Optimal Parameter Estimation for Model-Based Quantization

Alexey Ozerov 1 and W. Bastiaan Kleijn 2

1Institut Télécom, Télécom ParisTech, CNRS LTCI, Paris, France
2ACCESS Linnaeus Center, Electrical Engineering, KTH - Royal Institute of Technology, Stockholm, Sweden
e-mail: alexey.ozerov@telecom-paristech.fr, bastiaan.kleijn@ee.kth.se

Abstract

Two speech-coding schemes:

A. Ozerov and W. B. Kleijn, “Flexible quantization of audio and
We address

Practical optimization by Newton’s method:

A sequence of source vectors \( s = \{s^n\} \) is quantized using a sequence
of Gaussian models \( \theta = \{\theta_1, \ldots, \theta_n\} \), called hereafter model.

Under HR theory assumptions the (average) rate \( R \) (in bits per vector)
related to the (average) distortion \( D \) (per dimension) as:

\[
R = \frac{k}{2} \log_2 D + \psi(s, \theta),
\]

where

\[
\psi(s, \theta) = \frac{k}{2} \log_2 \left( \frac{2\pi e |C|}{k} \right) + \frac{k}{2} \log_2 \left( \frac{1}{\sqrt{2\pi e}} \prod_{i=1}^k \lambda_i \right),
\]

\[
\psi(s, \theta) = \frac{k}{2} \log_2 C - \frac{1}{2} \log_2 \prod_{i=1}^k N(s^n, \mu_i, \Sigma_i).
\]

with \( C = 1/12 \) and \( y^n = U^n \theta \), where \( \Sigma = U^T \Sigma U \).

Proposed Model Estimation Criterion

Maximum Likelihood (ML) criterion:

\[
\hat{\theta}_{ML} = \arg \max p(s|\theta) = \arg \max_{\theta} \prod_{n=1}^N N(s^n, \mu_n, \Sigma_n)
\]

is equivalent to \( \psi(s, \theta) \) minimization (consistent with the minimum
description length (MDL) principle), but not to \( \psi(s, \theta) \) minimization.

Proposed CR-MDL criterion (our work is related to [1]):

\[
\hat{\theta}_{CR-MDL} = \arg \min_{\theta} \psi(s, \theta).
\]

Practical optimization by Newton’s method:

\[
\theta^{n+1} = \theta^n - \gamma \left[ \nabla \psi(s, \theta^n) \right]^{-1} \nabla \psi(s, \theta^n).
\]

Toy Examples

Generalized Gaussian distributions with shape parameters 1.7 and 1.

Results

Two speech-coding schemes:

- AR model based scheme with KLT (AR-KLT) [2].
- MLT based scheme with a fixed frequency weighting (MLT-FFW).

The rate for the model should be constant [3] (not quantized here).

Conclusion

- Proposed CR-MDL criterion is optimal under HR theory assumptions.
- Compared to ML, CR-MDL improves both HR theory predicted
and practical CR quantization performances for a range of conditions.
- The larger the mismatch between the actual data distribution
and model distribution, the greater the performance improvement.

References

with Gaussian mixture components,” IEEE Trans. on Audio, Speech
and speech based on the autoregressive model,” in Proc. Asilomar