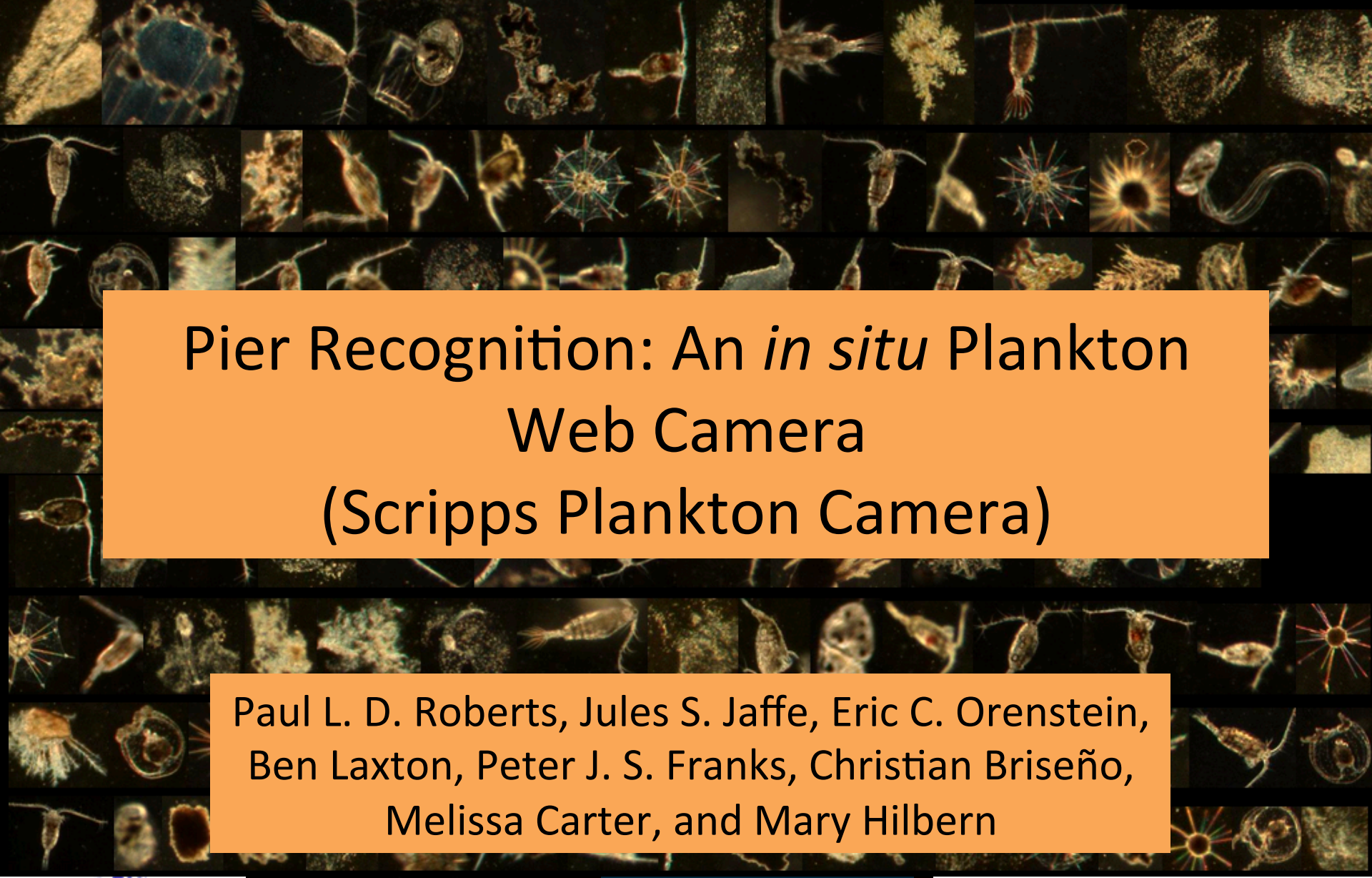


Pier Recognition: An *in situ* Plankton Web Camera (Scripps Plankton Camera)

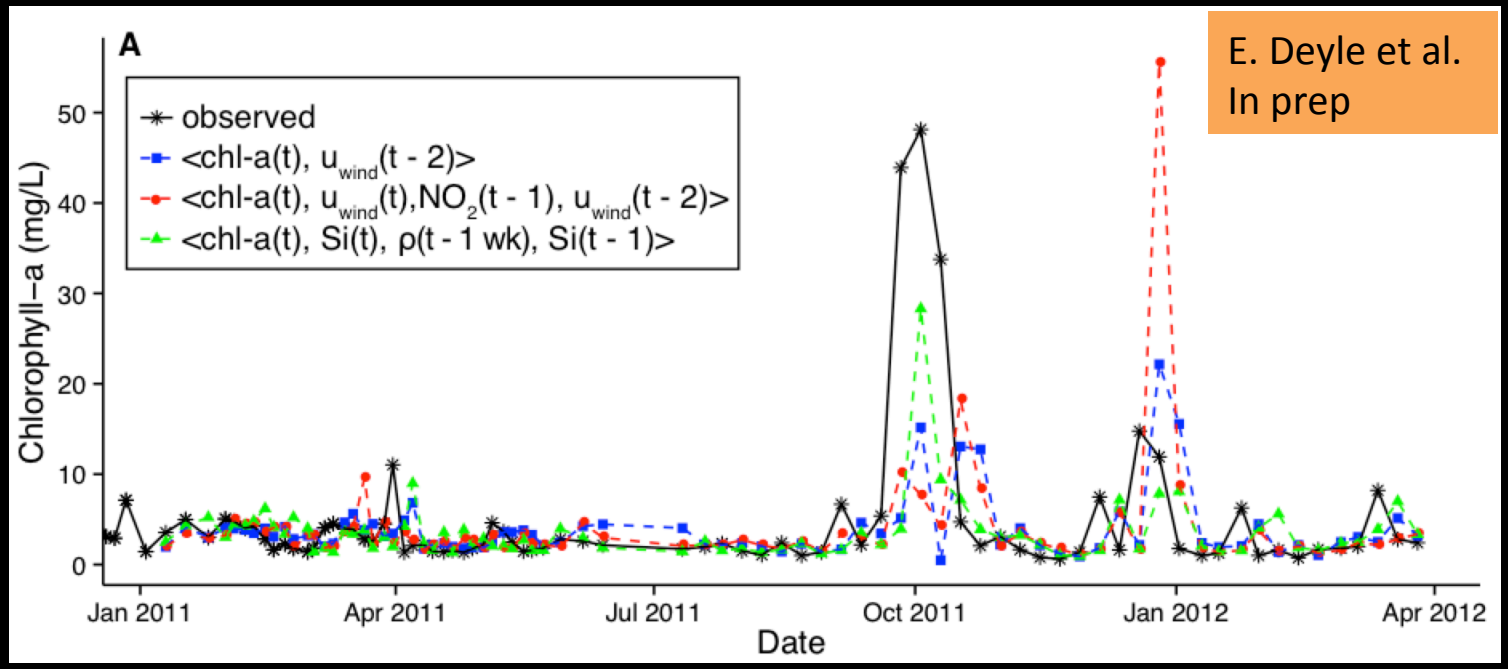
Paul L. D. Roberts, Jules S. Jaffe, Eric C. Orenstein, Ben Laxton, Peter J. S. Franks, Christian Briseño, Melissa Carter, and Mary Hilbern

MARINE PHYSICAL LABORATORY
UCSD
IOOS INTEGRATED OCEAN OBSERVING SYSTEM
SCRIPPS INSTITUTION OF OCEANOGRAPHY
UC San Diego

Outline

- Scripps Plankton Camera (SPC) Development
- Preliminary Automated Image Annotation
- Summary And Future Work

Motivation



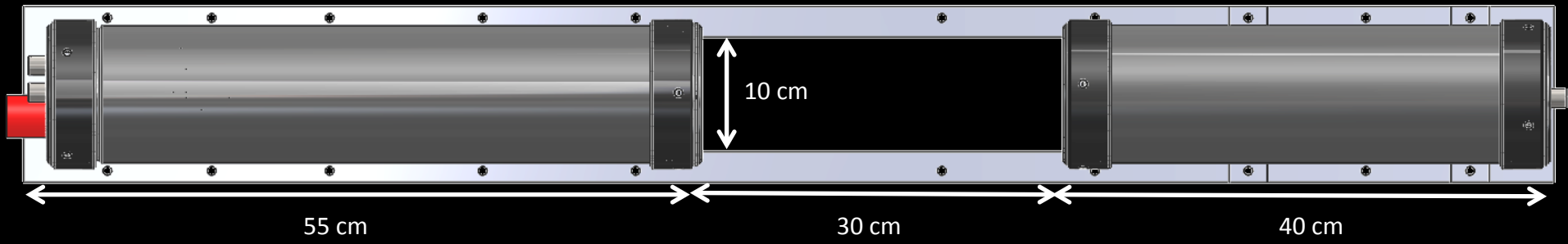
- Scripps Pier Time Series
  - A rich time series with a long history.
- Understanding and predicting the triggers of plankton blooms is still an open problem.
- There is a need for image data and rapid sampling/processing to improve our understanding of these processes.

SPC Imaging System Design Objectives

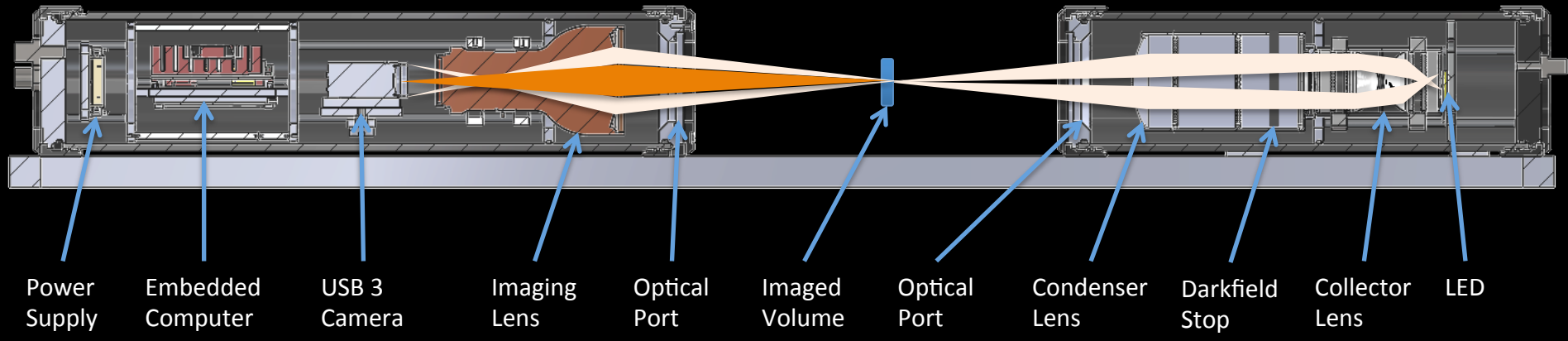
- A moderate cost ($10k-$30k) underwater, darkfield, microscope system.
- Use off-the-shelf parts supporting a wide range of customization of magnification and illumination.
- Image fragile plankton taxa without breaking them apart.
- Take advantage of modern embedded processors for low-power, real-time processing.
- **Real-time, interactive data available to all via a modern web application**.

SPC Hardware

Top View



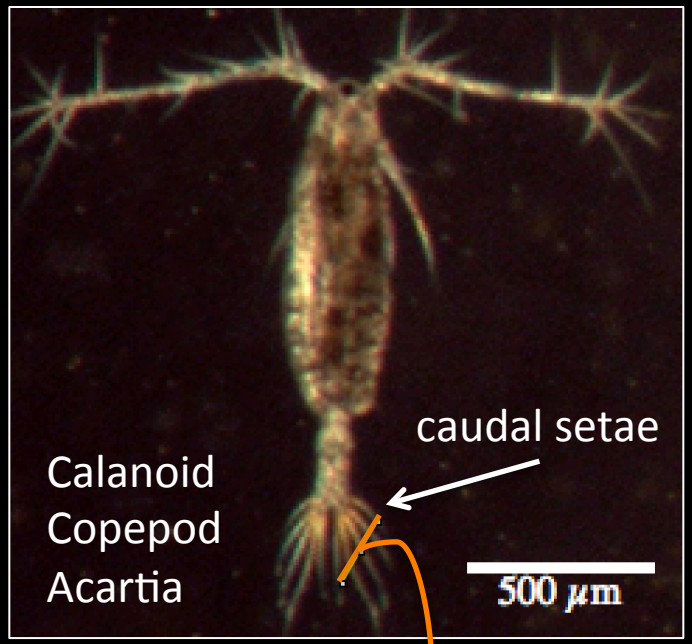
Section View

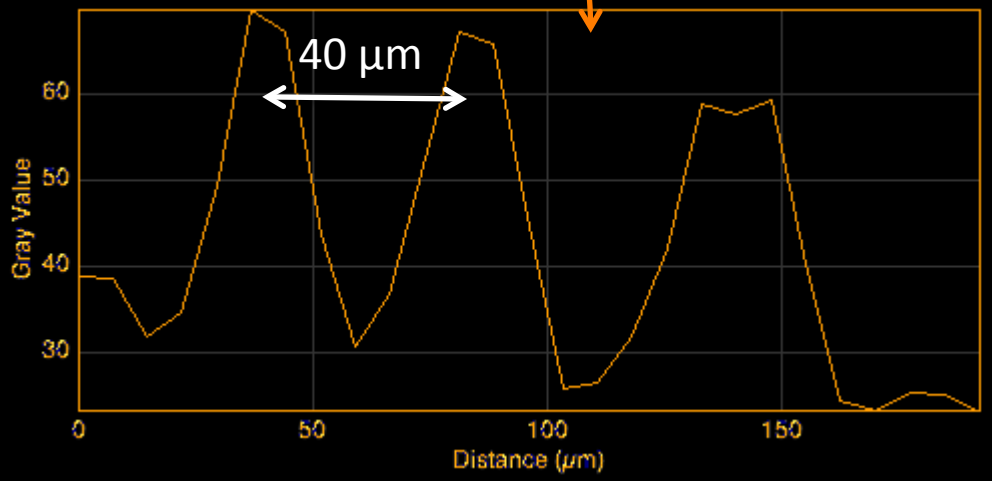


- Dual-housing design allows for wide range of working distances
- Imaging lens and Illumination NA can be changed easily.

SPC Imaging Performance

Property	Specification
Field of View	25 mm x 20 mm
Resolution	7.4 µm pixels 35 lp/mm @ 40 % contrast
Depth of Field	400 µm @ 35 lp/mm @ 20 % contrast
Hi-Resolution Volume	0.2 mL per Frame
Blob-Detection Volume	10 mL Frame
Data Rate	Up to 8 fps with ROI processing





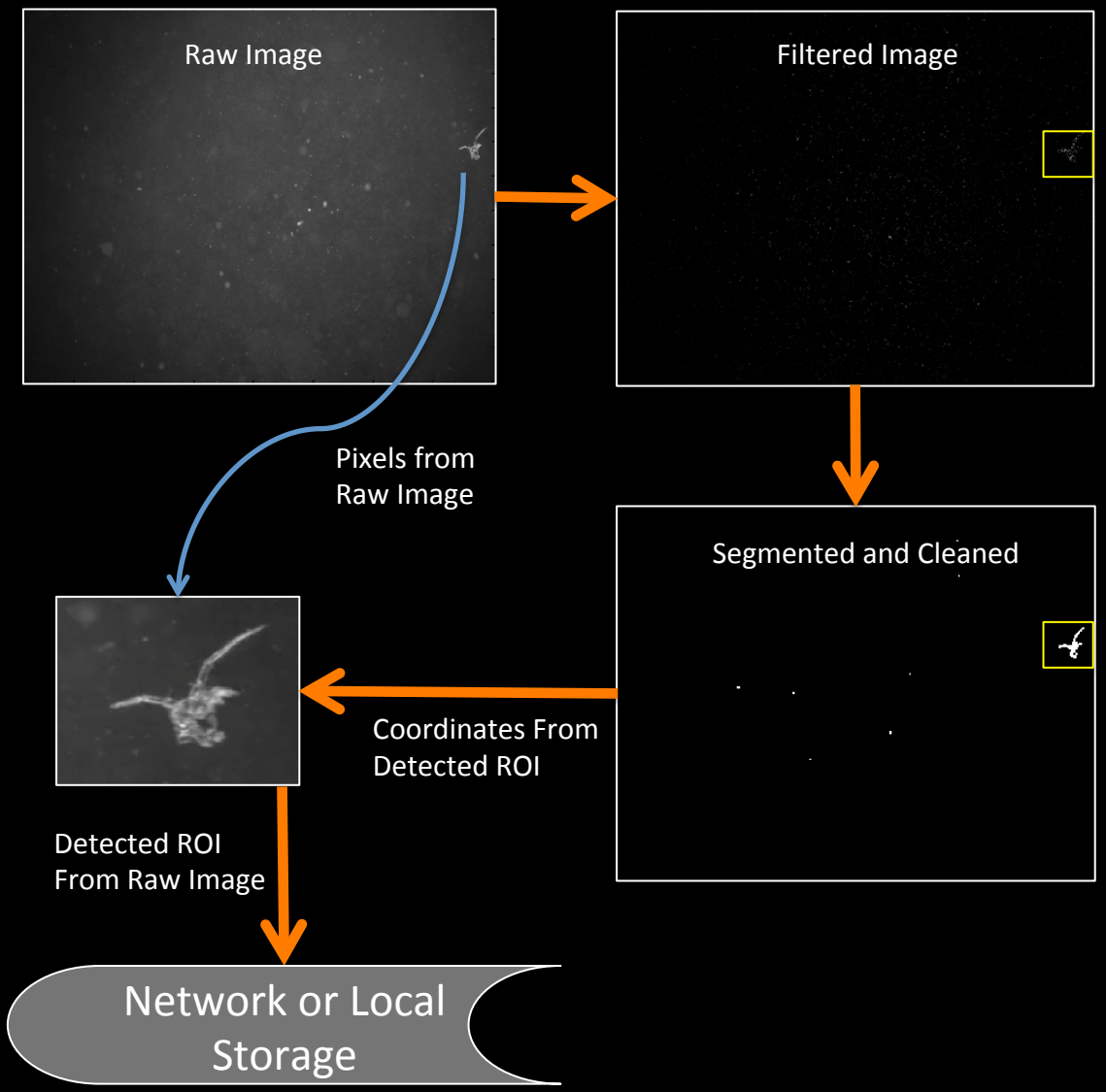
Comparison With Lab Microscopy

Scripps Plankton Camera (*in situ*)
https://jaffeweb.ucsd.edu:5001

Scripps Institution of Oceanography Zooplankton Guide
https://scripps.ucsd.edu/zooplanktonguide/species/acartia-acanthacartia-tonsa

SPC Acquisition Software

- Raw images are acquired continuously at a given frame rate (up to 8 Hz).
- 9 megapixel images are down-sampled by a factor of 4.
- Down-sampled images are filtered with a Laplacian filter.
- Output of the filtered images are segmented and binary regions are detected.
- The original raw pixel values from the detected binary regions are saved to local disk or over the network.



Below is an interactive tool for browsing images. Select a range of min and max lengths, number of images to display, and block of time. Loading status and database statistics are shown below.

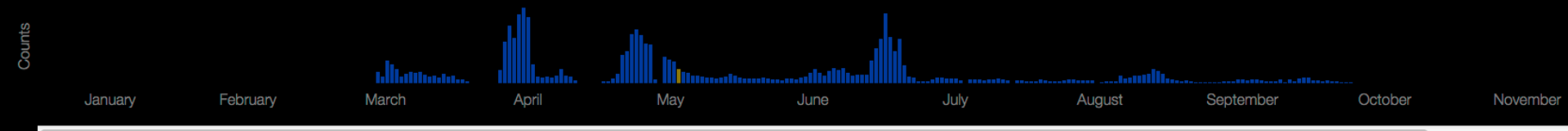
ALPHA v0.3: PRELIMINARY DATA, UNVERIFIED, MAY HAVE SIGNIFICANT ERRORS

Loaded 99 images.

Toggle Timeline

This timeline displays the number of counts recorded by the SPC per day.

SELECTED: 106016 counts on 5/17/14



Last 01 Hours | Last 02 Hours | Last 03 Hours | Last 12 Hours | Last 24 Hours | Last 36 Hours | Last 48 Hours | Last 7 Days | All Available | Show One Hour of Data From [ ] Hours Ago.

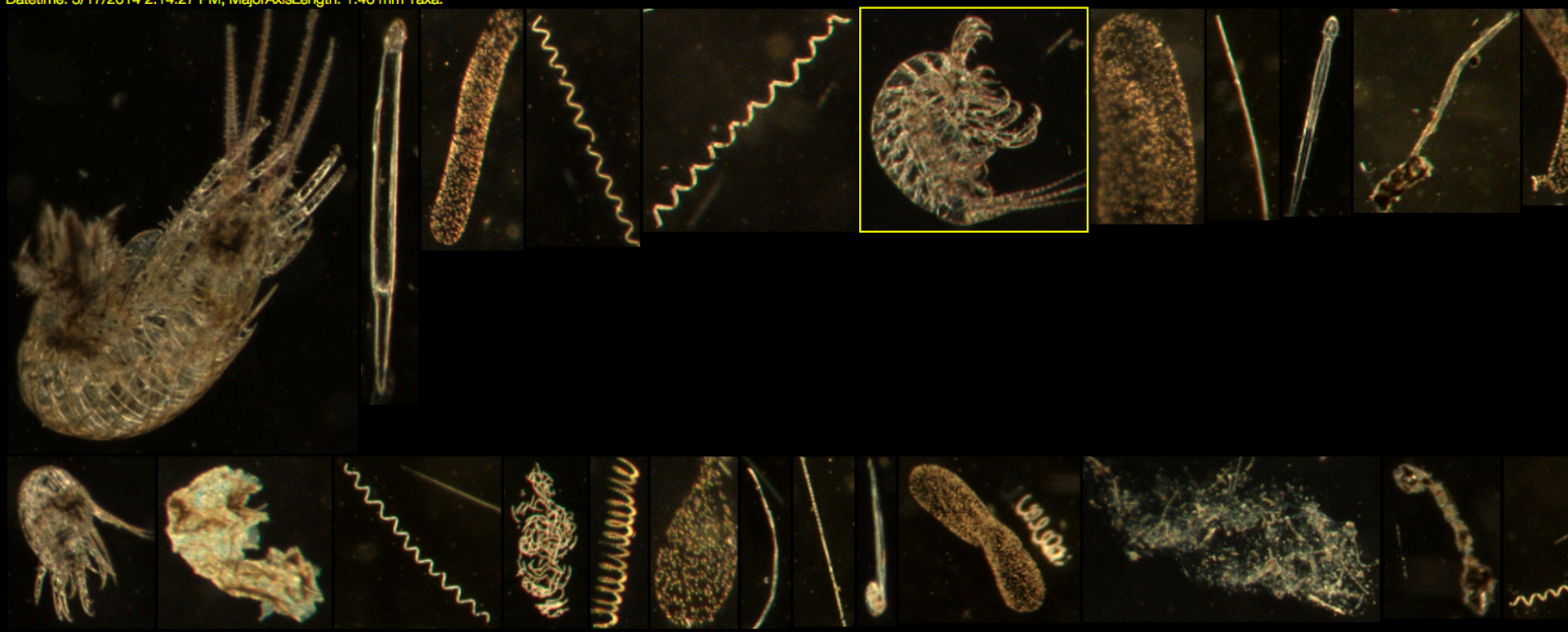
Images to Display [100] Minimum Length (mm) [1.04] Maximum Length (mm) [3.47] Minimum Aspect ([0,1]) [0.00] Maximum Aspect ([0,1]) [1.00]

User Taxa [All Taxa] Machine Taxa [All Taxa]

Classifier Selection Stuff: Classifier [Select Classifier] Taxa [Any Taxa] Percentile [0]

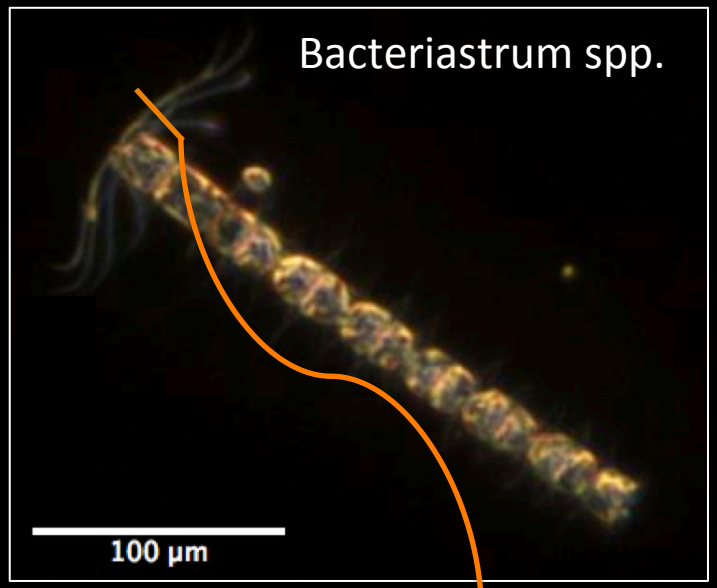
Selection Controls: [Clear Selected] [Tag Selected] [Helix] [Invert Selected]

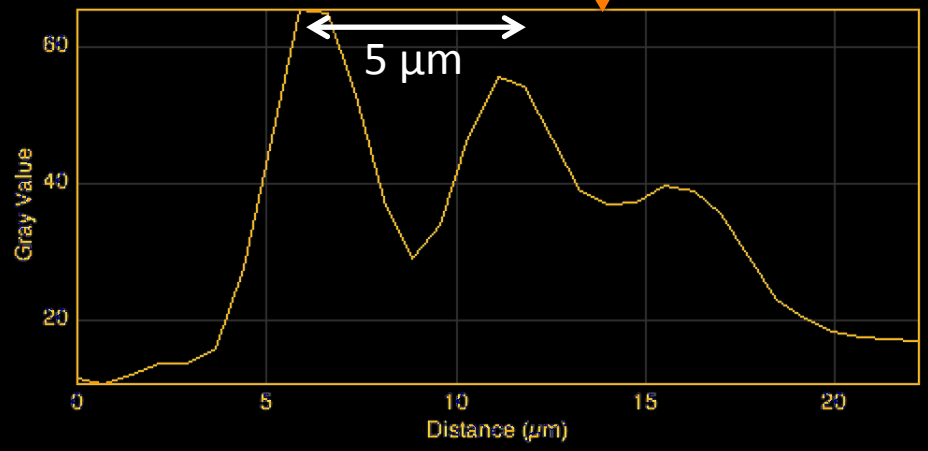
Datetime: 5/17/2014 2:14:27 PM, MajorAxisLength: 1.46 mm Taxa:



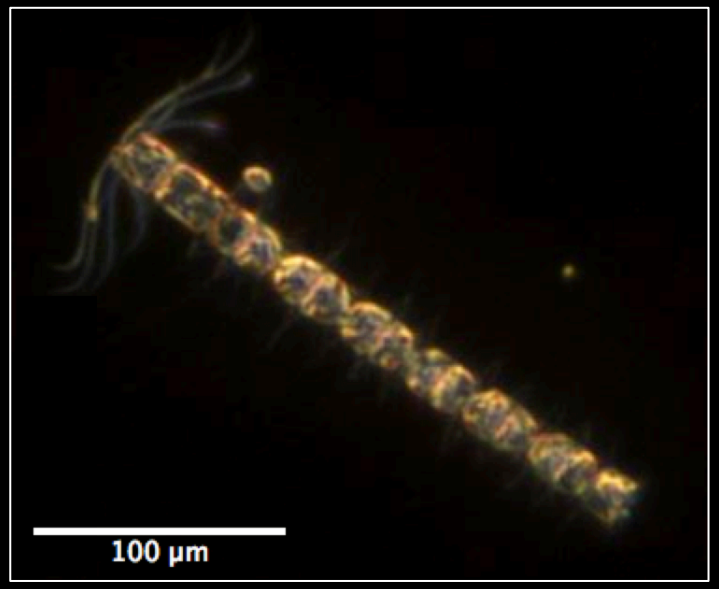
Recent Addition:
SPC-P: Microplankton Camera

Property	Specification
Field of View	2.5 mm x 2.0 mm
Resolution	0.74 µm pixels 200 lp/mm @ 40 % contrast
Depth of Field	20 µm @ 200 lp/ mm @ 20 % contrast
Hi-Resolution Volume	0.1 µL per Frame
Blob-Detection Volume	10 µL Frame
Data Rate	Up to 8 fps with ROI processing

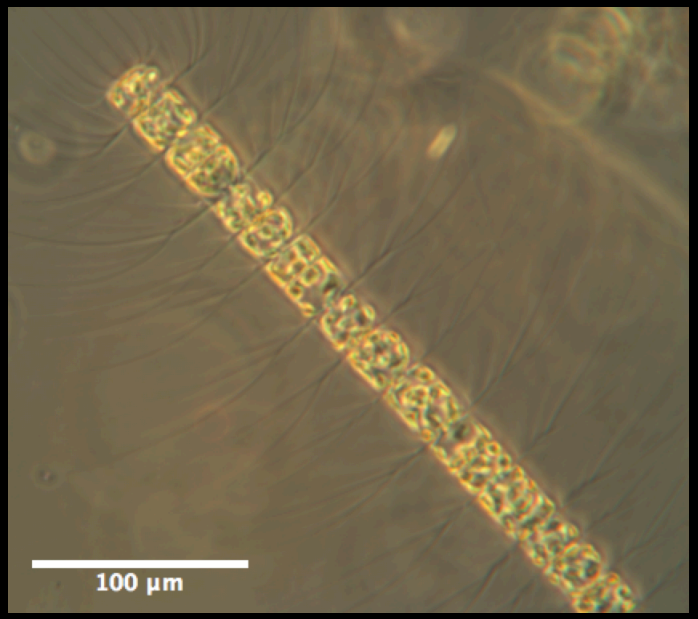




Comparison With Lab Microscopy



Scripps Plankton Camera - P (*in situ*)
http://jaffeweb.ucsd.edu:5055

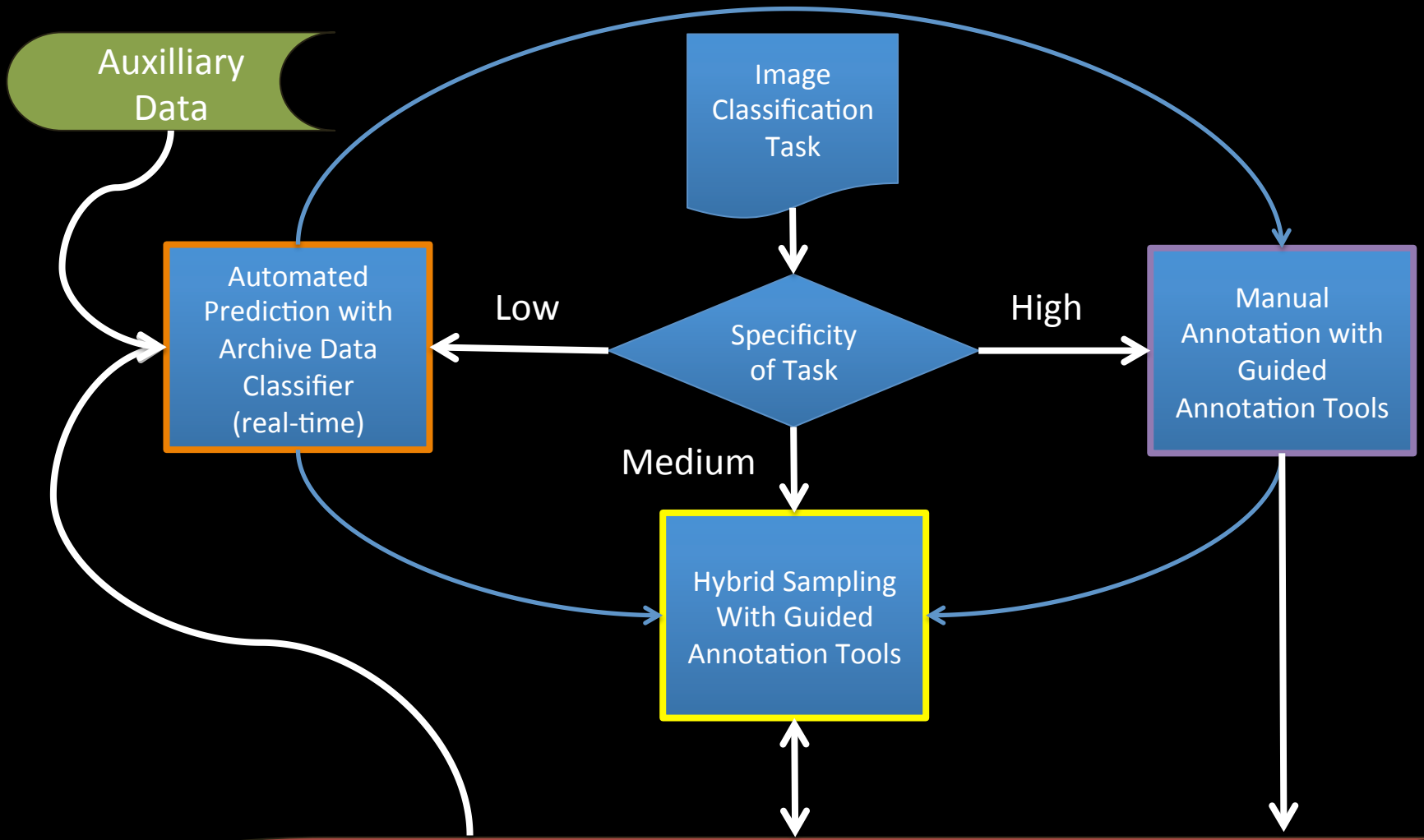


M. Hilbern and J. McGowan,
SCCOOS Harmful Algal Bloom Monitoring Program
http://www.sccoos.org/data/habs

Automated Image Annotation

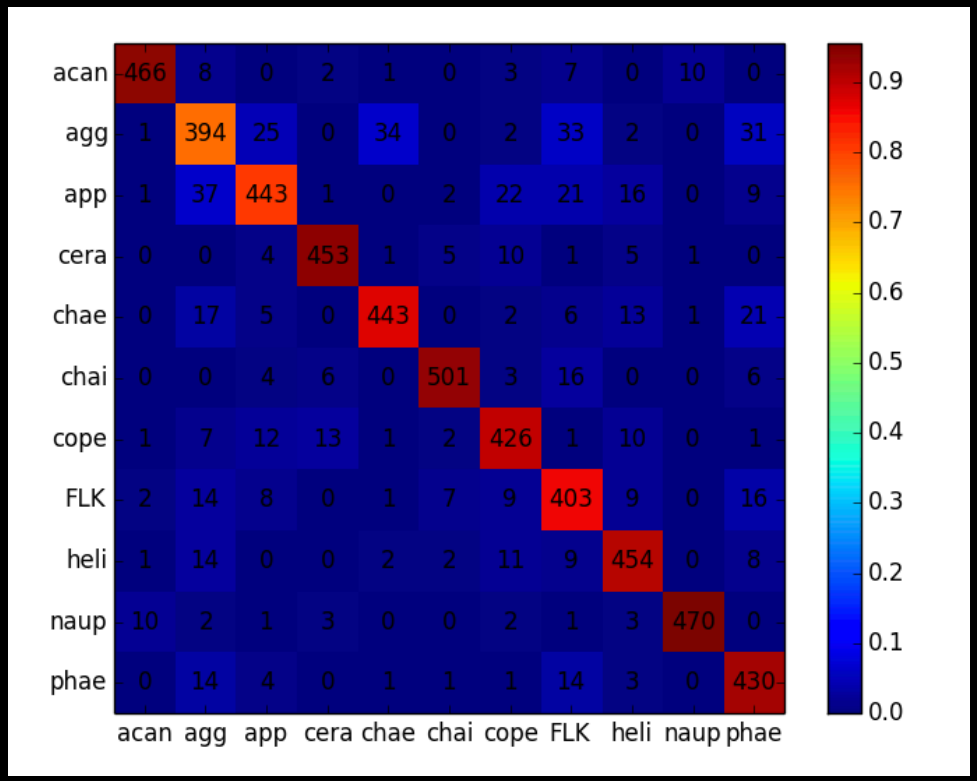
- The ease with which data is collected by the SPC systems gives rise to a significant annotation challenge.
- There has been excellent work in addressing the problem of annotating images from a test set using a large training set (Sosik & Olson L&O Methods, 2007, Gorsky et. al J. Plank. Res., 2010).
- Here, we consider the problem of applying these algorithms to data over a long time series where the relative abundance of classes can change dramatically from the training data.
- We aim to build a framework that combines humans, machines, and available auxiliary data to yield accurate predictions for relative abundance from the SPC system targeted at specific scientific questions.

Annotation Framework

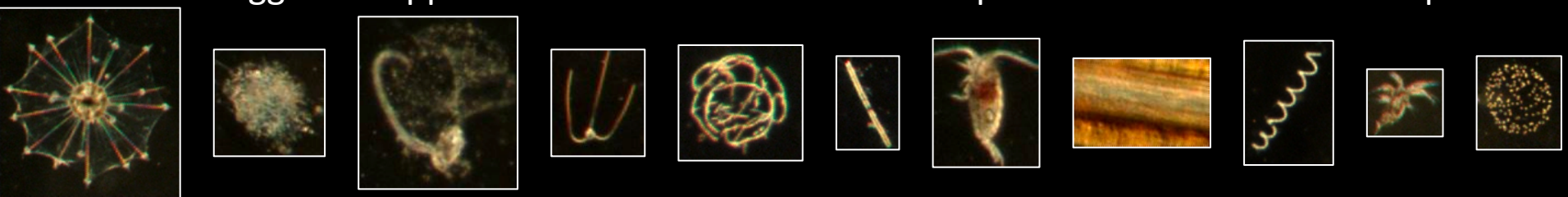


Archived Data Classifier

- Built from the archive data set labeled by users of the website.
- Linear, one-vs-one, SVM with probability outputs (sklearn implementation).
- 75 features are selected from morphological descriptors for the ROI and gray-level co-occurrence matrix.
- Average of 80 % to 90 % correct classifications on test set (500 images per class).
- A perfect algorithm yields a diagonal confusion matrix (right).



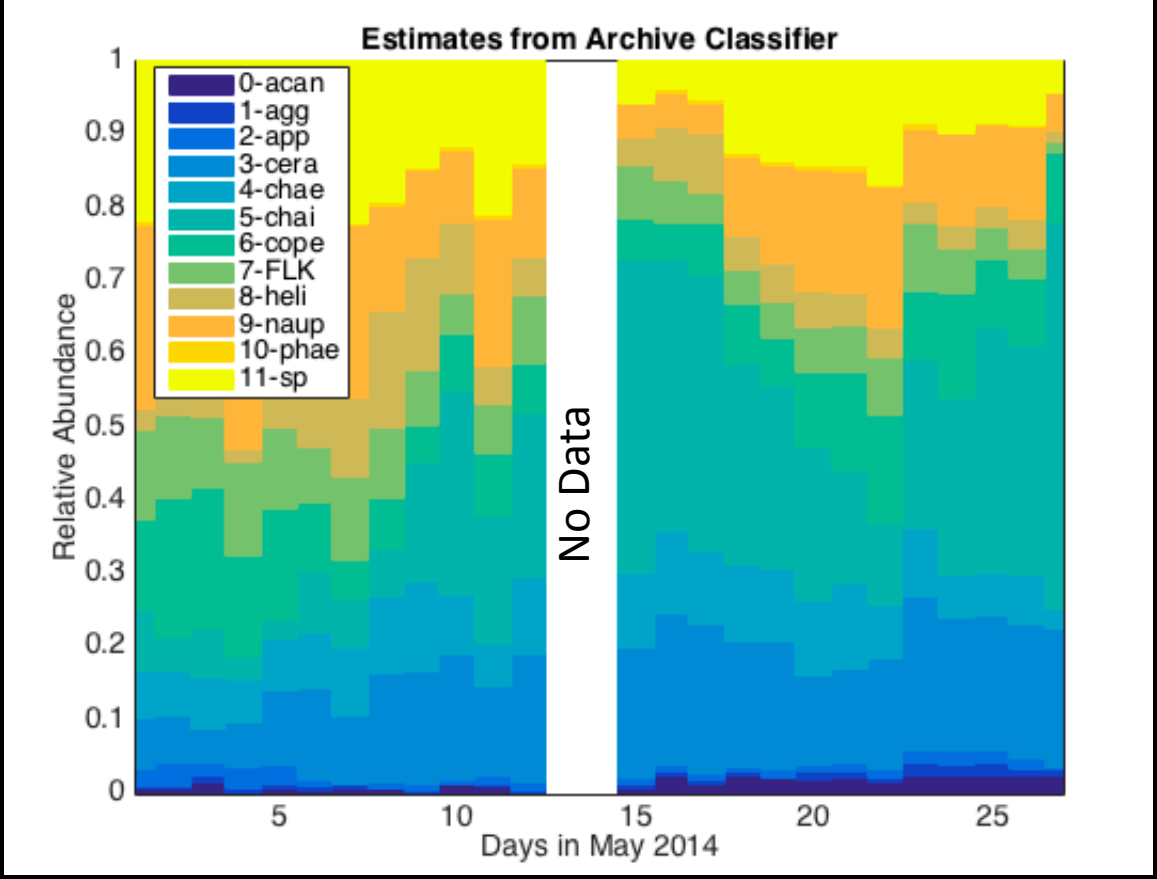
Acan. Agg. App. Cera. Chae. Chai. Cope. FLK. Heli. Naup. Phae.



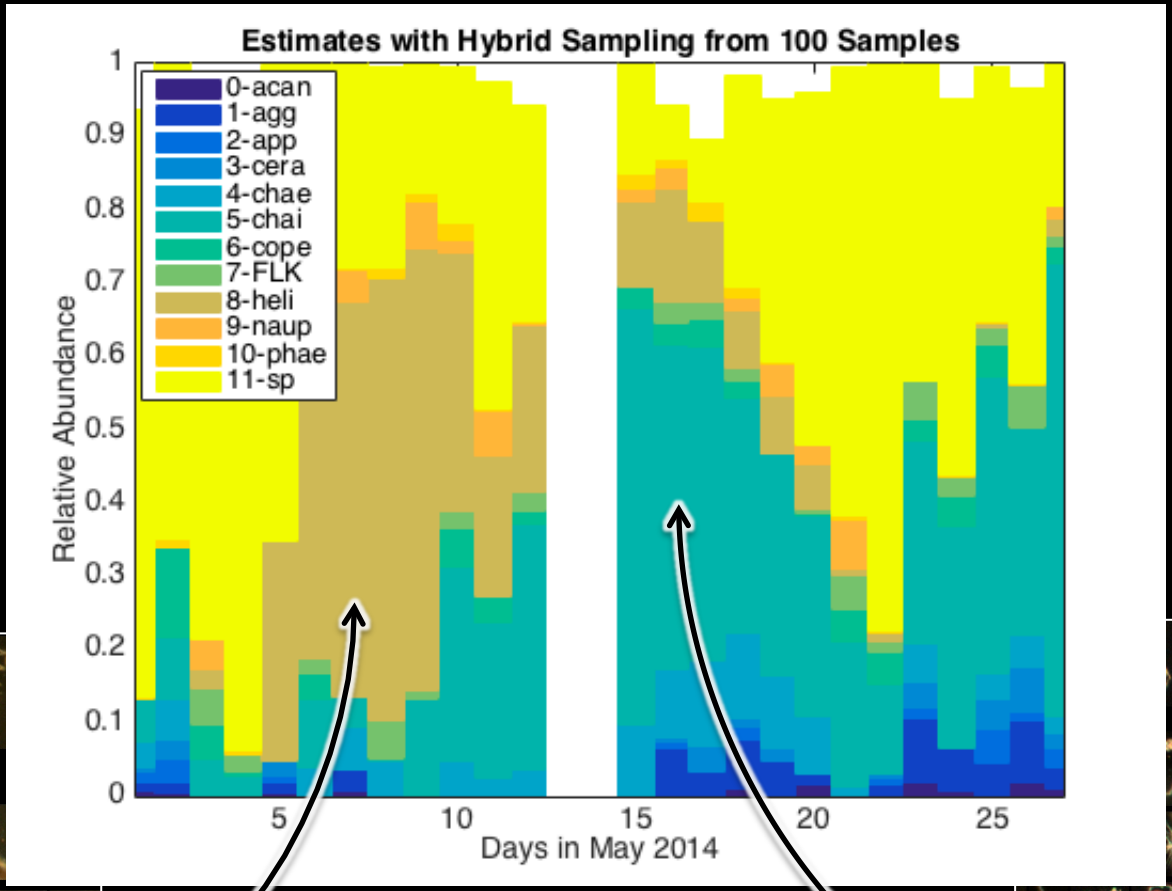
Hybrid Sampling

- Task: Highlight a bloom of Helical Diatom Chains in May 2014.
- Framework:
  - Archive Data Set of 11,000 annotated images drawn from March through October 2014.
  - Train the SVM on these 11,000 images and then apply the SVM to annotate 1000 images sampled randomly for each day in May 2014.
  - Of the 1000 images each day, 100 images are randomly selected and manually annotated.
  - The bias of the SVM output for a given day is estimated using the 100 manual annotations and applied to adjust the SVM output on the 1000 images.

Daily Relative Abundance with Archive Data Only



Daily Relative Abundance with Hybrid Sampling

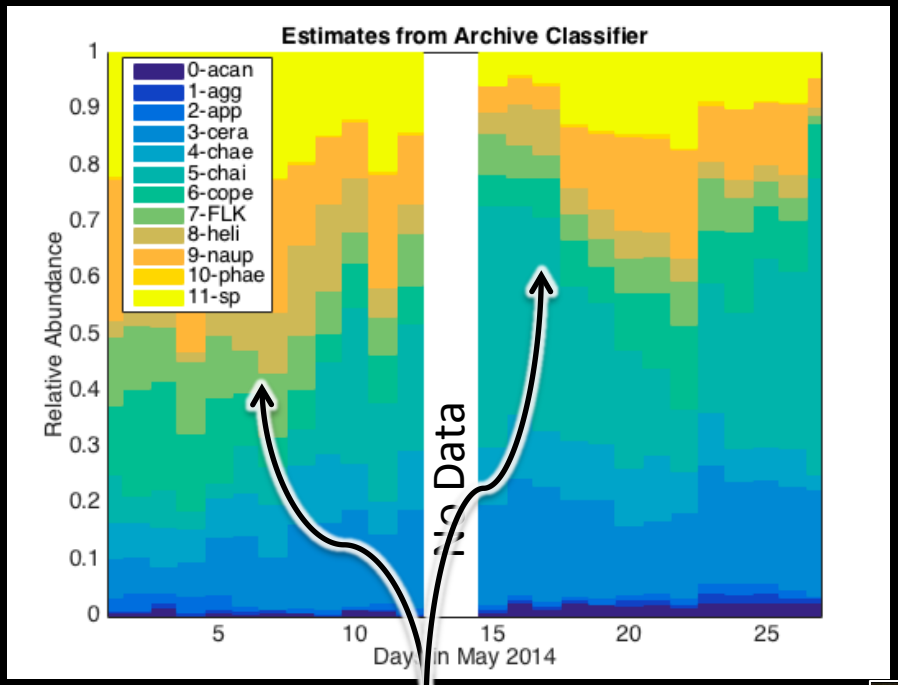


Helical Diatom Chain Bloom

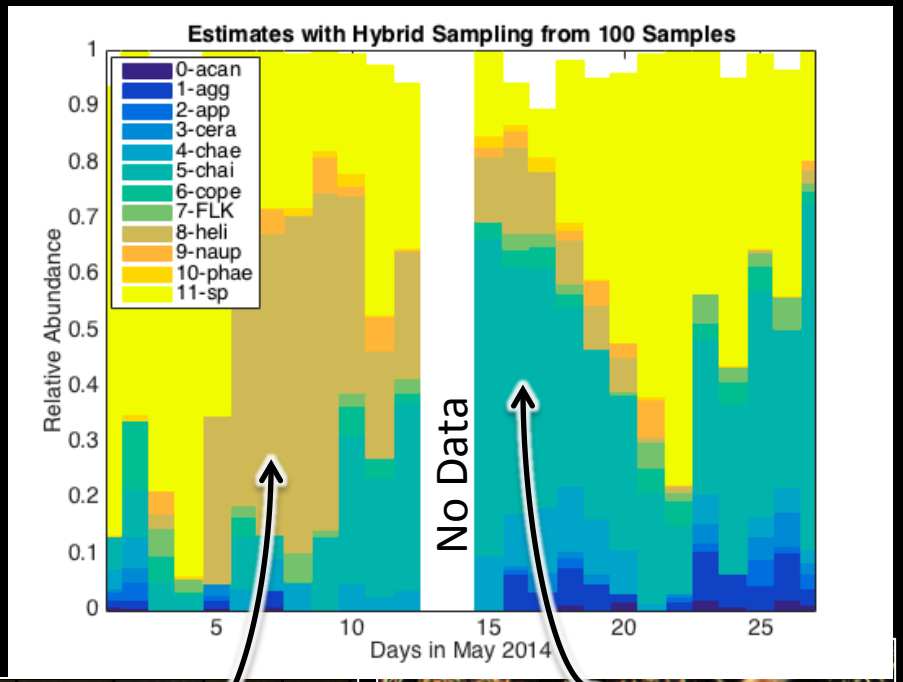
Diatom Chain Bloom

Comparison

Archive Data Only

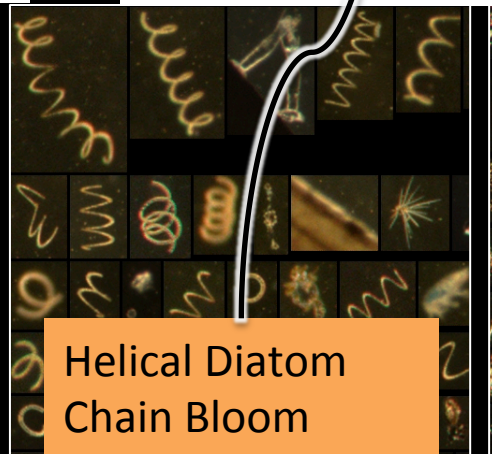


Hybrid Sampling



Blooms not clearly evident in SVM only estimates

Hybrid Sampling cost is only 10 % of manually annotating all 1000 images per day



Helical Diatom Chain Bloom

Diatom Chain Bloom

Summary and Future Work

- Two prototype imaging systems were designed and developed to study temporal changes in the plankton community off Scripps Pier.
  - In 2015 we will deploy permanent versions of these systems for one full year.
- The task of annotating image data from these systems was explored in a hybrid sampling framework.
  - Hybrid sampling can offer significant improvements in our ability to detect blooms in these data with only a small (~10 %) addition of manual annotations.
  - In 2015 we will further develop and test tools for rapid image annotation, prediction, and relative abundance estimation.

Acknowledgments

- Sponsors
  - W. M. Keck Foundation
  - Beyster Foundation
- Collaborators
  - Hybrid Sampling from Oscar Beijbom (UCSD, Computer Science). Please see: "*Cost-Effective Sampling for Pairs of Annotators*", Oscar Beijbom. [Preprint Availible on ArXiv from Oct 28 2014]
  - Southern California Coastal Ocean Observing System
  - Marine Science Development Center
  - Jaffe Lab Team
- Imaging FlowCytobot (Sosik & Olson)
  - Website: http://ifcb-data.whoi.edu/mvco
- SPC Websites (A work in progress; feedback is very welcome!)
  - SPC (Plankton): http://jaffeweb.ucsd.edu:5001
  - SPC-P (Microplankton): http://jaffeweb.ucsd.edu:5055