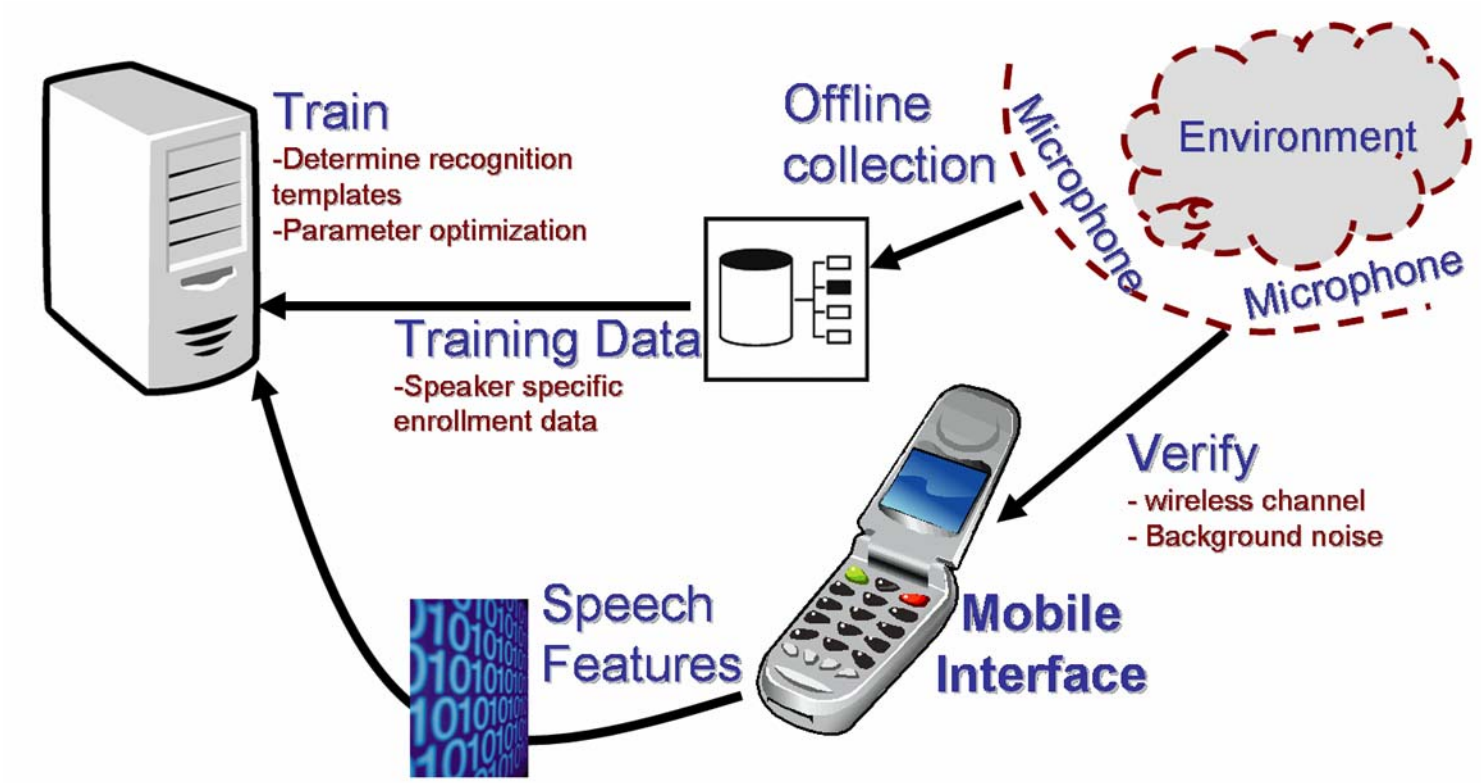
Key References

S. Chakrabartty*, Y. Deng and G. Cauwenberghs, Robust Speech Feature Extraction by Growth Transformation in Reproducing Kernel Hilbert Space, IEEE Transactions on Speech, Language and Acoustics, pp. 1842-1849, Vol. 15 Issue: 6, Aug. 2007

A. Fazel, S. Chakrabartty, Non-Linear Filtering in Reproducing Kernel Hilbert Spaces for Noise-Robust Speaker Verification, to appear in International Symposium on Circuits and Systems (ISCAS), Taipei, Taiwan, 2009

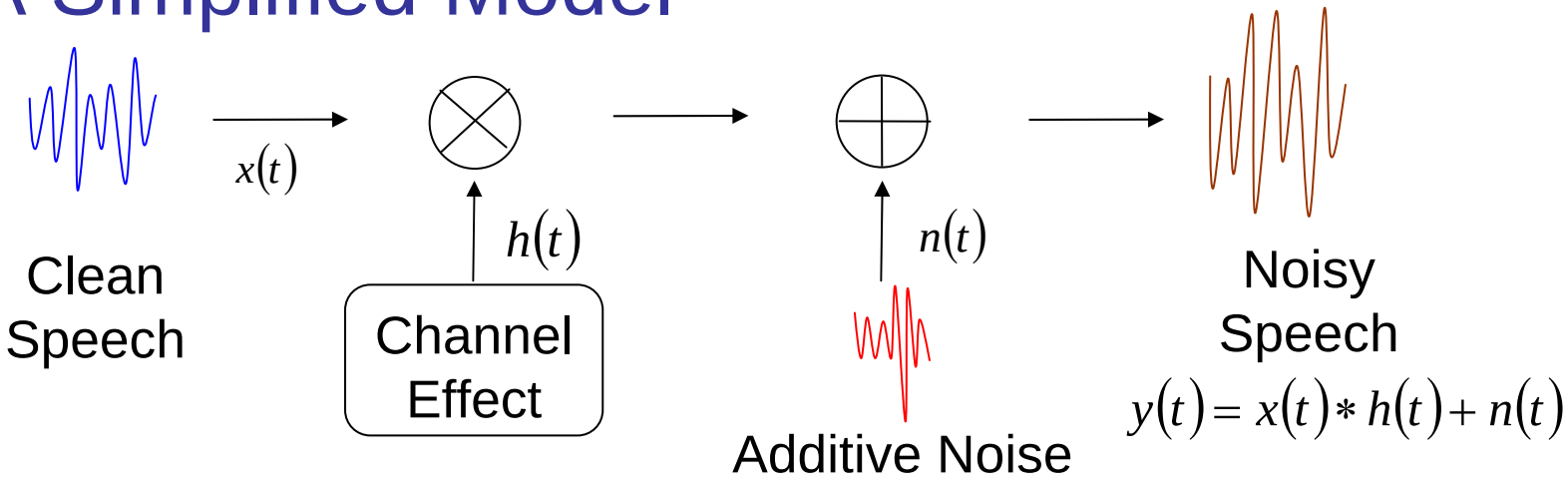
Robust Speech/Speaker Recognition

- Performance of existing speech/speaker recognition systems degrade significantly with noise.
- Training recognition systems for all noise conditions is impractical



Noisy speech Model

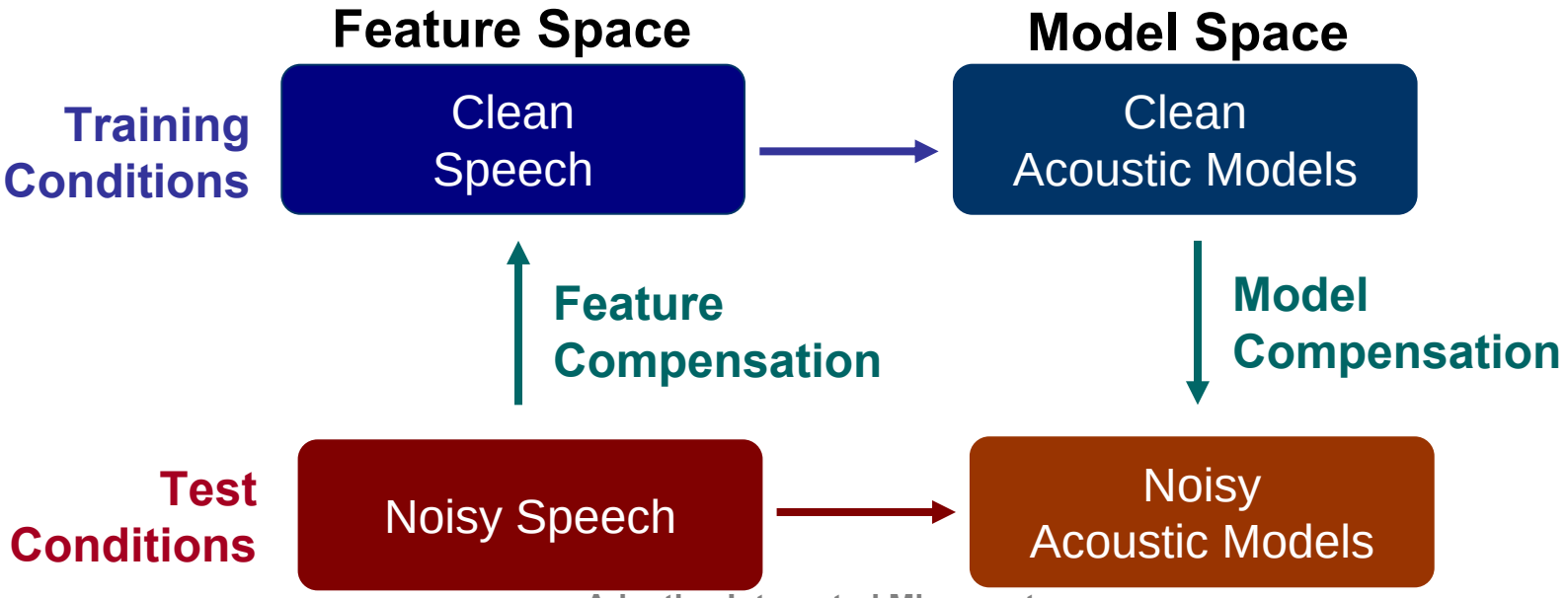
- A Simplified Model



- **Additive noise**
  - background noise such as air-conditioning
- **Channel noise**
  - difference between microphones in training and testing conditions

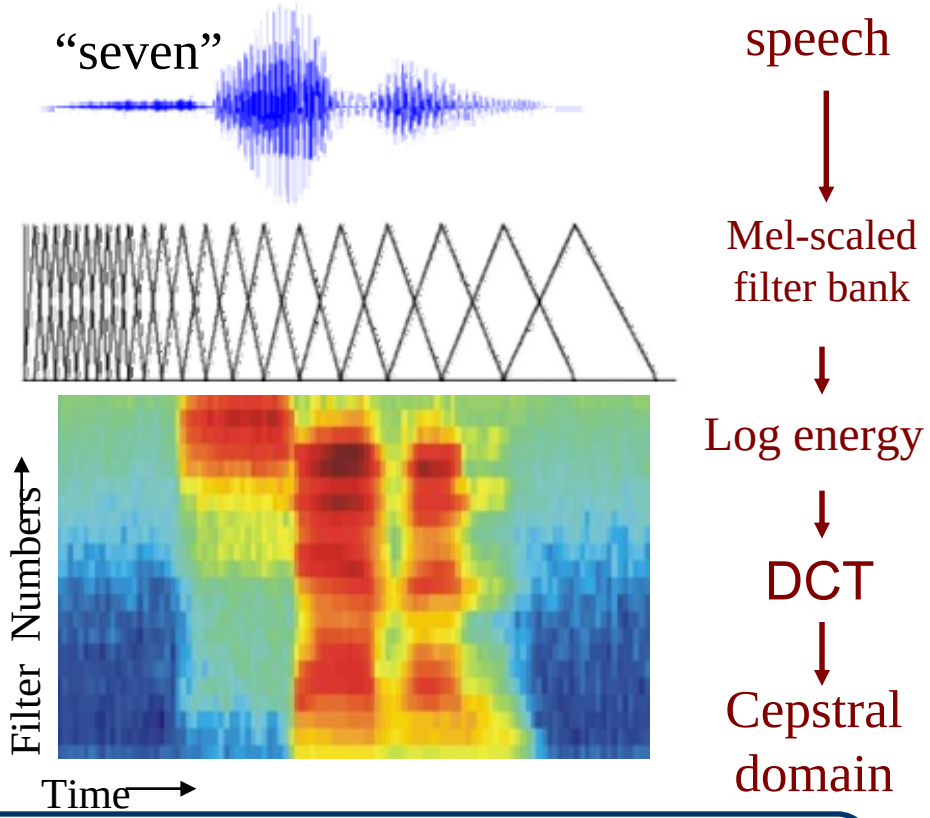
Robustness in Speech

- Two main approaches to improving noise robustness
  - Model Compensation
    - Adapt acoustic models to match corrupted speech features
    - Representative approaches: MLLR, MAPLR, etc.
  - Feature Compensation
    - Restore corrupted speech features to corresponding clean ones
    - Representative approaches: SS, CMN, CMVN, etc.



Speech Features (cont.)

- MFCC is the most widely-used speech features
- Robust MFCC
  - adding 1st and 2nd time derivation of feature to get the dynamic information
  - normalization/transformation
    - CMN, CMVN

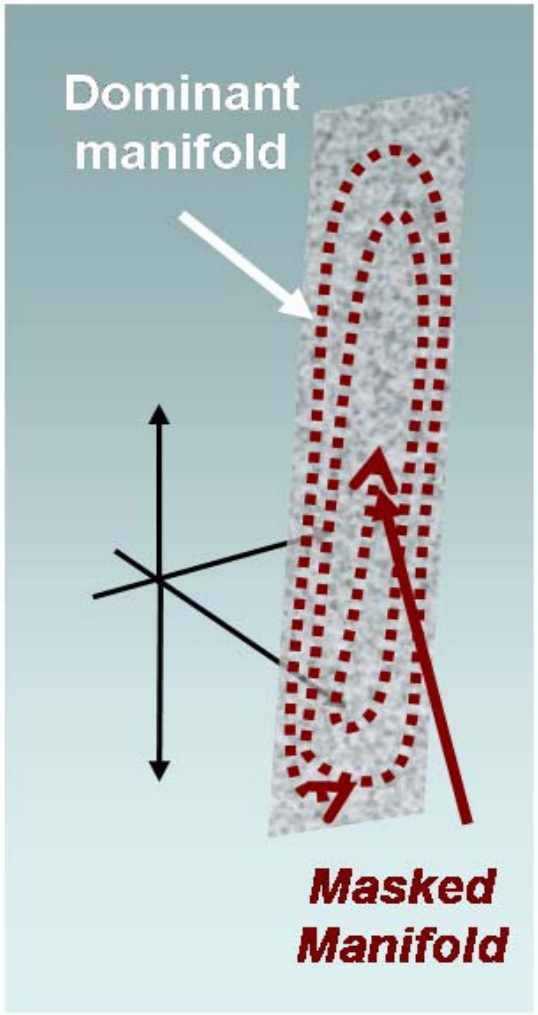


- Most state-of-the-art speech features
  - linear information
  - **Do not exploit** higher-order statistical characteristics of the signals

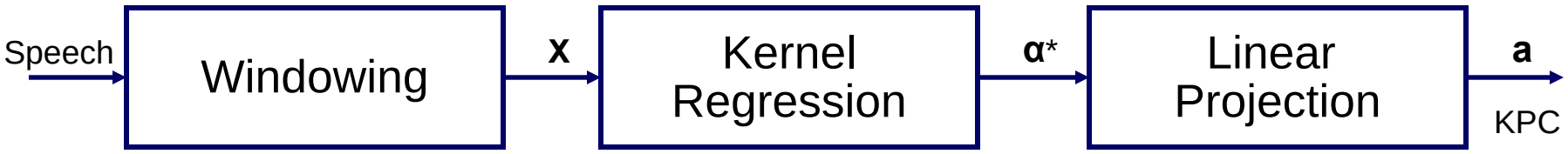
Kernel Filtering

Filter out noise in the higher-dimensional signal space so that robust non-linear features can be unmasked.

- Kernel filtering can be used to discover high-dimensional manifolds in empirical data.
- High-dimensional time-series features are difficult to learn because they are typically masked by dominant linear features.



KPC Feature Extractor



- doesn't make any prior assumption on noise statistics
- uses kernel methods
  - nonlinear
  - robust to corruption by noise
- regularization principles to ensure robust parameter estimation