Scalar Quantization with Lossy Binary Coding for Gaussian Sources

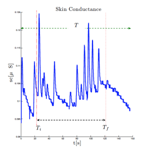
Melanie F. Pradier, Pablo M. Olmos, Fernando Perez-Cruz

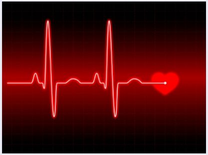
Universidad Carlos III of Madrid

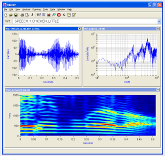
October 7, 2013

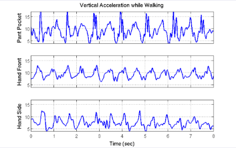
Motivation

Explosion of Data Adquisition

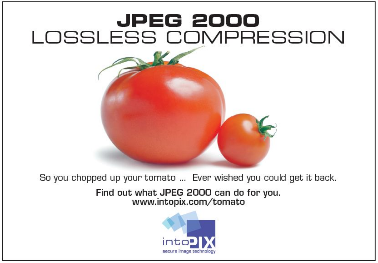








We need Data Compression!

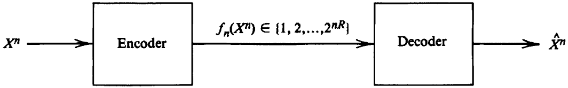


Focus: Lossy Source Coding

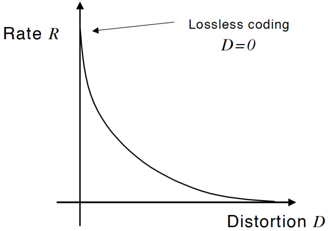
1. **Lossy Source Coding**
    - Rate Distortion Theory
    - State of the Art
2. Scalar Quantization
    - Traditional Approach
    - Our Approach
3. Results

Rate Distortion Theory

- Information theoretical bounds for lossy compression [Cover]



- $X_1, \ldots, X_n \sim p(x)$ i.i.d
- information rate R
- distortion $D = E\left[d\left(X, \hat{X}\right)\right]$
- $R(D) = \min_{p(\hat{x}|x): D \leq D_{max}} I\left(X, \hat{X}\right)$

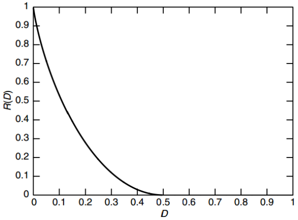


Rate Distortion Theory

Examples of RD bounds

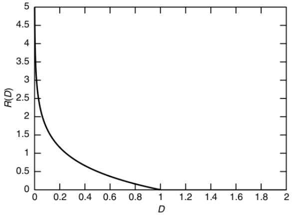
- For Bernoulli source:

$$R(D) = \begin{cases} H(p) - H(D), & 0 \leq D \leq \min\{p, 1-p\} \\ 0 & D > \min\{p, 1-p\} \end{cases}$$

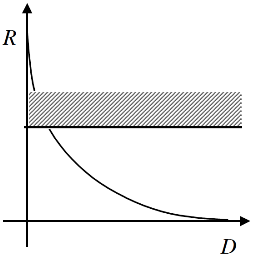


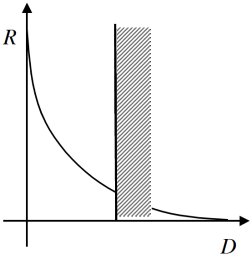
- For Gaussian source:

$$R(D) = \begin{cases} \frac{1}{2}\log\frac{\sigma^2}{D}, & 0 \leq D \leq \sigma^2 \\ 0 & D > \sigma^2 \end{cases}$$



Rate Distortion Theory





$$\min \; D$$
$$s.t. R \leq R_{max}$$

$$\min \; R$$
$$s.t. D \leq D_{max}$$

Rate Distortion Theory

$$R(D) = \min_{p(\hat{x}|x):D \le D_{max}} I\left(X, \hat{X}\right)$$

Interesting Facts

1. Duality between Channel Coding and Source Coding
   - Package Problem Vs Covering Problem
2. Vector Quantization optimal (even if inputs are independent!)
   - Reason: geometry, typicality

State of the Art

- Problem formally defined in 1959 by Shannon

In Theory: Achievability Results

- Inspired in Error Correcting Codes
- Most of works for the BSS
- Only some codes proved to achieve the RD bound:
  - Hamming codes
  - LDPCs, regular LDGMs
  - MN codes
  - ...

In Practice: Codes with good Performance

- Vector Quantization
  - Trellis-based quantizers
  - Lattice codes
  - Sparse codes
  - ...

  $\Rightarrow$ although optimal, very expensive
- **Scalar Quantization**

Is Scalar Quantization a Good Idea? [Ziv]

- any input distribution $p(x)$
- mean-square error constraint $d = (x - \hat{x})^2$
- uniform randomized quantizer + binary entropy encoder

Theorem 1: For any probabilistic n-vector source

$$H(Q_1(\boldsymbol{X} + \boldsymbol{Z})|Z) \leq H_n(\epsilon) + 0.754 \text{ bits/sample}$$

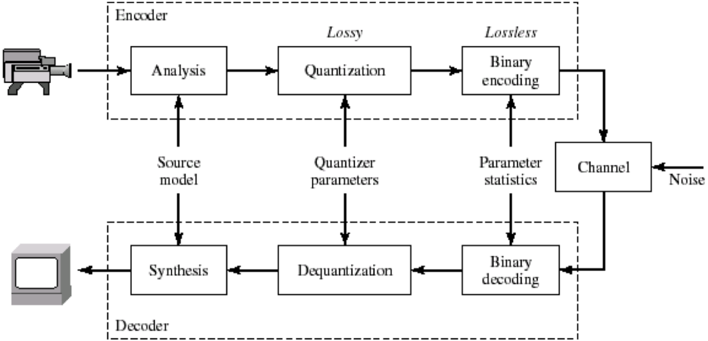
and hence, for stationary sources

$$\lim_{n \to \infty} H(Q_1(\boldsymbol{X} + \boldsymbol{Z})|Z) \leq R(\epsilon) + 0.754 \text{ bits/sample.}$$

	Performance	Complexity
Vector Quantization	:)	:(:(
Scalar Quantization	:|	:)

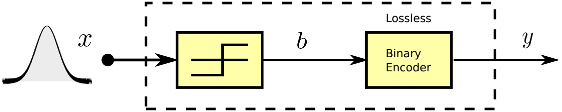
Scalar Quantization

Typical Approach

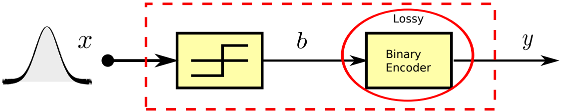


[Yao Wang, Brooklyn]

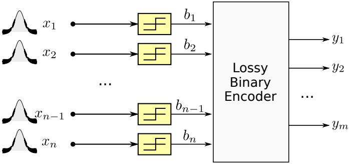
Our Approach

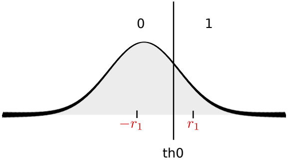


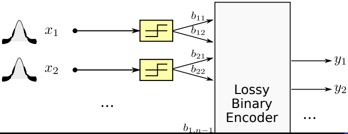
$\Downarrow$

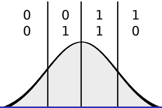


Examples


x_1
x_2
x_{n-1}
x_n
b_1
b_2
b_{n-1}
b_n
Lossy Binary Encoder
y_1
y_2
y_m


0
1
$-r_1$
r_1
th0

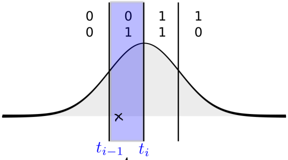

x_1
x_2
b_{11}
b_{12}
b_{21}
b_{22}
Lossy Binary Encoder
y_1
y_2


0 0
0 1
1 1
1 0

Theoretical Formulation

What is the distortion D?

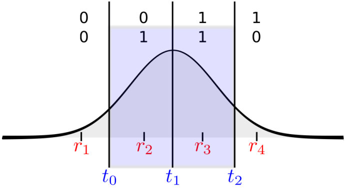
- Lossless: $D = \sum_{A_i} \int_{t_{i-1}}^{t_i} (x - r_i)^2 \phi(x)\, dx$
  with $\phi(x) \sim N(0, \sigma^2)$
- Lossy: $D = \sum_{A_i} \sum_{A_j} pr(A_i \rightarrow A_j) \cdot \int_{t_{j-1}}^{t_j} (x - r_j)^2 \phi(x)\, dx$



Theoretical Formulation

What is the rate R?

- $R = R_q \cdot R_b$
- We control $R_b \Rightarrow P_{err}$ for each bit
  $R_b = H(p) - H(P_{err})$



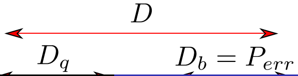
- If we have many bits, different rate allocation
  $R_b = \alpha_1 \cdot R_{b_1} + \ldots + \alpha_K \cdot R_{b_K}$

Theoretical Formulation

Optimization Problem

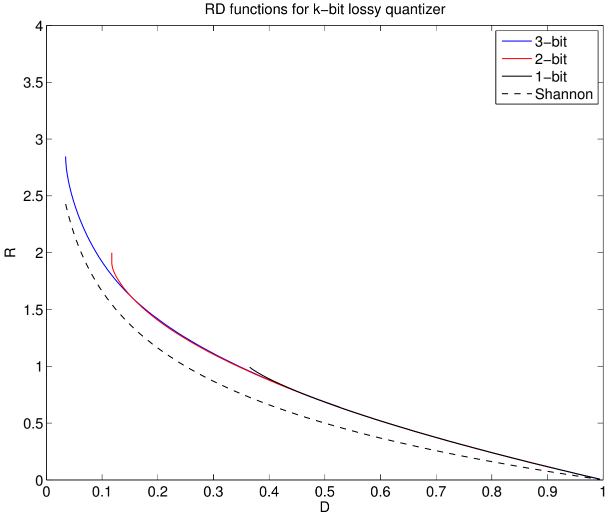
$$\min \; D = \sum_{A_i} \sum_{A_j} pr\left(A_i \rightarrow A_j\right) \cdot \int_{t_{j-1}}^{t_j} \left(x - r_j\right)^2 \phi\left(x\right) dx$$

$$s.t. R_{fixed} = R_q \cdot \sum_k \alpha_k \left[H\left(p_k\right) - H\left(D_{b,k}\right)\right]$$

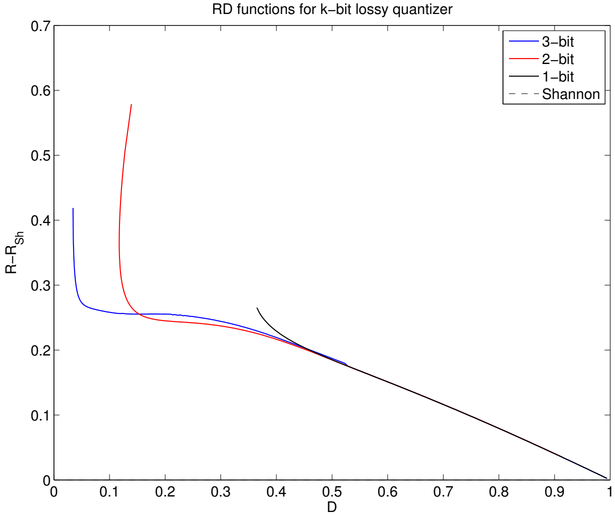


1. Lossy Source Coding
   - Rate Distortion Theory
   - State of the Art
2. Scalar Quantization
   - Traditional Approach
   - Our Approach
3. **Results**

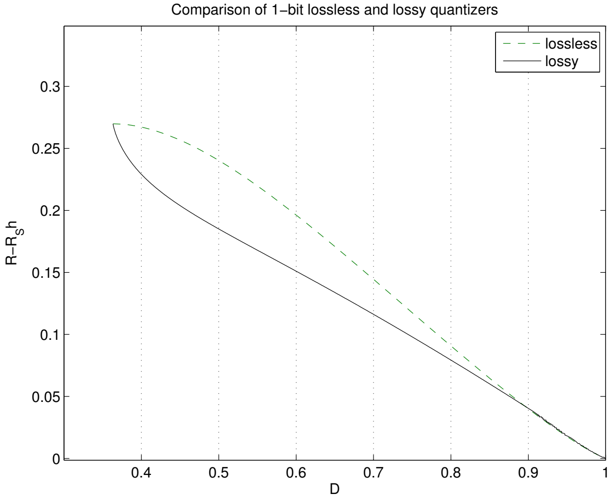
Results: Performance



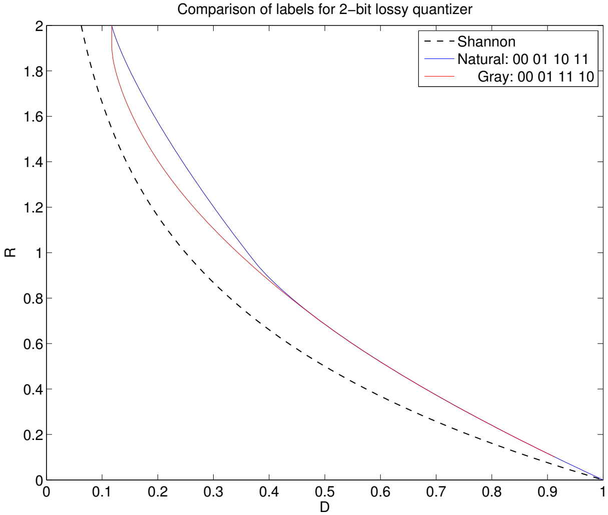
Results: Asymptotic Behavior



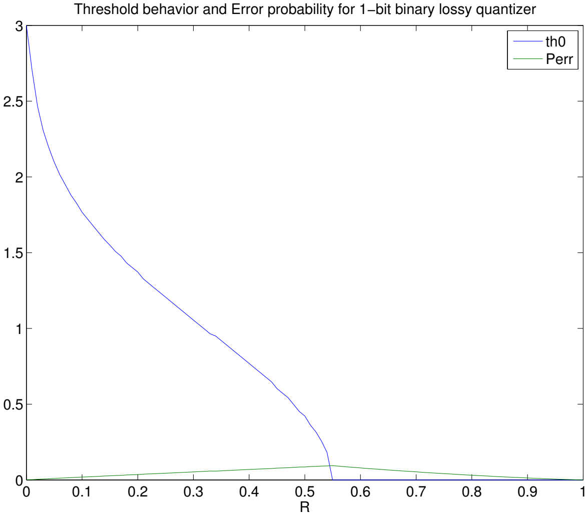
Results: Comparison with Lossless case



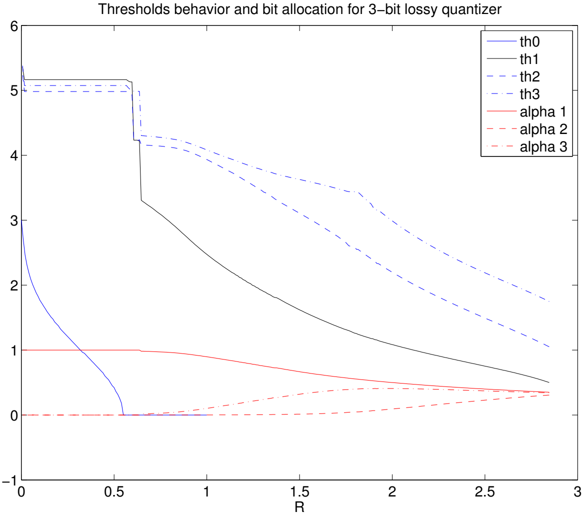
Results: Importance of Labels



Results: Optimal Threshold and Binary Error Probability



Results: Optimal Thresholds and Alphas



Conclusion

Outlook

- New approach: Scalar quantization + lossy binary compressor
  - Allowing binary distortion improves overall performance
  - Outperforms squema with lossless compressor
- Asymptotic behavior empirically
- Labeling matters, thresholds come from infinity

Further research topics

- Can we prove an upper bound for the whole D range?
- Extend approach for variable length quantizers
- Different input sources and distortion measures

Bibliography

- Thomas M. Cover and Joy A. Thomas, Elements of Information Theory, Chapter 13, 1991
- Toby Berger, Jerry D. Gibson, Lossy Source Coding, 1998
- Daniel Marco, David L. Neuhoff, Low-Resolution Scalar Quantization for Gaussian Sources and Squared Error, 2006
- Jacob Ziv, On Universal Quantization, 1985
- Herbert Gish and John N. Pierce, Asymptotically Efficient Quantizing, 1968
- Robert M. Gray, Vector Quantization, 1984

Thank you!

Looking forward to your questions...