Proceedings of The National Conference
On Undergraduate Research (NCUR) 2016
University of North Carolina Asheville
Asheville, North Carolina
April 7-9, 2016

A Practical Analysis of Bin Picking with Robotic Vision

Arin Preston, Kevin Gilmartin, Ashley Mills, Robert Swick
Mechanical Engineering
Eastern Washington University
Cheney, WA. 99224 USA

Faculty Advisor: Dr. Donald Richter

Abstract

Robotic vision is primarily used for identification, location, or inspection of components. Without vision, a robotic system must rely on complex, cumbersome, compliance devices and sensors to ensure proper placement of an item. Even so, the robot may not be able to know if the part is correct, or distinguish between different parts. This undergraduate research project was to create a continuously operating system using robotic vision system to perform the first two tasks; identification and location. To accomplish this, a vision system that would sort coins that were being fed unsorted and unordered by a conveyor was constructed. The coins provided a simplified type of "bin-picking," where the object's orientation was eliminated as a variable. The major obstacle to overcome was fine tuning the vision settings such that no coins were missed while maintaining high operating speed. To do this, each vision setting was analyzed and configured for each type of coin, which provided the necessary vision capabilities. Preliminary expectation for this research project was to have a 95% coin identification accuracy and able to operate the robot at over 75% of max speed. This research demonstrated the limitations, but also the practical applications associated with complex bin-picking in an industrial environment with a robotic vision system.

Keywords: Robotic Vision, Bin-Picking, Industrial Robotics.

1. Introduction

Vision systems are typically used in robotics for part identification, location, and inspection. When different items must be sorted, it can be done mechanically, based on the physical properties of the differing parts. But what if the parts have the same physical features, and only differ cosmetically, or the necessary mechanical apparatus for sorting would be too cumbersome to feasibly implement? These are situations where robotic vision can take over. The vision system requires no more space than the robot already fills; the only physical addition is that of the camera. The camera takes a picture of the items presented, the computer compares it to a database of images that have been taught, and identifies what it sees. The computer then calculates an offset (coordinate shift, including location (x,y) and rotational shifts) from the center point, to where the part currently is. The robot can then use this offset to retrieve the item, regardless of position or rotation. The vision system was new to our robotics department and it had not had a lot of work done with it. Part of our goal was just to learn about the system and to see what we could learn about it.

One of the more difficult challenges to robotic vision is the "Bin-Picking" problem; where multiple parts are presented unordered and unspaced. The robot has a difficult time accurately sorting the parts from this bin, even with vision. Multiple cameras may be needed to properly define the position and orientation of each part, depending on how large or complex the parts are. For this research, the bin picking problem was simplified to two dimensions by using coins. The flat discs eliminate orientation as a problem, as the rotational position of the coin does not matter. Also eliminated are the majority of the tilted axis issues one would find in typical bin picking problem. Eliminating these complexities allowed us to focus on certain aspects of the vision system, such as how well the camera discerns a single part from a mass of other acceptable targets.

2. Background

Robotic vision is becoming more feasible and cost effective with modern advances in computing. As such, it is gaining popularity in industry. This research focuses on the implementation of robotic vision, and determining when it should and should not be used. Additionally, in the situations where it is optimal to use vision, some key factors to consider.

Machine vision, in its most simple form, is comparing temporary photographs to others stored in permanent memory. These photographs are compared based on criteria set by the programmer, and the system determines whether or not the presented image qualifies as a match or not. Because of how this system operates, the settings must be carefully calibrated, and proper lighting must be maintained. If poor choices are made when calibrating the system, even perfect samples can be discarded as invalid targets.

While vision is one of the newest tools in robotic operation, one must keep in mind that not every tool is the best choice for every job. The intent for this research was to determine what types of project robotic vision is well suited to, and for what projects should vision be ruled out. The process that was chosen to evaluate the system is in fact a poor system for using vision, as mechanical sorting would prove more efficient and far more cost effective. The purpose, however, was not to make a better coin sorter, but to create a simplified version of a bin picking scenario.

3. Methods

The equipment used for this research project was provided in the robotics lab at Eastern Washington University. The robot used was a Fanuc M-3iA Delta, equipped with a Sony XC-56 Progressive Scan Camera and Fanuc's "iRVision" software. The physical system incorporated a conveyor belt, and separate containers for placing each demonstration of the sorted coins. These containers are bottomless, so the coins immediately fall back down onto the belt, which simulates an infinite supply of coins to sort.

American currency is incredibly varied in design; there are over 50 designs for the reverse side of a quarter alone. Teaching the vision to recognize each of these would have been incredibly time consuming, as well as cumbersome for the vision system to filter through for comparison each time it found a coin. Additionally, coins end up with different amounts of physical wear (see Figure 1); some older coins may be rubbed nearly smooth. All these factors combined would present great difficulty in identifying the coins by their designs. In order to bypass these issues, the system was taught to recognize the coins by perimeter size alone. When teaching the parts, the centers were simply masked out, leaving only the circumference as an identifying feature (see Figure 2 - red indicates "ignored" areas in the image).

Figure 1: New and worn coins.

Figure 2: Ignoring designs.

This effectively bypassed the design variation problem, while simultaneously introducing a different obstacle; if a small coin happened to be centered on top of a larger coin (see Figure 3), the system would occasionally recognize the perimeter of the larger coin and ignore the smaller coin. The robot would attempt to pick the large coin, but actually pick the small coin, then depositing this coin in the incorrect bin. It was found that if the system searched for coins in a particular order (smallest to largest), this was no longer a problem.

Figure 3: Concentrically Stacked Coins.

Another downfall to the vision system, in regards to coins, was the lighting problem. Lighting is one of the most important aspects of any vision system: too much causes washout, while too little causes shadows and prevents edge detection. Coins have an additional complication in that some are old and dull, while others are bright and shiny. This may be less of an issue in a production facility, where most things would be uniformly new, however part variation will always occur. Adjusting the lighting and vision parameters alone met the goal of 95% accuracy, but the occasional anomaly prevented the system from performing any better. If a mass of new, shiny coins presented in a tight knot, (see Figure 4) the vision would not recognize any single coin, but only an indistinct, undefined shape. Adjusting the lighting and settings to accommodate this situation caused more problems elsewhere. It was eventually determined that a mechanical solution was required to eliminate this problem. One option was to allow the mass of coins to pass undetected, and then cycle them back into the flow of coins, where they would be broken up. To avoid building this complicated apparatus, it was eventually decided that a simpler method could be used. After the system had determined there were no more valid coins to pick up, the suction gripper would pick a small sweeping device, which was designed to gently agitate the vision area. If any coins were massed together, the sweep would disperse them, allowing the vision to identify the individual coins.

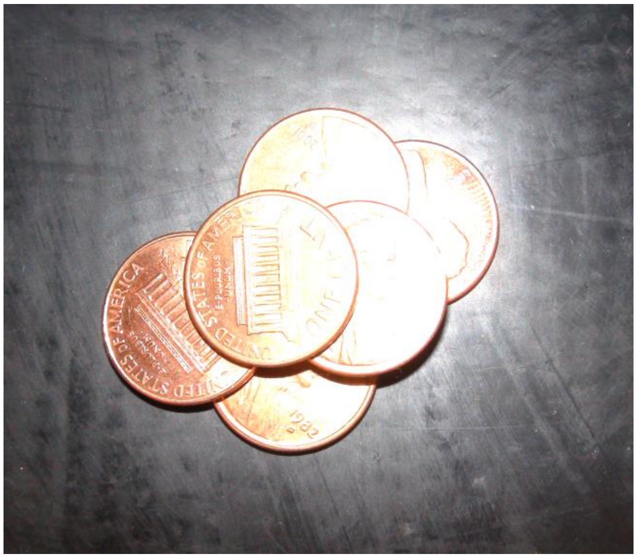

Figure 4: An indistinct mass of shiny coins.

4. Results

It was found that given a single type of object, where all the system has to do is locate and determine orientation, the location speed is incredibly high. There is a barely perceivable pause when the robot halts motion to take the picture, process the information, calculate offset, and resume motion. Adding a second item, the resulting lag time is noticeable, but minimal. With each additional object, the processing time is increased dramatically. Given six objects, and requiring a specific hierarchy for picking, the response times can be measured individually. The coin response times are listed in Table 1. Note that increasing the robot's operating speed does not change these times, as the speed only controls the physical motion, and not the processing ability. The response times given here are given for the specific pick order, as mandated by our system. That is, with priority given to the dime, then the penny, (from smallest to largest).

Table 1: Average delay times for coin identification.

Coin, Priority	Dime (1)	Penny (2)	Nickel (3)	Quarter (4)	50 Cent (5)	Dollar (6)
Delay (seconds)	0.5	1.75	3	4.5	5.25	6.25

The system was then ran the system for half an hour, placing the coins into bins, to record the overall rate of operation, and to determine the time for the physical motion. Using a team member's personal jar of unsorted coins from home, running at 70% speed, the following data was collected. (Note that no dollars or half dollars were present).

Table 2: Program run data.

30 Minutes Run Time	Dime	Penny	Nickel	Quarter	Total
Total Coins Picked	90	200	43	32	365
Amount Incorrect	14	0	2	0	16
Total Time Delay (sec)	45	350	129	144	668

From the data in Table 2, the reliability of the system can be determined, as well as the time spent on physical motion. For reliability, the system resulted in 95.62%, using equation (1).

$$\frac{Total\ coins\ picked - Amount\ incorrect}{Total\ coins\ picked} * 100 = Reliability\ percentage \tag{1}$$

$$\frac{1800 seconds - 668 seconds}{365\ Total\ coins\ picked} = 3.1\,\frac{sec}{coin} \tag{2}$$

This value is a little lower than shorter trials had previously indicated.
To determine the average time for physically moving the coins, one can take the 30 minutes of total run time (1800 seconds), subtract the total time delay for identifying the coins, and divide the remainder by the total amount of coins picked, see equation (2). Therefore the amount of time spent identifying coins is 11.13 minutes, while the actual time for physically moving the coins was 18.87 minutes.

In addition to recording the measurements in tables 1 and 2, the optimized camera settings for the coin setup were also recorded. Understanding these settings allows users to apply them more deliberately in future vision projects. Listed in Table 3 are the settings for the dime, other coins vary only slightly.

Table 3: Vision settings for dimes.

Exposure Time	Score Threshold	Contrast Threshold	Area Overlap
33.33 ms	87%	30	75%

Using these settings, our system could perform at approximately 95% reliability. Adding the mechanical compliance to break up any masses of coins pushed that number up to 98%.

5. Discussion

From the data presented in Table 2, we calculated reliability for the system. This value is considerably lower than when we evaluated the system originally. This is likely due to a higher volume of coins present on the conveyor. The majority of the errors came from picking dimes, which are the smallest coins in our system. It became obvious that the area overlap setting needed to be adjusted higher for dimes, as errors continued to accumulate. If the dime has any other coin overlapping it even slightly, the suction cup will pinch this second coin as it grabs the dime. Both coins are then deposited in the dime bin. Of the 14 errors in the dime bin, 10 of these were due to this double pick issue. If these had been omitted, reliability would have been 98.36%, which matches the previous estimations.

The final settings of some key vision parameters were recorded and listed in Table 3; these criteria will now be defined. The following are brief explanations of what each of these settings did, and how it affected the system.

Figure 5: White out due to overexposure.

5.1 Exposure Time:

This is the amount of time that the camera takes the picture. If this is too low, the image is dark, with shadows. Lengthening the exposure time makes the image lighter, but too long causes the image to blur, or experience "white-out". An example of this, using images from actual operation, can be seen in Figure 5. This parameter is one of the most important, as it sets the foundation for the rest.

5.2 Score Threshold:

This is the requirement that dictates how similar a part must be to the stored image to be considered valid. Setting this value too low allows the computer to see any coin as any other coin. If this value is too high, the computer will fail to recognize anything but a perfect sample. Depending on the values used, the penny shown in Figure 6 may or may not register as a valid penny. If the purpose for using vision is for quality control or inspection, this would be the key parameter to adjust. For this project, the score threshold parameter was set at 87% to allow a small amount of part variation.

Figure 6: A flawed penny.

5.3 Contrast Threshold:

This is the setting that adjusts the camera sensitivity. Lower values are actually higher sensitivity, so a value of 30 represents a relatively high sensitivity. This setting allows the camera to pick up smaller changes in contrast, which allows it to recognize edges easier. This helped the system to avoid the majority of grouped coins camouflaging with each other, but not entirely. Turning this value too low causes a lot of "noise" in the image, which begins to produce negative effects after a certain point. This is shown in Figure 7, with the dimes on the left simulating a high contrast threshold (70-90), and the dimes on the right simulating a low value (10-30). The upper images represent how the human eye would interpret the images, with the lower half showing how the computer sees it. The computer only registers what is highlighted in green. Notice how the image on the left doesn't register much, perhaps not enough to make proper identification, while the dime on the right has a lot of clutter. Neither is ideal, as valid targets will be ignored in either case. A proper middle ground must be found for reliable processing. For this system, since the centers were masked out (ignored) anyway, the setting could be very low (approximately 30 for all coins), as the vision was only looking for the edges.

Figure 7: Contrast, as seen by humans (top) vs. the computer (bottom).

5.4 Area Overlap:

This controls how much of a valid target must be exposed to allow identification. If a dime is half covered by any other coin, the suction cup would be unable to pick the dime alone. For this reason, area overlap had to be set relatively high, ignoring coins that were partially covered. This setting would be less important with larger objects, as a typical end effector may still be able to securely grip an object that was half covered. With the small coins, however, the system could only allow about 25% coverage. For example, the nickel shown in Figure 8 would be ignored, as it is covered just over our allowable amount.

Figure 8: Overlapping coins.

6. Conclusion

In the end, even with an excellent achieved accuracy of 98%, with the robot operating at 90%, this project is a perfect example of one where vision is not needed. From a pure efficiency standpoint, a mechanical sorting apparatus would beat the robot every time. It was assumed that this would be the case from the beginning, and the decision to sort coins was not made in hopes of reinventing the wheel, but rather as a simplified model for bin picking. An analysis of the vision system's ability suggests that it is ideal for use where only a few different parts are presented, and order constraints could be removed. In situations where many different items are presented all mixed together, and mechanical sorting is not feasible, one may have to simply accept the longer identification time.

Robotic vision is a great choice for quality control applications, where the parameters can be set high enough to discard all but the best samples of an item. Based on the issues encountered prior to the decision to use only the coin perimeter for identification, it is clear that this system would be capable of sorting out only the best quality coins. If the system were taught the front and back of a dime, for example, the parameters could easily be set high enough that only the best dimes would be identified as valid targets.

In the end, one must evaluate the task at hand, combined with the cost of each option. While robotic vision is a valid option for many situations, it may still be economically beneficial to stay low tech. Given that the system used for this project costs roughly *$70,000*[1](robot, vision software, and camera) with the camera itself accounting for $500, and the software approximately $5,000, vision is expensive. A lot of sensors can be bought and implemented for the price of the camera alone. Then considering that the actual process of coin sorting could be done by a purely mechanical system for under $100, this cost is astronomical. Until the cost of vision systems comes further down, the question one must ask is not "can vision do it?", but rather "is vision needed?"

7. . Acknowledgements

The authors would like to express their appreciation to Dr. Donald Richter, and Eastern Washington University for use of the Robotics lab, the Delta robot, and all the peripheral equipment that was necessary to implement this project.

8. References

1. Interview with Dr. Richter, December 2, 2015.
2. Tanya M. Anandan, “Robotic BinPicking – The Holy Grail in Sight,” Robotic Industries Associtation, http://www.robotics.org/content-detail.cfm/Industrial-Robotics-Industry-Insights/Robotic-Bin-Picking-The-Holy-Grail-in-Sight/content_id/6002
3. Ales Pochyly and Tomas Kubela, “Robotic Vision for Bin-Picking Applications of Various Objects,” ResearchGate, https://www.researchgate.net/publication/224232508_Robotic_Vision_for_Bin-Picking_Applications_of_Various_Objects
4. James Carroll, “3D vision system assists in robotic bin picking,” vision-systems, http://www.vision-systems.com/articles/2014/08/3d-vision-system-assists-in-robotic-bin-picking.html