Image Analysis for Automatic Phenotyping

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

Gerie van der Heijden, Gerrit Polder
Biometris, Wageningen UR

Wageningen, 7 March 2012
Manual phenotyping

Disadvantages:
- Slow and expensive
- Variation between observers
- Sometimes destructive
Phenotyping by Image Analysis

- Most image analysis systems for automatic phenotyping bring the plants to the camera.
Commercial tomato plants in Almeria, Spain

Pepper plants in our experiments

But for some crops, like pepper and tomato, this is not feasible
⇒ bring the cameras to the plants!
SPYSEE equipment
4* IR, Colour, Range (ToF) cameras
Plan

We aim to:

- Replace manual by automatic measurements
- Find new features, which are not possible or too difficult for manual measurement

Two approaches:
1. 3D
2. Statistical
1. 3D approach

3D information can be recovered from stereo pairs, because

\[
\text{Depth} = \frac{\text{constant}}{\text{disparity}}
\]
Stereo pair + ToF range image ⇒ detailed range image
Leaf in 3D ⇒ automatic measurement of size, orientation, etc
Validation trial (11 genotypes, 55 leaves):  Correlation = 98%  RMSE = 9.50cm²
Leaf size had a heritability of 0.70, three QTLs were found, together explaining 29% of the variation.
Leaf orientation:

- Angle between the leaf and the vertical axis.
Leaf orientation:

• Angle between the leaf and the vertical axis.
QTL analysis of automatically measured leaf orientation for 151 genotypes

- Heritability was 0.56, and one QTL explained 11% of the total variation
2. Statistical approach

Plant height estimated, from locations of ‘green’ pixels
Correlation 93% between automatic and manual plant heights
Total leaf area is a measure of how much solar radiation the plant can intercept.
Colour distribution

Counts how many pixels in the image have each red intensity
Colour distribution

![Colour distribution graph](image)
Colour histograms

Another example
Principal component regression

Call:
\[ \text{lm(formula = sep.leafarea} \sim \text{pr1}\$x[1:6], \text{na.action = na.exclude}) \]

Coefficients:

|        | Estimate | Std. Error | t value | Pr(>|t|) |
|--------|----------|------------|---------|----------|
| (Intercept) | 4.899e+03 | 6.288e+01  | 77.905  | <2e-16   *** |
| PC1    | 5.282e-02 | 5.473e-03  | 9.651   | <2e-16   *** |
| PC2    | 2.069e-01 | 1.875e-02  | 11.035  | <2e-16   *** |
| PC3    | -2.807e-01| 2.339e-02  | -12.002 | <2e-16   *** |
| PC4    | -5.750e-02| 3.477e-02  | -1.654  | 0.0997   .  |
| PC5    | 1.038e-02 | 3.686e-02  | 0.282   | 0.7785   |
| PC6    | 1.305e-01 | 5.607e-02  | 2.327   | 0.0209   * |

---

Residual standard error: 867.1 on 209 degrees of freedom
(1334 observations deleted due to missingness)
Multiple R-squared: 0.6419, Adjusted R-squared: 0.6317
F-statistic: 62.45 on 6 and 209 DF, p-value: < 2.2e-16

Number crunching to link colour histograms to manually measured total leaf area

Complex but standard methodology
Prediction vs manual

Correlation 80%
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 use bins of length 8 per component, leading to $16^3 = 4096$ variables
- Again do Principal Components regression
Multivariate histograms

Correlation 83%
The heritability of total leaf area was 0.55, and 20% of the variation was explained by QTLs. Two QTLs agree with 2 of 3 found from manual measurements.
Work in progress:

- Automatically find fruits
- Measure plant development
Summary

• The SPYSEE imaging setup records tall pepper plants while they are growing in a greenhouse

• Two approaches of automatic phenotyping have been explored:
  1. 3D
  2. Statistical

• QTLs have be found using our approaches, and good agreement with some manual measurements were achieved