Christian-Albrechts-Universität zu Kiel

Questions A

A1. Codebook definition

❑ What requirement is posed on the mean distance between feature vectors and codebook entries? Explain the respective equation.

A2. Feature extraction requirements

❑ What are the feature extraction requirements posed by the vector quantization?

A3. Outputs of the vector quantization

❑ What are the outputs of vector quantization?

❑ Which of these outputs is used for speaker recognition? Why?

❑ Which of these outputs is used for speech reconstruction and bandwidth extension? Why?

❑ How are the outputs of the vector quantization determined if both the codebook and the test feature vector are known?

Questions B

B1. Cost function

- ❑ The same distance function should be used both at the creation of a codebook and at the later application. Why is that?
- ❑ Why and in what way is a limitation of the cost function introduced?

B2. Codebook training

- ❑ Why is the codebook training done iteratively?
- ❑ Explain the different conditions for termination that can be used for codebook training.

B3. k-means and LBG

- ❑ Explain the k-means procedure for codebook training.
- ❑ What are the differences between k-means and LBG?
- ❑ Which of the methods is faster in terms of computational complexity?

Christian-Albrechts-Universität zu Kiel

Answers B

B1. Cost function

❑ Otherwise, a test vector could be assigned to a codebook vector c_i during the training, but assigned to a different codebook vector c_j during later application. As a result, the average distance is probably no longer minimized (see slide 6).

❑ logarithmic features may have very large negative values (slide 19) + slide 20.

B2. Codebook training

❑ Up to now, no optimal non-iterative procedure is known.

❑ Conditions for termination:
- The mean distance is below a certain threshold.
- The mean distance is only slightly decreased.
- The maximum number of iterations is reached.

B3. k-means and LBG

❑ See slide 23.

❑ See slide 24; *k-means* starts with *K* randomly chosen training vectors as a codebook. Whereas LBG starts with the overall average of the training vectors as single codebook vector and increases the number of codebook vectors in an iterative way.

❑ LBG starts with a lower number of codebook vectors, while the number of iterations is comparable to k-means; Thus the computational complexity of LBG is usually lower.

Answers A

A1. Codebook definition

❑ See bottom of slide 6. The average distance of the training vectors to the respective closest codebook vector is to be minimized.

A2. Feature extraction requirements

❑ Based on a simple distance function, the vector quantization should be able to calculate "reasonable" distances between feature vectors.

A3. Outputs of the vector quantization

❑ The vector quantization produces a codebook vector (or its index) as well as the distance between the test vector and the corresponding codebook vector.

❑ For speaker recognition, only the accumulated distance is needed. Based on this, a measure for the distance between the training set and the test set can be derived.

❑ For speech reconstruction and bandwidth extension, the codebook vectors themselves are used. They contain information about the wide-band envelope that is lost in the test vectors.

❑ The minimum distance between the test vector and all codebook vectors is determined.