The 9th International Symposium on Mobile Mapping Technology
MMT2015, 9-11 December 2015, Sydney, Australia

A long-range depth camera with configurable field of view and resolution

S. Montabone, D. Wood, R. Rafiei, M. Bishop
Ocular Robotics Limited, Unit F1, 13-15 Forrester Street, Kingsgrove, NSW 2208, Australia
{s.montabone, d.wood, r.rafiei, m.bishop}@ocularrobotics.com

Abstract

Building information rich 3D maps of the environment is an important task for mobile robotics, with applications in precision navigation, human-robot collaboration, mapping and telepresence. This paper presents for the first time a long-range configurable depth camera that is able to change its field of view and resolution on a frame by frame basis. The camera, which is based on a RobotEye RE05 3D laser scanner, overcomes key limitations present in current depth cameras by providing full control over its 360° × 70° field of view, down to 0.01°angular resolution, and the ability to image objects up to 250 m away from the camera at the highest resolution.

1 Introduction

Depth cameras, also known as time-of-flight (ToF) cameras, have been used in many robotics and computer vision applications to obtain a 3D model of the world through the depth maps they produce. By using this 3D information, autonomous systems can navigate, avoiding obstacles, and interact with the world. These cameras have the advantage of delivering depth information at high frame rates, however the resolution they provide is coarse, their field of view is fixed, and their imaging range is short [Foix *et al.*, 2011].

Building information rich 3D maps of the environment is critical for intelligent autonomous systems, with applications in precision navigation, human-robot collaboration, mapping and telepresence. In this paper, a depth camera sensor is presented based on a highly configurable 3D laser scanner which allows the user to define its scanning region along with the angular resolution on both horizontal and vertical axes. This long-range configurable depth camera has overcome the main limitations in currently available depth cameras.

In Section 2 the capabilities of commonly used depth cameras are presented. This is followed by a discussion on their limitations in Section 3. In Section 4 the long-range configurable depth camera is introduced and its 2D image acquisition is discussed in Section 5. Section 6 demonstrates how the configurable depth camera overcomes present-day limitations of commonly used depth sensors. This leads to Section 7 where novel applications for the configurable depth camera are presented and validated.

Sensor	Resolution	FoV	Range
SR4000-5m	176 × 148	69° × 55°	5 m
SR4000-10m	176 × 148	44° × 35°	10 m
CamCube 2.0	204 × 204	40° × 40°	7 m
Kinect	640 × 480	57° × 43°	5 m

Table 1: Summary of key features of commonly used depth cameras.

2 Commonly used depth cameras

Table 1 shows the resolution, field of view (FoV), and maximum imaging range for some of the most commonly used depth cameras. The SwissRanger SR4000 is a popular depth camera that uses the time-of-flight technique to obtain depth information. There are two options for the lens of the sensor, one that is able to measure up to 5 m and another which allows imaging at longer distances of up to 10 m. Even though there is an option for changing the focal length of the sensor, once the lens selection is made the field of view remains fixed. CamCube 2.0 provides the highest resolution in the time-of-flight type of depth cameras. The camera also has a fixed resolution and field of view, and is limited to imaging objects up to 7 m away. Kinect [Khoshelham and Elberink, 2012] projects an infrared pattern that is then detected by its infrared camera. By processing this information, Kinect is able to generate a depth map by triangulation which subsequently is interpolated to a fixed 640 × 480 depth image. As such, this depth resolution cannot be directly compared to other technologies such as time-of-flight based sensors. The system's motorized tilt allows the vertical field of view of the camera to be changed by $\pm 27^\circ$ for initial setup.

3 Limitations of current depth camera technology

The four limitations of depth camera technology are (i) low resolution of the depth maps produced, (ii) limited or fixed selection for the region of interest, (iii) fixed focal length, and (iv) short imaging range.

Depth cameras currently provide relatively low resolution images. There have been efforts to increase the resolution of the depth maps, such as in the case of Ref. [Park *et al.*, 2011] where they present a system that can upsample up to 8 times the original resolution of the depth map by adding an external high resolution color camera. The first step in that work is to calibrate and register both cameras. Once both cameras are calibrated, for each frame the input from both cameras is processed, which reduces the frame rate of the system. The maximum resolution subsequently depends on the resolution of the color camera used. This means that even after the depth map has been upsampled, the maximum resolution is limited to 1280×960.

Most depth cameras are tied to a particular field of view, and are generally unable to tilt or rotate to view any other location. Since mobile robots need to interact with their complex and changing environment, there has been a need to change the pose of the depth cameras. Ref. [Droeschel *et al.*, 2010] addresses this requirement by mounting a depth camera on top of a pan-tilt unit. While this solution increases the visibility of the camera, as the camera will not be rotating about its optical center parallax errors are introduced when techniques such as image registration or data fusion are part of the camera's application. In addition to the alignment problem, the dynamic performance of a pan-tilt camera system is limited by the speed and accuracy of the pan-tilt unit itself.

All commonly used depth cameras have a fixed focal length. Mesa Imaging address this problem by offering the SR4000 with a 5 m or a 10 m maximum range option with a field of view of $69^\circ \times 55^\circ$ and $44^\circ \times 35^\circ$, respectively. Once that decision is made, the focal length is fixed. With current depth cameras the field of view cannot be changed on-the-fly while the system is operating. Commonly used depth cameras are also short in imaging range [Foix *et al.*, 2011]. There are different types of technologies used to measure depth information, such as triangulation and time-of-flight. Depending on the technology used, the maximum imaging range is typically around 5 m, and can increase up to a maximum of 10 m in the case of the SR4000, albeit with a smaller field of view.

4 Configurable depth camera

The configurable depth camera is based on a long-range 3D laser scanner that allows full dynamic control over the scanning region as well as the angular resolution along both horizontal and vertical axes. The Ocular Robotics RobotEye RE05 3D laser scanner provides autonomous robotic platforms with a wide field of view of $360^\circ \times 70^\circ$ and a sensing range of up to 250 m (depending on the reflectivity of the targets). In addition to its long-range and wide field of view the RobotEye RE05 gives the user the unique ability to define the scan region, and to change its resolution along both axes down to the finest 0.01°angular granularity. The RE05 is pictured in Figure 1 and has been described in detail by Ref. [Wood and Bishop, 2012]. Because of its on-the-fly scanning adaptability, small size and IP67 rating, the RobotEye RE05 is currently being used for both navigation and high resolution imaging by mobile robotic platforms within complex environments. A recent application of the RobotEye RE05 is the autonomous navigation of robots within, and 3D visualization of, hard-to-access archeological sites such as catacombs [Ziparo *et al.*, 2013].

Figure 1: Ocular Robotics RobotEye RE05 3D laser scanning system pictured here is the basis of the hardware used for the configurable depth camera.

5 Generating a 2D depth camera image

This section describes the steps taken to generate a high resolution 2D depth camera image using the RE05 3D laser scanner.

5.1 Acquisition method of 3D point cloud

Most 3D laser scanners provide limited configurability over the resolution of the horizontal and (particularly) vertical axes. The RE05 provides the on-the-fly ability to change both horizontal and vertical resolution. This feature allows the creation of a near-uniform scan pattern over the entire $360^\circ \times 70^\circ$ field of view and at any desired angular resolution down to the finest granularity of $0.01^\circ \times 0.01^\circ$.

Given the desired RE05 horizontal field of view (HFoV), azimuth rotation (ω_{az}), laser frequency (f), and the number of averaged samples (N), it is possible to calculate the horizontal angular resolution of a scan:

$$Horizontal\ Resolution = \frac{f}{HFoV \times N \times \omega_{az}} \tag{1}$$

$HFoV$ may be set to any value up to 360° while the hardware sample averaging was set to unity for this measurement. The laser frequency has a maximum value of 10 kHz when intensity values are used, and 30 kHz when only range information is needed. The azimuth rotation has a maximum value of 15 Hz, however by reducing the azimuth rotation speed a finer horizontal resolution can be achieved. The vertical angular resolution of a scan is given by:

$$Vertical\ Resolution = \frac{numLines}{VFoV} \tag{2}$$

Here, $VFoV$ is the vertical field of view and $numLines$ is the number of horizontal lines in the scan. The $VFoV$ can range anywhere from 2° to 70° depending on the application and can be changed at any time. By matching the horizontal and vertical resolution the RE05 3D laser scanner becomes a 2D configurable depth camera.

Figure 2 is the 2D projection of a 3D point cloud. The 3D point cloud represents a wall segment which has been acquired by the RE05 using the uniform scan pattern condition of a matched horizontal and vertical resolution. After sampling the scene with the laser using a uniform pattern, it is then possible to reconstruct the depth and intensity images of the scene.

5.2 2D image reconstruction

After an initial projection of the generated point cloud to the desired field of view and image size, a combination of morphological operations and interpolation methods were performed to obtain a dense representation of the image. In order to achieve a fast image reconstruction, a closing morphological operation with a $n \times n$ square structuring element has been used. The size of the structuring element is a function of the focal length, field of view, and the angular resolution of the 3D laser scanner used to generate the projection. Figure 3 is a dense 2D image reconstruction of a 3D point cloud captured using the configurable depth camera. A frame captured using this camera is a two channel image. One channel contains the depth information (left), and the other channel represents the 905 nm laser intensity information (right) for the same scene.

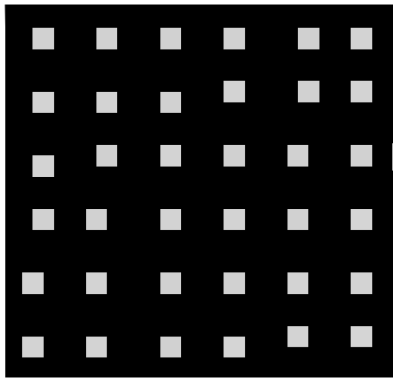

Figure 2: 2D projection of a 3D point cloud. The 3D point cloud represents a wall segment and is near-uniform along both axes.

6 Key features of the configurable depth camera

This section demonstrates how the configurable depth camera overcomes the limitations of current depth cameras by utilizing its customizable resolution and field of view, and long-range imaging capabilities.

6.1 High image resolution

The frame resolution of the depth camera is determined by the camera's field of view, and the angular resolution of the 3D laser scanner.

The angular resolution can be increased to obtain a more detailed image, or decreased to achieve faster frame rates. The maximum angular resolution that the underlying 3D laser scanner can provide is $0.01^\circ \times 0.01^\circ$ over the entire $360^\circ \times 70^\circ$ field of view. Figure 4 is a density map showing the relationship between the maximum resolution of the configurable depth camera and the camera's field of view (FoV). The density axis is the image resolution measured in megapixels. From Figure 4 the depth camera is capable of generating a 252 megapixels depth or intensity panorama image with a field of view of $360^\circ \times 70^\circ$. This resolution is orders of magnitude larger than that provided by current depth cameras as shown in Table 1, and even larger than that achieved by the upsampling techniques discussed in Section 3.

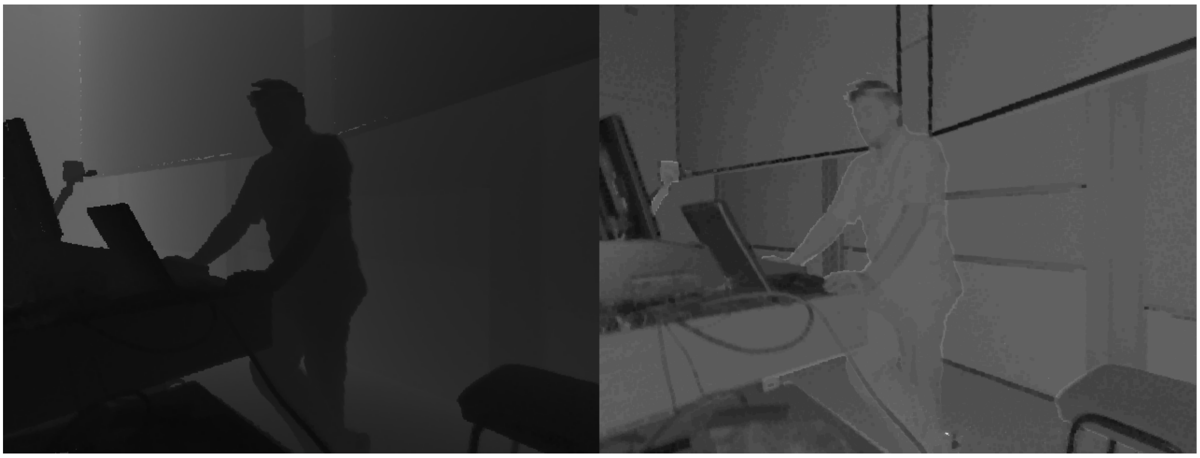

Figure 3: This image shows the two channel nature of a single depth camera frame. On the left is the depth map and on the right is the infrared intensity image, reconstructed from the range and intensity values, respectively. Each frame is constructed by first projecting the points gathered by the 3D laser scanner onto the user-defined field of view and subsequently reconstructing the image at the user defined resolution.

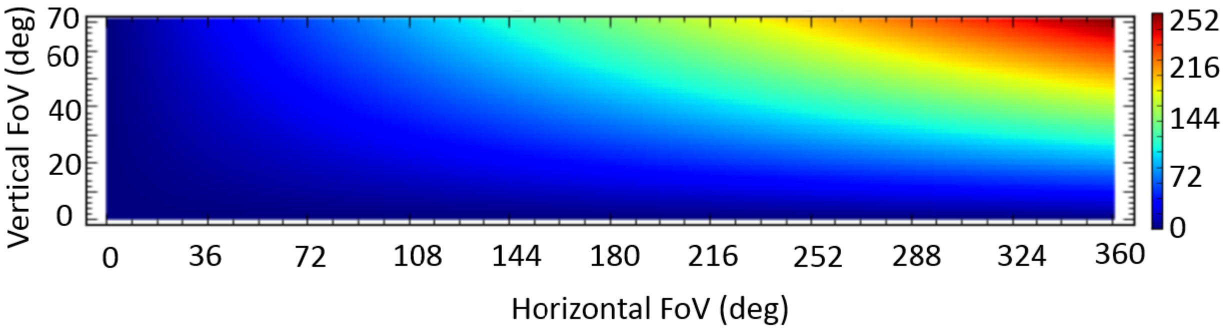

Figure 4: Maximum image resolution obtainable with the configurable depth camera for different fields of view of the depth camera. The density axis is the image resolution measured in megapixels. Depending on the field of view used for the depth camera, the image may contain up to 252 megapixels worth of information. Since the user can define his own image size, an even larger interpolated image may be generated per frame.

With the ability to change the angular resolution on-the-fly, the depth camera also provides a frame by frame user configurable image size. This is particularly useful for pairing this sensor with other cameras that may have different aspect ratios or image sizes, for example, to create a RGB-D dense point cloud.

6.2 Configurable region of interest

In many robotic applications such as human-robot collaboration, active gaze control is essential. With the configurable depth camera, each frame can be obtained from a different region in space. This is comparable to having a regular depth camera mounted on top of a high precision and fast pan-tilt system with a panoramic head, in addition to active focal length control (i.e. zooming in or out while preserving resolution).

As the underlying hardware is a 3D laser scanner, it can be used to constantly scan the full panoramic view, and if an object of interest appears on the scan the camera could change its orientation to that particular region to acquire a high resolution depth and intensity image for detailed analysis. For every frame, the sensor can image a $360° \times 70°$ view of its environment or the user may define a region of interest by specifying the camera's vertical angle of view, image aspect ratio (which defines the horizontal field of view) and orientation of the camera in the world. For example, the system may be configured to display four views of $90° \times 70°$ each, showing left, right, front, and back of the scene simultaneously, or just a single $40° \times 40°$ view to emulate a CamCube field of view.

Figure 5 shows two depth maps of a lecture theatre captured with the configurable depth camera. Both images utilize the full vertical field of view of the camera and have been captured with a 45° rotation in azimuth with respect to each other. The configurable depth cam-

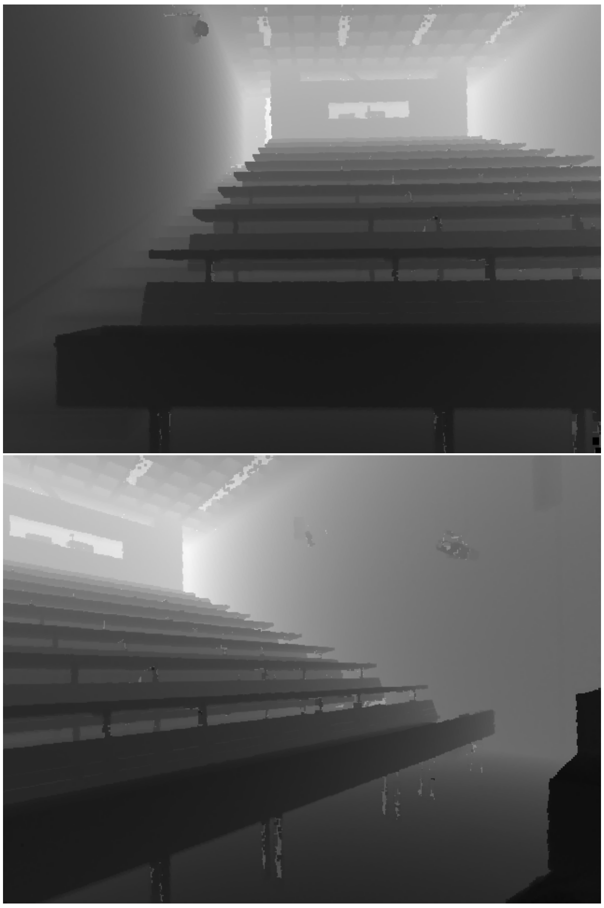

Figure 5: Top image shows a depth map of a lecture theatre (back wall 15 m away from camera) with a vertical field of view of 70° projected on a 640×480 image. The bottom depth map was captured by rotating the camera 45° in azimuth. The user can set the configurable depth camera horizontal angle of view by changing the aspect ratio of the projected image. This allows the user to specify the exact region of interest in the scene for each frame.

era can be rotated in azimuth and elevation similar to that achieved by using a pan-tilt system. However, in contrast to a pan-tilt system the configurable depth camera is capable of moving with very high speed and acceleration while simultaneously delivering precision pointing of the camera's field of view. Further, since the configurable depth camera rotates exactly through the optical center it can be used to create high resolution panoramic images of depth or intensity.

The configurable depth camera also provides the ability to move its origin to another location in space. This feature is useful when matching the field of view of the depth camera with another sensor. For example, a visible light camera can be combined with the configurable depth camera to generate a dense RGB-D point cloud. Here, registration can be achieved by combining the intensity channel of the depth camera and the visible image using normalized mutual information [Taylor, and Nieto , 2012].

6.3 Variable focal length

In security and surveillance applications the user may need to capture a higher resolution image of the region of interest, or zoom into the scene while preserving the existing resolution of the image. In these scenarios the on-the-fly ability to change the focal length of the camera is paramount. The configurable depth camera enables the user to change the desired angular resolution at any time, thus allowing capture of higher resolution imagery of the region of interest, or a zoom and enhance operation with a final resolution of up to 252 megapixels.

By changing the vertical angle of view, the user can zoom in or out of the scene while preserving the resolution of the frame. Figure 6 is a depth map of the back wall of the lecture theatre presented in Figure 5 captured using a vertical field of view of 20° and projected on a 640×480 image.

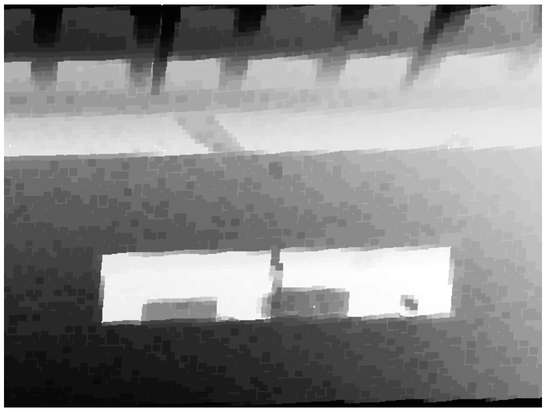

Figure 6: Depth map of the same scene as Figure 5, captured using a reduced vertical field of view of 20° projected on a 640×480 image. This has the effect of a zoom-in function for the camera.

With active control of angular resolution, the angular resolution of the scan used to generate the zoomed-in frame can be increased to achieve a higher resolution image. As the focal length of the depth camera is increased (zoom-in function), the vertical field of view becomes smaller. Keeping the number of horizontal lines per scan unchanged results in a higher angular resolution and translates into finer details in the zoomed frame. Alternatively, the user may also increase the resolution of the current field of view without the need of changing the focal length.

6.4 Long-range imaging

By taking advantage of the underlying 3D laser scanner, the configurable depth camera provides a long-range measurement capability of up to 30 m for common objects, and up to 250 m for high reflectivity objects. The top panel of Figure 7 is a depth map of a car park where the trees are at a distant of 47 m from the camera. In the bottom panel the imaging range has been increased to 60 m to capture the buildings at the back of the car park.

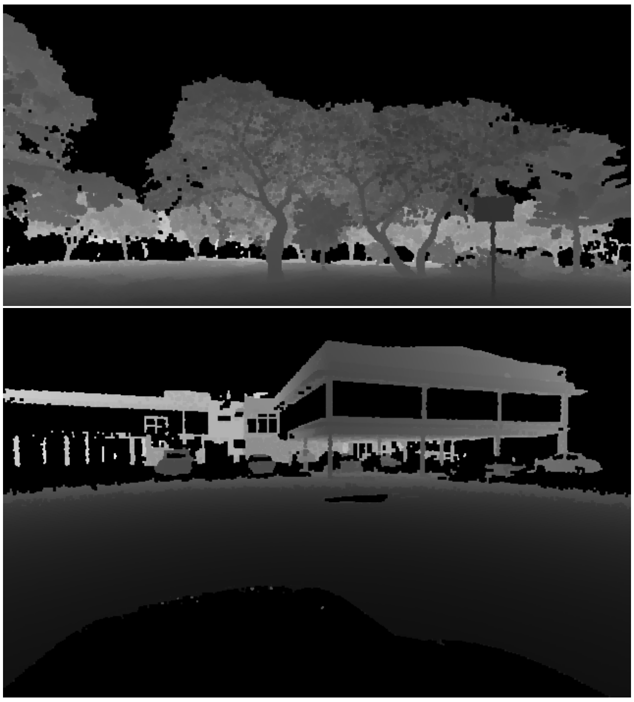

Figure 7: Depth map of a park where the trees are 47 m away from the camera (top panel). Depth map of a car park with the buildings at a distance of 60 m from the camera (bottom panel). Depending on the reflectivity of the objects in the scene the range of the camera can extend beyond 60 m.

7 Applications of a configurable depth camera

The adaptability of the configurable depth camera opens up a number of novel imaging applications, some of which are discussed in the following sections.

7.1 Image-based pose estimation

The configurable depth camera allows application of conventional image processing algorithms to a 3D laser scanner. One such example has been demonstrated in Figure 8 where a chessboard pattern has been imaged using the configurable depth camera. The image may be use to obtain the position and orientation of the underlying 3D laser scanner.

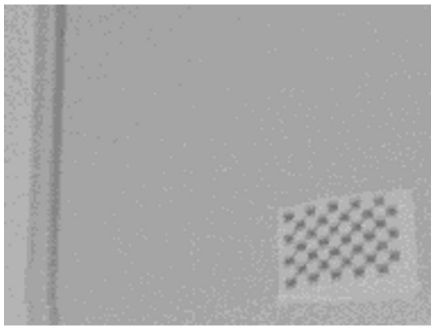

Figure 8: Intensity image of a typical chessboard pattern captured using the configurable depth camera. This image may be used to estimate the pose of the 3D laser scanner.

7.2 Depth and intensity panoramas

Because the configurable depth camera can be rotated exactly through its optical center, it is possible to generate high resolution panorama images up to 252 megapixels using the depth or intensity channel of the camera. An example of this capability is demonstrated in Figure 9. This $360° \times 70°$ cylindrical panorama has been captured using only four $90° \times 70°$ frames of the configurable depth camera. The top panel shows the depth information while the intensity channel was used to create the bottom panel.

7.3 Data fusion

The configurable depth camera can be set to match the position, orientation, field of view and resolution of another sensor. If the other sensor is a RGB camera, then for each pixel in the RGB image there will also be a depth measurement from the depth camera. This allows the user to create a RGB-D dense point cloud. Further, for every pixel on the depth image, there is an associated 905 nm intensity value which may be used to refine the initial alignment of the two sensors. Figure 10 is a RGB-D point cloud created by overlapping a single depth camera frame with multiple RGB images acquired by the RobotEye REV25 Vision System [Dansereau *et al.*, 2014]. The bottom half of this image has been generated from the IR intensity values alone.

7.4 3D object scanner

If the user needs a large number of samples in a specific region to create a high resolution frame, the region scan function of the underlying 3D laser scanner can be used.

Figure 9: This image shows two identical $360° \times 70°$ cylindrical panoramas created using only four frames of the depth camera ($90° \times 70°$ each). Top: panorama created using the depth channel. Bottom: panorama created using the 905 nm IR intensity channel.

This allows the system to scan a region in space delimited by a defined range in azimuth and in elevation. By using this region scan feature a high resolution depth map of the current pose of the object can be obtained. Once all the orientations are covered, a full high resolution 3D model of the object can be generated.

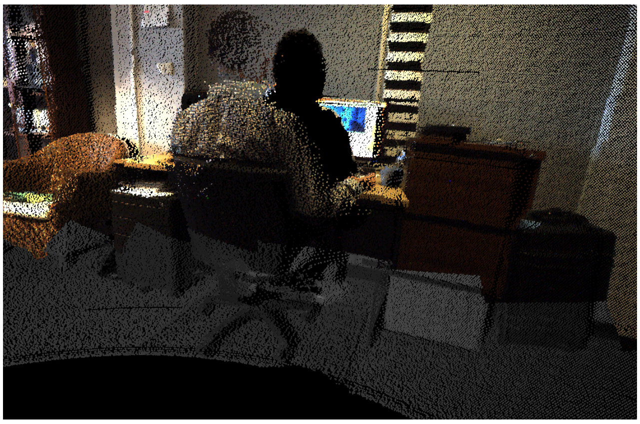

Figure 10: A RGB-D dense point cloud created by overlapping a single depth camera frame with multiple RGB images. The bottom half of this image has been generated from the IR intensity values alone.

8 Conclusions

This paper introduces for the first time a highly configurable depth camera. By taking advantage of the flexibility of the underlying 3D laser scanner, it has been possible to create a configurable long-range depth camera that can adjust its field of view, resolution, orientation, and position on a frame by frame basis. This depth camera overcomes the limitations of current depth cameras, namely low resolution, fixed field of view, and short-range detection.

The underlying RobotEye RE05 laser scanner allows full control over the vertical and horizontal resolution of the scan, as well as the region to be scanned up to $360° \times 70°$ and with a fine-grained angular resolution down to $0.01°$. By using a uniform laser scan, it is possible to reconstruct a 2D depth or intensity image for each scan at different angles of view.

This camera system enables many novel applications such as panoramas, scene exploration with zoom and enhance, RGB-D point clouds, and application of general image processing algorithms to a 3D laser scanner such as pose estimation using a chessboard pattern.

References

[Foix *et al.*, 2011] Sergi Foix, Guillem Aleny, and Carme Torras. Lock-in time-of-flight (ToF) cameras: a survey. *Sensors Journal, IEEE 11, no. 9: 1917-1926*, 2011.

[Khoshelham and Elberink, 2012] Kourosh Khoshelham and Sander Oude Elberink. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors Journal, IEEE 12, no. 2: 1437-1454*, 2012.

[Park *et al.*, 2011] Jaesik Park, Hyeongwoo Kim, Yu-Wing Tai, Michael S. Brown, and Inso Kweon. High quality depth map upsampling for 3d-tof cameras. *In Computer Vision (ICCV), 2011 IEEE International Conference on, pp. 1623-1630.* IEEE, 2011.

[Droeschel *et al.*, 2010] David Droeschel, Dirk Holz, J. Stuckler, and Sven Behnke. Using time-of-flight cameras with active gaze control for 3D collision avoidance. *In Robotics and Automation (ICRA), 2010 IEEE International Conference on, pp. 4035-4040.* IEEE, 2010.

[Wood and Bishop, 2012] David Wood and Mark Bishop. A Novel Approach to 3D Laser Scanning. *Australasian Conference on Robotics and Automation*, Wellington, New Zealand, December, 2012.

[Ziparo *et al.*, 2013] Ziparo, V. A., M. Zaratti, G. Grisetti, T. M. Bonanni, J. Serafin, M. Di Cicco, M. Proesmans et al. Exploration and mapping of catacombs with mobile robots. *In Safety, Security, and Rescue Robotics (SSRR), 2013 IEEE International Symposium on, pp. 1-2.* IEEE, October, 2013.

[Taylor, and Nieto , 2012] Zachary Taylor, and Juan Nieto. A mutual information approach to automatic calibration of camera and lidar in natural environments. *in Proc. Australasian Conference on Robotics and Automation, pp. 1-8.*, Wellington, New Zealand, December, 2012.

[Dansereau *et al.*, 2014] Donald G. Dansereau, David Wood, Sebastian Montabone, and Stefan B. Williams. Exploiting Parallax in Panoramic Capture to Construct Light Fields. *Australasian Conference on Robotics and Automation*, Melbourne, Australia, December, 2014.