16.0 Distributed Speech Recognition and Wireless Environment

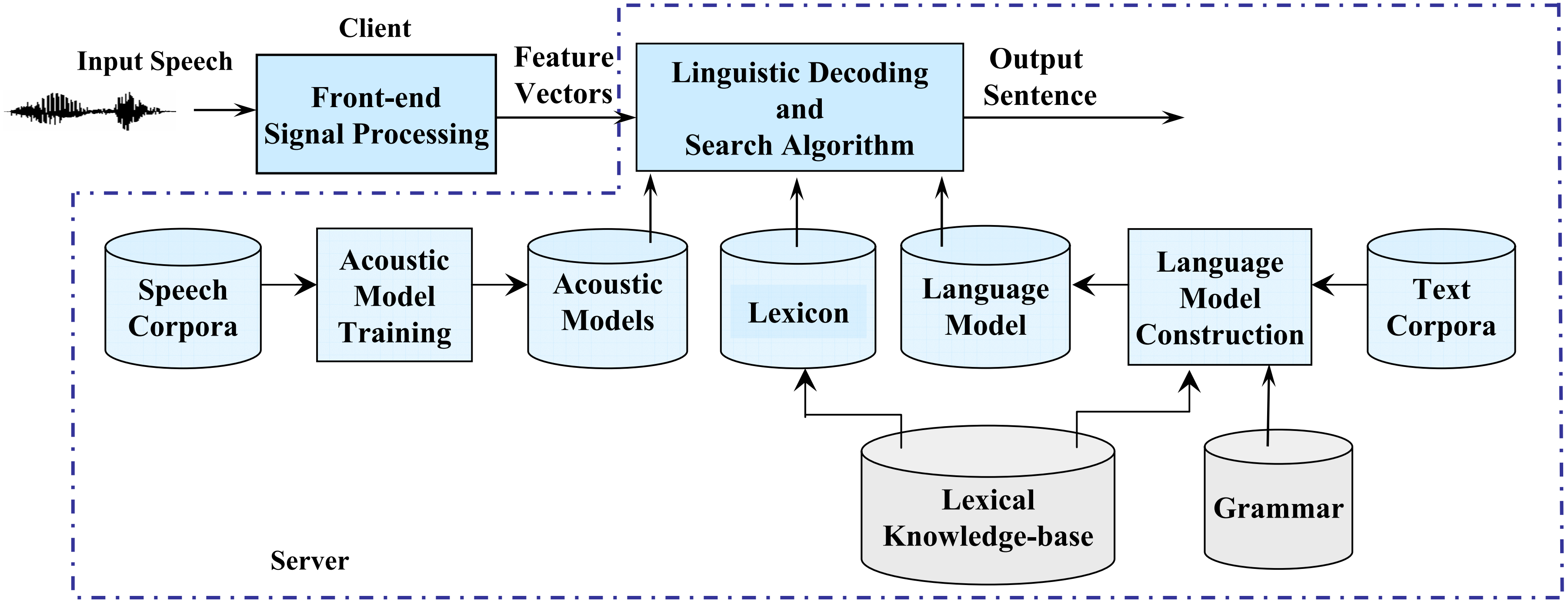
References: 1. “Quantization of Cepstral Parameters for Speech Recognition over the World Wide Web”, IEEE Journal on Selected Areas in Communications, Jan 1999

2. “A Bitstream-based Frond-end for Wireless Speech Recognition on IS-136 Communication Systems”, IEEE Trans. on Speech & Audio Processing, July 2001

3. “An Error Protected Speech Recognition System for Wireless Communications”, IEEE Trans. on Wireless Communications, April 2002

Distributed Speech Recognition (DSR) and Wireless Environment

- **An Example Partition of Speech Recognition Processes into Client/Sever**



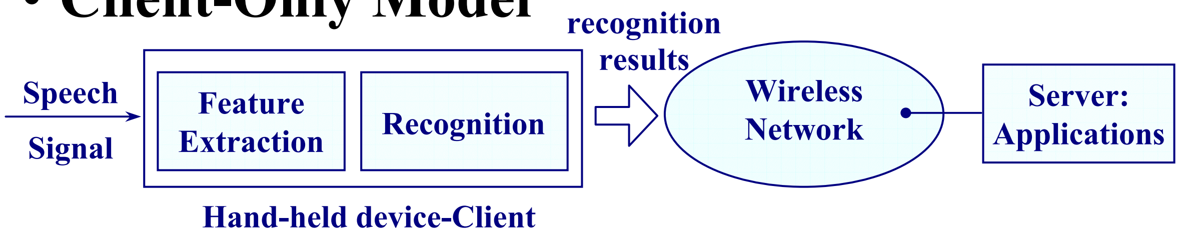
  – compressed and encoded feature parameters transmitted in packets

- **Client/Server Structure**

Possible Models for Distributed Speech Recognition (DSR) under Wireless Environment

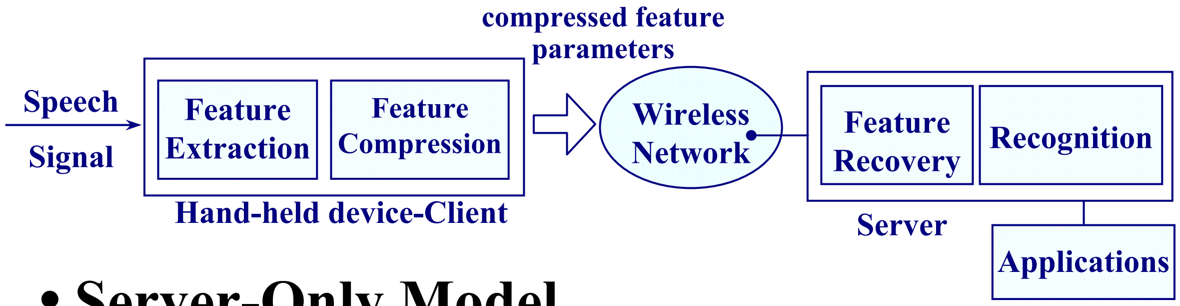
- **Problems with Wireless Networks**
  - limited/dynamic bandwidth, low/time-varying bit rates
  - higher/time-varying error rates, random/bursty errors

- **Client-Only Model**



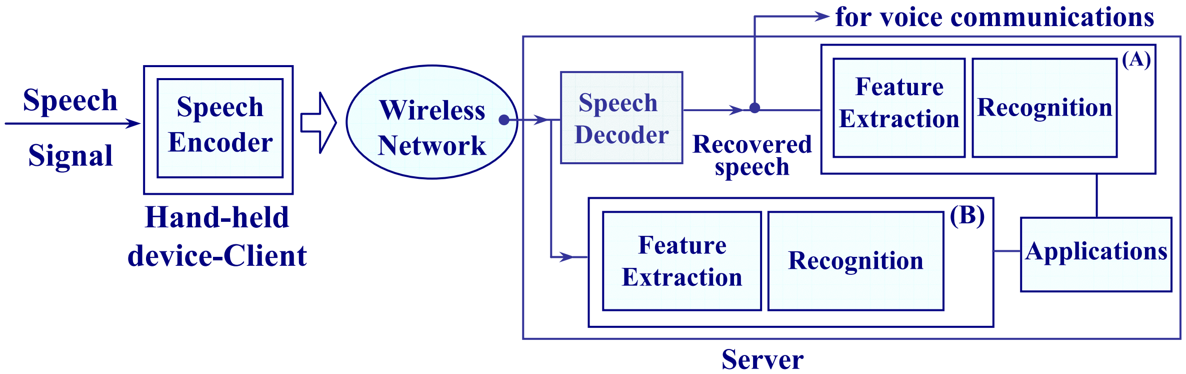
  - speech recognition independent of wireless environment
  - limitation by computation requirements for hand-held device

- **Client-Server Model**



  - proper division of computation requirements on client/server
  - bandwidth saving
  - not compatible to existing wireless voice communications
  - original speech can't be recovered from MFCC

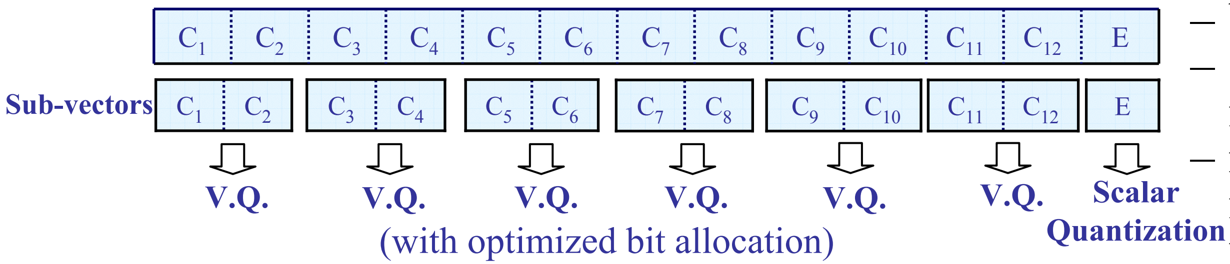
- **Server-Only Model**



  - compatible to existing voice communications
  - seriously degraded recognition accuracy (A)
  - need to find recognition efficient feature parameters out of perceptually efficient feature parameters (B)

Client-Server Model

- **Split Vector Quantization for Feature Parameters (as an example)**



  – delta parameters evaluated at server
  – bit rate minimized
  – computation requirements minimized
  – recognition accuracy degradation minimized with acoustic models trained by quantized parameters (matched condition)

- **Error Protection**
  – different error correction/detection schemes applied to V.Q. bit patterns for different parameters or different bits for scalar quantization
  – important bits well protected while extra bit rate for error protection minimized
  – correct identification of errors very helpful but at higher cost
  – compatibility with existing wireless networking platform needed

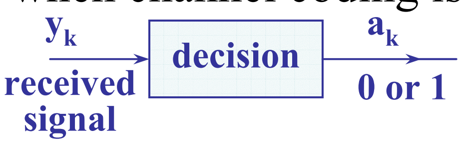
- **Error Concealment Examples**
  – extrapolation

$$\hat{x}_t = \beta \cdot \frac{1}{L} \sum_{k=1}^{L} x_{t-k} + (1-\beta) x_{t-1}$$

x_t: feature vector at time index t

  – interpolation also possible
  – performed with those sub-vectors with errors (if identifiable) only

- **Error Detection Example**
  – when channel coding is not able to detect errors or identify error bits



$$\rho(a_k; y_k) = \log\left(\frac{\text{Prob}[a_k = 1 | y_k]}{\text{Prob}[a_k = 0 | y_k]}\right) \qquad |\rho(a_k; y_k)| < th \ : \text{potential error}$$