**אוניברסיטת-אביב**

הפקולטה להנדסה ע"ש איבי ואלדר פליישמן

בית הספר לתארים מתקדמים ע"ש זנדמן-סליינר

**הערכת יבול אולטראסונית**

חיבור זה הוגש כעבודת גמר לקראת התואר "מוסמך אוניברסיטה" בהנדסת מכנית

על- ידי

**רועי פינקלשטיין**

העבודה נעשתה בביה"ס להנדסה מכנית בהנחית

**ד"ר גבור קושה, ד"ר יוסי יובל וד"ר אביטל בכר**

כסלו תשע"ז

תקציר

הספקטרום החוזר מהדים אולטראסונים של צמחים הראה במחקר מקדים כי הוא מכיל מידע לגביהם. את הפוטנציאל המלא שהדים אלו מכילים אפשר לראות בעטלפים שנעזרים בראייה ובביו-סונר מיליוני שנים לניווט ולחיפוש אחר מזון. עטלפים מסוגלים למפות את סביבות הגידול הצמחיות שלהם בדיוק רב ובצורה דינאמית ע"י מעבר על טווח של תדירויות אולטראסוניות. חושי העטלפים מתעלים על יכולות המערכות אולטראסוניות המתוחכמות ביותר שבשימוש האדם.

בעוד מערכות סונר תת מימיות נמצאות בשימוש נרחב במשימות צבאיות כמו מיפוי רצפת האוקיינוס ברזולוציה גבוהה, גילוי מוקשים ואפילו לשימושים חקלאיים כמו הערכת כמות דגה – הסונר בתווך אוויר משמש לניתוח מידע אטמוספרי ובשדה הרובוטיקה משמש בעיקר כאמצעי למדידת מרחק.

מחקר זה מתמקד בפיתוח מערכת אולטראסונית שתשולב בסופו של דבר במערכת רובוטית להערכת יבול. בבסיס המערכת עומד עקרון של ניתוח מאפיינים אולטראסונים שהתקבלו משימוש במערכת הסונר על צמחים ופיתוח אלגוריתמים להערכת מסת הפרי וכמות העלים בצמח. לצורך המחקר פותחה מערכת ניסוי אולטראסונית שנוסתה במעבדה, ובחממת פלפל ברמת שורת הצמחים, צמחים בודדים ואף פירות ועלים בודדים.

התוצאות מצביעות על כוחם של החיישנים האולטראסונים בהפרדת צמחים ותשתיות החממה כמו קירות. המחקר מראה את יכולות הסונר בגילוי שורות צמחים אשר מוסתרות מהחיישן בזמן אמת (אין לחיישן קו ראיה ישיר עם השורה) . כמו כן, המחקר מציג את היכולת להעריך את מסת הפרי וכמות העלים בצמחים. מודל רב משתנים המעריך את האנרגיה שחוזרת מצמח בודד הראה תוצאות מובהקות סטטיסטית עם ערכי 𝑅2 של 0.64 ו0.84 בתדירויות 28 קה"ץ עד 33 קה"ץ ו20 קה"ץ עד 28 קה"ץ בהתאמה. האלגוריתם לחיזוי מסת הפרי החזיר ערכי 𝑅2 של 0.34 ועבור מס' עלים 0.74. רשת נוירונים אומנה להערכת כמות פרי ומס' עלים והחזירה ערכי 𝑅2 של 0.86 ו 0.95 בהתאמה.

**TEL AVIV UNIVERSITY**

School of Mechanical Engineering

Mechanical Engineering

**Ultrasonic yield assessment**

A Paper submitted toward the degree of

MSc in Mechanical Engineering

By

**Roee Finkelshtain**

This research was carried out in The School of Mechanical Engineering

Under the supervision of

**Dr. Gabor Kosa, Dr. Yossi Yovel and Dr. Avital Bechar**

December 2016

Table of Content

[List of Figures 6](#_Toc469575355)

[Abstract and Objectives 7](#_Toc469575356)

[Nomenclature 8](#_Toc469575357)

[Symbols 10](#_Toc469575358)

[1 Introduction 11](#_Toc469575359)

[1.1 Brief background and importance 11](#_Toc469575360)

[1.2 Problem definition 12](#_Toc469575361)

[1.3 Research objective, innovations and contributions 12](#_Toc469575362)

[1.4 Scope and limitations 14](#_Toc469575363)

[1.5 Thesis structure 15](#_Toc469575364)

[2 Literature review 17](#_Toc469575365)

[2.1 Automatic yield assessment 17](#_Toc469575366)

[2.2 Air sonar fundamentals 19](#_Toc469575367)

[2.2.1 Transducers 20](#_Toc469575368)

[2.2.2 Speed of sound 21](#_Toc469575369)

[2.2.3 Attenuation 22](#_Toc469575370)

[2.2.4 Signal to noise ratio 23](#_Toc469575371)

[2.2.5 Directivity patterns 23](#_Toc469575372)

[2.3 Air sonar applications 25](#_Toc469575373)

[2.3.1 Improving directivity 25](#_Toc469575374)

[2.3.2 Time of flight classification 26](#_Toc469575375)

[2.3.3 Pseudo amplitude based classification 28](#_Toc469575376)

[2.3.4 Spectrum based classification 29](#_Toc469575377)

[2.4 Air sonar in agriculture 33](#_Toc469575378)

[2.5 Sonar Simultaneous Localization and Mapping 34](#_Toc469575379)

[2.5.1 Sonar and LiDAR SLAM 37](#_Toc469575380)

[2.5.2 Sonar and vision SLAM 37](#_Toc469575381)

[3 Materials and Methods 39](#_Toc469575382)

[3.1 Instrumentation 39](#_Toc469575383)

[3.1.1 Sensory apparatus 1.0 40](#_Toc469575384)

[3.1.2 Sensory apparatus 2.0 41](#_Toc469575385)

[3.1.3 Data acquisition apparatus 42](#_Toc469575386)

[3.1.4 Software 43](#_Toc469575387)

[3.1.5 Robotic platforms 44](#_Toc469575388)

[3.2 Experimental setup 45](#_Toc469575389)

[3.2.1 Single object experiment 46](#_Toc469575390)

[3.2.2 Qualitative experiment 46](#_Toc469575391)

[3.2.3 Prediction features extraction experiment 48](#_Toc469575392)

[3.3 Ordinary Least Squares attenuation and hidden row detection 48](#_Toc469575393)

[3.4 Applicative experiment 49](#_Toc469575394)

[3.4.1 Scanning procedure 50](#_Toc469575395)

[3.5 Statistical tools for results analysis 52](#_Toc469575396)

[3.6 Neural Networks estimation 54](#_Toc469575397)

[4 Experimental Results 55](#_Toc469575398)

[4.1 Qualitative experiments 55](#_Toc469575399)

[4.1.1 Attenuation compensation 55](#_Toc469575400)

[4.1.2 Ground effect 56](#_Toc469575401)

[4.1.3 Detection of crop rows 57](#_Toc469575402)

[4.1.4 Different cultivars 61](#_Toc469575403)

[4.1.5 Greenhouse wall and vegetation 62](#_Toc469575404)

[4.2 Applicative results 64](#_Toc469575405)

[4.2.1 Classification 64](#_Toc469575406)

[4.2.2 Mapping 66](#_Toc469575407)

[4.2.3 Hidden object detection 68](#_Toc469575408)

[4.3 Fruit and leaves statistical analysis and classification 69](#_Toc469575409)

[4.3.1 Single object experiment 69](#_Toc469575410)

[4.3.2 Prediction features extraction experiment 70](#_Toc469575411)

[4.3.3 Leaves experiment 77](#_Toc469575412)

[4.3.4 Neural Networks estimation 80](#_Toc469575413)

[5 Conclusions 83](#_Toc469575414)

[Reference List 85](#_Toc469575415)

[Publications 90](#_Toc469575416)

[Appendix 91](#_Toc469575417)

# List of Figures

[Figure 1: Beam pattern 24](#_Toc470387458)

[Figure 2: System’s Block Diagrams 39](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387459)

[Figure 3: UltraSroundGate 116hm Recording System 40](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387460)

[Figure 4: The Sensing System 42](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387461)

[Figure 5: Basic Data Processing Steps 43](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387462)

[Figure 6: Robotics Platforms 45](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387463)

[Figure 7: Experimental setups 46](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387464)

[Figure 8: Qualitative Expirment Ensonofication Positions 47](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387465)

[Figure 9: Real Time Application Setup 49](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387466)

[Figure 10 : Flow chart of the real time classification and hidden object detection Algorithm 51](#_Toc470387467)

[Figure 11: A view of the NN structure. 54](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387468)

[Figure 12: Attenuation values according to frequency. 55](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387469)

[Figure 13: Ground Effect 57](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387470)

[Figure 14: Detection of crop rows 58](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387471)

[Figure 15: Lidar Data 60](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387472)

[Figure 16 : Averaged amplitude as a function of frequency of pepper and cucumber plants. 61](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387473)

[Figure 17 : Slope and Frequencey Response comparason between two species. 62](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387473)

[Figure 18: Wall and plant spectrum at 1.0[m] 63](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387474)

[Figure 19 : Average support vectors 65](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387475)

[Figure 20: Mapping process 67](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387476)

[Figure 21: The sensing appartus’s correlation sum over spectrogram 68](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387477)

[Figure 22: Frequency spectrum of leaves and fruit. 69](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387478)

[Figure 23: Features Extraction Experiment 71](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387479)

[Figure 24: Sum of positive changes between spectrograms 72](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387480)

[Figure 25: The number of peaks per frequency for the 15 plant-orientation combinations 72](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387481)

[Figure 26: Predicted against experimental results in frequencies 28[kHz] - 32[kHz] 74](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387482)

[Figure 27: Predicted against experimental results in frequencies of 20[kHz] to 28[kHz] 75](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387483)

[Figure 28: Predicted against experimental mass 76](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387484)

[Figure 29: Quadratic model surface & Observations 78](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387486)

[Figure 30: Predicted against experimental number of leaves from 1.25[m] 78](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387485)

[Figure 31: Neuralnetworkprediction of fruit mass from 1.25[m] with no cross validation. 80](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387487)

[Figure 32: Neural network prediction of number of leaves from 1.25[m] with no cross validation 81](file:///E:\Roee%20Thesis%20Last%20Version\Roees%20Thesis%20LAST%20VERSION%20Chapter%204.2.docx#_Toc470387488)

# Abstract and Objectives

The spectrum of an ultrasonic return echo from plants has shown in previous research to contain information about them. The full potential that this information holds is observed in bats which have been using bio-sonar and vision in nature for millions of years for navigation and foraging. Bats are capable of dynamically mapping their dense vegetation habitats with great accuracy by sweeping through ultrasonic frequencies. Their senses improved throughout evolution to exceed any current in-air ultrasonic engineering systems. While underwater sonar systems are used in diverse military missions as mapping the sea floor in high resolutions, help in detection of mines and even agriculture applications such as evaluation of the tonnage of fisheries – air sonar is used for atmospheric inquiries and in the field of robotics it is mostly used as distance sensors.

This research focuses on developing an ultrasonic sensing system that would ultimately be used as the basis for a yield estimation robotic system. This is made possible through analyzing ultrasonic features obtained by ensonfying plants and developing algorithms for evaluation of fruit mass and leaf number. An ultrasonic sensor system was developed and tested in a lab, a pepper greenhouse and on single pepper plants, single leaves and fruit. The results showed the potential of ultrasonic sensors for such a robot in classifying plants and greenhouse infrastructures such as walls. It showed the sensor's ability to detect hidden plant rows and fruits in real time as well as making an estimation of the fruit mass and leaves number in single plants. A developed multi linear regression model for estimating the energy level was found to be highly significant with R2 of 0.64 and 0.84 for 28 kHz to 33 kHz and 20 kHz to 28 kHz ranges, respectively. The algorithm for prediction of fruit mass per plant yielded 𝑅2 of 0.34 and 0.72 for leaves. A neural network was trained for fruit and leaves estimation as an alternative evaluation method and it yielded a significance level of 0.86 and 0.95 respectively.

# Nomenclature

**ANOVA** - Analysis of Variance.

**ARL** - Agricultural Robotics Lab.

**RBM2S** - Robots and Bio-Medical Micro-System.

**CHIRP** - Compressed High Intensity Radar Pulse. A signal with a varying frequency over time. Despite the word "radar" in the acronym the term is commonly used in the field of ultrasonic sensing. (Also referred as frequency modulated signals and sweep signals).

**DAQ** – Data Acquisition.

**Echolocation** – Animal's bio-sonar navigation system.

**EKF** - [Extended Kalman filter](https://en.wikipedia.org/wiki/Extended_Kalman_filter)

**Ensonify** - To fill with sound.

**FFT** - Fast Fourier Transform.

**FrFT** - Fractional Fourier Transform.

**Frequency response** - Representation of the signal's intensity in the frequency domain. The frequency response is defined in this work as the sum of the echo amplitudes in a specific frequency divided by the mean CHIRP amplitude of the same frequency.

**HRTF** - Head Related Transfer Function.

**LDA** - Linear Discriminate Analysis.

**LIDAR** - Light Detection and Ranging.

**OLS** - Ordinary Least Squares.

**PA** - Precision agriculture.

**PAS** - Pseudo Amplitude Scan.

**PSD** - Power Spectrum Density.

**RANSAC** - Random sample consensus

**Reflectors** - objects that reflect the soun ds produced by sonar.

**SLAM** - Simultaneous Localization and Mapping.

**SNR** - Signal Noise Ratio.

**SONAR** **Artifacts** - pseudo-objects detected by sonar due to the complexity of the reflections in the environment.

**SoS** - Speed of Sound.

**Spectrogram** - Representation of a signal's intensity in the time-frequency space.

**TLS** - Total Least Squares.

**ToF** - Time of Flight.

**Thresholding** - Analyzing images using image thresholding techniques. Image thresholding is a simple, yet effective, way of partitioning an image into a foreground and background

# Symbols

𝒄𝑯 is the speed of sound

**DI** the directivity

𝑬𝒊𝒋 is a monotone function of the fruit mass / leaves number

𝒇𝒊𝒋 ensonified frequency

𝒉𝒓 is the relative humidity

𝒍𝒊𝒋 is the number of leaves

𝒎𝒊𝒋 is the fruit mass

**NL** the noise level

**PiOj –** i represents the plant number and j represents the orientation number

**SL** is the source amplitude level

𝐓𝐜 is the environment temperature in ℃

**TL** the transmission loss in air due to attenuation and spherical spreading pattern of the beam

**TS** the target strength is the amount of acoustic energy returned from an object

𝒙𝒊𝒋 is unknown and represent specific factors of 𝑃𝑖𝑂𝑗 that are unobserved

𝜶, 𝜷,  statistical model parameters

𝜹𝒊𝒋 indicator variable

𝝐𝒊𝒋 random error

# 1 Introduction

## 1.1 Brief background and importance

Precision agriculture (PA) is a field in agriculture studies concentrating on selective decision making based on processing of detailed farm information. The motivation behind PA is at improving aspects of the future farm such as crop profitability, farm productivity, and environmental influence by reducing amount of fertilizers, fuel, manual work, lease and crop insurance payments via technological means (Mulla 2013).

Specific yield assessment is essential for precision farming and agriculture in general. It is an important tool in agriculture for forecasting crop revenues, planning of budget and store capacity, labor management and compensation calculations (Fermont and Benson 2011). In several crops, such as deciduous trees, fruit thinning is done based on the yield estimation which conducted by manual sampling of selective trees.

An accurate and site specific yield assessment technique that will decrease the assessment cost and increase its accuracy has the potential to reduce the production costs, increase yield and profitability and save billions of dollars in tax money subsidies.

Governments world-wide are currently developing crop insurance programs as a main mechanism to support their farmers through subsidized risk management that is not subject to World Trade Organization constraints. Subsidized crop insurance has become the most important single support policy in agriculture at both the US and Israel. The program is immense in the US, currently insuring over $120 billion in agricultural values and costing US taxpayers approximately $10 billion each year (Glauber 2013, Goodwin and Smith 2013). In Israel, government subsidies for the crop and disaster insurance programs amount to more than 100 million NIS annually and the aggregate premium payments amount to 150 million NIS annually (Kanat 2015).

## 1.2 Problem definition

Traditional techniques for yield assessment include crop cuts of the field (statistical evaluation), and subjective assessment by farmers and professional appraisers. Farmer prediction relies on past annual yields which vary greatly between years. Crop cuts and assessment by professional appraisers are considered to be more reliable by the Food and Agriculture Organization, but require more labor and tend to be expensive.

Moreover, they are carried out manually by workers in the field and are based on crop sampling in small quantities and hence, in addition loose the variation information. There is a tradeoff between the amount of time invested in sampling the crop and the accuracy given the inhomogeneous nature of crop distribution (Fermont and Benson 2011).

To overcome difficult problems such as the variability in agricultural produce and continuously changing conditions, it is necessary to develop intelligent systems that will achieve successful task performance in such environments. Information acquisition systems, including sensors, fusion algorithms and data analysis need to be improved and adjusted to the dynamic conditions of unstructured agricultural environments (Bechar 2010):

Various sensing technologies were considered for yield estimation (Lee, Alchanatis et al. 2010) such as thermal imaging (Stajnko, Lakota et al. 2004), depth cameras (Andújar, Fernández-Quintanilla et al. 2015), optical methods (Wachs, Stern et al. 2010) and more. All of the above approaches have one thing in common: they require either a direct line of sight to all fruits or relay on past relations between visible yields to total yield.

## 1.3 Research objective, innovations and contributions

The objective of the research presented in this work is the development of a real-time automated system for detection of crop biomass and yield estimation in pepper plants suitable for robotic applications.

The proposed sensing system introduces an air sonar based method which is advantageous in the sense that sound waves are capable of propagating through foliage. This advantage is also the most innovative aspect of this work as it enables the detection of objects occluded by the plants foliage. With this technique all fruit can be observed and taken in account instead of just the fruit with visual contact to sensor. The classification of plant properties is possible due to sonars high-resolution distance measurement and differences in the return energy for each frequency which occur due to acoustic impedance, texture and mostly geometry of the ensonified object.

To examine the feasibility of an ultrasonic sensing system for yield assessment the following questions have guided the research:

* Is there a difference between the ultrasonic returns of greenhouse infrastructure and vegetation?
* What is the effect of fruit mass and foliage volume on the ultrasonic returns of the plants?
* Which frequencies, if any, hold the most information about the fruit mass and foliage?
* How significant is the single plants geometry on the ultrasonic return?
* Can the physical relationship between fruit mass and leaves number to the ultrasonic returns be reversed to obtain a prediction algorithm?

Answering the hypothesis questions and achieving the research objective will contribute a new approach in agricultural sensing and may be successful enough to replace traditional yield assessment techniques when mounted on a robot.

The difference between the ultrasonic returns of greenhouse infrastructure and vegetation was referred to in sec. (4.1.5) in which the frequency response of the objects was compared.

The effect of fruit mass and foliage volume was researched extensively in this study by approximating the constants of multivariate regression in Sec.(4.3.2) in which these two plant features were taken as variables and by training a neural network with these features as the output variables Sec.(4.3.4). The multivariate regression contained indicator variables that gave a measure to the significance of the single plant geometry. The frequency bands that were chosen for the evaluations were considered after reviewing the full frequency range for the largest differences between spectrograms see Sec.(4.3.2). Eventually the multivariate analysis and neural network models were used to reverse the physical relationship between fruit mass and leaves number to the ultrasonic returns to obtain evaluations of these quantities.

The main findings of this thesis have been presented (or are under review) in:

* Finkelshtain R., Bechar A., Yovel Y., Kósa, G., “Sonar-based Navigation of Agricultural Robot in Greenhouses: Preliminary Results”, the annual meetings of the Israeli Society of Agricultural Engineering, Israel, 2014. (Oral presentation)
* Finkelshtain R., Yovel Y., Kósa G., Bechar A., “Autonomous robot for yield assessment.”, The 33rd Israeli Conference on Mechanical Engineering, Israel, 2015 (Oral presentation).
* Finkelshtain R., Yovel Y., Kósa, G., Bechar A. “Autonomous robot for yield assessment.”, 10th Conference of the ECPA, ARO, Israel, July 2015. (Proceeding)
* Kósa G., Yovel Y., Bechar A., Finkelshtain R., "AGRICULTURAL ROBOT", PCT/IB2016/050303. (Patent submission).
* Finkelshtain R. , Kósa G., Yovel Y., Bechar A., "An agricultural robot for yield assessment using ultrasonic-based feature perception", The 5th Israeli Conference on Robotics, Israel, 2016, (Oral presentation).
* Finkelshtain R., Kósa G., Yovel Y., Bechar A., “Investigation and analysis of an ultrasonic sensor for specific yield assessment and greenhouse features identification” , Precision Agriculture, 2016, accepted for publication.

## 1.4 Scope and limitations

This thesis is focused on yield assessment in pepper greenhouses at the single plant level but same analysis may also hold for other crops.

The limitations of current air sonar sensors are the relatively low sampling rate, bad directivity patterns and attenuation with distance. In addition, the presented sensing system gives a one dimensional input due to current costs of high sampling rate data acquisition systems. Adding more microphones would help improving the predictions and may hold potential to solve other PA tasks boarder than the yield assessment problem such as single fruit localization.

Two key assumptions that are important for the understanding of the study are that the speed of a greenhouse monitoring robot is much smaller than 𝑐𝐻 for safety issues so that any consequences of the Doppler Effect can be neglected. This is important as the basis for this work is frequency response analysis.

Furthermore, the sum of energy of a plant's echo is assumed to be unbiased by fruit distance to the sensing system as the fruits are distributed normally (Bloch, Dgani et al. 2013).

## 1.5 Thesis structure

To meet the thesis objective and answer the research questions the work first presents a brief overview of the fundamentals of air sonar and related cutting edge research in section 2. Following is a description of the constructed real-time ultrasonic sensing system for yield assessment and its integration onto a greenhouse mobile robotic platform and in-lab robotic manipulator in section 3.1. Subsection 3.1.4 presents the basic signal processing algorithms that translate the extracted raw data gathered by the sensing system in experiments to the return energy at different frequencies. The experimental procedures practiced through the work and results are separated into three main subjects in chapters 3 and 4:

1. Experiments designed to gather ultrasonic data from the greenhouse environment and from single plants to examine qualitative results such as differences between plant rows and greenhouse walls, differences between different cultivars, detection of occluded plant rows, and detection of differences between plants with and without fruit.
2. An indoor lab demonstration showing the feasibility of real time ultrasonic Linear Discriminate Analysis (LDA) classification of plants and walls and detection of objects hidden by vegetation.
3. An experiment focused on showing the feasibility of a practical algorithm to estimate the fruit mass and number of leaves of a plant. The amplitude and frequency data returning from the ensonified plants is used to establish the basis for an automatic yield assessment system. A statistical analysis of the data was held to examine the significance of the relation between the measured data and the number of fruit and leaves with polynomial models. Predictors based on this statistical model and a Neural Network (NN) approach were established, cross validated and compared.

The summary and discussion over important results, observation regarding the research questions and future research ideas are presented to conclude the work in the last chapter.

# 2 Literature review

## 2.1 Automatic yield assessment

There are currently no commercial automatic yield assessments tools although steps in that direction have already been made. One of these steps is a digital processing technique of thermal imaging. This method relies on the different heat capacity of fruits in relation to the foliage.

Stajnko, Lakota et al. (2004) matched the number of manually counted apples to computer detected apples in 20 thermal images with an 𝑅2 of 0.83 to 0.88 by normalizing the color intensity and using a Normalized Difference Index (NDI) to discriminate plants from background.

Andújar, Fernández-Quintanilla et al. (2015) used the Kinect v1 sensor to model 3D plants offline by using statistical filters and manually stitching the 3D point clouds. Their research examined the best angles for biomass estimation. They have detected the height of the plants with an error of a few centimeters and high correlations between biomass and Kinect readings with 𝑅2 ranging 0.88 to 0.92.

Wachs J. et al. (2010) tried two approaches with an RGB and IR thermal cameras at detecting green apples. First, a high level approach in which the RGB image is translated to another color space was used. The K-means algorithm classifies the pixels in the new color space and the IR intensity dimension to the nearest predetermined class centroid (leaves/fruit/background). The new image is scanned for disk shaped apples using Hough transform which achieves 53% accuracy. The second, low level approach uses Haar classifiers with NN to achieve 71% accuracy when incarnated with the first approach. The combination of these methods improves the accuracy while lowering the number of false alarms.

The research described in Moonrinta J., et al. (2010) dealt with pineapple mapping using a monocular camera. An SVM classifier was created with a training set of images that tested combinations of SIFT, SURF and Harris features with SIFT or SURF descriptors. The points classified as positive for indicating fruits are identified as regions, if they were dense enough in a disk shaped structure elements. Such regions open new region trajectories or join previous trajectories if they overlap. Key frames are manually picked offline to create 3D fruit models using RANSAC and least squares ellipsoid fitting. 20 out of 20 fruits were detected in at least 80% of a moving trajectory of the system. However, the assessment of fruit size is less accurate and the ellipsoid models divert from the actual fruit size by more than 50%.

Nuske S., et al. (2011) predicted crop mass in vineyards using image processing with regular cameras. Since 90% of the variation in yield is embodied in the number of berries per cluster and clusters per plant, the researchers tried a new approach for yield estimation which is impractical for workers: counting the number of visible grape berries for the whole crop. The number of visible berries is later shown to have a linear relation with the crop yield. The identification of berries is first done with a radial symmetry transform: implying a Gaussian filter on the gradient of the image and using a radii variation on the Hough transform, edge pixels that are over a certain threshold are considered as places of radial symmetry. By segmenting these places of radial symmetry to 32 features and using the K-means algorithm with a training set - the radial symmetry places are sorted to positive/negative places for berries and positive places without at least 5 neighbors are removed. Results show that 53% to 74% of the visible berries were detected correctly depending on the grape specie and yet only a 9.8% error was measured in the yield estimation.

Improvements in berry detection can be done with multiple angles flash to detected radial edges better. Wang Q., et al. (2013) detected the number of apples in an orchard with a two camera stereo rig. The experiments were held at night with artificial light and the apples were relatively exposed due to the tall spindle planting system in the orchard. These controlled conditions allowed for segmentation in the hue, RGB intensity, and saturation channels by thresholding. The apples were segmented and counted as single apples by their roundness score. The roundness was determined by the signs of the intensity derivatives along lines close to the maximum specular reflections of the apple. The locations of the apples were determined by triangulation. The estimation errors for fruit thinned trees were less than 3 %.

## 2.2 Air sonar fundamentals

The studies above dealt with the development of an automatic crop assessment system based on thermal and optical sensors. Since the approaches suggested in these studies have not been matured to a commercial technology, in this thesis I study and promote the development of an automatic crop assessment system, which is based on air sonar. Hence, the basic ideas of air sonar systems are important for the understanding of the following cutting edge research and for the understanding of this work.

Kleeman and Kuc (2008) defined ultrasonic sensing as the use of acoustic energy propagation at higher frequencies than audible sound to extract information about the environment. This section talks about the basics of in-air ultrasonic transduction, attenuation, Speed of Sound (SoS), Signal to Noise Ratio (SNR), artifacts and the important directivity patterns. To this end, I review in this section the related literature on air-sonar.

### 2.2.1 Transducers

Ultrasonic transducers convert pressure waves into electric signals and vice versa. They differ in their operation from regular microphones and speakers in their nominal working frequencies. While regular speakers and microphones are designed for the audible hearing range of 20[𝐻𝑧] to 20[𝑘𝐻𝑧], ultrasonic transducers are designed to work at higher frequencies.

Two of the most common ultrasonic transducer technologies are based on the piezoelectric and electrostatic principles. These methods are exercised for both receivers and transmitters. Generally piezoelectric transducers have a larger quality factor but lower bandwidths (Nakamura 2012).

Electrostatic transducers implement the basic idea of a changing capacitor with movable plates. A constant electric charge is kept in the capacitor either by using permanently embedded static charged materials or by using a DC bias voltage. When applied as a receiver, the diaphragm which serves as the dynamic part of the capacitor, bends due to the vibrations of the air. Its distance to the stationary part called the back plate changes and the voltage varies according to:

𝑄 = 𝐶𝑉. (1)

where 𝑄, 𝐶, 𝑉 are the electric charge, capacity and voltage accordingly. Due to the constant electric charge, the voltage change is synonymous with the sound vibrations.

The standard loudspeaker belongs to a family of dynamic transducers. Dynamic transducers while similar to electrostatic ones in their design relay on the principle of magnetic induction. A constant rigid magnet interacts with a moving induction coil connected to diaphragm to create shifting oscillations in air pressure.

Piezoelectric materials are ferromagnetic materials which change their polarization under pressure. This relation is described by elasto-electric matrices which relate the stress, strain tensor to the electric fields that occur in the material. This phenomena is used "as is" in acoustics to transform the pressure oscillations in the medium of operation to a matching oscillating voltage on the piezoelectric material.

A main issue in using piezoelectric transducers in air is the low transmission coefficient resulting from the boundary between the piezoelectric material and the air due to a mismatch in the acoustic impedance order of magnitude (Nakamura 2012). The transmission coefficient represents the amplitude of the pressure wave passed at the interface of the materials.

Signal processing has improved the modern piezoelectric transducer greatly. A technique to overcome the poor SNR for distance sensing is to send a wideband CHIRP (Compressed High Intensity Radar Pulse**)** signal and cross-correlating it with the received signal. Another method to face the SNR problem is to try to fix the mismatch in impedance. This is done using new piezoelectric materials and by trying different material layers as a mid-layer between the air and piezoelectric material which have an intermediate impedance.

### 2.2.2 Speed of sound

In the medium of air, the SoS changes mostly due to atmospheric characteristics. Eq.(2) is derived from the ideal gas equation under the assumptions of a quasi-static, adiabatic thermodynamic state for air gas constants. The equation is accurate to 1% under most conditions. (Kleeman and Kuc 2008):

(2)

Where 𝑇𝑐 is the environment temperature. In the humid greenhouse environment, the SoS can be accurately determined by:

(3)

Where 𝑇𝑐 is the environment temperature in ℃, ℎ𝑟 is the relative humidity and 𝑐𝐻 is the SoS which under normal circumstances (dry air at 20 ℃) is taken as 343[𝑚/𝑠] (Siciliano and Khatib 2008).

Since the SoS is closely constant in the greenhouse environment, the Time of Flight (ToF) is proportional to the distance that a pulse travels to an object 𝑟0 and echoes back to the receiver as shown in:

(4)This simple relation combined with the low cost of sonar transducers explains their popularity in the field of robotics as simple proximity sensors.

### 2.2.3 Attenuation

The pressure magnitudes of ultrasonic signals are decreased in air due to two main mechanisms.

Viscous losses between particles turned to heat and to vibrational and rotational relaxation in the molecular level. The relation between pressure amplitudes in the medium of air is approximated by a power law (He 1998) in:

. (5)

is the atmospheric absorption coefficient, the frequency and the range between the points. The attenuation increases with frequency and varies greatly due to environmental changes in pressure, humidity and temperature. These environmental changes affect the air molecules relaxation mechanism of the signal. The values for vary around to at (Kinsler 2000, Jakevičius and Demčenko 2008). For comparison, the acoustic attenuation in water is four orders of magnitude lower, on average, therein lies the reason for the technological difference between in-air and underwater / medical technologies.

The relation in Eq.(5) is often used in sonar applications to normalize the amplitudes of echoes with range.

Ground attenuation, also known as the ground effect, is a result of the sounds reflected by the ground surface, which are interfering with the sounds propagating directly from the source to the receiver (ISO 9613-2). Porous ground such as in the greenhouse environment makes for a small ground effect in relation to hard ground as it is more easily penetrated by the sound waves which are absorbed by it.

### 2.2.4 Signal to noise ratio

The SNR in the bat's air sonar is described by the sonar equation:

𝑆𝑁𝑅 = 𝑆𝐿 − 2𝑇𝐿 + 𝑇𝑆 − (𝑁𝐿 − 𝐷𝐼). (6)

SL is the source amplitude level, TL the transmission loss in air due to attenuation and spherical spreading pattern of the beam, TS the target strength is the amount of acoustic energy returned from an object, NL the noise level and DI the directivity index is the reduction in noise due to the narrowing of the angle of view of the bat's ear. Møhl (1988) emphasized that the equation does not take all the parameters in account and should be considered as a check list of factors that influence the SNR.

### 2.2.5 Directivity patterns

Similarly to the bats SNR, the ultrasonic transmitters which determine the source level in engineering applications have a directivity patterns that result in transmission loss due to their spherical spreading. A simplified model of the beam pattern of a sonar emitter is modelled in (Kinsler 2000) with the assumption of a circular piston with a radius 𝑟 in an infinite baffle vibrating with a frequency 𝑓 with sound spreading only in the plane. The wave length is then represented as in:

(7)

and the angular wave number at Eq.(

(8)

If 𝑟 > 𝜆 the model predicts a pressure field for a distance and an angle 𝜃 that follows:

(9)

Where 𝛼 is a proportionality constant that takes in account the density of air and the source level, 𝐽 is the first order Bessel function

According to Eq.(9), the sound pressure spreads differently with respect to frequency with a main lobe that is described by Waters (2007) to scatter at an angle of 2𝛽 defined in:

(10)

Even for this simple case and despite of the complexion of the theoretical model, in reality it does not completely hold as the assumptions are too lenient. To get a realistic beam spreading pattern an experiment should be held. The general conclusion from this model, which is unavoidable, is that sonar sensors have bad directional resolution because of their divergence. The directivity pattern is not necessarily a bad thing. Fig.(1) illustrates how different frequencies react to the same geometry differently. This phenomenon is taken advantage of, for object classification in this work.

|  |
| --- |
| Beam pattern of a 10 𝑚𝑚 diameter transducer, producing a 40[k𝐻𝑧] signal (top**)** and a 100[k𝐻𝑧] signal (bottom) [ published by (Waters 2007)].  Figure **1**: Beam pattern |

## 2.3 Air sonar applications

In Sec.(2.2) the fundamentals of air sonar were introduced. This section deals with updated research about air sonar from improving the directivity and understanding it up to air sonar classification.

### 2.3.1 Improving directivity

To bypass the weakness of sonar directivity patterns and bad directional resolution described in Sec.(2.2.5), several researches were executed with the intent to increase sensors directivity by adopting a biomimetic approach.

Schillebeeckx, De Mey et al. (2008) have studied the echolocation system of bats by mimicking their pinnae. The echoes from different azimuths and elevation angles with different type of receivers and artificial pinnae were compared to one another and to the returns of standard receivers/transmitter setups in a range of frequencies. The directivity pattern of the sonar sensor which in the context of bio mimicry is called the HRTF (head related transfer function) has shown big differences in energy returns at different angles with and without the pinnae. The changes that were detected between the different receivers were also noticeable. In the second part of their work a comparison measure was defined between the absolute values of the HRTF in different configurations. This measure was used to find the difference between the monaural patterns of azimuth 0° as a function of elevation angle and frequency. The bigger the differences are the better the elevation detection will be. The results have proved that specific pinnae/receiver configurations are more suitable for elevation estimation.

Steckel, et al. (2011) have compared an array of 32 microphones using standard beam forming techniques (filtering the interactions of acoustic waves and then summing the amplitudes) to two higher quality microphones mounted to artificial pinnae. The comparison was done by a measure that considers the additive white noise and reflectors filtering for determining the amount of directivity information loss. The main conclusion is that both systems perform as well, but the development of a microphone array that could also incorporate control over its orientation might prove useful in the understanding of echolocation.

Linda and Manic (2011) presented the SOFAMap algorithm that was developed to offer an alternative solution to the traditional rotational multi sonar arrays that suffer of weak directionality and of low update rate. In order to improve the resolution of the sensor system, a neural network was implemented. The distances and angles were mapped on two kinds of neurons: base neurons that give the outline of the environment and represent the lowest resolution and adaptive neurons that update their quantity and position according to an accumulated local error. The speed and turning rate of the mobile robot are incorporated in the weights of a fuzzy controller so that the input resolution is improved by past samples. The haptic augmentation is implemented by directing a force opposite to the obstacles. The SOFAMap algorithm improves the operator's awareness of unstructured environments, making it applicable to wide range of mobile robot teleoperation systems.

### 2.3.2 Time of flight classification

Another drawback related to sonar sensors is reflections of sound waves created by interactions of the emitted signal with smooth surfaces such as walls. Such surfaces operate as 'acoustic mirrors'. This characteristic creates a problem in most applications as the beam is reflected from a wall resulting in identification of objects in the wrong place. These objects are referred to in sonar slang as 'artifacts'. However, some researchers exploited this problem for feature extraction of walls and edges. ToF classification relies on the distance differences in sensor arrays to calculate the probability of an indoor feature by simple geometrical means.

Akbarally and Kleeman (1995) classified a 16 different common 3D indoor features such as horizontal line corner and vertical line edge, using a 3 transceiver and 2 receivers array and an extension of a previous 2D sonar detection model as described: The 2D model classifies plane features by estimation of the bearing angle to the plane. The angle estimate and plane assumption are validated by the distance the sound waves pass to a second transmitter in a known location. If the estimate is close to the direct observation then the object is classified as a plane feature. By changing this method corner features can also be detected. The 3D model is an extension of the 2D model which assumes that in the indoor environment planes are usually horizontal and vertical to each other. More features can be detected such as edge features calculated from intersections of two planes and point features that are intersections of three planes. Only vertical\horizontal walls orientation assumption are possible, otherwise the complexity of detection rises and the accuracy decreases.

The conducted experiments show highly accurate results for indoor environment but the assumptions made are unreasonable for the greenhouse environment. (Lim, Kwon et al. 2012) identified edges and planes by using the ToF information from one ultrasonic transmitter and 5 receivers. Their algorithm calculates the range and bearing angle for plane and edge feature types using the data from each receiver separately. The average and standard deviation from the different receivers is calculated for both feature types. The system classifies an object by checking which of the orientations/objects the input fits best (the error is smallest). The feasibility of this sensor system was shown to work through simulation and an implementation was in progress.

### 2.3.3 Pseudo amplitude based classification

Kuc (2001) suggested Pseudo Amplitude Scan (PAS) information could help in identifying objects by echoes. PAS sonar is the use of a conventional sonar ranging module to acquire a measure of the returning echo amplitude. The conventional use of a sonar ranging module, sends a sinusoidal signal and awaits for an integrator of the rectified return to pass a certain threshold which indicates the echo has returned. The ToF is retrieved by counting the time to this threshold. PAS resets the ranging module after an echo is detected, and if the amplitudes of the return are strong enough the detection threshold is passed again.

Kuc (2004) illustrated an example by using PAS on a moving platform to scan two types of trees. The echoes were processed off-line, the strong returns from PAS sonar were compared to a simple model of the distance variation to obtain an accurate orthogonal passing distance from the trees. The comparison has shown that the size of sequences of readings close to the trees defined as glints change according to the smoothness of the material scanned.

lvarez, Kuc et al. (2011) used a PAS sonar which is able to operate at different frequencies to scan 7 types of fabrics. The fabrics were different in their porosity and material which has resulted in distinct acoustic impedance at different angles. The different frequencies have enriched the data as mentioned in Sec.(2.3.1). The fabrics were scanned from a range of angles and by discrete amount of frequencies. The PAS data for each frequency and angle was structured into 126 elements feature reduced to a 2D vector using Principle Component Analysis (PCA). The 2D plot of the data is divided into regions corresponding to types of fabrics and a NN was trained to obtain a 80 % classification success rate.

### 2.3.4 Spectrum based classification

Spectrum based classification of targets is the attempt to discriminate targets by the relation between frequency and energy retuned from it. In order to further inquire this relation the recorded data must contain information in a spectrum of frequencies. This type of method was already mention in Sec.(2.3.3) (lvarez, Kuc et al. 2011) for identifying fabrics in frequencies between 40[𝑘𝐻𝑧] to 65[𝑘𝐻𝑧] with jumps of 5[𝑘𝐻𝑧].

A common method in sonar spectrum based classification is to analyze objects response at a larger range of frequencies. Ideally an impulse excitation would be the best at testing a wide spectrum of frequencies. In practice it is impossible to emit a real impulse signal with a loudspeaker or any other real system. Therefore Müller and Massarani (2001) concluded that using CHIRPs is the most advantageous choice in terms of SNR for testing room impulse response (for characterizing room acoustic parameters) – due to the long duration extinction and high extinction energy contained in CHIRPs without the need to increase the peak power.

The recorded ultrasonic data is best represented by spectrograms (time-frequency-energy \ space-frequency-energy graphs) or by their instantaneous version, the Power Spectrum Density (PSD) graph. The recorded echo is usually transformed using the Fast Fourier transform (FFT) (Welch 1967). Cowell and Freear (2010) suggest a solution to a common problem with CHIRP signals called "time overlapping" which occurs when the duration of the signals is longer than the delay between echoes. The authors introduce a way to use the Fractional Fourier transform (FrFT) in order to separate the signals. The FrFT is a transform to a mixed time/frequency axis of which the FT, inverse FT and identity transform are unique cases of. After applying the FrFT the authors use a Hann window to isolate the signals followed by a concatenation of an Inverse FrFt and a normal FFT to get a PSD of the separated signals. The authors have implemented both a simulation of the method and an experiment which obtained good results in separating the signals from background noise despite its changing frequency.

Several works applying spectrum based classification in plants were published. Most of them with an intent to give a better understanding of echolocation by this Method: McKerrow and Lindsay Harper (1999) in their research present a way to extract geometric features of plants from the echo return and a machine perception method to differentiate between different plants. The authors claim that while vision methods have succeeded in recognizing plants in controlled environments, some of the problems encountered in them can be solved by a sonar machine perception system. It categorizes the plants by criterions that are easy to extract from the echo, invariant to orientation and reorient to a defined geometric characteristic of the plant. This way, plants can be recognized independently of their height, width and the distance to the system. Examples of properties that change the echo are: size, spatial orientation, number and orientation of leaves. Most of the acoustic features are influenced by multiple geometric features and vice versa which makes it difficult to find an inverse transformation from echo to plant geometry.

A fuzzy interference system was used to discriminate between the different plants. Müller and Kuc (2000) aimed their research at understanding bio sonar by analyzing the returns from foliage and extracting invariant features. An experiment with two plant species has been executed using a robotic arm to allow a specific orientation for each of the two plants. The data was processed to extract various echo properties such as total energy, max amplitudes, and impulsiveness. The echoes were found to be stochastic in nature so that the authors tried to estimate their probability density using Gaussian kernel estimators. A number of invariant features were ultimately discovered and are described together with the level of their confidence.

Yovel, Stilz et al. (2009) emphasize the important role of the vegetation echoes as the most encountered returns by bats. Studies have shown that bats change their behavior when nearby plants are covered and react to different plants. Because of this major role, an important step at understanding and deriving more data from sonar sensors would be to statistically analyze these echoes. Because of the complexity of a plant geometry and the width of the sonar beams it is futile to try to locate single reflectors and thus the analysis has to relay on statistics. Most of the energy returned from a plant is concentrated in proximity of the main stem where the density of leaves that act as the main sound reflectors is high. A Poisson model of the reflector (leaves) distribution and a more complex clustered Gaussian reflector "clouds" model, spread apart by a Poisson distribution were used to verify that peaks of energy in the Power Spectrum Density (PSD) correspond to the reflector distances. Plant models designed according to these statistical models were created and compared the echoes returning from them to the echoes returned from the real plants by using the Kullback-Leibler goodness of fit measure. The PSD graphs of the different plants were used to characterize the periodic structures of the plants because as mentioned before the energy returned for each frequency is depended of the geometry of the items. Leaves reflect most of the energy in plants but their spatial organization was found to be defined by branches that make up clusters of leaves which show in the PSDs as large scale periodic structures. In the PSDs some common periodic structures can be seen from any direction the plant is ensonified and are the consequence of common periodic distances of reflectors. The differences between 4 species of trees are also shown and how echoes returning from conifer trees show exceptional behavior in relation to the other three trees. The effect of a natural bat like emitter and receiver were also tested and showed higher energy return at frequencies more relevant to the bat's echolocation.

Balleri, Griffiths et al. (2012) stated that plants and their flowers represent a combination highly suited to the study of target classification for they are motionless and silent and habitat in densely cluttered environments (bats rely only on their own CHIRPs for localization). The authors decline the notion that bats use their sense of smell in near environment localization by presenting the example of the Mucuna holtonii plant – which releases the biggest amount of nectar on the first visit by a bat, behavioral experiments have shown that bats successfully detected unvisited flowers in darkness even though the scent of the flower remains unchanged before and after a visit. Studying the evolution of the pollination relationship between the bats and pollinated plants could turn useful in understanding the classification process using sonar. The data collected is the return from a 50[𝑘𝐻𝑧] to 250[𝑘𝐻𝑧] signal transmitted at the flowers from different orientations and ranges. The differences between pollinated and non-pollinated flowers were very clear both in the magnitude of the PSD's and graphs of magnitude vs range/angle. The power reflection from the flowers was tested and showed that full flowers returned the signals better then desiccated or modified flowers. Another aspect of the paper was classification of the visited flowers using two standard classifying algorithms: the k-nearest neighbors and naïve Bayesian classifier. Both algorithms were tested using a training data set derived using principle component analysis in order to reduce the data's dimensions. The k-Neural Networks (NN) algorithm gave better results, that improved by changing the resolution of the scans to up to 90% success rate. Another interesting conclusion was that most of the valuable data for classification of the plants can be derived by scanning only from a limited range of angles.

## 2.4 Air sonar in agriculture

There are only a few studies about air sonar classification in agriculture engineering. Andújar, Escola et al. (2011) presented a method for classifying three different plant types: broad leaved weeds, grass and the mixture of both. An ultrasonic sensor was directed downwards to the ground and can detect the height and biomass parameters of the plants and identify them. A static and dynamic experiment of both ultrasonic and manual measurements was made in two different dates to be later examined. Analysis of Variance (ANOVA) and linear regressions were used to decide which parameters are most important for discrimination of weeds (biomass, height) and how to classify different mixtures of plants. The success in prediction of weeds was better for the first date thus it is recommended by the authors to focus on early stages for weed discrimination. Main results show 80% success for detecting pure grass and 99% for broad leaved weeds using this low cost, robust ultrasonic system.

Sonar spectrum classification has also been used for classifying plants as landmarks for navigation such as Harper and McKerrow (2001) which have used their classification algorithm presented in sec. (2.3.4) and defined the different correlations and obstacles for a plant spectrum based navigation strategy.

Toda, et al. (1999) created a Fuzzy Logic Control (FLC) based robot using ultrasonic sensory to guide itself in lab simulated maize crop rows. The mapping process deals with sonar inaccuracies such as difference in temperatures and outliers caused by leaves that point out of the plant using simple statistical analysis. Data from the sonar array is filtered from outliers and used to derive the perpendicular distance from the middle of the robot and the heading angle by geometry. A training set was made for the FLC with 8 fuzzy sets concerning the angle and corresponding driving times and 16 linear fuzzy laws that they abide. Using least squared method the constants for the different laws are set and the FLC is able to translate the inputs (angle, distance) of this complex nonlinear system (mainly complex for the slips) into smooth directions for the robot so that the errors in finishing locations are small (0.03[m] and [3°]).

## 2.5 Sonar Simultaneous Localization and Mapping

Simultaneous Localization and Mapping (SLAM) is the acquisition and combination of acquired map data about an unknown environment and the process of keeping track of location in the acquired map while producing the navigation track. Sonar SLAM in contradiction to the navigation methods presented in the previous section does not necessarily involve the use of sonar sensors for spectrum based classification.

Elfes (1987) used a mobile robot with a ring of sonar distance sensors for one of the early works on SLAM. The robot scans the environment using its sonar array and the readings are interpreted to "empty" or "occupied" areas and serve to update the sonar map. To reduce error and uncertainties the next readings are integrated into the 2D map based on position and orientation using a probabilistic approach in order to get the absolute location and to identify landmarks. The software deduces from the sonar data three more data structures that help at different problem solving activities: 1. Abstraction axis- includes full data from sensors, a map with high certainty interesting points and a node map representing these areas. 2. Geographical axis: divided into local views that construct the basic building blocks of the map. Local views make local maps and these make up the global map. 3. Resolution axis: the mapping starts with the full data set (sensor map) and is decreased to ease computations. Finally A\* navigation is used to search the map with a cost function that considers the certainties of obstacles.

Kleeman (2003) introduces a mobile robot with two encoders connected to its wheels and two sonar transmitters that are attached to rotating pans and is controlled manually by a joystick. The sonar pans rotate in order to complete scans of the indoor environment. The range, and bearing from the scans is then clustered into the system's memory based on position with feedback from odometry. The statistics of the clusters is used in order to reject clutter inaccuracies. SLAM is implemented using an extended Kalman filter with an odometry error model based on his previous works and a sonar error model that takes in account vibrations and clutter tending sonar features described in the ToF classification section like edges.

A different fuzzy logic approach was used in Farooq, Abbas et al. (2012). Two simple sonar ultrasonic sensors stationed in a constant 45° angle from the front of the robot. The differences between the readings show the offset of the robot from the central axis of the corridor. The idea behind this approach is to avoid the creation of a kinematic model of the robot that could be non-linear or just very hard to derive. The authors implemented a zero order Sugeno FIS on a micro controller to control the robot with five trapezoidal input membership functions and five rules and two outputs for the motors. The experiments show good results for the four different corridor types the system encountered.

Schillebeeckx and Peremans (2010) state that air sonar systems have been abandoned in most works for localization tasks and are used mostly for obstacle avoidance only. Bats orientation has proved that there is much quality information to be extracted from air sonar. Their research focused on localizing reflectors using one emitter and two receivers. The receivers are Omni-directional thus an artificial bat pinnae was used to get directivity characteristics. The corresponding values from the response spectrum close to the local maximums of the amplitude/time graph are extracted and called the 'spectral code'. The left and right receiver spectra are a function of distance, elevation and azimuth of the reflectors to be localized. A database of the spectral codes of the reflector from a large interval of azimuth and elevation is created. The system uses the maximum likelihood function of the azimuth and elevation to find the closest spectral code to the current orientation. The same experiment was repeated with three reflectors and three databases for each of the distances. The overall results are mixed, in some of the cases the error is lower than 3°, and higher in others. Future research should incorporate a combination of consecutive measurements by changing the pinna configurations (like a bat that moves its ears).

Schillebeeckx and Peremans (2010) state that in early SLAM work sonar was used to create topological maps but nowadays other higher resolution sensors are used for this purpose. The authors argue that the noisy picture resulting from today's robotic sonar sensors are the results of misusing them as simple range detectors. The researchers fused a 3D biomimetic sonar sensor and a SLAM algorithm called RatSLAM which uses the data in a biologically inspired way, an innovative approach in relation to the existing probabilistic approach.

The original RatSLAM (Milford, Wyeth et al. 2004) is based on the pose cell structure, the local view structure and their interactions. Pose cells are interconnected cells that are associated with a specific robot position in the environment, given by its coordinates. An action, named activation is defined as triggering cells. A cell is triggered if either it is similar to other cells or if the interaction between the pose and the local view, called experience is similar to another experience. Activation of a cell activates the cells in a bounded neighborhood. A neighborhood of active cells is entitled a cluster. The cell position is updated using closed loop odometry. The pose cell with the highest activation rate is then used to determine the robot's position. The local view structure is a database consisting of low resolution pictures associated with the pose cells. This database is matched with the current picture and the associated pose cell is activated if the distance from previous to the current estimated pose cells is larger than some threshold. The experience map, a process that builds a spatially coherent node map using the concepts described above, connects between the two spaces. The experience node map contains recent local and pose views connected by time and movement information from one experience to another. The experience map is updated and corrected by considering both the difference between local views and the difference between the robots pose estimation. This way the experience map converges into a good spatial description. The difference between BatSLAM and RatSLAM is that the visual data is replaced with sonar collected data. Instead of regular images, BatSLAM uses 2D local view created from the 3D sonar data called a cochleograms. The data is derived via amplification of the small position differences and subtracting the left and right signals from each other.

### 2.5.1 Sonar and LiDAR SLAM

Tungadi and Kleeman (2012) added a Light Detection and Ranging (LiDAR) to the sonar sensors. While the laser scanner is good for localizing in most environments, at some places like corridors the map it produces looks the same which make it hard to find any differences between different environments for localizing. The sonar sensors compensate for the laser inability to find differences using the features described in (Akbarally and Kleeman 1995, Kleeman 2003) mentioned earlier.

Vornoi maps are used to transform the laser and sonar data into a node tree map. A Vornoi map is the division of the environment into areas where every point is closest to an obstacle originating in that area. A numerical algorithm with complexity of O(nlog(n)) is used to build the graph and the borders of the areas are described as branches while the intersections are described as nodes. The nodes are then used to make up trajectories for the robot. In order to map the environment the robot closes loops in the Vornoi graph to help updating the map and contribute to a more accurate robot orientation. The robot trajectory is planned to travel all the small loops in a map, and only then to travel all the rest of the nodes. Overall the experimental results show a 0.2[𝑚] improvement on both x and y axes in a 300[𝑚2] map for this system in comparison to a frontier based exploration algorithm.

### 2.5.2 Sonar and vision SLAM

An additional approach at sonar SLAM is the fusion of vision and sonar sensors. Jinwoo, Sunghwan et al. (2006) suggested that vision based SLAM, using SIFT measurement, is good for correcting the data association because of the abundant and salient visual information. However, the computational burden associated with this process is high and the SLAM update rate is low. Therefore this article suggests that adding sonar features into the SLAM would enable larger updating rate and could help in conditions of bad lightning. The sonar features to be extracted are point and line features. They are derived using ToF and simple geometrical heuristics and help in deciding whether certain measurements are either line or point features.

The features are collected from a number of samples and the difference in position between samples is compensated by the robot's position obtained by the SLAM algorithm. The object recognition algorithm is based on the SIFT method. A SIFT descriptor database is made and stored before using the object recognition algorithm. The probability of matching groups of descriptors is higher than finding a single feature, hence the authors apply geometric constraints on groups of descriptors to get better matching. The groups are matched using RANSAC clustering. After an object is identified, it's depth is determined by assuming that it is planar and by using a stereo camera to match a number of features and find their distances. Finally SLAM is implemented using an EKF model with the visual recognition given larger weights in determining the position.

# 3 Materials and Methods

Section 3.1 describes the ultrasonic and robotic hardware and their integration with the software. The instrumentation was used for the experiments designed and implemented according Sec.(3.2) and Sec.(3.4) and eventually analyzed using the algorithms and methods described in Sec.(3.3), (3.5) – (3.6).

## 3.1 Instrumentation

The sensor system has evolved via iterations through the research. As each system has its own frequency response it is crucial to use the same settings to get the exact same results. The sensors system has been changed to improve real time signal processing capabilities and improve SNR. It is worth noting that throughout the evolution of the sensor system, apparatus combinations of parts from the first to its final version were used together in experiments. The general block diagram for the system remains:

|  |
| --- |
| **Computer**  **Figure 2**: System’s Block Diagrams |

### 3.1.1 Sensory apparatus 1.0

The preliminary sensing apparatus presented in Fig.(3) is an off-the-shelf recording system (UltraSoundGate 116hm, Avisoft Bioacoustics, Germany) consisting of an Ultrasonic Dynamic Speaker Vifa capable of 10[w] output with a 70[𝑘𝐻𝑧] bandwidth, a CM16/CMP microphone which has a flat frequency response in the range of 10[𝑘𝐻𝑧] to 150[𝑘𝐻𝑧] with a sensitivity of 50[𝑚𝑉/𝑃𝑎], an UltraSoundGate 116 player amplifier with a 180[𝑘𝐻𝑧] bandwidth and digital to analog convertor at sample rates of 16[𝑏𝑖𝑡𝑠], 500[𝑘𝐻𝑧], and an UltraSoundGate 116hm recorder for amplifying the microphone output and sampling at 16 [𝑏𝑖𝑡𝑠], 1[𝑀𝐻𝑧]. In the qualitative experiment described in Sec.(3.2.2) a LiDAR (LMS 111, SICK, Germany) working at 50[𝑘𝐻𝑧] 270° at a resolution of 0.25° 0.012[m] and readings of distances up to 18[𝑚] was used to obtain ground truth range readings.

|  |
| --- |
| (d)  Ultrasonic Dynamic Speaker Vifa **speaker**  (c)  CM16/CMP **microphone**  (b)  UltraSoundGate  116 **recorder** |

(e)

SICK LMS 100 **laser** rangefinder

(

a

)

(a)

UltraSoundGate

116 **player**



**Figure 3**: UltraSroundGate 116hm Recording System

### 3.1.2 Sensory apparatus 2.0

|  |
| --- |
| 𝟔        𝟏    𝟐    𝟓  (A) Forward View   1. Microphone supply and **Amplifier** 2. Ultrasonic **Loudspeaker** 3. Ultrasonic **Microphone**, 4. Loudspeaker **Amplifier** 5. **Camera**     (B) Back View   1. **DAQ Card**.   𝟑  𝟒  **Figure 4:** The Sensing System |

The second sensing system is comprised of two parts encapsulated by 3D printed components: an ultrasonic sensing system and an RGB camera. The main components of the ultrasonic sensing system are a loudspeaker(2) and a microphone(3) in Fig.(4). The speaker (XT25SC90-04, Tymphany, Denmark) has a highly flat frequency response ranging 1[𝑘𝐻𝑧] to 40[𝑘𝐻𝑧] officially and as far as 120[𝑘𝐻𝑧] in practice.

The ultrasonic microphone is electro-static (CM16/CMP, Avisoft Bioacoustics, Germany). The signal to the ultrasonic loudspeaker has a custom electric circuit designed op-amp amplifier (PA12, Apex, USA) with a 2× to 7× gain and 160[𝑘𝐻𝑧] bandwidth. The microphone has an adjustable amplifier and power supply (CMPA405V, Avisoft Bioacoustics, Germany) supplying 200[𝑉] polarization voltage.

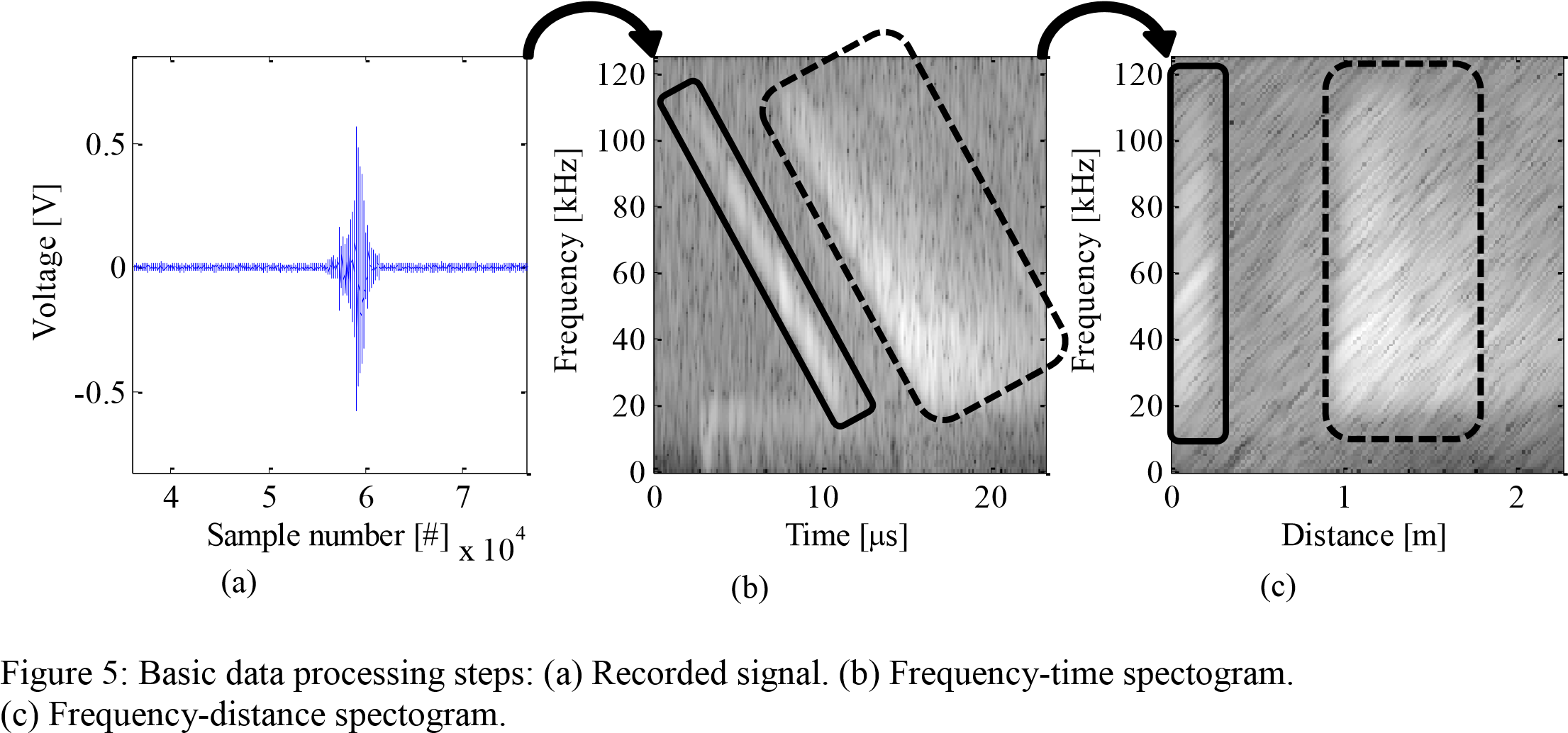
The output to the loudspeaker and the input from the microphone is processed by a Data Acquisition (DAQ) device (USB-6210, National Instruments, USA) capable of 16[bits], 250[kHz] analog input and output sample rate. The RGB camera is used for comparison to the sonar data. It is (HD Webcam C615, Logitech, Switzerland) working at 640 × 480 resolution in trigger mode.

### 3.1.3 Data acquisition apparatus

The data acquisition and signal processing is carried on by MATLAB (Mathworks, USA). For the sensory apparatus 1.0 system, the data was recorded continually to a 1 second ring buffer WAV file by USGH Recorder software (Avisoft Bioacoustics, Germany). The WAV file was sampled by MATLAB whenever it triggered for a new sample. This method has slowed down the real time capabilities of the system and affected the recorded data forcing a new solution for data acquisition. Using the NI-6218 DAQ together with the Data Acquisition Toolbox™ has granted the capability to sample the ultrasonic data directly to MATLAB at 10[Hz].

|  |
| --- |
|  |

### 3.1.4 Software



(a) Recording Signal (b) Frequency-Time Spectrogram (c) Frequency-Distance Spectrogram

**Figure 5**: Basic Data Processing Steps

As a multi spectral analysis approach was chosen for this work, the first step in signal processing is to acquire the ensonified objects frequency response. The frequency response is derived by transmitting an ultrasonic linear CHIRP (marked by a solid rectangle in Fig.(5𝑏) at objects and comparing it to the attenuated return echo (marked by a dashed rectangle) as following: The ensonification is recorded Fig.(5𝑎) and transformed into a spectrogram using the short time Fourier transform algorithm with a rectangular window of 100 time-steps, a 50 time-step overlap between windows and 256 frequency bins Fig.(5𝑏). The time axis is transformed according to the frequencies ensonification time so that the spectrograms horizontal axis represents distance from the sensor Fig.(5𝑐).

The CHIRP and echo spectrograms can then be extracted separately from the spectrogram and analyzed by operating over the frequency or distance dimensions using statistical, image processing or signal processing techniques. The frequency response is defined in this thesis as the sum of the echo amplitudes in a specific frequency divided by the mean CHIRP amplitude of the same frequency.

For in depth analysis, the detection of the return echoes is implemented automatically. This is done by cross-correlating the frequency-distance spectrograms with the emitted CHIRP. This cross-correlation is also known as a matched-filter, an optimal filter in the SNR sense. In the case of comparing two identical signals separated by a lag in time, the matched filter will peak in the ToF of interest. The implementation of the matched filter in this work is unique as it is done in the frequency-time domain of the spectrogram. This is not as efficient as in the time domain implementation but helps in visualizing noisy frequencies and allows the operator to correlate parts of the frequency range without any filtering additions and generates the spectrogram data so it is always accessible. The algorithm is attached in the Appendix.

### 3.1.5 Robotic platforms

The sensor apparatus was mounted in experiments on static holders. It was also mounted on two robotic platforms: one is a greenhouse sprayer (Chico , Degnia Sprayers, Israel) Fig.(6.1) with a 18 HP petrol engine, and 4𝑥4 articulated steering. The manual steering system was replaced by an autonomous windows PC, driver (AX 3500, RoboteQ , USA) and microcontroller (Arduino mega 2560, SmartProjects, Italy) based control system in the Agricultural Robotics Lab (ARL), ARO. The sensing system is held by a 3 axis gimbal (Ronin, DJI, China) capable of rotating at 0.25[rad/s] and controlled through a PWM to SBUS converter (RMELIC, China). The second robotic platform is 6 degrees of freedom manipulator Fig.(6.2) (VP-6242G, Denso, Japan). In all of the above cases, custom made end-effectors were designed to minimize the distance between the ultrasonic transducers.

|  |
| --- |
| 𝟑    𝟏    𝟐    **Figure 6:** Robotic Platforms     1. Chico (2) Denso manipulator (3) Ronin   Need  (A)    (B) |

## 3.2 Experimental setup

Three experiments were conducted: i) Single object experiment to determine the frequency response of single leaves and fruits – the basic building blocks of the full plant echoes Fig.(7a); ii) A qualitative greenhouse experiment to classify different objects such as plants, crop rows, common objects and greenhouse infrastructures in real environment using the ultrasonic sensor shown in Fig.(7b), and iii) A prediction features experiment, in which the influence of fruit mass and leaf number on the frequency response of different plants from different orientations and different distances was measured Fig.(7c).

|  |
| --- |
|  |

### 3.2.1 Single object experiment



(a) Single object (b) Qualitative (c) Prediction features

Experiment Experiment Experiment.

**Figure 7:** Experimental setups

The experiment was conducted at the ARL, ARO. The experimental setup presented in Fig.(7a) is using the sensory apparatus 2.0 system. Five different freshly picked leaves and fruit were ensonified 50 times by 5[ms] time linear CHIRP excitations ranging 5[kHz] to 120[kHz]. The distance to the ensonified objects was 1[m] and the orientations of the objects were randomly changed. In order to isolate their echoes, it was ensured that the fruit and leaves were the only objects in the vicinity of the sensing system by tying them up to the ceiling with a thin thread. The thread was ensonified and confirmed to return weak echoes below the sensing system detecting threshold.

### 3.2.2 Qualitative experiment

The experiment was conducted in the research greenhouse of the ARL at the ARO. Different objects were recorded at different distances and number of crop rows between the object and the sensor apparatus.

The experimental system was set in the following configurations: (a) Sensing system is perpendicular to an empty place in the greenhouse with no crop rows (measures the return echoes from the ground); (b) Sensing system is perpendicular to three successive Capsicum Annuum (bell pepper) crop garden beds; (c) Sensing system is perpendicular to the greenhouse plastic; and, (d) Sensing system is perpendicular to two cucumber. A 10[ms] linear CHIRP ranging 20[kHz] to 120[kHz] was selected as the excitation signal. The transmitting amplitude was set to minimize saturation by changing the speaker and microphone amplifiers gains. 30 repetitions of the acquired objects picture taken with the camera mentioned in Sec.(3.1.2), ultrasonic echoes and distance readings from the LiDAR referred to in Sec.(3.1.1) were taken for each configuration.

|  |  |  |  |
| --- | --- | --- | --- |
| **Distances [m]** | **Exp.(2)** | **Exp.(3)** | **Exp.(4)** |
| **0.5** | √ | **√** |  |
| **1** | √ | **√** | **√** |
| **1.5** | √ | **√** | **√** |
| **2** | √ | **√** |  |

|  |
| --- |
| (c) Greenhouse Wall            (a) Ground  (d) Cucumber  (b) Three Rows    **Figure 8**: Qualitative Expirment Ensonofication Positions |

### 3.2.3 Prediction features extraction experiment

This experiment was conducted at the ARL. Five fully grown pepper plants grown in the greenhouse were carefully uprooted and transplanted in pots, as shown in Fig.(7c) each bearing fruit with a total yield mass of 700[gr] to 1200[gr] per plant. The plants were held up to the ceiling with thin threads to replace the trellising while minimizing any noise to the return echoes caused by anything but the plants. The five plants were ensonified with a 20[kHz] to 120[kHz] in 5[ms] linear CHIRP for three fruit levels: full, intermediate and empty. Each plant at each fruit level was ensonified from three orientations, 120° apart – designated as PiOj where i represents the plant number and j represents the orientation: 1, 2 and 3 for 0°, 120° and 240° respectively.

Each PiOj was ensonified from distances of 0.625m and 1.25m with 30 repetitions. In post processing, a multi linear regression and a fruit mass estimation based on the experiment results was conducted.

## 3.3 Ordinary Least Squares attenuation and hidden row detection

The attenuation for each frequency was extracted using Ordinary Least Squares (OLS) on the frequency response results of the qualitative experiment. Sec.(3.2.2). The attenuation is used to normalize the energy returning from a row by using the attenuation values as weights to compensate for the energy loss due to distance as defined mathematically in Eq.(5). However, it amplifies the noise as well.

To overcome this situation, the noise is filtered prior to amplification. The return energies of the plants are relativity strong in the 40[kHz] to 60[kHz] band, therefore the detection process occurs in that range. Finally, the cross correlation of the CHIRP with the compensated return echo (also known as a matched filter) is used to maximize the SNR and detect the plants rows by thresholding.

## 3.4 Applicative experiment

The goals of the experiment were as followed: 1) map the environment of the robot, 2) classify walls into either plants or building, 3) scan for hidden objects behind plants. The experiment was done in the Robots and Bio-Medical Micro-Systems (RBM2S) lab at TAU. As at the time the experiment was held was off bell pepper growing season, six 0.2[𝑚] to 0.4[𝑚]diameter, 0.4[𝑚] to 0.6[𝑚] height *Vibrnum tinus* (Eve Price) plants were planted in pots and stationed to simulate a greenhouse plant wall as in Fig.(9).

|  |
| --- |
| 1. Styrofoam 2. Eve Price Plants 3. θ is the VP-6242 first joint angle   **Figure 9**: Real Time Application Setup |

The Eve Price specie was chosen for its dense foliage and relative similarly to pepper plants when not blossoming. A 1×1×0.05[𝑚] Styrofoam wall was stationed 0.2𝑚 behind the plants. The system work area was approximately 10[𝑚2]. The sensory apparatus 2.0 sensing system was mounted on the end-effector of the VP6242G manipulator. The end effector was held at a constant height against the center of mass of the foliage. The horizontal distance between the sonar transducers and base was held to a minimal 0.3[𝑚] distances. The manipulators starting position was chosen randomly. The indoor environment (as the greenhouse it simulates) is assumed to consist of only plant and building walls.

### 3.4.1 Scanning procedure

Fig.(10) presents the scanning procedure in the experiment. The procedure starts with applying standard proximity sensing for a high sample rate using the ultrasonic transducers at 40[kHz] with a band-pass filter to improve accuracy. The difference between samples is the change in the VP-6242G first joint variable θ. Utilizing ToF classification methods such as (Akbarally and Kleeman 1995, Lim, Kwon et al. 2012) the next Δθ movement is determined by fitting all n − k combinations from the last n samples into line regressions using Total Least Squares (TLS) method for detection of walls. Δθ is defined as the angle from current θ which would position the sensing system perpendicular to the best fitted line by TLS. The perpendicularity of the sensing system is important for classification as the data returning from the wall is the richest at that position as implied by the sonar Eq.(6).

The LDA classification is done over the frequency response vectors of the ensonified objects. The training set of the LDA classifier is composed of twenty samples of building and plant walls from random orientations and distances. In case the sensing system detected a plant wall it is scanned for a hidden object behind it by detecting energy peaks above a predetermined threshold in the time energy against time graph. If there is more than one energy peak (the first energy peak is the plant wall) the manipulator starts a sub routine of three 0.1m advances towards the object while validating the distance from the base of manipulator is the same. In case that the object has been validated three times the distance to the hidden object is printed. The procedure is stopped when the full perimeter of the room is classified to either building or plant wall. A workflow of the scanning procedure is presented in Fig.(10).

|  |
| --- |
| Legend:   * true * false       Figure : Flow chart of the real time classification and hidden object detection Algorithm |

## 3.5 Statistical tools for results analysis

Deducing from (Bloch, Dgani et al. 2013) ,the sum of energy of a plant's echo per frequency bin, contains information that is assumed to be unbiased by fruit distance to the sensing system [see Sec.(1.4)]. This individual frequency bin sum of energy returned from 𝑃𝑖𝑂𝑗 is denoted as 𝐸𝑖𝑗. To examine the hypothesis that 𝐸𝑖𝑗 is a monotone function of the fruit mass, 𝑚𝑖𝑗 and influenced by the ensonified frequency , a multiple linear regression model is defined. The actual physical relationships between the listed factors expressed in 𝐸𝑖𝑗 = (𝑚𝑖𝑗, 𝑓𝑖𝑗, 𝑥𝑖𝑗), is theoretically intractable. Where 𝑥𝑖𝑗 is unknown and represent specific factors of 𝑃𝑖𝑂𝑗 that are unobserved since the geometry of pepper plant is complex and specific for each specimen. However, given that (𝑚𝑖𝑗, 𝑓𝑖𝑗, 𝑥𝑖𝑗) is 𝑛-order differentiable, the Taylor theorem can be applied to establish a 𝑛-ordered polynomial approximation to the above relationships and estimate its parameters as a regression model. Eq.(11) shows the second order approximation of (𝑚𝑖𝑗, 𝑓𝑖𝑗, 𝑥𝑖𝑗) around , the mean mass and mean frequency respectively. For compactness reasons, and its derivatives are marked

|  |  |  |
| --- | --- | --- |
|  |  | **(11)** |

Assuming that the second order approximation is satisfying, eq. (11) leads to a linear regression model in:

|  |  |  |
| --- | --- | --- |
|  |  | **(12)** |

Assuming 𝜖𝑖𝑗 is the model's random and serially uncorrelated error, the model parameters denoted 𝛼, 𝛽, 𝛾 are estimable via OLS. 𝛿𝑖𝑗 are indicator variables defined in eq. (13) and are assigned to address and quantify the fixed influence of 𝑥𝑖𝑗.

|  |  |
| --- | --- |
|  | **(13)** |

Three estimation techniques based on the regression model were tested:

1. Solving the quadratic Eq.(14) for the mass term, an estimation of fruit mass depended in the specific specimen terms is derived (see Eq.(5)). As is unknown for new plants, the estimation function in its present form is of little use.

|  |  |  |
| --- | --- | --- |
|  |  | **(14)** |
|  |

1. The derivative of with respect to the frequency and mass is of special interest. The indicator variables fall by differentiating with respect to , which allows the isolation of an expression for mass using the regression model parameters as in Eq.(15). This comes with a cost of noisy numerical differentiation but it allows the prediction of fruit mass with the measured energy regardless of the unknown specific specimen factors .

|  |
| --- |
| (15) |

1. Minimizing the squared errors between measured energy and the quadratic model by differentiating over the predicted quantity. The roots of the equation minimizing OLS are the number of leaves/mass closet to the model:

|  |  |
| --- | --- |
| |  | | --- | | (16) | |

## 3.6 Neural Networks estimation

A linear regression model with the frequency response amplitudes as inputs was used iteratively to determine frequencies which are statistically significant to the mass of fruit and number of leaves:

|  |
| --- |
| **(17)** |

|  |
| --- |
| Figure  11  :  **Figure 11**: A view of the NN structure. |

Where 𝑚𝑖𝑗 is either the fruit mass or number of leaves, is the amplitude of the frequency response at frequency 𝑓𝑖𝑗, 𝛽𝑓𝑖𝑗 is the parameter fit by the regression and 𝜖𝑖𝑗 is the error.

These 18 frequency amplitudes were chosen as the inputs for a six hidden layer feedforward NN approach as depicted in Fig.(18).

The network was trained by the Levenberg-Marquardt backpropagation algorithm to predict the number of fruit and leaves. 2 samples out of 15 PiOj were used as test data for cross validation, the rest of the dataset was divided randomly so that 85% of the samples are used for training and 15% for validation. The average and standard deviation of the slope, intercept and R2 of the 2 test data sets were recorded and displayed.

# 4 Experimental Results

The results are separated to three conceptual subsections. The first deals with findings of qualitative nature, regarding the single object experiment and qualitative experiment for extracting ultrasonic features in the greenhouse environment. The second subsection focuses on results which are more applicative dealing with real time implementation of classifying plants from walls and detecting objects occluded by foliage with the sensor system. The third subsection focus on the analysis concerning the prediction of fruit mass and leaf number by ultrasonic means.

## 4.1 Qualitative experiments

### 4.1.1 Attenuation compensation

The returning echo amplitude data from the pepper rows at four different distances was extracted from the qualitative experiment Sec.(3.2.2) and fitted according to the power law in Eq.(5). The fitting was done by applying OLS method over the natural logarithm of the equation for each frequency. The adjusted attenuation is shown in Fig.(12):

|  |
| --- |
| **Figure 12**: Attenuation values according to frequency. |

The attenuation values above are higher by order of magnitude than the presented at Kinsler (2000) and [𝑑𝐵/𝑘𝑚] in the ISO 9613-1 standard. While the different temperature and humidity do not justify this difference as seen in Jakevičius and Demčenko (2008). These values are acceptable as Kinsler (2000) warns that the experimental values of the attenuation for gases can be significantly different than theoretical attenuation. Moreover, the theoretical values and the ISO standard attenuation formula are suited for pure tone frequencies, factors such as transient effects due to the CHIRP excitation and returns from the ground are ignored.

While adapting to these values is possible by applying passband filters it is not important for the scope of this work. As the purpose of this section is to calculate the right attenuation values that would compensate for small distance differences echoing from foliage, it is calculated differently than other works which measure the single tone attenuation directly from transmitter to receiver from hundreds of meter distance.

It is evident that [𝑑𝐵/𝑘𝑚] values are in general rising monotonically except for the 88[𝑘𝐻𝑧] to 97[𝑘𝐻𝑧] and 50[𝑘𝐻𝑧] to 70[𝑘𝐻𝑧] bands as expected from theory.

The return echoes can be compensated with the resulting 𝛼𝑎𝑏𝑠(𝑓) under the condition that noise at far distances are cleaned or else they will be amplified greatly as their amplitude will be compensated exponentially according to distance.

### 4.1.2 Ground effect

The sensing system was posted perpendicular to an empty area in the greenhouse with no crop rows Fig.(13𝑎) and measured the return echoes from the ground. Fig.(13𝑏) shows the averaged ground mean echo derived from 30 spectrograms.

The spectrogram indicates that there are substantial energy returns from a distance of 0.72[𝑚] up to 2.2[𝑚] away from the sensing system. At this range the echo is strong for frequencies between 20[𝑘𝐻𝑧] to 60[𝑘𝐻𝑧].

|  |
| --- |
| **Figure 13:** Ground Effect  (a) View from the sensor of an empty space in the greenhouse.  (b) Ground mean spectrogram  (a)  (b) |

The source of these returns could be only the ground, which comply with the cone shaped main lobe broadcasting characteristics of the speaker. Therefore, the ground signature should be considered in any implementation of the sensor system. The ground echo was cleaned from the next spectrograms assuming that the sensing system remains at the same height and orientation to the ground.

### 4.1.3 Detection of crop rows

Fig.(14𝑎) shows two parallel pepper garden beds, each garden bed consisting of two pepper rows. The mean echo spectrogram of the ensonifications taken from a distance of 0.5[𝑚] from the first row is shown in Fig.(14𝑏), with the correlation sum at 40[𝑘𝐻𝑧] to 60[𝑘𝐻𝑧] representing the echo's strength for each distance from the sensor. The correlation sum is calculated according to the methodology presented in Sec.(3.3). The four peaks in the correlation sum curve in Fig.(14𝑏) represent the two plant rows (separated into two garden beds) and are detectable by either thresholding or searching the local maxima of the plot. The second garden bed in each row is higher than the first as the amplitude values are normalized per frequency and in summing the correlation the error is accumulating.

The ability to detect and determine the distance to objects occluded by foliage (in this case, the second pepper garden bed) in the greenhouse environment is a unique feature of spectrum based ultrasonic sensing.

|  |
| --- |
| **Figure 14:** Detection of crop rows   1. Side view of two pepper plant rows. 2. Mean spectrogram and 40 [kHz] to 60 [kHz] correlation sum.   (a)  (b) |

Fig.(15) shows the LiDAR ranges as detected perpendicular to the pepper rows. The frequency response retains its shape and proportion at each distance. The distance from the first plant is changes proportionally to the LiDAR data (a constant 0.2[𝑚] distance separates them). In the spectrogram at 0.5[𝑚] there is clipping noise due to the microphones dynamic range. This behavior disappears at 1[𝑚], and is taken care of in sensory apparatus 2.0.

|  |
| --- |
| Distance[m] (perpendicular to row)  [Distance [m]  Figure 15 **(a)** : Combination of LiDAR, spectrogram and image from 0.49 [m] orthogonally as measured from the LiDAR. |
| **Figure 15:** Lidar Data  Figure 15 **(b)** : Combination of LiDAR, spectrogram and image from 0.85 [m] orthogonally as measured from the LiDAR.      Distance[m] (perpendicular to row)    [Distance [m] |
| Distance [m] (perpendicular to row)    [Distance [m]  Figure 15 **(c)** : Combination of LiDAR, spectrogram and image from 1.23 [m] orthogonally as measured from the LiDAR. |
| Distance [m] (perpendicular to row)  Distance [m]  Figure 15 **(d)** : Combination of LiDAR, spectrogram and image from 1.9 [m] orthogonally as measured from the LiDAR. |

### 4.1.4 Different cultivars

A comparison between pepper and cucumber rows from a distance of 1.0m and 1.5[m] is presented in Fig.(16):

|  |
| --- |
| **Figure 16** : Averaged amplitude as a function of frequency of pepper and cucumber plants |

The frequency responses of the two cultivars are similar but when examining the spatial response in the spectrograms the response of cucumbers continuously gets stronger and returns higher frequencies starting from 0.9[𝑚] up to a distance of 1.2[𝑚] where a blurry spike rising from 20[𝑘𝐻𝑧] to 120[𝑘𝐻𝑧] can be seen (dashed rectangle), wherever for pepper the returns start with a 20[𝑘𝐻𝑧] to 120[𝑘𝐻𝑧] strong spike.

The average spectrum shows similar patterns for each cultivar. This similarity in shape is expressed mathematically by calculating the slope of each frequency response and subtracting it for the two cultivars, an example is in Fig.(17). The slope differences are significantly higher when comparing different species and the frequency response seems unique for each of the two species.

|  |
| --- |
| Pepper at 1.5[m]slope difference from pepper and cucumber at 1.0[m].  **Figure 17**: Similar patterns for each cultivar in the average spectrum. |

Deducing from the last section and from the research mentioned in Sec.(2.3.4) concerning spectrum based classification in conifer and broad leaved trees: using the magnitude of the frequency response at frequencies which show high differences as features for classification is a logical step.

### 4.1.5 Greenhouse wall and vegetation

The sensing system was mounted at a distance of 2.0[m] perpendicular to the greenhouse plastic wall and to a pepper crop row. Fig.(18) illustrates the mean frequency response obtained from these samples.

|  |
| --- |
| **Figure 18**: Wall and plant spectrum at 1.0[m] |

The strongest differences between plant and wall is expressed at frequencies between 70[𝑘𝐻𝑧] to 105[𝑘𝐻𝑧]. Furthermore, as seen in the spectrograms the echo duration of the wall is shorter and stronger than that of the plants due to depth and porosity of the objects

To summarize, the plants serve as low-pass filters and the excitation returns stronger from walls.

## 4.2 Applicative results

The LDA classifier was trained by 7 plant and 13 wall samples from random orientations and distances. It was found that the ultrasonic return echoes are changed drastically with angle and for this reason the classifier should be trained by samples taken from different angles. The ensonified objects reflect the sound pressure normal from their surface. For this reason the receiver detects the most reflections when orthogonal to the object. The real-time algorithm was planned according to this observation – at the beginning, the algorithm detects a set of points that matches a plane and the classification takes place only after the robot is set orthogonally to that object.

### 4.2.1 Classification

The training set was taken from random orientations to improve the robustness of the real-time mapping. Each sample is the return echo normalized by the emitted CHIRP as detected by the algorithm in real-time. The samples are composed of 1428 pixels, 14×102 concerning 0.25[m] in the spatial dimension and 20[𝑘𝐻𝑧] to 120[𝑘𝐻𝑧] spectrum respectively. The individual pixels represent the different dimensions for the LDA classifier. The dimensionality of the data was important and wouldn’t reduce to 2D or 3D for visualization using PCA (Principle Component Analyses) without a significant loss of information (only 60% of the data is explained by the first three eigenvectors) the average support vectors are a good compromise to visualize the decision hyperplane. The decision echo, the vector normal to the decision hyperplane is also a good measure for the LDA classifier (Yovel, Franz et al. 2008). The decision vector is in practice the set of weights each pixel from the acquired samples will be multiplied to determine to which class it relates the most. The regions with higher absolute amplitudes in the decision vector are areas which relate to dimensions that are given more emphasize in the classification. Fig.(19) shows the two average support vectors and the decision vector for the LDA.

|  |
| --- |
| Figure : Average support vectors  (1)**Plant**  (2) **Wall**  (3) **Decision** vector for LDA. |

The main differences between plant and wall support vectors are first the consistency of power (represented by close colors) for walls in the 75[𝑘𝐻𝑧] to 120[𝑘𝐻𝑧] band in relation to the fluctuations of power for plants. Another difference is that the wall support vectors show higher overall values of power, especially at the 20[𝑘𝐻𝑧] to 50[𝑘𝐻𝑧] band. The higher values for walls start earlier in the echo distance wise with a peak that is stronger for plants at 45[𝑘𝐻𝑧]. This is compatible with Sec.(4.1.5) in which the wall returns are more intense than the plants returns.

It is evident from the decision vector that the classifier is more sensitive to changes in pixels from the lower frequencies, at the 20[𝑘𝐻𝑧] to 50[𝑘𝐻𝑧] band and at 100[𝑘𝐻𝑧] the weights are the largest. This fits the trends in the average support vectors of being stronger and sharper for the walls.

The experiment was run 10 times in which the sensor detected 183 orthogonal surfaces and classified them according to Table.(1):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 1:** Classification confusion matrix   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Known\Predicted | Wall | | Plant | | | Wall |  |  | | | Plant |  |  | | |

The real-time classification results reveal a reasonably reliable classifier which assures that the plant surfaces are classified correctly by moving the arm towards the target and classifying three times.

### 4.2.2 Mapping

A full scan of the room elapses between 2 to 5 minutes depending on the starting position. A result example map is plotted in the next figure:

The map simulates a 2D top view of the manipulator and the environment. The algorithm has had success at mapping the environment, classifying walls and plants and detecting the hidden Styrofoam distance.

The red arcs represent orientations that are unreachable or belong to orientations that cover already classified areas. The blue circles are the Cartesian xy coordinates obtained from the joint angle and sonar distance readings. The green points are the last 4 points scanned and the thin blue line represents the TLS output for best fit line– a suspected plane that will be classified if orthogonal to the robot.

|  |
| --- |
|  |

Subfigures (20𝑎) and (20𝑏) present the map after a surface orthogonal to the robot have been detected and classified as walls. The next subfigures are detections of the plant rows. The last subfigure (20d) shows that the entire perimeter was scanned which signals the end of the operation.



(c)

(a)

(b)

(d)

**Figure 20**: Mapping process

[from top left to bottom right.]

### 4.2.3 Hidden object detection

To ensure that this is in fact a hidden object behind the plants the manipulator scans for peaks over a threshold three times and return the average distance of the object:

|  |
| --- |
| **Figure 21**: The sensing appartus’s correlation sum over spectrogram when it is orthogonal to the plants wall. |

The compensated energy return show two sharp peaks, one corresponds to the plant row distance and the other at 0.85[m] is in correlation with the Styrofoam. The correlation sum is plot with arbitrary units to match the scale of the spectrogram.

## 4.3 Fruit and leaves statistical analysisand classification

### 4.3.1 Single object experiment

The mean frequency spectrum of 3 single leaves and fruit at random orientations is presented in Fig.(22).

|  |
| --- |
| **Figure 22**: Frequency spectrum of leaves and fruit. |

The results indicate that the spectrum of a single leaf and fruit were similar at all frequencies except for two peaks and their neighborhood at 33[kHz] in which the single leaves had higher intensities and 70[kHz] where the opposite is true.

McKerrow and Kristiansen (2006) developed a theoretical model explaining the echo returns from rough surfaces based on the assumption that the energy returns to the transmitting origin only from surfaces nearly perpendicular to the sensor system. This explains the behavior of the single leaves spectrum as their shape is approximately flat and return more energy as apparent from Fig.(22) at 33-55[kHz]. The single fruit second harmonic at 70[kHz] is higher than the leave’s. This could be explained by the increasing directivity of acoustic speakers with frequency. The increasing directivity could be the reason that the effects of area orthogonal to the ultrasonic sensor is turning place to other acoustic phenomenon. While the absolute standard deviation at 33[kHz] and 70[kHz] for a single fruit was 8.6 and 3.4 respectively, for leaves the standard deviation was 17 and 1.5, this difference in variance in the lower frequencies can be explained by the spherical axisymmetric shape of the fruit which cancels the influence of orientation in relation to leaves that return almost nothing when they are not perpendicular.

For this reason a solid criteria for differentiating the returns for single leaves and fruit may be the ration between the frequency response at 70[kHz] and 33[kHz]. This criterion for single leaves gives a ratio of 9 and 4 for fruit.

### 4.3.2 Prediction features extraction experiment

Five plants at three orientations (total of 15 𝑃𝑖𝑂𝑗) were ensonified 30 times each from a distance of 1.25[𝑚] and 0.625[𝑚], with different amounts of fruit. The mean spectrogram of the echoes of the 30 ensonifications for each 𝑃𝑖𝑂𝑗 at each distance was normalized by the CHIRP's intensity of each frequency as described Sec.(3.1.5) forming a normalized echo spectrogram. The mean echo spectrogram of the plants with no fruit 𝑃𝑖𝑂𝑗 are subtracted from the spectrogram of a fruit filled plant 𝑃𝑖𝑂𝑗 to highlight the differences in the echoes of the specific plant and specific orientation (𝑃𝑖𝑂𝑗) with different amount of fruits, e.g. Fig.(23c)=Fig.(23a)- Fig.(23b). Fig.(23) demonstrates this relationship for 𝑃1𝑂1.

|  |
| --- |
| **(a)** Mean normalized echo for unpicked  (b) Mean normalized echo for empty  (c) Difference between normalized mean echoes.  **Figure 23**: Features Extraction Experiment |

Positive differences between the spectrograms represent frequency-distance areas in which the unpicked plant returns higher amplitudes than the plant with no fruits and vice versa. To avoid a situation in which these differences are balancing each other, the positive and negative differences are summed separately for each frequency. The sum of the positive differences for each 𝑃𝑖𝑂𝑗 is displayed in Fig.(24).

The local maximum values of each plant and orientation in Fig(22) express special frequencies of interest that conceal more information about the difference between empty and unpicked plants. Fig.(25) illustrates the superposition of number of local maximums in the difference between full and empty plant at a resolution of 0.488 [𝑘𝐻𝑧] for all .

|  |  |
| --- | --- |
| **Figure 24**: Sum of positive changes between spectrograms  **Figure 25**: The number of peaks per frequency for the 15 plant-orientation combinations | |
|  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 2:** Summary of the regression parameters when estimating the model in Eq. (3) at the range of 28 [kHz] to 32 [kHz].   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Variable | Estimate | Standard Error | T-Statistics | P | | 𝑓𝑖𝑗 | 3.32E-12 | 1.68E-13 | 19.752 | 2.46E-85 | | 𝑚𝑖𝑗 | -7.79E-11 | 9.39E-12 | -8.2996 | 1.17E-16 | | 𝑚𝑖𝑗𝑓𝑖𝑗 | 1.55E-15 | 3.11E-16 | 4.9827 | 6.37E-07 | | 𝑚𝑖𝑗2 | -1.57E-16 | 5.48E-18 | -28.684 | 1.96E-174 | | 𝑓𝑖𝑗2 | 2.71E-14 | 8.90E-16 | 30.444 | 2.30E-195 | |

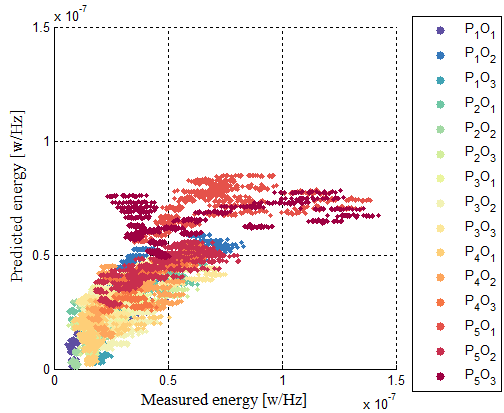
The number of peaks for frequencies 20[𝑘𝐻𝑧], 32[𝑘𝐻𝑧], 68[𝑘𝐻𝑧], 96[𝑘𝐻𝑧] and 116[𝑘𝐻𝑧] is over 10 (out of 15) which stands out in their frequency neighborhood (±10[𝑘𝐻𝑧]). For this reason, these frequencies are considered more carefully when specifying frequency bands for the statistical analysis.

A multi linear regression estimating the energy level 𝐸𝑖𝑗 was computed as described in the methods section over frequencies 28[𝑘𝐻𝑧] to 32[kHz] on 10800 observations. The results are displayed in Table.(2). It reveals a high level of significance as the null hypothesis is rejected for all variables at high probability (P << 0.001) and the adjusted 𝑅2 = 0.64, which suggests a reasonable goodness of fit. The fit parameters concerning energy-mass relation is monotonic and descending. This means that the existence of fruit attenuates the echoes from the plants in accordance to the single fruit and leaves experiment that showed that fruit returns lower amplitudes. The indicator variables, 𝛿𝑖𝑗, vary between −2.61 × 10−8 [𝑊/𝐻𝑧] to 2.99 × 10−8 [𝑊/𝐻𝑧] with high levels of significance (𝑃 ≪ 0.001). This suggests that the influence of the geometry of the single plant cannot be ignored and therefore an estimation of the mass is not possible in these frequencies without sidestepping the indicator variables problem as showed in Eq.(12). Fig.(26) compares the energy levels predicted by the model with the energy measured in the lab experiment, for each plant and orientation. Where each point in the figure represents the amplitude measured in the lab for a certain frequency and mass against the predicted amplitude. The colors of the points indicates to which 𝑃𝑖𝑂𝑗 it belongs to.

|  |
| --- |
|  |

Fig.(26) shows that most of the plant-orientation combinations lies around the 45° line representing complete fitness between the predicted and the measured energy. A similar analysis on 22950 observations at 20[kHz] to 28[kHz] using a quadratic model without indicator variables and using the robust regression, revealed that the calculated multi linear regression values Table.(3) maintain their high level of significance in rejecting the null hypothesis, adjusted 𝑅2 of 0.839 and a P << 0.001. The energy-mass relation is also monotonic and descending with a highly linear influence.

Figure 26: Predicted against experimental results in frequencies 28[kHz] - 32[kHz]



These high levels of significance without using the indicator variables in the model allows the extraction of the mass to frequencies and energy returns relation in Eq.(12) without the noisy outcome of numerical differentiation necessary in Eq.(13).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 3:** Summary of the regression parameters when estimating a quadratic model at the range of 20[kHz] to 28[kHz].   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Variable | Estimate | Standard Error | T-Statistic | P | | 1 | 8.32E-08 | 1.61E-09 | 51.526 | 0 | | 𝑓𝑖𝑗 | -8.14E-12 | 1.35E-13 | -60.108 | 0 | | 𝑚𝑖𝑗 | -1.24E-11 | 4.06E-13 | -30.553 | 5.49E-201 | | 𝑚𝑖𝑗𝑓𝑖𝑗 | 5.90E-16 | 1.64E-17 | 35.961 | 1.49E-275 | | 𝑚𝑖𝑗2 | 2.03E-16 | 2.82E-18 | 72.034 | 0 | | 𝑓𝑖𝑗2 | -4.47E-16 | 8.71E-17 | -5.1254 | 2.99E-07 | |

Fig.(27) shows the fitness level of the energy predicted by the model to the energy measured in the lab experiment for each plant and orientation.

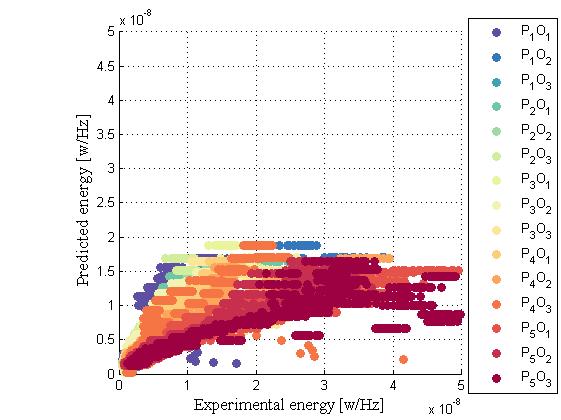


Figure 27: Predicted against experimental results in frequencies of 20[kHz] to 28[kHz]

|  |
| --- |
|  |

To overcome noise related to the quantization error and improve the results, an energy threshold is applied over the samples used for the prediction. The filtered data is averaged for each plant #𝑖 from all three orientations for each plant mass level to obtain an estimation of the mass displayed in the next figure:

|  |
| --- |
| Figure : Predicted against experimental mass |

The linear relationship between predicted and experimental mass is depicted in Fig.(28) and has an R2 of 0.34. The estimation accuracy can be improved by acquiring more samples of the plants from different orientations. Obtaining data from several orientations can be a rigorous task well suited for a robot.

### 4.3.3 Leaves experiment

At the time of year that the experiment was held, the five plants did not have any fruit. The plants were ensonified 30 times each from a distance of 1.25[[𝑚] and 0.625[𝑚] and three orientations (total of 15 𝑃𝑖𝑂𝑗), with different amounts of leaves. The energy differences between the plants with different amount of leaves were clearer and required less analysis to get better results.

A multi linear regression associating the normalized spectrum energy, frequency, and number of leaves (𝑙𝑖𝑗) was used to model the influence of leaves on the ultrasonic echo return. A total of 19016 observations were included in the estimation at ranges of frequencies from 20[𝑘𝐻𝑧] to 28[𝑘𝐻𝑧]. The models 𝑅2 stands on 0.639 and the F statistic P value is 0. The quadratic surface and the observations are plotted in Fig.(29) and the parameters are listed in Table(4):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | Variable | Estimate | Standard Error | T-Statistic | P | | 1 | 0.85 | 0.02267 | 37.743 | 1.10E-300 | | 𝑓𝑖𝑗 | -7.58E-05 | 1.89E-06 | -39.955 | 0 | | 𝑙𝑖𝑗 | -0.008 | 1.66E-05 | -51.312 | 0 | | 𝑙𝑖𝑗𝑓𝑖𝑗 | 4.19E-08 | 6.63E-10 | 63.17 | 0 | | 𝑙𝑖𝑗2 | 1.69E-09 | 3.96E-11 | 42.707 | 0 | | 𝑓𝑖𝑗2 | -7.24E-08 | 1.53E-08 | -4.715 | 2.43E-06 |   Table 4: Summary of the regression parameters when estimating a quadratic model at the range of 20 kHz to 28 kHz. |

The relation between number of leaves and energy return is expected to ascend as more leaves would form orthogonal planes that return more energy. Table.(4) shows that the linear relation between those properties is descending but as can be seen in Fig.(29) the other parameters of the surface compensate for this behavior and at frequencies of 20[k𝐻𝑧] to 28[k𝐻𝑧] the relation is ascending as expected.

|  |
| --- |
| **Figure 30**: Predicted against experimental number of leaves from 1.25[m] |

|  |
| --- |
| **Figure 29** : observations in black with the quadratic model surface at to from 1.25[m]. |

Fig.(29) shows every 5𝑡ℎ observation for visibility. The faded dots are observations concealed by the surface. The scattering of the observations intensifies with frequency as is apparent from the figure. To reduce the error in prediction, OLS over the error between observations energy and predicted energy is applied with the variable being the number of leaves see Eq.(14). By applying this method instead of averaging the predicted number of leaves for each energy the error is minimized.

The predicted number of leaves against the counted number is presented in Fig.(30). The linear regression of the prediction from a distance of 1.25[𝑚] has a ratio of 1 and an intercept close to 0 with a 𝑅2 value of 0.73. The results from closer distance are very close with a ratio of 0.96 and 𝑅2 of 0.72.

The results for the leaf estimation were cross validated by using a leave 1 out 5 plant validation technique to estimate the accuracy of the predictive model. Table.(5) gathers the statistics of the estimated parameters of the predicted vs experimental number of leaves curve.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Parameter | Mean | Standard deviation | | Slope | 1.16 | 0.49 | | Intercept | -13.3 | 51.75 | | 𝑹2 | 0.84 | 0.11 |   **Table 5**: Statistics of cross validation for 1.25[m] sample |

The cross validation slope is close to 1 with an intercept of -13 close to 0 leaves and high 𝑅2 which indicting overall good prediction capabilities in the scope of the sample. The standard deviations of the cross validation are not negligible but it seems that with more samples of plants the prediction would get better.

### 4.3.4 Neural Networks estimation

In the last two subsections it was established that the relations of fruit mass and leaf number to energy and frequency are statistically significant even for small scale Taylor approximations of the data in hands. Geometric plant specific properties were shown to influence the approximation and interfere with the estimate. In this subsection a neural network is experimented and cross validated in order to improve estimation beyond the suggested statistical method.

A six hidden layers feed forward neural network was trained using Levenberg- Marquardt algorithm. The inputs from the frequency response were chosen by iterating a linear regression model with the frequency response amplitudes as in Eq.(17). Each iteration the frequencies that were not statistically significant were excluded. The final frequencies that were used were close to 20[kHz], 33[kHz], 68[kHz], 96[kHz] and 116[kHz] – the frequencies that were found to contain data about difference between plants with and without fruits in Fig.(25). The weight of the fruit and leaves were naturally chosen as the output.

|  |
| --- |
| **Figure 31:** Neuralnetworkprediction of fruit mass from 1.25[m] with no cross validation. |

The samples were divided as follows: two PiOj samples were taken out as test data to leave 2 out of 13 cross validation, while the remaining data was chosen randomly so that 85% will be used for training and 15% for validation of the neural network convergence.

|  |
| --- |
| Figure 32: Neural network prediction of number of leaves from 1.25[m] with no cross validation |

An example of the neural network output over the whole sample is seen in Fig.(31). Each point in Fig.(31) represents a different PiOj with a specific weight of fruits and not the predicted data for a plant from all three orientations as was the case in the previous subsection. The output of the neural network estimate is more precise than the suggested statistical model but it is not accurate in the sense that it does not predict a 1:1 ratio between the predicted and experimental mass and a zero intercept. The problem also extends to the leave experiment, see Fig.(32):

Although the fit is precise it's important to remember this is a plot of predicted data with which the neural network was trained with. The cross validation results for the slope, intercept and R2 are presented in the tables for the fruit and leaves experiments from 1.25[m], the results are very similar for the 0.625[m] case:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Parameter | Mean | Standard deviation | | Slope | 0.847 | 0.26 | | Intercept | 60.9 | 108.59 | | 𝑹2 | 0.86 | 0.31 |   **Table 6:** Statistics of cross validation for 1.25[𝑚] fruit experiment |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Parameter | Mean | Standard deviation | | Slope | 0.94 | 0.12 | | Intercept | 10.55 | 19.61 | | 𝑹2 | 0.95 | 0.15 |   **Table 7**: Statistics of cross validation for 1.25[𝑚] leaves experiment |

The tables show that the slope and intercept of the cross validation stage converge to the whole sample space slope and intercept with high R2 values. This convergence allows for the compensation of the accuracy problem addressed earlier using the mean values of the intercept and slope to fix the predicted values. The standard deviation is higher for the fruit experiment; the standard deviation is 30% in comparison to 12%. This inaccuracy would most certainly be fixed with a larger sample space.

# 5 Conclusions

A sonar system for yield assessment and greenhouse features identification was developed and tested in lab and greenhouse environments. The acoustic signature of the greenhouse porous ground was constructed and can be regarded and filtered from the crop acoustic spectrogram signature.

The developed system in relation to sensors such as RGB-D cameras and LiDARs can detect and map crop rows without a direct line of sight using a matched filter and compensating for the acoustic attenuation in the greenhouse environment. The ability to detect and determine the distance to objects occluded by foliage in the greenhouse environment is a unique feature of spectrum based ultrasonic sensing.

The spectrum signature of walls and screens (greenhouse infrastructures) can be characterized and distinguished from the greenhouse vegetation at high ultrasonic frequencies. These differences have been implemented to classify between walls and plants in real-time.

Single fruits were found to return less energy than leaves because of their shape: the smaller surface area perpendicular to the sensing system returns less energy. At 30 [𝑘𝐻𝑧] single leaves return more energy than in other frequencies and more than a fruit.

A multi linear regression model for estimating the energy level was found to be highly significant and correlated for 20[𝑘𝐻𝑧] to 32[𝑘𝐻𝑧] range. Indicator variables concerning a constant energy level for the plants specific orientation show the significance of plant geometry. The results indicate that pepper fruits mass on the plants are correlated to energy and frequency at a number of frequency bands. The relation between fruit mass and energy returning from the plant is in general monotonic and decreasing. A return echo model to mass and number of leaves was developed with an estimation 𝑅2 of 0.34 and 0.72 respectively.

A feed forward 6 hidden layers neural network was trained to estimate the number of leaves and fruit mass with the return energy at frequencies that were the most sensitive to them. The neural network prediction after cross validation yields 𝑅2 values of 0.95 and 0.86 but their accuracy remains in question.

Using more frequency bands, ensonifying the plants from several directions, obtaining more samples from each plant and increasing the amplitude of the emitted signals will likely improve the estimations of the mass and energy evaluation models.

Future research directions which were not explored due to limited time for yield estimation with sonar include obtaining more data at high frequencies by increasing the emitted signal or by using a log-sweep CHIRP instead of linear CHIRP. Using the radon transform with a multi array of microphones to obtain 2D and 3D intensity images of the plants. Learning more about how the pressure flow of sound changes in vegetation by using the Schlieren flow visualization and using the current state of the art deep learning algorithms to obtain better estimation.

Furthermore, research regarding motion planning for a mobile platform for acquiring enough samples from plants to evaluate fruit mass and leaf number at a sufficient accuracy is also necessary for a complete solution of the problem.

# Reference List

1. Dams, R. A. and S. C. Pedersen (2000). “Ontogeny, Functional Ecology, and Evolution of Bats”, Cambridge University Press.
2. Akbarally, H. and L. Kleeman (1995) “A sonar sensor for accurate 3D target localization and classification”. ICRA.
3. Andújar, D., A. Escola, J. Dorado and C. Fernandez-Quintanilla (2011).”Weed discrimination using ultrasonic sensors”. Weed Research 51(6): 5.43-547
4. Andújar, D., C. Fernández-Quintanilla and J. Dorado (2015). "Matching the best viewing angle in depth cameras for biomass estimation based on poplar seedling geometry." Sensors 15(6): 12999-13011.
5. Balleri, A., H. D. Griffiths, C. J. Baker, K. Woodbridge and M. W. Holderied (2012). "Analysis of acoustic echoes from a bat-pollinated plant species: insight into strategies for radar and sonar target classification." IET Radar, Sonar & Navigation 6(6): 536544.
6. Bechar, A. (2010). "Robotics in horticultural field production." Stewart Postharvest Review 6(3): 1-11.
7. Bloch, V., A. Dgani and A. Bechar (2013). “The Effect of Location and Orientation of Agricultural Targets for the Design of an Optimal Robot”, The Israeli Society of Agricultural Engineering.
8. Cowell, D. M. J. and S. Freear (2010). "Separation of overlapping linear frequency modulated (LFM) signals using the fractional fourier transform. “Ultrasonic, Ferroelectrics, and Frequency Control, IEEE Transactions on **57**(10): 2324-2333.
9. Elfes, A. (1987)” Sonar-based real-world mapping and navigation. “Robotics and Automation, IEEE Journal of **3**(3): 249-265.
10. Farooq, U., G. Abbas, S. O. Saleh and M. U. Asad (2012). “Corridor navigation with fuzzy logic control for sonar based mobile robots”. Industrial Electronics and Applications (ICIEA), 2012 7th IEEE Conference on.
11. Fermont, A. and T. Benson (2011). "Estimating yield of food crops grown by smallholder farmers." International Food Policy Research Institute, Washington DC: 1-68.
12. Glauber, J. W. (2013). "The growth of the federal crop insurance program, 1990– 2011." American Journal of Agricultural Economics 95(2): 482-488.
13. Goodwin, B. K. and V. H. Smith (2013). "What harm is done by subsidizing crop insurance?" American Journal of Agricultural Economics 95(2.489-497 :)
14. Harper, N. and P. McKerrow (2001). "Recognising plants with ultrasonic sensing for mobile robot navigation." Robotics and Autonomous Systems 34(2–3): 71-82.
15. He, P. (1998). "Simulation of ultrasound pulse propagation in lossy media obeying a frequency power law." Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on 45(1): 114-125.
16. Jakevičius, L. and A. Demčenko (2008). "Ultrasound attenuation dependence on air temperature in closed chambers." Ultragarsas (Ultrasound) 63(1): 18.22-
17. Jinwoo, C., A. Sunghwan, C. Minyong and C. Wan Kyun (2006). Metric SLAM in Home Environment with Visual Objects and Sonar Features. Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on.
18. Kinsler, L. E. (2000). Fundamentals of acoustics, Wiley.
19. Kleeman, L. (2003). Advanced sonar and odometry error modeling for simultaneous localisation and map building. Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on.
20. Kleeman, L. and R. Kuc .)2008( Sonar Sensing. Springer Handbook of Robotics. B. Siciliano and O. Khatib, Springer Berlin Heidelberg, Germany: 491-519.
21. Kuc, R. (2001). "Pseudo-amplitude scan sonar maps." Robotics and Automation, IEEE Transactions on 17(5): 767-770.
22. Kuc, R. (2004" .)Neuro-computational processing of moving sonar echoes classifies and localizes foliage." The Journal of the Acoustical Society of America 116(3): 1811-1818.
23. Lee, W., V. Alchanatis, C. Yang, M. Hirafuji, D. Moshou and C. Li (2010). "Sensing technologies for precision specialty crop production." Computers and Electronics in Agriculture 74(1): 2-33.
24. Lim, Z., S. Kwon and M. Joo (2012). "Multi-object identification for mobile robot using ultrasonic sensors." International Journal of Control, Automation and Systems 10(3): 589-593.
25. Linda, O. and M. Manic (2011). "Self-Organizing Fuzzy Haptic Teleoperation of Mobile Robot Using Sparse Sonar Data." Industrial Electronics, IEEE Transactions on 58(8): 3187-3195.
26. lvarez, F. J., R. Kuc and T. Aguilera (2011). "Identifying Fabrics With a Variable Emission Airborne Spiking Sonar." Sensors Journal, IEEE 11(9): 1905-1912.
27. McKerrow, P. J. and B. E. Kristiansen (2006). "Classifying Surface Roughness With CTFM Ultrasonic Sensing." Sensors Journal, IEEE 6(5): 1267-1279.
28. McKerrow, P. J. and N. Lindsay Harper (1999). "Recognising leafy plants with in‐air sonar." Sensor Review 19(3): 202-206.
29. Milford, M. J., G. F. Wyeth and D. Prasser (2004). RatSLAM: a hippocampal model for simultaneous localization and mapping. Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on.
30. Møhl, B. (1988). Target Detection by Echolocating Bats. Animal Sonar: Processes and Performance. P. E. Nachtigall and P. W. B. Moore. Boston, MA, Springer US: 435450.
31. Møhl ,B. (1988). Target Detection by Echolocating Bats. Animal Sonar. P. Nachtigall and P. B. Moore, Springer US. 156: 435-450.
32. Moonrinta, J., S. Chaivivatrakul, M. N. Dailey and M. Ekpanyapong (2010). Fruit detection, tracking, and 3D reconstruction for crop mapping and yield estimation. Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on.
33. Mulla, D. J. (2013). "Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps." Biosystems Engineering 114(4): 358-371.
34. Müller, R. and R. Kuc (2000). "Foliage echoes: a probe into the ecological acoustics of bat echolocation." The Journal of the Acoustical Society of America 108(2): 836-845.
35. Müller, S. and P. Massarani (2001). "Transfer-function measurement with sweeps." Journal of the Audio Engineering Society 49(6): 443-471.
36. Nakamura, K. (2012). Ultrasonic transducers: Materials and design for sensors, actuators and medical applications, Elsevier.
37. Nuske, S., S. Achar, T. Bates, S. Narasimhan and S. Singh (2011). Yield estimation in vineyards by visual grape detection. Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on.
38. Schillebeeckx, F., F. De Mey and H. Peremans (2008). Bio-inspired sonar antennae: Enhancing directivity patterns for localization. Biomedical Robotics and Biomechatronics, 2008. BioRob 2008. 2nd IEEE RAS & EMBS International Conference on, IEEE.
39. Schillebeeckx, F. and H. Peremans (2010). Biomimetic sonar: 3D-localization of multiple reflectors. Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on.
40. Shi, R. Z. and T. K. Horiuchi (2007). "A Neuromorphic VLSI Model of Bat Interaural Level Difference Processing for Azimuthal Echolocation." IEEE Transactions on Circuits and Systems I: Regular Papers 54(1): 74-88.
41. Siciliano, B. and O. Khatib (2008). Springer handbook of robotics, Springer Science & Business Media.
42. Stajnko, D., M. Lakota and M. Hočevar (2004). "Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging." Computers and Electronics in Agriculture 42(1): 31-42.
43. Steckel, J., F. Schillebeeckx and H. Peremans (2011). Biomimetic sonar, outer ears versus arrays. Sensors, 2011 IEEE.
44. Toda, M., O. Kitani, T. Okamoto and T. Torii (1999). "Navigation method for a mobile robot via sonar-based crop row mapping and fuzzy logic control." Journal of Agricultural Engineering Research 72(4): 299-309.
45. Tungadi, F. and L. Kleeman (2012). "Autonomous loop exploration and slam with fusion of advanced sonar and laser polar scan matching." Robotica 30(1): 91-105.
46. Wachs, J. P., H. I. Stern, T. Burks and V. Alchanatis (2010). "Low and high-level visual feature-based apple detection from multi-modal images." Precision Agriculture 11(6): 717-73.5
47. Wang, Q., S. Nuske, M. Bergerman and S. Singh (2013). Automated crop yield estimation for apple orchards. Experimental Robotics, Springer.
48. Waters, D. (2007). "Echolocation in air: biological systems, technical challenges, and transducer design." Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science 221(10): 1165-1175.
49. Welch, P. D. (1967). "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms." Audio and Electroacoustics, IEEE Transactions on 15(2): 70-73.
50. Yovel, Y., M. O. Franz, P. Stilz and H.-U. Schnitzler (2008). "Plant Classification from Bat-Like Echolocation Signals." PLoS Computational biology 4(3): e1000032.
51. Yovel, Y., P. Stilz, M. O. Franz, A. Boonman and H.-U. Schnitzler (2009). "What a Plant Sounds Like: The Statistics of Vegetation Echoes as Received by Echolocating Bats." PLoS Computational biology 5(7): e1000429

# Publications

1. Finkelshtain R. , Yovel, Y. , Kósa, G., Bechar A. "Agricultural Robot in Greenhouses: Preliminary Results”, the annual meetings of the Israeli Society of Agricultural Engineering, Israel, 2014. (Oral presentation)
2. Finkelshtain R. , Yovel, Y. , Kósa, G., Bechar A. “Autonomous robot for yield assessment.”, The 33rd Israeli Conference on Mechanical Engineering, Israel, 2015 (Oral presentation).
3. Finkelshtain R. , Yovel, Y. , Kósa, G., Bechar A. “Autonomous robot for yield assessment.”, 10th Conference of the ECPA, ARO, Israel, July 2015. (Proceeding)
4. Kósa G., Yovel Y. , Bechar A. , Finkelshtain R. , "AGRICULTURAL ROBOT", PCT/IB2016/050303. (Patent submission).
5. Finkelshtain R. , Kósa G., Yovel Y., Bechar A. "An agricultural robot for yield assessment using ultrasonic-based feature perception", The 5th Israeli Conference on Robotics, Israel, 2016, (Oral presentation).
6. Finkelshtain R. , Kósa G., Yovel Y., Bechar A. “Investigation and analysis of an ultrasonic sensor for specific yield assessment and greenhouse features identification” , Precision Agriculture, 2016 (Journal, under review).

# Appendix

% getspecdist detected the ToF between a CHIRP emission and

% the first return

% echo.

% Input variables: lowf, highf - frequency bounds.

% P - "straitened" spectogram.

% pxlsightdist - maximum target distance translated to pixels.

% chirpmargin - width of the CHIRP at each frequency translate to pixels.

% sz - constant time between trigger translated to pixels.

% F, MET - Frequency and distance conversion vector to pixels.

% Output: Distance.

function [pxl d corvals] = getspecdist(lowf,highf,P,pxlsightdist,chirpmargin,sz,F,MET)

nsample=P(:,chirpmargin+sz+1:pxlsightdist+sz); %the spectogram without the CHIRP time

nsignal=P(:,chirpmargin:chirpmargin+sz); %isolate chirp

nsample=nsample/max(max(nsample)); %normalizing energy before correlation

nsignal=nsignal/max(max(nsignal));

[c range]=xcorr2D(F,lowf,highf,nsignal,nsample); %correlation of signals at chosen frequency range

corvals=nanmean(c(:,size(nsample,2):-1:sz+1),1); %the mean sums the correlation, the correlation is transformed to the time domain

[m i]=max(corvals); %the maximum correlation is analogous to the ToF

pxl=sz+1+i;

d=MET(pxl); %the distance is translated from pixel values to distance.

end

%xcorr2D correlates 2 matrices with the same number of rows, row by row.

function [c,range] = xcorr2D(F,F1,F2,signal,returnsignal)

range=find(F>F1 & F<F2);

j=0;

for i=1:length(range)

j=j+1;

c(j,:)=xcorr(signal(range(i),:),returnsignal(range(i),:));

end

end