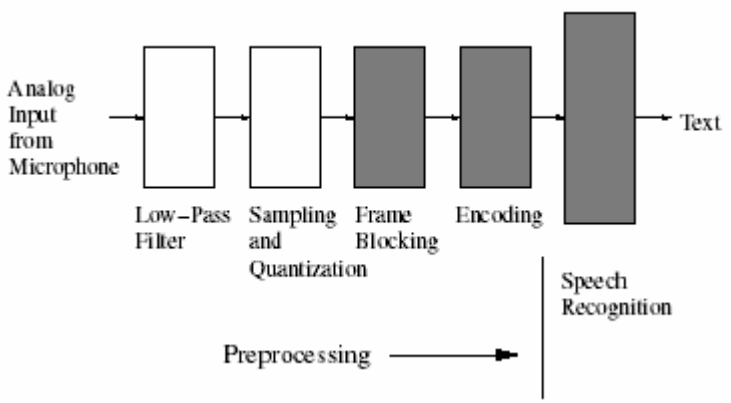
Pattern Recognition

Speech, Image, Handwriting, etc.

Wajahat Qadeer
Rebecca Schultz
Ernesto Staroswiecki

VOICE RECOGNITION

- Automatic conversion of speech into textual representation



- Preprocessing

  Partitioning and compression of speech into a stream of feature vectors

- Recognition

  Identification of words through an optimal path of a graph (most time consuming)

VOICE RECOGNITION

- Preprocessing
  - loop oriented with fixed bounds and no loop carried dependencies
  - High DLP with provision for TLP
  - Computationally intensive requiring floating point and integer operations
  - Small working set and memory foot print with regular data access patterns
  - High degree of spatial and temporal locality?

VOICE RECOGNITION

- Recognition
  - Large working set with highly irregular control and data access patterns
  - Big memory foot print during initialization requiring high bandwidth
  - Large caches and bigger block size reduce cache misses
  - Little ILP but TLP offers substantial gains
  - Algorithmic changes can exploit data locality

VOICE RECOGNITION

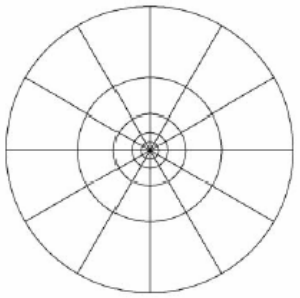
- Other Algorithms
  - Dynamic Time Warping, hidden Markov modeling, Neural Networks etc.
- Benchmarks
  - Common benchmarks are RASTA (pre-processing) and Sphinx (recognition)
- Scaling Trends
  - Complex search mechanisms requiring more computational resources
  - Large sets of databases requiring tremendous memory

IMAGE RECOGNITION

- Also a 3-step process:
    - **Edge detection:**
      Filtering
    - **Image processing / Characterization**
    - **Matching**

IMAGE RECOGNITION

- Processing / Characterization



  - We need to find image descriptors:
    Shape contexts, Fourier descriptors, etc.
  - Similar characteristics to voice recognition preprocessing except:
  - Not necessary to use floating point or excessive computation,
  - Yet more points to look at, which grow with the size of the image,
  - And although the memory access pattern is very regular, is important to remember that now we are looking at a 2D window.

IMAGE RECOGNITION

- Matching
  - Once again, similar to voice recognition, but problems really exacerbated!
  - Several algorithms: SVMs, Shortest Augmenting Path, etc
  - Remember that dictionary must be MUCH larger
  - Little ILP, some DLP, but mostly TLP
  - Topics to explore: CAMs, prefetching (but be careful!)

HANDWRITING RECOGNITION

- Special case of image recognition
- Similar algorithms for selecting descriptors and matching
  - Neural Nets, Hidden Markov Models, etc
- Matching library is small and fixed size
- Rarely done in hardware
  - Low data rate
- Scaling
  - Constant number of descriptor points irrespective of sample size
  - Limited opportunities for extensions