

# Photogrammetric Calibration of the SwissRanger™ 3D Range Imaging Sensor

Scott Robbins, Brigit Schroeder, Bryan Murawski, Nicholas Heckman, Jonathan Leung  
[robbins, bschroeder, bmurawski] @ mitre.org  
The MITRE Corporation, Bedford, Massachusetts, United States<sup>1</sup>

## ABSTRACT

Many robotic and industrial systems require 3D range-sensing capabilities for mapping, localization, navigation, and obstacle avoidance. Laser-scanning systems that mechanically trace a range-sensing beam over a raster or similar pattern can produce highly accurate models but tend to be bulky and slow when acquiring a significant field of view at useful resolutions. Stereo cameras can provide video-rate range images over significant fields of view but tend to have difficulty with scenes containing low or confusing textures. A new generation of active light, time-of-flight range sensors use a 2D array of sensor elements to produce a 3D range image at video rates. These sensors pose unique calibration challenges, requiring both the usual calibration of lens distortion (intrinsic calibration) and calibration of the time-of-flight range measurement (3D calibration). This paper presents our application of a photogrammetric calibration approach using inexpensive printed optical targets and off-the-shelf software to solve both intrinsic and range calibrations for the MESA Imaging SwissRanger™ 3100 range imaging sensor. We further identify specific calibration issues stemming from this sensor's correlation of reflectivity with measured range.

Keywords: Range Imaging, 3D Camera, SwissRanger, Photogrammetric Calibration

## 1. INTRODUCTION

Over the past decade, substantial improvements have been made in robotics-oriented 3D sensing, especially with laser scanning<sup>2</sup> and stereo vision technologies. These sensing modes are of particular interest in the field of robotics, with applications for mapping, localization, navigation, and obstacle avoidance. However, due to fundamental elements of these two sensing modes, substantial limitations inhibit their applications to real-time and real-world applications. Laser scanning, in which one or more laser ranging beams are scanned across a scene (either in a raster, Lissajous, or some other 2D pattern), tends to be bulky and is unable to practically produce data at video frame rates for large fields of view, which limits its suitability for use on moving platforms. Embedded stereo vision systems that operate at video rates have been developed; however, stereo vision technology has great difficulty sensing range for low-texture surfaces (such as blank walls).

Recently, a new generation of active, lightweight, time-of-flight video rate 3D sensors, utilizing two-dimensional imaging arrays and diffuse modulated light sources, have come on the market. These measure time-of-flight distance to each pixel simultaneously. This type of sensor is ultimately constrained by the  $1/R^4$  energy rule, which forces a trade-off between field of view and maximum measurable range for a given radiated power. In this work we studied an example of this type of sensor, the MESA Imaging SwissRanger SR3100, successor to the SwissRanger-2, which was studied in Kahlmann [1, 2]. As part of a larger body of work in the field of 3D robot mapping, we attempted to combine the SR 3100 with an embedded stereo sensing. This necessitated an accurate calibration of the SR 3100, both individually and with the other onboard sensors.

---

<sup>1</sup> Approved for Public Release, March 2008, MITRE Case Number: 08-0372. Distribution Unlimited.

<sup>2</sup> In this context, 3D laser scanning refers to laser beam scanning systems that produce a full depth image (sometimes known as a 2.5D image), such as the RIEGL LMS™, as opposed to 2D laser line scanners such as the SICK LMS™.

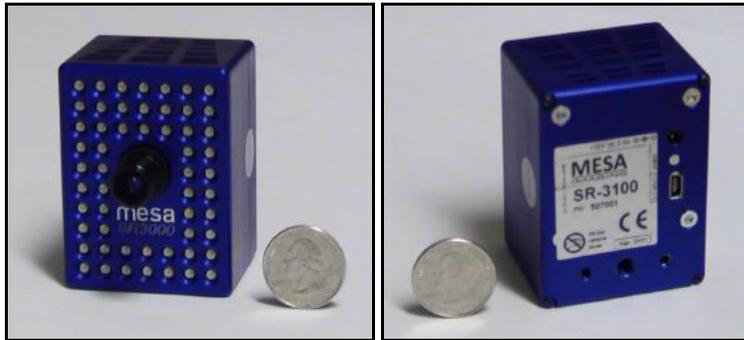


Fig. 1. The Mesa Imaging SwissRanger SR 3100

The SR3100 combines an integrated light source with sensor optics in a compact package measuring only 50 x 67 x 43 mm. The light source consists of 55 infrared light-emitting diodes operating at 850nm, with a total illumination power of approximately 1 watt over a roughly 50 degree wide cone. Internal circuitry modulates this light source at 20 MHz. This light bounces off nearby objects and after a short time is returned to the sensor through an integrated f/1.4 lens to a 176x144 pixel imaging array. The array then measures the phase offset caused by the travel time of the returned signal, as compared to the currently emitted signal yielding a measurement of the light's round-trip distance. With a basic modulation approach at 20MHz, the phase offset cycles after 15 meters, giving an ambiguity range of 7.5 meters (i.e., the phase angle offsets at 1.5m and 9m are the same). While this presents a problem in theory, in practice the illumination power is generally (though not always) too weak to effectively resolve distances of more than 6 meters. In addition to the distance, the total intensity of the returned signal is also measured, providing a simultaneous infrared grayscale image. This combination of grayscale 2D image and 3D range image enabled the use of a photogrammetric calibration approach using standard checkerboard optical targets.

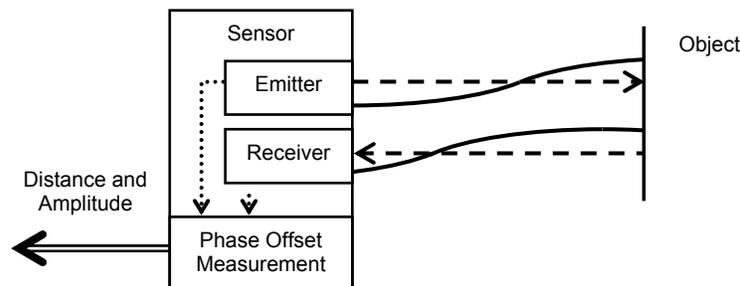


Fig. 2. The SwissRanger SR3100 measures the offset in phase angle of the returned light signal resulting from its time of flight, which is proportional to distance. More detail on this sensor process, its manufacture, and other applications can be found in [3] and [4].

The frame rate of the sensor is dependant on the integration time, a parameter roughly analogous to the exposure time of an optical camera. Longer distance imaging (4m or so) requires longer exposures on the order of 100-200ms, giving about 10-5 frames per second (respectively), while closer range imaging requires shorter exposures of 25-50ms for about 40-20 frames per second. An auto integration feature is also available through the supplied API that adjusts frame rate to maintain an optimal exposure.

## 2. SENSOR ACCURACY

While this sensor demonstrated relatively high precision for well-illuminated scenes, our initial experiments indicated systematic accuracy errors in range, when compared against external measurement. The regular nature of these systematic errors suggested they could be overcome with appropriate calibration.

A ground truth is needed to perform any sort of range calibration with the SwissRanger sensor. Kahlman [1] uses a total station consisting of a small target mounted on a sled that is moved along a track and measured with an interferometer. In addition to their expense, total stations such as these can only provide a ground truth measurement for a few pixels at a time. A more general calibration approach that would provide an external measurement of range for each pixel against which the sensed range could be compared was desired.

Our solution to this problem was to optically compute the ground truth range for each pixel using standard camera calibration techniques similar to those applied to the predecessor SwissRanger-2 in [5]. The premise behind this approach is to place the sensor in an orientation such that a single plane occupies the entire field of view of the sensor, whose relative position is then computed from one or more optical targets. Next, the plane is intersected with view vectors for each pixel in the sensor. The distance from the sensor's focal point to each plane/view vector intersection is defined as the ground truth range for that pixel. To do this, the view vectors needed to be defined through intrinsic calibration and the plane needed to be defined using extrinsic parameters.

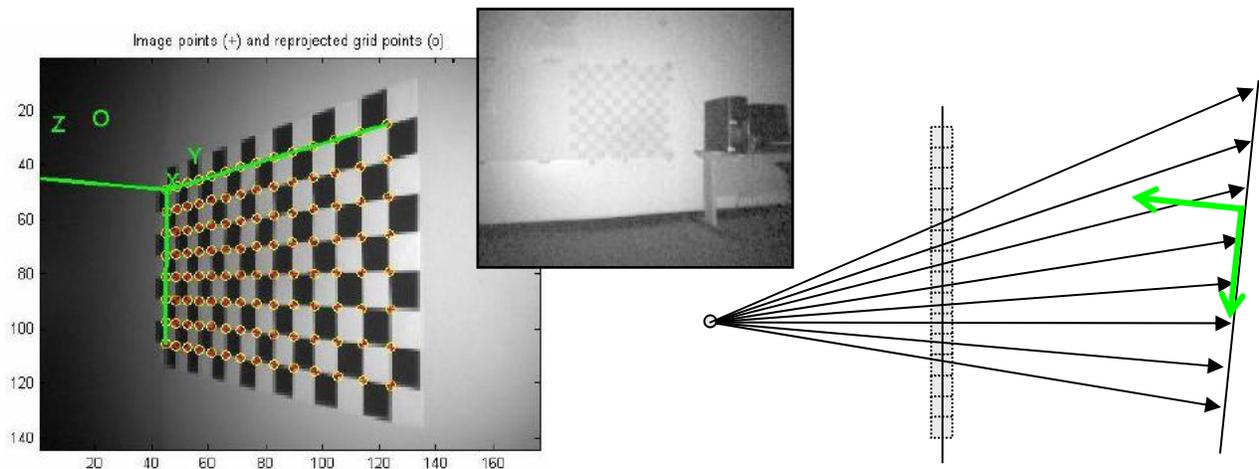


Fig. 3. Extrinsic parameters (left) of an optical target affixed to a flat surface that fills the sensors field of view is used to establish a ground truth value for each pixel in the camera field of view (right). Optical targets printed with an inkjet printer (inset) were barely visible under the 850mm infrared illumination, despite appearing black under visible light.

Intrinsic calibration is the process of computing an internal model of the camera that accounts for its focal length, center line, and the distortion caused by the sensor's lens [6]. These values were then used to provide the direction of view for each pixel. We used the Matlab Camera Calibration Toolkit [7] with a black and white checkerboard target. A side effect of that calibration, called the extrinsic parameter, gives the 3D position of a known optical target in the scene with respect to the imaging sensor (as is widely used in the calibration of stereo camera systems). Once coordinate transform from sensor to target is determined, the relative position of the entire planar surface to which it is attached can be determined.

Techniques for calculating intrinsic and extrinsic camera parameters are designed to use images captured from CCD cameras. While the SwissRanger does not produce an image in the typical sense, it does produce an image of the reflected infrared amplitude that can be used in place of a traditional CCD image. Since the image the SwissRanger produces is in the infrared spectrum, special care needs to be taken when creating optical targets. In our experiments, checkerboards printed with an inkjet poster printer on both glossy and matte poster paper yielded amplitude images where it was difficult to discern between black and white squares. Checkerboards printed with a laser printer on standard copier paper yielded the best contrast in returned light intensity between black and white squares. Additionally, applying a square root scaling to amplitude images improved the contrast between black and white squares.

### 3. DATA COLLECTION

In order to develop a calibration function that would work over a range of conditions, it was necessary to collect sample data that covered a wide variety of range values and amplitudes for which the ground truth values were known. Each setup consisted of a specific orientation of camera to planar surface, which we called a station. Each station consisted of multiple scenes with different textures (various flat materials) added to the surface to produce a variety of amplitudes at different ranges. Each station had at least one (but often more) calibration scenes in which the added texture consisted of calibration targets of differing sizes and orientations. These calibration scenes were used to establish the ground-truth measurements for their station. Ten stations in total were used.

For each scene at each station, we collected 100 range images, each consisting of reported range and amplitude (through the sensor API) at integration times from 25 to 250ms in 25ms intervals, as well as the auto integration time, as set through the sensors library API (for a total of 2200 range images per scene). Since the integration time was effectively analogous to the exposure time for an optical image, shorter integration times yielded lower amplitudes than longer integration times.

The goal of the data collection process was to cover a wide range of camera orientations and target distances. The camera was positioned so that some scenes contained high and low angles between the camera axis and the target planar surface. Targets were not limited to simple matte surfaces; a reflective marble floor, a dense, multi-colored carpet, and plastics tarps were all used. Other scene textures included construction, photographic, copy machine, and poster paper.



Fig. 4. Stations and scene textures were selected to cover a wide variety of incident angles, range values, and reflectances.

### 4. DATA ANALYSIS

Turning the data collected in Section 3 into useful plots required several processing steps. First, the ground truth range values based on extrinsic calibration were computed for each station. Second, the range offset was found for each scene at each integration time. This data was then put together in different ways to create plots that show the relationship to range offset.

The first step, computing the ground truth range values for each station, was done using the calibration scenes. For stations with multiple calibration scenes, a least squares fit plane was computed based on the 3D calculated positions of the corner points of the calibration targets. This “averaged” plane was then used to calculate the predicted range values for each pixel.

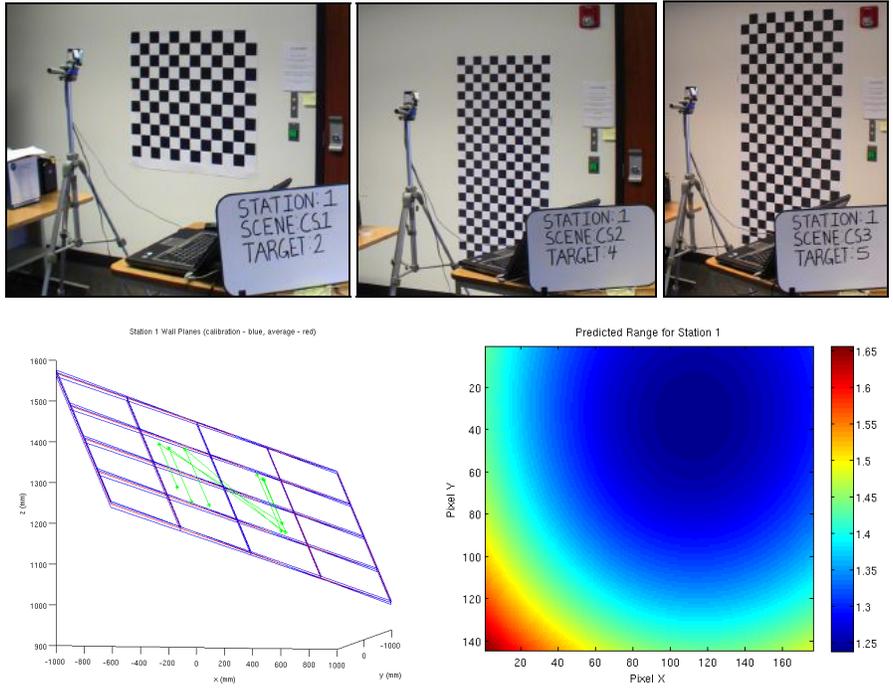


Fig. 5. Multiple calibration scenes (top) and their resulting planes for Station 1 (blue grids) and the best fit plane (red grid) based on the calibration target corners (green dots and lines) The predicted range image (left) from the computed best fit plane represents the “ground truth” range value for each pixel.

The next step was to compute the range offset values for each scene at each integration time. Each scene consisted of 100 range and amplitude images at each integration, as mentioned in Section 3. To reduce the effects of transient noise, each scene’s integration times range and amplitude images were averaged together, giving a single per-scene, per-integration time averaged range and amplitude image. These averaged images were what we used in our accuracy analysis.

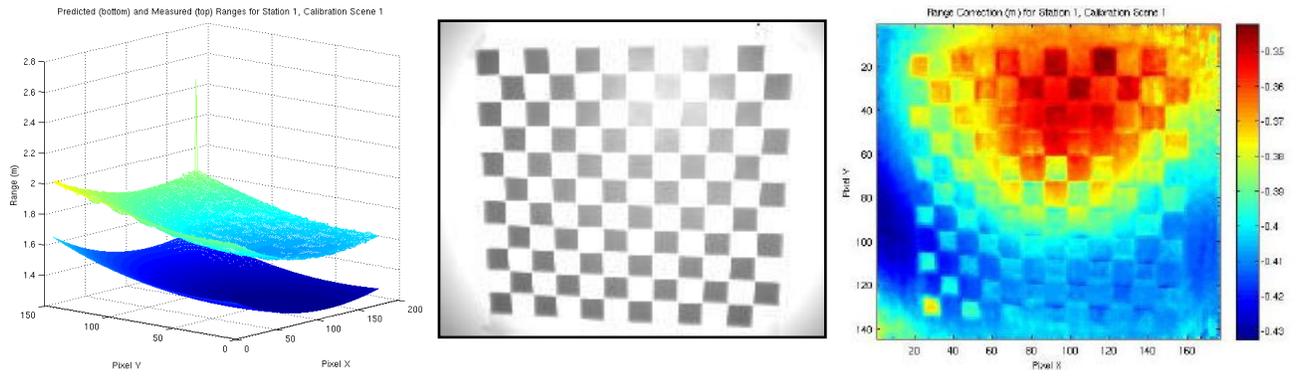


Fig 6. Left: Predicted (bottom) and Measured (top) Ranges for Station 1, Calibration Scene 1, as a 3D plot. A net range offset is clearly visible. Center: An image of the amplitude data returned by the sensor for that same scene. Right: An image of the range offset (difference between the two curves on the left). The affects of the amplitude on the range offset are clearly visible.

In Figure 6 the left image shows the range image (average of 100 range images) of Station 1, Calibration Scene 1, plotted along with the predicted range image computed previously. Both the predicted and indicated range images are clearly similar in shape, but there is also a general shift outward in the Z direction for the measured range data. The right image shows the range difference as a 2D image; note the effects of the calibration grid on the range data (as seen in the center amplitude image), where the dark squares affect the range values. The range offset values for each pixel are equal to the predicted range value minus the indicated range value. For analysis the data was aggregated into a histogram plot using the software in [8]. Our initial results demonstrated a profound effect on error for pixels indicating amplitudes less than 500 (out of a maximum of 65,536), and for this reason, pixels with amplitudes below this threshold were rejected.

First, a relationship between Range and Amplitude in the SR 3100 data will be examined. Figure 7 shows a range correction vs. amplitude plot for one station at one (left) and for all (right) integration times for a single station.

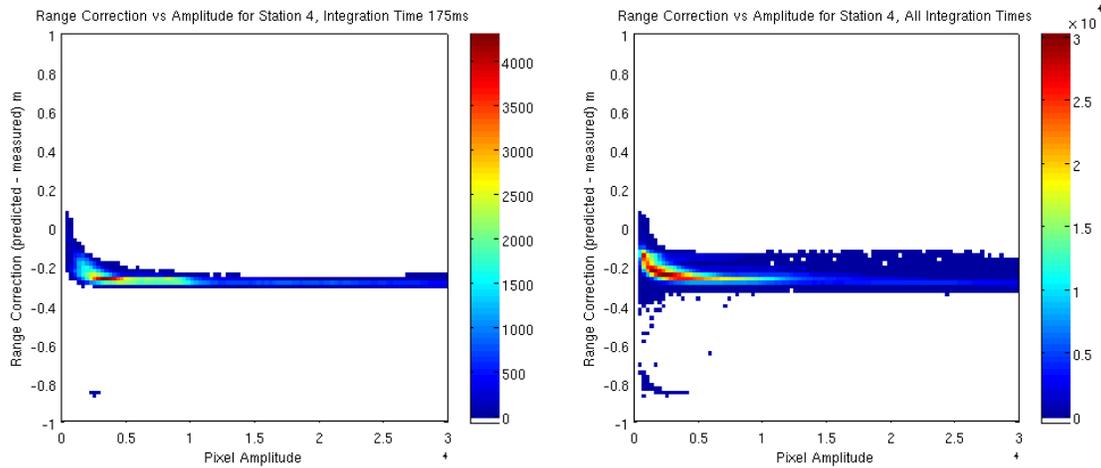


Fig. 7. Left: Range Correction vs. Amplitude for Station 4, Integration Time 175ms (Amplitude Scale is  $1 \times 10^4$ )  
 Right: Range Correction vs. Amplitude for Station 4, All Integration Times

There is a clear correlation between range offset and amplitude in this figure, which, in this data at least, is independent of integration time. The relation gets less crisp when data from all stations are included.

Figure 8 shows the range offset vs. amplitude graph for all stations at one (left) and for all (right) integration times.

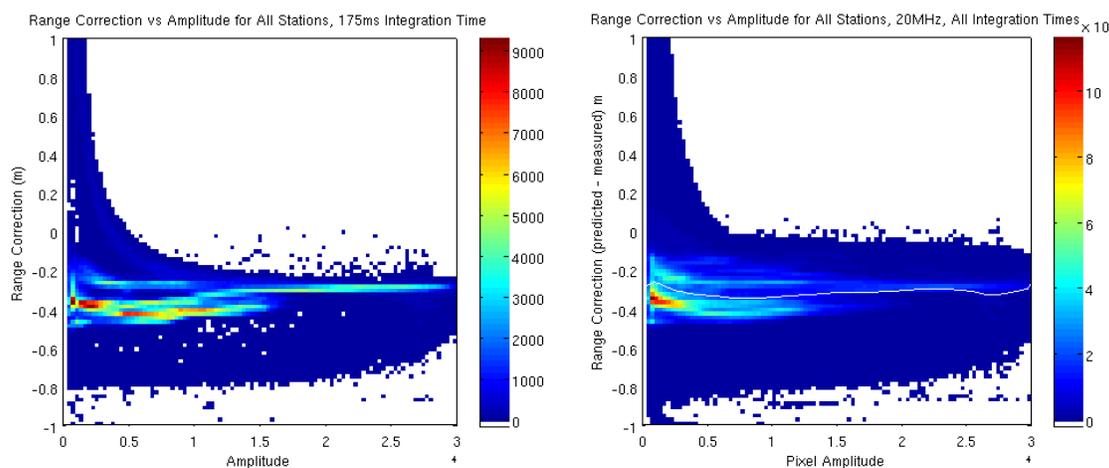


Fig. 8. Left: Range Correction vs. Amplitude for All Stations at 175ms. Right: Range Correction vs. Amplitude for All Stations, All Integration Times. The white line is the averaged offset for each amplitude bin.

In these figures the data begins to spread out a bit and not line as well as it had previously. This indicates that other data at the stations are having an effect on the correction function that needs to be applied. This result may, however, be small enough that the data could still be used for range correction.

The second relationship we looked at was range correction vs. indicated range. Figure 9 shows the same set of data plotted with indicated range as the independent variable. Because individual scene plots do not tend to cover a significant range of distances, only the accumulated data is shown.

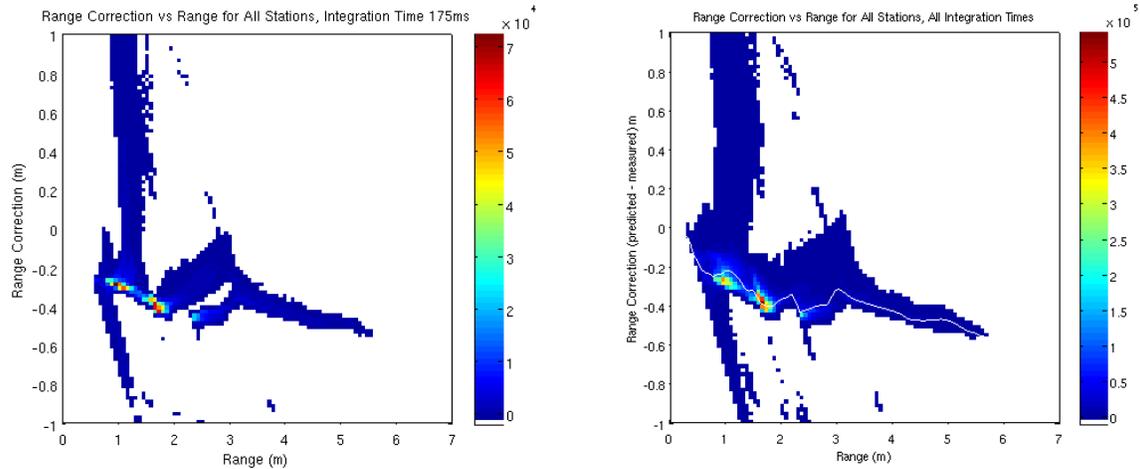


Fig. 9. Left: Range Correction vs. Range for All Stations, Integration Time 175ms. Right: Range Correction vs. Range for All Stations, All Integration Times. The white line is the averaged offset for each indicated range bin

It should be noted that the general trend shown in Figure 9 of increasingly negative range correction between one and two meters of indicated range is the opposite of what was indicated in Kahlmann [2]. Further investigation is warranted to resolve these discrepancies.

To validate these two correction plots (amplitude based and indicated-range based), we constructed an additive lookup table for each of the above white lines, and embedded them into our software interface for the SR3100. To provide an alternate reference of the quality of range adjustment, we calibrated a sensor mast containing both the SR3100 and a COTS stereo vision system (a Pt. Grey Bumblebee2). Extrinsic calibration was used to localize the relative positions of each sensor and project the stereo and SwissRanger data into a common reference frame. Both the amplitude-based offset function and the range-based offset functions were applied to the SwissRanger data. The results are shown below.



Fig. 10. Left: Picture of the Test Scene. The stereo camera is tilted 45 degrees to improve sensing to horizontal features.

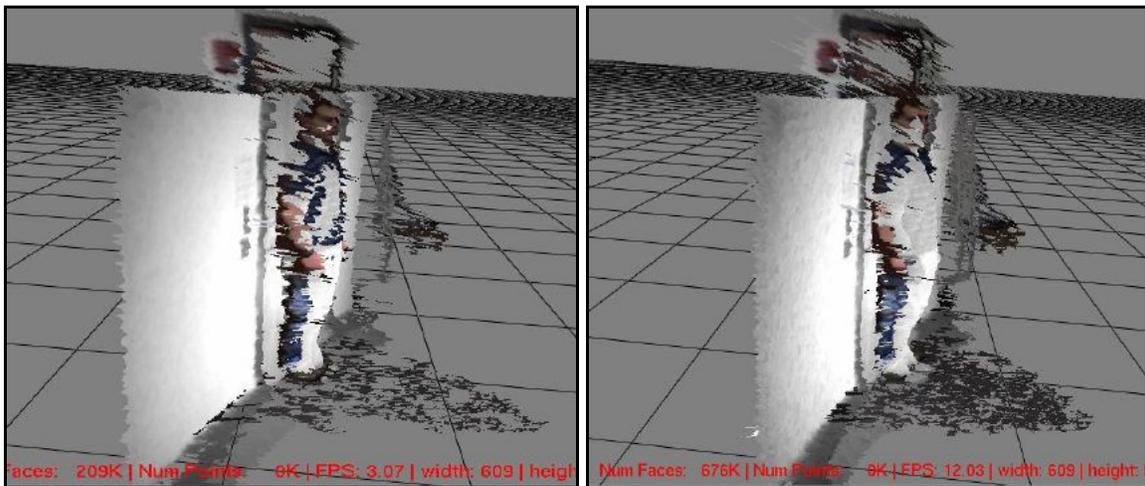


Fig. 11. Left: 3D scene after amplitude-based correction. Right: 3D scene after indicated range-based correction. Amplitude based correction matches the stereo vision far better, though neither are perfect.

The amplitude-based correction performed far better than the range-based correction on this as well as a variety of other test data. Neither method is perfect; however, both do improve the accuracy of the data. Ongoing enhancements will be the subject of future work.

## 5. NEXT STEPS

The correlation between amplitude and range offset was a good first step, but other variables are still affecting the data. Some candidates that could be examined in future work are integration time and the mean amplitude in a scene. May [9], working with the SwissRanger-2 identifies the effect of bright foreground objects influencing darker background objects, with the sensor consequently underreporting the distance to the background, an effect we have also noticed in practice with the SR3100.

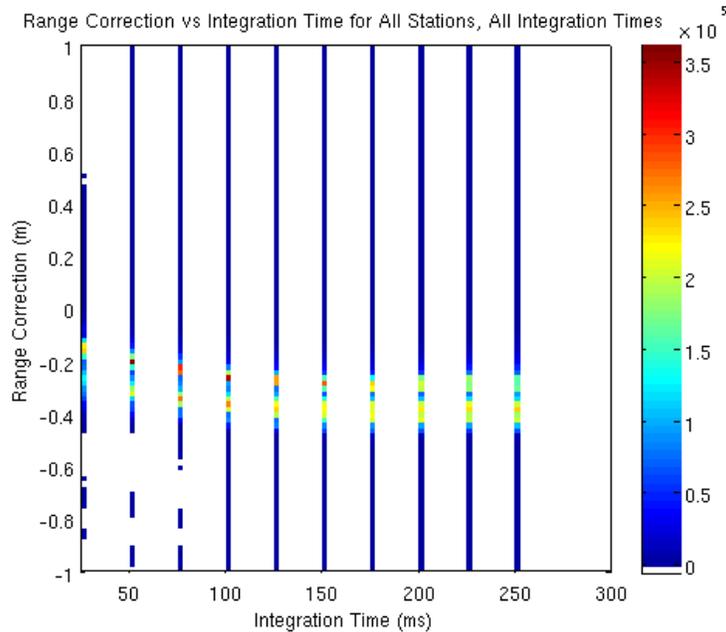


Fig. 12. Range Correction vs. Integration Time for All Stations, All Integration Times

Figure 12 shows range correction vs. integration time as a histogram. While there is definitely some correlation between integration time and offset, it should be noted that at lower integration times, scene brightness decreases substantially as well. This combined with the appearance of two shifted curves both on top of each other in this histogram indicates a more in-depth look at the data would be required before any strong conclusions could be drawn from it.

Additional future work will include an attempt to colorize the SR 3100 data using an optical camera calibrated in much the same manner as the stereo camera was in this work. Color data gives a much clearer view of the scene and allows someone viewing it to more easily determine what they are seeing. Results of calibrated and colorized SR3100 data will be presented in future work.

## 6. CONCLUSIONS

The SwissRanger 3100 is a compact range imaging sensor that produces relatively precise range images at video rates. However, a calibration is necessary to address accuracy issues related to indicated measurement correlation. We have shown the existence of specific range accuracy dependencies on the indicated amplitude, and provided a method and example for calibrating range given these dependencies. More work remains with this calibration to provide improved sensor accuracy. Nevertheless, this sensor, with its significant advantages over other modes of 3D sensing, shows great promise for a wide variety of applications.

## REFERENCES

- [1] Kahlmann, T., Ingensand, H., "Calibration and improvements of the high-resolution range-imaging camera SwissRanger™," Proc. SPIE Vol. 5665, 2005.
- [2] Kahlmann, T., Ingensand, H., "Calibration of the fast range imaging camera SwissRanger™ for use in the surveillance of the environment," Proc. SPIE Vol. 6396, 2006.

- [3] Oggier, T. et al., "SwissRanger SR3000 and First Experiences Based on Miniaturized 3D-TOF Cameras," available at: [http://www.mesa-imaging.ch/pdf/Application\\_SR3000\\_v1\\_1.pdf](http://www.mesa-imaging.ch/pdf/Application_SR3000_v1_1.pdf).
- [4] Oggier, T. et al., "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRangerTM)," available at [http://www.mesa-imaging.ch/pdf/CSEM\\_manuscript\\_5249-65.pdf](http://www.mesa-imaging.ch/pdf/CSEM_manuscript_5249-65.pdf).
- [5] Weingarten, J. W., Gruener, G., and Siegward, R., "A State-of-the-Art 3D Sensor for Robot Navigation," Proceedings of IROS, Sendai, September 2004.
- [6] Zhang, Z., "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000.
- [7] Strobl, K., Sepp, W., et al., "Camera Calibration Toolbox for Matlab," available at: [http://www.vision.caltech.edu/bouguetj/calib\\_doc/index.html](http://www.vision.caltech.edu/bouguetj/calib_doc/index.html).
- [8] Balkay, L., "Quick computation of two-dimensional histogram of bivariate data," available at: <http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=9896&objectType=file>.
- [9] Guomundsson, S., "Robot Vision Applications using the CSEM SwissRanger Camera," Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU ISBN/ISSN 87=643-0095-1, pp. 99, Ed. 1, August 2006.
- [10] May, S. et al., "3D Time-of-flight cameras for mobile robotics," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, Beijing, China, October 9-15, 2006