

General Sound Recognition and Similarity Tools

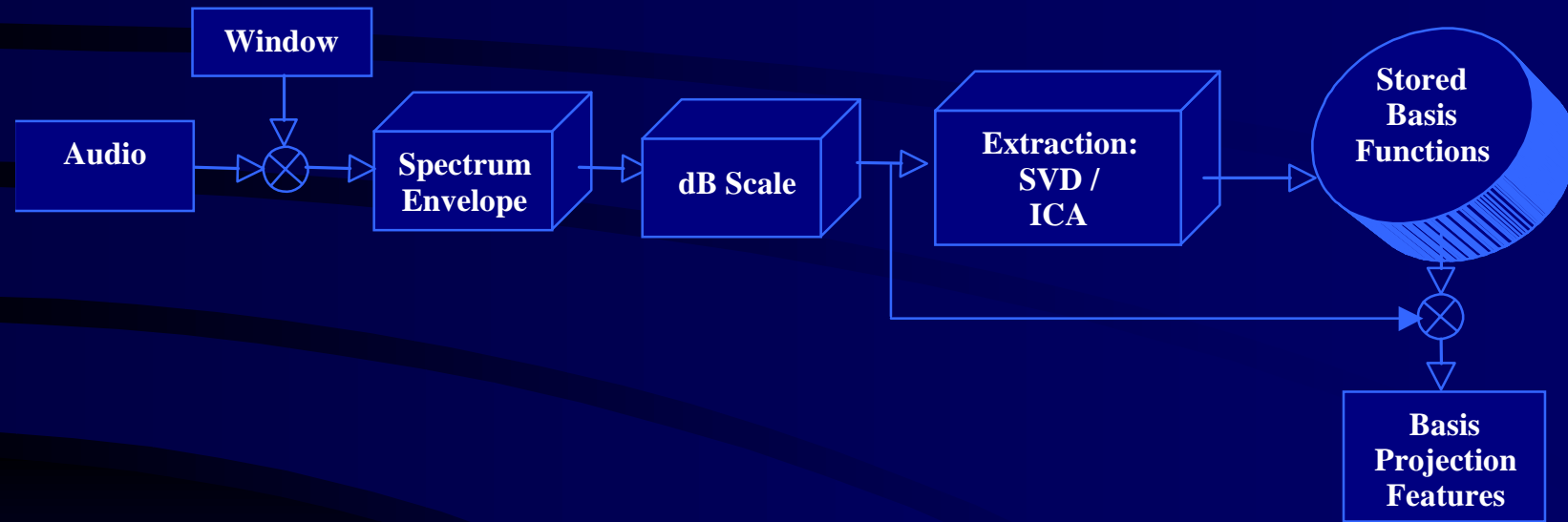
Michael Casey

MERL, Cambridge Research Lab

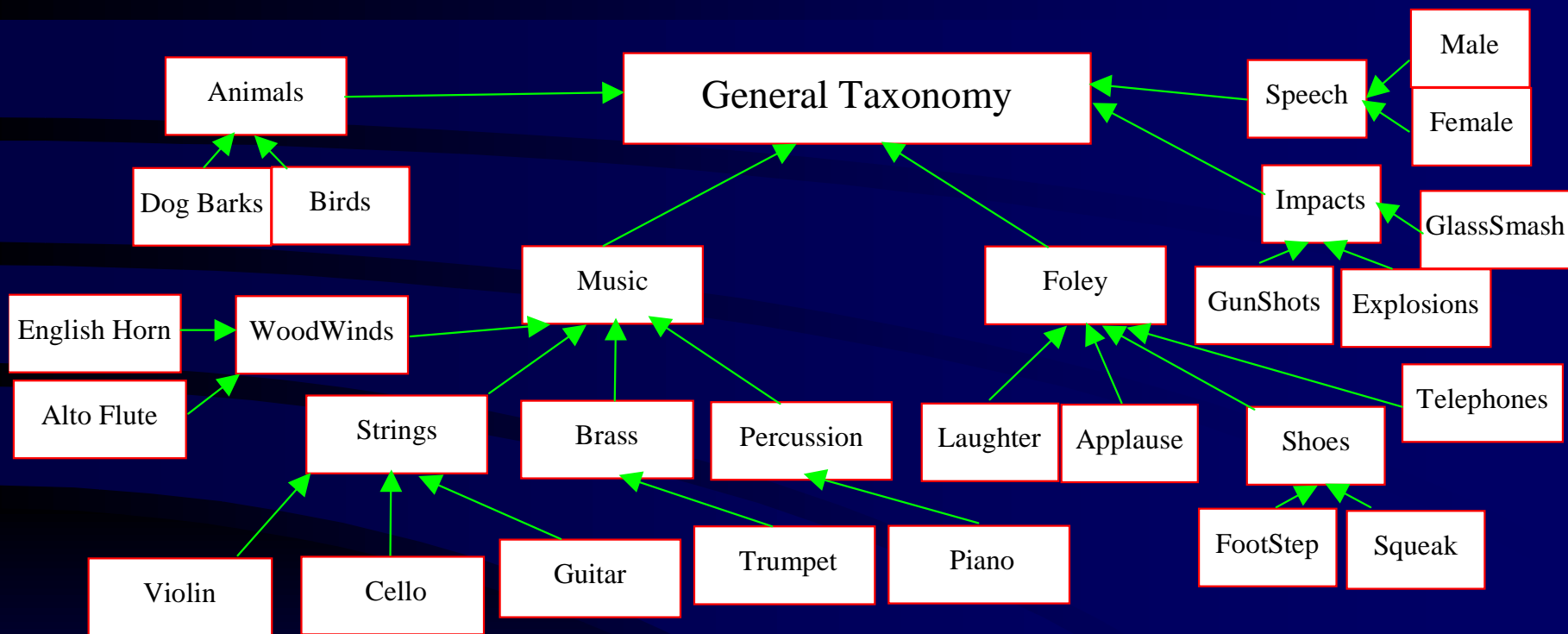
Searchable Audio

- **Sound Classification**
 - Robust Features
 - Taxonomy of Categories
 - HMM Classifier
- **Sound Indexing and Matching**
 - HMM Path Matching
 - Query by Example

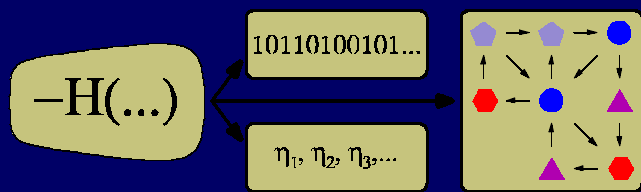
AudioSpectrumBasis SpectrumBasisProjection



Sound Taxonomy



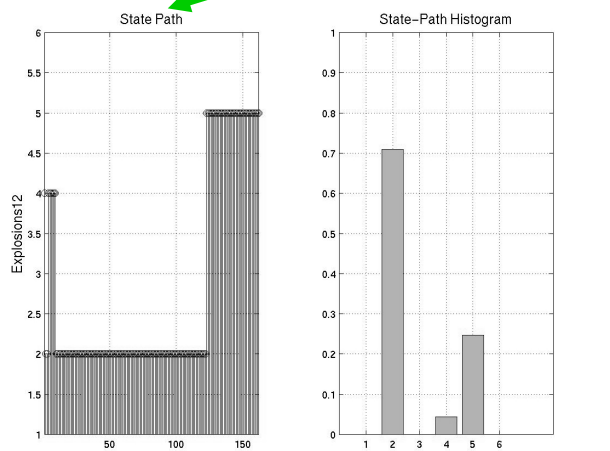
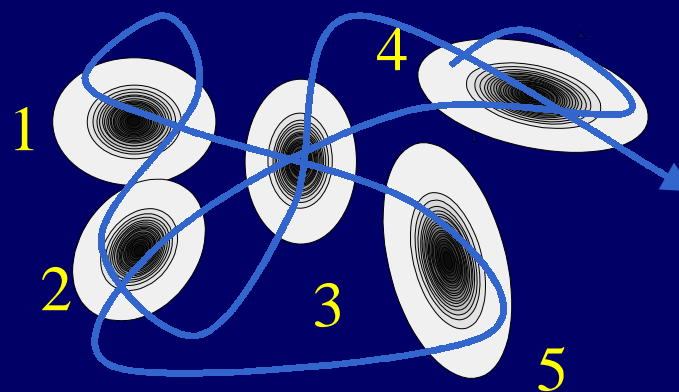
Sound Description = State Path



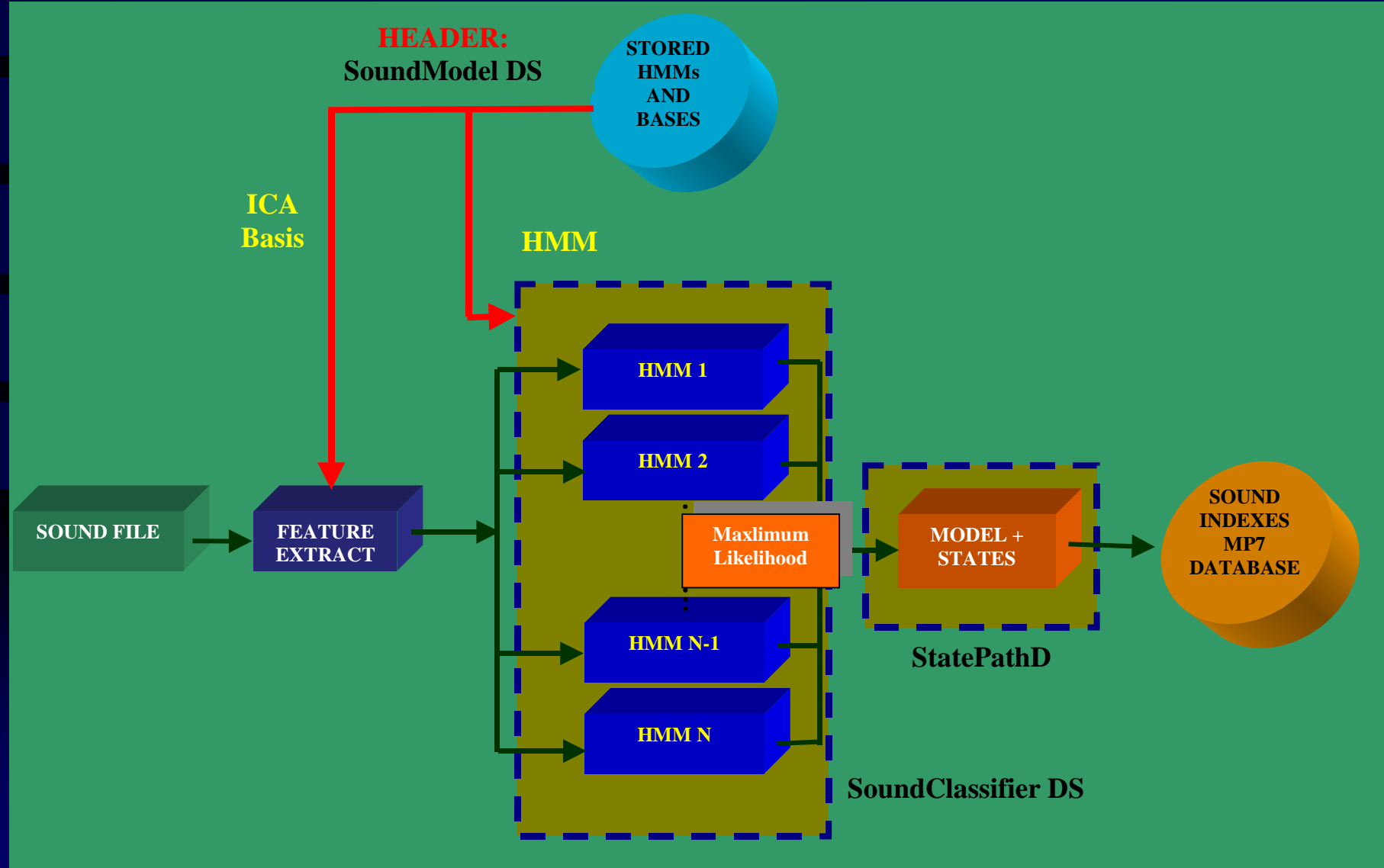
HMM

4 4 4 2 2 2 2 2 2 2 3 3 3 3 2 2 2 5 5 5 5

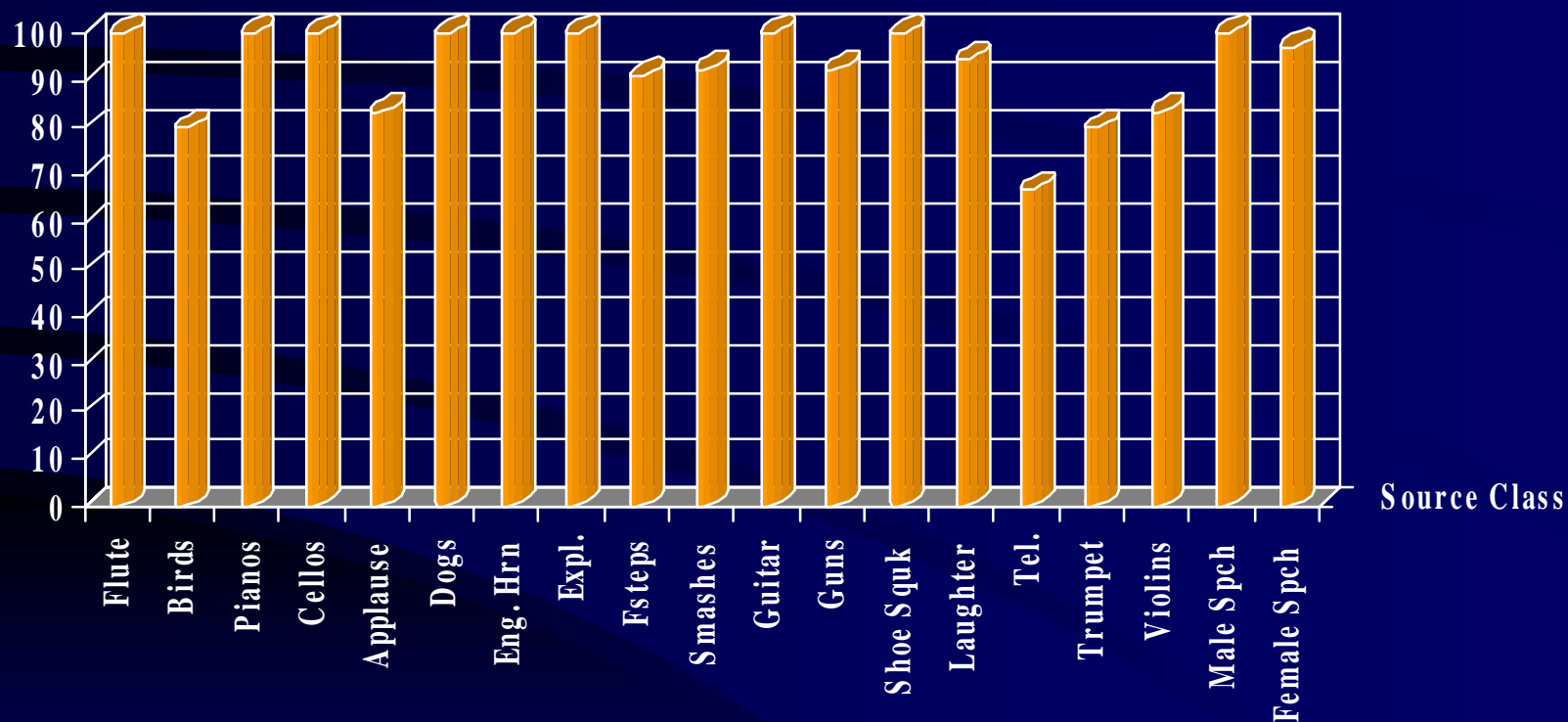
State Path



Sound Indexing

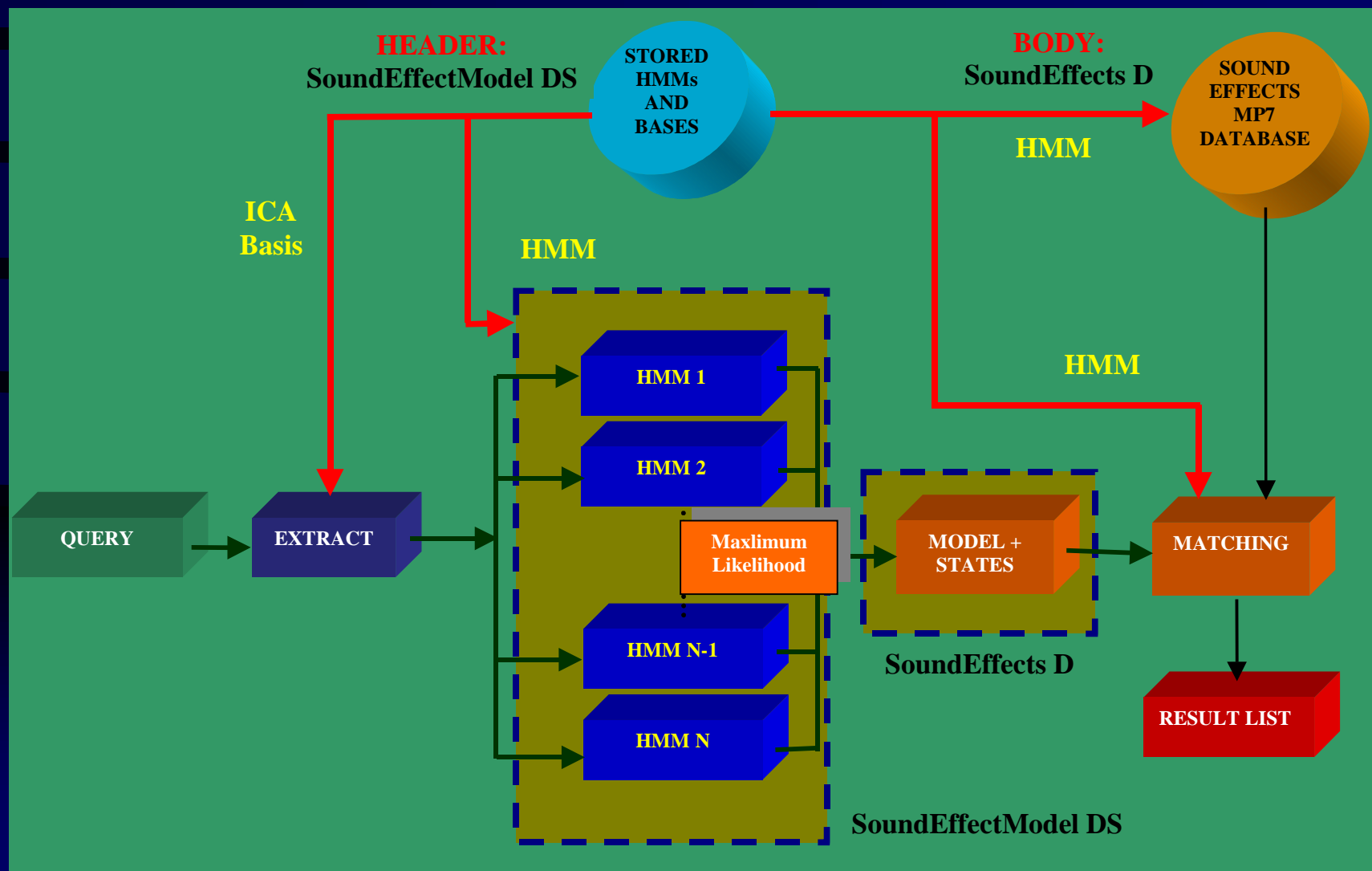


General Sound Classifier Performance



Mean Recognition = 92.646%

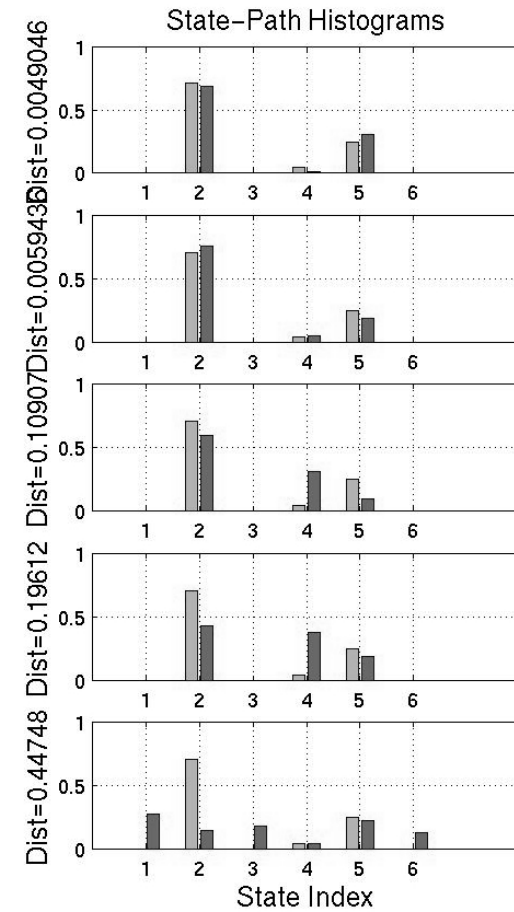
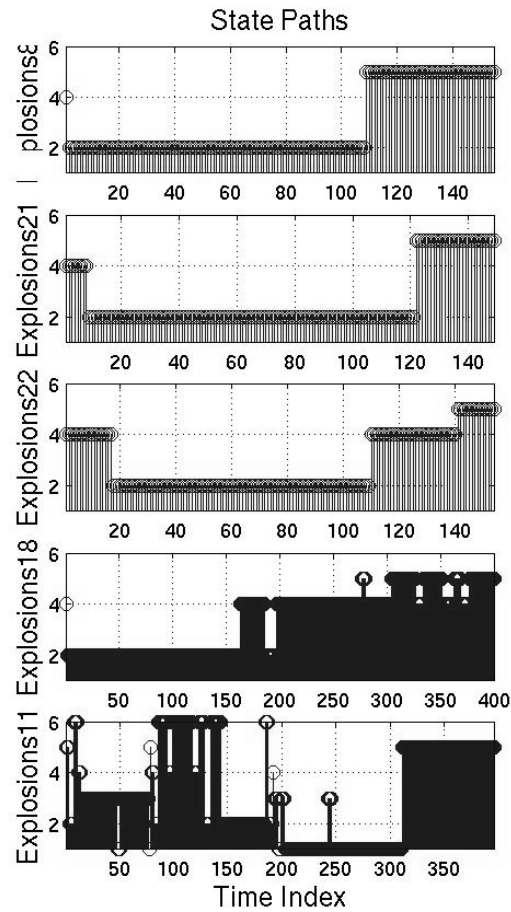
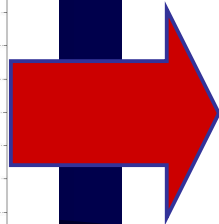
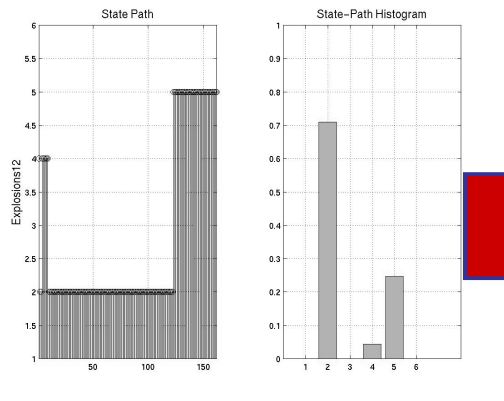
Audio Query by Example



Query-by-Example Matching

Audio Query

Query Result List



State-Path Matching = SSE / DTW

Applications

- Film / video indexing by audio

Index scenes with animals, gun fights, laughter, explosions

- Film / video search by audio

Find scenes with animals, gun fights, laughter, explosions

- Professional sound archiving

Place new sound in correct archive directory

- Semantic label => context

Infer scene type from audio labels:

Gun shot / Explosion => Action Scene