rig to improve tracking results.

Moreover, Holte *et al.* [29] propose a body gesture recognition algorithm that uses simultaneously the intensity and depth images. Gestures are characterised by motion primitives in the 3D data, represented compactly using *harmonic shape context*, a kind of spherical histograms. The use of 3D data delivered by a ToF sensor permits the definition of view-invariant motion primitives. Gestures are defined as a sequence of primitives, solving the problem of deciding when a gesture begins and ends. The intensity image is used to compute a Region of Interest (ROI) that excludes false readings in the edges of the body. As intensity and depth are already registered, depth filtering is straightforward.

Furthermore, motion can be transferred from a human to a humanoid robot in different ways. Most solutions to the motion-re-targeting problem are offline approaches based on pre-recorded human motion data collected with marker-based capture systems. In [8] an on-line solution is described with an algorithm not relying on markers placed on anatomical landmarks and not requiring special instrumentation, but on a single ToF camera. A re-targeting module enforces self-collision constraints and demonstrates its usefulness on a Honda ASIMO humanoid robot. Here again, the use of a single ToF camera simplifies data extraction in comparison with alternative stereo systems.

In the same group, at the University of Erlangen-Nuremberg, some research is being conducted on the use of ToF sensors in the area of medical applications, like patient positioning [55]. Recently, they have presented a non invasive method to detect the respiratory motion of humans in real time [49], with potential applications in the reduction of artifacts present in image-based medical techniques like tomography. The system allows to measure the motion of different areas at the same time, e.g. abdomen and torso, by fitting different planes to the different zones. This fitting process allows also to explicitly treat the noisy ToF images, and thus improve the stability of the process.

3 Conclusions

This survey has covered the topic of ToF cameras from different perspectives, including: underlying principle and characteristics, calibration techniques, applications where camera advantages are explicitly exploited, and potential for future research. Near one hundred publications in recognised conference proceedings and journals have been contrasted in order to give a comprehensive overview of key advances in the field, current research concerns, present-day applications and future lines of investigation. Over the last years, performance of ToF cameras has improved significantly; errors have been minimised and higher resolution and frame rates have been obtained. Although ToF cameras cannot yet attain the depth accuracy offered by other types of sensors such as laser scanners, plenty of research demonstrates that they perform better in many robotic applications. The application of ToF cameras in the wide range of scientific areas we have reviewed indicates their great potential, and widens the horizon of possibilities that were envisaged in the past for vision-based robotics research.

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

Finally, some broad challenges need to be mentioned. First, resolution is still generally low for ToF cameras, despite some efforts have already led to better resolutions as explained above. Second, short integration times produce strong noise ratio, and high integration times can result in pixel saturation [10]. Although some algorithms dealing with this problem have already been proposed, more research is needed in this direction. Third, an important issue for ToF cameras is the aliasing effect, a consequence of the periodicity of the modulated signal. Distances to objects that differ 360° in phase are indistinguishable. Use of multiple modulated frequencies can be a solution here, or lowering the modulation frequency since it would increase the unambiguous metric range.

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

References

[1] ZCam, http://www.3dvsystems.com. 3DV Systems, 2009.

[2] S. Acharya, C. Tracey, and A. Rafii. System design of time-of-flight range camera for car park assistand backup application. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1552–1557, Anchorage, June 2008.

[3] C. Beder, B. Bartczak, and R. Koch. A comparison of PMD-cameras and stereo-vision for the task of surface reconstruction using patchlets. In *Proceedings of the 21st IEEE Conference on Computer Vision and Pattern Recognition*, volume 1-8, pages 2692–2699, Minneapolis, June 2007.

[4] A. Bevilacqua, L. Di Stefano, and P. Azzari. People tracking using a time-of-flight depth sensor. In *Proceedings of the IEEE International Conference on Video Signal Based Surveillance*, pages 89–89, Sydney, Nov. 2006.

[5] Luca Bianchi, Riccardo Gatti, Luca Lombardi, and Paolo Lombardi. Tracking without background model for time-of-flight cameras. In *Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology (PSIVT)*, volume 5414, pages 726–737, Tokyo, Jan. 2009.

[6] Pia Breuer, Christian Eckes, and Stefan Müller. Hand gesture recognition with a novel IR time-of-flight range camera - A pilot study. In *Proceedings of the 3rd International Conference on Computer Vision/Computer Graphics*, volume 4418, pages 247–260, Rocquencourt, Mar. 2007.

[7] Ji-Ho Cho, Sung-Yeol Kim, Yo-Sung Ho, and Kwan Lee. Dynamic 3D human actor generation method using a time-of-flight depth camera. *IEEE Transactions on Consumer Electronics*, 54(4):1514–1521, Nov. 2008.

[8] B. Dariush, M. Gienger, A. Arumbakkam, C. Goerick, Youding Zhu, and K. Fujimura. Online and markerless motion retargeting with kinematic constraints. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1-3, pages 191–198, Nice, Sep. 2008.

[9] B. Dellen, G. Alenyà , S. Foix, and C. Torras. 3D object reconstruction from Swissranger sensors data using a spring-mass model. In *Proceedings of the 4th International Conferemce on Computer Vision Theory and Applications*, volume 2, pages 368–372, Lisbon, February 2009.

[10] P. Einramhof, S. Olu, and M. Vincze. Experimental evalutation of state of the art 3D-sensors for mobile robot navigation. In *Proceedings of the in the 31st Austrian Association for Pattern Recognition Workshop*, Krumbach, May 2007.

[11] Jan Fischer, Benjamin Huhle, and Andreas Schilling. Using time-of-flight range data for occlusion handling in augmented reality. In *Proceedings of the Eurographics Symposium on Virtual Environments*, pages 109–116, September 2007.

[12] S. Foix, G. Alenyà, J. Andrade-Cetto, and C. Torras. Object modeling using a ToF camera under an uncertainty reduction approach. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 1306–1312, Anchorage, May 2010.

[13] A. Frick, F. Kellner, B. Bartczak, and R. Koch. Generation of 3D-TV LDV-content with time-of-flight camera. In *Proceedings of the IEEE International Conference on 3DTV*, pages 45–48, Potsdam, May 2009.

[14] Stefan Fuchs and Stefan May. Calibration and registration for precise surface reconstruction with time of flight cameras. *International Journal on Int. Syst. Technology Applications*, 5(3-4):274–284, 2008.

[15] K. Fujimura and X. Liu. Sign recognition using depth image streams. In *Proceedings of the 7th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 381–386, Southampton, April 2006.

[16] S. Gächter, A. Harati, and R. Siegwart. Incremental object part detection toward object classification in a sequence of noisy range images. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 1-9, pages 4037–4042, Pasadena, May 2008.

[17] O. Gallo, R. Manduchi, and A. Rafii. Robust curb and ramp detection for safe parking using the Canesta ToF camera. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1558–1565, Anchorage, June 2008.

[18] P. Gemeiner, P. Jojic, and M. Vincze. Selecting good corners for structure and motion recovery using a time-of-flight camera. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5711–5716, Saint Louis, Oct. 2009.

[19] S. E. Ghobadi, K. Hartmann, W. Weihs, C. Netramai, O. Loffeld, and H. Roth. Detection and classification of moving objects-stereo or time-of-flight images. In *Proceedings of the International Conference on Computational Intelligence and Security*, volume 1, pages 11–16, Guangzhou, Nov. 2006.

[20] S. B. Gokturk and A. Rafii. An occupant classification system eigen shapes or knowledge-based features. In *Proceedings of the 19th IEEE Conference on Computer Vision and Pattern Recognition*, pages 57–57, San Diego, June 2005.

[21] S. B. Gokturk and C. Tomasi. 3D head tracking based on recognition and interpolation using a time-of-flight depth sensor. In *Proceedings of the 18th IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 211–217, Washington DC, June-July 2004.

[22] T. Grundmann, Z. Xue, J. Kuehnle, R. Eidenberger, S. Ruehl, A. Verl, R. D. Zoellner, J. M. Zoellner, and R. Dillmann. Integration of 6D object localization and obstacle detection for collision free robotic manipulation. In *Proceedings of the IEEE/SICE International Symposium on System Integration*, pages 66–71, Nagoya, Dec. 2008.

[23] Sigurjon Arni Guomundsson, Rasmus Larsen, H. Aanæs, M. Pardas, and J. R. Casas. ToF imaging in smart room environments towards improved people tracking. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1486–1491, Anchorage, June 2008.

[24] Sigurjon Arni Guomundsson, Rasmus Larsen, and Bjarne K. Ersboll. Robust pose estimation using the Swissranger SR-3000 camera. In *Proceedings of the 15th Scandinavian Conference on Image Analysis*, volume 4522, pages 968–975, Aalborg, June 2007.

[25] M. Haker, M. Bohme, T. Martinetz, and E. Barth. Geometric invariants for facial feature tracking with 3D ToF cameras. In *Proceedings of the International Symposium on Signals, Circuits and Systems*, volume 1-2, pages 109–112, Lasi, July 2007.

[26] M. Haker, M. Bohme, T. Martinetz, and E. Barth. Scale-invariant range features for time-of-flight camera applications. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1585–1590, Anchorage, June 2008.

[27] D. W. Hansen, R. Larsen, and F. Lauze. Improving face detection with ToF cameras. In *Proceedings of the International Symposium on Signals, Circuits and Systems*, volume 1-2, pages 225–228, Lasi, July 2007.

[28] G. Hedge and C. Ye. Extraction of planar features from Swissranger SR-3000 range images by a clustering method using normalized cuts. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4034–4039, Saint Louis, Oct. 2009.

[29] M. B. Holte, T. B. Moeslund, and P. Fihl. Fusion of range and intensity information for view invariant gesture recognition. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1531–1537, Anchorage, June 2008.

[30] B. Huhle, S. Fleck, and A. Schilling. Integrating 3D time-of-flight camera data and high resolution images for 3DTV applications. In *Proceedings of the 1st IEEE International Conference on 3DTV*, pages 289–292, Kos Isl, May 2007.

[31] Benjamin Huhle, Martin Magnusson, Wolfgang Strasser, and Achim J. Lilienthal. Registration of colored 3D point clouds with a kernel-based extension to the normal distributions transform. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 1-9, pages 4025–4030, Pasadena, May 2008.

[32] S. Hussmann and T. Liepert. Robot vision system based on a 3D-ToF camera. In *Proceedings of the 24th IEEE Instrumentation and Measurement Technology Conference*, volume 1-5, pages 1405–1409, Warsaw, May 2007.

[33] Timo Kahlmann, Fabio Remondino, and Sébastien Guillaume. Range imaging technology: new developments and applications for people identification and tracking. In *Proceedings of the 9th Videometrics Conference*, volume 6491, pages C4910–C4910, San Jose, Jan. 2007.

[34] S. Knoop, S. Vacek, and R. Dillmann. Sensor fusion for 3D human body tracking with an articulated 3D body model. In *Proceedings of the IEEE International Conference on Robotics and Automation*, volume 1-10, pages 1686–1691, Orlando, May 2006.

[35] A. Kolb, E. Barth, R. Koch, and R. Larsen. Time-of-flight sensors in computer graphics. In *Eurographics (State-of-the-Art Report)*, pages 119–134, 2009.

[36] J. U. Kuehnle, Z. Xue, M. Stotz, J. M. Zoellner, A. Verl, and R. Dillmann. Grasping in depth maps of time-of-flight cameras. In *Proceedings of the International Workshop on Robotic and Sensors Environments*, pages 132–137, Otawa, Oct. 2008.

[37] K. D. Kuhnert and M. Stommel. Fusion of stereo-camera and PMD-camera data for real-time suited precise 3D environment reconstruction. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1-12, pages 4780–4785, Beijing, Oct. 2006.

[38] Marvin Lindner, Martin Lambers, and Andreas Kolb. Sub-pixel data fusion and edge-enhanced distance refinement for 2D/3D. *International Journal on Int. Syst. Technology Applications*, 5(3-4):344–354, 2008.

[39] X. Liu and K. Fujimura. Hand gesture recognition using depth data. In *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*, pages 529–534, Seoul, May 2004.

[40] Z. Marton, R. Rusu, D. Jain, U. Klank, and M. Beetz. Probabilistic categorization of kitchen objects in table settings with a composite sensor. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 4777–4784, Saint Louis, Oct. 2009.

[41] S. May, S. Fuchs, D. Droeschel, D. Holz, and A. Nuechter. Robust 3D-mapping with time-of-flight cameras. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1673–1678, Saint Louis, Oct. 2009.

[42] S. May, K. Pervolz, and H. Surmann. *Vision systems: applications*, chapter 3D cameras: 3D computer vision of wide scope. I-Tech Education and Publishing, 2007.

[43] Stefan May, David Droeschel, Dirk Holz, Christoph Wiesen, and Stefan Fuchs. 3D pose estimation and mapping with time-of-flight cameras. In *Proceedings of the IEEE/RSJ IROS Workshop on 3D-Mapping*, Nice, Sep. 2008.

[44] Stefan May, Bjorn Werner, Hartmut Surmann, and Kai Pervolz. 3D time-of-flight cameras for mobile robotics. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1-12, pages 790–795, Beijing, Oct. 2006.

[45] K.; Nanda, H.; Fujimura. Visual tracking using depth data. In *Proceedings of the IEEE CVPR Workshops*, volume 3, pages 37–37, Washington, D. C., June 2004.

[46] Chayakorn Netramai, Melnychuk Oleksandr, Chanin Joochim, and Hubert Roth. Motion estimation of a mobile robot using different types of 3D sensors. In *Proceedings of the 4th International Conference on Autonomic and Autonomous Systems*, pages 148–153, Gosier, Mar. 2008.

[47] K. Ohno, T. Nomura, and S. Tadokoro. Real-time robot trajectory estimation and 3D map construction using 3D camera. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5279–5285, Beijing, Oct. 2006.

[48] K. Pathak, N. Vaskevicius, J. Poppinga, S. Schwertfeger, M. Pfingsthorn, and A. Birk. Fast 3D mapping by matching planes extracted from range sensor point-clouds. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1150–1155, Saint Louis, Oct. 2009.

[49] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert. Robust real-time 3D respiratory motion detection using time-of-flight cameras. *International Journal of Computer Assisted Radiology and Surgery*, 3(5):427–431, Nov. 2008.

[50] J. Poppinga, N. Vaskevicius, A. Birk, and K. Pathak. Fast plane detection and polygonalization in noisy 3D range images. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3378–3383, Nice, Sep. 2008.

[51] T.D.Arun Prasad, Klaus Hartmann, Wolfgang Weihs, Seyed Eghbal Ghobadi, and Arnd Sluiter. First steps in enhancing 3D vision technique using 2D/3D sensors. In *Computer Vision Winter Workshop*, pages 82–86, Prague, Feb. 2006.

[52] A. Prusak, O. Melnychuk, H. Roth, I. Schiller, and R. Koch. Pose estimation and map building with a time-of-flight camera for robot navigation. *International Journal on Int. Syst. Technology Applications*, 5(3-4):355–364, 2008.

[53] Ulrich Reiser and Jens Kubacki. Using a 3D time-of-flight range camera for visual tracking. In *Proceedings of the 6th IFAC/EURON Symposium on Intelligent Autonomous Vehicles*, Toulouse, Sept. 2007.

[54] A. Saxena, L. Wong, and A. Y. Ng. Learning grasp strategies with partial shape information. In *Proceedings of the 23th AAAI Conference on Artificial Intelligence*, pages 1491–1494, Chicago, Jul. 2008.

[55] C. Schaller, A. Adelt, J. Penne, and J. Hornegger. Time-of-flight sensor for patient positioning. In *Proceedings of the Medical Imaging: Visualization, Image-Guided Procedures, and Modeling*, volume 7261, Lake Buena Vista, Feb. 2009.

[56] R. Sheh, M. W. Kadous, C. Sammut, and B. Hengst. Extracting terrain features from range images for autonomous random stepfield traversal. In *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics*, pages 24–29, Rome, Sep. 2007.

[57] J. Shin, S. Gachter, A. Harati, C. Pradalier, and R. Siegwart. Object classification based on a geometric grammar with a range camera. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2443–2448, Kobe, May 2009.

[58] S. Soutschek, J. Penne, J. Hornegger, and J. Kornhuber. 3-D gesture-based scene navigation in medical imaging applications using time-of-flight cameras. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1538–1543, Anchorage, June 2008.

[59] J.A. Stipes, J.G.P. Cole, and J. Humphreys. 4D scan registration with the SR-3000 LIDAR. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2988–2993, Pasadena, May 2008.

[60] B. Streckel, B. Bartczak, R. Koch, and A. Kolb. Supporting structure from motion with a 3D-range-camera. In *Proceedings of the 15th Scandinavian Conference on Image Analysis*, volume 4522, pages 233–242, Aalborg, June 2007.

[61] Hartmut Surmann, Andreas Nüchter, and Joachim Hertzberg. An autonomous mobile robot with a 3d laser range finder for 3d exploration and digitalization of indoor environments. *Robotics and Autonomous Systems*, 45(3-4):181–198, 2003.

[62] A. Swadzba, N. Beuter, J. Schmidt, and G. Sagerer. Tracking objects in 6D for reconstructing static scenes. In *Proceedings of the IEEE CVPR Workshops*, volume 1-3, pages 1492–1498, Anchorage, June 2008.

[63] A. Swadzba, A. Vollmer, M. Hanheide, and S. Wachsmuth. Reducing noise and redundancy in registered range data for planar surface extraction. In *Proceedings of the 19th IAPR International Conference on Pattern Recognition*, volume 1-6, pages 1219–1222, Tampa, Dec. 2008.

[64] J.T. Thielemann, G.M. Breivik, and A. Berge. Pipeline landmark detection for autonomous robot navigation using time-of-flight imagery. In *Proceedings of the 22nd IEEE Conference on Computer Vision and Pattern Recognition*, volume 1-3, pages 1572–1578, Anchorage, June 2008.

[65] R. Valencia, E.H. Teniente, E. Trulls, and J. Andrade-Cetto. 3D mapping for urban service robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3076–3081, Saint Louis, October 2009.

[66] N. Vaskevicius, A. Birk, K. Pathak, and J. Poppinga. Fast detection of polygons in 3D point clouds from noise-prone range sensors. In *Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics*, pages 30–35, Rome, Sep. 2007.

[67] J.W. Weingarten, G. Gruener, and R. Siegwart. A state-of-the-art 3D sensor for robot navigation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 3, pages 2155–2160, Sendei, September 2004.

[68] F. Yuan, A. Swadzba, R. Philippsen, O. Engin, M. Hanheide, , and S. Wachsmuth. Laser-based navigation enhanced with 3D time of flight data. In *Proceedings of the IEEE International Conference on Robotics and Automation*, pages 2844–2850, Kobe, May 2009.

[69] Jiejie Zhu, Liang Wang, Ruigang Yang, and J. Davis. Fusion of time-of-flight depth and stereo for high accuracy depth maps. In *Proceedings of the 22nd IEEE Conference on Computer Vision and Pattern Recognition*, volume 1-12, pages 3262–3269, Anchorage, June 2008.

IRI reports

This report is in the series of IRI technical reports.
All IRI technical reports are available for download at the IRI website
`http://www.iri.upc.edu`.