Searching the World’s Herbaria: A System for Visual Identification of Plant Species

Chunxia Li
PhD Pre-Candidate
Construction Engineering & Management
Dept. of Civil & Environmental Engineering
University of Michigan

Silvio Savarese, PhD
Assistant Professor
Computer Vision Group
Dept. of Electrical and Computer Engineering
University of Michigan
Outline

- Introduction
- Dataset
- Segmentation & Shape Matching
- UI & Hardware
- Conclusion
Means to Access Data

- Digital on-line information:
  - Specimens in the US National Herbarium at the Smithsonian (9,000)
  - New York Botanical Garden (12,000)
  - Royal Botanical Gardens, Kew (50,000)
  - Missouri Botanical Garden (35,000)
  - The Encyclopedia of Life

- Visual Search
  - Neural network, centroid-contour distance
  - Hierarchical shape matching
  - IDSC

- Flow diagram of plant identification system
  - Photograph Leaf
  - Segment
  - Compute IDSC
  - Show Top Matches
Outline

- Introduction
- **Dataset**
- Segmentation & Shape Matching
- UI & Hardware
- Conclusion
Dataset

- **Current available dataset**
  - 15 species
  - 75 leaf images/species

- **Leaves processing:**
  - Collection
  - Flattened by pressing
  - Photographed with ruler and a color chart for calibration
  - Photographed with top/bottom lighting
  - Resize to a maximum dimension of 512 pixels
  - Automatic remove ruler, color chart and empty space
  - segmentation

- **Three datasets created:**
  - Flora of Plummers Island
  - Woody Plants of Baltimore-Washington, DC
  - Trees of Central Park
Outline

- Introduction
- Dataset
- Segmentation & Shape Matching
- UI & Hardware
- Conclusion
Segmentation

- Photograph on a plain white background
- Color-based EM algorithm
  - Map each pixel to SV
  - Separate pixels:
    - Discard pixels near boundary
    - Setting the background cluster to the mean of pixels near the boundary
    - Setting the foreground cluster to the mean of the central pixels

- SUM
  - Feature selection
  - Initialization
  - Sampling
  - Segment classification
Shape Matching

- IDSC (Inner Distance Shape Context) for matching
  - Sample points around along the boundary
  - 2D histogram descriptor at each point
    - Angle & distance from each pt to other pts

- To compare two leaves
  - Sample points
  - m different uniformly space locations

- Nearest neighbor classifier

Experimental results for two datasets
Outline

- Introduction
- Dataset
- Segmentation & Shape Matching
  - **UI & Hardware**
- Conclusion
User Interface & Hardware

- Mobile Computing Platform
  - Browse
  - Sample
  - Search results
  - History

- Augmented Reality
  - UMPC
  - ARToolkit
  - ARTag
  - Notebook USB2.0 camera
  - Head work display

- System evaluation (Rock Creek Park, Washington DC)
  - User studies
  - Field tests
Outline

- Introduction
- Dataset
- Segmentation & Shape Matching
- UI & Hardware
- Conclusion
Conclusion

- Complete Description of CV system & UI
- 3 datasets

Future plans:
- Expand the coverage of the system
- Develop a touch based version on an iPhone or Android-based device