TELEPHONE SPEECH RECOGNITION USING NEURAL NETWORKS AND HIDDEN MARKOV MODELS

DongSuk Yuk[†‡] and James Flanagan[†]

[†]CAIP Center, Rutgers University, Piscataway, NJ

[‡]Department of Computer Science, Rutgers University, New Br

{yuk,jlf}@caip.rutgers.edu

ABSTRACT

The performance of well-trained speech recognizers using high
quality full bandwidth speech data is usually de
used in real world environment
recogniti

the testing and the training environments. Since the low quality
speech feature vectors are transformed to high quality ones, it
can outperform the retrained recognizer that is train
quality speech. Therefore, the p
approach is

monophone HMM with 30 Gaussian distributions per state. In
total, the system has 3,630 Gaussian distributions. 1,344 utter-
ances from NTIMIT[2] test data are used for testing. When th
system is trained and tested under the sa
both using TIMIT co

a system which is comparable to the retrained recognizer, but
with much less training data.
The advantages of this approach are as follows. First, it does
not require retraining of the speech recognizer
task in terms of training
i