Adaptive Data Collection for Rapid Evaluation of New Plant Varieties

Katia Sycara (PI) and George Kantor
Carnegie Mellon University Robotics Institute

Carnegie Mellon THE ROBOTICS INSTITUTE

USDA United States Department of Agriculture National Institute of Food and Agriculture

This work was funded by the Agricultural Food Research Initiative competitive grant # 2017- 67007-26152 from the USDA National Institute of Food and Agriculture with the NSF Cyber Physical Systems program.

Challenges

- Ground-based phenotyping systems collect higher quality data than aerial systems, but have a lower coverage rate.
- Plant genotypes represent a non-spatial variable that must be considered when estimating phenotype distribution in a field.
- Measurements exhibit high stochasticity due to biological variations and sensor noise.
- Heterogeneous phenotype measurements have different levels of value and cost.

Approach

- Develop stochastic models (e.g., Gaussian Process (GP)) that capture the correlations between spatial and genetic variables.
- Develop informative path planning algorithms that optimally select sampling locations to update GP model.
- Develop multi-robot coordination algorithms that assign sampling positions for field coverage with limited communication between robots.

Scientific Impact

- Informative path planning provides a means of determining where and when to sample for a variety of scientific data collection tasks.
- Multi-robot coordination has relevance to decentralized decision making problems for static and mobile distributed systems.
- Employed models provide a potential framework for discovery of new phenotype-genotype associations.

Broader Impact

- Effectively increasing coverage rate of mobile phenotyping systems will improve crop breeding programs and provide producers with better crop monitoring for real-time management decisions.
- Educational impact: The project supports one full-time Ph.D. student (Wenhao Luo), one full-time M.S. student (Sumit Kumar), and PIs run K-12 education outreach (Girls of Steel)

Robotic Phenotyping

Approach:

- Wheeled ground robot autonomously navigates up and down rows of a breeding trial.
- Custom camera and processing collects images and extracts visual phenotypes while driving (stalk count, stalk width, light interception, leaf area, leaf necrosis, etc.)
- Robot occasionally stops to use arm to deploy contact measurements such as a stem penetrometer (stem stiffness) or a leaf spectrometer (chlorophyll, etc.)

Process:

- Current practice covers field exhaustively every two weeks.
- Measured phenotypes are associated with spatial plot locations and visualized in field maps.

Experimental Setting

- 4 acre sorghum breeding test plot in Clemson, SC (collaboration with Steve Kresovich, Clemson University)
- 250 accessions (varieties), 3 reps each, for a total of 750 plots
- Plots laid out in a rectangular field of 17 rows and 44 ranges (columns)

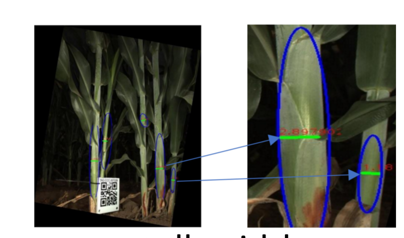
Ground Phenotyping Robot

Contact Measurements

leaf spectrometer

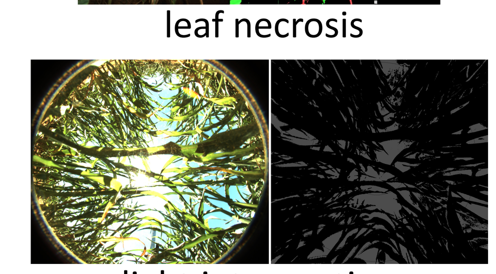
stem penetrometer

Example Visual Measurements



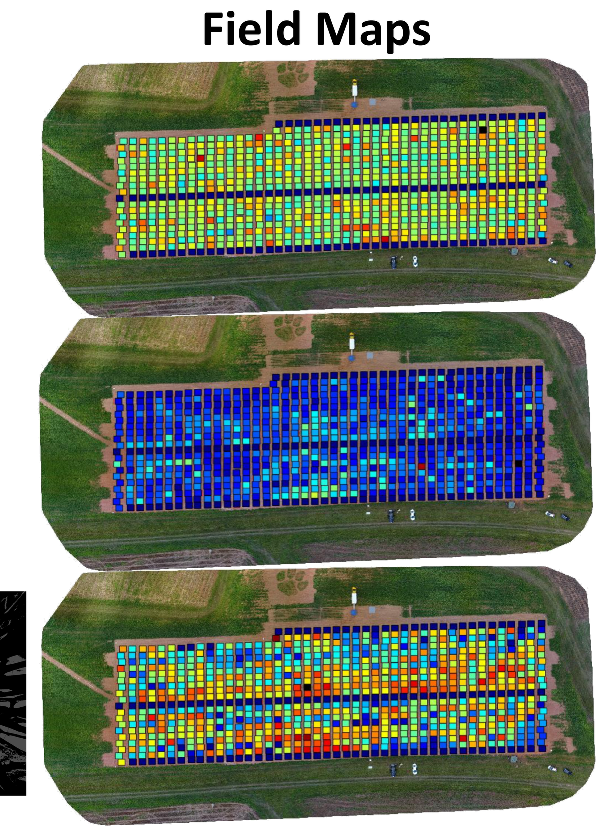
stalk width

leaf necrosis



light interception

Corresponding Field Maps



Informative Path Planning

Approach:

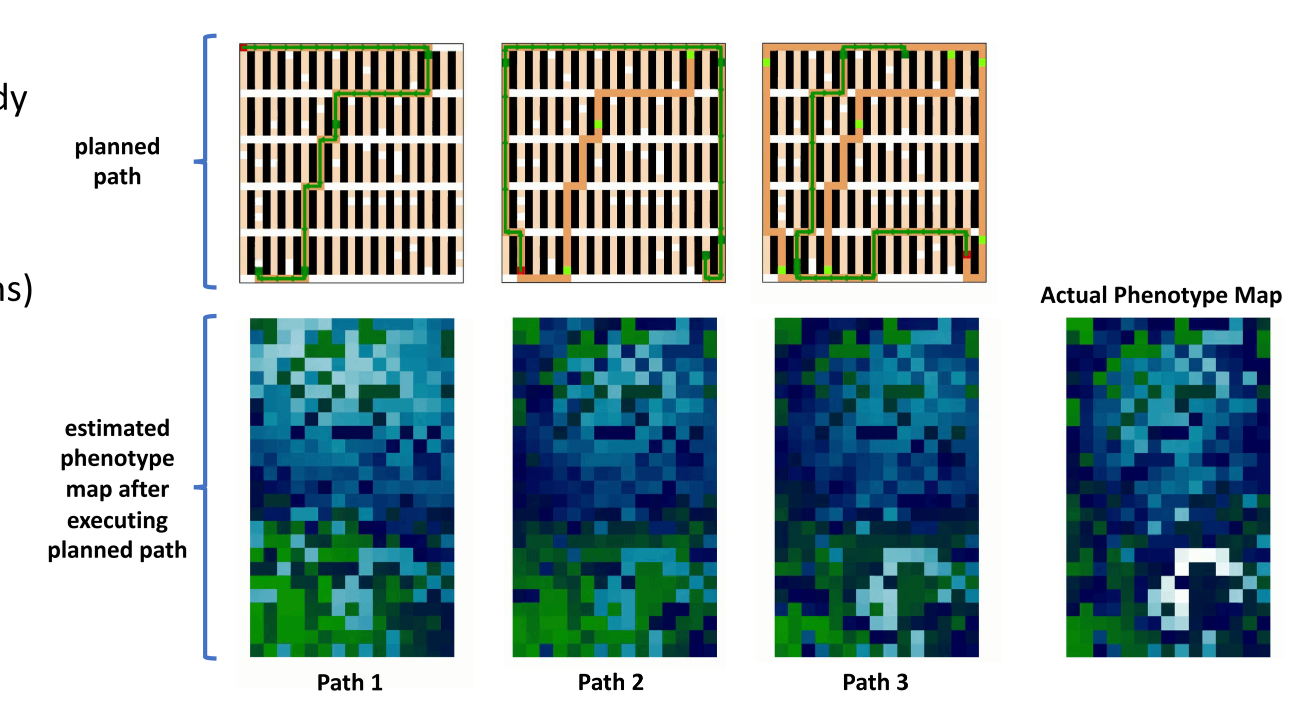
- Use a separate Gaussian Process (GP) spatial phenotype distribution model for each genotype.
- Identify a sequence of contact measurement sample locations using greedy information gain.
- Plan paths from one sample site to the next that maximize information gain (using visual measurements) while minimizing path length.
- Update GP estimate after each sample, using visual measurements along most recent path segment and contact measurement at sample location.

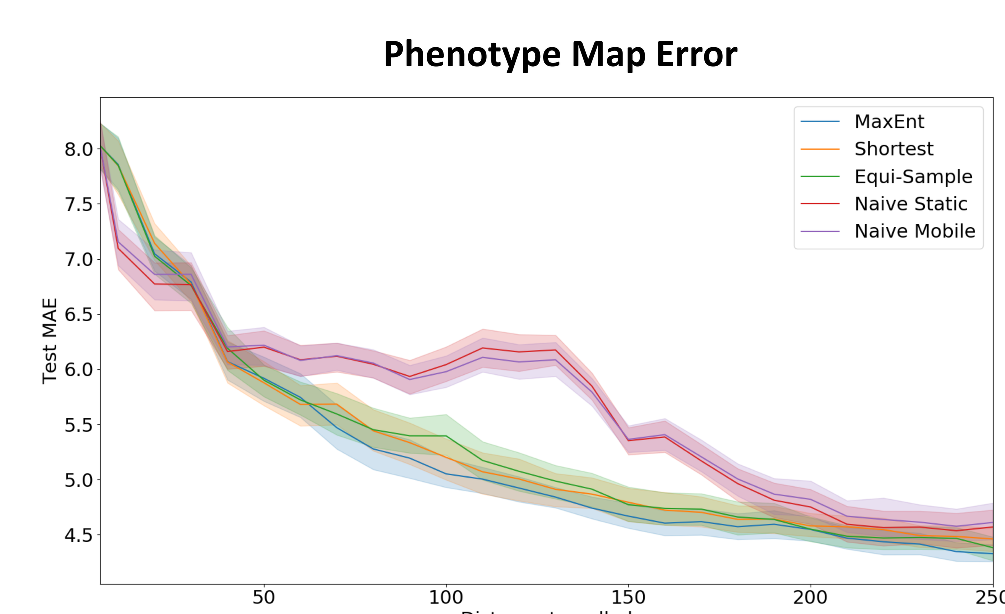
Process:

- Choose n sampling locations using greedy information gain.
- For each sampling location: plan path; execute path; sample; update GP.
- Repeat (choose n new sampling locations)

Setting (Simulated Data):

- 4 genotypes, spatial variation modeled by MOG for each genotype.
- Rectangular field, 15 rows, 25 ranges.
- Each plot is randomly assigned a genotype (~94 reps).
- Visual and contact measurements measure same phenotype, contact is higher fidelity.





Multi-robot Adaptive Sampling

Approach:

- Adaptive sampling with information-theoretic criterion for multi-robot coverage control.
- Efficient model learning and location optimization in an initially unknown environment.
- Collaboratively learning of the density function model using Mixture of GPs with hyper-parameters learned locally from each robot.
- Resulting GP mixture model provides improved prediction accuracy and reduced model uncertainty, increasing multi-robot coverage performance.

Process:

- Robots start a random initial positions.
- Initial distribution estimate is uniform.
- Robots follow adaptive coverage controller to learn model and approach Voronoi centroid.
- Robots take measurements at sample locations and update Mixture of GP model.
- Continue until robots converge at Voronoi centroids.

Setting and Assumptions:

- 54 node sensor network (Berkeley Intel Lab Data Set, 2004)
- 3 robots.
- Robots measure temperature when they are at a node location.
- Our approach is compared to uni-model GP approach.