The Perceptual Quantizer
Design Considerations and Applications

TIMO KUNKEL, PHD
DOLBY LABORATORIES, INC.
SAN FRANCISCO, CA
Introduction

• The Perceptual Quantizer or ‘PQ’ is an efficient signal non-linearity for HDR applications
• It is modeled after elements of human vision to efficiently and effectively encode a large absolute luminance range for content consumption
• PQ has developed into a foundation of HDR imaging

Overview

• Discuss the design considerations when developing PQ
• Outline the large field of applications
Example Application

- **Peak Brightness** @ 6700 cd/m²
- **Black** @ 0.0 cd/m²
- **Dark Textures** @ 0.5 cd/m²
- **Color Saturation** without clipping color channels

© 2022 DOLBY

Photo: Timo Künkel, Dolby  
Note: This image is a tone-mapped example

ICC HDR EXPERTS DAY - MARCH 15, 2022
Development of PQ
About 15 years ago, HDR displays used custom drivers and mapping algorithms had to be developed & applied which was not very efficient from a compatibility and scalability point of view.

Fully log or gamma functions encounter problems with contouring in certain regions when considering HDR luminance ranges.

- **Efficient but not Effective** for our intended Application

Linear light formats existed offering extreme luminance ranges (e.g., OpenEXR and Radiance HDR)

- Good for post processing and CG applications (e.g., Raytracing, SFX & Image Based Lighting - IBL)

- **Effective but not Efficient** for our intended Application

**Design considerations:**

- Don’t design for today’s display capabilities (limited performance)
- Nor for ranges offered by float linear luminance (limited efficiency).
- Avoid contouring artifacts (limited effectivity).
- **Base it on properties of human vision!**
Adaptive Processes

Apparently Simple Option:

Reduce Luminance Range to capability of Human Visual System

Note: This is for illustration only! The adaptive behavior of the HVS is more complex than shown here.
Content Luminance Ranges

• Content (Frame) Range does not have to be high!
  – Steady State typically sufficient
• But average Luminance might be **high (left) or low (mid)**
• Adaptation* adjusts sensitivity of HVS to **highest sensitivity at Adapting Luminance** $L_A$
• Option for HDR format: Encode and Transport the Steady State Range & Adapting Luminance?

• **Challenge:** Content contrast can be high (right)
• Now, adaptation is viewpoint and viewing environment dependent
• **Imaging encoding must facilitate all cases!**

* This is a simplification. Adaptive processes are more complex.
Viewing Conditions and Preferences

DEVELOPMENT OF PQ

Single Viewer

Screen

Light Content

Dark Content

Viewer’s Interest A

Viewer’s Interest B

Viewer

Multiple Viewers

Cinema Screen

Light Content

Dark Content

Viewer A

Viewer B

Audience
Summary (so far)

- Fully log or gamma functions only approximate perception in certain regions when considering HDR luminance ranges.
- They can work, but then, the bit depth requirements are too high.
- A smaller, floating range (steady state) with adaptation metadata (such as $L_A$) follows perceptual properties very well but is challenging to predict and robustly implement for the intended application.

Are there other options that:
- follow perceptual principles
- don’t require knowledge of $L_A$?
- Are both effective AND efficient?

Potential Approach:
- ‘Worst Case Engineering’ paradigm
- Intervals at Highest HVS Sensitivity
- Model discrimination thresholds through a useful luminance range
Worst Case Engineering

- **Fundamental question:** How to determine if a signal non-linearity retains quantization steps that are all equal or below the threshold of one JND over an absolute luminance range.
  - Approach in terms of sensitivity (the inverse of threshold).
  - Particularly: How does the sensitivity of the visual system vary with luminance level.
- One approach to answer this question is to follow **worst-case engineering design principles**.
- **Fundamental Concept:**
  - Worst-case engineering considers the most severe possible behavior of a system that can reasonably be projected to occur in a given situation.
  - In the context of imaging, the worst case can be defined by either contouring steps to become visible or by unnecessarily wasting bandwidth for a given luminance level.
Worst Case Engineering

Max. Lum.

- $L_{A,1} \ldots n$ spaced (quantized) in Just Noticeable Difference (JND) steps over intended luminance range
- JNDs based on Highest Sensitivity!

Note: Not all tone values simultaneously visible to the HVS

10$^{-6}$ cd/m$^2$

Steady State Lightness Scales

- Highest Sensitivity of $L_{A,1}$
- Highest Sensitivity of $L_{A,2}$
- Highest Sensitivity of $L_{A,3}$

HDR Signal Representation

- $L_{A,1}$
- $L_{A,2}$
- $L_{A,3}$

© 2022 DOLBY

Ramps & Curves are for illustration only!
DICOM

Gray Scale Display Function

- Created for medical imaging, adopted by VESA
- Covers 0.05 to 4000 cd/m$^2$ with 1023 steps
- Could be extended to a broader range by adding range above and below limits

Challenges

- Could see visible differences between adjacent bars on test HDR display
- Especially at darker levels
- Why?

References:

The Barten Model

- DICOM is based on the Barten Model
- The Barten model is not a fixed curve, but a model with many parameter values
- DICOM’s choice of parameters are not ideal for our use case

Reference:
Optimize Barten Model Parameters

- Highest Sensitivity at any light adaptation level, regardless of frequency
- Must work with zero noise imagery
- Angular size of 40° instead of DICOM’s 2°
- Recomputing the photon conversion factor for a D65 white point
- Tracking the peaks of contrast sensitivity instead of DICOM’s fixed 4 cycles/degree

Using Modulation to Set JND Steps

- Once we know the modulation threshold \( m_t \) at a given luminance level, we can calculate the next JND step up or down from that level

\[
m_t = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}} \quad \text{so:} \quad L_{j+1} = L_j \frac{1 + m_t}{1 - m_t} \quad \text{and} \quad L_{j-1} = L_j \frac{1 - m_t}{1 + m_t}
\]

- We can also use fractions of a JND step to cover a desired range with a desired bit depth

References:
Finding a Functional Representation

The perceptual model is a table built by iteration of fractional JNDs
- Awkward for standardization & specification

Identify functional form:
- Continuous function
- With no transitions required
- Invertible

Absolute Boundaries:

\( \text{PQ}_{\text{Max}} \)
- Accepted for PQ max: 10,000 cd/m\(^2\)
  - To fit to 12 Bits as usable by SDI
  - Also preferred by Hollywood studios & MovieLabs

\( \text{PQ}_{\text{Min}} \)
- Lowest Level of Visibility at 10\(^{-6}\) cd/m\(^2\)
- To simplify math, extended down to 0.0 cd/m\(^2\) (added negligible 'cost' to the accuracy)
Quantization Performance

![Graph showing quantization performance with legend]

**Legend**
- **Artifacts Very Likely Visible**
- **Artifacts Visible Under Some Circumstances**
- **Barten Threshold**
- **No Artifacts Visible - Efficient for Encoding**
- **Inefficient for Encoding - Good for Post-Processing**

Source: ICDM IDMS v1.1
Verification

Psychophysical Experiment: Detectability

- Test detectability thresholds using JND cross pattern & real images

Psychophysical Experiment: Preference Luminance Range

- In parallel to the mathematical modelling, a psychophysical experiment was carried out
- Objective: Identify the preference HDR luminance range
- Result: Reflected similar ranges for entertainment content

Vetting by Studios & Standardization Efforts

SMPTE standardization efforts involved extremely rigorous vetting of the technology using
- Professional, stable equipment, with careful calibration,
- Critical viewing, with all sorts of content and corner cases

References:

© 2022 DOLBY
Final Form of the Perceptual Quantizer

In the end, 10 years ago, in January 2012, PQ took on its finished form with a range of 0 to 10,000 cd/m² and the following formula:

\[
L = 10,000 \left( \max \left[ \left( \frac{V^{1/m_2} - c_1}{c_2 - c_3 V^{1/m_2}} \right), 0 \right] \right)^{1/m_1}
\]

\[
m_1 = \frac{2610}{4096} \times \frac{1}{4} = 0.1593017578125
\]
\[
m_2 = \frac{2523}{4096} \times 128 = 78.84375
\]
\[
c_1 = \frac{3424}{4096} = 0.8359375
\]
\[
c_2 = \frac{2413}{4096} \times 32 = 18.8515625
\]
\[
c_3 = \frac{2392}{4096} \times 32 = 18.6875
\]

Today, PQ is standardized in SMPTE ST.2084 and ITU-R Rec. BT.2100
Applications
Luminance Capabilities of Today’s HDR Imaging Pipeline
Capture, Postproduction & Deployment

• **Tools for capture, process and distribute HDR in PQ are deployed at scale**
• Both **professional and consumer level options** are available to create, process and distribute HDR PQ content
  • Most major video editing tools support PQ
  • Availability of higher end HDR-capable content editing and grading displays with high luminance capability of ~700 to 4000 cd/m\(^2\)
  • Support on computers, mobile devices, or tablets for UGC content
Display

- HDR capable displays are widely available in the market at different price points, and all support the PQ non-linearity
- Luminance capabilities
  - Deep blacks (<0.1 cd/m²) down to no light emission
  - Maximum luminance levels reaching 700-1000 cd/m² (some offer up to 2000 cd/m²).
  - Color gamut has extended from ITU-R rec. BT.709 to P3 or towards ITU-R Rec. BT.2020
- HDR TVs & mobile phones have been around for several years,
- Computer & tablet displays are also catching up.

Key Standards for Color Gamut:
APPLICATIONS

Availability

Displays
• Of 225 million TVs sold globally in 2020,
• 58% included HDR functionality
• 10% even provided 500 cd/m² peak luminance or higher (Source: OMDIA)
• Many also support dynamic metadata-enabled HDR formats.

HDR supported on PCs, mobile phones, tablets, as well as gaming platforms

HDR content is readily available
• Cinema, Blu-ray, broadcast, gaming & OTT
Summary
Summary

1. Designing PQ
   Complex considerations went into designing PQ. This included modelling as well as several psychophysical studies.

2. PQ is established!
   After its standardization in SMPTE ST.2084 and later ITU-R Rec. BT.2100, PQ has found wide-spread adoption with many HDR applications.

3. ‘Backbone’ of Consumer HDR
   PQ, together with HLG, form the backbone of today’s consumer HDR.

Continue enjoying the quality and fidelity improvements, HDR continues to provide!
THANK YOU!

Contact: timo.kunkel@dolby.com

Acknowledgement: Scott Daly & Scott Miller
REFERENCES
Standards Related to HDR

• ARIB STD-B67: Essential Parameter Values for the Extended Image Dynamic Range Television (EIDRTV) System for Programme Production
• CTA 861.3-A-2016: HDR Static Metadata Extensions
• ETSI TS 103 433-1: High-Performance Single Layer High Dynamic Range (HDR) System for use in Consumer Electronics devices
• SID ICDM IDMS 1.1: SID International Committee for Display Metrology (ICDM) Information Display Measurements Standard v1.1
• ITU-R Rec. BT.709-6: Parameter values for the HDTV standards for production and international programme exchange. ITU-R. June 2015
• ITU-R Rec. BT.1886: Reference electro-optical transfer function for flat panel displays used in HDTV studio production. ITU-R. March 2011
• ITU-R Rec. BT.2035: A reference viewing environment for evaluation of HDTV program material or completed programmes
• ITU-R Rec. BT.2100-2: Image parameter values for high dynamic range television for use in production and international programme exchange. July 2018
• ITU-T Rec. H.273: Coding-independent code points for video signal type identification, 2016
• OpenEXR: High-dynamic-range scene-linear image data and associated metadata format. www.openexr.com
• SMPTE ST 0196-2003: Motion-Picture Film - Indoor Theater and Review Room Projection - Screen Luminance & Viewing Conditions
• SMPTE ST 0431-1-2006: D-Cinema Quality - Screen Luminance Level, Chromaticity and Uniformity
• SMPTE RP 0431-2-2007: D-Cinema Quality - Reference Projector and Environment
• SMPTE ST 2086:2018: Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images
Appendix
What is a JND (Just Noticeable Difference)

Source: ICDM IDMS v1.1
Crispening Effect

- The visual system is more sensitive to lightness variations around the background luminance.
- This is known as the Crispening Effect.

References:
Contouring

Insufficient Intervals: Contouring visible

Sufficiently Small Intervals: ‘Smooth’ Ramp

10k cd/m²

10⁻⁶ cd/m²

Given Luminance Range
The Barten Model

The Barten model parameters as used for PQ

\[
CSF = \frac{1}{m_t} = \frac{M_{opt}(u)/k}{\sqrt[2]{T \left( \frac{1}{X_0^2} + \frac{1}{N_{\max}^2} \right) \left( \frac{1}{\eta p E} + \frac{\Phi_0}{1 - e^{-(u_0/u)^2}} \right)}}
\]

\[
M_{opt}(u) = e^{-2\pi^2 \sigma^2 u^2}
\]

\[
\sigma = \sqrt{\sigma_0^2 + (C_{ab} d)^2}
\]

\[
d = 5 - 3 \tanh \left( 0.4 \log \left( L X_0^2 / 40^2 \right) \right)
\]

\[
E = \frac{\pi d^2}{4} \left( 1 - \frac{(d/9.7)^2 + (d/12.4)^4}{L} \right)
\]

\[
k = 3.0
\]

\[
\sigma_0 = 0.5 \text{ arcmin}
\]

\[
C_{ab} = 0.08 \text{ arcmin/mm}
\]

\[
T = 0.1 \text{ sec}
\]

\[
X_0 = 40^\circ
\]

\[
X_{\max} = 12^\circ
\]

\[
N_{\max} = 15 \text{ cycles}
\]

\[
\eta = 0.03
\]

\[
\Phi_0 = 3 \times 10^{-8} \text{ sec deg}^2
\]

\[
u_0 = 7 \text{ cycles/deg}
\]

\[
p = 1.25 \times 10^6 \text{ photons/sec/deg}^2/\text{Td}
\]