Extracting and Plotting Features
A train/test scenario: Butterflies

Training images:
- *Polygonia c-album* (Comma)
- *Maniola jurtina* (Meadow brown)
- *Pyronia tithonus* (Gatekeeper)

Test image:
What label to give this butterfly?
How to describe a butterfly?
How to describe a butterfly?

- Colour?
- Patterns?
- Shape?
- Size?
- Number of Spots?
- Antennae?
- Legs?
- Location captured?
How to describe a butterfly?

- Colour?
- Patterns?
- Shape?
- Size?
- Number of...
  - Spots?
  - Antennae?
  - Legs?
- Location captured?
How to describe a butterfly?

Size?

Width

Height

Size?
How to describe a butterfly?

$x = (\text{width, height})$

Feature vector

$y = \text{species name}$

Feature (descriptor)

Target variable (label, class)
Visualising features

width = 6.4cm
height = 4.3cm
Visualising features
Visualising features
Visualising features
Longer feature vectors

\[ \mathbf{x} = (\text{width, height}) \]

2 dimensions

\[ \mathbf{x} = (\text{width, height, wing_area, latitude, longitude}) \]

5 dimensions

- We can’t plot this very easily…
- But everything else we show still applies
Getting started with the R practicals

1) Go to:

http://www0.cs.ucl.ac.uk/staff/O.MacAodha/ml_intro

2) Follow the dropbox link

3) Download as zip (or save to your dropbox):

4) Open RStudio

5) Navigate to the folder where the files are:
Loading and saving feature vectors
Loading and saving feature vectors

```r
flower_data <- read.csv('data/flower_dataset.csv')
```
Loading and saving feature vectors

flower_data <- read.csv('data/flower_dataset.csv')

Excel sheet with columns for Class labels and Feature vectors.
Loading and saving feature vectors
Loading and saving feature vectors

OSX: TextEdit
Windows: Notepad

csv = comma separated values
Loading CSV files into R

> butterfly_data <- read.csv('butterflies.csv')

> head(butterfly_data)

<table>
<thead>
<tr>
<th>Class</th>
<th>Width</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comma</td>
<td>6.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Comma</td>
<td>6.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Comma</td>
<td>6.3</td>
<td>4.4</td>
</tr>
<tr>
<td>Comma</td>
<td>6.5</td>
<td>4.4</td>
</tr>
<tr>
<td>Comma</td>
<td>6.4</td>
<td>4.4</td>
</tr>
<tr>
<td>Comma</td>
<td>6.7</td>
<td>4.8</td>
</tr>
</tbody>
</table>

> print(butterfly_data$Width)
Plotting data in R

```r
> plot(butterfly_data$Width, butterfly_data$Height, col=butterfly_data$Class, pch=16, asp=1)
```
Practical example

1_plot_data.R