ENEE739M Spring 2002
Lecture-9

Rate-Distortion Methods for MM Compression

Min Wu

Electrical & Computer Engineering

Univ. of Maryland, College Park

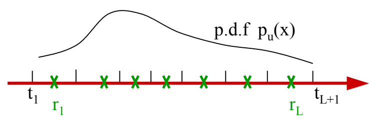
☞ *http://www.ece.umd.edu/class/enee739m/*

☞ *minwu@eng.umd.edu*

Review of Last Class

- Conclude 1st unit (data hiding)
  - Discussion on embedding capacity
    - *advanced Costa embedding ~ scaling-and-compensated enforcement*
- Start 2nd unit (MM coding and comm.)
  → Intro. to Rate-Distortion theory for Multimedia
  - Review of Scalar Quantization
- Today
  - Rate-Distortion func. for single r.v.
  - Optimal bit allocation for indep. r.v. w/ different variance
    ~ "reverse water-filling"
  - More on R-D methods for multimedia coding/compression

Last Time (1): Bring Prob. Distr. into Quantiz.



- Allocate more reconstruct. values in more probable ranges
- Minimize error in a probability sense
  - MMSE (minimum mean square error) $\mathcal{E} = E[(u-u')^2] = \int_{t_1}^{t_{L+1}} (x-u'(x))^2 p_u(x)dx = \sum_{i=1}^{L} \int_{t_i}^{t_{i+1}} (x-r_i)^2 p_u(x)dx$
    - *assign high penalty to large error*
    - *squared error gives convenience in math.: differential, etc.*
- An optimization problem
  - what $\{t_k\}$ and $\{r_k\}$ to use?
  - Last time: necessary conditions by setting partial differential to zero

Last Time (2): MMSE Quantizer (Lloyd-Max)

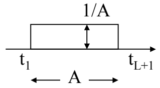
- Reconstruction and decision levels need to satisfy

$$\begin{cases} t_k = \frac{r_k + r_{k-1}}{2} & \text{(midpoint between reconstruction levels)} \\ r_k = \frac{\int_{t_k}^{t_{k+1}} x p_u(x)dx}{\int_{t_k}^{t_{k+1}} p_u(x)dx} = E(u \mid u \in [t_k, t_{k+1})) & \text{(conditional expectation within decision interval)} \end{cases}$$

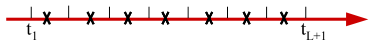
- Solve iteratively
  - Choose initial values of $\{t_k\}^{(0)}$, compute $\{r_k\}^{(0)}$
  - Compute new values $\{t_k\}^{(1)}$ and $\{r_k\}^{(1)}$ ……
  - Ultimately converge to locally-optimal solution
- For large number of quantization levels
  - Approx. constant p.d.f. within $[t_k, t_{k+1})$
  - Can obtain approximated closed-form solution
    Ref. A.K.Jain's image proc. book

Quantizer for Uniform Distribution

p.d.f. of uniform distribution



- Uniform quantizer

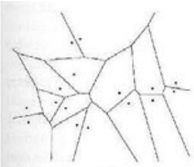


  - Optimal for uniform distributed r.v. in MMSE sense
  - MSE = $q^2 / 12$ with $q = A / L$
- SNR
  - Definition: SNR = 10 $\log_{10}$ sig_power / MSE
  - For uniformly distributed r.v.
    - *Variance = $A^2 / 12$*
    - *SNR = 20 $\log_{10} L$ = (20 $\log_{10} 2$)*B = 6B (dB) with $L = 2^B$*

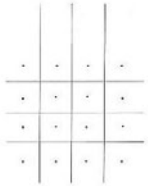
    ➔ *"1 bit is worth 6 dB."*

Vector Quantization

- Encode a set of values together
  - Find the representative combinations
  - Encode the indices of combinations
- Scalar vs. Vector quantization
  - VQ allows flexible partition of coding cells
  - VQ could naturally explore the correlation between elements
  - SQ is simpler in implementation
- Elements in VQ design
  - Codebook design
  - Encoder
  - Decoder



vector quantization of 2 elements

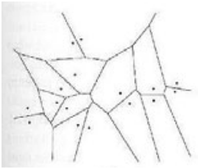


scalar quantization of 2 elements

From Bovik's Handbook Sec.5.3

Outline of Core Parts in VQ

- Design codebook
  - Optimization formulation is similar to MSE scalar quantizer
  - Given a set of representative points
    - *"Nearest neighbor" rule to determine partition boundaries*
  - Given a set of partition boundaries
    - *"Probability centroid" rule to determine representative points that minimizes mean distortion in each cell*
- Search for codeword at encoder
  - Tedious exhaustive search
  - Design codebook with special structures to speed up encoding
    - *E.g., tree-structured VQ*



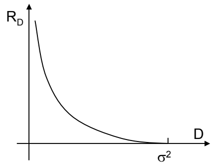
vector quantization of 2 elements

Quantization – A "Lossy Step" in Source Coding

- Quantizer achieves compression in a lossy way
  - Lloyd-Max quantizer minimizes MSE distortion with a given rate
- Need at least how many # bits for certain amount of error?
  - (information-theoretic) Rate-Distortion theory
- Rate distortion func. of a r.v.
  - Minimum average rate R_D bits/sample required to represent this r.v. while allowing a fixed distortion D
  - R(D) = min I(X;X*)
    - *minimize over p(X*|X) given a source p(X)*
  - For gaussian r.v. and MSE

$$R_D = \begin{cases} \frac{1}{2}\log_2(\sigma^2 / D), & D \le \sigma^2 \\ 0, & D > \sigma^2 \text{ (just use the mean)} \end{cases}$$

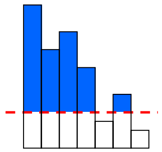
    - *1bit more cuts down distortion to ¼ => 6dB*



See Info. Theory course/books for detailed proof of R-D theorem

Bit Allocation Among Indep. r.v. with Different Var.

- How many bits to be allocated for each coeff.?
  - Determined by the variance of coeff.
  - More bits for high variance σ_k^2 to keep total MSE small



- Optimal solution for Gaussian: "Reverse Water-filling"
  - Idea: Try to keep same amount of error in each freq. band and no need to spend bit resource to represent coeff. w/ small variance than the "surface"
  - Results come from R-D func. and optimization via Lagrange multiplier approach

Details on Reverse Water-filling Solution

- Problem formulation
  - To encode a set of indep. Gaussian r.v. {X1, …, Xn}, $X_k \sim N(0, \sigma_k^2)$
  - Allocate R_k bits to represent each r.v. X_k, incurring distortion D_k
  - Total bit cost is R=R1+…+Rn
  - ➔ What is the best bit allocation {R1, …, Rn} such that R is minimized and the total distortion D1+…+Dn is kept below a bound D?
- Recall $R_k = \max(\frac{1}{2} * \log(\sigma_k^2 / D_k), 0)$
- Solving the optimization problem using Lagrange multiplier $\lambda \geq 0$

Details on Reverse Water-filling Solution (cont'd)

$$\min \sum_{i=1}^{n} \tfrac{1}{2} \ln \frac{\sigma_i^2}{D_i}, \text{ subj. to } \sum D_i = D$$

Construct func. using Lagrange multiplier

$$J(D) = \sum_{i=1}^{n} \tfrac{1}{2} \ln \frac{\sigma_i^2}{D_i} + \lambda \sum_{i=1}^{n} D_i$$

Necessary condition ➔ Keep the same marginal gain

$$\frac{\partial J}{\partial D_i} = -\frac{1}{2}\frac{1}{D_i} + \lambda = 0 \text{ for all } i \Rightarrow D_i = \lambda \text{ and } \sum D_i = \text{D}$$

$$\min \sum_{i=1}^{n} \max\left\{\frac{1}{2} \ln \frac{\sigma_i^2}{D_i}, 0\right\}, \text{ subj. to } \sum D_i = D$$

Necessary condition for choosing λ

$$\frac{\partial J}{\partial D_i} \begin{cases} = 0, \text{ if } D_i < \sigma_i^2 \\ \leq 0, \text{ if } D_i \geq \sigma_i^2 \end{cases} \Rightarrow D_i = \begin{cases} \lambda, \text{ if } \lambda < \sigma_i^2 \\ \sigma_i^2, \text{ if } \lambda \geq \sigma_i^2 \end{cases} \text{ s.t.} \sum D_i = D$$

Lagrangian Opt. for Indep. Budget Constraint

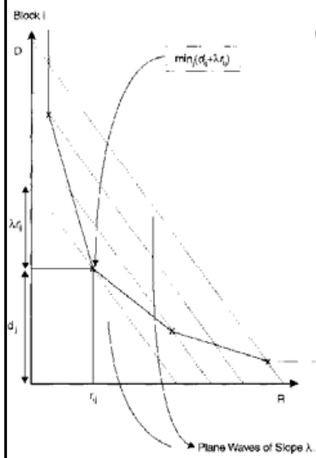
- Previous: fix distortion, minimize total rate
- Alternative: fix total rate (bit budget), minimize distortion
- (Discrete) Lagrangian optimization on general source

$$\min \left(\sum_{i=1}^{n} d_{i,q(i)} + \lambda r_{i,q(i)} \right)$$

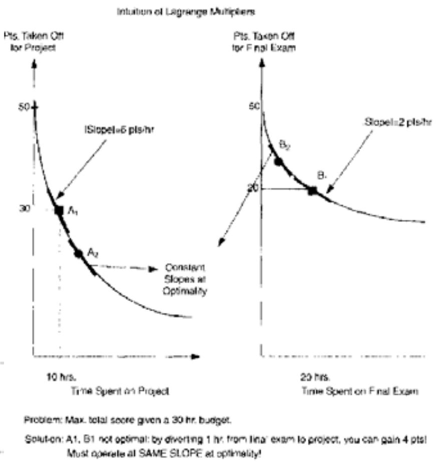
  - "Constant slope optimization" (e.g. in Box 6 Fig. 17 of Ortega's tutorial)
  - Need to determine the quantizer q(i) for each coding unit *i*
  - Lagrangian cost for each coding unit $d_{i,q(i)} + \lambda r_{i,q(i)}$
    - *use a line with slope -λ to intersect with each operating point* *(Fig.14)*
  - For a given operating quality λ, the minimum can be computed independently for each coding unit

$$\min \left(\sum_{i=1}^{n} d_{i,q(i)} + \lambda r_{i,q(i)} \right) = \sum_{i=1}^{n} \min \left(d_{i,q(i)} + \lambda r_{i,q(i)} \right)$$

=> Find operating quality λ satisfying the rate constraint



(from Fig. 14 of Ortega's tutorial)



(from Box 6 Fig. 17 of Ortega's tutorial)

Bridging the Theory and Ad-hoc Practice

- Operational R-D curves
  - Directly achievable with practical implementation
- Tradeoff! Tradeoff! Tradeoff! ☺
  - Rate vs. Distortion
    - *how close are the operational points to info. theory bounds?*
  - Also tradeoff among practical considerations
    - *Cost of memory, computation, delay, etc.*

  ➔ Often narrowing down to an optimization problem:
  ~ *optimize an objective func. subj. to a set of constraints*
- Model mismatch problem
  - How good are our assumptions on source modeling?
  - How well does a coding algorithm tolerate model mismatch?

Basic Steps in R-D Optimization

- Determine source's statistics
  - Simplified models: Gaussian, Laplacian, Generalized Gaussian
  - Obtain from training samples
- Obtain operational R-D points/functions/curves
  - E.g., compute distortion for each candidate quantizer
- Determine objective func. and constraints
- Search for optimal operational R-D points
  - Lagrange multiplier approach
    - *select only operating points on the convex hull of overall R-D characteristics* (*Fig.19 of Ortega's tutorial*)
    - *may handle different coding unit independently*
  - Dynamic programming approach
    - *not constrained by convex hull but has larger search space (higher complexity)*

R-D Examples in MM Compression Standard

- Syntax-constrained optimization of MM coding
  - Decoder is standardized ~ fixed syntax and transform
  - Encoder has a lot of freedom
    - *E.g., to determine quantization tables and entropy coding choices in JPEG*
- Optimizing quantizer: A budget constraint problem

Summary

- Rate-Distortion func. for single r.v.
- Optimal bit allocation for indep. r.v. w/ different variance
  - "reverse water-filling"
- Basic steps of R-D methods for multimedia coding
  - Source modeling problem
  - Optimization problem
- Next Time
  - More on R-D methods in MM coding standard

Suggested Reading

[See the reading list in course web page]

- Rate-Distortion
  - Cover-Thomas' info. theory book Chapt.13
  - R-D tutorial issue in IEEE Sig. Proc. Magazine 11/98
    - *Overview & framework by Ortega-Ramchandran*
    - *Applications in video compression by Sullivan-Wiegand*

Questions for Today

- Two possible ways to convey side info.
  - Attached separately in header or user-defined data segment
  - Hiding into the host data (esp. for multimedia)
- What are the pros and cons of data hiding vs. alternative?
- Is data hiding providing extra room to convey more info.?
  - E.g., Can we use data hiding to achieve better MM compression?

=> Hint: any useful insight from rate-distortion concept?

! Summarize your thinking into a one-page write-up and bring to Thur. class