WP 2.2 Development of an image analysis tool for large scale phenotyping

Chris Glasbey, Graham Horgan, Yu Song
Biomathematics and Statistics Scotland (BioSS)

Gerie van der Heijden, Gerrit Polder
Biometris, Wageningen UR
Wageningen, 6 March 2012

Plan of this talk:

1. Image analysis, locating and measuring plant parts (Yu)
   - Work done
   - Future plan

2. Statistical summary statistics (Graham)
   - Work done
   - Future plan

Previous Progress

- Automatically extract individual leaves by combining colour and Time-of-Flight images

Validation Trial in 2011

- To validate measurements on individual leaf area
- One double row in greenhouse
- 11 plots (one per genotypes)
- 8 leaves per plot were identified in image on 13 Sep 2011, and were harvested and measured on 14 Sep 2011.

# of automatically identified leaves: 59 (out of 88)
Correlation score: 0.98        RMSE: 9.50 cm
Second NL Trial in 2009

To find QTLs using leaf size and leaf angle of individual leaves.

Plant materials:
- 148 recombinant inbred lines from Yolo Wonder and CM 334.
- Standard double-row greenhouse setup in 4 compartments
- 151 genotypes, 264 plots in total

About 85,000 images were collected between 16 and 18 Sep 2009

Results

- Three vertical images per experimental plant were used, and there are overlapped views of experimental plants.
- All plots and genotypes were analysed.
- 11,790 leaves were automatically extracted in total, with a maximum of three leaves per image.
- Means per genotype were then calculated based on the experimental design.

Leaf size had a heritability of 0.70 (the ratio of genetic variance to the phenotypic variance)

Three QTLs were found, together explaining 29% of the variation.

Leaf angle:
- between the leaf and the vertical axis.
Leaf angle had only one QTL, which explained 11% of the total variation.

Future Plan
Combine developed methods in this WP to find fruits (with Graham)
- Identify individual fruits
- Number of fruits
- Their locations on plants
- Extract summary statistics on identified fruits
- Any relationship with QTLs of manual measurements

Image Statistics
- Instead of separately measuring plant parts, use statistical approach to derive features
  - having genotype differences / QTL or correlated with manual measurements
- Two features:
  - Plant height
  - Total leaf area

Plant height:
- Number of ‘green’ pixels (pepper plant) in a vertical direction.

Height Measurement for Tall Plants
- Several images to handle, filenames to match up etc.
- Top of plant might be in any of them
- Image overlap
- Sometimes plant in row behind is taller
- Occasional other problems
There are 4 images for each plant.
Top of plant might be any of the 4 images.
Estimating height

Max

1/3 of max

Estimated length

Sum of pixel greenness (G - 0.5B - 0.5R)

Height measurement for tall plants
Plant Height

- Image Analysis (IA) had a correlation of 0.93 with manual measurement (MANUAL) on plant height, and explained 43% of the total variation.

(Projected) Leaf area index

This is a measure of how much solar radiation the plant can intercept.

Colour histograms

Counts how many pixels in the image have each red, green, blue intensity.

Prediction vs manual

\[ R^2 = 0.64 \]

Total leaf area

Predicted project leaf area

Regression coefficients

Weight of each colour intensity count in predicting the leaf area index.

Multivariate histograms

- Count the number of times each combination of the three colour components occurs.
- Too many possibilities, so look at 8 or 16 bin ranges per component, leading to \( 8^3 = 512 \) or 4096 variables
- Principal component analysis to reduce the number of variables and avoid over-fitting
- Multiple regression on the first few components
The heritability of total leaf area was 0.55, and 20% of the variation was explained by QTLs.

2 QTLs agree with 2 of 3 found from manual measurements.

**Validation Trial in 2011**

- Frequent imaging (weekly and daily) and manual measurements
- Plant development over time can be analysed.

**Future work**

- Refine all of the current image statistics methods
- Look further into template matching for finding fruit (Yu Song / Graham)
- Look for QTLs and estimate heritability
- Publication

**Fruit Finding**

- Using some training images (templates) to find all the fruits in image.

**Record:**
1. Number of fruits
2. Fruit location
3. Summary statistics

- Relationship with QTLs of manual measurements
Produce an initial probability estimate of fruits.
Determine whether each region with a high probability is a fruit (or not).

Current problems:
1. Nearby fruits
2. Fruit behind leaf

Improvements will be made in the no-cost extension period.

The correlation between fruits seen in image and harvested fruits is 0.74.