Clustering Applications in Biometrics

BiometricWorks

SMU | Bobby B. Lyle School of Engineering

John Howard

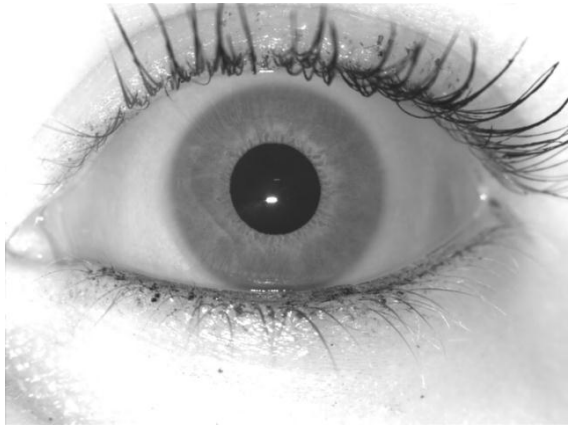
SMU Biometric Works Laboratory

Agenda

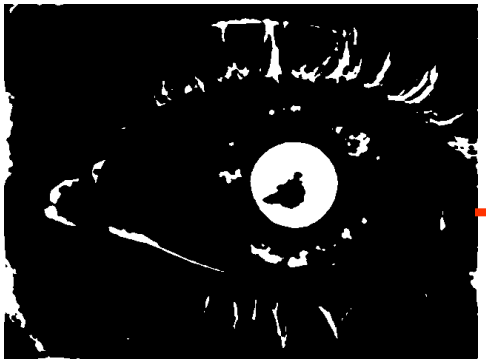
- Clustering in Iris Recognition
  - Binning
  - Hadoop and Biometric Databases
  - Pattern Recognition
- Clustering in Keystroke Dynamics
  - Three feature clustering

Iris Recognition

- Steps:
  - Locate Pupil
  - Locate Limbic Boundary
  - Generate Template
  - Compare Templates

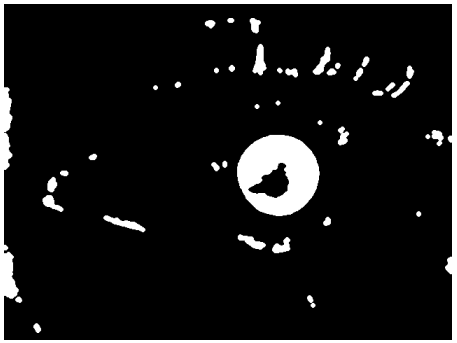


Pupil Location

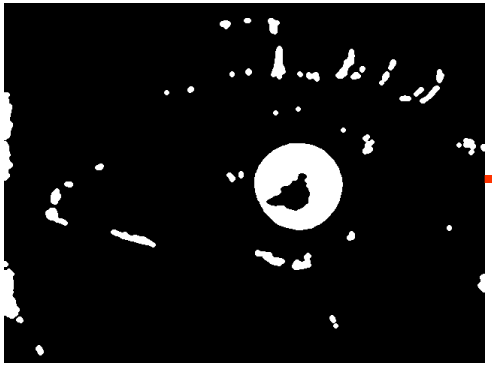


The original filtered image.

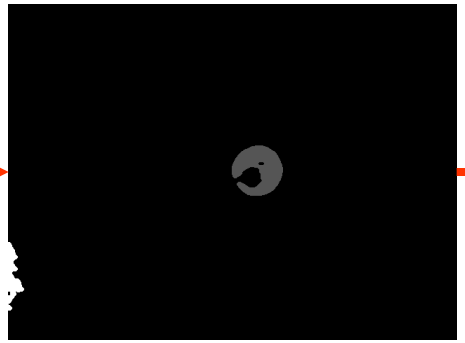
The eroded image.



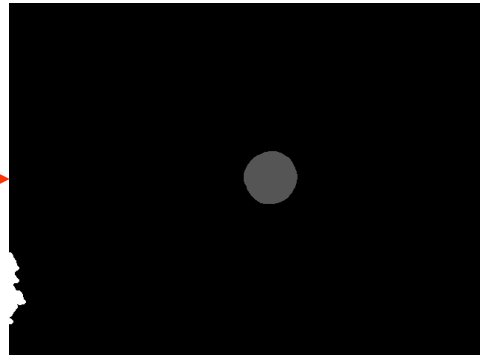
The image after erosion and dilation.



The image after erosion and dilation.



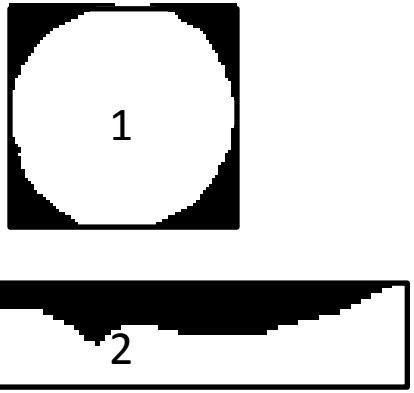
The image after labeling. Objects below a certain size are ignored.



The labeled image after being filled.

Pupil Location

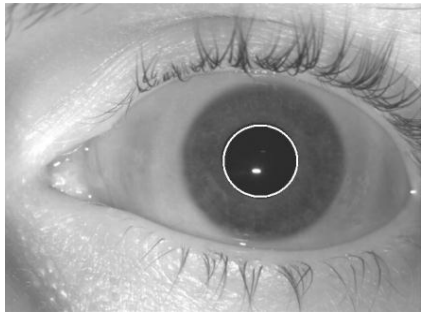
- Pupil Selection:



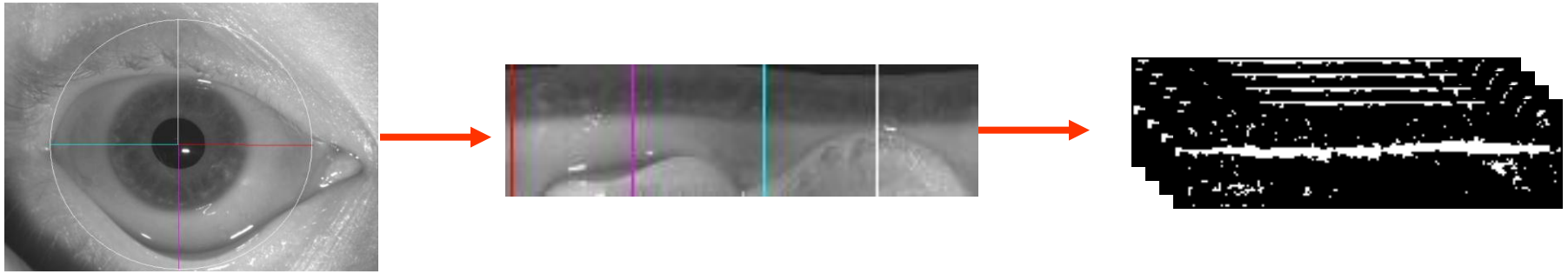
- Criteria:
  - Size
  - Eccentricity
  - Roundness

Limbic Location

- Given an image and a pupil location

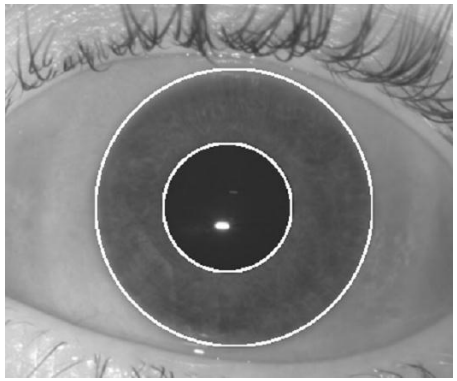


- “Unwrap”, filter, find straightest line

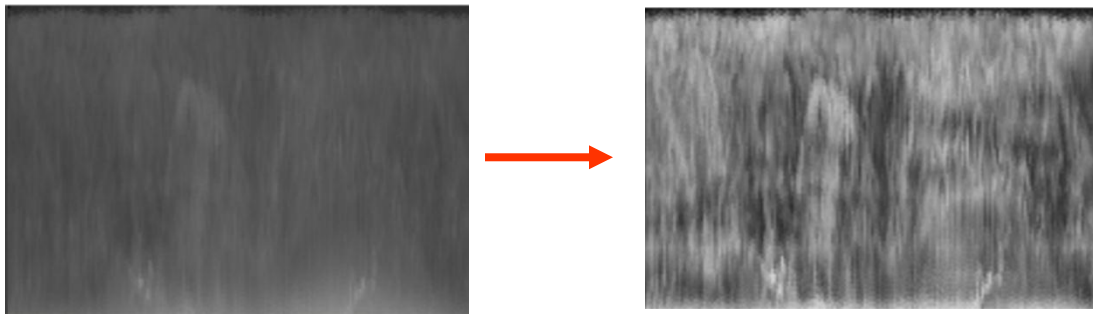


Template Generation

- Given an image, a pupil location, and a limbic location

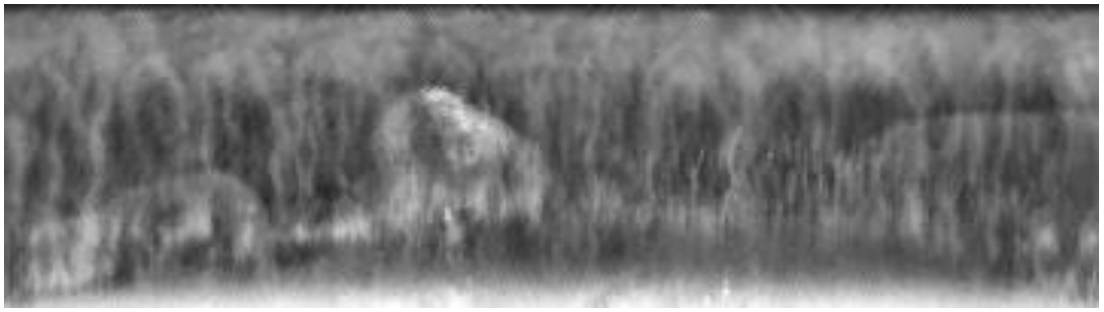


- Unwrap and contrast adjust



Template Generation

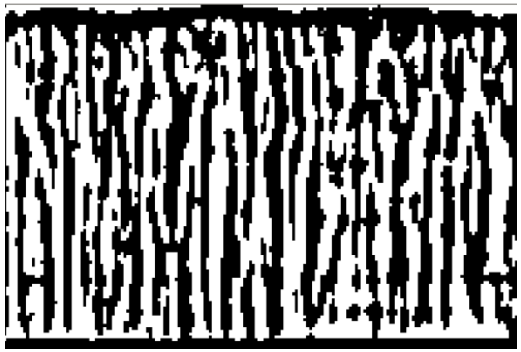
- Feature Extraction



- What features are algorithm specific (and highly proprietary)
- Extracted using feature kernels and convolution (wiki)

Template Generation

- Get a matrix of convolution responses
- Threshold and convert to binary



Template Comparison

- Every iris image becomes a template
  - Small
  - Easier to compare
  - Faster
- Comparing two templates = a distance metric
- Hamming Distance:

010111001
010101011

Template Comparison

- Imposter Comparisons:

"Bill"

"Joe"

- Should have a fractional Hamming distance of ~.5

Template Comparison

- Genuine Comparisons:

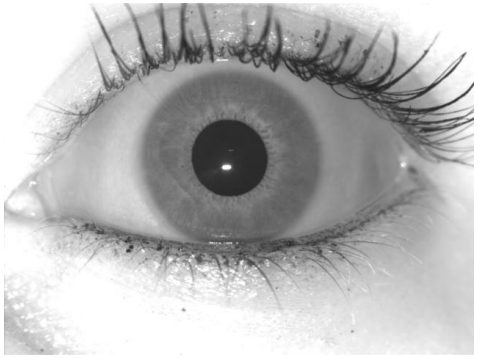
"Bill Yesterday"

"Bill Today"

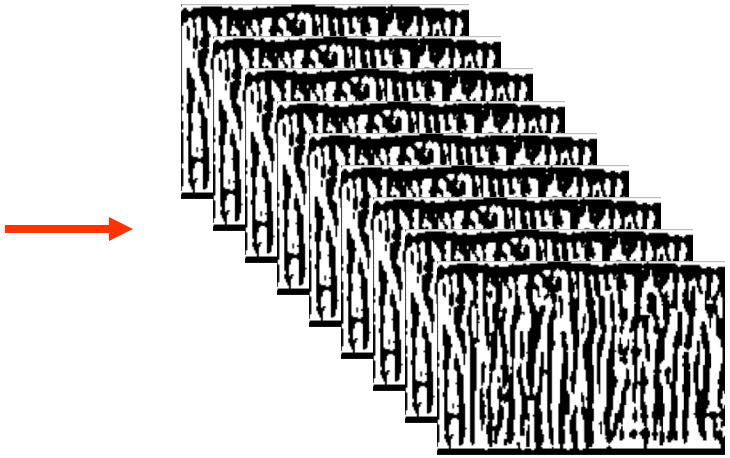
- Should have much lower fractional Hamming distance
- Can model both as a normal distribution

Iris Recognition

- Binning



Who is this?

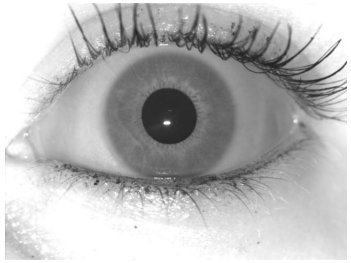


Search all these

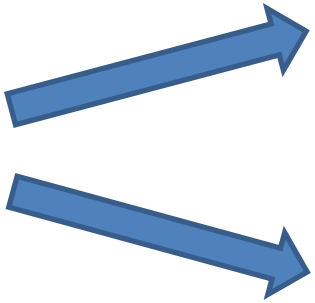
- Binning reduces the amount of templates you have to search (prune the search space)

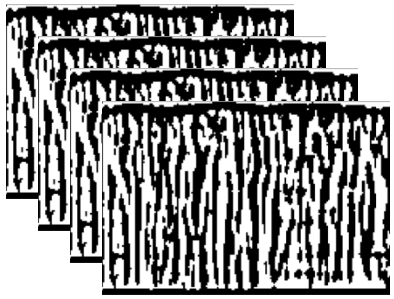
Iris Recognition

- Naive Binning



Left or Right Eye?





All left templates

All right templates

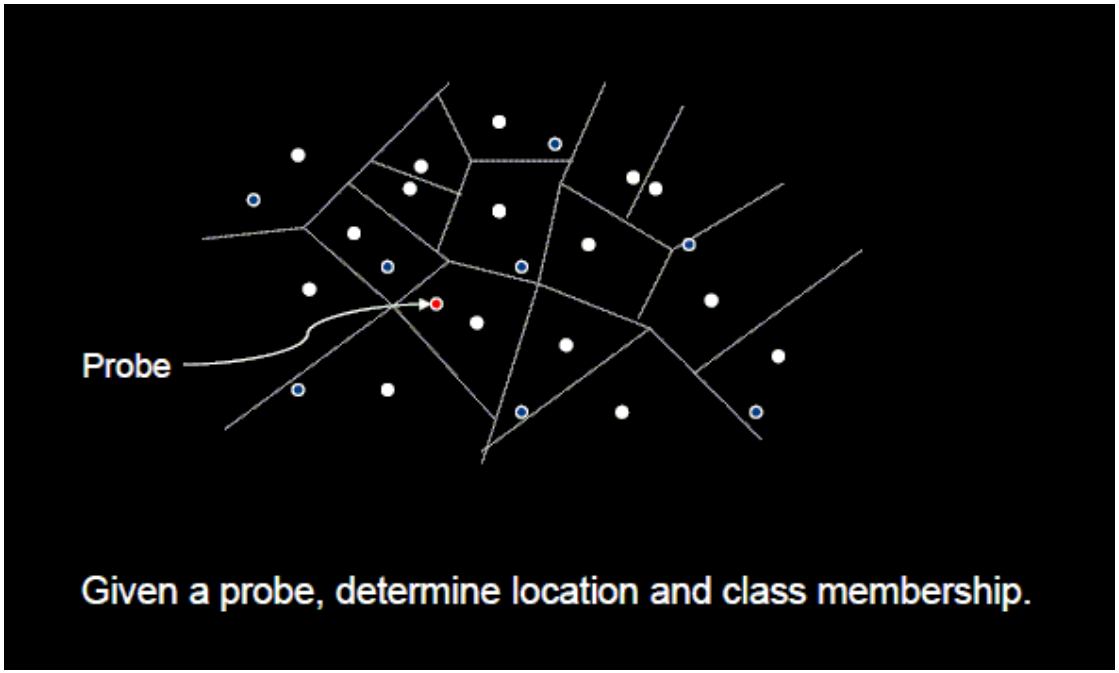
- No “natural” order
- False no-match if wrong bin is searched

Iris Recognition

- Feature Based Binning (Mehrotra, 2010)
  - 27 features of each template identified
    - Center of gravity
    - Area of black pixels
    - Etc.
  - Cluster an existing database based on these features
  - When a new probe image is introduced it gets “fuzzily” assigned to a cluster
  - The template comparison routine only searches hose clusters

Iris Recognition

- Similarity Based Binning (Becker, Potts, 2007)
  - Don’t bin on features
  - Bin on a similarity metric

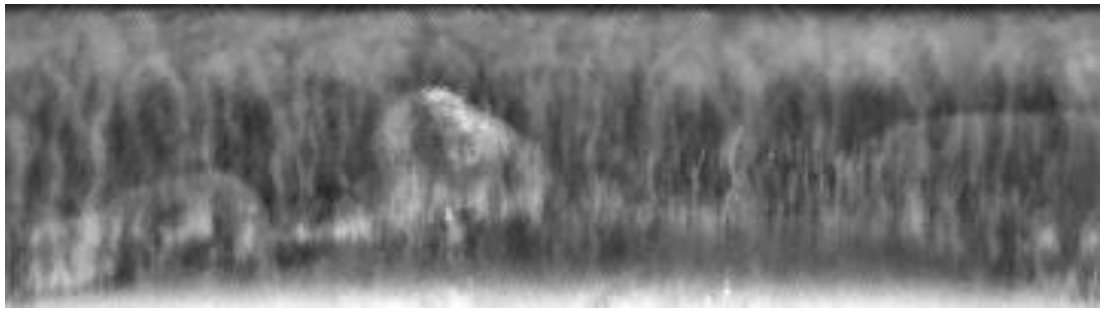


Hadoop and Iris Recognition

- Trost, 2010
- HDFS provides a framework for distribution, and redundancy of database
- Mahout/MapReduce to find clusters to reduce search space
- {Key, Value} stored in redundant SequenceFiles to allow for parallel search.

Iris Recognition

- Clustering to find patterns (pattern recognition)

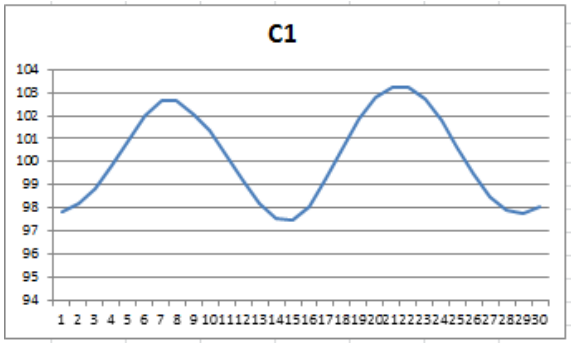


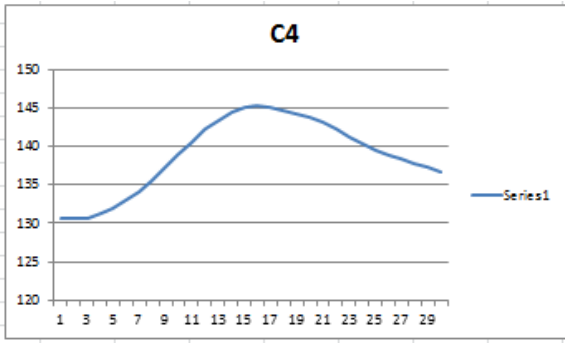
- Use sliding window of size k
- Generate 36,000 row vectors
- Cluster using k-means or k-medoids using a high number of clusters

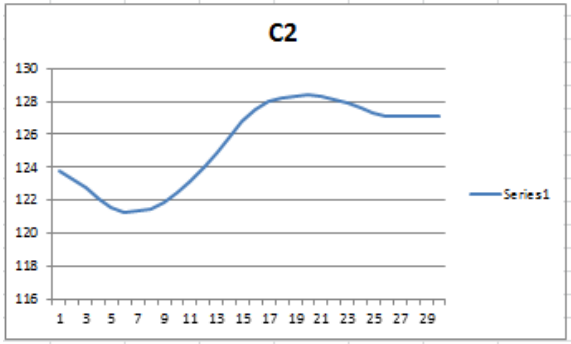
Iris Recognition

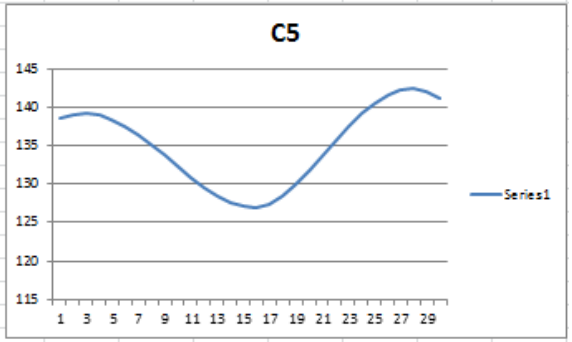
- Resulting Clusters should be most common patterns in the image
- Can then use these to generate kernels that get extremely high responses

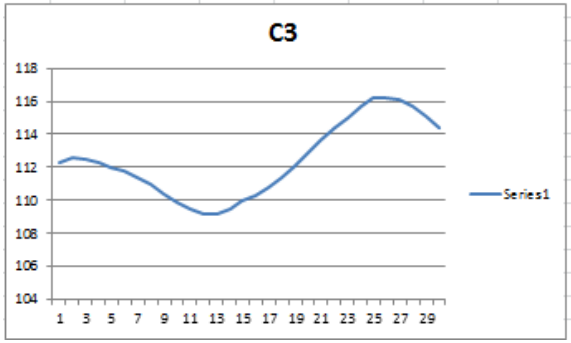
Iris Recognition

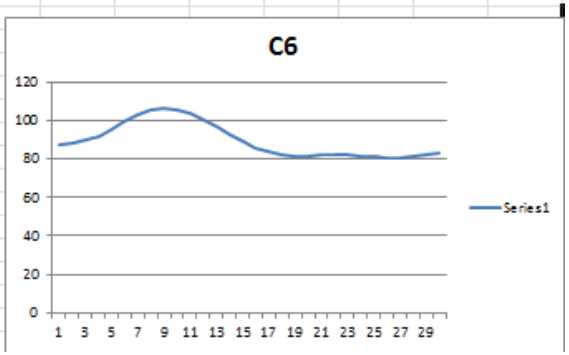






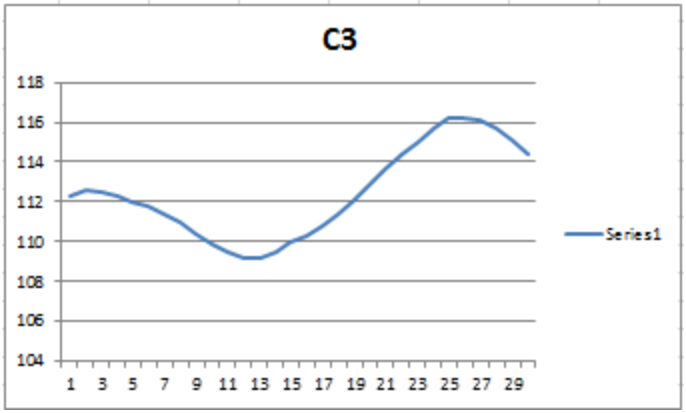




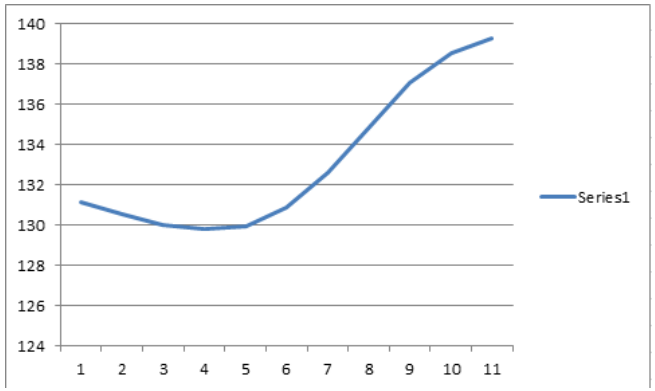


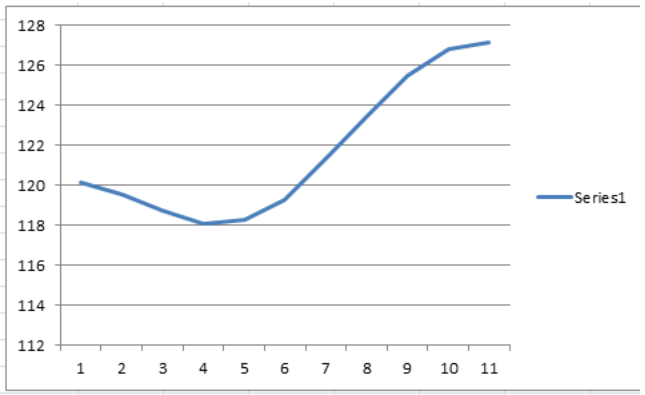
Iris Recognition

- Most common pattern not necessarily the most interesting
- Need to filter the vectors
  - Must have a certain % change



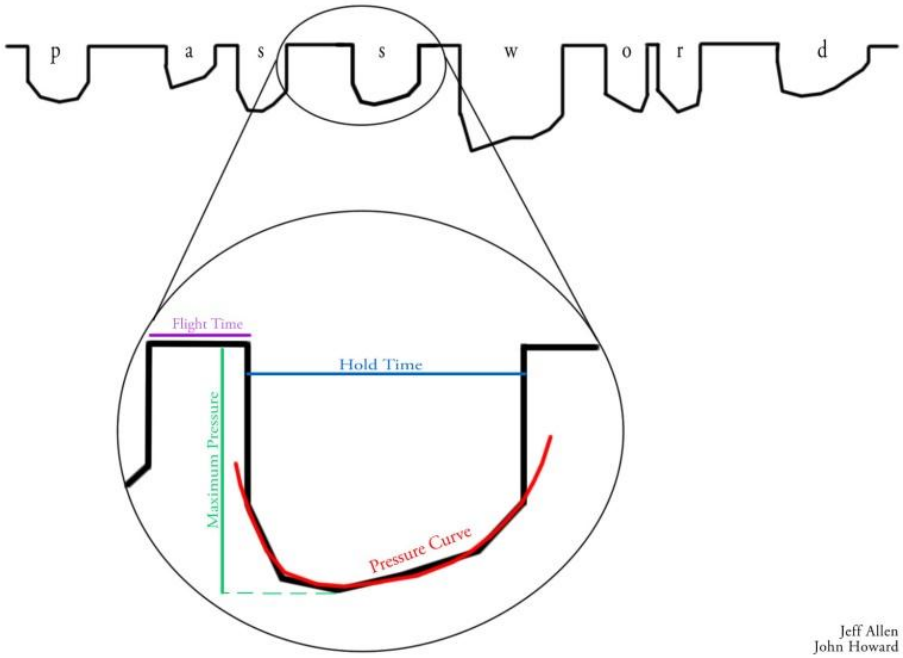
  - Must be normalized to the same starting value





Keystroke Dynamics

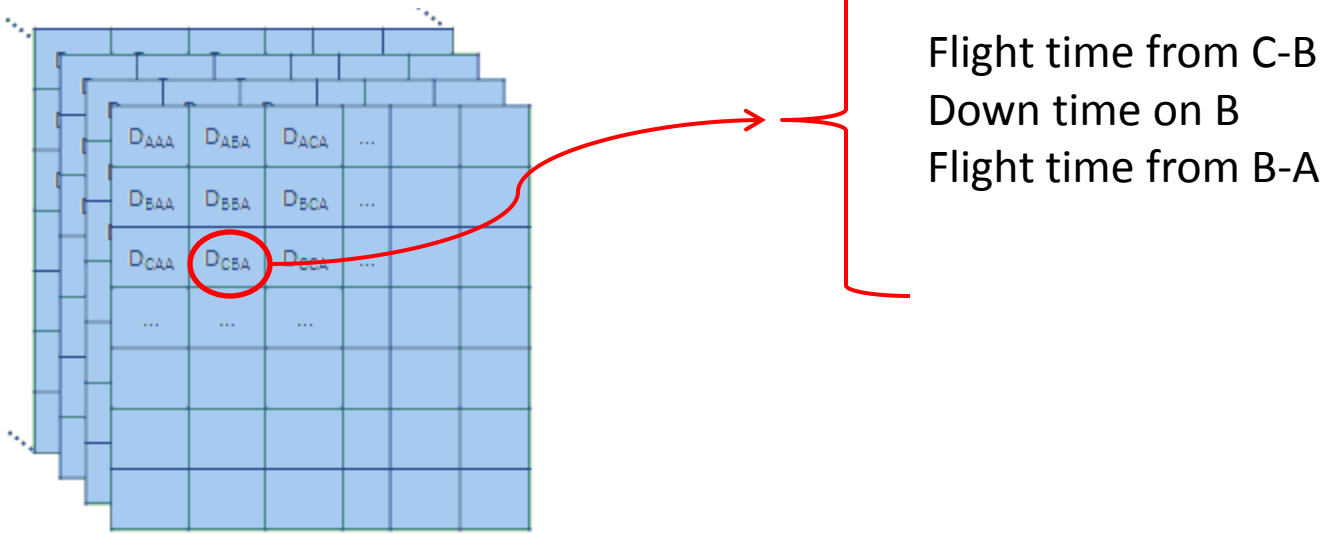
- Identify people based on their typing patterns



Jeff Allen
John Howard
Safelock - A Biometric Pressure-Sensitive Keyboard

Keystroke Dynamics

- Build up 3 dimensional matrices of timing information:



- Plot in 3D space
- Cluster (show example)

Clustering Applications in Biometrics

- Biometrics and Clustering are closely related
- Many applications across the field
- More uses in the future
- Question?

References

- Trost, J. Hadoop for Large Scale Biometric Databases. Hadoop Summit 2010.
- Becker, G. and Potts, M. Non-metric Biometric Clustering. Biometric Consortium Conference, 2007.
- Mehrotra, H., et. al. Feature Level Clustering of Large Biometric Databases.