Optimal Parameter Estimation for Model-Based Quantization

Alexey Ozerov [1] and W. Bastiaan Kleijn [2]

[1] *Institut Télécom, Télécom ParisTech, CNRS LTCI, Paris, France*

[2] *ACCESS Linnaeus Center, Electrical Engineering, KTH - Royal Institute of Technology, Stockholm, Sweden*

e-mail: alexey.ozerov@telecom-paristech.fr,bastiaan.kleijn@ee.kth.se

Abstract

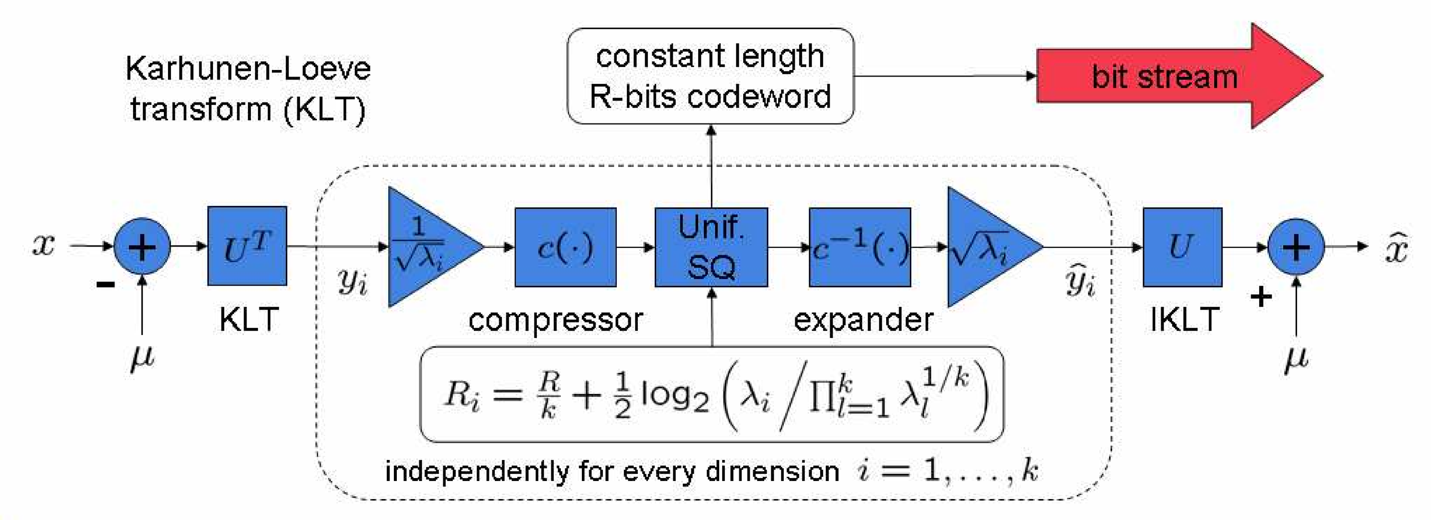
- We address optimal model estimation for model-based vector quantization for both constraint resolution (CR) (constant rate) and constraint entropy (CE) (variable rate) cases.
- Assuming a Gaussian model we derive under high-rate (HR) theory assumptions the rate-distortion (RD) relations for these two cases.
- Based on the RD relations we show that the maximum likelihood (ML) criterion leads to optimal performance for CE quantization, but not for CR quantization.
- We introduce a new model estimation criterion for CR quantization that is optimal in terms of the RD relation.
- Our experiments confirm that the proposed criterion for model identification outperforms the ML criterion for a range of conditions.

Model-Based Quantization

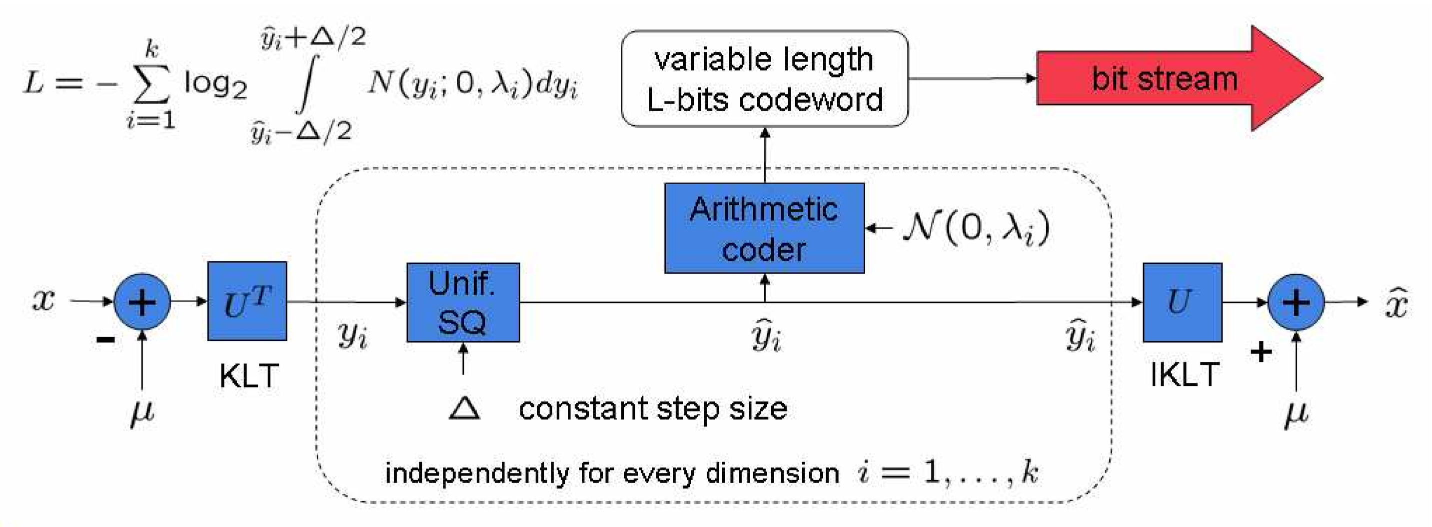
- Assumptions:
  - Source vector $x \in \mathbb{R}^k$ is a particular realization of a k-dimensional random Gaussian vector $X \sim \mathcal{N}(\mu, \Sigma)$.
  - Covariance matrix eigenvalue decomposition (EVD):

$$\Sigma = U \Lambda U^T, \qquad U^T U = I, \quad \Lambda = \mathrm{diag}\{\lambda_i\}_i.$$

- Constrained Resolution quantization (constant rate):



- Constrained Entropy quantization (variable rate):



Practical Rate-Distortion Relations

- A sequence of source vectors $\mathbf{s} = \{s^n\}_{n=1}^N$ is quantized using a sequence of Gaussian models $\theta = \{\theta_n\}_{n=1}^N$ ($\theta_n = \{\mu_n, \Sigma_n\}$), called here-after *model*.
- Under HR theory assumptions the (average) rate R (in bits per vector) is related to the (average) distortion D (per dimension) as:

$$R = -\frac{k}{2}\log_2 D + \psi(\mathbf{s}, \theta),$$

where

$$\psi_{\mathrm{CR}}(\mathbf{s}, \theta) = \frac{k}{2}\log_2\left(\frac{3(2\pi)^{\frac{2}{3}}C}{k}\right) + \frac{k}{2}\log_2 \frac{1}{N}\sum_{n=1}^{N}\left[\prod_{l=1}^{k}\lambda_{n,l}^{\frac{1}{k}} \sum_{i=1}^{k}\lambda_{n,i}^{-\frac{1}{3}} N(y_i^n; 0, \lambda_{n,i})^{-\frac{2}{3}}\right],$$

$$\psi_{\mathrm{CE}}(\mathbf{s}, \theta) = \frac{k}{2}\log_2 C - \frac{1}{N}\log_2 \prod_{n=1}^{N} N(s^n; \mu_n, \Sigma_n).$$

with $C = 1/12$ and $y^n = U_n^T(s^n - \mu_n)$, where $\Sigma_n = U_n \Lambda_n U_n^T$.

Proposed Model Estimation Criterion

- Maximum Likelihood (ML) criterion:

$$\theta_{\mathrm{ML}} = \arg\max_{\theta} p(\mathbf{s}|\theta) = \arg\max_{\theta} \prod_{n=1}^{N} N(s^n; \mu_n, \Sigma_n)$$

is equivalent to $\psi_{\mathrm{CE}}(\mathbf{s}, \theta)$ minimization (consistent with the minimum description length (MDL) principle), but not to $\psi_{\mathrm{CR}}(\mathbf{s}, \theta)$ minimization.

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

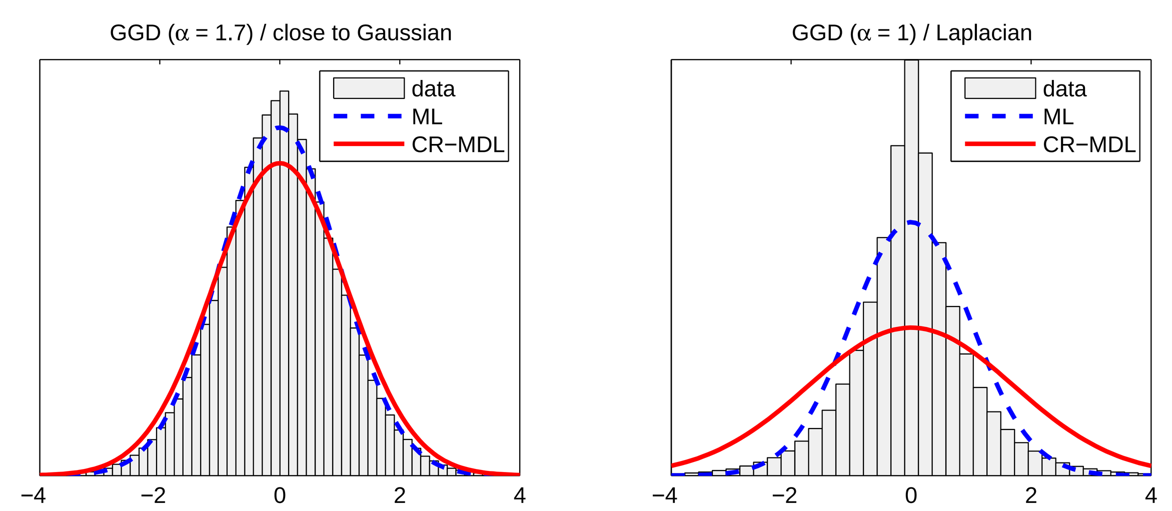
$$\theta_{\mathrm{CR\_MDL}} = \arg\min_{\theta} \psi_{\mathrm{CR}}(\mathbf{s}, \theta).$$

- Practical optimization by Newton's method:

$$\theta^{m+1} = \theta^m - \gamma\left[H_\theta \psi_{\mathrm{CR}}(\mathbf{s}, \theta^m)\right]^{-1} \nabla_\theta \psi_{\mathrm{CR}}(\mathbf{s}, \theta^m).$$

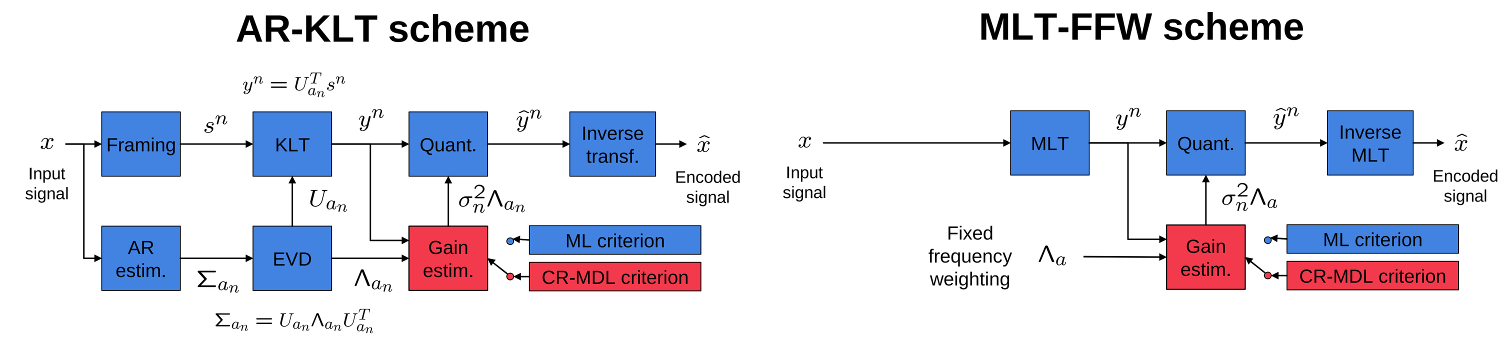
Toy Examples

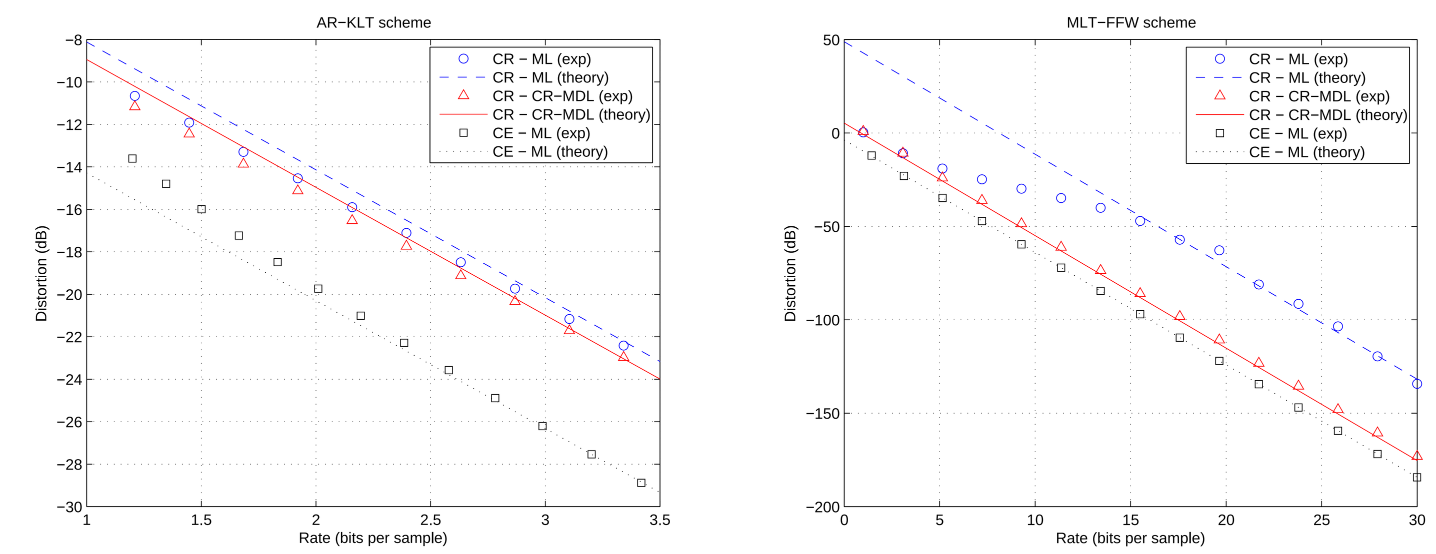
- Generalized Gaussian distributions with shape parameters 1.7 and 1.



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

[1] E. R. Duni and B. D. Rao, "A high-rate optimal transform coder with Gaussian mixture companders," IEEE Trans. on Audio, Speech and Lang. Proc., vol. 15, no. 3, pp. 770–783, Mar 2007.

[2] A. Ozerov and W. B. Kleijn, "Flexible quantization of audio and speech based on the autoregressive model," in Proc. Asilomar CSSC'07, Nov. 2007, pp. 535–539.

[3] W. B. Kleijn and A. Ozerov, "Rate distribution between model and signal," in Proc. IEEE WASPAA'07, 2007, pp. 243–246.