

Projects in Signals and Information Processing Systems (2008)

Offered by

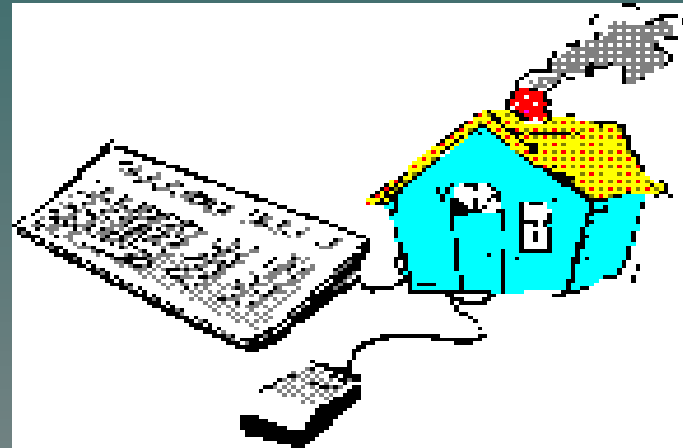
Dr. Roberto Togneri

Room 4.10, roberto@ee.uwa.edu.au

SIP FYP Projects

- IDEAL House (3A, 3H)
- Speech Processing (3B)
- Speaker Recognition (3C, 3D)
- Biomedical Engineering (3E)
- Advanced Signal Modelling (3F,3G)
- Music, Patterns and Images (3H)

3A. IDEAL House

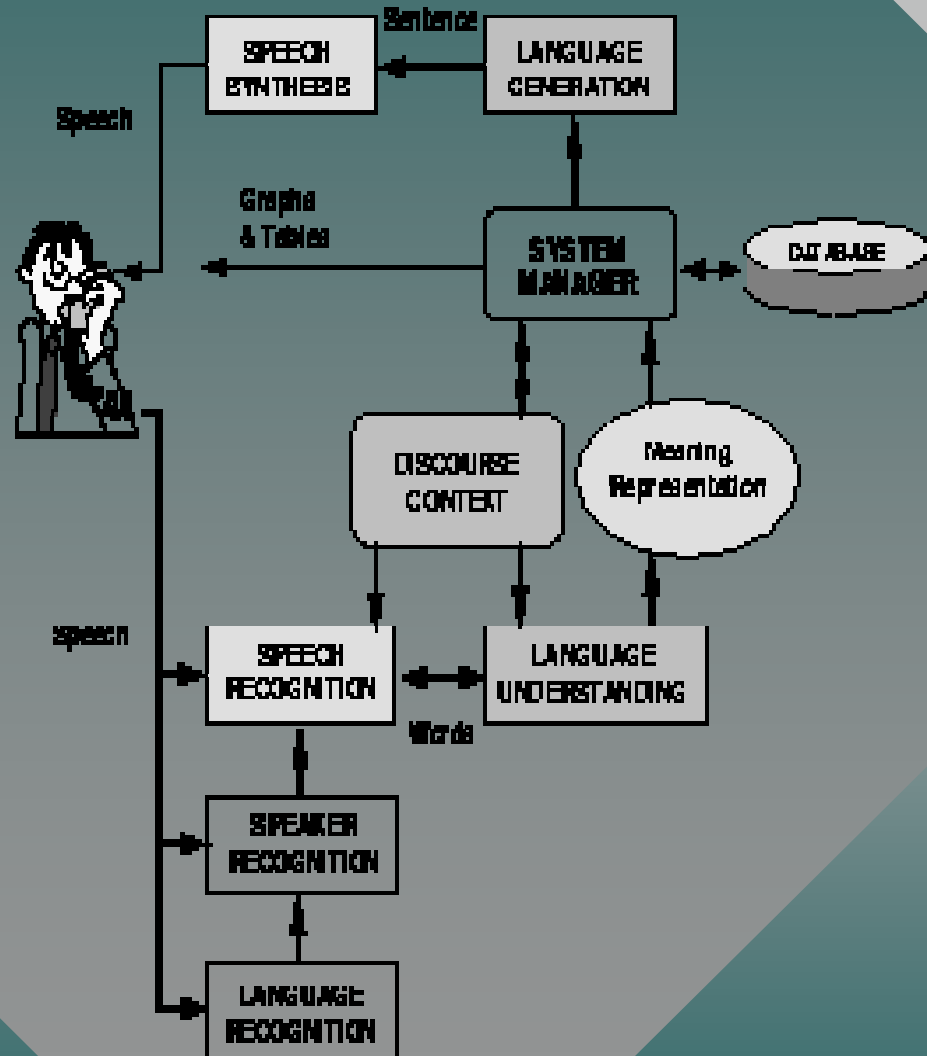


- Hands-Free and Assisted Control
 - Use the power of your voice
 - No need to find the light switch
 - No need to locate and use the remote control
 - Just say it and its done!
 - Use the power of your hands
 - Just use your arm and hands
 - Point with your arms and its selected
 - Use your hand gestures
 - Combine the above!
- Occupant and Intruder Detection and Identification
 - Use voice and sounds to identify owners and intruders
 - Use images and video to identify intruders and recognise owners
 - Combine the two!

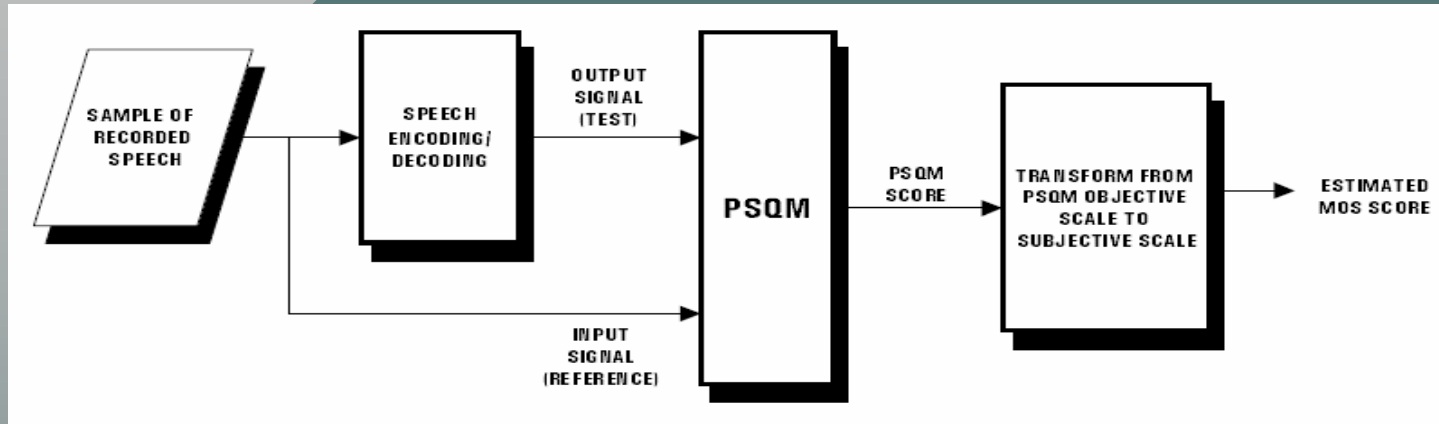
3A. ... and Beyond

Topics in Spoken Language Systems

- voice activity detection
- keyword spotting
- continuous speech recognition
- language understanding
- Dragon Naturally Speaking: how good is it?
- ***More than one student (on different sub-projects)***

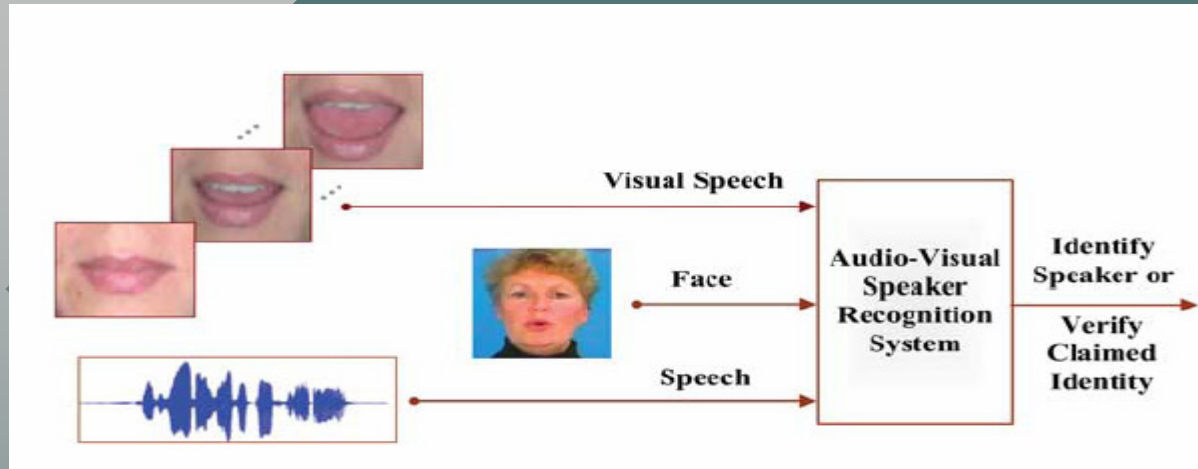


3B. Quality Measures for ASR



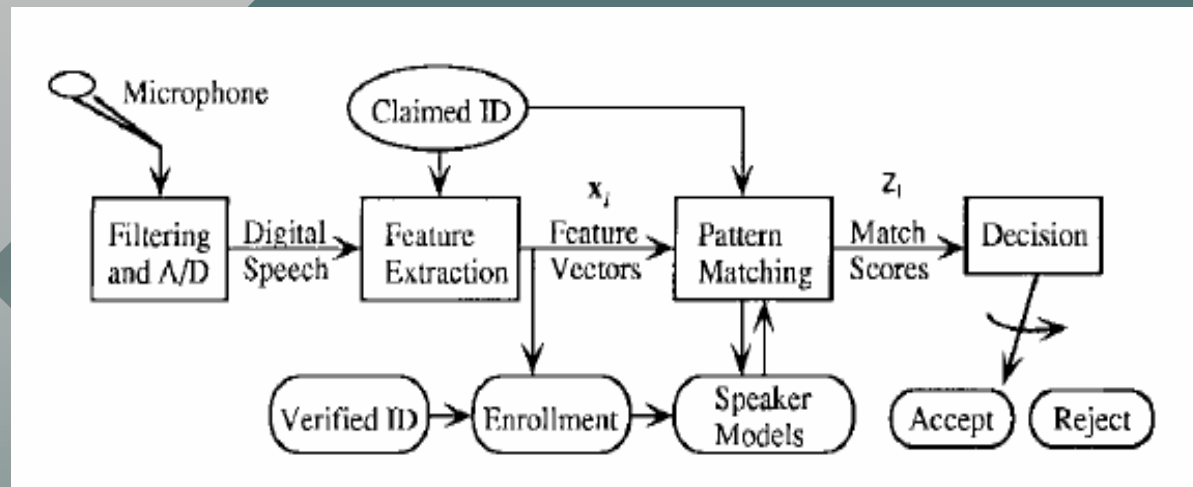
- Speech Enhancement for Robust Speech Recognition
 - Many speech enhancement / source signal separation algorithms exist
 - Optimised for telephony and speech intelligibility NOT speech recognition
 - Difficult to optimise speech enhancement for speech recognition directly
 - No “cost function” to be minimised
 - Speech intelligibility is subjective, speech recognition performance is objective
- Questions you can address
 - Is there a correlation between speech quality measures and speech recognition performance? Which speech quality measure is also a good indicator of speech recognition performance?
 - A new measure of speech quality assessed based on the speech recognition performance (objective) rather than intelligibility (subjective)

3C. Audio-Visual Speaker Recognition



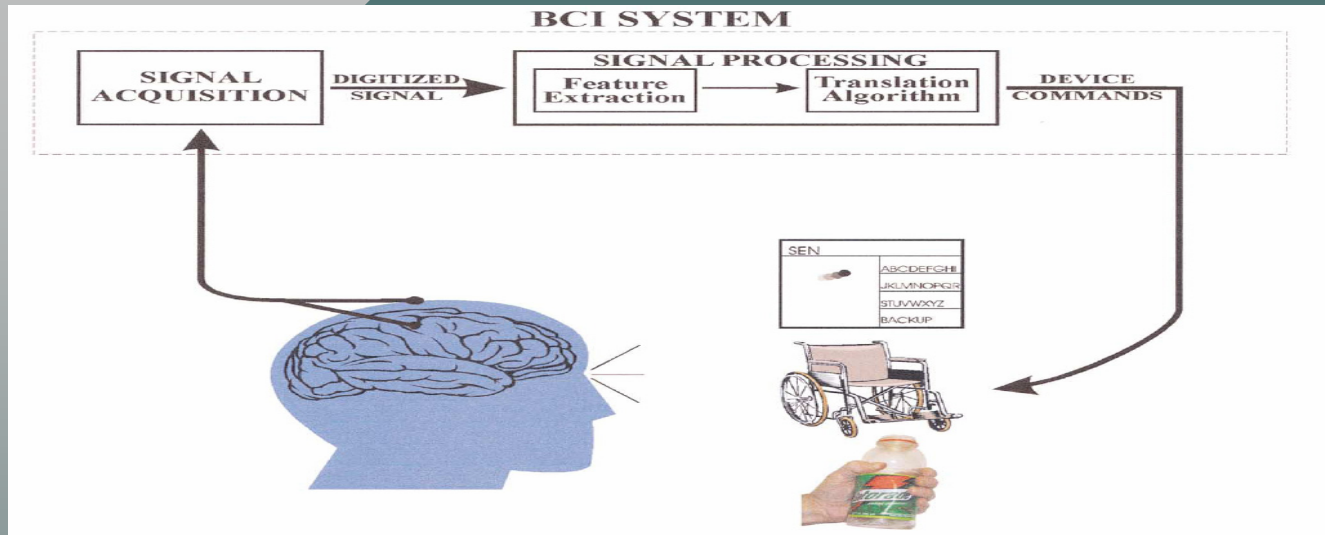
- Identifying persons and suspects (security, forensics, etc.)
 - One can forge a PIN or password but one can't easily forge one's looks or voice
- Possible tasks
 - Person identification by voice
 - Person identification by face
 - Combine the two!
- ***More than one student (EECE/CSSE)***
- ***Can be associated with the IDEAL House***

3D. Classification in Speaker Recognition



- Classifiers and then there are classifiers!
 - GMMs are the standard pattern classifier (if you don't believe in neural networks ☺)
 - SVMs are the new "kid on the block" and dominate speaker recognition classification
- Your objectives for this project
 - Compare and contrast GMM and SVM classifiers for speaker recognition
 - Investigate different training algorithm approaches
 - Explore novel combinations of training algorithms and architectures for both speaker recognition and beyond (faces? images? music? etc.)

3E. Using EEGs for BCI



- Control it just by thinking about it!
 - Investigate the latest in Brain-Computer Interface (BCI) research and findings
 - EEG data available from CCRN
 - Develop and evaluate algorithms to detect and identify the required responses from the EEGs
 - Project will be in conjunction with CCRN

3F/3G. Advanced Signal Processing and Modelling Processing

- Next generation speech recognition engine
 - Modelling of underlying physiological dynamics
 - A really hard problem, so in two parts
- 3F: Mapping of VTRs to Features
 - VTR tracking: $f(12d \text{ features}) = 4 \text{ VTRs}$
 - Generation model: $f(4 \text{ VTRs}) = 12d \text{ features}$
 - We have the VTR data (just released) and the feature data, now we need the mapping!
 - Possible investigations: linear, analytic (all-pole), neural network, something else?
- 3G: Estimation of non-linear state-space equations
 - Need to estimate both the parameters and state variables
 - Very hard if equations are non-linear and noise is non-Gaussian
 - Possible investigations: PF, UKF, QKF
 - Possible advances: UKS, QKS (smoothers are better than filters!)

3H. Make Your Own Project

- Do you have an interesting idea in:
 - **Audio or Video Processing**
 - A practical software implementation or hardware interface to the IDEAL house
 - Multi-microphone / Stereo camera processing for person localisation
 - **Signal Processing**
 - Modelling, Enhancement, Classification, Synthesis of speech, music, biomedical, etc.
 - **Image Processing**
 - Enhancement, Detection, Classification of images
 - **Pattern Recognition**
 - Recognition and classification of Audio, Music, Video, or Text?
 - **Music and Musical Instruments**
 - Music note recognition?
 - Music synthesis by physical modelling?
 - Music instrument identification?
 - Multimedia content summarisation?
- Then see me to discuss what you have in mind and lets see if we can turn it into a project!

!!! STOP PRESS !!!

- **IDEAL House Demo (TBC for 1pm, Oct 12)**
 - Demo of computer and vision system
 - Possible demo of wireless communications, voice activated system, etc.
 - You can see what the IDEAL House actually is!
- **New Project: 3I. Pause/Silence Detection**
 - Prof. Kim Kirsner for UWA School of Psychology is investigating importance of pauses in speech
 - Project involves use of pattern recognition (GMM, HMM, etc.) to classify or segment speech into non-speech (silence, pauses) and speech regions
 - Work could be extended to a front-end VAD for triggering a voice-activated control system (“kill two birds with the one stone?”)

Want to know more?

- SIP FYP 2008 Projects Page
 - <http://www.ee.uwa.edu.au/~roberto/research/projects2008.html>
- Contact me: roberto@ee.uwa.edu.au
- Interested students can get more information on each project including:
 - Reading list of the key articles and textbooks
 - Selected WWW pages and resources
 - Links to software and manuals
 - Contact emails of collaborators who are involved with the project