Development of an Autonomous Kiwifruit Harvester

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Industrial Automation

at Massey University, Manawatu,

New Zealand.

Alistair John Scarfe

2012

Abstract

The already intensive labour requirements within the New Zealand kiwifruit industry are increasing. Furthermore, ZESPRI Group Limited is targeting a threefold increase in industry return by 2025 (from approximately $NZ1Billion to $NZ3Billion). Development of autonomous mechanised solutions to assist manual labour is emerging as a strategic necessity.

The objective of this research was to develop a commercially viable autonomous kiwifruit harvester (AKH). The AKH must be capable of operating within variable and complex on-orchard environments to minimise manual labour requirements. Successful completion required development and integration of autonomous:

1. Fruit identification and localisation
2. Custom robotic arms with soft fruit extraction harvesting hands
3. Custom robotic arm for soft fruit handling
4. Transportation platform with navigational sensing and strategies
5. Storage bin collection and drop-off

The AKH has four robotic harvesting arms with hands specifically designed to mimic the human fruit harvesting action. Remotely mounted stereoscopic vision identifies and localises fruit. The fruit locations are mapped into the harvesting arms' coordinate space allowing fruit extraction. The presented system configuration resolves the slow harvest rates experienced by other systems. Practical on-orchard testing identified additional environmental complexities that present the greatest challenge to consistent fruit identification. These are mainly from natural lighting effects.

Stereoscopic machine vision (SMV) was investigated as the primary navigation sensor. However, diverse environmental conditions (lighting and structure appearance) made consistent object detection unreliable. Consequently, a light detection and ranging/SMV combination was used to achieve reliable navigational object detection and fruit storage bin identification.

Practical on-orchard testing and analysis verified AKH operational ability (testing was limited due to a vine killing bacterial (Psa-V) outbreak restricting orchard access):

1. Fruit identification (83.6% of crop) with combined localisation and extraction accuracy of 3.6mm in three-dimensional space
2. More gentle fruit harvesting and handling than humans harvesting
3. Reliable object detection and path planning for navigation. Over the twenty metre scanning range 96% of the in-row objects were correctly classified to reliably determine the drive path
4. Reliable fruit storage bin identification and localisation (98% correct classification)
5. Commercially viable manufacture cost less than $130,000 per unit
6. Although full commercial operation was not achieved, modifications are identified to rectify the limitations

Key system improvements are presented for:

1. High intensity artificial lighting for increased fruit identification rates. Natural sunlight variations affected identification ability, minimising this affect will increase identification rates
2. Alter the storage bin filling arm geometry to permit complete storage bin filling
3. Sensing the robotic arms' position to resolve positioning errors

Acknowledgements

Several individuals and organisations have contributed to and made this research project possible,

1. The author especially acknowledges Dr Rory Flemmer (SEAT, Massey University, Manawatu, New Zealand). Dr Flemmer provided core technology used to development the Autonomous Kiwifruit Harvester. This included:
    a. A patented robotic manipulator design (Flemmer, 2009). His concentric driveshaft design allowed low-cost stepper motor drive and position feedback using hall-effect sensors. It was used on both harvesting and fruit handling manipulators. He also provided skeleton control code which formed the basis for higher-level control strategies and customised picking motion. The provided control code contained methods to:
        i. Efficiently step each motor
        ii. Produce linear robot motion between two points

    Both manipulators kinematics (harvesting and bin filling) and custom movements were developed to increase functionality and efficiency. These were added to the initial control code by the author.
    b. Epipolar geometry for stereoscopic depth perception. This geometry included a method to calibrate and determine both extrinsic and intrinsic camera properties. This method allowed absolute distances to be calculated.
2. Dr Huub Bakker and Associate Professor Donald Bailey (supervisors) as well as Associate Professor Stephen Marsland and Professor Don Cleland, who supported, guided and shared knowledge with the author to complete his studies.
3. Plus-Group Horticulture Ltd funded the prototype robotic harvester from concept. Managing director Steve Saunders provided key industry knowledge and experience pertinent to successful on-orchard operation.
4. ZESPRI Group Ltd provided funding for the completion of dual robotic arm picking and storage bin filling (ZESPRI Project PE1059). ZESPRI Group Ltd also provided financial support and guidance through their Innovation Fellowship.
5. Dick and Mary Earle, who supported the author for three years with their technology scholarship
6. Richard Braithwaite (Stonar Consultants Ltd, NZ) for his patient and dedicated assistance editing the work presented

7. Dr Huub Bakker, Clive Bardell, Andrew Booth, Kyle Peterson, Colin Rose and Mike Lusby were involved with the project during development. Their involvement ranged from CAD modelling, fabrication, machining, wiring and assembly of components and systems. While working on the project they were under the author's supervision. The author directed design, development and management decisions during this process but wishes to acknowledge and thank them for their assistance.

Key Contributions

As a team helped with this project, the author has commented on the team members' contribution throughout this thesis. During development, the author managed and supervised the work of those other contributors. He directed design decisions and development through the process. This team assisted during the first year of development.
Within this team, the author's key contributions are the:

1. Fruit identification system
2. Improvement to the stereo vision system and its calibration
3. Improvement and implementation to the hand-eye coordination system. This includes the calibration device for the robotic arm and locating its position to sub-pixel accuracy with the vision system.
4. Development of the harvesting hands (each generation).
5. Modifications to the initial robotic harvesting arm design. This included: Improvements to the arms kinematic models and development of the custom motion paths; the parallel linkage system to keep the hand vertical, remote mounting of the hands actuators and some structural refinements.
6. Development of the collision detection, avoidance and harvesting scheduling systems.
7. Bin filling robot design and development.
8. Autonomous Transportation Platform. This included: programming the drive system; traction control; custom manoeuvres; chassis and bin lifter design; the initial vision work for navigation; development and implementation of all LIDAR based systems for navigation and bin fining; as well as the integration of the vision and LIDAR systems.
9. A significant amount of general fabrication, machining, construction, wiring, etc.

Contents

List of Figures

List of Tables

Worker harvesting kiwifruit from a T-bar growing system. Photo courtesy of Plus Group Horticulture.

1. Introduction and Commercial Drivers

Horticulture, the practice of plant growth for human advantage, has been recorded for 8000 years. Use of mechanical systems has made horticulture progressively more amenable and productive as humankind moved from digging stick, to mattock, to plough, to steam tractor and combine harvester. The time spent by horticultural labourers (an average seventy-hour workweek) has decreased to forty or fifty hours. Mechanical aids have also eased the heavy manual labour requirement. However, the labourer's work remains repetitive and physically demanding and is yet to experience the full advantages of significant mechanisations. Development of autonomous mechanised solutions is the next technological progression for horticulture. This thesis outlines development of an autonomous harvesting solution for kiwifruit, delivering solutions to the complex task of harvesting fruit on-orchard.

New Zealand currently produces approximately 380,000 tons (1.264 million eighteen bushel bins) of class one kiwifruit from 12,500 canopy hectares. Fruit is harvested from late March to mid-June and requires 2,500 people over fifty full harvesting days. 2011 domestic crop production increased by ten percent over the previous season. In comparison, Chile and Italy produce 280,000 and 480,000 tons respectively (Belrose, 2011). New Zealand fresh fruit and vegetables exports brought in $2.27 billion to the New Zealand economy in the 2008/09 financial year. Horticulture is an important part of our economy and accounts for $1 in every $13 (7.5%) of New Zealand's exports, where kiwifruit is the single highest earner (Aitken and Hewett, 2011)

Timely, efficient and careful harvesting of this fruit is critical to ensure optimum returns. Furthermore, horticultural labour demands are increasing, placing further pressure on harvesting quality fruit. ZESPRI Group Limited is targeting a threefold increase in industry return (from approximately $NZ1Billion to $NZ3Billion by 2025). Autonomous mechanised solutions to supplement manual labour is an emerging strategic necessity. The planned increased industry return is based upon increased plantings, higher yields per hectare (with current varieties), the introduction of licensed new varieties (with high taste and yield attributes), and delivery to market efficiencies. This corresponds with growth strategies identified in many other kiwifruit producing countries (an estimated 100,000Ha of kiwifruit exists worldwide). New Zealand, Italy

and Chile are the major exporters of kiwifruit and account for 50% of the worldwide-planted area and 90% of product exports.

Key issues for all domestic and international kiwifruit producers are: reliance on uncertain and increasingly expensive labour forces; increased productivity; efficiency, improving product quality and consistency; and health, safety and traceability. The kiwifruit industry, like most horticultural sectors, struggles to attract and retain labourers especially during seasonal high demand periods. This forces a temporary seasonal labour force migration to meet time critical on-orchard operations. However, seasonal workers can pose reliability, quality and socioeconomic challenges.

> There is a real concern within the existing labour paradigm that the necessity to drive down labour cost and increasing labour requirements (projection corresponding to industry growth rates) will also erode skills and capability within the industry (Mowat, 2011)

A common commercial objective is the improvement of grower return (ensuring the delivery of sustainable crop volumes for market access) by driving cost reduction through pre and post-harvest efficiency improvements, including labour.

Other robotic applications and markets (not limited to kiwifruit), including multiple arms (capable of pruning, cutting, slicing, pinching, etc.) and non-destructive sensor integration (capable of carrying out field assessments of crop, canopy density and crop quality) could be developed. Further capability demand will emerge as market requirements stimulate greater product ability. ZESPRI KIWISTART® and TASTE® ZESPRI programmes already require individual growers to develop and access specialist skills. The AKH and autonomous transportation platform (ATP, section 7) are desirable for enhancing orchard service, reducing on-orchard costs and delivering quality improvements to growers. This technology also supports competitive advantage.

Mass production of the AKH and on-going technological improvements will further drive down unit manufacture and ultimately the harvesting cost. ZESPRI's long-term 2025 growth strategy targets a tripling of total industry turnover and anticipates major labour shortages. Both Chile and Italy report rising labour costs and uncertain of supply. The presented robotic technology can provide growers and service providers with consistent high quality service capabilities for on-orchard operations, including harvesting.

The proposed business model is based upon developer ownership, unit manufacture and offering fully supported contractual harvesting services. The contract harvest rate of $21.00 per bin demonstrates long-term sustainable commercial and operational advantages to relevant stakeholders. Strategic domestic market entry requires large-scale deployment of commercialised AKH's. To maximise unit use, harvesting units could be shipped to and operate in Italy during the opposing harvest season (commencing in September), doubling total harvesting potential. The second phase envisages deployment of the ATP (section 7), and relevant implement attachments, to Chile for harvesting, pruning and other autonomous robotic activities.

Following the business model, focus shifts to assessing technological feasibility. To achieve autonomous on-orchard operation, most systems will require adaptation from structured and controlled indoor industrial environments into variable on-orchard environments. Examples of on-orchard variability include infrastructure placement and appearance (posts, vines etc.) where these structures' are not linear, or uniform. More importantly, natural lighting and weather conditions may seriously affect object detection. These variables are not often present in industrial type environments using automated systems. Developing operational systems with accurate and reliable interpretation within variable on-orchard environments is the greatest challenge (in comparison to projects demonstrating ability within synthetic environments).

Kiwifruit orchards use a pergola growing system that supports fruit and canes in an overhead roof-type structure. Posts, with near uniform spacing support the overhead structure, dividing the canopy area into rows, and bays (a bay is a row segment between posts). Row widths range from approximately 3.5 to 5.5m with the majority of fruit positioned anywhere from 1.5 to 1.9m above ground level.

Figure 1-1: An aerial photograph of several kiwifruit canopy areas showing some typical orchard layouts and structures – Image extracted from maps.google.co.nz

Rows, collectively known as blocks, are surrounded by sheltering hedges or windbreak cloth to reduce wind flow, creating on-orchard microclimates. Usually rectangular in formation (as shown in Figure 1-1, for operation ease) block shape is dictated by environmental features (e.g. creeks, land contours and orchard land shape) often producing blocks containing short or irregular rows, conforming to useable land area.

Variations in infrastructure are often found where old growing systems (T-bar) have been converted or adjacent blocks have been joined. This can produce irregular row widths, tapering row widths or rows that are not straight. Many orchards show these variations so any practical solution must correctly operate within these constraints. Areas at rows end (called the 'headland' - see Figure 1-1) provide row access and turning space. The headland area also contains the 'dead-men' posts, which anchor longitudinal row cables, supporting the entire pergola canopy structure. These short embedded 'dead-men' pose navigational obstacles.

This thesis presents autonomous robotic solutions developed to achieve autonomous kiwifruit harvesting, in environmentally diverse and variable conditions. On-orchard environmental conditions increase operational complexity when compared with structured, stable and controllable industrial type environments. While developed specifically for kiwifruit, the technological solutions presented could readily be adapted to other discrete fruit crop types. This is not the first discrete crop harvester development (see section 2). However, other developments have not achieved suitable performance (e.g. practical harvesting rates) to justify production or commercial use. The current lack of a commercially viable system indicates the complexity of product development.

A final system is required to replace human pickers. Therefore, the AKH is required to locate fruit, extract individual identified fruit from the vine, move the harvested fruit to a storage container and place those fruit within the storage container. These tasks must be carried out without damaging fruit. In addition, the storage container requires handling. This involves identifying, picking up and dropping off storage containers within the appropriate orchard area(s).

A core research consideration was commercial viability of system operation in diverse and variable on-orchard conditions. Pure laboratory based research often disregards these important but operationally relevant and complex factors. Practical commercial considerations directed many design and delivery outcomes. These included unit cost, robustness and reliability, health and safety and wider system use by designing the ATP (section 7) for other tasks.

Development of the total system is split into operational sub-sections (e.g. harvesting robot arms, bin filling robot, etc.). Product development cycles are conducted for each sub-section to ensure development meets targeted performance criteria (outlined in section 3). Although separate, development of these sub-sections must consider and resolve the interactions between them. After initial laboratory testing, integration of the sub-sections allows complete testing. This testing is carried out in real orchard environments to determine overall ability and viability. If necessary, the product development cycle will repeat to resolve any identified issues.

This thesis covers the product development process for the AKH. It begins with a review of current horticultural robotic systems. This provides a background of what has already been achieved, including performance and development status. The review also provides an overview of the different aspects required to develop a viable autonomous harvesting robot. AKH development began by outlining and establishing the systems requirements. This included mapping the systems operational schematic and defining operational and performance specifications (section 3). Initial development work is presented in section 4. This was developed using the system schematic and specifications from section 3. It supports the rational for certain development decisions and how those decisions progressed the system to methodology used. The implementation process for each operational sections is also discussed. Following this, the technical methodology is presented in the relevant sections. These sections are set out under the following numbers:

5. Fruit Identification and Localisation to Achieve Robotic Extraction: Details the process of stereo machine vision fruit identification, including three-dimensional space localisation. The process of mapping the fruit locations into the robotic arm coordinate space is also presented. This allows the harvesting arms to extract machine visions located fruit.
6. Robotic Manipulation: Covers the developed technology used to create both robotic arm designs. These are the four harvesting arms and the bin-filling arm.
7. Autonomous Transportation Platform: All task-performing systems are attached to this vehicle. Navigation, drive and control strategies are presented in the section.

Following the technical content, the final section provides the results and performance of the combined total system. Furthermore, conclusions relative to its operational performance are made. These conclusions have resulted in recommendations for future development to improve

system performance where limitations were identified. Furthermore, important observations have been made for developing mechanised solutions for horticultural and agricultural applications are included.

Important Note on Testing Limitations:

In November 2010, a highly virulent strain of Pseudomonas syringae pv. actinidiae (Psa-V) was discovered in New Zealand kiwifruit orchards. Psa is a kiwifruit vine killing bacteria that poses a serious industry threat and has resulted in many infected orchards being destroyed to prevent bacterial spread and contamination. Consequently, since November 2010, orchard access throughout the kiwifruit industry has been significantly limited and stringent hygiene protocols introduced where access is granted. As the robotic harvester is still a prototype, many integral systems are not suitably protected to meet the saturation cleaning and decontamination protocols. As a result, on-orchard testing was severely restricted (after November 2010) and this directly impacts on the ability to test the system. This is most evident in section 8, describing the autonomous transportation platform. For additional information, visit the Kiwifruit Vine Health website: www.kvh.org.nz.

Figure 1-2: Kiwifruit pergola growing system. Photograph captured in-row

2. Current Horticultural Harvest Robotics

The reviewed discrete crop harvesters provide background and context for this development. More specific literature will be reviewed in the appropriate sections.

There is considerable worldwide interest in the development of autonomous harvesting solutions. John Deere, currently invests 4-5% of net sales into research and development (including autonomous harvesting), equating to approximately $US2Million per day ($NZ2.5Million) (Lane, 2007). While primary development is focused on automated tractor piloting, their systems could potentially be adapted for autonomous harvesting system transportation, and automated produce handling.

Researchers and developers have considered the automated harvesting problem for several applications and approaches. Historical research into automated outdoor fruit harvesting is presented by Muscato (2005) and Sarig (1993) who highlight the main robotic problems associated with robotic fruit harvesting:

1. Accurately identify and localising fruit on-orchard
2. Grasping identified fruit, for extraction, without inflicting damage
3. On-orchard navigation sufficient to reach all fruiting areas

These areas indicate development focused on functionality, without any serious consideration of the commercial implications (e.g. cost and performance). This review assumes that the objective of a commercially viable robotic system is a necessity.

Over the last three decades, development has moved from the structured industrial environment into the less ordered environments of agriculture and horticulture. The majority of successful automation developments within these disciplines are associated with bulk harvesting on arable land. Here, the harvested crop allows a brute-force approach to extraction and handling. On the other hand, discrete fruit and vegetable crop harvesting requires individual extraction and handling procedures to maintain product quality.

A significant example is given by the use of combined harvesters for grain harvesting. The evolution of mechanised fruit picking from trees was much slower, doubtless, because it is

more complicated. Whereas with herbaceous crops, harvesting consists of cutting the crop and separating the useful part from the rest, with tree crops the fruit has to be separated from the tree without damaging its structure, which represents productive continuity (Muscato, et al., 2005).

Although the relationship (and subsequent technology) between bulk and discrete harvesting is not completely transposable, some areas, like navigation, have similarities. This allows technological sharing into the orchard environment (and vice-versa). The Claas Autopilot was the first commercially available navigational system for piloting a tractor in bulk cropping. In this system, sensors detect the crop edge (only suitable for crops with solid edging like maize) and actuators steer the tractor according to the determined crop row edge (Diekhans, 1999). Following the Class Autopilot, The Demeter System introduced passive crop edge tracking using machine vision (Pilarski, et al., 2002) in conjunction with RTK-GPS (real time kinematic-global positioning system) improving system reliability.

Edan, et al. (2000) notes significant advances automating the harvest of apples, asparagus, cit-rus, cucumbers, grapes, lettuce, tomatoes and melons. However, those advances have not yet produced a commercial harvesting machine. Primarily, the overall systems suffer from large build costs and substandard performance that discourages feasible commercial models.

2.1 An Orange Harvester

The most advanced discrete crop robotic harvester found within current literature is the prototype orange picking robot (Muscato, et al., 2005). Development was not only limited to fruit extraction, but included fruit handling and navigation. These inclusions not only expanded project scope, but also introduced complexity from increased operational interaction (e.g. the

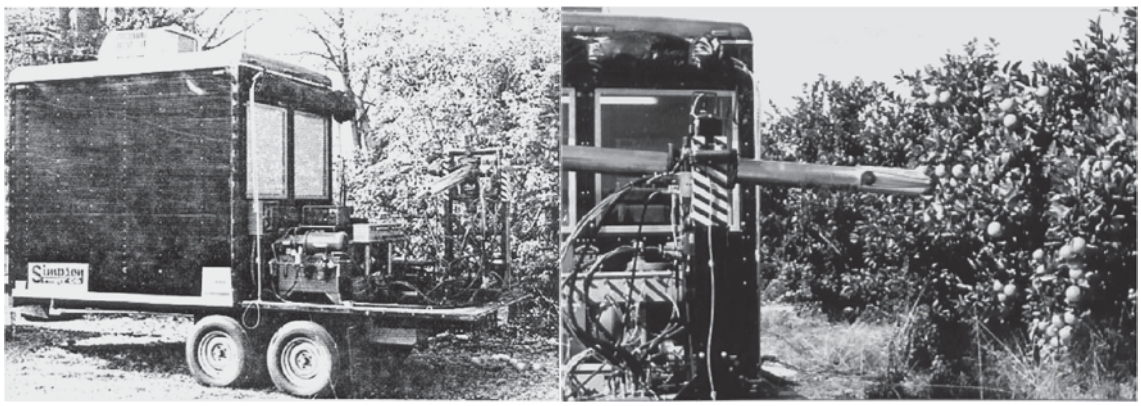


Figure 2-1 The first Prototype Orange Harvester (Muscato, et al., 2005)

picking arm requires fruit-handling capability to work in conjunction with the extraction process). Key project complexities are highlighted:

1. "More expensive design and realisation of the mechanical structures, because they have to be waterproof, powder proof and resistant to anything that is present in a non-protected environment.
2. The adoption of high-performance and outdoor-proof sensors and actuators; and
3. The development of elaborate, flexible control laws in order to cope with a great variety of conditions." (Muscato, et al., 2005)

To assist robotic operation and achieve higher functionality, modification to the growing system were considered.

The evolution of mechanical fruit picking systems would have been more intense if it had been possible to adapt the shape and structure of tree to the requirements of a picking machine (e.g. tree size and arrangement, rearing and pruning systems, etc) (Muscato, et al., 2005).

The first prototype, Figure 2-1, used a hydraulically driven harvesting arm, mounted on a trailer unit, towed behind a tractor. The harvester rate was too slow compared to manual harvesting (Muscato, et al., 2005). Issues with environment and lighting contributed to poor performance, requiring further research focusing on improving three main factors:

1. Robotic arm
2. Hand (for fruit extraction)
3. Visual identification system

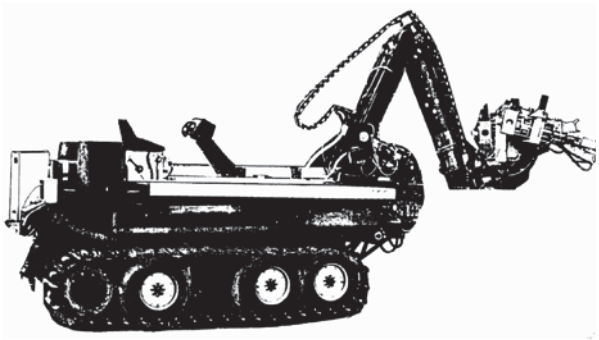


Figure 2-2 The second prototype orange harvester (Reece, et al., 1996)

Attempting to build a faster and more robust system resulted in the construction of a large hydraulic arm capable of harvesting an entire tree (Reece, et al., 1996), Figure 2-2. Although this version has an increased harvest cycle, it was too slow and had fruit handling shortfalls (imposed from the long arm members). It was however an improvement over the first attempt, but failed to meet the required performance criteria.

The third prototype, Figure 2-3, employed a similar tracked chassis as its power base. However, the introduction of an oblique scanning axis allowed two harvesting arms to operate concurrently. This implementation also allowed fruit extraction over the full tree height and the introduction of fruit handling systems.

Figure 2-3 The current orange harvester configuration (Muscato, et al., 2005)

The two thousand kilogram trolley incrementally moves along the row under GPS waypoint guidance (manually driven and positioned at the row's end). Upon reaching the desired harvesting position, the two arms scan for fruit with an integrated machine vision system. Altering the harvest arm's position on the oblique and horizontal axes changes the machine vision viewing vantage point. This allows fruit to be found over the entire tree. Upon fruit identification, the machine vision system maintains the fruit in the centre of its vision (camera in hand) as the arm extends towards the fruit. A Kalman filter enables the control algorithms to predict the distance to the fruit over several images. This allows the arm's control algorithms to adjust traversal speed and acceleration accordingly. When the hand, attached to the end of the harvesting arm, is correctly positioned relative to the identified fruit, the hand extracts the fruit, the harvesting arm retracts and the fruit is placed on a fruit-handling conveyor mounted within the oblique axis. The design supports harvesting flexibility and fruit handling.

Although this was an advance in discrete crop harvesting, the slow picking cycle (8.7 seconds per fruit) reduces commercial potential. This is primarily due to the integrated machine vision system positioned within the hand. This configuration promotes a 'hunting' cycle for fruit finding which requires time-consuming arm movements and high computer processing demands (from continually updating fruit position to orient the harvesting arm with respect to the fruit). It represents a limited solution.

2.2 A Robotic Apple Harvester

Baeton, et al. (2007) aimed to demonstrate the feasibility and functionality, of an autonomous fruit picking machine using existing industrial components (where possible).

Project scope was limited to fruit localisation and extraction without the complexities associated with fruit handling. A high-end Panasonic anthropomorphic arm is used for fruit extraction. An added seventh axis provides dynamic vertical motion, increasing the arm's operational envelope, allowing complete tree harvesting. A custom hand (section 6.2.1), fixed to the end of the arm, holds and extracts fruit without causing damage. Large arm inertia requires hydraulic stabilisation on the transportation tractor. This required detailed setup at every harvest location.

The hollow hand housed an eye-in-hand camera for fruit identification (as used in the orange harvester). Variable on-orchard lighting conditions adversely affected fruit identification. A tarpaulin was used to shroud the tree and reduce natural lighting variation (Figure 2-4). These additional structures required setup at every harvesting location, decreasing operational efficiency.

The fruit harvest cycle, described as a hierarchical structure, contains three levels. The first level actively stabilises the unit and shrouds the tree with the tarpaulin to normalise the illumination level. Secondly, the robotic arm moves through forty different viewing locations, capturing an image at each location. The captured images are used to construct a map of fruit on the tree. Finally, a schedule of identified apples provides the harvesting order. The harvesting cycle consists of:

1. An estimate of the fruit's position from the captured images.
2. Moving the arm to an offset position from the fruit's perceived position. The hand, with camera inside, moves vertically and horizontally until the apple is in the centre of the image.
3. The hand is moved towards the apple, altering its position to keep the apple in the centre of the image (corresponding to the hand's centre). By comparing several consecutive images, an estimation of the apple's distance/position is determined (this apparently used a Kalman Filter implementation).
4. When the apple is within the hand's grasping range, vacuum is applied and pressure indicates object attachment.

5. The apple is extracted by a rotating and tilting motion (Figure 2-4).

This system is slow, with minimal on-orchard operational flexibility. Furthermore, the necessity to stabilise the platform, shroud the tree and capture forty images from several positions increases setup time before fruit extraction. Cameras are a relatively low-cost sensor. System efficiency could be improved by positioning another camera mounted further away from the tree (as achieved by Edan, at al. (2000)) or even several cameras to gain different perspectives, allowing occluded fruit to be seen.

Field trials demonstrated system performance. However, with extraction requiring 8-10 seconds per fruit, the system is slow and consequentially commercially unviable. There is no indication if the stated harvesting rate includes setup. The use of a speciality high-cost industrial robotic arm and controller, a tractor locomotion unit (which could be otherwise used in the orchard during harvest) and labour to oversee and pilot the system, further reduce commercial viability in the presented configuration.

Figure 2-4 Robotic Apple Harvester (Baeten, et al., 2007)

2.3 The Robotic Melon Harvester

A prototype robotic melon harvester development capable of identifying and grasping large melons from the ground is presented by Edan, at al. (2000). Harvesting is conducted from a trailer behind a tow tractor (see Figure 2-5). Research focus is directed at development of:

1. Object detection in a natural environment
2. The design of an effective hand for fruit grasping and extraction
3. Arm design and motion control strategy

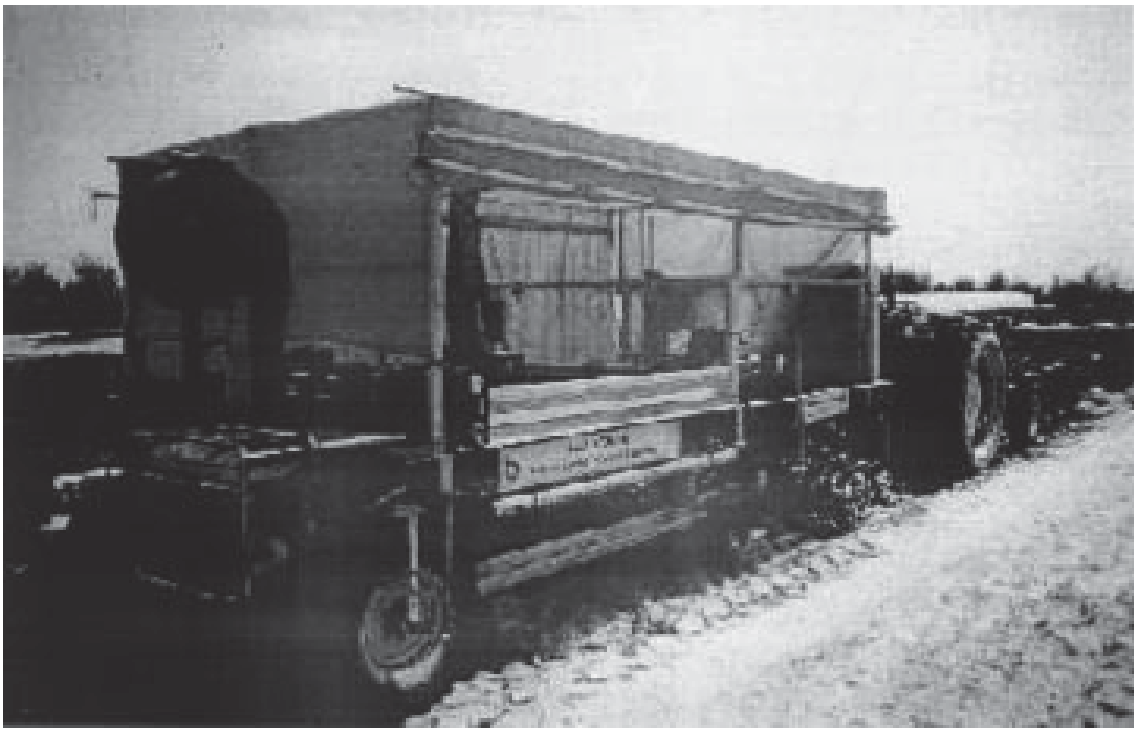


Figure 2-5 Robotic melon harvester (Edan, et al., 2000)

Melons are detected uses two black and white CCD cameras. One, mounted in the arm's hand (eye-in-hand configuration) assists in positioning the arm to the targeted fruit. The second camera, mounted in a higher viewing position, enables fruit scanning over the entire harvest area. This configuration provides approximate fruit position coordinates, where the arm's camera tracks the fruit during the harvesting cycle. The dual camera approach increases efficiency over other harvesting systems (Edan, et al., 2000), eliminating the fruit 'hunting' routine of single camera-in-hand configurations.

The harvesting cycle is performed as follows:

1. The Cartesian arm positions itself vertically over the fruits approximate position, as determined by the far camera
2. Using the camera-in-hand, the target fruit is located and the arm, with motion controlled by the fruits image location, proceeds towards the targeted fruit
3. Proximity sensors determine the hand's height from the ground
4. Upon reaching a predefined height from the ground, the fruit is clasped in the hand, lifted (pulling on the stem), where two knives cut the stem. A personal computer handles all control and sensing tasks

To increase fruit identification efficiency the system uses a plastic enclosure to shroud sunlight. Strong artificial lighting provides consistent lighting conditions within the shrouded trailer. Alternating air jets shift foliage, allowing occluded fruit to be detected. Both methods to aid fruit identification are only suitable in their specific operational environment. Sunlight shrouding is impractical in a kiwifruit orchard environment, due to the roofed canopy structure. Strong air currents (to shift occluding leaves) are likely to move the relatively light hanging fruit, increasing positioning complexity and potentially damaging fruit.

Harvesting arm movement is optimised using a superposition algorithm producing a smoother trajectory (bell-shaped fifth-order polynomial) as opposed to a trapezoidal velocity profile (Edan, et al., 2000). Harvest order was determined by a 'travelling salesman algorithm'. The cost function (for 'travelling salesman' computation, normally defined as the distance between individual fruits), includes the transition from the melon position to handling conveyor and back to the next fruit. The additional 'travelling salesman' scheduling algorithm reportedly reduced the total harvest cycle time by 49%, compared to harvesting the fruit in the order identified by machine vision.

Five thousand fruit were harvested during field-testing where:

1. 85.70% of the crop was successfully harvested (5.85% required a second attempt)
2. 7.02% of the fruit received damage
3. 19.88% of attempts were a result of incorrectly identified fruit and the harvesting attempt was aborted
4. A harvesting time of 15 seconds per fruit was achieved

For commercialisation, the authors concluded that:

1. Fruit detection rates must increase
2. Harvesting cycle times must decrease
3. The system must determine ripeness before picking

2.4 An Autonomous Cucumber Picking Robot

Van Henten's, et al. (2002, 2003) autonomous cucumber harvesting robot operates in an ordered, structured greenhouse environment. Here, controlled lighting and structures reduce system complexity (as outlined by Muscato, et al. (2005) in section 2.1). The need for specialty

outdoor rated sensors and interpretation of changing environmental conditions is removed. A rail system, setup throughout the greenhouse, reduces navigation complexity by defining structured locomotion paths.

Changes to traditional cucumber growing systems, through altered pruning and tying methods, presents fruit in a way that simplifies fruit detection and extraction. Manual removal of foliage and clustered fruit prevents occlusion and fruit grasping difficulties. Manual repositioning of fruit located behind vines was necessary for identification. Controlling the environment enables more capable systems to be developed rather than attempting to automate a standard growing system (identified by Muscato, et al. (2005)). Growing system modifications, although directly beneficial to robotic harvesting, can reduce automation benefits. Additional input to support automated harvest (structures, labour, etc.) offsets system benefits. It was noted that modifying the growing system increased fruit size and weight justifying additional inputs and the adaption of the robotic harvesting system. "Unfortunately, the higher labour demand required for crop maintenance has prevented a wide scale introduction of this cultivation system in Dutch horticultural practice until now" (Van Henten, et al., 2003).

The robotic cucumber harvester averaged harvesting 74.4% of the crop. The majority of failures arose from inaccurate hand positioning. The accuracy of hand positioning decreased for fruit further from the image centre. This is possibly due to incomplete calibration of the stereo vision system, and an inadequate camera model to compensate for the cameras' wide-angle lens tangential and radial distortions. To resolve these hand positioning inaccuracies, several picking attempts (to a maximum of three attempts), are performed from different locations.

Stereo machine vision algorithms determine fruit maturity and position. Fresh weight is the main criterion of cucumber ripeness (maturity). A geometric cucumber volume model achieved a 97% correlation to fruit weight, allowing machine vision to estimate ripeness and determine if a harvest cycle should be attempted. Upon determining fruit ripeness and position (imaging takes 10.8s, plus an additional 9.5s of analysis), an arm path is calculated (taking 6.9s), where arm traversal to the fruit occurs in two stages:

1. The arm moves the hand to within 100mm of the estimated fruit position, requiring 5.6s to complete
2. A slow motion approach to the cucumber commences, taking 7.1s

The hand secures and cuts the fruit from the vine, depositing the fruit in a storage container before returning to the home position (25.3s). This results in an average picking cycle time of 65.2s. Once the robotic harvester has attempted to remove all fruit in the viewing area, the chassis moves along the rails to a new viewing position. The 0.33 metre move provides three overlapping images of the vines allowing three picking attempts per fruit.

Although a single harvest cycle requires 65.2s per cucumber, this does not account for multiple harvest attempts, nor does it provide an average time for all functional actions (including fruit handling). Under normal harvest conditions, the average harvest rate is 124s per cucumber (not allowing for multiple harvest attempts). This extends to 142s per cucumber when individual cucumber visual positioning over consecutive images was included.

Future research focuses on:

1. Improving the success rate
2. Faster hardware and software for image processing and motion planning
3. Reducing the motion time of the arm

More efficient cycle times are necessary for commercialisation. However, detailed manual crop preparation negate many automation benefits. A system that operates within the presented environment (with occlusions, clustering etc.) or can itself prepare the environment pre-harvest would provide greater overall system and unit efficiency.

Figure 2-6 Cucumber harvesting robot (Van Henten, et al., 2003)

2.5 Strawberry Harvesting Robot

A robotic strawberry harvester was found later into the AKH's development. This system has been developed in the Unites States by Robotic Harvesting LLC.

Their harvester uses a stereo vision system to locate and differentiate between ripe and unripe fruit. A single robot arm, with customised hand, uses the vision system's fruit location to harvest the strawberry. Indoor testing showed that the arm, which is mounted under the platform, can pick and place a strawberry in two seconds. However, the three-fingered gripper has problems with the berries slipping out during extraction. Furthermore, the field trial videos show significantly slower harvest rate. This could be attributed to slow fruit identification. They noted that they are looking to progress to multiple synchronous arms in the future to increase performance.

Although they appear to have made improvements to the harvest rate but using stereo vision, fruit occlusion, handling and identification rates appear to be slowing performance. These issues require solving for commercial operation.

Figure 2-7: Robotic Harvesting LLC strawberry harvesting robot. Image from (Rose, 2010).

2.6 Conclusion

The most restricting factors to the commercial viability of the reviewed harvesting systems are slow fruit identification and extraction rates. Edan and Miles (1991) have partially resolved these issue by employing two cameras. One camera observes the overall scene and the second (located in the hand) refines the identified fruit position during harvest motion. Other reviewed systems rely on a single camera located within the hand. These systems often require finding the fruit from several vantage points to determine their locations. A method for accurate and fast fruit identification and localisation will be required to raise the AKH to a commercially acceptable level.

In addition to slow fruit identification and extraction rates, complications from natural on-orchard lighting variations were encountered and adversely affected fruit identification. The process of shrouding the environment to remove much of the natural light, and therefore variability, is not practically possible within a kiwifruit orchard due to the canopy 'roof' structure. The developed kiwifruit identification system will have to deal with the environmental conditions without shrouding.

Although off-the-shelf industrial arms were often used, custom fruit harvesting hands were always required and tailored to specific fruit extraction requirements. A hand capable of mimicking the human harvesting action will be required to ensure correct harvesting techniques are maintained.

3. System Overview and Specifications

The review of current harvest robotic systems (section 2) identified key developmental areas and specific approach limitations. Work by Sarig (1993) and Muscato, et al. (2005) assisted in defining necessary functional sections for the AKH (see Figure 3-1). The main sections or systems required are:

1. Fruit identification and localisation, to allow robotic extraction (section 5)
2. Fruit extraction, without damaging the fruit (section 6.2)
3. Fruit handling and storage, without damaging the fruit (section 6.3)
4. On-orchard transportation (section 7) of the robotic arms and other systems

These cooperative systems replace the current human harvesting activity. Human harvesting starts by extracting fruit from the vine. These fruit are then placed into harvest bags carried on the picker's front. When the bag is full, it is emptied into a storage bin on a nearby tractor/trailer unit parked under the vines. These trailers generally hold three bins. Once full, the tractor delivers the bins to a loading area and swaps them for empty bins before returning to the picking gang.

To achieve sufficient commercial payback the AKH is required to harvest three standard bins of Hayward kiwifruit per hour, each holding eighteen bushels, or approximately 3600 fruit, over a twenty hour day (equivalent to sixty bins per day). It will operate at, or above, the industry quality standards for fruit harvest and handling (outlined later in this section). A twenty-hour daily harvest period accommodates natural environmental conditions (e.g. dew, as wet fruit are not allowed to be harvested), unit relocation and maintenance. On average, a single human picker harvests eight to nine bins of fruit per day (in a ten-hour shift). Ten human pickers require the support of two tractors, two drivers and bin trailers (ten workers pick eighty to ninety bins of fruit per day). Consequently, 1.5 AKHs' could harvest the equivalent fruit volume of ten human pickers. No bin towing tractors, drivers and trailers are required to support the AKH.

Human pickers pick over a ten-hour day shift because night picking is not feasible. The AKH in any event requires artificial lighting to operate during the day due to variable lighting condi-

tions. Therefore, there is no additional cost associated with night operation. However, for humans, there are additional costs, setup, inefficiency and Occupational Safety and Health (OSH) considerations within poorly lit environments. Furthermore, pack houses limit the daily number of bins that can be harvested. Humans can only realistically sustain a certain number of daily working hours (a person will not be able to work two consecutive shifts). The labour required to harvest the required number of bins is calculated for a ten-hour day. Orchard management companies have found that human harvesting is more productive/cost effective using more staff in one day shift, rather than splitting them between a day/night shifts.

In the proposed business model, the customer is not required to pay the capital unit cost per AKH; the customer will only pay a per-bin harvest rate, as with human labour. Currently growers pay up to $33 per harvested bin. The labour component is $20 to $21 per bin, the balance comprising machinery, drivers and supervision costs. It is proposed that growers will be charged $21 per AKH harvested bin. Cost savings will be made by reducing supervision staff, bin towing tractors and trailers, gloves, tractor drivers and transportation (forklifts and follow up picking will still be required). A saving to growers of $6 to $9 per bin is expected (an average of 100 bins per Ha), resulting in savings of $600 to $900 per planted hectare. Additionally post-harvest operators will benefit from control efficiencies allowing better management of daily harvest requirements.

For an autonomous system, Sarig (1993) stated that "while this schematic system [components mentioned above] analysis may seem rather straightforward, it requires the integration of a host of technologies which are at the edge of our knowledge today". Although progress has been made since Sarig's review, no standardised or commercially available solutions for any of the functional sections exist, making the observation still relevant. Machine perception to identify and localise fruit (commonly machine vision) has historically impeded commercial operation due to slow identification speeds and limited crop identification ability. Davies (2005) noted that "the addition of visual perception to machines promises the greatest improvement and at the same time presents the greatest challenge". Advances in digital image capture and analysis has directly increased the potential for robotic harvesting systems using machine vision perception.

The interactions between the functional sections requires definition for integration planning. These functional sections are the picking arms, fruit locating system, bin filler and ATP. There

are two main interactions between these sections (in addition to physical mounting connections):

1. Communication
2. Relative physical position and orientation (coordinate space origin) of one system with respect to the other

Figure 3-1 shows the required interactions between functional sections. The four functional sections are shown in two task completion groups. The first (grey dashed box, Figure 3-1) manages fruit harvesting and handling, while the other system moves the fruit harvesting and handling systems (or potentially other devices) around the orchard so they can perform their task.

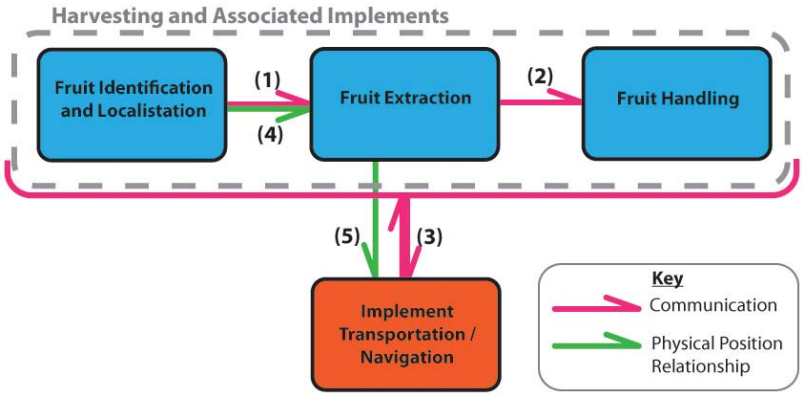


Figure 3-1: Autonomous kiwifruit harvester conceptual system schematic

Interactions (simplified, Figure 3-1) between the functional sections are conceptually detailed as follows:

1. The fruit localisation system communicates the ordered fruit positions to the fruit extraction arms.
2. The fruit extraction system informs the robotic fruit handling system that fruit are being harvested and to begin operating.
3. The transportation section informs the harvesting section when correctly positioned and to begin harvesting. The harvesting section informs the transportation section when harvesting is complete and control is returned to the transportation section.
4. The position and orientation of the machine vision system relative to the fruit extraction sections is determined. This allows the machine vision's coordinates to be transformed to the robotic fruit extraction system's coordinates so the located fruit can be accurately extracted.

5. The robotic fruit extraction systems position relative to the chassis. This allows the harvesting arms to be correctly positioned relative to the vines when ATP locomotion ends. Harvesting occurs after locomotion.

Full autonomous operation is needed to remove reliance on human input. Although using human operators could speed up development progress (e.g. driving the platform), full autonomous ability would remove this reliance and associated costs. The cost of human labour is not only financial (hourly wage and supporting services); people require regular breaks, have limited working hours and consistent task reliability can suffer. Although removing human operators requires additional sensing and environment interpretation (e.g. to achieve autonomous navigation decision making), cost savings and other functional advantages are gained (removal of manual controls and more platform space for other on-board systems). The cost and operational benefits of full autonomous ability have enough commercial justification to warrant further development investment.

Proof of concept (prototype) development began with a single AKH system. The prototype delivers system integration, enabling the methodology to be tested and commercial operation to be assessed. It is anticipated that autonomous harvesting units will operate cooperatively as swarms, with numbers dictated by harvesting demands (including orchard size and crop loading). While cooperative operation between AKH units is beyond the scope of current development, on-going development includes this requirement and the necessary supporting hardware is in place to allow future integration.

Technical specifications are derived from the commercial drivers discussed in Chapter 1. This delivered operationally functional specifications.

AKH units will be fully autonomous, with remote supervision if required. Each unit will have the following minimum overall system capabilities:

1. *Harvest at least 80% of the crop in conformance with industry fruit handling standards*
   Fruit positioning within the canopy ultimately limits extraction ability. Complete crop harvesting is not realistic as occlusions and fruit positioning behind structures or high in the canopy limits robotic extraction. Human harvesters also miss fruit and secondary 'clean ups' are often performed. The 80% harvest target is based upon relevant technical and commercial considerations. These considerations are based upon achieving a set financial return

with each AKH. There is a trade-off between AKH efficiency and 'clean up' labour requirement.

2. *Harvest rate of four fruit per second*

   A harvest rate of four fruit per second is required to achieve a commercially viable payback on the projected build cost. Four fruit per second equates to approximately four bins of fruit per hour. Taking some inefficiencies (moving the platform between harvesting areas and storage bin handling) into account, commercial outcomes are based upon a single AKH unit harvesting three bins per hour. This harvest rate is set as a balance between financial return from the AKH operating and a realistic performance metric that was considered achievable during technical conceptualisation of the system.

3. *Handle full and empty storage bins*

   To achieve full autonomy and continuous operation the AKH is required to drop off full fruit storage bins and locate, dock with and pick up empty bins.

4. *Gentle, even and complete storage bin filling in conformance with industry fruit handling standards*

   The on-board fruit storage bin requires filling with harvested fruit, without damaging fruit (see standards below). For efficiency, the bins require complete filling. Overfilling will damage fruit when bins are stacked for transportation to the pack house.

5. *Operate twenty-hours per day, seven days per week*

   Continuous twenty-hour operation is specified accounting for harvesting restrictions. Wet fruit cannot be harvested, therefore dew or rain may delay harvesting. This is a restriction put in place by the pack houses as wet fruit cannot be stored as long in the bins. Dew delays are factored into the operating period. Rain may halt harvesting but rain delays are not factored into the operation period and will extend the harvesting season. In the future, the robot will be required to determine suitable harvesting conditions. Currently these assessments are made manually. Sufficient fuel for continuous twenty-hour operation is required. Some servicing could occur in the non-operational period (e.g. checking wear and refuelling by hand).

6. *Autonomous on-orchard navigation*

   For full autonomy, environmental sensing and interpretation is required to allow navigation decision-making strategies to be made and performed. In-row, between row and around orchard navigation is required. This includes dynamic interpretation of many environmental (e.g. terrain and topographies) and infrastructural variations.

7. *Unit build cost less than NZ$130,000 (on a projected production of 10 units)*

   A projected target budget was established around a fully autonomous AKH capable of harvesting, navigating, handling storage bins and operating within the orchard row confines. AKH financial return dictated a projected build cost to allow a commercially acceptable business model. The targeted build cost was set as a trade-off between commercial return and a realistic technical build cost. This figure was agreed to between the commercialisation partner (Plus Group Ltd) and from build cost budgets during project conceptualisation.

In addition to overall system specifications, the on-orchard environment and industry fruit handling standards impose further requirements.

8. *Fruit drop height < 80mm*

   To follow industry fruit handling recommendations (to ensure fruit is not damaged during extraction and handling) fruit must not be dropped more than 80mm (free-fall, or equivalent of hitting a hard surface from an 80mm free-fall). This standard was set from recommendations from pack houses.

9. *Fruit pressure*

   The fruit grasping force for extraction, either by human or robotic system, must not damage fruit. Increasing the hands fruit contact area lowers the pressure experienced by the fruit (for a constant force). Grasping force will be minimised to work with the developed hand design. Grasping pressure will be verified by fruit damage analysis, as no standard currently exists.

10. *Fruit identification > 80% of total crop*

    More fruit than the targeted harvest rate must be identified to account for identified but non-harvested fruit. The required fruit identification rate will not be known until field trails are conducted to determine the relationship between identified and harvested fruit rates.

11. *Fruit localisation < 5mm in three dimensional space*

    Fruit positioning within ±10mm is required. This allows accurate hand positioning for extraction. This internal specification aims to achieve accurate fruit grasping to minimise mishandling and damage.

12. *Total machine width < 2.5m*

Kiwifruit orchard row widths vary from approximately 3.5 – 5.5m. Narrower row widths and general manoeuvrability requirements limit overall robotic system width to less than 2.5m.

13. *Total machine height < 1.4m*

    Kiwifruit growing structures positions fruit between 1.5 – 1.8m from ground level. This limits system height to avoid fruit contact while driving under canopy. A deflector plate may be necessary to gently guide low hanging fruit over the system to minimise fruit damage.

14. *Transportation platform four wheel drive*

    During the various seasons, especially harvest, constant ground traffic can quickly reduce traction (e.g. mud and ruts), hindering operation. A four-wheel drive transportation platform with suitable tyres will mitigate these effects.

15. *Transportation platform turning circle diameter < 3m*

    A tight turning circle is required for efficient manoeuvres within the restricted on-orchard environment (row-to-row, headland and other turns).

Overall system development will be successful if each operational section performs to the required performance specifications. Although the performance criteria apply to the individual sections, the interactions and integration of these sections is critical to overall operation. Therefore, these interactions and integration into the complete system will also be assessed as part of the AKH's viability. Testing will be conducted to assess the systems functional sections. This testing will also assess any critical interactions between systems (such as the interaction and integration of a sensor working with a mechanical system).

The development objective is to produce a complete system that is capable of autonomously harvesting fruit within all pergola styled kiwifruit orchards. Other practical applications of the system will also be considered (e.g. using the autonomous transportation platform for other orchard operations) during this development. This will allow the AKH to be easily modified/retrofitted with additional hardware and software for future on-going developments. However, these future developments are outside the scope of this thesis. With these defined aims, objectives, and specifications, development of the system can begin.

An orchard entrance with kiwifruit planted next to the driveway. Photo reproduced with permission from Plus Group Horticulture Ltd.

4. System Development

Before dealing with the development of the various sub-systems, some fundamental design, integration and implementation decisions need to be made. These deal with how the fruit should be identified and located, how it should be harvested and handled, and how the harvesting system should be transported around the orchard. This is dealt with in the following sections of this chapter.

4.1 Fruit Identification and Localisation

Autonomous operation requires automatic identification and localisation of fruit within the canopy. Observing the canopy from below, fruit are located over a relatively small height range (approximately 300mm). They have some distinctive visual characteristics (e.g. the blossom remnants and colour) that are consistent. Assessment of other harvesters (section 2) identified fruit identification and localisation as limiting factors of overall performance. These systems (further investigated in chapter 5.1) used machine vision to identify and localise fruit. However, the commonly used eye-in-hand configuration (camera mounted within the robotic hand) requires a fruit 'hunting' process. This process first observes the scene (requiring the robotic arm to move to a defined viewing distance), then targets and tracks a specific fruit until it is grasped. This is a time consuming process which requires large arm movements to find fruit before extraction begins. Mounting the camera(s) away from the arm removes the requirement for a 'hunting' routine. This may increase system efficiency, but will also increase complexity. Requiring depth perception and transformation of the observed object position into the harvesting arm's coordinate space to allow fruit extraction increases complexity. Stereo machine vision offers a viable identification and localisation option. This could exploit the unique visual characteristics of the fruit and was selected for:

1. Low cost sensing: supports the commercialisation model
2. High level of flexibility: the system can be programmed and adapted for the specific application, not available with other sensing options
3. Accurate passive depth perception

4. ‘Intelligence’ potential: a large amount of observed information is available for analysis, potentially increasing sensing awareness.

Other three-dimensional scanning systems were considered for this application. These included laser scanners, time-of-flight cameras and other colour/depth sensors like the Kinect system. There systems were each ruled out from slow scanning speeds (three-dimensional laser scanner) or their inability to reliably operate with the sunlight directed down at the sensor. The laser scanner system is reviewed in section 5.1. Like the stereo system, these other approaches would also suffer from occlusion issues. The impact of occlusion can be reduced by using more sensors, which is feasible with a lower cost sensor like cameras.

4.2 Fruit Extraction

After identifying and locating the fruit, robotic arm(s), each with an end-effector (hand), will extract the fruit. A proven and cost effective arm drive design was proposed by Flemmer (2009) as an alternative to current industrial arms. The stepper motor powered concentric drive system (section 6.1) was previously proven in a SCARA (Selective Compliant Articulated Robotic Arm) configured robot. The drive system required redesigning into a configuration suitable for the harvesting operation. In general, SCARA robots primarily operate through the horizontal plane (fixed Z-axis, articulated motion through the X-Y plane) with a distally attached actuator providing Z-axis range where required. Although kiwifruit generally lie within a horizontal plane, the harvesting arm requires positioning over a small vertical range (approximately 400mm). A three-axis articulated arm with a rotary base was developed for harvesting (chapter 6.2).

The arm positions a fruit extracting hand relative to the fruit. The hand is required to mimic the human fruit extraction technique. Humans rotate the fruit about the stem-fruit connection to ensure consistent and damage free fruit removal. Fruit must not be damaged by the grasping pressure during extraction. Development of the hand is presented in chapter 6.2.3.

4.3 Fruit Handling

Once fruit is harvested, it requires:

1. Conveyance to the storage bin
2. Gentle placement within the storage bin

A series of conveyors servicing the arms, with selected pocketed lifting sections, would satisfy the requirement of point 1 above.

Once fruit are at the storage bin, a method to gently place them into the bin is required. The low kiwifruit canopy height limits vertical size and operational motion options. The eighteen-bushel storage bins measure 650mm high. 150mm of ground clearance under the bin support slats (also 150mm), leaves 600mm of operating space above the bin. Complete bin filling is required (1200 x 1200 x 600mm) while operating under the canopy. The transportation platform and other on-board systems also restrict available operating space.

Experiments were initially performed with vertical tubes containing fruit decelerating baffles. Fruit was successfully slowed so as not to cause damage. However, space restrictions and the large filling volume made the system impractical. Variable on-orchard terrain limits bin filling system design options as gradients can cause fruit to move within the bin during operation. To stop fruit shifting and moving during bin filling the fruit is moved in 'pockets' to the desired placement position. This reduces potential fruit damage (fruit-fruit and fruit-mechanical contact) during fruit handling. A bin filling arm consisting of an articulated arm housing a pocketed conveyor capable of moving across and into the bin (chapter 6.3) is envisaged.

4.4 System Locomotion and Power Supply

An automatic guided vehicle is required for on-orchard transportation. This vehicle positions the harvesting arms' (or other implements from future developments) relative to the vines to allow task execution. This is presented in chapter 7 and requires:

1. A chassis
2. Propulsion and steering
3. Power source (for locomotion and other systems, e.g. computers, arms')
4. Obstacle sensing for navigation
5. Computer control (for locomotion and navigation decision making)
6. Kiwifruit storage bin handling
7. Weather proofing

A low slung and flat chassis platform supports modular system attachment and unobstructed operation (e.g. harvesting arms and machine vision). Although increasing overall machine mass (to achieve the same chassis strength as a three dimensional structure), it is easily fabricated. To

maintain all wheel contact (to maximise traction and stability) through varying terrain the steering and drive wheels are mounted to an articulating front end. This system is commonly found on agricultural tractors.

Hydraulic drive for chassis propulsion and steering was selected for its high power density and moderate cost. Both DC (brushed and brushless) and AC electric systems were considered (motor types outlined by Teat (2011)). DC motors are preferred in electric vehicles due to the relatively simple control electronics and reduced servicing needs (Larminie and Lowry, 2003). The size and weight, cost, large steering motors and mechanical brakes of a DC motor drive system outweighed its advantages (simplified control, no pumps, hoses etc.). A stationary petrol internal combustion engine provides power to the hydraulic pump and generates power for the other on-board systems. A petrol engine is only used in this proof of concept system. Future development could use hybrid petrol-electric or diesel-electric power systems running on bio-fuels to reduce environmental impact. This would help reduce the carbon footprint associated with fruit to market delivery.

Autonomous operation also required sensing of obstacles within the orchard. Stereoscopic machine vision was selected for navigation sensing. This allowed dynamic interpretation and control strategies to be developed. Stereoscopic vision has cost and flexibility advantages. However, on-orchard limitations and complexities could restrict performance (see section 7.3 for navigation sensor options and review). This navigational approach was assessed (section 7.4), and the complexities of reliable on-orchard operation are detailed. LIDAR sensors are integrated to provide reliable object detection and work with the stereoscopic machine vision system to provide increased interpretation ability.

The ATP requires on-board storage bin handling. Within the rear of the ATP a pneumatic forklift system with dynamic height control is planned. Pneumatic lifting actuators are preferred (over hydraulics) as they provide a cushioning effect. The cushioning helps maintain fruit integrity while the ATP negotiates rough terrain.

Unfortunately, the requirements of each sub-system mean that three different power systems are required (electrical, pneumatic and hydraulic). Although this increases complexity, it was a decision to get the best overall system performance. Electrical power is essential for the computing units and sensors. In addition, electrical motors also provide the best control for the robotic arms. Therefore, some form of electrical energy was always necessary. Even though it added

another drive system, hydraulic drive was determined to be the best option for vehicular propulsion and steering (as discussed in this section). The addition of a pneumatic system was initially added for the bin lifting system. This would provide the bin with some cushioning as the ATP traversed rough terrain. In addition, pneumatics were used to actuate the harvesting hand. This was desirable, as the actuator will allow for a change in length to the clasping cable without crushing the fruit.

4.5 Section Implementation

The process of implementing each conceptual design into an operational system is discussed in this section.

The fruit localisation system is initially implemented with a single stereo camera pair. The cameras are assisted by artificial lighting to help reduce natural lighting variation. These cameras capture and record a series of on-orchard images allowing the development of the fruit identification algorithms. Testing these algorithms in real world conditions verifies performance and system suitability. A stereo camera calibration process allows each fruit's three-dimensional position to be determined relative to the cameras. The fruit localisation system is then integrated with the robotic harvesting arms to allow robotic extraction of camera located fruit. This is achieved by calibrating the hand-eye coordination between the systems (section 5.3). This system is replicated over the four fruit harvesting arms.

The stereo cameras should not require recalibration during field operation. They are securely mounted within the vehicles structure to prevent being knocked. Ruggedised enclosures will also prevent any misalignment occurring from falling objects or from vibration and rough terrain during driving operation.

Implementation of the harvesting arms begins with fabrication and assembly of the initial design. The arm is designed for ease of fabrication with the available manufacturing resources. The main structural members are machined on a three-axis CNC machining centre. Design validation and manufacturing accuracy occur during and after assembly. Once the arm's components are assembled, the unpowered arm is manually moved to observe its operational envelope and the system's ability to reach into the canopy. This helps validate the design and manufacturing processes. Once assembled, the system's motors are wired to the drivers and controllers. Programming starts once the wiring is correctly installed. Each programmable action is

then tested to confirm operational ability (e.g. can each motor be driven as commanded and can all the sensors be monitored). This allows for programming and debugging of the complete system until desired control is achieved. Once programming is completed, the robotic arm is integrated and calibrated with the vision system (as discussed above). Testing confirms system performance and if necessary a redevelopment cycle is implemented to resolve any issues.

Bin filling arm implementation is similar to harvesting arm implementation. However, instead of integrating the bin filling arm with a machine vision system, it is integrated and tested with sensors to detect fruit in the bin. This allows a safe fruit layering height to be maintained during bin filling, ensuring fruit quality. Testing will determine system performance and if any redevelopment is required.

Implementation of the ATP (see section 7) begins with the design, fabrication and assembly of the ATP chassis and related power source. Wiring of controllers and drivers follows. This allows testing of each component to ensure correct operation (e.g. does each motor drive as directed and can all the sensors be correctly read?). Upon validated operation, programming to coordinate system operation and locomote the chassis is undertaken. Programming is tested by performing driving manoeuvres. Once the ATP drives as specified the navigation sensors and analysis algorithms are integrated to allow autonomous operation. Further testing will confirm this ability.

Hanging kiwifruit. Photo reproduced with permission from Plus Group Horticulture Ltd

5. Fruit Identification and Localisation to Achieve Robotic Extraction

Fruit identification and localisation has been a major obstacle to other discrete crop harvesters (see section 2). Slow identification and extraction methods limit their system's performance. This directly affected viability compared to using human labour.

Machine vision identification is commonly used for fruit crops (section 5.1); however, implementation has been limited. Using machine vision for finding and locating kiwifruit is desirable for its flexibility, low cost and 'intelligence' characteristics. The limited success of other reviewed harvesters (section 2) guided development to solutions which removed those issues. These issues include slow and low identification rates and inaccurate fruit positioning. Remotely positioning the fruit identification camera below the arm's hand allows fruit identification while harvesting over a larger canopy area. This improved system performance by allowing continuous fruit finding while the arms are operating. Successful extraction requires accurate three-dimensional localisation of each piece of fruit. A two-dimensional image from a single camera provides insufficient information for three-dimensional localisation. A pair of calibrated stereo cameras provides the system with depth perception. An alignment process determines the mathematical mapping between the camera and arm coordinate spaces. This enables the fruit position determined by the cameras to be transformed into the equivalent arm coordinate position. The arm can then be positioned to any reachable camera derived position.

Finding kiwifruit (with machine vision) is more complex than other fruit crops (such as oranges and apples). Many other crops have significant colour contrast between the fruit and the surrounding tree, making the fruit easier to find. Kiwifruit's green visual appearance merges with the surrounding canopy environment (limited hue colour difference) making it difficult to find. The variable colour differences within the canopy and leaves present areas of similar appearance, potentially confusing the fruit finding system. It is therefore necessary exploit the fruit's visual characteristics beyond colour segmentation. It must be capable of dealing with the variable visual phenomena within the canopy, where the most complex variable is sunlight illumina-

tion. Sunlight effects are increased as the stereo cameras observe the canopy from below (at times looking into the overhead sun). Other fruit identification systems operate in a horizontal orientation, reducing sunlight effects. Illumination variability occurs throughout the day (from daylight to darkness, compounded by changing weather conditions), and within individual images. Leaves, or the lack thereof, create localised illumination variations altering the fruit's visual appearance within individual images. These environmental changes require solution.

Hidden fruit (occlusion) also limits fruit finding ability. These occlusions include the fruit being partially visible in one or both cameras, only visible in one camera or not visible at all. Partial occlusion (in any form) makes definitive identification difficult. Occlusions can occur from several (potentially concurrent) canopy factors, including: other fruit, leaves, canes, structural support wires and canopy structure. Fruit-fruit occlusions are the most common presentation, but have limited impact on the systems overall fruit finding ability. This is because as fruit are removed during harvest, previously hidden fruit are exposed. Occlusion reduces harvesting efficiency, as multiple fruit finding routines and larger arm extraction movements are required.

This thesis section covers the necessary combination of technologies required to achieve on-orchard robotic fruit extraction. This includes locating fruit and a method to determine the mapping between the machine vision and arm systems.

Figure 5-1: A typical kiwifruit canopy with fruit hanging. Note: lighting conditions and fruit occlusion

5.1 Review of Fruit Identification Systems

Visual fruit detection techniques date back to 1968 with Schertz and Brown's (1968) paper. They suggested that fruit locations might be determined using changes in the light reflectivity between the fruit and surrounding foliage. Also discussed were the problems that may arise from inconsistent lighting and fruit occlusion.

Parrish and Goksel (1977) presented the first machine vision based apple identification system. This system consisted of a black and white camera with red optical filter to increase the visible contrast between red apples and green foliage. Information about overall detection rates etc. was not reported (Jimenez, et al., 2000).

Jimenez, et al. (2000) describes two common approaches for fruit detection using machine vision. These are:

1. Local analysis
2. Shape analysis

Local analysis commonly uses colour or intensity information to segment the desired image areas correlating to fruit. This represents an effective fruit identification method for fruit presenting colour differences from the surrounding environment (branches, leaves, etc). Kiwifruit do not present such obvious colour distinctions. Therefore, implementing such an approach for kiwifruit would require additional verification and image analysis. Colour distinctions (fruit to background scene) were often accentuated by band-pass filters. This increases the colour contrast between the fruit and background (D'Esnon, et al., 1987, Grasso and Recce, 1996, Levi, et al., 1988, Parrish and Goksel, 1977, Schertz and Brown, 1968, Sites and Delwiche, 1988, Van Henten, et al., 2003).

Shape analysis uses fruit edge information for detection. It is generally performed independently from colour analysis. This permits fruit detection at any maturity level (fruit can significantly alter in colour while maturing). Jimenez, et al. (2000) noted that shape analysis without colour or intensity information, can result in scene misinterpretation when features like sunlight and shadows are observed. The unwanted but observed features can present circular contours similar to fruit shape, leading to false positive identification.

Many variations of these two approaches exist. For example:

1. Morphological filter (for feature extraction): applied to smooth small dark regions and small light regions resulting in more uniform contrasts (Dobrousin and Edan, 1992, Parrish and Goksel, 1977, Sites and Delwiche, 1988).
2. Template matching: a template representing a desired identifiable feature is compared to the image at several (or even all) pixel locations. The template will produce higher values when correlated to the desired features (Levi, et al., 1988, Sites and Delwiche, 1988).
3. Circular Hough transform: performs analysis on identified pixel positions to determine the presence (and significance of the presence) of circular features. This can correlate to the boundary of round fruit (Benady and Miles, 1992, Grasso and Recce, 1996, Whittaker, et al., 1987).
4. Centre of mass: Finding the central mass of identified pixels provides the theoretical centre of the fruit. Its accuracy is unreliable for partially occluded fruit (Sogaard and Olsen, 2003).
5. Contour tracing: Muscato, at al. (2005) uses the visible peripheral fruit contour(s) to estimate its perimeter and determines the fruit's centre, diameter etc. (see Figure 5-2). This method can distinguish partially occluded fruit.

Figure 5-2 Image processing to locate the oranges in an image (Muscato, et al., 2005)

6. Sobel operator (Sobel, 1970): Uses gradient profiles to identify contour boundaries. Contours can be extracted that resemble fruit features (Levi, et al., 1988, Whittaker, et al., 1987).

Once fruit are 'found', further validation of their existence is usually required. Classification criteria validate fruit presence and help eliminate false positive identification. Three main classifiers are used:

1. Rule based classification: the object must comply with set rules, i.e. the diameter must be between x and y, the average colour should be z (Cardenas, et al., 1991, Dobrousin and Edan, 1992, Parrish and Goksel, 1977).
2. Bayessian classifier: Similar to a rule based classifier, but analyses individual properties to form classification probability (Slaughter and Harrell, 1989).
3. Circular Hough transform: This directly provides the level of support to circular features.

The classifiers provide a pass or fail fruit presence result.

The most common machine vision approach uses local analysis, based upon spectral imaging (Jimenez, et al., 2000). This identification system is generally only suitable for fruit crops with distinctive colour differences to the background environment (e.g. apples and oranges). As stated above, kiwifruit present a similar colour profile to its host environment. Therefore, a spectral imaging method is not independently sufficient.

Although machine vision is the most common fruit identification system, another system Jimenez, et al. (1997) uses a three-dimensional laser range finder to determine fruit positions. The range finder can be combined with machine vision to provide additional information for classification (Jimenez, et al., 1999). "The recognition strategy focused on general spherical object recognition since oranges, apples, peaches or similar fruit can be modelled as spheres" (Jimenez, et al., 1997). Laser beam reflectivity is proportional to the surface angle at the incidence point. Therefore, a spherical fruit produces a unique and identifiable reflectivity profile over its surface. Fruit with a waxy surface accentuates the reflectivity profile as the skin more closely represents a reflector. The fruit's skin reflects more laser light back to the range finder when compared with the surrounding plant, making the fruit more obvious in the reflectivity data.

The combination of colour and surface profiles (both distance and reflectivity for laser information) provide complimentary identification and classification results (Jimenez, et al., 1999). Empirical testing identified 80-90% of visible fruit within an artificial orange tree. The system's primary drawback is the long scanning time (25s) and small scanning area. Increasing scanning speed reduces system accuracy. With a current accuracy of 10mm (in three-dimensional space), further inaccuracy would compromise robotic extraction. Figure 5-3 pictorially represents the scanned laser range and reflectivity data.

Figure 5-3 Range/reflectance imaging by (Jimenez, et al., 1997). (Left) distance image, (Mid) reflectance image, (Right) detected fruit overlapping the range information.

Most machine vision systems for robotic fruit harvesting use a camera-in-hand configuration. This does not require fruit distance information. Fruit are initially identified by moving the arm away from the plant. At this point fruit distance is unknown and is determined from a sequence of images captured as the arm approaches the fruit from the arm's known positions (using a Kalman filter). This allows arm control and path prediction. Pneumatic vacuum levels or proximity data was used to indicate fruit placement within the hand, triggering extraction (Baeten, et al., 2007). Grasso and Recce (1996) and Kassay (1992) independently used stereo vision for finding the fruit's distance. However, their methods for stereo matching are not included in the literature.

5.1.1 Review Conclusions

It is apparent that machine vision techniques for finding and locating fruit are crop specific. A machine vision system for finding and locating kiwifruit requires specific development. The methodology identified in this review provides a developmental base which could be tailored to the unique kiwifruit environment.

Schertz and Brown (1968) have identified problems with inconsistent lighting conditions and fruit occlusions. Lighting inconsistencies' are accentuated with vertically oriented cameras (as for kiwifruit finding) in comparison to a camera looking horizontally into tree crops. The vertical viewing orientation looks directly towards the sun and does not allow shrouding to stop direct sun glare.

A combination of local and shape analysis presents viable analysis options. Kiwifruit do present some colour difference to the surrounding environment. However, filters (as used by Parrish

and Goksel (1977)) to enhance the fruits colour appearance are not practical as fruit colour is too similar to the surrounding canopy. The colour difference could be exploited for initial local analysis and eliminate most of the non-conforming image areas (those areas that do not appear to be fruit or similar to fruit). Within the segmented local analysis areas, shape analysis could determine actual fruit positions. The fruits blossom end presents a distinctive visual characteristic (the brown, near circular shape at the bottom end of the fruit). It is surrounded by a near uniform fruit colour. Shape analysis could identify this characteristic shape. A combination of local and shape analysis could therefore provide reliable fruit identification. A simple, yet functional solution is desirable for this application. Simpler solutions often allow for faster image processing, improving system efficiency. If this approach does not provide the necessary results, higher-end analysis will be further investigated. These methods could include such approaches as active shape analysis to find the fruits profile, or texture analysis to find the fruit's blossom end.

The developed analysis method is required to identify as many fruit as possible and accurately localise those fruit in three-dimensional space. This allows the identified fruit to be robotically harvested. In addition, this must occur within the variable orchard environment. Increased fruit identification increases overall harvesting ability and therefore commercial use. Development begins with the objective of identifying as many fruit as possible within real on-orchard canopy environments.

5.2 Stereo Machine Vision Fruit Identification and Localisation

Locating the fruit on-orchard is an essential part of the development. System efficiency, commercial viability and potential payback are increased by accurately locating more fruit for extraction. This also reduced the amount of manual labour required to harvest the remaining fruit missed by the AKH.

Mounting the camera within the hand simplifies coordination between camera and arm coordinate spaces. However, system inefficiencies result:

1. The arm must traverse from its operational area to allow fruit finding, increasing traversal distance and decreasing cycle times.
2. Path planning requires continual vision system updates, limiting arm traversal speed to the vision system's processing speed.
3. Shadowing occurs as the arm approaches fruit. Additional hand mounted lighting increases its penetrating profile, potentially causing fruit damage, or canopy entanglement.
4. An additional camera (and potentially lighting) adds mass distally to the arm, resulting in the most inertial increase for the given mass. This decreases arm performance and efficiency.
5. If fruit is not located from the initial search position, the arm is required to search the complete operation area (known as a 'hunting routine'). This reduces operational efficiency.
6. Requires Kalman filters (or similar) and mathematical descriptions of the systems performance to continually estimate and update fruit position on arm approach. This results in increased targeting and arm control complexity, decreasing operational performance.

Implementing remotely mounted cameras for fruit identification and localisation, mimics human visual perception. By observing the scene from a distance and employing depth perception, fruit can be localised in three-dimensional space. Fruit locations require transformation into the arm's coordinates to allow extraction. Although more complex, this configuration does not suffer the limitations of an eye-in-hand system, described above. Depth perception is achieved by using a pair of stereo cameras, where image analysis determines depths (section

5.2.3). This system is mathematically aligned to the robotic arm's coordinate space to achieve accurate hand-eye coordination (section 5.2.4).

5.2.1 Kiwifruit Identification

Figure 5-5 is a pair of captured stereo images observing the canopy (see section 5.2.3 for stereoscopic depth perception). Fruit within the canopy images are obvious to a human observer, even those which are partially occluded. A person can identify fruit, even when presented with size, shape and lighting variations. Machine vision identification aims to replicate this ability, finding all fruit from images that present appearance variations. To achieve this, key visual features, obvious to human interpretation, are exploited.

Observing the fruit in Figure 5-4, their noticeable features are:

1. A small colour difference from the surrounding scene
2. Brown blossom remnants on the fruit's lower end
3. Reflective bright area around the blossom end, making the blossom end more distinctive (contrast increase)
4. Distinguishable border around the fruit blossom end and often at the edge of the fruit

Figure 5-4: Expanded image from Figure 5-5

Figure 5-5: Stereo pair of images captured for finding fruit. Note the search area in the right hand image (inside black box)

First, the fruit's colour is exploited to remove most of the unwanted scene information. To achieve this, the image is segmented by subtracting a predefined reference colour range from each pixel, and selecting those pixels which are within a given threshold. Although not distinctive, this simple segmentation is able to eliminate significant proportions of the canopy from further processing. Figure 5-6 shows the surrounding canopy scene captured during segmentation. Reliable identification requires additional processing. The reference colour range is sufficient to cover the appearance of the fruit over the season and throughout the day. This is why fixed thresholds could be used. This makes it a reliable initial step as further analysis is performed to define the fruit positions.

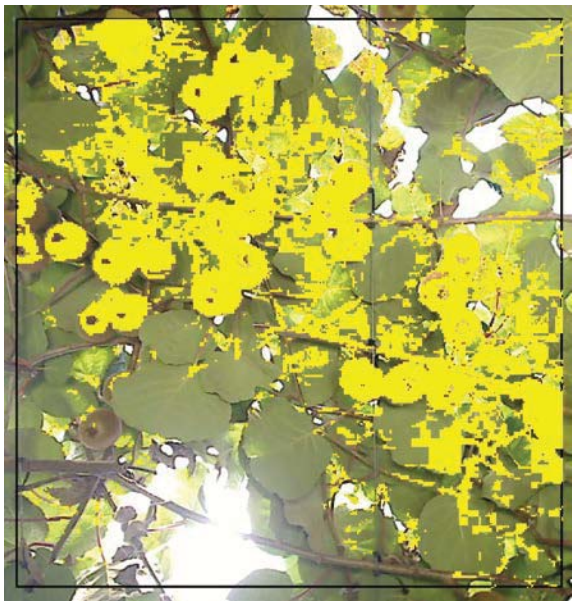


Figure 5-6: The result of segmentation to the operating envelope in Figure 5-5

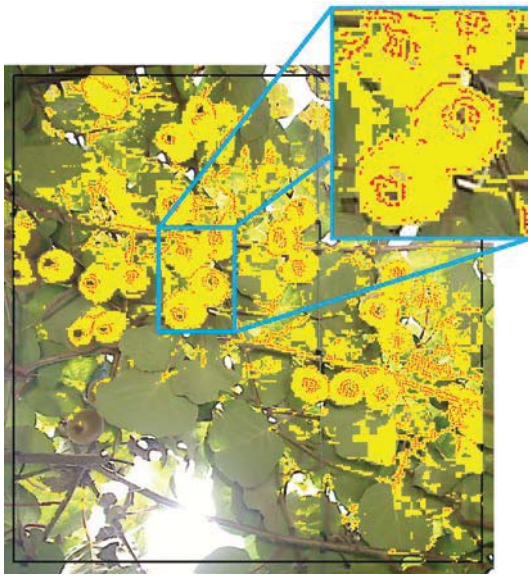


Figure 5-7: Result of Segmentation and Edge Detection

A Sobel filter (Sobel, 1970), responds strongly to the fruits blossom end. Artificial lighting helps enhance its appearance by increasing the contrast and intensity differential (light halo around blossom remnants). A Canny edge detection method was initially tested for this process. However, the Sobel filter provided better results over the sample images.

To increase the speed, filtering is only performed within the segmented areas (yellow points - Figure 5-7). This reduces data noise, minimising false positive identifications.

A circular template is scanned over the image looking at the segmentation and edge detection data. The template analyses thirty-two radial points around the circle's perimeter (blue ring in Figure 5-8). This template has a diameter of twenty-eight pixels to make it encompass a fruit's blossom end and initially only the periphery points are assessed. If more than thirty out of the thirty-two points corre-

spond to segmented points (yellow points in Figure 5-8) then the number of edge pixels within the circular template is counted. If there are thirty-eight or more edge pixels within the template then the centre of gravity of those points is calculated as the candidate blossom end location. The fixed size template function is to narrow down the search locations within the segmented data. As the fruit is hanging over a relatively small height range the templates size is set to encapsulate the fruits blossom end within the image. This blossom end does not change in size as much as the fruit does. This is because it is formed from the flowers stigmas that have dried out after pollination. Therefore, it is not subjected to the vast variability of the growing fruit.

When the template is positioned over a fruit in the image, the fruits blossom end edge points are the majority of points forming a pattern within the template. It is desirable to position the robot's hand to the centre of the fruit end. Therefore, the centre of the blossom end creates a well-defined target. The centre of the search template is therefore,

$$x_c = \frac{\sum x}{\sum m} \quad (1)$$

$$y_c = \frac{\sum y}{\sum m} \quad (2)$$

Figure 5-8: The search template (shown as a blue circle) at a positive fruit identification point

Blossom end edge points present a distinctive visual appearance. Mechanically, dispersing points further from a rotational centre increases the systems rotational inertia. Inertia (I (3)) of the edge points is assessed to determine point dispersion. Blossom end edge points are spread from centre of mass. False positive template 'hits' generally have points centrally clustered or more evenly spread (for the number of edge points presented), reducing the inertia for a given 'mass'.

$$I_k = \sum_{k=1}^{N} m_k r_k^2 \quad (3)$$

Since the mass of a detected pixel =1 therefore,

$$I_k = \sum_{k=1}^{N} r_k^2 \quad (4)$$

(3) shows that inertia, I, is proportional to the mass of the total system. Total system inertia is not of interest, relative inertia of the given mass is. Therefore the inertia is scaled by total mass, giving the radial variance, V_R, of the detected edge pixels. Points within the template require a V_R greater than twenty-one (determined through experimentation).

$$V_R = \frac{\sum r_k^2}{\sum m_k} \tag{5}$$

Final template point validation is determined from the uniformity of edge data variance, using a width/height ratio. The width/height ratio of the edge points is used to assess dispersal uniformity. As the width/height ratio approaches one (even horizontal to vertical point dispersal) the likelihood of desirable circular point dispersal increases. Meeting this width/height point dispersal criteria (and those discussed above), the potential fruit location is stored.

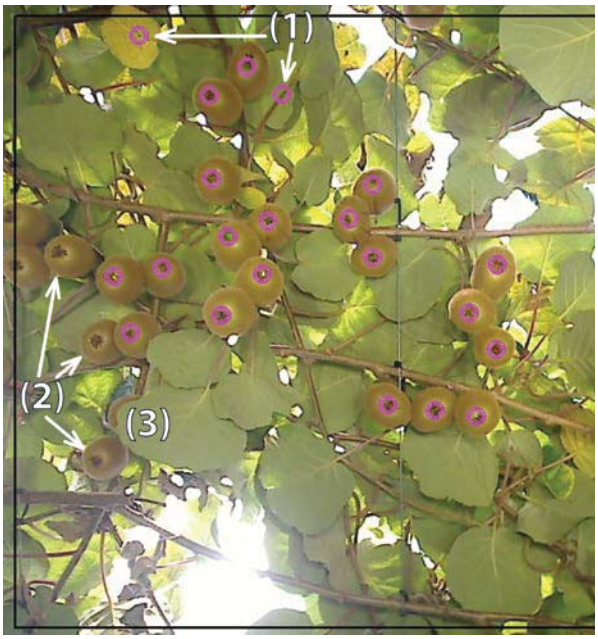


Figure 5-9: Identified 'fruit' from clustering the data 'hits'. Note the two false positives (1), three missed (2) and one partially occluded (3) fruit.

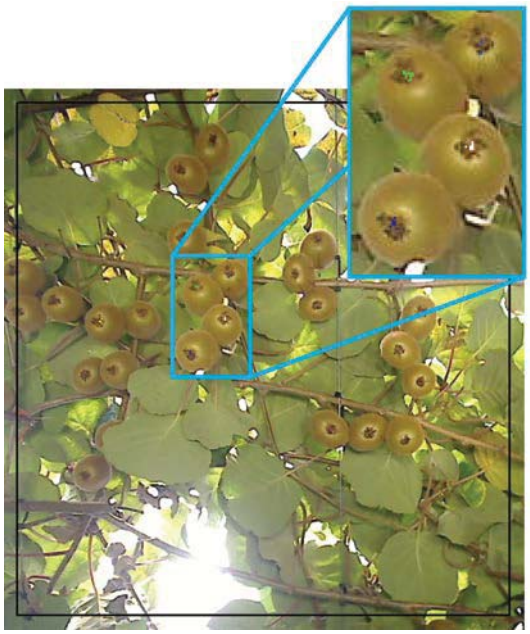


Figure 5-10: Result of clustering the data points. The different coloured points represent different clusters

Sweeping the template over the complete image produces multiple candidates for the location of a single fruit (Figure 5-10). The centre of mass of each cluster is determined ((1) and (2)). Experimentation determined that clusters correlating to fruit needed to be above a minimum 'mass' (number of data 'hits') and more than a minimum distance from another cluster. If two or more, clusters are too close, the most dominant cluster (having

the largest 'mass') is kept and the other rejected. Occasionally this produces a false negative result (rejection of an actual fruit position). The 'rejected' fruit is located on subsequent searches, following dominant fruit removal.

Some of the detected fruit are false positive at this stage. To minimise these a ratio of dark to light pixels is found over the candidate fruit's blossom end. Requiring a minimum number of pixels below a set brightness threshold reduced false positive occurrence. Figure 5-12 shows the removal of two false positives (shown by light blue rings).

After identifying the fruit, stereo image analysis localises the fruit in three-dimensional space to allow robotic extraction. The stereopsis equation is rearranged to define a small search area for the fruit in the second camera. This can be done as the fruit hangs between 1.4 and 1.8 m above the ground, constraining the expected disparity between the stereo images. The black boxes (Figure 5-11 and Figure 5-12) show the search area for fruit matching in the left hand image. This scan area was made sufficiently large enough to address any projective rectification or similar issues. Most of the other image distortion issues are accounted for in the stereo camera calibration, including: focal length and different pixel row and column scaling. The cameras were found to have minimal lens distortion; therefore, this is not accounted for. In addition, the fruit are found within a restricted area of the image, limiting any lens distortion effects which are

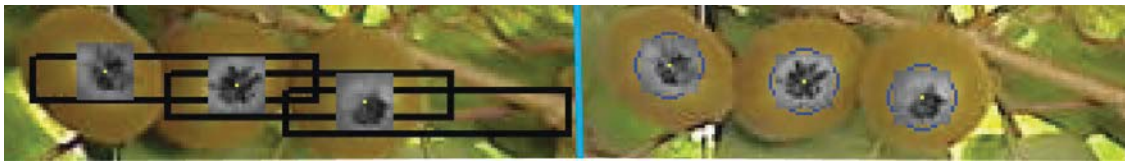


Figure 5-11: Stereo fruit matching. Result from finding the right hand fruit in the left image. The yellow dots, right hand image, are required to correspond to the same position in the left hand image.

Figure 5-12: The located fruit. Note the two false positives have been excluded (blue rings denotes point rejection in right hand image).

more noticeable further from the image's centre. These issues are discussed further in section 5.2.3.

To accurately locate the fruit between cameras, the distinctive blossom ends are again used. Unlike the bulbous fruit body, the blossom end has a relatively flat presentation (see Figure 5-11), providing a similar appearance between stereo images.

For fruit location in the second camera, a normalised black and white (brightness) template of the fruits blossom end is formed (Figure 5-11). This 16x16 pixel square template is normalised to reduce illumination variation effects between the cameras. Normalisation is achieved by shifting and scaling the template's pixels brightness values to make them cover the full 0-255 brightness range. Reliable matching was achieved with a black and white template. This is due to the shape of the blossom end provides the most significant match. Therefore, including colour was not required and would only slow the matching process. This template is scanned over the search area in the left image (which is also normalised) to find the minimum sum of squared differences with the template. However, this assumes the fruit is visible in the second image and not occluded (see Figure 5-13, fruit 1, 3, 4 and 5). If occluded, the resultant location could be a neighbouring fruit or a random position. Due to the blossom ends distinctive appearance a threshold of the sum of the absolute differences is set. This rejects most false positive identifications as the fruit's blossom end appears to be as unique as a finger print. A sample of fifty fruit blossom end templates were compared to each other to find if false positive matches could occur. Even when the templates are rotated and shifted to try and achieve a match, none occurred. Therefore, it is highly unlikely that a false positive match will occur with two neighbouring fruit, when the best match is always selected.

The disparity in detected fruit location between the two images localises each fruit in three-dimensional space (section 5.2.3). Calculated positions are transformed into the harvesting arm's coordinate space to allow extraction (see hand-eye coordination, section 5.2.4). Images from both stereo cameras are captured in rapid succession (within 10ms) to ensure accurate scene representation. Ideally, the cameras would be synchronised, however the USB cameras used do not support this functionality, unlike FireWire cameras. Fruit movement, from wind or canopy disturbance during fruit extraction, can shift fruit location between subsequent images, causing errors in depth estimation.

Identification limitations predominately arise from occlusions and lighting inconsistencies'. Figure 5-13 shows the limitations caused by occlusion:

1. Tight fruit clustering (Fruit 3)
2. Leaves (Fruit 2, 4 and 5)
3. Plant canes
4. Deflections stopping the fruit hanging vertically (Fruit 1)
5. Canopy structure: wire, AgBeam (canopy structure material), timber, etc.

Encouraging fruiting in the lower more accessible canopy areas would minimise occlusion. The additional crop management to achieve this is not necessary for human harvesting as foliage is

Figure 5-13: Stereo image pairs. Top Pair - demonstrates issues arising from occlusions as fruit are visible in only one camera. The vision system is required to handle these scenarios. Bottom Pair – The result from computer analysis.

easily moved or observation points altered, locating most fruit. Robotic harvest technology will encourage changes in pruning and layering of canes to maximise harvest efficiencies.

Inconsistent canopy lighting restricts fruit identification. Modifications to current artificial lighting could reduce variable lighting effects. High intensity lighting, from camera flashes, HID (High Intensity Discharge) or LED arrays could be used to saturate the canopy with light. This would reduce noticeable natural lighting effects through canopy gaps. With the additional artificial lighting the cameras exposure time needs to be reduced to ensure clear images are still captured.

Figure 5-14: Canopy image showing the variation in illumination from natural lighting effects. Note the difference in fruit appearance across the image.

Another method to help maintain image detail with the variable natural lighting is to use high dynamic range cameras. These cameras use multiple image exposures and merge them into a single high dynamic range image allowing over or under exposed image compensation. However, this requires additional images processing to extract the information from the high dynamic range images to compensate for the changes in light levels. Furthermore, each of the images taken at different exposure levels are captured sequentially. This means that the target could potentially move between images and affect accuracy. The idea of this technology is to provide images that more closely represent human visual interpretation. Humans automatically perform this process, allowing us to view shaded areas, from a brightly lit vantage (or vice-versa), or simultaneously interpret bright and dark areas of a scene. This technology might become viable in future applications as it progresses.

5.2.2 Fruit Identification – Testing and Results

To assess fruit identification performance forty pairs of stereo images were captured:

1. Mid-season over two days (twenty pairs each day)
2. At different times of the day (morning and afternoon)

3. On a sunny and overcast day respectively
4. Over several rows and row positions
5. With varying canopy environments

This ensured a representative range of lighting and real-world operational conditions were presented. 821 fruit were observed by the author within the stereo images search area (an average of 20.5 fruit per image pair). The total count for each of the criteria (in Table 1) was determined by counting the analysed machine vision output images. These analysed stereo images showed the fruit located with the developed algorithm. Unfortunately, due to restricted orchard access after Psa the initial test set that was captured was used for testing and validation. This is because it captured before the Psa outbreak. Therefore, no repeatability testing was performed. Although this was the case, the test set does cover the majority of resemble orchard variation. Therefore, it provides a strong correlation to actual expected performance. It also showed the fruit's stereo correspondence position in the second image (as shown in Figure 5-12). Machine vision identification provided:

Table 1: Fruit identification results using machine vision

Criteria	Count	Percentage
Positively identified	571	69.55%
False positive	15	1.83%
Unidentified	247	30.09%
False negatives	3	0.37%

For this assessment a false negative result occurs when a fruit is initially identified but is later rejected during verification (see section 5.2.1). False negative correlations can be grouped with unidentified fruit. This generally occurs during stereo matching when the blossom end cannot be found in the second image. If the position cannot be matched in the stereo image, then the identified position is rejected. Additional verification helps reduce the number of false positives.

Fruit obscuring fruit (fruit-fruit) occlusions impaired positive identification. Identified and possibly occluding fruit were not removed during algorithm development and testing. If fruit were removed, comparison between different finding algorithms was not possible. An estimation exercise (by the author's assessment) of potential visual fruit appearance after 'removing' occluding fruit, was conducted. For this, the author observed each analysed image looking at

the unidentified fruit. A judgement was made as to whether the unidentified fruit would be identified if the surrounding fruit were removed. An experienced kiwifruit horticulturist was also asked if they thought the appearance of the initially unidentified fruit would be similar to those surrounding it, if the surrounding identified fruit were removed. This helped validate the author's judgement process and increased positive machine vision identification to 83.6%.

Table 2: Estimated fruit identification results using machine vision. Achieved by using a theoretical fruit removal process.

Criteria	Count	Percentage
Positively identified	686	83.56%
False positive	15	1.83%
Unidentified	132	16.08%
False negatives	3	0.37%

Higher positive identification rates (than those reported) were achieved during development by altering the thresholds and other requirements of each analysis step. However, false positive results also increased. Commanding robotic extraction of false positive fruit could entangle the hand in canopy environment, or damage surrounding fruit. Therefore, a solution balancing positive to false positive fruit identification was sought.

Comparing the achieved identification rates (first pass machine vision fruit identification) to the estimated results (see above) increased positive identification rates by 14%. This indicates that most fruit are positioned in clusters, partially or even fully occluding some fruit from detection. The standard deviations of the identification rates are presented in Figure 5-15. These show greater consistency of identification when fruit-fruit occlusion occurrences are removed (from the estimated identification rate). The majority of remaining unidentified fruit result from inconsistent lighting across the captured images (unidentified fruit are shown in Figure 5-14). This alters the fruits visual appearance beyond the algorithm's identification ability.

The machine vision fruit identification process currently requires between 15.2 – 28.7 seconds per pair of stereo images (using an Intel Core 2 Duo – E6750 processor at 2.66GHz and 960 x 720 image resolution). The algorithms were non-compiled and written in Visual Basic 6 (VB6) and also included drawing analysis information to the screen. Testing was focused on fruit identification rates not processing time. While VB6 has some advantages as a development environment, it is not suitable for end application. Significant performance increase will be gained

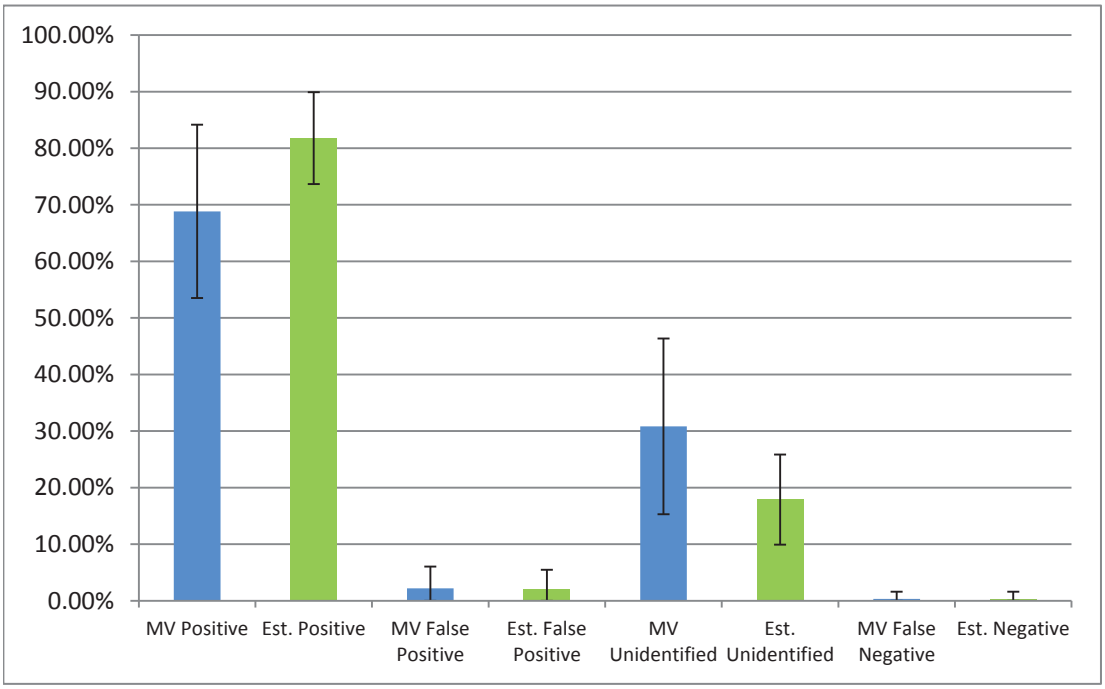


Figure 5-15: Average fruit identification results - Machine vision (MV, Table 1) and Estimated rates (Est., Table 2) accounting for fruit removal with standard deviation bars shown.

from translating the developed algorithms into a more suitable language (one of the C family of languages). Translation will potentially reduce computation time to 0.15 – 0.28 seconds for fruit identification per stereo pair of images. This estimate is based on the improvement achieved with the translation of other image analysis algorithms into C++. A multithreaded implementation, devoting a thread to each pair of stereo cameras would better use individual processing cores of modern multicore processors (preferably at least a four-core processor, one core per stereo camera pair). This would allow simultaneous analysis of each pair of stereo images, sharing the processing burden. For the analysis, it was noted that processing time is dependent upon the number of fruit identified. During analysis, approximately 92% of time is spent finding the fruit's location in the second stereo camera. This is due to the significant number of comparisons required to match the complete blossom end template at every possible pixel position in the second stereo image.

If desired performance criteria (four fruit extracted per second, including identification time) is not achieved from algorithm translation to C (described above) a CUDA (NVidia or similar) parallel 'co-processing' architecture development using parallel Graphics Processing Unit (GPU) could be investigated. Lengyel, et al. (2011) report considerable speed increases with CUDA based machine vision analysis, up to 2.7x for some algorithms. Performance increases will strongly depend upon the implemented algorithms ability to be processed in parallel.

Consequently, the required performance criteria (speed and identification rate) can be met by using existing modern hardware.

5.2.3 Stereoscopic Depth Perception

Stereopsis uses two images of a scene, observering it from slightly differenent vantage points, to determine depth. The disparity in horizontal position (of an object in the scene) between the two stereo cameras (horizontally offset), directly correlates to the object's distance. Animals (including humans) percieve depth in a similar manner. Our brains analyse the complex visual information, subconsciously perceiving depth. Animals effectively perceive relative distances, while stereo machine vision achieves absolute distance measurement. The process of stereo visual perception was first described by Charles Wheaton in 1838. It is now recognised as "one of the most active research areas in computer vision" (Scharstein and Szeliski, 2002).

The following stereo camera analysis and derivation of distance from the two images was provided by Dr Rory Flemmer, Massey University. This method was provided for the AKH.

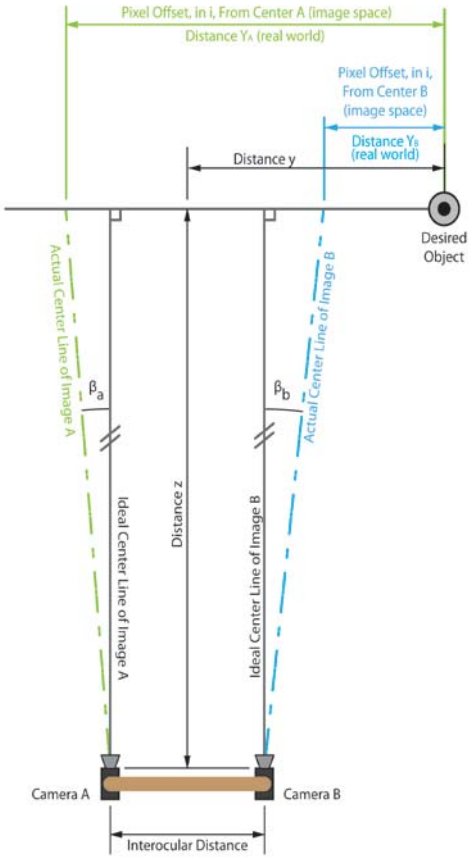


Figure 5-16: Stereo camera geometry for canonical camera configuration

Picture elements (pixels) represent an area's averaged colour value. The closer the object is to the camera the smaller the area the pixel represents. This characteristic, defined as metres of coverage (horizontally or vertically on the sensor) of each pixel, for every metre an object is from the camera, is assigned the term *mpppm* (meters per pixel per metre). Experimentation determined that the sensors pixels are approximately square (*mpppm* horizontally or vertically are equal, a direct characteristic from the sensors fabrication), allowing simplified calculation. Relating this back to the standard pinhole camera model, the term *mpppm* is equivalent to *1/f* in pixels (Figure 5-17). In this model, *C* is the camera's centre, P_P is the principal point and *P* is the

projected height in pixels of the point on the image plane. Therefore by rearranging *P* to show *mpppm* is equivalent to *1/f*,

$$Y = \frac{PZ}{f} \tag{6}$$

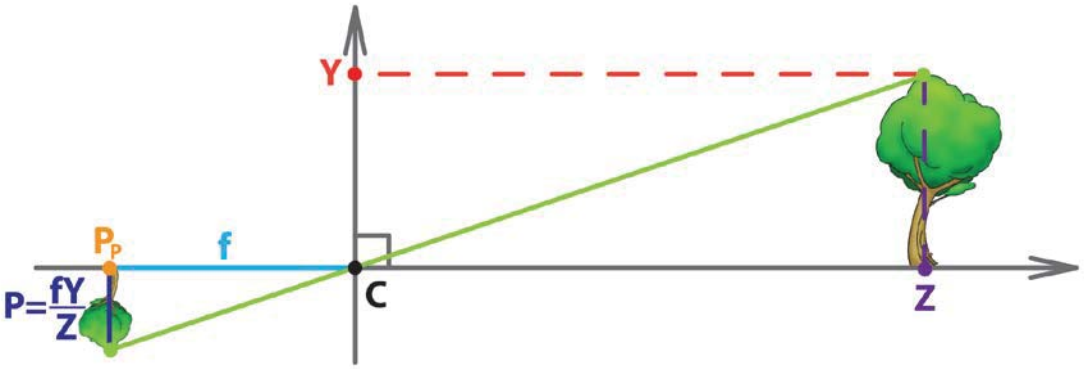


Figure 5-17: Pinhole camera geometry looking at the Z-Y plane

In stereopsis, the pair of stereo cameras are physically offset by a defined displacement, *IOD* (interocular distance, or distance between lens centres, often referred to as baseline distance). The horizontal offset allows depth perception from the horizontal disparity between stereo images. Vertical misalignment offsets do not affect the system's ability to accurately determine depth (assuming only minimal camera rotation about the optical axis achieved with mechanical housing). Desired ideal camera alignment (Figure 5-16) is difficult to achieve and in reality, for the accuracy required, compensation for any camera misalignment or offsets are required. Offsets are determined via a calibration process (section 5.2.3.1) enabling their compensation.

Once object depth is known, its three-dimensional position can be determined from its image position (with respect to one camera). Calculating the objects horizontal and vertical pixel offsets, from centre (i = 480, j = 360) and substituting into (7) and (8) gives the three-space position of the object, with respect to the camera.

$$y = z(mpppm \times Pixels_i) \tag{7}$$

$$x = z(mpppm \times Pixels_j) \tag{8}$$

The stereopsis depth analysis presented is an alternative to standard epipolar geometry. Hartley and Zisserman (2010) and Cyganek and Siebert (2009) (amongst others) provide more comprehensive epipolar geometry analysis and derivation. However, this analysis normally relates

to cameras of any pose, or even for two images taken from a single camera, from two viewing positions. Simplifications are possible for the canonical camera configuration used. Further simplification was possible for the level of accuracy required. This is because the optical axes of both cameras are close to parallel (β_a and β_b) < ±2°, Figure 5-16) and minimal rotation about the optical axes are present (the horizontal axes of both images are close to parallel). Therefore, for small rotations from the cameras ideal optical axis (as shown in Figure 5-16), the difference between the objects perpendicular distance (from the actual optical axis) and the distance derived from the ideal optical axis is small. This results in a small and acceptable distance error. This applies equally for small rotations about the optical axis. The presented simplified stereo analysis method achieved acceptable depth accuracy over the entire operating area.

From Figure 5-16, y is the actual distance of an object from the centre of an image, taking into account objects depth. Balancing distances assuming perfect camera alignment:

$$y_A - y_B = IOD \tag{9}$$

Accounting for camera misalignment:

$$y_A - y_B - z\tan\beta_a - z\tan\beta_b = IOD \tag{10}$$

Substituting into (10):

$$z(mpppm \times Pixels_A) - z(mpppm \times Pixels_B) - z\tan\beta_A - z\tan\beta_B = IOD \tag{11}$$

Rearranging for common term, z:

$$z[(mpppm_A \times Pixels_A) - (mpppm_B \times Pixels_B) - (\tan\beta_a + \tan\beta_b)] = IOD \tag{12}$$

Rearranging for depth, z:

$$z = \frac{IOD}{[(mpppm_A \times Pixels_A) - (mpppm_B \times Pixels_B) - (\tan\beta_a + \tan\beta_b)]} \tag{13}$$

Simplifying, assuming identical camera characteristics ($mpppm_A = mpppm_B$):

$$Depth = \frac{IOD}{(mpppm \times Disparity - AlignmentErrors)} \tag{14}$$

Logitech Webcam Pro 9000 cameras are used and configured to a resolution of 960 x 720 (*i, j*). Their lens provides a viewing angle of 60°. These cameras have a spatial resolution of 640 pixels per metre at the maximum operating range of 1.3m. The stereo pair has an IOD of 150mm giving a depth resolution of 13.4mm at the maximum operating range. However, most fruit will be less than one metre from the cameras, where a depth resolution of 7.9mm is achievable.

Accuracy, over a given range is proportional to IOD. As the objects distance increase to infinity, image disparity trends to zero (Figure 5-18 and Figure 5-19). Depth accuracy decreases as small errors in disparity produce large distance errors. IOD is determined by three key elements:

1. Depth perception accuracy, plot Figure 5-19. As depth accuracy is proportional to the slope of the curve it is desirable to operate in the near linear section approaching the horizontal asymptote where distance = 0. This section produces increased distance resolution for a given disparity. Increasing IOD increases accuracy for further distances.
2. Corresponding viewing area. As IOD increases, the overlap between images decreases. Observed objects(s) must be visible in both cameras and image area coverage defined (grey area Figure 5-18).
3. Increased IOD changes the angle of view between the two images making it harder to match the corresponding object in the two views.

An IOD of 150mm is used for locating fruit. This IOD provides a balance between positioning precision and operational range, ensuring optimal harvest areas are identified, increasing system efficiency.

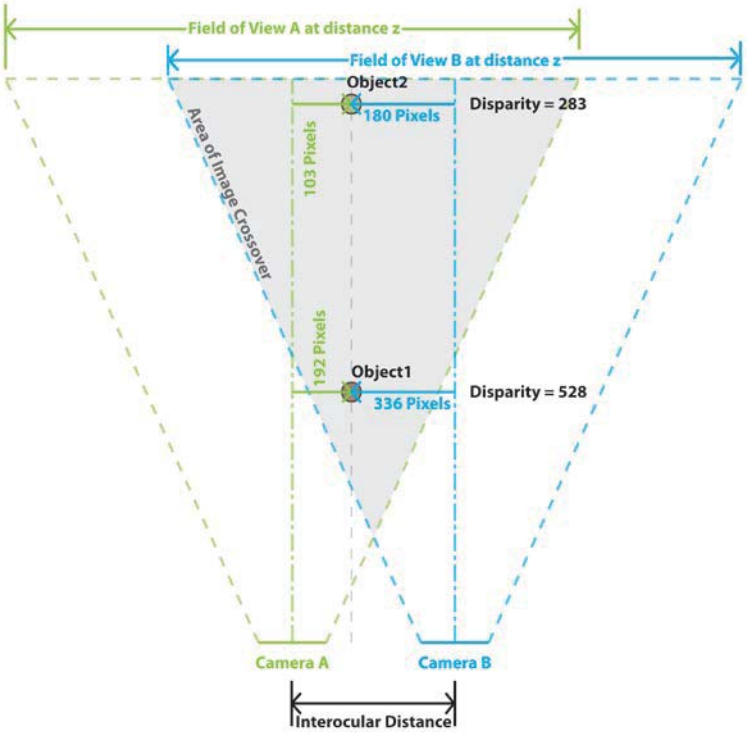


Figure 5-18: Interocular Distance Interaction, assuming idealised camera alignment

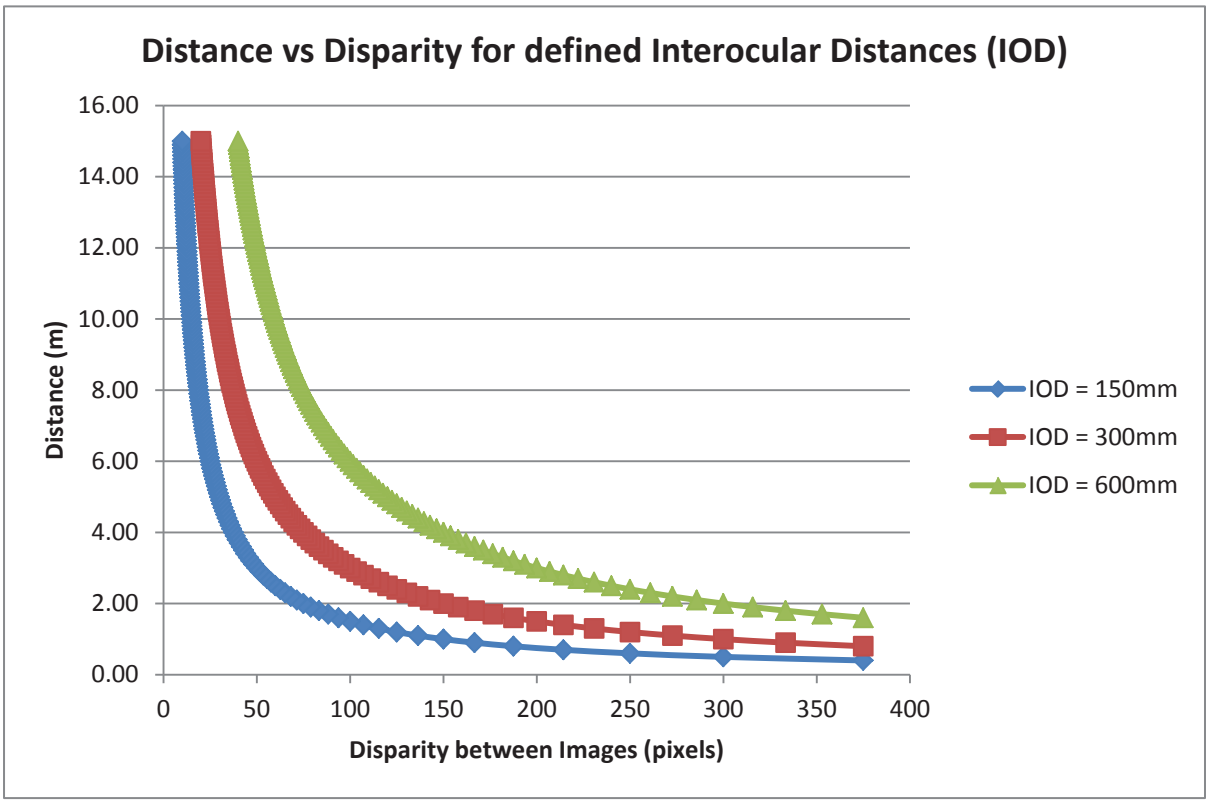


Figure 5-19: Distance/Disparity relationship for a given Interocular Distance (assuming perfect camera alignment)

5.2.3.1 Stereoscopic Camera Calibration

To achieve accurate stereoscopic object localisation, both intrinsic and extrinsic system characteristics are required. Manual measurement of translation and rotations between the two cameras does not deliver the accuracy required. It does however provide a suitable starting estimate for automated calibration. The system's accuracy is limited by how accurately these intrinsic and extrinsic characteristic are determined, along with the cameras baseline distance and camera resolution. In addition, the stereopsis equation provided by Dr Flemmer does not account for some camera rotations (discussed in section 5.2.3), further limiting possible distance measurement accuracy. The system's accuracy decreases as the object's distance increases. Stereo camera calibration was carried out by the author and was performed as part of the hand-eye coordination process that is explained in section 5.3.

To calibrate the cameras, a calibration hand attaches to the robotic arm, replacing the harvesting hand (Figure 5-25, page 87). The calibration ring moves through 350 known positions. This provides a large data set over the robotic arm's entire operational area. At each arm position, image analysis accurately determines the calibration ring's image position (in both stereo cameras, section 5.3.2) and records it against the ring position as set by the arm.

A simplex regression uses the known values of:

1. Distance to object (measured or known)
2. Object position, Image A (in pixels)
3. Object position, Image B (in pixels)

To optimise the unknown system values (see section 5.2.3 for variables):

1. $mpppm_A$
2. $mpppm_B$
3. IOD
4. $\gamma = (\tan \beta_a + \tan \beta_b)$, camera misalignment constant

This simplex regression calculates the theoretical position of the hand using the estimated variables being optimised for every image. Each of these is compared to the actual hand position and an error is calculated to obtain the sum of squares error for the entire data set. The simplex

regression alters the system variables to minimise the sum of squares error and produced the system characteristics. Outliers are inevitable within the data set and are removed during optimisation. This is achieved by optimising the system until the sum of squares error shows little change for progressive variable changes. Next, the variance of each error is assessed and those outlier points are removed. The regression process is continued to better optimise the system characteristics. This process is repeated to remove any outliers. These outliers are from incorrectly locating the hand in the images due to the calibration ring being partially occluded from both cameras.

This was a customised calibration process using the robotic arms (as explained in section 5.3) during hand-eye coordination. This is to achieve accurate depth perception with the cameras in their mounted operating position and removes any misalignment/distortion issues that might arise from fixing the cameras to their frame. Calibrating the cameras in this manner also allows for on-orchard calibration if the camera modules ever need replacing during operation.

If local minima exist, the simplex regression can erroneously optimise system characteristics. 'Best estimate' starting conditions are measured to help reduce a false alignment result.

1. mpppm values are obtained by measuring, in pixels, an known objects length, at a known distance.
2. IOD is measured
3. γ is set to 0, as this is a small offset.

5.2.4 Stereoscopic Depth Perception – Testing and Results

Depth accuracy is determined during calibration. A comparison between actual and machine vision perceived object positions is performed during calibration providing system accuracy. The same images are used for both calibration and performance evaluation. However, once the system is calibrated the performance is measured. This is achieved by rounding the sub-pixel accurate calibration ring image position to pixel level accuracy. This provides a performance value that represents real-world performance for fruit localisation as they are only positioned to pixel level accuracy.

The stereo cameras used for fruit finding (IOD of approximately 150mm, operating over a range of 0.5–1.3m) achieved an absolute average distance error of 2.7mm with standard deviation of 1.2mm.

During calibration, some abnormally high errors were observed and were eliminated from the calibration regression. The high errors were attributed to errors in determining the object position in the stereo image pair. Any positioning error (when matching the object's image position in both stereo images) will produce a distance measurement error. This error arises from the robotic arm's structure blocking complete viewing of the calibration ring. When this occurs the calibration ring is sometimes incorrectly located (see Figure 5-25, page 87). Figure 5-20 contains a plot of distance error with respect to object distance, for disparity errors of 1, 2 and 3 pixels. A one pixel positioning error of an object at 500mm (from the cameras) produces a distance error of 2.2mm. This error increases to 14.3mm at an object distance of 1.3m. Although calibrated over a large distance range (for accurate determination of system rotation and translation offsets, see section 5.2.4) the majority of harvestable fruit is within a one-metre range of the cameras, providing an 7.9mm maximum fruit position resolution (with no disparity error).

The calibrated system's accuracy is greater than expected. Both intrinsic and extrinsic camera variables were more accurately found using sub-pixel image accuracy (for object positioning). This resulted in better determination of the system variables, increasing system accuracy (when compared to pixel accurate calibration).

5.2.5 Conclusions: Stereo Machine Vision

The achieved stereoscopic depth accuracy is insufficient for fruit localisation over the entire operational distance range. Although the calibration process indicated a suitable accuracy was achieved, the stereo cameras depth accuracy at maximum range does not meet the accuracy requirements. Increased system accuracy will be achieved by upgrading the stereo analysis algorithms (provided by Dr Flemmer) to account for all camera rotations. This will be achieved using the standard camera pinhole model and analysis for a canonical stereo camera setup (Figure 5-17). In this model, a compromise between accuracy and field of view must be sought. Increasing the interocular distance between the stereo cameras to achieve better accuracy over the desirable distance range decreases the common viewing area between them (see Figure 5-18), especially at closer ranges. If increased accuracy is required, four main adaptations are feasible:

1. Increase camera resolution. However, image acquisition rates may decrease and object identification algorithms would require modification.

2. Sub-pixel object positioning: as used in the calibration process would effectively increase the distance resolution without requiring hardware to be modified. The process of matching the object between the two stereo images to sub-pixel accuracy is significantly more complex. This is due to the organic softened fruit edges.

3. Converging the two cameras optical axes ('cross-eyed'): in addition to an increased inter-ocular distance, this will increase the common viewing area between the two cameras. This offsets some disadvantages of an increased camera interocular distance. For close fruit, the increased IOD may increase object occlusions between stereo cameras, as its viewing position is shifted further from the reference camera. In addition, this camera configuration requires more complex calibration to correct for perspective distortion.

4. Tri-nocluar (three camera) vision (could use more cameras as required): could provide dual interocular distances, where one primary reference camera is used for object detection and the other two provided their own individual stereo image. It is analogous to two pairs of stereo cameras with different interocular distances without the additional hardware and alignment requirements.

Each implementation possibility exhibits advantages and disadvantages that require assessment.

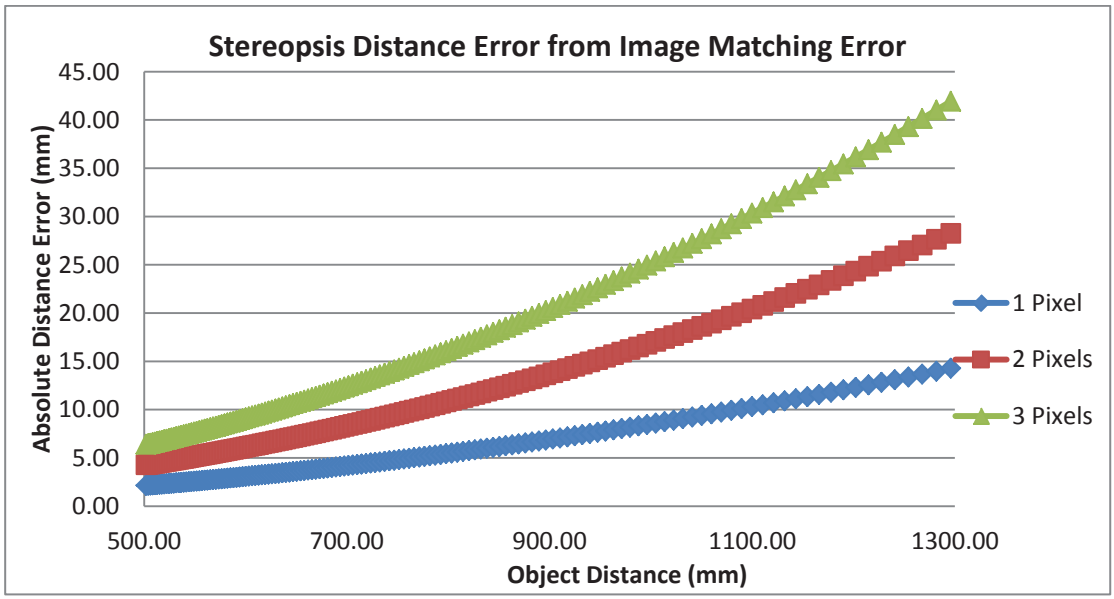


Figure 5-20: Distance determination error from an error in matching the objects position in the second stereo image. Interocular distance of approximately 150mm.

In addition to insufficient accuracy, slow stereo matching (between the images) resulted. Implementing complete stereo camera epipolar geometry (Cyganek and Siebert, 2009) (also derived below) will not only increase the accuracy, but also decrease the time taken to locate the fruit. With the current analysis, fruit position is determined in the first camera and searched for in the second, over a 2D search area. A full epipolar geometry implementation provides a projected epipolar line across the second stereo image. The identified fruit will be situated on this line. Currently, approximately 92% of the fruit localisation time is required for stereo matching. Stereo matching could be achieved in 1/20th the current time (scanning 1 line rather than 20) with the proposed modification. The time required to locate the fruit would potentially decrease to 1.9 – 3.6 seconds (approximate) from the current setup (from 15.2 – 28.7 seconds).

5.3 Hand-Eye Coordination

Mathematical alignment of the camera and arm coordinate spaces is required to position the robotic arm relative to a visually identified object.

Accurate 'hand-eye' coordination results from mathematically modelling the arm's coordinate system relative to the visual coordinate system. Determining the relationships (offsets and rotations of one origin with respect to the other) allows a mathematical translation of a point found in the visual coordinate system into the arm's coordinate system. A calibration hand replaces the harvesting hand. The calibration ring's centre (Figure 5-21) corresponds to the harvesting hand's centre. It is also at the required fruit base position for extraction. Therefore, when the fruit's base is localised (for arm extraction), the correct height will be determined for the harvesting hand.

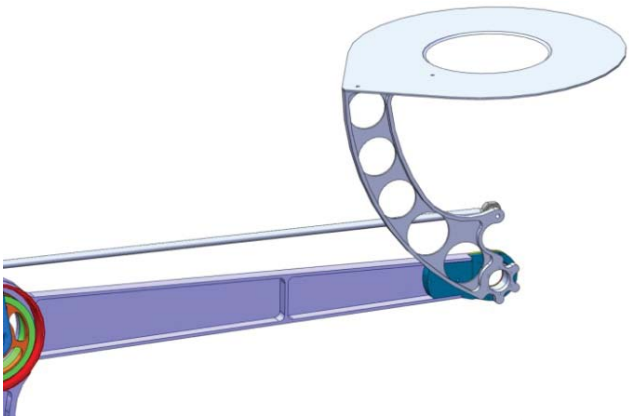


Figure 5-21: Calibration ring CAD image used for camera/robotic arm coordinate space alignment

Alignment of only one stereo camera to arm coordinate space is required. This is due to the stereo positioning method provided by Dr Flemmer. Using standard epipolar geometry both cameras provide the fruit three-dimensional position and therefore the pair of cameras requires alignment. The fruit is located in three-dimensions with one camera, once the depth is known. The object's pixel distances ((i, j), from centre) can be scaled using (7) and (8) to provide an actual

three-dimensional position after determining distance. Rotating and translating the perceived position (by the rotation and translation offsets between camera and arm coordinate spaces) produces the object's position in the arm's coordinate space.

5.3.1 Calibrating Hand-Eye Coordination

Translational (*x*, *y* and *z*, Figure 5-22) and rotational (pan, tilt and rotation) offsets must be determined to achieve accurate hand-eye coordination. A simplex regression determines these system offsets. Three hundred and fifty calibration hand ring (Figure 5-21) images (as used in section 5.2.3.1) provide arm positions over the complete arm-vision range. The arm positions are known and recorded against each captured image. This allows a comparison between the perceived arm position (machine vision) and the actual arm position. The difference between actual arm position and the calculated estimated object position is used to find an alignment error. The estimated image position is calculated by taking the actual arm position and mapping it into the cameras image space using the rotation and translation variables. Alignment error is calculated as the difference between this estimated image position and actual calibration ring image position. Calculating this for every observed arm position delivers an average alignment error over the entire operational space and a total sum of squares error is calculated. The simplex regression alters the system variables (rotations and translations) to minimise the sum of squares alignment error. This provides the necessary coordination between visual and arm systems.

To achieve this, mathematical alignment of the geometric spaces is required. The alignment needs to account for three-space translation, $t=[X,Y,Z]^T$, and rotation offsets ($R_x(\rho)$, $R_y(\theta)$, $R_z(\sigma)$) between the coordinate spaces. The key aspect of this is mapping a camera derived position into robotic arm coordinates.

Rotation Matrices for camera alignment

$$R_x(\rho)=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos(\rho) & -\sin(\rho)\\ 0 & \sin(\rho) & \cos(\rho)\end{bmatrix} \tag{15}$$

$$R_y(\theta)=\begin{bmatrix}\cos(\theta) & 0 & \sin(\theta)\\ 0 & 1 & 0\\ -\sin(\theta) & 0 & \cos(\theta)\end{bmatrix} \tag{16}$$

$$R_z(\sigma) = \begin{bmatrix} \cos(\sigma) & -\sin(\sigma) & 0 \\ \sin(\sigma) & \cos(\sigma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{17}$$

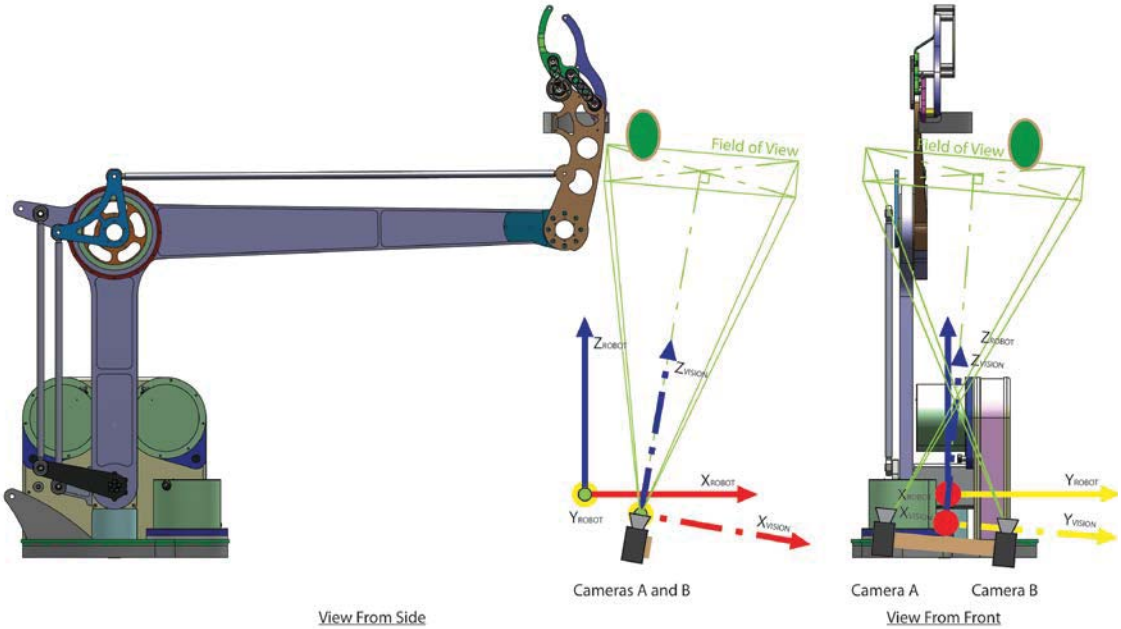


Figure 5-22: Hand-Eye coordination: Camera to robot coordinate space alignment

The relationship between robotic arm coordinates, $A=[X,Y,Z]^T$, and camera coordinates, $c=[x,y,z]^T$, is described as,

$$A = R_{Ac}c + t_{Ac} \tag{18}$$

where, R_{Ac} is the resultant 3 x 3 orthogonal rotation matrix,

$$R_{Lc} = R_x(\rho) \times R_y(\theta) \times R_z(\sigma) \tag{19}$$

Equation (18) provides a position in robotic arm coordinates from a translated and rotated stereo positioned object. This allows the robotic arms to be positioned relative to camera derived coordinates.

5.3.2 Calibration Ring Position

Accurate hand-eye coordination requires accurately measuring the hand ring image position. Template matching, sub-pixel edge determination and ellipse fitting are used to achieve this.

A circular template formed with two concentric rings (corresponding to the outer and inner edges of the calibration ring), is scanned over the captured calibration image. The template's concentric ring sizes are calculated from the arm's height (distance and image mpppm are known). This delivers a tailored template for every arm position imaged. The template is used to search for complete black rings, corresponding to the black calibration ring. The image position with the most black template points is used. Template matching provides an approximate calibration hand ring central position (see red circle Figure 5-25). This position requires refinement.

Next, the inner and outer edge points of the calibration ring are found to sub-pixel accuracy. This is achieved by searching radially from the template found centre position and looking for brightness threshold crossings indicating an edge (corresponding to the calibration hand ring edges). By performing this search over small incremental angles the entire ring edges are found. These edge positions are refined to sub-pixel accuracy using linear interpolation (Figure 5-23). This method finds where the edge brightness threshold is crossed between two consecutive pixels, relative to the distance *d* between them (as they could be diagonally offset). The edge positions are stored for further analysis.

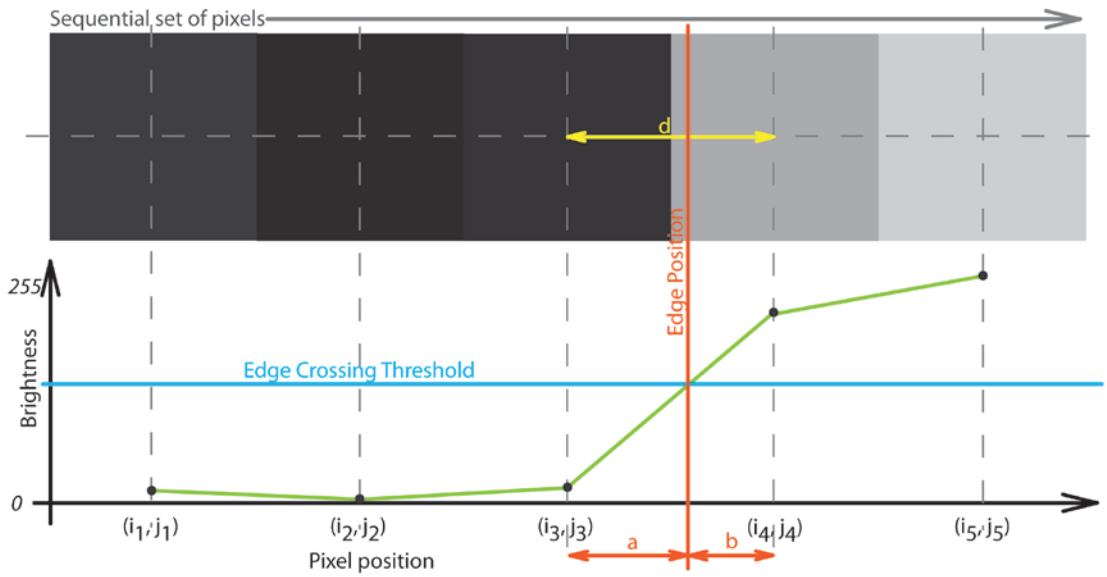


Figure 5-23: Linear interpolation to achieve sub-pixel accurate edge positioning

False positive edge points can be detected when the arm occludes some of the ring. These unwanted edge points are removed during ellipse fitting (see below). The blue points (Figure 5-25) represent the determined calibration hand ring edges, where the removed false positive points have a yellow circle around them.

The round calibration hand ring is mainly viewed at an angle. This results in the rings appearing (in the images) as ellipses. Although only slightly distorted from circular in projection, circular data matching (using a circular Hough transform, or similar) to the edge points is inaccurate. To accurately determine both ring's centre, two ellipses are fitted to the located edge data using a simplex regression. The regression optimises the sum of squares error of the radial distance to the identified edge points when compared to their theoretical elliptical radial distance. In this process, both the inner and outer ellipses (calibration ring edges) have the same centre and axis rotation, but are scaled in size.

The regression process iteratively alters (refer to Figure 5-24):

1. Ellipse centre position in *i* and *j*, to sub-pixel accuracy.
2. Major axis *a* of outer ellipse
3. Major axis *a* of inner ellipse
4. Minor axis *b* of outer ellipse
5. Minor axis *b* of inner ellipse
6. Ellipse rotation, φ.

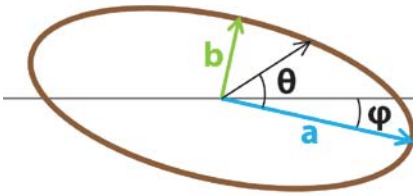


Figure 5-24: Ellipse nomenclature

Starting at the template defined central position, each identified edge points angle and radius are calculated. Using the rotation, φ, major axis, *a*, minor axis, *b,* (specified in the regressions estimate) and the angle to the point of interest, θ, a theoretical radius can be calculated (from the simplex regression estimate). The theoretical radius (from regression) is compared against

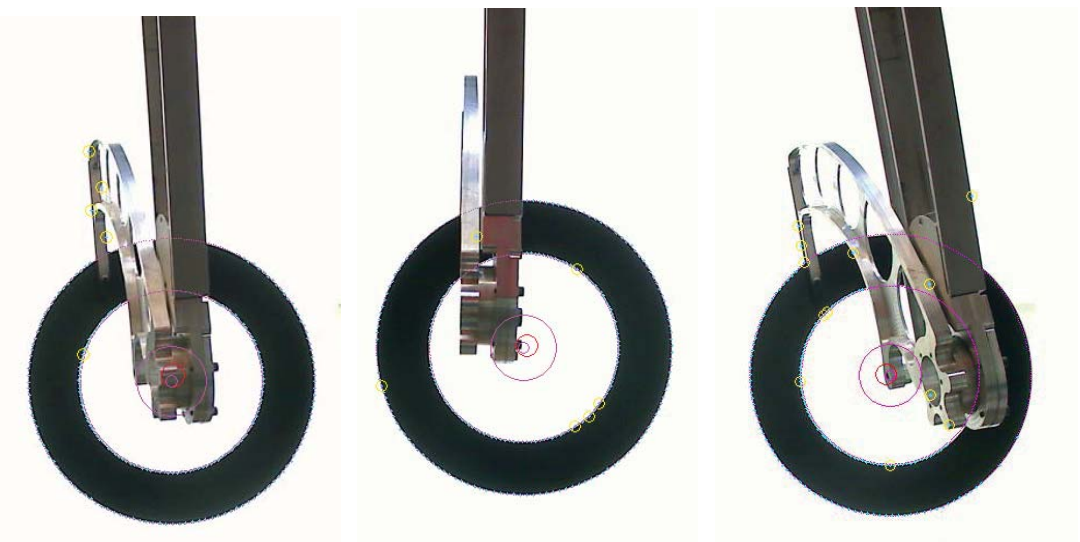


Figure 5-25: Hand-eye coordination calibration disc analysed images

the measured radius, providing an error between the estimated and actual positions. This process is applied over all inner and outer edges points, calculating a total dual ellipse 'fit' error to the edge point data sets.

In polar coordinates, where the ellipses origin is central and rotation, φ (Figure 5-24), is present, the radius, $r(\theta)$, can be described:

$$r(\theta) = \frac{ab}{\sqrt{(b\cos(\theta-\varphi))^2 + (a\sin(\theta-\varphi))^2}} \tag{20}$$

Outlier image data points (not correlating to the determined dual ellipse 'fit'), affect positioning accuracy. After five hundred and later one thousand regressions iterations, outliers with large radius errors are removed from the regression (shown highlighted in yellow circles, Figure 5-25) providing higher positioning precision.

5.3.3 Reconciliation of Multiple Coordinate Spaces

Multiple arms are required to achieve the necessary harvesting rate. However, operating these arms close together can lead to collisions or inefficient operation if one or more arms have less work and are waiting. To help overcome this, the eight coordinate spaces (four arm (R_n) and four camera (C_n)) require reconciliation into a single coordinate space. This allows harvest cycle optimisation and arm/arm collision avoidance strategies to be implemented. Harvest rate opti-

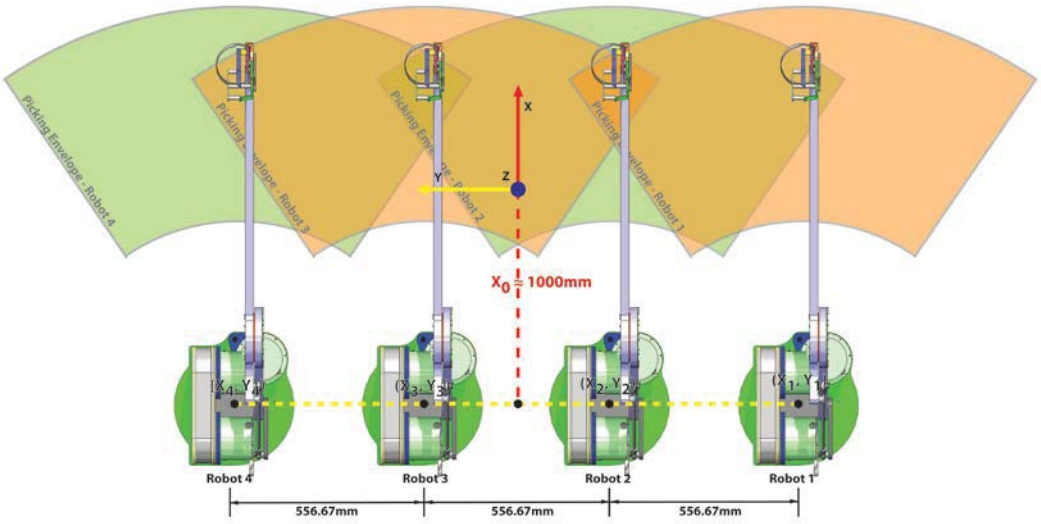


Figure 5-26: Pan view of robotic arms'. Note the overlapping operational envelope

misation is achieved by operating an arm within an adjacent arm's camera space. This requires all identified fruit to be located within a single coordinate space, allowing harvesting demand to be shared between arms. Mathematical alignment determines the mapping between coordinate spaces, allowing a unified coordinate system.

A method to accurately reconcile the coordinate spaces using stereo machine vision is described. The coordinates of robotic arm two (R_2, Figure 5-26) are defined as the reference coordinate system. All other coordinate systems are defined relative to this arm's coordinates. A middle arm (R_2 or R_3) was selected as the reference coordinate space to minimise error propagation across sequential alignments.

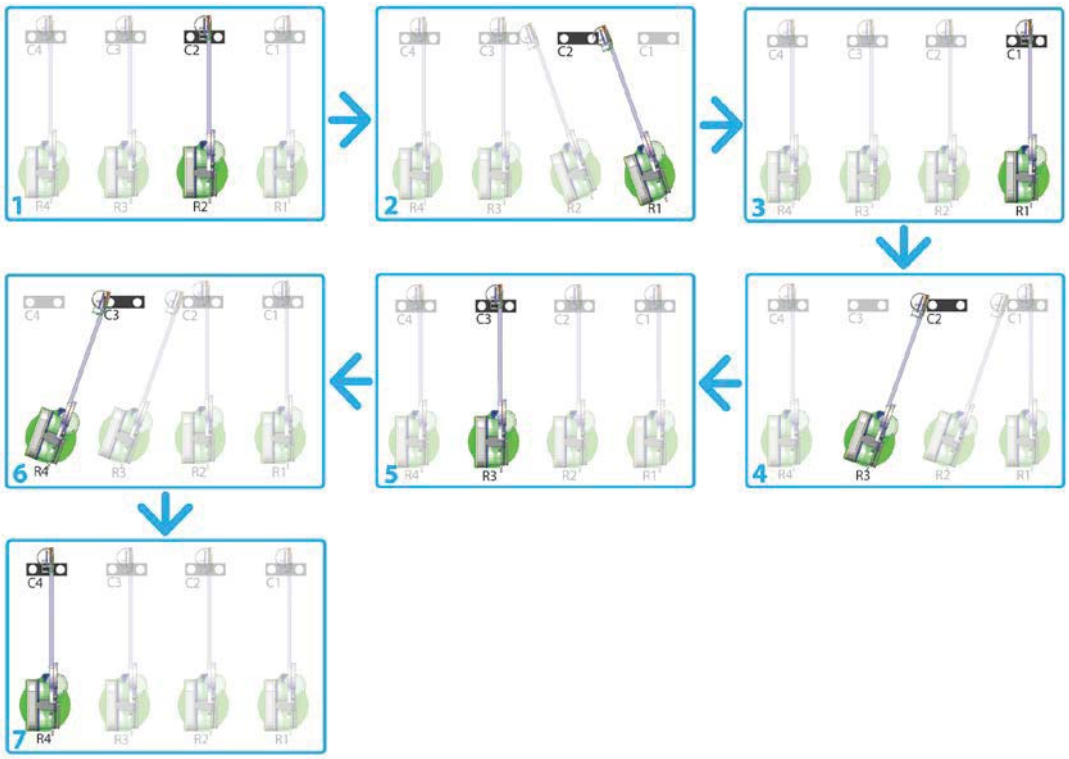


Figure 5-27: Directed graphic of the process used to reconcile the robotic arm coordinate spaces (viewed from above)

Coordinate space reconciliation process:

1. Hand-eye coordination is determined for R_2, see section 5.2.4, aligning $R_2 - C_2$.
2. Hand-eye coordination is determined for R_1, using R_2 cameras. Aligning $R_1 - C_2$, mapping R_1 into the reference coordinate system (R_2).
3. Hand-eye coordination is determined for $R_1 - C_1$, aligning C_1 into the reference coordinate space.

4. Hand-eye coordination is determined for R_3, using R_2 cameras (C_2). Aligning $R_3 - C_2$, mapping R_3 into the reference coordinate system (R_2).
5. Hand-eye coordination is determined for $R_3 - C_3$, aligning C_3 into the reference coordinate space.
6. Hand-eye coordination is determined for $R_4 - C_3$, aligning R_4 into the reference coordinate space (R_2).
7. Hand-eye coordination for $R_4 - C_4$ completes the alignment of all coordinate spaces.

5.3.3.1 Coordinate Space Reconciliation Calibration

The regression process for aligning multiple coordinate systems follows the hand-eye coordination regression (section 5.2.4). This is performed using camera and arm combinations described in section 5.3.3 above. Camera position and offsets are estimated by measurement, providing initial variables for optimisation and are accurately determined by simplex regression.

5.4 Testing and Results

Accurately transforming fruit position (from machine vision to arm coordinate space) depends upon accurate resolution of the rotation and translation variables between the two respective spaces.

Three hundred and fifty calibration hand stereo images (section 5.2.4) were captured. A data set of hand position was recorded referencing the captured images. The stereo images were analysed to accurately locate the calibration hand (complete methodology, section 5.2.4) within the images. The hand's image position was recorded and referenced against the actual hand position. Recording both actual and perceived hand positions allowed the rotation and translation variables between both coordinate spaces to be determined by regression. The mathematical alignment provided an accuracy of ±3.6mm (average in three-dimensional space) with a 1.1mm standard deviation, over the full operational envelope. Capturing data over the complete arm's operational range (in horizontal and vertical planes) ensures rotation and translation offsets are accurately determined, as any variation is accentuated at the hand's furthest positions from the origin.

After five hundred regression iterations, the regression algorithm identified and removed any alignment outlier errors (corresponding to incorrectly locating of the calibration hand in the stereo images). Sixty-four stereo images were automatically removed and stored (for manual miss positioning confirmation to assess the algorithms ability to determine outliers). All sixty-four removed outliers were manually confirmed as incorrect and excluded from the alignment process. Poor identification (in the excluded analysed images) resulted from the arm occluding parts of the hand calibration ring, providing a large number of incorrect edge points.

5.5 Section Conclusions

A substantial advancement to fruit identification and localisation has been achieved. This system is accurate enough to support commercial operation and viability. However, this accuracy was not achieved over the system's full range. In addition, the current identification time limits viable operation. Both of these issues will be recitified by using the standard canonical camera analysis and pinhole camera model. By accurately mapping all camera rotations higher localisation accuracies will be achieved. This analysis method will also reduce object correspondance time between the two images as a smaller search area will be used. In addition to the analysis modification, altering the software to a more efficient language (e.g. C) will provide significant performace gains (refer to sections 5.2.2 and 5.2.4). Environmental variabilities affecting fruit identification require further development to ensure consistant and viabile operation. This includes integration of high intensity ligting (conceviably large, high output LED arrays) to saturate the canopy with light and reduce the effects of natural lighting. Natural white light (5000K - white) can be achieved and have an intensity similar to a camera flash. Operation on-demand will minimise power consumption.

Faster harvesting cycle speeds have been achieved (in comparison to other developments, section 2) by adapting remotely mounted stereoscopic cameras. The additional alignment complexity does not detract from the real operational benefits delivered. This systems allows conitinual fruit identification during fruit harvesting and does not require the time-consuming fruit 'hunting' process used in the 'eye-in-hand' configuration.

During on-orchard testing the arms occasionally entangled or collided with the canopy and supporting structures (discussed further in section 6.2). This temporarily halted operation, to allow the system to recover. Additional development to identify kiwifruit canes, supporting wires and bracing timber will help to achieve higher-level picking strategies and outcomes.

Overall system performance for identifying and locating fruit will achieve commercially viable fruit extraction rates as outlined in specifications. Although some specifications are not met, tangible modifications are outlined to achieve the necessary performance. These outcomes are significant when taking into account the operational complexities associated with the on-orchard environment.

Hanging kiwifruit, photo reproduced with permission from Plus Group Horticulture Ltd.

6. Robotic Manipulation

Most tactile tasks require direct physical human involvement. Undertaking intensive, mundane and repetitive tasks has encouraged people to develop mechanical aids to help. Robotic manipulators ('arms') are modern versions of these mechanical aids. The rapid advance of digital computing and control circuitry in the late twentieth century has advanced robotics miniaturisation and capability. This advance has been widely adopted in industrial applications.

Robotic arm deployment is widespread. Robots have improved productivity and reduced physical demands on workers. Benefits include: lower labour inputs/cost; ability to operate in dangerous environments; and increased efficiency and precision. As robotics progresses further, increased integration into 'normal' social, engineering and industrial applications will occur. This progression of artificial intelligence will further blur the dividing line between human and robot capability.

Robotic control and perception improvements have surpassed human capabilities in many areas. Robots are immune to fatigue or boredom and can repetitively perform tasks without error. Manufacturing industries have embraced these core benefits for the mass production of high volume, low cost products. Like the manufacturing industry, kiwifruit harvesting is highly repetitious with large labour requirements. Introducing robotic harvesting will reduce labour input and potentially add value to the sector.

Two custom arm designs are required for fruit extraction and handling. These were developed to minimise production costs and increase efficiency, as industrial arms are often expensive and not designed specifically for discrete fruit harvesting. In addition, control of commercial arms is normally limited. Custom controller development enabled task-specific optimisation to increase efficiency. This is beyond the capabilities of current standard commercial arms.

Both fruit extraction and handling present unique implementation challenges. The fruit harvesting arms requires:

1. Sufficient movement speed to harvest one fruit per second, where the average distance between consecutive fruit is approximately 150mm.
2. Three-dimensional positioning accuracy to ±2mm. Accuracy needs to be better than the targeted harvesting accuracy of ±5mm. This allows for some additional inaccuracy in the stereo localisation (section 5.2.3) and hand-eye coordination (section 5.3) systems
3. Sufficient output torque on each axis for acceleration to achieve the required movement speeds. The torque also needs to resist any fruit removal forces
4. Hand to grasp and remove fruit without inflicting damage
5. To reach into the canopy structure to reach higher positioned fruit

For robotic fruit extraction, the on-orchard growing systems and autonomous transportation platform (ATP - section 7) restrict the design. Canopy height, the need for multiple arms, and the mounting platform restrict allowable size. Although restricted, it needs to reach fruit and operate on and around the ATP's structures. An end-effector ('hand') (distally attached to the arm) is designed to mimic human hand harvesting. This enables fruit to be handled and extracted without being damaged. The hand is an integral component of the arm, designed to perform the specific harvesting function. Standard add-on hands are generally limited to specific industrial tasks and are not available for this application. This is due to their task specific design limiting flexibility. Kiwifruit are not like mass produced parts, they have significant variations in size, shape, canopy position and firmness that the hand needs to accommodate. In addition, kiwifruit require a specific and unique harvesting action. Hand requirements are discussed further in section 6.2.2 - End-Effector Development.

An average harvesting rate of four fruit per second is required for operational efficiency and commercial return. This necessity governed many design aspects of the arms. Four arms were considered necessary because of the large harvesting area, fruit density and performance criteria. Therefore, each arm has been designed to be capable of extracting an average of one fruit per second. To achieve the required performance in the operational environment an appropriate design is required. There are several standard robot configurations (and variations of) considered, including:

1. Articulated, or anthropomorphic, robots have a structure that can dynamically reach over or around obstructions. They can generally move to any position and orientation

within a working envelope and are widely used in industry. These tasks include welding, spraying, etc. where complex head angles and positions are required.

2. Cartesian, or gantry, robots have three principle prismatic axes (X, Y and Z) that are at right angles to each other. This design provides high rigidity, accuracy and repeatability; however, they have a large structure and struggle to reach around objects.
3. Cylindrical robots rotate about their primary axis and have additional prismatic axes. This design struggles to reach around objects due to the cylindrical primary axis.
4. SCARA (Selective Compliant Assembly Robot Arm) are commonly used for pick and place operations in relatively flat environments. The horizontal compliant axes limit this designs use around other structures (like adjacent arms) as they are more likely to interfere with each other (collisions).
5. Parallel robots: their design provides increased accuracy, stability and rigidity. They can also achieve higher cycle speeds as each manipulator member is lighter than a serial system. However, their base is large and the parallel link structure does not allow for canopy penetration or large operating envelopes.

An articulated robot design was adopted for the harvesting arm for its ability to move around obstacles, move to any position and orientation within the working envelope and its structure will not interfere with other systems/sensors.

The bin filling arm requires:

1. Linear motion of 100mm per second. The bin filler operates slower as it will likely be heavier than the harvesting arms and will need to shift several fruit at a time
2. Increased axis torque to support and move the increased arm mass
3. Three-dimensional positioning accuracy of ±10mm. This is because this arm will be placing items and not interacting with static objects.
4. Fruit cannot be dropped more than 80mm. This ensures that the fruit will not be damaged during the handling process
5. Shift four fruit per second into the bin to keep up with the harvesting rate
6. Reach the complete bin area. A storage bin measures 1200 x 1200 x 600mm and this system must completely fill it.
7. Operate within the confined ATP structures, under the kiwifruit canopy and on uneven ground

To meet these criteria an arm capable of layering fruit into a storage bin is required. Integrated fruit sensing will allow the arm to track fruit already in the bin and maintain a safe layering height. Current industrial systems are designed to operate in unrestricted spaces, not under kiwifruit canopies. The space restriction under the canopy was the biggest influence on the overall design.

The bin filling robot went through several conceptual ideas of how to best approach the design. A gravity based filling system was attractive for its mechanical simplicity, however it was not practical with the limited head height above the bin. A standard arm design (those listed above) was not practical as an internal conveyor system is required to move the fruit along each axis section, into the bin. A modified articulated arm design was identified as it could be adapted to incorporate an internal conveyor. From this concept two prismatic axes (pseudo Cartesian set-up) were considered to allow the articulated arm to reach the entire bin. After further development a design was achieved, this only required a single prismatic axis to move the articulated arm transversely across the bin. The articulated arm was capable of reaching out into the bin to layer fruit.

To support the requirements or both arms several design approaches were used. These are presented in section 6.1 and include methods to: move to a desired position; track the systems position; and remove positioning backlash. Although both systems are different, a technological cross-over exists within the respective motion control and drive systems. This supports a shared technology approach to shorten development time.

With a design for both arms conceptually decided on, Flemmer (2009) proposed using a arm drive system and provided skeleton control code. This technology was previously used in a SCARA fruit packing robot. His drive system was adapted for each design. It uses concentric drive shafts to transfer power to the articulating axes. Stepper motors through reduction spur gears drive each axis. The gear ratio provides the desired output torque and speed characteristics. This drive system is shown in Figure 6-5 (balloon (9) and (11)) and although modified, the concept is used in both the harvesting and bin filling arms. Development focused on a fruit extracting hand, and a bin filling system for harvesting and handling respectively as a drive system was provided.

6.1 Robotic Arm Design Features

The arms have several features to achieve the performance necessary to support commercial operation.

Stepper motors were selected to provide power to each arm axis for their low cost and functionality. Stepper motors are approximately 1/10th the cost of an equivalent standard brushless servomotor and harmonic gearbox combination. Although less expensive, stepper motors do not have internal position tracking encoders or gearboxes to deliver the required output torque/speed characteristics from the motors. Therefore, gearbox and position monitoring systems need to be incorporated into a stepper motor based arm design. These are incorporated for each arm in their respective sections (6.2 and 6.3).

The first requirement is to control the motors to achieve a desired arm position (section 6.1.1). Through better motor control (section 6.1.2), robotic arm motion can be achieved more efficiently than other control strategies like trapezoidal motion profiles. From the motor a reduction drive system is required. This reduction drive provides suitable torque to each axis using a smaller motor. It also increases positioning accuracy by reducing the output shafts motion for a given motor 'step'. Standard gearboxes were considered, however, they were considerably more expensive than gear or belt drive systems. Low backlash gearboxes were even more expensive. Spur gears were selected for this task in the harvesting arms as they provide better repeatability than belts as they do not flex, slip, extend, or stretch when torque is repetitively applied. The bin filling arm does not require high positioning accuracy, therefore toothed belts were sufficient in that application. Although spur gears present advantages, machining inaccuracies and tooth wear can introduce backlash into the reduction drive. This leads to positioning inaccuracies. A solution to this problem is presented in section 6.1.3 to minimise these effects. Standard module 1 spur gears are used for axis power transmission. These gears are made from AISI 3140 surface hardened alloy steel. The manufacturer recommended this tooth profile and material for its strength, wear resistance, service life and cost of manufacture. The tooth module is a standard size that is commonly used and therefore cost effective to manufacture. In addition, the smaller teeth provide smooth power transmission and sufficient gear meshing for the large 10:1 drive ratio.

With the motors providing suitable torque to each axis, the position of each of these axes requires tracking. Incremental or absolute encoders are generally used to achieve this require-

ment. However, these encoders are expensive for the update speeds required. The encoders would have to update at 10k Hz if every motor step were to be measured. Therefore, an accurate method of tracking each axes position has been developed using Hall effect sensors (section 6.1.4).

6.1.1 Stepping to Position

Stepper motors provide arm motion control. They are directed (by the controller) to perform small incremental steps in a given rotational direction. The output steps are geared to deliver a proportional arm member rotation. For example, the harvesting arms have motors configured to 3200 steps per revolution and use a 10:1 gear reduction. This configuration produces a 0.01125° output shaft rotation for every motor step.

A homing routine moves the arm axes to a known (home) position using on-board sensors. The step count is incremented or decremented dependant on motor direction. By tracking this step count for each axis (with respect to the known home position) the real-time arm position is calculated. A desired arm position (*x,y,z*) is translated into a step count for each axis by calculating their respective angles for each particular mechanical structure (Harvesting arm - section 6.2.4 and Bin filling arm – section 6.3.4). Each axis can be controlled proportionally by the number of steps required to move it from current to desired position. This produces linear point A-B movement (synchronised motion on all axes). Other custom arm motion is performed to optimise the harvesting cycle (section 6.2.5).

6.1.2 Stepper Motor Control

A novel motion trajectory solution was developed to achieve the required motion performance. This motion trajectory solution was provided by Dr Flemmer for use in the AKH. For stepper motors, the characteristic curve of breakout torque versus speed declines smoothly and strongly. This is maximum at rest and about 10% of maximum at 1 kHz. A linear operating profile was used to describe this characteristic while staying within the breakout torque envelope (Figure 6-1). This line intercepts the torque axis at the starting torque, *C*, with a slope (negative), *M*. The line must lie below the curve at every point. This ensures that the motor does not exceed its performance capability.

Thus,

$$T = Mn + C \tag{21}$$

where n is the number of steps per second and T is the pull-out torque.

Then, by Newton,

$$T = \frac{d^2\omega}{dt^2} I \tag{22}$$

where I is the rotational inertia of the system experienced at the motor shaft and ω is the rotational angle of the motor.

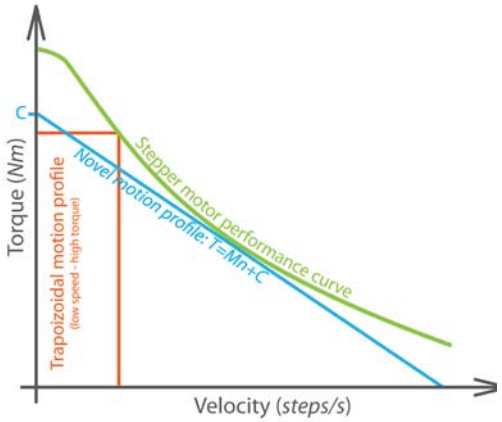


Figure 6-1: Typical stepper motor performance curve and motion control strategies

Then, in order to operate on our approximating line,

$$\frac{d^2\omega}{dt^2} I = Mn + C \tag{23}$$

or,

$$\frac{d^2\omega}{dt^2} I = M\alpha \frac{d\omega}{dt} + C \tag{24}$$

where α is a constant of proportionality. Integrating twice,

$$\omega = \frac{\alpha M}{I} t + \frac{C}{2I} t^2 + At + B \tag{25}$$

where A and B are integration constants.

(25) permits a simple and smooth trajectory using the following algorithm:

1. Calculate the total number of motor steps required for the move
2. Accelerate for half the step count, then decelerate for the remainder
3. At each time increment (step frequency), calculate the desired number of motor steps from (25). A time interval of 10k Hz is small enough to provide smooth acceleration.
4. If the actual number of steps taken is less than the total steps required, then take another

This algorithm provides smooth stepping motion and strong acceleration (positive and negative). This delivers progressively smaller acceleration mid-move as the maximum speed of the motor is reached. It is unnecessary to limit motor speed. This is because when the control line, (21), approaches the speed axis, the torque tends to zero and the motor no longer accelerates. When compared with constant acceleration (based on the cruise speed for a trapezoidal move), a tenfold increase in torque at the moves start and end is possible. Overall, this technique more than halves the move time (compared to a trapezoidal motion profile).

6.1.3 Backlash Minimisation

Precise arm positioning requires backlash reduction (ideally elimination). Spur gear driven arm axes are prone to backlash from poor gear meshing. The system play is excentuated along arm axes, producing positioning errors. Minimising this play increases their accuracy.

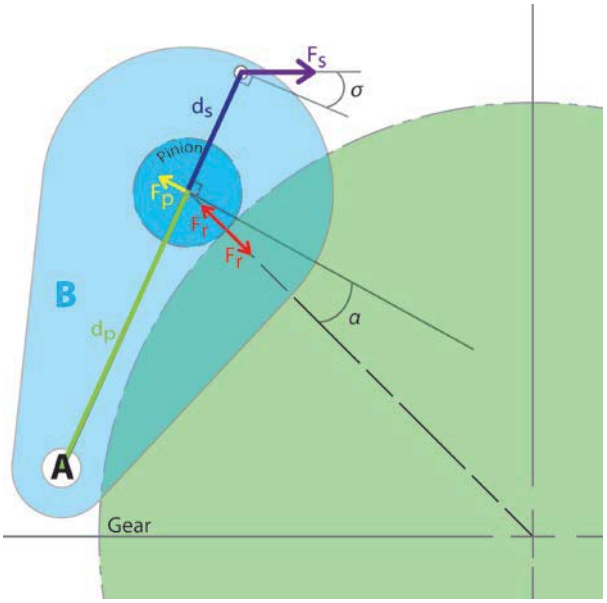


Figure 6-2 Spur gear preloading arrangement to eliminate backlash. The pinion rotates about point 'A'.

Backlash is minimised by introducing a spur gear meshing force. This force needs be greater than the motor introduced seperation force (F_r in Figure 6-2). Force F_r is the radial component introduced from the motor as it rotates the pinion. The meshing force makes the motors pinion gear teeth contact both adjacent gear teeth surfaces. This removes any tooth contact gaps (even with tooth wear), and minimises backlash.

The motor with attached pinion is mounted on plate *B* (Figure 6-2). This assembly is free to rotate about point *A* on a plane parallel to the gear face. The motor applies torque, *T*, to the pinion. This is a result of the involute tooth profile. An opposing force greater than F_P is required to keep the teeth in mesh. Force F_P is the perpendicular component of F_r. A spring applies this opposing force. It is represented by F_s (Figure 6-2).

A module 1 spur gear is used for the arms. This gear profile has a pressure angle, ϕ, of 20°. The radial component separating the two gears is,

$$F_r = \left(\frac{T}{d}\right)\sin\phi \tag{26}$$

Where, T is the motor's holding torque (6.8Nm, as this is at its maximum) and d is the pinions pitch diameter (0.03m). A radial force of 155N is the maximum force separating the meshing gears. This force acts to rotate plate B, about point A. F_r has a component, F_P, acting perpendicular to plate B of 148.8N (α = 16.3°). The perpendicular component of springs force, F_S, is required to (at minimum) balance the applied torque from F_P (acting about A). Therefore, the pinion separation torque, T_P, equals the applied spring torque, T_S, giving,

$$F_P d_P = \left(F_s \cos\sigma\right) d_s \tag{27}$$

solving for spring force F_S,

$$F_s = \frac{F_P d_r}{\cos(\sigma) d_s} \tag{28}$$

where: d_r = 0.1m, d_s = 0.175m, σ = 36.9° and recalling that F_P = 148.8N.

Therefore, a minimum spring tension, F_S, of 106.3N is required to maintain gear meshing and minimise backlash. Manufacturing tolerances and imperfections will produce some backlash. Assuming perfect teeth, applying a spring tension greater than F_S will have negligible system effect. This system and involute tooth profile ensures adjacent tooth face contact. This balances the resultant tangential force from the spring to the gear. A minimal increase to frictional resistance will result.

6.1.4 Hall Effect Sensor Position Feedback

Stepper motors do not always perform their directed steps. Increasing motor demands (e.g. speed, acceleration) increases these errors. Currently, approximately 1 out of every 1000 steps are not performed during unloaded linear motion. This results from the motors operating too close to their limits. It is recommended that stepper motor operation should have a 30% safety

factor from the published torque/speed curve. This ensures that steps are not lost. However, this requires larger motors that require more power and cost more. Although larger motors could be used, they do not account for increased loading from external orchard and environmental factors. Therefore, it is necessary to keep track of the step error to achieve accurate positioning.

A Hall effect sensor is used to track step error. This sensor detects drive gear holes that are machined into the gear's face. The gap between each subsequent hole increases by a small angle. This gives a unique step distance between consecutive holes. During setup the arm scans and records the drive gear holes. This is done at low speed to minimise step error. The step count for each axis is stored in a lookup table.

During axis rotation, the step count is recorded, when a rising edge is detected. When a falling edge is sensed the step count and lookup table are compared. Any error between expected and current step count is calculated. This updates the current step count removing the associated error.

Figure 6-3: Parallel linkage maintaining a vertical hand throughout range of motion

6.2 The Harvesting Arm

The harvesting arm is a three axis articulated robotic arm. Its three-degrees of freedom allows the arm to reach any canopy position within a working envelope. This is achieved with minimum impact to surrounding systems as the arm can reach out and around them. The arms motion also allows only the hand to penetrate the canopy.

A harvesting hand (section 6.2.1) attaches distally to the arm's third axis. The hand extracts the fruit from the canopy and provides some fruit handling ability. The fruit handling ability increases system efficiency. This is achieved by taking the harvested fruit from the hand without requiring the arm to move to a predetermined position. If the hand structure does not stay vertical, it would present a smaller opening to grasp the fruit. Consequently, fruit extraction would suffer. Therefore, a parallel linkage keeps the hand vertical throughout motion (Figure 6-3 and Figure 6-5). The linkage ensures the fruit are approached from the same direction every time and helps maintain correct fruit grasping.

Stepper motors provide axis motion. A kinematic model (section 6.2.4) and stepper motor control provides precise arm positioning. Axis motion is achieved by:

1. Axis 1, The Base: A 10:1 reduction spur gear drive provides rotation to the selective compliant upper and lower members. The large gear is fixed to the mounting plate and remains stationary. This reduces arm rotational inertia, increasing performance.
2. Axis 2, Lower Member: A pinion mates to a one third spur gear section (e.g. 1/3 circle segment). Torque is transferred through the outer concentric driveshaft directly mated to the lower member.

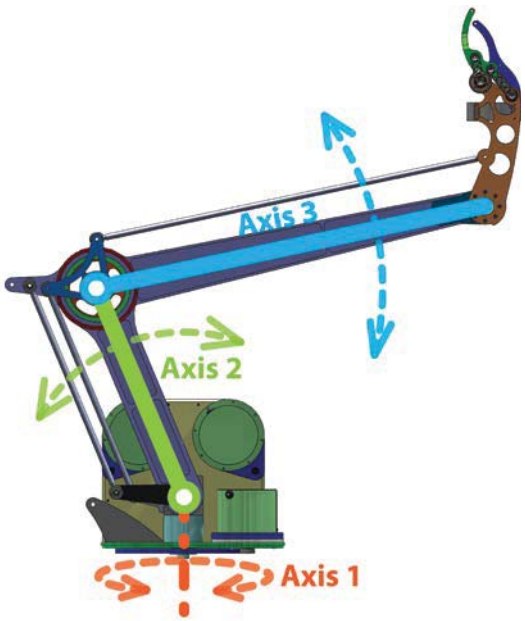


Figure 6-4: Robotic arm axis nomenclature

3. Axis 3, Upper Member: Another one third spur gear section/pinion, transfers torque through the inner concentric driveshaft. This is coupled to a crank and tie rod connected to the upper member.

Integrating developed control strategies and mechanical design features aims to produce a kiwifruit harvesting arm capable of efficient on-orchard operation.

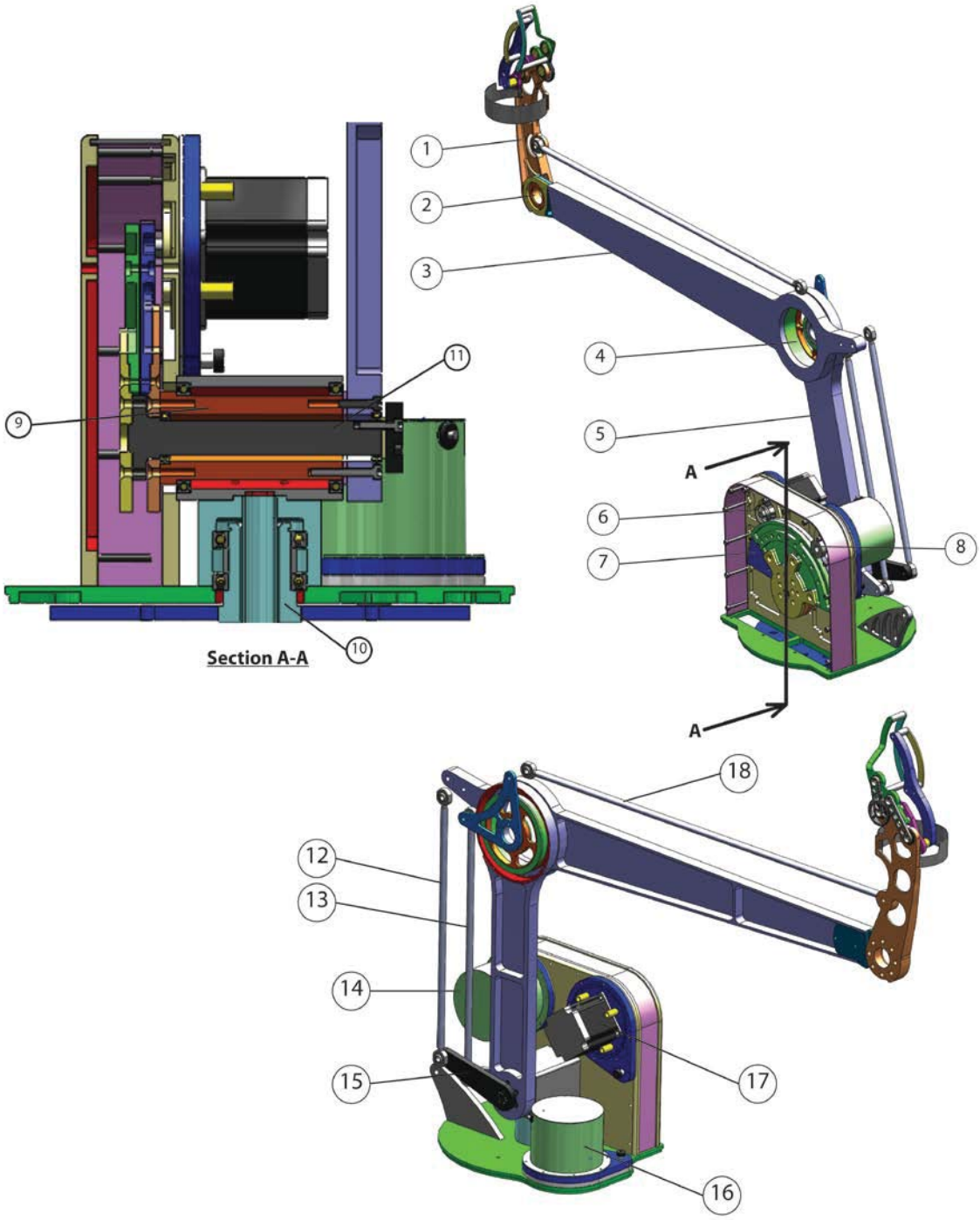


Figure 6-5: Functional layout and description of the robotic harvesting arm

(1) Hand, (2) Wrist Joint, (3) Upper Member, (4) Elbow Joint, (5) Lower Member, (6) Lower Member Pinion, (7) Lower Member Spur Gear, (8) Upper Member Pinion, (9) Lower Member Drive Shaft, (10) Shoulder Joint, (11) Upper Member Drive Shaft, (12) Upper Member Tie Rod, (13) Proximal Hand Linkage, (14) Upper Member Stepper Motor Unit, (15) Upper Member Crank, (16) Shoulder Motor Unit, (17) Lower Member Stepper Motor, (18) Distal Hand Linkage

6.2.1 A Review of Fruit Harvesting End-Effectors

An end-effector ('hand' or 'gripper') attaches distally to the robotic arm. Once positioned it performs a designated task. For harvesting, the hand extracts the fruit and provides some handling capability. However, few robotic applications currently exist for discrete fruit harvesting. Sarig (1993) comments that hands are one of the most complex mechanisms for automated fruit harvesting. It is a key component governing project success (Baeten, et al., 2007). A mechanical solution is required that can cope with the variability presented by the growing system.

Initial attempts at mechanised fruit extraction used thrashing implements. The Universities of California and Florida investigated the problem of fruit detachment. However, they applied techniques more typical of the agricultural industry (Muscato, et al., 2005). These harvest techniques are not suitable as they would cause fruit damage. They were:

1. Mechanical beaters and suckers
2. Heavy shaking
3. Chemicals sprays to cause fruit fall

Baeten, et al. (2007) recognises that "An apple by apple picking system does, however, require an adequate fruit gripper. The gripper is the key element in the success of automated apple by apple harvesting". This is a significant step beyond the initial attempts outlined above. Discrete harvesting can help maintain fruit quality.

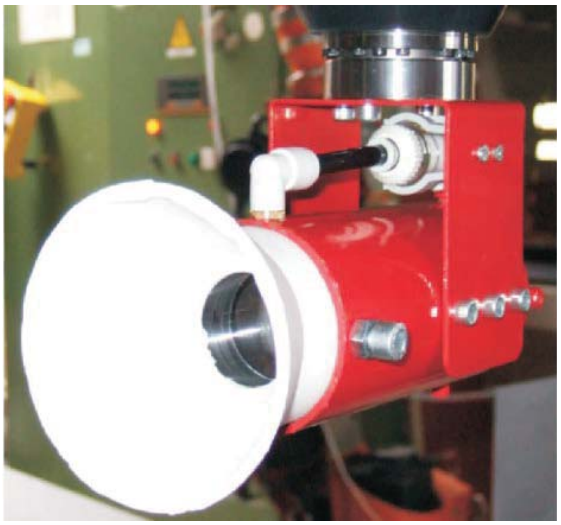


Figure 6-6: Apple harvesting robot silicone suction cup hand (Baeten, et al., 2007).

Edan, et al. (2000) comments on the hand's need to accommodate fruit variability (fruit size, shape, colour, texture and firmness). Harvesting hands are recognised as one of the most complex harvesting mechanisms (Sarig, 1993). Hands are generally used in industrial style environments where they are not required to handle the level of variation found on-orchard. Additionally, hand size and profile must be small. This helps limit surrounding fruit damage and canopy entanglement during extraction.

No standard fruit harvesting hands are available as every crop has different fruit extraction characteristics. Kiwifruit do not require stem cutting for extraction (unlike apples, oranges and mandarins). Therefore, kiwifruit extraction requires a specific hand design. This hand must rotate the fruit about its stem/fruit connection to ensure correct detachment. This is important as growers are penalised if fruit have stems attached.

The flexible hand for the autonomous apple harvester (Baeten, et al., 2007) provided a solution for grasping and extracting apples (Figure 6-6). This design allows small hand positioning inaccuracies without compromising fruit grasping. Vacuum draws and retains the fruit onto a silicon funnel (white cone, Figure 6-6). The soft funnel conforms to the fruit's shape, sealing the vacuum, providing suitable fruit holding.

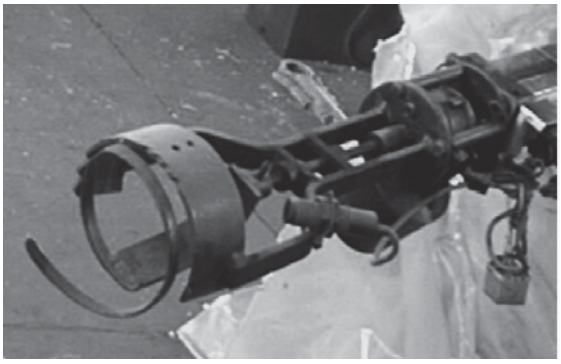


Figure 6-7: Helical citrus hand, (Muscato, et al., 2005)

The orange harvester's hand (Muscato, et al., 2005), required several redevelopments. The three main progressions, Figure 6-7, Figure 6-8 and Figure 6-10, are summarised.

Their first hand used a helical insert to catch the fruit's stem when it rotates. This places the stem into the correct position for cutting. Although mechanically simple the design does not allow for fruit handling after extraction (Muscato, et al., 2005).

Three prehensile fingers were used for fruit grasping in the next design (Figure 6-8). Application of pneumatic pressure closes the padded fingers against the fruit, where grasping force is controlled by air pressure. Once the fruit is grasped, the arm retracts, applying tension to the stem. A small force sensor is mounted on the hands wrist. This locates the stem's position allowing a micro saw to cut it (Muscato, et al., 2005). The results showed that it was capable of meeting all their required fruit extraction specification. However, fruit handling capabilities were not included and its "implementation cost

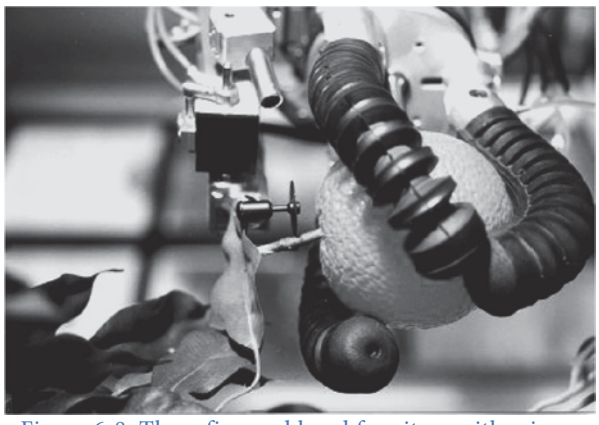


Figure 6-8: Three fingered hand for citrus with micro-saw stem cutting, (Muscato, et al., 2005).

was comparable to those estimated for the whole robot" (Muscato, et al., 2005). A solution that allowed fruit handling was required.

Figure 6-10: Hand implemented on the latest prototype harvester, (Muscato, et al., 2005).

Their last hand (Figure 6-10) is undesirably bigger and more complex. The larger size makes it harder to penetrate the tree. This increases the chance of damaging surrounding fruit and tree entanglement. Muscato, et al. (2005) attempted to "make the impact section as small as possible", however a suitable size has not been achieved. The large size is needed to house the sensors and actuators, while incorporating fruit handling.

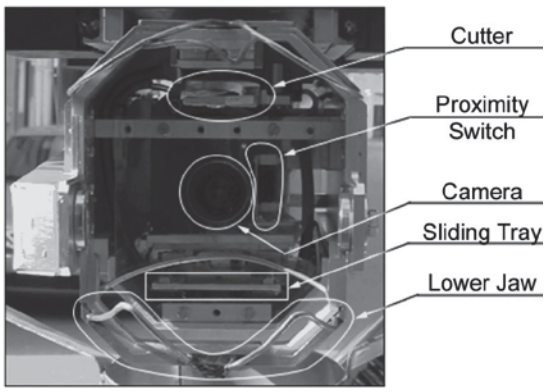


Figure 6-9: Orange harvesting hand – view from front (Muscato, et al., 2005)

Although large, this hand has simple operation. A camera within the hand (Figure 6-9) guides the hand to the fruit. While approaching the fruit a proximity sensor indicates when the fruit is within range. When in range, the lower 'jaw' (Figure 6-9) raises, lifting the fruit, where a pneumatic cutter separates the fruit from the tree.

Van Henten, et al. (2002) describe a hand for cucumber harvesting. This device incorporates fruit handling after it is cut from the vine. The fruit's stem is held by an industrial gripper (Figure 6-11, a modified Mitsubishi 1E-HM01), then stabilised by a suction cup (attaching to the fruit's side). A hot wire cutter separates the fruit from the vine. In addition, the cutter cauterises the cut to prevent viral transmission between plants. It seals the wound on both vine and fruit. This prevents water loss, minimises plant trauma and extends fruit shelf life. The thermal cutting device and method is patent protected (Van Kollenburg-Crisan, et al., 1999). Although beneficial for cucumbers, this is not suitable for kiwifruit harvesting where no stems are allowed.

The Kubota Company has developed an experimental hand for citrus harvesting (discussed by Sarig (1993)) (Hayashi and Ueda, 1991). Mechanical operation is shown in Figure 6-13, however no capability or performance information is provided.

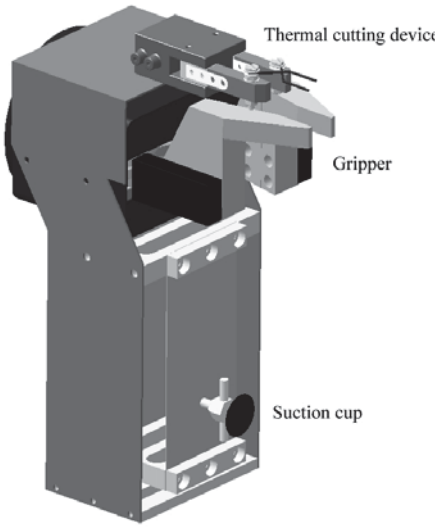


Figure 6-11: Hand for harvesting cucumbers

A central suction cup is used for holding fruit. Once attached, the cutter and cover extends over the fruit. Its shape guides the stems into the pneumatic cutter, which cuts the fruit from the tree. The cover and cutting assembly that moves over the fruit has many sharp edges. As there is not performance information, no fruit damage conclusions can be made. However, there appears to be a significant risk.

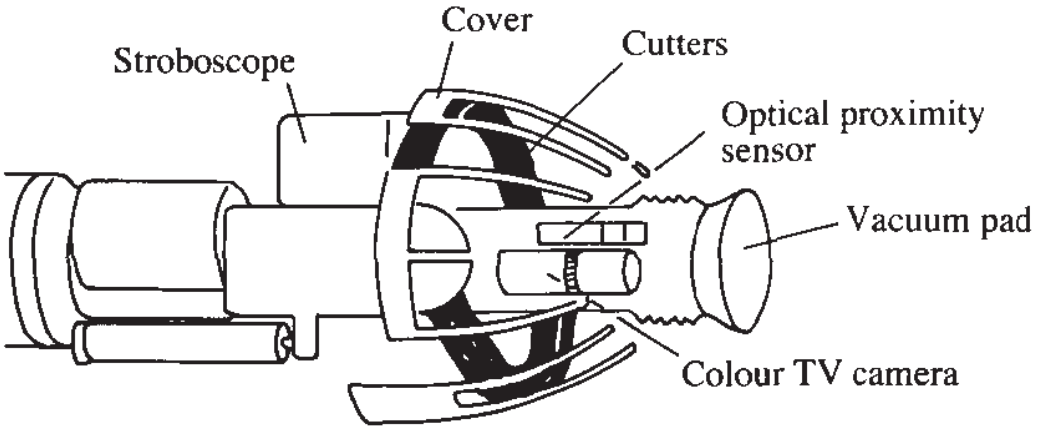


Figure 6-12: Experimental citrus hand (Hayashi and Ueda, 1991)

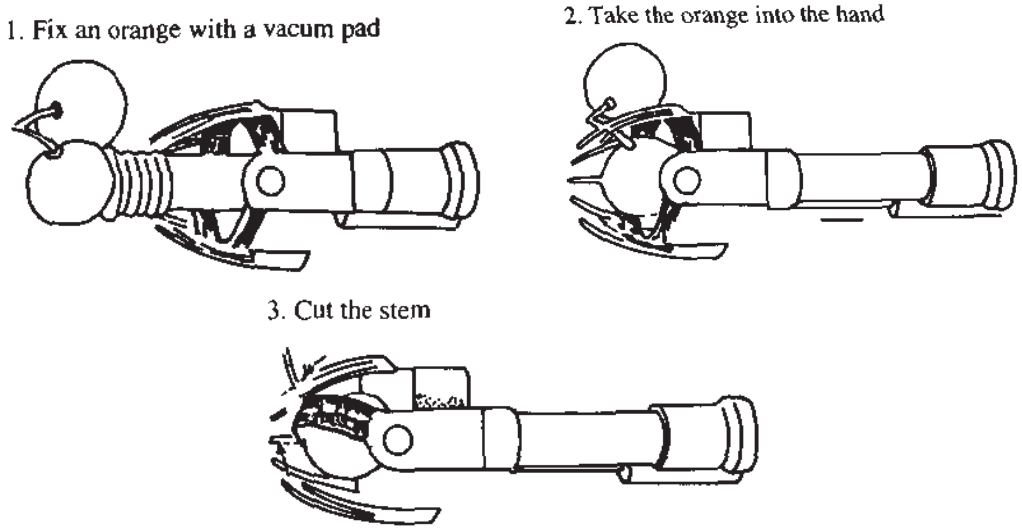


Figure 6-13 Operation of the experimental citrus hand (Hayashi and Ueda, 1991)

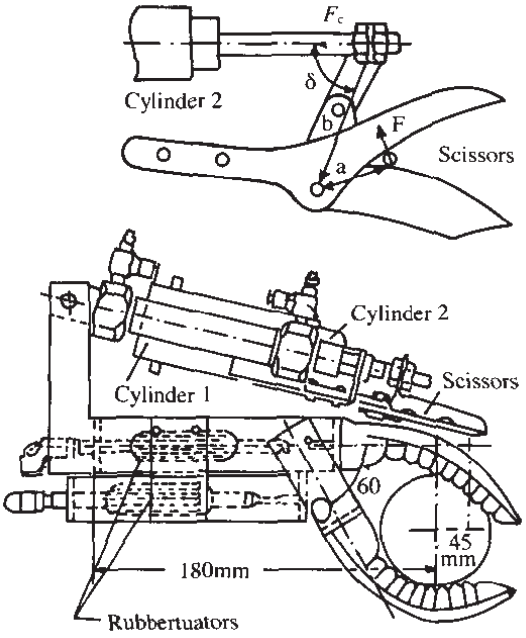


Figure 6-14 Pneumatic three fingered orange harvesting hand, developed at Kyoto University (Sarig, 1993).

Another orange harvesting hand was developed by Kyoto University (Sarig, 1993). This design has three rubber fingers to grasp the fruit. Bridgestone Co. rubber pneumatic actuators ('Rubbertuators', Figure 6-14) provide grasping. These actuators provide a controllable grasping force to prevent fruit damage. After the fruit is held, a pneumatic cylinder shifts the cutting scissors forward fifty millimetres. A second pneumatic cylinder controls the scissors and cuts the fruit's stem. Stem cut length is the only performance data reported. This was approximately 30mm too long for citrus fruit targeted for fresh consumption.

6.2.1.1 End-effector Conclusions

Unlike the reviewed hands, kiwifruit do not require cutting from the vine. The hand must mimic a human hand and rotate the fruit to detach it from its stem. Efficient fruit handling is another key consideration. Most of the reviewed devices did not have this ability. Where fruit handling was included, it required large arm movements to drop the fruit at a predefined position. This is not an efficient system and slows harvesting cycles.

The prehensile fingers configuration presents a practical grasping design. This design easily accommodates a variety of fruit sizes and it presents a small penetrating profile. Handling fruit variation is a key design consideration (Edan, et al., 2000) and can limit design practicality. Suction cups are a common holding device but they are not suitable in this application due to the rotation action required. Suction cups are suitable for pulling/holding actions but not rotations. The fruit's furry surface would further reduce their holding ability.

The required hand needs to include:

1. The ability to handle all kiwifruit variations. These include: size, shape and firmness.
2. Present a small cross sectional area for canopy penetration. This limits damage to fruit and entanglement

3. Integrated fruit handling for efficient harvesting cycles. This allows the fruit to be moved to a storage system after extraction. This increases efficiency as the arm does not have to move over to and place fruit onto a conveyor, or similar.
4. Lightweight construction to improve arm acceleration

To achieve acceptable arm motion performance, hand weight and size reduction is necessary. Removing the camera and pneumatic cylinders from the hand will reduce its size and mass. Better arm motion performance will result.

6.2.2 End-Effector Development

The hand must follow the extraction and handling techniques to help maintain fruit quality. The main extraction technique specifies that the fruit must be rotated to create a bending moment at the stem fruit connection point. This is to ensure stemless fruit removal. Additional fruit handling specifications are outlined in section 3.

Fruit quality must be maintained, including the quality of the surrounding fruit still in the canopy. In addition, fruit of any size, shape, profile and position require extraction. The specific extraction techniques, large-scale variables and fruit quality requirements increase design complexity. Several development cycles were required to achieve a suitable design that fulfilled all

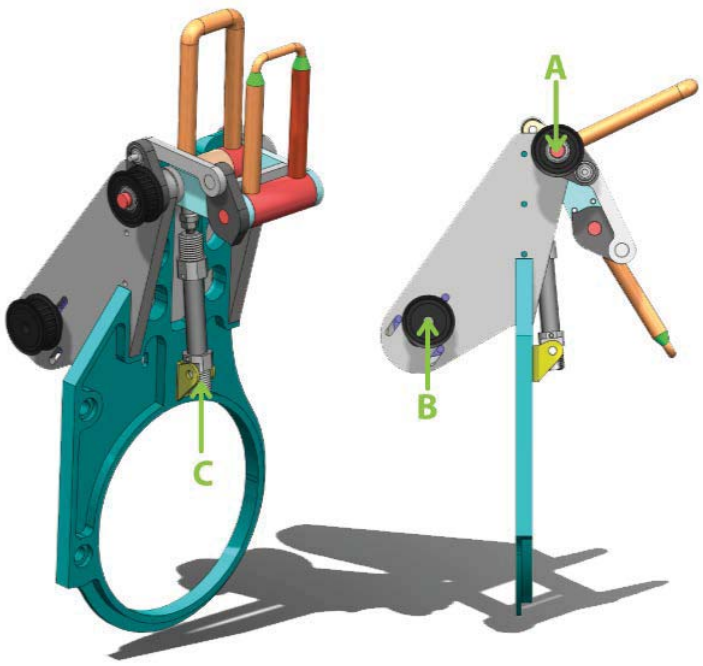


Figure 6-15: The first hand (isometric and side views). A - fruit rotation point, B – Clasping pneumatic rotary actuator and C – Fruit rotation pneumatic cylinder.

requirements.

Industrial style environments often develop hands for specific tasks. These controlled conditions are not subject to the variations present in orchard environments. To cope with the harvesting variations, large fruit clasping motion and rotation is required. This is conceptually easily achieved however, an hand with increased open profile results. The larger profile reduces canopy penetration and risks damaging neighbouring fruit and entanglement within the canopy.

The first design used two fingers to grasp the fruit. Although modified, this design feature is consistent throughout all redevelopments. The fingers small cross sectional area allowed easy canopy penetration. Vine entanglement and surrounding fruit damage was also reduced. The initial finger concept used a rubber covered metal arch for fruit grasping. However, this design does not spread the grasping force uniformly across the fruit's surface, where pressure damage could result. Subsequent designs use silicon tubing to spread the grasping force evenly over a larger fruit area. Pressure on the fruit was reduced, which helped maintain quality.

In addition, the first hand concept (Figure 6-15) had a significant extraction limitation. Insufficient rotation about the stem-fruit connection caused stems to remain connected. The fruit rotation point used (point A, Figure 6-15), below the grasping finger applies some bending moment about the stem-fruit connection point. However, an increased tensile force is also applied. This led to stems remaining attached, which would incur grading penalties. To achieve stemless fruit extraction, the tension applied to the stem during extraction needed reducing. The redevelopment cycle set out to correct this issue.

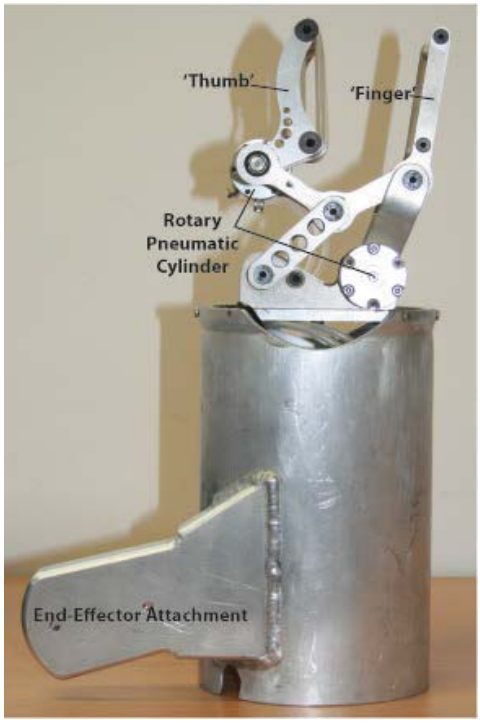


Figure 6-16: Third hand development, with the first four bar linkage implementation

The second development rotates the fruit about the fruit-stem connection point. This achieved stemless fruit extraction. A toothed belt and pulley system (driven by a pneumatic rotary actuator) rotates the mechanism and fruit, removing the fruit. However,

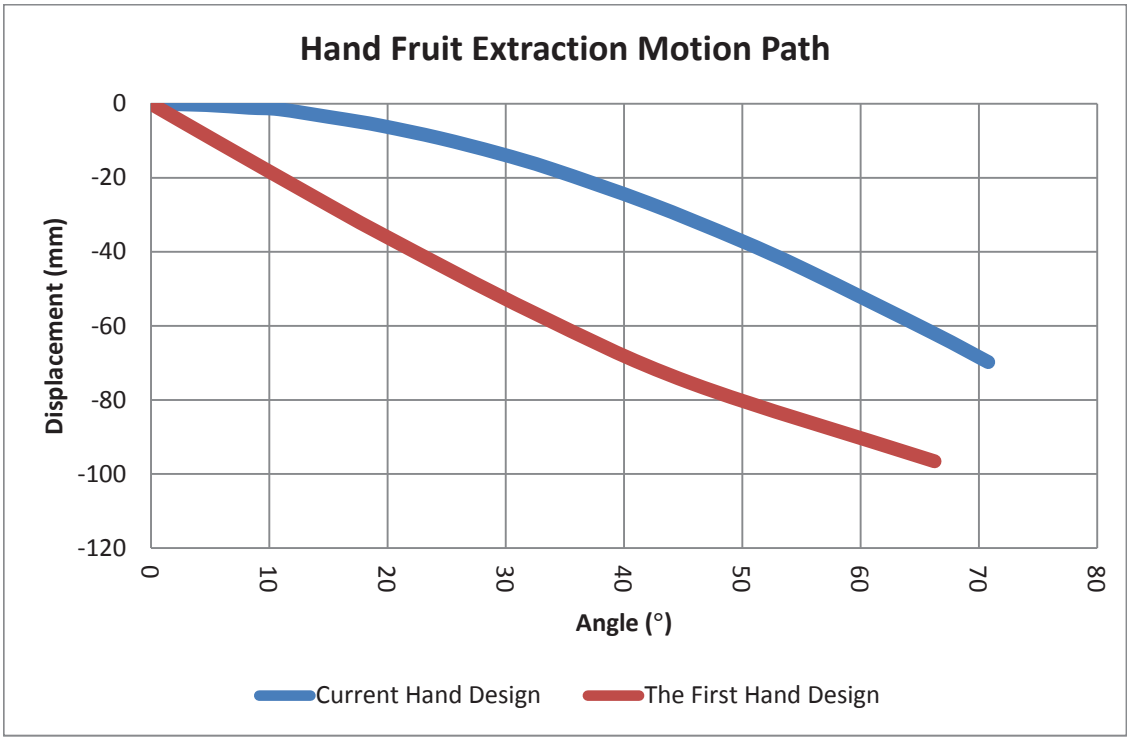


Figure 6-17: Hand Fruit Extraction Motion Path

an increased cross-sectional area resulted in canopy penetration issues and damage to surrounding fruit from collision. A small linear pneumatic cylinder applies the grasping force by rotating the thumb. However, the protruding cylinder body created fruit damage and canopy entanglement issues.

A four-bar linkage is used to resolve the issues presented with the first and second designs. The linkage provides sufficient rotation about the stem-fruit connection to provide clean fruit extraction. The stem tensile force gradually increases with fruit rotation. This cleanly separates the fruit once sufficient rotation is applied. This design reduces cross-sectional area for canopy penetration by positioning the linkage below the grasping fingers. Two pneumatic rotary actuators move the four-bar linkage (Figure 6-16). Another pneumatic rotary actuator opens and closes the fruit grasping thumb. When the thumb opened it was intended that the extracted fruit could fall through. The released fruit would be caught by a soft tube, with gravity delivering the fruit to the fruit handling conveyor. This design would provide efficient harvesting cycles and fruit handling capability. However, the thumbs geometry did not always allow larger fruit fall as required.

The four-bar linkage concept performed well. However, some design modifications were required to improve performance to a commercial level. These included:

1. Reducing the canopy penetrating cross-sectional area. The fingers were too wide and could damage neighbouring fruit.
2. Thumb rotation geometry to achieve fruit handling.
3. Weight reduction to increase arm motion performance. The three pneumatic rotary actuators increase hand mass.

There are several industrial type gripper actuators available on the market that we investigated for their potential to hold fruit. These ranged from electric, pneumatic and hydraulic systems; however, these actuators had at least one of three primary flaws. These were:

1. Mass: The weight of the some of the actuators was too heavy to be mounted to the end of the arm. For example, an actuator(s) of 0.3kg would result in an additional 2.6Nm reaction torque on the arm's third axis. This will require an approximate 42% increase in axis motor size, potentially for all axes.
2. Size: It is important to reduce the size of the hand penetrating the canopy. This is to avoid entanglement and damaging other fruit.
3. Stroke: A lot of the actuators are built for factory made products that do not vary in size like fruit. Therefore, most gripper type actuators have a limited stroke for these common applications.

These considerations drove development through the several design cycles to the final pneumatic/cable actuated system. This system allows for a small profile, large stroke and the ability to position the actuators mass at the axes rotation points to minimise arm inertia. This system provided the greatest benefits and operation and is presented in the next section.

The latest hand design (section 6.2.3) addresses these design changes.

Figure 6-18: Left: Image of the robotic arm mounted on the chassis in the harvesting position. Right: Image of the hand and arm assembly showing all of the external components and linkages.

6.2.3 Fruit Extracting Hand

The revised hand retains the four bar linkage design. This linkage accentuates a bending moment at the stem-fruit connection for stemless fruit extraction. A fixed finger (green component, Figure 6-20) and rotating thumb (purple component, Figure 6-20) provide fruit grasping. Bowden cables connected to pneumatic actuators provide grasping and rotation force, F_{C2} and F_{C1} (Figure 6-20). The actuators are mounted near the arm's axis 1 and 2 rotation points to reduce system inertia. This increases robotic arm motion performance.

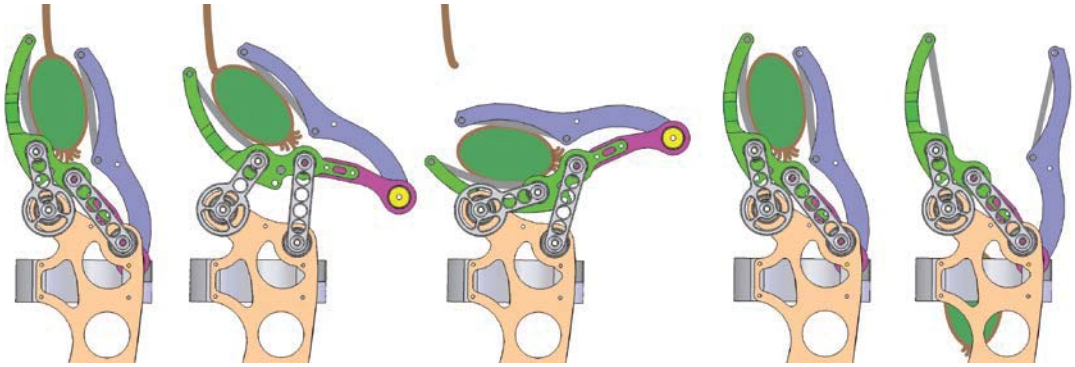


Figure 6-19: Demonstrates harvesting end-effector picking motion

Curving the finger and thumb minimises fruit damage from contact and canopy entanglement. Winding silicon rubber tube around the finger and thumb (grey lines in Figure 6-20) provides a large soft contact surface. The tubing conforms to the fruit's shape and disperses the grasping force over a large area. This reduces pressure on the fruit and damage. A human hand, grasping fruit during extraction, may cause thumb pressure damage. This results from the thumb having a smaller contact area than the four fingers applying the opposing force. The hand balances finger and thumb contact areas and provides even pressure distribution over the fruit/hand contact surfaces, reducing damage.

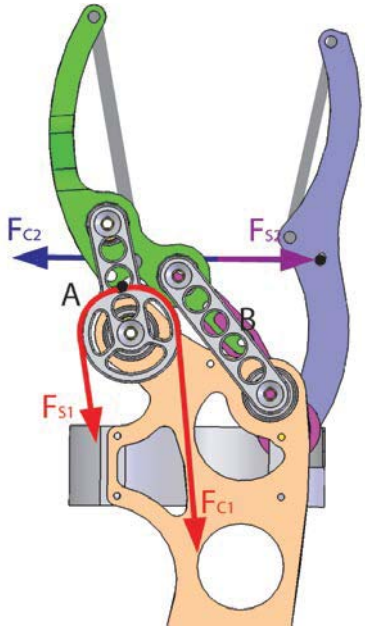


Figure 6-20: Diagram of the functional operation of the harvesting hand

The Bowden cable that rotates the four-bar linkage lies within a circular housing (represented by F_{C1}, Figure 6-20). This provides a constant rotational torque throughout extraction motion. A tension spring pulls the linkage upright after fruit extraction. This spring pulls against the same

cable providing reversed rotation to linkage A (F_{S1}, Figure 6-20).

Rubber cuffs, mounted between the finger and thumb, guide extracted fruit downward when released. Spiral wound rubber tubing cushions fruit descent to the fruit handling conveyors. By loosely suspending the tubing, the fruit's impact time constant is increased. This reduces impact pressure and fruit damage.

Difficulties with fruit extraction arose during on-orchard testing (see results – section 6.2.8). The fruit's base (the localised fruit point) must sit low in the hand to accommodate all fruit lengths. Therefore, the stem of smaller fruit sit lower in the hand's grasping mechanism. For smaller fruit, the thumb can contact and bend the fruit's stem during extraction. This reduces the bending moment at the stem/fruit connection point and can compromise clean fruit extraction. To counter this, a V shaped slit was added to the thumb which allows the fruit's stem to pass through the thumb during fruit rotation. The modification ensured correct fruit rotation and clean fruit removal.

6.2.4 Harvesting Arm Kinematics

Each arm axis angle must be calculated to position the hand in three-dimensional space. Axes angles are calculated relative to the arm's home position. This is defined as:

1. Axis1 (A_1): Axis 2 and 3 parallel to the ATP's x axis (Figure 7-5)
2. Axis2 (A_2): Perpendicular to the arm's base plane (nominally vertical)
3. Axis3 (A_3): Perpendicular to axis2 (nominally horizontal)

Angles α, β and γ (Figure 6-21) are calculated from the home position using the required x, y and z end location.

The arm's coordinate origin is 1000mm forward of A_2's rotation point. This closely aligns its coordinate geometry with the localisation cameras positioning the fruit for extraction. Therefore, $x+1000$ is used for any commanded arm position. This converts the supplied (x,y,z) position to a coordinate in arm geometry.

Position values x and y determines the angle of A_1, α (Figure 6-21):

$$\alpha = \tan^{-1}\left(\frac{x+1000}{y}\right) \tag{29}$$

A_2 (L_2) and A_3 (L_3) form a triangle linked by radius, r, with internal angles σ and φ (Figure 6-21). Where the Pythagorean theorem gives length r,

$$r = \sqrt{(x^2+1000)+(z-D)^2} \tag{30}$$

Rearranging the law of cosines, to determine internal angle φ,

$$\varphi = \cos^{-1}\left[\frac{r^2 - {L_3}^2 - {L_2}^2}{-2L_3L_2}\right] \tag{31}$$

Likewise for angle σ,

$$\sigma = \cos^{-1}\left[\frac{{L_3}^2 - {L_2}^2 - r^2}{-2L_2r}\right] \tag{32}$$

The radius is at angle, θ, to the horizontal, where,

$$\theta = \tan^{-1}\left(\frac{z}{x}\right) \tag{33}$$

To obtain β, A_2's angle from its home position,

$$\beta = 90 - \theta - \sigma \tag{34}$$

A_3's angle, γ, from home, from the sum of internal angles,

$$\gamma = \varphi - \beta - 90 \tag{35}$$

All required axis angles for a given end position are now known. The controller requires these angles as step counts for each axis. The controller will step each axis from its current position until the correct end step count is reached (see section 6.1.1). This corresponds to the desired end position.

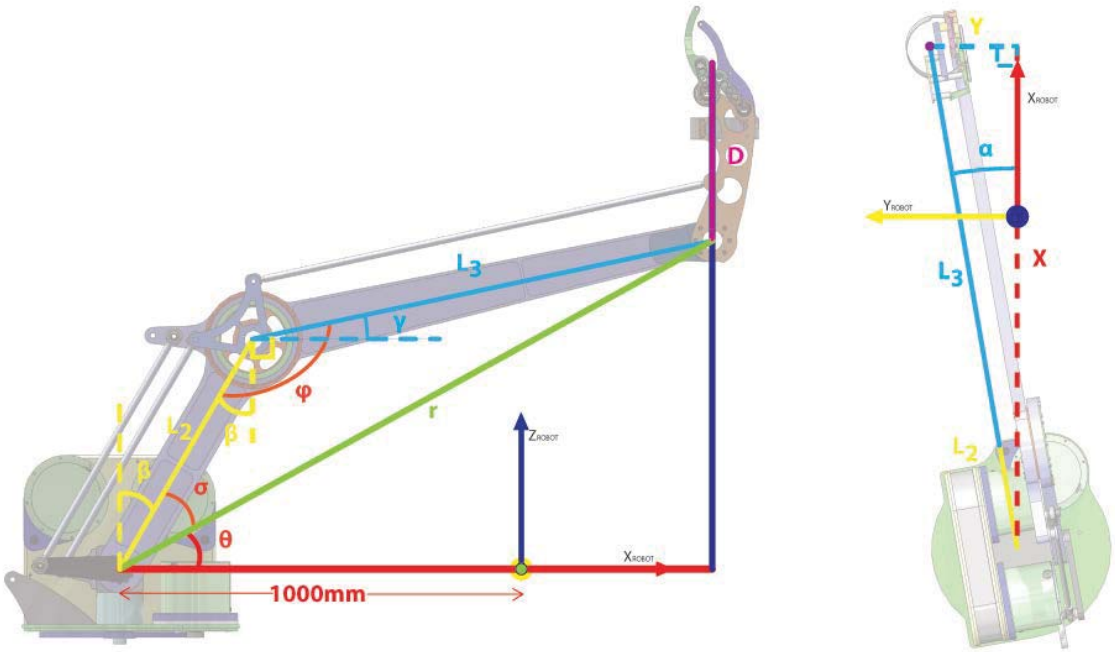


Figure 6-21: Harvesting Manipulator Kinematic Nomenclature

At the output shaft the number of degrees for every step (ω_S) is,

$$\omega_S = R \times S_R \tag{36}$$

where, S_R is the step resolution (S_R = 360°/3200) and R is the gear ratio (R=1/10). Therefore, for every motor step the output shafts will rotate 0.01125°.

Multiplying ω_S by the end position angle gives the required step count, η_S, for that axis. The controller knows the current step count. The difference between the desired step count and the current step count gives the number of steps and direction for the axis.

6.2.5 Custom Picking Trajectory: The 'U-Move'

Scattered fruit positions and clustering limit direct linear transition between fruit. To avoid fruit damage the hand approaches the fruit from below. A novel motion trajectory optimises fruit-to-fruit transitions and minimises fruit damage. By combining harvest scheduling (section 6.2.6) and arm-arm collision detection (section 6.2.7) fruit are extracted without collision, in an ordered and efficient manner.

The grey path (Figure 6-22) shows a series of linear transitions between two fruit locations. This ensures vertical movement at the start and end of the move. Using a series of linear traversal

requires several time-consuming acceleration and deceleration movements. Harvest cycle performance would consequently suffer. By starting the next linear section during the current linear section (e.g. start section *B-C*, part way through section *A-B*) produces a seamless transition. This reduces fruit-to-fruit move times, increasing harvest efficiency.

The transition path (blue line, Figure 6-22) is achieved by:

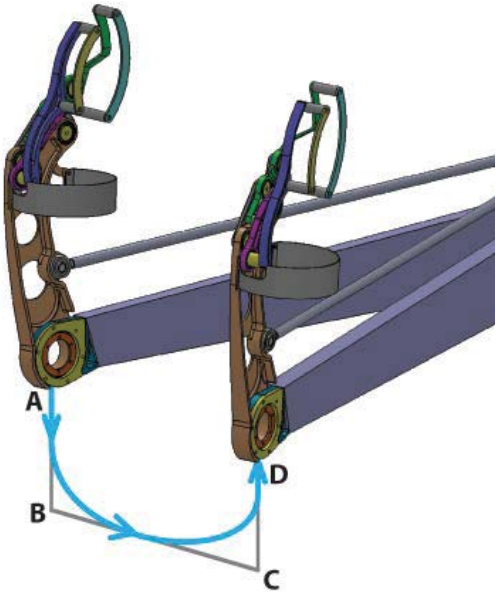


Figure 6-22: U-Move custom trajectory

1. The height of point *B* and *C* is set 100mm below the lowest fruit height. This ensures fruit are approached from below
2. Linear distances *A-B*, *B-C* and *C-D* are calculated referencing the start and end positions
3. The shorter transition distance of *A-B* or *B-C* is calculated and halved (η_1 - in 'steps')
4. The shorter transition distance of *B-C* or *C-D* is calculated and halved (η_2 - in 'steps')
5. Starting at *A*, the arm moves down towards *B*. At η_1 'steps' from the end of move *A-B*, transition *B-C* starts.
6. Move *C-D* begins at η_2 'steps' from the end of transition *B-C*
7. Stops when *D* is reached

This places the hand at the next fruit position, from its current position. Points 3 and 4 above ensure the curved sections of motion are restricted to no more than half of any one section. Without this restriction, transition *C-D* could start before *A-B* had completed. This could damage fruit and alter the arm's end position.

6.2.6 Fruit Extraction Scheduling

Fruit extraction scheduling is used to minimise damage to fruit and increase extraction efficiency. Hand design and operation impose specific extraction rules. These rules are associated

with closely positioned fruit and mean that fruit clusters are harvested in a specific order. Figure 6-19 shows the hand's fruit extraction method. Fruit damage can occur due to:

1. The open thumb: this can impact surrounding fruit on canopy entry
2. Fruit rotation: the thumb's structure rotates upward, potentially colliding with surrounding fruit
3. Four-bar linkage rotation: could potentially impact surrounding fruit during operation

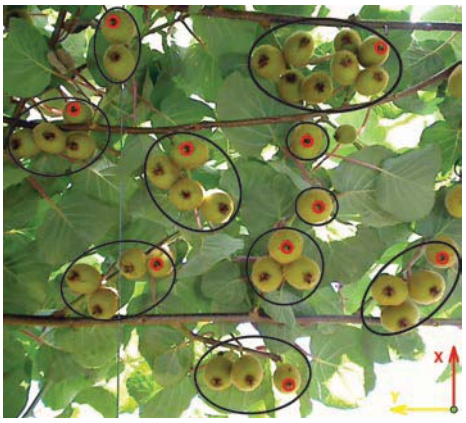


Figure 6-23: Fruit cluster harvesting order. Black rings define cluster grouping. The red circles show the first fruit to be harvested in each group.

Fruit are located with all four camera pairs. These fruit locations are mapped into the reconciled coordinate space (section 5.3.3). Fruit clusters are determined by the distance between them. The clusters are grouped and recorded (black ellipses, Figure 6-23). The fruit are divided approximately evenly between the arms where complete clusters are assigned to each arm. This allows the harvesting rules to be followed.

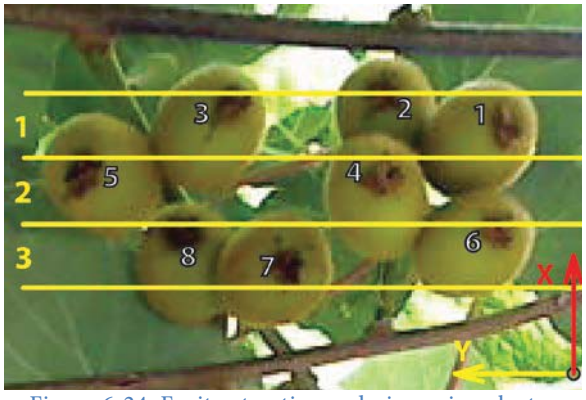


Figure 6-24: Fruit extraction ordering using cluster banding

The first harvested group (1, Figure 6-23), for each arm is the cluster nearest the top right hand corner. This helps avoid arm/arm collision by moving all arms in the same direction. The shortest path to traverse the groups defines their removal order (travelling salesman problem). The distance between the groups is the distance between the first fruit in each group to be harvested (red circles, Figure 6-23).

Each cluster group requires harvesting in a specific order. Clusters are divided into 50mm wide bands in the Y direction (Figure 6-24). The fruit with the largest X value (fruit 2, Figure 6-24), within the cluster, is used to define the first band. More bands are added until all fruit are assigned to a band. The fruit are harvested band-by-band from the lowest Y value to the highest. Harvesting bands are also assigned over the vertical (Z) range. The lowest fruit bands are harvested first, progressing upwards in 50mm intervals. Once all fruit in the first band are re-

moved, the next band (over the complete height range) are harvested. This cluster harvesting method limits potential fruit damage (points 1-3 above) and supports fruit quality.

This harvesting technique assumes that all fruit within a target area are identified. Hand to fruit collision can occur with unidentified fruit. However, these fruit are generally located higher, reducing this risk. This was evident during development of the fruit identification algorithms and is mainly due to leaves occluding the fruit. There is greater chance of this occurring the higher the fruit is.

6.2.7 Arm-Arm Collision Detection

Adjacent arms can collide as they share overlapping operating areas. Detecting a potential collision before it occurs allows the collision to be avoided. The common method of collision detection has been to use geometric primitives. These simple geometric shapes (often blocks) are used to describe each structure of the arm. "Thereby a collision is detected by intersections between the geometry of obstacles and robot" (Henrich and Cheng, 1992)

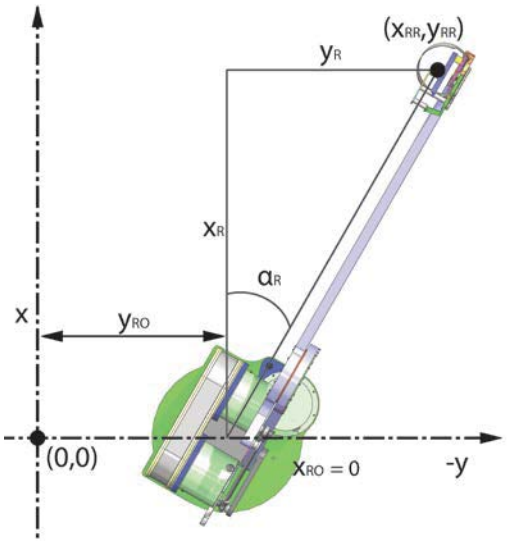


Figure 6-25: Robotic arm geometric layout: x-y plane.

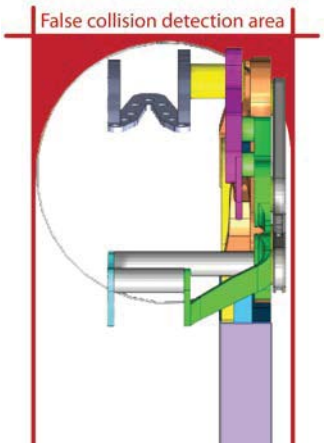


Figure 6-26: False positive collision area

In this system, each arm's physical structure and operation supports simplification to collision detection calculation. Collision detection is calculated in the arm's *X-Y* plane only. The long hand, constant vertical harvesting motion (section 6.2.5) and low canopy profile makes overlapping arms impractical. Therefore, arm overlapping is not allowed.

The reconciled coordinate space (section 5.3.3) is used for collision detection and avoidance. A 'keep-out' zone is defined around the arm's upper member and hand (Figure 6-28). The rectangular zone configuration simplifies detection calculation. A small area of 'false positive collision zone' is in-

cluded around the hand (Figure 6-27). This is the area where the system identifies a collision will occur, that would not actually occur. However, false positive collision detection is insignificant due to the small 'false positive collision zone'. A fruit carrying rubber tube runs parallel to the upper member, filling the remaining 'keep-out' space. The tube is not shown in the diagrams.

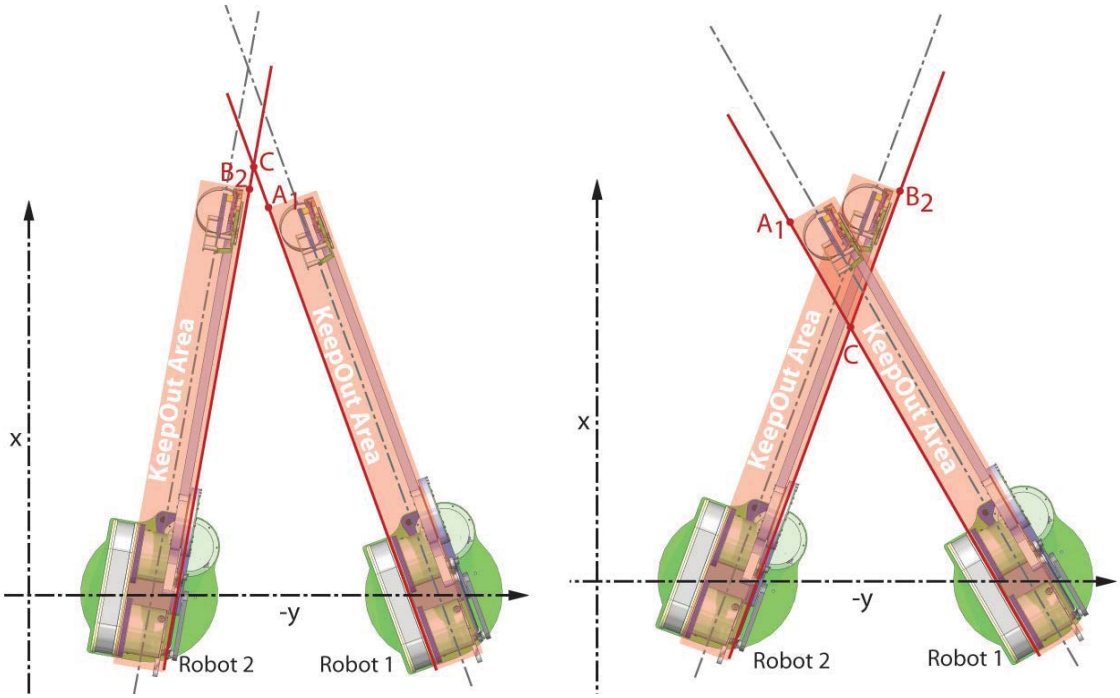


Figure 6-27: Adjacent arm collision detection method. Left – safe operational condition; Right – a collision will occur if this move is performed

The intercept between adjacent keep-out area borders (point *C*) is calculated to determine if a collision will occur. If x_C is less than both x_{A1} and x_{B2} (Figure 6-27) a collision will occur. If x_C is less than 0, the arms are pointed away from each other. This is a safe operational condition. Values y_{RO}, x_{RO}, x_R, y_R and α_R (Figure 6-25) are known. y_{RO} and x_{RO} are the offsets of the base in the global coordinates. x_R and y_R are the respective components of the arm's end position (x_{RR}, y_{RR}).

To calculate intercept *C*, the position of A_1 and B_2 (Figure 6-27) and intercepting border lines are required. k_w and k_h are the arms' respective keep-out width and height (Figure 6-28). Where,

$$h = \sqrt{2} \times k_h \tag{37}$$

therefore, points *A* and *B* for robotic arm *R* (A_{Rx},A_{Ry}) and (B_{Rx},B_{Ry}),

$$x_{RA} = x_R + h\sin(45 - \alpha_R) \tag{38}$$

$$x_{(R+1)B} = x_{(R+1)} + h\sin(45 + \alpha_R) \tag{39}$$

The lines through points A_R and $B_{(R+1)}$ are needed (red lines, Figure 6-27). These lines are parallel to the arm (line between (x_{RO}, y_{RO}) and end-effecter position (x_R, y_R)) with slope, m_R,

$$m_R = \frac{(y_R - y_{RO})}{(x_R - x_{RO})} \tag{40}$$

Equations for the edges of the adjacent arms are therefore,

$$y_{RA} = m_R x_{RA} + c_{RA} \tag{41}$$

$$y_{(R+1)B} = m_{(R+1)} x_{(R+1)B} + c_{(R+1)B} \tag{42}$$

To find the intercept of the two lines, $y_R = y_{(R+1)}$,

$$m_R x_R + c_R = m_{(R+1)} x_{(R+1)} + c_{(R+1)} \tag{43}$$

At the intercept $x_R = x_{(R+1)} = x_C$, therefore,

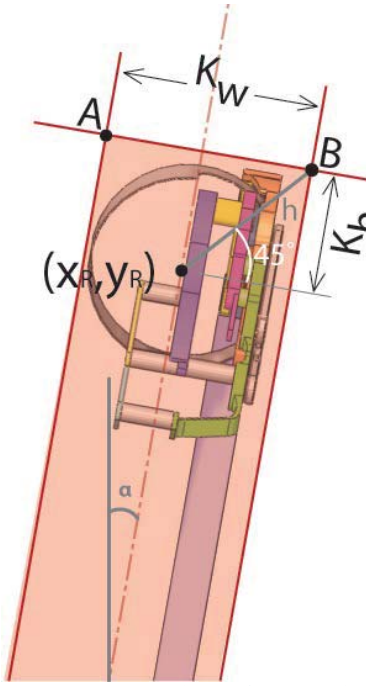


Figure 6-28: Defining 'keep-out' areas

$$x_c = \frac{c_{(R+1)} - c_R}{(m_R - m_{(R+1)})} \tag{44}$$

where x_c is used to determine if a collision will occur.

This describes static collision detection. The arms move linearly between fruit when viewed from above during harvesting. As collision detection ignores the height range this linear path is used for dynamic collision detection. This dynamic collision detection works on a master-slave system. Harvesting arms *1* and *3* are set as master arms and are not interrupted if a collision is detected. The remaining arms (*2* and *4*) alter their operation to avoid any collisions. As the fruit harvesting order is known, along with the time required to perform each fruit extraction, the fruits' static positions are used to determine if collision will occur between neighbouring arms. While the master arm is operating the slave arm's positions are used to monitor for any collisions for future fruit locations. A single, or consecutive, fruit are skipped if a collision is going

to occur. The slave arm harvests fruit further through its schedule and returns to the skipped fruit when the master arm is out of the way. This approach is used, as potential collisions do not occur very often due to the way the harvesting order is scheduled (section 6.2.6), sending the arms in the same offset direction to begin.

6.2.8 Testing and Results

The harvesting arm is accurate to ±0.3mm in three-dimensional space with the configured micro stepping resolution of 3200 steps per motor rotation. However, this assumes that every step is taken when commanded. To achieve ±1mm positioning accuracy each axis can lose a set number of steps. This number is proportional to the axis length (at worst-case position):

1. Axis 1 – 3 steps
2. Axis 2 – 10 steps
3. Axis 3 – 5 steps

Therefore the tracking system needs to be able to maintain each axis step error to be less than or equal to these values. Although backlash was removed from each axis, the axes can still be deflected by a small amount. Axis deflection to find any backlash was measured using a micrometre with a magnetic base. However, the axis returns to its original position when the deflecting force is removed. This means that the axis will be offset with acceleration forces, but will settle to its correct position upon completing the move. These deflections were measured with a micrometre by applying a force equal to the greatest motor force. They produced momentary errors of less than 1.5mm in three-dimensional space. As these errors are only during acceleration or loading, and settle to a neutral point during deceleration, the error is not significant.

Initial testing was conducted in the laboratory to assess the arms ability to harvest fruit. This process used hand-eye coordination (section 5.2.4) to localise the fruit for harvesting. Ten randomly positioned fruit were hung from a wooden ‘canopy’ structure. Stereopsis (section 5.2.3) located and scheduled all fruit for extraction. Harvest time was recorded and the test repeated ten times (one hundred fruit in total). Different fruit positions were used with varied spacing. All one hundred test fruit were correctly identified, localised and extracted. An average harvesting cycle of 1.3 fruit/second/arm was achieved. This exceeded the required harvesting rate of

one fruit/second/arm (four fruit per second total). The combined machine vision and arm system accuracy is reported in section 5.3 and is sufficient for fruit harvesting.

Two harvesting arms were used for on-orchard testing. Cooperative, asynchronous and collision free operation was achieved. The control code supports operation of all four arms to this level. Currently, the control computer and Visual Basic 6 cannot support concurrent USB connection of all eight cameras. Modular micro form-factor computers will be used to resolve this connection issue. Implementing more computers in parallel (perhaps one for each stereo pair of cameras) will also decrease analysis time. Fruit positions will be relayed to a main scheduling and control computer. This computer will assign commands to the harvesting arms.

Five consecutive orchard bays were used for harvest trials (approximately 4 x 25m total). The fruit covered the range of variability expected with real operation. They were grouped in clusters, randomly scattered and hung throughout the canopy height range. Difficulties with extraction were quickly obvious. Natural fruit-stem attachment was much stronger than the artificial fruit attachment used in the laboratory. Modifications were made to the hands overcome this issue. These were:

1. Increased rotational speed and force
2. A 'pause' delay after fruit grasping and rotation. This ensured the fruit was fully rotated before downward arm motion was applied

These modifications ensured successful fruit extraction (for correct arm positioning). However, harvesting rates slowed and issues with arm positioning resulted. The additional extraction forces could alter the arm's position. This position error was often uncorrected for short arm movements (see section 6.1.4). These errors could accumulate into significant positioning errors. The error correcting system (section 6.1.4) could not always correct for these larger errors and resulted in poor arm positioning and missed fruit. To reduce this problem arm movement acceleration was decreased. This increased axis torque to resist the applied fruit extraction forces. These modifications reduced the harvest rate to an average 0.45 fruit/second/arm. This was less than half the desired harvesting speed. The operational video (referenced in 9.1.1) shows footage of fruit extraction.

Laboratory testing provided few harvesting complications in comparison to the on-orchard environment. On-orchard variability and complexity resulted from a combination of environ-

mental conditions. These included canopy infrastructure, growing system and fruit occlusion. Extracting clustered fruit in laboratory conditions did not fully represent on-orchard conditions. Natural fruit clusters are often tightly packed, contain a large number of fruit and have a more varied height range, making them more difficult to harvest. The harvesting order rules (section 6.2.6) were refined to address and resolve some of these issues.

Another issue that occasionally occurred was hand entanglement. This would halt harvesting until the system recovered. Entanglement was automatically detected when there was a large discrepancy with the arm's expected position (refer to section 6.1.4 for arm position tracking). The additional entanglement forces on the hand cause the arm to lose track of its current position. However, it would not completely stop operation. Smooth-surfaced covers for springs and entrapment surfaces will help resolve this.

The hand has demonstrated successful extraction and handling results. The slim design and four-bar linkage improved canopy penetration and fruit grasping capabilities. Collision and entanglement were also reduced with this design. The modified V shaped slot to the grasping thumb ensured clean fruit extraction for all fruit sizes (assuming correct arm positioning). Improvements to the harvesting arms will deliver performance at the required commercial level. These improvements include:

1. Encoder position feedback to remove the issues of the arms losing track of their current position.
2. Low backlash gearboxes. Cost effective gearboxes have been sourced that deliver the required accuracy and lower cost (than previous versions sourced). Gearboxes remove assembly time, cost and complexity as well as minimise the number of components for manufacture.
3. Monocoque structures for the articulating arm sections. These will be laser cut and folded from sheet metal to provide reduced weight and increased torsional rigidity
4. Rubber covered hand sides. This will remove the chance of damaging fruit from any hard or sharp edges.

Fruit can move during harvesting. Removing one fruit contacting another may result in the remaining fruit shifting. This change in fruit position is not updated by the current machine vision system. Consequently, the remaining fruit may not be harvested or could be damaged. Rubber sides to the hand's thumb and finger could help correct for small movements. This would move fruit into the middle of the hand's grasp, providing clean extraction. Machine vision fruit tracking during harvest is not practical as the hand and arm often block the camera's view.

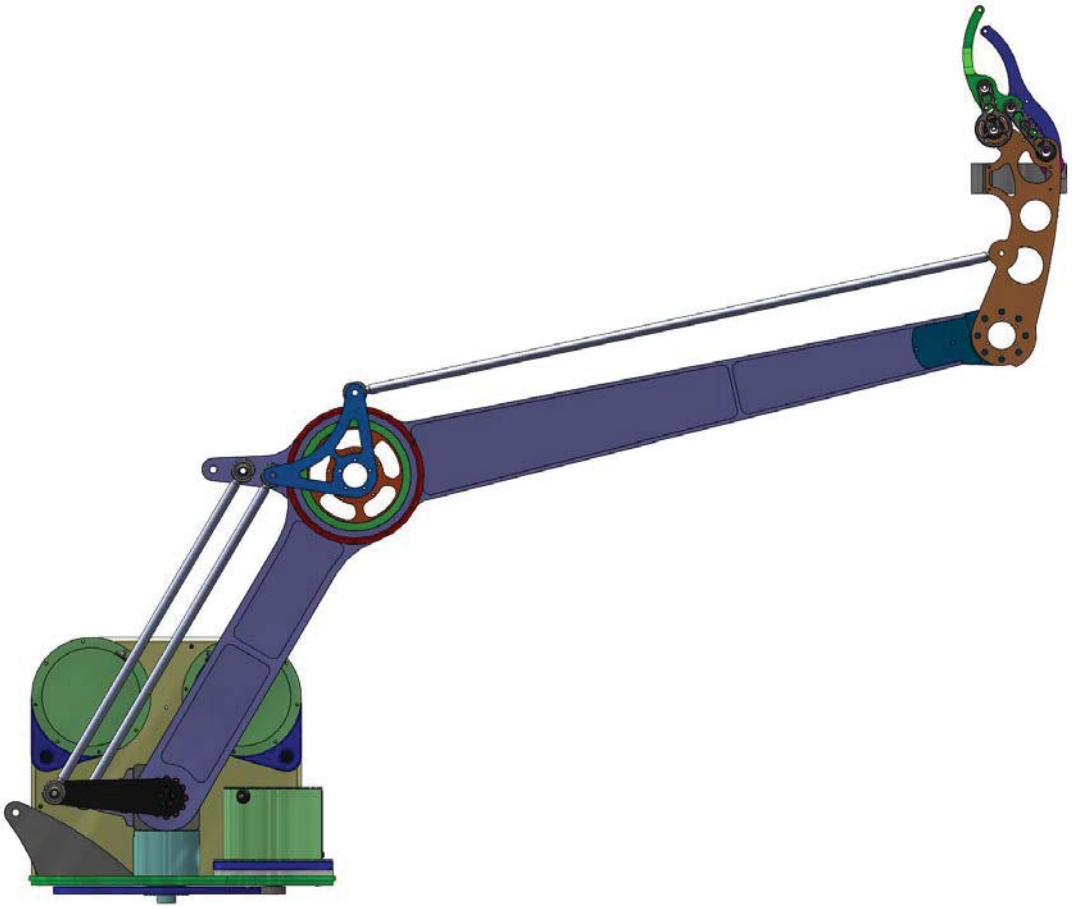


Figure 6-29: CAD model of the robotic harvesting arm

6.3 Robotic Storage Bin Filling Arm for Fruit Handling

A gentle robotic fruit handling system is required to fill the fruit storage bins. When people harvest, fruit damage can occur when their harvested bagged fruit are released into the bin. A robotic system could maintain fruit quality throughout this process and create a commercial point of difference.

Bin filling systems are currently available for packhouse environments. Generally, they need a large operational space, are heavy and are designed to fill stationery bins on stable and level ground. The size of the current systems is the biggest limitation to their on-orchard use. This is because the bin filling system is required to operate under canopy and within the ATP's structure. In addition, binned fruit will move during locomotion, especially on uneven terrain. Therefore, a dynamic fruit layering and sensing system is required. This dynamic system needs to be capable of placing fruit anywhere in a bin. The fruit from the four harvesting arms are shifted to the bin filling arm using a series of conveyors.

The developed arm has a pocketed conveyor system to hold and move fruit. It uses distance-detecting sensors to determine its height from fruit within the bin. The arm moves over the fruit (or bin bottom), maintaining a safe height above. A drop height of less than 80mm helps ensure fruit quality. No collisions with fruit or bin structures are allowed. This arm also has to keep up with the harvesting rate of four fruit per second. However, unlike harvesting, it can also operate while the vehicle is moving to store fruit.

6.3.1 Depositing Arm and Placement Conveyor

The bin filling arm has three degrees of freedom. This allows it to move transversely across, build up fruit height within and reach out into the bin. The internal conveyor moves fruit from

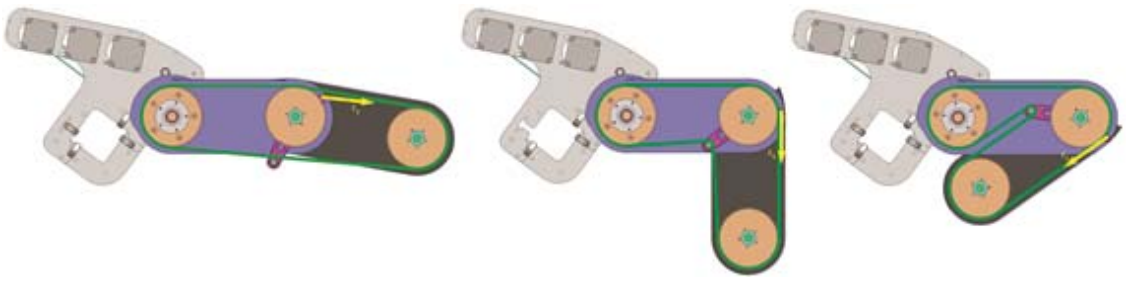


Figure 6-30: Cross-section of the robotic depositing arm. The chains path is shown in green, with tensioning arm (pink) and spring force F_s.

a delivery conveyor to the arm's distal end. Fruit leaves the bin filling arm and falls into the bin at a designated point. The fall height is kept less than the maximum allowable fruit drop height of 80mm. If the drop distance becomes greater than 80mm the conveyor stops until the correct layering height is achieved. This helps maintain fruit quality. A single belt forms the internal pocketed conveyor. The belt is draped between, and fixed to, support rails. This forms the pocket's shape. Two chains provide drive to the conveyor. The chains run in parallel and around the internal conveyor sprockets (Figure 6-30). The required chain length alters with arm position. Any chain slack is removed with a dynamic tensioning arm. This prevents jamming and maintains constant drive. Each chain runs around three sprockets (three pairs). Two pairs of sprockets rotate about the arm's axes pivot points (axis 2 and 3). The final pair are mounted at the arm's distal end.

Chain tensioning (shown in Figure 6-30), is performed by a short pivoting arm. This has a distally mounted idler sprocket. A tension spring applies torque to the tensioning arm. This in-turn applies a force to tension the chain. For small spring extensions the force applied to the tensioning arm will remain relatively constant. Spring extension is minimised by mounting its fixed end to third axis. The tensioning arm remains relatively consistently positioned with respect to the third axis. This allows shorter springs to be used to maintain consistent chain tensioning.

Figure 6-31 Depositing arm shown without its top fruit retaining enclosure

Fast arm movements are not required for bin filling. Therefore, arm mass was not critical during design. Design focused on fruit handling and maintaining fruit quality.

6.3.1.1 Transverse motion slide

'Axis 1' is linear motion only (prismatic joint). This transversely moves the arm across the bin. The transverse motion slide requires torsional rigidity to support the depositing arm's mass. The applied torque is maximised when the arm is horizontally outstretched. Standard slide and

guide systems are designed for vertical loading applications. Units with sufficient torsional rigidity are expensive. A novel slide design satisfies the strength and cost requirements (Figure 6-32).

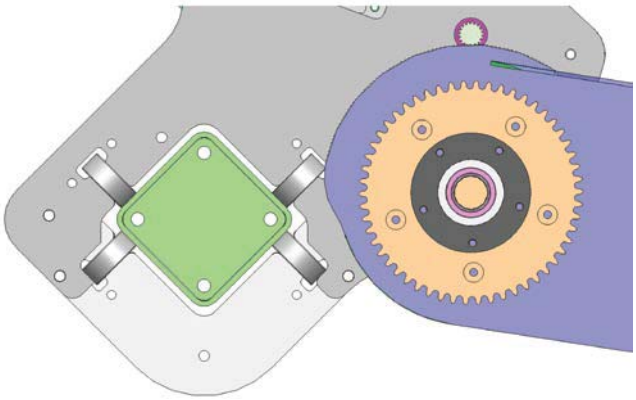


Figure 6-32: Cross-section of the transverse slide system. Stainless steel square hollow section shown as green (endplate also green). Note bearing contacting outer edge.

The slide design uses standard square hollow section (SHS) tube. This provides load and torsional rigidity. The SHS member is oriented at 45° to the horizontal, allowing debris to easily clear. Eight bearings (two groups of four) contact the SHS tube. These are positioned furthest from the tube's centre to minimise their loading (from torsional reaction force). A small interference fit (0.1mm) between the bearings and SHS removes any torsional backlash. The SHS tube accommodates this small elastic deformation. These combine to provide a rigid and stable slide platform for Axis1.

6.3.2 Power Transmission to Axes'

A stepper motor drives each axis. Reduction gearing provides the required output torque and speed.

Axis One (A_1), Transverse Motion: A split belt arrangement moves the bin filler's carriage (Figure 6-33). Either end of this belt is connected to the carriage. Driving the belt in either direction moves the carriage. A large tension spring in line with the belt removes any belt slack. The spring's tension is higher than that applied from the motor. Therefore, it stretches minimally while moving the carriage. It settles back to a constant neutral position upon completing motion, keeping accurate position.

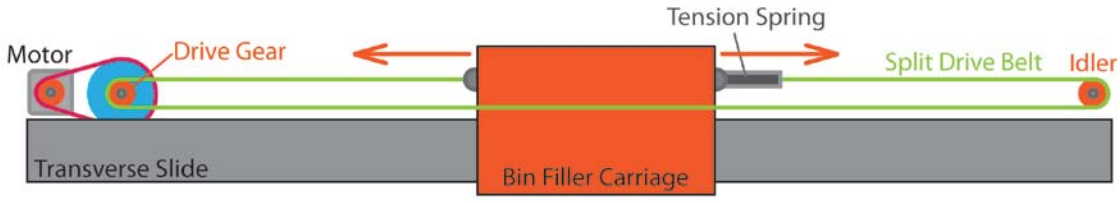


Figure 6-33: Axis One split belt drive

Axis Two (A_2), The Upper Member: This axis experiences the greatest reaction torque from the depositing arm's mass. Therefore, it requires the largest drive gear reduction. A 5M timing belt/pulley reduction (3.6:1) in conjunction with a pinion/spur gear reduction (11.1:1) is used. This provides a 40:1 gear reduction. The arm's mass preloads this axis, minimising positioning backlash.

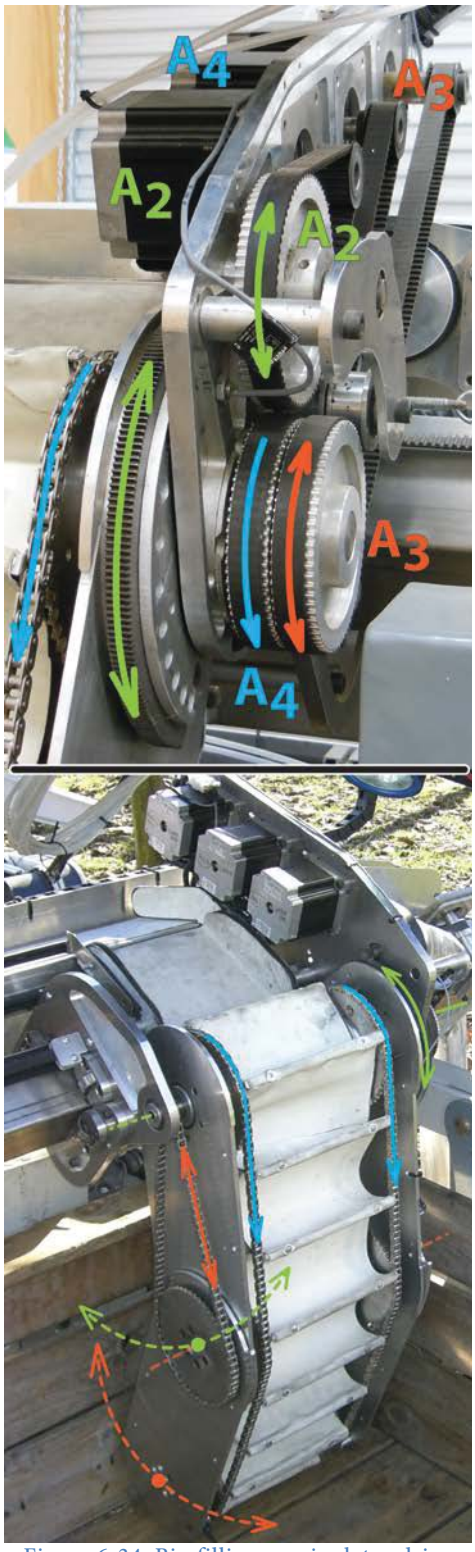


Figure 6-34: Bin filling manipulator drive

Axis Three (A_3), The Lower Member: Another double gear reduction drives A_3. A 5M timing belt drives the shaft concentric to A_2's rotation point (3.6:1). The shaft has a sprocket attached at its opposite end. This sprocket drives a fixed sprocket on A_3 via a chain (6.75:1). A 24.3:1 gear reduction results. A spring loaded chain tensioner aids backlash minimisation. This combined with the arm's mass helps remove backlash and increases positioning accuracy.

Axis Four (A_4), The Conveyor Drive: A 5M timing belt reduction drives the conveyor chains (Gear Ratio 3.6:1). The driveshaft is concentric to A_3's drive shaft (Figure 6-35, Figure 6-34). Motion accuracy is not required. Therefore, no backlash elimination is used.

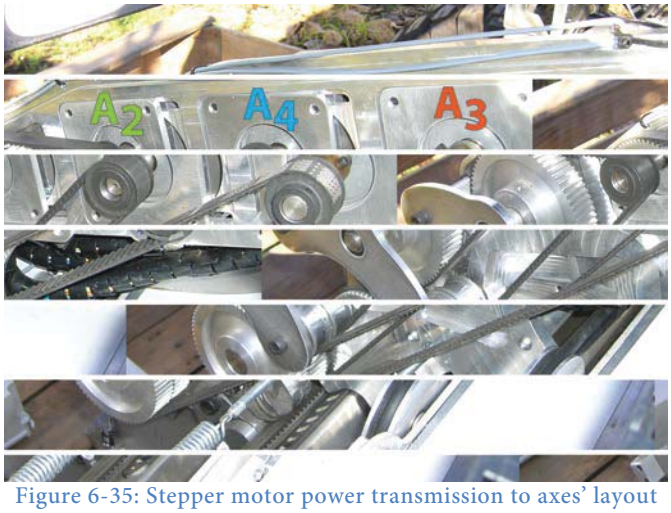


Figure 6-35: Stepper motor power transmission to axes' layout

6.3.3 Fruit Detection Antennae

Four distance detecting sensor arrays are mounted to the arm's end. These sensors are switching devices. When an object is within a predefined distance their output switches. Each array has sensors mounted at 30mm intervals. A distance array is mounted on each depositing arm side. This allows a safe fruit depositing height to be maintained for any arm motion direction. A parallel linkage arrangement (pulleys P_1, P_2 and P_3, with cable, Figure 6-36) keeps the sensor arrays parallel to the bin's bottom. Pulley P_1 is fixed, where as P_2 and P_3 are free to rotate. P_2 is a double pulley. Two distance thresholds allows for three possible distance states:

1. Too close
2. Too far
3. Safe fruit layering height (distance between states 1 and 2)

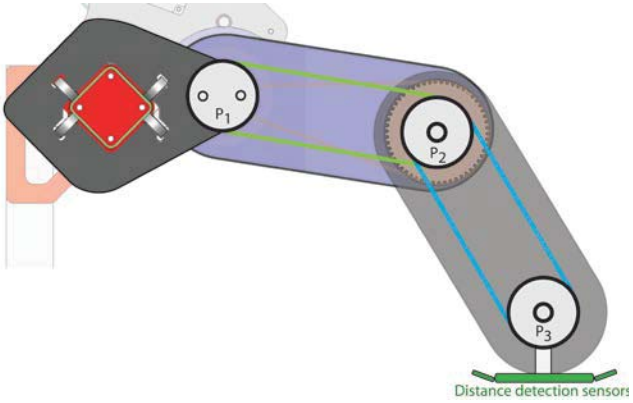


Figure 6-36: Fruit detection sensor arrangement

The arm's controller adjusts the depositing arm height to maintain a safe distance. This is dependent on the current distance state. Fruit is added to the bin while in the 'safe' state.

6.3.4 Depositing Arm Kinematics

Three-dimensional arm space positioning is required. To achieve a specific end position, each axis angle requires calculation. The arm's home position is defined as:

1. Axis1 (A_1): Far right of the transverse axis
2. Axis2 (A_2): Parallel to the arms y axis
3. Axis3 (A_3): Perpendicular to A_2 (parallel to the Z axis)

A_1 is a linear motion along the transverse slide. Motion is constrained to the x direction. The required step count, η_1, is determined by,

$$\eta_1 = \frac{x}{d_\eta} \tag{45}$$

where, d_η is the distance travelled per stepper motor step. The gear ratio and number of stepper motor steps per revolutions sets this value.

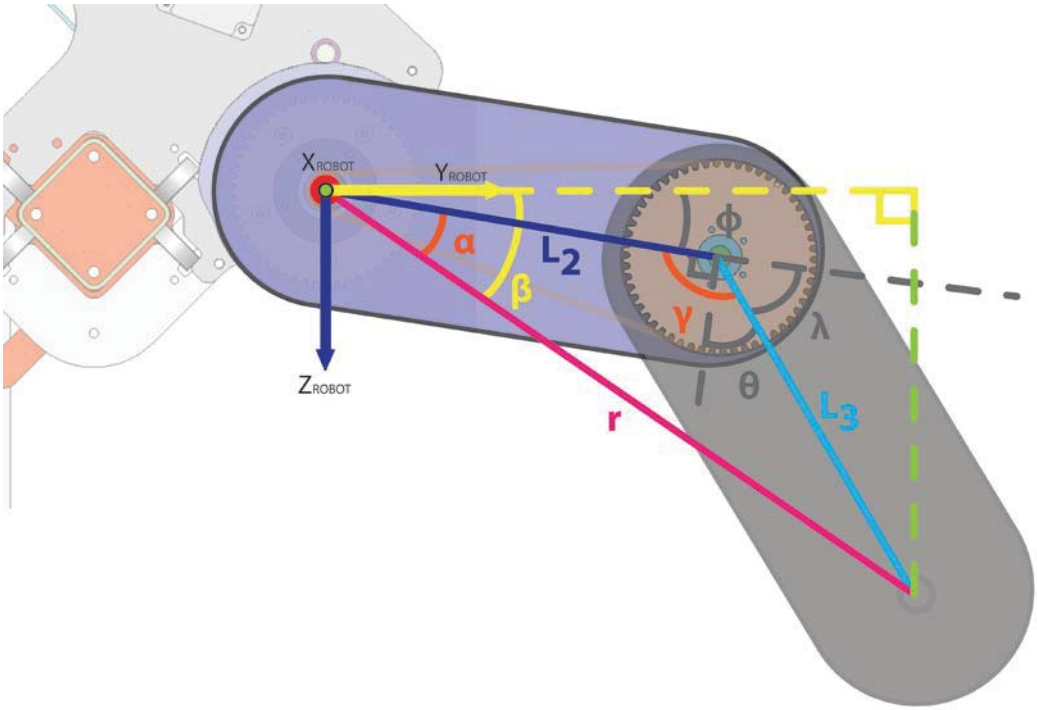


Figure 6-37: Bin filling arm kinematic nomenclature

A_2 and A_3 move in unison through the y – z plane. Axes angles (ϕ and θ) are calculated using the axes lengths (L_2 and L_3). First, the radius, r, of the end position is required. Where,

$$r = \sqrt{z^2 + y^2} \tag{46}$$

this is at an angle to the horizontal β, when $y \neq 0$,

$$\beta = \tan^{-1}\left(\frac{z}{y}\right) \tag{47}$$

if $y = 0$,

$$\beta = 90° \tag{48}$$

The first internal angle, γ, is found by rearranging and substituting into the law of cosines,

$$\gamma = \cos^{-1}\left[\frac{r^2 - {L_3}^2 - {L_2}^2}{-2L_3L_2}\right] \tag{49}$$

where $L_2 = L_3$,

$$\gamma = \cos^{-1}\left[\frac{r^2 - 2{L_2}^2}{-4{L_2}^2}\right] \tag{50}$$

The second internal angle, α, is found by rearranging and substituting into the law of cosines,

$$\alpha = \cos^{-1}\left[\frac{{L_3}^2 - {L_2}^2 - r^2}{-2L_2r}\right] \tag{51}$$

where $L_2 = L_3$,

$$\alpha = \cos^{-1}\left[\frac{r}{2L_2}\right] \tag{52}$$

Knowing angles α and β, angle ϕ can be calculated. This is the angle of A_2 from its 'home' position.

$$\phi = \beta - \alpha \tag{53}$$

and angle θ is,

$$\theta = 90 - \lambda \tag{54}$$

Axes angles are converted to step counts, allowing arm positioning. The step count for A_2 is η_2,

$$\eta_2 = \frac{\phi}{\omega_{\eta 2}} \tag{55}$$

where, $\omega_{\eta 2}$ is A_2 degrees rotated per motor step, η_2. Driving gear ratio and motor configuration determine this value.

A_3 step count, η_3, is also determined by the motor configuration and gear ratio. However, an interaction from A_2 exists. As A_2 rotates, the drive sprocket of A_3 remains fixed. This imposes a

rotation in A_3 proportional to any A_2 rotation. The angle error, θ_{E3}, is proportional to the sprocket/chain drive ratio (6.75:1).

$$\theta_{E3} = \phi_2 \times \frac{8}{54} \tag{56}$$

Accounting for the interaction, A_3 step count, η_3, is determined by,

$$\eta_3 = \frac{(\theta - \theta_{E3})}{\omega_{\eta 3}} \tag{57}$$

where $\omega_{\eta 3}$ is A_3 degrees of rotation for a single motor step. Axes step counts are used to position the arm to the desired position. See section 6.1.1.

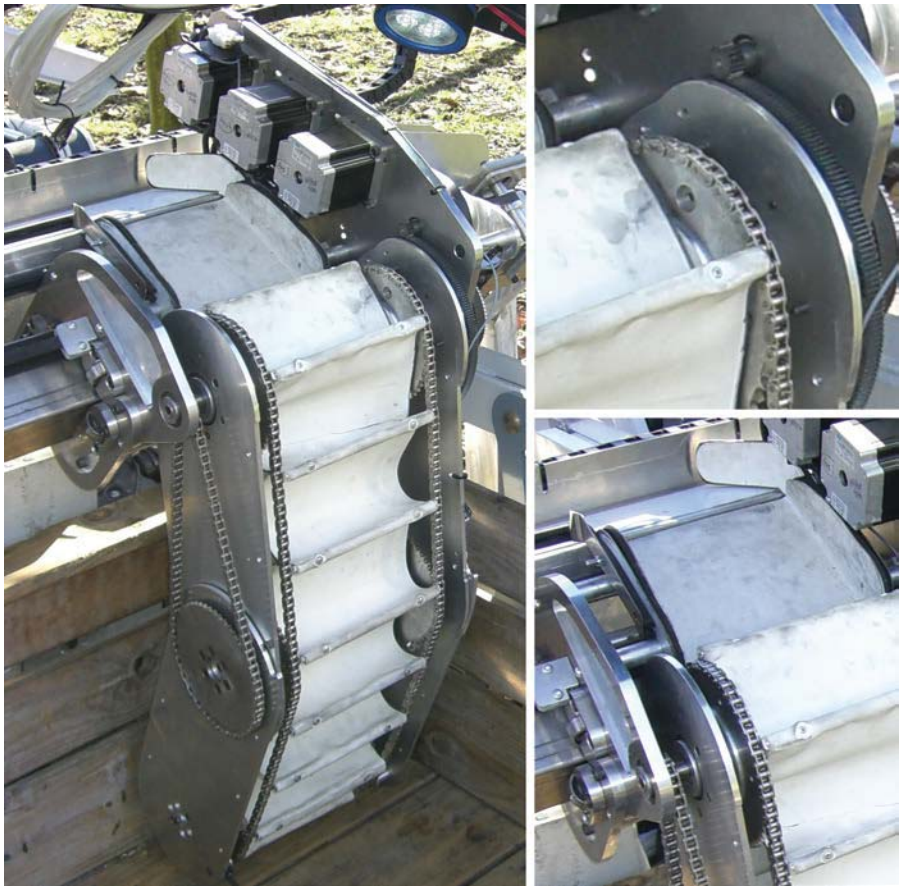


Figure 6-38: Left: Depositing arm. Right Top: Gear and pinion drive for power transmission to axis two. Right Bottom: Feed chute to depositing arm

6.3.5 Testing and Results

Testing of the bin filling arm was performed outside, but not on-orchard. The arm's sensors (monitoring safe depositing height) are not affected by variable orchard conditions, therefore, on-orchard testing was not required. Four people manually filled the fruit supply conveyors with fruit. This was done to approximately match actual fruit supply (four fruit per second). Four soft rubber rods were added to extend 80mm from the arm's end. This allowed the layering height to be visually monitored by people during testing. Successful fruit depositing was achieved. Fruit were not dropped more than eighty millimetres as set out in the handling specifications. However, issues with the arm's geometry were discovered.

The bin filling arm can only fill approximately three quarters of an eighteen-bushel storage bin. The pocketed conveyor drive chains restrict A_3 rotation (restriction to minimising γ, Figure 6-37). This stops fruit being deposited in the front-top of the storage bin, where the front of the bin is under the arms linear axis (A_1). Modification to the arm's geometry and mounting position will allow complete bin filling. An additional prismatic axis is the main modification (An_3, Figure 6-39). This will be integrated into the upper articulated arm and rotate to replace A_1 in the current design. The lengths of the articulated axes are also altered. In the modified design (Figure 6-39):

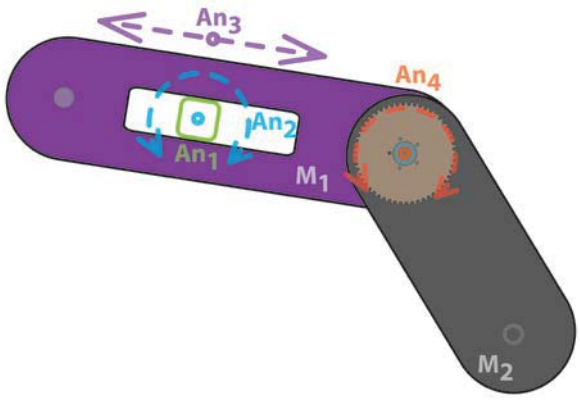


Figure 6-39: Modified bin filler arm design

1. An_1 is a linear axis (in and out of the page, similar to A_1 in the current design)
2. An_2 is a rotation of An_1's structure to change the angle of the upper arm member (M_1)
3. An_3 is the new prismatic axis to alter the effective length of M_1
4. An_4 rotates the lower arm member (M_2)

A predefined filling cycle is used to achieve uniform and progressive fruit layering. This process fills the bin in bands and helps minimise fruit damage. The arm currently runs at a constant speed while depositing fruit. This method achieved filling speeds faster than the targeted harvesting rate of four fruit per second. However, moving the arm at a constant speed can cause

fruit voids in the bin. This occurs when there is a break in the supply of fruit. An additional light beam sensor will monitor the fruit entering the system. This will allow the arm to halt motion when there are no fruit ready. A reduction to arm motion and fruit voids will result, providing even fruit filling and increase its efficiency.

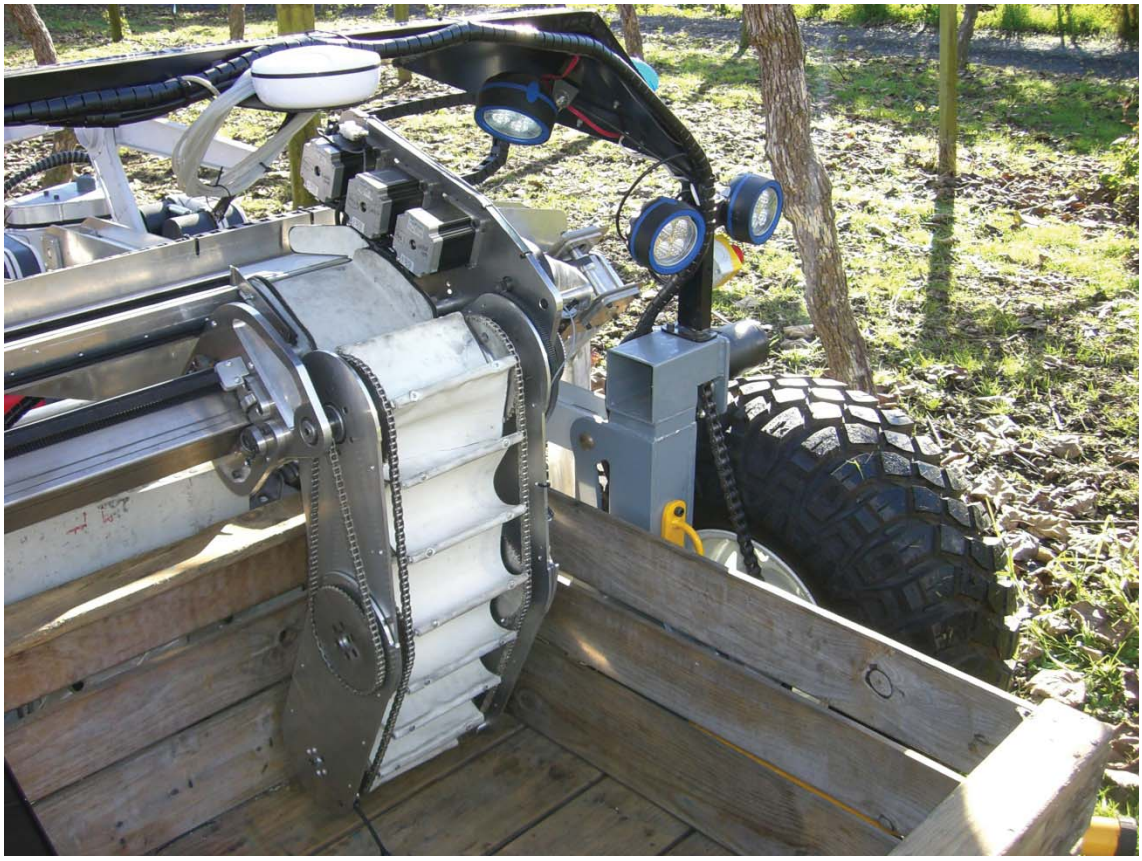


Figure 6-40: The bin filling arm on the AKH on-orchard

6.4 Fruit Handling Analysis

The AKH's handling ability has been compared to human harvesters, which are the current benchmark.

6.4.1 Methodology

Thirty fruit were taken from each key operational area (e.g. from the hand and conveyors). For comparison, thirty fruit were taken from a human's comparable process. Thirty fruit were used for each process as this equates to a storage tray. The trail took place outside the normal harvest season. This meant that the fruit were softer than normally expected, accentuating any damage found. Therefore, total damage is not as relevant as the comparison between man and machine.

Table 3: Man versus Machine comparable tasks

Human Action:	Robot Action:
Hand Extracting Fruit	Harvesting Hand Extracting Fruit
Picked Fruit into Picking Bag	Fruit Delivery Tube and Conveyor System
Picked Bag into Bin	Bin Filling Robot

For trial tasks, that were not fruit extraction, fruit were carefully removed by hand. These fruit were individually stored in fruit trays in preparation for testing. Testing was performed within minutes of their extraction. This allowed damage to be traced to a specific operational section. Running the fruit through all sections would not define damage causing areas.

The fruit was box stored in Plix trays after testing. Plix trays are the industry's standard storage tray. The fruit were kept at approximately 0°C. This follows industry storage recommendations. The storage method allowed fruit damage analysis of fruit damage over a three-month period, as damage can take time to become noticeable. An independent fruit quality expert inspected the fruit. Damage focused on:

1. Soft Fruit (e.g. normally not handling damage, can be over ripe or degrading fruit)
2. Soft Spots on Fruit (e.g. pressure damage from dropping, or fruit beginning to over ripen)
3. Physical Damage (e.g. puncture marks or surface scuffing)

An independent and experienced quality controller at Mount Pack and Cool Limited carried out quality control inspections. Industry quality standards were followed.

6.4.2 Results

The AKH handled more fruit during testing due to its additional handling process. Therefore, results are analysed on a percentage basis. Potential AKH fruit damaging elements have been identified, including:

1. Hard and sharp edges on the harvesting hands causing physical fruit damage
2. Uncontrolled fruit drop height from the harvesting hand into the fruit tube. This can cause pressure damage when the fruit contacts the tube wall
3. Fast fall descent down the fruit tube can cause some soft spots when the fruit decelerates at the conveyor
4. Hard conveyor corners require more padding – physical damage
5. Bin filling arm's open sides to the conveyor pockets. This exposes some of the mechanical conveyor system to the fruit

These elements will be modified to rectify the problem. Modification will further accentuate the AKH's ability to outperform people. They will also increase fruit quality outcomes. People damaging fruit during harvesting is a constant problem. This is because they are incentivised by the 'per bin' payment system to pick rapidly, not carefully. Fruit quality can further suffer when considering other human flaws. We are prone to fatigue and boredom while performing long repetitious tasks. This potentially reduces concentration and careful completion. Robots do not

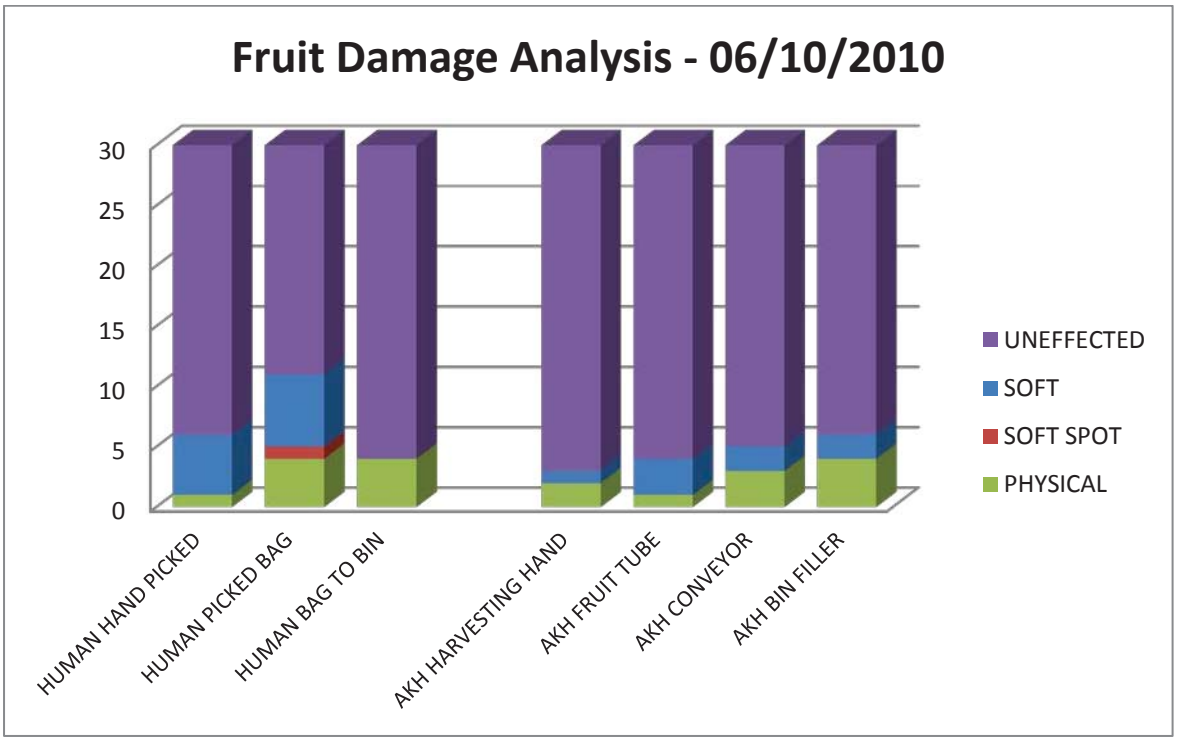


Figure 6-41: Total damage percentages across all man and machine processes

suffer these flaws. Once a standard is achieved, a robotic system will maintain that standard throughout operation. Trials that are more extensive will be conducted once the identified modifications are made.

The main areas for improvement are:

1. Harvesting Hands: Silicone mouldings will remove fruit exposure to hard/sharp edges during extraction.
2. Fruit Tube:
    - A reduction in the decline angle of the tube
    - Added integral slowing baffles to assist in slowing fruit descent to the conveyors
3. Conveyors: Further padding the transition areas between conveyors
4. Bin Filling: Improved fruit height tracking, making the drop height even smaller and fully enclosed conveyor pockets to stop fruit contact with mechanical systems.

Although a limited trial was conducted, the following results were obtained:

1. Areas for robotic improvement
2. Operational fruit handling comparisons between man and machine

The results showed that AKH inflicted less damage than human harvested fruit harvested by humans. This was achieved across all damage categories and most sections.

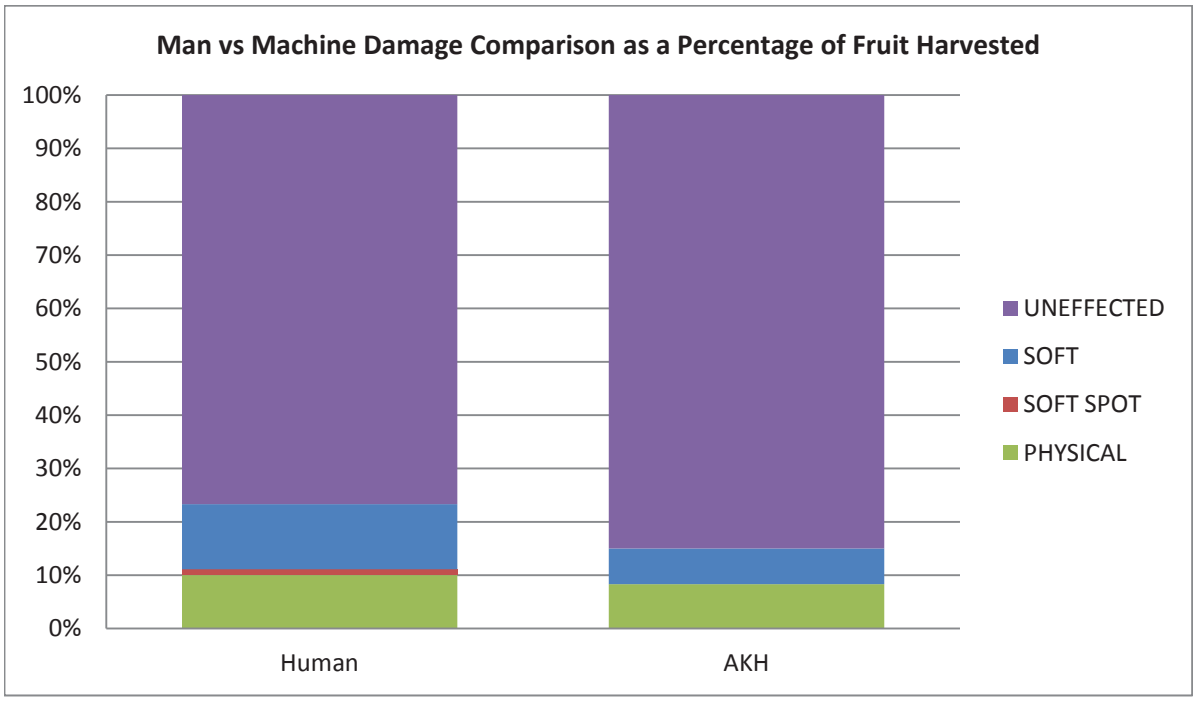


Figure 6-42: Accumulated fruit damage as at the final test (06 October 2010)

6.4.3 Conclusion

Both harvesting and bin filling arms have demonstrated successful core functionality. The fruit handling trial has delivered satisfactory initial results. The consistently gentle fruit handling (compared to humans) is a beneficial commercial point of difference. This is a significant step towards commercial uptake where growers will benefit from higher financial return. Most significantly, these results are delivered in the variable on-orchard environment. Although the system has demonstrated successful operation, some limitations have been identified. These will be resolved in future development with the following modifications:

1. Rubber covered hand sides. This will encapsulate the fruit during clasping and will accommodate small position errors from fruit swing (see comments above). It will also remove any hard and sharp edges, reducing potential fruit damage.
2. Arm position feedback for both systems. Incremental or absolute encoders will eliminate position loss. Position tracking with the slot array (section 6.1.4) was successful in the synthetic laboratory environment. However, it did not perform adequately when presented with on-orchard complexities. These arose from the increased fruit extraction force and some vine entanglement.
3. Low backlash arm gearboxes (for both systems): the presented drive system has operated successfully and reliably. However, machining and assembly complexity increase production cost and complexity. Low backlash gearboxes are increasingly available and have significantly reduced in price. This is a practical and cost effective design modification and simplify arm production.
4. Harvest arm monocoque structural design. This will increase the torsional rigidity of the arms and reduce system inertia. Laser cut and folded sheet metal parts will be used to achieve this design. Production cost reduction will result.
5. Bin filling arm fruit monitoring. Monitoring fruit entering the pocketed conveyor (with light beam sensor) will minimise arm movement. It will also provide even fruit distribution.
6. Bin filling arm geometry alteration. Modification to this arm's mounting position and articulation geometry is required to achieve complete bin filling.
7. Integration of the fruit handling improvements, outlined in section 6.4.2.

Arm modifications to achieve suitable performance for commercial operation have been identified. Further harvesting arm efficiency will be gained by improving the scheduling algorithms. Although the harvesting rules, set to avoid fruit damage, limit the amount of optimisation possible, gains can still be made. More efficiently dividing the fruit between the arms will give the greatest efficiency gains. The travelling salesman solution will be adapted to have four 'salesmen' (arms), with limited 'sales area' (arm operating area). This will restrict what fruit each arm can reach, increasing the optimisation problem's complexity. Each cluster will have a weighting associated, which relates to the number of fruit. This will dictate the amount of time required to harvest the cluster and the demand on the 'salesman'. Large fruit clusters might also need to be split between arms. In this case, one of the arm's will need to harvest its part of the cluster first. This will avoid breaking the harvesting rules. An improvement to operational efficiency and commercialisation potential will result.

Although some operational results have not met the requirements, it is expected that the outlined modifications would deliver realistic solutions. Further development, including improved scheduling algorithms, will deliver commercial level performance and operation. A better understanding of the requirements and complexities has been obtained by testing in real-world conditions. This provides realistic conclusions based on actual current performance, where modifications are planned to overcome the identified issues. The performance differences between laboratory and orchard environments shows the importance of real-world testing. Proving ability in the laboratory, does not guarantee that same ability in real-world conditions. The testing to date has been an important step to commercialisation. Although arm performance is not at full commercial levels, the testing has identified areas for improvement to obtain the necessary performance. Future development will implement these changes to achieve full operation at commercial levels.

Figure 6-43: The AKH on-orchard

7. Autonomous Transportation Platform

Autonomous movement and navigation to harvesting sites is required for complete on-orchard operation. The autonomous transportation platform (ATP) provides this ability by moving the harvesting equipment to their on-orchard operating positions. Additionally, the ATP provides fruit storage bin handling. Combined, these functions will allow efficient autonomous operation.

Autonomous operation requires robust and reliable real-time interpretation of the orchard environment. This benefits operation and commercialisation by removing the human driver or operator, eliminating the costs of:

1. Human labour and error factor
2. Specialist driver or operator training
3. The requirement of legislated breaks and operating hour limitations. This allows continuous operation, increasing efficiency.

Future development will evolve the ATP's operation beyond fruit harvesting. This allows wider system use during the offseason to increase commercial payback. Intelligent devices will attach to the ATP to enable automation of many labour-intensive tasks. These tasks include pruning, fruit thinning, flower pollination and crop analysis.

Autonomous navigation is fundamental to the ATP. Reliably sensing and interpreting the environment poses the greatest challenge. Orchard navigation has added complexity when compared to structured and ordered industrial environments. Although there is some ordered layout, orchards are changeable, have unstructured clutter, and objects can blend and merge into the background or surrounding objects. This limits assumptions the amount of environmental assumptions that can be made. For example, if a boundary is detected in an industrial environment, it can often be assumed as a vertical wall, or that the floors are flat and even. These boundaries are also more distinctive in industrial environments. On-orchard edges can blend into the background making them hard to detect. Furthermore, industrial environments allow a wider range of sensing options to be used. These include local-position markers (often on or in

the floor) to allow predefined paths to be followed. This is not practical on-orchard as the AKH will be deployed in many different environments and this method requires large setup costs.

Although customised local-positioning markers are not feasible, kiwifruit orchards do offer some unique markers. The posts supporting the kiwifruit canopy are positioned with nearly regular spacing and can be used to determine the ATP's position. From this orchard position drive actions can be calculated.

Real-time locomotion calculation is necessary as no two orchards are the same. Although orchards generally follow a standard layout, there is significant variation from an ideal layout. The ideal layout would have all orchards conforming to exact uniform and linear spacing etc. Variability is not limited to post placement. The posts can be short, surrounded by grass, hidden by other objects (e.g. plant vines), curved, slanted or have defects. The environment differences and extent of variation significantly increases the complexity of autonomous navigation, limiting the range of suitable sensors and sensory interpretation strategies. Any sensor used must cope with environmental variations, must also be cost effective to maintain commercial viability of the AKH and must be safe.

Solutions requiring large setup infrastructure were not considered (e.g. navigational markers, buried wires and high cost base stations). Likewise, cost and reliability concerns ruled out the commonly used RTK-GPS (real time kinematic global positioning system, see section 7.3.2 for more comments), directing development to use the posts as local positioning markers.

Navigation using stereo machine vision fits the system requirements. Investigation into machine vision navigation (section 7.4) showed the difficulties due to variations in the orchard environment. The most significant comes from natural lighting. This often altered the object's appearance beyond the envelope of reliable machine vision recognition. When combined with object occlusion issues, which are commonplace, machine vision navigation was not going to provide reliable object detection. This being a key requirement, it was determined that another solution was required. To ensure reliable object detection, the machine vision system was combined with LIDAR (light detection and ranging) sensors. In the combined solution, LIDAR provides accurate and reliable object detection, where machine vision will identify the detected objects. This leads to better decision-making for navigation.

Although RTK-GPS is not suitable for navigation due to signal dropout (see section 7.3.2), standard GPS with a high gain antenna provides reliable positioning to ±3m. This allows navigation of the greater kiwifruit orchard area using GPS maps. These maps are created for crop management purposes. The maps provide position information of canopy areas, hedgerows, shade cloth, access points, roads and other orchard structures. This provides suitable orchard information to guide the ATP around the orchard. As accuracy and object detection is limited, this navigation will be refined using the LIDAR and machine vision sensors.

The extent of variability in environmental conditions was required to understand the environment and guide development. Industry consultation and on-orchard observation built a broad overview of expected conditions. Several design requirements were identified:

1. Twenty-four hour operation. The system is required to cope with all on-orchard illumination conditions, from bright sunlight to complete darkness. Although only twenty-hour operation will be used, the four-hour break period could occur at any time. This break is also weather dependent. Operation is limited to twenty hours per day because of settling dew overnight and to allow any necessary machine servicing of shifting between orchards.
2. Handle all orchard row variations including:
    a. Width and post spacing variation: rows widths alter between rows and along the same row. Rows can also taper in width. Row widths range from 3.5 to 5.5m
    b. Missing plants and posts within the rows
    c. Uneven row lengths
    d. Objects within rows
    e. Terrain variations
3. Four wheel drive with mud-grip tyres. Although most orchards are well drained, some orchard areas can become waterlogged. Also, areas of high use become muddy and slippery, creating traction difficulties.
4. Detection of random objects and structures

To achieve these requirements, a wheeled chassis structure and power source is required. In addition, systems for vehicular propulsion, sensing and drive calculation and control are required. The combination of these systems will allow for autonomous orchard navigation. These systems are presented in the following section of this chapter.

7.1 Chassis and Power Plant Overview

The chassis primarily needs to provide:

1. Locomotion: manoeuvrability, stability and traction
2. Attachment and transportation of other devices, for example robotic arms.
3. Bin lifting, holding and placement structures
4. Power generation

To achieve these requirements, the main chassis is fabricated from rectangular hollow section (RHS) steel tube. An RHS structure provides a rigid platform that is cost effective and easily fabricated for prototype development. The design was selected to provide ample space for additional systems, for example the arms, to operate. A monocoque, tubular or similar design was considered, however it could restrict operation of other systems. This includes obstructing or limiting the placement of their sensors (e.g. obstructing the cameras' viewing area or LIDAR sensors' scanning plane). These alternative designs are also likely increase fabrication complexity and production cost. Although the RHS chassis design has several advantages, the chassis mass and material costs increased. However, this cost is offset against reduced fabrication time. Once all task performing systems (and their required sensor positions) are confirmed, an alternative production chassis could be designed to reduce mass and allow unrestricted operation to all systems. The chassis structure can be changed without compromising its functionality and therefore it was not optimised for the prototype ATP.

Figure 7-1: Chassis during construction, pre finishing

Several laser-cut, flat-steel parts are used to simplify fabrication. These accurately laser-cut parts help align the necessary components. The rear bin lifter, rear wheel assembly, front steering trunnion assemblies and arm mounting stage benefitted from this process. The front steering trunnion assemblies are folded and rolled to form strong monocoque structures (see Figure 7-2). Standard hub axles are welded to the trunnion structures to support the wheels and drive. Similar hubs are also installed for the rear wheels. The hubs have sprockets attached to provide wheel drive.

Figure 7-2: (Left) Rear wheel and drive assembly, (right) Front wheel trunnion and steering assembly

The front wheel assemblies are attached to a pivoting front mount (blue unfinished steel, Figure 7-1). This ensures constant four-wheel ground contact, adding stability and traction. The mount pivots about a point in the middle of the two front wheels. Positioning this pivot point in line with the wheels' centre removes wheel side loading during rotation. Similar pivoting designs are commonly found on agricultural and horticultural tractors.

A four-wheel drive design was used to ensure adequate traction and stability. In addition, the wide wheels beneficially spread the vehicles mass over a larger area to reduce ground compaction. Vehicle ground compaction has been shown to adversely affect plant growth and production (Deurer, et al., 2010, Goh, et al., 2000). The rear wheels are larger to support the increased

weight of the fruit storage bin. These larger wheels also increase traction and smooth traversal of rough terrain.

The ATP is required to support fruit-storage-bin handling. This includes lifting, lowering and carrying a bin around the orchard. A pneumatic forklift mechanism is integrated at the rear of the chassis to provide this ability. Pneumatic actuators were selected for their load cushioning characteristic. This helps with gentle fruit handling by cushioning the bin while driving over uneven terrain. A four-bar-linkage lifting mechanism ensures the bin is kept level to the chassis throughout motion. The storage bin's centre of mass is set forward of the rear wheels' rotation axis. This helps maintain traction for drive and steering by ensuring the bin will not unload the front wheels.

All systems are powered from a 6.5kW Honda generator. In addition to providing electrical power, the generator's motor also drives a hydraulic pump which provides chassis propulsion and steering. A swash plate pump was selected to maximise system efficiency, only providing pressurised fluid on demand.

Figure 7-3: Assembled chassis after powder coating.

7.2 ATP Locomotion and Control

Control of propulsion and steering is required to achieve on-orchard driving. Navigation decisions (section 7.5) are converted to manoeuvre commands. A controller carries out these instructions, converting the commands into autonomous drive motion. Keicher and Seufert (2000) outline the three sections required to achieve autonomous driving.

1. A sensor to report deviation from the desired position (covered in later sections)
2. A controller to generate a correction signal according to the deviation
3. An actuator that manipulates the heading of the vehicle, so that the position error can be corrected using vehicular motion

Hydraulic systems for drive and steering actuation were selected for the ATP as they provide reliability, large power density and cost-effectiveness. An internal combustion engine provides power for the hydraulic swash plate pump. The pump provides fluid at constant pressure, altering fluid volume on demand. Altering fluid flow on demand reduces power consumption. The fluid from the pump is distributed by the valve block to the relevant motor(s). The block's solenoid valves are computer-controlled.

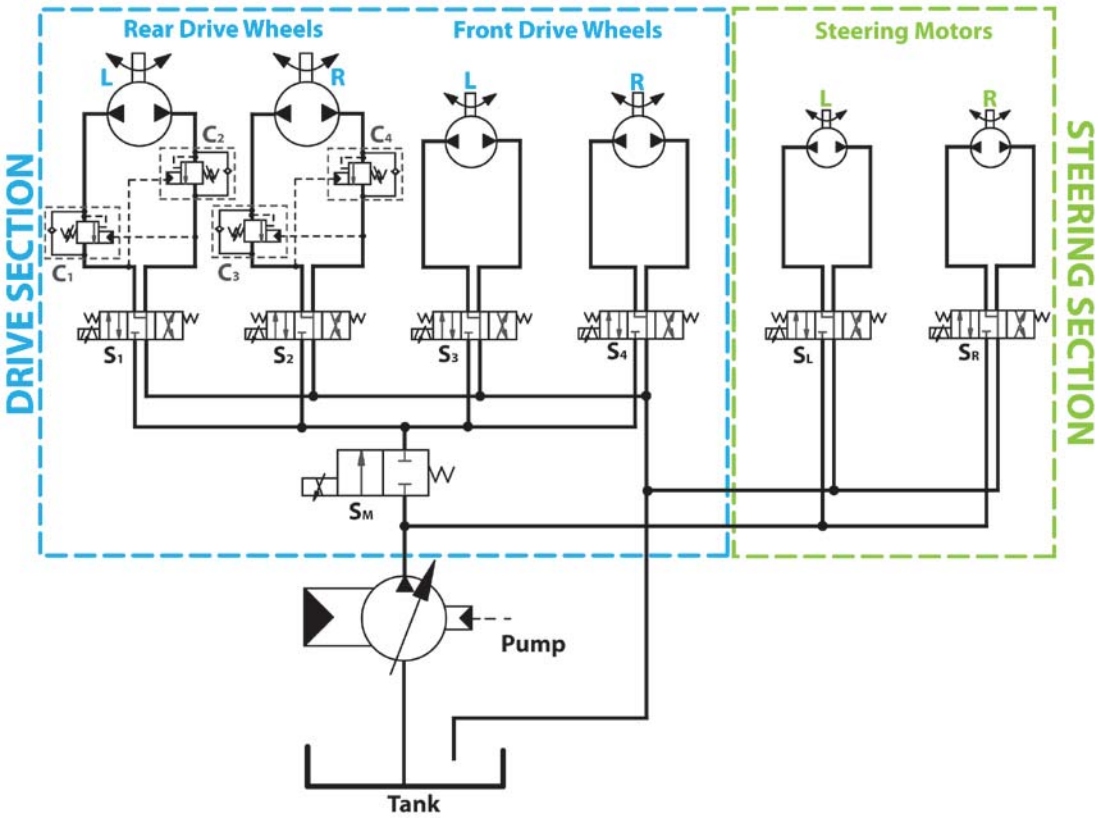


Figure 7-4: Simplified hydraulic drive system schematic

Steering is achieved using proportional solenoid valves to drive the motors rotating the front wheels (S_L and S_R, Figure 7-4). To achieve accurate steering control, 12-bit, absolute encoders report both wheel positions to the controller. The ATP can follow curved paths by calculating the direction of each of the front steering wheels. Each wheel's direction is calculated independently for a specific turn manoeuvre (see section 7.2.1), eliminating the need for differentials.

A bi-directional proportional valve controls each drive wheel (S_1 – S_4, Figure 7-4). Having proportional valves allows individual wheel speed control in low traction conditions. By comparing each wheel's actual rotational speed, to its desired speed, 'slippage' is detected. Wheel speed is determined from each wheel's incremental encoder. If 'slippage' is detected, action is taken to resolve the issue and maintain constant drive (see section 7.2.3). However, for general operation the hydraulic valves operate as tri-state devices to control direction (forward, off and reverse). A large proportional valve controls fluid flow to all drive wheels (S_M, Figure 7-4). The hydraulic fluid follows the path of least resistance, therefore, during a turning manoeuvre, when the outer wheels are rotating faster, the fluid flow increases to the outer wheels as they provide less resistance.

Counterbalance valves (C_1 - C_4, Figure 7-4) are installed in parallel to the drive motors on the rear wheels, these reduce motor 'free-wheeling' (e.g. the ATP rolling down a slope, unpowered). However, the counterbalance valves have a limited holding effect and the ATP can still slowly 'creep' on inclines. Therefore for parking, the front wheels splay, preventing any undesired rolling.

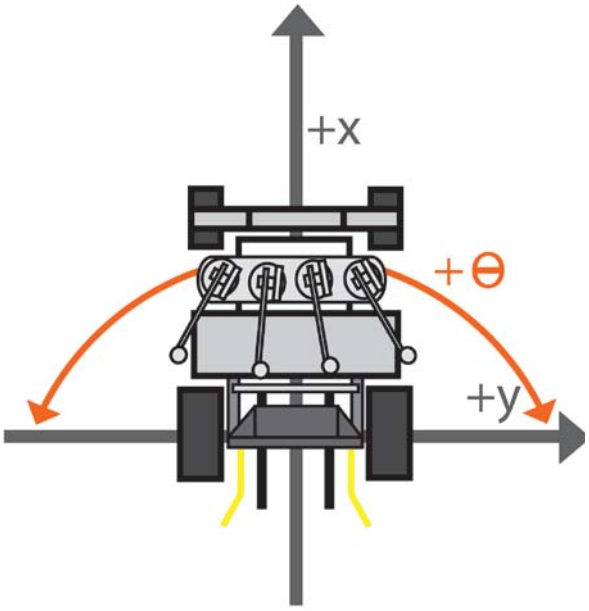


Figure 7-5: Chassis nomenclature for coordinates of motion (as viewed from above).

7.2.1 Steering Geometry Calculation

Accurate control of wheel direction is required to perform ATP manoeuvres. Unlike a motor vehicle, the ATP's steering wheels are independently set. Most vehicles use a steering linkage arrangement to set both steering wheels' angles with a single actuator. Having independent front steering wheels, without a mechanical linkage, allows tighter turning. Each wheel's direction is calculated from the desired manoeuvre's radius-of-curvature (R_C - Figure 7-6).

Steering geometry is calculated in real-time for every radius-of-curvature command. The wheel angles (σ_1 and σ_2, Figure 7-6) are determined from the ATP's chassis geometry. The ATP manoeuvres about the rear wheel axis. The distance of the rotation point from the centre of the rear wheels is the manoeuvre's radius-of-curvature. From this centre of rotation (*COR* - Figure 7-6) the steering angles are calculated. Each steering wheel's direction (σ_1 and σ_2) is perpendicular to a line projected from *COR* to the wheel's steering rotation point (blue and orange lines, Figure 7-6).

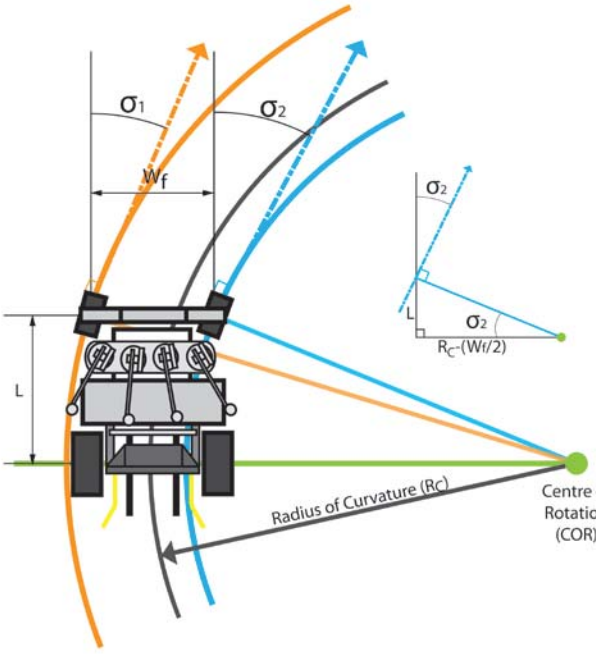


Figure 7-6: Steering Geometry Calculation

Therefore, each wheel's angle (σ_1 and σ_2, Figure 7-6) is,

$$\sigma_1 = \tan^{-1}\left(\frac{L}{R_C + \frac{W_f}{2}}\right) \tag{58}$$

$$\sigma_2 = \tan^{-1}\left(\frac{L}{R_C - \frac{W_f}{2}}\right) \tag{59}$$

The tighter the turn the more these angles deviate from each other. Once the angles are known, a PID control system keeps the steering wheels' pointed in the desired direction

7.2.2 Hydraulic Solenoid PID Control

Each solenoid value is independently computer-controlled so as to remove the need for an expensive hydraulic controller. A pulse width modulated (PWM) signal is used to control each solenoid valve through a pair of smart high-side power switches (Infineon, 2004); one for forward operation, the other for reverse. The smart FETs are designed to drive inductive loads and do not require free-wheeling diode protection like other FETs. They also provide current limiting, thermal shutdown, and overload and short-circuit protection.

Control of solenoid valves is more difficult than servo valves. This is due to their nonlinear output characteristic (shown in Figure 7-7) and non-responsive starting dead-zone. This dead-zone can slow system response when starting output, or changing direction. To overcome this issue the PWM has an offset value to ensure it is never within this dead-zone. Servo valves do not exhibit these flaws, however, they are approximately ten times the cost of solenoid valves. Consequently, a PID control system was designed and implemented to control the standard solenoid valves. This control program monitors feedback from wheel and steering encoders to provide feedback to the controller.

There are seven PID control loops concurrently operating on the control computer. However, only three of these are required in normal operating conditions. One for each steering wheel and one to control fluid flow to the drive wheels. The remaining four control loops are for each drive wheel. These are only used when traction control is required.

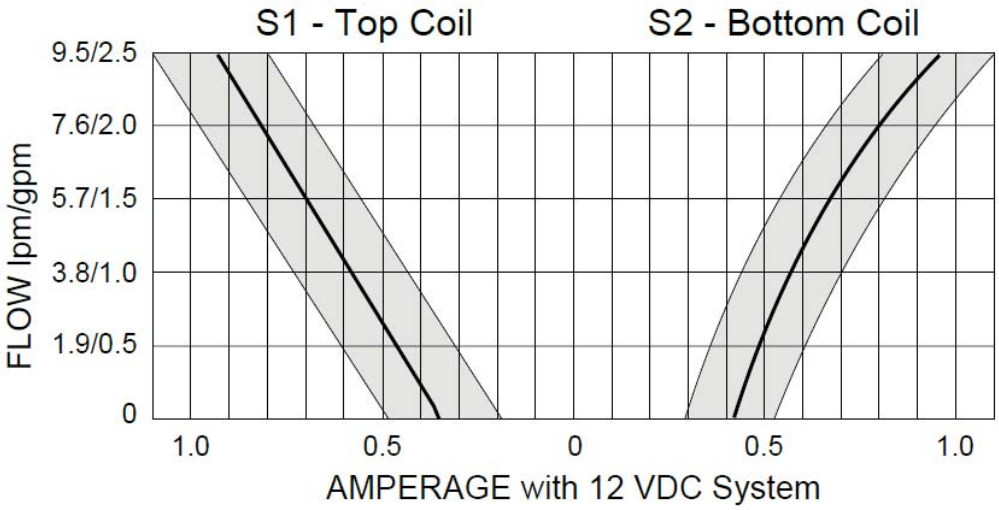


Figure 7-7: Solenoid valve fluid flow characteristics for an applied current (Robert, 2006)

7.2.3 Traction Control

In low traction conditions all fluid can flow though a single wheel. The fluid follows the path of least resistance, therefore a wheel that has lost traction ('slipping') provides a low resistance path. This would impede, or even stop, locomotion. Therefore, traction control was required. A system was designed and implemented to monitor and control each wheel independently. This ultimately increases the ATP's ability to drive in low traction and uneven terrain. Currently, orchard contractors' vehicles often suffer loss of traction on-orchard. The commonly used Atom sprayers (Andreoli, 2005) also have hydraulic drive, but do not have traction control. This has led to accidents, vehicles becoming stuck and inefficient operation.

To implement traction control, the speed of each drive wheel is required. This is calculated from an encoder mounted to each wheel. This speed is compared to its theoretical speed to detect any loss of traction. Theoretical wheel speeds are calculated from their circumferential path (Figure 7-6) relative to the ATP's set speed. While driving in a straight line (i.e. $R_C = \infty$) the wheel velocities are all the same. Their proportional rotational rates during a turning move are calculated by,

$$FL_C = 2\pi\left(\frac{L}{\sin(\sigma_1)}\right) \tag{60}$$

$$FR_C = 2\pi\left(\frac{L}{\sin(\sigma_2)}\right) \tag{61}$$

$$RL_C = 2\pi\left(R_C + \frac{W_r}{2}\right) \tag{62}$$

$$RR_C = 2\pi\left(R_C - \frac{W_r}{2}\right) \tag{63}$$

Where the subscript C denotes a circumference measurement and FL denotes the front left wheel etc. C_C is the chassis circumferential path distance. Where,

$$C_C = 2\pi R_C \tag{64}$$

The individual wheel required velocities, v_w, are proportional to the chassis setpoint velocity, v_{sp}, where W_c is a specific wheels circumferential distance (e.g. FL_C),

$$v_W = v_{SP} \times \frac{W_C}{C_C} \tag{65}$$

From this, an error in wheel velocity, v_{EW}, is calculated using (66). The actual velocity, v_A, is determined from wheel encoder information.

$$v_{EW} = v_A - v_W \tag{66}$$

This provides an error relative to ATP overall velocity. Traction control is only required when one, two or three wheels are rotating faster than 5% of its setpoint. This indicates a loss of traction to those wheels. The corresponding wheel(s) valve is briefly closed to its lowest response level (allowing minimal fluid flow and increasing resistance). This prevents the counterbalance valves locking the wheel and resisting motion. The fluid to the slipping wheel is diverted to the non-slipping wheel(s). When the non-slipping wheel(s) speed increase, the slipping wheel(s) valve is gradually reopened. The slipping wheel(s) are individually PID controlled until the speed error is corrected and traction regained. Only PID control of the main flow solenoid (S_M, Figure 7-4) is required under 'normal' drive conditions. This is because the hydraulic fluid follows the path of least resistance and resolves any differential balancing of wheel rotational speeds for driving curved paths.

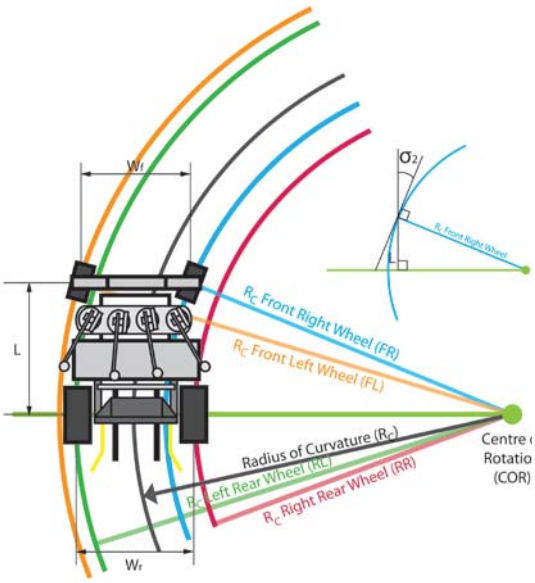


Figure 7-8: Depicts the functional layout for the calculation of individual wheel speed with respect to set point speed

7.2.4 Chassis Manoeuvres

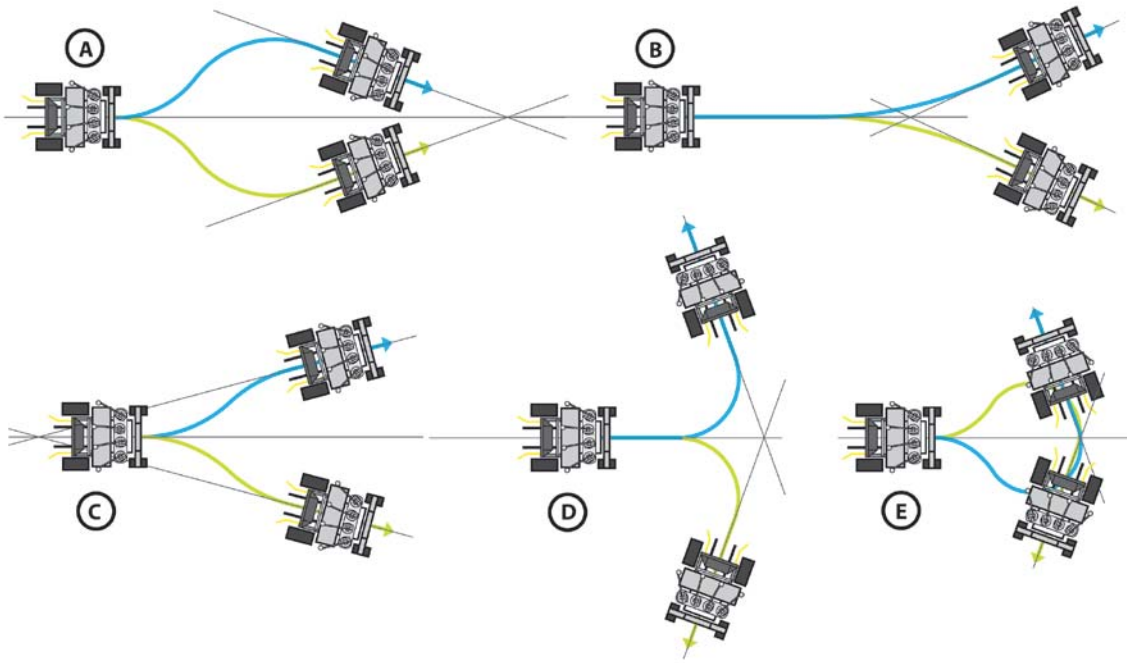


Figure 7-9: Shows the five possible chassis manoeuvres. A – Positive intercept serpentine, B – Single curve, C – Negative offset serpentine, D – Double back single curve, E – Double back serpentine.

Movement in the orchard generally involves line-following routines of lines inferred by environment sensing. For line following the drivelines are inferred from sensor analysis (see section 7.5). However, there are instances where speciality manoeuvres are required. These are mainly used for object avoidance and docking with fruit storage bins. Figure 7-9 shows five manoeuvres (and their mirrors) used for point A to B manoeuvres. An additional three-point-turn manoeuvre is also described (section 7.2.4.6). This three-point-turn move allows turning at the end of a row, allowing navigation back along the same row.

To determine which manoeuvre to use, the heading difference (at the start and end of the move) and the intercept of these headings are used. A tangential approach is often required to the end position. This ensures the ATP is in-line with the desired end heading upon completing the manoeuvre. A tangential approach is also desirable for bin docking as it ensures the bin lifting forks are aligned before engaging the bin. In the manoeuvre examples (in the following sections), all manoeuvres are shown with tangential end approaches included.

7.2.4.1 A - Positive Intercept Serpentine

A positive intercept serpentine is used when the intercept of current and desired headings is in front of point *e*. The absolute difference between current and destination heading also has to be less than 90°. Point *f* (Figure 7-10) is the desired end position. A tangential approach segment (*e* to *f*) ensures that the ATP is aligned to the correct end-heading when point *f* is reached.

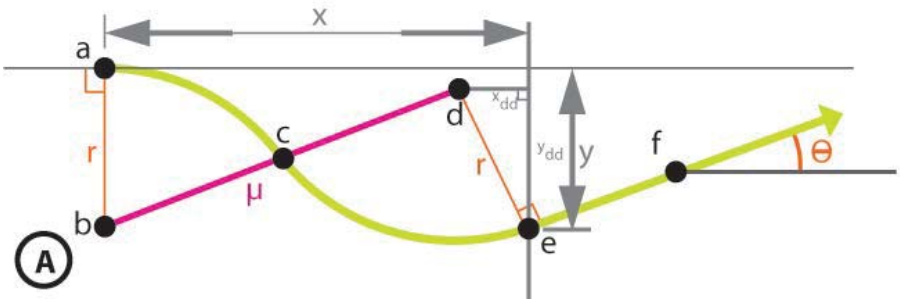


Figure 7-10: Positive Intercept Serpentine calculation (from Figure 7-9 – A).

Where length μ is (defining *a* as the origin, (Figure 7-10),

$$2r = \sqrt{(x_d)^2 + (y_d - r)^2} \tag{67}$$

The position of point *d* is the centre of the second arc segment, where

$$y_d = y - r\cos\theta \tag{68}$$

$$x_d = x - r\sin\theta \tag{69}$$

(68) and (69) are substituted into (67) to give,

$$2r = \sqrt{(x - r\sin\theta)^2 + (y - r\cos\theta - r)^2} \tag{70}$$

Rearranging and expanding,

$$4r^2 = x^2 + r^2\sin\theta - 2rx\sin\theta + y^2 + r^2\cos^2\theta + r^2 \\ -2ry\cos\theta - 2ry + 2r^2\cos\theta \tag{71}$$

factorising for *r*, noting that the expansion of (71) is a quadratic, producing,

$$2r^2(1-\cos\theta)+2r(x\sin\theta+y\cos\theta+y)-x^2-y^2=0 \tag{72}$$

using the quadratic equation solution,

$$r=\frac{-(x\sin\theta+y\cos\theta+y)\pm\sqrt{(x\sin\theta+y\cos\theta+y)^2+2(1-\cos\theta)(x^2+y^2)}}{2(1-\cos\theta)} \tag{73}$$

where r is the radius of curvature for the required manoeuvres two arc sections.

7.2.4.2 B - Single Curve

A single curve is performed when the end heading intercepts the current heading between the ATP's current and final x position. An absolute heading difference of less than 90° is also required.

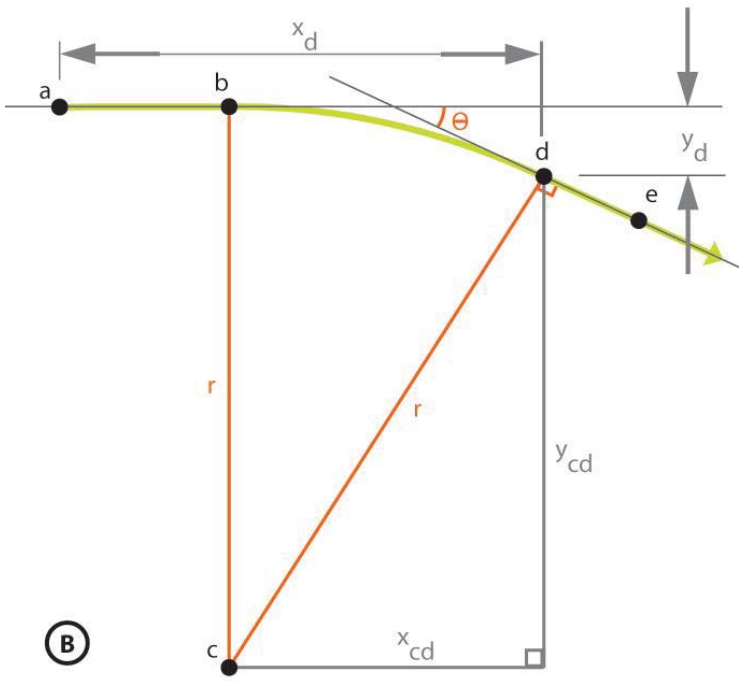


Figure 7-11: Single Curve path calculation (from Figure 7-9 – B)

Calculating the x and y components between the arc's centre, c. The centre is positioned perpendicular to the curves start b, and arcs end, position, d. Therefore,

$$y_c = y_d + r\cos\theta \tag{74}$$

substituting $y_c=r$, and solving for radius of curvature, r,

$$r = \frac{y_d}{(1-\cos(\theta))} \tag{75}$$

Where the linear approach section to get to position *b*, before performing the curve has a distance,

$$x_b = x_d - r\sin\theta \tag{76}$$

7.2.4.3 C - Negative Intercept Serpentine

A negative intercept serpentine manoeuvre is defined when the intercept of the current and end headings is behind the ATP's current position. An absolute heading difference of less than 90° is also required (see Figure 7-12).

Defining known variables, from manoeuvre conditions (Figure 7-12), where, $x_b = 0$, $y_b = r$. Therefore, from Pythagorean Theorem,

$$2r = \sqrt{(x_d)^2 + (y_d - r)^2} \tag{77}$$

Defining the components of D as,

$$y_d = y - r\cos\theta \tag{78}$$

$$x_d = x + r\sin\theta \tag{79}$$

Substituting into (77),

$$2r = \sqrt{(x + r\sin\theta)^2 + (y - r\cos\theta - r)^2} \tag{80}$$

Squaring both sides and rearranging to quadratic equation form,

$$2r^2(1-\cos\theta) + 2r(y + y\cos\theta - x\sin\theta) - x^2 - y^2 = 0 \tag{81}$$

Solving the quadratic, to get radius *r*,

$$r = \frac{-(y + y\cos\theta - x\sin\theta) \pm \sqrt{(y + y\cos\theta - x\sin\theta)^2 + 2(1-\cos\theta)(x^2 + y^2)}}{2(1-\cos\theta)} \tag{82}$$

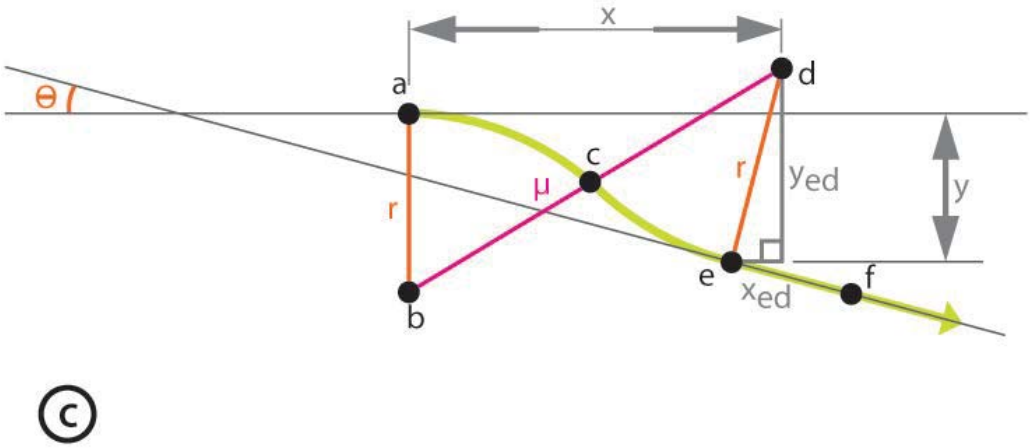


Figure 7-12: Negative Intercept Serpentine (from Figure 7-9) path calculation.

7.2.4.4 D - Double Back Single Curve

The double back single curve (Figure 7-13) is a manoeuvre performed when the absolute heading difference (current minus destination headings) is greater or equal to 90°. The headings intercept must also be beyond the *x* value of the end position.

Defining lengths,

$$y_{dc} = r\cos\theta \tag{83}$$

$$x_{dc} = r\sin\theta \tag{84}$$

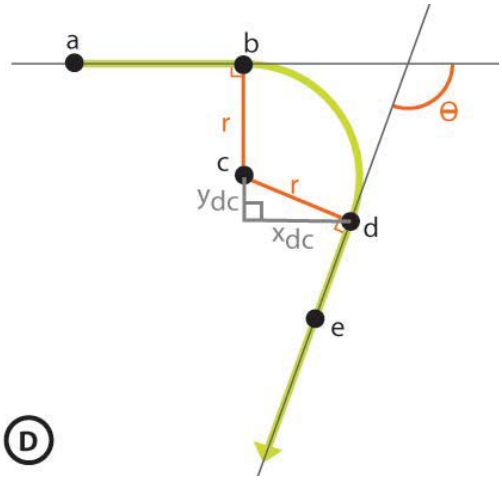


Figure 7-13: Double Back Single Curve path calculation (from Figure 7-9, D)

Then,

$$y_c = y_d + r\cos\theta \tag{85}$$

$$x_c = x_d - r\sin\theta \tag{86}$$

Since $r = y_c$,

$$r = y + r\cos\theta \tag{87}$$

Therefore,

$$r = \frac{y}{1-\cos\theta} \tag{88}$$

With a linear approach to the radius (transition *a* to *b*, Figure 7-13) of distance d_{ab},

$$d_{ab} = x_d - x_{dc} \tag{89}$$

7.2.4.5 E - Double Back Serpentine

The Double Back Serpentine manoeuvre is used when the absolute heading difference is greater than or equal to 90° and the desired end points *y* value is less than the ATP's smallest turning circle.

Radius *r*, for the manoeuvre is determined from length μ,

$$2r = \sqrt{(x_d - x_b)^2 + (y_d - y_b)^2} \tag{90}$$

where $x_b = 0$ and $y_b = -r$ giving,

$$2r = \sqrt{(x_d)^2 + (y_d + r)^2} \tag{91}$$

This has components y_d and x_d,

$$y_d = y + r\cos\theta \tag{92}$$

$$x_d = x - r\sin\theta \tag{93}$$

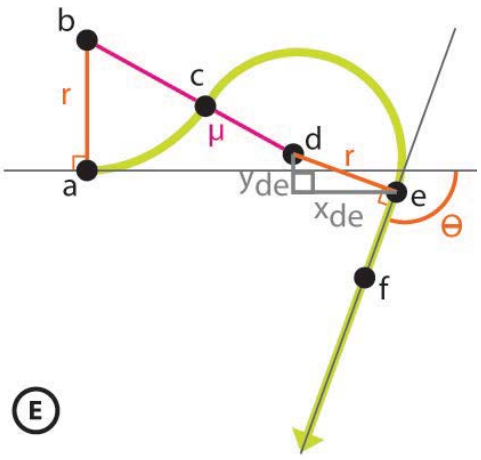


Figure 7-14: Double Back Serpentine path calculation (from Figure 7-9, E)

Substituting into (91) produces,

$$2r = \sqrt{(x - r\sin\theta)^2 + (y + r\cos\theta - r)^2} \tag{94}$$

As before, solving for radius *r*,

$$r = \frac{-(x\sin\theta + y - y\cos\theta) \pm \sqrt{(x\sin\theta + y - y\cos\theta)^2 + 2(1+\cos\theta)(x^2 + y^2)}}{2(1+\cos\theta)} \tag{95}$$

7.2.4.6 Three Point Turn

Insufficient headland space or off-centred row driving often requires a three-point turn manoeuvre. This allows the ATP to navigate back along the same orchard row (Figure 7-15).

When the end-of-row is found (section 7.5.1), the row's width and headland space is calculated. Analysis of the LIDAR scanner's data provides these values. Collisions with orchard objects are avoided by adjusting length *A* - *B* and the turn radius.

Starting at position *A* (Figure 7-15), a linear section is driven to ensure the ATP is clear of any support structures or bracing. This is followed by a constant radius curve to rotate the ATP until it is perpendicular to the row. The on-board compass module is used during transition *B* – *C* to find when the chassis is perpendicular to the row. Point *C* is reached when the chassis is perpendicular to the row heading. The estimated distance of the reversing section (transition *C* – *D*) is calculated from (96). The ATP reverses for distance d_{CD}, to reach point *D*. While driving forward from *D* through curve *D* – *E*, the new row driveline is interpreted from the LIDAR scanner data (section 7.5.1). When the ATP's heading is within 10° of the new driveline the PID row drive system takes control. This system navigates the ATP along the row.

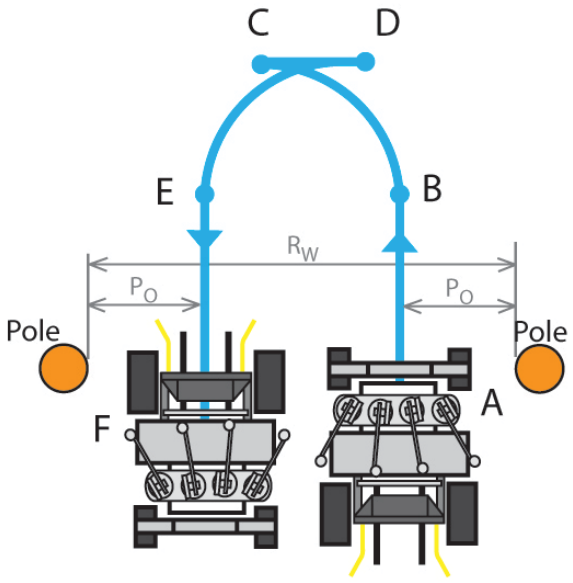


Figure 7-15: showing the three point turn manoeuvre. Note that traversal C – D is a reverse section

$$d_{CD} = 2 \times (P_O + r) - R_W \tag{96}$$

7.3 Review of Autonomous Navigation Systems

A large body of work is devoted to autonomous robotic navigation. However, most of this work is for structured, relatively controllable and ordered industrial environments. The variability within horticultural and agricultural environments imposes significant additional navigation complexity. Most successful navigation within agriculture and horticulture has been in arable cropping. However, the open-field cropping environment is less complex than an orchard. This is because cropping has no canopy cover and the large open area allows a wider range of environment sensors to be used.

Within the outdoor environment, the most significant advancement in autonomous navigation is within the DARPA Grand and Urban challenge vehicles. Like most of the more successful competitors the Volkswagen/Stanford team (Montemerlo, et al., 2006) used a combination of LIDAR (5x multi-plane units), machine vision, GPS, RADAR and a six degrees of freedom Inertial Measurement Units (IMU) on their autonomous vehicle Stanley. The LIDAR, RADAR and vision units are used to map the surrounding ground to provide enough information to make drive path decisions. These vehicles require the wide range of sensing due to the vast amount of terrain experienced. The outlined systems provide a base for investigation into their application for kiwifruit orchard navigation. With the relatively ordered and repetitive orchard structure it is unlikely that the AKH will require this number of sensors. This is because these structures are able to be exploited as local positioning markers and their detection will not be as detailed as the vehicles entire surroundings.

Figure 7-16: Stanley, the 2006 DARPA Grand Challenge winner

There are many navigational sensors available with potential for orchard application. These are described by Keicher and Seufert (2000) and categorically sorted into two areas:

1. Local or relative positioning (outlined by Spencer (2004))
2. Absolute, or global, positioning

Relative positioning requires a reference datum landmark(s) where a current position is calculated with respect to the landmark(s) using triangulation. Humans tend to navigate using rela-

tive navigation (when absolute navigational tools are not used) by observing landmarks and other structures. Absolute positioning also uses reference points but they are typically much further away and correspond to mapped positions. The popular global positioning system (GPS) is an absolute positioning method using a satellite constellation and fixed ground locations, conceptually derived from traditional celestial navigation.

An autonomous driving system usually consists of three components. These are independent of the underlying system (local or absolute positioning). Keicher and Seufert (2000) outline the three component requirements:

1. A sensor to report deviation from the desired position
2. A controller to generate a correction signal according to the deviation
3. An actuator that manipulates the heading of the vehicle. This allows any position error to be corrected using vehicular motion.

This review focuses on the sensing system to determine the vehicle's position as it presents the greatest challenge to autonomous navigation. Common position sensors include mechanical sensors, global navigational satellite systems (GNSS), machine vision, laser triangulation, ultrasonic and geomagnetic, attitude and direction-of-movement information sensors (Keicher and Seufert, 2000).

7.3.1 Relative Positioning

Attempts at relative navigation began in the 1930's. These systems used mechanical attachments to alter a tractor's steering position, allowing it to follow a ploughed furrow (Figure 7-17). This mechanical system uses a guide wheel within the furrow. The mechanical linkage alters the tractor's steering when the wheel deviates within the furrow. This kept the tractor on its intended path.

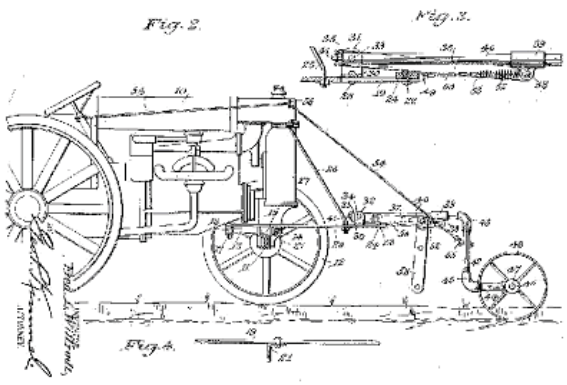


Figure 7-17: Schematic of the Willrodt tractor steering system (Willrodt, 1930)

Progression of mechanical attachments allowed tractors to steer in different ways.

One idea used a drum of piano wire mounted to the tractor, with the end of the wire anchored to the middle of the field. The tractor drove in progressively smaller circles by shortening the wire's length, covering the working area (Sissons, 1939). Although these systems still required operator assistance they demonstrate the beginnings of autonomous navigation. As technology progressed, the scope of autonomous solutions grew.

The 1970's yielded the next generation of automated guidance systems. Rushing's (1971) driverless tractor followed the magnetic field of energised buried wires. This system was the first significant progression in automated navigation. Later, in 1978, the Claas Autopilot was released. This system used sensors to detect the crop edge, but was only suitable for crops with solid edging like maize. The sensor information was used to steer a tractor or chopper unit relative to the crop's edge (Diekhans, 1999).

The evolution of more sophisticated navigation sensors and techniques progressed with computing technology, allowing higher levels of integrated sensor analysis and communication.

7.3.1.1 Machine Vision

In agriculture and horticulture, the most commonly employed methods of visual navigation are crop row tracking and positioning systems. These systems work out a vehicle's position relative to the detected row(s). Although kiwifruit orchards do not present such visibly obvious rows making this approach less applicable, there are some similarities between the environments. Therefore, techniques used for crop row identification have some potential for orchard navigation. To date, horticultural automation has focused on harvesting as opposed to navigation. This limited orchard specific applications. Where navigation has been considered, precision GPS is generally used (Muscato, et al., 2005, Reece, et al., 1996). The GPS system works off pre-defined waypoints. However, the unique kiwifruit environment does not permit this approach due to the canopy, hedges and trees blocking high accuracy RTK-GPS (discussed in section 7.3.2).

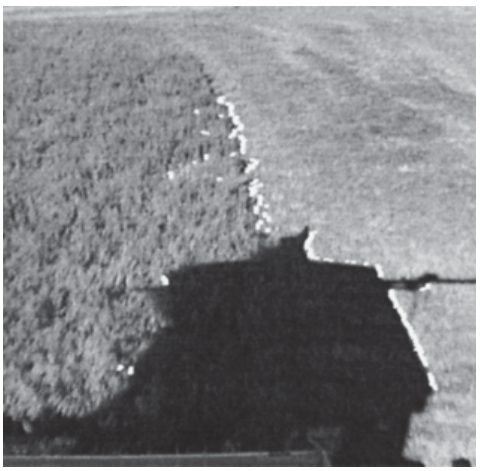


Figure 7-18: The Demeter System for visual navigation. The erroneous points from the tractor's shadow are shown (Pilarski, et al., 2002).

Ollis and Stentz (1997), Pilarski, et al. (2002), Benson, et al. (2003) and Gerrish, et al. (1997) all used

machine vision sensing for navigation of their respective robotic harvesters. These systems focused on detecting and tracking the crop cut edge. This edge can range from distinctive, for tall crops like maize, to less visible for shorter crops like grass (see Figure 7-18). Although these developments are not directly related to orchard navigation they showed significant implementation issues, including:

1. the effects of shadowing from general objects and the harvest unit
2. the crop shadow shifting over the day. This is especially noticeable with taller crops
3. edge detection issues associated with bare patches of crop (no edge)

The added complexity introduced from shadowing within the uncontrolled environment is significant. This complexity could cause positioning errors and incorrect navigation. Pilarski, et al. (2002) reports on the Demeter system. He concluded that their visual navigation was less accurate and reliable than an RTK-GPS system they trialled. These inaccuracies occurred where the observed crop was patchy or shorter. The camera viewed open areas of soil that affected analysis. Some of these issues could have been resolved by lowering the camera's mounting position. This would minimise the viewable gaps between plants, allowing more plant length variation without the ground becoming visible. This modification would also increase the viewable area ahead of the harvesting system. However, a reduced field of view either side of the harvesting system would result. This could limit the system's ability to track the crops edge when the crop's edge deviates from directly in front of the harvester. This would also pose an issue if the harvester veered off path, as it may not be able to regain the crop's edge and therefore navigate.

Gerrish, et al. (1997) report another crop edge tracking system. They achieved instances where their system navigated more accurately than a human operator. Although some success was achieved harvesting corn, soybeans, red peppers and alfalfa, autonomous unsupervised operation was not achieved.

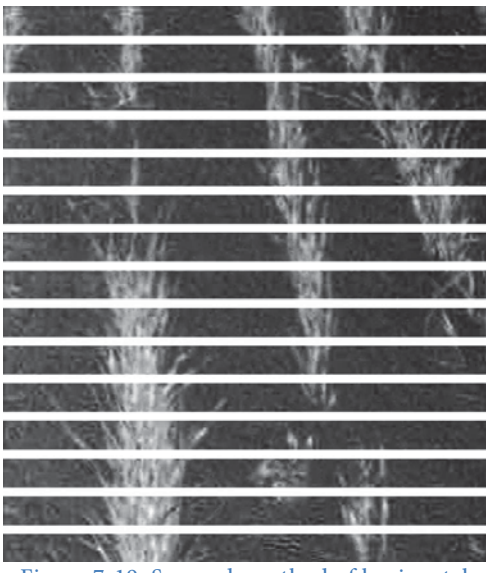


Figure 7-19: Sogaards method of horizontal banding to find the plants location (Sogaard and Olsen, 2003).

Crops cultivated in rows have a distinctive visual appearance. This makes visual navigation solutions more reliable and easier to implement. A kiwifruit row does not exhibit these distinctive

visual characteristics. However, in-row vines and posts do present a near linear pattern. Therefore, the identification of individual row elements (plants or structures) is required for in-row kiwifruit navigation. The plants and structures act as relative positioning markers. Although the identification of these elements differs from row crops, the analysis of the identified elements to find the row has similarities. Therefore, row finding techniques were investigated. Several authors (Astrand and Baerveldt, 2002, Bakker, et al., 2008, Leemans and Destain, 2006, Slaughter, et al., 2008, Sogaard and Olsen, 2003) have focused on the identification of individual plants to determine row line for navigational purposes.

Some authors (Astrand and Baerveldt, 2002, Leemans and Destain, 2006, Sogaard and Olsen, 2003) use the Hough Transform (Duda and Hart, 1972) to determine plant placement correlating to the row.

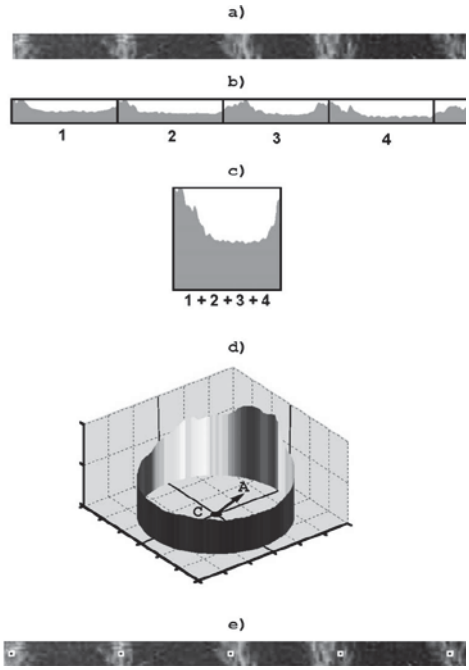


Figure 7-20: Band extraction and application of centre of mass determination for localisation of plants in row (Sogaard and Olsen, 2003)

Sogaard and Olsen (2003) have developed another method to find the row's position. This technique uses horizontal banding, splitting the image into horizontal sections (see Figure 7-19). The area of the image corresponding to a plant is identified and its 'centre of mass' determined (Figure 7-20, *e*). Repeating this process for each band provides a series of data points along the crop row. A Hough Transform is used to detect the linear placement of these plants (see Figure 7-19 and Figure 7-20).

Another visual navigation method is presented by Morimoto, et al. (2005). This system recognises visual markers positioned within the operational environment. Implementation within a single operational environment is potentially feasible using visual markers. However, placement within many orchards of varying sizes would make this approach impractical, costly, time consuming and labour intensive. Visual recognition of established on-orchard infrastructure and vines could possibly serve as local navigation reference markers (similar to Morimoto, at al. (2005) system). Although identification of current on-orchard infrastructure and vines is significantly

more complex, it removes the setup, maintenance and other costs associated with tailor made markers.

7.3.1.2 Stereo Machine Vision

Stereoscopic machine vision (stereopsis) allows an object's distance to be measured using two cameras observing a scene from slightly different positions. From these stereo images, distance is calculated from the object's image disparity. Shanshan, et al. (2009) states that stereo machine vision is one of the most accurate passive methods for distance measurement.

Rovira-Mas, et al. (2009) reports increased system accuracy over a larger depth range in comparison to a standard binocular stereo vision. This was achieved by using two pairs of stereo cameras with different baseline distances. Their previous work used a single pair of cameras but limited distance measurement to a specific distance band. This distance band was governed by the interocular (or baseline) distance between the lenses and the lens focal length. By implementing two pairs of stereo cameras the operating band is increased. The same result is achievable using three cameras. One camera works as the reference camera and the other two independently determine distance. This provides an obvious cost advantage and reduces image analysis requirements. However, it requires the cameras to be calibrated instead of buying off-the-shelf stereo cameras.

Figure 7-21: RASCAL (Kogler, et al., 2006)

Kogler, et al. (2006) used stereopsis for "RASCAL" (Robust Autonomous Sensor-Controlled All-terrain Land-vehicle), developed for the DARPA Grand Challenge. The autonomous unit

completed 29km of the 212km route using stereo vision for autonomous obstacle detection and path planning.

Stereo vision is not limited to short range systems. Applications for terrestrial navigation and sensing are under development by Cui, et al. (2006). Far off land features and terrain are more accurately measured using a pair of cameras with a large interocular distance. However, this approach presents problems when objects are closer to the cameras. As an object moves closer to the camera pair, its disparity in the images increases. Therefore, if an object is too close its distance cannot be calculated as it will only be visible in one of the cameras.

Two main analysis approaches are used to find the distance information:

1. Feature-based and
2. Dense disparity

Feature-based distance measurement requires an object(s) to be identified within the image. This object(s) is then located in the second stereo camera's image, where its distance is calculated from the disparity between the two images. The dense-disparity method (Moghadam, et al., 2008, Rovira-Mas, et al., 2009) determines distance information for an entire image area. This provides additional three-dimensional scene information allowing higher levels of object recognition. Feature-based measurement is computationally faster, but it relies on identifiable features being present. The dense disparity method however provides more scene information to help locate areas of interest.

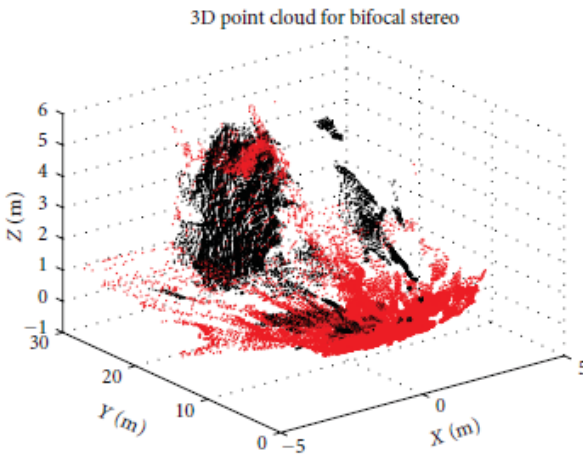


Figure 7-22: Three-dimensional point cloud from stereo vision (Rovira-Mas, et al., 2009)

Three-dimensional point cloud plots are used to visually represent dense-disparity data (Figure 7-22). This method has application for the formation of three-dimensional environments and general object detection. However, system performance will suffer from the amount of analysis required. Coincidently, Rovira-Mas (2009) analysed an orchard row. This work showed the large amount of information captured requiring further analysis to interpret the environment. The large computa-

tional requirements render real-time application unrealistic and compromise feasible navigational ability.

A catadioptric stereo system (Bailey, et al., 2006) uses mirrors to split a single camera's image. This allows a single camera to view the scene from different positions to find distances. A single dimensional comparison theory was applied to find the disparity of pixels in the split image. From this, a pixel-by-pixel comparison (from both image halves) was used to build a two-dimensional disparity map. This map represented the likeness between the two compared pixels. A minimum cost path was used to traverse the disparity map. This path represented the real world distance to every observed pixel point for the analysed image line. By performing this analysis for every image line, a three-dimensional point cloud is created. This technique could be used over a small height range to produce similar information as a two-dimensional range finder (section 7.3.1.4). To achieve this, the image height intercepting the in-row objects is required. Although computation time results published by Bailey, et al. (2006) indicate this method is not feasible for real-time or near real-time implementation, Morris, et al. (2009) developed their real-time dense stereo implementation using an FPGA. However, this was done within a laboratory environment. Dense stereo matching within an orchard environment is more complex that in a laboratory environment due to complex environmental fluctuations. In addition, tracking the minimum cost path (as used by (Bailey, et al., 2006)) for large distance step changes, as presents on-orchard, requires further analysis and investigation. This is required to determine object disparity information and therefore distance.

The stereopsis development with characteristics closest to those required for kiwifruit orchard navigation is presented by Huertas, et al. (2005). This research project aimed to find the entire tree based obstruction in an image. From this, each tree's navigational threat was interpreted using stereo vision. The navigational threat is a judgment of whether the robotic vehicle could push through the obstacle. To achieve this, the system needs to identify and localise all the trees within the natural environment. These relatively straight pruned trees have similar appearance characteristic to the posts found on-orchard that can be used for navigation. Although their tree diameter measurement for navigational threat analysis is not directly relevant, their methods for identifying trees are of interest.

Huertas, et al. (2005) used a feature-based system to find the trees by exploiting their distinctive visual characteristic. The main characteristic is the uniformly contrasted tree area (in compari-

son to the background). However, the trees also have an opposing background contrast at the tree trunks edge. "It [the system] assumes that the portions of the trunk appear to be either brighter, or darker, than the background, and thus the boundaries have opposite contrast" (Huertas, et al., 2005). Although some success was reported, difficulties were experienced in identifying trees when the trunks were observed against a background with similar contrasts. For example, if the scene has trees with thick vegetation, mountains in the background or lower substantial ground vegetation. Fortunately, these scene presentations were not commonplace within their environment.

Figure 7-23 (b) shows the analysis from their single dimensional edge detection method. Edge information is grouped into contour lines allowing the complete tree edge to be found. The single dimensional edge detection array also provides contour line polarity. This indicates whether the detection array has transitioned from light to dark contrasts (or vice-versa), allowing the identified edge information to be grouped according to polarity. This step helps group edge information belonging to a single edge. From these segments, contour fragment length and angle are used to assess if several edge groups belong to a single fragmented edge.

Discussion on stereopsis range accuracy was not included. However, the results presented for trunk diameter measurement indicate that accurate distance measurement was achieved. Over a distance range of 2.5-30m, the average diameter error was 50mm, for trunk diameters of 100-650mm.

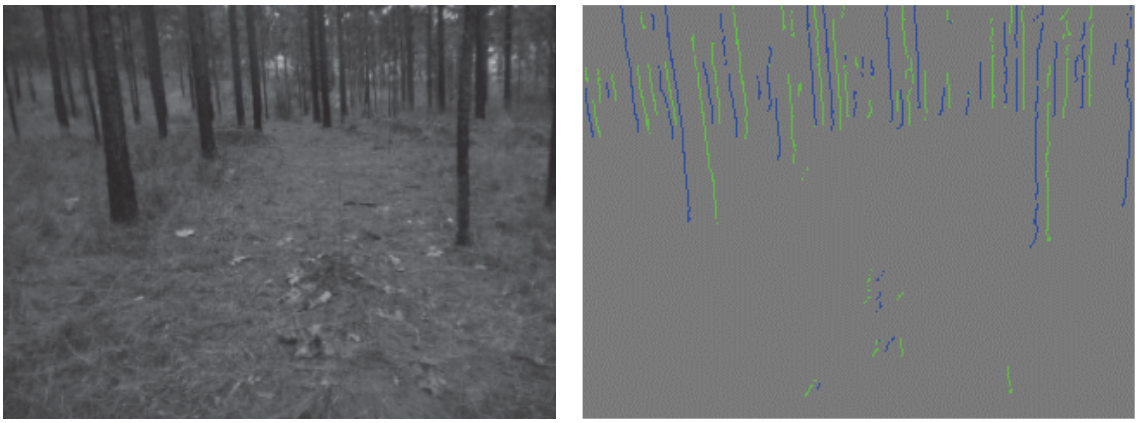


Figure 7-23: Contour detection for tree location. Green and blue lines indicate polarity (Huertas, et al., 2005).

7.3.1.3 Laser Triangulation

Keicher and Seufert (2000) outline laser triangulation for relative positioning using an Arnex navigation system. This system operates using simple passive reflectors in known positions. The

lasers find the distance to each visible reflector, where the unit's relative position is found by triangulation. This requires three or more reflectors to be visible at any one time.

> "The navigator has a rotating laser-optic sensor unit that should be mounted on an elevated support on board the moving vehicle. A rotating narrow disc emits an eye-safe laser light, and a sensor detects the light returned from the reflectors placed along the field's boundaries. It measures the horizontal as well as the vertical angle to the passing reflector. These angles are determined in a co-ordinate system fixed to the vehicle" (Keicher and Seufert, 2000)

This system's complex setup requirement limits it commercial potential. Furthermore, there is significant potential for signal blocking on-orchard. Signal blocking could result from many sources, including: on-orchard infrastructure, terrain undulation and the low canopy structure. To overcome this, the system would require many reflectors to maintain system accuracy. This further increases setup complexity and cost. Therefore, for this on-orchard application, where commercial potential is a critical factor, this system is impractical.

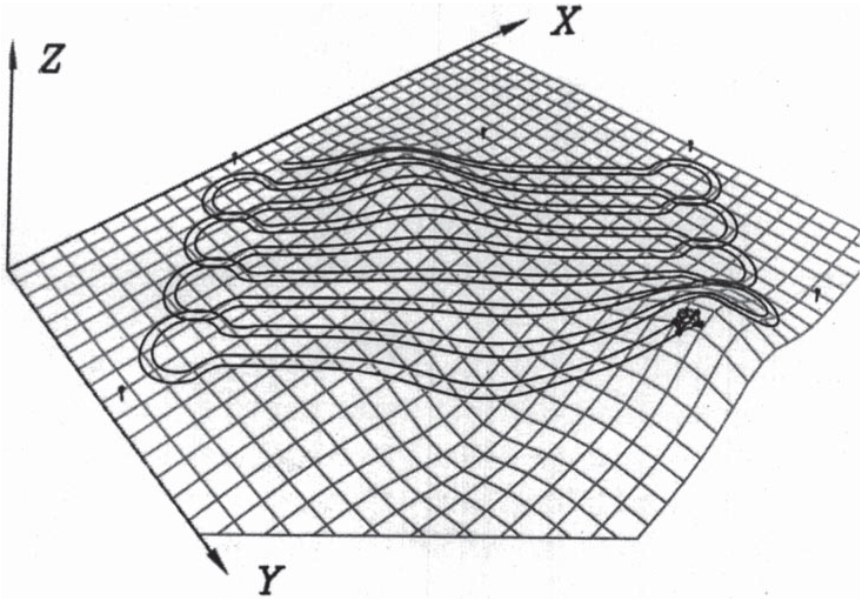


Figure 7-24: The Amex 3D position determination (Holmqvist, 1993)

7.3.1.4 Laser Range Scanner

A variety of laser range scanners are commonly available for object sensing applications. These are commonly known as LIDAR scanners. Single dimensional, two-dimensional, multi-plane two-dimensional and three-dimensional units are available. Each has a variety of data capture rates, scanning areas and range configurations. For these, unit cost and data acquisition time significantly increases with additional scanning dimension, range and area (Aliakbarpour, et al.,

2009). Multi-plane two and three-dimensional units are beyond current commercial feasibility, where multi-plane two-dimensional units are priced from NZ$22,000 each. Three-dimensional versions are further impeded by slower data acquisition rates and scanning times. They are currently not suited to real-time applications (Zhang and Pless, 2004).

LIDAR provides accurate range measurement of incident objects. It can scan over large angular fields, with very fast data acquisition rates (Baltzakis, et al., 2003). Baltzakis, et al. (2003) also comment on their reliable object detection results achieved using LIDAR scanners. These scanners are available in robust, waterproof, outdoor rated housings, making them suitable for orchard environments. Their fast scanning and pre-processing (supplying range information without analysis) allows faster environment interpretation when compared to an inferred scene analysis system like machine vision.

The scanner's detection ability is dependent on range, size and surface reflectivity. Therefore, the probability of detection decreases as distance increases (Huang and Barth, 2009). This also assumes that the scanner's beam will intercept the required objects. Implementation of a two dimensional LIDAR scanner cannot guarantee object detection on-orchard. This occurs when low or hanging objects are not within the scanning plane. This makes those objects invisible to the LIDAR scanner, potentially allowing collision to occur. Consequently, LIDAR is commonly implemented in a complimentary capacity to other navigational systems (Durrant-White, 2006, Kogler, et al., 2006, Lumelsky and Skewis, 1990, Moghadam, et al., 2008). However, standalone systems are still implemented (Barawid, et al., 2007). These systems are predominately found in commercial industrial type environments (e.g. buildings). In industrial environments objects are more likely to coincide with the LIDAR scanning plane(s) (Travis, et al., 2005).

To date orchard navigation strategies have predominately focused on map-based applications. This allows predefined paths to be followed (Barawid, et al., 2007). However, this method requires initial and often time-consuming technical setup. In addition, a map-based navigation system does not dynamically adjust to environmental changes. Therefore, any alteration to the orchard requires navigation paths to be updated. A LIDAR based navigation system would allow real-time dynamic navigation decision-making without comprehensive mapping. This helps support on-orchard operation in dynamic environments, where items, like fruit storage bins, could be placed in the ATP's path. To achieve this, the navigation algorithms are required

to account for any physical on-orchard environmental changes (e.g. the relocating of orchard structures).

Barawid, et al. (2007) and Ryo, et al. (2004) both use two-dimensional LIDAR scanners for navigating orange orchards. The LIDAR's scanning plane is set at a height of 700mm parallel to the ground. This plane intercepts the trees' foliage and/or trunks along the orchard rows (Figure 7-25). The scanner's data is analysed with a Hough Transform algorithm to "automatically detect the orchard tree lines and provides lateral offset and heading measurement" (Barawid, et al., 2007). However, variations in tree size, positioning or missing trees significantly impaired orchard row determination. "In order to get the most accurate data, a specified portion in the experiment site where there were no apparent breaks between trees were taken and considered as the basis for evaluating the autonomous accuracy of the robot tractor" (Barawid, et al., 2007). This does not represent a true orchard environment, where complex variation is commonplace and a system that successfully operates within all presented variation is required.

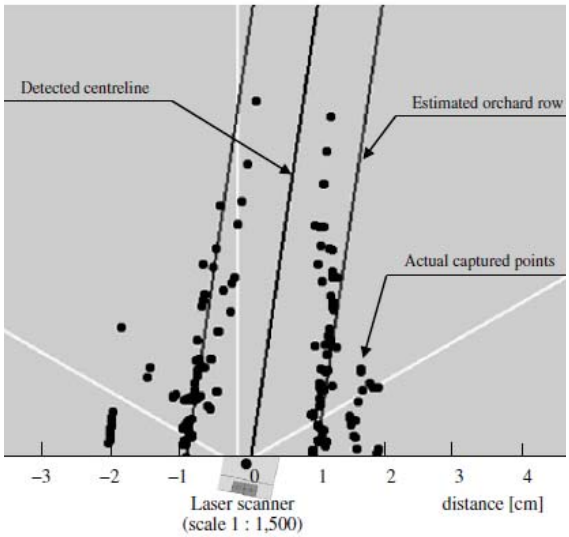


Figure 7-25: Actual captured data points from the LIDAR scanner of the on orchard environment from (Barawid, et al., 2007)

Despite the ideal orchard section, Barawid, et al. (2007) conclude that their system is suitable for real-time on-orchard navigation. An analysis of their system's mean lateral and heading error was carried out during testing to make this conclusion. This conclusion was also based on a 0.36m/s drive speed, which may not be fast enough for efficient operation. An RMS error of 0.11m and 1.5˚ in lateral position and heading respectively was recorded at this drive speed. However, individual errors did become as large as 0.48m and 7˚ respectively while operating within the ideal orchard section. Therefore, outside the ideal orchard area their errors would potentially become significantly worse when presented with expected orchard variation.

7.3.1.5 Laser Radar and Machine Vision Combination

Investigation into LIDAR systems highlighted their combined use with machine vision (Baltzakis, et al., 2003, Holmqvist, 1993, Moghadam, et al., 2008, Subramanian and Burks, 2009,

Subramanian, et al., 2006, Whittaker, 2005). A famous example of the combined technology is the automated vehicle 'BOSS'. This vehicle used long and short range LIDAR sensors together with machine vision (Whittaker, 2005) to win the 2007 DARPA Grand Challenge.

Subramanian and Burks (2009) and Subramanian, et al. (2006) report success navigating orange orchards using a combination of LIDAR and machine vision. They performed row navigation and headland turns with this combination. A LIDAR – machine vision fusion was adopted as "GPS cannot be effectively used for positioning in citrus applications, since the vehicle frequently moves under tree canopy, which blocks the satellite signal to the GPS receiver" (Subramanian, et al., 2006). They mounted their LIDAR scanner above the tractors cab, this mounting position created a steep LIDAR viewing angle. The angle meant that only a short section of row was scanned. This configuration produced large positioning variations or even an inability to determine the drive path. The positioning errors were due to local in-row tree variation or even missing trees. Since the scanner is only observing a short row section it cannot average these variations over many trees. If a drive path is not found with LIDAR data, the machine vision system takes over navigation control. To find the rows, the machine vision algorithms exploit the visible contrasts between fruit tree foliage and the grass row. This is an environmental condition unique to tree crops. Kiwifruit orchards do not present this, or any similar, observable characteristic. Although machine vision is the backup navigation system it too experiences positioning difficulties. These are caused by varied lighting conditions. They note that "many trees were found on one side only: early morning or evening" (Subramanian, et al., 2006). Difficulties with machine vision analysis from lighting variation is also described by Huang and Barth (2009).

By combining the two technologies and their sensing abilities, it is necessary to align their geometric spaces (mapping LIDAR and machine vision coordinate spaces). This allows a cooperative navigation implementation. Huang and Barth (2009) outline the three techniques for calibrating mapping between the respective devices geometric spaces:

1. Visible beam calibration: Using infrared or near-infrared cameras to observe the LIDAR's laser line. Huang and Barth (2009) state that this requires high-cost infrared cameras sensitive to the spectral emission band of the LIDAR unit.

2. Three-dimensional LIDAR based calibration techniques: Captures various features with LIDAR and machine vision, i.e. corners and edges of specific calibration objects in the environment.
3. Two-dimensional planar-based calibration: as proposed by Zhang and Pless (2004), observe the plane of an object with a single plane LIDAR.

Huang and Barth (2009) also propose an additional calibration method for a multi-plane two dimensional LIDAR. They use an inclined board with a chequered patterned. The chequer board is observed and positioned in both geometric spaces, allowing alignment.

7.3.2 Absolute Positioning

Global navigational satellite systems (GNSS) support most common absolute positioning systems. The American GPS is the most common of these. Accuracies to within 0.02m are achievable using the GPS variant, RTK-GPS. However, the orchard environment makes RTK-GPS unsuitable due to unreliable signal reception. RTK-GPS relies on a local base station to provide error correction from standard GPS levels. This is used as standard GPS has insufficient positioning accuracy to provide a standalone navigational system. Durrant-White (2006) published the relevant issues:

> "it is essential to recognize and understand that GPS has many failure modes that make it inappropriate for stand-alone use in autonomous systems: In particular GPS may fail due to loss of line-of-site, from multi-path of local terrain, and from active RF jamming" (Durrant-White, 2006)

RTK-GPS satisfy precision navigational requirements in many applications. It improves on standard GPS by providing a decimetre-accurate position every second. Unfortunately, as indicated by Durrant-White (2006), RTK-GPS is generally not suitable for on-orchard application due to unreliable signal reception. The orchard environment hinders reliable operation as the hedgerows and crop canopy impede the reception of the error-correcting signal (loss of line-of-sight). Even if these problems could be overcome, navigation with RTK-GPS alone is risky as it cannot detect potential unmapped obstacles (people, bins, machinery etc.). The combination of signal unreliability, cost and additional base station make RTK-GPS impractical.

GPS and its variants (RTK-GPS and DGPS) have assisted successful navigation solutions in environments where line-of-sight is achievable and obstructions are unlikely to occur. Murakami, et al. (2008) describe a GPS guided manure spreader under operator remote control. It is capable of spreading manure on a field up to one kilometre away. Pilarski, et al. (2002), Muscato, et al. (2005), Reece, et al. (1996) and Durrant-White (2006) also report successful GPS guided navigation solutions.

Although impractical as a standalone system, standard GPS (accurate to approximately 3m) with high gain antennas could be a suitable navigational aid. The high gain antenna receives reliable reception and there are no issues with needing an error-correcting signal. GPS would provide an approximate absolute position that will be referenced against orchard maps. In comparison to the waypoint maps required for GPS navigation, these maps only provide the orchard features, like canopy areas hedges, etc. and are made for orchard management purposes. Other navigational sensors would further refine the GPS provided position for accurate and collision free operation.

7.3.3 Conclusions

Many navigational systems exist for differing environments. However, no off-the-shelf navigation solution is available for orchards. High accuracy GPS, commonly used for navigation in cropping, is impractical due to signal dropout under the kiwifruit canopy. In addition, this method of navigation does not detect moveable threats (anything that is not mapped) making it unsafe for humans to be near operating machines. A navigational sensor, capable of real-time on-orchard operation and detecting any object, is required for successful development.

As absolute positioning (GPS) as a standalone system is not viable, a method of relative positioning is required. Relative positioning requires reference landmarks for position calculation (relative to the landmarks). For this, custom datum landmarks are often used. However, this method requires large setup and maintenance costs (as required for laser triangulation). Progressing relative positioning to use the canopy structures as the landmarks, increases development complexity (structures are harder to reliably detect) but eliminates setup and maintenance costs. This will also raise the system's commercial viability by reducing costs.

A machine vision system that is robust and capable of reliable canopy infrastructural identification would fulfil the desired system requirements. It also presents as a lower cost and flexible system. Machine vision has been successfully implemented in several cropping navigation systems. Positioning accuracy better than a human operator was reported in one of these systems (Gerrish, et al., 1997). Although success has been reported, the machine vision developments highlighted interpretation issues with lighting variations. Shadowing, shifting lighting during the day and patchy crop areas all affected system accuracy and reliability. Expansion to stereo machine vision will allow the system to determine object distance and calculate optimal drive paths. Initial development explored a stereo machine vision navigation solution for the presented benefits.

7.4 Machine Vision Navigation

Several sensing systems are identified with navigation potential. However, machine vision is selected for:

1. Additional system 'intelligence'. Further development of the machine vision system could allow it to interpert more of the environment. This would deliver better awareness and allow for higher-level navigation decision making.
2. Flexibility: By developing the system in a variety of orchards this system would be adaptable to several orchard layouts.
3. Low hardware setup cost (in comparison to other suitable sensory technologies).

A near uniform and repetitive structural post layout forms the orchard rows. These structures provide local positioning markers along the rows. By identifying these markers the ATP could navigate along the rows. To achieve this, a reliable identification system is required to accurately identify and locate these in-row objects.

Real-time or near real-time scene interpretation is required for efficient operation. This means that the computational time of any developed analysis algorithms is important. In addition, these algorithms must be reliable to achieve safe, consistent and efficient operation. Failure to meet these criteria directly affects the ATP's commercial potential. Therefore, they are an important consideration.

The posts often present a desirable uniformly coloured and straight appearance. This unique appearance can make them obvious in the environment. However, significant visual variations are still presented, including:

1. Square, quarter round, half-round and round posts are all used.
2. Moss, lichen etc. grows on the posts blending their appearance into the background scene.
3. Knots and bowing in the posts degrade their uniform appearance.
4. Kiwifruit vines can grow against the posts, occluding them.
5. Speckled and inconsistent scene illumination, caused by shadowing from overhead leaves, neighbouring objects (e.g. hedges) etc.

On-orchard post identification is significantly more complex than the same task performed in a controlled industrial environment. The level of consistency and uniformity found in controlled environments is not present on-orchard. This variability poses the greatest obstacle to achieving a reliable in-row visual navigation system. Reliability is essential for safe operation and ultimately determines whether the system will be deployed on-orchard.

7.4.1 Post Identification

The in-row posts need to be found to achieve a machine vision based in-row navigation system. These posts are generally presented in pairs with one on either side of the row. These pairs are positioned progressively further from the ATP's current position along the row (Figure 7-26). The closer pairs of posts are easier to identify as they appear bigger and more distinctive in the captured images. Ideally, these posts present two distinctive vertical edges that are parallel and visible over each post's entire height. In reality, many variations exist to the ideal model. However, it does provide a starting point for post-identification analysis. The post's surface has more variation than the edges, therefore, the edges are targeted. However, significant appearance variations are still evident, so the final solution must address all presented variations.

An edge extraction method is required. This must identify all post edges even when variations are presented. Variations include: Non-linear appearance, gaps, background blending, partial and complete shadowing and occlusions. The starting approach is to identify vertical or near-vertical line segments within the image. These edge segments are then 'stitched' together to cre-

Figure 7-26: An orchard row as observed from the on-board navigation cameras (cropped to show relevant image section)

ate a single, continuous edge. This allows for gaps, shadows, etc. to hide edge sections without compromising the system's ability to identify the overall post edge. The edge segments are grouped by their image position (radius) and angle, as described by a Hough transform (Duda and Hart, 1972) applied to the data points.

> "Ideally, edge detection should yield sets of pixels lying on edges. In practice, these pixels seldom characterise edges completely because of noise, breaks in edges due to non-uniform illumination, and other edge effects that introduce spurious discontinuities in intensity values" (Gonzalez and Woods, 2008).

To help with this detection a higher camera viewing angle was used. This places more visual grass area behind the viewed posts and increases the viewable contrast of the post's edge. A lower viewing angle placed more kiwifruit canopy behind the posts. This often merged the posts' appearance with the canopy. For this reason the cameras are set to 1.3m from ground level.

The post edge has "an important geometric property" (Gonzalez and Woods, 2008), the gradient vector. This vector identifies the direction and magnitude of the greatest rate of change relating to the edge. This vector is perpendicular to direction of the edge and represents a step change in intensity level. Several gradient detecting operators are available. These operators compute the partial derivatives $\partial f/\partial i$ and $\partial f/\partial j$ (where f is the gradient vector) over every pixel in the image's search area (Gonzalez and Woods, 2008). This identifies edges with the required magnitude and is known as a differential gradient technique. Template matching is another commonly used edge detection technique. Template matching can be used when specific edge characteristics are known. To implement this method a template that exploits the unique visual characteristics to find the intensity step changes is created. When the template coincides with the desired edge, it returns a higher value than non-edge areas. Both template matching and differential gradient based edge detection methods

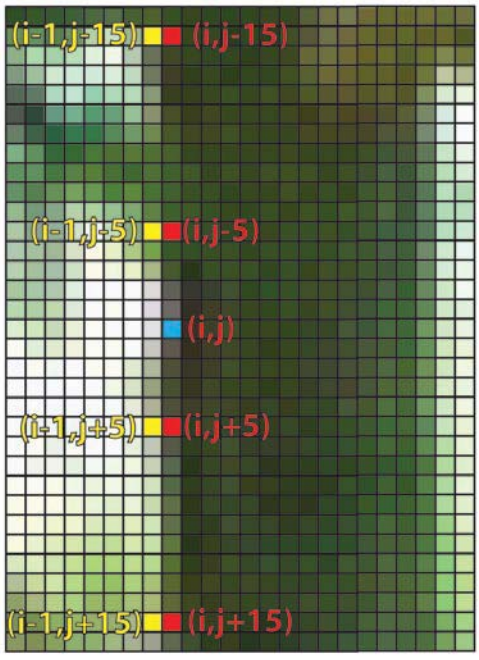


Figure 7-27: Pixel view of the template points used for edge detection. This instance coincides to a post edge producing a positive edge correlation at point (i, j). The red and yellow points represent the template points to coincide with each side of the edge.

were used to find the most reliable method for post edge detection. This covers the two main methods of edge detection as described by Davies (2005).

Template matching is generally a more intuitive approach. The template (e.g. Figure 7-27) is formed from observations of characteristics over the desired edge's intensity step change. The template matching aims to highlight the desired areas while producing minimum false correlations to other image information.

The developed template finds the normalised difference between the average green (image channel) values of the red and yellow template points (Figure 7-27). This is a two dimensional template and provided less false positives than a single dimensional system as described by Huertas, et al. (2005). The template's components *a* and *b* for the yellow and red tempalte points respectively are (Figure 7-27),

$$a=\frac{\Sigma G_{Yn}}{4} \quad (97)$$

$$b=\frac{\Sigma G_{Rn}}{4} \quad (98)$$

Figure 7-28: Resulting identified points (purple) from iteratively applying the template mask in Figure 7-27

Where G_{Yn} are the green colour values (0-255) of the four yellow template points and G_{Rn} for the red template points. *Gr*, (99), is the normalised gradient for a specific pixel location (*i, j*), Figure 7-27.

$$G_r = Abs\frac{(a-b)}{(a+b)} \quad (99)$$

When G_r is greater than 0.05 an edge point of the desired characteristic is found and its image position is recorded. These are represented in Figure 7-28 as purple points on the image.

Figure 7-28 shows the result from the vertical edge-detecting template created to detect the posts' edges (Figure 7-27). The template spans a small vertical range. This allows some post variation to be present. This technique is used because, even posts that present with variation, are made up of small linear segments that can be linked into a single continuous edge. The template

is rotated ±5° from vertical at each search point to detect small variations in the post's angle. True vertical is found using the on-board fluxgate compasses three-axis accelerometer.

Figure 7-29: Result from applying a Hough Transform to the data set in Figure 7-28

An improved Hough Transform (Duda and Hart, 1972) is used to identify linear placement of data points from the edge detection process. The improvement is achieved by performing the Hough transform on clusters of data points instead of all the data points located in the image. It achieves detection while ignoring erroneous data noise (Zhang and Du, 2005). The post is found in the edge data by finding the two most significant Hough Space correlations (Figure 7-29). These edges must be near parallel and positioned within a set width range. Meeting these criteria, the two edges are classified as belonging to the nearest post in the visual range. This is set as the nearest post as it appears taller in the image and therefore has more edge data points, giving it a stronger weighting in Hough space. Once a potential post position is found, its corresponding position in the second stereo camera is found. This allows the stereoscopic system to determine the post's distance and position (section 5.2.3). By knowing the post's distance, its actual width is calculated from (100). This allows the identified object to be positively validated, as the posts are all within a specific width range. This process helps to eliminate false positive identifications.

$$Size = Pixels(mpppm \times distance) \tag{100}$$

When tested, the template matching method was only effective in favourable lighting and orchard conditions (as in Figure 7-26). In these conditions this system guided the ATP along a 15m section of orchard row using dead reckoning once the next post in the row was localised. However, for in-row images, like those in Figure 7-33 to Figure 7-35, with variable lighting conditions, the template matching method identifies large amounts of noise. This makes the post edge sections indistinguishable during analysis. Therefore, operating in an orchard environment with large fluctuations in illumination and object appearance (as highlighted in Figure 7-26 and Figure 7-33 to Figure 7-35) requires a detection method capable of managing any presented variation.

In order to develop a more reliable post edge detection system, differential gradient techniques were used. A Sobel operator (Sobel, 1970) was used as it presents better noise-suppression (smoothing) characteristics over other techniques (Gonzalez and Woods, 2008). "Such methods come into their own in applications like general scene analysis, where vision systems are required to cope with largely unconstrained image data" (Davies, 2005). The Sobel operator was used as it provided better results than the more commonly used Canny edge detector. Many of the other edge detection methods are not fast enough to achieve real-time operation when consideration is given to the additional algorithms for classify post location.

Figure 7-30: An example of a post mostly indistinguishable from the background scene.

Local maxima are found within the edge data to identify where the edge actually lies. A global threshold can cause edges to be missed as the threshold has to be set high enough to remove most of the noise. This approach also helps with consistent identification with varying illumination intensities. If a global threshold was set, changes in illumination could make the desired edges less significant and therefore not detected. Therefore,

Figure 7-31: Edge identification from Sobel operator – where vertically spanning clusters of a targeted size are identified and drawn. Image converted to black and white to emphasise results.

this method aids post edge detection in individual images with variations across the image as well as across several sequential images.

Figure 7-31 shows the results from implementing a Sobel edge detection operator. The results show the vertical or near vertical clusters of identified edge points. Extracting clusters in this way eliminates most of the undesirable edge noise that would otherwise be present.

7.4.2 Development Conclusions

Although promising results were initially achieved, the significant on-orchard variation makes reliable operation unrealistic with machine vision. Not only do the posts' appearances vary, they are often indistinguishable within the images due to partial occlusion. This occlusion blends them into the surrounding orchard. Although the majority of posts are relatively conspicuous and are detectable with the presented techniques, inconspicuous posts limit reliable identification. A navigation system for a commercial AKH is restricted by the requirement for reliable operation under all conditions. Other techniques and processing methods (e.g. filtering techniques like smoothing (Gonzalez and Woods, 2008), phase congruency (Kovesi, 1995), blurring (Davies, 2005), scale-invariant feature transforms (Lowe, 1999), morphology (Haralick, 1987)) and post processing to help remove shadowing effects are available to aid in the detection process (beyond those outlined in this development). These techniques were not tested for this application as the posts are commonly indistinguishable or occluded. These conditions hinder reliable operation.

Figure 7-32: Sequential row posts difficult to distinguish against background scene.

When a post is occluded, edge sections are often still visible. Although these edges could be identified, they do not provide reliable post identification or verification. Therefore, false positive identification could result. This could lead to collisions and inaccurate row driving.

This conclusion guided development to investigate alternative navigation methods. The complimentary combination of LIDAR and machine vision was selected as the next development step. LIDAR was added to gain reliable object detection. Development of this system is presented in the following section.

Figure 7-33: Example of the variations in on orchard illumination. Note the diminished edge contrast in low illumination areas or where the background is a similar colour to the posts

Figure 7-34: Adjacent row to image above – opposite direction of travel. Note the indistinguishable posts hidden in the shadowing to the left as well as the post occluded by the kiwifruit vine (right).

Figure 7-35: Shadowing, background colouring and vine placement (next to posts) make post identification difficult for a robust implementation.

7.5 Navigation using LIDAR, Machine Vision and a Compass

The combination of LIDAR and machine vision has produced more accurate and reliable navigation in several applications (Baltzakis, et al., 2003, Barawid, et al., 2007, Moghadam, et al., 2008, Subramanian, et al., 2006) (see section 7.3.1.5). This synergistic combination of machine vision and LIDAR allows the strengths of each technology to be used to complement the weaknesses of the other. LIDAR's main benefit is reliable and accurate object detection and positioning. This was lacking in the machine vision only system. Although LIDAR is predominately used indoors (Baltzakis, et al., 2003, Moghadam, et al., 2008, Travis, et al., 2005), adaptations to outdoor application are increasingly more common (Barawid, et al., 2007, Holmqvist, 1993, Huang and Barth, 2009, Subramanian, et al., 2006).

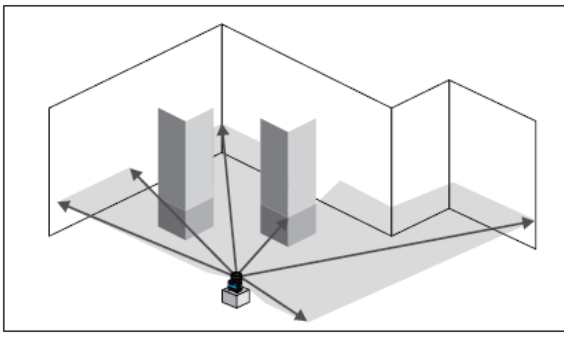


Figure 7-36: LIDAR measuring principle (Sensors, 2008)

LIDAR provides more than reliable and accurate object detection. Wide scanning areas and lower data processing requirements are also gained (Huang and Barth, 2009). However, a compromise between cost and functionality is required to allow commercial use (see section 7.3.1.4). Single plane two-dimensional LIDAR units were selected for their cost, functionality, reliability and robustness. This model is significantly cheaper than the four-plane models ($4,200 versus $20,000) and provides a larger scanning area (270° versus 180°). LIDAR scanners measure the time-of-flight of laser light reflected from a target object, where the time-of-flight directly correlates to the object's distance. The measurement beam sweeps over a flat plane (two-dimensions), capturing distance measurement readings at small incremental angles. SICK LMS111 LIDAR units were selected as they are 'outdoor' rated (IP67), are cost effective ($4,200 each) and scan over a 20m range through 270° every 0.25° (angular resolution) at 25Hz. For LIDAR scanning, object detection depends on both target size and reflectivity. Therefore the probability of detection decreases as distance increases (Huang and Barth, 2009).

In this combined approach, LIDAR is the primary object detection and positioning device. However, LIDAR units can only identify objects if they present a unique and distinctive profile to the scanner. Therefore, most of the detected objects cannot be identified with the LIDAR

data. To overcome this, machine vision will add system intelligence by identifying the detected objects. This will enable higher levels of navigation decision-making to be achieved. For example, the machine vision could identify if a detected in-row object is a branch or a more substantial object. If identified as a branch, the ATP can ignore the threat and drive through it, increasing task efficiency. This is not achieved with LIDAR alone to the required level of reliability. To gain this functionality the LIDAR-detected positions must be mapped into camera image space. Furthermore, in the ultimate commercial implementation, it is envisaged that the machine vision system will detect and identify objects that do not intercept the LIDAR's scanning plane. Stereoscopic machine vision (see section 5.2.3) will localise these object positions. For example, the 'dead men' at a row's end are short and their position is currently estimated. Machine vision could be used to more accurately identify and locate these objects. This will ultimately deliver greater awareness and navigation decision making.

Barawid, et al. (2007), Subramanian, et al. (2006) and Subramanian and Burks (2009) have demonstrated success with their respective LIDAR/machine vision complimentary systems. These are developed for navigating citrus orchard rows. Subsequent work by Subramanian, et al. (2006) showed their system's ability to perform a row-row headland turn. This was performed using machine vision to identify the row's end. Machine vision was also used for navigation where fruit trees were missing. Machine vision operation was required due to the LIDAR's high viewing angle (see section 7.3.1.5), limiting scanning to a short section of the row. The option of using machine vision was also possible due to the specific on-orchard environment (Barawid, et al., 2007, Subramanian, et al., 2006). In this environment, a large visible contrast between fruit trees and the ground exists. This allows reliable visual interpretation. However, reliable visual interpretation was not achieved for kiwifruit orchards (section 7.4). This places a reliance on the LIDAR unit to detect the orchard rows, even when presented with variation.

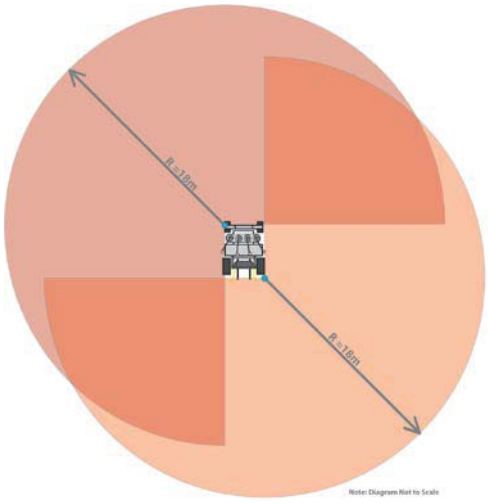


Figure 7-37: Dual LIDAR configuration for the ATP

Two LIDAR units will detect on-orchard objects for the ATP. The two LIDAR units are set to scan horizontally, at a height to intercept the posts and plants (approximately halfway up) with-

in the rows. By using two scanners the ATP's entire periphery can be scanned (Figure 7-37). This helps detect all objects, even when performing tight turning manoeuvres. A kiwifruit orchard is very different to a citrus orchard. In a citrus orchard, tree foliage encapsulates the acquired range data to the current row as the dense trees form hedge like structures, keeping the laser beams within the current row. A kiwifruit orchard is relatively open, allowing the scanners to detect the neighbouring rows and surrounding areas that are within the scanners' range. To achieve on-orchard operation, the unresolved issues presented by both Barawid, et al. (2007) and Subramanian, et al. (2006) need to be addressed. These were:

1. Row position when fruit trees were small or missing in a row
2. Row ends with LIDAR data alone

Machine vision was used to resolve these situations for the other systems. This is not an option within kiwifruit orchards (as previously explained) and has to be resolved using LIDAR data.

Although navigation using a single forward facing LIDAR unit is possible, the dual LIDAR configuration provides safer operation. Humans, animals or other objects could move or be placed into the ATP's path. These would be invisible to a forward facing LIDAR unit while the ATP is reversing. The dual LIDAR configuration also increases storage bin finding efficiency (section 7.5.3).

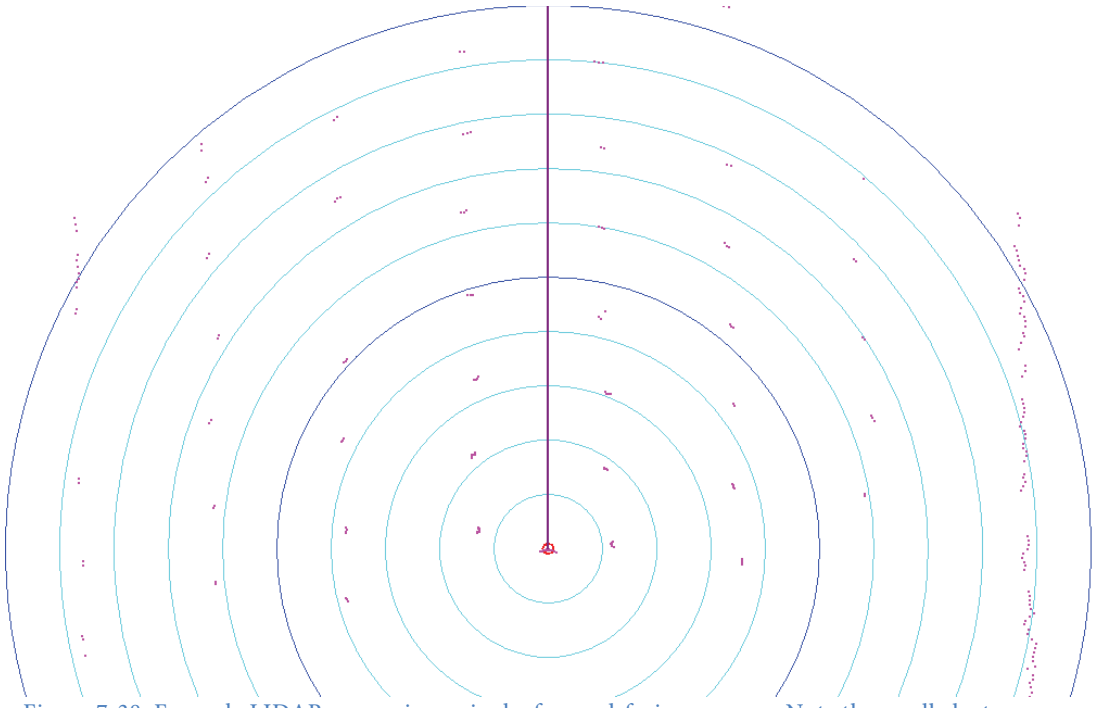


Figure 7-38: Example LIDAR scan using a single, forward-facing scanner. Note the small clusters resembling posts and trunks. A hedgerow is detected (far right). The outer circle (blue) is radially 20m from centre

A fluxgate compass provides a heading direction to assist orchard row interpretation. This informs the analysis of an approximate row direction. The row's direction is interpreted from a supplied GPS orchard map. A fluxgate compass ($1,800) was chosen as it is considerably cheaper than a fibre optic gyroscope (between 1/10 to 1/100 of the cost, as used by Barawid, et al. (2007)), or inertial measurement unit (used by Subramanian and Burks (2009)). In addition, these compasses do not suffer the high level of error introduced from engine vibration, like the cheaper Micro-Electro-Mechanical (MEMs) compasses do. These MEMs compasses suffered from ±10˚ (approximate) heading error due to oscillations introduced by engine vibration despite averaging, damping and shielding.

7.5.1 Orchard Row Navigation using LIDAR

A novel method to determine the desired driving line along the row was required. It is necessary to identify posts and vines to achieve this and allow row direction and position to be determined. These row elements are common to pergola-style kiwifruit orchards. It requires a custom interpretation method as layout is unique. The vines' position, relative to the posts, alters between orchards and sometimes within an orchard or block. For example, the trunks can be positioned directly next to the post, or midway between them. This alters the orchard's appearance to the LIDAR scanner. Therefore, the interpretation algorithms must be capable of dynamically handling these, or any other, variation. Consistent, uniform and linear post spacing is obviously desirable; however, in reality variation is always present. This is most evident where alterations to the orchard structures have occurred. The most common variations include:

1. Width variation: rows widths alter between rows and along the same row. Rows can also taper in width. Row widths range: 3.5 – 5.5m
2. Inconsistent row widths and post spacing (even within adjacent rows)
3. Missing plants and posts within the rows
4. Uneven row lengths
5. Objects within a row
6. Terrain variations

A real-time, autonomous navigation system that dynamically adapts to these variations is therefore required.

Orchard row navigation begins by finding the orchard row's heading. This is interpreted from the supplied GPS orchard map. The on-board fluxgate compass tracks row direction. The compass automatically corrects for pitch and roll (accuracy ±1° with engine vibration). The row heading is projected onto the LIDAR data (Figure 7-39, *1*) to help row finding. LIDAR data interpretation starts using the known row direction. Firstly, clusters of data points are grouped using the linear distance between them. As the LIDAR scans in an anticlockwise arc, only sequential points require assessing. This is because points correlating to a single object are represented sequentially in the data from the scanner.

The current row's posts and vines must be identified once the data clusters are grouped. The rest of the LIDAR information is ignored for row driving. This additional information is used for collision detection and avoidance. From the projected compass heading each cluster's size (spread of data points), weight (number of data points) and perpendicular distance (from the heading) are assessed. The in-row objects were reliably detected when:

1. Size: There was greater than two and less than twenty-six LIDAR data points representing an object. The large range is required to detect far off, as well as close posts.
2. Spread: The object did not span more than 250mm and was bigger than 100mm. This covered the range of trunk and post widths.

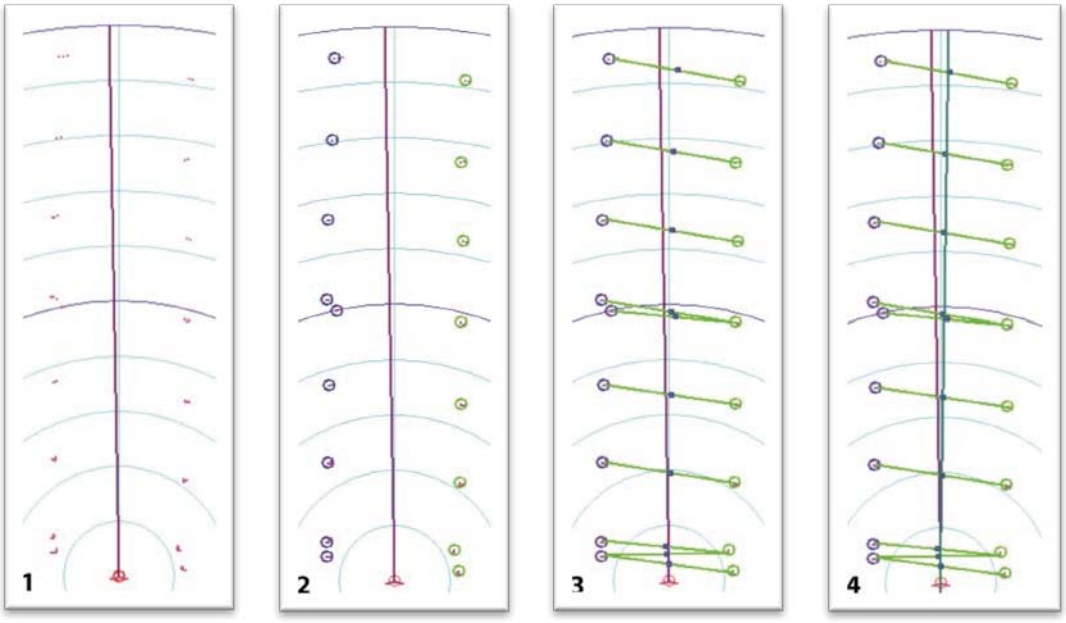


Figure 7-39: Row drive interpretation stages (a section of a complete scan showing relevant information). 1 – Raw scan with compass heading projection (purple line) and current heading (light blue), 2 – Row information identification, 3 – Small blue circles define the row's centre, at that point in the row, 4 – Row drive path fit to row information (Teal). Outer concentric arc is radially 20m from centre.

3. Perpendicular distance: Less than, or equal to four metres. This range was to account for the varying row widths and any error in compass heading.

Figure 7-39 - image *2*, shows an actual orchard scan with the analysis identifying the necessary row elements.

To achieve this classification the following calculations are performed. Size and spread characteristics for each cluster are determined using the distance between the data points (by Pythagoras). To find the perpendicular distance to the cluster, the projected compass heading is described in the form $Ax + By + C = 0$ where the slope,

$$-\frac{A}{B} = \frac{\Delta y}{\Delta x} = \frac{y}{x} \tag{101}$$

and $C = 0$ (as the line is projected through the origin $(0,0)$). This allows the perpendicular distance, d, of point (m, n) to be calculated by,

$$d = \frac{|Am + Bn + C|}{\sqrt{A^2 + B^2}} \tag{102}$$

Once these groups of points are classified, those identified as row objects are stored. The side these data points lie on, relative to the heading, is found. This allows the groups to be paired across the row (blue and green circles, Figure 7-39, *2*). Next, the centre of mass of each group is found.

For each left-hand object (blue circles, Figure 7-39, *2*) the distance to every right-hand object (green circles) is found. Each left object is paired with its closest right object. If the linear distance between them is greater than 3.5m and less than 6m (green line, Figure 7-39, *3*) the midpoint of the two objects is calculated (small blue circles, Figure 7-39, *3*). These midpoints points are used to determine the ATP's driveline. However, they do not need to be at the midpoint. These points would be shifted from centre if offset row driving was required. From these points, a least squares regression produces a line-of-best-fit to the centre positions (Figure 7-39, *4*) to compute the desired driveline. A large shift in the position(s) of a post(s) or vine(s) can produce an error in the best-fit line. This is removed during the regression by assessing each midpoint's perpendicular distance to the line. Any outlier points (large variance) are removed to ensure an accurate driveline is found.

The row's end must also be detected while driving. Turning and other manoeuvres are performed by finding the row's end. This allows continuous operation. To achieve this, the pairs of row objects in front of the ATP are counted (green lines, Figure 7-39, *3-4*). The end of row is classified as the furthest pair when there are less than five pairs and the furthest pair is less than six metres away. A row-to-row turn (section 7.5.2) is used when the end of the row is reached.

While driving along a row, uneven or changing terrain gradient can cause the LIDAR units to detect the ground or kiwifruit canopy (Figure 7-40). Although kiwifruit orchards have a variety

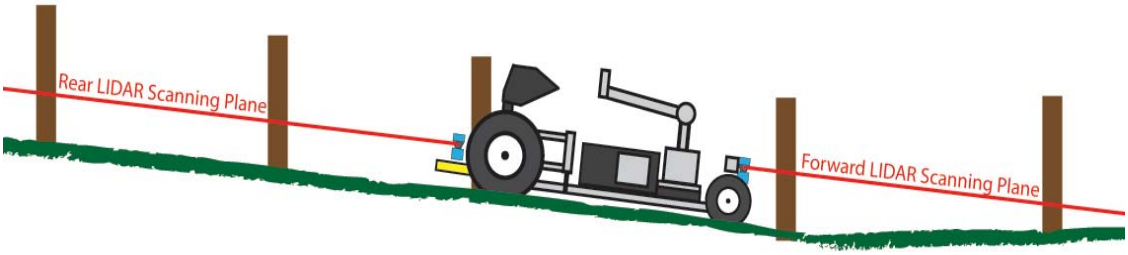


Figure 7-40: Dual LIDAR configuration for reducing gradient change effects for on-orchard navigation

of physical fluctuations, they are relatively smooth. This means that they do not have significant sudden changes in gradient, where the LIDAR's scanning distance would be limited as it would intercept the ground or canopy. This would impede row driving. Any gradient change effects are minimised by using a forward and a rear facing LIDAR unit (Figure 7-40). This provides sufficient viewed row distance during gradient changes to allow continuous operation.

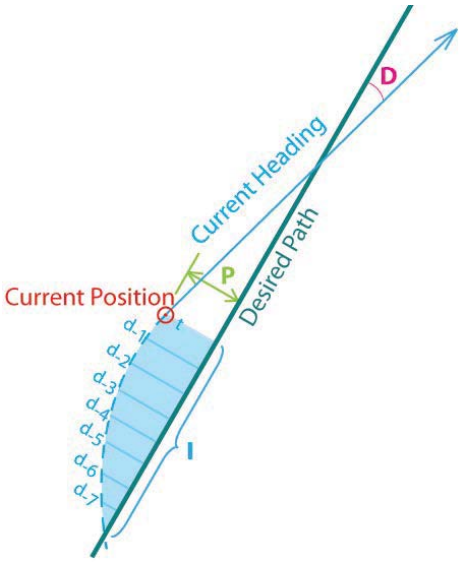


Figure 7-41: PID control variables determination

Now that the driveline is known, it needs to be followed. To achieve line following a PID control system is implemented. To calculate the control values (*P*, *I* and *D*) the driveline acts as the set point. The integral is determined at distance intervals to ensure consistent control when driving at different speeds. The PID controller manipulates the ATP's driving radius of curvature to keep it on the desired path.

A Ziegler Nichols tuning method (Ziegler and Nichols, 1942) for a closed loop system was used to find the appropriate control variables. Control with zero steady state error and minimal overshoot was sought. Subramanian and Burks (2009) also successfully used this tuning method. However, their control variables were found using a simulation model in Simulink.

7.5.1.1 Orchard Row Drive Collision Detection and Avoidance Strategy

It is inevitable that objects will be found in the ATP's drive path. These objects will most likely be vines, storage bins, random posts or even people. Unlike the solid objects, dangling vines do not pose a danger to driving. On the other hand, avoiding colliding with people is critical for safe operation. However, the LIDAR cannot reliably detect the difference between these objects. To achieve object identification the LIDAR data points will be translated into the camera's image space. This will allow for high-level, navigation decision making using machine vision analysis.

Objects positioned near the row's centre restrict in-row avoidance. If there is enough space to drive around a solid object, the driveline is offset until the object has passed. Detection of unavoidable in-row objects automatically stops row driving. If this occurs, the ATP will reverse out of the row and use an adjacent row to avoid the object. The object's location is recorded for future reference to optimise operations. For unidentified objects, an image of the object will be sent to a human supervisor's electronic device (iPad or similar). The object will be highlighted from the LIDAR data to clearly identify the object of concern. From this information, the supervisor will tell the ATP to (from the remote device):

1. Ignore the object and continue along the row ('push through')
2. Back out of the row and navigate around.

Although a human supervisor will reduce the systems autonomous nature. This was considered necessary, especially in the early operational stages to ensure there are no operating issues. The supervisor is also required to load/unload the units from transportation trucks, carry out any servicing and direct units to their appropriate operating block. The cost of this labour unit has been considered in the commercialisation model.

These unidentified objects will be recorded and used to improve the machine vision's interpretation ability. The LIDAR profile and orchard position of the object will also be recorded to allow automatic avoidance strategies to be carried out on future encounters. This information will be shared with the other ATP units on the orchard (if any) through their wireless communication link. Upon return to their base, this information will be uploaded to a base station to update all the ATP units.

7.5.2 Row-to-Row Transition

One of four manoeuvres will be used when the end of the row is reached:

1. Left turn into adjacent row
2. Right turn into adjacent row
3. Perform a three-point turn and go back up the same row (for offset row driving)
4. Stop operation and drive along the headland (for finding storage bins or similar)

The first two manoeuvres only alter by direction. They both need the adjacent row to be identified (in the appropriate direction) to allow the move to be calculated and performed. However, the three-point turn does not require any adjacent rows to be identified as it is performed within the same row (see section 7.2.4.6).

The required manoeuvre starts when the end of the row is identified (as outlined in section 7.5.1). Next, for manoeuvres 1 and 2 (above) the adjacent row is determined. This is achieved by offsetting a line parallel to the current driveline by the row width (depicted in Figure 7-42), where the row width is the average distance between the paired row objects. Once this estimated row line is defined, row identification is calculated using the same method as in section 7.5.1. This estimated row line becomes the reference line to calculate the adjacent row's position. Once the adjacent row is found, its end is determined as the pair of row objects furthest from the ATP (Figure 7-42). The midpoint between the end row objects is used to define the end of the move (Figure 7-42, (*1*)). This also defines the moves radius *R*. The manoeuvre radius, *R*, is set to half the perpendicular distance from the current desired path to point (*1*). This is calculated using (102).

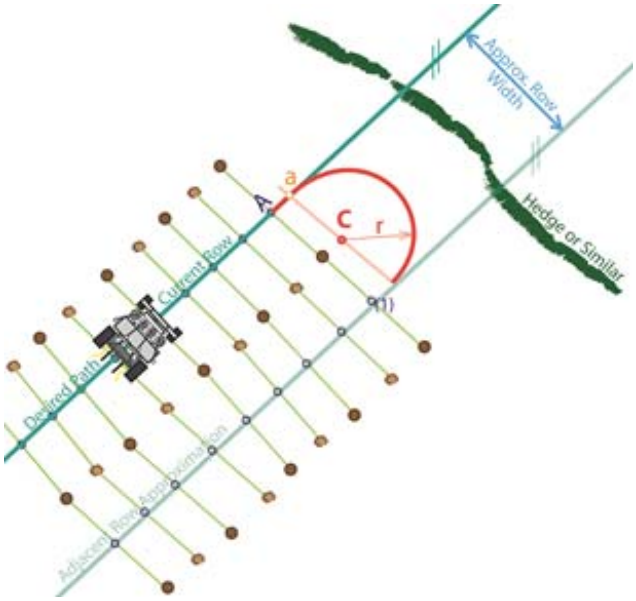


Figure 7-42: Row-to-Row transition procedure

To ensure the ATP does not collide with any structures or obstacles an adjustable offset from *A*, Figure 7-42, is included (length *A*-*a*). Point *a* is initially set to 0.5m from *A* to avoid the dead

men supporting the canopy structure. Length *A-a* is adjusted to give position *a* using the method in section 7.5.2.1 to avoid any collisions. However, the 'dead men' are often lower than the LIDARs scanning height so they are not detected. Their position is estimated from the cable connecting them to the main canopy structure. This cable is detected with the LIDAR scanner and the machine vision system will confirm the position of the 'dead men' in future development.

When position *a* is reached (Figure 7-42) the ATP starts to drive at the specified radius of curvature *R*. The ATP continues to drive this curve until its heading is within 10° of the new row's heading. When this condition is met, the process of finding the new row begins. This is the same as outlined in section 7.5.1. This method produces a smoother transition into the new row as the actual driveline will be slightly different to its estimated position.

To successfully perform this manoeuvre the centre of the row-to-row transition (*C*, Figure 7-42), is required. This needs to be consistently positioned over consecutive LIDAR scans whilst approaching position *a*. By referencing the projected drive path and the row's end (*A*) the turn's centre can be calculated. Unfortunately, the driveline is not guaranteed to intercept point *A* as it is a 'line-of-best-fit' and *A* is only one centre location. Therefore, *A* is projected onto the driveline to determine point *a* using point A_A. This is done by forming a line that intercepts point *A* and is perpendicular to the driveline. The point where the driveline and the line through *A* intercept is defined as the centre of the end of the row (A_A). This is calculated by expressing the driveline as a line, where the perpendicular line intercepting *A* is expressed as,

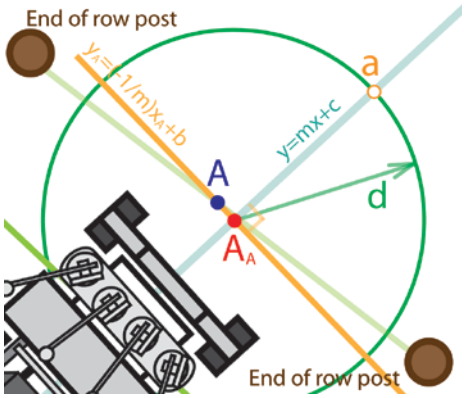


Figure 7-43: Method to work out the end of row position on the driveline (A_A)

$$y_A = Mx_A + b \tag{103}$$

Where gradient *M* is the negative reciprocal of the drive line gradient,

$$M = \frac{-1}{m} \tag{104}$$

providing an expression of the line through A, perpendicular to the drive line. Intercept b is determined by rearranging (105) and inputting (x_A, y_A).

$$y_A = \left(\frac{-1}{m}\right) x_A + b \tag{105}$$

At the intercept of the driveline and the line through A, x and y must correspond, therefore the straight-line equation for the driveline and (105) are combined. Solving for x to obtain the intercept and simplifying, where c is the y intercept of the driveline,

$$x = \frac{\left(\frac{-1}{m}\right) x_A + b - c}{m} \tag{106}$$

$$x = \frac{-1}{m^2} x_A + \frac{b}{m} - \frac{c}{m} \tag{107}$$

Inserting x into the straight-line equation provides the corresponding y intercept value. Now the adjusted value of A is known (A_A). A_A is the position of A shifted to the driveline.

From A_A, position a is required. A circle, with radius, d (Figure 7-43), centred at A_A, intercepts the driveline at a. The value of d is set to the required length A_A-a, this has an initial value of 0.5m. However, this length will dynamically alter if a potential collision is detected (section 7.5.2.1). To compute point a, both intercepts of the circle and driveline are required. Where the circle with radius d is described as,

$$d^2 = (x_a - x)^2 + (y_a - y)^2 \tag{108}$$

where the driveline's slope m is,

$$m = \frac{y_a - y}{x_a - x} \tag{109}$$

rearranging (109), solving for y_a,

$$y_a = y + m(x_a - x) \tag{110}$$

substituting (110) into (108) and simplifying,

$$d^2 = \left(x_a - x\right)^2 \left(1 + m^2\right) \tag{111}$$

rearranging for x_a,

$$x_a = x \pm \frac{d}{\sqrt{\left(1 + m^2\right)}} \tag{112}$$

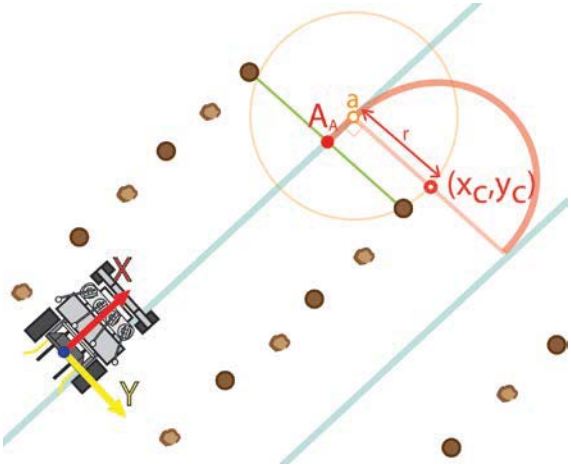


Figure 7-44: End-of-row turn, centre of arc calculation (x_C, y_C).

With this solution, the value of x_a larger than x is used as the coordinates are relative to the ATP. By substituting $x_a = x$ into the straight-line equation provides y_a. Position a, (x_a, y_a), is now known.

From position a a line, perpendicular to the drive path is projected. The desired drive arc's centre (x_C, y_C) is on this line at distance r from a. Therefore, where a circle of radius r intercepts the projected line, that is perpendicular to the driveline, provides the desired curve's centre (x_C, y_C). This circle is described as,

$$r^2 = \left(x_c - x_a\right)^2 + \left(y_c - y_a\right)^2 \tag{113}$$

The line that intercepts the circle at C has a gradient M which equals the negative reciprocal of the driveline gradient from (104). This gradient is equivalent to,

$$M = \frac{-\left(x_c - x_a\right)}{\left(y_c - y_a\right)} \tag{114}$$

rearranging to solve for x_C,

$$x_c = x_a - M\left(y_c - y_a\right) \tag{115}$$

substituting (115) into the projected circle (113),

$$r^2 = \left(x_a - M\left(y_c - y_a\right) - x_a\right)^2 + \left(y_c - y_a\right)^2 \tag{116}$$

simplified and factored to,

$$r^2 = \left(1 - M^2\right)(y_c - y_a)^2 \tag{117}$$

rearranging to solve for y_c(the drive arc's centre),

$$y_c = y_a \pm \frac{r}{\sqrt{\left(1 - M^2\right)}} \tag{118}$$

Two solutions are possible for (118). The solution chosen is dependent on the direction of the turn. Value x_c is found by substituting y_c into the straight-line equation representing the line perpendicular to the driveline, through *a*. With a consistently positioned row-to-row centre, the navigation system can calculate when to execute the manoeuvre. Furthermore, object collision avoidance (section 7.5.2.1) can be achieved by consistently working out the manoeuvre's centre.

7.5.2.1 Collision Detection and Avoidance Strategy for Row-to-Row Transitions

Collision with objects is possible during row-to-row transitions. A collision detection and avoidance strategy is consequently implemented. This strategy uses the predefined row-to-row drive path (orange semicircle, Figure 7-45) determined in the previous section. The drive path is the proposed path of the ATP's coordinate space origin. However, the ATP's body will cover an area either side of this line. Therefore, a drive area is defined from the inner and outer ATP body extremities (blue and green curves respectively, Figure 7-45). These are placed on either side of the main path. Any object between R_a and R_b ($R_a \leq R_P \leq R_b$, Figure 7-45) over the 180° span of Φ_P is in the ATP's drive path. This area is called the exclusion zone. Position *a* (Figure 7-42) is adjusted along the drive line to ensure the ATP does not collide with any detected objects. The aim is to keep the exclusion zone clear.

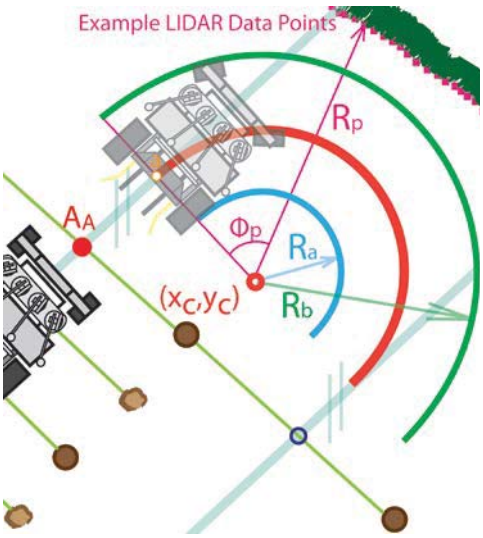


Figure 7-45: Row transition collision avoidance. Refer to Figure 7-42 for additional detail.

The centre of the drive path's arc (x_C, y_C) is referenced to determine if an object is within the exclusion zone. Point (x_C, y_C) is calculated in the previous section. Once the move is calculated the subse-

quent LIDAR scans are used to look for potential collisions. If a collision is detected, length A_A-a (Figure 7-42) is altered to avoid it. This is updated for every LIDAR scan until the end of the row is reached. A collision is detected by assessing its radial distance, R_p, and angle, Φ_p from the centre of the turn (Figure 7-45). This is calculated with respect to the curves centre (x_C, y_C), and base line α (Figure 7-46). Where (x_P, y_P) is the LIDAR data point's position, P, with radius R_P,

$$R_p = \sqrt{(x_p - x_c)^2 + (y_p - y_c)^2} \tag{119}$$

The angle of this point is also required. This is calculated with respect to the baseline α. Therefore from the Law of Cosines,

$$\beta^2 = \alpha^2 + R_P{}^2 - 2\alpha R_P \cos\Phi_P \tag{120}$$

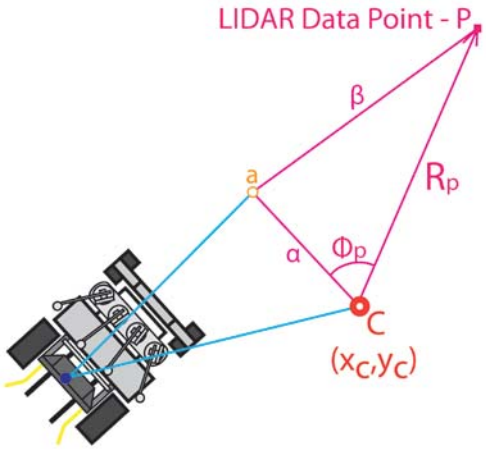


Figure 7-46: Calculation of the data point angle, Φ_P. Elaborated from Figure 7-45

Solving for data point angle Φ_P,

$$\Phi_p = \cos^{-1}\left[\frac{\beta^2 - \alpha^2 - R_P{}^2}{-2\alpha R_P}\right] \tag{121}$$

As positions a, P and C are known, lengths α and β are calculated by,

$$\beta = \sqrt{(x_p - x_a)^2 + (y_p - y_a)^2} \tag{122}$$

$$\alpha = \sqrt{(x_a - x_c)^2 + (y_a - y_c)^2} \tag{123}$$

Point a (Figure 7-42) is shifted along the drive path by distance D if the detected object is within the exclusion zone. The shift distance, D, for point a, when the conflicting LIDAR object is on the inside of the drive curve of radius r,

$$D = (R_p - R_a)\cos\Phi_p \tag{124}$$

otherwise,

$$D = (R_b - R_P)\cos\Phi_P \tag{125}$$

This provides a clear row-to-row manoeuvre path for safe and continuous operation.

7.5.3 Fruit Storage Bin Detection

Fruit storage bins are used to store and transport harvested fruit. They are an essential part of the harvesting operation. Therefore, to achieve fully autonomous operation the ATP must be capable of handling these bins. Typical operation involves dropping off a full bin, then locating, docking with and picking up an empty bin. Empty bins are left in the orchard's headland area to be collected for transportation to the packhouse. Achieving this functionality allows continuous operation, however it requires the bins to be located.

The LIDAR sensors are primarily used to locate the bins. Reliable LIDAR detection is achievable as the storage bins present a unique profile to the LIDAR scanner. This profile is formed from the bin's flat 1200x1200mm faces that are between 550 and 650mm high. It is proposed that machine vision will be integrated to advance this bin identification system. Machine vision will confirm the object is a bin and whether it is full or empty. This will occur when the ATP is within a few metres of the bin when sufficient detail is evident. This will reduce time spent docking and lifting the bin to find out if it is empty.

To detect bins in the LIDAR data (purple points, Figure 7-47), the data points are first grouped. This is achieved by assessing the distance between them. Data points belonging to a single object are presented sequentially in the scanner's data as the scanner sweeps in an anti-clockwise direction (when observed from above). This means that only neighbouring data points need to be assessed for grouping.

The unwanted groups are removed once all the points are grouped. This process starts by removing those data groups that are unlikely to represent a bin. A specific range of data points will represent a bin as it is a set size. The number of points is dependent on its distance from the scanner and the number of visible sides. Therefore, the number of data points within the group is measured with respect to the group's distance from the scanner. This discards the bulk of information representing the surrounding environment and is not relevant for bin finding.

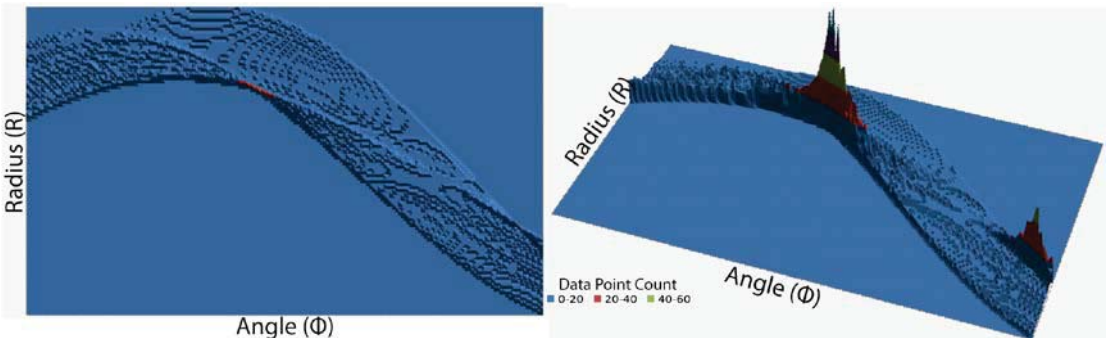


Figure 7-48: Hough Space plots for a cluster of data points representing a fruit storage bin. Note the two peaks 90 degrees apart correlating to the two bin edges visible to the LIDAR scanner.

Once the groups of interest are identified, further analysis confirms if they represent a bin. The confirmation process begins by analysing the identified groups with a Hough transform (Duda and Hart, 1972) to assess the linearity of the points. As one or two bin faces can be viewed at any one time, the analysis measures the linearity of the points for up to two lines. The scanner's position relative to the bin determines whether one or two faces are visible. This is shown in Figure 7-47 where only a single bin face is visible to the LIDAR scanner. Although only a single bin face may initially be detected, two faces become visible as the ATP approaches the bin. This is due to the offset LIDAR mounting positions (Figure 7-37). Having two faces visible is beneficial to confirming bin identity and provides more reliable detection. The two faces allow both edges to be measured as well as the angle between them. This provides further correlation that the object is a bin. The angle between the two faces is detected with the Hough transform,

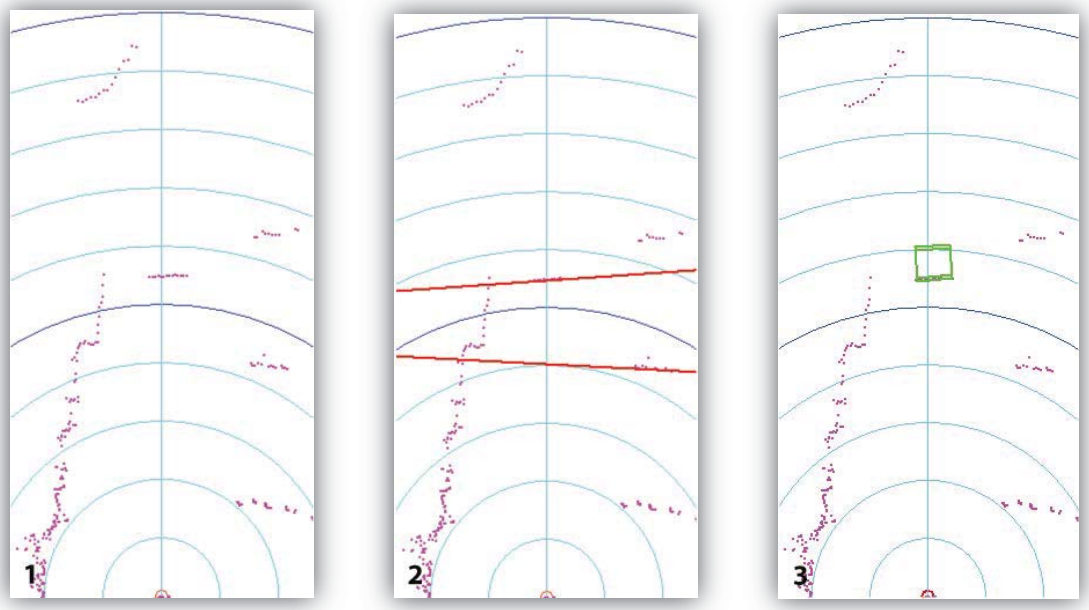


Figure 7-47: On orchard LIDAR data scans: 1 – Unanalysed scan, 2 – Lines from Hough Transforms on the identified groups (lower line later rejected) and 3 – the determined bins position (bin profile drawn in green).

where two correlations in Hough space are expected to present 90° apart. Figure 7-48 shows a graphical representation of two edges shown in Hough space from a LIDAR identified group representing a bin.

In Hough space a line is shown in the data as a local maxima. The height of this maxima represents the significance of the line within the data set. Hough space analysis (Duda and Hart, 1972) assesses the radial distance and angle of each data point over a given range. Therefore, the data points representing a line must be linear to achieve a strong Hough space correlation. As the LIDAR data points are supplied to millimetre accuracy but are only accurate to ±20mm, the data points do not all sit perfectly in a line. This results in a blurred maxima in Hough space and could result in misinterpretation. Therefore, each data point's radial distance is rounded to the nearest 10mm help achieve linear data. This is represented as a large accumulator value in Hough space. Next, the local maxima are identified in the data set. These maxima must be over a set threshold to ensure the correlation is significant enough to be a line. This threshold is proportional to the total number of data points in the group. This ensures that the significance of the maxima is related to the number of points presented. Setting this ensures the bin is still detected when it is further away as there are fewer data points representing it.

Detecting one or two perpendicular lines does not confirm that the item is a bin. At this point the item could be a section of hedge or the edge of a small building. Therefore, further validation is required. As the LIDAR scanner supplies relatively accurate range information, all the detected data points should be very close to the inferred line. If it was a hedge row that was detected the data points would be more variable than on a flat surfaced bin. To test this, the perpendicular distance of each data point to the determined line is assessed for every point within the group. The variance of these points is assessed to find how well they match the line. Where two lines are identified, the perpendicular distance of all points to both lines is calculated and the shorter of the two distances is used to assess the variance of the points to the lines.

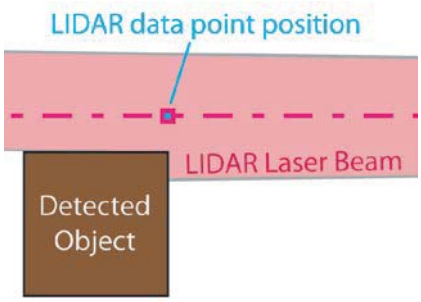


Figure 7-49: LIDAR data point position with a diverging laser beam. Note the data points position on the partially incident object

If linearity testing is confirmed, the object's dimensions are measured. Although the LIDAR scanner accurately measures range its dimensional measurement ability is not as accurate. This is due to the conical LIDAR beam

shape making dimensional measurement accuracy decrease with object distance. The LIDAR detected distance is reported as relating to the middle of the laser beam. In reality, the laser beam might only partially intercept the object, but still return enough reflection to produce a distance measurement. This produces a data point that is not on the actual object (Figure 7-49). Consequently, the object's detected size may differ from its actual size. As the laser beam gets bigger with distance the size error is also proportional to distance. Therefore, some tolerance in the bin's LIDAR measured dimensions, proportional to distance, must be included.

To calculate the object's dimensions the data points are perpendicularly projected onto the Hough transform inferred edge(s). This reduces some measurement variation. Next, the span and centre of the points is computed to measure and position the bin. If the detected object fits within the predefined size range, the object is confirmed as a bin for that LIDAR scan.

If the object is confirmed as a bin, its position and orientation are required. Interpreting this information allows manoeuvres to be calculated to engage and dock the ATP with the bin. The bin's position is derived from the centre position of data points spanning the incident line(s) (determined above). Its orientation is provided from the slope of the line representing the bin's face found with the Hough transform. Although its position is now known, the bin must be engaged on the appropriate side. This is governed from the under bin support rails which restricts ATP engagement to two sides. Fortunately, the headland areas where the bins are left are relatively narrow. Because of this, tractors/trailers are restricted to drive approximately perpendicular to the rows. This, combined with the trailer's rollers, dictates that the bins must slide off the delivery trailer in the same orientation. Therefore, the bin's docking face is the face that is closest to parallel with row direction. Now that the bin has been localised, the appropriate manoeuvre is required to direct the ATP to engage with the bin. The manoeuvre is determine in section 7.2.4 by setting the end position and direction.

The LIDAR unit scans the environment at 25Hz. For every scan the bin is re-identified and localised to obtain a new ATP engagement path. This accounts for any inaccuracies during ATP driving (due to ruts etc). It also more accurately positions and identifies the bin as it gets closer. Therefore, the small localisation inaccuracies discussed in this section are minimised to acceptable levels as the ATP approaches the bin to dock with it and pick it up.

7.5.4 Alignment of LIDAR, Machine Vision and ATP Coordinates

To allow LIDAR and machine vision to work together the spatial relationship between their respective coordinate systems is required. These are mapped into the ATP's coordinate system to allow any detected objects to be positioned relative to the ATP.

To achieve this, mathematical alignment of the geometric spaces is required. The alignment needs to account for three-space translation, $t=[X,Y,Z]^T$, and rotation offsets (Pan (ρ), Tilt (θ) and Rotation (σ)) between each coordinate space. The key aspect of this is mapping the LIDAR data points into the correct image position. This allows LIDAR detected objects to be positioned in the camera's corresponding image, *(i, j)* for higher level object analysis. The same method is used to align the LIDAR and ATP coordinate spaces. This allows the LIDAR detected objects to be positioned relative to the ATP, allowing navigation decisions to be made.

The relationship between LIDAR coordinates, $L=[X,Y,Z]^T$, and camera coordinates, $c=[x,y,z]^T$, is described as,

$$c = R_{Lc}L + t_{Lc} \tag{126}$$

where, R_{Lc} is the resultant 3 x 3 orthogonal rotation matrix,

$$R_{Lc} = R_x(\sigma) \times R_y(\theta) \times R_z(\rho) \tag{127}$$

using,

$$R_x(\sigma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\sigma) & -\sin(\sigma) \\ 0 & \sin(\sigma) & \cos(\sigma) \end{bmatrix} \tag{128}$$

$$R_y(\theta) = \begin{bmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{bmatrix} \tag{129}$$

$$R_z(\rho) = \begin{bmatrix} \cos(\rho) & -\sin(\rho) & 0 \\ \sin(\rho) & \cos(\rho) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{130}$$

and t_{Lc} is a 3-vector representing camera position with respect to LIDAR. This uses the standard camera pinhole model (Davies, 2005), and gives a position in camera coordinates, $c=[X_c, Y_c, Z_c]^T$. This is not the required pixel, *(i, j)*, position.

To transform the camera coordinate, c, into image pixel coordinate *(i, j)*, the stereoscopic depth equation (13) is referenced. By rearranging (13) the required object's image space, *(i, j)*, position is found,

$$i_c = \frac{Y_c}{X_c} \times mpppm \tag{131}$$

$$j_c = \frac{Z_c}{X_C} \times mpppm \tag{132}$$

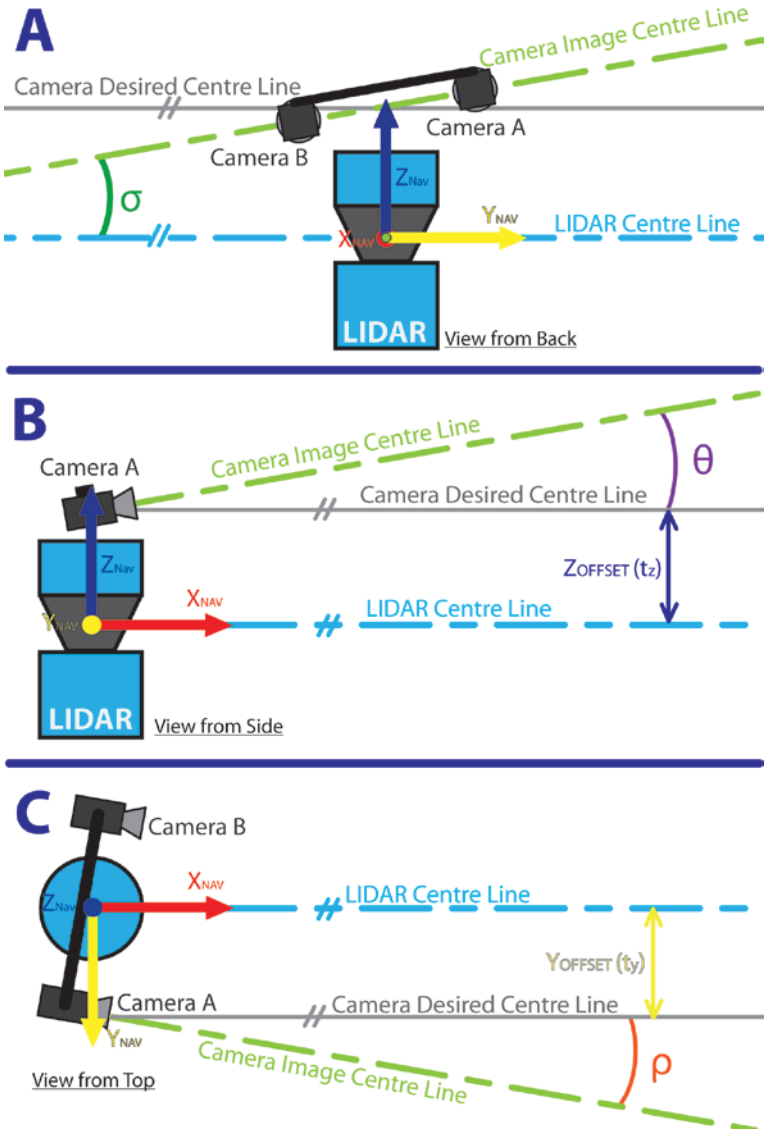


Figure 7-50: Rotational offsets between LIDAR and Machine vision coordinated spaces. A – Rotation (σ), B – Tilt (θ) and C – Pan (ρ)

A similar equation is used to describe the LIDAR's coordinates, $L=[X_L, Y_L, Z_L]^T$, with respect to the ATP, $A=[X_A, Y_A, Z_A]^T$. This is also required for each LIDAR unit, where the conversion of a LIDAR coordinate into ATP coordinates is given by,

$$A = R_{AL}L + t_{AL} \tag{133}$$

where, R_{AL} is the resultant 3 x 3 orthogonal rotation matrix and t_{AL} is a 3-vector representing the translation of the scanner with respect to the ATP's coordinates.

Once this is done the spatial relationship between all coordinate spaces is known. This allows a machine vision data point to be directly mapped into the ATP's coordinates. This is achieved by adding the translations and rotations by,

$$A = \left[R_{AL} + R_{Lc}\right]c + \left[t_{AL} + t_{Lc}\right] \tag{134}$$

7.5.5 Calibrating Alignment Variables between ATP, LIDAR and Machine Vision

The rotation and translation values between the coordinate spaces are required to map the sensors to the ATP. This is required for both LIDAR sensors and the cameras with respect to the ATP's coordinate space.

Eight calibration devices are used to determine each system's alignment. These devices are made from 1.2m x 2.4m timber sheets (Figure 7-53) making them distinctive for LIDAR analysis. In addition, they are painted with a high contrast pattern making them easily and accurately locatable within the machine vision system. These calibration devices are propped to stand approximately vertically within the LIDARs' scanning area (Figure 7-51). The relationship between the coordinate systems is found by identifying these calibration boards with each sensor over several positions.

The LIDAR scanners use the same method to identify the boards as used for bin finding (section 7.5.3). This is done as the boards present the same unique profile as a single bin face. However, there is one difference between the two methods. A line-of-best-fit is used to fit the line to the identified data points representing the board's face. This increases accuracy. The line-of-best-fit is determined with a simplex regression. In the bin finding method the data points are rounded to the nearest 10mm and the line is found with a Hough transform. The line-of-best-

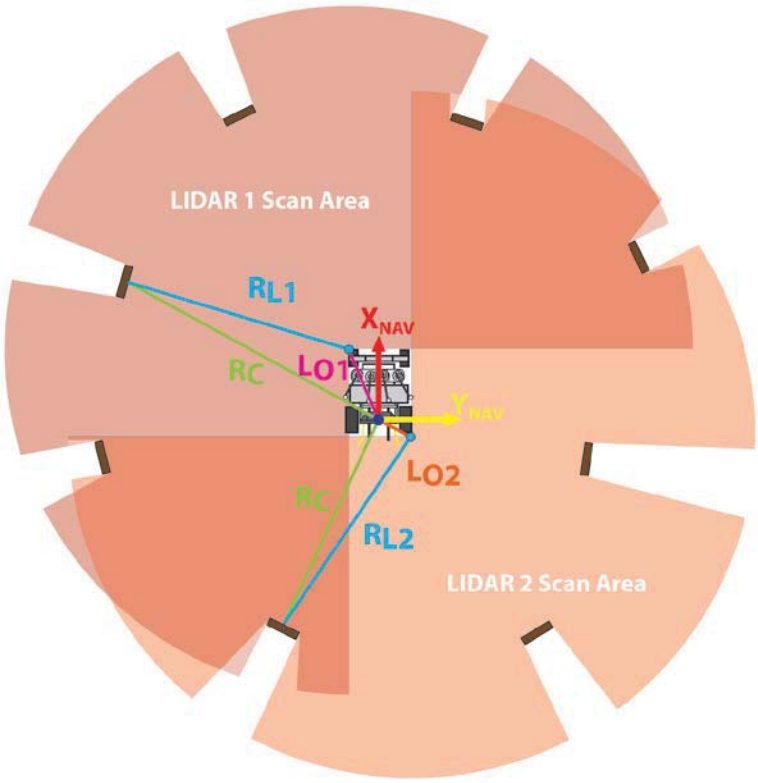


Figure 7-51: LIDAR to chassis space calibration. Note that the LIDAR scanners have overlapping operational areas. Calibration boards shown as brown rectangles.

fit averages the inherent range errors associated with LIDAR measurement. This provides a more accurate measurement of the board's position. In addition, as calibration is performed in an open area, not an orchard environment, it is unlikely that irrelevant data points will be included in the data group representing the calibration board. Therefore, noise within the data points is not expected that could affect the analysis.

To obtain the same measurements in ATP coordinates, manual radial measurements are taken from the ATP's origin (R_G, Figure 7-51). These measurements provide the actual distances to calibrate the sensors' coordinate geometry.

The calibration boards' position in each system's coordinate space is used to map the translation and rotations between them. The first calibration aligns both LIDAR scanners to the ATP's coordinates. Each scanner's translation is manually measured and their rotation estimated from their mounting position. This provides a starting point for the alignment regression. Without this, the regression could optimise the alignment to a local minima, not the global minima corresponding to the best optimisation. To align the coordinate spaces the simplex regression uses the radial distance measurements for each calibration board (ATP, r_A – LIDAR, r_{Lx}, where x is

the LIDAR's number). A theoretical radial distance is calculated with respect to the ATP's co-ordinates to allow comparison to the actual radius, r_A, in ATP space.

The LIDAR unit's three-dimensional translation, T_{Lx}, and rotation, R_{Lx}, values are used to calculate the theoretical radial distance, *b*, Figure 7-52. Using the theoretical radial distance, *b*, the LIDAR unit's translation, T_{Lx}, and rotation, R_{Lx}, values are adjusted to make *b* and the measured radial distance, r_A, the same. This is done with the simplex regression for all relevant boards (those detected by the scanner). The regression is developed to minimise the error between ATP measured radii, r_A, and the inferred radii from the LIDAR, *b*. A data set of calibration board positions is created for the optimisation. The calibration boards are moved radially outwards from six to twenty metres, in approximately two-metre increments. At each point the LIDAR measured and manually measured (in ATP coordinates) board positions are recorded to create the data set. Next, the regression algorithms produce the relationship between the LIDAR and ATP coordinates. This is done with both LIDAR scanners.

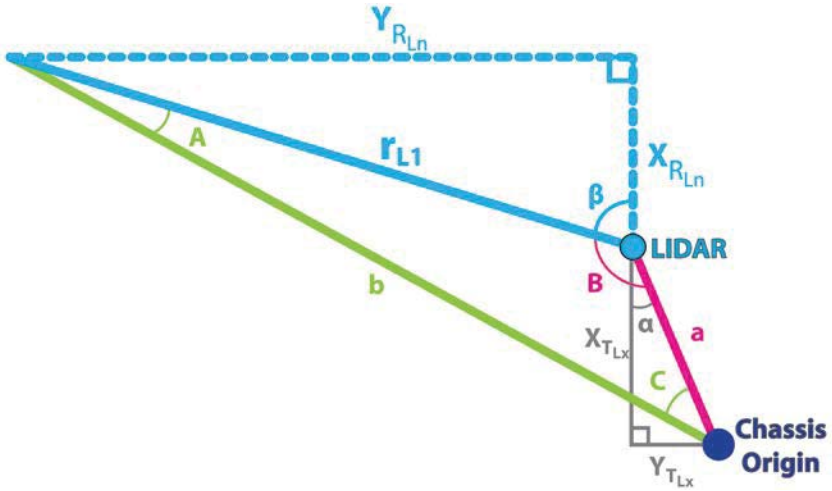


Figure 7-52: Computation of theoretical radial distance of calibration device (b) (from Figure 7-51)

The LIDAR measured distance, r_{Lx}, is used to calculate the theoretical radial distance, *b*, for LIDAR to ATP alignment. Length *b* is calculated using the Law of Cosines, where *a* is the horizontal component of the LIDAR's translation position, T_{Lx}, giving,

$$b^2 = a^2 + r_{Lx}{}^2 - 2ar_{Lx}\cos(B) \tag{135}$$

where angle *B* is defined as,

$$B = (180 - \beta) + \alpha \tag{136}$$

with components,

$$\beta = \tan^{-1}\left(\frac{Y_{RLn}}{X_{RLn}}\right) \tag{137}$$

$$\alpha = \tan^{-1}\left(\frac{L_{nY}}{L_{nX}}\right) \tag{138}$$

Next, the camera's coordinates are mapped to the LIDAR's coordinates. Currently, there are only forward facing cameras. Therefore, the cameras are only mapped into the front LIDAR's coordinate space. However, the process is the same to integrate rear-facing cameras when required. For this alignment, a calibration device is moved through approximately 60 machine vision image positions, over a 5–18m range. These calibration board positions must be scattered throughout the complete image space. Placing the calibration board throughout the entire image, over an extended range, ensures the rotation values are more accurately determined.

Once the calibration board positions are found with the LIDAR scanners (see above), the same boards must be located within the machine vision system. Estimated rotation and translation values are used to turn the LIDAR detected board positions into approximate machine vision (i_t, j_t) space locations using (130). The board's position is further refined by identifying its unique visual markings (Figure 7-53).

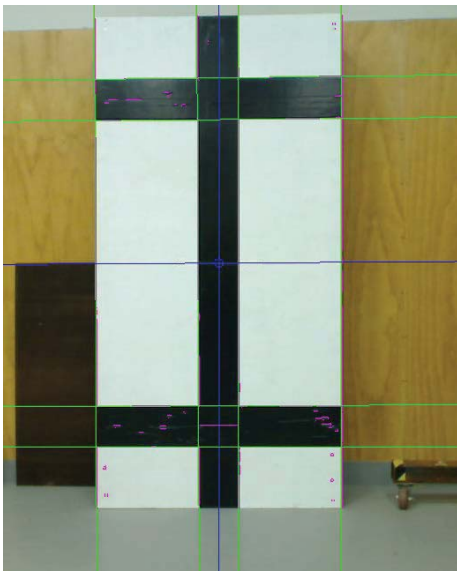


Figure 7-53: Machine Vision analysed image of the calibration board. Defined boards centre (blue lines and circle).

Firstly, a template is used to refine its position. The template is scaled according to the LIDAR determined distance and detects the white parts of the board and the vertical black stripe. Once this approximate position is known, a Sobel operator (Sobel, 1970) is used to find all the visible edges. The Sobel operator uses a set threshold to identify the edges of interest. From the identified edges, a Hough Transform (Duda and Hart, 1972) finds the straight lines that make up the calibration board's pattern and edges (green lines, Figure 7-53).

The resulting Hough space is searched to find the edge lines. There are two sets of lines, vertical and horizontal. The 'vertical' lines are parallel and approximately vertical, and the 'horizontal' lines are parallel and approximately horizontal. Therefore, only two small angle ranges are searched over in Hough Space. Within the search range, the identified maxima must have the same Hough space angle. This shows that the lines are parallel.

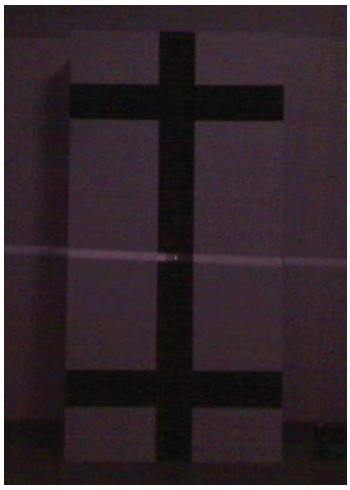


Figure 7-54: Infrared image of the LIDAR laser line on calibration board.

Once found, the calibration board's centre is located. This is done by averaging the sets of lines (result - blue lines, Figure 7-53) where the board's centre is at the intercept of the two averaged lines (blue circle, Figure 7-53). From this, the board's *(i, j)* image space centre is recorded for both stereo cameras. These *(i, j)* image positions are used to find the calibration board's three-dimensional position in camera space, $c=[x,y,z]^T$, using stereo depth perception from equation (13).

Now the board's image *(i, j)* and LIDAR positions are known. Unfortunately, the LIDAR is only a two-dimensional scanner. Therefore, rotation and tilt rotations (see Figure 7-50) cannot be found. To find these rotations the height at which the LIDAR's beam intercepts the calibration board is required. The right hand stereo camera is modified so it can operate in infrared (IR) or visible light spectrums. This is achieved by removing the IR blocking filter in the standard camera. When a normal image is required a polarising and IR blocking filters are fitted over the camera lens. To take IR images an IR band pass filter replaces the polarising filter. Figure 7-54 shows the resulting IR image of the laser line on the calibration board. Therefore, for every visible spectrum image there is a corresponding IR image. Next, the position of the laser beam is required to allow complete mapping between the coordinate spaces. The width and position of the board is known from the visible spectrum image analysis (Figure 7-53). Therefore, the search area for the laser line is known, as it is the same position as presented in the standard image. A Sobel operator (Sobel, 1970) and Hough transform (Duda and Hart, 1972) are used to find the top and bottom edges of the laser line. The height of the LIDAR beam is used to find the vertical offset between the camera and LIDAR spaces for every calibration board image. This allows the rotations and Z_C offset to be more accurately found. Other LIDAR/machine vision calibration techniques (Zhang and Pless, 2004) use multi-plane LIDAR scanners or specialist IR cameras. By modifying a standard camera these expensive sensors are not needed.

A simplex regression determines the offsets between the two geometric spaces using the camera and LIDAR calibration board positions. This is done by using the LIDAR board positions as the reference geometry. By using initial estimated variables of the rotation, $R_{L\omega}$ and translation, $T_{L\omega}$ between the two coordinate systems, the camera located calibration board positions are transformed into LIDAR coordinates. From this, an error between the LIDAR and camera translated position is calculated. This is done for the entire data set to work out a total error. The simplex regression iteratively alters the three-space rotation, $R_{L\omega}$ and translation, $T_{L\omega}$ values to minimise the error across the data set and produce the unknown variables.

With the relationship between the coordinate spaces (ATP, LIDAR1, LIDAR2 and machine vision) defined, any detected point found in one sensor can be mapped into any of the other sensors (see section 7.5.4). This delivers complete cooperative operation and information sharing between the sensing elements. Ultimately, this delivers the potential for higher-level orchard interpretation by combining the strengths of each sensor.

Figure 7-55: Photo of an orchard row

7.6 Greater Orchard Navigation

Navigation under and around the canopy area is outlined above. However, the ATP must also be capable of navigating the greater orchard area. This is navigation between canopy areas or to loading and unloading areas, but still on the orchard. An on-board GPS is referenced to achieve this navigation. This unit provides a rough absolute ATP position (±3m, within the world). The GPS position is referenced against the supplied digital orchard maps (from high accuracy GPS) to get a position relative to the orchard. These supplied maps outline canopy areas, hedgerows and other environmental information like on-orchard structures. Knowing the position of these and the ATP's current position items allows navigation decisions to be made.

Navigational nodes (red dots, Figure 7-56) are positioned on the orchard maps as waypoints for greater orchard navigation. Each node resembles the root of a linked tree, with children node(s) indicating allowable traversal paths. The node linkages are linear paths that the ATP can follow from GPS coordinates. For example, if the ATP was at node *F* and needed to get to *A* (referencing Figure 7-56), a lookup table is referenced. Node *F* has three child nodes *E*, *G* and *I* and each of those has their associated child nodes. The linked tree is traversed where the shortest path is selected as the navigational path to *A*. This is a form of Dijkstra's algorithm (Dijkstra, 1959). To determine this path the distance between each node is required. This is calculated and stored as part of the lookup table from their coordinates. In this scenario the algorithm determines traversals *F* – *E*, *E* – *D* and *D* – *A*.

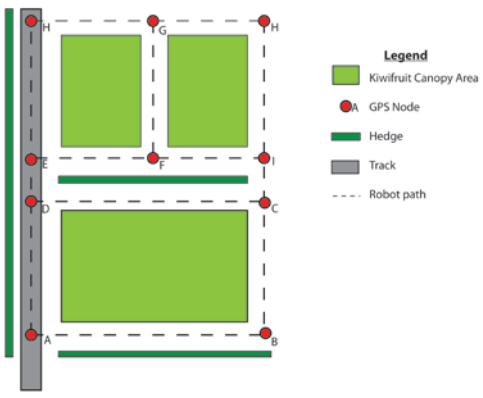


Figure 7-56 Schematic layout for GPS navigation

The ATP's current GPS position is used to drive between nodes. From the current position a heading is projected to the node using the flux-gate compass module. This path is followed to navigate to the node (as in section 7.5.1). As there are significant GPS positioning errors, the on-board LIDAR scanners are used to avoid obstacles and refine the ATP's position relative to near objects.

7.7 Testing and Results

Autonomous Transportation Platform testing and performance results are presented in four categories:

1. ATP hydraulic drive
2. LIDAR/machine vision orchard interpretation
3. LIDAR/machine vision alignment
4. LIDAR storage bin localisation

The outbreak of Psa-V (see introductory comments) affected testing of the drive and navigation systems. Although there has been sufficient testing to prove the performance of each independent system in the ATP, integrated systems testing has not been performed. The risk of spreading the Psa-V bacteria was too great for any integrated on-orchard testing of the prototype AKH because the prototype systems cannot be cleaned to meet the industries decontamination standards.

7.7.1 ATP Hydraulic Drive

Accurate and smooth operation where the primary specifications for the hydraulic drive system. A reliable turn radius within 50mm was required. This is because the system rapidly updates its drive path; therefore, some errors are allowable as they will be compensated for with subsequent sensor analysis. The ATP's manoeuvres are made up of a series of arcs with specific radii. Therefore, the ATP's manoeuvre accuracy is assessed by how accurately it drives in set radius circles. This also tests the control system's ability to set and maintain the front steering wheels' direction. Ten different diameter circles were driven five times each. Circle diameters ranged from two to fifteen metres. Testing was performed in a paved car park where the surface was relatively flat and the ATP's coordinate origin could be marked. The ATP was instructed to maintain a set radius and drive through 180˚. When this point was reached it stopped and its origin was marked. This allowed the diameter to be measured. Once the diameter was measured the ATP was instructed to complete the circle to ensure it returned to the same start position. An average absolute accuracy error of 52mm was achieved over the test range. This average had a standard deviation of 34mm. These diameters were measured with a measuring tape. On every manoeuvre the ATP appeared to return to its starting position (within ±20mm). This was

visually verified using a ruler. The drive system was tuned to give smooth acceleration from standstill and to limit oscillation of the steering wheels. This was achieved and was a subjective visual assessment. However, this tuning parameter slowed the systems response time.

In addition, the ATP is required to travel at 1.4 m/s on the flat and climb a 25° incline. A top speed just under 1 m/s was achieved with the current hydraulic pump. It was initially intended that hydraulic fluid would not be sent to the front wheels during high speed operation. This would increase the available fluid to the rear wheels allowing for faster drive. However, the front drive motors do not allow freewheeling. Therefore, an upgraded pump is required to meet this speed specification. This specified incline was considered the steepest operating slope. Although this was not performed on-orchard the system is designed to deliver sufficient torque to meet this criteria. Furthermore, the ATP is capable of driving onto the transportation trailer that has this required incline.

Testing in a synthesised environment was considered. However, testing the system's ability to interpret real orchard environments was considered more important, since accurate hydraulic drive was achieved. This was due to the amount of environmental variability that can be experienced and not replicated in a test environment. The secondary reason for not setting up a test environment, was the amount of space required and the need to setup infrastructure. Neither of these was readily available.

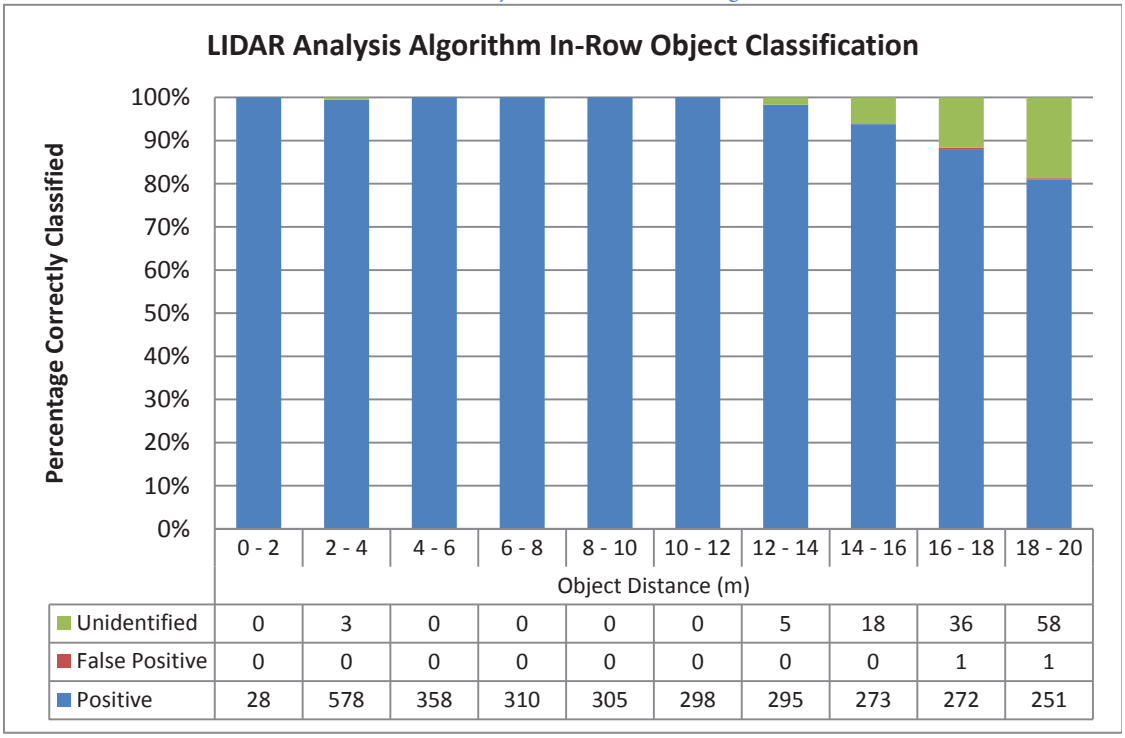
The LIDAR system was not tested for emergency stop conditions. This will be conducted during future testing.

7.7.2 LIDAR Orchard Navigation

For other orchard applications LIDAR-based navigation testing has focused on lateral and direction error from a defined path (Barawid, et al., 2007, Ryo, et al., 2004, Subramanian and Burks, 2009, Subramanian, et al., 2006). In these applications, testing was performed in either ideal or in synthesised orchard environments. The ideal orchard sections were selected where there were no gaps or significant variation between trees (Barawid, et al., 2007). This however, removes the significant variation that any system must handle. Subramanian, et al. (2006) constructed their synthesised environment with hay bales to form solid walls. Therefore, this testing methodology does not provide an accurate representation of their system's actual perfor-

mance. Testing with real orchard data is important to assess actual on-orchard ability. This orchard data must represent all on-orchard operating conditions and variation. This is important as interpreting environments where real-world variations are not present is significantly easier. Furthermore, using an ideal or synthesised environment allows an ideal drive path to be plot-

Table 4: Results from in-row object classification using LIDAR information



	0 - 2	2 - 4	4 - 6	6 - 8	8 - 10	10 - 12	12 - 14	14 - 16	16 - 18	18 - 20
Unidentified	0	3	0	0	0	0	5	18	36	58
False Positive	0	0	0	0	0	0	0	0	1	1
Positive	28	578	358	310	305	298	295	273	272	251

ted. This allows the navigation systems ability to be measured from its offset and heading errors with respect to the ideal line. In an actual orchard environment, this ideal path is difficult to determine due to infrastructure and plant placement variation. This means that the testing would not reflect on the system's ability to interpret the orchard environment or determine the correct drive path. However, the LIDAR scanners are proven accurate and reliable detection devices. Therefore, this testing focuses on the analysis algorithms ability to interpret and classify the detected orchard objects used to determine the drive path. A drive path is plotted from this analysis. The interpreted drive paths suitability was manually verified by analysing its placement relative to the identified row objects and data points.

For testing, orchard rows were scanned with a single forward facing LIDAR unit. This LIDAR unit was mounted to a quad bike. A quad bike was used as it could be cleaned to meet the industries decontamination standards and still gather moving scanner data. The LIDAR scanner was mounted to replicate its position on the ATP. Although the quad bike provided a suitable platform to gather row scans, it has a shorter wheelbase and track (quad bike – 1260 x 890, ATP

– 2450 x 1950) and soft suspension. The shorter wheelbase and track accentuates any pitching and rolling from in-row ground undulations, impressions etc. The ATV's operator also introduces pitching and rolling by leaning to avoid overhead obstacles. This introduces additional variation in the data set that would not be present with the ATP. The additional pitching and rolling can cause the scanner to detect the ground or canopy and miss the posts and trunks of the orchard row. This can affect row analysis. In addition, to ensure a range of driving scenarios were captured, the ATV was purposefully swerved during row driving. The swerving is shown in the analysis/testing video (referenced in section 9.1.3). This was done to ensure the analysis system could still interpret the row even if large drive errors occurred. Therefore, if the analysis system achieves suitable performance on the ATV with a single LIDAR (only half the row information) with accentuated variations and introduced drive errors, it will perform better onboard the ATP.

For this analysis, a data set of LIDAR scans was captured throughout a one-hectare canopy area. From this data set, two hundred scans were randomly selected to test the navigation system. 3090 in-row objects of interest (post and trunks) were captured in the scanned information. From these, the analysis algorithms correctly classified 96.1% of all objects over the full scanning range (20m). The remaining results comprised of 0.06% false positive and 3.9% unidentified. However, most of the errors were in the 12 to 20 metre range. Correct classification increased to 99.84% when assessing the results under 12m. There was an average of 15.3 in-row objects per scan over the 20m range. At the achieved identification rate there will be 14.7 in-row objects detected per scan. This has a standard deviation of 0.7 objects per scan. Therefore, when both LIDAR scanners are fitted it is likely that 28 in-row navigation objects will be detected to determine the drive path.

Figure 7-57: Two consecutive LIDAR scans showing occlusion hindering correct classification (left). The LIDAR scanner is positioned in the lower left hand edge of each scan image.

As anticipated, the results show that the probability of detection decreases with distance (Table 4). Within this, most of the detection errors occurred beyond twelve-metres. This is because there is a bigger gap between consecutive LIDAR beams. Therefore, there are fewer captured

data points for an object of the same size at a greater distance. This makes it harder to identify the same object when it is further away. However, positive identification also drops over the two to four metre range. This is due to the partial occlusion of an object (see Figure 7-57). In this case, incorrect classification occurs as there are insufficient data points representing the occluded object. This makes valid identification improbable. An example of this is shown in Figure 7-57 where the left hand scan has several unidentified data points (purple) behind a classified cluster (green circle denotes an identified cluster of points). These occlusions do not affect operation as the closer occluding object still contributes to determining the driveline. Furthermore, they do not occur very often and only happen over a short distance (approximately 50-200mm) until more of the object is detected.

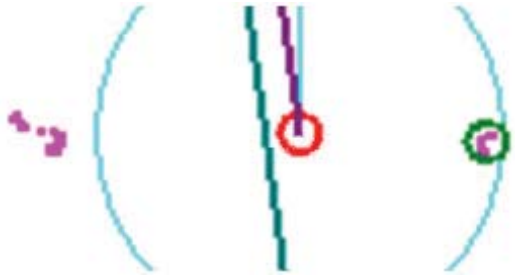


Figure 7-58: LIDAR navigation misclassified data points (purple points, left) correlating to a neighbouring post and vine.

Another notable result is the low object count in the zero to two-metre range. This is primarily due to the row width. The row objects only enter this range when the ATV was purposely deviated off course. This was done to assess the algorithm's ability to track the row objects for any in-row pose possibility (see video referenced in section 9.1.3).

7.7.3 LIDAR Storage Bin Localisation

To test locating of bins, two hundred and sixty LIDAR scans were captured. The LIDAR scanner was mounted to an ATV to capture the data (see comments in 7.7). These scans captured a bin throughout the sensor's scanning range, where the bin's orientation to the scanner was also altered. This was repeated over four different orchard environments to ensure there was environmental variability within the test. One hundred and eighty eight of the captured scans had a storage bin present, whilst the remaining scans only captured the orchard headland environment. This tested the identification algorithms for false positives.

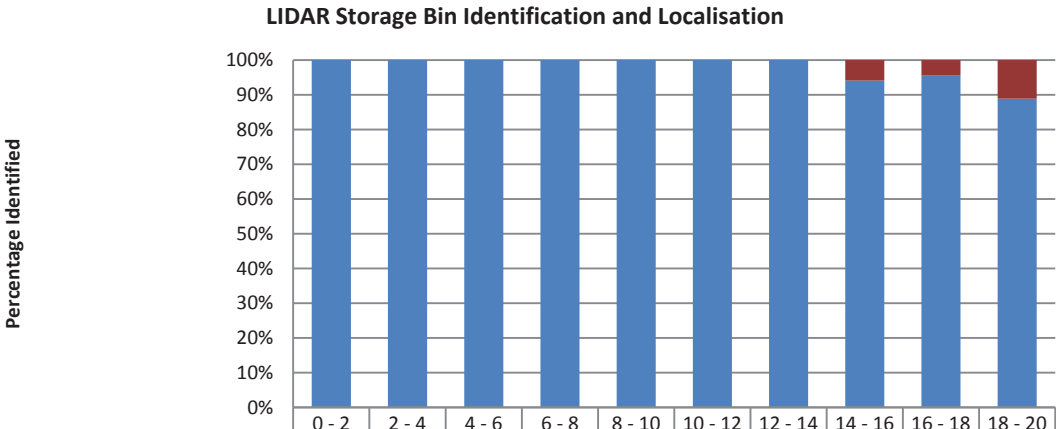
The analysis method correctly identified 98% of the bins. This was achieved with no false positive identifications. Eliminating false positive identifications ensures safe and efficient autonomous operation. Although a high correct classification rate was achieved, it was noted that a small vertical rotation of 1.5° was sufficient to cause the LIDARs scanning plane to miss the bin

at the sensor's maximum range. Driving over uneven terrain and general terrain variation introduces this rotation. As the captured scans were taken using an ATV its shorter wheelbase and track accentuate any terrain undulation (as discussed in previous the previous test). The ATP's longer wheelbase will help reduce this issue, however, terrain variation may still hinder storage bin detection.

In the final configuration, a rear facing LIDAR unit will be used for bin finding. To assist with detection the rear scanner's height will be set to half the bins height. This will help keep the scanning plane on target. Although scanning plane variation occurred, it was observed that the scanning plane reliably intercepted the bin out to ten metres during testing. The effect of undulation on bin finding will not significantly affect operation, because:

1. The headland areas are narrow (4-5m). Therefore, it is unlikely that the ATP will drive past a bin without detecting it.
2. As the LIDAR is scanning at 25Hz the bin is often momentarily detected during the undulations. This position is stored and the ATP will begin driving towards the recorded position. The bin's position is updated if it comes into the scanners plane while driving to that detected position.
3. When the bins are dropped off in the headland area, their position will be GPS marked. Therefore, the ATP will know an approximate bin position (±3m) and can drive towards it until the LIDAR detects its actual position. The LIDAR determined position will be used to dock the ATP with the bin.

Table 5: LIDAR bin localisation results



	0 - 2	2 - 4	4 - 6	6 - 8	8 - 10	10 - 12	12 - 14	14 - 16	16 - 18	18 - 20
False Positive Classification	0	0	0	0	0	0	0	0	0	0
Incorrect Classification	0	0	0	0	0	0	0	2	1	2
Correct Classification	18	20	38	30	18	28	32	32	22	16

7.7.4 LIDAR/Machine Vision Alignment

Accurate and consistent alignment between LIDAR and machine vision is necessary for machine vision identification of LIDAR detected objects. The alignment methodology (section 7.5.4) and subsequent calibration process determined alignment accuracy, i.e. how well a LIDAR detected point is mapped into (*i, j*) image space. The alignment process captures both LIDAR and image information over the entire image range. This provides a data set to determine the alignment between the coordinate spaces and measure alignment accuracy.

An average absolute alignment error of 1.0 and 0.77 pixels in image space was achieved. The errors had standard deviations of 0.86 (*i*) and 0.84 (*j*) from the 60 captured images/scans. This represents an alignment of a LIDAR detected point into camera space to ±1 pixel in either *i* or *j* image directions and is sufficient for the desired operation.

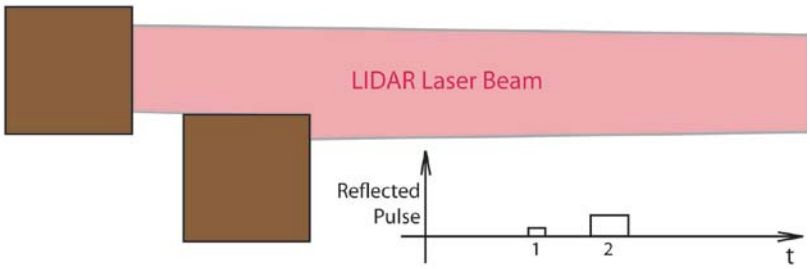


Figure 7-59: LIDAR second-pulse technology

For testing and on-orchard operation the LIDAR unit was configured to return the first reflected pulse. This is shown in Figure 7-59. While scanning, a single laser beam can reflect off several objects. This returns multiple pulses to the scanner. In the 'first reflected pulse' configuration the first detected object is shown. Figure 7-60 shows the LIDAR data overlaid onto a concurrently captured row image. Within the shown data points, it can be seen where the centre of the data point does not correspond to the row object. This is because the scanner is set to 'first reflected pulse'. Therefore, it returns any object that the diverging laser beam intercepts. If the scanner was set to 'second reflected pulse' this would not occur. However, the current setup ensures the whole object's width is detected along with narrow objects (like wires).

7.8 Conclusion and Future Work

The development process identified the limitations of a standalone, machine-vision-based navigation system (section 7.4). These were overcome by integrating LIDAR scanners to detect on-orchard objects. The combination of LIDAR and machine vision allows future development to integrate machine vision object identification. This will complement the reliability and accuracy of LIDAR to gain higher-level orchard interpretation.

LIDAR has also allowed a shift from dead-reckoning manoeuvres. This will increase navigation and drive system efficiency and safety with real-time analysis and fast update speeds (25Hz). The high refresh rate constantly corrects for any inaccuracies from environmental factors like ruts, holes and inaccuracies within the steering system. This adds further safety and reliability when compare to the standalone machine vision navigation system. In conclusion, the cost of including additional LIDAR sensors is fully commercially justified.

Desired future development aims to integrate the full potential of machine vision to complement LIDAR. The alignment methodology to map LIDAR found positions into machine vision image positions is sufficiently accurate (±1 pixel in image space) to support object identification. This will allow machine visions 'intelligence' characteristic to be fully supported. Future machine vision development will focus on on-orchard object identification, including identification of:

1. Hanging canopy vines which do not pose navigational concern to the ATP
2. In-row poles and trunks to ensure that the in-row objects used for navigation are not false positive identifications
3. Random on-orchard objects to determine their navigational 'threat'. These objects include posts outside of the rows or hedge branches.

In conclusion, the identification methods and strategies for navigation and bin finding have been demonstrated. The achieved interpretation ability supports commercial operation with row object and bin classification at 96% and 98% respectively. This will allow the ATP to operate autonomously on-orchard once upgraded to meet industry decontamination standards. This will ensure mitigation of bacterial transmission risks. Although the identification systems have not run on the ATP on-orchard, the ATV based testing shows reliable detection and operation. Furthermore, the system is likely to perform better when integrated onto the ATP as

there will be less variability in the scanned information gathered. This is because the ATP provides a more stable platform. Future development will fully integrate and test these systems in a variety of orchard environments to test and prove operational viability. The results achieved meet all required specifications and demonstrates advancement to orchard navigation. This advancement comes from the total system's ability to successfully operate even when presented with orchard variation. This is required for practical on-orchard and commercial operation, but has not been achieved with previous navigation systems.

Figure 7-60: LIDAR determined row information superimposed onto a corresponding in-row image. The red circles represent the laser beams size and incident position at detection distance, where the green circles are the centre of the LIDAR data cluster, used to determine the driveline (superimposed in darker green)

Figure 7-61: The AKH on-orchard

8. AKH: The Overall System

The separate functional AKH sections have been presented. This section assesses the overall system considering the product's development and commercial operation. Comments on the lessons learnt and future modifications are included for their relevance to the future product development cycle. These are included for their significance to the overall development of discrete crop harvest robotics.

The systems are not fully integrated due to the risk of Psa-V bacterial spread. Access to orchards limited how much of the complete system could be tested. However, the main interactions have been tested. The main interactions occur where one section is dependent on another to operate. For example, the harvesting arms require the machine vision system to locate and map the fruit into harvesting arm coordinates. The other section interactions that are not fully tested operate independently. Therefore, their interaction is insignificant to overall operation. This allows conclusions to be made on the performance of the complete system despite limited testing.

As outlined below, the developed AKH did not meet all required performance specifications. Considering this, modifications and future development are presented to evolve the prototype AKH into a commercially viable, fully operational system.

8.1 Results

The AKH, as a whole, has extended the current state of the art and demonstrated capability across all of its functional areas. Although some areas require improvement, the ability has been demonstrated in actual orchard environments. These orchards were not tailored for the AKH, nor were ideal orchard environments selected to operate in. The orchards used for testing all had typical variability, complexity and challenges. Some operational aspects of the AKH did not meet the desired specification. Practical solutions and/or modifications are presented where specifications were not met. These changes will allow the system to meet, and in some instances exceed, the required performance specifications. This will allow the AKH to be developed into a commercially operational machine.

The only functionality not demonstrated on-orchard is the transitioning between tasks. This is mainly changing from harvesting to driving (or visa-versa). As these are only instructions from one section to another, this process does not affect core functionality.

The key performance criteria are presented in section 3. The AKH's performance is assessed against these specifications:

1. *Harvest at least 80% of the crop in conformance with industry fruit handling standards.*

During on-orchard harvest trials, it was observed that if a fruit was located by the machine vision system, the robotic arms could harvest that fruit. This is because the hand enters the canopy from below, which is the approximate vantage point as the cameras. Therefore, if there is an unobstructed line of sight to identify the fruit, there is sufficient space below that fruit to harvest it. Identification rates greater than 80% were achieved (section 5.2.1). However, modification to the harvesting arms (section 6.2) is required to ensure this translates into the required harvest rate. The arms incurred positioning errors resulting from forces during of fruit extraction. This limited their ability to harvest sequences of fruit. However, this will be solved by integrating industrial position encoders. Furthermore, to increase identification rates the addition of high intensity lighting is proposed. This will help limit image illumination variability and increase fruit identification rates.

2. Harvest rate: Four fruit per second (one fruit per second per arm).

A harvest rate of 1.3 fruit per second per arm was achieved during laboratory testing. Although this exceeds system requirements it was not achieved on-orchard. The on-orchard environment was more physically rigorous, causing arm positioning errors (see section 6.2). As a result, harvesting rates were slowed to minimise positioning errors. An average of 0.7 fruit per second per arm was achieved with asynchronous operation. The integration of industrial encoders to the arms will allow increased operating speed to meet the required harvesting rate. The encoders will allow errors to be corrected for without affecting operation.

The harvesting arms must operate cooperatively. Although asynchronous harvesting was only performed with two harvesting arms, the inter-arm interactions and harvesting cycle speeds meet the necessary performance criteria. The presented methodology demonstrated efficient and collision-free operation. As this method was created and implemented to operate with four arms this performance will replicate across all four arms once they are operational.

3. Handle full and empty storage bins.

The outbreak of Psa-V restricted on-orchard access for testing storage bin handling. Although the complete system was not tested, the key aspect of bin handling is reliable location of bins. Bin localisation was successfully achieved to commercially acceptable levels (section 7.7.3). Directing the ATP to drive and engage with the located bin is not presented due to the above mentioned restrictions. Although this was not directly tested, the required system attributes to achieve bin handling have been achieved. These are:

1. Reliable and accurate storage bin localisation
2. Accurate ATP manoeuvrability
3. Fast sensory refresh rates to quickly correct for any driving errors

Future work will assess and validate complete commercial system performance.

4. Gentle, even and complete storage bin filling in conformance with industry fruit handling standards.

The bin filling arm has achieved gentle and even bin filling in conformance with industry fruit handling standards. In-fact, this system handled fruit more gently than the comparative human process of releasing bagged fruit into the bin. The distance sensors mounted to the arm allow a safe fruit drop height (less than 80mm) to be maintained, ensuring fruit quality. Although this arm met most of the specification, it could not completely fill the bin. The internal pocketed conveyor limited axis motion, preventing complete bin filling. This issue will be resolved by altering the arms mounting position and articulation geometry. Most importantly, the bin filling system achieved the presented ability within a confined operational space. The space restriction (mainly in height) was the most significant limitation to a bin filling system's development. By implementing the above-mentioned modifications, it is expected that this system will operate at commercially acceptable levels.

5. Operate twenty hours per day, seven days per week.

Operational performance was comparable day or night. In-fact, night time operation is beneficial as sunlight illumination variations are not present. This simplifies the fruit identification process. To assist daytime fruit identification the integration of high intensity artificial illumination will minimise natural lighting effects. This will conceivably be solved with a large high-

powered LED array producing daylight coloured light (5000k colour temperature). The selected colour temperature will help maintain a constant canopy appearance day or night.

For on-orchard navigation LIDAR sensing is consistent and independent of lighting conditions. This provides day and night consistent operation. Future development will employ a machine vision system fully integrated with the current navigation system to provide further 'intelligence'. To assist this, high intensity lighting will also be added. This 'intelligence' will allow the vision system to identify the LIDAR detected objects and determine if they pose a threat to navigation.

6. Autonomous on-orchard navigation.

The Psa-V outbreak limited comprehensive on-orchard testing. Although limited, a reliable and accurate in-row object identification and path planning system is presented. This identification and path planning system forms the primary aspect of the ATP's navigation system. To complete navigation these paths are mapped into driving commands. The drive system has been tested independently of the navigation interpretation system. This hydraulic drive system was sufficiently accurate to reliably perform the required manoeuvres. Therefore, having an accurate and reliable navigation system and an accurate drive system provides suitable evidence that combined system will operate effectively on-orchard. Future work will assess and refine the complete navigation and drive systems for commercial application once orchards can accessed without risking orchard contamination.

7. Unit build cost less than $NZ130,000+GST (on a projected production of 10 units at a time).

The projected unit production costs is estimated at $NZ122,800+GST per unit (based on the current prototype and producing ten concurrent units). Altering the system to meet required performance criteria is unlikely to increase the production cost beyond the specified build cost. This is because some of the modifications will also result in decreased production costs. Prototype development has always considered watertight enclosures to allow for heavy rain during operation. Although these enclosures are not all installed their cost has been included in the budgets. Therefore, the complete AKH will be sufficiently protected to allow decontamination clean down. This will reduce the risks of spreading Psa-V between orchards. This cost is for

building each machine and only includes materials and labour. Any other costs, like the cost of development, commercialisation or marketing, are not included in this figure.

Overall, the integration of each section was critical to operational success. Furthermore, this integration had to meet the physical rigours of on-orchard operation. The real-world orchard environment posed significant additional complexity over the laboratory environment. Some added complexity was expected, however, the level of variability was more significant than had been predicted. This affected development by forcing a redevelopment loop in the product development cycle to account for the environmental variability. Every developed section required alteration to account for presented variability. Other developments have operated within synthesised or pre-prepared environments, or have been able to limit the most significant variability. Within the environment, the most significant variability is from sunlight illumination affecting fruit identification. To overcome this issue most other harvesting attempts have shrouded the targeted plant. This is not possible with the kiwifruit growing system. Therefore, considering the AKH's current operational ability was achieved in a typical orchard environment, its performance has demonstrated an evolutionary advancement in autonomous harvest technology.

Reduction or management of all variability will be critical to the system's on-going success. However, the potential to reduce variability is limited as it is present within the on-orchard environment. Therefore, the system must be sufficiently flexible to meet any presented variability. This was considered during development and each operational section has been refined to meet, or be easily adapted to meet any new variability. Machine vision sensing was most affected by variability. This was evident in the transition from synthetic laboratory to on-orchard environment. Fruit and growing system variability are an operational reality. Therefore, the fruit identification algorithms were developed to manage presented variations. However, beyond the growing system, the most significant variation effect is from inconsistent sunlight illumination (see section 5.2). The effects of this variation will be minimised with the addition of a high intensity lighting system. The transition from laboratory to orchard environment taught valuable lessons. Primarily, not to underestimate the effect of environmental conditions. These were more diverse and changeable than initially expected. This realisation forced the adaption of the analysis algorithms and addition of a high intensity lighting system to manage variability. This lesson was applied when later developing the LIDAR based navigation system. When the analysis algorithms were developed, actual orchard rows were scanned to create the develop-

ment dataset. Several orchards were scanned to ensure that as much variability as possible was represented in the dataset. Therefore, very few modifications were required when the system was tested on-orchard. Ultimately, this saved development time and provided real-world results.

On-orchard environment variability affected more than machine sensing. Although successful in the laboratory, the fruit harvesting arms did not meet their performance specifications on-orchard. The arms, including the hand, required significant alterations to achieve suitable on-orchard operation. Fruit positioning, attachment strength and size variations all detrimentally affected harvest operations (discussed in section 6.2.8). Four design cycles were required to achieve a hand suitable for on-orchard operation to overcome fruit and growing system variation. The hand design alterations caused arm position inaccuracies (see section 6.2.8). Although this affected harvesting, the issue will be resolved by integrating rotational encoders to each axis. Therefore, the lessons emerging from robotic arm development are similar to those for sensing. Although laboratory testing is helpful to test initial conceptual ideas and implementation, it is not suitable to determine orchard performance on the basis of laboratory performance. Development must quickly move from laboratory to orchard to ensure a complete understanding of the environment is gained and the results implemented into the system. This will help ensure that the variability affecting performance is understood and managed. A broad understanding of the environment is important.

Any redevelopment loops within the product development cycle will evolve the system to manage variability as cause and effect are better understood. For example, the stereo camera setup will be reassessed with future development. The fruit identification system currently has one pair of stereo cameras for each arm. Fruit occlusion was a significant limitation to identification. However, since an accurate method to map the camera and arm coordinate spaces has been developed and proven, it is feasible to use more cameras. More fruit could be observed by mounting these additional cameras in different positions and orientations. These multiple cameras are able to be configured into multiple stereo pairs (see Figure 8-1) using the stereo camera calibration system (section 5.2.3.1). In this figure the current setup is represented by camera 1 and 2 (*C1, C2*). In this setup, fruit could be found in *C2* but might be hidden from *C1.* The fruit

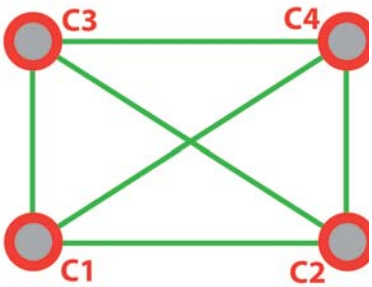


Figure 8-1: Plan view of multi-camera setup. Green lines represent possible stereo camera pairing, where C1 – C4 are cameras

needs to be visible in the second camera to find its position. However, in the proposed system the hidden fruit might be visible in *C4*. Therefore, the fruit can now be located. Likewise, fruit might be hidden from *C1* and *C2*, but can be viewed by *C3* and *C4*. The new viewing angles/positions will allow more fruit to be located. However, the number and position of the cameras will need to be optimised. In addition, the suitability and benefits of this system will need to be evaluated against cost before possible commercial implementation. However, it is envisaged that this will help increase system efficiency and the percentage of crop harvested.

8.2 Conclusions and Future Work

Development of a fully autonomous kiwifruit harvester is a significant step beyond other discrete crop harvesters (see section 2). In those projects, the added complexities of navigation, bin handling and fruit handling were not all included. The integration of these abilities during product development is required for a commercially viable system, yet they are not always considered. Including this additional functionality requires additional complexities to be managed. This complexity is beyond dealing with environmental variability. The systems have to work synergistically together. This is more than in a control sense. For example, the mechanical structures (if required) cannot interfere with the other sections or sensors cannot be blocked by any structures. This increases design complexity, however, it is critical to overall commercial operation. Furthermore, all system testing has been performed within unaltered and 'typical' orchard environments. This ensures the expected orchard variability is present during testing giving an accurate assessment of system performance. This important testing environment is often ignored in other parties testing.

Fruit identification and extraction are at the core of this project. These functional elements have demonstrated significant advancement over previous developments (see section 2). This system achieved better identification rates, increased harvest speeds and low cost robotics when compared to the other reviewed developments. The achieved performance is set to increase further when the presented modifications are implemented. These advancements result from implementing an alternative camera setup to locate fruit and from using customised robotics. Advanced control strategies will increase the custom robotics operational ability, including their ability to operate cooperatively, to further increasing efficiency. The use of and strategies to implement multiple arms for discrete fruit harvesting has only been used by Muscato, et al. (2005). There, only two arms were used. These were on a fixed oblique plane, limiting their interaction.

The four arm's configuration presented in this document, has faster harvesting cycles and co-operative operation within overlapping working envelopes. This provides significantly increased harvest rates and future task potential.

Consistently gentle fruit handling is required throughout operation. In its current prototype configuration the AKH has demonstrated the required handling standards (see section 6.4). The AKH's fruit handling ability was consistently better than the human counterpart. In addition, unlike humans, robotic systems are not susceptible to financial incentives to harvest fruit rapidly, nor do they suffer fatigue or boredom. The human flaws present an increased risk to fruit quality over extended working periods. The AKH consequently supports and delivers consistent quality produce. This helps increase financial return to growers. It also provides a strategic point of difference for commercial uptake.

The ATP was required to gain complete autonomous functionality. The importance of an autonomous vehicle is overlooked in many other developments. Although the ATP's full navigational capabilities were not tested (because of the Psa-V outbreak), robust and reliable object detection and path-planning capability has been achieved. Most importantly, this was achieved over several varying orchard environments encompassing significant variability. Therefore, this result is likely to represent actual on-orchard performance.

Development of the ATP goes beyond harvesting. This platform will allow the automation of other on-orchard tasks. Ultimately, this increase its commercially viability by maximising its potential use beyond the relatively short harvest season. Furthermore, industry and commercial benefits are extended by applying robotic efficiency and operational benefits to other orchard tasks. These are likely to include pollination, pruning, crop counting and thinning. Having modular section (discussed below) allows these task-performing units to be retrofitted as required.

The technology and methodologies developed are not limited to kiwifruit. These systems have direct application within most horticultural sectors. Although the physical structures will likely be different for other crops, the methods presented can be adapted into specific configurations for other crops. The advantages of this technology are similar throughout the horticultural sector and industry is becoming increasingly aware of the benefits mechanisation offers. The AKH (or similar system for other crops) delivers a long-term sustainable commercial advantage. This is achieved by having a static, or even declining, pricing structure in comparison to increasingly

expensive and inconsistent human labour. In addition, it is becoming harder to attract people with the necessary skill sets into menial jobs. There is concern within the horticultural industry that the need to reduce labour costs will force the industry to use untrained labour. This could further affect production. Robotics could alleviate these industry issues by reducing high labour requirements, costs and skill shortages. Furthermore, robotics can add competitive advantage by guaranteeing quality, adding value and ensuring sustainable future productivity.

During development it was inevitable that systems would require modification. These modifications reinforced the importance of modular design. A centralised computer control system requires many cables to control and power each section. Considering long-term commercial operation, this design will slow production, servicing and maintenance of the units. Therefore, a modular approach will be implemented. This will allow individual sections to be manufactured and tested without requiring implementation into the complete system. Furthermore, if a section develops a fault during orchard operation, that section can be quickly replaced to allow continued operation. This approach removes the requirement for centralised computer controllers. These will be replaced with micro-controller based custom controllers for the robotics sections and small form factor computers for the visions systems. At production levels, individual controllers have lower cost than a complete PC with added input/output control cards and complex wiring. This will allow the arms to become relatively plug-and-play. All they will require is power and communications line to become operational. After attachment, an automatic calibration processes would configure the system if required. These modular controllers will be used for each arm (four harvesting arms and the bin filler) as well as for the ATP's hydraulic drive.

Necessary modifications have been identified to progresses system capability to meet the required specifications. Future development will initially focus on implementing these modifications as well as complete system testing. Primary focus will be:

1. Encoder position feedback for the robotic arms. This will ensure accurate on-orchard operation
2. High intensity lighting to reduce natural lighting variation for fruit identification
3. Navigation and storage bin handling - complete system testing and refinement
4. Storage bin filling arm geometric modification
5. Complete system interaction, testing and commercial validation

Following the discovery of the virulent Psa strain in the New Zealand kiwifruit industry, design considerations to mitigate risk of bacterial transfer between orchards are now critical. Therefore, before any further future testing is possible the systems must be protected from decontamination wash-down. The use of modular controllers will help ensure this is achieved. The AKH was developed to conform to ingress protection ratings level four to protect against water splash from any direction. This protection level was to be introduced after prototype development on all systems. However, it is determined that a minimum of IP5 (IP6 desirable) is necessary to protect all AKH systems from water jets. This will become the standard protection rating for future development. The increased protection will allow bactericide pressure washing before and after orchard use. Furthermore, enclosures and cowlings will be designed to prevent the AKH entrapping organic matter. The organic matter could potentially spread the Psa bacteria even during road transportation. Removing entrapment areas will also help speed up the clean down processes.

The presented development has achieved technology and methodology capable of performing the tasks required to robotically harvest kiwifruit. Furthermore, those components are integrated into a single system that will have the capability of complete autonomous operation at commercially viable rates. This means that the system has sufficient performance and payback to make its operation and manufacture commercially viable. Although some of its systems are not fully completed or require some modification there is suitable evidence that with those modifications, full commercial performance will be achieved. This will allow the technology to be adapted to other crops to benefit New Zealand's horticultural industry. Within this technology base, the ATP has the greatest potential. Development to better utilise this platform will expand commercial potential. Considering the delivered system, the AKH's development has achieved a significant advancement to horticultural automation. Most importantly, this has been achieved in real-world environments that have not been tailored to suit the system. This makes the total AKH and the integration of each functional area a significant step beyond other discrete crop systems.

In summary, the main results achieved were:

1. Fruit identification (83.6% of crop) with combined localisation and extraction accuracy of 3.6mm in three-dimensional space
2. More gentle fruit harvesting and handling than humans harvesting

3. Reliable object detection and path planning for navigation. Over the twenty metre scanning range 96% of the in-row objects were correctly classified to reliably determine the drive path
4. Reliable fruit storage bin identification and localisation (98% correct classification)
5. Commercially viable manufacture cost less than $130,000 per unit
6. Although full commercial operation was not achieved, modifications are identified to rectify the limitations.

Figure 8-2: The ATP chassis, during early development

9. Appendices

9.1 Operational and Analysis Videos

The attached disc contains operational and systems analysis video of the autonomous kiwifruit harvester and sensory systems in real on-orchard environments.

9.1.1 Autonomous Kiwifruit Harvester Operational Video

Autonomous_Kiwifruit_Harvester_Operation.mp4 shows the prototype AKH working on-orchard.

9.1.2 Created Here Video

Created Here_Shot.mov is a shortened version of the movie commissioned by Trade and Enterprise New Zealand, show casing New Zealand innovation. The AKH was selected to be part of this innovation campaign.

9.1.3 LIDAR Navigation Video

LIDAR_Navigation.mpg is the sequential compilation of on-orchard LIDAR scans and subsequent analysis for row drive path determination. Deviations in the drive path were purposefully performed during data capture to simulate an ATP pose error and the algorithms ability to accurately determine the correct navigational path, with the presented pose error.

9.1.4 LIDAR/Machine Vision Navigation Video

LIDAR_Navigation_Vision.m2v captures sequential row LIDAR and machine vision information. The LIDAR data points are transformed into machine vision space and graphically rep-

resented as circles to demonstrate alignment accuracy. LIDAR analysed scans are embedded for reference.

9.1.5 LIDAR Bin Localisation Video

LIDAR_Bin_Localisation.mpg shows the sequential LIDAR scans of a headland area with a storage bin present. The raw captured scans are first shown, followed by the analysed scans which show the identified storage bin and position it with respect to the scanner.

9.2 Fruit Handling Analysis

Fruit analysis data supplied from the independent monitoring and assessment company (Mount Pack and Cool Ltd) used to assess the system's ability to handle fruit gentle, when compared to human harvesting. Fruit were removed when damage was found (Table 9 below) and therefore the total number of fruit altered.

Table 6: Raw fruit analysis data provided by Mount Pack and Cool Ltd for evaluation

Monitored Fruit Damage									
DATE	DAMAGE TYPE	HUMAN PICKED BAG	HUMAN BAG TO BIN	HUMAN HAND PICKED		ROBOT BIN FILLER	ROBOT CONVEYOR	ROBOT HARVESTING HAND	ROBOT FRUIT TUBE
20/07/2010	SOFT	1	0	0		1	0	1	0
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	0	0	1		1	0	1	0
	UNEFFECTED	28	30	29		28	30	28	30
5/08/2010	SOFT	3	0	2		0	0	0	1
	SOFT SPOT	0	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	2	1	1
	UNEFFECTED	21	26	26		24	28	27	28
18/08/2010	SOFT	0	0	1		0	0	0	1
	SOFT SPOT	0	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	2	1	1
	UNEFFECTED	17	22	24		20	26	26	26
8/09/2010	SOFT	2	0	2		1	2	0	1
	SOFT SPOT	0	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	11	18	21		15	21	24	24
22/09/2010	SOFT	0	0	0		0	0	0	0
	SOFT SPOT	0	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	7	14	20		11	18	22	23
6/10./2010	SOFT	0	0	0		0	0	0	0
	SOFT SPOT	0	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	3	10	19		7	15	20	22

Table 7: Accumulated fruit analysis data

Accumulated Monitored Fruit Damage									
DATE	DAMAGE TYPE	HUMAN PICKED BAG	HUMAN BAG TO BIN	HUMAN HAND PICKED		ROBOT BIN FILLER	ROBOT CONVEYOR	ROBOT HARVESTING HAND	ROBOT FRUIT TUBE
20/07/2010	SOFT	1	0	0		1	0	1	0
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	0	0	1		1	0	1	0
	UNEFFECTED	28	30	29		28	30	28	30
5/08/2010	SOFT	4	0	2		1	0	1	1
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	2	1	1
	UNEFFECTED	21	26	27		25	28	28	28
18/08/2010	SOFT	4	0	3		1	0	1	2
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	2	1	1
	UNEFFECTED	21	26	26		25	28	28	27
8/09/2010	SOFT	6	0	5		2	2	1	3
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	19	26	24		24	25	27	26
22/09/2010	SOFT	6	0	5		2	2	1	3
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	19	26	24		24	25	27	26
6/10/2010	SOFT	6	0	5		2	2	1	3
	SOFT SPOT	1	0	0		0	0	0	0
	PHYSICAL	4	4	1		4	3	2	1
	UNEFFECTED	19	26	24		24	25	27	26

Table 8: Fruit Damage Analysis 20/07/2010

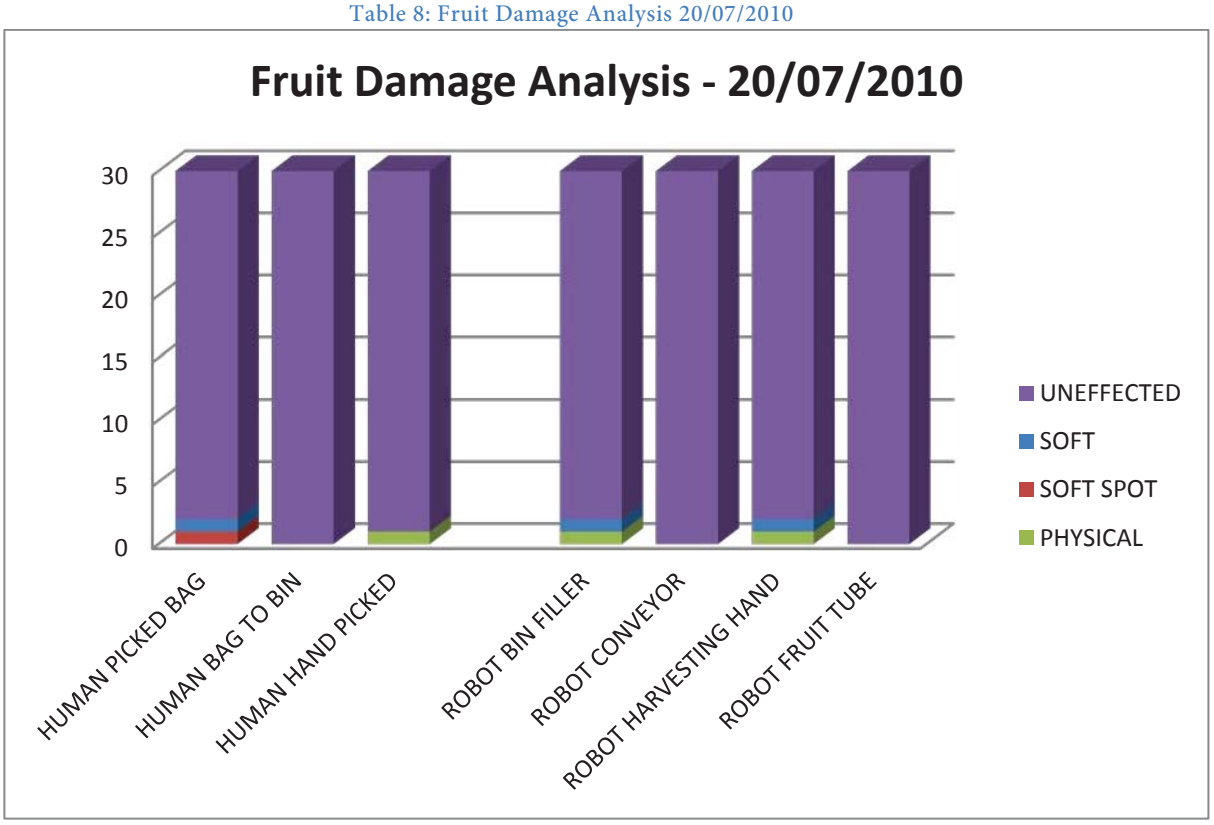


Table 9: Fruit Damage Analysis 05/08/2010

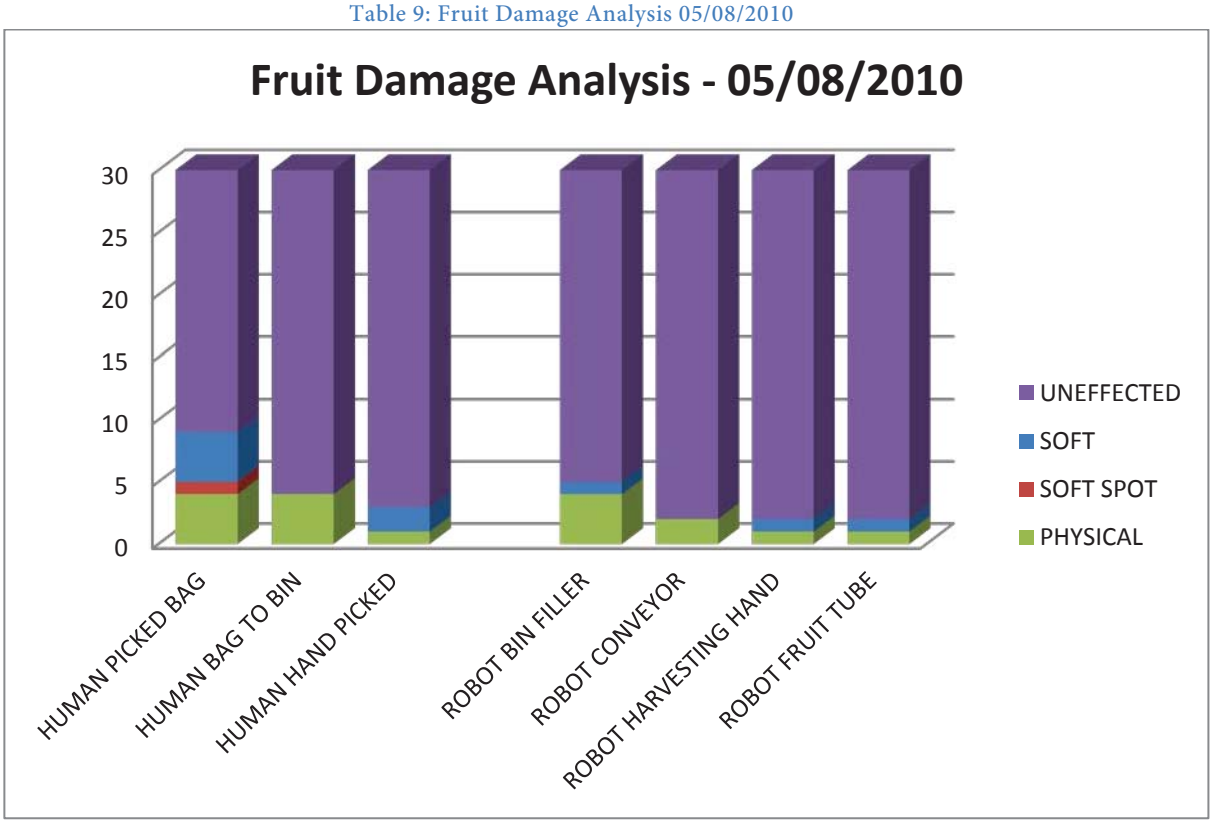


Table 10: Fruit Damage Analysis 18/08/2010

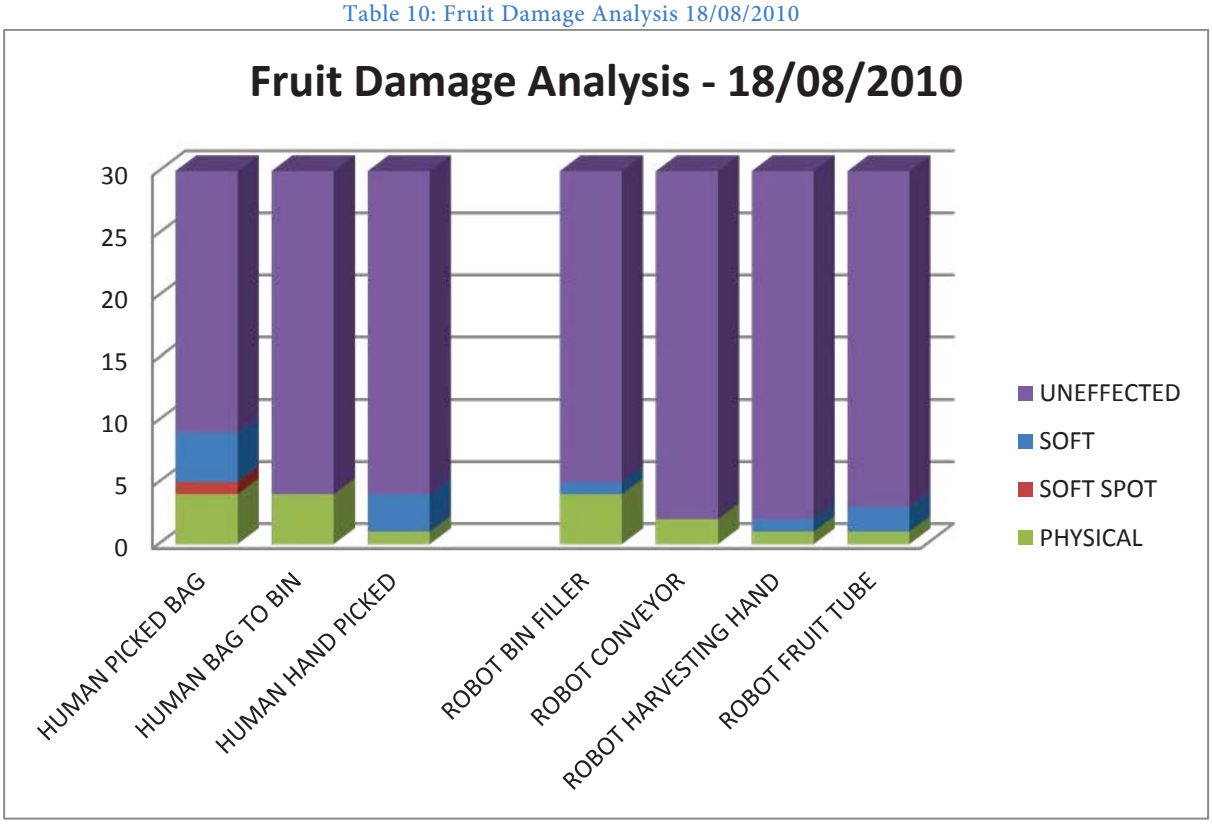


Table 11: Fruit Damage Analysis 08/09/2010

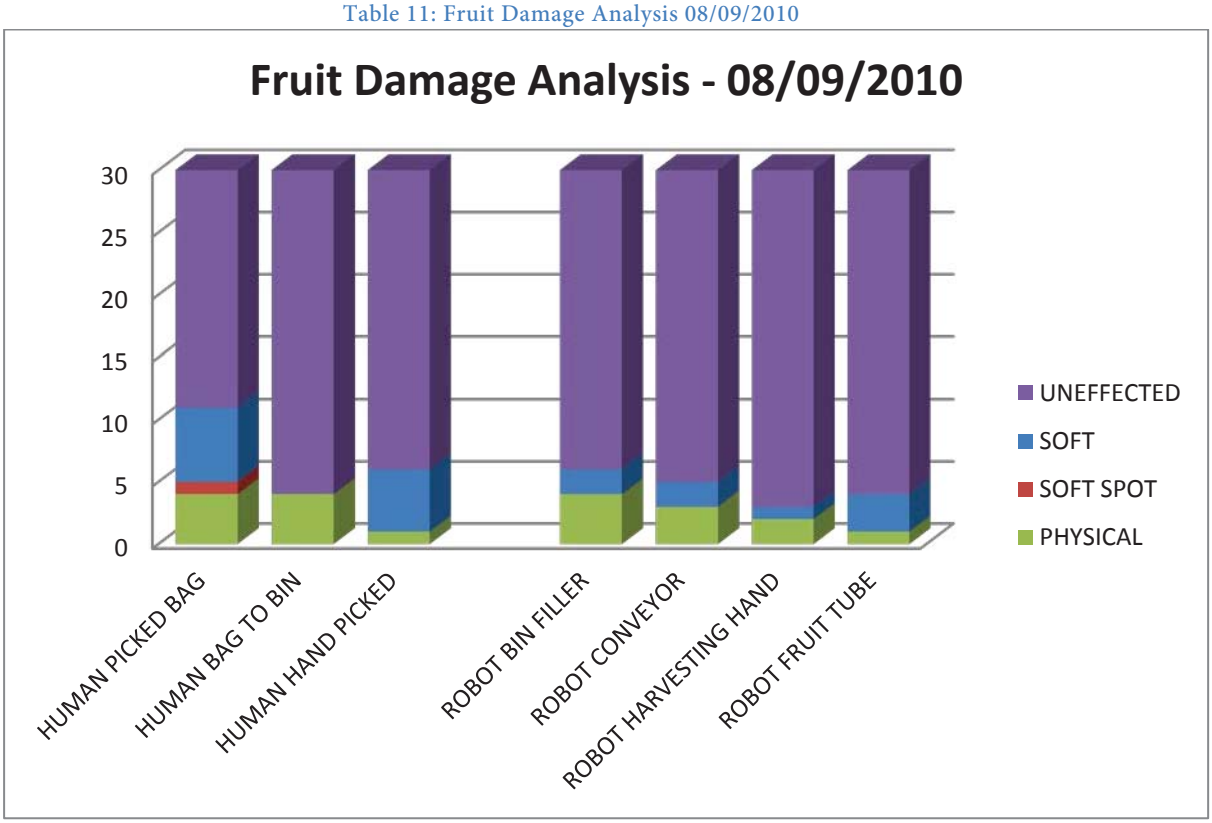


Table 12: Fruit Damage Analysis 22/09/2010

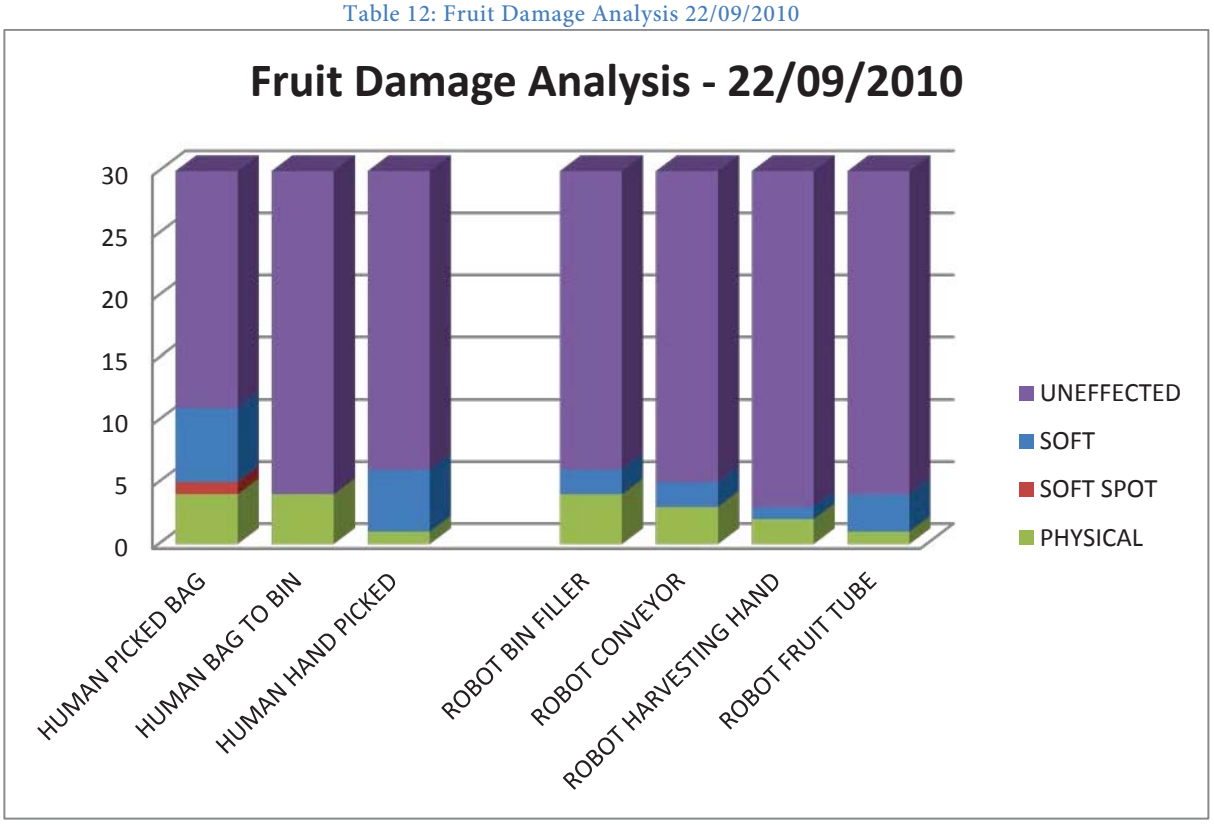
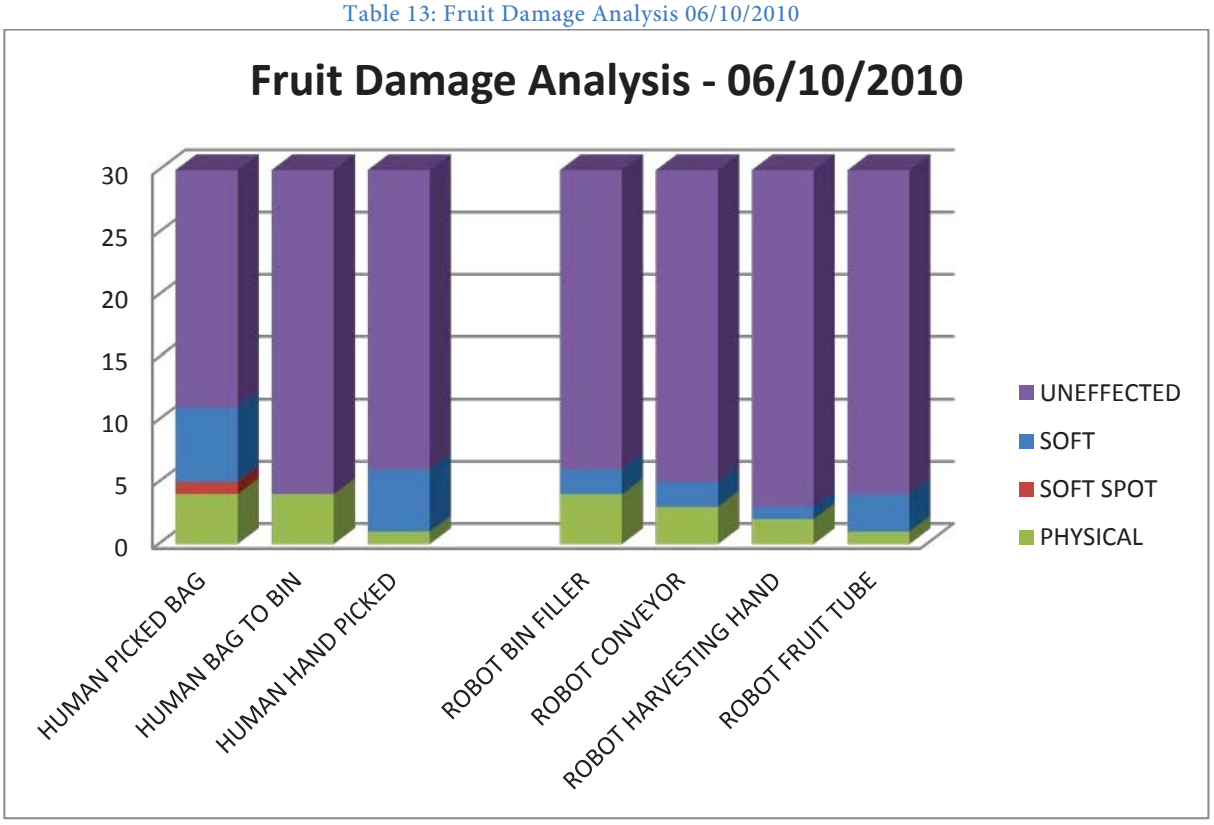


Table 13: Fruit Damage Analysis 06/10/2010



10. References

Aitken, A. G. & Hewett, E. W. (2011). *Fresh Facts, New Zealand Horticulture 2011.* Available: http://www.freshfacts.co.nz/file/fresh-facts-2011.pdf

Aliakbarpour, H., Nuez, P., Prado, J., Khoshhal, K., & Dias, J. (2009), "An Efficient Algorithm for Extrinsic Calibration Between a 3D Laser Range Finder and a Stereo Camera for Surveillance," *3rd International Conference on Advanced Robotics (ICAR)*, pp. 1-6.

Andreoli, E., (2005), Accessed 10/12/2011. *Andreoli Engineering.* Available: http://www.andreoliengineering.com

Astrand, B. & Baerveldt, A. (2002), "An Agricultural Mobile Robot with Vision-Based Perception for Mechanical Weed Control," *Autonomous Robots 13,* pp. 21-35.

Baeten, J., Donne, K., Boedrij, S., Beckers, W., & Claesen, E. (2007), "Autonomous Fruit Picking Machine: A Robotic Apple Harvester," *6th International Conference on Field and Service Robotics,* pp. 63-72.

Bailey, D. G., Seal, J., & Goupta, G. S. (2006), "Validation of a Catadioptric Stereo System," *3rd International Conference on Autonomous Robots and Agents (ICARA),* pp. 539-545.

Bakker, T., Wouters, H., van Asselt, K., Bontsema, J., Tang, L., Muller, J., & van Straten, G. (2008), "A Vision Based Row Detection System for Sugar Beet," *Computers and Electronics in Agriculture,* pp. 87-95.

Baltzakis, H., Argyros, A., & Trahanias, P. (2003), "Fusion of Laser and Visual Data for Robot Motion Planning and Collision Avoidance," *Machine Vision and Applications,* pp. 92-100.

Barawid, O., Ishii, K., Mizushima, A., & Noguchi, N. (2007), "Development of an Autonomous Navigation System using a Two-dimensional Laser Scanner in an Orchard Application,," *Biosystems Engineering,* pp. 139-149.

Belrose, I. (2011), "World Kiwifruit Review - 2011 Edition," Belrose Inc., Pullman, WA 99163-3806, USA,

Benady, M. & Miles, G. E. (1992), "Locating Melons for Robotic Harvesting using Structured Light," *American Society of Agricultural Engineers,* Paper No.: 92-7021.

Benson, E. R., Reid, J. F., & Zhang, Q. (2003), "Machine Vision-Based Guidance System for an Agricultural Small-Grain Harvester," *Transactions of the American Society of Agricultural Engineers,* pp. 1255-1264.

Cardenas, M., Hetzroni, A., & Miles, G. E. (1991), "Machine Vision to Locate Melons and Guide Robotic Harvester," *Transactions of the American Society of Agricultural Engineers (ASAE),* No. 91-7006 pp. 21-27.

Cui, P., Yue, F., & Cui, H. (2006), "Stereo Vision Based Motion Estimation for Lunar Rover Navigation," *5th International Conference on Machine Learning and Cybernetics,* pp. 3847-3852.

Cyganek, B. & Siebert, J. (2009), *An Introduction to 3D Computer Vision Techniques and Algorithms*. West Sussex: John Wiley & Sons Ltd.

D'Esnon, A. G., Rabatel, G., & Pellenc, R. (1987), "Magali: A Self-Propelled Robot to Pick Apples," *American Society of Agricultural Engineers (ASAE),* Paper No. 87-1037.

Davies, E. R. (2005), *Machine Vision: Theory, Algorithms, Practicalities*, Third Edition. San Francisco: Elsevier Ltd.

Deurer, M., Grinev, D., Young, I., Clothier, B., & Muller, K. (2010), "The Impact of Orchard Management on Macro-pore Topology and Function," *Proceedings of the 19th World Congress of Soil Science,* pp. 102-105.

Diekhans, N. (1999), "Claas Industrietechnik Product Information," *CIT Development Department*.

Dijkstra, E. W. (1959), "A Note on Two Problems in Connexion with Graphs," *Numerische Mathematik 1,* pp. 269-271.

Dobrousin, Y. & Edan, Y. (1992), "Real-Time Image Processing for Robotic Melon Harvesting," *American Society of Agricultural Engineers (ASAE),* Paper No.: 92-3515.

Duda, R. & Hart, P. (1972), "Use of the Hough Transform to Detect Lines and Curves in Pictures," *Communications of the Association for Computing Machinery,* pp. 11-15.

Durrant-White, H. (2006), "Autonomous Land Vehicles," *Australian Centre for Field Robotics (ACFR),* p. 33.

Edan, Y. & Miles, G. E. (1991), "Design of an Agricultural Robot for Harvesting Melons," *American Society of Agricultural Engineers (ASAE),* pp. 831-835.

Edan, Y., Rogozin, D., & Miles, G. E. (2000), "Robotic Melon Harvesting," *IEEE Transactions on Robotics and Automation,* 16 pp. 831-835.

Flemmer, R. C. (2009), "A Low Cost Hybrid SCARA Robot," Intellectual Property Office of New Zealand (IPONZ), New Zealand Patent 580787.

Gerrish, J. B., Fehr, B. W., Van, G. R., & Welch, D. P. (1997), "Self-Steering Tractor Guided by Computer Vision," *Applied Engineering in Agriculture,* 13 pp. 559-561.

Goh, K. M., Pearson, D. R., & Daly, M. J. (2000), "Effects of Apple Orchard Production Systems on some Important Soil Physical, Chemical and Biological Quality Parameters," *Biological Agriculture & Horticulture,* 18 pp. 269-292.

Gonzalez, R. C. & Woods, R. E. (2008), *Digital Image Processing*, Third Edition. New Jersey: Prentice Hall.

Grasso, G. & Recce, M. (1996), "Scene Analysis for an Orange Picking Robot," *International Congress for Computer Technology in Agriculture,* ICCTA.

Haralick, R. M. (1987), "Image Analysis using Mathematical Morphology," *IEEE Transactions on Pattern Analysis and Machine Vision Intelligence,* vol. 4 pp. 532 - 550.

Hartley, R. & Zisserman, A. (2010), *Multiple View Geometry*, Second Edition. New York: Cambridge University Press.

Hayashi, M. & Ueda, Y. (1991), "Orange Harvesting Robot," Kubota Co, Sakai, Japan (Mimeo.),

Henrich, D. & Cheng, X. (1992), "Fast Distance Computation for On-line Collision Detection with Multi-Arm Robots," *IEEE - International Conference on Robotics and Automation,* vol.13 pp. 2514-2519.

Holmqvist, R. (1993), *A Laser-Optic Navigation System for Driverless Vehicle* Västra Götalands län: Arnex Navigation.

Huang, L. L. & Barth, M. (2009), "A Novel Multi-Planar LIDAR and Computer Vision Calibration Procedure Using 2D Patterns for Automated Navigation," *IEEE Intelligent Vehicles Symposium,* pp. 117-122.

Huertas, A., Matthies, L., & Rankin, A. (2005), "Stereo-Based Tree Traversability Analysis for Autonomous Off-Road Navigation," *7th IEEE Workshop on Application of Computer Vision,* pp. 210-217.

Infineon (2004), "Smart High-Side Power Switch," Infineon Technologies AG, Munchen,

Jimenez, A. R., Ceres, R., & Pons, J. L. (1999), "A Machine Vision System using a Laser Radar Applied to Robotic Fruit Harvesting," *IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications (CVBVS '99)*, pp. 110-119.

Jimenez, A. R., Ceres, R., & Pons, J. L. (2000), "A Survey of Computer Vision Methods for Locating Fruit on Trees," *Transactions of the American Society of Agricultural Engineers (ASAE),* 43 pp. 1911-1920.

Jimenez, A. R., Cerez, R., & Pons, J. L. (1997), "Shape-base Methods for Fruit Recognition and Localization using a Laser Range-Finder," *Bio-Robotics,* 97 p. 5.

Kassay, L. (1992), "Hungerian Robotic Apple Harvester," *The American Society of Agricultural Engineers,* No. 92-7042 p. 14.

Keicher, R. & Seufert, H. (2000), "Automatic Guidance for Agricultural Vehicles in Europe," *Computers and Electronics in Agriculture,* 25 pp. 169-194.

Kogler, J., Hemetsberger, H., Alefs, B., Kubinger, W., & Travis, W. (2006), "Embedded Stereo Vision System for Intelligent Autonomous Vehicles," *Intelligent Vehicles Symposium,* pp. 64-69.

Kovesi, P. (1995), *Image Features from Phase Congruency*. Nedlands, Western Australia: Robotics and Vision Research Group.

Lane, R. W. (2007), "Avoiding Commodity Hell: How John Deere is Driving Growth through Innovation," I. R. I. A. Meeting, California: John Deere.

Larminie, J. & Lowry, J. (2003), *Electric Vehicle Technology Explained*. West Sussex: John Wiley & Sons Ltd.

Leemans, V. & Destain, M. F. (2006), "Application of the Hough Transform for Seed Row Localisation using Machine Vision," *Biosystems Engineering,* 94 pp. 325-336.

Lengyel, T., Gedarovich, J., Cusano, A., & Peters, T. (2011). "GPU Vision: Accelerating Computer Vision Alogrithms with Graphics Processing Units." C13 Software.

Levi, P., Fella, R., & Pappalardo, R. (1988), "Image Control Robotics Applied to Citrus Fruit Harvesting," *Proceedings of the 7th International Conference on Robot Vision and Sensory Controls,,* 12 pp. 327-338.

Lowe, D. G. (1999), "Object Recognition from Local Scale-Invariant Features," *The Proceedings of the 7th IEEE International Conference on Computer Vision,,* 2 pp. 1150-1157.

Lumelsky, V. & Skewis, T. (1990), "Incorporating Range Sensing in the Robot Navigation Function," *IEEE Transaction on System, Man, and Cybernetics,* pp. 1058-1069.

Moghadam, P., Wijesome, W. S., & Feng, D. J. (2008), "Improving Path Planning and Mapping Based on Stereo Vision and LIDAR,," *Conference on Control, Automation, Robotics and Vision,* 10 pp. 384-390.

Montemerlo, M., Thrun, S., HDahlkamp, H., Stavens, D., & Strohband, S. (2006), "Winning the DARPA Grand Challenge with an AI Robot," *American Association of Artifical Intelligence National Conference on Artificial Intelligence*, Boston, USA.

Morimoto, E., Suguri, M., & Umeda, M. (2005), "Vision-based Navigation System for Autonomous Transportation Vehicle," *Precision Agriculture,* 6 pp. 239-254.

Morris, J., Khurram, J., Gimel'farb, G., & Khan, T. (2009), "Breaking the 'Ton': Achieving 1% Depth Accuracy from Stereo in Real Time," *International Conference Image and Vision Computing New Zealand (IVCNZ 2009),* pp. 142-147.

Mowat, A. (2011), "Autonomous Kiwifruit Harvester: The Future for Industry", Published by Southern Cross Robotics Ltd, for ZESPRI International Ltd, Mount Manuganui.

Murakami, N., Ito, A., Will, J. D., Steffen, M., Inoue, K., Kita, K., & Miyaura, S. (2008), "Development of a Teleoperations System for Agricultural Vehicles," *Computers and Electronics in Agriculture 63,* pp. 81-88.

Muscato, G., Prestifilippo, M., Abbate, N., & Rizzuto, I. (2005), "A Prototype of an Orange Picking Robot: Past History, the New Robot and Experimental Results," *Industrial Robot: An International Journal,* pp. 128-138.

Ollis, M. & Stentz, A. (1997), "Vision-Based Perception for an Automated Harvester," *Intelligent Robotics and Systems (IROS '97),* 3 p. 6.

Parrish, E. & Goksel, A. K. (1977), "Pictorial Pattern Recognition applied to Fruit Harvesting," *American Society of Agricultural Engineers,* 20 p. 5.

Pilarski, T., Happold, M., Pangels, H., Ollis, M., Fitzpatrick, K., & Stentz, A. (2002), "The Demeter System for Automated Harvesting," *Autonomous Robots 13,* pp. 9-20.

Reece, M., Taylor, J., Plebe, A., & Tropiano, G. (1996), "Vision and Neural Control of an Orange Harvesting Robot," *Department of Anatomy and Developmental Biology,* p. 9.

Robert (2006). "Hydraforce Model Numbers." *Hydraforce*. Available: www.hydraforce.com/,

Rose, C. (2010), "Fruit Ripe for Mechanisation," *HDC News,* August 2010 pp. 10-14.

Rovira-Mas, F., Wang, Q., & Zhang, Q. (2009), "Bifocial Stereoscopic Vision for Intelligent Vehicles," *International Conference on Vehicular Technology,* 2009 p. 9.

Rushing, K. (1971), "Developing Driverless Tractor," *Agricultural Engineering,* 52 pp. 260-270.

Ryo, T., Noboru, N., & Akira, M. (2004), "Automatic Guidance with a Laser Scanner for a Robot Tractor in an Orchard," *Automation Technology for Off-Road Equipment,* p. 9.

Sarig, Y. (1993), "Robotics of Fruit Harvesting: A State-of-the-art Review," *Journal of Agricultural Engineering Research,* 54 pp. 265-280.

Scharstein, D. & Szeliski, R. (2002), "A Taxonomy and Evaluation of Dense Two-Frame Stereo Corrospondence Algorithms," *International Journal of Computer Vision,* vol. 47 pp. 37-42.

Schertz, C. E. & Brown, G. K. (1968), "Basic Considerations in Mechanizing Citrus Harvesting," *American Society of Agricultural Engineers (ASAE),* pp. 343-346.

Sensors, S.-I. (2008), "LMS100/111/120/151 Laser Measurement Systems," SICK AG Waldkirch,

Shanshan, C., Wuheng, Z., & Lijun, Z. (2009), "Camera Calibration via Stereo Vision Using Tsai's Method," *First International Workshop on Education Technology and Computer Science,* vol. 3 pp. 273-276.

Sissons, R. (1939), "Ploughing in Circles Saves Time," *Prarie Farmer,* 20.

Sites, A. & Delwiche, M. J. (1988), "Computer Vision to Locate Fruit on a Tree," *American Society of Agricultural Engineers (ASAE),* 85-3039.

Slaughter, D. C., Giles, D. K., & Downey, D. (2008), "Autonomous Robotic Weed Control Systems: A Review," *Computers and Electronics in Agriculture,* pp. 63-78.

Slaughter, D. C. & Harrell, R. C. (1989), "Discriminating Fruit for Robotic Harvest using Color in Natural Outdoor Scenes," *Transactions of the American Society of Agricultural Engineers,* 32 pp. 757-763.

Sobel, I. (1970), "Camera Models and Machine Perception," PhD, Department of Computer Science, Stanford University, California, 1970.

Sogaard, H. T. & Olsen, H. J. (2003), "Determination of Crop Rows by Image Analysis Without Segmentation," *Computers and Electronics in Agriculture,* 38 pp. 141-158.

Spencer, S. (2004), "The Development of an Autonomous Vehicle for use in Agriculture," Masters of Science, Biological Engineering, North Carolina State University, North Carolina, 2004.

Subramanian, V. & Burks, T. F. (2009), "Headland Turning Maneuver of an Autonomous Vehicle Navigating a Citrus Grove using Machine Vision and Sweeping Laser Radar," *ISHS Acta Horticulturae 824: International Symposium on Application of Precision Agriculture for Fruits and Vegetables,* pp. 321-328.

Subramanian, V., Burks, T. F., & Arroyo, A. A. (2006), "Development of Machine Vision and Laser Radar Based Autonomous Vehicle Guidance Systems for Citrus Grove Navigation," *Computers and Electronics in Agriculture,* 53 pp. 130-143.

Teat, J. (2011), "Short Range Electric Vehicle Motor Controller," Master of Science, Department of Physics, University of Dunedin, Dunedin, 2011.

Travis, W., Simmons, A. T., & Bevly, D. M. (2005), "Corridor Navigation with a LiDAR/INS Kalman Filter Solution," *IEEE - Proceedings from the Intelligent Vehicles Symposium,* pp. 343-348.

Van Henten, E. J., Hemming, J., Van Tuijl, B. A., Kornet, J. G., Meuleman, J., Bontsema, J., & Van Os, E. A. (2002), "An Autonomous Robot for Harvesting Cucumbers in Greenhouses," *Autonomous Robots vol. 13,* pp. 241-258.

Van Henten, E. J., Hemming, J., Van Tuijl, B. A. J., Kornet, J. G., & Bontsema, J. (2003), "Collision-free Motion Planning for a Cucumber Picking Robot," *Biosystems Engineering,* 86 pp. 135-144.

Van Henten, E. J., Van Tuijl, B. A. J., Hemming, J., Kornet, J. G., Bontsema, J., & Van Os, E. A. (2003), "Field Test of an Autonomous Cucumber Picking Robot," *Biosystems Engineering,* 86 pp. 305-313.

Van Kollenburg-Crisan, L. M., Sjwed, D. A., Wennekes, P. C., & Kornet, J. G. (1999), "Method and Device for Cutting Stalks of Fruits," Netherlands Patent 1008161.

Whittaker, Miles, Mitchell, & Gaultney (1987), "Fruit Location in a Partially Ocluded Image," *American Society of Agricultural Engineers (ASAE),* vol. 30 p. 7.

Whittaker, W. R. (2005), "Red Team DARPA Grand Challenge 2005 Technical Paper," in *CSC Technical Report* Pittsburgh: Carnegie Mellon University.

Willrodt, F. L. (1930), "Steering Attachment for Tractors," United States Patent 1776940.

Zhang, Q. & Pless, R. (2004), "Extrinsic Calibration of a Camera and Laser Range Finder (Improves Camera Calibration)," *IEEE International Conference on Intelligent Robots and Systems (IROS 2004).* , pp. 2301-2306.

Zhang, W. & Du, S. (2005), "Hough Transform Applying in Vision Navigation of Farming Machine," *Chinese Journal of Scientific Instrument vol. 26,* pp. 706-707.

Ziegler, J. G. & Nichols, N. B. (1942), "Optimum Settings for Automatic Controllers," *American Society of Mechanical Engineers vol. 65,* pp. 433-444.