Locality-Sensitive Hashing

- phoneme proximity
  - extract features from the whole TIMIT corpus (462 speakers, 10 sentences/speaker)
    * 39MFCCs, frame=25ms, step=10ms
  - train general UBM with all the features extracted from the TIMIT corpus
  - create files for each phonemes of each speaker (≈ 462 · 61) → perform MAP adaptation

- analysis of the results
  - figures 1-3 present results for phoneme classification and positive example detection using 4, 8 and 16 gaussian mixtures, respectively
  - results are inconclusive, as it seems that there is not enough training material for performing an adequate MAP adaptation

(a) Phoneme classification.

(b) Positive examples detection.

Figure 1: Detection results, 4 gaussian mixtures.
Results of E2LSH on phones (% correct classification)

Results of E2LSH on phones (% positives found)

(a) Phoneme classification.
(b) Positive examples detection.

Figure 2: Detection results, 8 gaussian mixtures.

Results of E2LSH on phones (% correct classification)

Results of E2LSH on phones (% positives found)

(a) Phoneme classification.
(b) Positive examples detection.

Figure 3: Detection results, 16 gaussian mixtures.