

© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Evaluation of Fast, High-detail Projected Light 3D Sensing for Robots in Construction

Jae-Sang Hyun*, Marc G. Carmichael[†], Antony Tran[†], Song Zhang*, Dikai Liu[†]

*School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

[†]Centre for Autonomous Systems, University of Technology Sydney, Ultimo, NSW 2007, Australia

Abstract—Robots used on-site in construction need to perceive the surrounding environment to operate autonomously. This is challenging as the construction environment is often less than ideal due to changing lighting conditions, turbid air, and the need to detect fine details. In this work we evaluate a custom made projected light 3D sensor system for suitability and practicality in enabling autonomous robotics for construction. A series of tests are performed to evaluate the sensor based on ability to capture environmental details, operate robustly in challenging lighting conditions, and make accurate geometric measurements. Analysis shows that high fidelity measurements with accuracy in the order of millimeters can be obtained, making the technology a promising solution for robots operating in construction environments.

Index Terms—Structured Light Sensing, Robotics, Construction

Robots are beginning to be utilized more and more for construction, with motivation coming from the potential improvements in quality, productivity, health and safety. In the past, automation in construction has often resembled a traditional manufacturing paradigm, with prefabricated building elements being made off-site and then transported for final installation. An alternative and desirable paradigm is to have construction robots on-site to perform tasks in-situ. Such robots could be used to perform various on-site tasks including abrasive blasting [1], erection of walls [2], installation and removal of panels [3], form work construction [4], and more.

Before robots can be used for general on-site construction activities there are significant challenges to be overcome. The construction of large-scale buildings requires robots to perform tasks outside of their immediate workspace, hence a robot needs to move around the work-site whilst simultaneously localizing with respect to its environment. Another challenge is to achieve robust operation in the dynamic and often non-ideal construction environment. Such environments can be particularly challenging with factors such as changing lighting conditions and turbid air making the use of conventional sensors difficult. To facilitate the practical use of robotic in construction, sensors capable of performing accurate measurements and capturing detail of suitably high fidelity are needed. Additionally, sensors are needed to be robust and have a large operating range to work in environments of varying sizes.

In this work we explore the practicality of using projected light 3D sensing for construction tasks. It is proposed that such a sensor is a suitable candidate as a primary sensor for performing various construction tasks with materials such as bricks and cement. We focus on tasks requiring the sensing

of cement, however considerations are not limited to this application. A custom developed sensor [5], shown in Fig. 1 is used to measure materials common on construction sites. Evaluation is made with regard to the detail obtained by the sensor, robustness to varying light conditions, measurement artifacts due to motion, and overall practicality for use in the construction environment.

The rest of the paper is organized as follows. Section I provides an overview on projected light sensing and details the custom sensor evaluated. Section II summarizes the key requirements of a sensor used in construction robotics, and the methods used to evaluate the projected light sensor. Section III presents the results from the analyses performed. In Section IV we discuss the performance of the sensor and implications for robotics in construction.

I. STRUCTURED LIGHT SENSING

Due to high spatial resolution, high accuracy and robustness to noise, digital fringe projection (DFP) methods are widely used in many fields [6]. The method uses continuous sinusoidal patterns which are called as fringe patterns. Instead of using intensity values of the captured images, the methods utilize phase information to find the corresponding pairs between the camera and the projector. A set of continuous fringe patterns is used for converting from intensity domain to phase domain pixel by pixel, and we can achieve camera pixel spatial resolution. Also, it is robust to noise such as ambient light, since the phase is used to reconstruct 3D instead of intensity which is easily affected by the noise.

A. Phase-shifting algorithm

Among many DFP methods a phase-shifting algorithm is extensively used in fringe analysis for 3D shape measurement, since the algorithm is accurate and robust to noise, and has high spatial resolution [7]. If the number of steps used in phase-shifting is N , the k -th fringe image can be mathematically defined as,

$$I_k(x, y) = I'(x, y) + I''(x, y) \cos(\phi + 2k\pi/N), \quad (1)$$

where $I'(x, y)$ is the average intensity, $I''(x, y)$ the intensity modulation and $\phi(x, y)$ the phase to be solved for. A least-square method is applied to achieve the phase by solving these N equations,

$$\phi(x, y) = -\tan^{-1} \left[\frac{\sum_{k=1}^N I_k \sin(2k\pi/N)}{\sum_{k=1}^N I_k \cos(2k\pi/N)} \right]. \quad (2)$$

We used the arctangent function to obtain $\phi(x, y)$, which is often referred to as the *wrapped* phase, since there are 2π discontinuities on every period which ranges $(-\pi, \pi]$. To remove the discontinuities and make the wrapped phase continuous, a *phase-unwrapping* procedure is performed by adding the required multiples of 2π for each pixel. We refer to the map of these multiples as fringe order, $K(x, y)$ and the unwrapped phase can be acquired as,

$$\Phi(x, y) = \phi(x, y) + 2\pi \times K(x, y). \quad (3)$$

For this paper, to minimize the projector defocusing error and discrete camera sampling error, we used the gray coding algorithm for fringe order determination developed by Wang et al. [8]. After unwrapping, the phase $\Phi(x, y)$ becomes continuous and uniquely defined, which is referred to as the absolute phase.

B. DFP system model

A DFP system typically consists of one camera and one projector which follow the well-known pinhole model [9]. The pinhole model of each optical device defines the relationship between 3D world coordinate (x^w, y^w, z^w) and 2D image coordinate (u, v) mathematically as,

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \begin{bmatrix} x^w \\ y^w \\ z^w \\ 1 \end{bmatrix}, \quad (4)$$

where s represents the scaling factor; \mathbf{A} is the 3×3 intrinsic matrix which has the information of the focal length, the skew factor, the principal points which is the intersection of the optical axis with the imaging plane; $\mathbf{R} \in SO(3)$ is the rotation matrix and $\mathbf{t} \in \mathbb{R}^3$ is the translation vector. Both the camera and the projector follow the same pinhole model, so two sets of equations can be achieved as follows:

$$s^p \begin{bmatrix} u^p \\ v^p \\ 1 \end{bmatrix} = \mathbf{A}^p \begin{bmatrix} \mathbf{R}^p & \mathbf{t}^p \end{bmatrix} \mathbf{X}^w, \quad (5)$$

$$s^c \begin{bmatrix} u^c \\ v^c \\ 1 \end{bmatrix} = \mathbf{A}^c \begin{bmatrix} \mathbf{R}^c & \mathbf{t}^c \end{bmatrix} \mathbf{X}^w. \quad (6)$$

Here $\mathbf{X}^w = [x^w, y^w, z^w, 1]^T$, superscript p means projector, superscript c means camera. \mathbf{A} , \mathbf{R} and \mathbf{t} of the projector and camera can be estimated by calibration [10]. Although Eqs. (5) and (6) has six equations, the number of unknowns is seven ($s^c, s^p, x^w, y^w, z^w, u^p, v^p$). Therefore, we need one more equation to uniquely solve the unknown variables. By using the phase information, we can calculate an unknown variable, a projector pixel line u^p , since the phase Φ is proportional to u^p as,

$$u^p = \Phi \times T / (2\pi). \quad (7)$$

Here T is the fringe period of projector image. With the number of equations equal to the number of unknown variables, we can solve for 3D information of the object \mathbf{X}^w .

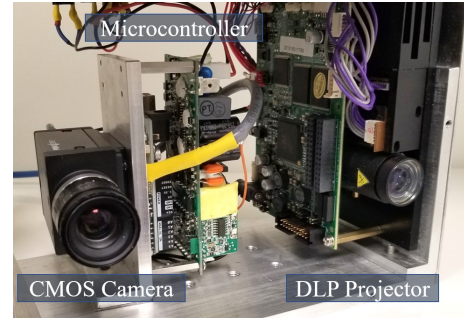


Fig. 1. Experimental setup. The system includes a DLP projector (model: LightCrafter 4500) and a CMOS camera (model: Grasshopper 3). The projector resolution is 912×1140 and the camera resolution is 1920×1200 .

II. SENSOR EVALUATION METHOD

For the projected light sensor to be suitable for use with construction robots it needs have several key requirements. This section details the key requirements evaluated, along with the methods used. Results from the tests are detailed in Section III.

A critical requirement of the sensor is the ability to capture fine detail. This is especially true for construction tasks such as interior finishing operations where the surface finish needs to meet certain quality standards. To evaluate the ability to create high fidelity reconstructions of the scene, the sensor is used to measure scenes containing objects typical of construction environments. This includes brick walls, steel girders, sand and foam. The level of detail captured but the sensor is evaluated by visually comparing the measurements obtained with the scenes from which they were taken. Results are shown in Sections III-A and III-B.

A construction robot is required to reconstruct a scene so features such as the straightness, perpendicularity, and accuracy of building elements can be determined. The limited field of view of sensors requires that multiple sensor measurements be stitched together to construct a representation of a larger scene. This scenario is tested by mounting the sensor to the end of a robot manipulator to move the sensor to know locations so that several measurements of the scene from different perspectives can be stitched together. Results are shown in Section III-C

A known challenge with using optical sensors is varying light conditions. A construction robot must be capable of making measurements in the presence of difficult and changing lighting conditions. A second, perhaps less obvious challenge is the varying optical properties of materials. The construction site is by nature an environment that undergoes significant variation during construction. Temporary woodwork may be replaced by glass windows, surfaces may collect dust, objects become reflective when sprayed with water. A situation of interest is the curing of cement. When wet, cement can be reflective, then as it cures its optical properties change. It is unclear how projected light sensors will be affected by such variations in material properties. To examine this, an experiment is conducted with the sensor taking measurements

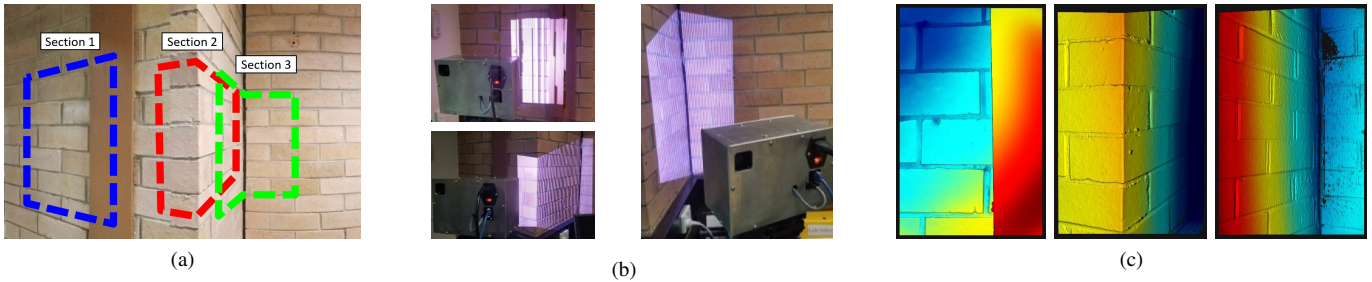


Fig. 2. (a) The brick wall scene measured by the sensor. Three different sections were chosen for testing. (b) Each section was measured individually with the sensor. (c) Results obtained from the three measurements. Colors represent the distance of each pixel from the sensor.

of wet cement as it cures. Measurements are taken under ambient lighting conditions, and with an external light source shone on the cement as well as straight into the sensor lens. Results are shown in Section III-D.

Another critical requirement of the sensor is to make accurate measurements. Evaluation of the sensor accuracy is made by measuring a scene containing objects of known geometry, and comparing the size of the objects in the measurements to their physical geometry. The sensor is evaluated at longer range with it calibrated for 1.8 m, and again at shorter range with it calibrated for 1.0 m. For each calibration the scene is measured at several distances so that measurement accuracy versus distance can be analyzed. Results are shown in Sections III-E and III-F.

III. SENSOR EVALUATION RESULTS

A. Scene Detail: Brick Wall

Figure 2a shows a brick wall scene measured with the sensor. The scene is separated into three sections; 1) a flat surface perpendicular to the sensor that includes both brick and flat steel girder; 2) an external corner of brick wall; 3) an internal corner of brick wall. The three sections were scanned with the sensor individually as shown in Fig. 2b.

The results from scanning the three scenes are shown in Fig. 2c. Upon examination of the capture depth measurements fine details of the brick surface can be seen. Observed are small

imperfections in the bricks and mortar which were found to be consistent upon examining the physical scene. In comparison, the steel girder is measured to be flat as indicated by the smooth transition of colour representing depth. It is proposed that the high level of detail capable of being captured with the sensor makes it likely to meet the requirements needed for automating many industrial tasks.

B. Scene Detail: Sand Pit

A second scene, this time smaller in size but also containing fine details was measured. The scene shown in Figures 3a and 3b contains clumps of dry sand, porous rock and cut foam. Added in the scene was a ruler to provide a baseline from which distance measurements could be compared. Results from scanning this scene are shown in Figures 3c and 3d. Fine detail in the scene such as the sand texture and the pores of the rock are visible. These results again demonstrate the high detail that the sensor is capable of capturing.

Points from the surface of the ruler aligning with its markings were manually chosen and their 3D coordinates obtained as seen in Figure 3e. The Euclidian distance between the two left-most, two right-most, and the two outer points are 10.06 mm, 9.67 mm and 19.73 mm respectively. These calculated distances closely matching the 10 mm markings on the ruler. This indicates that the sensor is capable of

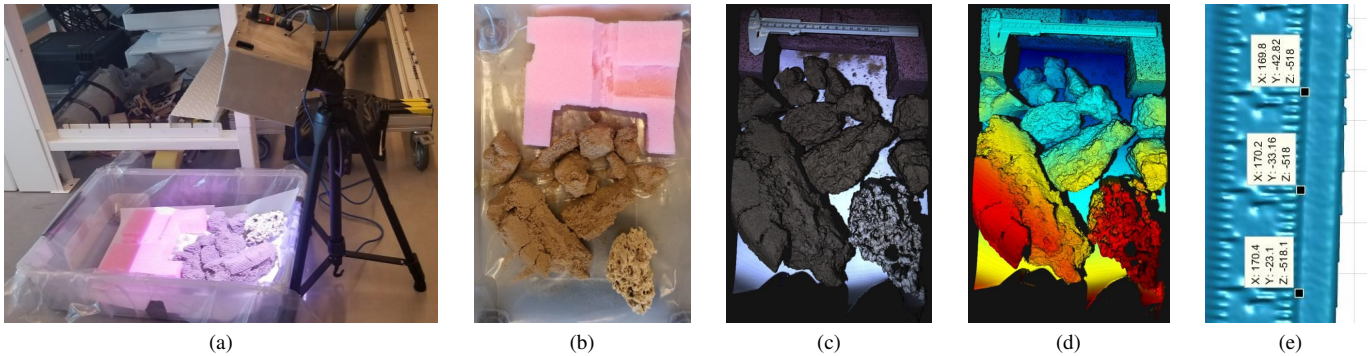


Fig. 3. (a) Setup of the scene scanned containing elements such as sand, porous rock and foam. (b) Top down view of the scene. (c) Colour reconstruction of the scene. (d) Reconstruction with colours used to represent the distance from the sensor. (e) Coordinates of points in the scene aligning with the ruler markings. Calculated distances between the points closely match the expected distances based on the ruler.



Fig. 4. Setup during the measurement stitching experiment. The sensor is mounted on the end-effector of a robot manipulator and translated over the scene below.

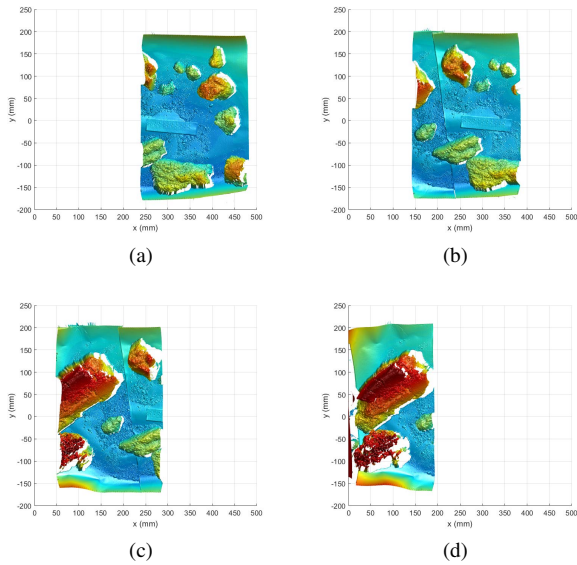


Fig. 5. Four measurements taken moving the sensor to different locations above the scene with the robotic manipulator.

not only capturing high scene detail, but also accurate scene measurements.

C. Measurement Stitching

Using the manipulator to translate the sensor above the scene, four separate measurements are taken. Figure 4 shows the setup with the scene containing paper, sand, rock and foam. Using the estimated pose of the sensor obtained from the manipulator forward kinematics, the individual scans shown in Figure 5 are stitched together to form a representation of the larger scene. The stitched image, shown in Figure 6a, depicts the scene in its entirety. However it is noticed that misalignments caused the stitching to contain errors. Misalignment in a ruler that was placed in the scene was significantly noticeable.

Upon examining the misalignment, a correction factor was applied which improved the result as shown in Figure 6b. These results highlight the need for accurate calibration when mounting the sensor to the robotic system. Given the high level of detail capable of being measured, small misalignments can

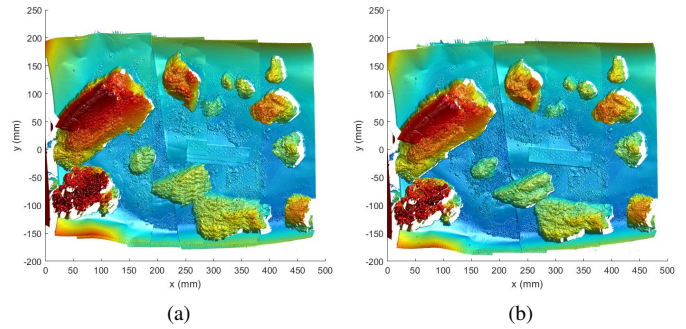


Fig. 6. (a) Individual measurements combined to reconstruct the larger scene. Due to slight misalignments of the sensor mounting, artifacts in the reconstruction are observed. (b) Applying a correction to the sensor positions allowed the reconstruction to be improved.

lead to significant artifacts when stitching together multiple measurements.

D. Changing lighting conditions and material properties

Figure 7a shows the setup with a container of wet cement allowed to cure whilst being measured with the sensor. Sensor measurement were taken 10 minutes and 20 minutes after the cement was poured. By the third measurement, about 30 minutes after pouring, the cement was dry to the touch and no further changes in visual properties were noticeable by eye. Figure 7b shows the external light source that was shone on to the cement and into the sensor lens during the experiment. The objective was to observe variation in the quality of the sensor output during the curing process.

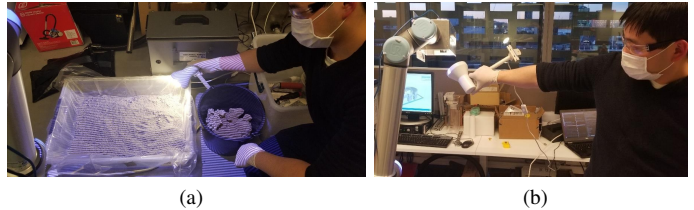


Fig. 7. (a) Container of cement that is mixed with water and allowed to cure during the experiment. (b) An external light source is shone on the cement to reflect into the sensor, and also shone directly into the camera.

The results shown in Figure 8 demonstrate negligible differences between the sensor measurements during the experiment. Some small differences are observed where plastic was used to line the container. The lining was not measured as well with the external light source used due to its highly reflective surface finish.

E. Sensor accuracy with 1.8 m sensor calibration

The setup of the experiment is shown in Figure 9. The scene contains several objects with simple geometry. A large, smooth sphere of approximately 200 mm diameter is chosen as the primary object used to evaluate measurement accuracy. With the sensor calibrated for a nominal operating range of 1.8 m,

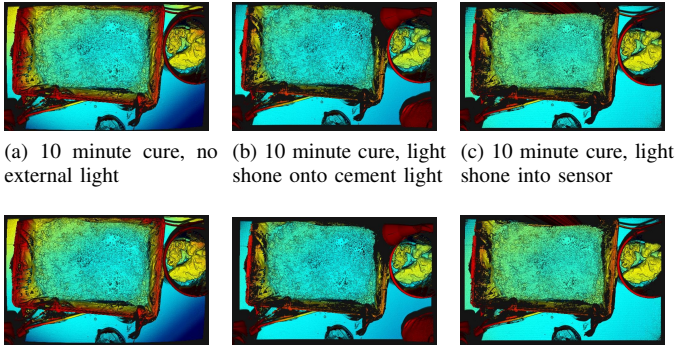


Fig. 8. Results captured during the experiments showed no significant differences with the light source off, being shone on to the cement, or shone directly into the sensor. Negligible differences was also observed as the cement cured.

the sensor is moved using a robot manipulator between 1.0 m and 2.2 m from the scene.

During repositioning of the sensor the large spherical object is kept within view. Individual measurements are shown in Figure 10. The measurements corresponding to the sphere are isolated for analysis. Figure 11 shows the data analyzed with a sphere of best fit at sensor distances of 1.8 m and 1.0 m. Accuracy is evaluated by calculating the RMS error between the measured spherical shape and an ideal sphere of best fit. Table I shows the calculated RMS error of the fitted sphere to

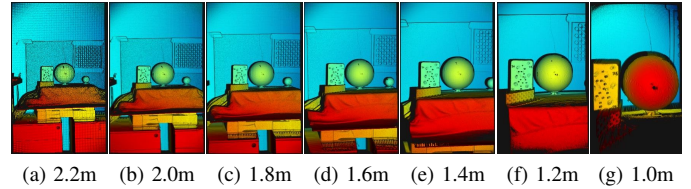


Fig. 10. Measurements of the scene collected at distances between 2.2m and 1.0m. Sensor calibrated for 1.8m range.

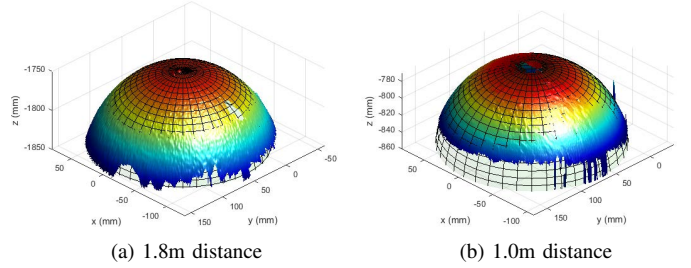


Fig. 11. Measurements with sphere of best fit shown.

TABLE I
ACCURACY OF THE SPHERE MEASUREMENTS VERSUS DISTANCE TO THE SCENE, CAMERA CALIBRATED FOR 1.8M DISTANCE

Distance to scene	RMS Error	Sphere best fit diameter
2.2 m	4.1519 mm	201.7526 mm
2.0 m	6.2298 mm	203.5877 mm
1.8 m	3.0127 mm	201.6306 mm
1.6 m	2.2485 mm	201.8684 mm
1.4 m	1.6882 mm	201.3326 mm
1.2 m	0.6470 mm	200.2948 mm
1.0 m	0.6351 mm	200.1048 mm

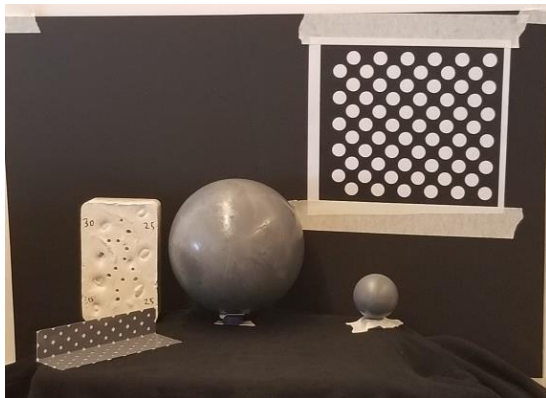
the data versus sensor distance.

The results show that the RMS errors tends to increase at greater distances. Despite being calibrated for a range of 1.8m, RMS error is reduced by taking measurements closer to the scene. The sphere of best fit was consistently calculated to be within a few millimeters of the nominal diameter, demonstrating the reliability making of geometric measurements despite changes to the distance from the scene.

F. Sensor accuracy with 1.0 m sensor calibration

A shorter range version of the experiment is performed, with the sensor calibrated for a nominal operating range of 1.0 m, and located at distances ranging between 1.5 m and 0.75 m from the scene. Figure 12 shows the measurements taken from the sensor.

Using the same method as the previous experiment, the measurements corresponding to the sphere are isolated and analyzed. Figure 13 shows the data analyzed with a sphere of best fit at sensor distances of 1.2 m and 1.0 m. Table II shows the calculated RMS error of the fitted sphere to the data versus sensor distance.



(a)



(b)



(c)

Fig. 9. Experimental setup showing the scene being measured and the sensor mounted on the robot manipulator used to position the sensor.

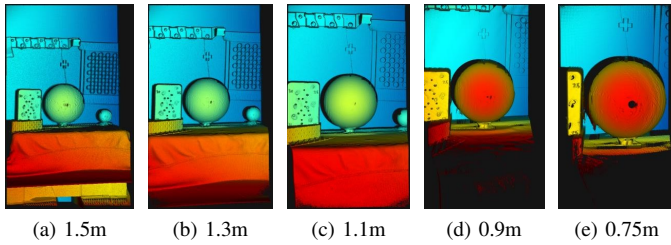


Fig. 12. Measurements of the scene collected at distances between 1.5 m and 0.75 m. Sensor calibrated for 1.0 m range.

Similar to the previous experiment, these results show that the RMS errors tends to increase at greater distances. In this case the increase in error was less significant. The sphere of best fit was again consistently calculated close to the nominal 200 mm diameter, this time with less variation in the calculated values.

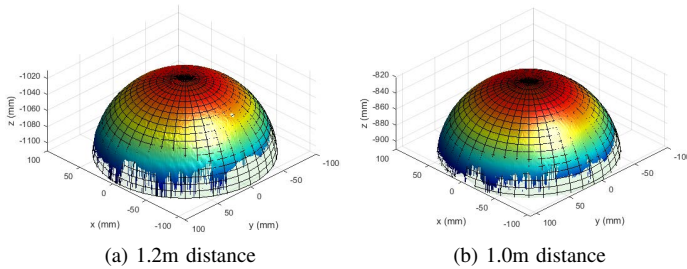


Fig. 13. Measurements with sphere of best fit shown.

TABLE II
ACCURACY OF THE SPHERE MEASUREMENTS VERSUS DISTANCE TO THE SCENE, CAMERA CALIBRATED FOR 1.0M DISTANCE

Distance to scene	RMS Error	Sphere best fit diameter
1.4 m	0.7670 mm	199.8470 mm
1.3 m	0.5492 mm	199.9022 mm
1.2 m	0.4373 mm	199.9470 mm
1.1 m	0.4613 mm	199.8776 mm
1.0 m	0.4274 mm	199.8846 mm
0.9 m	0.4377 mm	199.8608 mm
0.8 m	0.5269 mm	199.7292 mm

IV. DISCUSSION

The results obtained in the experiments indicate that the sensor is a promising solution for construction robotics. The measurements obtained contain high detail, likely more detail than needed for many construction tasks to be automated. Robustness of the measurements against challenging lighting conditions was also shown, however further testing in alternative lighting conditions should be performed.

As well as being functional, there are other requirements for sensors to be practical for use in construction. The construction environment can be harsh, requiring sensors to be suitably

protected. This requires the sensor to be protected from ingress, which the evaluated sensor in its current form is not. Ingress protection may be added at the expense of size and weight. This also needs to be considered as the form of the sensor can be a critical factor, especially when mounted on the end-effector [4].

The sensor demonstrated only relatively small reductions in accuracy when operating outside of its nominal calibrated range. This characteristic is beneficial for construction applications where the size of the environment, and hence the operating range may vary from site-to-site, or even within the same site. Future work is necessary to characterize the reduction in sensor accuracy when operating outside of the nominal calibrated range.

V. CONCLUSION

In this work we analyzed the suitability of a custom projected light 3D sensor for use with construction robotics. The custom made sensor was evaluated with regards to its ability to capture detailed measurements robustly and accurately. Using scenes containing materials commonly found on construction sites, the sensor was found to be capable of capturing high fidelity measurements of the scene. Accuracy of the sensor was evaluated and shown to be capable of taking accurate geometric measurements of objects in the scene at different distances. The results indicate that the evaluated sensor is a promising solution for construction robotics.

REFERENCES

- [1] P. Manamperi, P. Brooks, W. Kaluarachchi, G. Peters, A. Ho, S. Lie, A. To, G. Paul, D. Rushton-Smith, S. Webb *et al.*, "Robotic grit-blasting: Engineering challenges," *Sustainable Bridges: The Thread of Society*, 2011.
- [2] E. Gambao, C. Balaguer, A. Barrientos, R. Saltaren, and E. A. Puente, "Robot assembly system for the construction process automation," in *Proc. Int. Conf. Robotics and Automation*, vol. 1, Apr. 1997, pp. 46–51 vol.1.
- [3] S. Ashizawa, Y. Kuromiya, T. Watanabe, Y. Suzuki, O. Takeo, and J. Maeda, "Development of dismantling robot for ceiling boards human-robot cooperative system and system design of the robot," in *Proc. IEEE/SICE Int. Symp. System Integration (SII)*, Jan. 2009, pp. 30–35.
- [4] M. Lussi, T. Sandy, K. Drfler, N. Hack, F. Gramazio, M. Kohler, and J. Buchli, "Accurate and adaptive in situ fabrication of an undulated wall using an on-board visual sensing system," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, May 2018, pp. 1–8.
- [5] J.-S. Hyun, B. Li, and S. Zhang, "High-speed high-accuracy three-dimensional shape measurement using digital binary defocusing method versus sinusoidal method," *Optical Engineering*, vol. 56, no. 7, p. 074102, 2017.
- [6] S. Zhang, "High-speed 3d shape measurement with structured light methods: A review," *Optics and Lasers in Engineering*, vol. 106, pp. 119–131, 2018.
- [7] C. Zuo, S. Feng, L. Huang, T. Tao, W. Yin, and Q. Chen, "Phase shifting algorithms for fringe projection profilometry: A review," *Optics and Lasers in Engineering*, vol. 109, pp. 23–59, 2018.
- [8] Y. Wang, S. Zhang, and J. H. Oliver, "3d shape measurement technique for multiple rapidly moving objects," *Optics express*, vol. 19, no. 9, pp. 8539–8545, 2011.
- [9] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, 2000.
- [10] S. Zhang and P. S. Huang, "Novel method for structured light system calibration," *Optical Engineering*, vol. 45, no. 8, p. 083601, 2006.