FINAL PROJECT REPORT

Project Title: Robot scout for tree fruit

PI: Tony Koselka
Organization: Vision Robotics Corporation
Telephone: 858-523-0857
Email: tkoselka@visionrobotics.com
Address: 11722 Sorrento Valley Road
Address 2: Suite H
City: San Diego
State/Zip: CA 92121

Cooperators: Sanjiv Singh, Carnegie Mellon University

Other funding sources

Agency Name: USDA
Amount awarded: $250,000
Notes: Part of the SCRI, CASC project

Total Project Funding: $200,000

Budget History:

Item	2009
Salaries	$117,985
Benefits	$45,000
Wages	
Benefits	
Equipment	
Supplies	
Travel	$19,061
Materials and CMU Sub-contract	$17,661
Miscellaneous	$293
Total	**$200,000**

OBJECTIVES

2009 was the first year of a three-year project to develop a Scout system for crop load estimation, with the goals of improving the crop load estimates and creating a market-ready system; the project builds upon earlier work where VRC demonstrated a proof of concept Scout, detecting and sizing apples in production orchards. Vision Robotics' 2009 goals were to advance the project in three directions:

1. Refine and improve the detection performance
2. Build the next generation Scout prototype (hardware)
3. Integrate with other systems under development as part of the Specialty Crop Research Initiative's (SCRI) Comprehensive Automation of Specialty Crops (CASC) project.

The task of improving the detection has many parts, the specifics of which were not all known at the beginning of the project. However, prior work had identified the need to enhance the lighting system to enable robust operation regardless of the environmental conditions. Thus, a goal in 2009 was to introduce enhancements to both the hardware and the software to combat difficult lighting situations. Additionally, the software was to be refined to, in general, improve the detection algorithms.

The main 2009 objective for the next generation Scout mechanical prototype was to design and build a system which was rugged and robust like true farming equipment, and would be towed by typical farm machines. This includes a start at weatherization. The next generation Scout should:

1. Operate at speeds up to approximately 1 mphFunction in all lighting conditions Mechanically reflect the production design, with the capability of two-sided operation; however only a single mast mounted on one side would be used in 2009.

The main goal for integration with other CASC projects in 2009 was to couple the Scout with Carnegie Mellon University's Autonomous Prime Mover (APM). The APM would serve as the towing vehicle for the Scout and would issue basic commands to the Scout, directing the start and stop of scanning. Additionally, the APM would provide GPS information to the Scout so that GPS referenced data could be produced for inclusion in a GIS database.

This project will enable the creation of a pre-production Scout in 2010 – 2011. The Scout will provide valuable data for use in precision farming and is a significant step towards mechanization.

SIGNIFICANT FINDINGS

During the course of the project, Vision Robotics made substantial progress with respect to the three main 2009 goals and thus towards its ultimate goal of a production Scout. More specifically, at the conclusion of this year's project:

- The crop estimation performance has improved through the use both software and hardware refinements, in particular through the use of flash units
    - In the Jazz apple scan, computer derived crop load estimates had an error of less than 2%
    - Nearly 100% of the apples were visible to the human eye in images captured by the Scout cameras
- A robust, full Scout prototype has been constructed
- Basic integration with the Carnegie Mellon University's Autonomous Prime Mover has been achieved.
- All aspects of a commercial apple Scout are now proven

RESULTS & DISCUSSION

Improved Estimation Performance

To achieve the 2009 goal of improving performance of this system's crop load estimation, both the Scout hardware and software underwent modifications. Experiments were then conducted to test the effectiveness of the enhancements and evaluate the system as a whole. This section describes the refinements and the setup used to evaluate the system. It then presents and analyzes performance results. Finally, some comments on processing times are presented.

Hardware Refinements

One class of modifications to the Scout hardware to improve estimation performance was introduced to improve the camera and lighting systems. First, a new generation of cameras was developed to increase image quality and facilitate the use of different contexts, where a context is a set of camera settings which can be dynamically controlled by the Scout software; the system utilizes multiple contexts to adapt to varying lighting conditions. Figure 1 shows results of preliminary testing done to validate the potential of the use of multiple contexts, showing the difference in visibility that can be achieved.

Figure 1: Two pairs of pictures taken at high and low exposure levels.

Next, a visor consisting of a flat plate with rectangular viewing holes was placed in front of each camera to shield the camera lenses from oblique and direct sunlight, which can cause flaring. Figure 2 shows the results of preliminary testing done to validate the use of the visor, showing the achieved reduction of flaring. When coupled with the camera mast's pivoting capability, the amount of time in which the sun directly strikes the camera lenses can be significantly reduced.

Figure 2: Effect of flaring on image quality.

The final and key enhancement to the Scout hardware was the introduction of flash units to provide active illumination of the fruit to be counted. A flash normalizes the lighting throughout an entire image and provides a consistent lighting situation that is largely independent of the ambient light. Figure 3 gives an example of the difference between a non-flash and a flash image. As seen in the figure, the use of a flash when imaging green apples makes the apples stand out significantly against the remainder of the tree. It is clear visually that this result has major impact in improving the performance of crop load estimation for green apples.

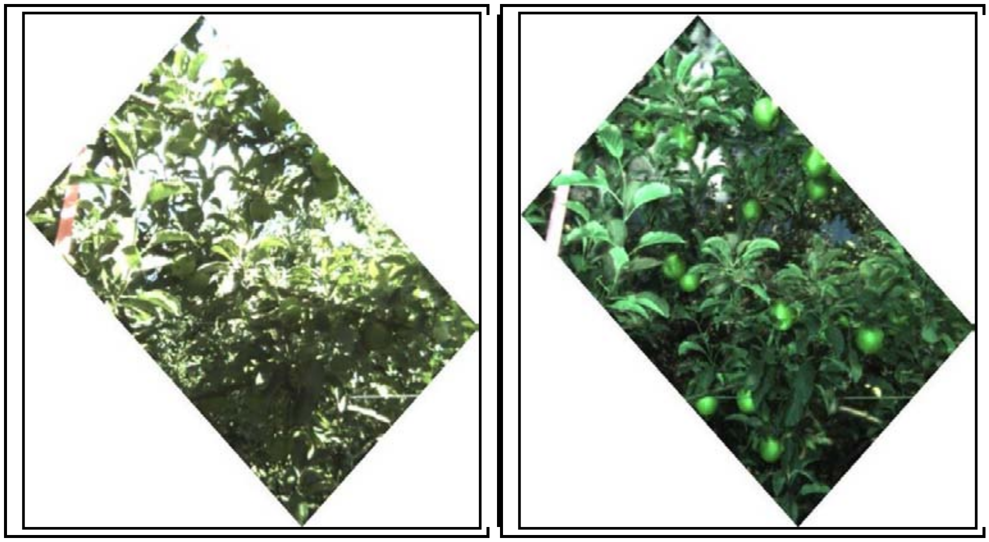

Figure 3; Similar pictures taken without and with a flash.

Software Refinements

To estimate a crop load, the Scout software analyzes the images captured by the mast cameras, searching for apples contained within. This detection consists of two main software components, which are referred to as the back and and the front end of the detection algorithm. The front end considers images on an individual basis and classifies regions which could potentially be parts of apples. Combining information from left and right stereo image pairs, it passes a list of candidate apples to the back end portion of the software. The back end portion tracks these candidate apples between frames and across cameras, and ultimately decides which candidate apples are actual apples. Both the front and back ends of the software were improved during 2009. For the case of green apples, the introduction of the flash units allowed the front end processing to use the same approach for both red and green apple classification, as opposed to an alternate, slower method that was used in past years for green apples.

Setup of Evaluation

To determine the performance of the Scout system, a field test in each of July and September was conducted in central Washington. In each case, the Scout primarily scanned two, one hundred foot sections of rows for each two varieties, Fiji and Red Delicious in July and Jazz and Granny Smith in September. Figure 4 shows representative pictures of the trees scanned during the second field tests, which were conducted at the Allan Brothers' orchard in Othello; the first tests were conducted at the Valley Fruit Orchards near Royal City. The Jazz sections used vertical trellises and contained relatively sparsely growing apples in a thin canopy. Conversely, the Granny Smith sections used the angled-V configuration and contained fruit which were growing in tight clusters with a dense canopy.

Figure 4: Trees used for field tests.

To enable a concrete analysis, a team from the WTFRC collected data from 100' sections in each of the four rows. Each 100' section was divided into segments, and the number and size of apples contained within each was recorded. Table 1 gives the collected yield information.

Table 1: Hand collected yields for the four 100' test sections.

Row Description	**Number of Apples**
Jazz – Row 1	393
Jazz – Row 2	678
Granny Smith – Row 1	1748
Granny Smith – Row 2	2209

Evaluation Results and Analysis

The Scout scanned each of the four test sections multiple times in order to experience different lighting conditions as well as examine the effects of various system configuration parameters. During each scan, the scout simultaneously (between cameras) and continuously (in time) captured images with eight pairs of stereo-cameras (16 cameras total) at approximately eight frames per second. Up to five different camera contexts were employed, some of which included the synchronized flash. Upon further analysis of the data, it was determined that the uppermost cameras on the mast were tilted too far downwards to properly image the tallest portions of the trees. This issue should be solved by simply tilting these cameras correctly, and thus the comparisons with the hand collected data made in this section exclude this region.

The first aspect of the Scout system which was evaluated was the ability of the cameras to capture images which were deemed to contain enough information for the software to provide crop load estimates. Images from an approximately 15' long portion of one Jazz and one Granny Smith section were examined by hand to determine a count of the number of apples that were human-identifiable in the images captures by the Scout. The hand examination consisted of reviewing pictures from all eight camera pairs and distinguishing unique from duplicate apples between cameras (since the cameras have overlapping fields of view), as well as following apples in successive images captures by individual camera pairs as the Scout moves down the row. Consequently, this evaluation was limited to 15'. Table 2 tabulates the number of apples identified by a human from the images in each of the five approximately 3' sections comprising the 15' portion for the Jazz section, as well as the true hand counts. Table 3 contains the corresponding results for 15' of the Granny Smith section. Averaging over the five sections, it is seen that the Scout system collected images in which a human could identify 91.53% and 98.77% of the actual fruit for the Jazz and Granny Smith sections, respectively. Consequently, with strong software performance, the Scout should be capable of producing accurate crop load estimates.

Table 2: Count of apples identified by a human in captured images and true count for the Jazz section.

	Section 0 - 1	Section 1 - 2	Section 2 - 3	Section 3 - 4	Section 4 - 5
Apples Counted In Images	9	19	11	4	11
True Counts	9	20	13	4	13

Table 3: Count of apples identified by a human in captured images and true count for the Granny Smith section.

	Section 0 - 1	Section 1 - 2	Section 2 - 3	Section 3 - 4	Section 4 - 5
Apples Counted In Images	26	37	36	39	22
True Counts	27	37	36	39	23

The next portion of the Scout system to be evaluated was the front end portion of the estimation software. To do so, 33 images were selected at random from scans of one of the Jazz sections and another set for scans of one of the Granny Smith sections. For each image, a human identified the number of human-visible fruit which were correctly detected by the software and the total number of human-visible apples. A representative image for this analysis is given in Figure 5. Here, squares denote regions of the image which were marked by the front end processing as candidate portions of apples. As can be seen in the image, the front end admits many more regions than are truly apples (since the back end will determine which are true apples based on information across multiple frames) with very few missed fruit. The total numbers of fruit and correctly detected fruit are given in Table 4.

Figure 5: Representative image showing regions identified by the front end processing (squares) and regions remaining after the back end processing (circles).

Table 4: Counts of total and number of fruit correctly detected by the front end for the Jazz section.

Apple Variety	Total Apples	Number Correctly Detected in Front End
Jazz	169	169

The next stage in analyzing the system performance was to determine the number of human-identifiable apples that were correctly identified by the back end portion of the software as well as the number false positives produced by the system, as identified by a human. Due to time considerations, analysis was restricted to the Jazz section, with the same 33 randomly selected images. A representative images is given in Figure 5. Here, solid circles represent apples identified by the back end portion of the software. The total numbers of fruit, correctly detected fruit, and falsely detected fruit in the analyzed images are given in Table 5. Falsely detected fruit are primarily caused by misalignment between camera pairs and the fact that the background dirt is red in color; both issues are easily resolved.

Figure 5: Representative image showing apples identified by the (back end) software.

Table 5: Counts of total, correctly detected, and falsely detected fruit for the Jazz section.

Apple Variety	Total Apples	Number Correctly Detected in Back End	Number Falsely Detected in Back End
Jazz	169	155	27

The final analysis performed on the crop load estimation aspect of the Scout system was to compare the load estimates produced by the software with the true values determined by the WTFRC team. Table 6 gives the load estimate and hand-collected yield for the first five of the approximately 3' wide segments of one of the Jazz test sections, while Table 7 gives the corresponding results for one of the Granny Smith test sections. The average relative error

$$\frac{(estimated\ yield - true\ yield)}{true\ yield}$$

in the crop load estimate over the entire 100' sections segments was 1.77% the Jazz section. For the Granny Smith section, only the first five segments have currently been analyzed, however, the average relatively error in crop load was -14.81%. For the Jazz sections, the estimates were observed to be above and below the true counts per 3' region, suggesting a close match to the true yield distribution. For the Granny Smith sections, the estimates tended to be lower than the true counts, in part due to the software missing individual fruit in the centers of large clusters. These observations help guide the planned improvements to the estimation software. It is believed that, in addition to enhancements to both the front and back ends of the software, the incorporation of a statistical model will improve load estimates, particularly for fruit growing in clumps. The potential value of accurate crop load estimates is illustrated by observing the variation in true yield from row to row as given in Table 1; thus, hand counting a small region of an orchard may not be representative of the crop load as a whole.

Table 6: Crop load estimate and actual crop load for first five 3' regions from the Jazz section.

	Section 0 – 1	Section 1 – 2	Section 2 – 3	Section 3 – 4	Section 4 – 5
Yield Estimate	7	22	10	5	12
True Yield	9	20	13	4	13

Table 7: Crop load estimate and actual crop load for first five 3' regions from the Granny Smith section.

	Section 0 – 1	Section 1 – 2	Section 2 – 3	Section 3 – 4	Section 4 – 5
Yield Estimate	23	29	36	31	19
True Yield	27	37	36	39	23

An evaluation of the sizing ability of the Scout software was also performed. More specifically, a distribution of apple sizes was determined for both the 100' Jazz and Granny Smith sections. The estimated means and standard deviations, in meters, for the Jazz and Granny Smith sections were 0.0303 and 0.0062, and 0.0294 and 0.0043, respectively, while the corresponding values for the true data were 0.0369 and 0.0039, and 0.0370 and 0.0027. These distributions are plotted in Figures 6 and 7 for the Jazz and Granny Smith sections, respectively, along with the distribution of the hand collected data. Comparing the estimated distribution with the true curve illustrates that the sizing estimates appear slightly low. It is believed that, in addition to refinements to the underlying sizing algorithms, the future introduction of statistical models into the sizing estimation will improve the accuracy of the estimates.

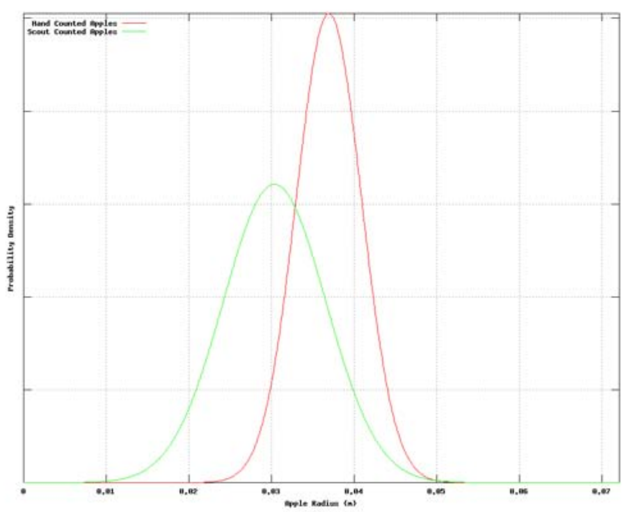

Figure 6: Estimated and true distributions of apple sizes for the Jazz section.

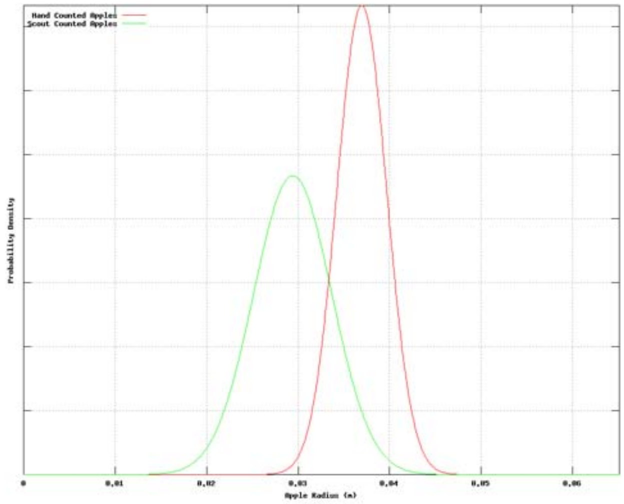

Figure 7: Estimated and true distributions of apple sizes for the Granny Smith section.

Processing Time

Currently, the Scout captures data in real-time to be processed offline. In production, processing will occur in near real-time. Currently, the analysis software is approximately an order of magnitude slower than real-time. This factor, however, represents a substantial performance improvement over past years for the case of green apples, which was made possible by the use of the flash units. Our analysis concludes that through major software optimization near real-time operation should be possible with current computers. However, Intel, AMD, ASUS and other semiconductor manufacturers have recently introduced and announced processors specifically aimed at this type of analysis. For example, the Intel Larrabee processor is approximately five times as powerful for this type of analysis as the computers currently being used. This new class of processors significantly reduces the amount of optimization required to reach production, saving both time and development cost.

Construction of Full Scout Prototype

The 2009 goal of designing and building a new Scout prototype, as shown in Figure 8, was achieved. This section describes some of the attributes incorporated into the new design. A goal in the development was to create a prototype which was a solid evolutionary step between the proof-of-concept prototype and a finished production design. The basic platform will remain the same for the remainder of the project, and new and refined sub-systems will be incorporated as they become available.

The platform is a trailer that is easy to convert from four-wheeled (for scouting) to two-wheeled (for road towing). As requested by the Commission, the Scout is electrically powered with a generator integrated into the design. While VRC only used a single mast in 2009, the new hardware has mountings for masts on both sides for future use. The mast can be moved in-and-out, tilted and pivoted through adjustable steel bracing.

The production Scout will be fully weatherized, and many of the production features were implemented this year. The camera mast is the biggest exception because the adjustability required this year made it impractical to seal, but it will be weatherized in the future. Also, active cooling was specifically not included in this year's project. As such, the computers experienced overheating in the high temperatures during the summer field tests. This issue will be addressed in 2010.

During the July field tests in Royal City, the Scout completed several runs being towed at speeds in excess of 1 mph and as fast as 2.5 mph. The images at 1 mph were of quality sufficient for crop load estimation.

The flash units used in the new prototype were selected for their adjustability and high power intensity, which were desired for design and testing purposes. However, these units are designed for photography and their use on the Scout exceeded their design capabilities with respect to flash rate. Fortunately, their full intensity was not utilized and, moreover, the intensity which proved most effective is within the capability of industrial LED flashes, which are practical for incorporating into a production Scout.

Figure 8: 2009 Scout Prototype.

Integration with CASC Project

As part of the CASC project, the Scout was integrated with the Carnegie Mellon APM. The APM was used exclusively and successfully to tow the Scout during the three day field test in Royal City in July.

Control information, signals to direct the Scout to either start or stop collecting data, was sent from the APM to the Scout using an Ethernet connection. This information allowed the Scout to only capture and analyze data when the APM was moving down a row, and allowed the Scout to be paused and resumed if warranted.

GPS information was also passed from the APM to the Scout to enable accurate land based reference of the data collected for each orchard. This data can also be used to better determine the positions of the Scout's cameras, which is relevant information when detecting fruit. Additionally, GPS data allows the crop load estimates output by the Scout software to be more easily tied to particular locations within an orchard, through the use of real-world coordinate information. Moreover, such information will more easily allow the inclusion of the data output by the Scout in GIS databases. Also as part of the CASC project, CMU is working on such a GIS database, and Scout will output data in a format that will allow statistics to be observed on a 0.5 meter scale.

Summary of Results

The three main goals in 2009 for the robotic Scout for tree fruit of improving estimation performance, constructing a full hardware prototype, and integrating the Scout with other members of the CASC project have all had significant success this past year. Basic integration with the CMU APM was achieved, and the groundwork for further information transfer has been laid. The new Scout prototype has demonstrated robust operation and is a strong step towards a production versions. While numerically the crop load and sizing estimation performance has room for improvement, the refined lighting system, in particular due to the significant benefits of the flash units, along with accompanying software enhancements yielded a vast step forwards from previous years. Moreover, limitations of the current approaches have been identified and a plan for both software and hardware refinements is already in place. Thus, it is believed that the Scout system will achieve its ultimate goal of providing valuable, accurate information that can assist both in areas such as harvest timing and precision farming.

EXECUTIVE SUMMARY

2009 was the first year of a three-year project to develop a crop load estimation system. Previous work demonstrated a proof of concept Scout, detecting and sizing apples in production orchards. This year's goals were to advance the project in three directions:

1. Refine and improve the detection performance including operation in difficult lighting conditions.
2. Build the next generation Scout prototype (hardware).
3. Integrate with other CASC project sub-system.

During the course of the project, Vision Robotics made significant progress towards its ultimate goal of a production Scout including improved apple detection and sizing including operation in all lighting conditions; a robust, full Scout prototype; and integration with the Carnegie Mellon University's Autonomous Prime Mover. In particular:

- The crop estimation performance has improved through the use both software and hardware refinements, in particular through the use of flash units
  - In the Jazz apple scan, computer derived crop load estimates had an error of less than 2%
  - Nearly 100% of the apples were visible to the human eye in images captured by the Scout cameras
- A robust, full Scout prototype has been constructed
- Basic integration with the Carnegie Mellon University's Autonomous Prime Mover has been achieved.
- All aspects of a commercial apple Scout are now proven

With the inclusion of a flash and other modifications, the Scout now uses the same detection algorithms for red and green apples and achieves similar performance. In operation, the "front-end" of the detection system uses classifiers to identify potential regions in an image which may be apples. Each candidate apple is compared against previously identified candidate apples where the goal is to identify any possible apple. The "back-end" analysis prunes this large number of candidate apples to a much smaller subset of actual apples. The more often a candidate apple is identified, the more likely it is an actual apple. By only making final decisions on which candidate apples are actual apples once all data have been analyzed, the system is achieves better performance.

The field tests consisted primarily of multiple data collection runs through 100' sections of rows with hand-counted data. By conducting multiple runs, VRC was able to test a variety of different hardware parameters and to experience different lighting conditions. The hand-counted data enabled accurate comparison of the visibility of apples in the images captures during scanning with ground truth. Images displaying the fruit detection results where then analyzed by hand to determine the percent of fruit seen in each image that were detected. Refinement of the detection algorithms is primarily related to improved apple visibility, a more sophisticated process to track possible apples between multiple images and improved filters for detecting fruit.

The Scout will determine the size distribution for all apples for which it sees enough of the perimeter to accurately estimate the size. Even in orchards with lightly clustered apples, a percentage of the apples will be obscured enough to prevent accurate sizing. However, the distribution should be accurate when applied across the full crop. The actual sizing results will be the size distribution plus the number of fruit sized.