Recognition Technologies

These have a generic structure
      Seen most fully in speech reco but valid for others

Analog front end:
      Camera
      Pen
      Microphone
      Galvanic skin response, … biometrics

Features are extracted
      Example: speech every 10 ms extract 39 feature amplitudes

"Hypothesis search"
      informed by databases, which capture context and user training
            lexicon
            language model
            acoustic model
      earliest hypothesis search is simple pattern reco – "templates"
            Speech found this inadequate, but you have to start there
            Pen, image, and emotional reco still at this stage
                  Graffiti quite effective, even tho user must adapt
      Dynamics first appeared as "dynamic time warping"
            Now generalized as HMMs
                  Good algorithms exist for training

Padmanabhan and Picheny article gives a good overview of speech status – conservative, estimating still many years to go. For constrained tasks, speech reco seems good, yet still 10x more errors than humans make.

Reco in the user interface

Schneiderman article – can't speak and think at the same time (??)
      Mechanism is overcomittment of 'acoustic' shared memory
      Psychophysics evidence for this
      Error cascades when using an all-speech tool
      Add emotion to the inputs by recognizing user state

"Affective computing" community (esp. Roz Picard group at MIT)
      analog inputs are many:
            gsr, ekg, emg, bvp, …
      and they correlate strongly with task-induced stresses
      Transitions between emotional states (if you agree on what
these are and what they mean) can be modeled with HMM's
      Most credible results are in the areas of:
            User frustration
            Driver stress level

A side effect – programs which attempt to respond to frustration are becoming more common, although not particularly effective yet.