Polytechnic University

http://wireless.poly.edu

Rate, Distortion and Quality: A Tractable Abstraction of the Human Sensory System

Virgilio Rodriguez

ECE Department
Polytechnic University
Brooklyn, NY
email: vr@ieee.org

Outline of research program

- Analytical Core
- Single-user applications: choose power, data rate and/or coding rate for data, image, video transmission
- Decentralized multi-user applications:
  - Game formulation
  - Mechanism design
- Centralized data throughput maximization
  - Without noise
  - With noise and media terminals present (yesterday)

Overview of our analytical framework

- Many radio-resource optimizations share a common analytical core, which enables robust and tractable analysis and provides clear answers in fairly general scenarios
- It involves
  - A tractable abstraction of the physical layer (yesterday)
  - A tractable abstraction of the human visual system
  - A fundamental result: maximize $f(x)/x$ with f an "S-curve".
- Problems to which this framework applies:
  - Power and coding rate choice for media files (images, video)
  - Choosing the "right amount" of media distortion
  - Decentralized power control for 3G CDMA
  - Data rate and power allocation for maximal cell throughput when data and media terminals share a CDMA cell (yesterday)

Outline of the remainder of this presentation

- Motivation (sample problems)
  - "Right amount" of distortion when fidelity is expensive
  - Energy optimization for image transmission
  - Coding rate and power for video streaming
- Quality-Distortion Theory
- Rate-distortion theory
- Rate-quality theory
- Discussion/extensions

The “right amount” of distortion

- Equally valuable media files are offered
- Each file can be acquired at any desired distortion, $y \in [0, \bar{D}]$
- Cost of file ($$, Joules, time, etc) is $c(y)$, which is *decreasing* in y
- Consumer has fixed budget, B ($$, Joules, etc)
- trade-off: more media quality $\longrightarrow$fewer files acquired
- In order to solve this problem we need a model of how the human sensory system evaluate an “imperfect” copy of a signal

Energy policy for media files

- Terminal faces:
  - many equally important media files to transfer (assume simple b/w images, 1 bit per pixel)
  - limited energy, E
  - fixed transmission bit rate R bps
  - N packets per file, L info bits, M total bits
  - I watts of interference (noise)
  - NO retransmissions. Code can correct up to m errors.
- Uncorrected bit errors treated as non-errors $\Rightarrow$ distortion!
- More transmission power $\rightarrow$ smaller BER $\rightarrow$ less distortion
  BUT more transmission power $\rightarrow$ fewer transferred files
- To solve this problem we need a model of how the human visual system evaluate an "imperfect" copy of an image

Coding rate & power for video streaming

- Each T-secs of video yields "scalable" file
  (i.e., file can be truncated and decoded; e.g., MPEG-4, SPIHT-3D)
- Energy E is limited!
- File for given segment must be transferred in a deadline of Δ secs.
- Files will be split into small packets for transmission purposes; ECC bits will be added and an ARQ system will be available
- Transferring each file complete $\Rightarrow$ maximal quality per segment BUT short total viewing time with available energy. Transferring few bits per file $\Rightarrow$ long running time BUT low quality per segment.
- Problem: how many bits per file to transfer (where to truncate) AND at which power to transmit?
- We need a model of how the "perceptual quality" of a decoded video segment varies with the number of bits in the truncated file (coding rate)

Quality-distortion theory

- The perceptual quality of an "imperfect" signal is determined by the human sensory system
- Distortion is a very simple measure of the difference between a signal and its copy (e.g., the original vs. the reproduced video segment)
- It is reasonable to assume that the perceptual quality of an "imperfect" copy of a signal is determined by distortion; i.e., that a function $Q(D)$ that translates distortion into perceptual quality can be found.
- $Q(D)$ must be decreasing,..., but with which "shape" (convex, concave, linear, step, etc.) ?
- $Q(D)$ can be obtained by psychophysical experiments for a specific user.

Plausible Q-D relations

- Some plausible simple relations are:
  - (i) fidelity equals quality (red dashed line)
  - (ii) hard threshold (step) (often assumed by resource-management literature)
  - (iii) ramp (blue broken line). The ramp includes as special case the threshold ($a = b = c$) and the linear relation ($a = 0$, $b = D_{\mathrm{MAX}}$).
- By assuming that all that is known about $Q(D)$ is that it is a "reversed" S-curve, the “hard threshold” and many plausible curves (“almost” convex, “almost” concave, “almost” linear, etc.) are contained as special cases.

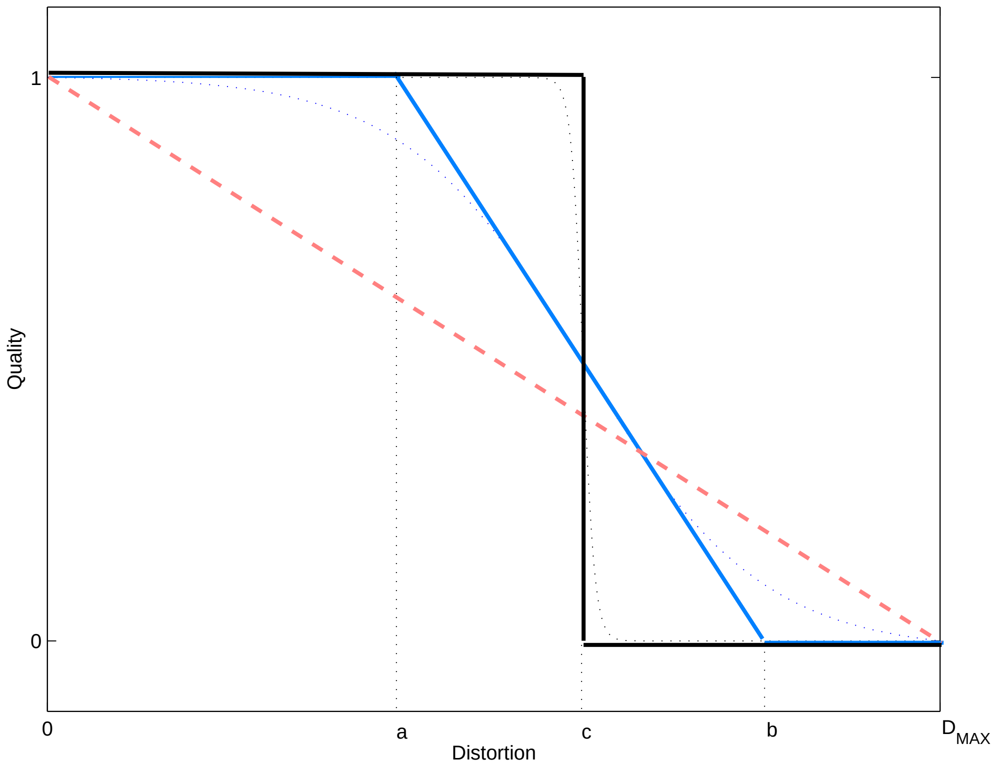
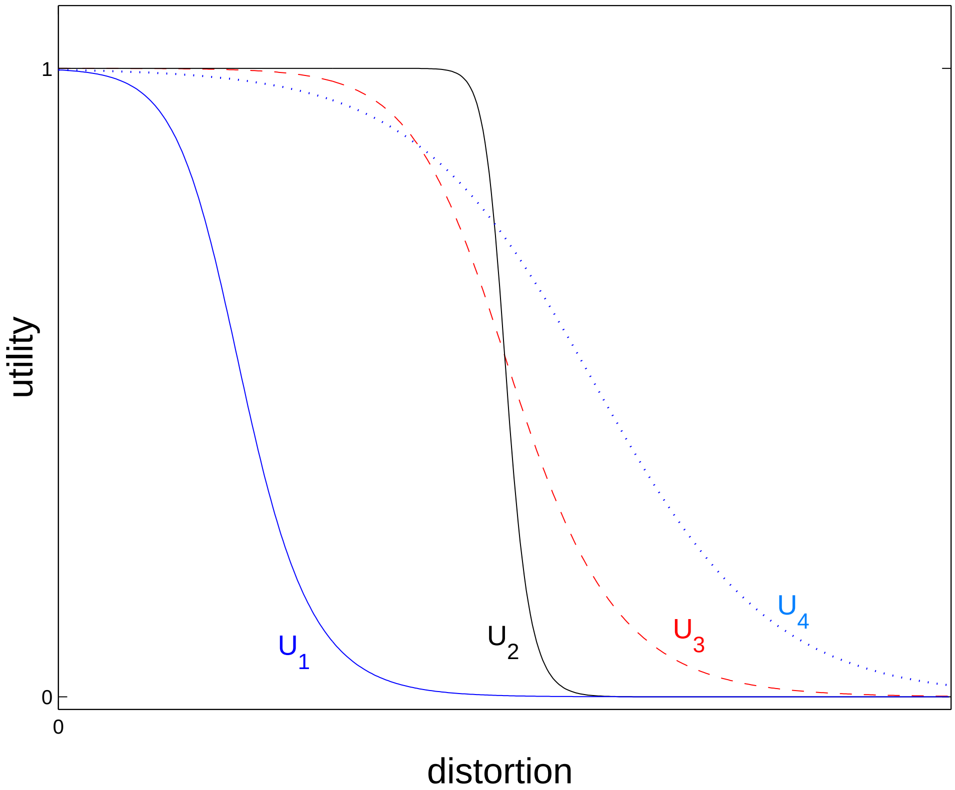


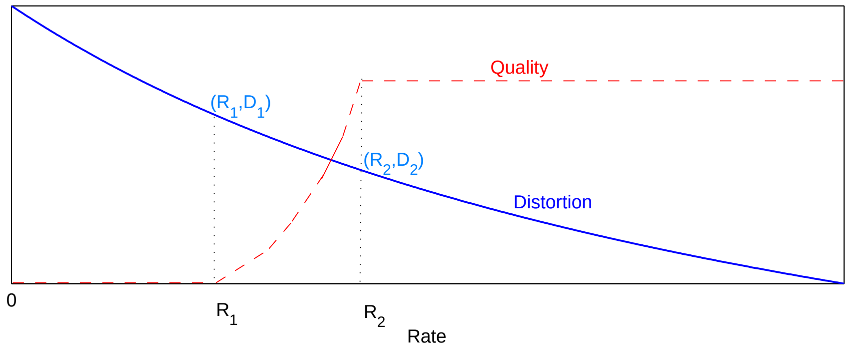
Figure 1: Quality vs. distortion: Some plausible simple relation

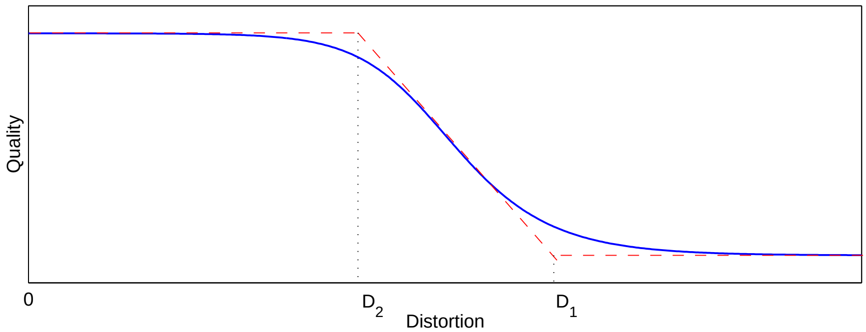


Beauty is in the eyes of the beholder

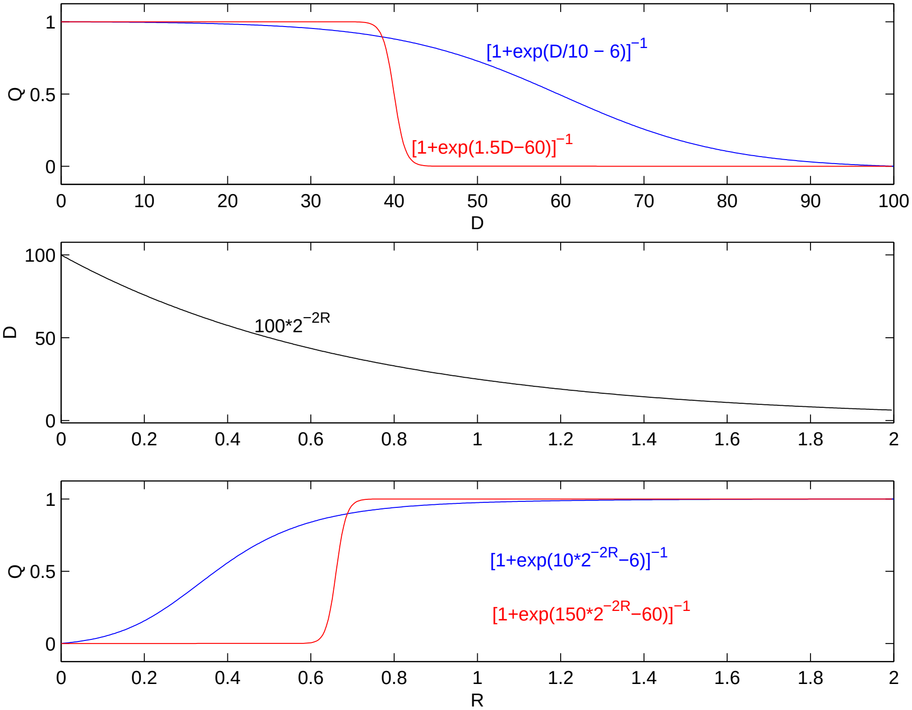
From rate-distortion to rate-quality theory

- We want a quality/rate function. But there is a well developed theory relating rate to distortion.
- It is generally accepted that the function $D(R)$ giving distortion as a function of the coding rate is decreasing and convex.
- For the memoryless Gaussian source, $D(R) \propto 2^{-2R}$
- Given a quality-distortion function $Q(D)$ and a distortion-rate function $D(R)$ the desired quality-rate function is $Q(D(R))$
- Question: if all that is known about $D(R)$ is that it is decreasing and convex, and all that is known about $Q(D)$ is that it is a "reversed" S-curve, what can be said about $Q(D(R))$?





For $Q(D)$ a reversed S-curve, we expect $Q(D(R))$ to be an S-curve



For a memoryless Gaussian source, $D(R) \propto 2^{-2R}$. For the $Q(D)$curves at the top, $Q(D(R))$ are S-curves (bottom).

Discussion

- In order to formulate and solve certain interesting problems, we need to model how the human sensory system assigns "perceptual quality" to an "imperfect" copy of a signal.
- Two specific sample problems involving "quality vs. quantity" trade-offs (1, 2) have been cited (solutions available elsewhere).
- Distortion is a relatively simple, well understood, mean square measure of the difference between a signal and its copy. It seems reasonable to assume that a function $Q(D)$ that translates distortion into perceptual quality can be found.
- If we assume that all we know about $Q(D)$ is that it is a "reversed" S-curve, we include as special cases many plausible Q-D relations (step, "ramps", "mostly" convex, "mostly" concave, etc) .

- This level of generality is important, because the “true” Q-D curve can only be obtained by psychophysical experimentation, and should depend on the chosen subject- application pair.
- To solve another sample problem (coding rate optimization for video streaming), a function giving perceptual quality in terms of the coding rate is needed.
- A well-developed rate-distortion theory holds that the function $D(R)$ relating distortion and coding rate is generally decreasing and convex. The composite function $Q(D(R))$ relates perceptual quality to coding rate.
- Our analysis indicates that when $Q(D)$ is a “reversed” S-curve, and $D(R)$ is decreasing and convex, $Q(D(R))$ is a “regular” S-curve.
- Because of its generality, this model and any analysis derived from it should hold for many user/application/data source combinations.

THANK YOU!!!

WIRELESS.POLY.EDU