BugID: Rapid-Throughput Insect Population Counting Using Computer Vision


Faculty: R. Paasch, A. Moldenke, D. A. Lytle, E. Mortensen, L. G. Shapiro, T. G. Dietterich

Oregon State University
University of Washington
Ecological Sciences Need Better Instruments

- **Scientific Questions:**
  - How do communities of organisms interact?
  - Why does increased biodiversity create more resilient ecosystems?
  - How do invasive species succeed? How can they be stopped?

- **Instrumentation Requirements:**
  - What is the temporal and spatial distribution of organisms?
  - What is the population size of each species in a region?
  - What are all of the interactions (e.g., eating, mating) between pairs of organisms?
Ecological Sciences Need Better Instruments

- **Scientific Questions:**
  - How do communities of organisms interact?
  - Why does increased biodiversity create more resilient ecosystems?
  - How do invasive species succeed? How can they be stopped?

- **Instrumentation Requirements:**
  - What is the temporal and spatial distribution of organisms?
  - What is the population size of each species in a region?
  - What are all of the interactions (e.g., eating, mating) between pairs of organisms?
Example 1: Stonefly populations in freshwater streams

- differentially sensitive to many pollutants
- live in river substrate – reliable indicator of stream health
- difficult and expensive for people to classify (particularly to genus or species levels)
Example 2: Small arthropods in soil
“soil mesofauna”
Soil Arthropods

- Measure biodiversity of soils
  - response of soil biodiversity to
    - forest/agricultural practices
    - disease
    - climate change

- ~2000 possible species
  - ~100 species in any single sample
Goal: Rapid-Throughput Automated Arthropod Population Counting

- Specimens are manually collected in the field following standard protocols
- Automated (robotic) devices photograph the specimens
- Images are then classified to family, genus, or species level
Computer Vision Challenges (1)

- Highly-articulated objects with deformation
Computer Vision Challenges(2)

- Huge intra-class changes of appearances due to development and maturation

![Images showing changes in appearance from tergites to wings](image-url)
Computer Vision Challenges (3)

- Small between-class differences

Calinueria

Doronueria
Human Test

- Identifying them (from whole-specimen images) is hard even for humans

For Calinueria vs. Doronueria:

- Mean Classification accuracy: 78.6%
- Standard Deviation: 8.4
Modern Computer Vision Approaches

- **Image based methods**
  - Assumes objects are aligned across all images
  - Eigenfaces

- **Model based**
  - "inverse graphics": parameterized model that can generate a predicted image
  - adjust parameters to obtain best match with the actual image
  - computation is very slow and expensive

- **Patch based**
  - identify "interesting regions"
  - extract and describe these regions
  - learn to recognize based on these regions
Machine Learning

Training Examples

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Calineuria</td>
</tr>
<tr>
<td></td>
<td>Calineuria</td>
</tr>
<tr>
<td></td>
<td>Doroneuria</td>
</tr>
<tr>
<td></td>
<td>Doroneuria</td>
</tr>
</tbody>
</table>

Learning Algorithm

Classifier

New Examples

Doroneuria
Machine Learning Requires Fixed Number of “Features”

- To apply machine learning algorithms, we need to convert the image of a bug into a fixed number of “features”
- Idea: Create a “visual dictionary” of “parts” and then count parts
Count each detected “part” into a “feature table”

<table>
<thead>
<tr>
<th>abdomen</th>
<th>head</th>
<th>eye</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>1</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Stonefly Identification System

- Semi-automated specimen handling and photography
- Computer Steps:
  1. Dorsal view detection (not yet implemented)
  2. Region detection
  3. Region description
  4. Region mapping into features
  5. Combination into a “feature table”
  6. Classification
Semi-Automated Specimen Handling
Semi-Automated Specimen Handling (2)

- Specimens inserted here
- Infrared detectors
- Mirror assembly produces image pair
- Trapped by side jet
- Jet spins them intermittently
- Ejected here after classification
Semi-Automated Specimen Handling (3)

- Photographed by a 5 megapixel QImaging camera
- Leica microscope at 0.63x magnification
- Mirrors capture two views per image
Region Detectors

Hessian-Affine Detector  Kadir Entropy Detector  PCBR Detector
Scale-Independent Feature Transform
SIFT (Lowe, 1999)

- Compute intensity gradient at each pixel in 16x16 region
- Weight them by a gaussian distribution (indicated by circle)
- Collect into histograms within each 4x4 region (gives 16 histograms)
- Result: 128-element vector normalized to have Euclidean norm 1
Cluster Analysis of the SIFTs

- Automatically group together similar regions to build a “visual dictionary” for each species

100 clusters
Count each detected “part” into a “feature table”

| 4 | 2 | 0 | 4 | 2 | 5 | . | . | . | 1 | . | . | 2 | 1 | 2 |
Learning Algorithm

- Bagged Logistic Model Trees

\[
P(Cal) = \logit(w_1 \cdot X) \quad P(Cal) = \logit(w_2 \cdot X) \quad P(Cal) = \logit(w_3 \cdot X) \quad P(Cal) = \logit(w_4 \cdot X)
\]
Experimental Study

Calineuria

Doroneuria

Hesperoperla

Yoroperla
Experiment Design

- Data set

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Specimens</th>
<th>Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calineuria</td>
<td>85</td>
<td>400</td>
</tr>
<tr>
<td>Doroneuria</td>
<td>91</td>
<td>463</td>
</tr>
<tr>
<td>Hesperoperla</td>
<td>58</td>
<td>253</td>
</tr>
<tr>
<td>Yoraperla</td>
<td>29</td>
<td>124</td>
</tr>
</tbody>
</table>

- Split into 3 parts by specimens, balanced by taxon
  - clustering
  - training
  - testing
Results

- Overall accuracy: 82.42% ($\pm 2.14$)
- Confusion matrix

<table>
<thead>
<tr>
<th>True Class</th>
<th>Predicted Class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cal.</td>
</tr>
<tr>
<td>Calineuria</td>
<td>313</td>
</tr>
<tr>
<td>Doroneuria</td>
<td>87</td>
</tr>
<tr>
<td>Hesperoperla</td>
<td>24</td>
</tr>
<tr>
<td>Yoroperla</td>
<td>1</td>
</tr>
</tbody>
</table>
Discussion

- Discriminating Cal. vs. Dor. is very difficult
  - Human study: 26 students and faculty from OSU Entomology program
    - trained on 50 Cal and Dor images
    - tested on 50 Cal and Dor images
    - 78.6% correct
  - Comparison:
    - our system trained only on Cal. Vs. Dor. is 79.37% correct (±2.70)
  - If we combine the Cal. and Dor. classes, then our system is 95.40% correct (±1.20)
Why Three Region Detectors?

- The combination of all three gave better results than any other configuration:

<table>
<thead>
<tr>
<th>Hessian Affine</th>
<th>Kadir Entropy</th>
<th>PCBR</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>√</td>
<td></td>
<td></td>
<td>73.14</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td></td>
<td>70.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>√</td>
<td>71.69</td>
</tr>
<tr>
<td>√</td>
<td>√</td>
<td></td>
<td>78.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>√</td>
<td>80.48</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>√</td>
<td>78.31</td>
</tr>
<tr>
<td></td>
<td>√</td>
<td>√</td>
<td>82.42</td>
</tr>
</tbody>
</table>
The Same Method Works On Other Tasks

- Caltech 2-class data sets

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Airplanes</td>
<td>99.2</td>
<td>93.7</td>
<td>88.9</td>
<td>98.8</td>
<td>98.0</td>
</tr>
<tr>
<td>Faces</td>
<td>98.4</td>
<td>91.7</td>
<td>93.5</td>
<td>99.5</td>
<td>99.5</td>
</tr>
<tr>
<td>Motorbikes</td>
<td>98.3</td>
<td>96.7</td>
<td>92.2</td>
<td>99.5</td>
<td>96.7</td>
</tr>
<tr>
<td>Leopards</td>
<td>98.0</td>
<td>89.0</td>
<td></td>
<td>91.0</td>
<td></td>
</tr>
<tr>
<td>Cars (Rear)</td>
<td>95.5</td>
<td>91.2</td>
<td>91.1</td>
<td></td>
<td>94.5</td>
</tr>
</tbody>
</table>

- UIUC cars versus no-car background

<table>
<thead>
<tr>
<th>Dataset</th>
<th>IDC-BDL</th>
<th>Fergus (IJCV 2007)</th>
<th>Opelt (PAMI 2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cars (Side)</td>
<td>92.7</td>
<td>88.5</td>
<td>83.0</td>
</tr>
</tbody>
</table>
Other Datasets (2)

- GRAZ-01 2-class problems

<table>
<thead>
<tr>
<th>Dataset</th>
<th>IDC-BDL</th>
<th>Opelt (PAMI 2006)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bikes</td>
<td>76.5</td>
<td>73.5</td>
</tr>
<tr>
<td>Persons</td>
<td>71.7</td>
<td>63.0</td>
</tr>
</tbody>
</table>
New Results: 9 Stonefly Taxa

- **Classifier Adaboost M1**
  - 65 Iterations
  - Base Classifier: Pruned C4.5 trees

<table>
<thead>
<tr>
<th>Classified as</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cal</td>
<td>75.61</td>
<td>12.60</td>
<td>0.41</td>
<td>7.11</td>
<td>0.41</td>
<td>0.61</td>
<td>0.81</td>
<td>0.00</td>
<td>2.44</td>
</tr>
<tr>
<td>Dor</td>
<td>12.41</td>
<td>79.51</td>
<td>2.63</td>
<td>4.70</td>
<td>0.00</td>
<td>0.00</td>
<td>0.75</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Hes</td>
<td>1.22</td>
<td>3.87</td>
<td>92.46</td>
<td>0.41</td>
<td>0.00</td>
<td>0.61</td>
<td>1.22</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Iso</td>
<td>3.00</td>
<td>1.60</td>
<td>0.20</td>
<td>89.60</td>
<td>0.40</td>
<td>0.00</td>
<td>3.40</td>
<td>0.00</td>
<td>1.80</td>
</tr>
<tr>
<td>Mos</td>
<td>5.22</td>
<td>2.61</td>
<td>0.87</td>
<td>2.61</td>
<td>49.57</td>
<td>3.48</td>
<td>3.48</td>
<td>1.74</td>
<td>30.43</td>
</tr>
<tr>
<td>Pte</td>
<td>0.00</td>
<td>0.90</td>
<td>4.05</td>
<td>0.45</td>
<td>0.45</td>
<td>73.42</td>
<td>6.76</td>
<td>4.05</td>
<td>9.91</td>
</tr>
<tr>
<td>Swe</td>
<td>2.71</td>
<td>1.04</td>
<td>0.21</td>
<td>5.01</td>
<td>0.42</td>
<td>1.25</td>
<td>86.01</td>
<td>1.88</td>
<td>1.46</td>
</tr>
<tr>
<td>Yor</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.61</td>
<td>2.64</td>
<td>93.09</td>
<td>3.25</td>
</tr>
<tr>
<td>Zap</td>
<td>1.20</td>
<td>0.00</td>
<td>0.20</td>
<td>0.60</td>
<td>0.20</td>
<td>1.20</td>
<td>1.41</td>
<td>3.21</td>
<td>91.97</td>
</tr>
</tbody>
</table>

Acc: 3269 85.6%

Tot: 3821
Future Work

- Dorsal view detector
- Mechanical sorting on the output
- Reject unknown species
- Propose field testing to EPA
- License for commercialization
Soil Mesofauna

- ~2000 species possible, but only ~100 in any given sample
- motorized X-Y stage and focus
- step focus through 12 levels and combine to produce high depth-of-field synthetic image
- vary background lighting to compute transparency
- employ shape in addition to appearance
- robot arm with pipette will extract each identified specimen and place in 96-well plate