Arboris folium

Image Recognition of Plant Leaves for Species Identification

by Hugh Pearse

Contact

Email: hughpearse@gmail.com
Phone: 0872167673

Abstract

The world contains countless hybrids and unusual combinations of living things. There are interspecific hybrids such as the Liger and even an interkingdom hybrid between the sea slug "Elysia chlorotica" and algae. Hybrids are not under threat from extinction but true species face the damages of environmental impact every day and there is an urgency to document their numbers and locations so that a methodical conservation plan can be devised. We have entered an age where identifying and tracking these endless combinations can be accelerated with software. In this document is described a series of techniques for extracting plant leaf biometrics from a photograph and indexing them in a search engine.

Approach

The biometrics can be extracted from leaves using a combination of reliable image manipulation techniques. Segmentation is a difficult step that reliably separates the leaf from the background of the image. In a scientific environment this process is made easier because we can rely on the users to take images in a way that suits the software by using a uniform background. Cropping the image so that the border fits against the side of the leaf is made easier by a clean segmentation. Any noise that has not been removed from the segmentation could potentially cause cropping to give a false result. After the segmentation and cropping the biometrics can easily be extracted without much additional effort.

There are several identification schemes commonly used by biologists such as a diagnostic key, a multi-access key and the comparison method but there are two principal techniques. The monothetic identification methods are methods where one characteristic is used at a time in a sequence to identify a species. Polythetic methods are ones where several characteristics are used simultaneously. For any polythetic identification method which calculates a measure of similarity, it is possible to use a numerical weighting for each biometric in the hope of improving the results.

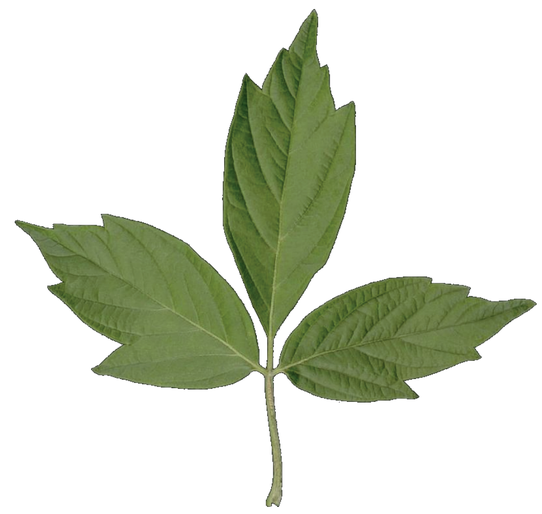
Conclusion

There are countless improvements that could be made to this project to increase accuracy and improve performance. The addition of more metrics would reduce the number of false positives.

In terms of accuracy, flower colour information and seed colour information to the database would be more true to real life as living organisms are classified by their reproductive systems. In terms of reliability, derived deviation information could be added to the parent tables to show the reliability of the averaged information in the temporal tables, this would be a deviation of the deviation data. In terms of performance, the weighting algorithm for calculating the total difference between a search request and a result could be improved by adding a precalculated hash table so that search requests do not need to compute the difference at the time of retrieval. This would mean that the database software is performing read operations instead of mathematical operations which means it can be done without using the CPU. Finally the reliability of the system is improved by the quantity of data in it. If the database contains thousands of leaves per variety then this will be averaged up into the parent tables.

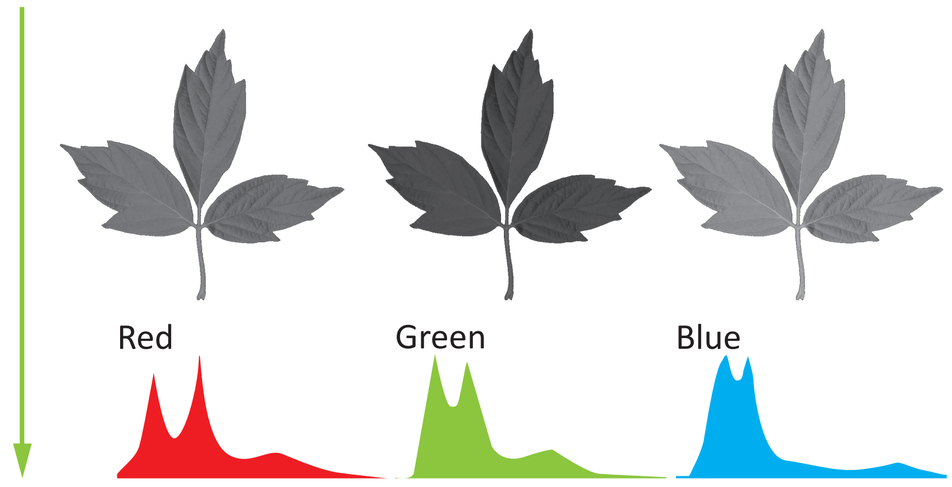
The system could also be modified to accommodate different priorities of requests. A request with an accuracy priority could perform computational calculations on the leaf table using the Crunch algorithm. This could take several hours depending on the size of the database. Similarly a request with a performance priority could perform the Skim algorithm on the averaged temporal tables which would minimize the amount of reads that need to be performed and minimizes the amount of computation that needs to be performed.

Input

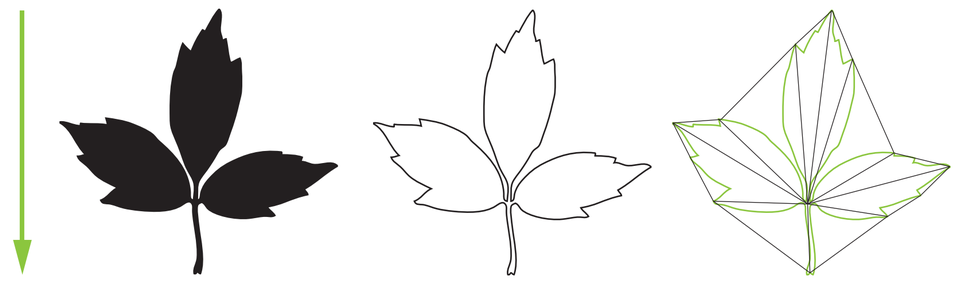


Extraction

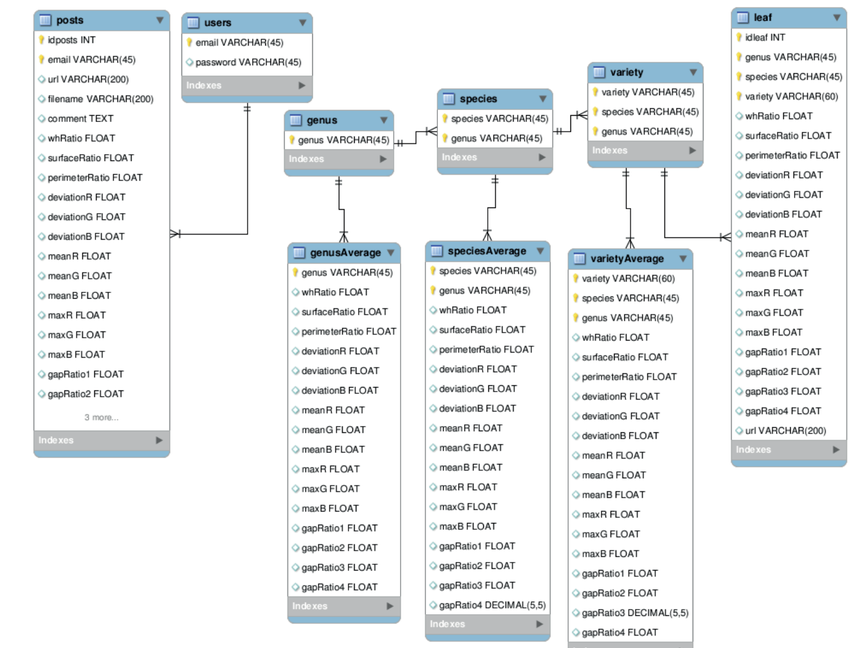
RGB Extraction



Surface Area, Perimeter, X Y Boundary Extraction



Database Search



Results

Searching for:

Search Results:

Genus	Species	Occurrences
Acer	negundo	14
Acer	griseum	9
Carya	ovata	9
Aesculus	flava	9
Eucommia	ulmoides	8
Cornus	kousa	8