Mini Survey of Ji Li

Title

Real-time crop recognition system for mechanical weeding robot using time-of-flight 3D camera and advanced computation based on field-programmable gate array (FPGA)

Introduction:

Weed control is one of the most labor-intensive tasks for organic vegetable farms, and a real-time crop recognition system is the main challenge to develop the automatic mechanical weeding system. One solution using time-of-flight (TOF) 3D camera and advanced computation technique based on FPGA is proposed to address this problem. It has great advantages and can work reliably in the real harsh weeding environment.

Broader Impact:

The proposed machine vision system for weeding robot will bring several great potential impacts as following:

- As the computational intelligence system, it is the most critical part to lead the weeding system to evolve to the practical automatic mechanical weeding robot.

- This study will greatly help to reduce the labor cost for weed control of organic vegetable productions, thus will potentially promote its production abilities. People will be benefited with more sufficient organic vegetable supplies with lower prices.

- The new technology will make it easier to grow organic vegetables, and this will potentially lead more farms to its production. This will reduce the usage of herbicide, and help to prevent the environment and people from the pollutions.

- Forceful manual weeding work may lead to musculoskeletal disorder, this research will prevent farmers from it, and benefit them better health.

- As this study is seeking a way to do the advanced image computation using hardware design on FPGA rather than normal software, it will prevent the intelligence property of the machine vision technique from being pirated like the normal software on computer.

Current Knowledge:

Sensors:

Automated weeding system calls for reliable imaging sensors and robust image processing algorithm. Radar and spectral image sensors are widely installed on satellite for crop recognition.

A paper reviewed the technique applying spectral inputs to crop identification and condition assessment (Bauer, 1985). It introduced the spectral-temporal profile models for crop identification. And the same model form also was used for estimating crop development stage, leaf area index, and canopy light interception. The spectral image sensor is mainly used by satellite for remote sensing. As this method requires a set of spectral image at the same view point, it requires the weeding machine stop when it is capturing images. As it is desirable to make the weeding machine keeping on moving with no stop, spectral image sensor is not a good choice for our project.

Some machine vision systems for automated agricultural application were developed based on two dimensional (2D) color cameras and corresponding image processing algorithms. For example, corn plant detection system were developed in this approaches (Shrestha & Steward, 2003, 2005; Shrestha, Steward, & Birrell, 2004; Tang & Tian, 2008a, 2008b). Their system mounted 2D cameras on a mobile platform to get the top-view 2D color image of plant canopies. The image frames were mosaicked together to form long crop row image, which was later processed to do the corn plant identification. However, their works have some shortcomings because of the 2D color camera. Firstly, their systems rely on the image processing algorithm based on the color information, but the 2D color camera is vulnerable to the changing outdoor lighting conditions. It needs the direct sun light to be blocked or be diffused. Otherwise they may fail to work because the image color of part of objects is washed out due to the mirror reflection or the strong sun light. Even so, the changing light conditions due to the different weather conditions and time of day will greatly affect the color rendering of objects. The inconsistent image color is a great inherent drawback of the sensor, and is hard to overcome to make the system robust to changing sky conditions and different time of day. Secondly, it is hard to reliably separate the overlapped canopies to individual plant when they have similar color. Thirdly, while the same crop may have very different 2D shapes at different viewpoints, the weed and crop may have the similar ones at some viewpoints. This challenge is also hard to overcome for 2D image sensor to do the crop recognition task. Therefore, additional information such as the depth information is necessary. With the introduction of the depth information, the crop plant is sensed in three-dimensional (3D) space.

Stereo vision is one approach to get the 3D image. There are two basic techniques, which are passive stereo vision and active stereo vision (Chen & Zheng, 1995). Passive stereo vision applies two or more calibrated 2D cameras in distinct locations to sense the objects. The mechanism of passive stereo vision is similar with human eyes. It estimates the depth information through searching the corresponding pixels of the same point of objects in the images captured by different cameras and doing triangulating. A method of corn plant detection based on passive stereo vision was reported by Jian and Lie (2009). For their work, the corn plant were reliably segmented from the background of weed and earth, and the stem position of the crop is well identified in order to estimate spacing conditions of corn plants to help the improvement of planter design. Compared with the passive stereo vision approach, active stereo vision system has a well calibrated projector in addition to cameras. It projects structural light pattern, for example, parallel lines or grids, on the surface of the sensed objects, and the 3D shape is recovered by analyzing the distortion image of the light pattern due to the 3D geometry captured by the camera. Veeken et al., 2006 reported a corn plant sensing system, which estimated corn plant 3D structure by projecting parallel lines on them. But their system is vulnerable to the overlapped canopies formed complex conditions.

Both passive and active stereo vision systems have some drawbacks for crop plant sensing. The passive stereo vision system relies on the correspondence matching between the images of two or more calibrated cameras, and it requires adequate textural information to ensure its accuracy and reliability (Trucco & Verri, 1998). However, for many cases, the leaves of plant are lack of texture information, and their uniform texture will lead to unreliable correspondence matching, thus result in none or false depth information. On the other hand, the passive stereo camera requires good illumination conditions. Both weak and strong sun light may result in lack of exposure or over exposure, thus lead to false correspondence matching and depth information because of the loss of texture information. For the active stereo vision, effective depth information of only the places where the structural light pattern is projected can be estimated. If the structural light patterns are parallel lines or grids, only the depth information of where the lines are projected can be achieved. In that case, its special resolution is low, and is not suitable for complex plant canopies. At last, both the passive and active stereo vision systems are vulnerable to the vibration to the machine platform in the field. They need their cameras and projectors to be well calibrated in order to do the later matching calculation. Once they are calibrated, the relative directions and positions of their components, including cameras and projectors, should remain constant. Otherwise, they need to be recalibrated. However, the great vibration of the machine platform, such as weeding system in this case, will lead to change of their components' relative positions and directions, thus make the depth information ineffective.

An innovative 3D time-of-flight camera (TOF camera) enables us greater power to do the crop recognition task. TOF camera has a light source and light detector. Its light source transmits light pulse or laser to sensed objects, and the light detector is utilized to receive the light pulse or laser reflected back. The distance information is achieved by calculating the time of flight of light. Its working principal is similar with radar. There are several bands of TOF camera in the market. Thanks to the strong light power, the products of PMD are robust to the direct sun light, and their error rate of depth information ranges from 3mm to 7mm depend on different products. Thus, the robustness to the outdoor sunlight and the high accuracy level of PMD camera benefits us a better sensor than those mentioned above. While none plant sensing related research tried to use this technique, the proposed research plans to apply it for the outdoor crop recognition task.

Processors:

Apart from the image sensor, the image processing processor is also critical for real-time crop recognition task. There are several types of processors, including computer, Digital Signal Processor (DSP), Graphical Processing Unit (GPU), FPGA, and ARM are used for image processing. Among them, computer is a standard and widely used processor, and it is relatively easier to do image processing programming for it. Most plant sensing system were based on computer (Shrestha & Steward, 2003, 2005; Shrestha, Steward, & Birrell, 2004; Tang & Tian, 2008a, 2008b, Jin & Tang 2009). But their systems were not able to do real-time image processing. Their images data were processed in the lab after being captured in the field. Their image processing algorithm is heavy load calculation task for computer they used. As a general purpose processor, the CPU of computer implements the instructions serially. That means it only can carry out calculations one by one. Compared to other processor, such as GPU, FPGA, and DSP, which can do many calculations simultaneously, computer is not a good solution for image processing task as the algorithm allows parallel calculation heavily. As the weeding system requires real-time crop recognition, the heavy load calculation tasks requires a parallel-calculation-enable processor, such as DSP, FPGA, or GPU.

DSP is a specialized microprocessor with an architecture optimized for the fast operational needs of digital signal processing (Yovits & Marshall C., 1993, Liptak & Béla G., 2006). DSP has multiple calculators to do the multiple calculations simultaneously. As digital image processing consists of a large number of mathematical operations to be performed parallel, DSP is a choice for it. None research related to crop recognition is found based on DSP. However, there are a large number of machine vision systems for robotics reported successfully applied DSP. A vision guided dual arms robot (Huang & Huang, 2011) employed a DSP to do the complicate computation of 3D image processing. And their machine vision system can detect and track the objects in 3D space to direct arms to pick and place objects or track the specified moving targets.

FPGA is a potential optimal processor for crop recognition task for weeding system. It is an integrated circuit chip. Designers can design a customized digital system by programming its rich hardware resource using a hardware description language (HDL). The main idea of FPGA is to customize a circuit system to do the task, rather than making program runs on general purpose CPU or microprocessor. Thus, developers are empowered the ability to design whatever architecture which is best fit for a specific task. Thanks to its parallelism characteristics, the optimal customized hardware based on FPGA can carry out hundreds or more calculations simultaneously. Compared to the computer and DSP, which can do only one or several calculations simultaneously, FPGA is a better choice for some applications in the area of advanced computation, machine vision, digital signal processing, etc. Savarimuthu, et al., 2011, reported a real-time medical video processing system which applied FPGA to do the correlation-based algorithm for blood vessel detection task. Their FPGA based system can process 215 frames of 640 x 480 pixel size images per second, while the high speed computer only can process 0.25 frames per second maximum. Thus, their FPGA based system is over 800 times faster than computer. This suggests that FPGA is better choice for the machine vision system for the automated weeding machine, and much faster speed can be achieved than computer. On the other hand, FPGA features low power consumption, and does not need any thermovent as computer. This advantage make it possible to develop a portable and completed isolated device based on FPGA from the dust environment in weeding field, and make it much more robust to the harsh environment compared to computer.

- Gap: While FPGA is a potential ideal platform for crop recognition, many image processing functions necessary for this task are still left to be studied. Compared to the mature 2D image processing techniques based on general purpose computer, those based on FPGA are much more challenging to design and much less studies of them have been reported. A FPGA based image processing skeletons were reported for common image processing tasks (Benkrid, Crookes, & Benkrid, 2002). And parallel neighborhood operations based functions for 2D images were provided as examples. Additionally, Pearson, 2009, reported a FPGA-based image processing system for high-speed inspection of grains. However, its function implemented on FPGA is a simple algorithm based on intensity analysis of 2D images. For all their works, only some simple basic image processing functions, such as filter, edge detection, and intensity adjustment were realized. As implementing image processing algorithm on FPGA is much more complicated than on computer, few researches were reported about other important functions necessary for crop recognition task, including segmentation, feature extraction, and other complicated morphological operations. To successfully develop a real-time crop recognition system based on FPGA, the optimal solutions for above problems need to be studied, and these studies can promote the development of related area. Furthermore, the 3D matching engine, which can be used for object recognition task by matching the captured 3D image data with 3D model in database, is a potential optimal method of crop recognition, as it can fully take the advantage of depth information of 3D camera. Our research will be novel to realize 3D matching engine on FPGA as none related research was found.

Gap as an Important Problem

- As only some basic image processing algorithm were reported to be implemented on FPGA, and the optimal design of other more complicated ones which are always necessary for machine vision have not been reported, it is important to study them in order to promote the development of FPGA-based machine vision technique. As the functions are necessary for the crop recognition algorithm, they have to be solved. On the other hand, 3D model matching engine can do crop recognition task by matching the data of leaves with the 3D model in the database, it can take the full advantages of the 3D camera, and it should be a potential optimal solution. Additionally, as the calculation load of 3D matching engine is heavy for computer, and none research based on FPGA was found, implementing it on FPGA is an ideal real-time solution, and it will be a novel work in related areas.

Object Classification Algorithm

The pattern recognition algorithm is important for crop recognition system based on machine vision technique. The soft calculation algorithm, including support vector machine (SVM), Artificial Neuron Network (ANN), and Support Vector Machine (SVM) are widely used for pattern recognition task. ANN was applied for remote sensing estimation of impervious surface, which is significant in monitoring urban development and determining the overall environment health of a watershed (Wang, et al, 2008). In their work, ANN is trained to classify the impervious surface, and reasonable result is achived. Brown, et al., 2000, discussed applying SVM for remote sensing. And the landform classification task was given as an example of SVM.

Other techniques based on statistics were also widely used for pattern recognition. For example, Principal Component Analysis (PCA) was applied to face recognition. And Capobianco, et al., 2009, reported a target detection method with semisupervised kernel orthogonal subspace projection. Their technique needs manually labeled samples, and then it can do the classification task. They used their algorithm to classify 15 land-cover types of environment on the satellite images, and demonstrate good capabilities.

A paper reported a face recognition system using 3D model matching engine. In their research, they recovered the 3D model of human face using a set of sequential 2D video images, and then the 3D model achieved were feed into a 3D model matching engine to compare them with the ones in database. The idea of implementing a 3D model matching engine on FPGA to do the crop recognition is inspired by their work. This engine can be used to do the crop recognition by classifying their leaves as the shapes of leaves belong to same type of plants are similar.

Long-Term Goal:

- The long term goal of the proposed project is to develop a real-time crop recognition system for automated weeding robot using TOF 3D camera and FPGA. To do that, an interface will be designed on FPGA to make it be able to directly get access to the 3D camera. Additionally, a systematical FPGA-based image processing functions will be developed to do the plant segmentation from the background and individual separation. Moreover, a 3D model matching engine will be developed to do the final crop recognition task by matching the 3D data of leaves with the model in the database. Once the proposed project is done, a customized optimal crop recognition system which features real-time performance and reliability will be achieved. On the other hand, it will be a novel systematic solution of object recognition system based on optimal 3D image processing architecture on FPGA, thus a new research area will be started, and this technique can be applied to many other application areas. Finally, the control system of our weeding robot will be integrated to the proposed crop recognition system. In that case, the single FPGA will take charge of both tasks parallel. It will do the crop recognition task, while controlling the mechanical weeding operation simultaneously according to the recognized crop positions.

References

Shrestha, D. S., & Steward, B. L. (2003). Automatic corn plant population measurement using machine vision. Transactions of the ASAE 46(2), 559 – 565.

Bauer M. E. (1985). Proceedings of the IEEE. 73(6), 1071-1085.

Shrestha, D. S., & Steward, B. L. (2005). Shape and size analysis of corn plant canopies for plant population and spacing sensing. Applied Engineering in Agriculture, 21(2), 295 – 303.

Shrestha, D. S., Steward, B. L., & Birrell, S. J. (2004). Video processing for early stage maize plant detection. Biosystems Engineering, 89(2), 119 – 129.

Tang, L., & Tian, L. (2008a). Real-time crop row image reconstruction for automatic emerged corn plant spacing measurement. Transactions of the ASABE, 51(3), 1079 – 1087.

Jin J., & Tang, L. (2009). Corn plant sensing using real-time stereo vision. Journal of Field Robotics, 26(6–7), 591–608

Tang, L., & Tian, L. (2008b). Plant identification in mosaicked crop row images for automatic emerged corn plant spacing measurement. Transactions of the ASABE, 51(6), 2181 – 2191.

Chen, C., & Zheng, Y. F. (1995). Passive and active stereo vision for smooth surface detection of deformed plates.IEEE Transactions on Industrial Electronics, 42(3), 300 – 306.

Huang, S., & Huang, J. (2011). Vision guided dual arms robotics system with DSP and FPGA integrated system structure. Journal of Mechanical Science and Technology 25 (8) 2067~2076

Savarimuthu T. R., Kjær-Nielsen A., & Sørensen A. S. (2010). Real-time medical video processing, enabled by hardware accelerated correlations J Real-Time Image Proc, 6, 187–197

Veeken, M. V., Tang, L., & Hofstee, J. W. (2006, July). Automated corn plant spacing measurement at early growth stages using active computer vision. In ASABE 2006 Annual International Meeting, Portland, OR (paper 063059). St. Joseph, MI: ASABE.

Trucco, E., & Verri, A. (1998). Introductory techniques for 3-D computer vision (pp. 95–122, 145–149, 326–328). Englewood Cliffs, NJ: Prentice Hall.

Yovits, Marshall C. (1993). Advances in computers. 37. Academic Press. pp. 105–107.

Liptak, Béla G. (2006). Instrument Engineers' Handbook: Process control and optimization. 2. CRC Press. pp. 11–12.

Pearson, T. (2009). Hardware-based image processing for high-speed inspection of grains. Computers and Electronics in Agriculture, 69, 12–18.

Benkrid, K., Crookes, D., & Benkrid, A. (2002). Towards a general framework for FPGA based image processing using hardware skeletons. Parallel Computing, 28, 1141–1154

Capobianco L., & Garzelli, A. (2009). Target Detection With Semisupervised Kernel Orthogonal Subspace Projection. IEEE Transactions on Geoscience and Remote Sensing 47(11), 3822-3833