Efficient Search Strategies in Hierarchical Pattern Recognition Systems

by

Neeraj Deshmukh	Joseph Picone	Yu-Hung Kao
Dept. of EE	Inst. for Signal & Info. Proc.	Systems & Info. Sci. Lab.
Boston University	Mississippi State University	Texas Instruments Inc.
Boston, MA 02215	MS State, MS 39762	Dallas, Texas 75243
neeraj@engc.bu.edu	*picone@ee.msstate.edu*	*yhkao@csc.ti.com*

SSST, '95

March 13, 1995

The Problem of Statistical Pattern Recognition

♠ **Mathematical representation**

$$p(\widehat{W}/A) = \max_{W} p(W/A)$$

Bayes' Theorem:

$$p(\widehat{W}/A) = \arg\max_{W} p(A/W)p(W)$$

♠ **Automatic Speech Recognition**

- Acoustic Model
- Language Model
- Search

♠ **Hidden Markov Models**

♠ **Complex Applications ⇒ Hierarchical Modeling**

- Large Vocabulary Continuous Speech Recognition (LVCSR)
- Target recognition in radars, SAR imagery
- Intelligent database access (audio-video)

Speech Recognition System

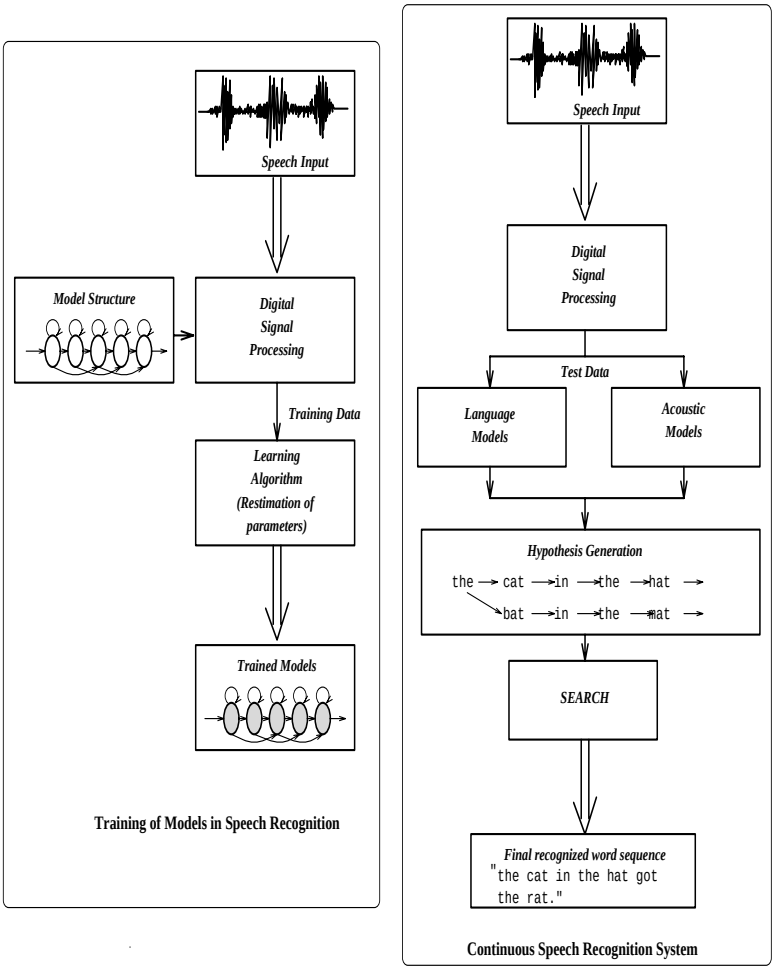


Figure 1: Training and Recognition

Schematic of training and recognition systems

Search Strategies

♠ **Search Paradigm**

To choose a pattern with the highest likelihood score for our feature models given the observed data.

♠ **Motivation**

The number of hypotheses (choices for the correct pattern) grows exponentially with number of feature models. Hence a strategy that saves on computation and storage requirements is sought.

♠ **Popular search techniques**

- Viterbi Search
- Viterbi Beam Search
- A* Stack Decoding
- N-best Search
  - ▷ Maintains *all* hypotheses within specified beam.
  - ▷ Propagates top N hypotheses at each state.
  - ▷ N is independent of Viterbi beam.
  - ▷ Tool to integrate information from multiple sources.
  - ▷ Partial towards shorter hypotheses.
- Generalized N-best Search
  - ▷ Forward-Backward Search
  - ▷ Progressive Lattice Search

Application to Hierarchical Systems

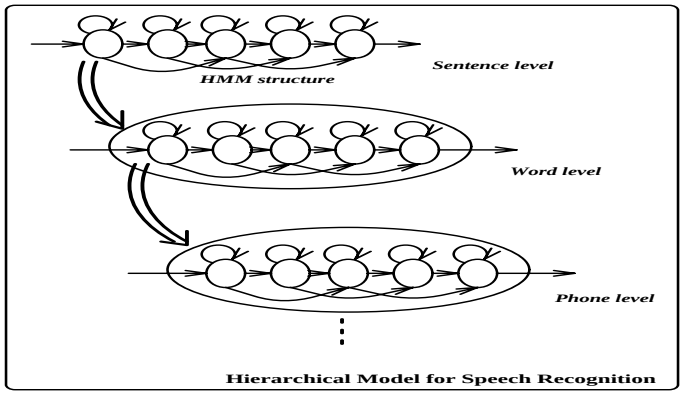


Figure 2: Hierarchical structure of feature models

- Multi-level computation.
- Information flow both across and along layers of model framework.
- Excessive requirements on computation and storage capacity.
- For N-best, N different paths are to be traced back ⇒ degradation of performance of Viterbi search.
- Integration of different N-best hypotheses obtained from different levels.

Frame-Synchronous Viterbi Search

♠ **Motivation**

Reduce computational and memory requirements by limiting the number of valid hypotheses to be processed.

♠ **FSVS Algorithm**

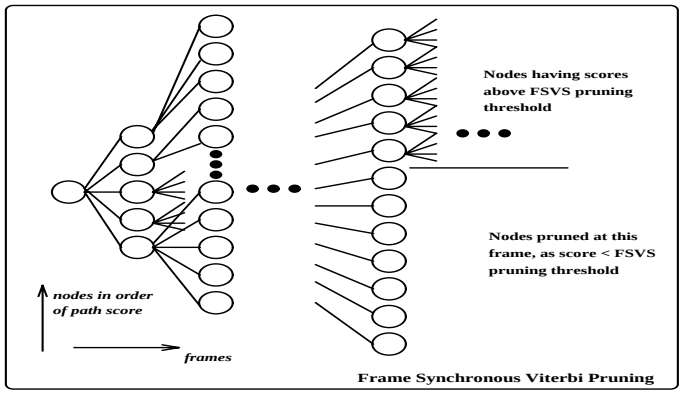


Figure 3: FSVS pruning algorithm

During Viterbi beam search, at the end of each frame of input data

1. Sort all active hypotheses in decreasing order of scores.
2. Keep only a few top-scoring hypotheses based on some pruning threshold. This threshold can be fixed or dynamic, depending on the application.

Practical Issues

♠ **Computational overhead**

♠ **The perils of over-pruning**

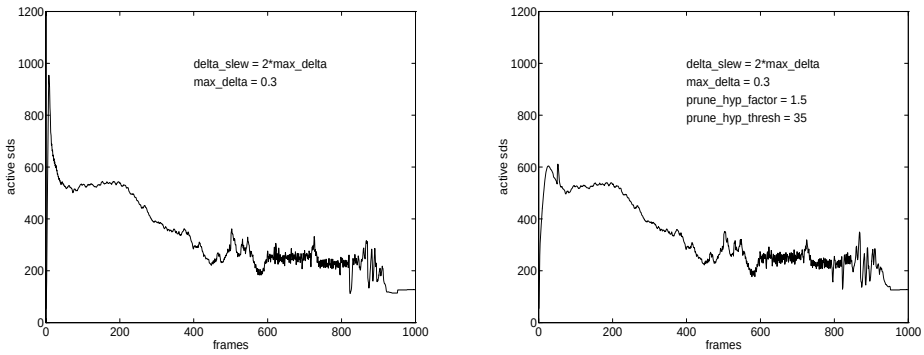
♠ **Goal: Generalized N-best Algorithm**

- should identify models associated with output symbols.
- should keep N highest scoring paths for such models at every level in the model hierarchy.
- should trace all these paths to obtain N best hypotheses at the top level.
- The generalized N-best search for a hierarchical system thus maintains N-best hypotheses at every level and not only at the top.

♠ **Advantages**

- Reduction in problem space
- Very useful in memory-critical applications
- Some gain in computational efficiency

Experimental Results



a: Viterbi beam search b: FSVS

Figure 4: Memory usage for (speech) sentence recognition

Type of pruning	Sent. over flow	Comptn. (fracn of realtime)	Mem. slots/ frame	Word error %	Sent. error %
Viterbi	308	0.289	590	24.2	36.1
FSVS	0	0.274	424	3.5	21.2

Results of Frame-Synchronous Viterbi Search

Summary

- ♠ Hierarchical pattern recognition systems are required to solve complex pattern matching applications.

- ♠ Such systems have excessive requirements of memory and computational power for search.

- ♠ Frame-synchronous Viterbi Search algorithm is a step in reducing the problem space.

- ♠ FSVS algorithm is particularly attractive to memory-critical recognition tasks.