

ICASSP '98 Key Papers

Matthew Siegler
ECE Department
Speech Group

Carnegie
Mellon

Recognition Confidence

- SP7.1 **Experiments in Confidence Scoring Using Spanish Callhome Data** J.Vaver
- SP7.2 **A New Decoder Based on a Generalized Confidence Score**
M.Koo,C.Lee,B.Juang
- SP7.5 **Using Word Probabilities as Confidence Measures**
F.Wessel,K.Macherey,R.Schluter
- SP7.7 **A Fast Vocabulary Independent Algorithm for Spotting Words in Speech**
S.Dharanipragada,S.Roukos

Other Papers

- **SP13.5 LVCSR Rescoring with Modified Loss Functions: A Decision Theoretic Perspective** V.Goel,W.Byrne,S.Khudanpur
- **SP17.7 Separation of Spontaneous and Non-Spontaneous Speech** O.Kenny,D.Nelson,J.Bodenschatz,H.McMonagle
- **SP29.7 Baby Ears: A Recognition System for Affective Vocalizations**
M.Slaney,G.McRoberts

Important Points

- SP7.1 Call-Home database.
Tried various smoothing functions.
Phone Duration is **not** a good predictor.
- SP7.2 Build a decoder from confidence
Attempted to integrate predictors.
Exponential smoothing functions.
“Generalized Confidence Score” = product.

Experiment: car rental task -- Keyword bound.
Instead of ML formulation, was able to
perform recognition pass.

Important Points

- SP7.5 New estimator for word probability.
Estimating posteriors from lattices.
Convert into word probabilities.
Experiment: NAB, h194
Reduced Lattice-WER by 20%
- SP7.7 Goal: Wordspotting, OOV detect for IR
Coarse-to-Fine structure.
Fast phone-based then word-based.
Combined with IR using word or n-phone.
Experiment: TREC6-test (10 hours ~100kw)
Able to detect 50% OOV with 10 false alarms.

Important Points

- SP13.5 Use Bayesian decision criterion.
Various optimization of loss functions.
Rescoring n-best lists.
Experiment: LVCSR task
Small improvements in WER. <- why?
- SP17.7 Analyze lots of features, wavelet.
Classify stress: Fast, Loud, Lombard, Neutral,
Question, Slow, Soft, Cond50/70 (computer)
Beat out MFCC by 9-14%.
Used “SUSAS: Speech Under Simulated and
Actual Stress Database”

Important Points

- SP29.7 Detect emotion in parent-child.
Approval, Attention, Prohibition.
Features: Pitch, formants.
Gaussian discriminator.
Experiment: actual parent-infant conversations
Data reviewed by adults and classified.
Depending on analysis, ~60% correct.