www.cad-journal.net
Computer-Aided Design AND APPLICATIONS™

Evaluation of a Stereo Depth Camera for Part Monitoring For CNC Machining

Samuel Delattre[1] , Sanjeev Bedi[2] , Stephen Mann[3] , Allan Spence[4]

[1]University of Waterloo, samuel.delattre@uwaterloo
[2]University of Waterloo, sbedi@uwaterloo.ca
[3]University of Waterloo, smann@uwaterloo.ca
[4]Niagara College,

Corresponding author: Stephen Mann, smann@uwaterloo.ca

Abstract. A stereo-depth camera is proposed to be used in conjunction with fiducial markers on a calibration plate and a fine-tuning alignment algorithm for part monitoring in a CNC machine. Together, a selected pyramid-shaped part within the machine is monitored. The position, orientation, geometry and surfaces of the pyramid part are measured and compared with the pyramid's desired model.

This system can monitor the position and geometry within 1mm of accuracy, orientation within 1 degree of accuracy, and surface fitting within 2mm of accuracy, which closely aligns with the advertised accuracy of the stereo depth camera. While the accuracy is not enough to verify that the part was machined to an industrial tolerance, this accuracy is sufficient to show that a part roughly matches the position and expected geometry of the model.

Keywords: CNC Machining, Smart Machining, Stereo Depth Camera
DOI: https://doi.org/10.14733/cadaps.2025.91-106

1 INTRODUCTION

Computer Numerically Controlled (CNC) machines are used to remove material either in large batches of similar parts or for quick manufacturing of smaller number of custom parts. The workflow requires the generation of G-codes to program the machine. For large batches, the G-codes are typically optimized to reduce air cutting. For small batches or custom parts, the G-codes are not optimized as the optimization time may be larger than time spent in air cutting. The unoptimized G-code necessitates the presence of an operator to ensure no untoward tool motion takes place as such a move may cause damage to the machine, spindle, part, or all of them. The ever presence of an operator adds to the cost of operating a CNC machine. This cost is significant as it is a recurring cost and adds up over the life of a CNC machine [1]. This cost was the reason for the offshoring (wherein industry moved their machining operations overseas in search for lower labor cost) seen in the machining industries. As machining went offshore, the interest among youth to pursue a career in

machine related jobs also declined. Many who enter the field lack manufacturing experience and are unaware of practical skills and knowledge that would aid their work [10].

Recent years has seen a reversal in the offshoring trend in the machine related industries. The delay experienced by industries during the pandemic, the rising cost of labor overseas, and problems with the supply chain are some of the reasons behind this reshoring trend. The issues impeding the reshoring trend include the lack autonomous ability in CNC machines and the lack of skilled machine operators with knowledge and experience in efficient toolpath planning and awareness of machining conditions that may lead to chatter. One way to assist in the return of this industry is by reducing the dependence on experienced labor. The solution involves developing systems and embedding specific knowledge and experience into a more autonomous CNC machine, such as proposed by Poon et al. [11], who suggested a system that is more aware, using a CAD model of the workpiece existing inside of the CNC controller. The CAD model existing inside of the controller would allow for real-time tool position generation and simulation before sending the commands to be run on the CNC.

While some modern CNC machines may include some sensors for determining if the tooling is broken or if the machine is attempting to move the gantry further than the machine is capable, these solutions add additional movements and time during the machining process. And while these potential issues can often be remedied by taking slower or more shallow cuts, this remedy leads to longer manufacturing times.

The overarching issues are the requirement and loss of specific knowledge required by CNC machinists, a general lack of quick and useful autonomous feedback to inform the CNC if issues may have occurred, and the inability of current machines to attempt to fix issues themselves without direct assistance from an operator.

In this paper, we explore the viability of using a stereo depth camera for observing a part within a CNC workspace. Our solution uses three physical components to monitor the part while it is being machined and validate that the machining was completed to a certain tolerance. The components consist of an Intel RealSense stereo depth camera, an Epson LCD Projector H550A, and three or more fiducial markers placed within the CNC machine. Our tests show the feasibility of the idea, with our system matching the part to within the resolution of the camera. However, additional work is required to make our system commercially viable.

The CNC workspace is used by the part, jigs, fixtures, tool, tool changer and sometime by robotic pallet changers. The workspace is made difficult to observe by flying and accumulating chips and coolants of different types. In this environment, observation can take-on many forms depending on the accuracy and the observation system. The observation system can be used to check that the part is loaded in the correct orientation; is located in the desired part of the space; the jigs and fixtures are in place; the desired tool and tool type are in the spindle; digitize the rough shape of the raw stock to reduce in air cutting; the correct tool is in the spindle; the tool dimensions, the part location and coordinate system; the shape of the stock during the cutting process; etc. All the tasks mentioned above can aid in safe and accurate machining of parts, but do not require the same level of sophistication in the workspace observation system. This paper explores the accuracy of observation that can be achieved with the current low cost camera systems and the tasks the observations can be used for. In the future the camera systems will improve and increase the list of task that can be accomplished.

2 BACKGROUND

Part positioning and geometry verification are areas of exploration to make machines more automated and aware. Dimensional metrology technologies such as Computer Measuring Machines (CMM) are more common in the manufacturing and machining industries. These typically involve the use of a probe to determine the position of the part or features. CMMs can be programmed to automate inspection and improve productivity [8]. Software is used to compare the position and geometry of the probed parts. These machines rely on physically probing the part, which limits their speed. Alternatively, 3D scanning technologies have been

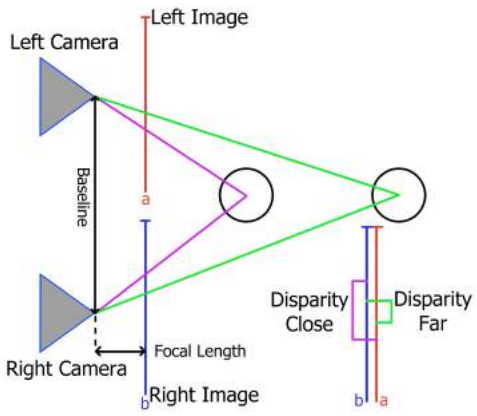


Figure 1: Diagram of stereo depth perception. Each camera captures its own image of the scene. The *correspondence problem* must be solved to find the matching points between the two images. By superimposing the left and right images, the *disparity* is the pixel positional difference between the matching points found in each image. Objects that are closer to the cameras have a larger disparity compared to objects that are further away.

experimented with for non-contact dimensional metrology [5].

3D scanning technologies and cameras have been primarily used in the robotics industry to make robots more aware of their surrounding. Many groups such as Tadic et al. [12] have applied depth cameras in robotics to detect and extract key geometry such as obstacles.

An application of 3D scanning technologies is automating part positioning to make machines more aware without human intervention. Pajor et al. [9] discuss a vision system for quick matching a workpiece reference point using a non-contact 3D scanning method, based on structural light patterns.

One method of improving the resulting scans taken using stereo depth scanning technologies was explored by many groups, including Madeira et al. [7]. Fiducial markers were used to better refine transformations rather than relying on aligning features that may appear differently between multiple scans or how they are expected to appear. The fiducial markers can be robustly detected which helps to overcome incorrectly matching features.

In this work, a stereo depth camera scans a part within the workspace to determine its position, rotation, and geometry. ArUco markers are used to initially align the camera's coordinate system with the CMM or CNC machine workspace. An additional algorithm is used to further improve the precision and accuracy of the scans for comparison with either the desired position of the stock material or the final machined part.

3 TEST SETUP

Our system consists of a stereo depth camera, a calibration plate with fiducial markers, a projector, and a pyramid-shaped part. To test our system, these components were mounted on a CMM, as shown in Figure 3.

At the core of our system is a stereo depth camera. The depth camera we used was the Intel RealSense D405 (Figure 2), a stereo depth camera with two cameras, designed for close-range applications. In the stereo camera, the two cameras are positioned apart from one another, and perceive different views. Given the same point in each image, to determine the distance the point in the camera's view, the pixel positional difference between each pair of matching points is found as shown in Figure 1. The depth calculation is done

Figure 2: Intel RealSense D405 Stereo Depth Camera used for experimentation

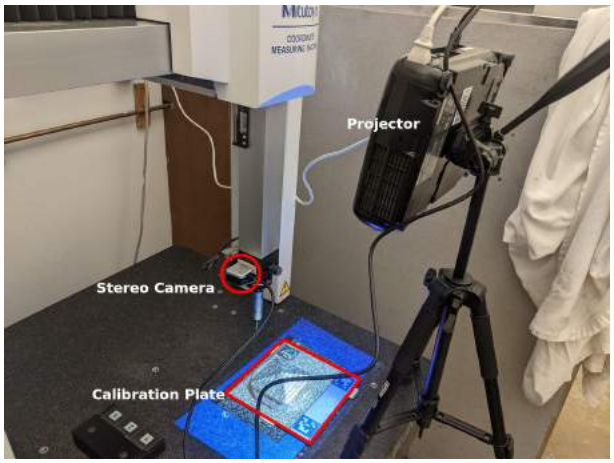


Figure 3: Test setup overview for applying the stereo depth camera and algorithms to detect a pyramid-shaped part. The stereo camera is mounted to the head of the CMM for camera location consistency. The calibration plate is mounted in the workspace, and one of the pyramid parts is mounted above. The projector projects a pattern onto the pyramid surface to improve scan quality.

by multiplying the baseline by the focal length divided by the disparity [4].

$$Depth = \frac{Baseline \times FocalLength}{Disparity} \tag{1}$$

The correspondence between points in the two images are computed in hardware by the the Intel camera, which matches up to 36 million depth points/second using a custom variant of a Semi Global Matching algorithm [4], with an advertised 0.05 sub-pixel accuracy for well-textured passive targets.

To associate the physical part and the model for comparison, the stereo camera must find the physical part in the coordinate system of the CNC machine. However, the stereo camera, the CNC machine, and model have different coordinate systems. The connections between the coordinate systems must be determined.

Our method introduces an intermediary coordinate system created by a plate with three precisely located fiducial markers on it, which is used to calculate an initial rotation and displacement to be applied to align the camera's coordinate system with the plate. The rotation and displacement are then fine-tuned using a pyramid that is scanned by the camera. The scanned points are fit to the four faces of the pyramid to determine the pyramid's location and rotation to be applied to future scans.

The plate is fixtured inside the CMM. The stereo depth camera was positioned above the workspace,

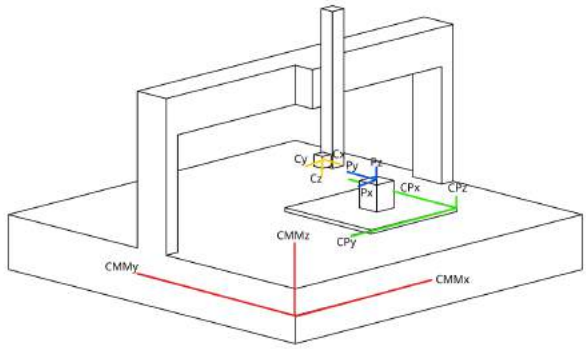


Figure 4: Coordinate systems to be aligned with alignment algorithms. The CMM coordinate system (CMM) is shown in red. The calibration plate (CP) coordinate system is shown in green. The pyramid part (P) coordinate system is shown in blue. The camera (C) coordinate system is shown in yellow. Since all coordinate systems are mounted to the CMM, the CMM was used to verify the location of each coordinate system.

facing downwards, having a view of the plate and the markers engraved onto it. The camera's orientation was chosen to face down towards the part, assuming that all required features would be only on the top face of the part. The CMM's coordinate system is calibrated to be aligned with the edge of the plate, and the stereo camera identifies the marker's location in the 2D image created by the stereo camera. The pixel locations of the markers in the 2D image can be converted to their location within 3D space. The camera then uses these markers to represent the coordinate system of the plate and roughly aligns point clouds from scans taken to the plate's coordinate systems. Finally, the pyramid part is placed in the center of the camera's field of view, and an external projector is used to project a pattern on the pyramid's surface to help improve the scan quality.

The choice to calibrate the CMM's coordinate system with the edge of the plate was used to verify the location of the plate's coordinate system and allow the CMM to be used to verify the locations of the markers on the plate that the camera detects in addition to the pyramid part coordinate system and the pyramid's features. The location of the stereo depth camera's coordinate system did not need to be verified because we used the markers and alignment algorithms to match the camera's coordinate system with the calibration plate and pyramid part coordinate systems. However, verification was completed by mounting the camera to the head of the CMM's probing tool. This configuration allowed the precise location of the camera to be measured by the CMM with respect to the plate's coordinate system that the CMM was aligned with.

4 CALIBRATION

The coordinate systems used in our system are shown in Figure 4. The coordinate systems are used to derive transformations that are applied to the points from the camera's coordinate system, C, to align them with the pyramid part coordinate system, P. Additionally, transformations from the CMM's coordinate system, CMM, to the pyramid part coordinate system, P, were used to verify the pyramid part features. The point cloud created by the stereo depth camera is initially in the camera coordinate system, C. The point cloud is then mapped to be relative to the orientation and position of the plate in the calibration plate's coordinate system, CP, and then is further tuned using a pyramid calibration piece in coordinate system P. Similarly, when the CMM measures the object, the measurements are mapped from the CMM's coordinate system, CMM, to be relative to the orientation and position of the plate's coordinate system, CP. By aligning the CMM with the plate's coordinate system, the CMM was used to verify coordinate systems and pyramid features. The alignment of the camera's coordinate system with the plate and pyramid coordinate systems allows for using either existing point registration algorithms or other methods to compare the desired model and the point

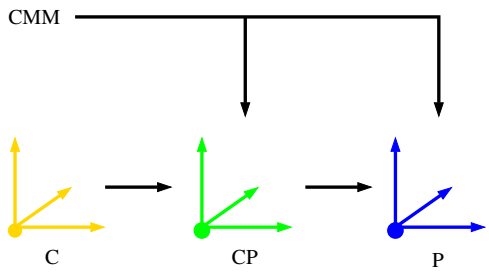


Figure 5: Mapping between coordinate systems. The calibration plate (CP) coordinate system, green. The pyramid part (P) coordinate system, blue. The camera (C) coordinate system, yellow. Since all coordinate systems are mounted to the CMM, the CMM was used to verify the location of the CP and P coordinate systems.

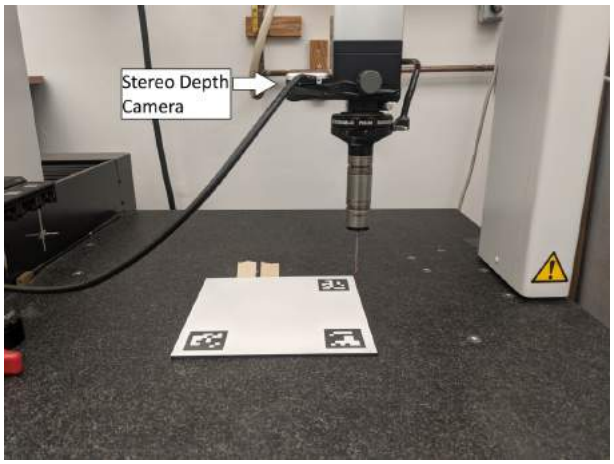


Figure 6: Stereo depth camera is mounted to the CMM head. The calibration plate is mounted in the workspace below. The CMM and stereo depth camera can be calibrated to the same coordinate system as the plate.

cloud created from scanning the machined part. The comparisons were then be verified by the CMM. Figure 5 illustrates these mappings.

The calibration of the CMM with the plate was handled by the CMM. We then used a two-step method to calibrate the stereo depth camera with the plate that first uses fiducial markers precisely engraved upon the plate to be found by the stereo depth camera as shown in Figure 6. In the second step, the alignment is fine-tuned using an algorithm for finding and fitting a calibration pyramid.

The scanning and alignment objective is to determine the geometry and location of the pyramid part within the CNC or CMM workspace. The algorithm was split into two separate steps, one for scanning, a rough alignment, and cropping the point cloud and another for aligning the point cloud with the desired pyramid model and outputting results. The scanning algorithm is used to scan the pyramid part and roughly align it using the fiducial markers on the calibration plate, similar to the method used by Madeira et al. [7]. The alignment algorithm fine-tunes the alignment based on the planes fitted to the scanned points and the apex that is calculated based on the planes.

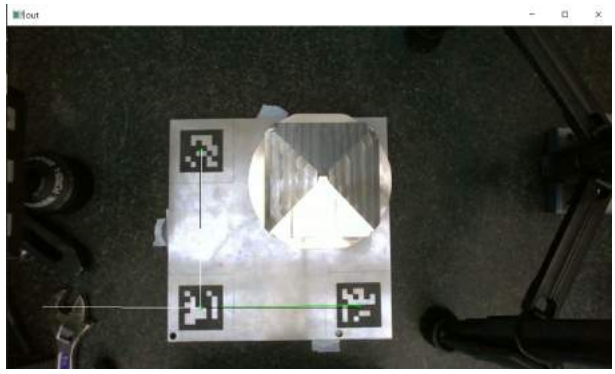


Figure 7: Stereo depth camera colour view with ARuCO marker detection for camera calibration

4.1 Application of Fiducial Markers

Stereo depth cameras can be used to find fiducial markers (QR codes). By designing a plate that can be aligned inside of the machine, markers can be precisely added to the plate to create three points at known offsets and orientations from each other to symbolize the origin of the plate and two other points. The markers are used to define two vectors on the surface of the plate, representing the X and Y axes of the plate. These markers allow the stereo depth camera to capture a 2D image of the markers, and the marker's pixel position within the 2D image can be determined using fiducial pose estimation algorithms.

The marker's pixel position are used to determine the position of the markers within 3D space in the camera's field of view. The three marker positions are then used to construct the X and Y axes of the plate. The cross product of these axes gives the Z axis of the plate. Together, these three axes, where X and Y are shown in Figure 7, are found in the camera's coordinate system and are compared to the desired axes found in the plate's coordinate system. This comparison results in a transformation matrix that defines a rough transformation between the stereo camera's coordinate system to the coordinate system of the plate. In our tests, this transformation locates the point clouds to an accuracy of a few centimetres of the real world, as measured by the CMM. This accuracy is several orders of magnitude worse than Intel's minimum object detection claim for the camera of 0.1mm. Therefore, we performed a second calibration step to improve the accuracy. However, the initial rough tuning is useful to crop the point cloud around the desired area, and the cropped area are used to further tune the alignment.

4.2 The Calibration Pyramids

A calibration pyramid is used to improve upon the calibration results found after using the calibration plate. Two pyramids were used during testing. One pyramid was 10cm×10cm with a 2cm apex height, machined from aluminum shown in Figure 8a. This pyramid has some scallops on the surface from machining. The second was 5cm×5cm with a 20-degree incline for the faces, machined from a blue Ferris File-A-Wax shown in Figure 8b. The height of the apex was determined by the angle of the faces, resulting in a height of around 9.1mm.

4.3 Fine-Tuning Calibration Algorithm

The stereo depth camera creates a cloud of points. To align this point cloud with the pyramid, our fine tuning algorithm first splits the points into four groups based on their position relative to the position of the pyramid faces. Four planes were calculated based on the points using a least square fit. The distance from each

(a) (b)

Figure 8: (a) Blue Ferris File-A-Wax Pyramid. (b) Aluminum Pyramid

point to each of the four planes was found, and the points were redistributed among the four groups based on which plane they were nearest. This process was repeated until the size of the groups did not change, to a maximum of 50 iterations. This part of the algorithm is illustrated in Figure 9, where to illustrate the algorithm, we sampled the pyramid at random points, and then rotated the points. The scanned data is initially in closer alignment to the pyramid than the points are in this figure; we have exaggerated the error in the initial alignment so that the adjustments are visible in the figure.

Once the groups and points were established and their planes were calculated, the pyramid's apex was found using the pyramid's geometry, where the four planes overlap should be the apex. Due to both the noise in the scans and errors depending on the position of the pyramid within the camera's view, the planes did not always overlap where they should. In Figure 9 (c)–(f), you can see the misalignment; while the misalignment is not visible in parts (g) and (h) of this figure, higher accuracy is required to have an alignment approaching the tolerance of the camera.

To improve the alignment, the previously calculated apex was compared with the desired apex and the points were offset by the difference. The four groups of points were then rotated around the Z axis passing through the apex to put them on the same face of the pyramid. All of the points together were used to calculate a new plane. By using all points together to calculate one plane that could be rotated to use for all faces of the pyramid, the error from the sides further from the center of the camera's view and noise in the point cloud is reduced, resulting in a better representation of the pyramid. Additionally, the new plane is compared with the actual plane of the pyramid face to determine two rotational offsets: the first rotated the plane around the Z axis and the second around the Y axis.

The points were then rotated based on the two angles found, and then the plane was rotated around the Z axis passing through the apex to form all four faces. The apex was recalculated based on these faces. The offset of the points to align with the desired apex and the two angles found while aligning the face of the pyramid were used further to calibrate the stereo depth camera's point clouds. This fine-tuning calibration improved the accuracy of the point clouds produced from within centimetres to within millimetres of the real-world part.

4.4 Validation Metrics

To determine the viability of the stereo camera and algorithms used, two validation metrics were chosen. These were part position and part geometry. The part position encompasses how well the location and any rotations applied to the part can be determined, while part geometry includes how well key geometric features of the part can be found and how well the points fit to the desired surface. The validation metrics are shown in Figure 10, which is a top view of the calibration plate and pyramid on the CMM.

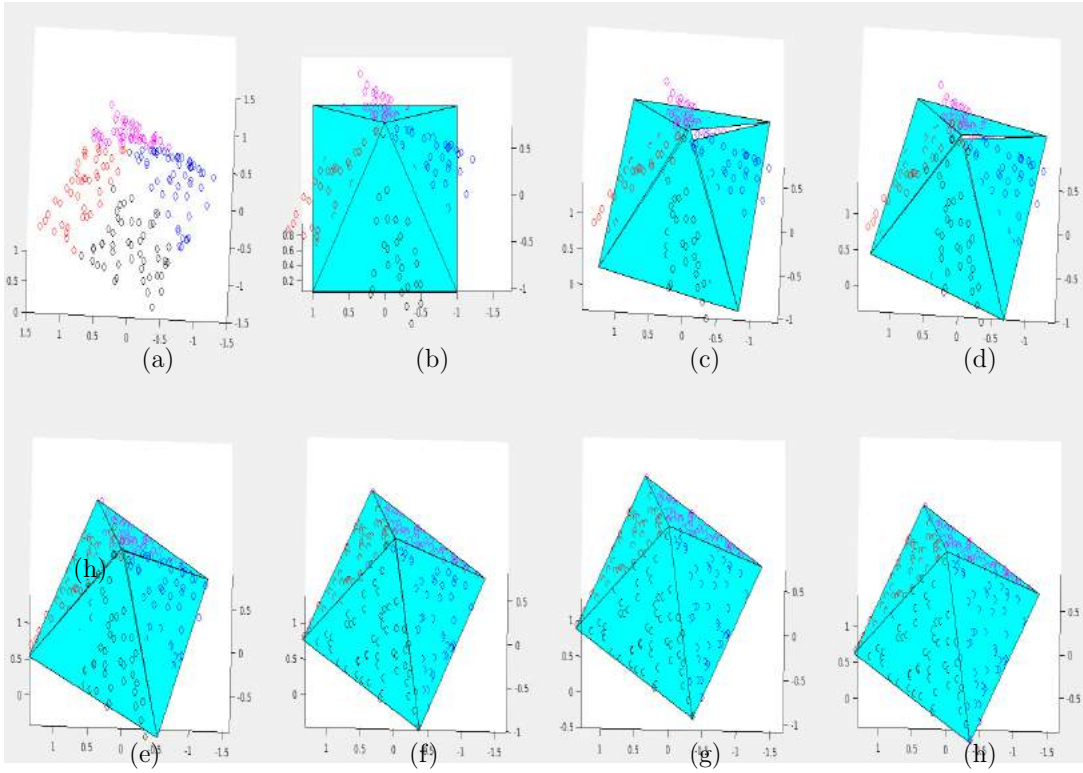


Figure 9: Aligning the pyramid with the point cloud through repeated plane fitting. Points are colored based on which face they are closest to. (a) The point set. (b) The pyramid aligned using just the fiducial markers. (c)–(h) Iterations 1 to 6 of the algorithm.

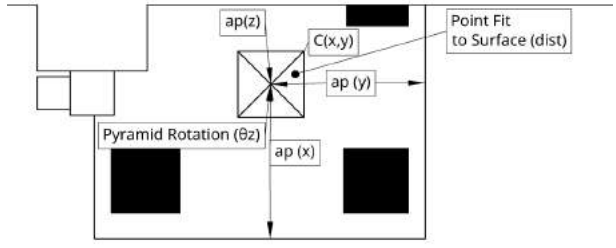


Figure 10: Top view sketch of validation metrics on the CMM. The part position is determined by $ap(x)$, $ap(y)$, $ap(z)$ and pyramid rotation θ_z. The part geometry is determined by $ap(z)$, $C(x, y)$, and the point fit to surface distance.

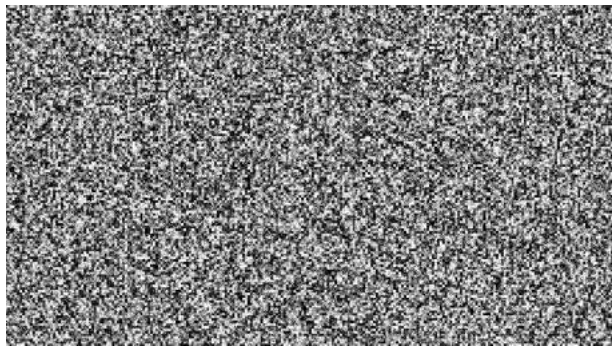


Figure 11: Semi-random pattern projected on less-textured objects to improve point cloud quality.

Since a pyramid was used for testing, part position will be determined by the location of the pyramid's apex and any rotation around the Z axis running normal to the bottom of the pyramid through the apex. Part geometry will be determined by the location of the apex and corners of the pyramid and the point fitting with the surface of the pyramid.

Due to the expected noise found in the scans, five or more scans were taken, and the results averaged to determine a more accurate result. Additionally, these scans were used to determine the precision of the metrics.

5 RESULTS

The mapping from the camera's coordination system to both the calibration plate's coordinate system and pyramid part coordinate system were verified using a CMM. Rather than using the CMM to verify the other coordinate system mappings, we verified these mappings by testing the mappings of the point clouds created on the camera and their mapping to the part and part features.

The algorithms were implemented using C++ to capture the point clouds and Matlab to process the resulting point clouds. Although the methods developed were for pyramid-shaped calibration parts, the algorithms can be generalized for other designs by performing multiple least square fits for the points on the desired faces of the part. The Intel RealSense D405 stereo depth camera has advertised object detection of 0.1mm at 7cm [6] in ideal conditions; our goal is for our algorithm to have an accuracy within an order of magnitude of this advertised accuracy.

We begin in Section 5.1 by giving our experiments on testing various lighting conditions to find lighting that worked well with our setup. In Section 5.2, we tested the impact of the part's position relative to the camera. And in Section 5.3, we present the results of our plane fitting algorithm, which demonstrates the accuracy we could obtain in locating the pyramidal part with our setup.

5.1 Lighting and Pattern

The effects of lighting are a significant consideration for the quality of any scans taken. In our work, the use of the pattern projected on any part to be scanned and proper lighting were found to significantly impact the quality and accuracy of the scans produced. Three separate tests were performed, with the part positioned in the center of the camera's field of view. The camera and part positions were not varied between tests. Each test consisted of five scans that were taken and averaged. In the first test, scans were taken with normal room lighting. The second test used the projector to light the part with a white-coloured light. The final test projected a pattern onto the surface of the pyramid.

Maximum Overall Offset	Minimum Overall Offset	Maximum Apex Offset	Minimum Apex Offset	Maximum Corner Offset	Minimum Corner Offset
16.280	0.974	-2.560	0.173	7.034	0.029
Found at position 5-5	Found at position 3-3	Found at position 5-5	Found at position 4-3	Found at position 5-5	Found at position 4-2

Table 1: Maximum and minimum deviation (in mm) of Pyramid Overall Offset, Apex Location, and Corner Location. Overall offset was found by the vector length created by the XYZ error.

A consumer-grade projector, an Epson LCD Projector H550A, was used to project a semi-random pattern that avoided periodic arrangements of dots while also providing better lighting for the scene. Some tests were completed without the pattern to observe the improvements when the pattern was projected. The pattern was obtained from Intel RealSense documentation [3] and is shown in Figure 11.

With average room lighting, the resulting point cloud was noisy, to the point that the results did not represent the physical model of the pyramid. When the projector was used to light the scene, the results were more representative of the dimensions of the pyramid. The data was still too noisy due to the pyramid's relatively untextured surface, making it difficult for the correspondence algorithm to find the correct distance for all of the points.

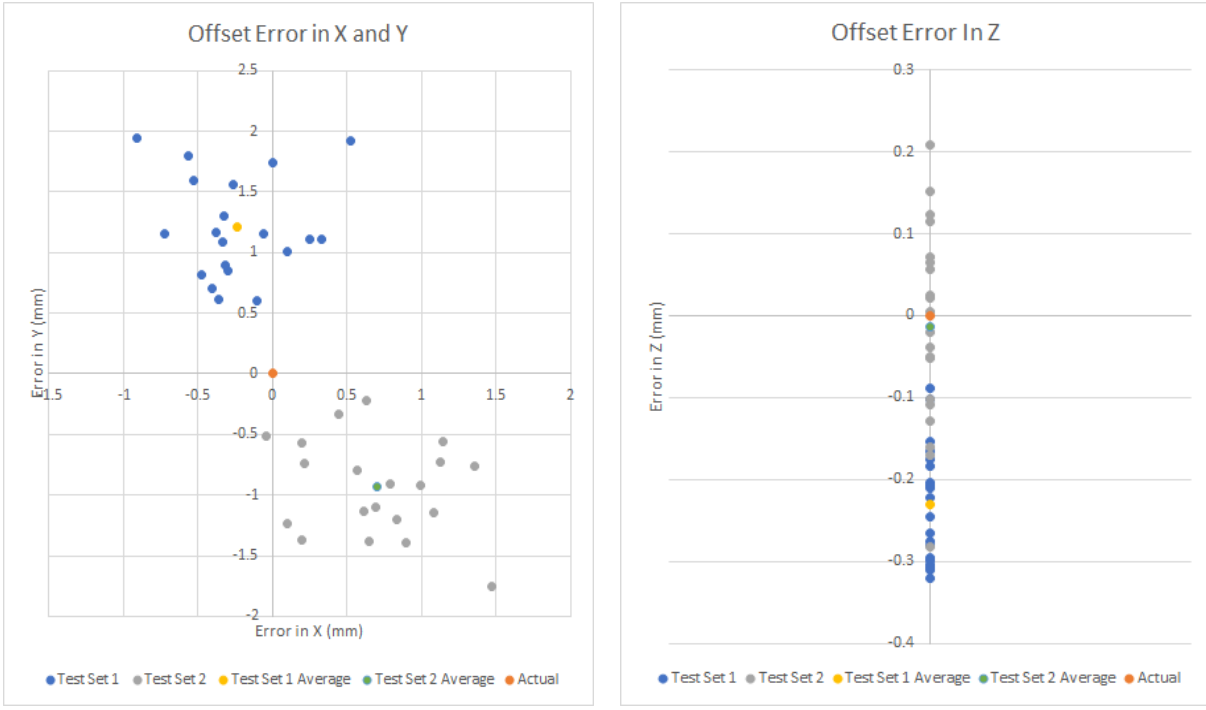
Finally, when the pattern was projected on the pyramid's surface, the results were closer to the part dimensions, with deviation being 5 to 10 times less than just projecting light. The pattern provided an artificial texture to the pyramid, while also lighting the scene better in the process. Since the pattern used was made from a pixel-like pattern, the ideal size of the pixels should be near matching the stereo depth camera's pixel resolution. If the pattern's pixels are smaller than the pixel resolution of the camera, they will not improve the results of the correspondence problem, and therefore the results of the point cloud scans. However, if the pattern's pixels are larger than the pixel resolution of the camera, the correspondence problem will have more issues precisely matching the corresponding points within the left and right camera images. This will result in either more noise in the point cloud or incorrect depth values.

5.2 Part Position and Rotation

The part's location within the camera's field of view also significantly affected the results. The camera was mounted on the probe of the CMM, allowing the camera to be moved accurately. The camera was moved such that the part was scanned in 24 locations across the camera's field of view, from the top left of the field of view to the bottom right. Five scans were taken at each location, and the results of the offset, apex height and corner positions were averaged between the five scans.

The metrics observed are the apex height and corner positions to inform the part's geometry and the offset to inform the position of the part. Overall, the most accurate results for both the geometry and offset of the part were found when the part was positioned closer to the center of the camera's field of view. The smallest overall positional offset was found in the center position. In contrast, the most accurate apex height was found just below the center, and the most accurate corner position was found just to the bottom left of the center. Conversely, the most significant errors were found around the edges, specifically the right edge and bottom right corner. However, better results tended to be found in the center of the field of view, making up approximately 70% of the camera's field of view. These results are summarized in Table 1.

The position offset precision can vary and is determined by the location of the part within the camera's field of view. Two separate test sets of 20 scans each were taken. Both tests positioned the part around the center of the image. However, while the ranges of both tests' offset values were similar, the tests produced two



(a) (b)

Figure 12: (a) Offset error in X and Y from two sets of 20 scans, (b) Offset error in Z from two sets of 20 scans. These show the average accuracy and overall precision capabilities of the depth camera and algorithm for finding the position of the Pyramid in XYZ

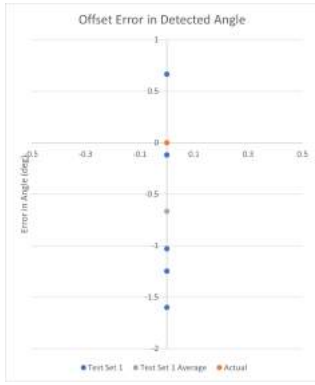


Figure 13: Offset error (in degrees) of pyramid rotated about the Z axis. This shows the average accuracy and overall precision capabilities of the depth camera and algorithm for finding the rotation applied to the Pyramid about the Z axis running through normal to the bottom of the pyramid and through the apex

different sets of values, shown in Figure 12. In Test Set 1, the offsets in X and Y ranged from approximately -1 to 0.5 mm in X and 0.5 to 2 mm in Y, Test Set 2 had its offsets in X and Y range from approximately 0 to 1.5 mm in X and -1.75 to -0.25 mm in Y. The precision was around 1.5 mm in both tests, while the sets of values for both tests differed, shown in Figure 12a. In Z, Test Set 1 had an offset between -0.3 and 0 mm, whereas Test Set 2 ranged between -0.3 and 0.2 mm, shown in Figure 12b.

The rotation of the pyramid part observed by the stereo depth camera was precise and accurate. When the part was rotated around the Z axis running through the apex by a known amount, the rotation error ranged between -1.6 to 0.7 degrees, averaging around -0.7 degrees, shown in Figure 13. The rotational accuracy greatly depended on how well the pyramid faces were fit. This is due to the equation of the plane representing the faces being used to determine the rotation. The plane fitting was not as effective around the edges of the camera's field of view and was better towards the center. Therefore, the rotation of the pyramid could be more accurately observed by the camera when the pyramid was located around the center of the camera's field of view.

The higher accuracy near the center of the camera's field of view is not surprising, since the region in the center of the field of view is closer (giving better resolution in the depth), and the pixels in the center represent smaller areas (given better resolution in xy). Thus, this result (of better accuracy near the center of the field of view) should generalize to non-pyramidal parts.

5.3 Plane Fitting

After the points had been translated and rotated to best align with the desired orientation of the pyramid, the distance between the newly aligned points and the desired planes of the pyramid were measured. Table 2 shows the maximum, minimum, and average distance between the aligned points and the desired planes. The distance was calculated using the equation of the plane and the points, which is shown in Equation 2, where $[A, B, C, D]$ defines the equation of the plane in the form $Ax + By + Cz + D = 0$, and $[x_p, y_p, z_p]$ is the coordinates of the point to be measured:

$$Distance = \frac{Ax_p + By_p + Cz_p + D}{\sqrt{A^2 + B^2 + C^2}} \tag{2}$$

Figure 14, Figure 15, and Figure 16 show the results of the points and planes when the part is in the center

Position 1-5 Max Error	Position 1-5 Min Error	Position 1-5 Avg Error	Position 3-3 Max Error	Position 3-3 Min Error	Position 3-3 Avg Error
2.01	-5.78	-1.28	2.42	-1.57	0.48

Table 2: Maximum, minimum, and average distance (in mm) between aligned points and desired pyramid planes for scans taken at locations 1-5 and 3-3.

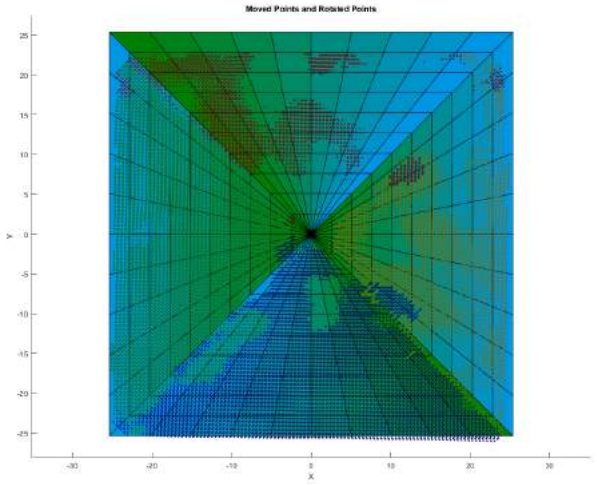


Figure 14: Top view of fitted points and desired planes for pyramid at location 3-3

position 3-3. The points are relatively evenly distributed above and below the desired planes. On average, for position 3-3, the distance between the points and the plane is less than 0.5 mm.

The plane fitting is completed after the points are transformed. Therefore, a similar correlation is found between the location of the part within the camera's field of view and the error in plane fitting. Parts scanned at the edge of the camera's field of view, such as location 1-5, have lower precision and worse accuracy than parts scanned at the center of the camera's field of view, such as location 3-3. Location 1-5 has point distances between -5.78 mm and 2.01 mm compared with location 3-3, with point distances between -1.57 mm and 2.42 mm. Additionally, the average point distance is more accurate at location 3-3, less than 0.5 mm from the desired plane, compared with location 1-5, which is around -1.28 mm from the desired plane.

6 CONCLUSIONS

A method was developed for testing the viability of using the stereo depth camera for observing a part within a CNC workspace, and tested on a pyramid part. An Intel RealSense D405 stereo depth camera was used,

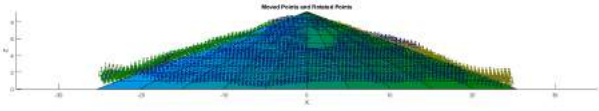


Figure 15: Side view of fitted points and desired planes for pyramid at location 3-3

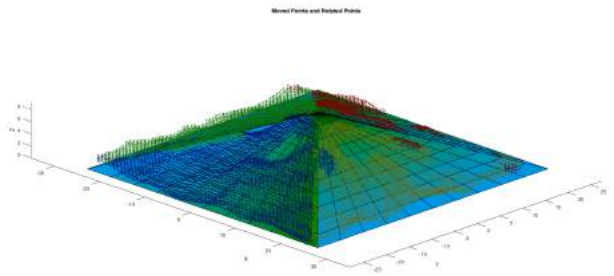


Figure 16: Orthographic view of fitted points and desired planes for pyramid at location 3-3

which alongside a calibration plate with fiducial markers, was first used to determine a transformation to align the camera's coordinate system with the calibration plate and CNC/CMM. The stereo depth camera then scans the work volume and pyramid part to create point clouds, which are transformed by the calculated transformation to roughly align the point clouds. A fine-tuning the alignment is then applied to the roughly aligned point clouds to determine the pyramid part position, orientation, geometry and plane fitting.

The depth camera has advertised object detection of 0.1mm at 7cm [6] in ideal conditions. In our method, the camera and algorithms were able to determine pyramid position and geometry within 1mm accuracy and orientation within 1 degree accuracy in experimental conditions. The points were on average 2mm away from the desired pyramid faces, with good distribution above and below showing a reasonable fitting with the desired planes. These results are comparable with the claimed camera accuracy provided by Intel. Additionally, in addition to the studies discussed in this paper, we did a comparison of our method to Geomagic, which showed that the two methods had similar accuracy [2].

However, the resulting accuracy found would not allow this solution to be used for autonomously determining if a part has been machined properly within a reasonable tolerance for most machined parts. The results could be used to show that a part roughly matching the geometry of the desired model has been manufactured, and could determine if the raw stock was positioned somewhat correctly. In addition, any large errors or obstructions would be easily detected by this system. The largest bottleneck holding back the potential of this project is the point cloud quality scanned by the stereo depth camera. At present, the scans have a large amount of noise. More powerful and expensive scanning methods or future improvements to the technology may improve the results found in this work.

The millimeter level accuracy that can be achieved with the current system can be used to determine the orientation of the part; the location of the jigs and fixtures; the shape of the tool; the location of the part reference frame; etc. The knowledge of these can reduce the potential for accidental damage to part, fixtures, machine and operator. In addition, the camera observations could be used to determine if the path of a tool pass is clear of jigs, fixtures and other unexpected materials, to determine that the part has not moved and is held in place, the fixtures and jigs have not loosened and fallen in the path of the tool and much more. The authors believe that a CNC workspace observation system with cameras can be used to make the machine operation safe and prepare the machine for complete autonomous operation. The above features are not hampered by the difficulty posed by coolant and chips. Most machines have chip removal systems, but chips still collect on and around the part and will effect the 3D scan of the workspace by introducing errors in measurement. The influence of chips in scanning can be reduced with better chip removal and or algorithmically using AI techniques. Similarly, methods to resolve the obstruction caused by the coolant, such as intermittent flow, are required.

Additional details on this system can be found in [2].

ACKNOWLEDGEMENTS

This research funded in part by the Natural Sciences and Engineering Reach Council of Canada (NSERC). The authors are grateful for the assistance of the Niagara College Research & Innovation for laser etching and measurement equipment use.

Samuel Delattre, https://orcid.org/0009-0006-6861-1707
Sanjeev Bedi, https://orcid.org/0000-0002-6993-8502
Stephen Mann, https://orcid.org/0000-0001-8528-2921
Allan Spence, https://orcid.org/0000-0002-6835-6498

REFERENCES

[1] Bengtsson, M.; Kurdve, M.: Machining equipment life cycle costing model with dynamic maintenance cost. Procedia CIRP, 48, 102–107, 2016. ISSN 2212-8271. http://doi.org/10.1016/j.procir.2016.03.110.

[2] Delattre, S.: Application of Stereo Depth Camera and Alignment Algorithms for Part Monitoring During Machining. Master's thesis, University of Waterloo, 2023.

[3] Grunnet-Jepsen, A.; Sweetser, J.N.; Winer, P.; Takagi, A.; Woodfill, J.: Projectors for d400 series depth cameras. https://dev.intelrealsense.com/docs/projectors#5-the-dot-pattern, 2023.

[4] Grunnet-Jepsen, A.; Sweetser, J.N.; Woodfill, J.: Stereo depth cameras for mobile phones. https://dev.intelrealsense.com/docs/stereo-depth-cameras-for-phones, 2020.

[5] Horaud, R.; Hansard, M.; Evangelidis, G.; Ménier, C.: An overview of depth cameras and range scanners based on time-of-flight technologies. Machine Vision and Applications, 27(7), 1005–1020, 2016. ISSN 1432-1769. http://doi.org/10.1007/s00138-016-0784-4.

[6] Intel Corporation: Depth camera d405. https://www.intelrealsense.com/depth-camera-d405/, 2023.

[7] Madeira, T.; Oliveira, M.; Dias, P.: Enhancement of RGB-D image alignment using fiducial markers. Sensors, 20(5), 2020. ISSN 1424-8220. http://doi.org/10.3390/s20051497.

[8] Modern Machine Shop: 5 ways automated on-machine probing improves productivity. https://www.mmsonline.com/articles/5-ways-automated-on-machine-probing-improves-productivity, 2019.

[9] Pajor, M.; Grudziński, M.: Intelligent machine tool – vision based 3D scanning system for positioning of the workpiece. Solid State Phenomena, 220-221, 497–503, 2015. http://doi.org/10.4028/www.scientific.net/SSP.220-221.497.

[10] P.E., M.M.U.; Stufflestreet, B.A.; Johnson, K.V.: Identifying the need for trained machinists in the greater tri-cities area: A survey of employers to evaluate the future of machining. In 2021 ASEE Virtual Annual Conference Content Access. ASEE Conferences, Virtual Conference, 2021. https://peer.asee.org/37271.

[11] Poon, G.; Gray, P.J.; Bedi, S.; Miller, D.E.: Architecture for direct model-to-part cnc manufacturing. Journal on Systemics, Cybernetics and Informatics, 4, 14–18, 2006.

[12] Tadic, V.; Odry, A.; Kecskes, I.; Burkus, E.; Királ, Z.; Odry, P.: Application of Intel RealSense Cameras for depth image generation in robotics. WSEAS Transactions on Computers, 18, 107–112, 2019.