Apr 10th, 1:00 PM - 5:00 PM

Leaf Identification Using Image Processing and Automatic Pattern Recognition

Thomas E. Humbert
Cedarville University, tehumbert@cedarville.edu

Joseph E. Niemiec
Cedarville University, jniemiec@cedarville.edu

Joshua M. Kaster
Cedarville University, jkaster@cedarville.edu

Leaf Identification Using Image Processing and Auto Pattern Recognition

Authors:
Thomas Humbert
Joshua Kaster
Joseph Niemiec

Consulted:
Timothy R. Tuinstra Ph.D.

Sample Images

Data from Sample Set of Images

Leaf Type	Image #	Perimeter	Rectangularity	Convex Ratio	Spectral Peak Location
Ginkgo	1	2906	1.39	1.26	4898
	2	2818	1.38	1.26	4789
	3	3009	1.37	1.30	4866
	4	3613	1.40	1.31	7727
	5	3107	1.40	1.23	5374
	6	3313	1.38	1.30	7011
	7	3421	1.43	1.44	8925
	8	2661	1.34	1.20	4492
Sweet Gum	9	3666	1.77	1.88	6375
	10	3479	1.69	2.21	9753
	11	3251	1.55	1.66	3970
	12	3477	1.82	2.33	6342
	13	3734	1.67	2.37	6874
Fig Tree	14	3417	1.51	1.31	5953
	15	2163	1.46	1.14	6628
	16	3666	1.41	1.33	7384
	17	2799	1.52	1.36	7611
Tulip Tree	18	2126	1.36	1.25	4471
	19	2809	1.42	1.62	7532
	20	2664	1.44	1.46	4559
	21	3554	1.37	1.31	5286
	22	2403	1.36	1.38	4831
	23	2931	1.38	1.52	4853
Sugar Maple	24	3664	1.46	1.69	6556
	25	3572	1.43	1.64	6018
	26	3096	1.40	1.61	5684
	27	3966	1.38	1.49	7080
	28	3499	1.50	1.66	5933
White Birch	29	3080	1.36	1.25	5438
	30	2596	1.31	1.18	4445
	31	2781	1.33	1.26	4519
	32	2707	1.34	1.22	5842
	33	2436	1.35	1.22	4937
	34	2954	1.34	1.26	5368
White Oak	35	3780	1.56	1.47	6664
	36	4093	1.88	1.86	7930
	37	3949	1.74	1.63	6659
	38	4829	1.65	1.48	8433
	39	4604	1.77	1.70	8241
	40	3209	1.45	1.27	5693
	41	3510	1.40	1.24	6031
Shagbark Hickory	42	4731	1.78	2.03	10647
	43	3868	1.81	1.65	11255
	44	4935	2.03	1.94	9600
	45	4824	1.97	2.04	8230
	46	5237	2.11	2.29	9281
Sassafras	47	2796	1.40	1.23	4891
	48	3253	1.57	1.28	5893
	49	2851	1.41	1.25	5252
	50	2961	1.67	1.41	5236
	51	2905	1.68	1.38	4993
	52	2890	1.60	1.33	5308
Flowering Dogwood	53	2353	1.30	1.08	4081
	54	2379	1.30	1.08	4551
	55	2049	1.29	1.05	3640
	56	2139	1.29	1.05	4133
	57	2348	1.29	1.05	4150

Introduction: We are fearfully and wonderfully made in the image of God. This is evidenced in how the human brain is able to process and make decisions based on what it sees. In this project we designed a system that identifies and classifies leaves just like a human would.

Problem: In order to distinguish between different leaves, a certain set of criteria had to be established. We needed to use as few leaf parameters as possible in order to correctly identify each leaf.

Materials and Methods: Using MatLab as a base, we utilized various image processing techniques to create the desired criteria. In order to accurately identify each leaf, we used a combination of four parameters. Rectangularity, convex area, perimeter, and the spectral peak location made up our identification criterion.

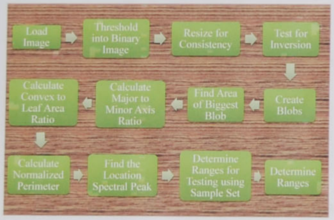
Rectangularity: This multiplies the longest possible combination of length and width of the leaf. It then takes this area and divides it by the actual leaf area. This tells us if how closely a leaf reflects the shape of a rectangle.

Perimeter: This measures the pixel distance around a leaf.

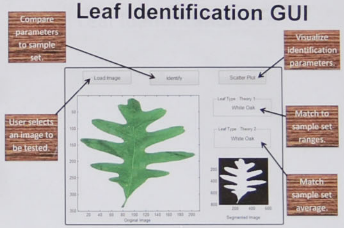
Convex Area: We first fill in all concavities native to each leaf so that we are left with a convex version of the leaf. We then measure the area of this convex leaf and divide it by the original leaf area. This gives us a ratio of convex leaf area to the original leaf area.

Spectral Peak Location: Visualize placing a point at the center of a leaf and then measuring the distance to the edge of the leaf. This distance will change as you rotate around the entire leaf. The distance rate of change is represented as a value in the frequency domain – a larger value for the spectral peak correlates to a higher rate of change.

Sampling Process


Load Image
Threshold into Binary Image
Resize for Consistency
Test for Inversion
Create Blobs
Find Area of Biggest Blob
Calculate Major to Minor Axis Ratio
Calculate Convex to Leaf Area Ratio
Calculate Normalized Perimeter
Find the Location Spectral Peak
Determine Ranges for Testing using Sample Set
Determine Ranges

Leaf Identification GUI



Results: With this project we were able to lay the essential ground work for an accurate leaf identification system. Building from this foundation, the system could be further developed to identify leaves for a mobile application.

The test image was categorized using two different theories: Theory 1 uses the minimum and maximums, determined from the sample set of images, to establish a range. After determining the different features, those values are then sorted within those ranges.

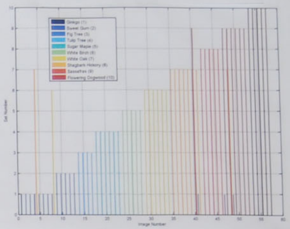
Theory 2 uses the same ranges as Theory 1 to determine the mean. From this mean, the smallest distance to the test image is determined. The closest one is the resulting category.

Even with two classification processes, we were only able to identify leaves with mediocre success. We found four features that gather a lot of information from the leaves, but we believe one more unique feature is needed to simulate the human perception. We realized there is not one "golden parameter" that can adequately categorize each leaf.

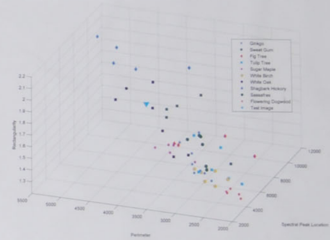
Despite using different theories to narrow down the selection, ties often occurred between each category of leaf. Creating an effective and robust tie-breaker would increase our accuracy.

We also utilized a GUI for easy interaction with the leaf identification system. By simply loading in the leaf and clicking 'Identify,' the user can know what the program determines.

Sample Set Output



3D Scatter of Sample Leaves with Test Image



Overall Results