Digital Video Image Quality and Perceptual Coding

H. R. Wu and K. R. Rao

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Overview

I. Picture Coding and Human Visual System Fundamentals

II. Picture Quality Assessment and Metric

III. Perceptual Coding and Processing of Digital Pictures

© CRC Press, Wu & Rao Eds, 2006

I. Picture Coding and Human Visual System Fundamentals

1. Digital Picture Compression and Coding Structure

2. Fundamental of Human Vision and Vision Modelling

3. Coding Artifacts and Visual Distortions

© CRC Press, Wu & Rao Eds, 2006

II. Picture Quality Assessment and Metric

4. Video Quality Testing
5. Perceptual Video Quality Metrics – A Review
6. Philosophy of Picture Quality Scale
7. Structural Similarity Based Image Quality Assessment

© CRC Press, Wu & Rao Eds, 2006

II. Picture Quality Assessment and Metric

8. Vision Model Based Digital Video Impairment Metrics
9. Computational Models for Just-Noticeable Difference
10. No-Reference Quality Metric for Degraded and Enhanced Video
11. Video Quality Experts Group

© CRC Press, Wu & Rao Eds, 2006

III. Perceptual Coding and Processing of Digital Pictures

12. HVS Based Perceptual Video Encoders

13. Perceptual Image Coding

14. Foveated Image and Video Coding

© CRC Press, Wu & Rao Eds, 2006

III. Perceptual Coding and Processing of Digital Pictures

15. Artifact Reduction by Post-Processing in Image Compression
16. Reduction of Colour Bleeding in DCT Block-Coded Video
17. Error Resilience for Video Coding Services
18. Critical Issues and Challenges

© CRC Press, Wu & Rao Eds, 2006

I. Picture Coding and Human Visual System Fundamentals

1. Picture Coding and Human Visual System Fundamentals

J.J. Hwang[†], H.R. Wu[‡] and K.R. Rao[§]

†*Kusan National University, Korea*

‡ *Royal Melbourne Institute of Technology, Australia*

§ *University of Texas at Arlington, USA*

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 1 Outline

1.1 Introduction to Digital Picture Coding
1.2 Characteristics of Picture Data
1.3 Compression and Coding Techniques
1.4 Picture Quantization
1.5 Rate-Distortion Theory
1.6 Human Visual System
1.7 Digital Picture Coding Standards and Systems
1.8 Summary

© CRC Press, Wu & Rao Eds, 2006

1.1 Introduction to Digital Picture Coding

- Digital versus Analogue
  - Advantages
    - High Picture Quality
    - Portability/mobility
      - Storage and transmission over existing networks
  - Disadvantages
    - Large storage and transmission requirements
      - e.g. NTSC motion video 720 x 480 pixels @ 30 fps with 4:2:2 format requires 20.8 Mb/s bandwidth.
      - A 30 second video clip occupies 624 Mb (standard CD).

© CRC Press, Wu & Rao Eds, 2006

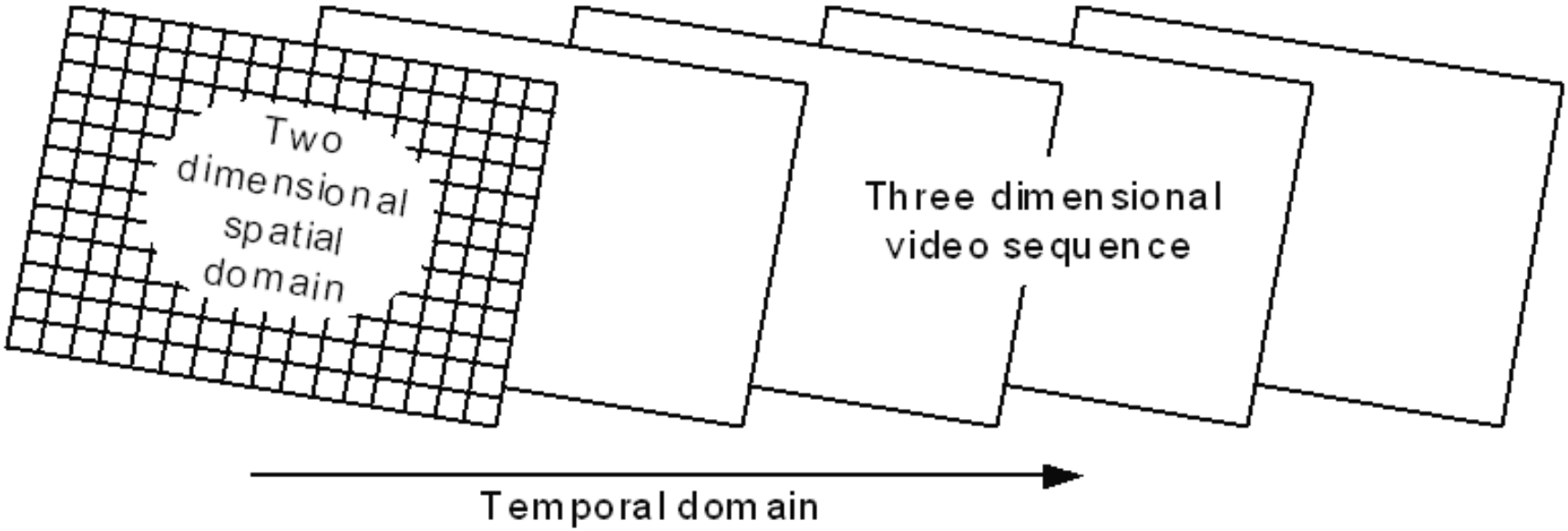
- Compression: a practical necessity for digital pictures
  - Reduces storage and transmission requirements
- How?
  - Compression reduces redundancies in digital pictures

© CRC Press, Wu & Rao Eds, 2006

- Statistical redundancies may appear
  - Spatially in adjacent pixels
  - Temporally in adjacent picture frames
  - Spectrally in different frequency and orientation bands
- Psychovisual Irrelevancy
  - Exists due to the limitations of the Human Visual System (HVS)
    - Optical (human eyes)
    - Cortical (the human brain)

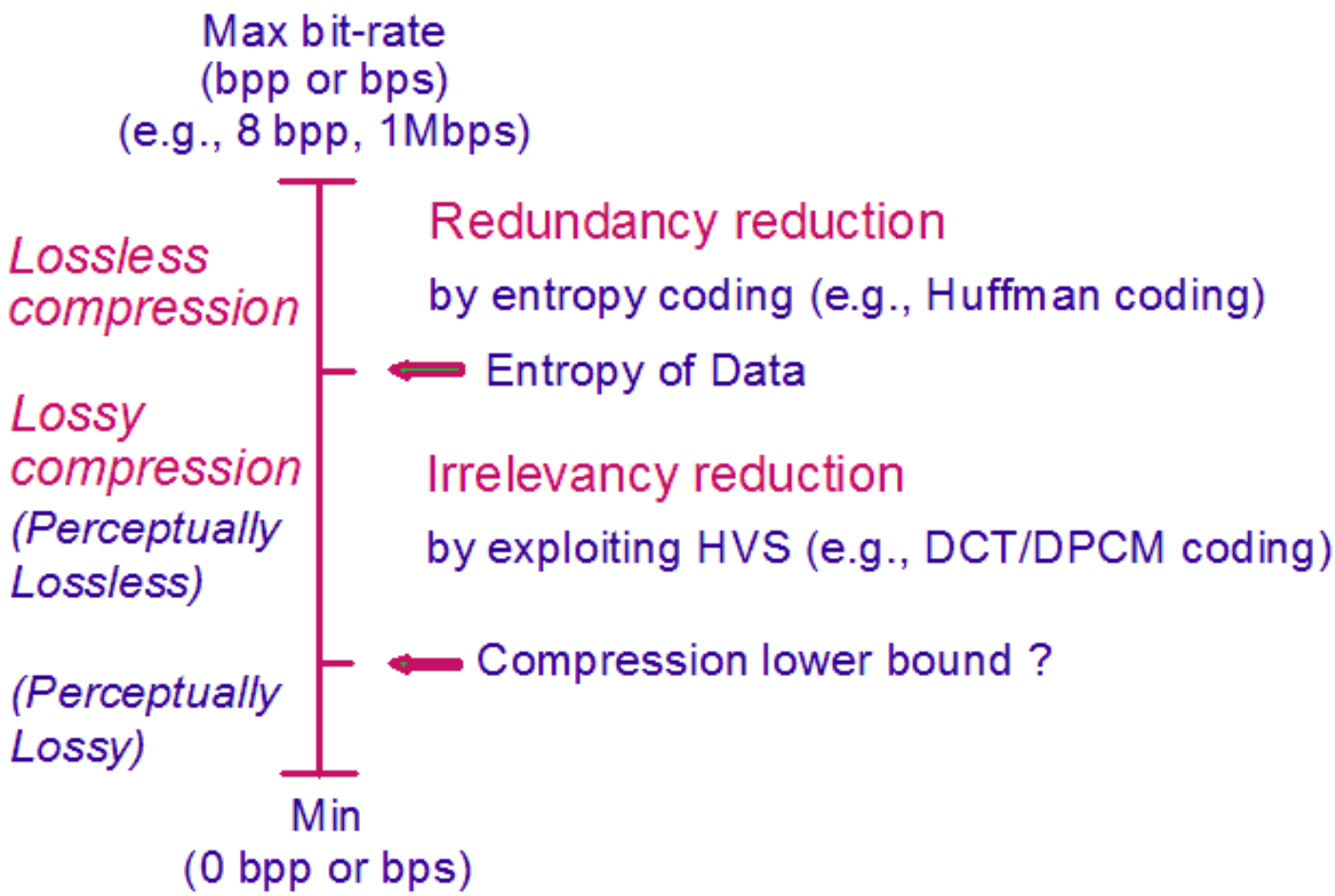
© CRC Press, Wu & Rao Eds, 2006

Spatial and Temporal Domains



© CRC Press, Wu & Rao Eds, 2006

Digital Picture Compression Fundamental Concepts



© CRC Press, Wu & Rao Eds, 2006

- Lossy and Lossless Compression
  - Lossless
    - ***No*** information loss, ***low*** compression
  - Lossy
    - ***Some*** information loss, ***high*** compression
- Quality/Distortion measures
  - Quantify coding quality/errors in lossy compression schemes
  - Objective and subjective measures

© CRC Press, Wu & Rao Eds, 2006

- Subjective distortion measure
  - Employs human observers, gauges perceived quality/distortions
  - Impractical for general purposes
- Objective distortion measure
  - Utilises functional metric to determine quality/distortions
  - Statistical metric. e.g. MSE, PSNR
  - HVS bases metric. e.g. PBIM, VQM

© CRC Press, Wu & Rao Eds, 2006

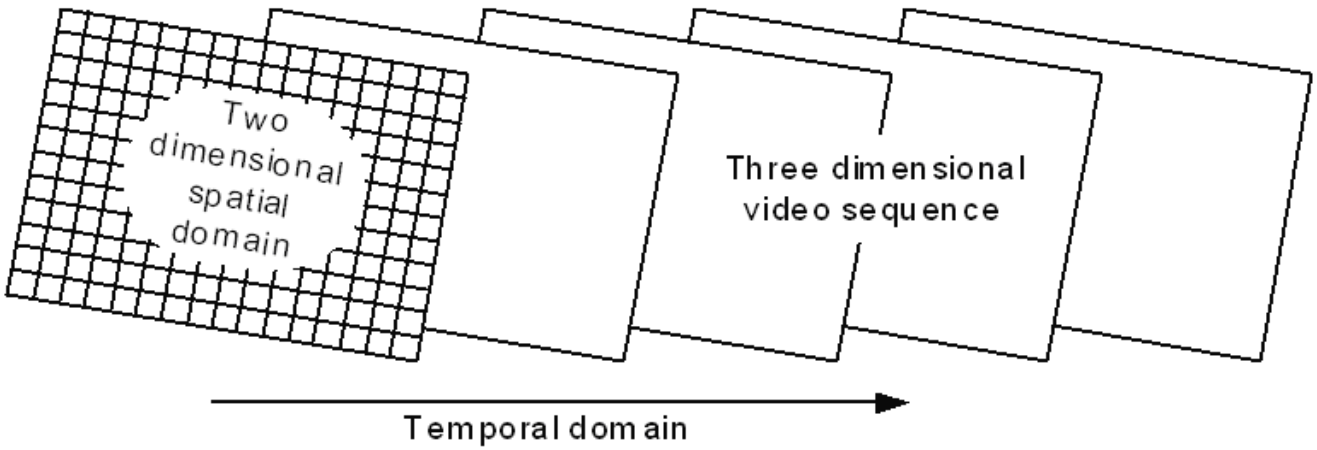
1.2 Characteristics of Picture Data

- Digital Image Data
  - Derived from sampling continuous image field data over an area of *M x N*
  - Individual digitized element is referred to as a pixel (or pel)
  - Each pixel has 2^q discrete levels per colour channel
  - q is typically 8 which corresponds to one byte (256 levels) per colour channel per sample

© CRC Press, Wu & Rao Eds, 2006

1.2.2 Digital Video Data

- Digitization of video
  - Sampled spatially as with digital image
  - Sampled at regular temporal interval to obtain multiple image frames



© CRC Press, Wu & Rao Eds, 2006

- Colour Spaces
  - *RGB* (Red, Blue and Green)
  - $Y_d C_b C_r$ (1 Luminance, 2 Chrominance)
- RGB Colour Space
  - For source capture and display
  - Corresponds to the Red, Blue and Green elements in display devices. e.g. CRT, LCD
  - Each channel has equal bandwidth – inefficient for transmission and storage

© CRC Press, Wu & Rao Eds, 2006

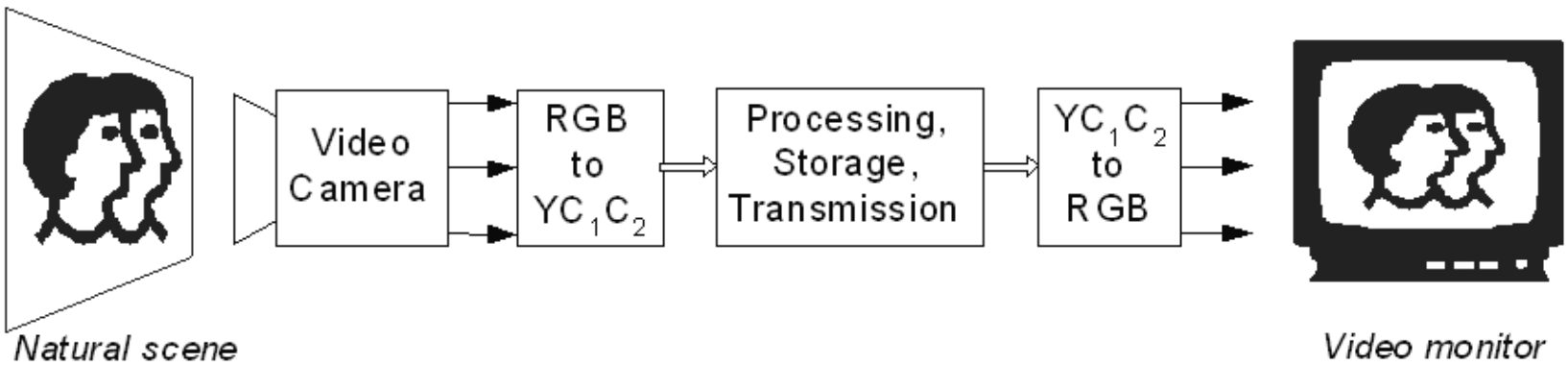
- $Y_d C_b C_r$ Colour Space
  - For storage, transmission and processing
  - Isolates luminance component (Y)
    - Human vision is more sensitive to luminance than colour

$$
\begin{aligned}
Y_d &= 219[0.229R' + 0.587G' + 0.144B'] + 16 \\
C_b &= 224[-0.169R' - 0.331G' + 0.500B'] + 128 \\
C_r &= 224[0.500R' - 0.419G' + 0.081B'] + 128
\end{aligned}
$$

R', ***G'*** and ***B'*** are normalised to 1

© CRC Press, Wu & Rao Eds, 2006

Digital representation of colour format conversion of natural video stream



© CRC Press, Wu & Rao Eds, 2006

- $Y_d C_b C_r$ Operational Ranges
  - $16 \leq Y_d \leq 235$
  - $-112 \leq C_b, C_r \leq 112$
- $Y_d C_b C_r$ Component Resolutions
  - Full resolution (4:4:4)
  - Half resolution (4:2:2)
  - Quarter resolution (4:2:0)

© CRC Press, Wu & Rao Eds, 2006

1.2.2 Statistical Analysis

- Design of coding system is dependent on statistical property of target images
- Correlation implies similarity
  - High correlation in data → high compression
- Variance implies difference
  - High variance in data → low compression

© CRC Press, Wu & Rao Eds, 2006

1.3 Compression and Coding Techniques

- Entropy Coding
  - Shannon's first theorem
    - Describes the minimum bits required, ***H(x)***, to represent some source data, ***x***, with ***m*** number of symbols having a probability distribution, ***p***.

$$H(x) = -\sum_{i=1}^{m} p(x_i) \log_2 p(x_i)$$

© CRC Press, Wu & Rao Eds, 2006

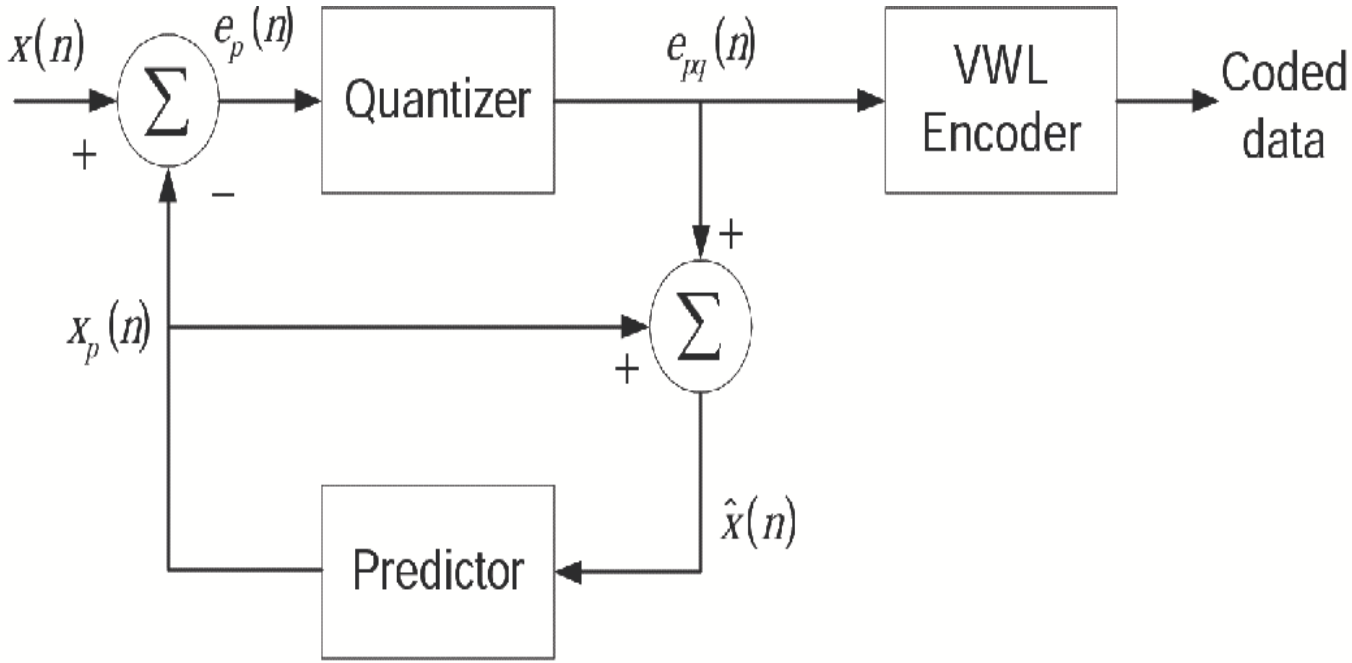
- Entropy coders
  - Huffman – block oriented
  - Arithmetic – stream oriented
  - Utilise variable length codes
    - More efficient than fixed length codes

© CRC Press, Wu & Rao Eds, 2006

1.3.2 Predictive Coding

- Rationalè for predictive coding
  - high correlation → low prediction errors
  - low prediction errors → lower entropy
- Applicable to lossy and lossless coding
  - Lossless coder
    - LOCO (JPEG-LS core), CALIC
  - Lossy coder
    - Differential Pulse Code Modulation (DPCM) for motion compensated video
    - Pulse coding of DC coefficients in JPEG/MPEG

© CRC Press, Wu & Rao Eds, 2006



Closed loop DPCM encoder with Variable Word Length (VWL) coding

© CRC Press, Wu & Rao Eds, 2006

1.3.3 Transform Coding

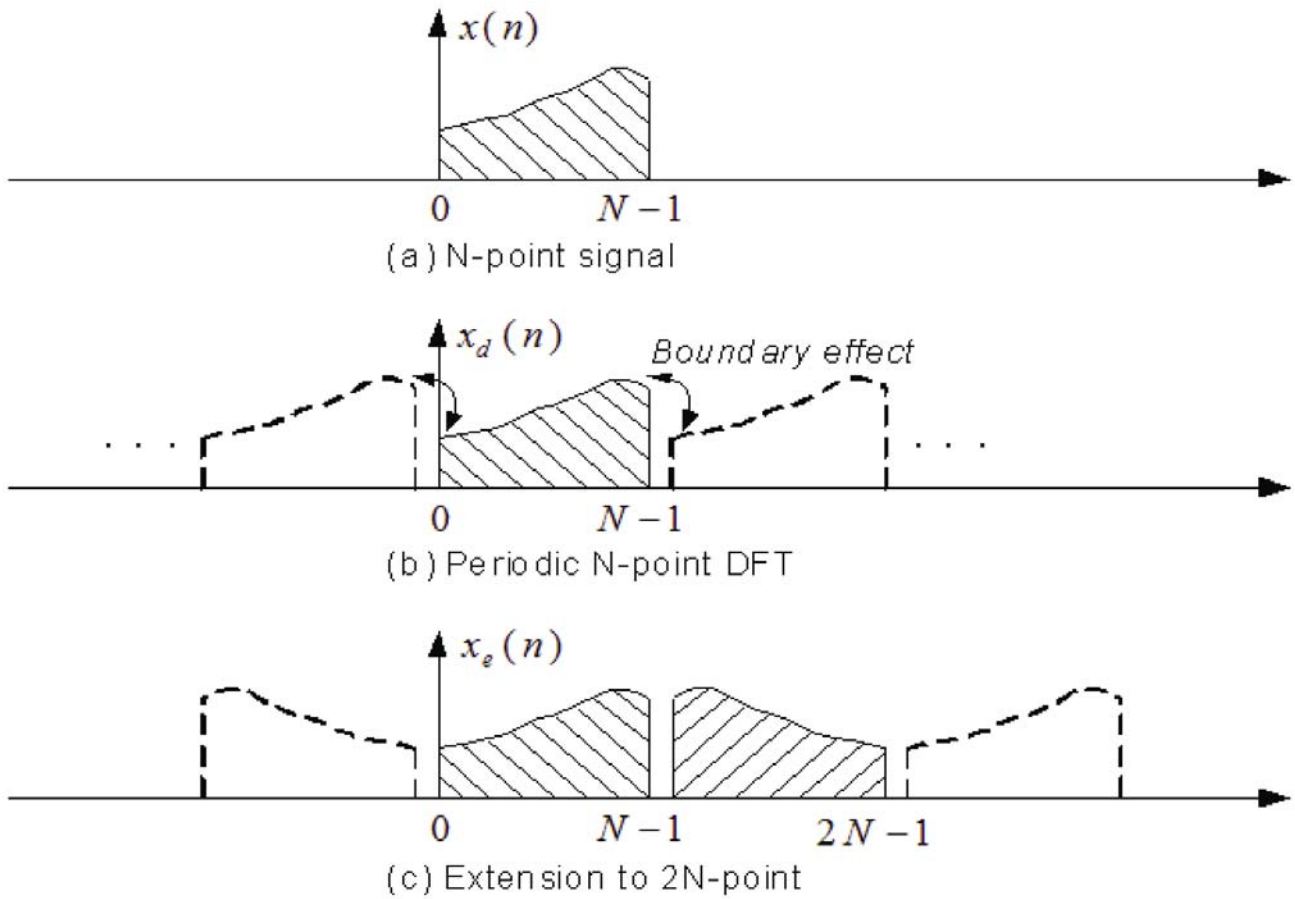
- Discrete Cosine Transform (DCT)
  - Derived from DFT
  - N-point can be extended to 2N-point (Mirroring)
  - Functional definition of 2-D DCT

$$X_{DCT}[u,v] = \frac{4}{N \times M} c[u]c[v] \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x[n,m] \cos\left[\frac{(2n+1)u\pi}{2N}\right] \cos\left[\frac{(2m+1)v\pi}{2M}\right]$$

$$c[u], c[v] = \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0; \\ 1 & otherwise. \end{cases}$$

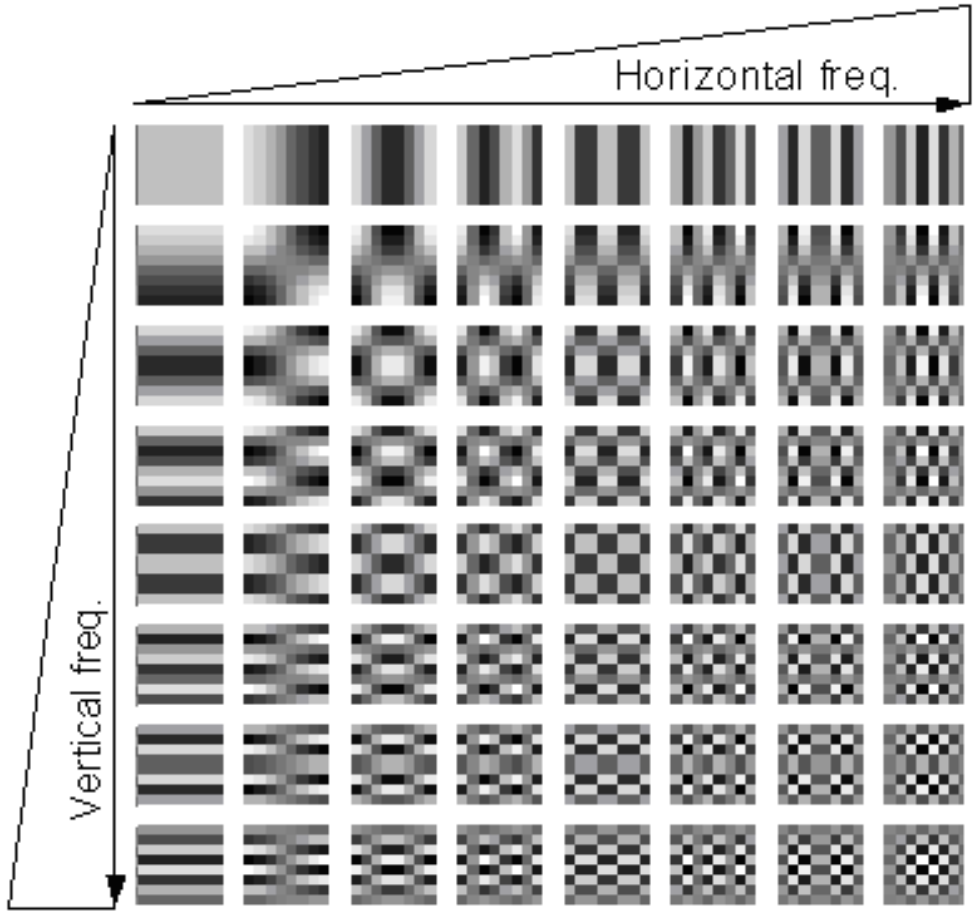
© CRC Press, Wu & Rao Eds, 2006

Signal extension. *A)* input signal. *B)* N-point DFT boundary effect. *C)* symmetric even extension (mirroring)



© CRC Press, Wu & Rao Eds, 2006

8x8 DCT Basis Image

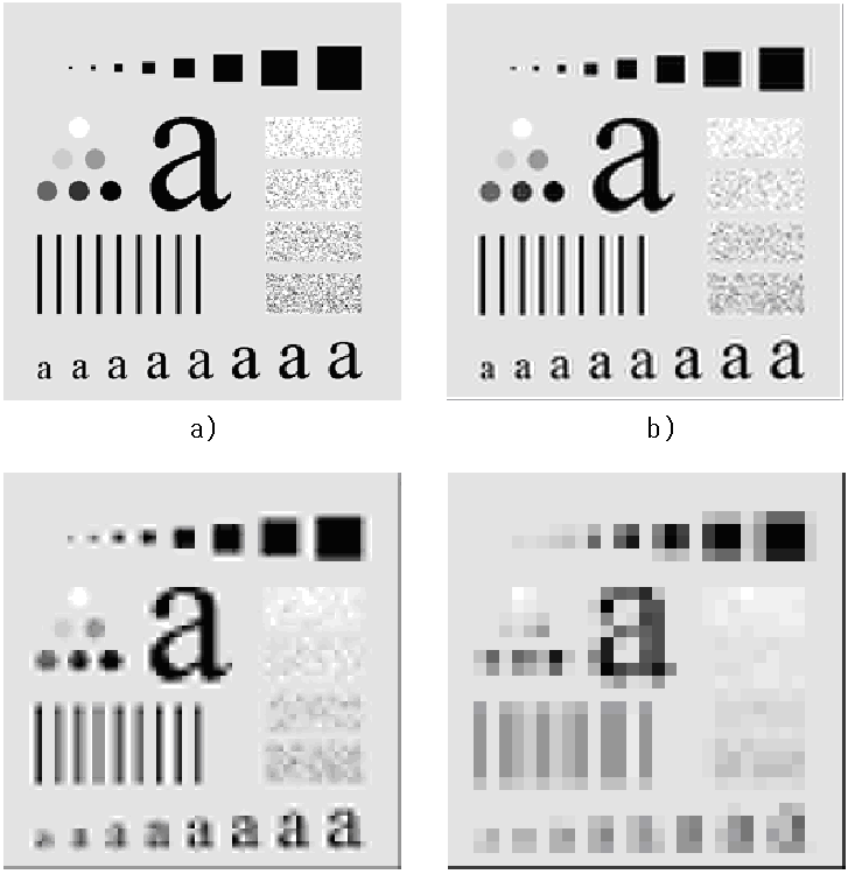


© CRC Press, Wu & Rao Eds, 2006

- Separable transform
  - 2-D DCT obtained through cascade of 1-D DCT
- Fast algorithm implementation
  - Only for 2^n-point DCTs
- Produces visually undesirable blocking artifacts at low bitrates.

© CRC Press, Wu & Rao Eds, 2006

Blocking artifacts from DCT



a) Test image

b) Inverse DCT with 4x4 low frequency coefficients

c) Inverse DCT with 2x2 low frequency coefficients

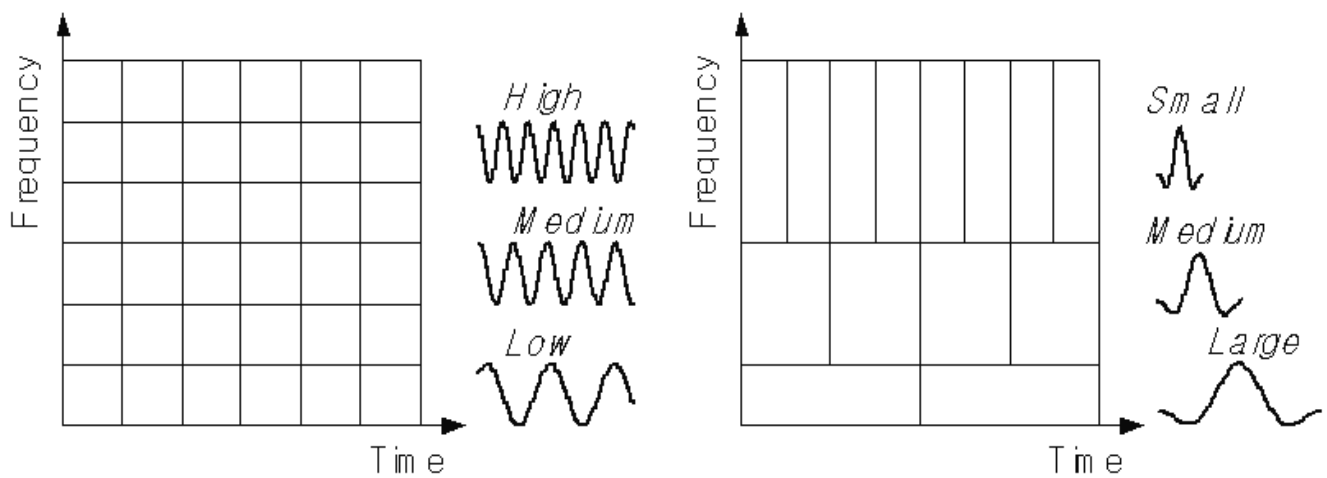
d) Inverse DCT with DC coefficients only

© CRC Press, Wu & Rao Eds, 2006

1.3.3.2 Discrete Wavelet Transform

- DWT is localised in space
- DWT more suited than DFT in dealing with non-stationary signals

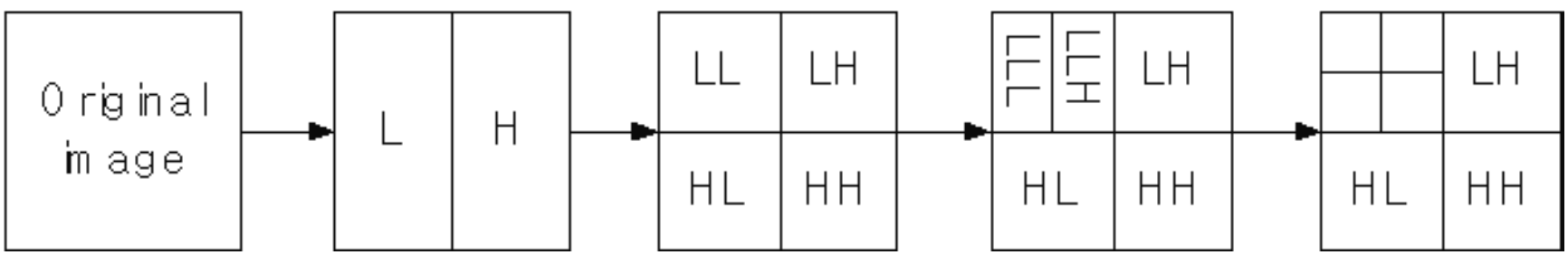
Below: time-frequency resolution of DFT and DWT



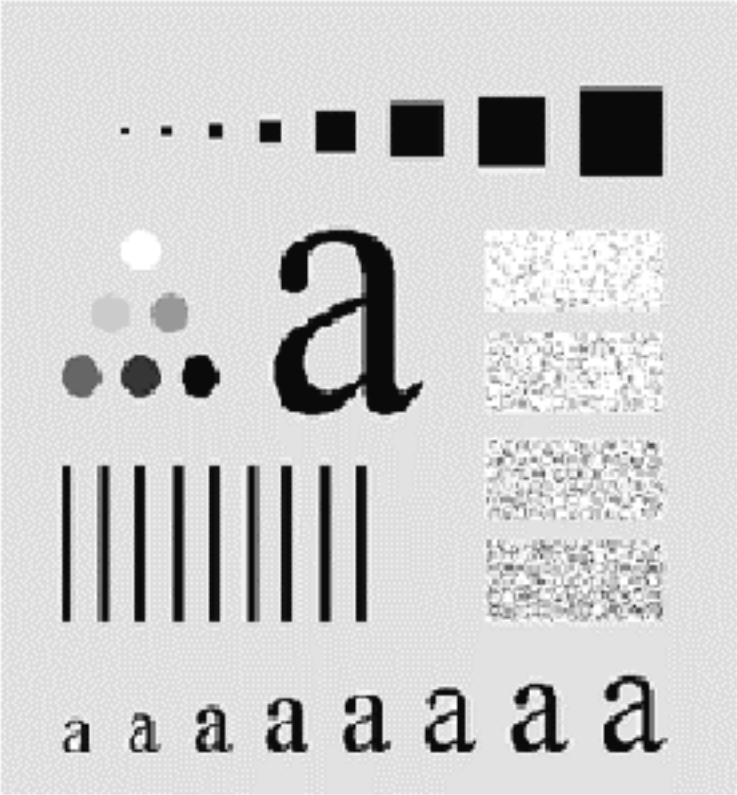
© CRC Press, Wu & Rao Eds, 2006

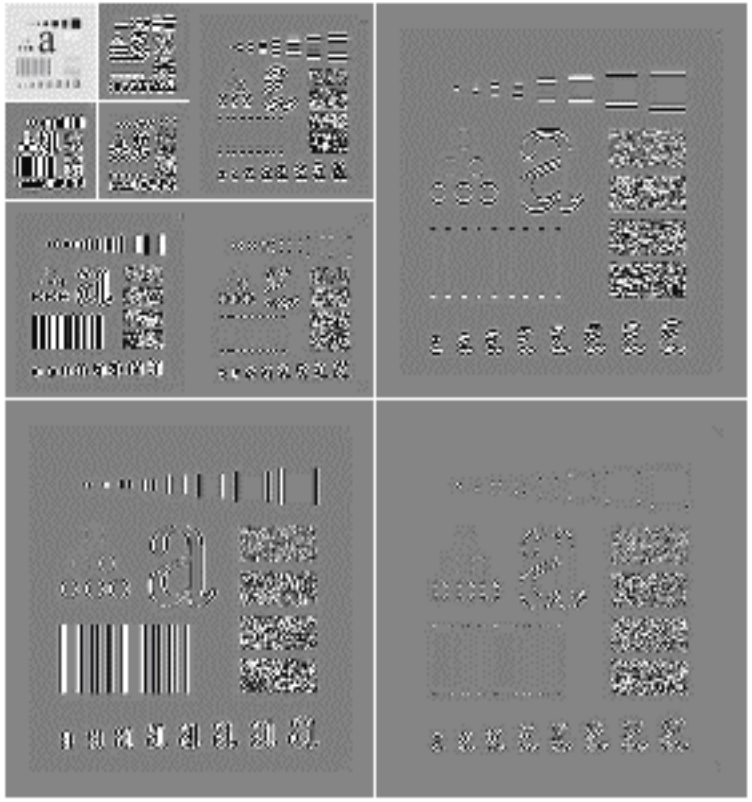
DWT

- Generally implemented in dyadic structure
  - 2 band filters, high-pass and low-pass
- Multi-resolution representation
  - Obtained through successive Iterative filtering of low-pass band



© CRC Press, Wu & Rao Eds, 2006



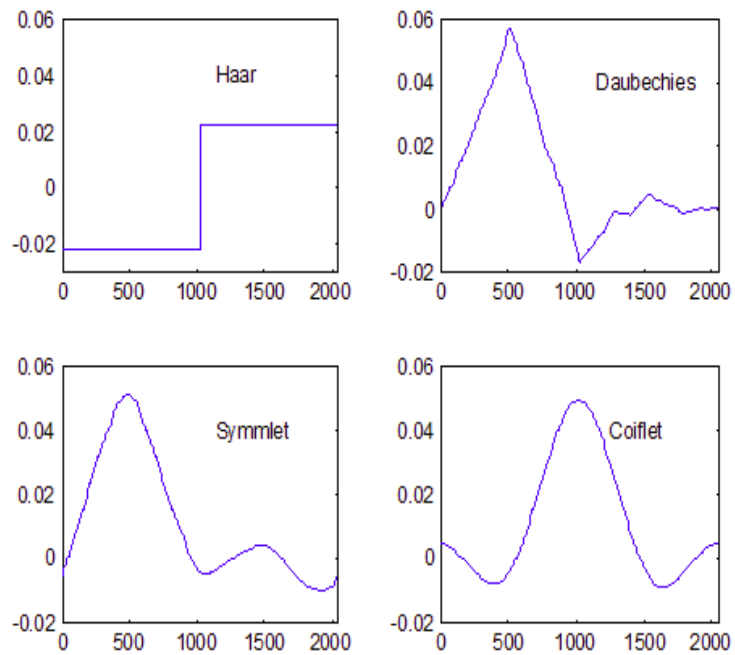


Left: Test Image
Right: 3 level spectral image

© CRC Press, Wu & Rao Eds, 2006

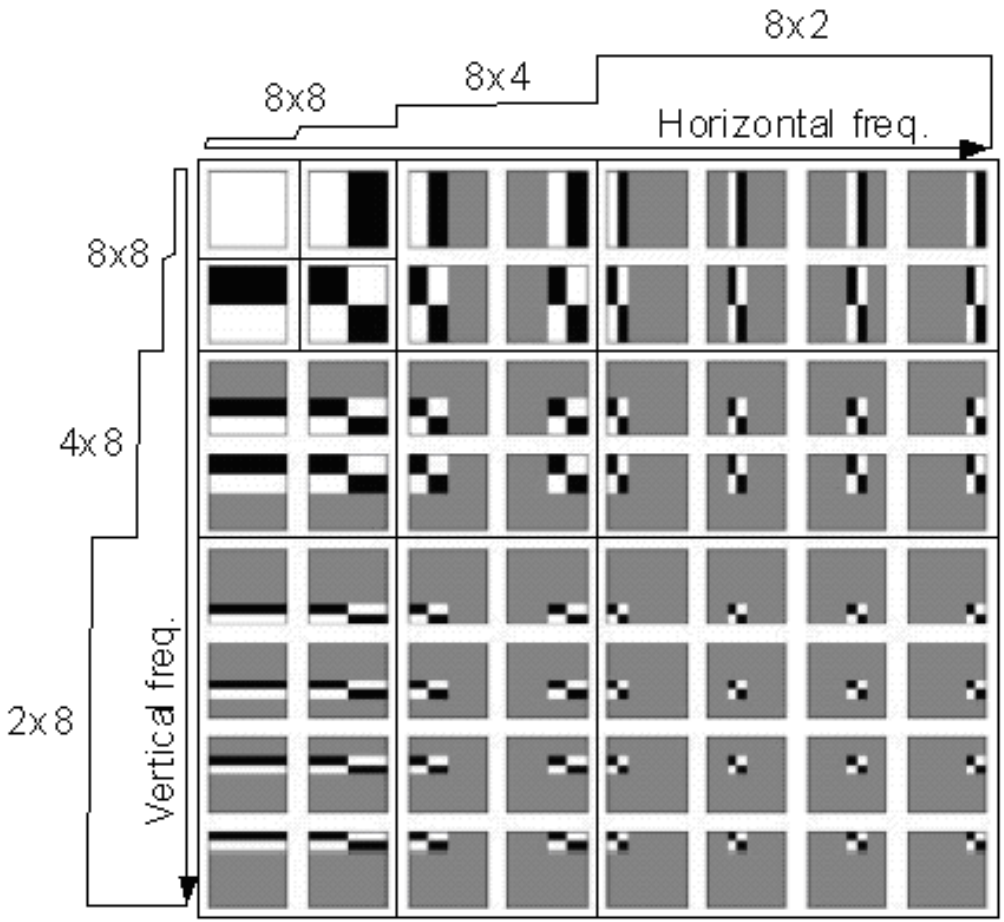
DWT

- DWT may use any number of filters (right)
- *Haar* filters - Simplest filter set
- For picture coding, spline filters are commonly employed



© CRC Press, Wu & Rao Eds, 2006

Discrete Haar Transform 8x8 basis image



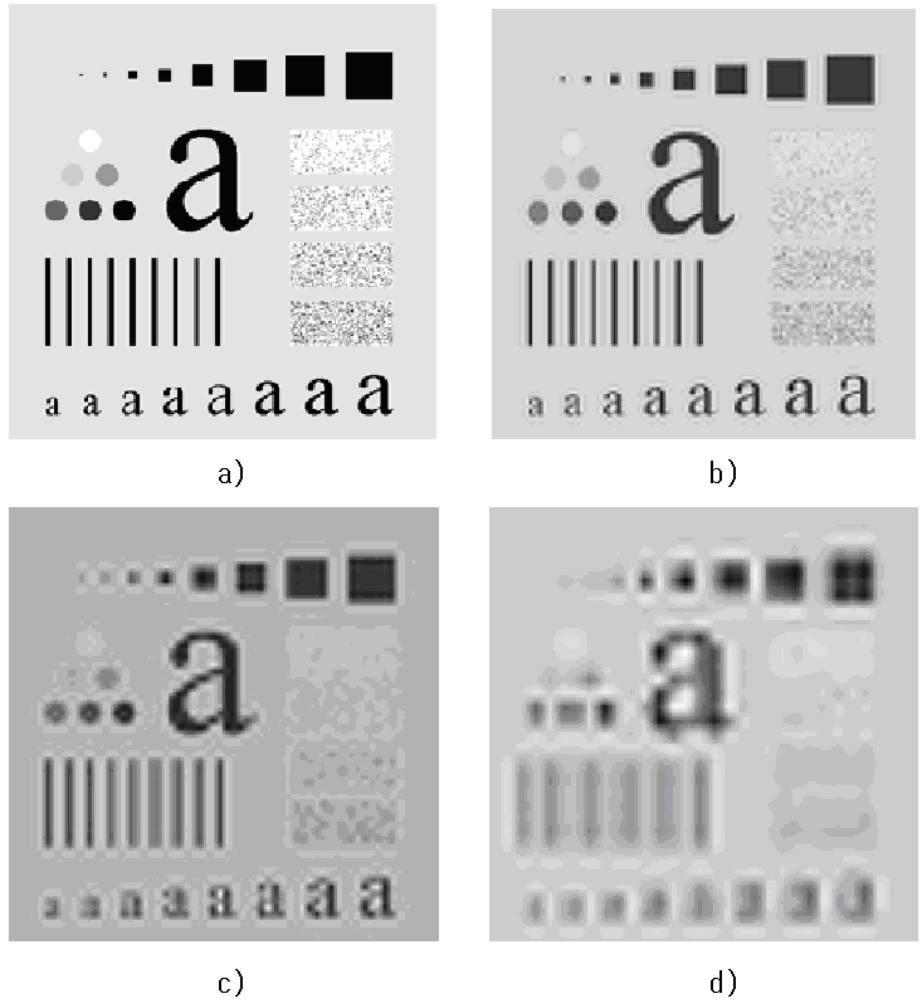
© CRC Press, Wu & Rao Eds, 2006

- Dyadic wavelet transform matches well with psychoviual models
  - Produces coded images with superior visual quality at medium and high bitrate
- Distortions in DWT coded images
  - blurring and ringing artifacts at low bitrates

© CRC Press, Wu & Rao Eds, 2006

Blurring artifacts in 3 level wavelet transform

a) Test image
b) DWT image without level 1 coefficients
c) DWT images without levels 1 and 2 coefficients
d) DWT images with only the low-pass (LL) band coefficients



© CRC Press, Wu & Rao Eds, 2006

1.4 Picture Quantization

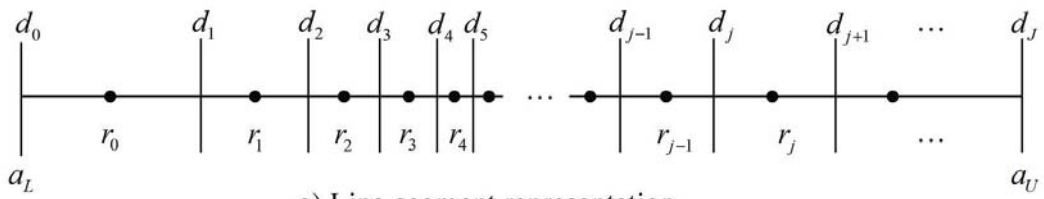
- Maps *K* inputs to *L* outputs
- Memory and Memoryless
  - Quantization is dependent and independent of other causal sample data, respectively
- Deadzone quantiser
  - Located about the mean value of an image, typically zero
  - Larger step size within the deadzone reduces variance and improves entropy coding

© CRC Press, Wu & Rao Eds, 2006

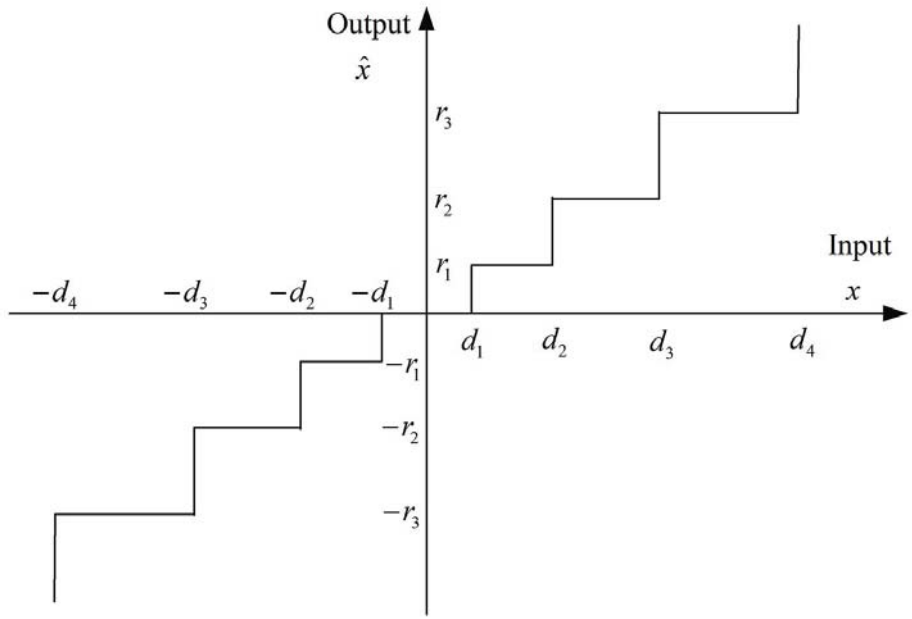
- Uniform Quantizer
  - Constant step size
  - Mid-rise or mid-tread
- Non-uniform quantizer
  - Variable step size
  - Outputs are dependent on the PDF of inputs

© CRC Press, Wu & Rao Eds, 2006

Non-uniform quantizer



a) Line-segment representation



b) Staircase representation

© CRC Press, Wu & Rao Eds, 2006

1.4.2 Optimal Quantizer Design

- Minimize quantization errors
  - For K number of inputs (d) and L number of outputs (r)
  - $q_{L-1} \leq d < q_L \rightarrow r_L$
  - Determine the best set of quantization intervals, $q = \{q_0, q_1, q_2, \ldots, q_L\}$, that produces minimal quantization errors.

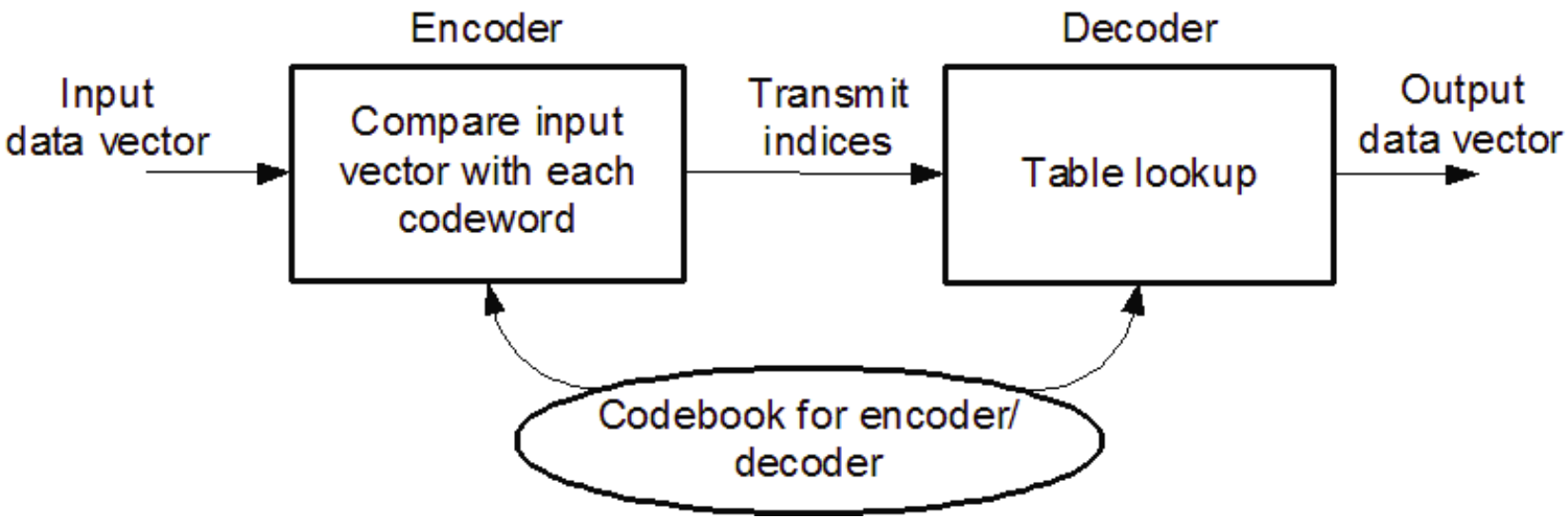
© CRC Press, Wu & Rao Eds, 2006

1.4.3 Vector Quantizer

- Input vector is represented by a symbol within a codebook
- Minimal distortion criteria for selecting symbol
- Optimal vector quantizers are impractical
  - Require large codebooks
- Scalar quantizers are vector quantizers of dimension 1

© CRC Press, Wu & Rao Eds, 2006

Block diagram of a simple vector quantizer



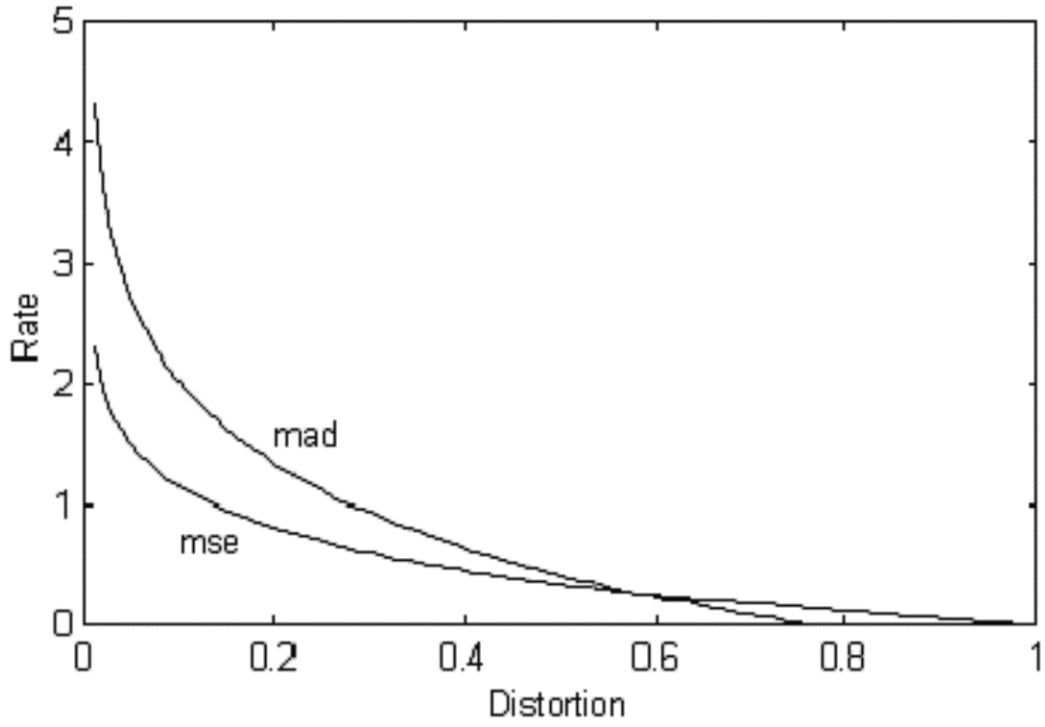
© CRC Press, Wu & Rao Eds, 2006

1.5 Rate-Distortion Theory

- Minimum bitrate required to reproduce a source at various distortion levels
- May utilise different error measures
  - MSE, MAD, etc
- Employed to control quality of digitally coded pictures
- Accuracy of R-D estimation increases with the sophistication of algorithms

© CRC Press, Wu & Rao Eds, 2006

Rate-Distortion curve for MSE and MAE



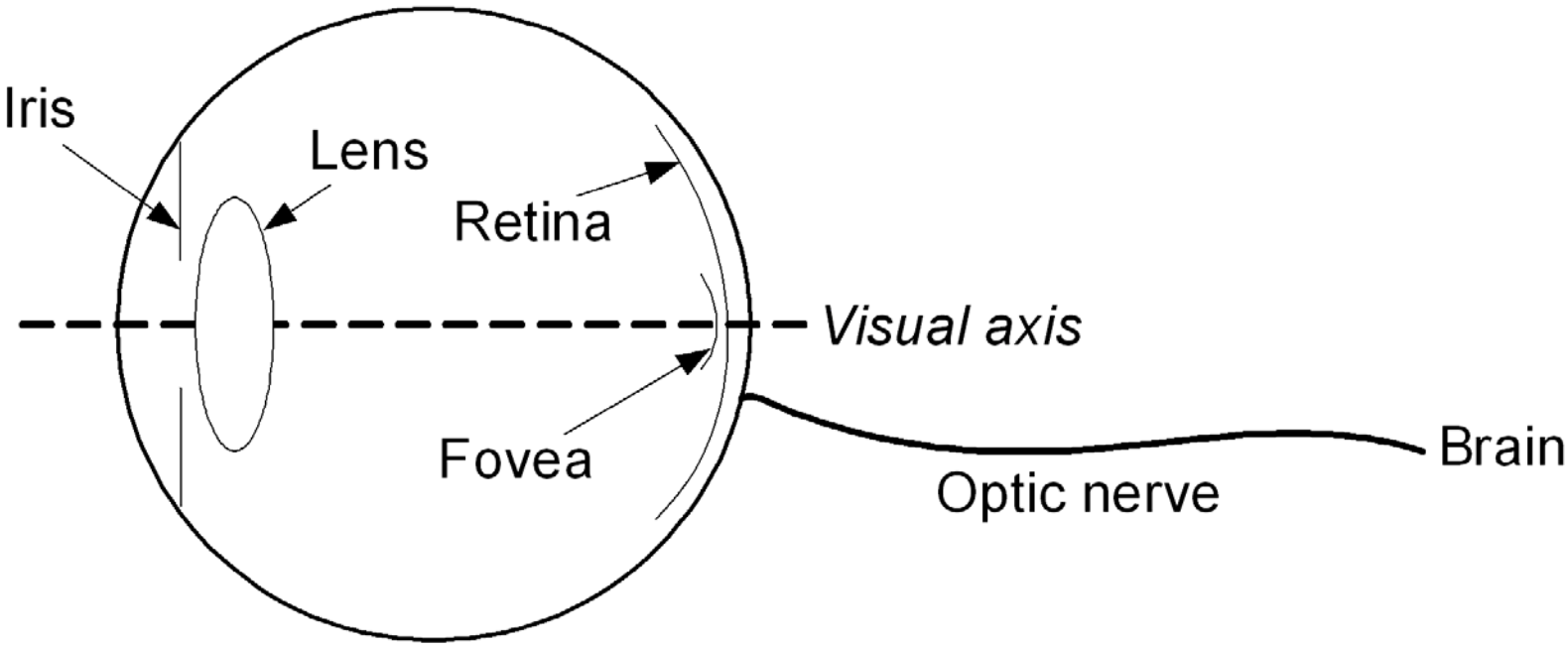
© CRC Press, Wu & Rao Eds, 2006

1.6 Human Visual System

- HVS Physiological make-up
  - Eyes
    - Cornea and lens for focusing
    - Photoreceptors in retina (cornea) to capture light
    - Optic nerve for transmission of receptor signals
  - Visual pathway
    - Connects the eye to the cortex
  - Visual cortex
    - Processing centre – detection, recognition, etc

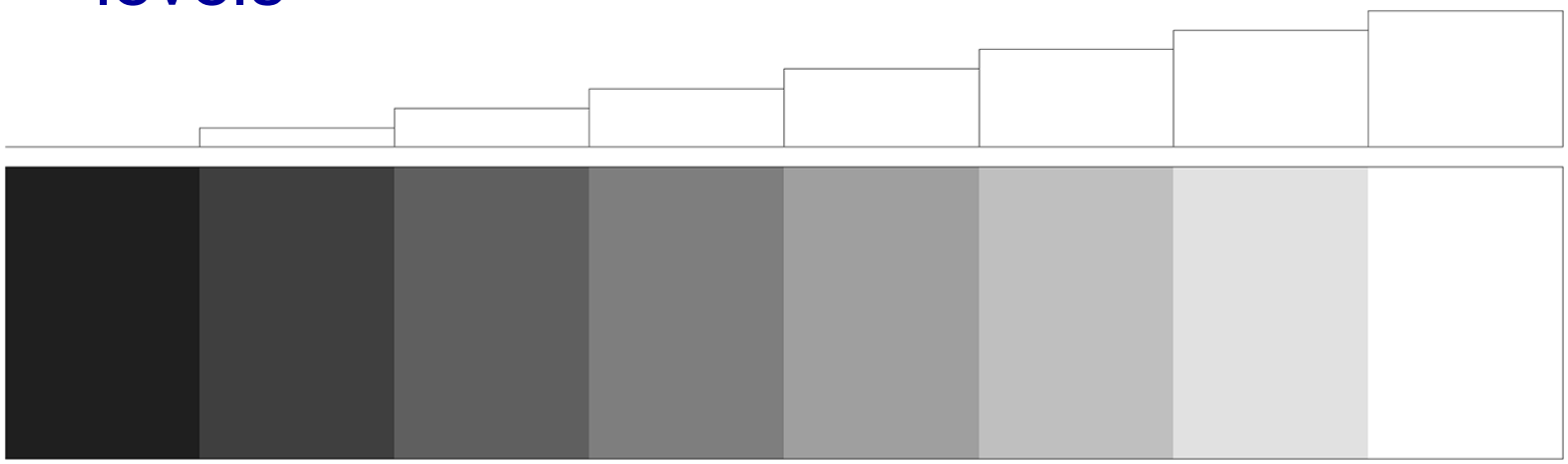
© CRC Press, Wu & Rao Eds, 2006

Simplified human eye – cross section



© CRC Press, Wu & Rao Eds, 2006

- Contrast
  - HVS operates in relative manner. i.e., difference
  - Additionally, the HVS is sensitive to intensity levels



© CRC Press, Wu & Rao Eds, 2006

- Weber's fraction: $\Delta I / I$
  - HVS sensitivity relative to difference and intensity

© CRC Press, Wu & Rao Eds, 2006

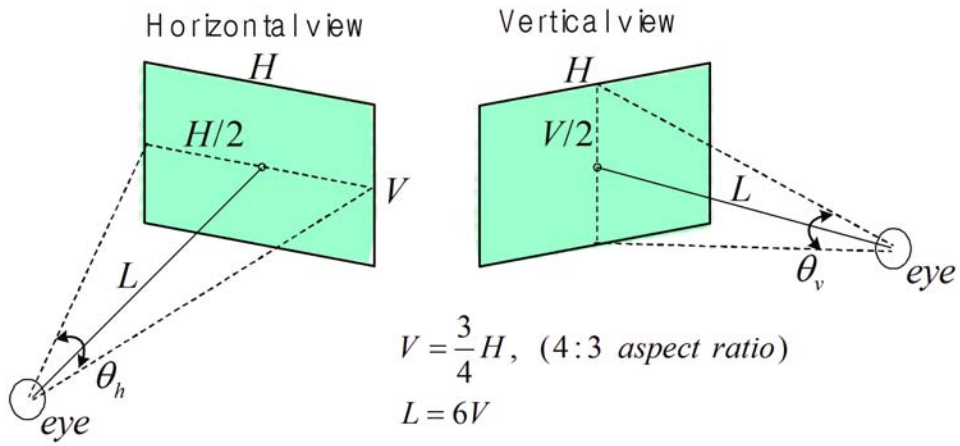
- Spatial Frequency
  - Visual acuity (cyc/deg) of the human optics
  - Expressed by the modulation transfer function (MTF)
  - MTF Formulation (Mannos and Sakrison):

$$F_{MTF}(f_r) = 2.6(0.0192 + 0.144 f_r)e^{-(0.114 f_r)^{1.1}}$$

f_r denotes spatial frequency normalised by viewing angles

© CRC Press, Wu & Rao Eds, 2006

Viewing environment and spatial frequency concept



Horizontal angle

$\theta_h = 2\arctan((H/2)/6V)$

Vertical angle

$\theta_v = 2\arctan((V/2)/6V)$

© CRC Press, Wu & Rao Eds, 2006

- Masking
  - Occurs during the interaction of neural signals
  - Results in the suppression and enhancement of certain details in images

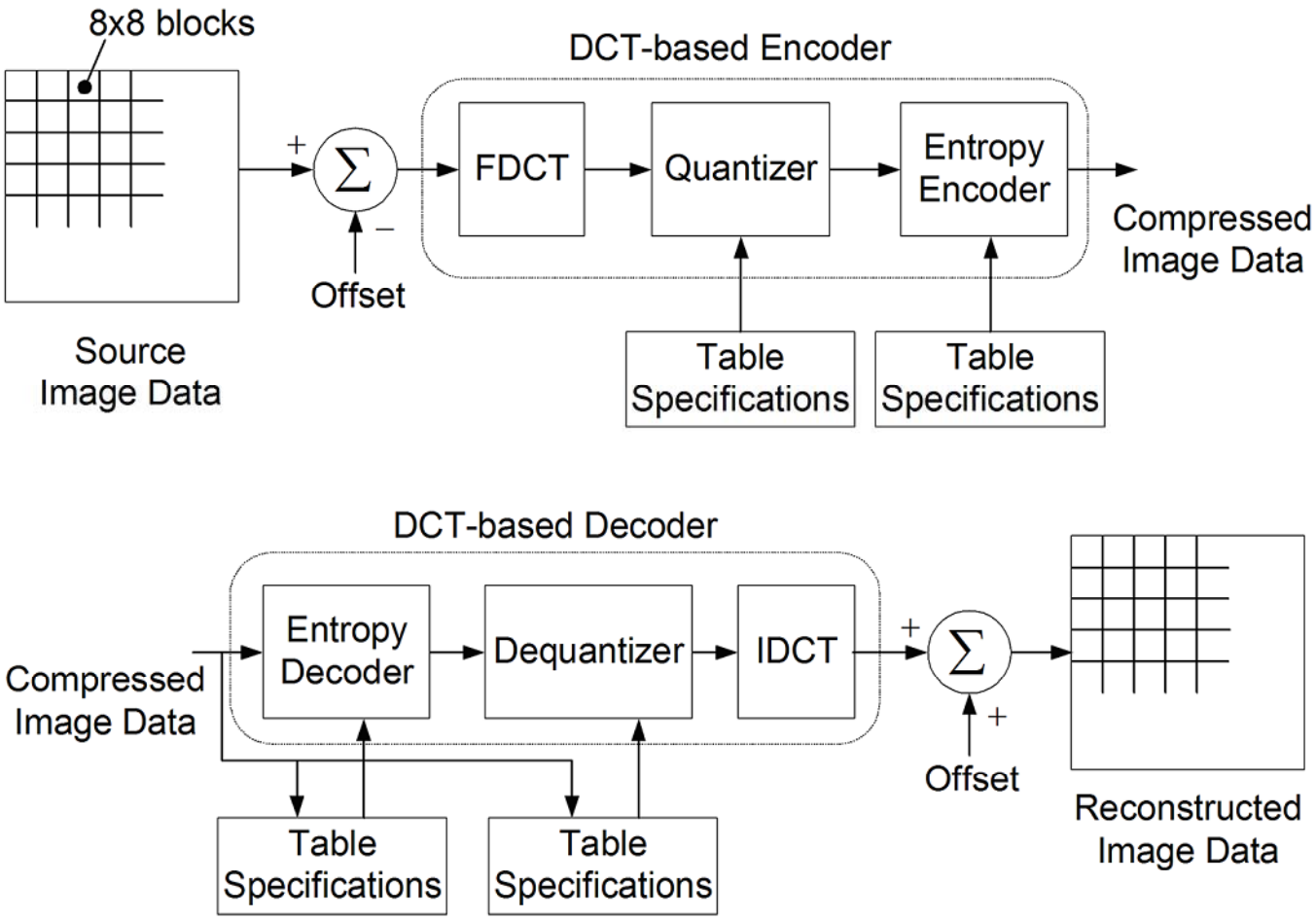
© CRC Press, Wu & Rao Eds, 2006

1.7 Digital Picture Coding Standards and Systems

- JPEG Still Image Coding Standard
  - Block Based DCT
    - 8x8 DCT
  - Scalar quantizer
    - For luminance and chrominance channels
  - Run-length and huffman coding
    - Zigzag scanning order

© CRC Press, Wu & Rao Eds, 2006

JPEG Coding Structure



© CRC Press, Wu & Rao Eds, 2006

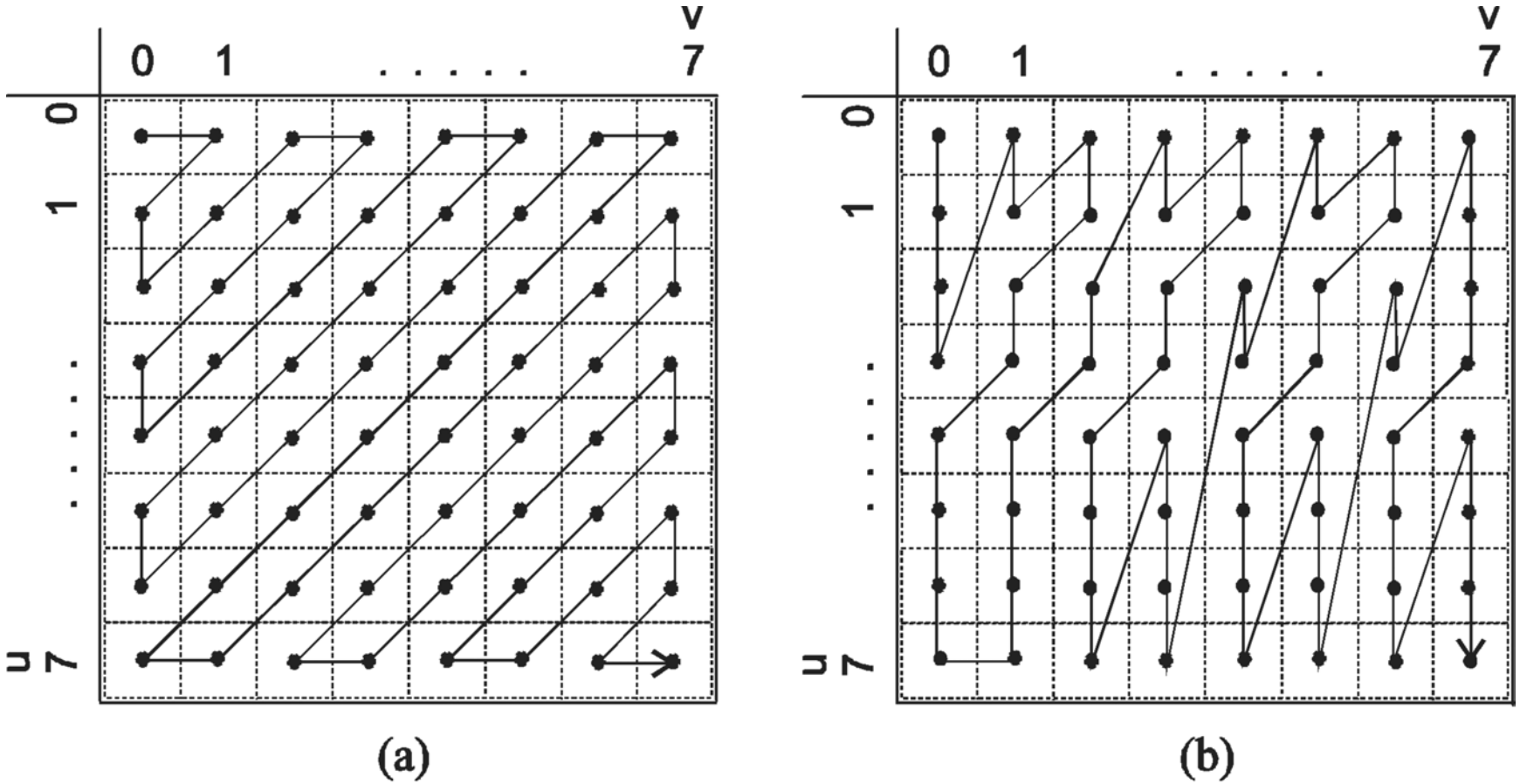
Quantization matrix for luminance (left) and chrominance (right) channels

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

© CRC Press, Wu & Rao Eds, 2006

Scanning order for DCT Coefficients

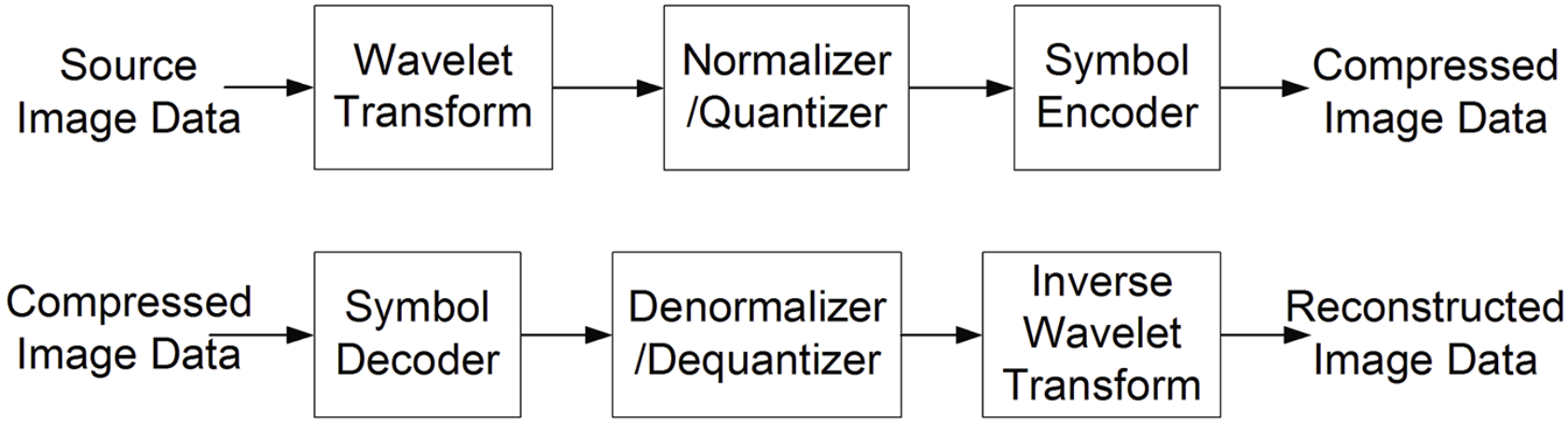


© CRC Press, Wu & Rao Eds, 2006

- JPEG2000 Standard
  - Wavelet transform
    - Multi-resolution
    - Any bi-orthogonal filters
      - Typically 9/7 or 5/3 integer filter set
  - Progressive bitplane quantization
  - Run-length and arithmetic coding
  - Scalable coding
    - Rate
    - Resolution

© CRC Press, Wu & Rao Eds, 2006

Simplified JPEG2000 coding structure



© CRC Press, Wu & Rao Eds, 2006

Comparative performance of JPEG and JPEG2000 with *Girl* image

	Compression Ratio	Error (rms)	Compression Parameter
JPEG	21.85	7.63	Quality factor = 4
JPEG2000	22.53	6.26	M = 8, e = 8.4

© CRC Press, Wu & Rao Eds, 2006

Left column: JPEG

Right column: JPEG2000

Bottom row: Difference image (between original and coded)

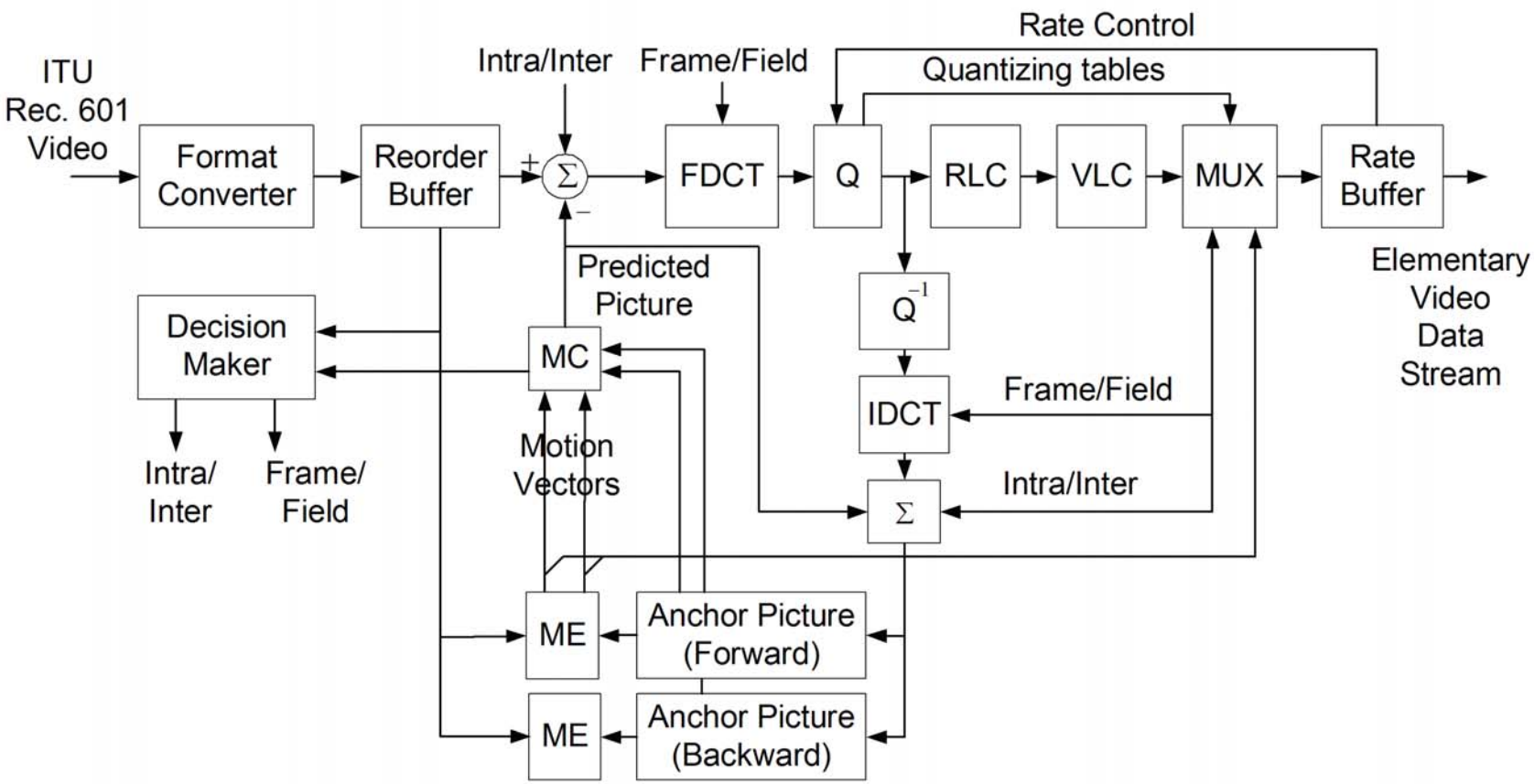
© CRC Press, Wu & Rao Eds, 2006

MPEG-X Standards

- Basic Framework
  - Block based DCT coding
  - MC/DPCM for inter-frame coding
  - Scalar quantisation
    - Different quantisation strategies for inter- and intra-frame coding
    - Different matrix for different colour channel
  - Run-length and huffman coding
    - Zigzag scanning of DCT coefficients

© CRC Press, Wu & Rao Eds, 2006

MPEG-1 Coding Structure

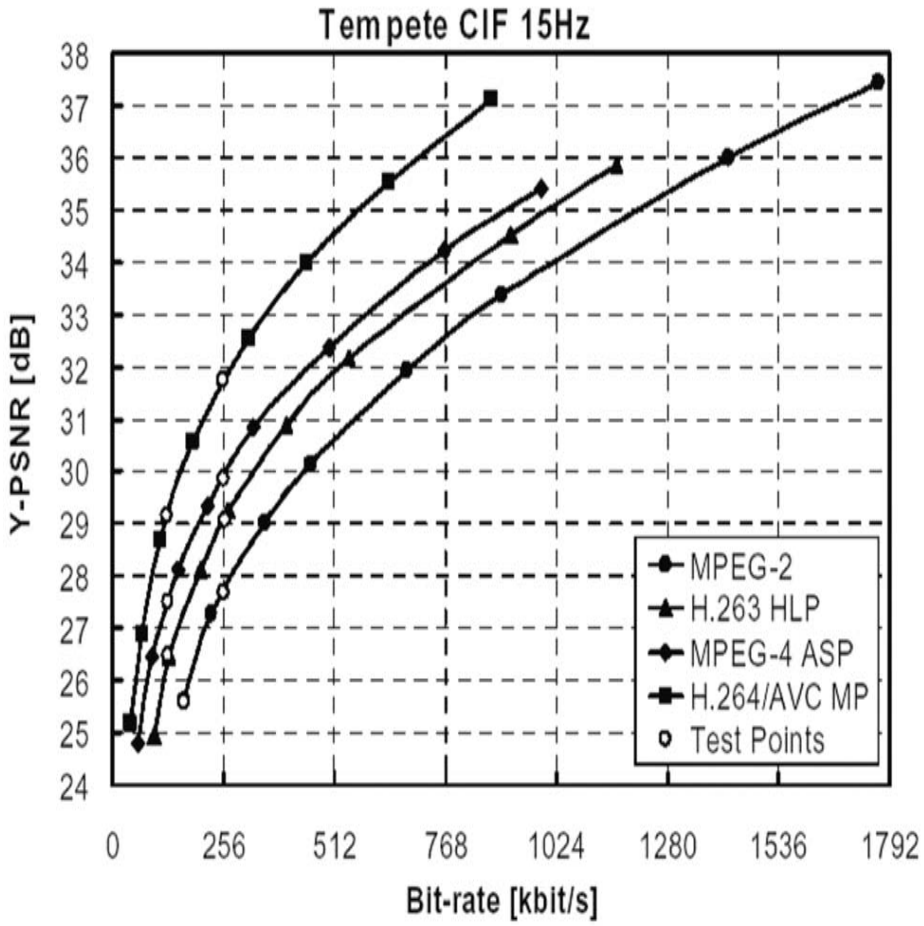


© CRC Press, Wu & Rao Eds, 2006

Video coding standards

- Chronology
  - MPEG1/H.261
  - MPEG2 → H.263
  - MPEG4/ASP
  - H.264/AVC

Left: standardized video codecs performances



© CRC Press, Wu & Rao Eds, 2006

Comparison of standards: MPEG1, MPEG2, MPEG-4 and H.264/MPEG-4 Part10

Feature/ Standard	MPEG-1	MPEG-2	MPEG-4	H.264/MPEG-4 Part 10
Macroblock size	16x16	16x16 (frame mode) 16x 8 (field mode)	16x16	16x16
Block size	8x8	8x8	16x16, 8x8, 16x8 8x8, 16x8	8x8, 16x8, 8x16, 16x16, 4x8, 8x4, 4x4
Transform	DCT	DCT	DCT/Wavelet transform	4x4 Integer transform
Transform size	8x8	8x8	8x8	4x4
Quantization step size	Increases with constant increment	Increases with constant increment	Vector Quantization	Step sizes increase at rate of 12.5%
Entropy coding	VLC	VLC(different VLC tables for Intra and Inter modes)	VLC	VLC, CAVLC and CABAC
Motion estimation & compensation	Yes	Yes	Yes	Yes, more flexible up to 16 MVs per MB
Pel accuracy	Integer 1/2-pel	Integer 1/2-pel	Integer 1/2-pel, 1/4-pel	Integer 1/2-pel, 1/4-pel
Profiles	No	5 profiles. Several levels within a profile	8 profiles. Several levels within a profile	3 profiles. Several levels within a profile
Reference frame	Yes, One frame	Yes, One frame	Yes, One frame	Yes, Multiple frames(as many as 5 frames allowed)
Picture types	I, P, B, D	I, P, B	I, P, B	I, P, B, SI, SP
Playback & random access	Yes	Yes	Yes	Yes
Error robustness	Synchronization concealment	Data partitioning, redundancy, FEC for important packet transmission	Synchronization, data partitioning, header extension, reversible VLCs	Deals with packet loss and bit errors in error-prone wireless network
Transmission rate	Up to 1.5Mbps	2-15 Mbps	64Kbps-2Mbps	64Kbps-150Mbps
Encoder complexity	Low	Medium	Medium	High
Compatible with previous standards	Yes	Yes	Yes	No

© CRC Press, Wu & Rao Eds, 2006

Average bitrate savings for video streams [J+02]

	Average bitrate saving relative to		
Coder	MPEG-4 ASP	H.263	MPEG-2
H.264/MPEG-4 Part 10 AVC	39%	49%	64%
MPEG-4 ASP	-	17%	43%
H.263	-	-	31%

- Substantial improvement in H.264

© CRC Press, Wu & Rao Eds, 2006

1.8 Summary

- Basic introduction to digital picture coding
  - Characteristics of picture data
  - Picture coding techniques
  - Quantisation
  - Rate-distortion
  - Human visual system and picture coding
- Brief description of existing image and video coding standards

© CRC Press, Wu & Rao Eds, 2006

Ber71	T. Berger. *Rate-Distortion Theory*. Upper Saddle River, NJ: Prentice Hall, 1971
Bov05	A. Bovik (Ed.). *Handbook of Image and Video Processing*. Orlando, FL: Academic Press, 2nd Ed., 2005
CCI90	CCITT Study Group XV. Recommendations of the H-Series, CCITT Study Group XV – Report R37. Technical report, CCITT, August 1990
Cho95	C.H. Chou. Adaptive transform coding of images based on removing just noticeable distortion. In *Proceedings of SPIE Visual Com. Im. Proc.*, 2501:607-618, Taipei, Taiwan, May, 1995
Cla85	R.J. Clarke. *Transform Coding of Images*. New York: Academic Press, 1985
CSE00	C. Christopoulos, A. Skodras and T. Ebrahimi. The JPEG2000 still image coding system: An overview. *IEEE Trans. Consumer Elect.* 46:1103-1127, November 2000
Ger77	A. Gersho. Quantization. *IEEE Communications Society Magazine*, 15:16-29, September, 1977

© CRC Press, Wu & Rao Eds, 2006

Gah03	M.Ghanbari. *Standard Codecs: Image Compression to Advanced Video Coding*. Londonm UK: IEE, 2003
Gib97	J.D. Gibson (Ed.). *Multimedia Compression: Applications and Standards*. San Francisco, CA: Morgan Kaufmann, 1997.
Gla95	A.S. Glassner. *Principles of Digital Image Synthesis*. San Francisco, CA: Morgan Kaufman, 1995
GW02	R.C. Gonzalez and R.E. Woods. *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall, 2nd Ed., 2002
Hec24	S. Hecht. The visual discrimination of intensity and the weber-frechner law. *Journalof General Physiology*. 7:241, September 1924
ISO92	ISO/IEC JTC1/S2/WG9. CD11544, Progressive Bi-Level Image Compression, Revision 4.1. Technical report, ISO/IEC, September 1992
ISO93a	ISO/IEC. ISO/IEC 1172 Information Technology: Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s. Technical report. ISO/IEC, 1993

© CRC Press, Wu & Rao Eds, 2006

ISO93b	ISO/IEC JTC1 10918-1. ITU-T Rec. T.81, Information Technology – Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines. Technical report, ISO/IEC, 1993
ISO93c	ISO/IEC JTC1/SC29/WG11. CD 13818, Generic Coding of Moving Pictures and Associated Audio. Technical report, ISO/IEC, November 1993
ISO00	ISO/IEC FCD 15444-1. JPEG 2000 Image Coding System. Technical report, ISO/IEC, March 2000
ISO01	ISO/IEC. ISO/IEC 14496, Information Technology – Coding of Audio Visual Objects. Technical report, ISO/IEC, 2001
ITU82	ITU-R. ITU-R Recommendation BT.601, Encoding Parameters of Digital Television for Studios. Technical report, ITU-R, 1982
ITU96	ITU-T. ITU-T Recommendation H.263, Video Coding for Low Bit Rate Communication. Technical report, ITU-T, 1996
ITU03	ITU-T. ITU-T Recommendation H.264, Advance Video Coding for Generic Audio Visual Services. Technical report, ITU-T, 2003

© CRC Press, Wu & Rao Eds, 2006

J+02	A. Joch *et al*. Performance comparison of video coding standards using lagrangian coder control. In *Proceedings of International Conference on Image Processing*, 2:501-504, September 2002. Special Session: The emerging JVT/H.26L video coding standard
JN84	N.S. Jayant and P. Noll. *Digital Coding of Waveforms, Principles and Applications to Speech and Video*. Upper Saddle River, NJ: Prentice Hall, 1984
LBG80	Y. Linde, A. Buzo and R.M. Gray. An algorithm for vector quantization desgin. *IEEE Trans. on Commun.*, 28:84-95, January 1980
LOW91	A. Leger, T. Omachi and G.K. Wallace. JPEG still picture compression algorithm. *Optical Engineering*, 30:949-954, July 1991
LR78	J.O. Limb and C.B. Rubenstein. On the design of quantizer for DPCm coders: A functional relationship between visibility, probability and masking. *IEEE Trans. on Commun.*, 26:573-578, May 1978
Max60	J. Max. Quantizing for minimum distortion. IRE Trans. Information Theory, IT-6:16-21, March 1960.

© CRC Press, Wu & Rao Eds, 2006

MPFL97	J.L. Mitchell, W.B. Pennebaker, C.E. Fogg and D.J. LeGall. *MPEG Video Compression Standard*. New York: Chapman Hall, 1997
MS74	J.L. Mannos and D.J. Sakrison. The effects of a fidelity criterion on the encoding of images. *IEEE Trans. Information Theory*. IT-20:525-536, July 1974
NH95	A.N. Netravali and B.G. Haskell. *Digital Pictures-Representation, Compression, and Standards*. New York and London: Plenum Press, 2nd Ed., 1995
O'N76	F.B. O'Neal, Jr. Differential pulse code modulation with entropy coding. *IEEE Trans. Information Theory*, IT-21:169-174, March 1976
PC00	A. Puri and T. Chen (Eds.). *Multimedia Systems, Standards and Networks*. New York: Marcel Dekker, 2000
PE02	F.Pereira and T. Ebrahimi (Eds.). *The MPEG-4 Book*. Upper Saddle River, NJ: IMSC Press, 2002
RH96	K.R. Rao and J.J. Hwang. *Techniques and Standards for Image, Video and Audio Coding*. Upper Saddle River, NJ: Prentice Hall, 1996

© CRC Press, Wu & Rao Eds, 2006

Ric03	I.E.G. Richardson. *H.264 and MPEG-4 Video Compression*. Hoboken, NJ: Wiley, 2003
RY90	K.R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. San Diego, CA: Academic Press, 1990
Sal00	D. Saloman. *Data Compression: The Complete Reference*. Berlin: Springer-Verlag, 2000
Say00	K. Sayood. *Introduction to Data Compression*. San Francisco, CA: Morgan Kaufmann, 2000
Sch83	R. Schafer. Design of adaptive and nonadaptive quantizers using subjective criteria. *Signal Processing*. 5:333-345, July 1983
Sch86	W.F. Schreiber. *Fundamentals of Electronic Imaging Systems*. Berlin: Springer-Verlag, 1986
SN77	D.K. Sharma and A.N Netravali. Design of quantizers for DPCm coding of picture signals. *IEEE Trans. on Commun*. 25-1267-1274, November 1977

© CRC Press, Wu & Rao Eds, 2006

SW96	D. Shen and S. Wang. Measurements of JND property of HVS and its applications to image segmentation, coding and requantization. In *Proceedings of SPIE Dig. Comp. Tech. Sys. For Vudeo Com.*, 2952:113-121, October 1996
VH92	M. Vetterli and C. Herley. Wavelets and filter banks: theory and design. *IEEE Trans. On Sig. Proc.*, 40:2207-2232, November 1992

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

2. Fundamentals of Human Vision and Vision Modelling

E.D. Montag and M.D. Fairchild

Rochester Institute of Technology, USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 2 Outline

2.1 Introduction
2.2 Brief Introduction of the Visual System
2.3 Colour Vision
2.4 Luminance and the Perception of Light Intensity
2.5 Spatial Vision and Contrast Sensitivity
2.6 Temporal Vision and Motion
2.7 Visual Modelling
2.8 Conclusion

© CRC Press, Wu & Rao Eds, 2006

2.1 Introduction

- Picture display and coding systems are adapting to the human visual system
- The need to understand the HVS
  - Characteristics
  - Limitations
  - Sensitivity to spatial and temporal changes
  - Colour perception

© CRC Press, Wu & Rao Eds, 2006

2.2 Brief Introduction of the Visual System

- Human Optics
  - Lens for focusing
  - Cornea
    - 3 cellular and 2 synaptic layers
    - Fovea - Primary focal point of the visual field
  - Photoreceptors
    - Cones for photopic vision – 7 million
      - Short (S), medium (M) and long (L) wavelengths
    - Rods for scotopic vision – 120 million

© CRC Press, Wu & Rao Eds, 2006

- Optic nerve
  - Integrates receptor signals via ganglion cells
    - 100 rods to 1 ganglion
      - Increase sensitivity of scotopic vision
    - 1 cone to 1+ ganglion
- Light adaptation - localized in the retina
- Opponent colour
  - Due lateral ganglion inter-connection

© CRC Press, Wu & Rao Eds, 2006

- Visual Cortex
  - Neurons are orientation and frequency selective
- Medial Temporal region of the Cortex
  - Sensitive to motion

© CRC Press, Wu & Rao Eds, 2006

2.3 Colour Vision

- Colour sensation
  - of the mind not of worldly objects
- Functionally
  - Colour signals are linearly related to cone signals
  - No standardized method for specifying cone signals

© CRC Press, Wu & Rao Eds, 2006

2.3.1 Colorimetry

- All colours are derived from 3 primaries
  - Red, Blue and Green

$$C_1 = r_1R + g_1G + b_1B$$

- Colour matching – *Tristimulus*
  - Follows linear algebraic rules
    - additive

$$C_2 = r_2R + g_2G + b_2B$$

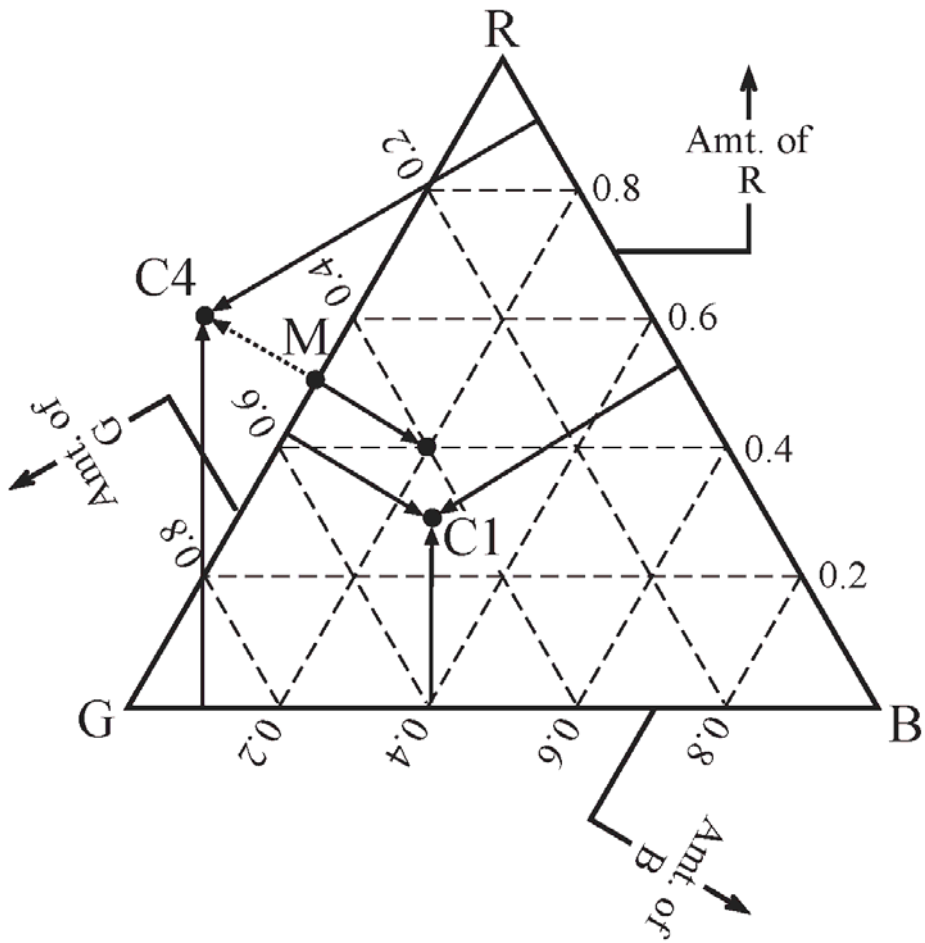
$$C_3 = C_1 + C_2 = (r_1+r_2)R + (g_1+g_2)G + (b_1+b_2)B$$

    - Proportion/product

$$kC_3 = kC_1 + kC_2$$

© CRC Press, Wu & Rao Eds, 2006

Maxwell diagram for colour matching with 3 primaries. $C_1 = 0.3R + 0.45G + 0.25B$



© CRC Press, Wu & Rao Eds, 2006

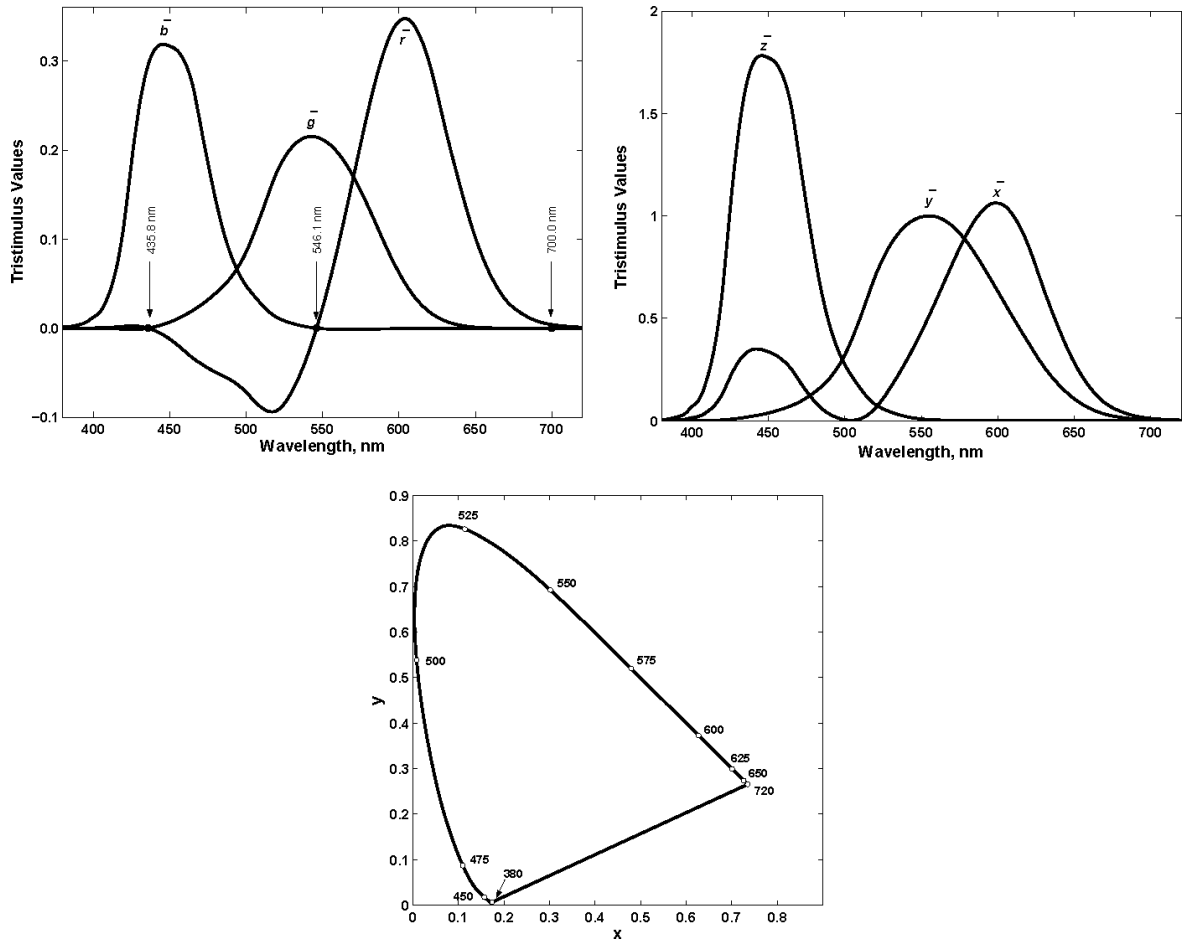
- Adding primaries to test light
  - Resulting colour outside gamut

$$C_4 = rR + gG - bB$$

- Colorimetry - matching experiments
  - real R, G and B primaries
  - CIE 1931 standard observer system
    - Imaginary primaries X, Y and Z

© CRC Press, Wu & Rao Eds, 2006

CIE 1931 2°. Top: Colour Matching for RGB (left) and XYZ (right). Bottom: Chromaticity diagram



© CRC Press, Wu & Rao Eds, 2006

2.3.2 Colour Appearance, order system and difference

- Depends on
  - Adaptation, local and global
  - Stimulus in visual field
    - Size, location and configuration (shape)
  - Objects
    - Colour and location
  - Background/surrounding colour
- Same chromaticity and luminance → may have different appearances

© CRC Press, Wu & Rao Eds, 2006

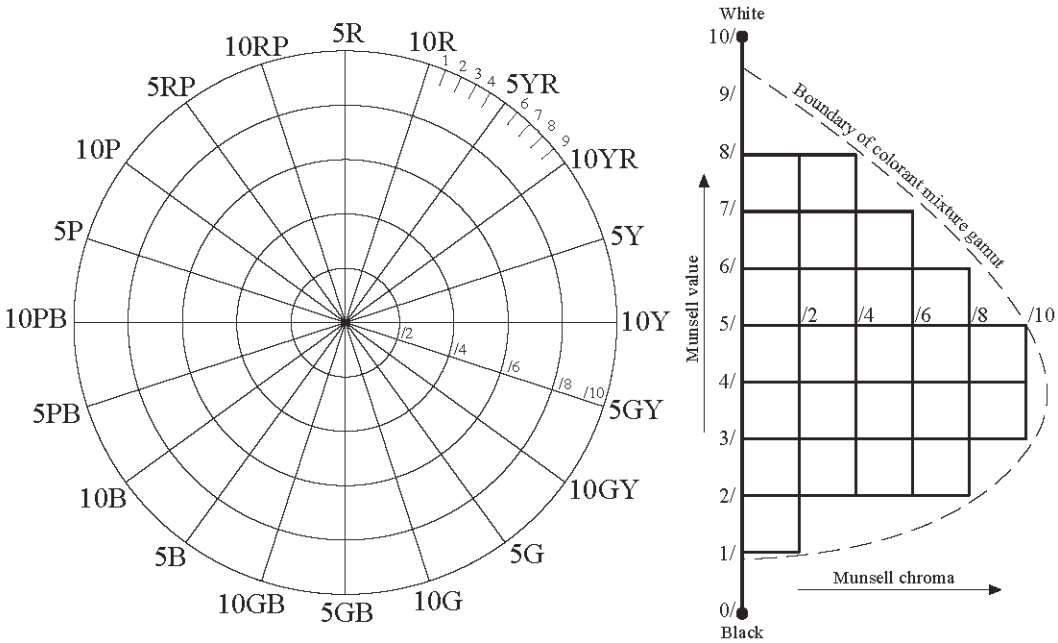
- Colour has 5 attributes
  - Brightness → intensity
  - Lightness → intensity relative to white
  - Colourfulness → chromaticity
  - Chroma → chromaticity relative to white
  - Hue → attribute of a colour
- Brightness and Colourfulness - Absolute
- Lightness and Chroma - Relative

© CRC Press, Wu & Rao Eds, 2006

- Unique Hues
  - No mixture of other hues
  - Red, Blue, Green and Yellow
- Opponent colour (hue) pair
  - Cannot co-exist in the same stimulus
  - Red-Green and Blue-Yellow
- Need for uniform description of colour appearance
  - Colour order system (Munsell book of colour)

© CRC Press, Wu & Rao Eds, 2006

Left: Hue arrangements in constant lightness plane. Right: A Hue leaf with variation in lightness and chroma.

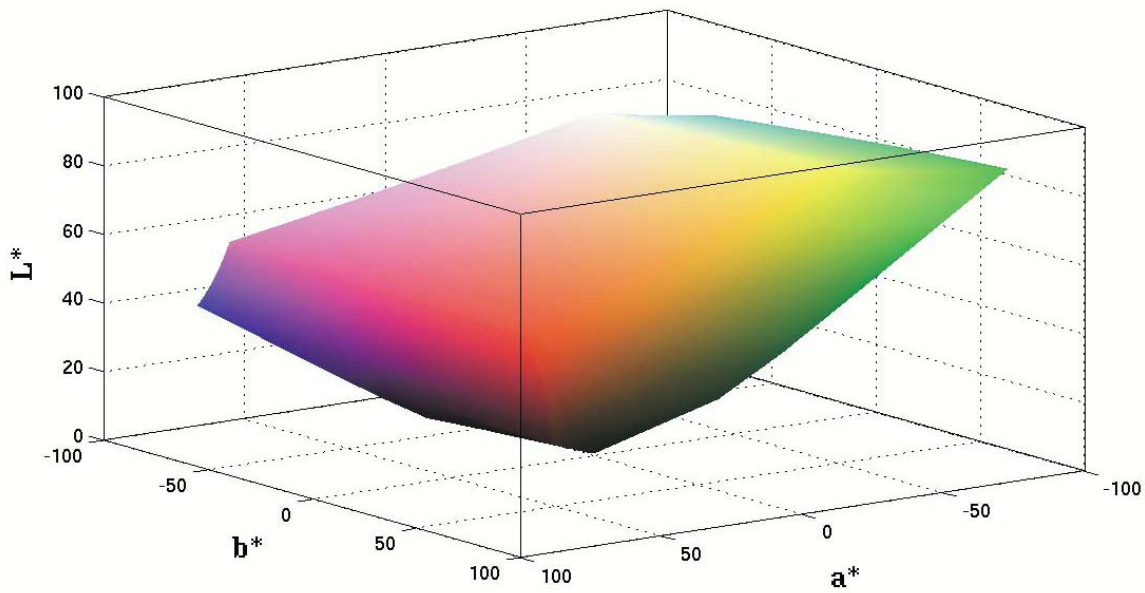


© CRC Press, Wu & Rao Eds, 2006

- Perceived colour difference
  - CIE 1976 *L*, a*, b** (CIELAB) colour space
    - Quantify perceived difference between two colours
- Colour appearance models
  - Dependent on viewing conditions
  - Calculates relative colour attributes
  - Predict brightness and colourfulness
  - may adjust for changes in colourfulness and contrast

© CRC Press, Wu & Rao Eds, 2006

CRT gamut plot in CIELAB space



© CRC Press, Wu & Rao Eds, 2006

2.4 Luminance and the Perception of Light Intensity

- Luminance
  - Different definitions
  - Leads to confusion
  - Described as "**apparent intensity**"
  - Used to label achromatic (luminance) channel in visual processing
- Actual definition
  - Effectiveness of light of different wavelengths in specific photometric matching task

© CRC Press, Wu & Rao Eds, 2006

2.4.1 Luminance

- CIE functional definition of luminance

$$L = k \int_{\lambda} P_{1\lambda} V_{\lambda}(\lambda) d\lambda$$

Where *k* is a constant, *P1λ* and *V(λ)* are the radiance and photopic spectral luminous efficiency function

V(λ) = ratio of radiant flux of wavelengths λ_m and λ

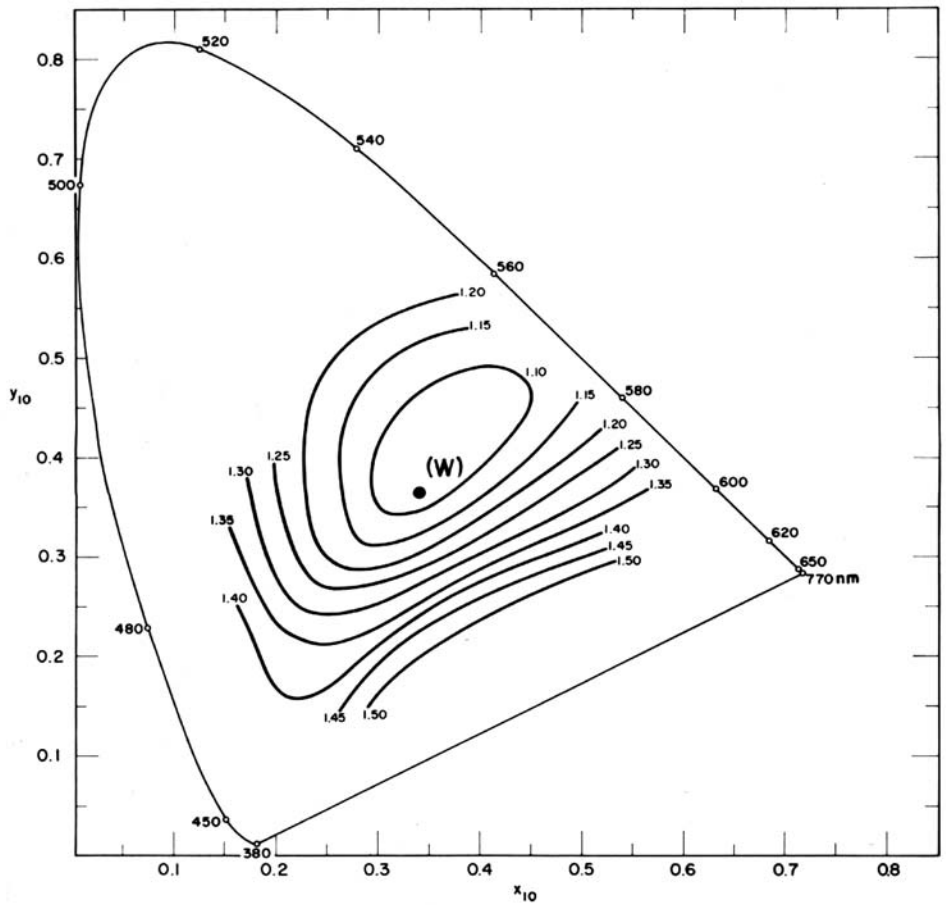
- Luminance measured in candelas (cd/m^2)
- CIE heterochromatic brightness matching
  - Additive (Abney's) law fails
    - Brightness not always additive in coloured stimulus

© CRC Press, Wu & Rao Eds, 2006

- Alternate matching methods
  - Heterochromatic Flicker Photometry (HFP)
    - Luminance intensity adjusted to match test and flickering light
  - Minimally Distinct Border (MDB)
    - Intensity adjusted to minimize the distinctness between the borders of bipartite luminance fields
  - HFP and MDB obeys additivity rule

© CRC Press, Wu & Rao Eds, 2006

Helmholtz-Kohlrause effect. The apparent isoluminant chromatic colours increases with chroma.



© CRC Press, Wu & Rao Eds, 2006

- Luminance efficiency function
  - Relative to the sensitivity of cone types
  - Dependent on adaptation, esp., chromatic
    - Can only be defined under conditions for which it is measured
  - HFP and MDB
    - Luminosity taken as weighted sum of L and M cones
    - Contribution of S cones is negligible under normal conditions
  - Affected colour blindness (L and M cones)

© CRC Press, Wu & Rao Eds, 2006

2.4.2 Perceived Intensity

- Weber's Law : $\Delta I / I = k$
  - General rule of thumb
  - Measures contrast (relative intensity)
  - States that the ratio of background (*I*) and incremental (*ΔI*) intensities is constant (*k*)
  - As *I* increase, *ΔI* also increase to meet detection threshold – Just Noticeable Difference (JND)
  - Does not hold for full range of intensities
  - Reliable for intensities above 100 cd/m^2

© CRC Press, Wu & Rao Eds, 2006

- Fechner's Law
  - Logarithmic relationship between stimulus intensity and associated sensation, S

$$S = k \log(I)$$

I is the intensity

k is a constant dependent on stimulus condition

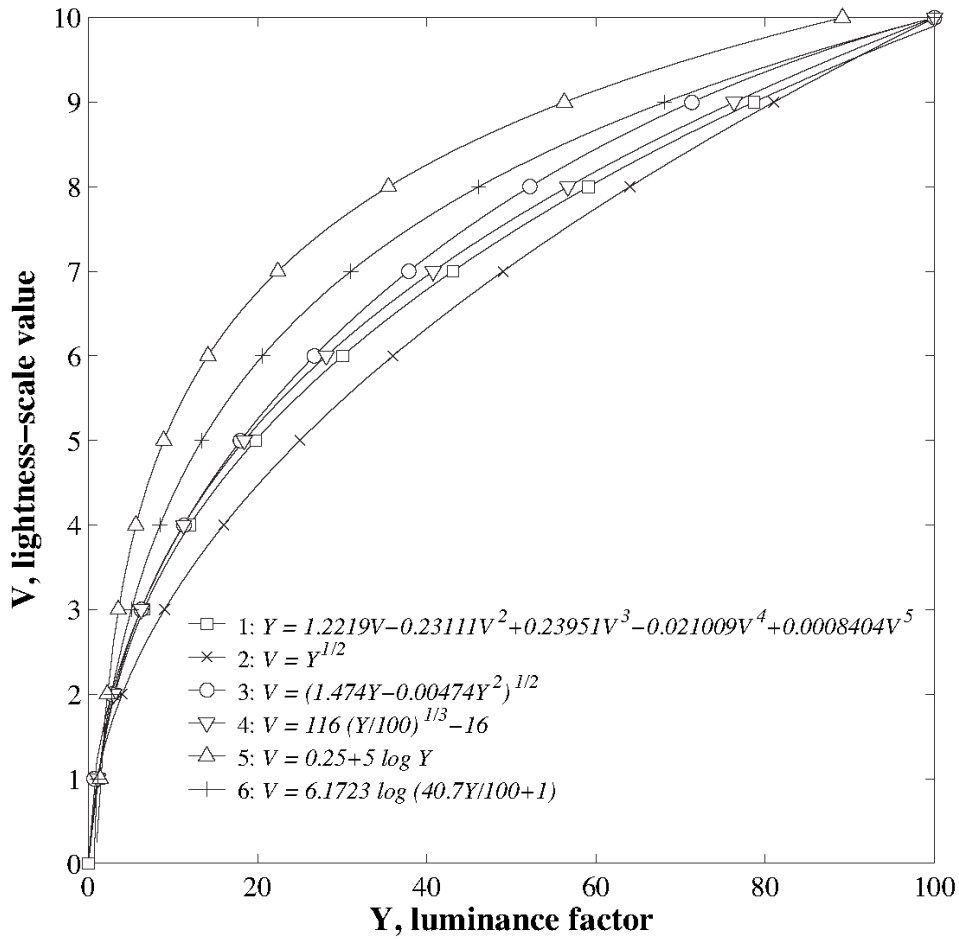
  - Does not match data from direct brightness experiments
    - Data exhibit power function profile: $S = k I^a$

© CRC Press, Wu & Rao Eds, 2006

- luminance and lightness relationship
  - Non-linear general profile
  - Specific profile depends on
    - Exact stimulus conditions
    - Data collection methods

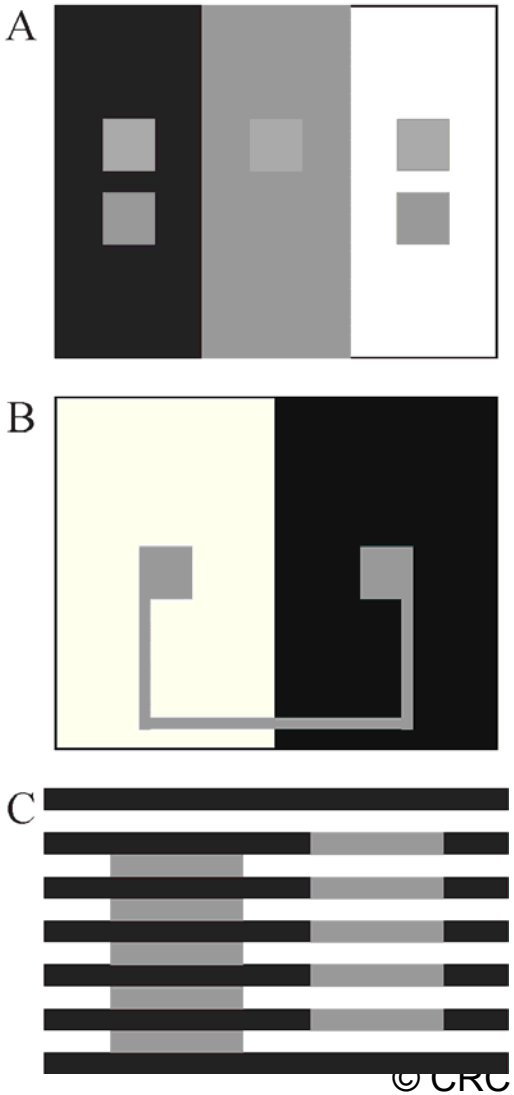
© CRC Press, Wu & Rao Eds, 2006

Luminance vs Lightness



© CRC Press, Wu & Rao Eds, 2006

Lightness illusions



A. Simultaneous contrast
B. Crispening
C. White's illusion

© CRC Press, Wu & Rao Eds, 2006

2.5 Spatial Vision and Contrast Sensitivity

- Spatial information processing
  - Visual acuity of the human optics
    - Modelled by
      - Modulation transfer function (MTF)
    - Able to detect spatial variation and spatial alignment (Vernier acuity or hyperacuity)
  - Sampling by photoreceptor array
    - Sampling rate of 120 samples/degree → nyquist frequency of 60 cpd (cycles per degree)
    - Optics degrades images above 60 cpd, no aliasing

© CRC Press, Wu & Rao Eds, 2006

2.5.2 Contrast Sensitivity

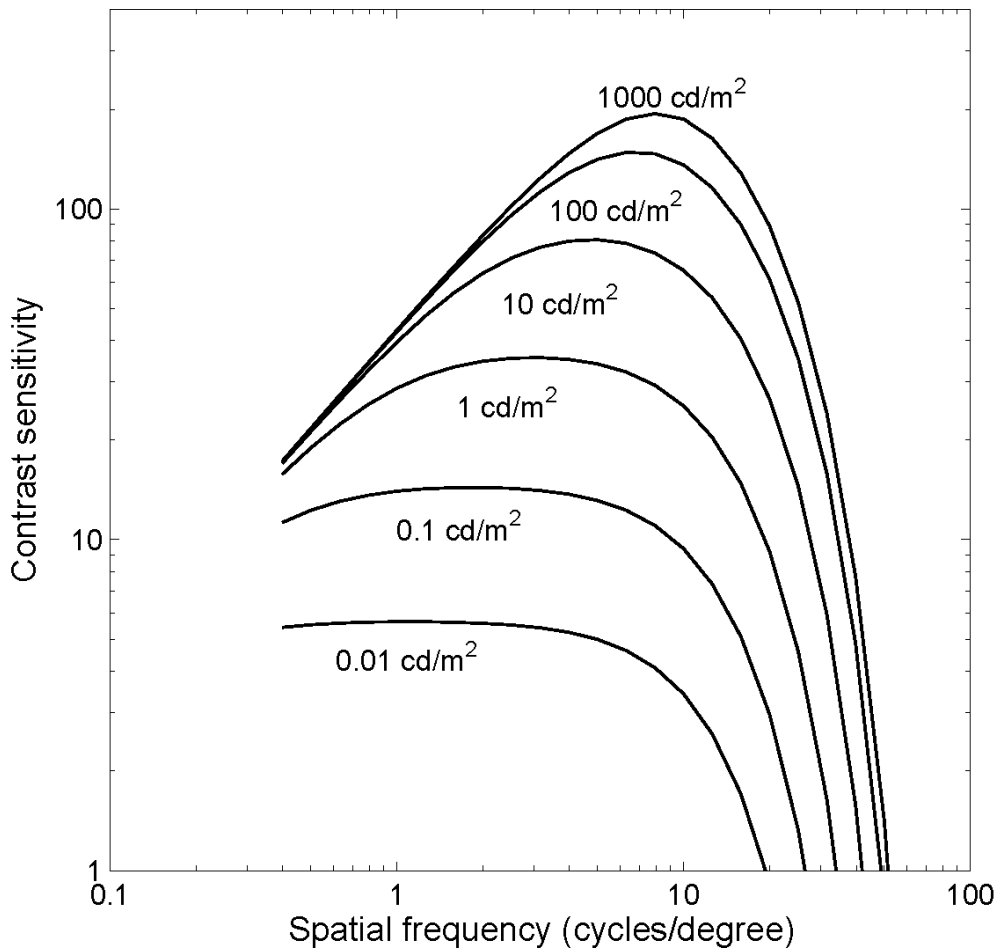
- Studied with harmonic stimuli (sine wave)
- Definition
  - Reciprocal of the threshold needed to detect stimuli of different frequencies
- Contrast sensitivity function
  - Measures visual system sensitivity to varying sinusoidal frequencies
- Michelson's Contrast

$$C_m = (L_{max} - L_{min})/(L_{max} + L_{min})$$

L_{max} and L_{min} are peak and trough luminance

© CRC Press, Wu & Rao Eds, 2006

Contrast Sensitivity Function for different mean luminance (Barten's Empirical Model)



© CRC Press, Wu & Rao Eds, 2006

- Contrast Sensitivity depends on
  - Mean luminance
  - Orientation
  - Temporal frequency
  - Spatial position in retina
    - Maximal sensitivity at the fovea
- Oblique effect
  - Reduced contrast sensitivity for obliquely oriented gratings (2-3 times)

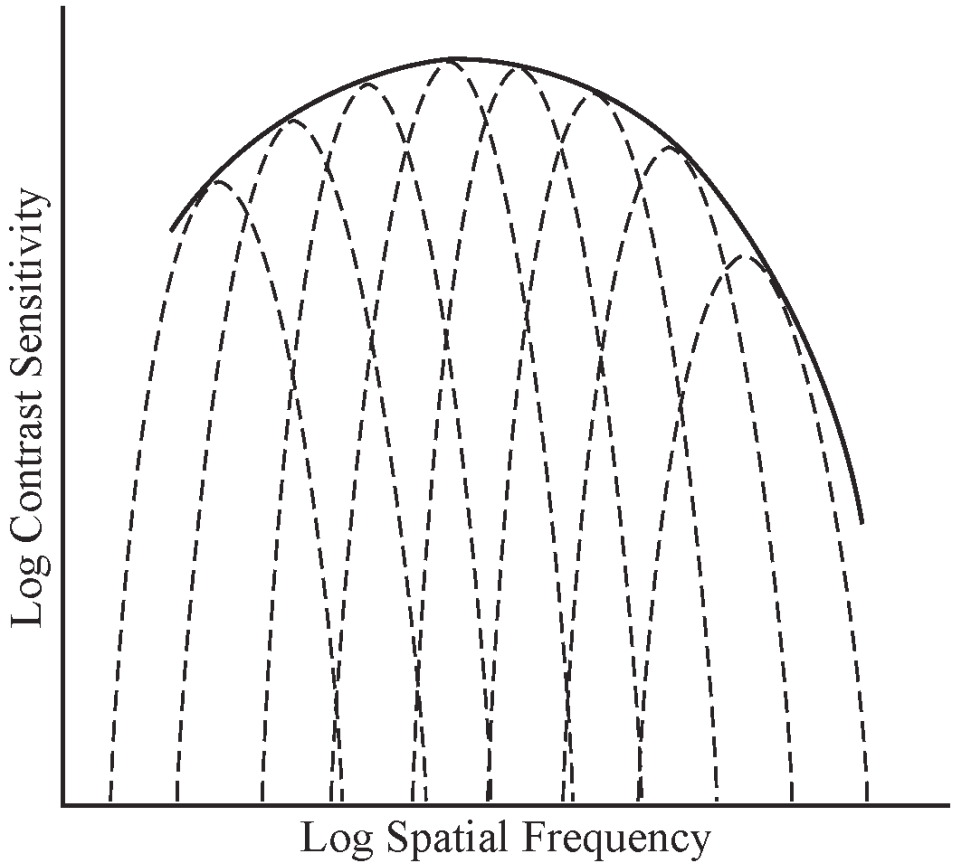
© CRC Press, Wu & Rao Eds, 2006

2.5.3 Multiple Spatial Frequency Channels

- CSF represent multiple narrowly tuned spatial frequency channels
  - Supported by physiological, psychophysical and anatomical evidence
- Concept of multi-channel structure for human vision
  - narrow frequency bands
  - narrow orientation bands

© CRC Press, Wu & Rao Eds, 2006

Envelop of multiple CSFs tuned to narrow spatial frequencies channels



© CRC Press, Wu & Rao Eds, 2006

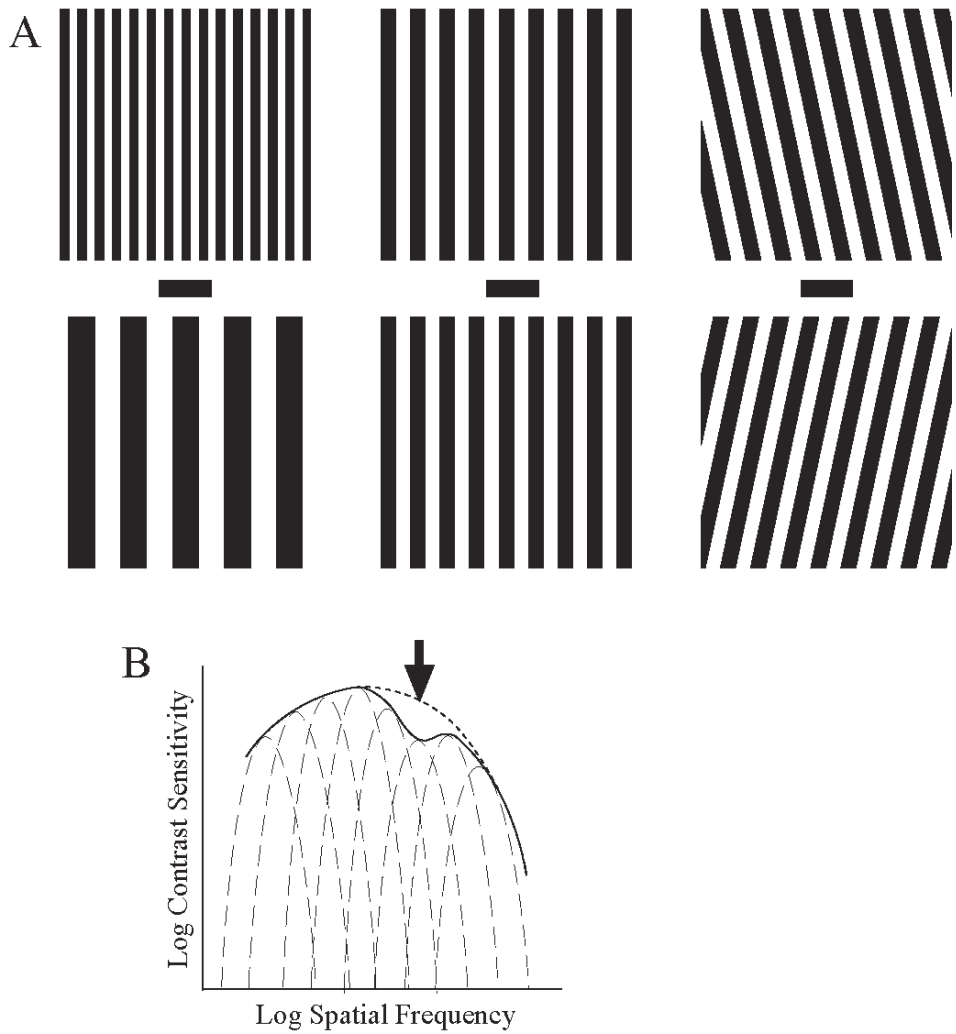
- Cortical receptive fields
  - Are size specific
  - Indicative of multi-resolution structure
- Frequency bandwidths
  - broader at low frequencies
  - Narrower at high frequencies

© CRC Press, Wu & Rao Eds, 2006

- Pattern Adaptation
  - Vision system takes inputs from multiple spatial frequency channels
  - Channels most sensitive to adaptation stimuli responds less vigorously
  - Once adaptation sets in, the distribution of inputs responses is skewed
  - Also affects orientation channels

© CRC Press, Wu & Rao Eds, 2006

Pattern and orientation adaptation. *A)* adaptation test. *B)* skewed sensitivity due to adaptation

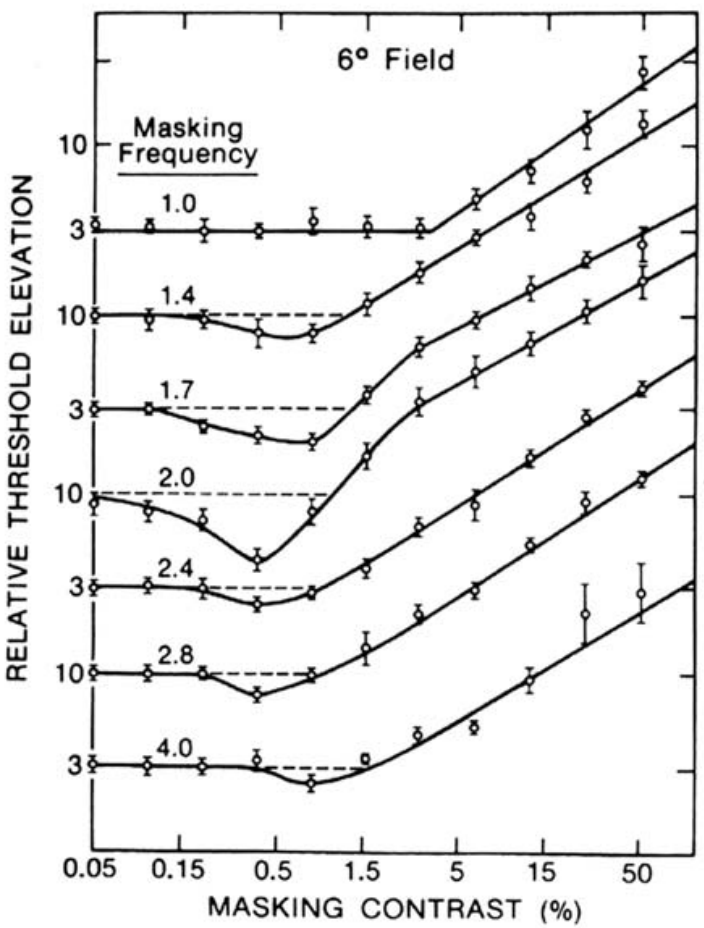


© CRC Press, Wu & Rao Eds, 2006

- Pattern Detection
  - Dependent on contrast of fundamental components, not overall contrast
- Masking and Facilitation
  - Masking hinders detection
  - Facilitation aids detection.
  - Depends spatial frequency, orientation and phase

© CRC Press, Wu & Rao Eds, 2006

Masking and facilitation



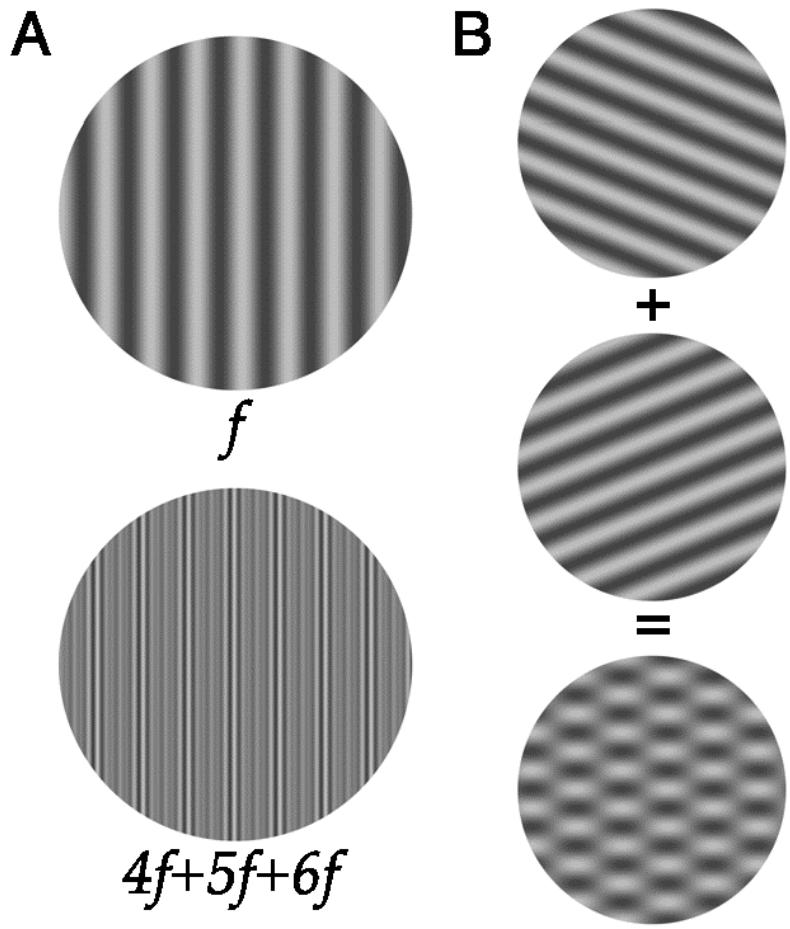
- Target frequency 2.0
- “dips” on curves signifies facilitation
- Facilitation is strongest at low contrast when target and masking frequencies are closest

© CRC Press, Wu & Rao Eds, 2006

- Dependencies in spatial frequencies and orientation
  - Non-linear interactions between separate spatial frequency and orientation bands exist
  - Physiological and psychophysical evidence supports inter-channel dependencies
  - However, Some Vision Models assumes independence between spatial frequency and orientation channels

© CRC Press, Wu & Rao Eds, 2006

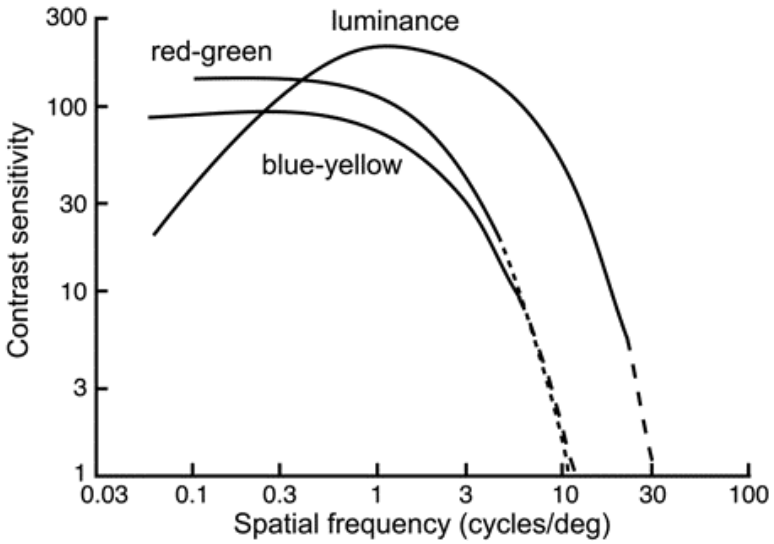
Inter-channel dependencies



© CRC Press, Wu & Rao Eds, 2006

Chromatic contrast sensitivity

- Opponent colour
- Chromatic and achromatic profile differs
- Luminance contrast operates at higher frequencies

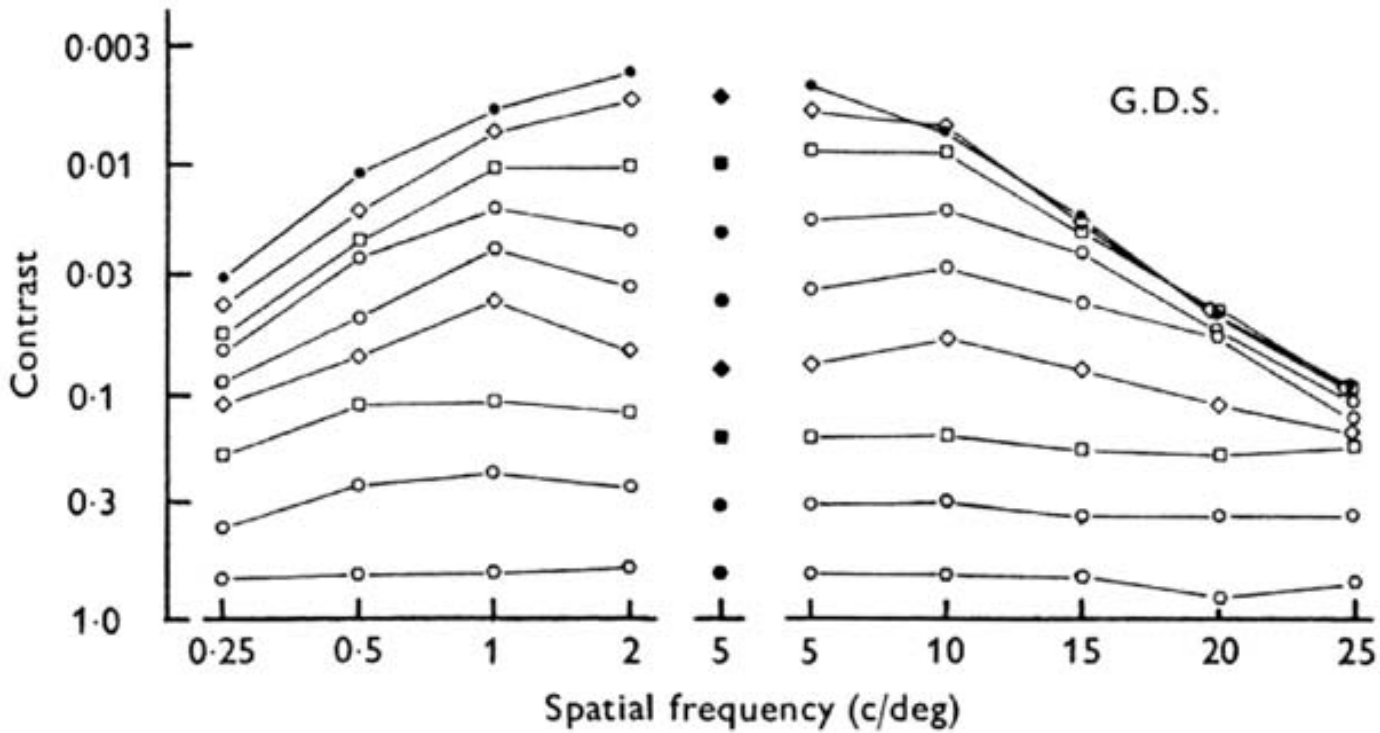


© CRC Press, Wu & Rao Eds, 2006

- Supra-threshold contrast sensitivity
  - Effects of threshold and supra-threshold are different
  - Flattening effect or “non-constancy”
    - Occurs as contrast of gratings exceed far beyond that of the threshold levels
- Hypothesis (Georgeson and Sullivan)
  - Spatial frequency channels adjust gains independently to operate at supra-threshold

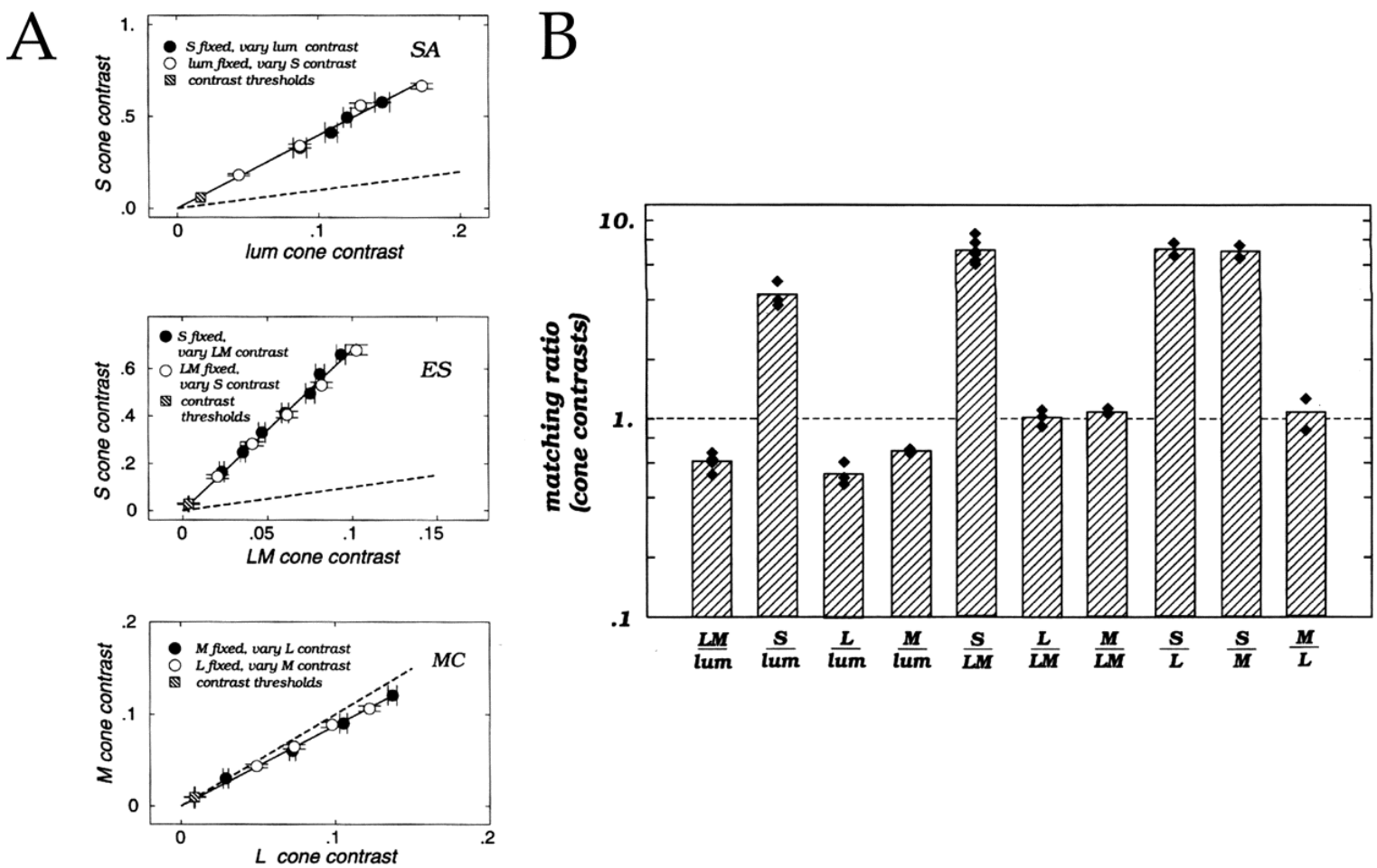
© CRC Press, Wu & Rao Eds, 2006

Supra-threshold contrast matched to 5 cpd gratings



© CRC Press, Wu & Rao Eds, 2006

Contrast matching. A) achromatic and chromatic matching. B) Contrast ratio of colour matching.



© CRC Press, Wu & Rao Eds, 2006

- Image Compression and Image Difference
  - Incomplete understanding of human vision
  - Have some operational knowledge of visual system
    - Improve image processing/compression system
      - Perceptually lossless coding
      - Perceptual coding (JPEG2000)
      - S-CIELAB computes visible colour difference

© CRC Press, Wu & Rao Eds, 2006

2.6 Temporal Vision and Motion

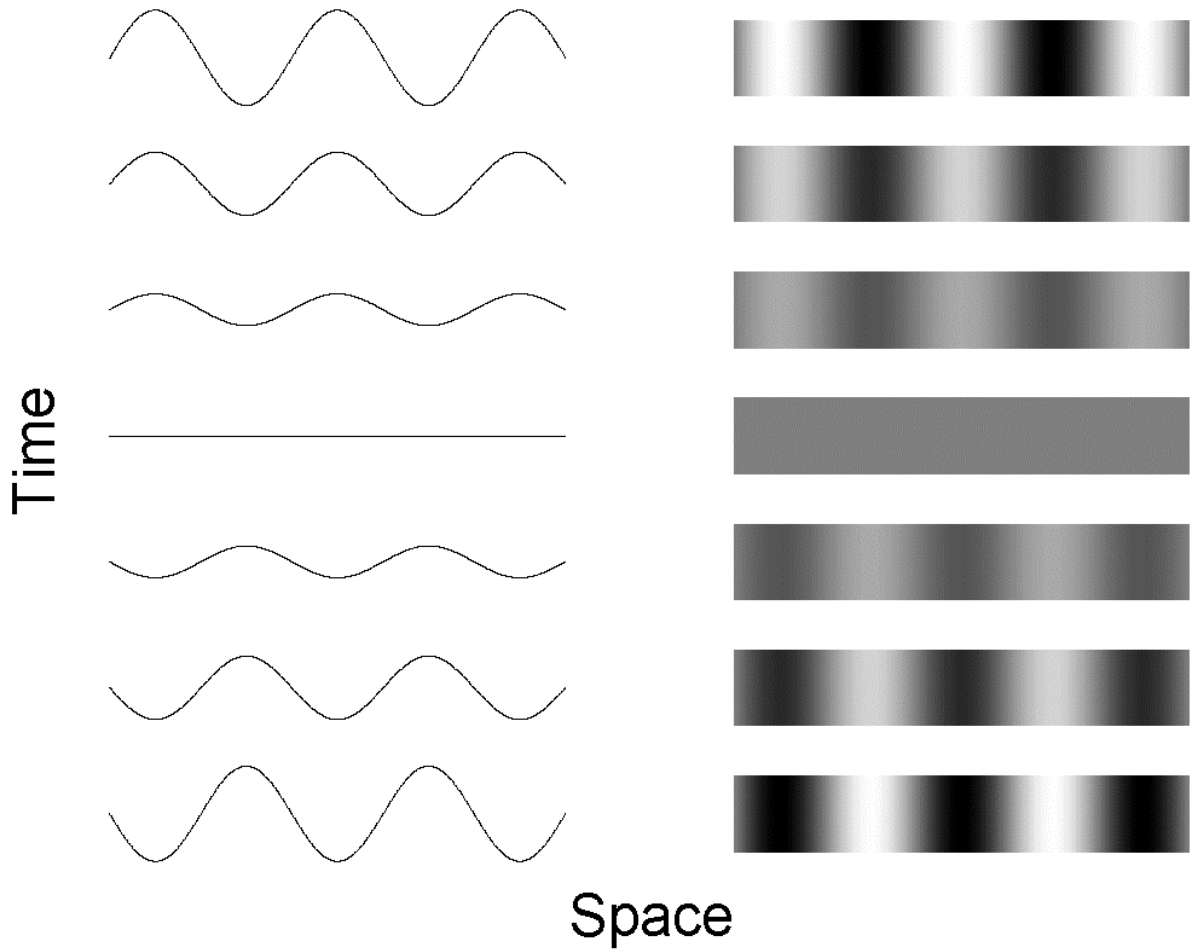
- Temporal CSF
  - Critical Flicker Frequency (CFF)
  - Ferry-Porter Law
    - T-CSF increase linearly with average background intensity (in log scale)
  - Talbot-Plateau Law
    - Perceived intensity of a fused periodic stimulus is equivalent to a steady stimulus of the same time-average intensity

© CRC Press, Wu & Rao Eds, 2006

- Counter phase grating for experiments
  - Contrast reverses over time
- Temporal contrast threshold for detection, a function of
  - temporal frequency of the modulation
  - spatial frequency of the pattern

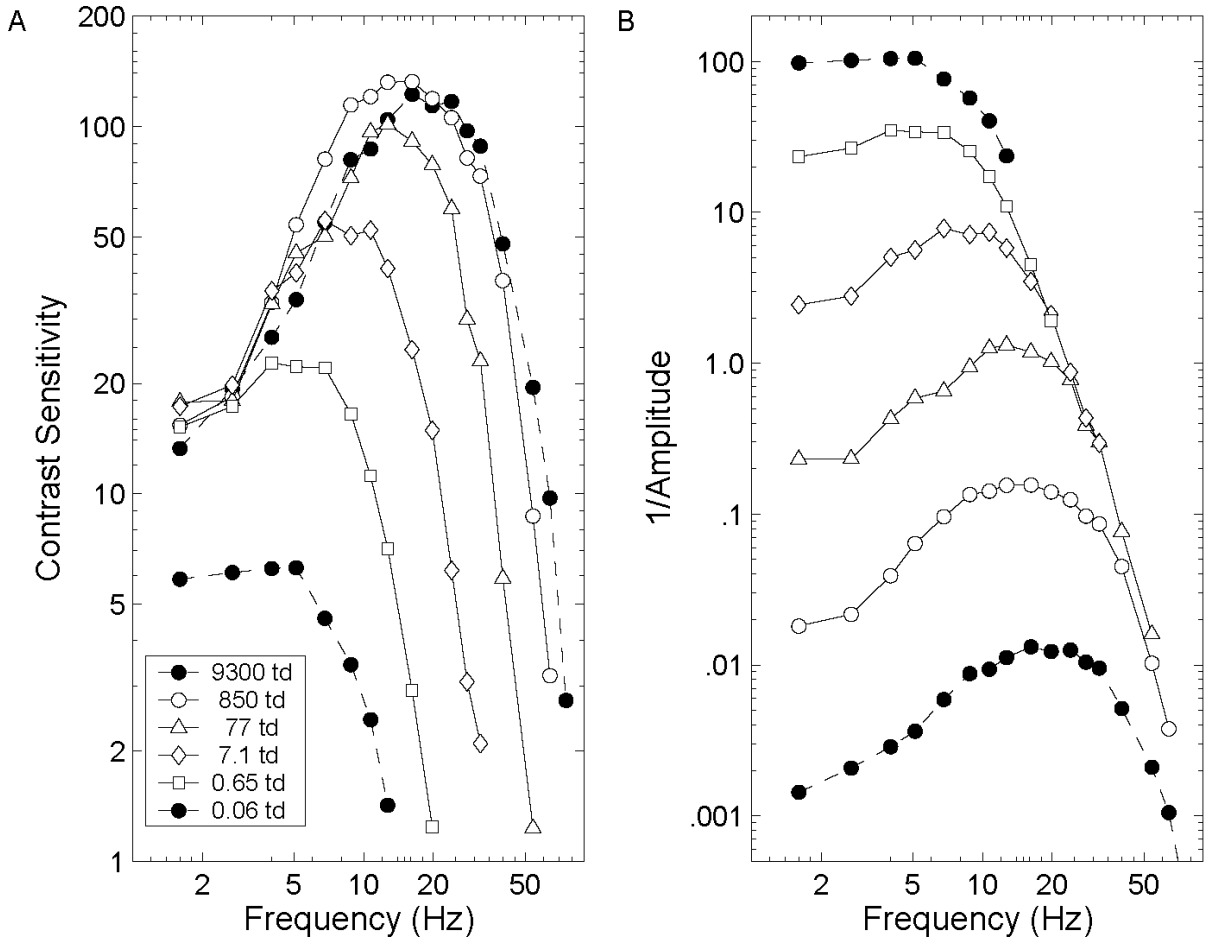
© CRC Press, Wu & Rao Eds, 2006

Temporally counter phased grating



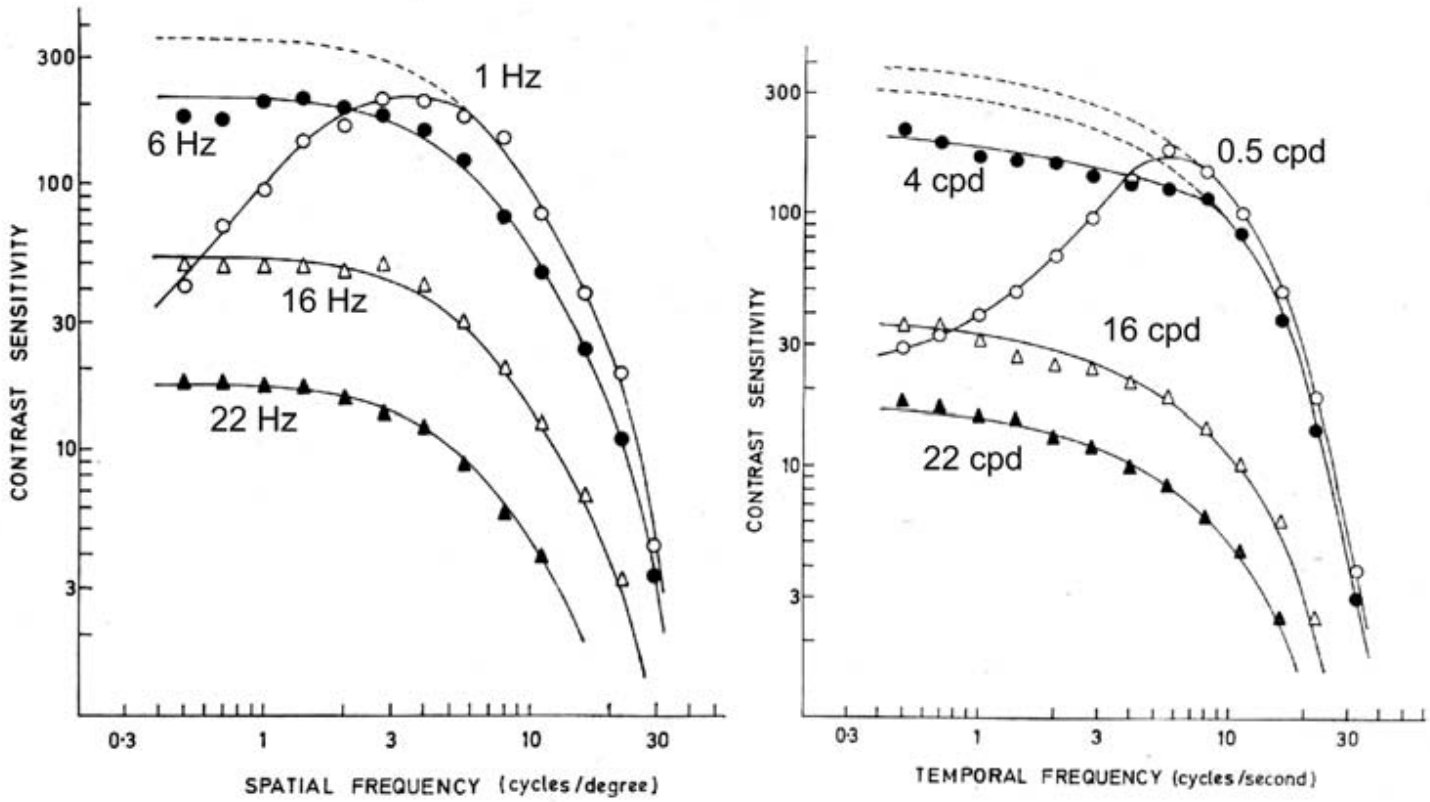
© CRC Press, Wu & Rao Eds, 2006

Temporal CSF with mean luminance measured in trolands



© CRC Press, Wu & Rao Eds, 2006

Spatial contrast sensitivity at various temporal frequencies (left) and vice versa (right)



© CRC Press, Wu & Rao Eds, 2006

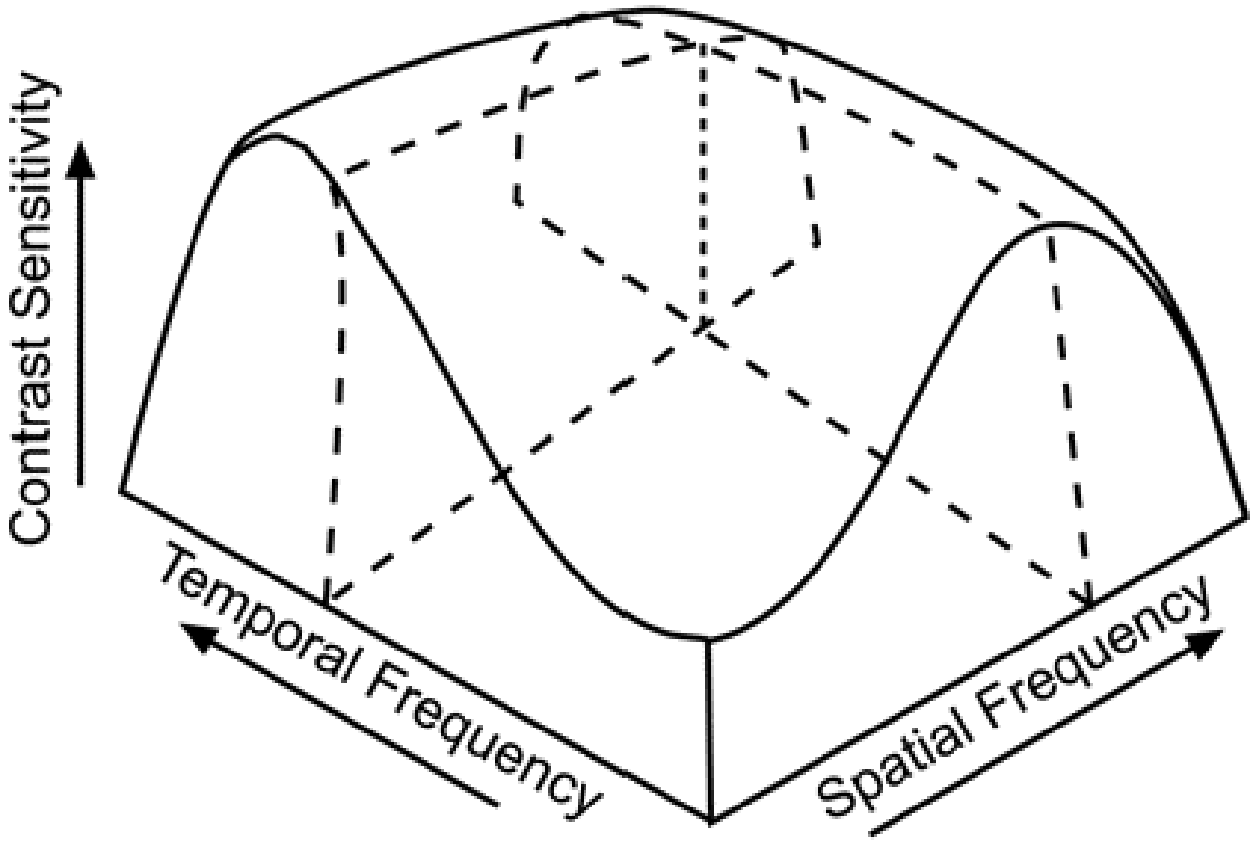
- Apparent Motion
  - motion after-effect due to adaptation
    - e.g. the moon can appear to move behind clouds
- Inter-Stimulus Interval (ISI)
  - Interval between stimulus
  - Short ISI (< 30 msec), no motion observable
  - Perceived motion between 30 – 60 msec (ISI)
  - Continuous motion with > 100 msec (ISI)

© CRC Press, Wu & Rao Eds, 2006

- Korte's Law
  - Motion Detection
    - higher contrast for larger positional change
    - Higher contrast for shorter ISI
    - Larger positional separations for long ISI

© CRC Press, Wu & Rao Eds, 2006

Spatio-temporal contrast sensitivity



© CRC Press, Wu & Rao Eds, 2006

2.7 Visual Modelling

- Discoveries made in small steps
- Purpose of vision modelling
  - Automate tasks
  - Improve reproduction of real world images
- Applications
  - Image/video processing
    - Quality/fidelity measures
  - Image/video compression
    - Improve quality and/or compression performance

© CRC Press, Wu & Rao Eds, 2006

2.8 Conclusion

- Overview of the visual system
- Colour vision
  - Colorimetry, colour appearance
- Luminance and intensity
- Spatial vision
  - Contrast sensitivity
- Temporal vision

© CRC Press, Wu & Rao Eds, 2006

And03	B.L. Anderson. Perceptual organization and white's illusion. *Perception*, 32(3):269-284, 2003
ATI01	ATIS. Objective perceptual video quality measurements using a JND-based full reference technique. Technical Report T1.TR.PP.75-2001, Alliance for Telecommunications Industry Solutions Technical Report, 2001
Bar99	P.G.J. Barten. *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. Bellingham, WA: SPIE Press, 1999
Bar04	P.G.J. Barten. Formula for the contrast sensitivity of the human eye. In Y. Miyake and D.R. Rasmussen, Eds., *Image Quality and System Performance, Proceedings of SPIE-IS&T Electronic Imaging*, 5294:231-238, 2004
BC69	C. Blakemore and F.W. Campbell. On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *J. Physiology*, 203:237-260, 1969
Ber00	R.S. Berns. *Billmeyer and Salzman's Principles of Color Technology*. New York: John Wiley & Sons, Inc., 2000

© CRC Press, Wu & Rao Eds, 2006

BK80	C.A. Burbeck and D.H. Kelly. Spatiotemporal characteristics of visual mechanisms: Excitatory-inhibitory model. *J. Opt. Soc. Am.*, 70(9):1121-1126, 1980
Boy96	R.M. Boynton. History and current status of a physiologically based system of photometry and colorimetry. *J. Opt. Soc. Am. A*, 13(8):1609-1621, 1996
BW94	R.W. Bowen and H.R. Wilson. A two-process analysis of pattern masking. *Vision Research*, 34:645-658, 1994
CG66	F.W. Campbell and R.W. Gubisch. Optical quality of the human eye. *J. Physiology*, 186:558-578, 1966
CIE86	CIE No. 15.2. *Colorimetry*. Commission Internationale de l'Éclairage, Vienna, Austria, 2nd Ed., 1986
CR68	F.W. Campbell and J.G. Robson. Application of fourier analysis to the visibility of gratings. *J. Physiology*, 197:551-556, 1968
Dal93	S. Daly. The visible difference predictor: An algorithm for the assessment of image fidelity. In A.B. Watson, Ed., Digital Images and Human Vision, 179-206, Cambridge, MA: MIT Press, 1993

© CRC Press, Wu & Rao Eds, 2006

DD90	R.L. De Valois and K.K. De Valois. *Spatial Vision*. New York: Oxford University Press, 1990
de54	H. de Lange. Relationship between critical flicker frequency and a set of low frequency characteristics of the eye. *J. Opt. Soc. Am.*, 44:380-389, 1954
de58	H. de Lange. Relationship between critical flicker frequency and a set of low frequency characteristics of the eye. *J. Opt. Soc. Am.*, 48:777-784, 1958
DH89	A.M. Derrington and G.B. Henning. Some observations on the masking effects of two-dimensional stimuli. *Vision Research*, 28:241-246, 1989
Fai98	M.D. Fairchild. *Color Appearance Models*. Reading, MA: Addison-Wesley, 1998
FBM03	J. Fiser, P.J. Bex and W. Makous. Contrast conservation in human vision. *Vision Research*, 43(25):2637-2648, 2003
FC99	J.M. Foley and C.C. Chen. Pattern detection in the presence of maskers that differ in spatial phase and temporal offset: Threshold measurements and a model. *Vision Research*, 39:3855-3872, 1999

© CRC Press, Wu & Rao Eds, 2006

FJ02	M.D. Fairchild and G.M. Johnson. Meet iCAM: A next-generation color appearance model. In *Proceedings of IS&T/SID 10th Color Imaging Conference*, 33-38, Scottsdale, AZ, 2002
FJ03	M.D. Fairchild and G.M. Johnson. Image appearance modeling. In B.R. Rogowitz and T.N. Pappas, Eds., *Human Vision and Electronic Imaging VIII, Proceedings of SPIE-IS&T Electronic Imaging*, 5007:149-160, 2003
Gle93	W.E. Glenn. Digital image compression based on visual preception. In A.B. Watson, Ed., *Digital Images and Human Vision*, 63-72. Cambridge, MA: MIT Press, 1993
GN71	N. Graham and J. Nachmias. Detection of grating patterns containing two spatial frequencies: A comparison of single channel and multichannel models. *Vision Research*, 11:251-259, 1971
Gra89	N.V.S. Graham. *Visual Pattern Analyzers*. New York: Oxford University Press, 1989
GS75	M.A. Georgeson and G.D. Sullivan. Contrast constancy: Deblurring in human vision by spatial frequency channels. *J. Physiology*, 252(3):627-656, 1975

© CRC Press, Wu & Rao Eds, 2006

Hel11	H. von Helmholtz. *Treatise on Physiological Optics*. Rochester, New York: Optical Society of America, 1866/1911. Translated form 3rd German ed., J.P. Southall, Ed.
Her64	E. Hering. *Outlines of a Theory of the Light Sense*. Cambridge, MA: Harvard University Press, 1920/1964
HF86	D.C. Hood and M.A. Finkelstein. Sensitivity to light. In K.R. Boff, L. Kaufman and J.P. Thoma, Eds., *Handbook of Perception and Human Performance*, *Vol. 1: Sensory Processes and Perception*, (5-1)-(5-66), Vienna, 1963, 1986. New York: John Wiley & Sons.
JF02	G.M. Johnson and M.D. Fairchild. On contrast sensitivity in an image difference model. In *Proceedings of the IS&T PICS Conference*, 18-23, Portland, OR, 2002
JF03	G.M. Johnson and M.D. Fairchild. Measuring images: Difference, quality and appearance. In B.R. Rogowitz and T.N. Pappas, Eds., *Human Vision and Electronic Imaging VIII, Proceedings of SPIE*, 5007:51-60, 2003
JT75	R.M. Jones and U. Tulunay-Keesey. Local retinal adaptation and spatial frequency channels. *Vision Research*, 15:1239-1244, 1975

© CRC Press, Wu & Rao Eds, 2006

KN96	P.K. Kaiser and R.M. Boynton. *Human Color Vision*. Washington, DC: Optical Society of America, 2nd Ed., 1996
Kel61	D.H. Kelly. Visual responses to time-dependent stimuli, I. Amplitude sensitivity measurements. *J. Opt. Soc. Am.*, 51(4):422-429, 1961
Kel83	D.H. Kelly. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *J. Opt. Soc. Am.*, 73(11):1784-1793, 1983
Kel89	D.H. Kelly. Opponent-color receptive-field profiles determined from large-area psychophysical measurements. *J. Opt. Soc. Am. A*, 6(11):1784-1793, 1989
KNM84	S.W. Kuffler, J.G. Nicholls and A.R. Martin. *From Neuron to Brain*. Sunderland, MA: Sinauer Associates Inc., 2nd ed., 1984
LCR01	M.R. Luo, G. Cui and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Application*, 26(5):340-350, 2001
LF80	G.E. Legge and J.M. Foley. Contrast masking in human vision. *J. Opt. Soc. Am.*, 70(12):1458-1471, 1980

© CRC Press, Wu & Rao Eds, 2006

LR87	M.R. Luo an B. Rigg. BFD (1:c) colour-difference formula, Part 1 – Development of the formula. *J. Soc. Dyers. Colour*, 103:86-94, 1987
Lub93	J. Lubin. The use of psychophysical data and models in the analysis of display system performance. In A.B. Watson, Ed., *Digital Images and Human Vision*, 162-178. Cmabridge, MA: The MIT Press, 1993
Lub95	J. Lubin. A visual discrimination model for imaging system design and evaluation. In E. Peli Ed., *Vision Models for Target Detection and Recognition*, 245-283, Singapore: World Scientific, 1995
MF85	N.M. Moroney and M.D. Fairchild. Color space selection for JPEG image compression. *J. Electronic Imaging*, 4(4):373-381, 1985
MFH+02	N. Moroney, M.D. Fairchild, R.W.G. Hunt, C. Li, R. Luo and T. Newman. The CIECAM02 color appearance model. In *Proceedings of IS&T/SID 10th Color Imaging Conference*, 23-27, Scottsdale, AZ, 2002
MM93	W.H. Merigan and J.H.R. Maunsell. How parallel are the primate visual pathways. *Annual Review of Neuroscience*, 16:369-402, 1993

© CRC Press, Wu & Rao Eds, 2006

Mul85	K.T. Mullen. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *J. Physiology*, 359:381-400, 1985
Nic76	D. Nickerson. History of the Munsell color system. *Color Research and Application*, 1:121-130, 1976
Pal99	S.E. Palmer. *Vision Science: Photons to Phenomenology*. Cambridge, MA: The MIT Press, 1999
PS68	A. Pantle and R.W. Sekular. Size-detection mechanisms in human vision. *Science*, 162:1146-1148, 1968
PW84	G.C. Phillips and H.R. Wilson. Orientation bandwidths of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A*, 1(2):226-232, 1984
PW95	A.B. Poirson and B.A. Wandell. The appearance of colored patterns: Pattern-color separability. *J. Opt. Soc. Am. A*, 12:2458-2471, 1995
PW03	O. Packer and D.R. Williams. Light, the retinal image and photoreceptors. In S.K. Shevell. Ed., *The Science of Color*, 41-102. New York: Optical Society of America, 2004

© CRC Press, Wu & Rao Eds, 2006

Rob66	J.G. Robson. Spatial and temporal contrast-sensitivity functions of the visual system. *J. Opt. Soc. Am., 56(8):1141-1142, 1996*
SC99	E. Switkes and M.A. Crognale. Comparison of color and luminance contrast: Apples versus oranges? *Vision Research*, 39(10):1823-1832, 1999
SD83	E. Switkes and K.K. DeValois. Luminance and chromaticity interactions in spatial vision. In J.D. Mollon and L.T. Sharpe, Eds., *Color Vision*, 465-470. London: Academic Press, 1983
SS99	A. Stockman and L.T. Sharpe. Cone spectral sensitivities and color matching. In K.R. Gegenfurtner and L.T. Sharpe, Eds., *Color Vision*, 52-88. Cambridge, UK: Cambridge University Press, 1999
SW64	C.L. Sanders and G. Wyszecki. Correlate for brightness in terms of CIE color matching data. In *CIE Proceedings 15yh Session*, Paper P-63.6. CIE Central Bureau, Paris, 1964
SWB93a	N. Sekiguchi, D.R. Williams and D.H. Brainard. Aberration-free measurements of the visibility of isoluminant gratings. *J. Opt. Soc. Am. A*, 10(10):2105-2117, 1993

© CRC Press, Wu & Rao Eds, 2006

SWB93a	N. Sekiguchi, D.R. Williams and D.H. Brainard. Efficiency in detection of isoluminant and isochromatic interference fringes. *J. Opt. Soc. Am. A*, 10(10):2118-2133, 1993
VE03	D.C. Van Essen. Organization of the visual areas in macaque and human cerebral cortex. In L.M. Chalupa and J.S. Werner, Eds., *The Visual Sciences*, 507521. Cambridge, MA: The MIT Press, June 2003
Vim00	R.L.P. Vimal. Spatial color contrast matching: Broad-bandpass functions and the flattening effect. *Vision Research*, 40(23):3231-3244, 2000
WA85	A.B. Watson and A.J. Ahumada, Jr. Model of human visual-motion sensing. *J. Opt. Soc. Am. A*, 2(2):322-342, 1985
WAF86	A.B. Watson, A.J. Ahumada, Jr. and J.E. Farrell. Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. *J. Opt. Soc. Am. A*, 3(3):300-307, 1986
Wan95	B.A. Wandell. *Foundations of Vision*. Sunderland, MA: Sinauer Associates, Inc., 1995

© CRC Press, Wu & Rao Eds, 2006

Wat86	A.B. Watson. Temporal sensitivity. In K.R. Boff, L. Kaufman and J.P. Thoma Eds., *Handbook of Perception and Human Performance, Vol. 1: Sensory Processes and Perception*, (6-1) – (6-64). New York: John Wiley & Sons, 1986
Wat98	A.B. Watson. Toward a perceptual video quality metric. In *Human Vision and Electronic Imaging III*, 3299:139-147. Bellingham WA: SPIE Press, 1998
WBMN94	D.R. Williams, D.H. Brainard, M.J. McMahon and R. Navarro. Double-pass and interferometric measures of the optical quality of the eye. *J. Opt. Soc. Am. A*, 11(12):3123-3135, 1994
WHM01	A.B. Watson, J. Hu and F.J. McGowan. DVQ: A digital video quality metric based on human vision. *J. Electronic Imaging*, 10:20-29, 2001
Wil88	D.R. Williams. Topography of the foveal cone mosaic in the living human eye. *Vision Research*, 28(3):433-454, 1988
WLM+90	H.R. Wilson, D. Levi, L. Maffei, J. Rovamo and R. DeValois. The percfeption of form: Retina to striate cortex. In L. Spillman and J. Werner, Eds., *Visual Perception: The Neurophysiological Foundations*, 231-272. New York: Academic Press, 1990

© CRC Press, Wu & Rao Eds, 2006

WM77	G. Westheimer and J. McKee. Spatial configurations for visual hyperacuity. *Vision Research*, 17:941-947, 1977
WMP83	H.R. Wilson, D.K. McFarlane and G.C Phillips. Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23:873-882, 1983
WR84	H.R. Wilson and D. Regan. Spatial frequency adaptation and grating discrimination: Predictions of a line element model. *J. Opt. Soc. Am. A*, 1:1091-1096, 1984
WS82	G. Wyszecki and W.S Stiles. *Color Science Concepts and Methods, Quantitative Data and Formulae*, New York: John Wiley and Sons, 2nd ed., 1982
YM95	J. Yand and W. Makous. Modeling pedestal experiments with amplitude instead of contrast. *Vision Research*, 35(14):1979-1989. 1995
ZDL02	W. Zeng, S. Daly and S. Lei. An overview of the visual optimization tools in JPEG2000. *Signal Processing: Image Communication*, 17:85-104, 2002

© CRC Press, Wu & Rao Eds, 2006

Zek91	S. Zeki. Cerebral akinetopsia (visual motion blindness). *Brain*, 114(2):811-824, 1991
ZW88	X. Zhang and B.A. Wandell. Color image fidelity meetrics evaluated using image distortion maps. *Signal Processing*, 70(3):201-214, 1988
ZW96	X. Zhang and B.A. Wandell. A spatial extension to CIELAB for digital color image preproduction. In *Socienty for Information Display Symposium Technical Digest*, 27:731-734, 1996

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

3. Coding Artifacts and Visual Distortions

M. Yuen

ESS Technologies Inc., China

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 3 Outline

3.1 Introduction

3.2 Blocking Effect

3.3 Basis Image Effect

3.4 Blurring

3.5 Colour Bleeding

3.6 Staircase Effect

3.7 Ringing

© CRC Press, Wu & Rao Eds, 2006

Chapter 3 Outline

3.8 Mosaic Patterns

3.9 False Contouring

3.10 False Edges

3.11 MC Mismatch

3.12 Mosquito Effect

3.13 Stationary Area Fluctuation

3.14 Chrominance Mismatch

© CRC Press, Wu & Rao Eds, 2006

Chapter 3 Outline

3.15 Video Scaling and Field Rate Conversion

3.16 De-interlacing

3.17 Summary

© CRC Press, Wu & Rao Eds, 2006

3.1 Introduction

- Hybrid MC/DPCM/DCT coding
  - Classification and evaluation coding artifacts
  - Leads to improvement of coding systems
    - Constant quality video
- Artifacts identification
  - Attributes
    - Visual manifestation
    - Causes
    - Relationships

© CRC Press, Wu & Rao Eds, 2006

- Classification limited to MPEG-1 artifacts
  - Derived through various coding rates
- Artifacts are difficult to isolate
  - unable specify coding rate which leads to the manifestation of any particular artifact
  - Multiple artifacts may co-exist together

© CRC Press, Wu & Rao Eds, 2006

3.2 Blocking Effect

- Inherent in block based (DCT) coding
- Visual Characteristic
  - Discontinuity at block boundary
- Cause
  - Quantization of DCT coefficients
- Intra-frame mode
  - Prominent smoothly textured regions
  - DC coefficients has highest impact
  - Coarse quantization of AC coefficients

© CRC Press, Wu & Rao Eds, 2006

Example of Blocking Effect

© CRC Press, Wu & Rao Eds, 2006

- Inter-frame mode
  - External blocking
    - Boundary of 16x16 macro-block
    - Occurs around moving objects
      - Resulting from poor MC, block mismatch
  - Internal blocking
    - Boundaries of 8x8 DCT blocks within macro-block
    - Generally appears in mildly textured area with sufficiently high prediction error – errors not quantized to zero

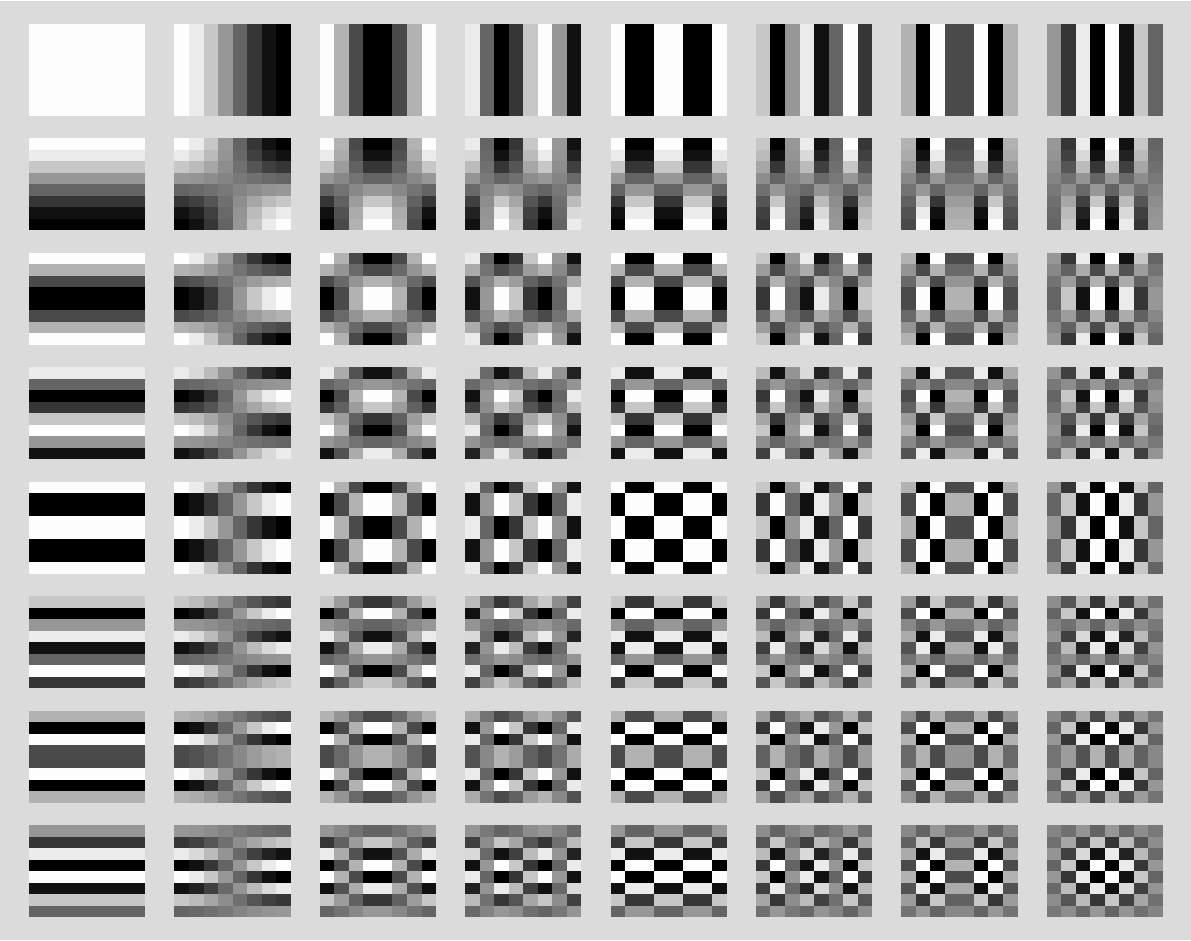
© CRC Press, Wu & Rao Eds, 2006

3.3 Basis Image Effect

- Block bearing distinct DCT basis image
  - AC basis with horizontal and vertical patterns
- Cause
  - Coarse quantization of low magnitude DCT coefficients
- Relationship
  - May lead to other artifacts
    - Blocking and mosaics

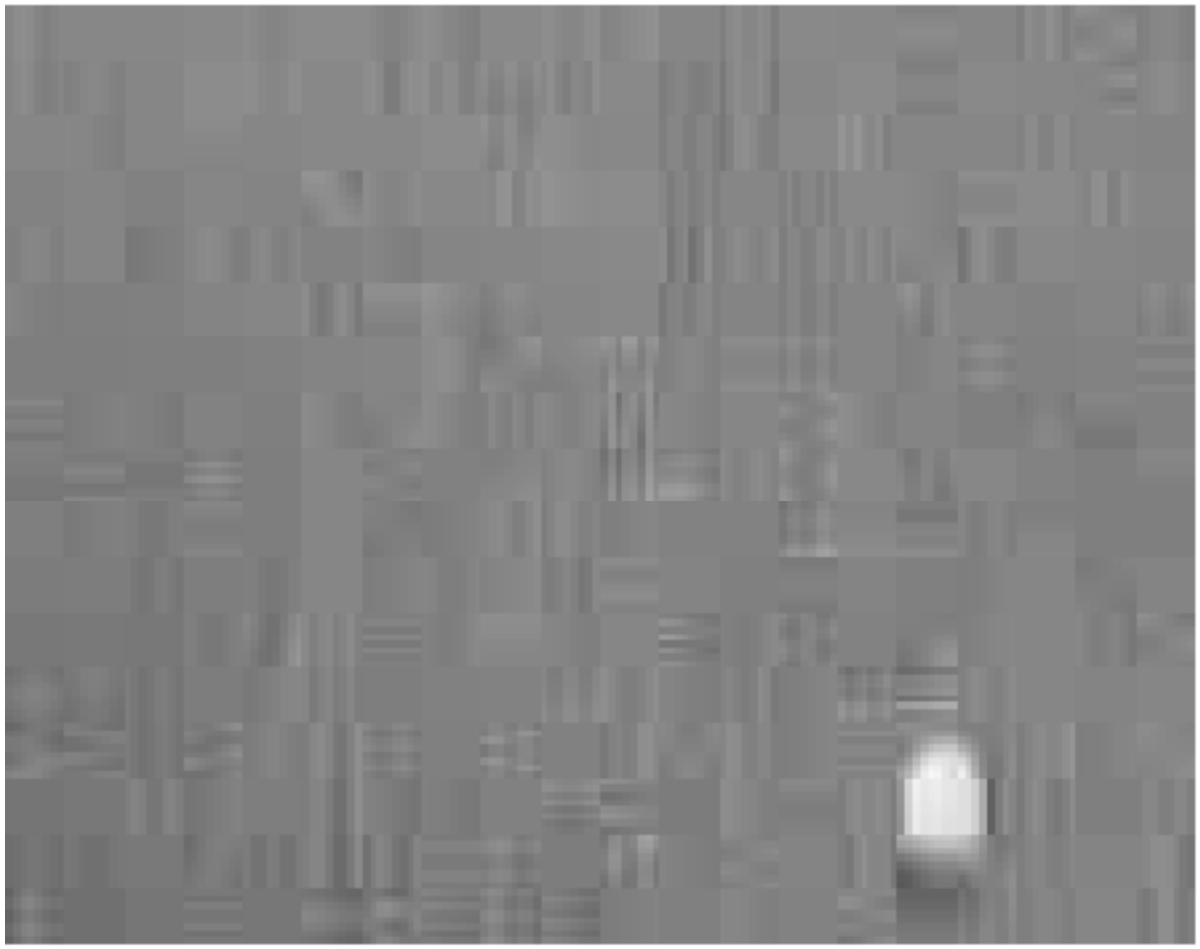
© CRC Press, Wu & Rao Eds, 2006

DCT Basis Function



© CRC Press, Wu & Rao Eds, 2006

DCT Basis Image Effect



© CRC Press, Wu & Rao Eds, 2006

- Visual significance of basis function
  - Different basis has different significance
  - Proportional to AC energy concentration in each coefficient
- Inter-frame coding
  - Appears in high spatial activity areas due to coarse quantization

© CRC Press, Wu & Rao Eds, 2006

3.4 Blurring

- Loss of spatial detail
- Cause in DCT coding
  - Coarse quantization of mid to high frequency AC coefficients
- Associated artifacts
  - Blocking and mosaic effects if sufficient spatial information has been blurred
- Inter-frame coding
  - Prediction from macro-blocks with spatially degraded details

© CRC Press, Wu & Rao Eds, 2006

3.5 Colour Bleeding

- Effectively colour blurring
  - Smearing of colour over areas of strong contrasting chrominance
- Cause
  - Coarse quantization of AC components
- Manifestation extends to macro-block boundaries
  - Due to sub-sampling of chroma channels (4:2:2 or 4:2:0)

© CRC Press, Wu & Rao Eds, 2006

Colour bleeding in coded (0.6Mbps) chroma channels

Top left: C_b channel
Bottom left: C_r channel
Right: Decoded frame

© CRC Press, Wu & Rao Eds, 2006

3.6 Staircase Effect

- Primary reason - no diagonal DCT basis function
- Cause by coarse quantization of AC components associated with diagonal edges
- Related to blocking and mosaic
- Appears around block boundaries

© CRC Press, Wu & Rao Eds, 2006

Staircase Effects

© CRC Press, Wu & Rao Eds, 2006

3.7 Ringing

- Occurs along high contrast edges in areas of generally smooth texture
- Appears as ripples moving outwards from the edge
- Chrominance Ringing
  - Related to colour bleeding/mismatch
  - Spans whole macro-block

© CRC Press, Wu & Rao Eds, 2006

Ringing Artifact: along the edge of the table and the arm

© CRC Press, Wu & Rao Eds, 2006

Low Ringing: *Claire* sequence @ 1Mbps. PSNR: approximately 45 dB

© CRC Press, Wu & Rao Eds, 2006

3.8 Mosaic Patterns

- Mismatch between all or parts of adjacent blocks
  - Analogy from mismatch of mosaic tiles
- Cause
  - Coarse quantization of AC components
  - Subsequent contour and/or texture difference between adjacent blocks

© CRC Press, Wu & Rao Eds, 2006

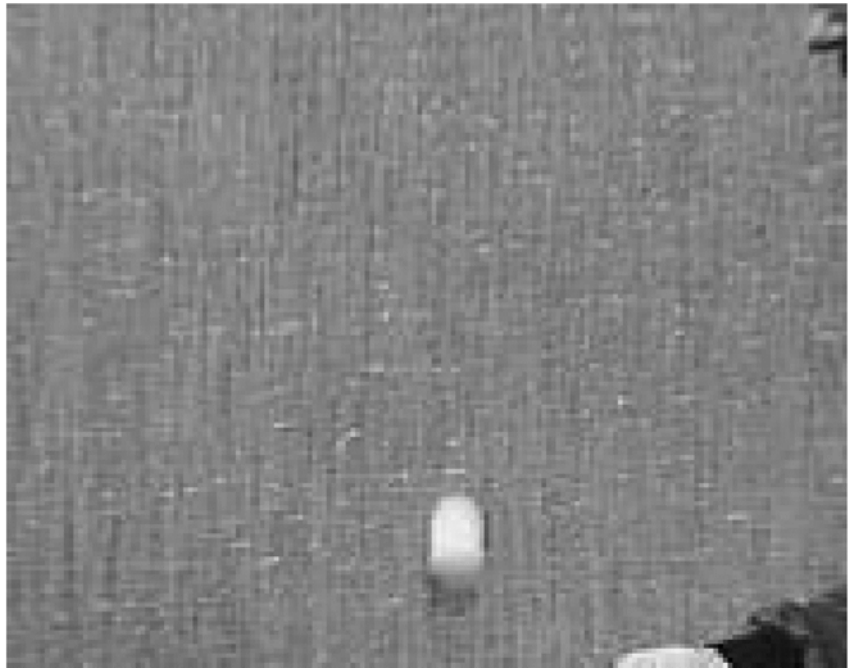
Mosaic Effect post de-blocking filter

© CRC Press, Wu & Rao Eds, 2006

- Intra-frame mode
  - Generally occurs in areas of high spatial activity
  - Related to basis image effect
  - Mosaic effect more pronounced if mismatched patterns have different orientations

© CRC Press, Wu & Rao Eds, 2006

Mosaic effect and basis image effect. *left:* original frame. *Right:* coded frame

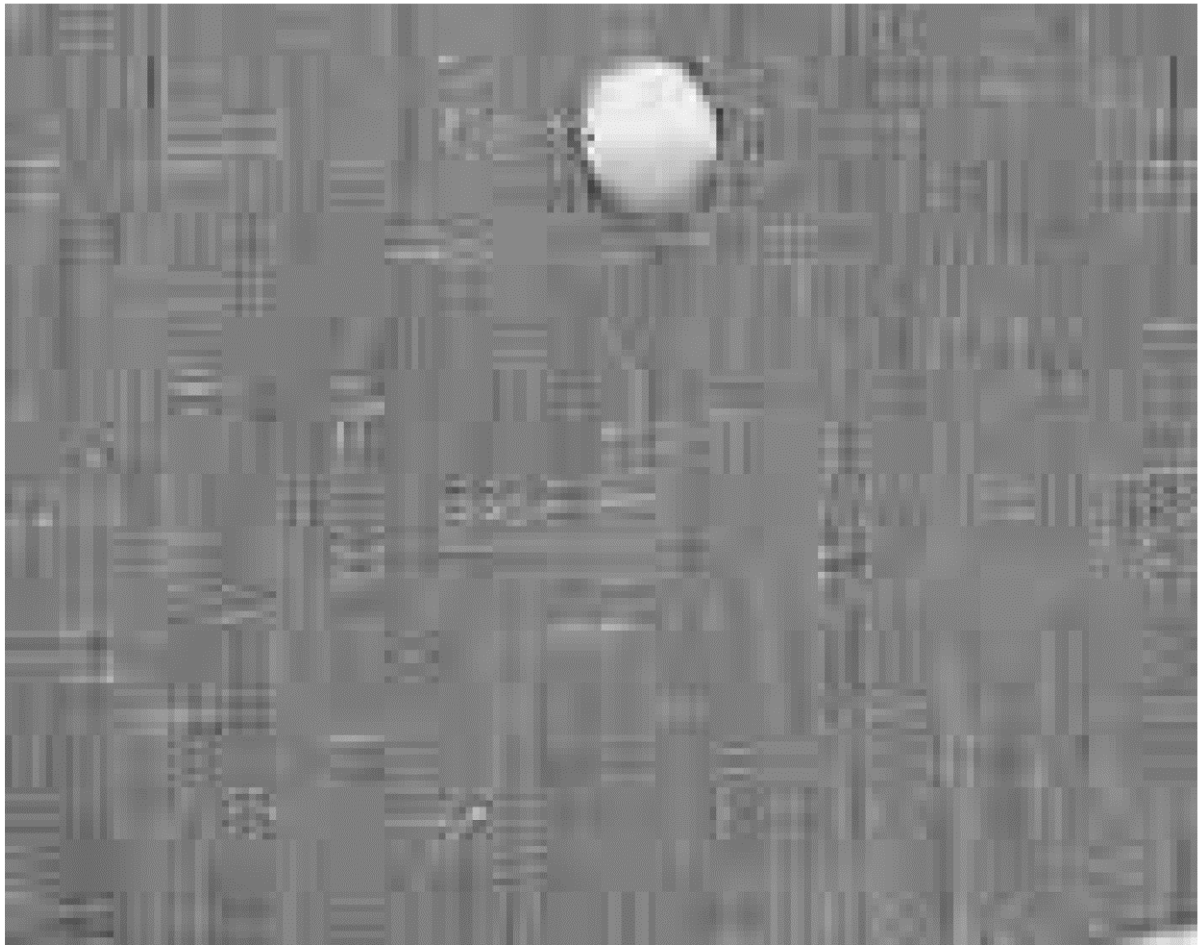


© CRC Press, Wu & Rao Eds, 2006

- Inter-frame mode
  - MC mismatch
  - Occurs in high spatial activity areas
  - Successive prediction affects the degree of the mosaic effect
    - Successive prediction improves/refine spatial information

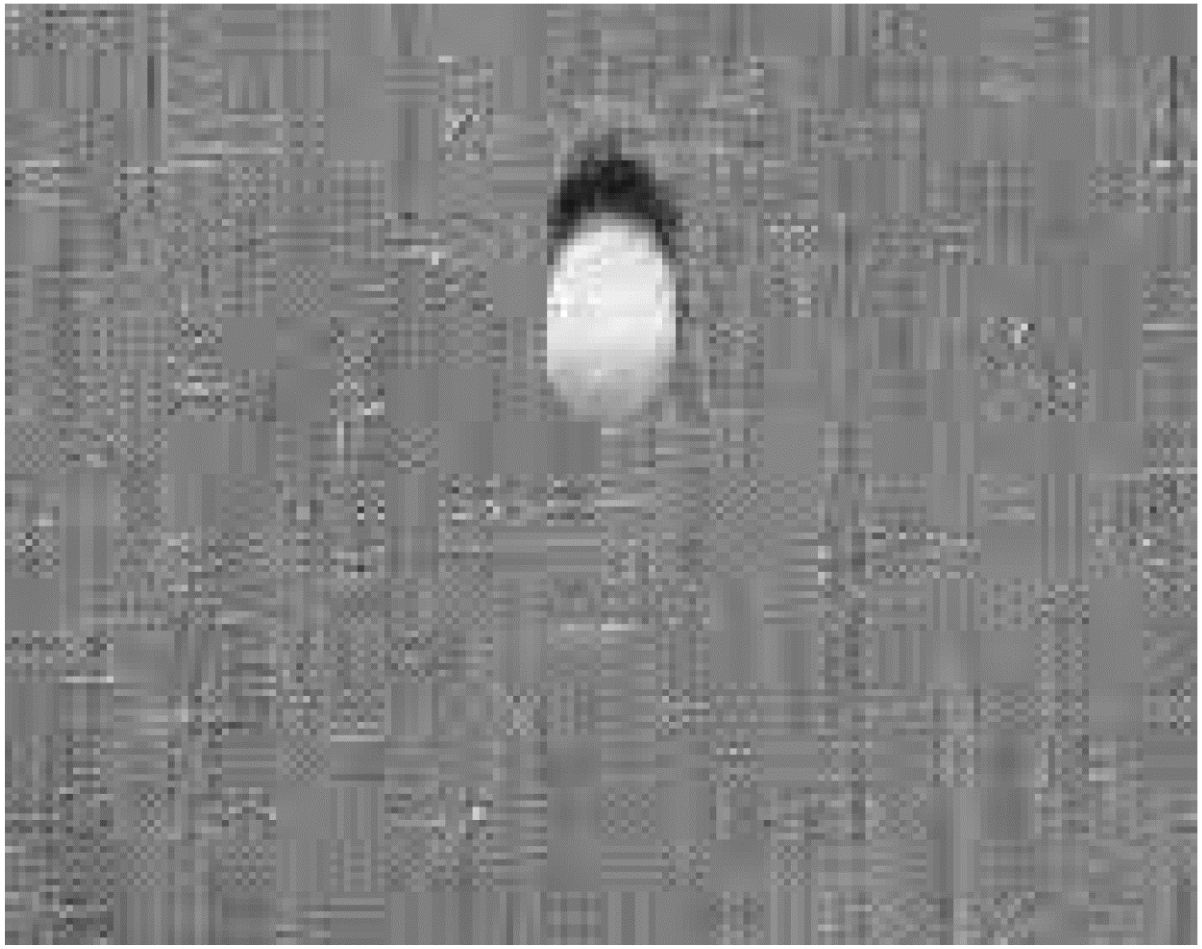
© CRC Press, Wu & Rao Eds, 2006

Mosaic effect in I-frame



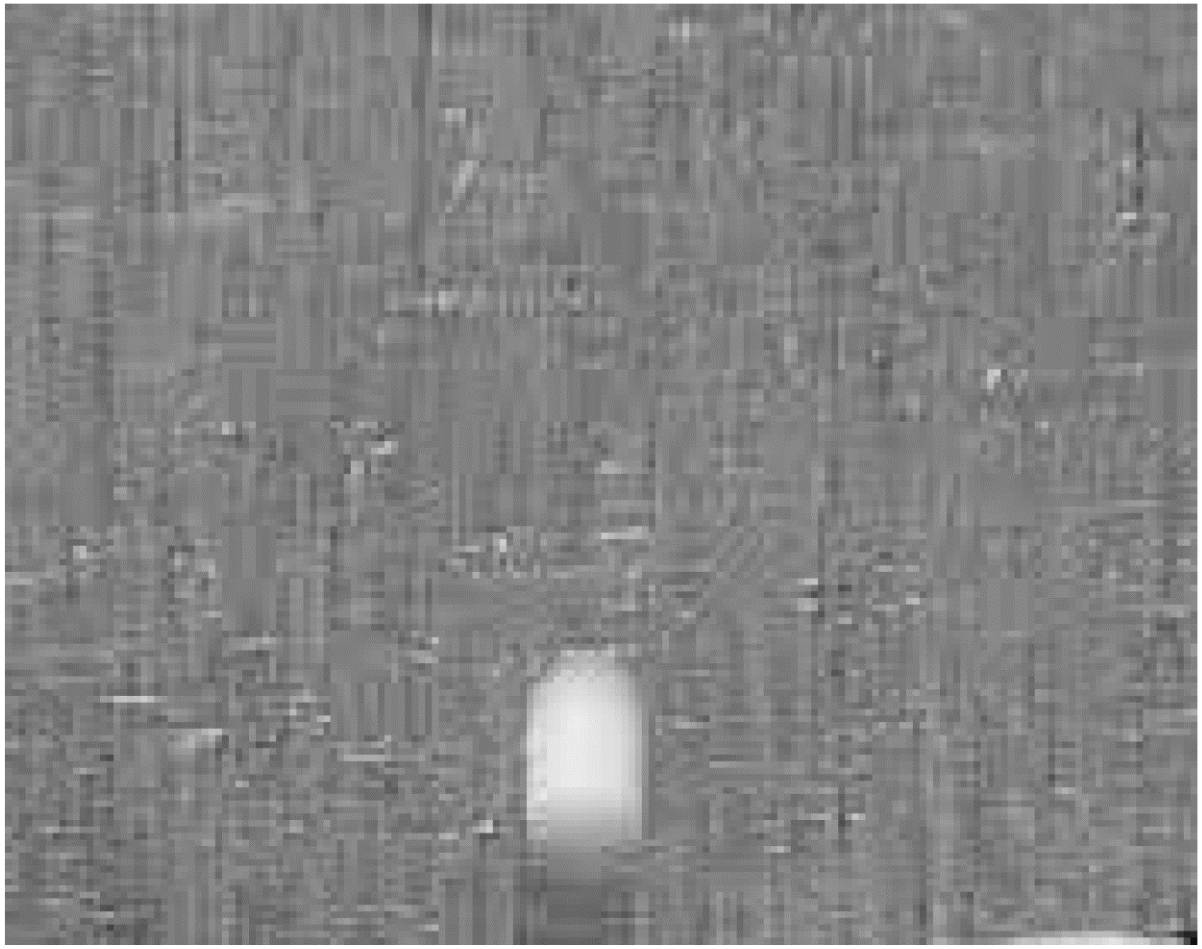
© CRC Press, Wu & Rao Eds, 2006

Mosaic effect in P-frame predicted from I-frame



© CRC Press, Wu & Rao Eds, 2006

Deterioration on Mosaic effect in P-frame predicted from P-frame



© CRC Press, Wu & Rao Eds, 2006

3.9 False Contouring

- Occurs in smoothly textured area with gradual transition in pixel intensity/value
- Appears as step-like gradation in affected areas
- Cause
  - Inadequate number of quantization levels
  - Irregular distribution of AC energy due to quantization

© CRC Press, Wu & Rao Eds, 2006

False Contouring

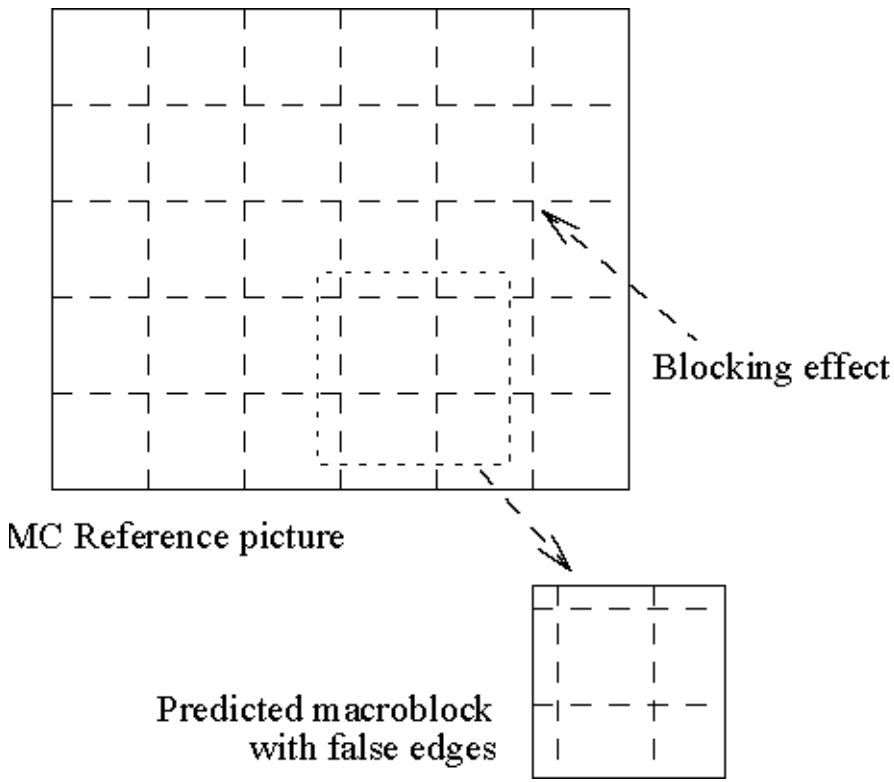
© CRC Press, Wu & Rao Eds, 2006

3.10 False Edges

- Occurs within inter-frame mode
- Propagated through blocking effect
  - Block-edge discontinuity used as reference of motion estimation
- Most prominent in smooth areas
- Usually localized to neighbouring frames

© CRC Press, Wu & Rao Eds, 2006

Propagation of blocking effect to false edges



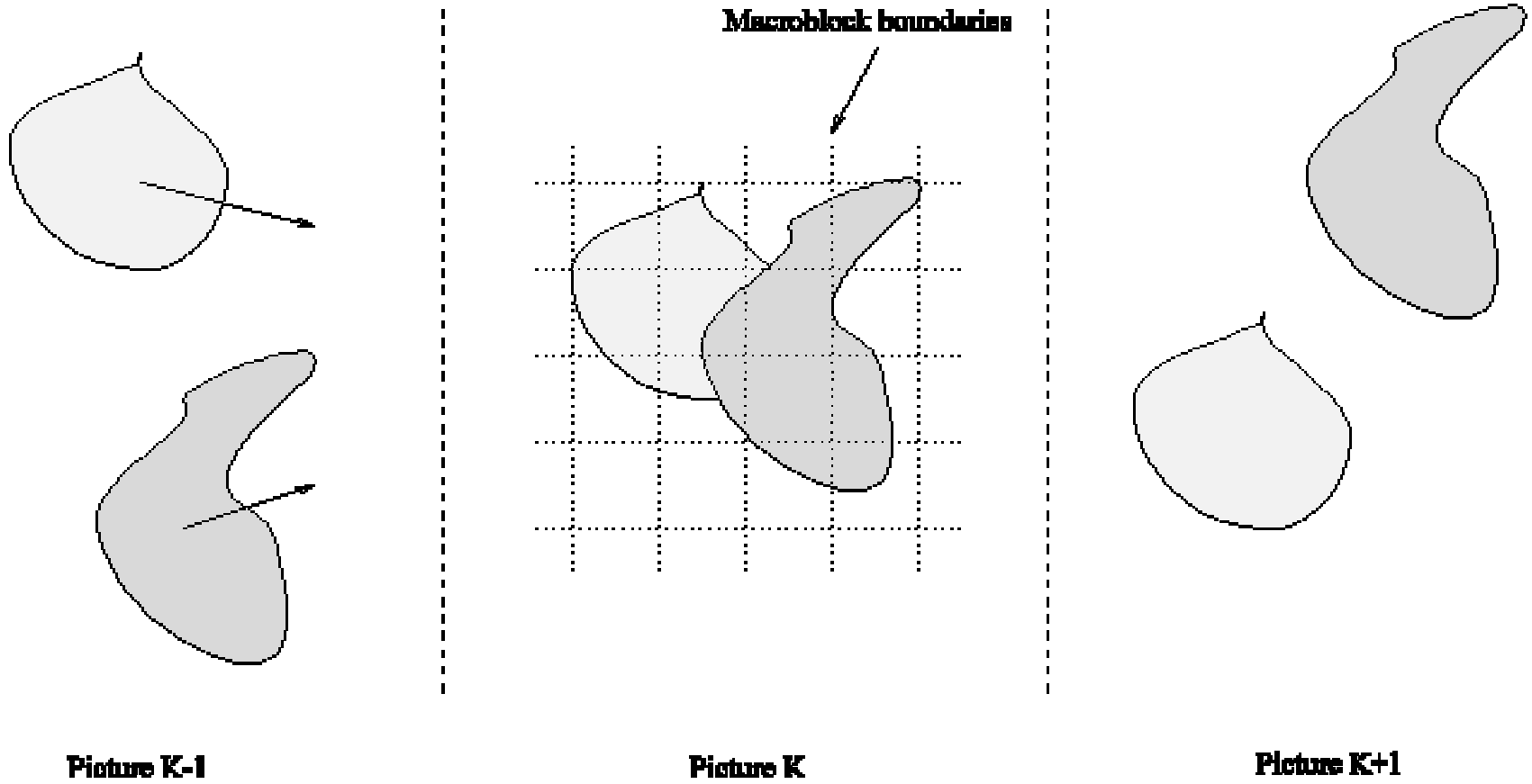
© CRC Press, Wu & Rao Eds, 2006

3.11 MC Mismatch

- Occurs when objects overlap in macro-blocks
- Subsequent MC will be unable to find a satisfactory match for overlapping objects
- Resulting in high prediction errors
- Errors ineffectively coded due to quantization
- Mismatch becomes visible

© CRC Press, Wu & Rao Eds, 2006

Macro-blocks encompassing the region of two objects with diverging motion



© CRC Press, Wu & Rao Eds, 2006

High frequency noise induced by motion mismatch

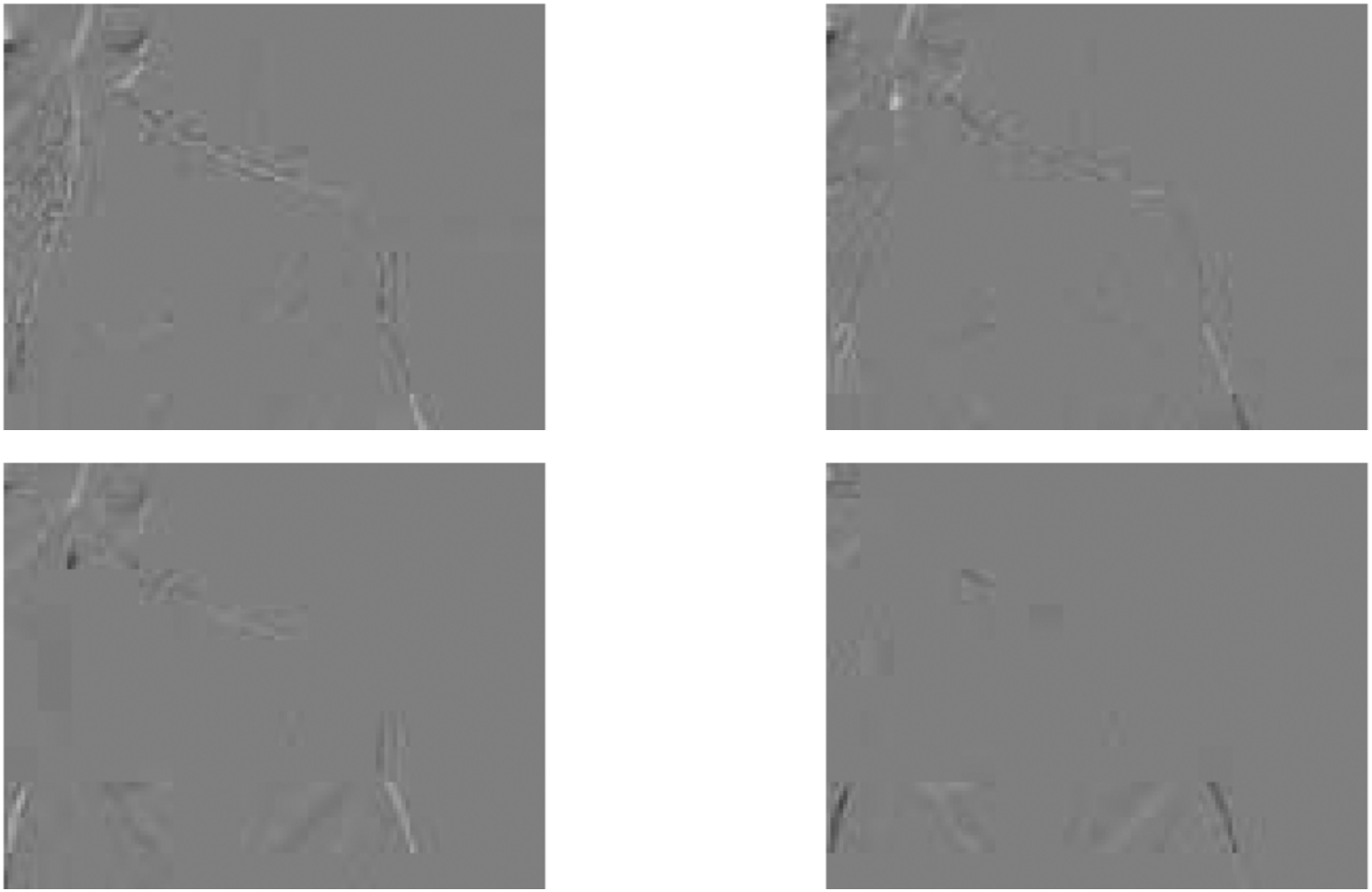
© CRC Press, Wu & Rao Eds, 2006

3.12 Mosquito Effect

- Due to temporal fluctuations
- Related to
  - Block mismatch
  - Ringing effect
- Ringing related mosquito effect
  - Occurs when coarse quantization is applied to high prediction errors in adjacent frames
  - Equivalent to fluctuating ringing

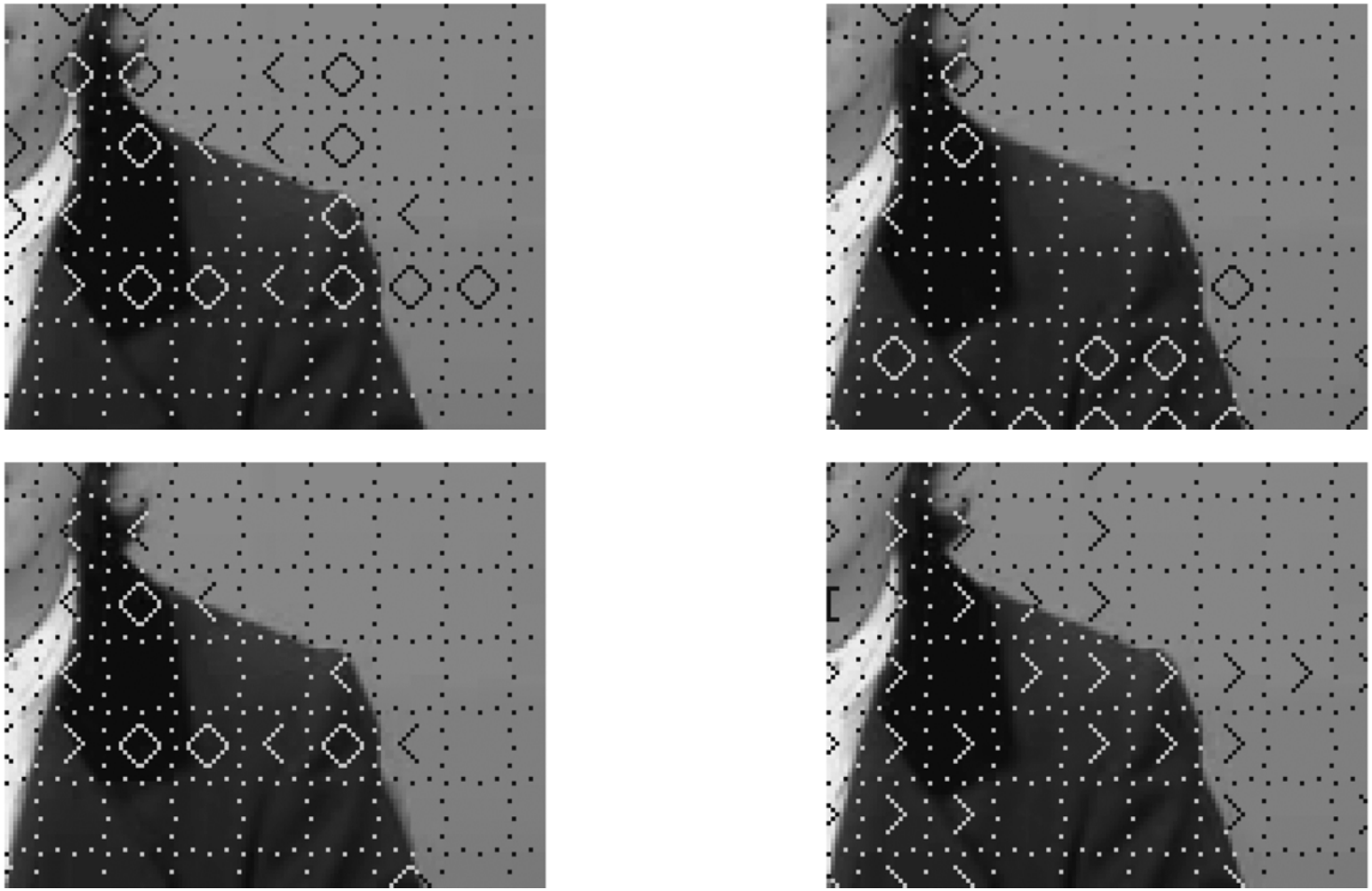
© CRC Press, Wu & Rao Eds, 2006

Difference sub-images of 5 successive *Claire* sequence. Order of coded frame: clockwise from top left.



© CRC Press, Wu & Rao Eds, 2006

MC prediction type for 4 successive frames. Forward (>), backward (<) and bi-directional (<>) predictions.



© CRC Press, Wu & Rao Eds, 2006

- Mismatched related mosquito effect
  - Enhanced mosquito effect due to MC mismatch of macro-blocks
  - Equivalent to fluctuating MC mismatch

© CRC Press, Wu & Rao Eds, 2006

3.13 Stationary Area Fluctuation

- Occurs around high contrast edges in the temporal domain
- Visible even around areas with high details (e.g. highly textured areas)
- Primary cause
  - Quantization of high frequency prediction errors (DPCM)
    - Resulting in fluctuation of high spatial frequency

© CRC Press, Wu & Rao Eds, 2006

Sons and Daughters Sequence (frame 40 and 41). Static background with some foreground motion

© CRC Press, Wu & Rao Eds, 2006

Difference Image of frame 40 and 41 of *Sons and Daughters* sequence.

© CRC Press, Wu & Rao Eds, 2006

3.14 Chrominance Mismatch

- Misplacement of a macro-block
  - relative to it's own general colour and surrounding colour
- Cause
  - Use of luminance channel motion vectors (MV) for chrominance
    - Optimum luminance and chrominance MVs for a given macro-block may not be the same

© CRC Press, Wu & Rao Eds, 2006

Chrominance mismatch at the top of the tree trunk

© CRC Press, Wu & Rao Eds, 2006

3.15 Video Scaling and Field Rate Conversion

- Scaling
  - Decimation → aliasing
  - Interpolation → blurring
- Field Rate Conversion
  - 60 Hz (NTSC) → 50 Hz (PAL) and vice versa
  - Jerky motion from downsampling
    - Temporal interpolation remove jerky motion, but produces ghosting artifacts

© CRC Press, Wu & Rao Eds, 2006

3.16 De-interlacing

- Interlaced (dual) fields → frame
- Some de-interlacing methods
  - Line repetition and averaging
    - Aliasing at high vertical frequencies and some blurring problems
  - Field repetition
    - Serration artifacts around edges of moving objects
  - Motion adaptivity - hybrid methods
    - Luminance difference, median filters, motion compensation, etc

© CRC Press, Wu & Rao Eds, 2006

3.17 Summary

- Classification of distortions in hybrid MC/DPCM/DCT coders
  - Based on consistent characteristics
- Quantization errors
  - Primary source of distortions
  - Spatial distortions
    - Within blocks and between blocks
  - Temporal distortions
    - Due to MC mismatch and quantized predicted errors

© CRC Press, Wu & Rao Eds, 2006

- Distortions manifestation
  - Multiple types of distortions may co-exists
  - Some distortions are related
    - e.g. blocking and ringing
  - Distortions are difficult to isolate
  - Dependent on
    - Source content – distortion levels varies with picture content
    - Coding parameters – MC, DCT, quantization, etc

© CRC Press, Wu & Rao Eds, 2006

AH94	A.J. Ahumada and R. Horng. De-blocking DCT compressed images. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 2179:109-116, May 1994
Ahu92	A.J. Ahumada. Luminance-model-based DCT quantization for color image compression. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 1666:365-374, 1992
Cla95	R.J. Clarke. *Digital Compression of Still Images and Video*. London: Academic Press, 1995
DM95	F. Dufaux and F. Moscheni. Motion estimation techniques for digital TV: A review and a new contribution. *Proc. IEEE*, 83(6):858-876, June 1995
Gir89	B. Girod. The information theoretical significance of spatial and temporal masking in video signals. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 1077:178187, October 1989
GW92	R.C. Gonzales and R.E. Woods. *Digital Image Processing*, Reading, MA: Addison-Wesley, 1992

© CRC Press, Wu & Rao Eds, 2006

HB98	G.D. Haan and E.B. Bellers. Deinterlacing – an overview. *Proc. IEEE*, 86(9) September 1998
HLS93	H. Hwang, M.H. Lee and D.I. Song. Interlaced to progressive scan conversion with double smoothing. *IEEE Trans. Consumer Electronics*, 39:241-246. August 1993
ITU90	ITU. *ITU-T Recommendation H.261: Video Codec for Audiovisual Services at p x 64 kbits/s*, 1990
ITU95	ITU. *Draft ITU-T Recommendation H.263: Video Coding for Low Bitrate Communication*, 1995
Jay92	N. Jayant. Signal compression: Technology targets and research directions. *IEEE J. Sel. Areas Commun.*, 10(5):796-818, June 1992
KHK87	M. Kaneko, Y. Hatori and A. Koike. Improvements of transform coding algorithm for motion-compensated inferframe prediction errors – DCT/SQ coding. *IEEE J. Sel./ Areas Commun.*, SAC-5(7):1068-1078, August 1987

© CRC Press, Wu & Rao Eds, 2006

KK95	S.A. Karunsekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Trans. Image Processing*, 4(6):713724, June 1995
KSC92	S.A. Klein, A.D. Silverstein and T. Carney. Relevance of human vision to JPEG DCT compression. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 1666:200-215, 1992
Le91	D. Le Gall. MPEG: A video compression standard for multimedia applications. *Communications of the ACM*, 34(4):47-57, April 1991
NH89	A.N. Netravali and B.G. Haskell. *Digital Pictures: Representation and Compression*, New York: Plenum Press, 1989
PAW93	H.A. Peterson, A.J. Ahumada and A.B. Watson. Improved detection model for DCT coefficient quantization. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 1913:191-201, September 1993
Plo89	R. Plompen. Motion video coding for visual telephony, 1989

© CRC Press, Wu & Rao Eds, 2006

RG86	B. Ramamurthi and A. Gersho. Nonlinear space-variant postprocessing of block coded images. *IEEE Trans. Acoust., Speech and Signal Processing*, ASSP-34(5):1257-1268, October 1986
RL83	H.C. Reeve and J.S. Lim. Reduction of blocking effects in image coding. In *Proc. IEEE ICASSP*, 1212-1215, Boston, MA, April 1983
RY90	K.R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. San Diego, CA: Academic Press, 1990
Sec91	Secretariat ISO/IEC JTC1/SC29. *ISO CD11172-2: Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s*. ISO, November 1991
Sec92	Secretariat ISO/IEC JTC1/SC29/WG11. *Test Model 1*. ISO , May 1992. Description of MPEG-2 test model
Tzo88	K.H. Tzou. Post-filtering of transform-coded images. In *Proc. SPIE Human Vision Visual Processing and Digital Display*, 2178:99-108, May 1994

© CRC Press, Wu & Rao Eds, 2006

YGK95	Y. Yang. N.P. Galatsanos and A.K. Katsaggelos. Projection-based spatially adaptive reconstruction of block-transform compressed images. *IEEE Trans. Image Processing*, 4(7):896-908, July 1995
YWR95	M. Yuen, H.R. Wu and K.R. Rao. Performance evaluation of POCS loop filtering in generic MC/DPCM/DCT video coding. In *Proc. SPIE Visual Communication and Image Processing*, 2501:65-75, Taipei, Taiwan, May 1995

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

II. Picture Quality Assessment and Metric

4. Video Quality Testing

Philip Corriveau

Intel Media and Acoustics Perception Lab, USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 4 Outline

4.1 Introduction
4.2 Subjective Assessment Methodologies
4.3 Selection of Test Materials
4.4 Selection of Participants - Subjects
4.5 Experimental Design
4.6 International Testing Methods
4.7 Objective Assessment Methods
4.8 Summary

© CRC Press, Wu & Rao Eds, 2006

4.1 Introduction

- Proliferation of audio/visual media
- Audio/visual services
  - Important to maintain quality of services
  - How to measure picture quality?
  - Difference between analogue and digital media
- Two approach to Quality measure
  - Subjective assessment
  - Objective assessment

© CRC Press, Wu & Rao Eds, 2006

4.2 Subjective Assessment Methodologies

- Utilize human subjects for assessment
- Most reliable way for quantify picture quality
- Different assessment methodology for different conditions

© CRC Press, Wu & Rao Eds, 2006

4.3 Selection of Test Materials

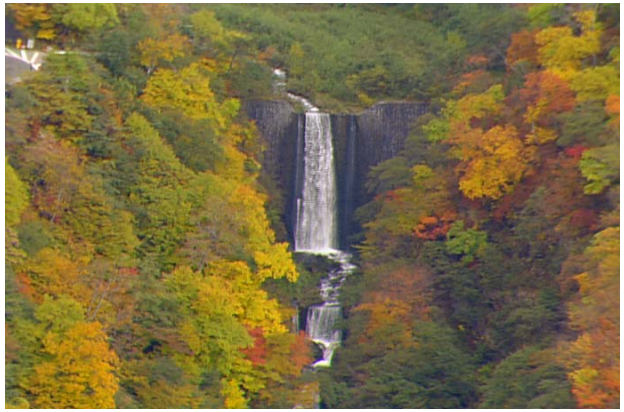
- Variations in picture quality
  - is subjected to compression, transmission and manipulation operations
- Picture content also affects picture quality
  - video sequences with minimal motion can maintain sufficiently high picture quality at low bitrates (*Susie*)
  - At similar bitrates, other types (content) of video sequences exhibit noticeable distortions. e.g. *Mobile & Calendar*

© CRC Press, Wu & Rao Eds, 2006

Susie Sequence

© CRC Press, Wu & Rao Eds, 2006

Common test sequences: clockwise from top left: *Waterfall, Tree, Mobile & Calendar and Tempete*



© CRC Press, Wu & Rao Eds, 2006

- Robust picture quality test for motion video equipment and system should
  - Have multiple sources
  - 8 to 16 different source material types
- Source materials provided through standards and co-operative bodies:
  - IEEE
  - SMPTE
  - VQEG (Video Quality Experts Group)

© CRC Press, Wu & Rao Eds, 2006

4.4 Selection of Participants - Subjects

- Either experts or non-experts
- Screening
  - Visual acuity
  - Colour blindness
  - Other visual impairments
- Subject numbers
  - 16 to 24 subjects to maintain statistic relevance

© CRC Press, Wu & Rao Eds, 2006

- Experts
  - Advantages
    - Experience
    - More precise evaluation – good for evaluating technical performance
  - Disadvantages
    - Bias evaluation
    - Critical evaluation – does not reflect the generalized view of the average consumer

© CRC Press, Wu & Rao Eds, 2006

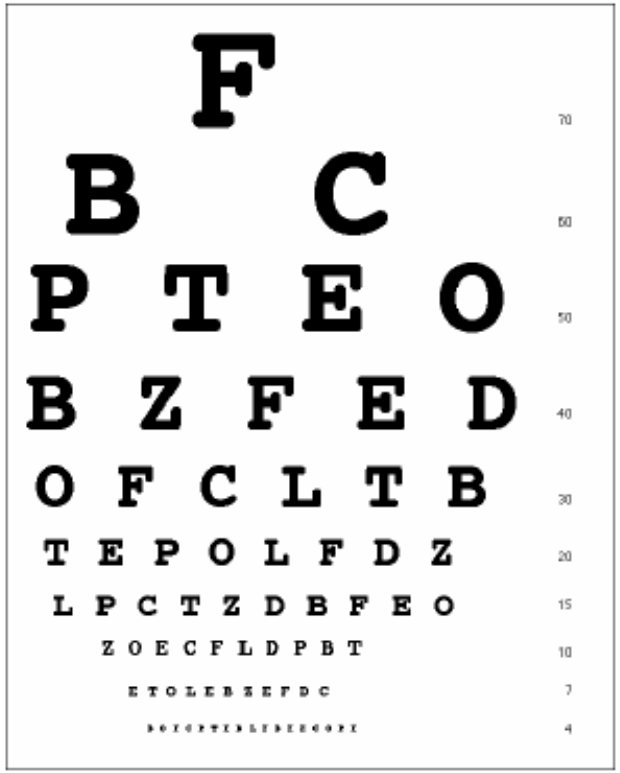
- Non-experts
  - General public, average user
  - May see artifacts that experts missed
    - Since they generally have a non-structured approach to evaluating picture quality/impairment
  - Generally reliable if tests are structured properly
  - Factors influencing non-expert
    - Age, gender, etc

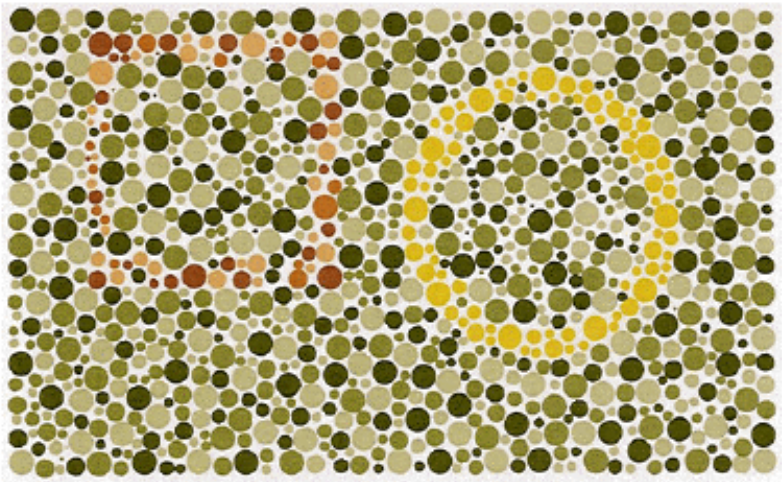
© CRC Press, Wu & Rao Eds, 2006

- Screening
  - For visual defects
  - Standardized tests
    - Snellen Eye Chart ©
      - Screens for 20/20 normal or corrected vision
    - Ishihara ©
      - Screens for colour blindness

© CRC Press, Wu & Rao Eds, 2006

Snellen eye chart (left) and Ishihara test plate (right)





© CRC Press, Wu & Rao Eds, 2006

4.5 Experimental Design

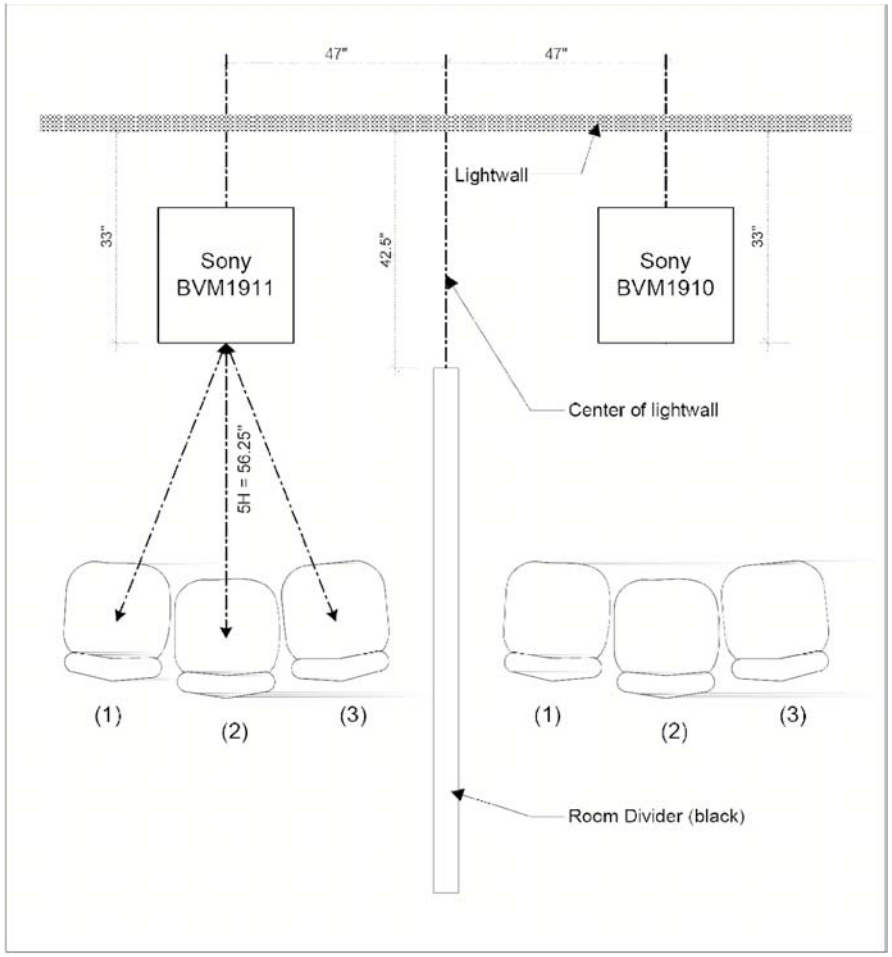
- Factors to consider for tests
  - Time
    - Excessive test duration leads to fatigue
    - Commonly 30 mins session with warm-up trials
  - Double/Single presentation
    - Affects test duration and limit number of test conditions covered
  - Repeat tests
    - For stabilizing results.
    - Past experience has shown little gain in data stability with repeats

© CRC Press, Wu & Rao Eds, 2006

- Factors affecting testing environment
  - Lighting
  - Ambient noise
  - Quality and calibration of display device
- Environmental specifications
  - ITU-T BT.500 standard
  - Exemplar testing environment
    - Communications Research Centre (CRC), Canada

© CRC Press, Wu & Rao Eds, 2006

Testing room layout at CRC



© CRC Press, Wu & Rao Eds, 2006

- Common experimental mistakes
  - Limited set of test material
  - Inadequate coverage of the range of conditions to be tested
    - Make sure that which is to be tested is tested thoroughly
  - Excessively long test durations → fatigue
  - Neglecting warm-up and re-set trials
    - Allows adjustment to testing procedures

© CRC Press, Wu & Rao Eds, 2006

4.6 International Testing Methods

- Double Stimulus Impairment Scale (DSIS)
- Double Stimulus Quality Scale (DSQS) Method
- Comparative Test
- Single Stimulus Continuous Quality Evaluation (SSCQE)
- Clear instructions on “how to do the test” must be given to subjects prior to test

© CRC Press, Wu & Rao Eds, 2006

4.6.1 Double Stimulus Impairment Scale (DSIS)

- Impairment measurement
- A 5-point discrete scale is commonly used
  - *Imperceptible* – perfect quality
  - *Perceptible, but not annoying*
  - *Slightly annoying*
  - *Annoying*
  - *Very Annoying* – unwatchable

© CRC Press, Wu & Rao Eds, 2006

DSIS 5-point rating scale

Imperceptible	☐
Perceptible, But Not Annoying	☐
Slightly Annoying	☐
Annoying	☐
Very Annoying	☐

© CRC Press, Wu & Rao Eds, 2006

DSIS test instructions

INSTRUCTIONS FOR IMPAIRMENT TESTS

In this test, we ask you to judge the visibility and severity of impairments in the television picture resulting from imperfections or interference in broadcasting.

Each trial will be announced verbally by number. The first presentation of a trial will be announced as "Reference", and the second presentation will be announced as "Test".

The test consists of a series of judgement trials, each consisting of two presentations of the same piece of video material. The first presentation shows the video material from a Reference signal (a broadcast not subjected to interference). The second presentation shows the same material from a Test signal (a broadcast subjected to a measured amount of interference).

We will now show you four demonstration trials. Across the four demonstrations, the test material will exhibit different levels of visible impairment to illustrate the kinds of impairments you will see during the test session.

DEMONSTRATION TRIALS

You are asked to examine the Reference presentation and note any deficiencies present, and then to assess the visibility and severity of any additional impairment in the Test material.

In judging the visibility and severity of transmission impairments, we ask you to use judgement scales like the sample shown below.

IMPERCEPTIBLE	☐
PERCEPTIBLE, BUT NOT ANNOYING	☐
SLIGHTLY ANNOYING	☐
ANNOYING	☐
VERY ANNOYING	☐

SAMPLE JUDGEMENT SCALE

The judgement scale offers five response options. These options are: "imperceptible", "perceptible, but not annoying", "slightly annoying", "annoying", and "very annoying". You are asked to place a single "X" in the box that best corresponds to your judgement (as shown in the example).

IMPERCEPTIBLE	☐
PERCEPTIBLE, BUT NOT ANNOYING	☒
SLIGHTLY ANNOYING	☐
ANNOYING	☐
VERY ANNOYING	☐

SAMPLE JUDGEMENT SCALE

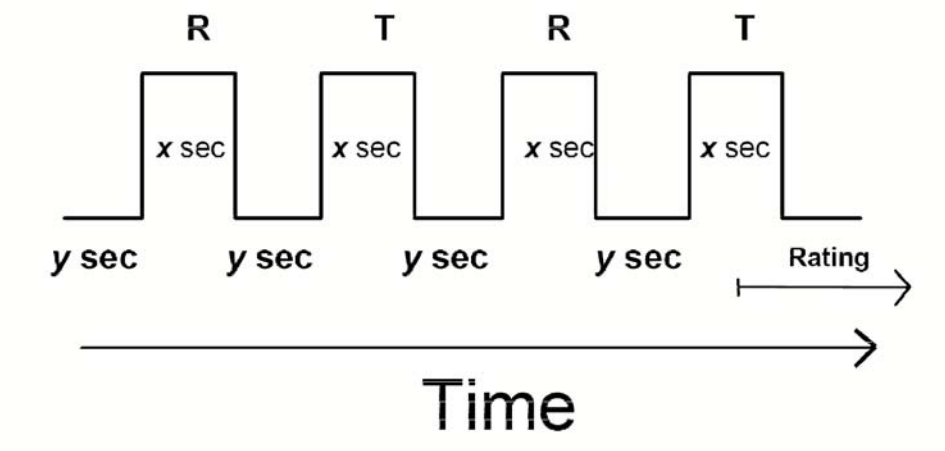
We ask you to use the judgement scales to express your judgement of the visibility and severity of the incremental impairment in the Test material. Please note, in some trials, the Reference material itself may have visible impairments. These are not of interest in this test.

Please refrain from recording your judgement until the final grey period of the trial.

© CRC Press, Wu & Rao Eds, 2006

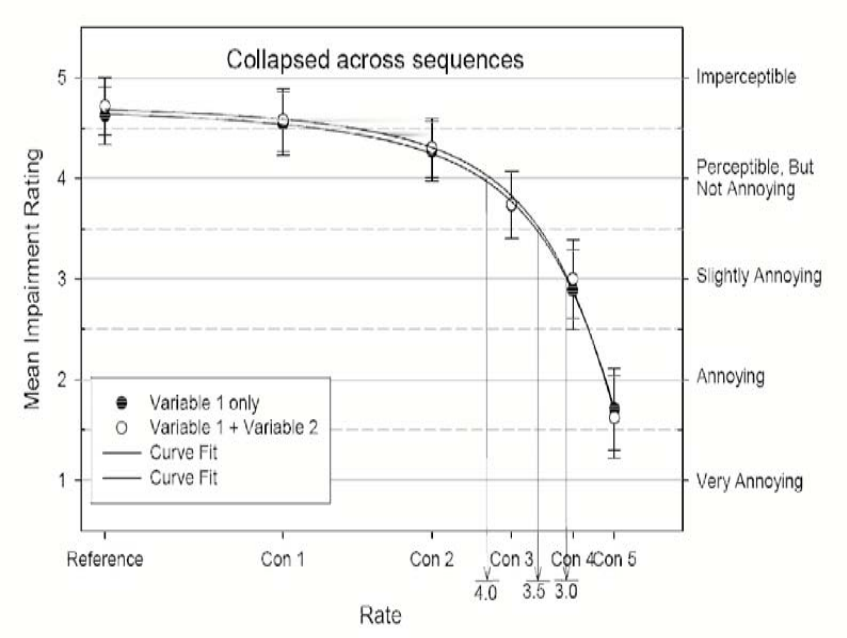
- DSIS measures failure characteristics
  - e.g. minimum coding rate relative to quality

Below: DSIS trial structure with reference (*R*) and test (*T*) material. Impairment is recorded during the rating period



© CRC Press, Wu & Rao Eds, 2006

Typical outcome of the overall impairment results

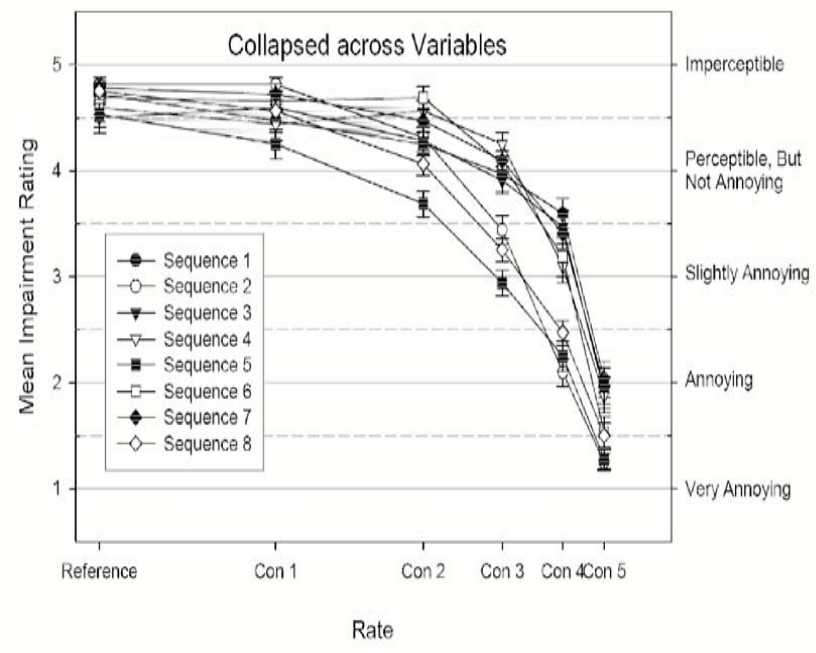


- Overall analysis
  - Testing various conditions (con1, con2, etc)
    - Graph signifies behaviour
  - Test methods with 1 and 2 variables
    - Simultaneous assessment of two systems under multiple conditions

© CRC Press, Wu & Rao Eds, 2006

Individual sequence results for a typical DSIS test

- Specific Analysis
  - Evaluate the response of individual sequence
    - Identify extremities, thus operational boundaries of systems
  - Identify system behaviour relative to specific source type
    - Poor with what type?
    - Good with what type?



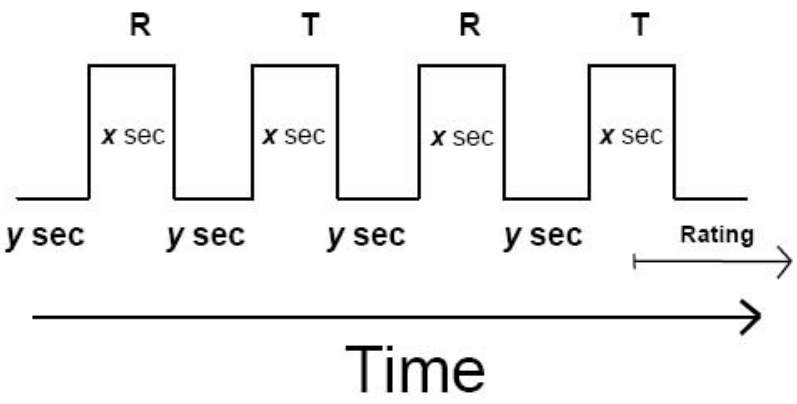
© CRC Press, Wu & Rao Eds, 2006

4.6.2 Double Stimulus Quality Scale (DSQS) Method

- Quality measurement
- Double scale for grading quality
  - Measurement taken as difference between two scores
- Continuous scale (0-100)
  - Minimizes quantization errors
  - 5 categories: Excellent (80-100); Good (60-79); Fair (40-59); Poor (20-49) and bad (0-19)

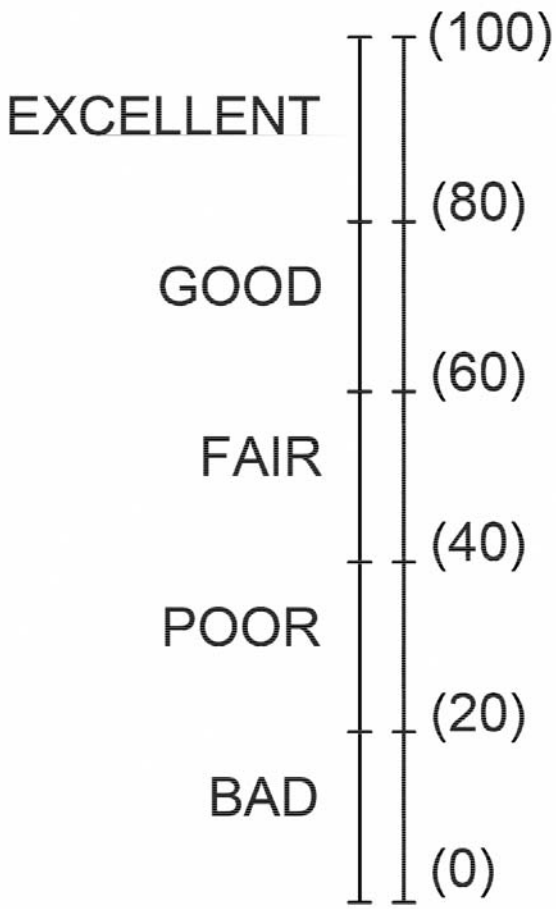
© CRC Press, Wu & Rao Eds, 2006

- Subjects rate both the first (*R*) and the second (*T*) sequence.
- Double blind test
  - *R* and *T* may be in reverse order
  - Sequence order is unknown to both subject and tester



© CRC Press, Wu & Rao Eds, 2006

DSQS scale



- Properly design tests
  - Score difference between 1st and 2nd sequence rarely exceed 40 units
  - Often under 25
- Reference may be rated lower than test sequence

© CRC Press, Wu & Rao Eds, 2006

DSQS test instructions

INSTRUCTIONS FOR QUALITY TESTS

In this test, we ask you to evaluate the overall quality of the video material you see. By overall quality, we mean the quality of the appearance of the video, not the desirability of the material itself.

Possible problems in quality include:

- poor, or inconsistent, reproduction of detail;
- poor reproduction of colours, brightness, or depth;
- poor reproduction of motion;
- imperfections, such as false patterns, or "snow".

The test consists of a series of judgement trials, each consisting of four presentations of the same piece of video material.

Each trial will be announced verbally by number. The first presentation of a trial will be announced as "A", and the second as "B". This pair of presentations will then be repeated, thereby completing a single trial. Please note that one of the presentations, A *or* B, will be a reference picture that shows the video material in a high quality format, with the other being a test picture that shows the same video material in the format being tested.

We will now show you four demonstration trials.

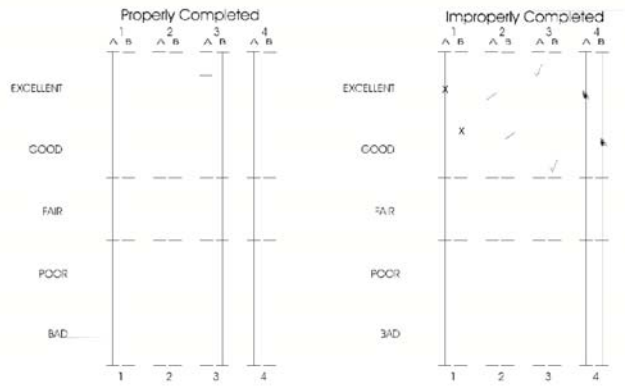
DEMONSTRATION TRIALS

In judging the overall quality of the presentations, we ask you to use judgement scales like the samples shown below.

SAMPLE QUALITY SCALE

As you can see, there are two scales for each trial, one for the "A" presentation and one for the "B" presentation, since both the "A" and "B" presentations are to be judged.

The judgement scales are continuous vertical lines that are divided into five segments. As a guide, the adjectives "excellent", "good", "fair", "poor", and "bad" have been aligned with the five segments of the scales. You are asked to place a single horizontal line at the point on the scale that best corresponds to your judgement of the overall quality of the presentation (as shown in the example).



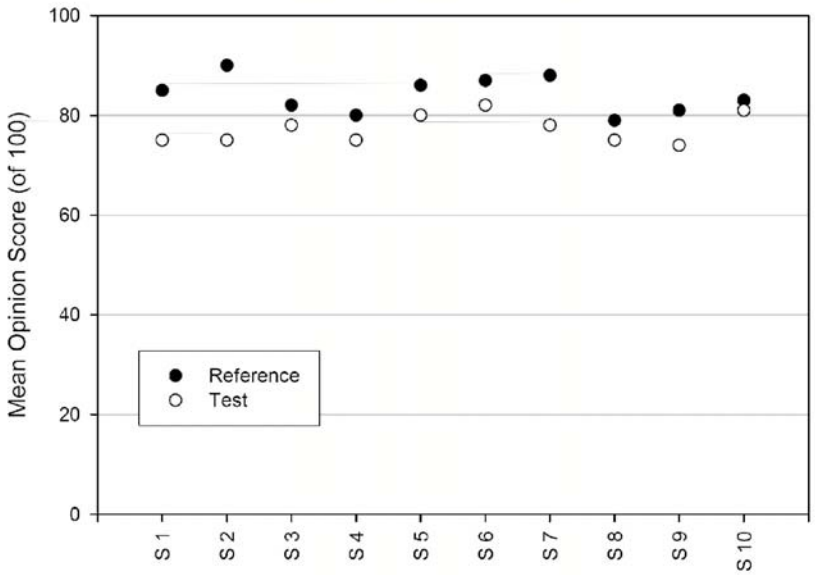
You may make your mark at any point on the scale which most precisely represents your judgement.

In making your judgements, we ask you to use the first pair of presentations in the trial to form an impression of the quality of each presentation, but to refrain from recording your judgements. You may then use the second pair of presentations to confirm your first impressions and to record your judgements in your Response Booklet.

© CRC Press, Wu & Rao Eds, 2006

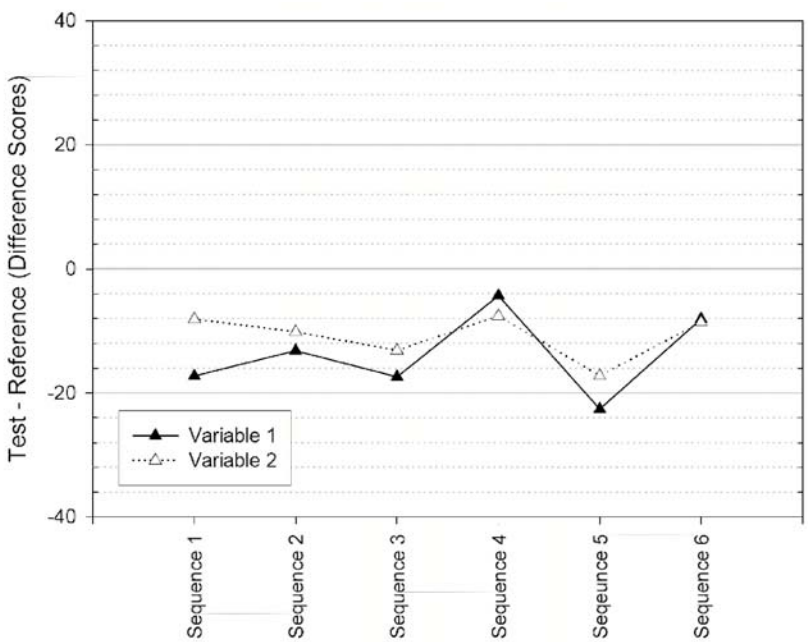
Typical DSQS raw results

- Mean Opinion Score (MOS)
  - Purposes
    - Check stability of reference score
    - Rough indication of differences between reference and test sequences



© CRC Press, Wu & Rao Eds, 2006

Typical DSQS difference results

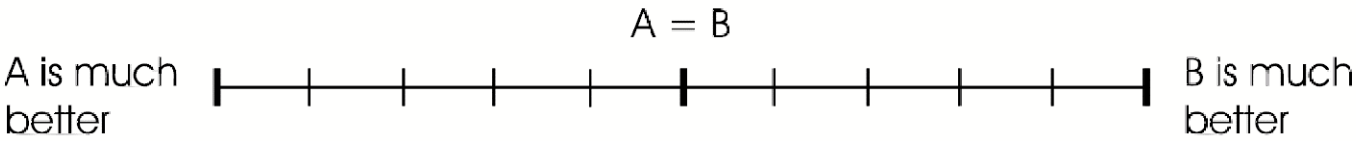


- Minus test
  - Check if test sequence has been rated higher than reference
- Analysis
  - Test methods
    - Ascertain the difference between 1 and 2 variables
  - Is quality dependent on sequences?

© CRC Press, Wu & Rao Eds, 2006

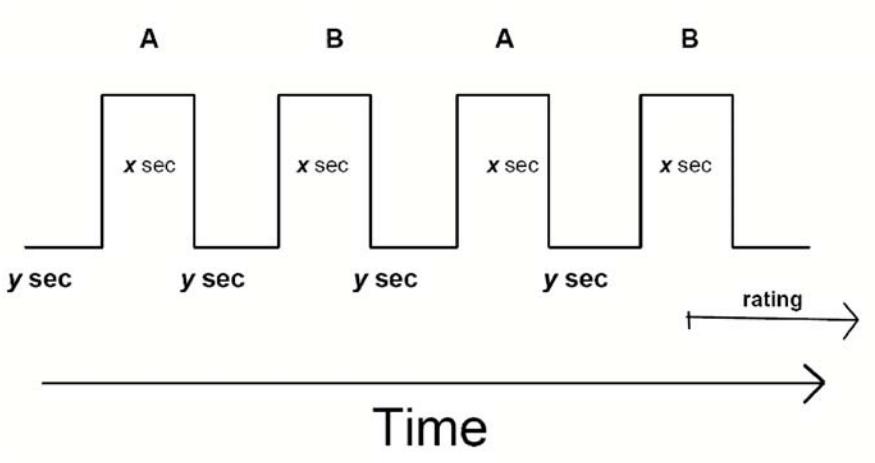
4.6.3 Comparison Scale Method

- For comparing systems with equal capability
- Comparative Scale (below)
  - 100 units: -50(A) → 0 (A=B) → 50(B)



© CRC Press, Wu & Rao Eds, 2006

Comparative test trial structure

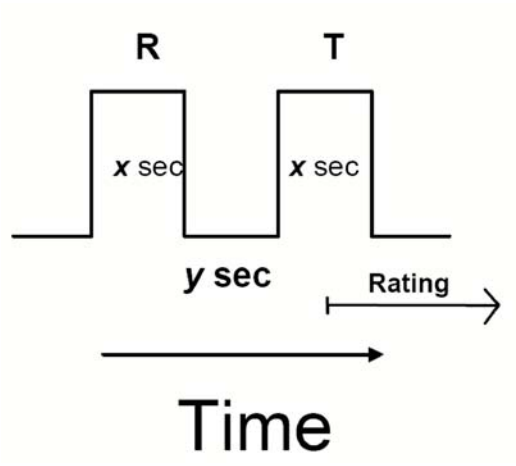


- Double blind test
- Single or double stimulus

© CRC Press, Wu & Rao Eds, 2006

4.6.4 Single Stimulus Method

- One stimulus per set of sequence
  - One opportunity to evaluate, no repeat
- Shorter testing time
- Applicable to DSQS, DSIS, CS, etc



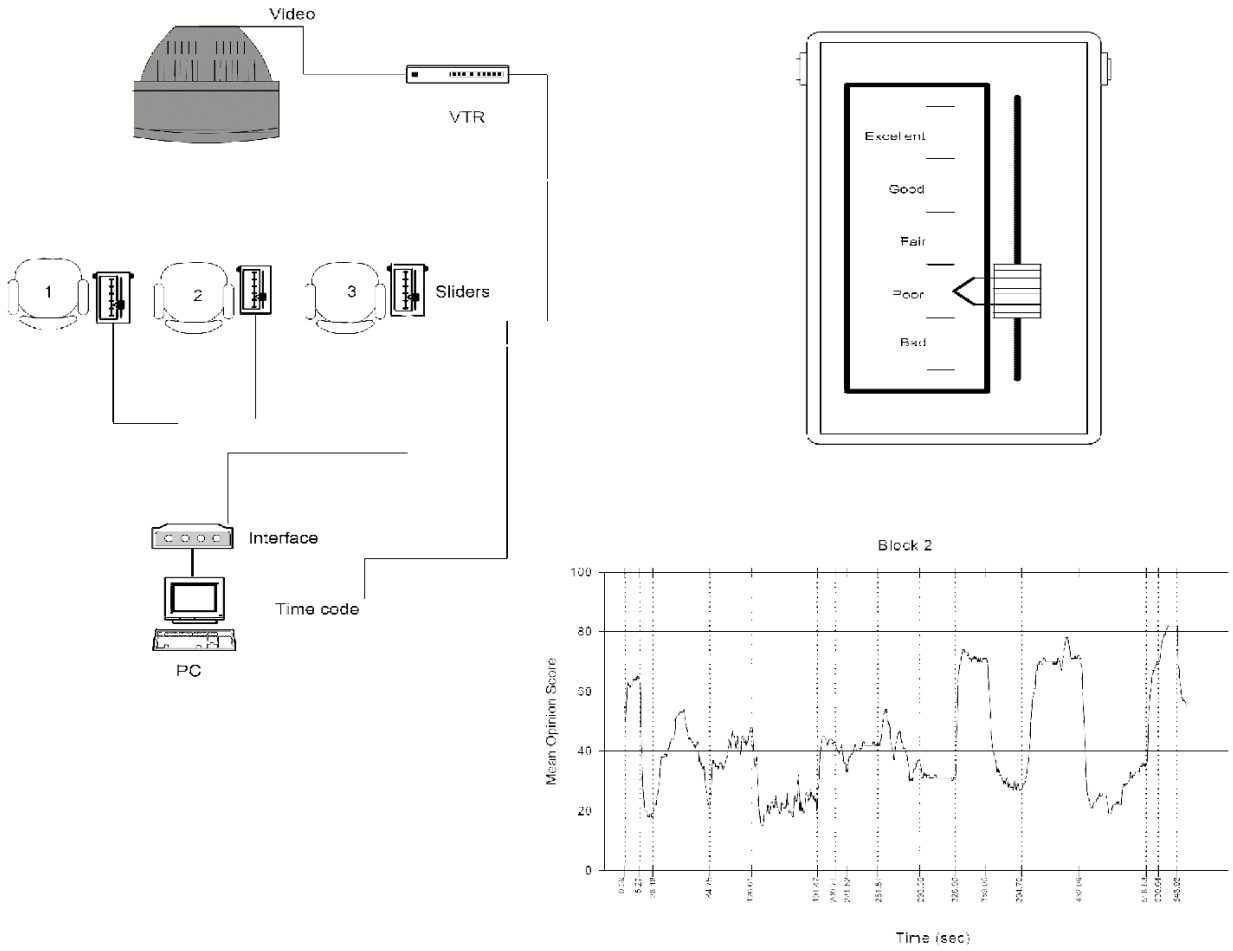
© CRC Press, Wu & Rao Eds, 2006

4.6.5 Continuous Quality Evaluations

- Single Stimulus Continuous Quality Evaluation (SSCQE) method
- No reference sequences
- Continuous rating
  - Sequences under evaluation are rated for the duration of the sequence
    - Record variation in picture quality within sequences in addition to between sequences
  - Continuous recording with slider (i.e. mouse or cursors) in computer

© CRC Press, Wu & Rao Eds, 2006

SSCQE evaluation layout (left), slider (right) and typical continuous score (bottom)



© CRC Press, Wu & Rao Eds, 2006

4.6.6 Discussion of SSCQE and DSCQS

- Primary difference
  - SSCQE
    - Continuous opinion scoring
    - Computer recorded with slider
  - DSCQS
    - Single opinion score per set of test sequences
    - Recorded on score sheet
- Slider vs Score Sheet
  - Marking errors on score sheets
    - Leads to discarding of data

© CRC Press, Wu & Rao Eds, 2006

- Continuous score vs single score
  - Pinpoint error segments in sequences
    - SSCQE can, DSCQS cannot
  - Real-time application
    - Evaluation of pictures during broadcast
  - Reflects subjective opinion with better precision
- SSCQE more applicable than DSCQS for quality assessment of video

© CRC Press, Wu & Rao Eds, 2006

- Comparatively, SSCQE and DSCQS results are statistically indistinguishable
  - Within the limits of what could be compared
- Comparative study of SSCQE and DSCQS [PW03]
  - M.H. Pinson and S. Wolf, “Comparing subjective video quality testing methodologies”, In *SPIE Video Comms. And Image Processing Conference,* pp. 146-157, 2003
- DSCQS improvement
  - apply continuous (slider) score

© CRC Press, Wu & Rao Eds, 2006

- Primacy and Recency Effect
  1. Strong impression of distortion (Primacy)
  2. Impression lingers and subsequently weights heavily in picture quality for a certain duration of time (Recency)
  - SSCQE prone to this problem
  - Uni-opinion score DSCQS has no problem
- Uneven changes with continuous score
  - Sharp degradation with poor quality frames
  - Slow recovery with good quality frames

© CRC Press, Wu & Rao Eds, 2006

4.6.7 Pitfalls of different methods

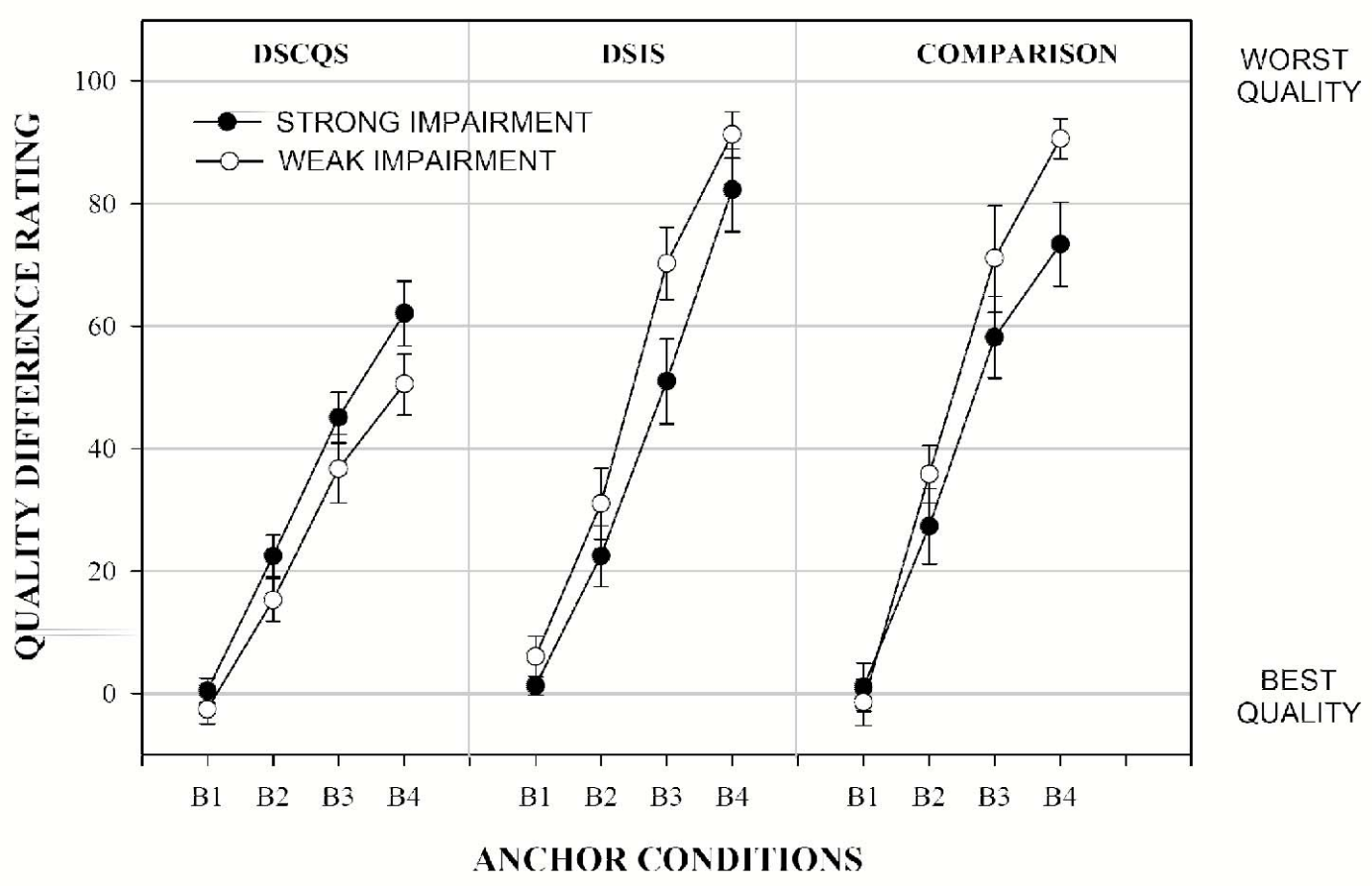
- Contextual effect –
  - Manifests according to operational context
  - Dependency between sequences
    - Perceived quality is at times influenced by previous sequences
  - Ceiling or floor (range) effect
    - Perceived quality does not properly match the range of quality score.
      - e.g., “excellent” is matched to “good” due to inadequate quality range for test materials
      - Partially remedied with anchors

© CRC Press, Wu & Rao Eds, 2006

- Joint investigation into contextual effect
  - Labs involved
    - CRC (communications Research Centre), Canada
    - CCRTT (France Telecom)
    - IRT (Institut für Rundfunktecknik), Germany
    - SPTT (Swiss Telecom)
  - Assessment method studies
    - DSCQS
    - DSIS
    - Comparative

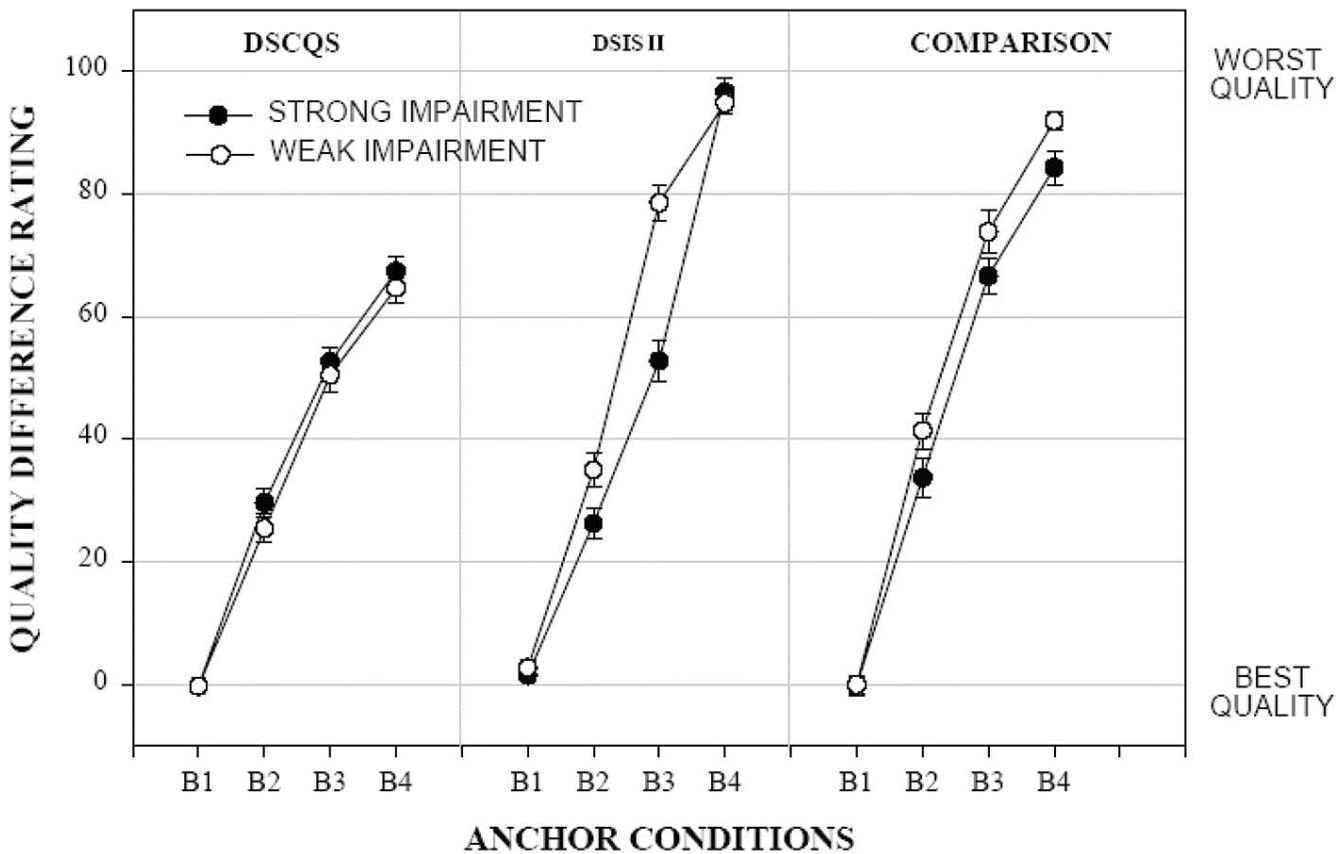
© CRC Press, Wu & Rao Eds, 2006

CRC results for various anchor conditions in DSCQS, DSIS and Comparison methods



© CRC Press, Wu & Rao Eds, 2006

Collapse (combined) results from CRC, CCETT, IRT and SPTT [ITU97] for various anchor conditions



© CRC Press, Wu & Rao Eds, 2006

- CRC results
  - Moderate contextual effect for all methods studied at most conditions except **B1**
- Combined results
  - Slight contextual effect overall except DSIS under condition **B3**
- Noticeable difference between CRC and combined results
  - Variations unavoidable, but acceptable if within limits

© CRC Press, Wu & Rao Eds, 2006

4.7 Objective Assessment Methods

- Automated assessment
  - Without direct human input
- Classical analogue objective measures
  - Peak Signal-to-Noise Ratio (PSNR)
  - Mean Square Error (MSE)
- Analogue measures inadequate for digital pictures
  - Different characteristics between digital and analogue pictures

© CRC Press, Wu & Rao Eds, 2006

- For digital pictures
  - Quality is source dependent
  - Distortions are structured
    - According to coding techniques
- Different objective measures required for digital pictures
  - Should account for the HVS
  - Able to handle a variety of picture types
    - Natural, animated, synthetic, etc

© CRC Press, Wu & Rao Eds, 2006

- Industrial requirements
  - Full reference
    - Assessment based on an original reference
  - Reduced reference
    - Assessment based on a partial/reduced original reference
  - No reference
    - No point of original reference for quality/distortion assessment

© CRC Press, Wu & Rao Eds, 2006

4.8 Summary

- Subjective Assessment
  - Design test to answer question one is seeking
    - Assessment method depends on
      - Quality or Impairment measure
      - Comparison or discrimination
      - Double or single stimulus
    - Selection of test material
    - Selection of subjects
      - Experts or non-experts

© CRC Press, Wu & Rao Eds, 2006

All80	J.W. Allnatt. Subjective assessment method for television digital codecs. *Electronics Letters*, 16(12):450-451, June 1980
Cea99	P. Corriveau *et al*. All subjective scales are not created equal: The effect of context on different scales. *Signal Processing*, 77:1-9, 1999
CW99	P. Corriveau and A. Webster. VQEG evaluation of objective methods of video quality assessment. *SMPTE Journal*, 108:645-648, 1999
IL93	A.F. Inglis and A.C. Luther. *Video Engineering*. New York: McGraw-Hill, 1993
ITU97	ITU-R. *Investigation of Contextual Effects*. ITU-R, March 1997. ITU-R Document 11E/34-E, 24
ITU98	ITU. ITU-R BT. 500-9, methodology for the subjective assessment of the quality of television pictures. *ITU-R BT*, 1998
MD82	I.F. MacDiarmid and P.J. Darby. Double-stimulus assessment of television picture quality. *EBU Technical Review*, E78-A(192):70-78, April 1982
Nar94a	N. Narita. Consideration of subjective evaluation method for quality of image coding. *Electronics and Communications in Japan, Part 3*, 77(7):84-103, 1994

© CRC Press, Wu & Rao Eds, 2006

Nar94b	N. Narita. Proposal of a modified ebu method for reliable evaluation of picture quality. *NHK Laboratories Note*, E78-A(431):1-10, December 1994
Nar97	N. Narita. Effect of impairment ranges on reliability of the modified ebu method. *IEICE Tans. Fundamentals*, E78-A(11):1553-1555, Nov. 1997
NS97	N. Narita and Y. Sugiura. On an absolute evaluation method of the quality of television sequences. *IEEE Transaction on Broadcasting*, 43(1):26-35, 1997
PW03	M.H. Pinson and S. Wolf. Comparing subjective video quality testing methodologies. In *SPIE Video Communications and Image Processing Conference*, 146-157, Lugano, Switzerland, July 2003
Ris86	D.R. Riskey. Use and abuse of category scales in sensory measurement. *Journal of Sensory Studies I*, 217-236, 1986
Sar67	V. Sarris. Adaptation-level theory: Two critical experiments on Helson's weighted average model. *America Journal of Psychology*, 80:331-334, 1967
VQE03	VQEG. *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. VQEG, August 2003

© CRC Press, Wu & Rao Eds, 2006

WA80	T. White and J. Allnatt. Double Stimulus quality rating method for television digital codecs. *Electronics Letters*, 16(18):714-716, 1980
ZC96	E. Zou and P. Corriveau. Methods for evaluation of digital television picture quality. In *SMPTE 138th Technical Conference and World Media Expo*, 146-457, October 1996

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

5. Perceptual Video Quality Matrix – A Review

Stefan Winkler

Genista Corporation, Montreux, Switzland

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 5 Outline

5.1 Introduction
5.2 Quality Factors
5.3 Metric Classification
5.4 Pixel-Based Metrics
5.5 The Psychophysical Approach
5.6 The Engineering Approach
5.7 Metric Comparison
5.8 Conclusion and Perspective

© CRC Press, Wu & Rao Eds, 2006

5.1 Introduction

- Evaluation and Optimization of Picture Coding/Processing Systems
  - Reliant on the Human Visual System (HVS)
    - Determines visual distortion → perceived quality
- Subjective Assessment
  - Most reliable for quantifying perceived quality
  - Complicated and time consuming to implement
  - Impractical for real-time applications

© CRC Press, Wu & Rao Eds, 2006

- Objective Assessment
  - Mathematical metrics
    - MSE, PSNR, etc
    - Simple and fast
    - Unreliable predictors of perceived quality/distortion
  - HVS based metrics
    - Give a more consistent account of perceived quality and distortion
- Still image and video metrics
  - Discussions limited to video

© CRC Press, Wu & Rao Eds, 2006

5.2 Quality Factors

- What influences quality?
  - Interest and expectation
    - of services (DVD, VHS, mobile broadcasts)
    - of programs (movie, sports, etc)
  - Display type and properties
    - CRT, LCD, Plasma, etc – affects contrast
  - Viewing Condition
    - Distance, size, resolution, etc
    - Luminance from ambient lighting

© CRC Press, Wu & Rao Eds, 2006

- What else influences quality?
  - Fidelity of reproduction
    - Visible distortions (e.g. blurring)
    - Enhancement (e.g. sharpness, clarity)
  - Accompanying soundtrack
    - Synchronization with video
    - Clarity of audio
- “Interests” (of viewers) are the most difficult factor to account for

© CRC Press, Wu & Rao Eds, 2006

5.3 Metric Classification

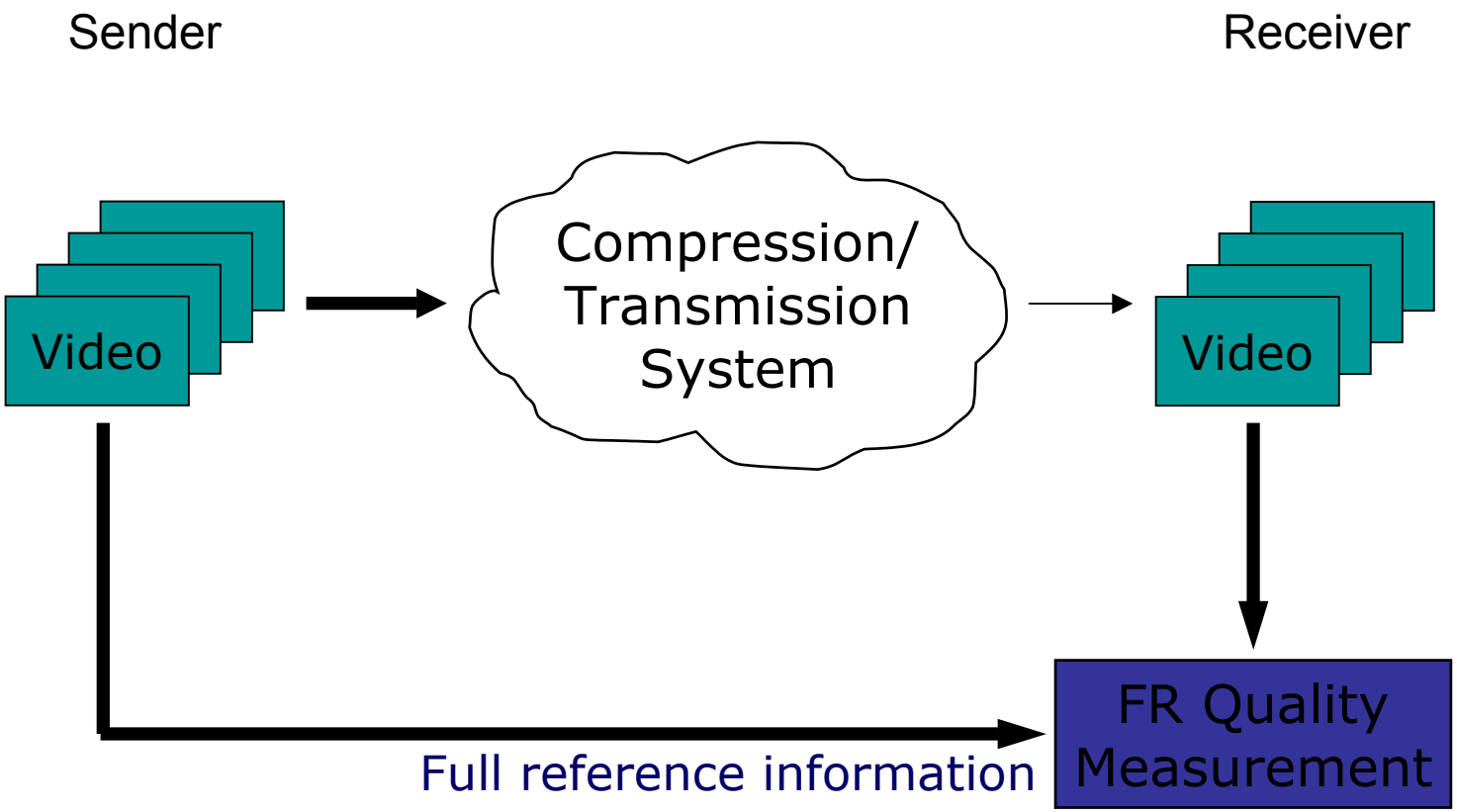
- Psychophysical – HVS modelling
  - Quantify picture quality relative to HVS
    - Simple models with contrast, colour modelling
    - More complicated with neurological modelling
  - Suitable for a wide range of video applications
- Engineering – signal analysis
  - Quantify quality relative to structural features and/or distortions
  - May consider HVS characteristics

© CRC Press, Wu & Rao Eds, 2006

- Metric referencing types
  - Full reference (FR)
    - Frame by frame comparison between reference and test video
  - Reduced reference (RR)
    - Only some references are available
  - No reference (NR)
    - Only has test video for measurement
      - Problem with distinguishing distortions from content
  - FR, RR and NR are driven by applications

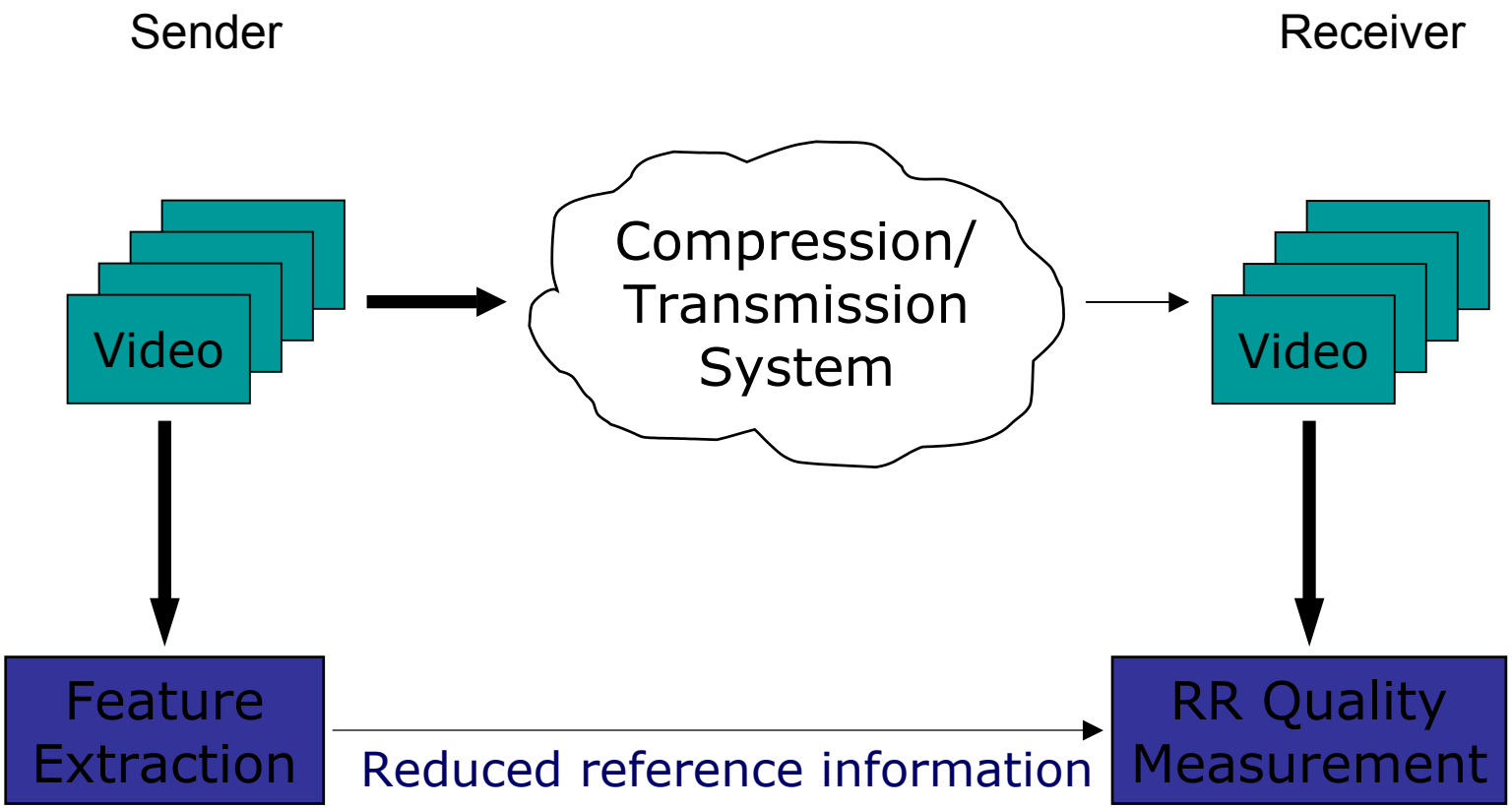
© CRC Press, Wu & Rao Eds, 2006

Full-Reference Metrics



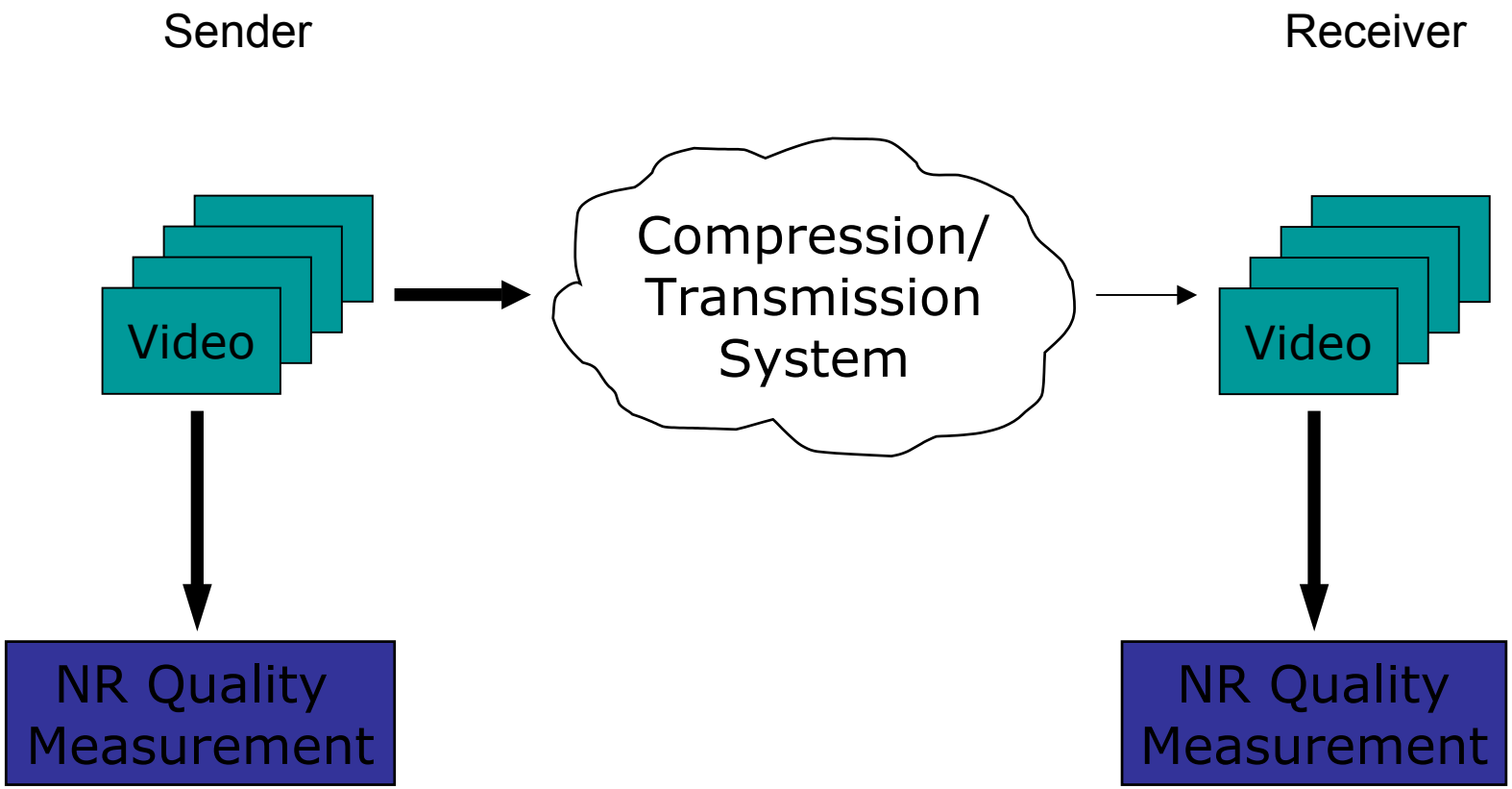
© CRC Press, Wu & Rao Eds, 2006

Reduced-Reference Metrics



© CRC Press, Wu & Rao Eds, 2006

Non-Reference Metrics



© CRC Press, Wu & Rao Eds, 2006

5.4 Pixel-Based Metrics

- The Classics
  - Mean Square Error (MSE)

$$MSE = \frac{1}{TMN} \sum_t \sum_m \sum_n [x_o(m,n,t) - x_r(m,n,t)]^2$$

  - Peak Signal-to-Noise Ratio (PSNR)

$$PSNR = 10 \log_{10} \frac{I^2}{MSE}$$

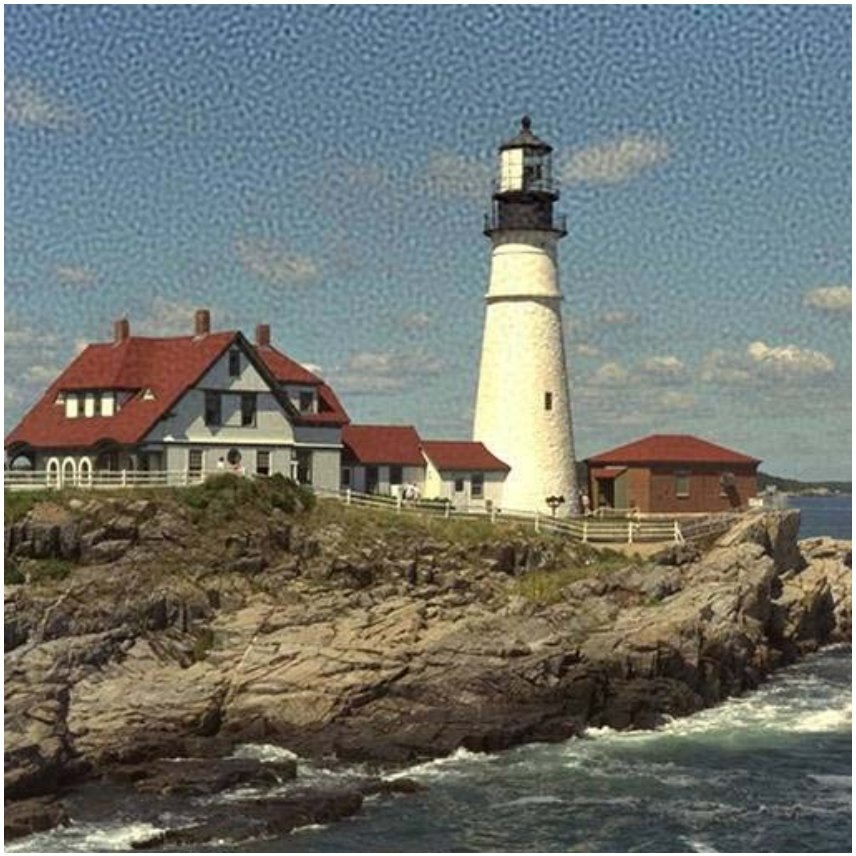
Temporal (t) and spatial (m,n) axis with maximum luminance/chrominance magnitude (I)

© CRC Press, Wu & Rao Eds, 2006

- MSE measures picture difference
- PSNR measures picture fidelity
- Fast but limited approximation of perceived distortion/quality
  - similar MSE/PSNR does not always equate to similar perceived distortion/quality
- Effectiveness of Pixel-based metric
  - Tailored to specific type of distortion
    - MSE good for additive noise, poor for coding artifacts

© CRC Press, Wu & Rao Eds, 2006

Images polluted with high frequency noise (left) and band-pass filtered noise (right) having identical PSNR.



© CRC Press, Wu & Rao Eds, 2006

5.5 The Psychophysical Approach

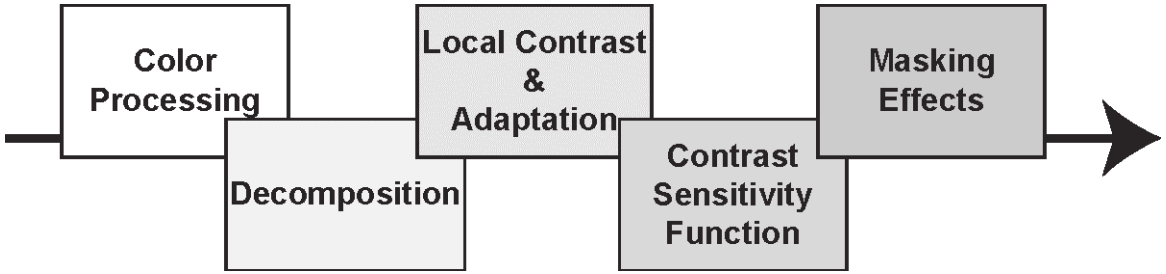
- Basic components of vision models
  - Colour processing
    - Colour space conversion
  - Multi-channel decomposition
    - Spatial and temporal frequencies
  - Contrast sensitivity and adaptation
    - HVS operates in contrast (difference) levels
    - Affected by adaptation to luminance level
  - Masking (facilitation)
    - Suppression and enhancement of visual stimuli

© CRC Press, Wu & Rao Eds, 2006

– Pooling

- Integration of (all) visual stimuli
  - Spanning spatial, temporal and colour channels

Below: Block diagram of a typical HVS model



© CRC Press, Wu & Rao Eds, 2006

5.5.2 Single Channel Model

- Single channel spatial filter
- Simple to implement
- Computationally efficient
- Modelling
  - Contrast Sensitivity Function (CSF)
  - Distortion detection depends on some threshold criterion

© CRC Press, Wu & Rao Eds, 2006

- Examples of CSF
  - For Image
    - Mannos and Sakrison's Modulation Transfer Function (MTF)
      - Contrast sensitivity relative to spatial frequencies
  - For Video
    - Lukis and Budrikis's spatio-temporal CSF model
      - Incorporates masking through spatial-temporal activity weighting function

© CRC Press, Wu & Rao Eds, 2006

5.5.3 Multi-Channel Model

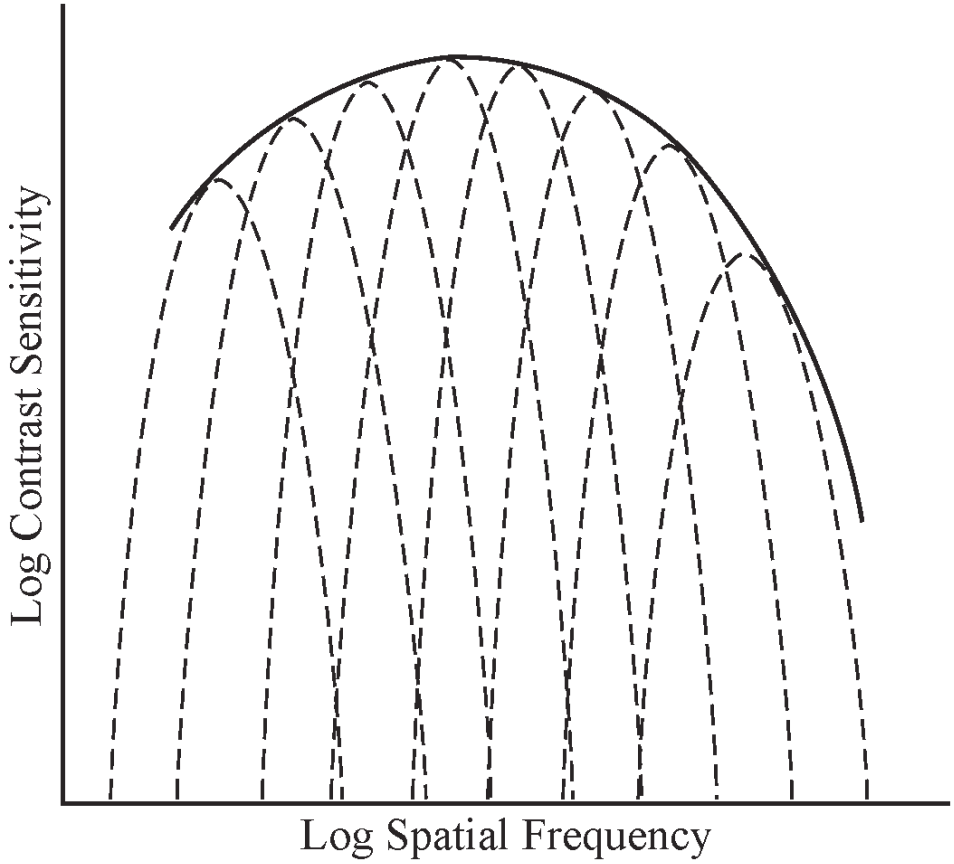
Multiple channels

- Each channel deals with one frequency band
- Overall CSF is equivalent to the envelop of sensitivities of all channels
- Detection is independent in each channel when threshold is reached

- Relative to single channel models
  - More complex in implementation
  - Better prediction performance

© CRC Press, Wu & Rao Eds, 2006

Envelop of multiple CSFs



© CRC Press, Wu & Rao Eds, 2006

- Some Multi-channel models
  - Still images
    - Visible Difference Predictor (VDP)
    - Visual Discrimination Model (VDM)
  - Video
    - Just Noticeable Difference (JND) metric
    - Moving Picture Quality Metric (MPQM)
    - Normalization Video Fidelity Metric (NVFM)
    - Continuous Video Quality Evaluation (CVQE)
      - Designed for low bitrate video

© CRC Press, Wu & Rao Eds, 2006

5.6 The Engineering Approach

- Specialized metrics
  - Ad hoc techniques
  - Operates well within a predefined scope
  - Good computational efficiency
- Some Full Reference (FR) metrics
  - DCTune – for still images
  - Perceptual Video Quality Measure (PVQM)
  - Structural Similarity Index Metric (SSIM)

© CRC Press, Wu & Rao Eds, 2006

- A Reduced Reference (RR) metric
  - Video Quality Metric (VQM)
    - Partial reference derived from extracted features
- No Reference (NR)
  - More specialized in application
  - Used for measuring specific artifacts
    - Blurring, jerkiness, ringing, clipping, blocking
      - Blocking Impairment Metric (BIM)
    - May operate in time or spectral domains

© CRC Press, Wu & Rao Eds, 2006

5.7 Metric Comparison

- Evaluation criterions
  - Prediction accuracy
  - Monotonicity
    - ordering of pictures according to quality
  - Consistency
    - Number of outliers
- Quantified with regression analysis
  - Correlation most commonly used

© CRC Press, Wu & Rao Eds, 2006

5.7.2 Video Quality Experts Group

- VQEG evaluations
  - Source materials
    - Production and distribution videos
    - 20 scenes with 16 test conditions
  - DSCQS Subjective Assessment
  - 2 test phases
- Phase I
  - Evaluated 10 metrics including PSNR
  - Conclusion: no clear winner

© CRC Press, Wu & Rao Eds, 2006

- Phase II
  - Focused on digitally encoded video
  - Covers wider range of distortions
  - New sources and conditions
    - 128 sequences produced
  - Evaluated 6 metrics including PSNR
  - Each metric calibrated to reference videos
  - Mean Opinion Score (MOS) correlation
    - 70% MOS for PSNR
    - 94% MOS for the best metric

© CRC Press, Wu & Rao Eds, 2006

- Current VQEG activity
  - RR and NR metrics evaluations
    - for television and other multi-media services
    - At low bitrates and small frame sizes

© CRC Press, Wu & Rao Eds, 2006

5.7.3 Limitation of prediction performance

- Statistically factor
  - Subjective scores rest on statistical relevance
    - Large sample size needed
- Performance of metrics
  - Phase I below expectation
    - Correlation of MOS across all labs is 90-95%
      - Indicative of the performance limit of metrics
    - Best metric has correlation of 80-85%
  - Phase II
    - Best metric not statistically equivalent to MOS

© CRC Press, Wu & Rao Eds, 2006

5.8 Conclusion and Perspective

- Quality metrics
  - Applicable in many fields
    - e.g. broadcasting services
  - Many issues and problems to be resolved
    - Comprehensive evaluation of quality metrics
    - Better modelling of human vision
      - attention area (region of interest)
    - Perceived quality in addition to picture fidelity
      - Colour, sharpness, brightness, etc, affect quality
    - What is the Impact of audio in video quality?

© CRC Press, Wu & Rao Eds, 2006

ABE98	A.J. Ahumada, Jr., B.L. Beard and R. Eriksson. Spatio-temporal discrimination model predicts temporal masking function. In *Proc. SPIE*, 3299:120-127, San Jose, CA, Jan. 26-29, 1998
Ahu93	A.J. Ahumada, Jr.. Computational image quality metrics: A review. In *SID Symposium Digest*, 24:305-308, 1993
AN93	A.J. Ahumada, Jr. and C.H. Null. Image quality: A multidimensional problem. In A.B. Watson, Ed., *Digital Images and Human Vision*, 141-148. Cambridge, MA: The MIT Press, 1993
ASS02	İ. Avcıbaş, B. Sankur and K. Sayood. Statistical evaluation of image quality measures. *J. Electronic Imaging*, 11(2):206-223, April 2002
Bar99	P.G.J. Barten. *Contrast Sensitivity of the Human Eye and Its Effects in Image Quality*. Bellingham, WA: SPIE, 1999
BP99	V. Baroncini and A. Pierotti. Single-ended objective quality assessment of DTV. In *Proc. SPIE*, 3845:244-253, Boston, MA, Sep. 19-22, 1999
C+00	T. Carney *et al*. Modelfest: Year one results and plans for future years. In *Proc. SPIE*, 3959:140-151, San Jose, CA, Jan. 23-28, 2000

© CRC Press, Wu & Rao Eds, 2006

C+02	T. Carney *et al*. Extending the Modelfest image/threshold database into the spatio-temporal domain. In *Proc. SPIE*, 4462:138-148, San Jose, Jan. 20-25, 2002
CGDC01	F.X. Coudoux, M.G. Gazalet, C. Derviaux and P. Corlay. Picture quality measurement based on block visibility in discrete cosine transform coded video sequences. *J. Electronic Imaging*, 10(2):498-510, April 2001
CO03	J. Cavuedes and F. Oberti. No-reference quality metric for degraded and enhanced video. In *Proc. SPIE*, 5150:621-632, Lugano, Switzerland, July 8-11, 2003
CS04	A. Cacallaro and S. Winkler. Segmentation-driven perceptual quality metrics. In *Proc. ICIP*, 3543-3546, Singapore, Oct. 24-27, 2004
CWT+98	G.W. Cermak, S. Wolf, E.P. Tweedy, M.H. Pinson and A.A. Webster. Validating objective measures of MPEG video quality. *SMPTE J.*, 107(4):226-235, April 1998
Dal93	S. Daly. The visible difference predictor: An algorithm for the assessment of image fidelity. In A.B. Watson, Ed., *Digital Images and Human Vision*, 179-206. Cambridge, MA: The MIT Press, 1993

© CRC Press, Wu & Rao Eds, 2006

Dal98	S. Daly. Engineering observations from spatiovelocity and spatiotemporal visual models. In *Proc. SPIE*, 3299:108-191, San Jose, CA, Jan 26-29 1998
Dau80	J.G. Daugman. Two-dimensional spectral analysis of cortex receptive field profiles. *Vision Research*, 20(10):847-856, 1980
dRBF95	H. de Ridder, F.J.J. Blommaert and E.A. Fedorovskaya. Natyralness and image quality: Chroma and hue variation in color images of natural scenes. In *Proc SPIE*, 2411:51-61, San Jose, CA, Feb. 6-8, 1995
DS98	I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):245-267, 1998
EAV98	R. Eriksson, B. Andrén and K Brunnström. Modelling the perception of digital images: A performance study. In *Proc. SPIE*, 3299:88-97, San Jose, CA, Jan. 26-29, 1998
EF95	A.M. Eskicioglu and P.S. Fisher. Image quality measures and their performance. *IEEE Trans. Comm.*, 43(12):2959-2965, Dec. 1995

© CRC Press, Wu & Rao Eds, 2006

Fau79	O./D. Faugeras. Digital color image processing within the framework of a human visual model. *IEEE Trans. Acoust. Speech Signal Process.*, 27(4):380-393, Aug. 1979
FBC95	D.R. Fuhrmann, J.A. Baro and J.R. Cox, Jr. Experimental evaluation of psychophysical distortion metrics for JPEG-coded images. *J. Electronic Imaging*, 4(4):397-406, Oct. 1995
FH98	R.E. Fredericksen and R.F. Hess. Estimating multiple temporal mechanisms in human vision. *Vision Research*, 38(7):1023-1040, 1998
Gh73	E.M. Granger and J.C. Heurtley. Visual chromaticity-modulation transfer function. *J. Opt. Soc. Am.*, 63(9):1173-1174, Sep. 1973
GZR02	P. Gastaldo, R. Zunino and S. Rovetta. Objective assessment of MPEG-2 video quality. *J. Electronic Imaging*, 11(3):365-374, July 2002
$H^{+}02$	A.P. Hekstra *et al*. PVQM – a perceptual video quality measure. *Signal Processing: Image Communication*, 17(10):781-798, Nov. 2002
Han04	D.S. Hands. A basic multimedia quality model. *IEEE Trans. Multimedia*, 6(6):806-816, Dec. 2004

© CRC Press, Wu & Rao Eds, 2006

HMG+03	Y. Horita, T. Miyata, I.P. Gunawan, T. Murai and M. Ghanbari. Evaluation model considering static-temporal quality degradation and human memory for SSCQE video quality. In *Proc. SPIE*, 5150:1604-1611, Lugano, Switzerland, July 8-11, 2003
HMM99	T. Hamada, S. Muyaji and S. Matsumoto. Picture quality assessment system by three-layered bottom-up noise weighting considering human visual perception *SMPTE J.*, 108(1):20-26, Jan. 1999
HS92	R.F. Hess and R.J. Snowden. Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research*, 32(1):47-59, 1992
ITU02	ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union, Geneva, Switzerland, 2002
Jac95	R.E. Jacobson. An evaluation of image quality metrics. *J. Photographic Sci.*, 43(1):7-16, 1995
KCBST97	S.A. Klein, T. Carney, L. Barghout-Stein and C.W. Tyler. Seven models of masking. In *Proc. SPIE*, 3016:13-24, San Jose, CA, Feb. 8-14, 1997

© CRC Press, Wu & Rao Eds, 2006

Kel83	D.H. Kelly. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *J. Opt. Soc. Am.*, 73(6):745-750, June 1983
Kle93	S.A. klein. Image quality and image compression: A psychophysicist's viewpoint. In A.B. Watson, Ed., *Digital Images and Human Vision*, 73-88, Cambridge, MA: The MIT Press, 1993
LB82	F.X.J. Lukas and Z.L. Budrikis. Picture quality prediction based on a visual model. *IEEE Trans. Comm.*, 30(7):1679-1692, July 1982
LF97	J. Lubin and D. Dibush. Sarnoff JND vision model. T1A1.5 Working Group Document #97-612. ANSI T1 Standards Committee, 1997
LLO+03	Z. Lu, W. Lin, E. Ong, X. Yang and S. Yao. PQSM-based RR and NR video quality metrics. In *Proc. SPIE*, 5150:633-640, Lugano, Switzerland, July 8-11, 2003
LMK98	B. Li, G.W. Meyer and R.V. Klassen. A comparison of two image quality models. In *Proc. SPIE*, 3299:98-109, San Jose, CA, Jan. 26-29, 1998
LPB02	S. Lee, M.S. Pattichis and A.C. Bovik. Foveated video quality assessment. *IEEE Trans. Multimedia*, 4(1):129-132, March 2002

© CRC Press, Wu & Rao Eds, 2006

Lub95	J. Lubin. A visual discrimination model for imaging system design and evaluation. In E. Peli, Ed., *Vision Models for Target Detection and Recognition*, 245-283. Singapore: World Scientific Publishing, 1995
Lun93	A.M. Lund. The influence of video image size and resolutionon viewing-distance preferences. *SMPTE J.*, 102(5):407-415, May 1993
LvdBL96	P. Lindh and C.J. van den Brenden Lambrecht. Efficient spatio-temporal decomposition for perceptual cprocessing of video sequences. In *Proc. ICIP*, 3:331-334, Lausanne, Switzerland, Sep. 16-19, 1996
Mar86	H. Marmolin. Subjective MSE measures. *IEEE Trans. Systems, Man and Cybernetics*, 16(3):486-489, May 1986
MDWE02	P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. A no-reference perceptual blur metric. In *Proc. ICIP*, 3:57-60, Rochester, NY, Sep. 22-25, 2002
MDWE04	P. Marziliano, F. Dufaux, S. Winkler and T. Ebrahimi. Perceptual blur and ringing metrics: Application to JPEG2000: Signal Processing: Image Communication, 19(1), Jan. 2004

© CRC Press, Wu & Rao Eds, 2006

MEC98	A. Mayache, T. Eude and H. Cherifi. A comparison of image quality models and metrics based on human visual sensitivity. In *Proc. ICIP*, 3:409-413, Chicago, IL, Oct. 4-7, 1998
MH04	M.A. Masry and S.S. Hemami. A metric for continuous quality evaluation of compressed video with severe distortions. *Signal Processing: Image Communication*, 19(1):133-146, Feb. 2004
MS74	J.L. Mannos and D.J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Trans. Information Theory*, 20(4):525-536, July 1974
Mul85	K.T. Mullen. The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *J. Physiology*, 359:381-400, 1985
MW00	T.S. Meese and C.B. Williams. Probability summation for multiple patches of luminance modulation. *Vision Research*, 40(16):2101-2113, July 2000
BRK02	M.J. Nadenau, J. Reichel and M. Kunt. Performance comparison of masking models based on a new psychovisual test method with natural scenery stimuli. *Signal Processing: Image Communication*, 17(10):807-823, Nov. 2002

© CRC Press, Wu & Rao Eds, 2006

OR01	W. Osberger and A.M. Rohaly. Automatic detection of regions of interest in complex video sequences. In *Proc. SPIE*, 4299:361-372, San Jose, CA, Jan, 21-26, 2001
Pel90	E. Peli. Contrast in complex images. *J. Opt. Soc. Am. A*, 7(10)2032-2040, Oct. 1990
Pel97	E. Peli. In search of a contrast metric: Matching the perceived contrast of Gabor patches at different phases and bandwidths. *Vision Research*, 37(23):3217-3224, 1997
PS00	T.N. Pappas and R.J. Safranek. Perceptual criteria for image quality evaluation. In A. Bovik, Ed., *Handbook of Image & Video Processing*, 669-684. New York: Academic Press, 2000
Qui74	R.F. Quick, Jr. A vector-magnitude model of contrast detection. *Kybernetik*, 16:65-67, 1974
R+00	A.M. Rohaly *et al*. Video Quality Experts Group: Current results and future directions. In *Proc. SPIE*, 4067:742-753, perth, Australia, June 21-23, 2000

© CRC Press, Wu & Rao Eds, 2006

Rih96	S. Rihs. The influence of audio on perceived picture quality and subjective audiovideo delay tolerance. In *MOSAIC Handbook*, 183-187, Jan 1996
Rou89	J.A.J. Roufs. Brightness contrast and sharpness, interactive factors in perceptual image quality. In *Proc. SPIE*, 1077:209-216, LA, CA, Jan. 18-20, 1989
SB65	A.J. Seyler and Z.L. Budrikis. Detail perception after scene changes in television iamge presentations. *IEEE Trans. Information Theory*, 11(1)”61-43, 1965
Sch93	W. Schreiber. *Fundamentals of Electronic Imaging Systems*. New York: Springer, 1993
SEL00	A.E. Savakis, S.P. Etz and A.C. Loui. Evaluation of image appeal in consumer photography. In *Proc. SPIE*, 3959:111-120, San Jose, CA, Jan. 23-28, 2000
Ste96	R. Steinmetz. Human perception of jitter and media synchronization. *IEEE J. Selected Areas in Comm.*, 14(1):61-72, Jan. 1996

© CRC Press, Wu & Rao Eds, 2006

SW04	S. Süsstrunk and S. Winkler. Color image quality on the Internet. In *Proc. SPIE*, 5304:118-131, San Jose, CA, Jan. 19-22,2004. Invited paper
TGP98	K.T. Tan, M. Ghanbari and D.E. Pearson. An objective measurement tool for MPEG video quality. *Signal Processing*, 70(3):279-279, Nov. 1998
TH94	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. SPIE*, 2179:127-141, San Jose, CA, Fab. 8-10, 1994
THvdBL99	X. Tong, D. Heeger and C.J. van den Brenden Lambrecht. Video quality evaluation using ST-CIELAB. In *Proc. SPIE*, 3644:185-196, San Jose, CA, Jan. 23-29, 1999
vdBL96	C.J. van den Branden Lambrecht. *Perceptual Models and Architectures for Video Coding Applications*. Ph.D thesis, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, 1996
cdBLV96	C.J. van den Branden Lambrecht and O. Verscheure. Perceptual quality measure using a spatio-temproal model of athe human visual system. In *Proc SPIE*, 2668:450-461, San Jose, CA, Jan. 28 – Feb. 2, 1996

© CRC Press, Wu & Rao Eds, 2006

vdHB69	G. van der Horst and M. A. Bouman. Spatiotemporal chromaticity discrimination. *Journal of the Optical Society of America*, 59:1482-1488, 1969
Vka00	T. Vlachos. Detection of blocking artifacts in compressed video. *Electronics Letters*, 36(13):1106—1108, June 2000
VQE00	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, April 2000.
VQE03	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment – Phase II, Aug. 2003
Wat87	A.B. Watson. The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics and Image Processing*, 39(3):311-327, 1987
Wat95	A.B. Watson. Image data compression having minimum perceptual error. U.S. Patent 5,426,512, 1995
Wat97	A.B. Watson. Image data compression having minimum perceptual error. U.S. Patent 5,429,780, 1997

© CRC Press, Wu & Rao Eds, 2006

Wat98	A.B. Watson. Toward a perceptual video quality metric. In *Proc. SPIE*, 3299:139-147, San Jose, CA, Jan. 26-29, 1998
WBE00	Z. Wang, A.C. Bovik and B.L. Evans. Blind measurement of blocking artifacts in images. In *Proc. ICIP*, 3:981-984, Vancouver, Canada, Sep. 10-12, 2000
WBT97	A.B. Watson, R. Borthwick and M. Taylor. Image quality and entropy masking. In *Proc. SPIE*, 3016:2-12, San Jose, CA, Feb. 8-14, 1997
WC03	S. Winkler and R. Campos. Video quality evaluation for Internet streaming applications. In *Proc. SPIE*, 5007:104-115, Santa Clara, CA, Jan. 21-24, 2003
WD03	S. Winkler and D. Dufaux. Video quality evaluation for mobile applications. In *Proc. SPIE*, 5150:593-603, Lugano, Switzerlandm July 8-11, 2003
WF05	S. Winkler and C. Faller. Audiovisual quality evaluation of low-bitrate video. In *Proc. SPIE*, 5666:139-148, San Jose, CA, Jan. 16-20, 2005
WHM01	A.B. Watson, J. Hu and J.F. McGowan III. DVQ: A digital video quality metric based on human vision. *J. Electronic Imaging*, 10(1):20-29, Jan. 2001

© CRC Press, Wu & Rao Eds, 2006

Win98	S. Winkler. A perceptual distortion metric for digital color images. In *Proc. ICIP*, 3:399-403, Chicago, IL, Oct. 4-7, 1998
Win99a	S. Winkler. Issues in vision modeling for perceptual video quality assessment. *Signal Processing*, 78(2):231-252, Oct. 1999
Win99b	S. Winkler. Quality metric design: A closer look. In *Proc. SPIE*, 3959:37-44, San Jose, CA, Jan. 23-28, 2000
Win01	S. Winkler. Visual fidelity and perceived quality: Towards comprehensive metrics. In *Proc. SPIE*, 4299:114-125, San Jose, CA, Jan. 21-26, 2001
WLB04	Z. Wang, L. Lu and A.C. Bovik. Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, 19(2):121-132, Feb. 2004
WP99	S. Wolf and M.H. Pinson. Spatial-temporal distortion metrics for in-service quality monitoring of any digital video system. In *Proc. SPIE*, 3845:266-277, Boston, MA, Sep. 19-22, 1999

© CRC Press, Wu & Rao Eds, 2006

WR89	J.H.D.M. Westerink and J.A.J. Roufs. Subjective image quality as a function of viewing distance, resolution and picture size. *SMPTE J.*, 98(2):113-119, Feb. 1989
WS97	A.B. Watson and J.A. Solomon. Model of visual contrast gain control and pattern masking. *J. Opt. Soc. Am. A*, 14(9):2379-2391, Sep. 1997
WSM01	S. Winkler, A Sharma and D. McNally. Perceptual video quality and blockiness metrics for multmedia streaming applications. In *Proc. International Symposium on Wireless Personal Multimedia Communications*, 547-552, Aalborg, Denmark, Sep. 9-12, 2001. Invited paper
WV99	S. Winkler and P Vandergheynst. Computing isotropic local contrast from orientated pyramid decompositions. In *Proc. ICIP*, 4:420-424, Kobe, Japan, Oct. 25-28, 1999
WY97	H.R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 4(11):317-320, Nov. 1997

© CRC Press, Wu & Rao Eds, 2006

YBdR99	S.N. Yendrikhovskij, F.J.J. Blommaert and H. de Ridder. Towards perceptually optimal colour reproduction of natural scenes. In L.W. MacDonald and M.R. Luo, Eds., *Colour Imaging: Vision and Technology*, 363-382. New York: John Wiley & Sons, 1999
YM94	J. Yang and W. Makous. Spatiotemporal separability in contrast sensitivity. *Vision Research*, 34(19):2569-2476, 1994
ZW96	X. Zhang and B.A. Wandell. A spatial extension of CIELAB to predict the discriminability of colored patterns. In *SID Symposium Digest*, 27:731-735, 1996

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

6. Philosophy of Picture Quality Scale

M. Miyahara† and R. Kawada‡

†Japan Advance Institute of Science and Technology, Japan

‡KDDI R&D Laboratories Inc., Japan

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 6 Outline

6.1 Objective Picture Quality Scale for Image Coding

6.2 Application of PQS to a Variety of Electronic Images

6.3 Various Categories of Image Systems

6.4 Study at ITU

6.5 Conclusion

© CRC Press, Wu & Rao Eds, 2006

6.1 Objective Picture Quality Scale for Image Coding

- Objective measures
  - Useful for systematic design of coders
    - Allows for rapid evaluation of coding performance
  - Traditional scales (PSNR and MSE)
    - Effective for qualifying random errors
    - Ineffective for structured coding errors
  - need to incorporate human visual system characteristics to objective measures

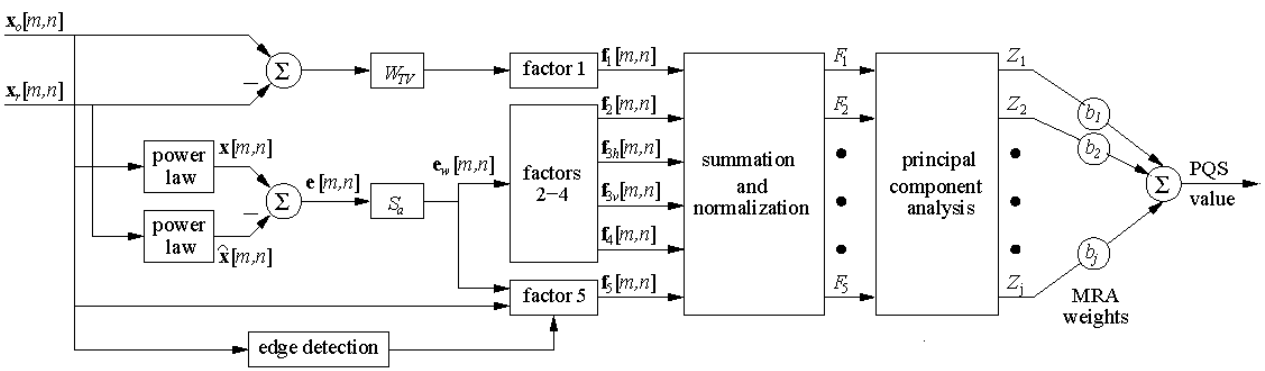
© CRC Press, Wu & Rao Eds, 2006

- Picture Quality Scale (PQS)
  - Defined as a measurement of difference between coded and original images
  - Two representations of measurement
    - 5 grade quality scale
    - 7 grade deterioration or improvement scale

© CRC Press, Wu & Rao Eds, 2006

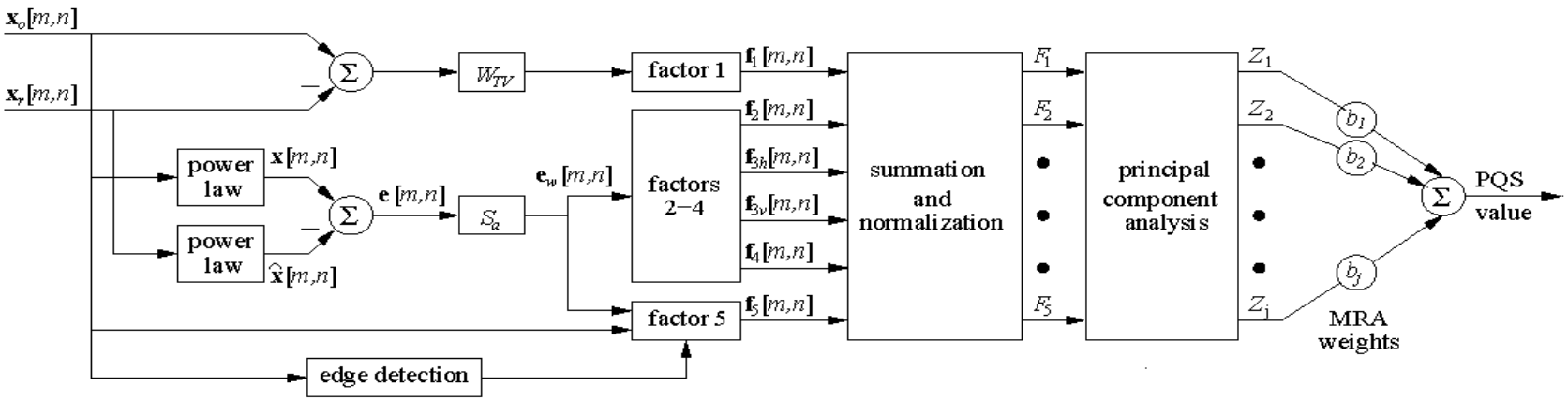
6.1.3 Construction of a Picture Quality Scale

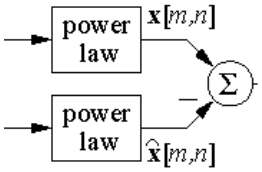
- PQS methodology
  - Compute local distortion map $\mathrm{f}_i[m,n]$
    - from distortion factor F_i
  - Quality derived from combined distortions
    - Via regression methods



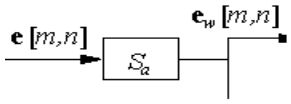
© CRC Press, Wu & Rao Eds, 2006

PQS components

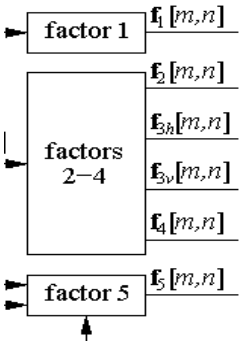




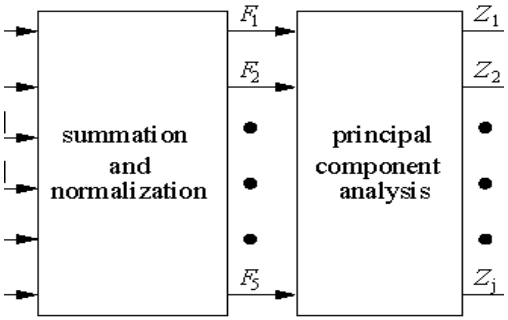
Luminance coding error



Spatial frequency weighting error



Distortion factors F_1 to F_5



Principal Component Analysis

© CRC Press, Wu & Rao Eds, 2006

- Luminance coding error
  - Transform input to uniform scale

$$\mathrm{x}[m,n] = k \bullet \mathrm{x}_o[m,n]^{1/2.2}$$

(Weber-Fechner Law approximation) with the $\mathrm{x}_o[m,n]$ original image

  - Computation of contrast adjusted error

$$\mathrm{e}[m,n] = \mathrm{x}[m,n] - \mathrm{x}_c[m,n]$$

$\mathrm{x}_c[m,n]$ is the contrast adjusted compressed image

© CRC Press, Wu & Rao Eds, 2006

- Spatial frequency weighting of errors
  - Spectral sensitivity of contrast

$$S(\omega) = 1.5e^{-(\sigma^2\omega^2)/2} - e^{2\sigma^2\omega^2}$$

$\sigma=2$, $\omega=(2\pi f)/60$, $f^2 = u^2 + v^2$

u and v are horizontal and vertical frequencies

  - High frequencies attenuation needed to account for CRT transfer function
  - High frequency response is anisotropic

© CRC Press, Wu & Rao Eds, 2006

– High frequency contrast sensitivity

$$S_a(\omega) = S(\omega) \bullet O(\omega,\theta)$$

with

$$O(\omega,\theta) = (1+e^{-\beta(\omega-\omega_o)}cos^4 2\theta)/(1+e^{-\beta(\omega-\omega_o)})$$

where

$$\theta=tan^{-1}(v/u),\ \omega_o=2\pi f_o\ ,\ \beta=8,\ f_o=11.13\ \text{cyc/deg}$$

– Weighted error is contrast adjusted error filtered with spectral sensitivity function

© CRC Press, Wu & Rao Eds, 2006

Random errors and disturbances

- Distortion factor F_1
  - from coding noise

$$e_i[m,n] = x_o[m,n] - x_r[m,n]$$

$$f_1[m,n] = (e_i[m,n] * w_{TV}[m,n])^2$$

$$w_{TV}[m,n] = 1/(1+(f/f_c)^2)$$

f_c=5.56 cyc/deg at viewing distance 4H (height)

  - Computation of F_1

$$F_1 = \frac{\sum_{m,n} f_1[m,n]}{\sum_{m,n} x_o^2[m,n]}$$

© CRC Press, Wu & Rao Eds, 2006

Random errors and disturbances

- Distortion Factor F_2
  - Due to incremental noise in slowly varying regions of images

$$f_2[m,n] = I_T[m,n](e_o[m,n] - s_a[m,n])^2$$

with $s_a[m,n]$ the spatial contrast sensitivity and $I_T[m,n]$ threshold indicator for perceptual visibility

  - Distortion F_2

$$F_2 = \frac{\sum_{m,n} f_2[m,n]}{\sum_{m,n} x_r^2[m,n]}$$

© CRC Press, Wu & Rao Eds, 2006

Structured and localized errors and disturbances

- Distortion Factor F_3 (blocking errors)
  - Along horizontal (h) and vertical (v) edges

$$f_{3h}[m,n] = I_h[m,n](e_w[m,n] - e_w[m,n\text{-}1])^2$$

$$F_{3h} \frac{1}{\sum_{m,n} I_h[m,n]} \sum_{m,n} f_{3h}[m,n]$$

  - Final distortion composed of horizontal and vertical factors

$$F_3 = \sqrt{F_{3h}^2 + F_{3v}^2}$$

© CRC Press, Wu & Rao Eds, 2006

Structured and localized errors and disturbances

- Distortion Factor F_4 (correlated errors)
  - Localized to a windowed area (w)

$$f_4[m,n] = \sum_{(k,l)\in W; l=k\neq 0} |r(m,n,k,l)|^{0.25}$$

$$r(m,n,k,l) = \frac{1}{w-1}\left[\sum_{i,j} e_w[i,j]e_w[i+k,j+l] - \frac{1}{w}\sum_{i,j} e_w[i,j]\sum_{i,j} e_w[i+k,j+l]\right]$$

  - Distortion F_4

$$F_4 = \frac{1}{MN}\sum_{m,n} f_4[m,n]$$

© CRC Press, Wu & Rao Eds, 2006

Structured and localized errors and disturbances

- Distortion Factor F_5 (masking)
  - In horizontal (S_h) and vertical (S_v) directions

$$S_h[m,n] = e^{\left\{-0.04\left(\frac{|x_o[m,n-1]-x_o[m,n+1]}{2}\right)\right\}}$$

$$f_5[m,n] = I_M[m,n]\cdot|e_w[m,n]|\cdot(S_h[m,n] + S_v[m,n])$$

  - Final distortion

$$F_5 = \frac{1}{N_K}\sum_{m,n} f_5[m,n]$$

N_K, number of 3x3 Kirsch edge $k(m,n) \geq K$, K=400

© CRC Press, Wu & Rao Eds, 2006

Principal component analysis

- Quantify combined distortion factors
  - With covariance (COV) matrix

$$C_F = COV(F) = E\{(F - \mu_F)\,(F - \mu_F)^T\}$$

where $F = (F_1, F_2, F_3, F_4, F_5)^T$ is a vector of distortion factors with μ_F is the its mean

  - Derivation of principal components (Z)
    - Obtain eigenvectors of C_F to determine individual contribution of distortion factors (F_1 to F_5)
    - De-correlation of eigenvectors of C_F
      - More effective and robust for quality assessment

© CRC Press, Wu & Rao Eds, 2006

Computation of PQS

- PQS defined as linear combination of principal components

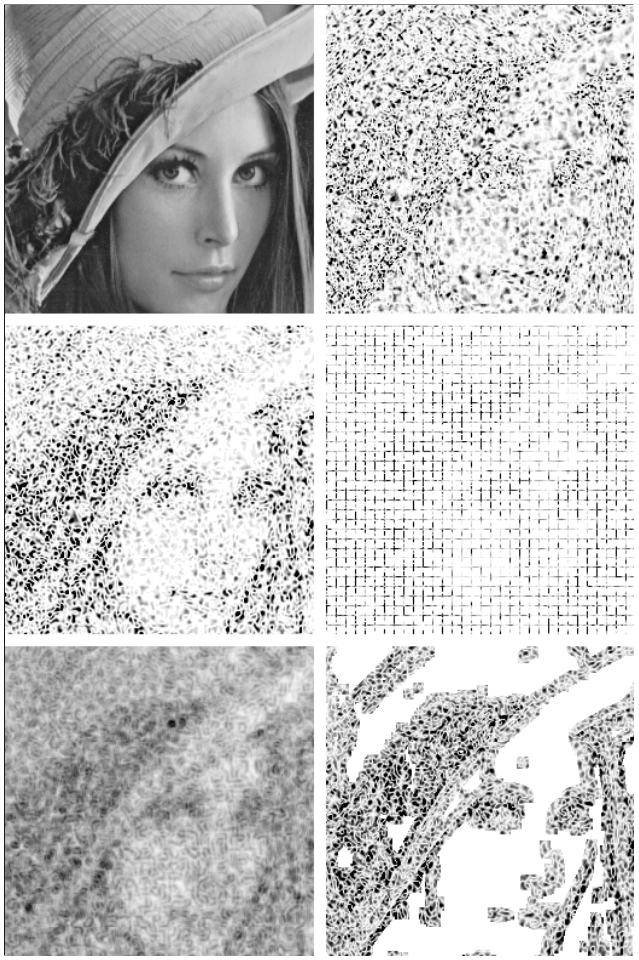
$$PQS = b_0 + \sum_{j=1}^{J} b_j Z_j$$

where b_j is the partial regression coefficient of the Multiple Regression Analysis (MRA)

- Regression analysis
  - performed on mean opinion score (MOS)
    - collected through subjective experiments

© CRC Press, Wu & Rao Eds, 2006

PQS distortion factor images from a JPEG coded image with quality factor = 15



From left to right, top to bottom:

1. Original
2. $f_1[m,n]$
3. $f_2[m,n]$
4. $\sqrt{f_{3h}^2[m,n]+f_{3v}^2[m,n]}$
5. $f_4[m,n]$
6. $f_5[m,n]$

© CRC Press, Wu & Rao Eds, 2006

6.4.1 Visual Assessment Test

- Follow ITU-R bt.500 recommendations
  - General experimental conditions
  - Subjective scale
  - Specific conditions
    - Test image resolution 256 x 256 pixels
    - 4 times image height viewing distance
    - Trained non-expert subjects
    - Grading of quality in half step increments
  - 9 subjects with 675 evaluations from 75 encoded images

© CRC Press, Wu & Rao Eds, 2006

Impairment scale (left) and experimental conditions (right)

Scale	Impairment
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

Ratio of viewing distance to picture height	4
Room illumination	None
Peak luminance on the screen	42.5 (cd/m^2)
Lowest luminance on the screen	0.23 (cd/m^2)
Time of observation	unlimited
Number of observers	9 (expert observers)

© CRC Press, Wu & Rao Eds, 2006

Test images. Clockwise from top left: *Church, Hairband, Weather, Barbara and Cameraman*

© CRC Press, Wu & Rao Eds, 2006

- Coding engines
  - DCT, Subband and Wavelet coders
- Collation of data
  - Mean Opinion Score (MOS)
    - Mean score for each of the 75 encoded images
- Data analysis
  - Compute covariance matrix
  - Compute eigenvalues and eigenvectors

© CRC Press, Wu & Rao Eds, 2006

Covariance matrix (top) and eigenvalues with corresponding eigenvectors (bottom)

	F_1	F_2	F_3	F_4	F_5
F_1	1.0000	0.9967	0.9753	0.8391	0.5823
F_2	0.9967	1.0000	0.9743	0.8231	0.5682
F_3	0.9753	0.9743	1.0000	0.8749	0.5696
F_4	0.8391	0.8231	0.8749	1.0000	0.6714
F_5	0.5823	0.5682	0.5696	0.6714	1.0000

Principal component analysis

λ_1	λ_2	λ_3	λ_4	λ_5
4.19165	0.59144	0.19021	0.02392	0.00278

ℓ_1	ℓ_2	ℓ_3	ℓ_4	ℓ_5
0.47500	-0.24449	0.27083	-0.38709	0.70100
0.47198	-0.26829	0.32384	-0.31044	-0.70993
0.47526	-0.24723	0.01921	0.84271	0.04966
0.45058	0.10598	-0.86219	-0.20065	-0.04604
0.35030	0.89213	0.27936	0.05793	-0.00216

Multiple Regression Analysis

© CRC Press, Wu & Rao Eds, 2006

- Observations of covariance matrix
  - High correlation between F_1 and F_2
    - Expected since both evaluate random errors
  - High correlation between F_3 and F_2/F_2
    - Unexpected between blocking and random errors
    - Appears to track one another as coding parameters change
- Observation of Eigenvalues
  - 99.5% of energy concentrated in λ_1, λ_2 and λ_3
    - Image quality mainly depends on F_1, F_2 and F_3

© CRC Press, Wu & Rao Eds, 2006

- Principal Component Analysis
  - Transformation from $\mathrm{F_i}$ to Z_i

$$\begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \\ Z_5 \end{bmatrix} = \begin{bmatrix} 0.47500 & 0.47198 & 0.47526 & 0.45058 & 0.35030 \\ -0.24449 & -0.26829 & -0.24723 & 0.10598 & 0.89213 \\ 0.27083 & 0.32384 & 0.01921 & -0.86219 & 0.27936 \\ -0.38709 & -0.31044 & 0.84271 & -0.20065 & 0.05793 \\ 0.70100 & -0.70993 & 0.04966 & -0.04604 & -0.00216 \end{bmatrix} \begin{bmatrix} \mathrm{F}_1 \\ \mathrm{F}_2 \\ \mathrm{F}_3 \\ \mathrm{F}_4 \\ \mathrm{F}_5 \end{bmatrix}$$

© CRC Press, Wu & Rao Eds, 2006

- Multiple Regression Analysis
  - Two expressions

$$PQS = 5.632 - 0.068Z_1 - 1.536Z_2 - 0.0704Z_3$$

$$PQS = 5.797 + 0.035F_1 + 0.044F_2 + 0.01F_3 - 0.132F_4 + 0.135F_5$$

  - Existing regression coefficient lack robustness
    - Consequently, may not operated well outside test images

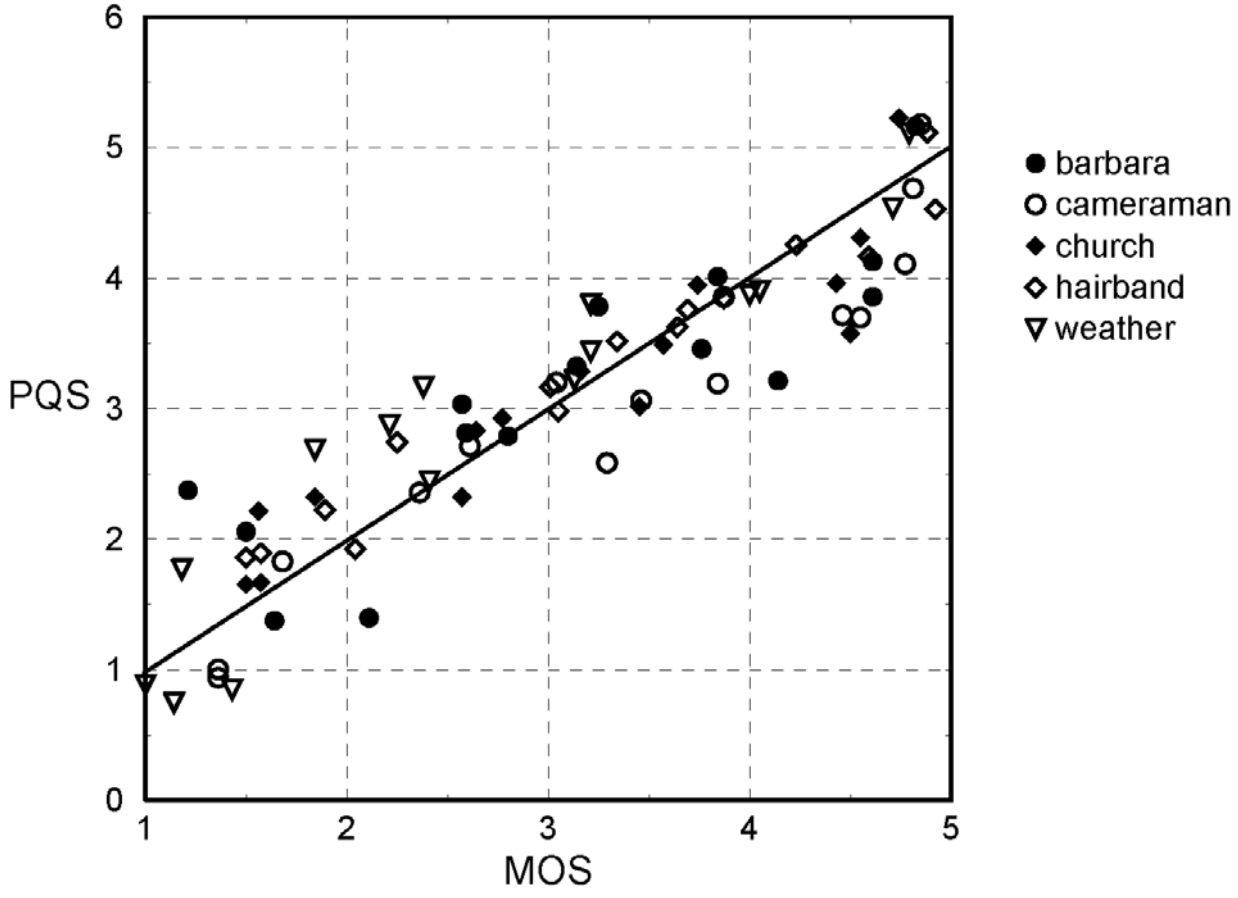
© CRC Press, Wu & Rao Eds, 2006

Evaluation of PQS

- PQS matched well with MOS
  - Better fit in mid-range MOS than extremeties
  - Correlation coefficient $\rho = 0.928$
  - For WMSE and MOS, $\rho = 0.57$
    - Calculated with F_1 alone
  - PQS absolute error about the regression line
    - Within 0.5 with 70% probability

© CRC Press, Wu & Rao Eds, 2006

PQS vs MOS scatter diagram



© CRC Press, Wu & Rao Eds, 2006

- PQS vs PSNR
  - PQS captures differences missed by PSNR between JPEG and wavelet coders
- PQS Generality and robustness
  - Some variations relative to coding technique
    - PQS evaluation slightly dependent on coding techniques

© CRC Press, Wu & Rao Eds, 2006

6.1.6 Key Distortion Factors

- Analyse distortion factors F_i
  - Determine if individual F_i is a key distortion factor
- Overall distortion characterized by
  - Z_1: the amount of error
  - Z_2: the location of error
  - Z_3: the structure of error
- PQS a linear combination of 3 errors

© CRC Press, Wu & Rao Eds, 2006

Correlation between PQS and MOS

Factors					ρ	ρ^*	$E\{\|error\|\}$
1					0.5665	0.5582	0.8347
	2				0.5373	0.5282	0.8510
		3			0.5714	0.5633	0.8269
			4		0.7426	0.7384	0.6551
				5	0.8985	0.8971*	0.4268
1	2				0.5547	0.5461	0.8418
1		3			0.5729	0.5564	0.8289
1			4		0.7499	0.7417	0.6362
1				5	0.9001	0.8972	0.4206
	2	3			0.5779	0.5616	0.8222
	2		4		0.7539	0.7459	0.6336
	2			5	0.8991	0.8962	0.4236
		3	4		0.7600	0.7522	0.6302
		3		5	0.9015	0.8986	0.4144
			4	5	0.9180	0.9156*	0.3773
1	2	3			0.5716	0.5550	0.8265
1	2		4		0.7516	0.7435	0.6352
1	2			5	0.8996	0.8967	0.4220
1		3	4		0.7600	0.7522	0.6302
1		3		5	0.9015	0.8986	0.4144
1			4	5	0.9271	0.9239	0.3536
	2	3	4		0.7600	0.7522	0.6302
	2	3		5	0.9015	0.8986	0.4144
	2		4	5	0.9302	0.9271*	0.3460
		3	4	5	0.9279	0.9247	0.3551
1	2	3	4		0.7600	0.7522	0.6302
1	2	3		5	0.9015	0.8986	0.4144
1	2		4	5	0.9285	0.9253*	0.3501
1		3	4	5	0.9279	0.9247	0.3551
	2	3	4	5	0.9279	0.9247	0.3551
1	2	3	4	5	0.9279	0.9247	0.3551

- Different combination of PQS distortion factors and their associated correlation to MOS

© CRC Press, Wu & Rao Eds, 2006

- Contribution of distortion factors
  - F_5 may operate by itself
    - Perhaps for evaluating high quality pictures
  - F_1, F_2 and F_3 are insufficient by themselves
- Other distortion factors to consider
  - Jagged distortions

© CRC Press, Wu & Rao Eds, 2006

6.1.7 Discussion

- Limitations
  - Assigning single quality score may not reflect picture quality clearly
- ITU-R bt.500 scales and methods
  - 5 point scale lack resolution
  - Need for anchor images
- Modelling human vision
  - Importance of visual characteristics for quality metrics

© CRC Press, Wu & Rao Eds, 2006

- Specialization of PQS
  - Can calibrated to evaluate specific coding systems
- PQS of colour picture coding
  - Evaluates colour distortions in addition to the usual distortions
- Applications for PQS
  - Adaptive coders with quality feedback loop

© CRC Press, Wu & Rao Eds, 2006

6.2 Application of PQS to a Variety of Electronic Images

- Categories of electronic media
  - Conventional TV
  - High Definition TV
  - Extra high quality images
    - Digital cameras, specialized image capture devices
  - Handheld devices
    - cellular phones, PDA
  - Computers (computer graphics)

© CRC Press, Wu & Rao Eds, 2006

- Subjective picture evaluation
  - Purpose of evaluation. Measure what?
    - Fidelity, impairment or picture quality
  - Picture spatial resolution and viewing distance
  - Constancy of viewing distance
    - Viewing distance as a function of display height
  - Viewing angle between adjacent pixels
  - Other viewing conditions
    - Ambient light, noise, etc

© CRC Press, Wu & Rao Eds, 2006

6.3 Various Categories of Image Systems

- Standard TV
  - NTSC, PAL, SECAM
- HDTV and Super HDTV
  - 1000 x 1600 pixel HDTV
  - 2000 x 4000 pixel super HDTV
- Extra High Quality Images
- Cellular Phone type
- Personal computer

© CRC Press, Wu & Rao Eds, 2006

Encoding parameters of digital TV for studios

Encoding parameter values for the 4:2:2 of the family

Parameters	525-line, 60 field/s systems	625-line, 50 field/s systems
1. Coded signals	*Y, R-Y, B-Y*	
2. Number of samples per total line: - luminancd signal (*Y*) - each colour-difference signal (*R-Y, B-Y*)	 858 429	 864 432
3. Sampling structure	Orthogonal, line, field and picture repetitive. *R-Y*and *B-Y* samples co-sited with odd(1st, 3rd, 5th, etc.) *Y* samples in each line.	
4. Sampling frequency: - luminancd signal - each colour-difference signal	 13.5 MHz 6.75 MHz	
5. Form of coding	Uniformly quantized PCM, 8bits per sample, for the luminance signal and each colour-difference signal	
6. Number of samples per digital active line: - luminance signal - each colour-difference signal	 720 360	
7. Correspondence between video signal levels and quantization levels: - luminance signal - each colour-difference signal	 220 quantization levels with the black level corresponding to level 16 and the peak white level corresponding to level 235 224 quantization levels in the centre part of the quantization scale with zero signal corresponding to level 128	

© CRC Press, Wu & Rao Eds, 2006

6.4 Study at ITU

- Some recent Recommendations
  - J.143
  - J.133
  - J.147
  - J.148

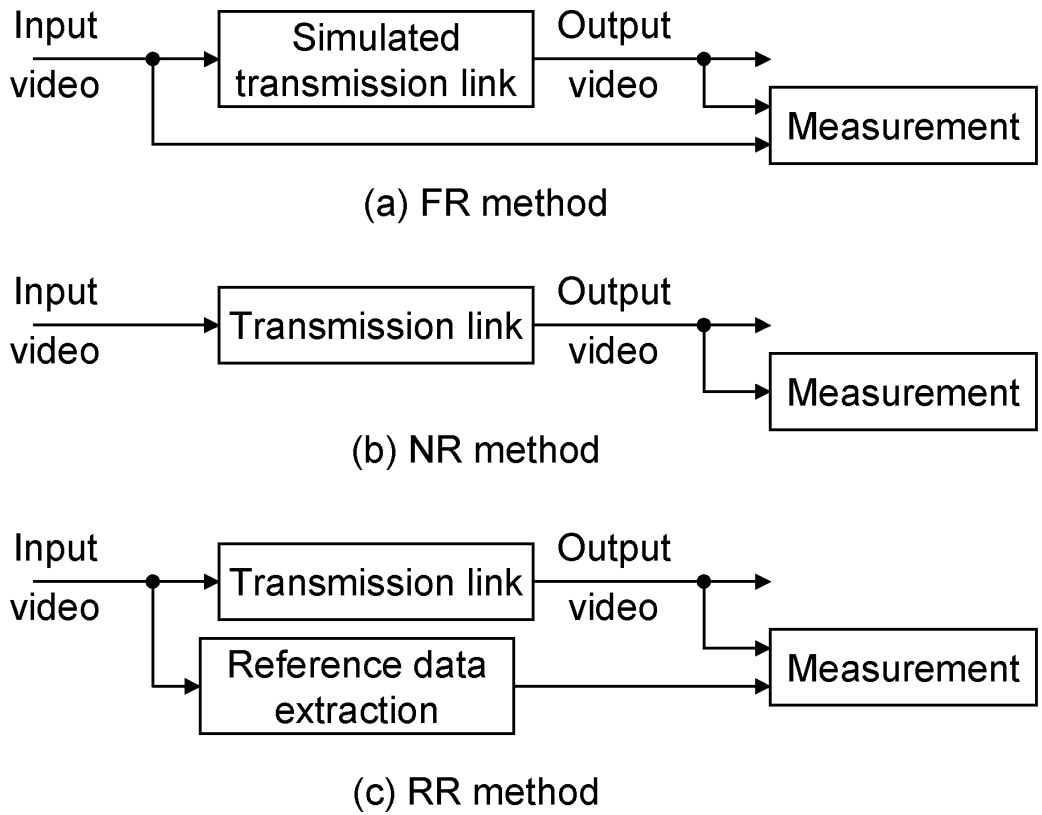
© CRC Press, Wu & Rao Eds, 2006

Recommendation J.143

- Move towards standardization of assessment models
  - Full Reference (FR)
  - Reduced Reference (RR)
  - No Reference (NR)
- FR
  - Compares between processed and reference
  - Intended for TV standard evaluations
  - May be deploy on handheld devices in future

© CRC Press, Wu & Rao Eds, 2006

Concept of FR, NR and RR methods



© CRC Press, Wu & Rao Eds, 2006

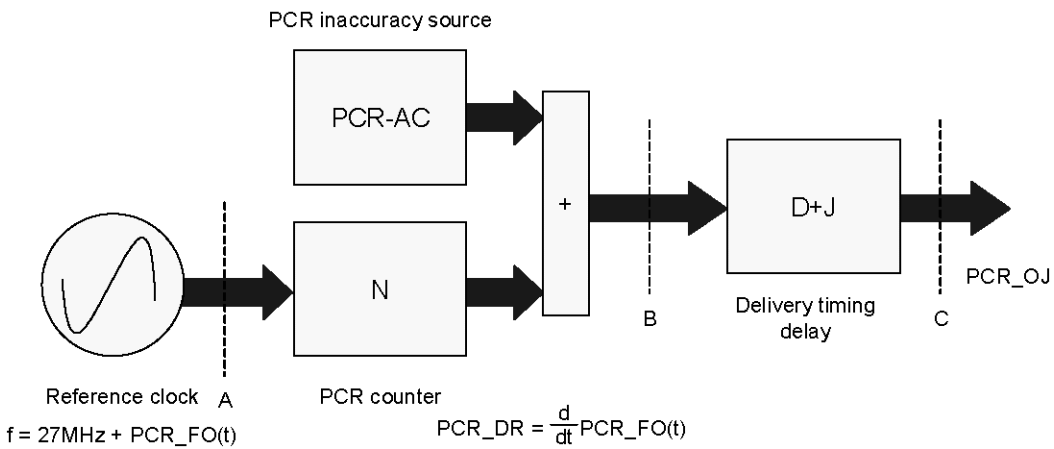
- NR
  - Evaluates based on processed only
  - No effective generic NR metric has been established
- RR
  - Evaluation based on extracted feature of reference pictures
  - Intended for monitoring TV transmission

© CRC Press, Wu & Rao Eds, 2006

Recommendation J.133

- Measurement of MPEG-2 transport stream in networks
  - For television transmission

Below: concept of J.133. PRC (program reference clock)

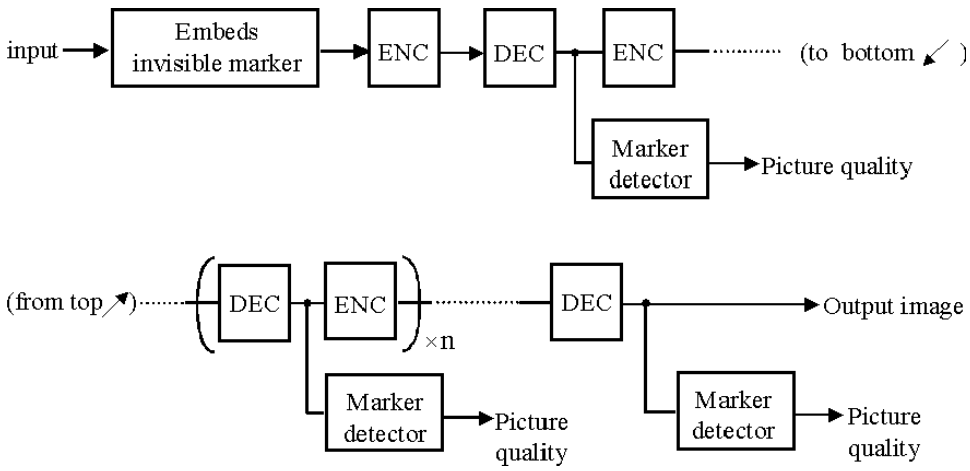


© CRC Press, Wu & Rao Eds, 2006

Recommendation J.147

- Objective picture quality measurement method by use of in-service test signals
  - Can operate in either FR, RR or NR modes

Below: Concept of J.147

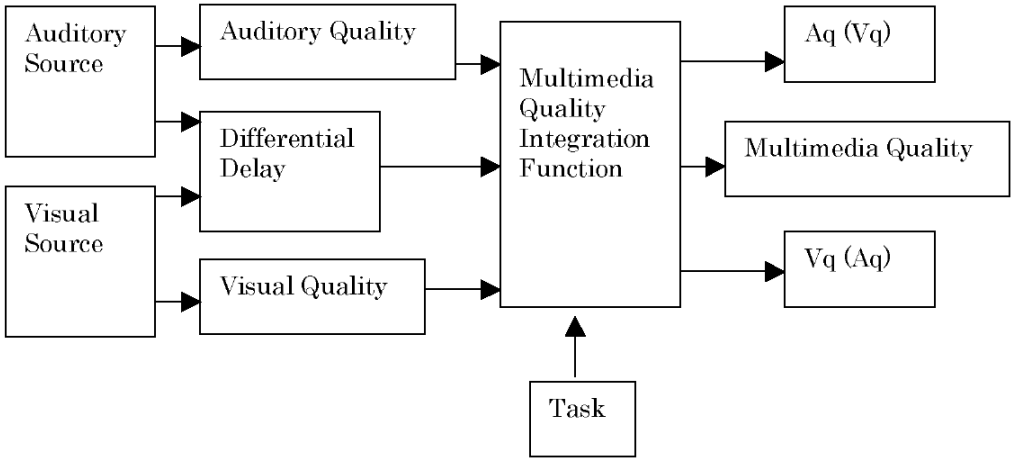


© CRC Press, Wu & Rao Eds, 2006

Recommendation J.148

- Requirements for an objective perceptual multimedia quality model
  - Considers both audio and video

Below: Concept of J.148



© CRC Press, Wu & Rao Eds, 2006

6.5 Conclusion

- Picture quality evaluation is a difficult task
  - Subjectively - *perception, viewing conditions, time factor, etc*
  - Objectively – *Modelling/prediction issues*
- PQS: an objective quality assessment
  - For still images, but extendable to video
  - Operates on 5 or 7 point scale
  - Quantifies different types of distortions
  - Correlates well with MOS

© CRC Press, Wu & Rao Eds, 2006

- Moves in standardization bodies
  - Moves towards perceived picture quality
  - Enhancing picture evaluation tools
    - According to applications
      - FR, NR and RR metrics

© CRC Press, Wu & Rao Eds, 2006

AA96	N. Avadhanam and V.R. Algazi. Prediction and measurement of high quality in still image coding. In *Proceedings of SPIE Very High Resolution and Quality Imaging Conference, EI 96*, 2663:100-109, San Jose, CA, 1996
AFC95	V.R. Algazi, G.E. Ford and H. Chen. Linear filtering of images based on properties of vision. *IEEE Transaction on Image Processing*, 4(10):1460-1464, Oct. 1995
AFMN93	V.R. Algazi, G.E. Ford, M. Mow and A. Najmi. Design of subband coders for high quality based on perceptual criteria. In *Proceedings of SPIE, Applications of Digital Image Processing XVI*, 2028:40-49, SPIE. San Diego, CA, 1993
All83	J. Allnatt. *Transmitted-picture Assessment*, New York: John Wiley & Sons, 1983
AN93	A.J. Ahumada, Jr. and C. H. Null. Image quality: A multidimensional problem. In A.B. Watson, Ed., *Digital Image and Human Vision*. Cambridge, MA: The MIT Press, 1993

© CRC Press, Wu & Rao Eds, 2006

ANS96	ANSI T1.801.03-1996. *Digital transport of one-way video signals-parameters for objective performance assessment*, February 1996
AV98	N. Avadhanam and V.R. Algazi. Anchored subjective quality assessment scale for still image coding. *Submitted to Signal Processing*, Feb. 1998
Bed79	R.A. Bednarek. On evaluation impaired television pictures by subjective measurements. *IEEE Trans. On Broadcasting*, BC-25(2):41-46, 1979
C.65	C.H. Graham *et al*. *Vision and Visual Perception*. New York: John Wiley & Sons, 1965
CCI82	CCIR. Rec. 567-1. Transmission performance of television circuits designed for use in international connections, pl-38. In Recommensations and reports of the CCIR and ITU, Geneva, 1982
CKL66	F.W. Campbell, J.J Kulikowsky and J.Z. Levinson. The effect of orientation on the visual resolution of gratings. *J. Physiology*, 187:427-436, 1996
CP77	S. Chatterjee and B. Price. *Regression Analysis by Example*. New York: John Wiley and Sons, 1997

© CRC Press, Wu & Rao Eds, 2006

Dal93	S. Daly. The visible difference predictor: An algorithm for the assessment of image quality. In A.B. Watson, Ed., *Digital Image and Human Vision*, Cambridge, MA: The MIT Press, 1993
Dau88	I. Daubechies. Orthonormal bases of compactly supported wavelets. *Commun. On Pure and Applied Math.*, 41:909-996, Nov. 1988
ETS01	ETSI TR 101 290, *Digital Video Broadcasting (DVB); Measurement guidelines for DVB systems*, V.1.2.1, May 2001
Fuj73	T. Fujio. Two-dimensional processing of TV signal. Technical Report Vol. 25, NHK Technical Research Lab., Mar. 1973
G+88	B. Girod *et al*, A subjective evaluation of noise-shaping quantization for adaptive intra/interframe dpcm coding of color television signals. *IEEE Trans. on Communications*, COM-36(3):332-346, Mar 1988
HA84	T.A. Hentea and V.R. Algazi. Perceptual models and the filtering of high0contrast achromatic images. *IEEE Trans. On Systems, Man. And Cybernetics*, SMC-14(2):230:246, Mar. 1984

© CRC Press, Wu & Rao Eds, 2006

HM87	Y. Horita and M. Muyahara, Image coding and quality estimation in uniform perceptual space. IECE Technical Report IE87-115, IECE, Jan. 1987
HMM99	T. Hamadam S. Miyaji and S, Matsumoto. Picture quality assessment system by three-layered bottom-up noise weighting considering human visual perception. *SMPTE Journal*, 108(1):20-26, Jan.1999
HPN97	B.G. Haskell, A. Puri and A.N. Netravali. *Digital Video: An Introduction to MPEG-2*. New York: Chapman & Hall, 1997
Ind	ITE. *ITE Television System Chart (version II). Digital Standard Picture*. ITE, December 1987
ITU00a	ITU. Rec. ITU-R BT.500-10, ITU Rec. BT.601-5, ITU-R BT.709-3 (details are described in the text.). 2000
ITU00b	ITU-R Recommendation BT.500-10. *Methodology for the subjective assessment of the quality of television pictures*, 2000 (Question ITU-R 211/11)
ITU00c	ITU-R Recommendation BT.500-10. *Methodology for the subjective assessment of the quality of television pictures*, may 2000

© CRC Press, Wu & Rao Eds, 2006

ITU00d	ITU-R Report BT.2020-1. *Objective quality assessment technology in a digital environment*, Sept. 2000
ITU00e	ITU-R Recommendation J.143. *User requirements for objective perceptual video quality measurements in digital cable television*, May 2000
ITU00f	ITU-R Recommendation J.144. *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*, Mar. 2000
ITU02a	ITU-R Recommendation J.133. *Measurement of MPEG-2 transport streams in networks*, July 2000
ITU02b	ITU-R Recommendation J.146. *Loop latency issues in contribution circuits for conversational TV programmes*, July 2002
ITU02c	ITU-R Recommendation J.147. *Objective picture quality measurement method by use of in-service test signals*, July 2002
ITU03	ITU-R Recommendation J.148. *Requirements for an objective perceptual multimedia quality model*, May 2003

© CRC Press, Wu & Rao Eds, 2006

I.U82	I. Uyama. II fundamental requirement for high-definition television systems, II-1 large screen effects. *NHK Monograph*, 32:14-20, June 1982
Jai89	A.K. Jain. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989
J.B98	J. Baïna *et al*, Quality of MPEG-2 Signal on a simulated digital terrestrial television. *IEEE*, BR-44(4):381-391, Dec. 1998
JJS93	N. Jayant, J. Johnston and R. Safranek. Signal compression based on models of human perception. *Proceedings of the IEEE*, 81(10)”1385-1422, Oct. 1993
JM86	B.L. Jones and P.M. McManus. Graphics scaling of qualitative terms. *SMPTE Journal*, 95:1166-1171, Nov. 1986
Joh80	J.D. Johnston. A filter family designed for use in quadrature mirror filters. In *Proceedings of IEEE ICASSP*, 291-294, 1980

© CRC Press, Wu & Rao Eds, 2006

K+85	M. Kunt *et al*, Second-generation image-coding techniques. *Proceedings of the IEEE*, 73:549-574, April 1985
Ken75	M. G. Kendall. *Multivariate Analysis*. London: Charles Griffin, 1975
KGMA95	K. Kotani, Q. Gan, M. Miyahara and V.R. Algazi. Objective picture quality scale for color image coding. In *Proceedings of International Conference on Image Processing*, 3:133-136, 1995
KK95	S.A. Karunasekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transaction on Image Processing*, 4(6):713-724, June 1995
KM96a	V. Kayargadde and J.B. Martens. Perceptual characterization of images degraded by blur and noise: experiments. *Journal of the Optical Society of America A*, 13(16):1166-1177, June 1996
KM96b	V. Kayargadde and J.B. Martens. Perceptuakl characterization of images degraded by blur and noise, 13(6):1178-1188, June 1996
LA65	N.W. Lewis and J.A. Allnutt. Subjective quality of television pictures with multiple impairments. *Electronic Letters*, 1:187-188, July 1965

© CRC Press, Wu & Rao Eds, 2006

LAE95	J. Lu, V.R. Algazi and R.R. Estes, Jr. A comparison of wavelet image coders using a picture quality scale. In *Proc. of SPIE Wavelet Applications Conference*, Orlando, FL, April 1995
LAE 96	J. Lu, V.R. Algazi and R.R. Estes, Jr. A comparative study of wavelet image coders. *Optical Engineering*, 9:2605-2619, Sept. 1996
LB82	F.X.J. Lukas and Z.L. Budrikis. Picture quality prediction based on a visual model. *IEEE Trans. on Communications*, COM-30(8):1679-1692, July 1982
Lim78	J.O. Limb. On the design of quantizer for dpcm coder: A functional relationship between visibility, probability and masking. *IEEE Transaction on Communications*, COM-26(5):573-578, May 1978
Lim79	J.O. Limb. Distortion criteria of the human viewer. *IEEE Trans. on System, Man. and Cybernetics*, SMC-9(12):778-793, Dec. 1979
MIS+98	M. Miyahara, T. Ino, H. Shirai, S. Taniho and R. Algazi. Important factors to convey high irder sensation. *IEEE Trans. on Communications*, E81-B(11):1966-1973, Nov. 1998

© CRC Press, Wu & Rao Eds, 2006

Miy88	M. Miyahara. Quality assessments for visual service. *IEEE Communications Magazine*, 26(10):51-60, Oct. 1988
MK96	J.B. Martens and V. Kayargadde. Image quality prediction in a multidimensional perceptual space. In *Proc. of ICIP*, Lausanne, Switzerland, 1:877-880, 1996
MKA98	M. Miyahara, K. Kotani and V.R. Algazi. Objective picture quality scale (PQS) for image coding. *IEEE Trans. on Communications*, 46(9):1215-1226, Sept. 1998
MY88	M. Miyahara and Y. Yoshida. Mathematical transformation of (R,G,B) color data to munsell (H,V,C) color data. In *SPIE Visual Communications and Image Processing III*, 1001-1118, Nov. 1998
Net77	A. N. Netravali. Interpolative picture coding using a subjective criterion. *IEEE Transactions on Communications*, COM-25(5):503-208, May 1977
O+87	A. Oosterlick *et al*, *Image Coding Using the Human Visual System*, Chap. 5. Belgium: Katholieke Universiteit, 1987

© CRC Press, Wu & Rao Eds, 2006

ORMS01	O. Sugimoto, R. Kawada, M. Wada and S. Matsumoto. A method for estimating PSNR of coded pictures using embedded invisible markers without reference pictures. In *Proc. of ISPACS*, 42-46, Nashville, Tennessee, USA, November 2001
Pro95	Proc. MOSAIC Workshop: Advance methods for the evaluation of television picture quality. Proceedings of the mosaic workshop, Institute for Perception Research, Eindhoven, Netherlands, 1995
PW84	G.C. Phillips and H.R. Wilson. Orientation bandwidth of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A*, 1:226-231, 1984
Roh00	A.M. Rohaly *et al*. Video Quality Experts Group: Curren results and future directions. In *Proc. SPIE*, 4067:742-753, Perth, Australia, 2000
Sak77	D.J. Sakrison. On the role of the observer and a distortion measure in image transmission. *IEEE Trans. on Communications*, COM-25(11), Nov. 1977
S.O97	S. Olsson *et al*. Objective methods for assessment of video quality: start of the art. *IEEE*, BR-43(4):487-495, Dec. 1997

© CRC Press, Wu & Rao Eds, 2006

TH94	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. ICIP*, 2:982-986, 1994
vDMW95	A.M. van Dijk, J.B. Martens and A.B. Watson. Quality assessment of coded images using numerical category scaling. In *Proc. SPIE Advanced Image and Video Communications and Storage Technologies, Amsterdam*, 2451:90-101, 1995
VQE	VQEG, *The video Quality Experts Group Website*. http://www.its.bldrdoc.gov.vqeg/.
E+83	R. Wilson *et al*, Anisotropic non-stationary image estimation and its applications: Part II – Predictive image coding. *IEEE Trans. on Communicaitons*, COM-31(3):398-406, Mar. 1983
Wat93a	A.B. Watson. DCT quantization matrices visually optimized for individual images. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 1913:202-216, 1993

© CRC Press, Wu & Rao Eds, 2006

Wat93b	A.B. Watson, Ed. *Digital Image and Human Vision*, Cambridge, MA: The MIT Press, 1993
WLB95	S.J.P. Westen, R.L. Lagendijk and J. Biemond. Perceptual image quality based on a multiple channel HVS model. In *International Conference on Acoustics, Speech and Signal Processing*, 4:2351-2354, 1995
YH68	M. Yasuda and K. Hiwatashi. A model of retinal neural networks and its spatio-temporal characteristics. *Japanese Journal of Medical Electronics and Biological Engineering*, 6(1):53-62, 1968
ZH89	C. Zetzsche and G, Hauske. Multiple channel model for the prediction of subjective image quality. In *SPIE Proceedings*, 1077:209-216, 1989

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

7. Structural Similarity Based Image Quality Assessment

Z. Wang[†], A. Bovik[‡] and H.R. Sheikh[§]

† University of Texas at Arlington, USA

‡ University of Texas at Austin, USA

§ Texas Instrument Inc., USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 7 Outline

7.1 Structural Similarity and Image Quality

7.2 The Structural SIMilarity (SSIM) Index

7.3 Image Quality Assessment Based on the SSIM Index

7.4 Discussions

© CRC Press, Wu & Rao Eds, 2006

Introduction

- Properties of images
  - Highly structured
    - Due to strong spatial dependency of pixels
    - Holds important information of objects
- Structural Similarity assessment
  - Assumes the HVS is adaptive to structural information
  - Consequently, quantifying structural similarity should account for perceived quality

© CRC Press, Wu & Rao Eds, 2006

7.1 Structural Similarity and Image Quality

- Discussions centred on Full Reference (FR) metric
  - Test image evaluated based on a reference
    - Reference image is assumed to have perfect quality
- Classification of statistics based quality measures
  - Error sensitive approach
  - Feature/structural extraction

© CRC Press, Wu & Rao Eds, 2006

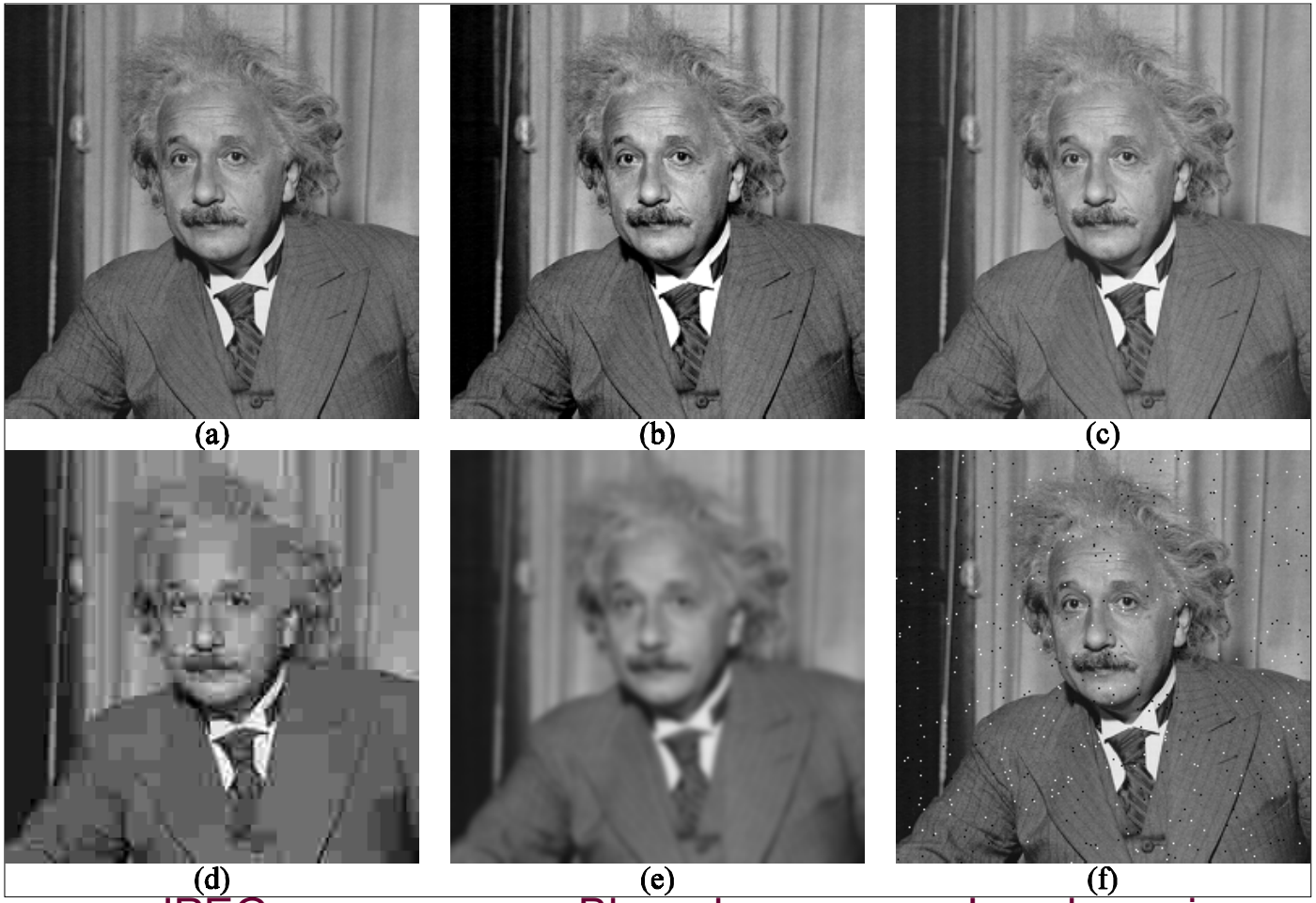
- Error sensitive
  - Error signal defined as the difference between test and reference image. e.g. MSE
  - Assumes quality loss as a linear function of error signal
- Feature/structure extraction
  - More attenuated to the HVS
  - Distortion as a function of structural changes
  - Measure structural deterioration

© CRC Press, Wu & Rao Eds, 2006

Original Einstein

Contrast stretched
MSE = 144
MSSIM = 0.9884

Mean shifted
MSE = 144
MSSIM = 0.9884



JPEG
MSE = 142
MSSIM = 0.6624

Blurred
MSE = 144
MSSIM = 0.6940

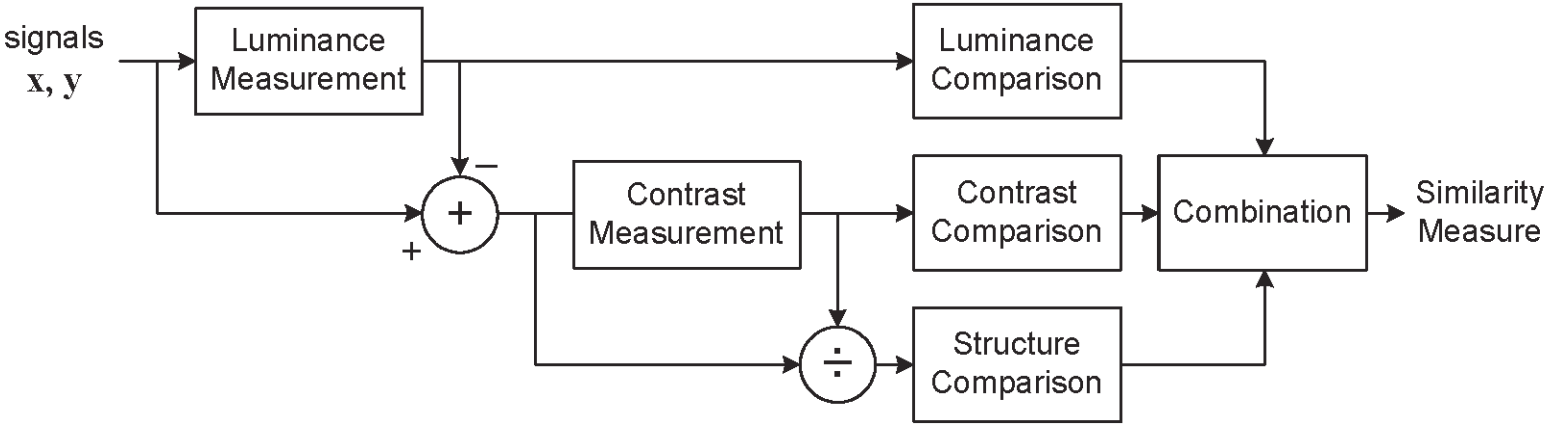
Impulse noise
MSE = 144
MSSIM = 0.8317

© CRC Press, Wu & Rao Eds, 2006

7.2 The Structural SIMilarity (SSIM) Index

- Similarity measurement in 3 components
  - Luminance, Contrast and structure

Below: Illustration of the SSIM system



© CRC Press, Wu & Rao Eds, 2006

- Luminance similarity
  - Mean intensity $$\mu_x = \bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$
  - Luminance comparison $l(\mathbf{x},\mathbf{y})$ of test and reference signals $\mathbf{x}$ and $\mathbf{y}$, respectively, is

$$l(\mathbf{x},\mathbf{y}) = l(\mu_x,\mu_x)$$

$$\mathbf{x} = \{x_i \mid i = 1,2,\ldots, N\},\ \mathbf{y} = \{y_i \mid i = 1,2,\ldots,N\}$$

© CRC Press, Wu & Rao Eds, 2006

- Contrast similarity
  - Remove mean intensity $\mathbf{x} = \mathbf{x} - \mu_x$
    - Projects of vector $\mathbf{x}$ onto hyperplane
  - Estimate contrast with standard deviation, σ

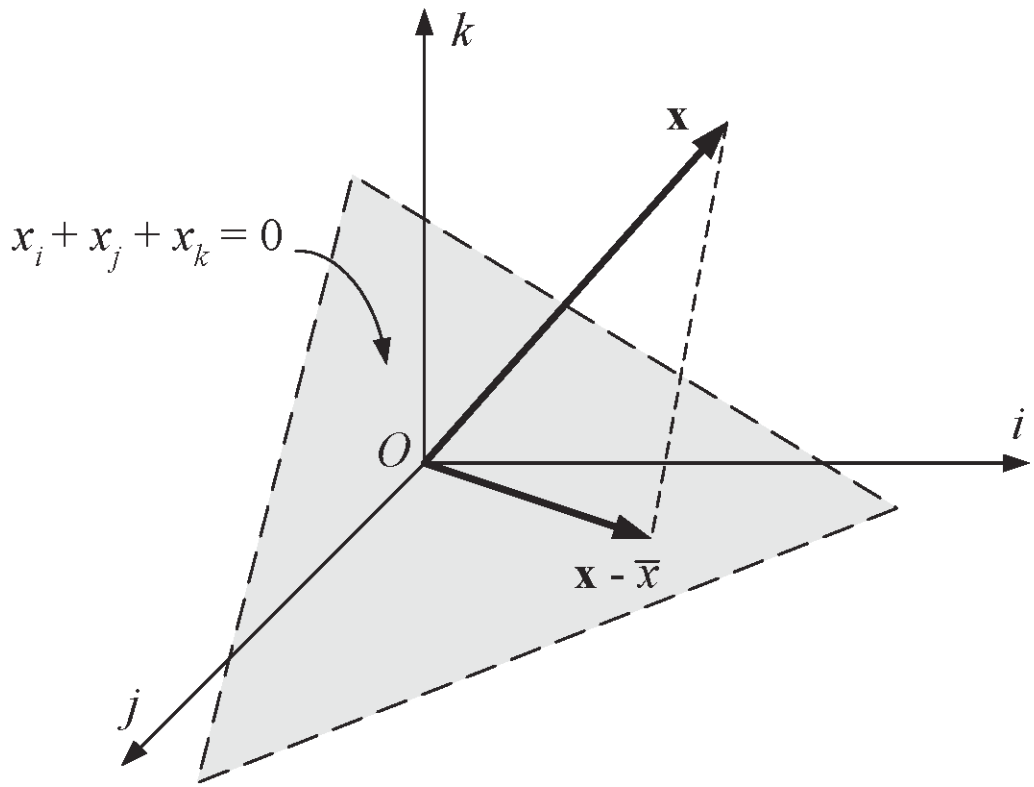
$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)^2 \right)$$

  - Contrast comparison given by

$$c(\mathbf{x}, \mathbf{y}) = c(\sigma_x, \sigma_y)$$

© CRC Press, Wu & Rao Eds, 2006

Projection onto hyperplane



© CRC Press, Wu & Rao Eds, 2006

- Structural similarity
  - Normalization by σ (range bounded to 0 - 1)
  - Structural comparison

$$s(\mathbf{x},\mathbf{y}) = s\left(\frac{\mathbf{x}-\mu_x}{\sigma_x}, \frac{\mathbf{y}-\mu_y}{\sigma_y}\right)$$

- Overall similarity measure

$$S(\mathbf{x},\mathbf{y}) = f(l(\mathbf{x},\mathbf{y}),\ c(\mathbf{x},\mathbf{y}),\ s(\mathbf{x},\mathbf{y}))$$

© CRC Press, Wu & Rao Eds, 2006

- Similarity definition $f(\mathbf{.})$
  - Must satisfy 3 conditions
    - Symmetry: $S(\mathbf{x},\mathbf{y}) = S(\mathbf{y},\mathrm{x})$
      - Exchanging order of inputs should not affect output
    - Boundedness: $S(\mathbf{x},\mathbf{y}) \leq 1$
    - Unique maximum: $S(\mathbf{x},\mathbf{y}) = 1$ if and only if $\mathbf{x} = \mathbf{y}$

© CRC Press, Wu & Rao Eds, 2006

- Luminance definition

$$l(\mathbf{x},\mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

  - C_1 Controls instability when both μ_x and μ_y are close to 0

$$C_1 = (K_1 L)^2$$

  - L is the dynamic range of pixel values (255) and $K_1 << 1$, a small constant

© CRC Press, Wu & Rao Eds, 2006

- $l(\mathbf{x},\mathbf{y})$ is related to Weber's law
- Contrast definition

$$c(\mathbf{x},\mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

where $C_2 = (K_2 L)^2$; $C_2 \geq 0$ and $K_2 << 1$

© CRC Press, Wu & Rao Eds, 2006

- Structural definition

$$s(\mathbf{x},\mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$

where σ_{xy} is defined as

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y)$$

© CRC Press, Wu & Rao Eds, 2006

- SSIM definition

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = [l(\mathbf{x},\mathbf{y})]^{\alpha} \cdot [l(\mathbf{x},\mathbf{y})]^{\beta} \cdot [l(\mathbf{x},\mathbf{y})]^{\gamma}$$

Where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are gains for individual components

- Setting conditions $\alpha=\beta=\gamma=1$ and $C_3=C_2/2$

$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = \frac{2(\mu_x\mu_y + C_1)(\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

© CRC Press, Wu & Rao Eds, 2006

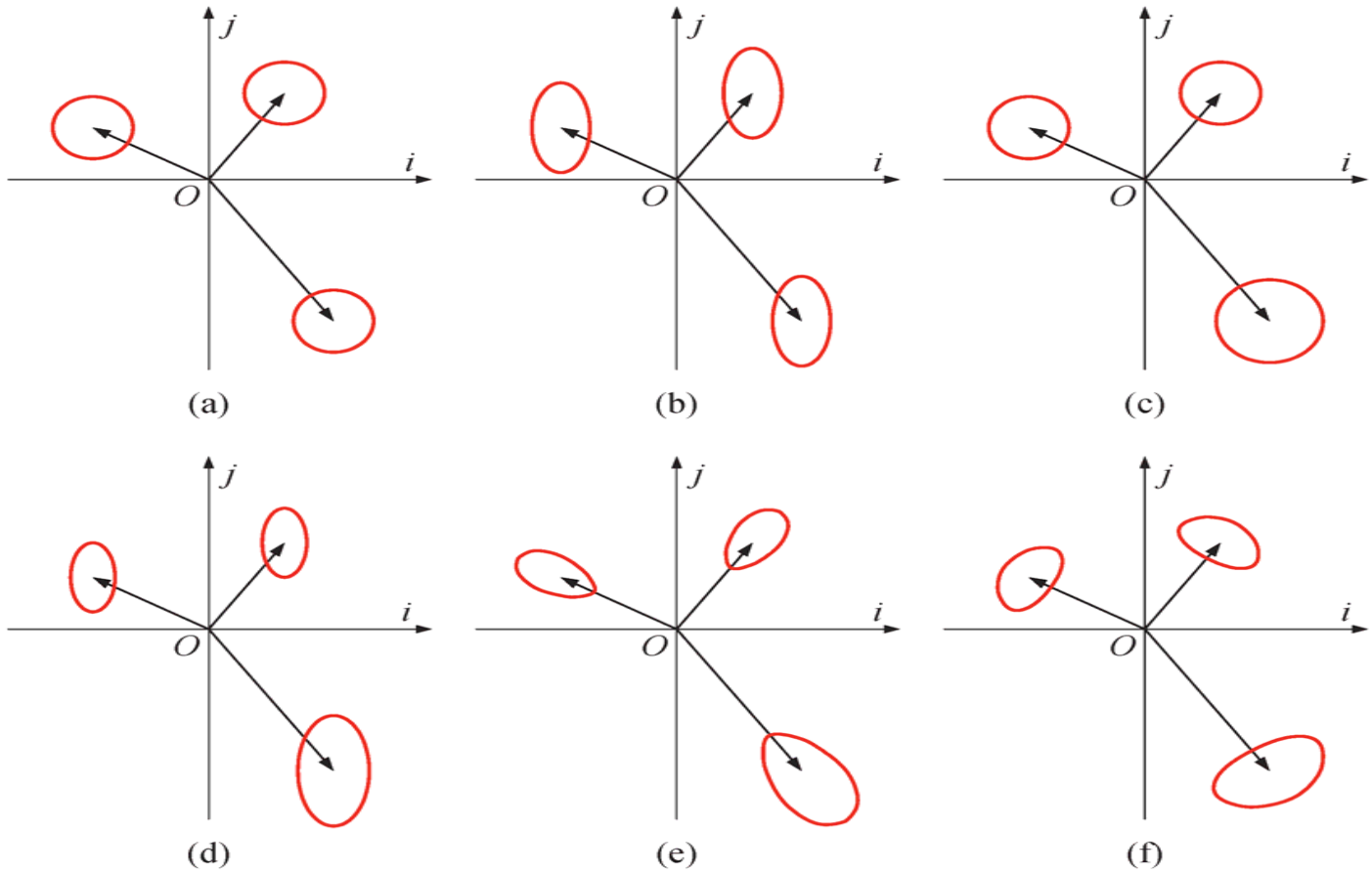
- Distortion Contour
  - Geometric projection of distortion in vector space
  - Contours area represents error
  - Distortion contours changes according to the profile of the error functions
    - Shape, size/volume

© CRC Press, Wu & Rao Eds, 2006

Distortion Contour for Minkowski error

Magnitude weighted Minkowski

Component weighted Minkowski



Magnitude and component weighted Minkowski

SSIM with component emphasis

SSIM with structure emphasis

© CRC Press, Wu & Rao Eds, 2006

7.3 Image Quality Assessment Based on the SSIM Index

- SSIM more effective for localized evaluation
  - Statistical features of images are highly non-stationary
  - Image distortion may be space-variant
  - Focus of the HVS localized on an area
    - Due to viewing angle of the foveal area
  - Provides spatially varying quality maps
    - Different areas with different quality levels

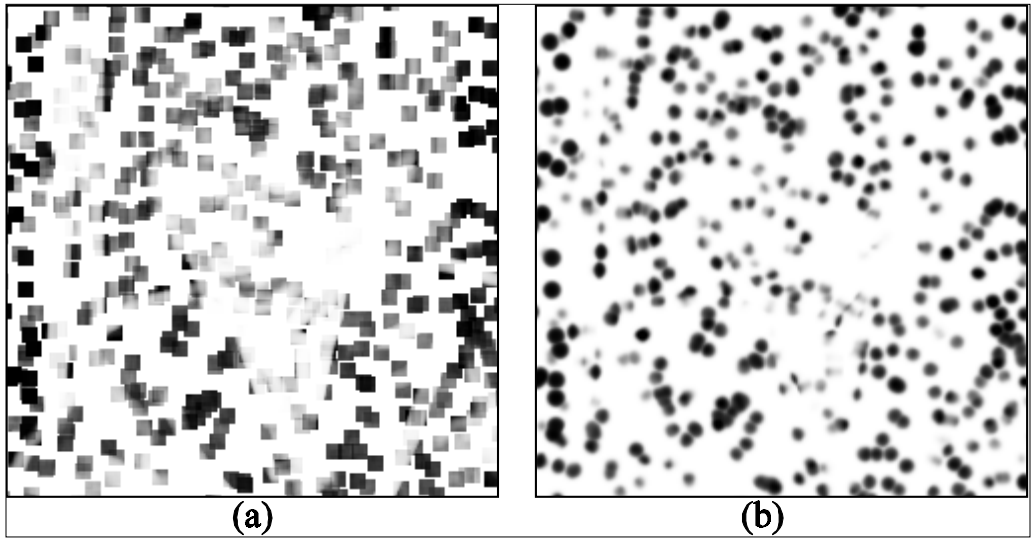
© CRC Press, Wu & Rao Eds, 2006

- SSIM windows
  - Squared 8 x 8 windows
    - Leads to blocking in index map
  - Weighted circular Gaussian window
    - Alleviates blocking problem
    - weights $\mathbf{w} = \{w_i \mid i = 1,2,\ldots, N\}$ bounded by

$$\sum_{i=1}^{N} w_i = 1$$

© CRC Press, Wu & Rao Eds, 2006

SSIM index maps of impulse corrupted *Einstein* image. a) square window. b) Circular window



© CRC Press, Wu & Rao Eds, 2006

- Weighted mean

$$\mu_x = \sum_{i=1}^{N} w_i x_i$$

- Weighted standard deviation

$$\sigma_x = \left(\sum_{i=1}^{N} w_i (x_i - \mu_x)^2 \right)^{\frac{1}{2}}$$

$$\sigma_{xy} = \sum_{i=1}^{N} w_i (x_i - \mu_x)(y_i - \mu_y)$$

© CRC Press, Wu & Rao Eds, 2006

- Overall quality measure for SSIM

$$\mathrm{MSSIM} = \sum_{j=1}^{M} W_j \cdot \mathrm{SSIM}_j$$

Where the sample weight W_j follows

$$\sum_{j=1}^{M} W_j = 1$$

  – Weighting strategies
    - uniform
    - Region of interest (ROI)
    - Foveal area

© CRC Press, Wu & Rao Eds, 2006

Practical evaluation MSSIM

- Evaluate the performance of SSIM
  - Measure against subjective scores
- Source material
  - 29 colour images from LIVE database
  - Compressed with JPEG and JPEG2000
    - JPEG bitrates 0.150 - 3.336 bpp
    - JPEG2000 bitrates 0.028 – 3.150 bpp
  - Generating 175 coded images

© CRC Press, Wu & Rao Eds, 2006

- Subjective evaluation
  - Quality assessment with 5 point scale
    - Excellent, good, fair, poor and bad
  - Subject numbers
    - 13 to 20 for JPEG coded images
    - 25 for JPEG 2000 coded images
  - Screening for outliers in data
  - Data normalized to range of 1 -100

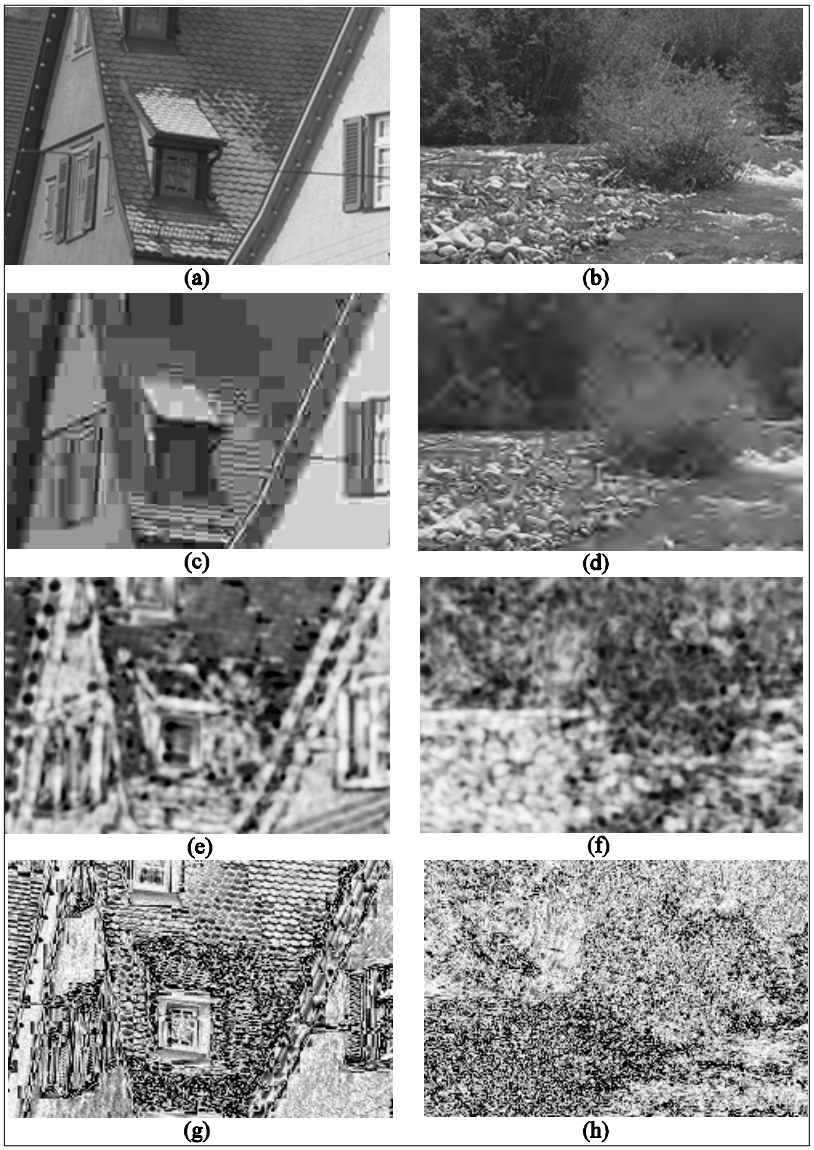
© CRC Press, Wu & Rao Eds, 2006

- MSSIM operation
  - Calculation based on downsampled image
    - ¼ resolution image
- Performance observations
  - Colour components has little impact on MSSIM performance
  - MSSIM more consistent with perceived quality
  - Better prediction than PSNR

© CRC Press, Wu & Rao Eds, 2006

Original *Buildings* image

Original *Stream* Image



JPEG Coded @ 0.2673 bpp

JPEG2000 Coded @ 0.1896 bpp

SSIM map

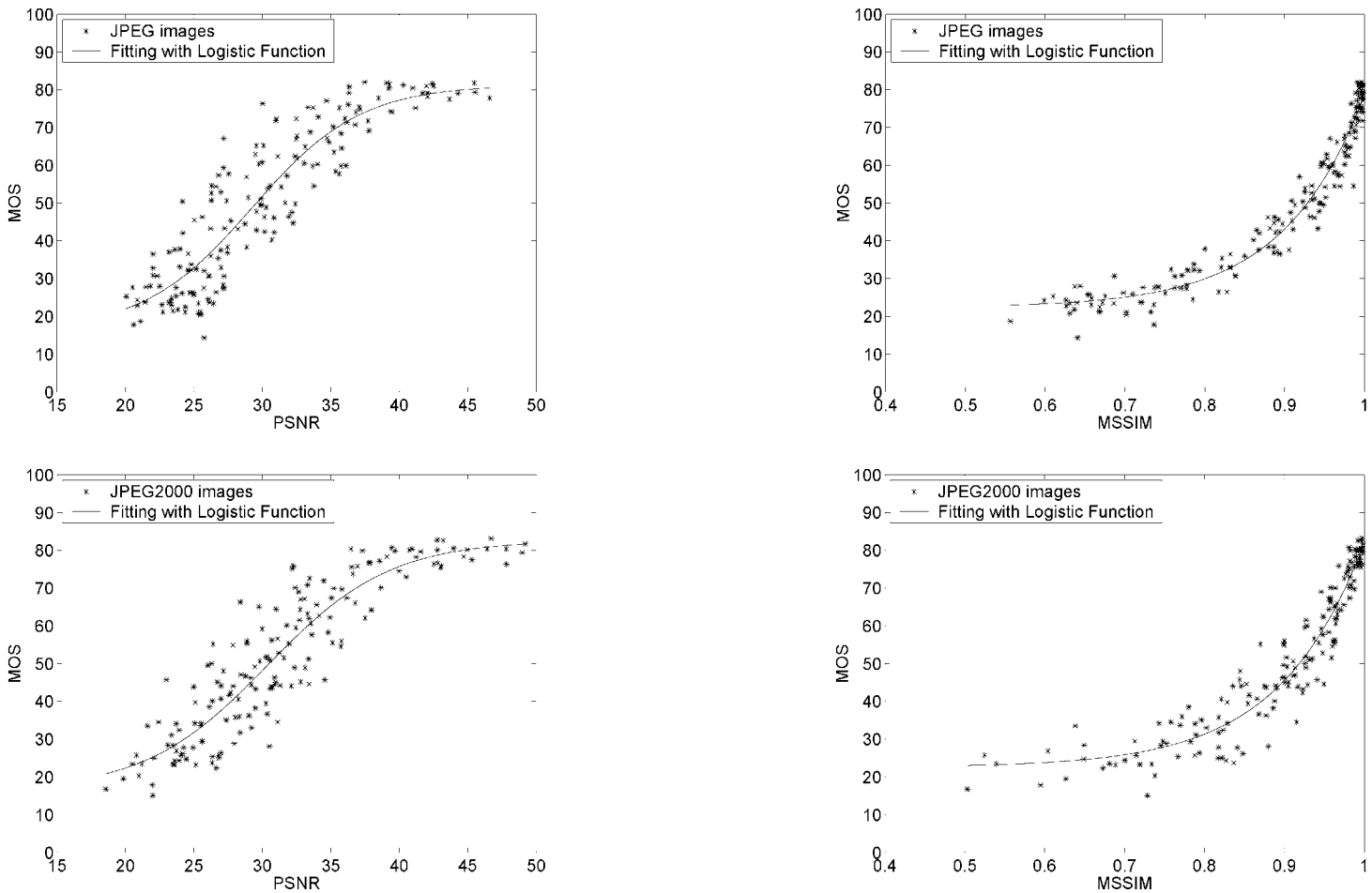
SSIM map

Absolute error map

Absolute error map

© CRC Press, Wu & Rao Eds, 2006

Scatter plots of MOS vs model prediction



© CRC Press, Wu & Rao Eds, 2006

- Quantitative performance evaluation
  - Correlation coefficient (CC)
  - Spearman rand order correlation coefficient (SROCC)
- MSSIM clearly better

Database	JPEG		JPEG2000	
Model	PSNR	MSSIM	PSNR	MSSIM
CC	0.904	0.978	0.910	0.958
MAE	6.769	3.324	6.300	4.352
RMS	8.637	4.176	8.062	5.540
OR	0.200	0.006	0.095	0.024
SROCC	0.893	0.973	0.906	0.955

© CRC Press, Wu & Rao Eds, 2006

7.4 Discussions

- Structural Similarity measurement
  - Different to traditional error sensitive method
    - MSE/PSNR
  - Measures perceived structural information
    - Not perceived error
  - Top-down approach
    - Mimics hypothesized overall HVS functionality

© CRC Press, Wu & Rao Eds, 2006

- SSIM
  - A practical application of the structural similarity philosophy
  - Simple implementation
  - Good prediction results
    - Compared to PSNR

© CRC Press, Wu & Rao Eds, 2006

Dal93	S. Daly. The visible difference predictor: An algorithm for the assessment of image quality. In A.B. Watson, Ed., *Digital Image and Human Vision*, Cambridge, MA: The MIT Press, 1993
EGM03	I. Epifanio, J. Hutirrez and J. Malo. Linear transform for simultaneous diagonalization of covariance and perceptual metric matrix in image coding. *Pattern Recognition*, 36(8):1679-1923, Aug. 2003
GB95	W.S. Geisler and M.S. Banks. Visual Performance. In M. Bass, Ed., *Handbook of Optics*, 25.1-25.55. New York: McGraw-Hill, June 1995
Lub93	J. Lubin. The use of psychophysical data and models in the analysis of display system performance. In A.B. Watson, Ed., *Digital Images and Human Vision*, 163-178. Cambridge, MA: The MIT Press, 1993
MNE+00	J. Malo, T. Navarro, I. Epifanio, F. Ferri and J.M. Artifas. Non-linear invertible representation for joint statistical and perceptual feature decorrelation. *Lecture Notes on Computer Science*, 1876:658-667, 2000
PS00	C.M. Privitera and L.W. Stark. Algorithm for defining visual regions-of-interest: comparison with eye fixations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(9):970-982, Sep. 2000

© CRC Press, Wu & Rao Eds, 2006

RCB03	U. Rajashekar, L.K. Cormack and A.C. Bovik. Image features that draw fixations. In *Proc. IEEE Inter. Conf. Image Processing*, 3:313-316, Sep. 2003
SO01	E.P. Simoncelli and B.A. Olshausen. Natural image statistics and neural representation. *Annu. Rev. Neuroscience*, 24:1193-1216, May 2001
SWBC03	H.R. Sheikh, Z. Wang, A.C. Bovik and L.K. Cormack. *Image and video quality assessment research at LIVE*. University of Texas at Austin, 2003. (http:live.ece.utexas.edu.research/quality/)
TH94	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Human Vision, Visual Processing and Digital Display V*, *Proc. SPIE*, 2179:127-141, 1994
VQE00	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Mar. 2000.
VQE03	VQEG. *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. VQEG, August 2003
Wan01	Z. Wang. *Rate scalable foveated image and video communications*. Ph.D. thesis, Dept. of ECE, The University of Texas at Austin, Texas, 2001

© CRC Press, Wu & Rao Eds, 2006

WB01	Z. Wang and A.C. Bovik. Embedded foveation image coding. *IEEE Trans. Image Processing*, 10(10):1397-1410, Oct. 2001
WB02	Z. Wang and A.C. Bovik, A universal image quality index. *IEEE Signal Proc. Letters*, 9(3):81-84, Mar. 2002
WBSS03	Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. *The SSIM Index for Image Quality Assessment*, New York University and Univerisity of Texas at Austin, 2003 (http:www.cns.nyu.edu/~lcv/ssim/)
WBSS04	Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. Image quality assessment: From error measurement to structural similarity. *IEEE Trans. On Image Processing*, 13(4):600-612, April 2004
WLB04	Z. Wang, L. Lu and A.C. Bovik. Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, 19(2):121-132, Feb. 2004. Special issue on video quality metrics
WS04	Z. Wang and E.P. Simoncelli. Stimulus synthesis for efficient evaluation and refinement of perceptual image quality metrics. In B.E. Rogawitz and T.N. Pappas, Eds., *Human Vision and Electronics Imaging IX, Proc. SPIE*, 5292:99-108, Jan. 2004

© CRC Press, Wu & Rao Eds, 2006

WSB03	Z. Wang, E.P. Simoncelli and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *Proc. IEEE Asilomar Conf. Signals, Systems and Computers*, 1398-1402, Nov. 2003
WYSV97	A.B. Watson, G.Y. Yang, J.A. Solomon and J. Villasenor. Visibility if wavelet quantization noise. *IEEE Trans. Image Processing*, 6(8):1164-1175, Aug. 1997

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

8. Vision Model Based Digital Video Impairment Metrics

Z. Yu† and H.R. Wu‡

† National Information Communication Technology Australia (NICTA)

‡ Royal Melbourne Institute of Technology, Australia

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 8 Outline

8.1 Introduction

8.2 Vision Modelling for Impairment Measurement

8.3 Perceptual Blocking Distortion Metric

8.4 Perceptual Ringing Distortion Measure

8.5 Conclusion

© CRC Press, Wu & Rao Eds, 2006

8.1 Introduction

- Proliferation of digital services and applications
  - Digital broadcast, multimedia, camera, etc
  - Rely upon compression
    - Leads to distortions
    - Therefore need to quantify quality of services
- Traditional metrics
  - MSE/PSNR
  - Do not always reflect perceived quality or distortion

© CRC Press, Wu & Rao Eds, 2006

- Vision Model based metrics
  - Emulate mechanism of the *Human Visual System* (HVS)
  - More accurate measurement of perceived quality/distortion
  - Generalized metric
    - Measures overall quality
  - Specialized metric
    - Measures specific impairment. e.g. blurring

© CRC Press, Wu & Rao Eds, 2006

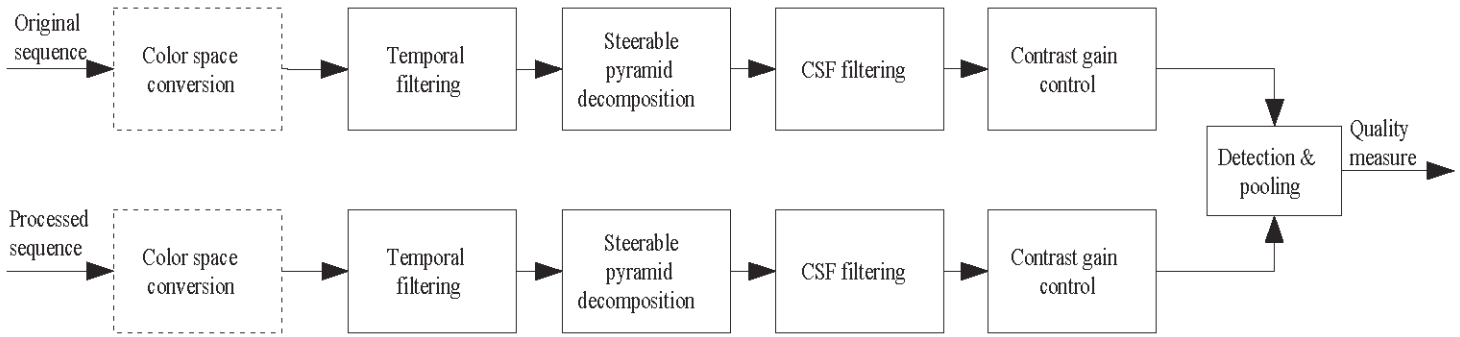
Impairment metrics

- Useful for studying the characteristics of individual impairment
  - May lead to improvements of coding systems
    - e.g. blocking impairment used to design de-blocking algorithms

	impairment	vision model	reference	image/video
Karunasekera et al [KK95]	blocking	Y	full	image
ANSI T1.801.03 [Sec96]	blocking	N	reduced	both [1]
GBIM [WY97]	blocking	Y	no	image
PBDM [YWC00a, YWWC02]	blocking	Y	full	video
PRDM [YWC00b, ZWY+03]	ringing	Y	full	video
Wang et al [WBE00]	blocking	Y	no	image
Bovik et al [BL01]	blocking	Y	no	image

© CRC Press, Wu & Rao Eds, 2006

8.2 Vision Modelling for Impairment Measurement



- Key element of vision model (*above*)
  - Colour sensitivity
  - Spatio-temporal contrast sensitivity
  - Contrast gain control

© CRC Press, Wu & Rao Eds, 2006

8.2.1 Colour Space Conversion

- Colour Spaces
  - *RGB*, YC_bC_r, *CIE XYZ*, etc
  - For typical video - YC_bC_r
- *Opponent Colour Space* (OCS)
  - HVS operated in OCS
    - Black-White (BW), Red-Green (RG), Blue-Yellow (BY) colour channels
  - Conversion from YC_bC_r to OCS needed
  - *Y* and *BW* channels have similar properties

© CRC Press, Wu & Rao Eds, 2006

- Sensitivity of colour channels
  - *BW* and *Y* channels contains most information
    - Possible for vision models to operate on colour pictures with luminance channel only with minimal error

© CRC Press, Wu & Rao Eds, 2006

8.2.2 Temporal Filtering

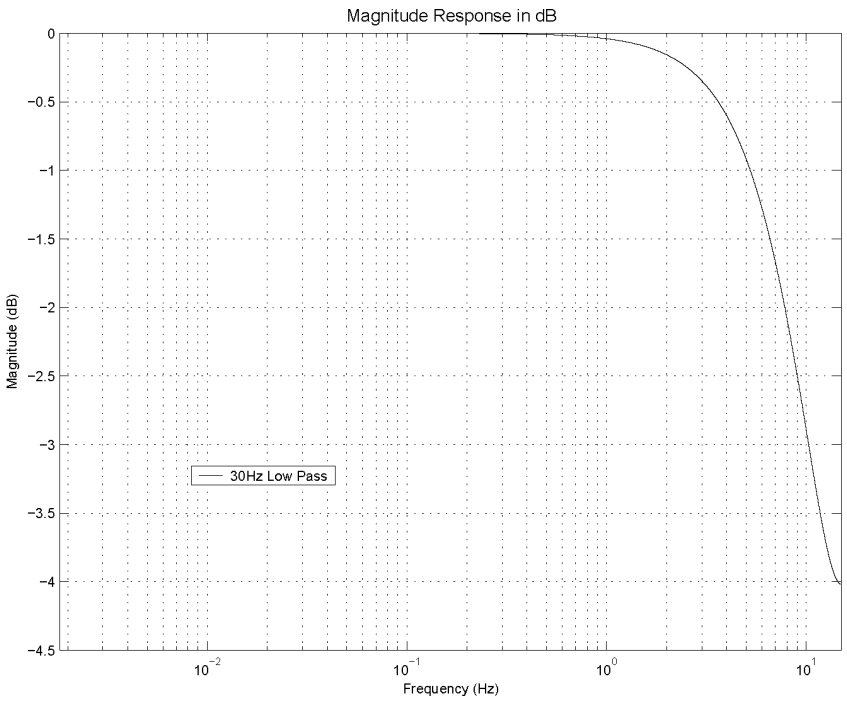
- Temporal channels of HVS
  - Sustained (low pass)
    - Contains majority of information (and distortions)
  - Transient (band pass)

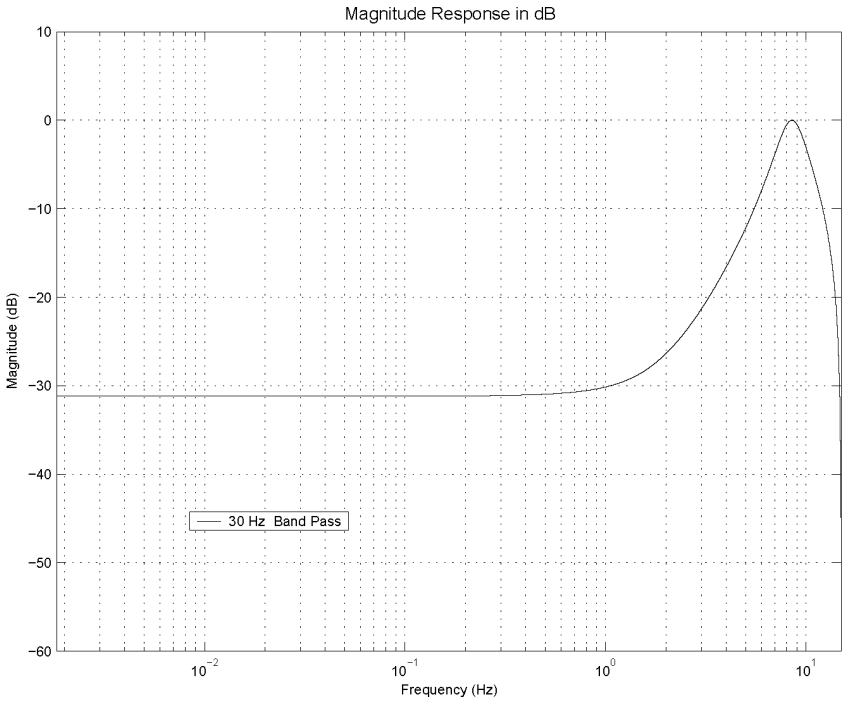
Below: 30 Hz filter coefficients

Filter	a[2] ... a[n_a+1]	b[1] b[2] ... b[n_b+1]
low pass 30Hz	-0.080000	0.800000, 0.120000
band pass 30Hz	0.264560, 0.532900	0.205877, -0.079503, -0.182592, 0.105776
low pass 25Hz	-0.080000	0.859813, 0.060187
band pass 25Hz	0.599392, 0.348100	0.222765, -0.210861, -0.099928, 0.182667

© CRC Press, Wu & Rao Eds, 2006

Temporal filters (30 Hz). *Left*: low pass. *Right*: band pass





© CRC Press, Wu & Rao Eds, 2006

8.2.3 Spatial Filtering

- Spatial sensitivity of HVS
  - 4 to 7 bands
  - Frequency with 1 to 2 octaves bandwidth
  - Orientation with 30 to 60 degrees bandwidth
- Practical implementation
  - *Steerable Pyramid Transform* (SPT)
    - 1 isotropic low-pass level
    - 4 band pass levels
      - 4 orientations (0°, 45°, 90° and 135°)
    - 1 isotropic high pass level

© CRC Press, Wu & Rao Eds, 2006

Steerable Pyramid Image of *Claire*

© CRC Press, Wu & Rao Eds, 2006

- Contrast Sensitivity realizations
  - Convolution through filter banks
  - Multiplier through weighting coefficients
    - Simple but coarse approximation
    - One weight for each frequency level

© CRC Press, Wu & Rao Eds, 2006

8.2.4 Contrast Gain Control

- Accounts for masking phenomenon
  - Consist of excitatory and inhibitory given by

$$\tilde{\mathbf{X}}_R[i,f,\theta,k,l] = \sum_{i=1}^{4} w_i \frac{(\mathbf{X}[i,f,\theta,k,l])^2}{\sum_{\phi}(\mathbf{X}[i,f,\phi,k,l])^2 + \gamma_i^2}$$

where $\mathbf{X}[j,f,\theta,k,l]$ is an SPT coefficient at specific frame (j), frequency (f), orientation (θ), vertical (k) and horizontal (l) location

w_i is the scaling constant and γ_i the saturation coefficient. $\Phi = \{0^o, 45^o, 90^o, 135^o\}$ with $i = \{1,2,3,4\}$ representing 4 contrasts mechanisms

© CRC Press, Wu & Rao Eds, 2006

- Multiple contrast mechanism
  - Needed to overcome the rapid saturation of individual mechanism
    - Occurs when STP coefficient $\mathbf{X}[j,f,\theta,k,l]$ is considerably larger than γ_i

Model Parameters	
C_i (CSF)	0.41, 1.25, 1.2, 0.4097, 0.083, 0.001
w_i	1, 1, 1, 1
γ_i	4.4817, 7.3891, 12.1825, 54.5982

© CRC Press, Wu & Rao Eds, 2006

- Isotropic low-pass band
  - Raw magnitude of coefficient are often significantly larger than γ_i
  - Apply mean subtraction to reduce magnitude energy

© CRC Press, Wu & Rao Eds, 2006

8.2.5 Detection and Pooling

- Decimation factor
  - Normalize sensor output for uniform presentation

$$\mathbf{X}_{R_o}[j,f,\theta,k,l] = \frac{\tilde{\mathbf{X}}_{R_o}[j,f,\theta,k,l]}{X_f}$$

$$\mathbf{X}_{R_p}[j,f,\theta,k,l] = \frac{\tilde{\mathbf{X}}_{R_p}[j,f,\theta,k,l]}{X_f}$$

  - $X_f = \{1,4,16,64,256\}$

© CRC Press, Wu & Rao Eds, 2006

- Difference and Pooling operator
  - Squared difference between original (R_o) and processed (R_p)
  - Summation over all variable domains

$$\Delta \mathbf{X}_R = \frac{\sum_{j,f,\theta,k,l} |\mathbf{X}_{R_o}[j,f,\theta,k,l] - \mathbf{X}_{R_o}[j,f,\theta,k,l]|^2}{N}$$

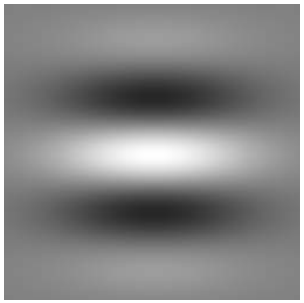
© CRC Press, Wu & Rao Eds, 2006

8.2.6 Model Parameterization

- Calibrate Vision Model Parameters
  - Ensures proper operation of vision model
- Two approaches
  - Vision research experiments
  - Video quality experiment

© CRC Press, Wu & Rao Eds, 2006

Vision research experiments



- Utilise synthetic test patterns →→
- Target specific aspect of the HVS
  - e.g. contrast sensitivity
- Testing generally at threshold level
  - Appropriate for threshold level assessment
  - Unsuitable for quality assessment
- Modelfest data set
  - Collection of test data for common stimuli

© CRC Press, Wu & Rao Eds, 2006

Video Quality Experiments

- Utilize natural video sequence
- Two phases
  - Collect subjective quality scores associated with test materials
    - e.g. VQEG and ANSI T1A1 sequences
  - Find model parameters which yields the best correlation between prediction and subjective data (Optimization process)

© CRC Press, Wu & Rao Eds, 2006

- Optimization
  - Cost function to measure correlation
    - Spearman rank correlation (r_s)
      - Measures the order of quality/distortion ranking between subjective data and prediction

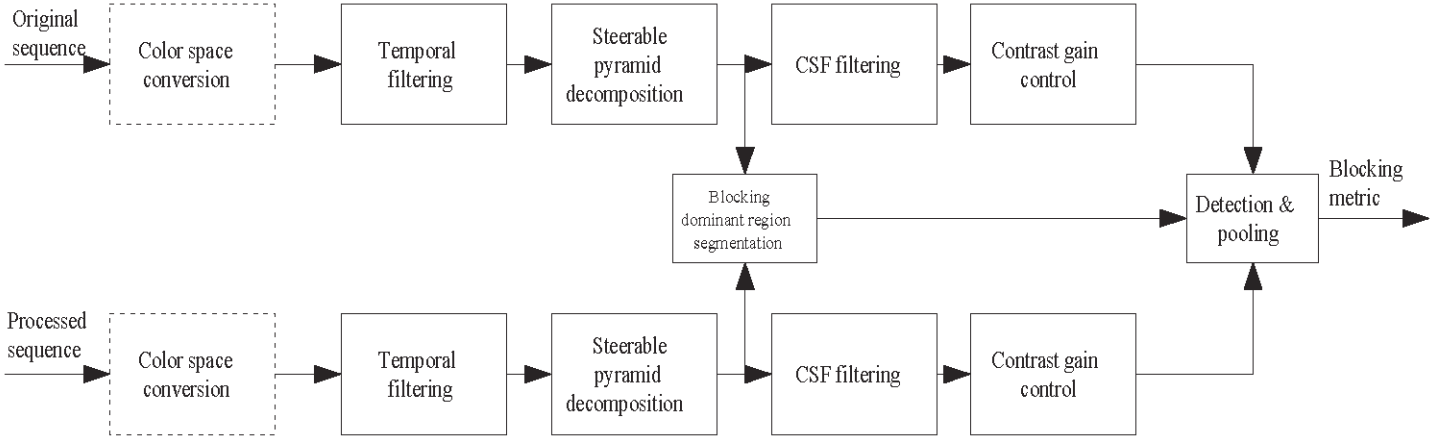
$$r_s = 1 - \frac{6\sum_l D_l^2}{N(N^2-1)}$$

  - Two tier optimization
    - Coarse search for general search space
    - Refinement stage to locate local optima

© CRC Press, Wu & Rao Eds, 2006

8.3 Perceptual Blocking Distortion Metric (PBDM)

- Blocking distortion from DCT coders
  - Seen as discontinuities at block boundaries
  - Manifest in smooth regions of pictures
  - Quantified with PBDM (below)

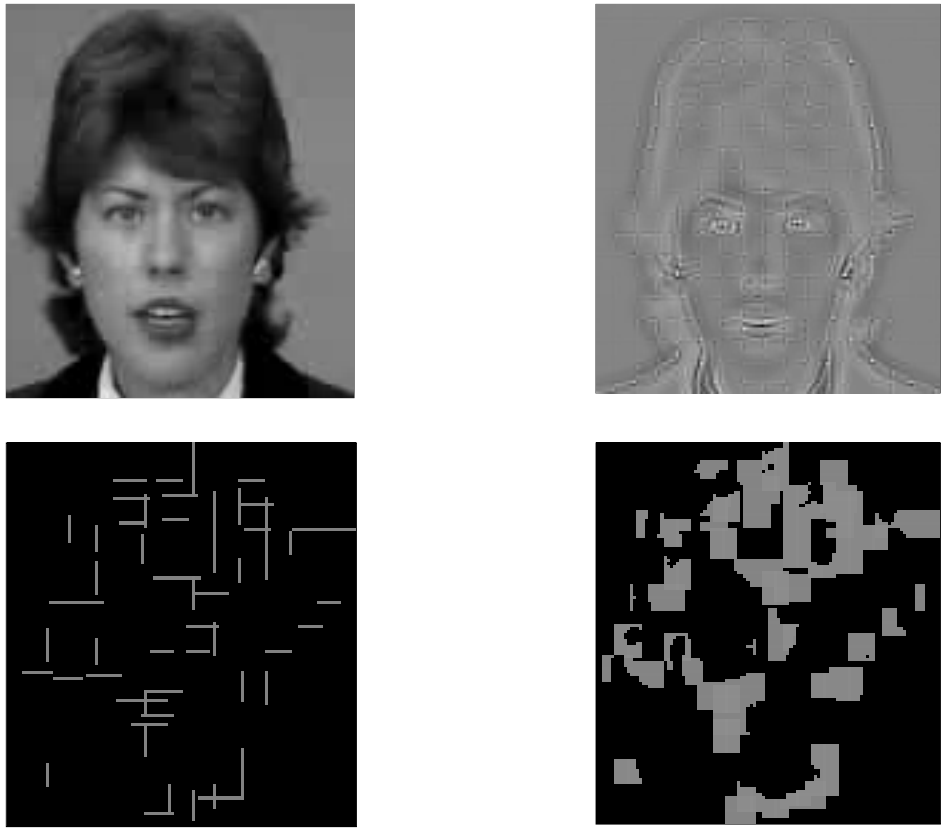


© CRC Press, Wu & Rao Eds, 2006

- PBDM Components
  - Blocking dominant region segmentation
    - Vertical and horizontal edge detection
    - Isolate blocking edges from ringing distortions
    - Generate blocking map
  - Quantify distortions in blocking region
    - With vision model in section 8.2
    - Covert distortion to Objective Blocking Rating (OBR) scale

© CRC Press, Wu & Rao Eds, 2006

Clockwise from top left. Coded *Claire* frame, high pass SPT image, detected block edges and blocking map



© CRC Press, Wu & Rao Eds, 2006

8.3.3 Performance Evaluation of the PBDM

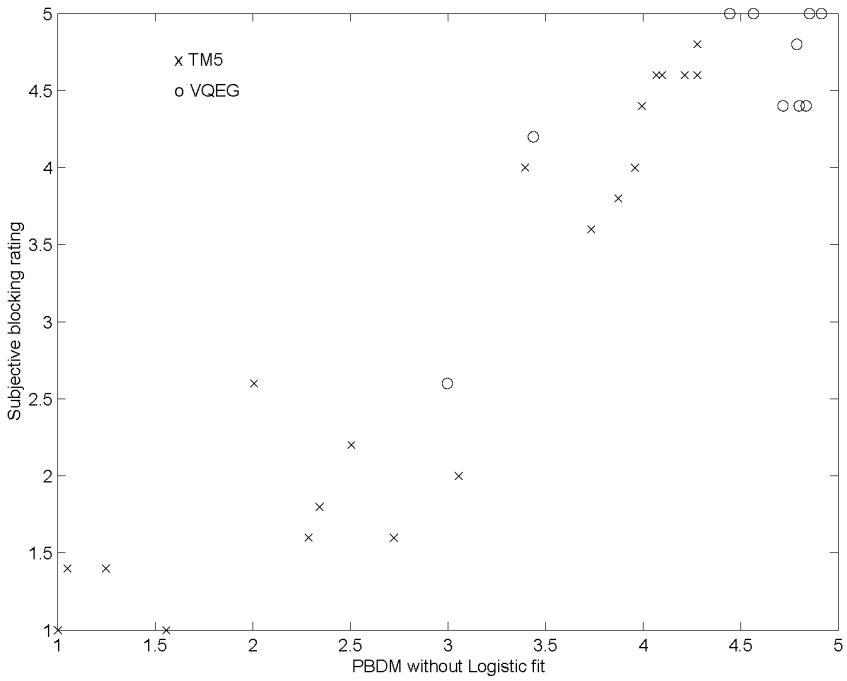
- Test material
  - ANSI T1A1
    - 5 sequences
    - Coded with MPEG-2 coder at 768 kbps, 1.4 Mbps, 2Mbps, 3Mbps and 5 Mbps
  - VQEG
    - 5 sequences
    - Coded with MPEG-2 coder at 3Mbps and 2Mbps (3/4 resolution)

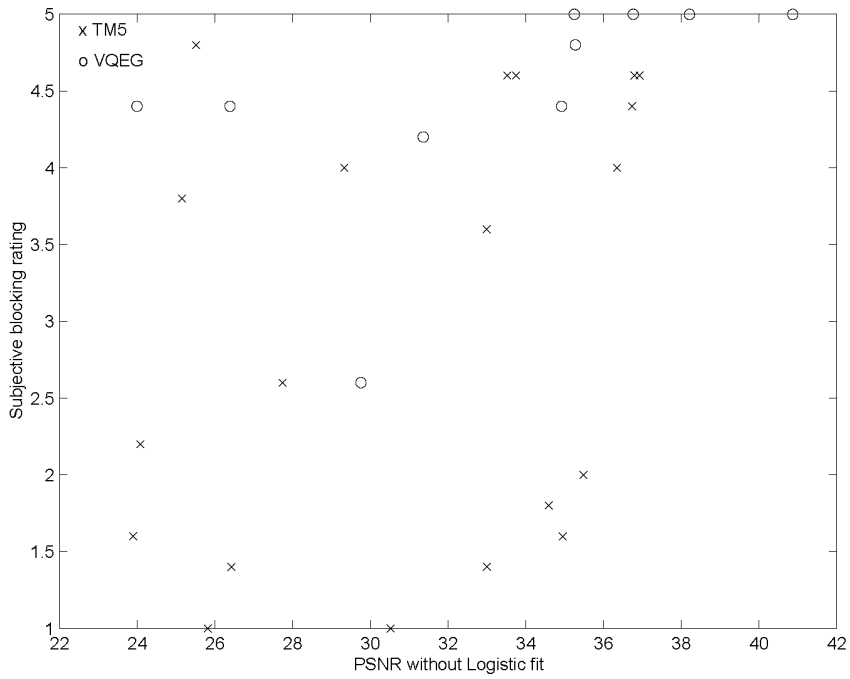
© CRC Press, Wu & Rao Eds, 2006

- Subjective assessment
  - ITU-R bt.500 recommended test setup
  - Double Stimulus Impairment Scale variant II (DSIS-II) method
  - 5 expert test subjects
- Evaluation with 3 correlation measures
  - spearman rank order correlation
  - Pearson's correlation coefficient
  - Logistic fit

© CRC Press, Wu & Rao Eds, 2006

MOS vs PBDM (left) and MOS vs PSNR (right)





© CRC Press, Wu & Rao Eds, 2006

PBDM Evaluation

Metric	Pearson-logistic metric			Spearman	E|*error*|	Average standard deviation in the subjective test data
	Correla-tion	Upper bound	Lower bound			
PBDM	0.961	0.982	0.918	0.910	0.296	0.455
PSNR	0.489	0.726	0.149	0.517	1.038	

- PBDM has excellent correlation to MOS
- PSNR correlation to MOS is poor

© CRC Press, Wu & Rao Eds, 2006

5row1 sequence: MOS: 3.6; PBDM: 3.79; PSNR: 32.99 dB

© CRC Press, Wu & Rao Eds, 2006

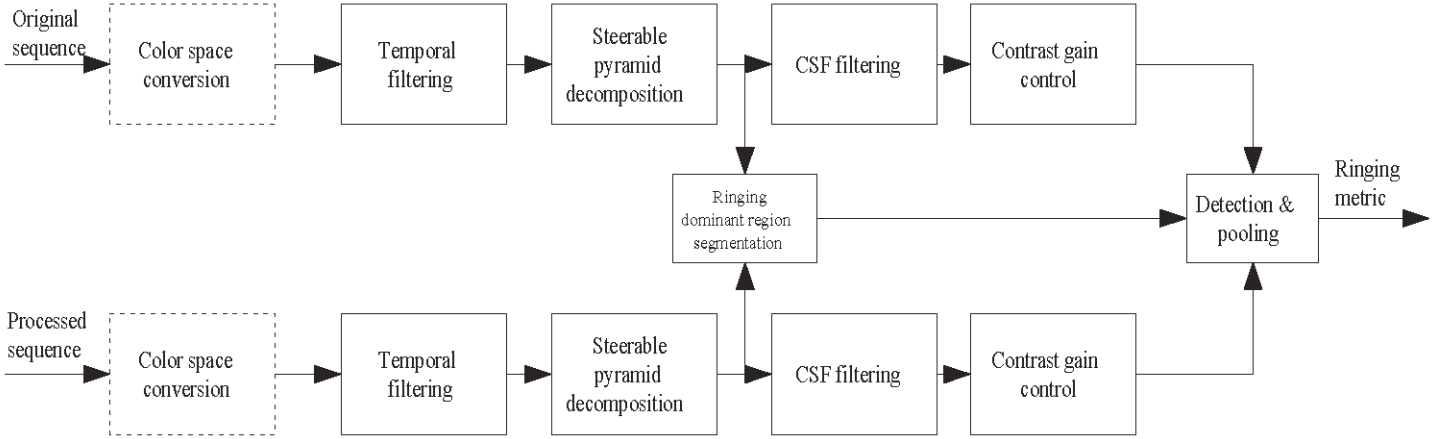
Inspec sequence. MOS: 1.4; PBDM: 1.22; PSNR: 33.00 dB

© CRC Press, Wu & Rao Eds, 2006

8.4 Perceptual Ringing Distortion Metric (PRDM)

- Ringing distortion
  - Related to Gibb's phenomenon
    - Quantization of transform coefficients leads to high frequency irregularities

Below: PRDM diagram



© CRC Press, Wu & Rao Eds, 2006

- PRDM components
  - Ringing region segmentation
    - Initial classification of regions into 3 categories based on variance
      - Smooth, complex and unknown regions
    - Detection of smooth and complex regions
      - Unknown regions are re-classified into either smooth or complex regions
    - Boundaries of smooth and complex regions are seen as ringing
  - Vision model detection and pooling

© CRC Press, Wu & Rao Eds, 2006

8.4.3 Performance Evaluation of PRDM

- Test material
  - Original sequences
    - Mobile & Calender (MC)
    - Table Tennis (TT)
  - Both sequences coded with MPEG-2 at
    - 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 4, 5 and 7 Mbps
- Subjective Assessment
  - DSIS-II
  - 5 expert and 2 non-expert viewers

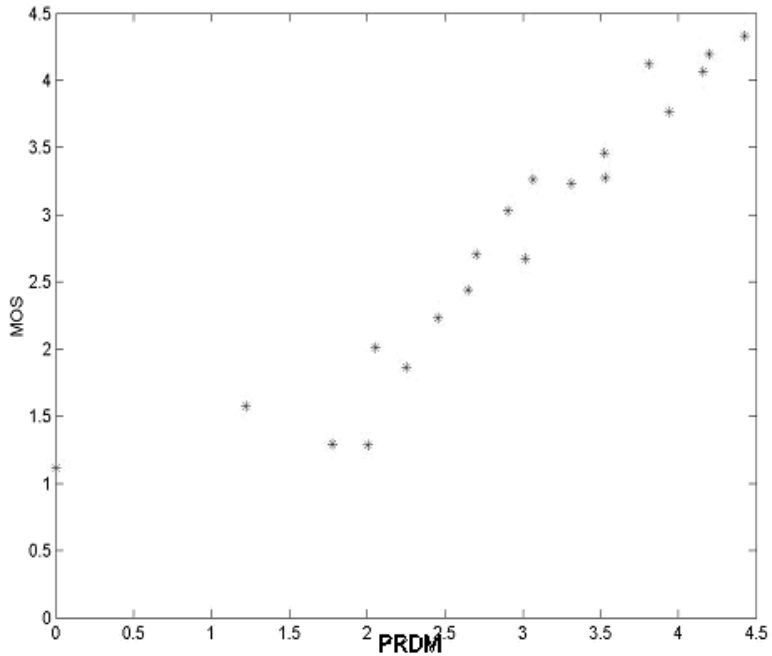
© CRC Press, Wu & Rao Eds, 2006

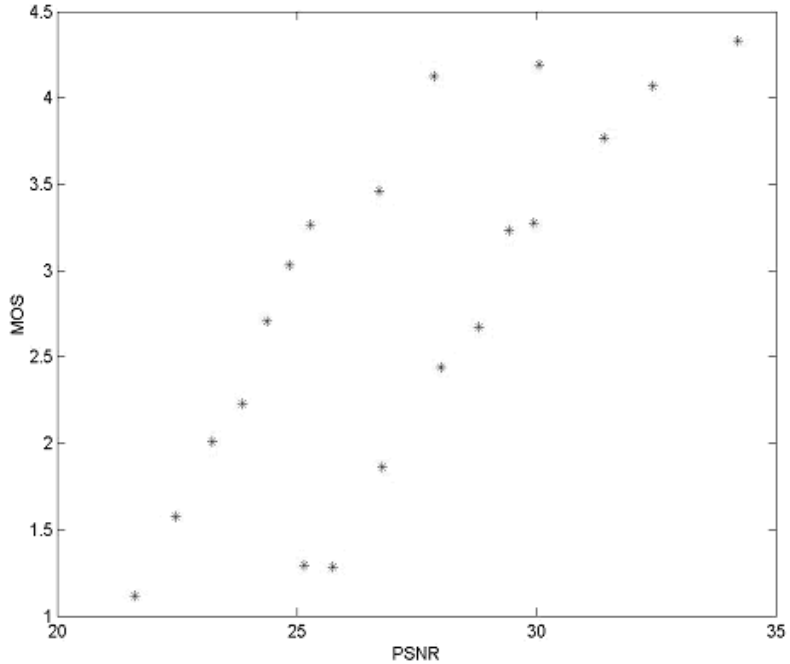
Metric	Pearson	Spearman	E$\lvert error \rvert$	Average standard deviation in the subjective test data
PRDM(both)	0.9397	0.9805	0.2651	0.4402
PSNR(both)	0.7674	0.7489	0.5358	
PRDM(MC)	0.9642	1	0.2457	0.4219
PSNR(MC)	0.9341	1	0.6659	
PRDM(TT)	0.9945	1	0.2845	0.4649
PSNR(TT)	0.9709	1	0.4056	

- Results shows excellent correlation between PRDM and MOS
- PSNR correlation to MOS is also good
- DCT coder exhibit limited ringing artifacts

© CRC Press, Wu & Rao Eds, 2006

MOS vs PRDM (left) and MOS vs PSNR (right)





© CRC Press, Wu & Rao Eds, 2006

8.5 Conclusion

- Progress have be made in vision model based objective metric
- However, still more work ahead
  - Limited understanding of HVS
    - e.g. temporal masking
  - Other impairment metrics
    - Blurring, temporal fluctuation, etc
  - Improve vision model for NR and RR metrics
  - Application of impairment metrics for picture coding systems development

© CRC Press, Wu & Rao Eds, 2006

ANS93	ANSI Accredited Standards Working Group T1A1 contribution number T1A1.5/94-118R1. *Subjective Test Plan (10th and Final Draft)*. Alliance for Telecommunications Industry Solutions, 1200 G Street, NW, Suite 500, Washington DC, Oct. 1993
ANS95	ANSI T1.801.01-1995. *ANSI T1.801.01-1995. American National Standard for Telecommunication – Digital Transport of Video Teleconferencing/Video Telephony Signals – Video Test Scenes for Subjective and Objective Performance Assessment*. American National Standards Institute, 1995
Bal98	I. Balasingham *et al*. Performance evaluation of different filter banks in the jpeg2000 baseline system. In *Proc. IEEE ICIP*, 2:569-573, 1998
BL01	A.C. Bovik and S. Liu. DCT-domain blind measurement of blocking artifacts in DCT-coded images. In *Proc. IEEE ICASSP*, 3:1725-1728, 2001
Car99	T. Carney *et al*. The development of an image/threshold database for designing and testing human vision models. In *Human Vision, Visual Processing and Digital Display IX, Proc. SPIE*, 3644:542-551, 1999
Fol94	J.M. Foley. Human luminance pattern-vision mechanisms: Masking experiments require a new model. *J. Opt. Soc. Am*, 11:1710-1719, 1994

© CRC Press, Wu & Rao Eds, 2006

Gir93	B. Girod. What's wrong with mean-squared error. In A.B. Watson, Ed., *Digital images and human vision*, 207-220. Cambridge, MA: The MIT Press, 1993
HJ57	L.M. Hurvich and D. Jameson. An opponent-process theory of color vision. *Psych. Rev.*, 64(2):384-404, 1957
HS92	R.F. Hess and R.J. Snowden. Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research*, 32(1):47-59, 1992
ITU95	ITU-R. *BT.601-5, Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios*. ITU, Geneva, Switzerland, 1995
ITU98	ITU-R. 500-9. *ITU-R BT. 500-9, Methodology for the Subjective Assessment of the Quality of Television Pictures*. ITU, Geneva, Switzerland, 1998
KK95	S.A. Karunasekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transaction on Image Processing*, 4(6):713-724, June 1995
LB82	F.X.J. Lukas and Z.L. Budrikis. Picture quality prediction based on a visual model. *IEEE Trans. on Communications*, COM-30(8):1679-1692, July 1982

© CRC Press, Wu & Rao Eds, 2006

LB02	S. Liu and A.C. Bovik. Efficient DCT-domain blind measurement and reduction of blocking artifacts. *IEEE Trans. Circuits and Systems for Video Technology*. 12(12):1139-1149, Dec. 2002
Lin96	P. Lindh. *Perceptual Image Sequence Quality metric Using Pixel Domain Filtering*. Masters Thesis, Linkoping Institute of Technology, Linkoping, Sweden, April 1996.
MKA98	M. Miyahara, K. Kotani and V.R. Algazi. Objective picture quality scale (PQS) for image coding. *IEEE Trans. on Communications*, 46(9):1215-1226, Sept. 1998
MPEG94	MPEG Software Simulation Group. *Mpeg2encode/mpeg2decode version 1.1*, Jun. 1994 (ftp.netcom.com)
OS89	A.V. Oppenheim and R.W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989
PW84	G.C Philips and H.R. Wilson. Orientation bandwidth of spatial mechanisms measured by masking. *J. Opt. Soc. Am. A*, 1(2):226-232, 1984

© CRC Press, Wu & Rao Eds, 2006

PW96	A.B. Poirson and B.A. Wandell. Pattern-color separable pathways predict sensitivity to simple colored patterns. *Vision Research*, 36(4):515-526, 1996
Sac84	L. Sachs. *Applied Statistics*. Berlin and Heidelberg: Springer-Verlag, 1984
SC89	G.W. Snedecor and W.G. Cochran. *Statistical Methods*. Ames, IA: The Iowa State University Press, 1989
Sec96	Secretariat, Alliance for Telecommunication Industry Solution. *T1.801.03-1996, American National Standard for Telecommunications – Digital Transport of One-Way Video Signals – parameters for Objective Performance Assessment*, ANSI, 1996
SF95	E.P. Simoncelli and W.T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proc. IEEE ICIP*, 175-184, 1995
TH94a	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. ICIP*, 2:982-986, 1994
TH94b	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. SPIE*, 2179:127-141, 1994

© CRC Press, Wu & Rao Eds, 2006

Val82	R.L.D. Valois *et al*. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):545-559, 1982
Van96a	C.J. van den Branden Lambrecht. *Perceptual Models and Architectures for Video Coding Applications*. Ph.D thesis, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, 1996
Can96b	C.J. van den Branden Lambrecht. Color moving pictures quality metric. In *Proc. ICIP*, 1:885-888, Lausanne, Switzerland, 1996
VQE99	VQEG, *VQEG Subjective Test Plan Ver. 3*. VQEG, Jul. 1999 (http://www.vqeg.org/)
VQE00	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, April 2000.
Wan95	B.A. Wandell. *Foundations of Vision*. Sunderland, MA: Sinauer Associates, Inc., 1995
Wat83	A.B. Watson. Detection and recognition of simple spatial forms. In A.C. Slade, Ed., *Physical and Biological Processing of Images*, 100-114. Berlin: Springer-Verlag, 1983

© CRC Press, Wu & Rao Eds, 2006

WBE00	Z. Wang, A.C. Bovik and B.L. Evan. Blind measurement of blocking artifacts in images. In *Proc. IEEE ICIP*, 3:981-984, 2000
Win99a	S. Winkler. Issues in vision modeling for perceptual video quality assessment. *Signal Processing*, 78(2)231-252, 1999
Win00	S. Winkler. Quality metric design: A closer look. In *Proc. SPIE*, 3959:37-44, San Jose, CA, Jan. 23-28, 2000
WJP+93	A.A. Webster, C.T. Jones, M.H. Pinson, S.D. Voran and S. Wolf. An objective video quality assessment system based on human perception. In *Proc. SPIE Human Vision, Visual Processing and Digital Display IV*, 1913:885-888, San Jose, CA, Feb. 1993
WR84	H.R. Wilson and D. Regan. Spatial-frequency adaptation and grating discrimination: Predictions of a line-element model. *J. Opt. Soc. Am.*, 1:1091-1096, 1984
WS97	A.B. Watson and J.A. Solomon. A model of visual contrast gain control and pattern masking. *J. Opt. Soc. Am. A*, 14:2379-2391, 1997

© CRC Press, Wu & Rao Eds, 2006

WY97	H.R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 4(11):317-320, Nov. 1997
YGvV97	I.T. Young, J.J. Gerbrands and L.J. van Vilet. Image processing fundamentals. In V.K. Madisetti and D.B. Williams, Eds., *The Digital Signal Processing Handbook*. Boca Raton, FL: CRC Press, 1997
YW98	M. Yuen and H.R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing*, 70:247-278, 1998
YWC00a	Z. Yu, H.R. Wu and T.C. Chen. Perceptual blocking distortion measure for digital video. In *Proc. Of the SPIE Visual Communication and Image Processing*, 4067L768-779, June 2000
YWC00b	Z. Yu, H.R. Wu and T.C. Chen. A perceptual measure of ringing artifact for hybrid MC/DPCM/DCT coded video. In *Proc. of the IASTED International Conference on Signal and Image Processing*, 94-99, Nov. 2000
YWWC02	Z. Yu, H.R. Wu, S. Winkler and T. Chen. Objective assessment of blocking artifacts for digital video with a vision model. *Proc. IEEE*, 90(1):154-169, jan. 2002. (An invited paper in special issue in translating human vision research into engineering technology)

© CRC Press, Wu & Rao Eds, 2006

ZWY+03	Z. Zhe. H.R. Wu, Z. Yu, T. Ferguson and D. Tan. Performance evaluation of a perceptual ringing distortion metric for digital video. In *Proc. Of ICASSP*, 3:613-616, April 2003

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

9. Modeling Just-noticeable Difference in Visual Signal

Weisi Lin

Institute for Infocomm Research, Singapore

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 9 Outline

9.1 Introduction

9.2 JND with DCT Subbands

9.3 JND with Pixels

9.4 JND Model Evaluation

9.5 Conclusions

© CRC Press, Wu & Rao Eds, 2006

9.1 Introduction

- Digital pictures are ultimately viewed by the *Human Visual System* (HVS)
  - Regardless of operations or applications
- The HVS has limited sensitivity
  - Unable to sense all changes in pictures
    - Note: not all changes equate to distortions
      - e.g. post processing, edge enhancement, etc
  - Advantageous to consider HVS characteristics in picture coding systems

© CRC Press, Wu & Rao Eds, 2006

Just-Noticeable Difference (JND)

- JND definition
  - The visibility threshold below which any change cannot be detected by the HVS
- Factor affection JND
  - local characteristics of images
  - Viewing conditions
  - Individual subject

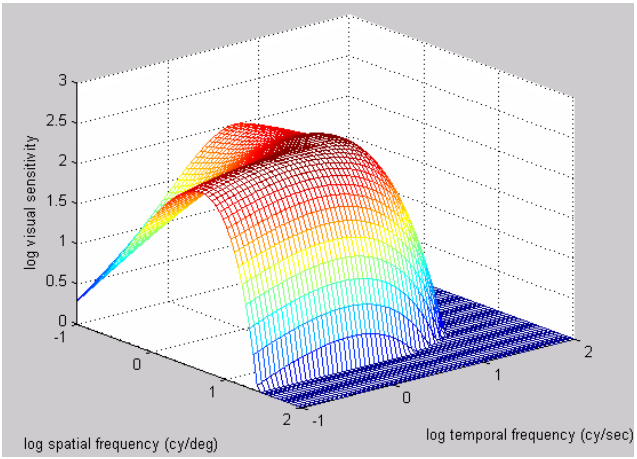
© CRC Press, Wu & Rao Eds, 2006

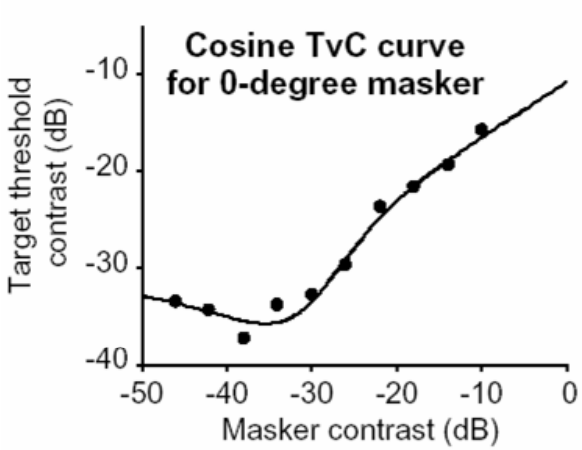
9.1.1 Single Stimulus JND Test

- Based upon simple stimuli
  - Varying foreground stimulus over a uniform static background
- Follows Weber-Fechner's Law
  - Just noticeable foreground threshold increases with background luminance above 10 cd/m^2

© CRC Press, Wu & Rao Eds, 2006

Factors affecting JND threshold levels





– Contrast Sensitivity (*top left*: CSF)
  - JND varies with spatial and temporal frequencies

– Masking and facilitation (*top right*)
  - JND dependent on the interactions between stimulus and its surroundings

© CRC Press, Wu & Rao Eds, 2006

9.1.2 JND Tests with Real-World Images

- Test the JND levels relative to distortions
  - i.e., $x_t = x_o + h(x_d - x_o)$

    where x_o, x_d and x_t original, distorted and test images with h the weighting factor
  - Distortions could be of any form including:
    - Blocking, ringing, blurring, fluctuations, etc
  - JND level determined if 75% of observes could perceive the distortion $(x_d - x_o)$
  - JND can be scaled, i.e., 2-JND, 3-JND, etc

© CRC Press, Wu & Rao Eds, 2006

9.1.3 Applications of JND Models

compression & coding
- quantization/rate control
- motion search
- pre-processing
- filtering of residues/coefficients
- inter-frame replenishment

edge enhancement
- optimum sharpness

quality/distortion metrics
- differentiation
  - near-JND
  - supra-JND
- quantification

watermarking:
- authentication
- error resilience

other visual processing, for
resource savings/allocation
bandwidth, computing power, memory, display/printing resolution
and/or
performance enhancement
picture quality

© CRC Press, Wu & Rao Eds, 2006

Existing Works with JND models

image/video compression	quantizer and rate control Watson'93 Hontsch & Karam'00,02, Yang, *et al.*'05 motion search Yang, *et al.*'03 inter-frame replenishment Chiu & Berger'99 filtering of residues/coefficients Yang, *et al.*'05

visual communication	watermarking Wolfgang, *et al.*'99 self-embedment for error correction unequal error protection Jiang, *et al.*'99 joint source-channel coding

enhancement/ reconstruction	synthesis Ramasubramanian, *et al.*'99 edge-enhancement Lin, *et al.*'05

© CRC Press, Wu & Rao Eds, 2006

JND Estimation

DCT subbands

Ahumada & Peterson'92, Watson'93, Hontsch & Karam'00,02, Zhang, et al.'05

wavelet subbands

Watson, et al.'93

pyramid subbands

Ramasubramanian, *et al.*'99

pixel domain

Chou & Li'95, Chiu &Berger'99, Yang, *et al.*'03

contrast masking

Tong&Venetsanopoulos'98, Zhang, et al.'05

temporal effect

eye motion: Daly'98
frame difference: Chou & Chen'96
temporal CSF:for subbands- Daly'98, Watson, *et al.*'01
for pixel-Zhang'04

Visual-attention modulation

Lu, *et al.*'05

© CRC Press, Wu & Rao Eds, 2006

9.2 JND with DCT Subbands

- Step for determining JND for DCT subbands
  - Base threshold
    - Spatial CSF
  - Luminance adaptation
  - Contrast masking
    - Intra- and inter-band masking
  - Other factors
    - Temporal and colour considerations

© CRC Press, Wu & Rao Eds, 2006

- General formulation in DCT subband

$$s(n,i,j) = t_{s-csf}(n,i,j)\prod_{\xi}\alpha_{\xi}(n,i,j)$$

where n and (i,j) are the DCT block and subband index

- $t_{s\text{-}csf}(n,i,j)$ Base threshold due to spatial CSF
- $\prod_{\xi}\alpha_{\xi}(n,i,j)$ Elevation parameter due to adaptation and masking

© CRC Press, Wu & Rao Eds, 2006

9.2.1 Formulation of Base Threshold

- Base threshold defined as

$$t_{s-csf}(n,k,l) = \frac{G}{\phi_k \phi_l (L_{\max} - L_{\max})} T^o(n,k,l)$$

where $T^o(n,k,l)$ is the DCT domain CSF threshold and

$$\phi_u = \begin{cases} \sqrt{1/N}, & u = 0 \\ \sqrt{2/N}, & \text{otherwise} \end{cases}$$

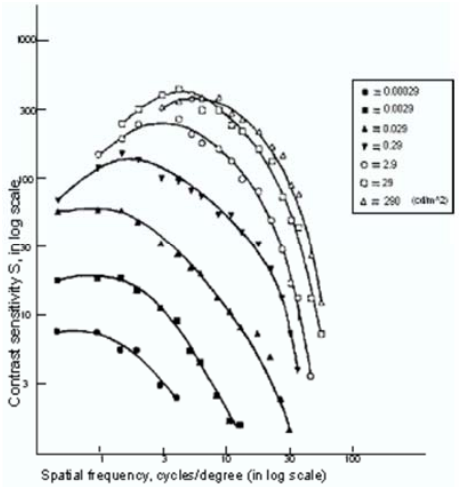
© CRC Press, Wu & Rao Eds, 2006

Modelling spatial CSF

- CSF threshold defined as

$$\log T = \log T_{\min} + k(\log f - \log f_p)^2$$

  – From van Nes and Bouman's data (below)



Van Nes & Bouman'67

© CRC Press, Wu & Rao Eds, 2006

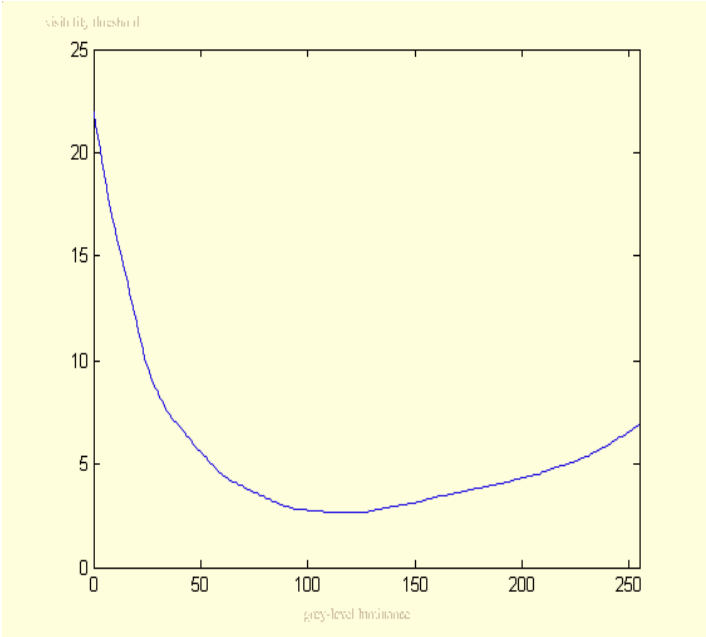
9.2.2 Luminance Adaptation Considerations

- Simple luminance adaptation formulation

$$\alpha_{lum}(n) = \left(\frac{X[n,0,0]}{X_L} \right)^{\varepsilon}$$

$X[n,0,0]$ is DC coefficient at n^{th} block

X_L is the background luminance of DC band



Threshold vs Luminance

© CRC Press, Wu & Rao Eds, 2006

- Realistic adaptation model
  - Needed due to non-monotonic characteristic

$$\alpha_{lum}(n) = \begin{cases} k_1\left(1 - \dfrac{2 \cdot \mathrm{X}[n,0,0]}{G \cdot N}\right)^3, & \mathrm{X}[n,0,0] \le \dfrac{G.N}{2} \\ k_2\left(\dfrac{2 \cdot \mathrm{X}[n,0,0]}{G \cdot N} - 1\right)^2, & \text{otherwise} \end{cases}$$

G is the maximum grey level with K_1 and K_2 fitted parameters

© CRC Press, Wu & Rao Eds, 2006

9.2.3 Contrast Masking

- Intra-band masking (within subband)

$$\alpha_{intra}(n,k,l) = \max\left(1, \left|\frac{X[n,k,l]}{t_{s-csf}(n,i,j)\cdot\alpha_{lum}(n)}\right|^{\zeta}\right)$$

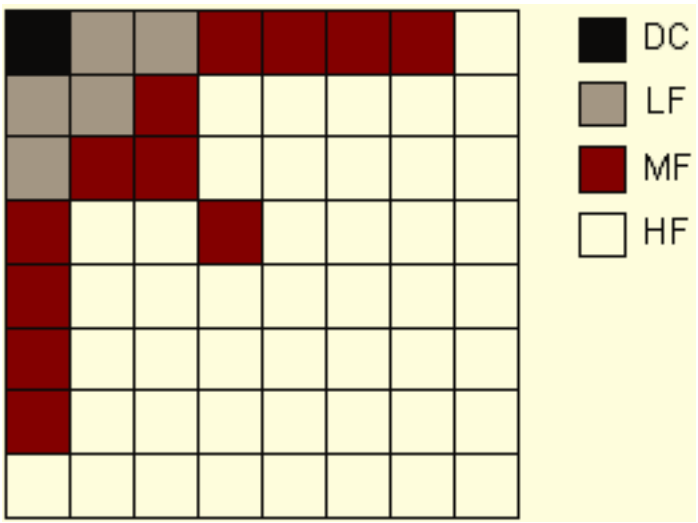
with $0 \le \zeta \le 1$

- Inter-band masking considerations
  - Classification of DCT coefficients
  - Classify subblocks into masking group

© CRC Press, Wu & Rao Eds, 2006

Inter-band masking classification

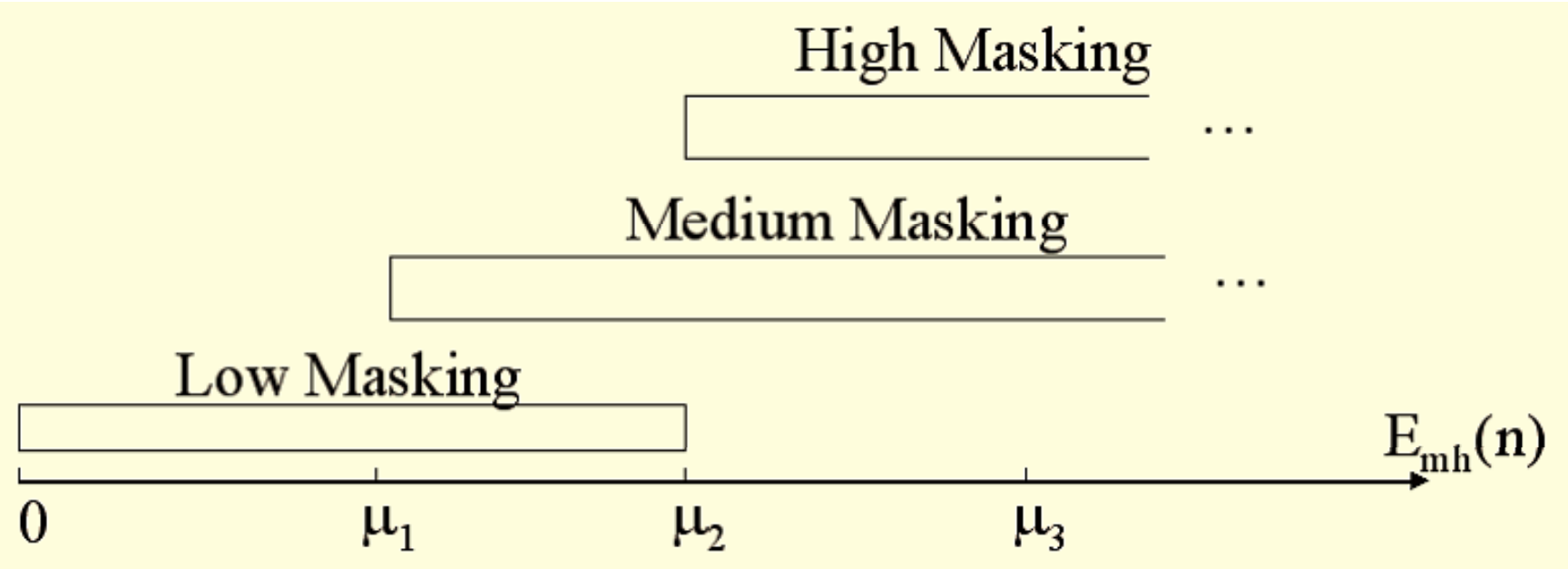
1. Segregate DCT coefficients into low (LF), mid (MF) and high (HF) frequency groups
2. Calculated absolute energy of each group
3. Classify DCT block to masking group based on energy of HF and MF, (E_{mh}), relative to predefined parameters μ_1, μ_2 and μ_3



DCT Coefficient grouping

© CRC Press, Wu & Rao Eds, 2006

Illustration of low (LM), medium (MM) and high (HM) masking classification of DCT subblocks



μ_1=125, μ_2=290 and μ_3=900

© CRC Press, Wu & Rao Eds, 2006

- Inter-band masking (across subband)

$$\alpha_{inter}(n) = \begin{cases} 1, & LM \\ \delta_1 & MM, E_{ml} \le R_0 \\ \delta_2 & MM, E_{ml} \ge R_0 \\ \left(1 + \dfrac{E_{mh}(n) - \mu_2}{2 \cdot \mu_3 - \mu_2}\right), & otherwise \end{cases}$$

E_{ml} absolute energy of mid and low frequency coefficients, δ_1=1.125, δ_1=1.25 and R=400 are model parameters

© CRC Press, Wu & Rao Eds, 2006

9.3 JND with Pixels

- Considers
  - Luminance adaptation
  - Texture masking
  - Spatial contrast sensitivity
- Spatial JND

$$P(i,j) = T^L(i,j) + T^t(i,j) - C^{Lt}(i,j) \cdot \min\{T^L(i,j), T^t(i,j)\}$$

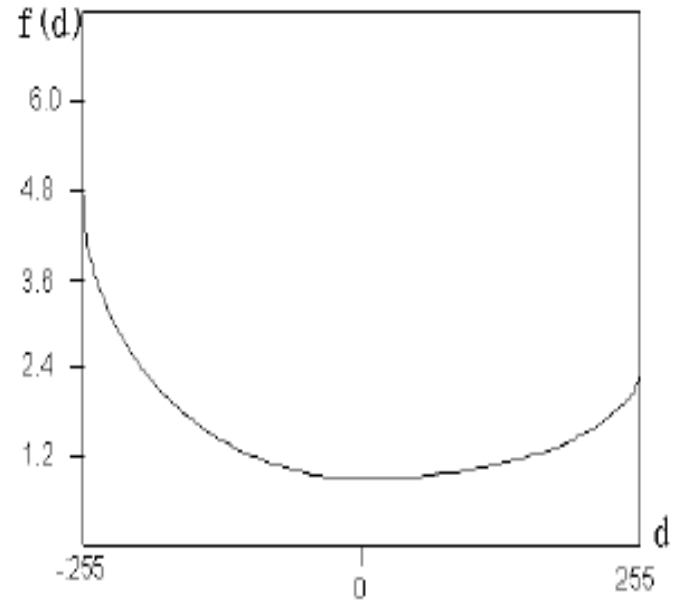
with $T^L(i,j)$ and $T^t(i,j)$ visibility threshold for luminance adaptation and texture masking.

$0 < C^{Lt} \leq 1(i,j)$ accounts for overlapping masking effect

© CRC Press, Wu & Rao Eds, 2006

Temporal Masking Effect

Right: masking f(d) vs frame change (d)



- Modelled as

$$P_v(i,j) = f(a(i,j)) \cdot P(i,j)$$

where $f(a)$ is the temporal masking factor defined as

$$a(i,j) = \frac{\mathrm{x}[i,j] - x^p[i,j]}{2} + \frac{L_0 - L_0}{2}$$

© CRC Press, Wu & Rao Eds, 2006

9.3.2 Conversion between Subband and Pixel Based JNDs

- Pixel-wise JND derived from summation of subband JND

$$t_p(n,x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} \phi_u \phi_v \cos\left(\frac{u(2x+1)\pi}{2N}\right)\cos\left(\frac{v(2y+1)\pi}{2N}\right)\cdot t'(n,i,j)$$

$$t'(n,i,j) = \begin{cases} sign(\mathrm{X}[n,k,l]\cdot s(n,k,l), & |\mathrm{X}[n,k,l]| \geq s(n,k,l) \\ 0 & otherwise \end{cases}$$

  - Truncation of threshold t' avoids over estimation of smooth region
  - Sign operator eliminates blocking effect

© CRC Press, Wu & Rao Eds, 2006

- Estimation of subband JND

$$s(n,k,l) = [E_p(n) \cdot z(k,l)]^{1/2}$$

with $E_p(n)$, the pixel-wise JND energy in N x N image neighbourhood

$$E_P(n) = \sum_{\zeta=0}^{N-1}\sum_{\eta=0}^{N-1} P^2\left(\zeta + \left(\zeta + \left\lfloor \frac{n \cdot N}{W} \right\rfloor \frac{W}{N}\right) \cdot N, \eta + \left\lfloor \frac{n \cdot N}{W} \right\rfloor \cdot N\right)$$

and z containing the parameter weights (*below*)

$$\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} z(u,v) = 1$$

0.0211	0.0197	0.0273	0.0445
0.0211	0.0261	0.0365	0.0581
0.0396	0.0464	0.0626	0.0954
0.0865	0.0990	0.1286	0.1865

© CRC Press, Wu & Rao Eds, 2006

9.4 JND Model Evaluation

- Noise shaping in image/video
  - i.e., inject noise into images/videos and see if the noise level is noticeable
  - Pixel domain noise shaping

$$\tilde{\mathrm{x}}[i,j] = \mathrm{x}[i,j] + q \cdot S^{random}(i,j) \cdot P(i,j)$$

  - Transform domain noise shaping

$$\tilde{\mathrm{X}}[n,k,l] = \tilde{\mathrm{X}}[n,k,l] + q \cdot S^{random}(n,k,l) \cdot s(n,k,l)$$

S^{random} takes ± 1 with $P(i,j)$ and $s(n,k,l)$ the pixel and transform domain JND

q < 1 shapes the noise at perceptual lossless quality

© CRC Press, Wu & Rao Eds, 2006

Noise injected *Cameraman* images. MSE = 144.
Left: DCT domain JND. *Right*: Pixel domain JND

© CRC Press, Wu & Rao Eds, 2006

- For same total error energy
  - Accuracy of JND model improves with perceptual quality of noise-injected images/video
- For same perceptual visual quality
  - Accuracy of the JND model dictates the amount of noise that could be embedded into images/video

© CRC Press, Wu & Rao Eds, 2006

9.5 Conclusions

- JND models
  - Pixel domain
  - Transform (DCT) domain
- JND domain conversion
  - from pixel to DCT and vice versa
- Evaluation of JND models
  - Systematic JND-guided noise injection approach

© CRC Press, Wu & Rao Eds, 2006

- Applications of JND models
  - Perceptual quality/distortion metric
  - Picture compression
  - Data hiding
  - Error concealment

© CRC Press, Wu & Rao Eds, 2006

AK01	A.J. Ahumada and W.K. Krebs. Masking in color images. *Proc. of SPIE Human Vision and Electronic Imaging VI*, 4299:187-194, Jan. 2001
AP92	A.J. Ahumada and H.A. Peterson. Luminance-model-based DCT quantization for color image compression. *SPIE Proc. Human Vision, Visual Processing and Digital Display III*, 3665-374, 1992
Bar99	P. Barten. Contrast Sensitivity of the Human Eye and Its Effects on Image Quality. *SPIE Press*, 1999
CB99	Y.J. Chiu and T. Berger. A Software-only Videocodec Using Pixelwise Conditional Differential Replenishment and Perceptual Enhancement. *IEEE Trans. Circuits and Systems for Video Technology*, 9(3):438-450, April 1999
CC96	C.H. Chou and C.W. Chen. A Perceptually Optimized 3-D Subband Image Codec for Video Communication over Wireless Channels. *IEEE Trans. Circuits and System for Video Technology*, 6(2):143-156, 1996
CL95	C.H. Chou and Y.C. Li. A perceptually tuned subband image coder based on the measure of Just-Noticeable-Distortion Profile. *IEEE Trans. Circuits Systems for Video Technology*, 5(6):467-476, Dec. 1995

© CRC Press, Wu & Rao Eds, 2006

Cor97	Sarnoff Corp. Sarnoff JND vision model. *Contribution to IEEE G-2.1.6 Compression and Processing Subcommittee*, Aug. 1997
Gha03	M. Ghanbari. Visibility of wavelet quantization noise. *Standard Codecs: Image Compression to Advanced Video Coding*, Aug. 2003
HK02	I. Höntsch and L.J. Karam. Adaptive image coding with perceptual distortion control. *IEEE Trans. on Image Processing*, 11(3):214-222, Mar. 2002
HM98	P.J. Hahn and V.J. Mathews. An analytical model of the perceptual threshold function for multichannel image compression. *Proc. IEEE ICIP*, 3:404-408, 1998
ITU00	ITU. Methodology for the subjective assessment of the quality of television pictures. *ITU-R Recommendation 500-10*, Geneva, Switzerland, 1998
JJS93	N. Jayant, J. Johnston and R. Safranek. Signal compression based on models of human perception. *Proc. IEEE*, 81:1385-1422, Oct. 1993
LDX03	W. Lin, L. Dong and P. Xue. Discriminative Analysis of Pixel Difference Towards Picture Quality Prediction. *Proc. IEEE ICIP*, 3:193-196, Sept. 2003

© CRC Press, Wu & Rao Eds, 2006

LF80	G.E. Legge and J.M. Foley. Contrast masking in human vision. *J. Opt. Soc. Am.*, 70(12):1458-1471, 1980
MGE+01	J. Malo, J. Gutierrez, I. Epifanio, F. Ferri and J.M. Artigas. Perceptual feedback in multigrid motino estimation using an improved DCT quantization. *IEEE Trans. Image Processing*, 10(10):1411-1427, Oct. 2001
Nad00	M.J. Nadenau. *Integration of Human Color Vision Models into Hight Quality Image Compression*. Ph.D. thesis, EPFL, Lausanne, 2000
NH88	A.N. Netravali and B.G. Haskell. Digital Pictures-Representation, Compression. *Proc. SPIE Human Vision, Visual Processing and Digital Display VI*, 1998
OT86	L.A. Olzak and J.P. Thomas. Seeing spatial patterns, vol. 1. In *Handbook of Perception and Human Performance*. New York: John Wiley and Sons, 1986
PAW93	H.A. Peterson, A.J. Ahumada and A.B. Watson. An improved detection model for DCT coefficient quantization. *Proc. SPIE Human Vision, Visual Processing and Digital Display VI*, 1913-191-201, 1993

© CRC Press, Wu & Rao Eds, 2006

PPMP91	H.A. Peterson, H. Peng, J.H. Morgan and W.B. Pennebaker. Quantization of color images components in the DCT domain. *Proc. SPIE Human Vision, Visual Processing and Digital Display II*, 1453:210-222, 1991
RPG99	M. Ramasubramanian, S.N. Pattanaik and D.P. Greenberg. A perceptually based physical error metric for realistic image synthesis. *Computer Graphics*, 73-82, Aug. 1999
Saf94	R.J. Safranek. A JPEG compliant encoder utilizing perceptually based quantization. *Proc. SPIE Human Vision, Visual Processing and Digital Display V*, 2179:117-126, Feb. 1994
SJ89	R.J. Safranek and J.D. Johnston. A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression. *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing*, 1945-1948, 1989
TS96	T.D. Tran and R. Safranek. A locally adaptive perceptual masking threshold model for image coding. *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing*, 4:1883-1886, 1996

© CRC Press, Wu & Rao Eds, 2006

TV98	H.Y. Tong and A.N. Venetsanopoulos. A perceptual model for JPEG applications based on block classification, texture masking and luminance masking. *Proc. IEEE ICIP*, 3:428-432, 1998
vNB67	F.L. van Nes and M.A. Bouman. Spatial modulation transfer in the human eye. *Journal of the Optical Society of America*, 57:401-406, Mar. 1967
Wat93	A.B. Watson. DCTune: A technique for visual optimization of DCT quantization matrices for individual images. *Society for Information Display Digest of Technical Papers XXIV*, 946-949, 1993
Wat00	A.B. Watson. Proposal: Measurement of a JND Scale for Video Quality. Submission to the IEEE G-2.1.6 Subcommittee on Video Compression Measurements Meeting on August 7th, 2000
WHMI01	A.B. Watson, J. Hu and J.F. McGowan III. DVQ: A digital video quality metric based on human vision. *Journal of Electronic Imaging*, 10(1):20-29, January 2001.
WPD99	R.B. Wolfgang, C.I. Podilchuk and E.J. Delp. Perceptual Watermarks for Digital Images and Video. *Proc. IEEE*, 87(7):1108-1126, July, 1999

© CRC Press, Wu & Rao Eds, 2006

WYSV97	A.B. Watson, G.Y. Yang, J.A. Solomon and J. Villasenor. Visibility of wavelet quantization noise. *IEEE Trans. on Image Processing*, 6(8):1164-1175, Aug. 1997
YLL+03a	X. Yang, W. Lin, Z. Lu, E.P. Ong and S. Yao. Just-noticeable-distortion profile with nonlinear additivity model for perceptual masking in color images. *Proc. IEEE ICASSP*, 3:609-612, 2003
YLL+03b	X.K. Yang. W. Lin, Z.K. Lu, E.P. Ong and S.S. Yao. Perceptually-adaptive Hybrid Video Encoding based On Just-noticeable-distortion Profile. *SPIE 2003 Conference on Video Communication and Image Processing*, 5150:1448-1459, 2003
YW98	M. Yuen and H.R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing*, 70:247-278, 1998
Zen99	W. Zeng. Visual Optimization in Digital Image Watermarking. *Proc. ACM Multimedia Workshop on Multimedia and Security*, 1999
Zha04	X.H. Zhang. Just-noticeable distortion estimation for images. Master's thesis, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Feb. 2004

© CRC Press, Wu & Rao Eds, 2006

ZLX03	X.H. Zhang, W.S. Lin and P. Xue. A new DCT-based just-noticeable distortion estimator for images. *Proc. Joint Conference of 4th International Conference on Information Communication and Signal Processing and 4th IEEE Pacific Rim Conference on Multimedia*, 2003, Singapore

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

10. No-Reference Quality Metric for Degraded and Enhanced Video

J.E. Caviedes[†] and F. Oberti[‡]

†*Intel Corporatrion, USA*

‡ *Philips Research, The Netherlands*

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 10 Outline

10.1 Introduction

10.2 State-of-the-Art for No-Reference Metrics

10.3 Quality Metric Components and Design

10.4 No-Reference Overall Quality Metric

10.5 Performance of the Quality Metric

10.6 Conclusions and Future Research

© CRC Press, Wu & Rao Eds, 2006

10.1 Introduction

- Objective Image Quality (OIQ) Metric
  - Predict standard subjective quality ratings
  - Substantial work carried out by VQEG
    - Supported by ITU and broadcast industry
    - Work focused on quality measures for broadcasting applications
      - Should consider enhancements as well as distortions
      - Must be applicable along any point of broadcast chain
        - » Including points with no references
        - » Need for No-Reference Metric

© CRC Press, Wu & Rao Eds, 2006

- No-Reference Objective Quality Metrics (NROQM)
  - Application areas
    - Video coding system development
    - Embedded monitoring and control of quality
    - Evaluation of picture display (coding) systems
  - Evaluations picture features
    - e.g. sharpness, contrast, distortions, etc
  - Incorporates aspects of the HVS

© CRC Press, Wu & Rao Eds, 2006

10.2 State-of-the-Art for No-Reference Metrics

- Digital Video Quality (DVQ) Analyzer
  - Measure blocking distortions in video
    - Measurements account for motion and spatial masking
- Picture Appraisal Rating (PAR)
  - Estimates PSNR from quantized data
- Tektronix's PQM300
  - Blocking distortion metric for video

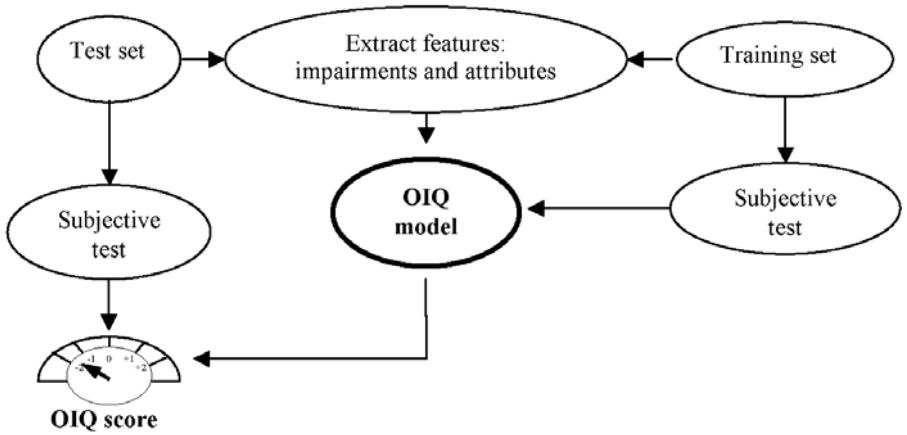
© CRC Press, Wu & Rao Eds, 2006

10.3 Quality Metric Components and Design

- Centred on feature measurement
- Two aspects to feature quantification
  - Desirables which adds to quality
    - Sharpness, contrast, etc
  - Undesirables that impedes quality
    - Impairments, e.g., blurring, blocking, etc

© CRC Press, Wu & Rao Eds, 2006

Framework for the development of no-reference metric (*above*)



- Train the model to predict subject responses
  - Gather subjective data associated with images/videos
  - Optimize model with test images/videos to produce the best prediction match with subjective data

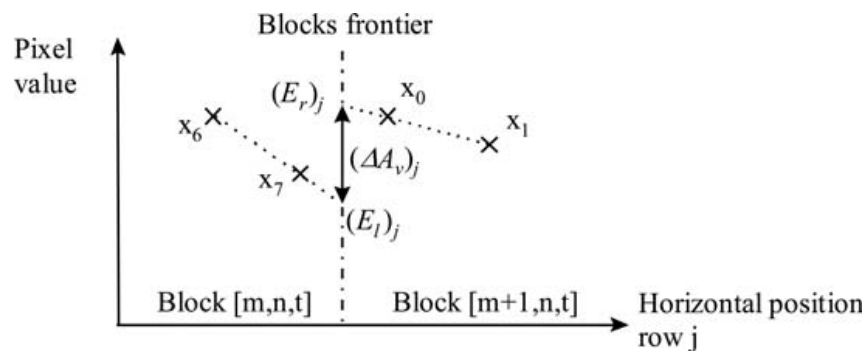
© CRC Press, Wu & Rao Eds, 2006

10.3.1 Blocking Artifacts

- Due to coarse quantization of DCT coefficients
- Occurs along block boundaries of blocks
- Need to differentiate between blocking edges and content edges
  - Threshold measurements
    - Assumes blocking edges below certain threshold
      - Measure magnitude of discontinuities or
      - Calculate extrapolated discontinuities of blocks

© CRC Press, Wu & Rao Eds, 2006

Extrapolation for blocking detection



- $(E_r)_j$ and $(E_l)_j$ are extrapolated values from adjacent blocks
- Discontinuity is the difference between all extrapolated values $(E_r)_j$ and $(E_l)_j$ along the block boundary

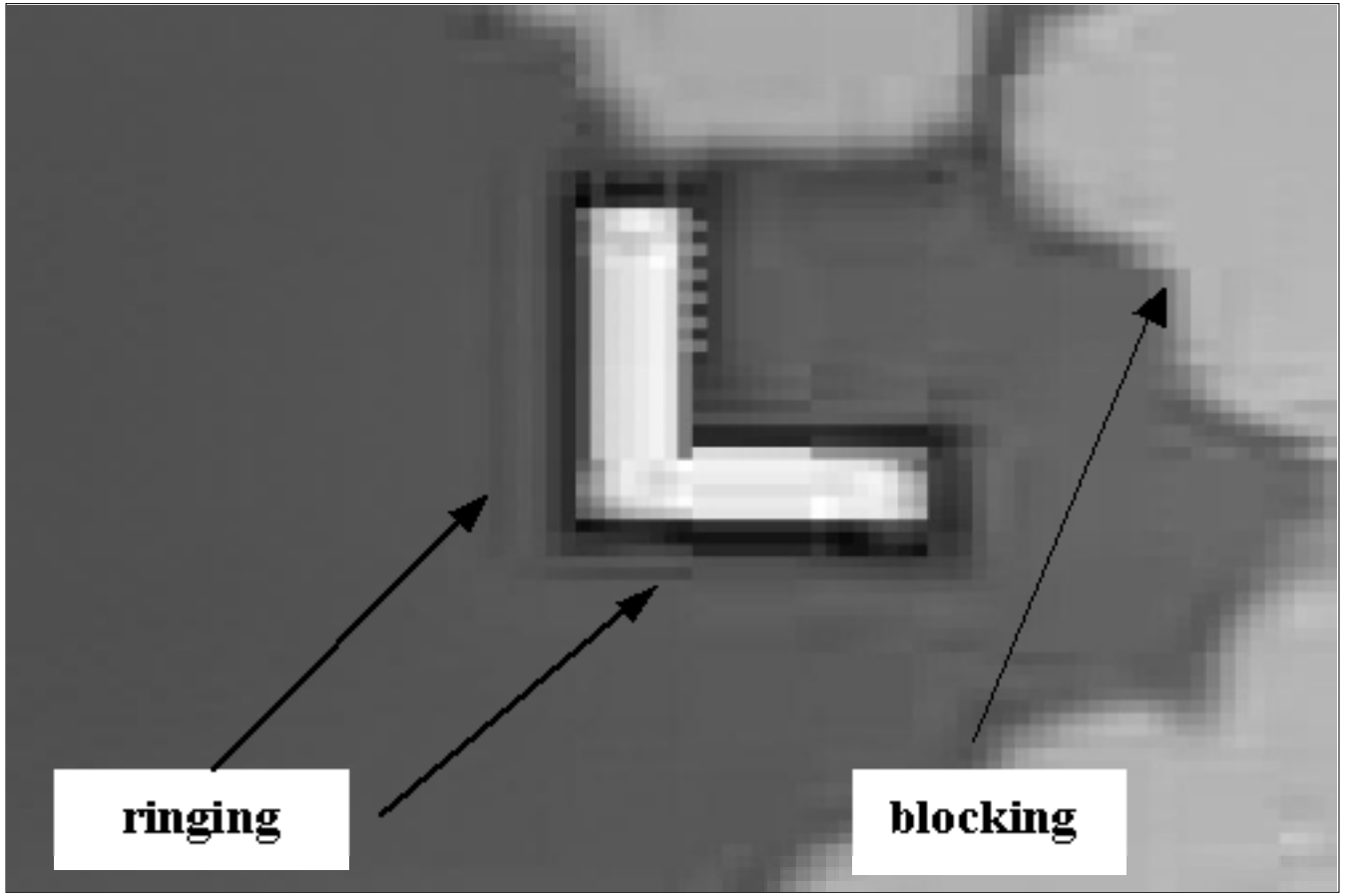
© CRC Press, Wu & Rao Eds, 2006

10.3.2 Ringing Artifacts

- Ringing definition
  - Shimmering effect around high contrast edges
- Manifests as edge doubling
- Ringing and blocking not necessarily correlated
  - Ringing depends on the strength of edges in images
  - Blocking depends on presence smoothly varying areas

© CRC Press, Wu & Rao Eds, 2006

Blocking and ringing artifacts



© CRC Press, Wu & Rao Eds, 2006

- General ringing detection strategy
  - Edge detection
  - Spatial activity measurement around edges
    - Usually based on the variance of a windowed area
  - Classify ringing pixels relative to variance
  - Examples metric
    - Visible Ringing Measure (VRM)
      - Also employ morphing operations for detection

© CRC Press, Wu & Rao Eds, 2006

10.3.3 Clipping

- Truncation of pixel/coefficient magnitude due to the arithmetic precision process
  - Leads to aliasing artifacts
  - Introduced during picture coding and processing
    - e.g. “peaking”, a sharpness enhancement process may lead to clipping when enhancement pictures overshoots its dynamic operational range
- General clipping measures
  - Quality the number of clipped pixels

© CRC Press, Wu & Rao Eds, 2006

10.3.4 Noise

- Random variation in spatial and temporal dimension
  - Introduced during transmission or picture generation
  - May enhance or degrade picture quality
  - Most noticeable in relatively smooth regions
    - Appears as speckle in spatial domain
    - Fluctuation in temporal domain

© CRC Press, Wu & Rao Eds, 2006

- Noise measurement
  - Generally reliant of variance
    - To classify noisy regions
  - Assumes noise to be considerably weaker then image signals
    - Isolated noise (or signal) with low pass or band pass filtering
    - Quantify isolated noise

© CRC Press, Wu & Rao Eds, 2006

10.3.5 Contrast

- The difference in luminance between a pixel and its background
- Contrast dependent on
  - Object in question, background, mean luminance and colour
- Simple Measurement of contrast levels
  - Compute luminance histogram
  - Calculated the difference (in luminance energy) between the upper and lower halves of the histogram

© CRC Press, Wu & Rao Eds, 2006

10.3.6 Sharpness

- Quantify by the clarity of details in images
  - Through preservation/definition of edges
- Sharpness depends on
  - Resolution, content, contrast and noise
- Sharpness metric with local edge kurtosis
  - Generate edge image
    - Partition edge image into blocks
  - Apply 2-D kurtosis to individual blocks
    - Sharpness is the average 2-D kurtosis

© CRC Press, Wu & Rao Eds, 2006

Edge detection and block assignment of *Doll* sequence for sharpness measurement

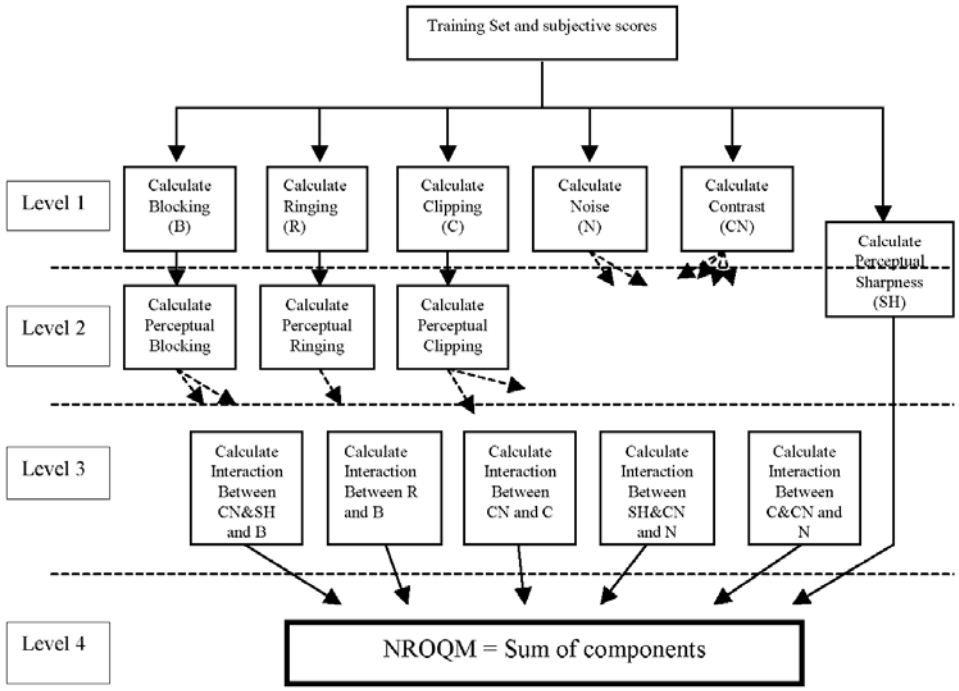
© CRC Press, Wu & Rao Eds, 2006

10.4 No-Reference Overall Quality Metric (NROQM)

- Consists of components described in section 10.3
  - Contrast, noise, clipping, sharpness, blocking and ringing
- Model Optimization
  - Find a set of NROQM parameters that best match subjective data
  - Incremental heuristic knowledge approach
    - More precise calibration of metric components

© CRC Press, Wu & Rao Eds, 2006

Operational flowchart of the NROQM



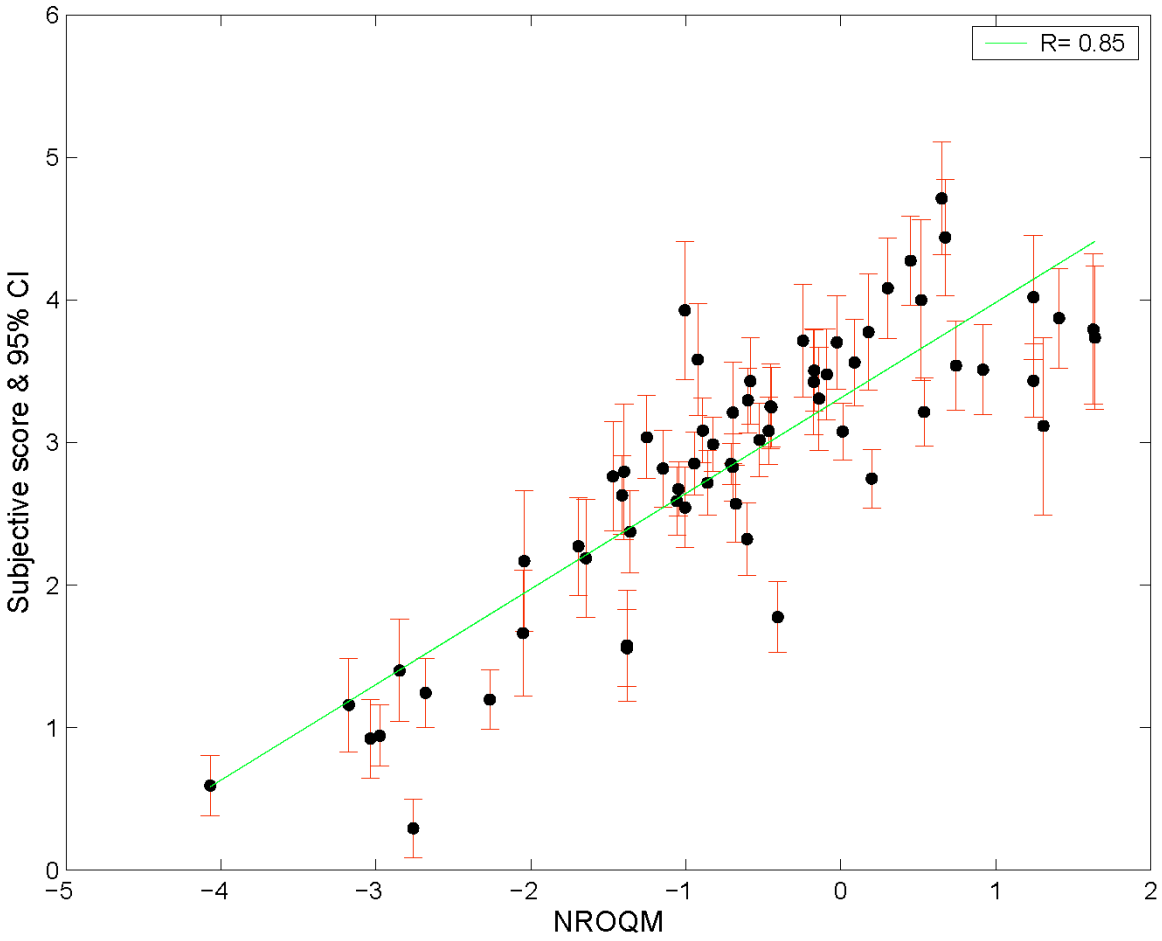
© CRC Press, Wu & Rao Eds, 2006

10.5 Performance of the Quality Metric

- Evaluated with 65 sequences
  - With various level of processing
    - Coding, sharpness enhancement and noise reduction
  - Derived from 3 sequences
    - *Doll, Dolphins and Lawnmowerman*
  - Subjective assessment
    - 20 non-expert subjects
    - 7-point scale comparative test
  - Correlation: NROQM=0.85, PSNR=0.399

© CRC Press, Wu & Rao Eds, 2006

Performance of NROQM



© CRC Press, Wu & Rao Eds, 2006

- Additional evaluations
  - With different sequences (*Fade* and *Boat*)
    - Correlation: 0.848 (*Fade*) and 0.943 (*Boat*)
  - Expert viewers
    - Extreme enhancement of test material
      - NROQM operating outside operational range
        - » Not calibrated for extreme enhancement
    - Resulting evaluation uncovers some deficiencies
      - High clipping
      - Sharpness in high contrast areas
      - Over sharpening

© CRC Press, Wu & Rao Eds, 2006

10.6 Conclusions and Future Research

- NR metric has practical applications
  - picture quality and impairment measurement
  - Quality assessment for coding system development
  - Real-time monitoring and control of picture quality
- NROQM an example of NR metric
  - Demonstrated operability/applicability
  - Potential for improvement

© CRC Press, Wu & Rao Eds, 2006

- Areas for future development
  - Motion field smoothness
  - Mosquito noise
  - IBP frame drift
  - Judder
    - Unnatural modulation of smooth motion

© CRC Press, Wu & Rao Eds, 2006

ATI01	Methodological framework for specifying accuracy and cross-calibration of video quality metrics, Sept. 2001. Technical Report No. T1.TR.PP.72.2001, ATIS T1A1.1
Can86	T. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis Machine Intelligence*, 8(6):679-698, Nov. 1986
CDGR00	J.E. Caviedes, A. Drouot, A. Gesnot and L. Rouvellou. Impairment metrics for digital video and their role in objective quality assessment. In *Proc. VCIP 2000*, 791-800, perth, Australia, June 2000
CG02	J.E. Caviedes and S. Gurbuz. No-reference sharpness metric based on local edge kurtosis. In *Proc. ICIP*, 3:53-56, Rochester, NY, Sept. 2002
CJ01	J.E. Caviedes and J. Jung. No-reference metric for a video quality control loop. In *Proc. SCI2001*, 290-295, Orlando, FL, July 2001
dH00	G. de Haan. *Video Processing for Multimedia Systems*. Eindhoven: University Press Facilities, 2000

© CRC Press, Wu & Rao Eds, 2006

Eng00	P.G. Engeldrum. *Psychometric Scaling: A Toolkit for Imaging Systems Development*. Winchester, MA: Imcotek Press, 2000
Fai99	M.D. Fairchild. A victory for equivalent background – on average. In *Proc. IS&T/SID 7th Color Imaging Conference*, 87-92. Scottsdale, AR, 1999
GEN	Genista Corp. (http://www.genista.com)
GPLL99	C. Glassman, A. Peregoudov, A. Logunov and V. Lichahov. Video compression artifacts: predicting the perceptual ratings. In *Proc. IBC*, 560-564, Amsterdam, 1999
Inc97	T. Inc. A guide to picture quality measurements for modern television systems, Beaverton, OR, 1997. (http://www.tektronix.com)
ITU00	ITU. Methodology for the subjective assessment of the quality of television pictures. *ITU-R Recommendation 500-0*, Geneva, Switzerland, 2000
ITU03	Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference, 2003. Draft New Recommendation ITU-T BT.[Doc. 6/39]

© CRC Press, Wu & Rao Eds, 2006

JF00	G.M. Johnson and M.D. Fairchild. Sharpness rules. In *Proc. IS&T/SID 8th Color Imaging Conference*, 24-30, Scottsdale, AR, November 2000
KK95	S.A. Karunasekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transaction on Image Processing*, 4(6):713-724, June 1995
Ogu99	S.H. Oguz. *Morphological post-filtering of ringing and lost data concealment in generalized lapped orthogonal transform based image and video coding*. Ph.D. thesis, University of Wisconsin, Madison, WI, 1999
PIX	Pixelmetrix Corporation (http://www.pixelmatric.com)
ROH	Rohde & Schwarz (http://www.rohde-schwarz.com)
SDZ96	B. Spehar, J.S. DeBonet and Q. Zaidi. Brightness induction from uniform and complex surrounds: A general model. *Vision Research*, (36):1893-1906, July 1996
SNE	Snell & Wilcox (http://www,snellwilcox.com)
TCC02	D.S. Turaga, Y. Chen and J.E. Caviedes. No-reference PSNR estimation for compressed pictures. In *Proc. ICIP*, 3:61-64, Rochester, NY, Sept. 2002

© CRC Press, Wu & Rao Eds, 2006

VQE	The Video Quality Experts Group (http://www.vqeg.org)
VQE03	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment – Phase II, Aug. 2003
vZA01	K. von Zon and W. Ali. Automated video chain optimization. *IEEE Trans. on Consumer Electronics*, 47(3):593-603, Aug. 2001
WL01	A. Worner and J. Lauterjung. A real time single ended algorithm for objective quality monitoring of compressed video signals. In *Proc. Of SCI2001*, 13:329-334, Orlando, FL, July 2001
WY96	H.R. Wu and M. Yuen. Quantitative quality metrics for video coding blocking artifacts. In *Proc. of Picture Coding Symposium 1*, 23-26, Melbourne, Australia, 1996
WYSV97	A. Watson, G. Yang, J. Solomon and J. Villasenor. Visibility of wavelet quantization noise. *IEEE Trans. on Image Processing*, 6:1164-1175, Aug. 1997

© CRC Press, Wu & Rao Eds, 2006

YHTN01	S. Yang, Y. Hu, D. Tull and T. Nguyen. Maximum likelihood parameter estimation for image ringing artifact removal. *IEEE Trans. On CSVT*, 11(8):963-973, August 2001
YWC00	Z.Yu, H.R. Wu and T. Chen. Perceptual blocking distortion measure for digital video. In *Proc. VCIP2000*, 4067:4067:768-779, Perth, Australia, 2000

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

11. Video Quality Experts Group

Philip Corriveau

Intel Media and Acoustics Perception Lab, USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 11 Outline

11.1 Formation

11.2 Goals

11.3 Phase I

11.4 Phase II

11.5 Continuing Work and Directions

11.6 Summary

© CRC Press, Wu & Rao Eds, 2006

1.1 Formation

- VGEG formed in 1997
  - To address standardized objective methods for picture quality measurement
  - From ITU Tele-communication and Radio-communication sectors
    - ITU-R SG-11
    - ITU-T SG-12 and SG-9

© CRC Press, Wu & Rao Eds, 2006

11.2 Goals

- To develop recommendation for objective picture quality assessment
  - Basic requirement
    - Objective metric should provide adequate equivalent subjective assessment of picture quality
    - Must be effective for evaluating standardized video coding schemes, e.g. MPEG-2, H.263, etc
      - At various bitrates from 768 kbps to 36 Mbps
      - At varying spatial and temporal information
        - » interlaced/non-interlaced, PAL/NTSC, etc

© CRC Press, Wu & Rao Eds, 2006

- Plan
  - Model verification process
    - Solicit objective models to be include for verification leading to recommendation
    - Develop of objective and subjective test plan
    - Objective test for evaluating proposed models
    - Subjective test for acquiring subjective data
    - Analyze objective test results and subjective data
  - Present report to ITU

© CRC Press, Wu & Rao Eds, 2006

11.3 Phase I

- Selection of source material
- Selection of test conditions (HRC)
  - e.g. coding rate
  - *Hypothetical Reference Circuits* (HRC)
- Subjective test
  - Follows ITU-R bt.500 recommendations
    - Double Stimulus Continuous Quality Scale (DSCQS) method
  - 20 sequences x 16 HRCs

© CRC Press, Wu & Rao Eds, 2006

- Subjective test (continue)
  - Undertaken is 4 parts
    - 50 Hz Low and High Quality
    - 60 Hz Low and High Quality
  - Run in multiple laboratories
    - CRC (Canada), RAI (Italy), CCETT (France), NHK (Japan), DCITA (Australia), Berkom (Germany) and FUB (Italy)

© CRC Press, Wu & Rao Eds, 2006

- Objective test
  - Ran by individual proponent of each model
  - Proponent results (10%) verified by independent labs
    - CRC (Canada), FUB (Italy), IRT (Germany) and NIST (USA)

© CRC Press, Wu & Rao Eds, 2006

Proponents of objective measures

- [po] Peak Signal to Noise Ratio (PSNR)
- [p1] CPqD (Brazil)
- [p2] Tektronix/Sarnoff (USA)
- [p3] NHK/Mitsubishi Electric Corp. (Japan)
- [p4] KDD (Japan)
- [p5] EPFL (Switzerland)
- [p6] TAPESTRIES (European Union)
- [p7] NASA (USA)
- [p8] KPN Research (The Netherlands)/Swisscom CIT (Switzerland)
- [p9] NTIA/ITS (USA)

© CRC Press, Wu & Rao Eds, 2006

- Comparison metrics
  - Measures a number of attributes between subjective and objective data
    - Prediction accuracy
      - Correlation measures, e.g. Pearson's, Spearman, etc
    - Monotonicity
    - Consistency
    - Variation

© CRC Press, Wu & Rao Eds, 2006

Phase I conclusions

- No objective measurement system could replace subjective testing
- No one objective model outperforms the other in all cases
- Objective assessment methods have improved considerably, but
- No method for ITU recommendation can be made
- Data set generated is uniquely valuable

© CRC Press, Wu & Rao Eds, 2006

11.4 Phase II

- More discriminating than Phase I test
  - More precise area of interest
  - Focused on secondary distribution of digitally coded television quality video
  - Two experiments
    - 525-line and 625-line video covering a wide range of qualities
  - Proponents required to submit video calibration parameters
    - Spatial and temporal registration, gains, etc

© CRC Press, Wu & Rao Eds, 2006

- Laboratories
  - 525-line test - CRC (Canada) and Verizon (USA)
  - 625-line test - FUB (Italy)
- Proponents
  - 10 initially
  - 8 began testing
  - 6 completed testing
  - PSNR included as baseline metric

© CRC Press, Wu & Rao Eds, 2006

- Final proponents
  - A. NASA (USA)
  - D. British Telecom (UK)
  - E. Yonsei University/Radio Research Lab./SK Telecom (Korea)
  - F. CPqD (Brazil)
  - G. Chiba University (Japan)
  - H. NTIA/ITS (USA)

© CRC Press, Wu & Rao Eds, 2006

- Results
  - 625 and 525 results were similar, not identical
  - For 525
    - Models D and H performed better than other models
    - D and H statistically equivalent in performance
  - For 625
    - Models A, E, F and H are statistically equivalent in performance and are better than remaining models
  - PSNR performance is mid range

© CRC Press, Wu & Rao Eds, 2006

- ITU directions
  - Draft standard for estimating perceived video quality
    - 4 model following VQEG Phase II results
      - British Telecom
      - Yonsei University/Radio Research Lab./SK Telecom
      - CPqD
      - NTIA/ITS
  - Eventually recommend only one normative full reference method

© CRC Press, Wu & Rao Eds, 2006

11.5 Continuing Work and Directions

- Continuing work
  - with Full Reference Metric
- Future
  - Reduced Reference Metric
  - No-Reference Metric

© CRC Press, Wu & Rao Eds, 2006

11.6 Summary

- VQEG
  - Purpose
    - Develop a unified an objective video quality measurement method for standardization
  - Existing work
    - Phase I and II test
      - has a number of proponents
      - for applications in broadcasting
      - Produced valuable data for future quality metric work
  - Future
    - Unified FR metric, work in NR and RR metric

© CRC Press, Wu & Rao Eds, 2006

ISO94	ISO/IEC 13818. Information technology – generic coding of moving pictures and associated audio information. Technical report, ISO/IEC, Nov. 1994
ITU98a	ITU. ITU-R BT.500-9, methodology for the subjective assessment of the quality of television pictures. *ITU-T BT*, 1998
ITU98b	ITU-T Video Coding Experts Group. Video coding for low bit rate communication. ITU-T recommendation H.263, Jan. 1998
RH96	K.R. Rao and J.J. Hwang. *Techniques and Standards for Image, Video and Audio Coding*. Upper Saddle River, NJ: Prentice Hall, 1996
VQE99	VQEG, *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment*. VQEG, Dec. 1999
VQE00a	VQEG, *VQEG meeting report for the Gaithersburg meeting*. VQEG, Jan. 2000
VQE00b	VQEG, *VQEG Objective Test Plan*. VQEG, Jan. 2000
VQE00c	VQEG, *VQEG Objective Test Plan*. VQEG, Jan. 2000

© CRC Press, Wu & Rao Eds, 2006

VEQ03	VQEG. *Final Report form the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. VQEG, August 2003
Wor03	Working Group 6Q. *Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference*, ITU-R, Oct. 2003. Working Group 6Q draft new recommendation ITU-R BT.[DOC. 6/39]

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

III. Perceptual Coding and Processing of Digital Pictures

12. HVS Based Perceptual Video Encoders

A. Pica, M. Isnardi and J. Lubin

Sarnoff Corporation Inc., USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 12 Outline

12.1 Introduction

12.2 Noise Visibility and Visual masking

12.3 Architectures for Perceptual based Coding

12.4 Standard-Specific Features

12.5 Salience/Maskability Pre-Processing

12.6 Application to Multi-Channel Encoding

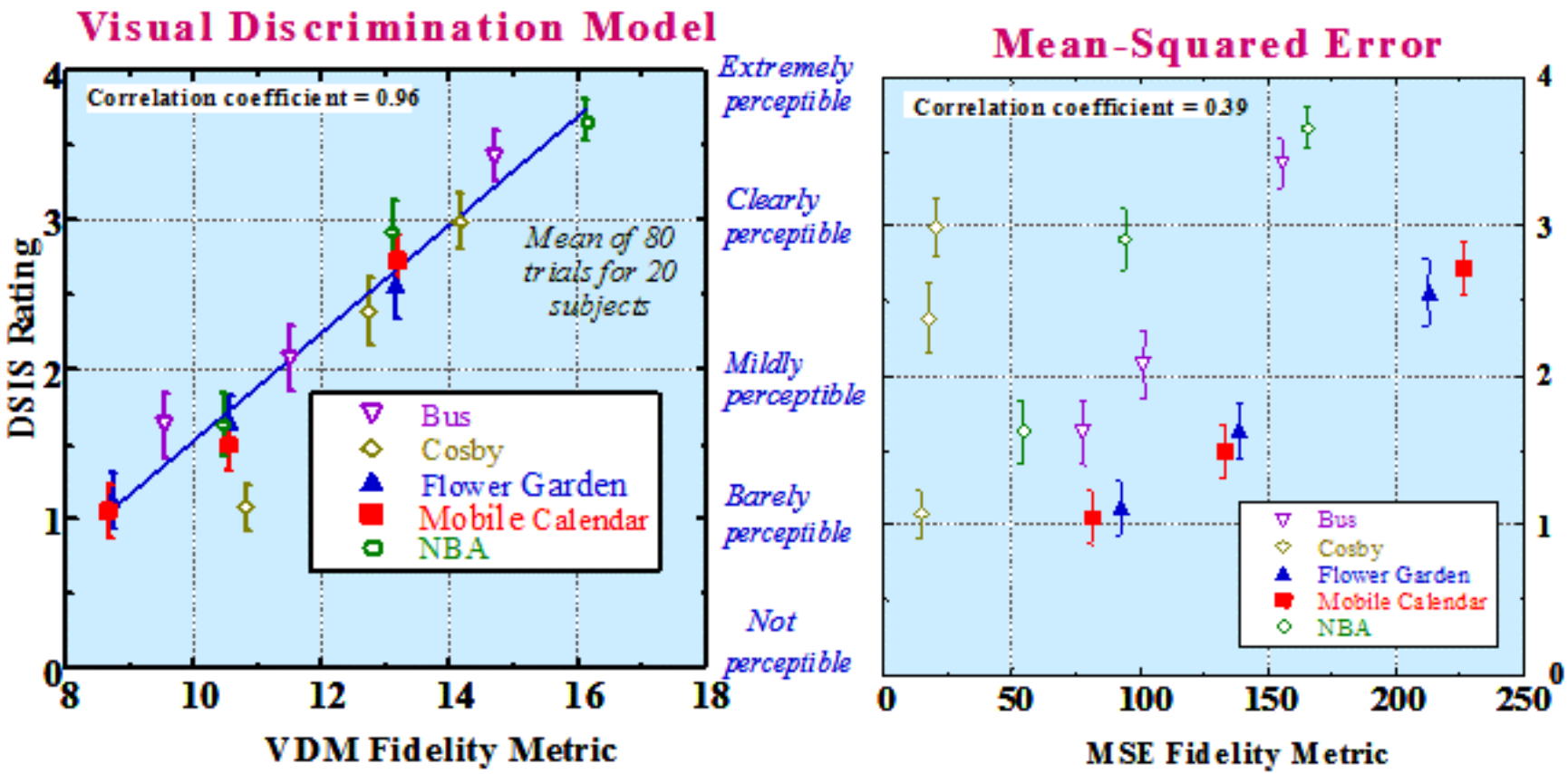
© CRC Press, Wu & Rao Eds, 2006

12.1 Introduction

- Video encoders
  - Traditional approach with MSE type metric
  - HVS based utilizing perceptual metric
- MSE metric
  - Fast computation, simple implementation
  - Poor correlation with perceived distortion
- Perceptual metric
  - Good correlation with subjective data
  - Computationally intensive

© CRC Press, Wu & Rao Eds, 2006

Performance of HVS (*left*) and non-HVS (*right*) based objective metrics



© CRC Press, Wu & Rao Eds, 2006

12.2 Noise Visibility and Visual masking

- Lossy coding systems
  - Introduce distortion to pictures
  - So, how to minimize visible distortions?
- Characteristics of distortions
  - Structural in appearance
    - Manifestation has consistent structure
  - Appearance dependent on image content
    - Certain area of images may hide distortions

© CRC Press, Wu & Rao Eds, 2006

Visibility of noise. *Left*: Gaussian (MSE: 27.1). *Right*: DCT quantization (MSE: 21.3)

© CRC Press, Wu & Rao Eds, 2006

- Block based DCT coders
  - Used for video coding (MPEG standards)
  - Blocking distortions prominent
    - Due to quantization
- Discrete Wavelet Transform (DWT) coder
  - No blocking distortions
  - Ringing distortions at low bitrates

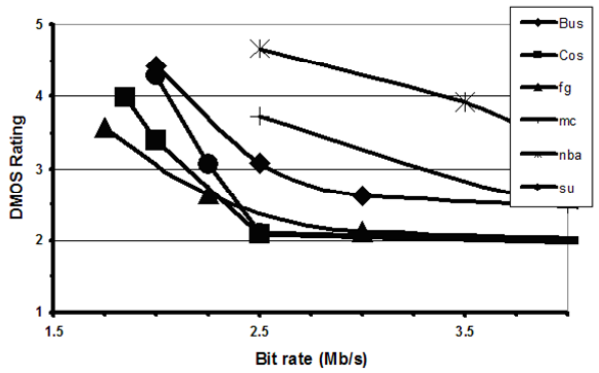
© CRC Press, Wu & Rao Eds, 2006

Images coded at 0.25 bpp. *Left*: DCT (JPEG). *Right*: DWT (JPEG2000)

© CRC Press, Wu & Rao Eds, 2006

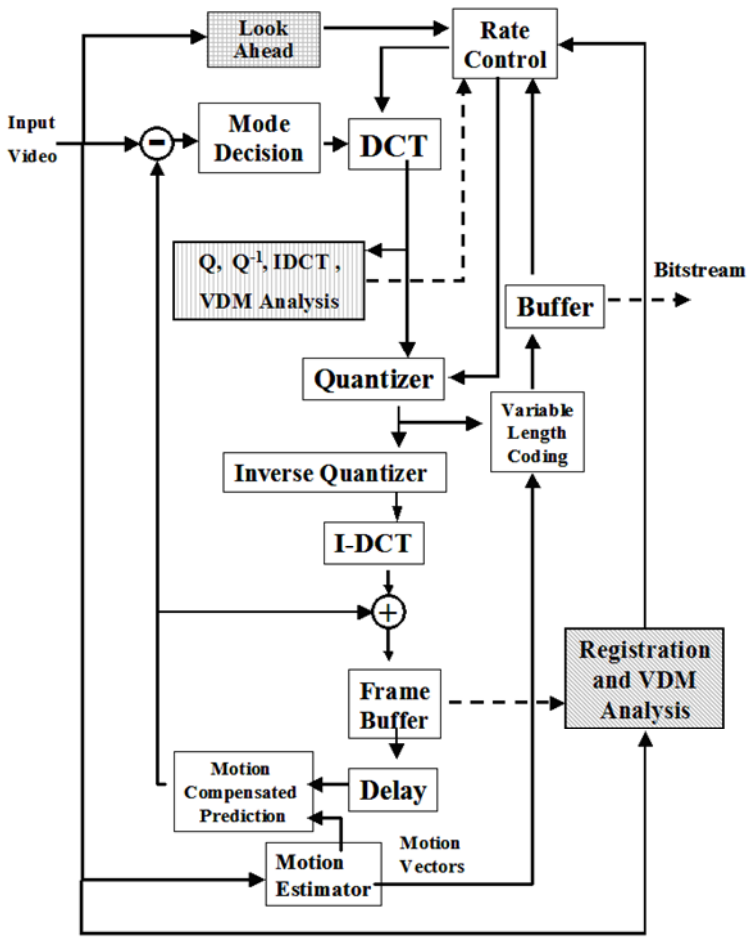
12.3 Architectures for Perceptual based Coding

- The Rate-Distortion (R-D) Curve (*Below*)
  - Regulates distortions relative to bitrates
  - Profile of R-D curves are content dependent
    - i.e., different picture have different R-D curve



© CRC Press, Wu & Rao Eds, 2006

MPEG-2 coder with 2 possible application points (shaded boxes) for R-D optimizer



- Estimated R-D
  - Look-ahead technique
  - Outside coding loop
- Computed R-D with multi-pass coding scheme
  - 1st pass for R-D curve
  - 2nd pass for coding
  - Inside coding loop

© CRC Press, Wu & Rao Eds, 2006

- Latency issue
  - Estimated R-D outside coding loop has high latency
    - may not cope well with rapidly changing content
  - Computed R-D inside coding loop has low latency
    - Distortion levels can be controlled by in-built distortion metric
    - Coding rate is then adjusted to the desired quality level with the R-D curve

© CRC Press, Wu & Rao Eds, 2006

12.3.1 Masking Calculations

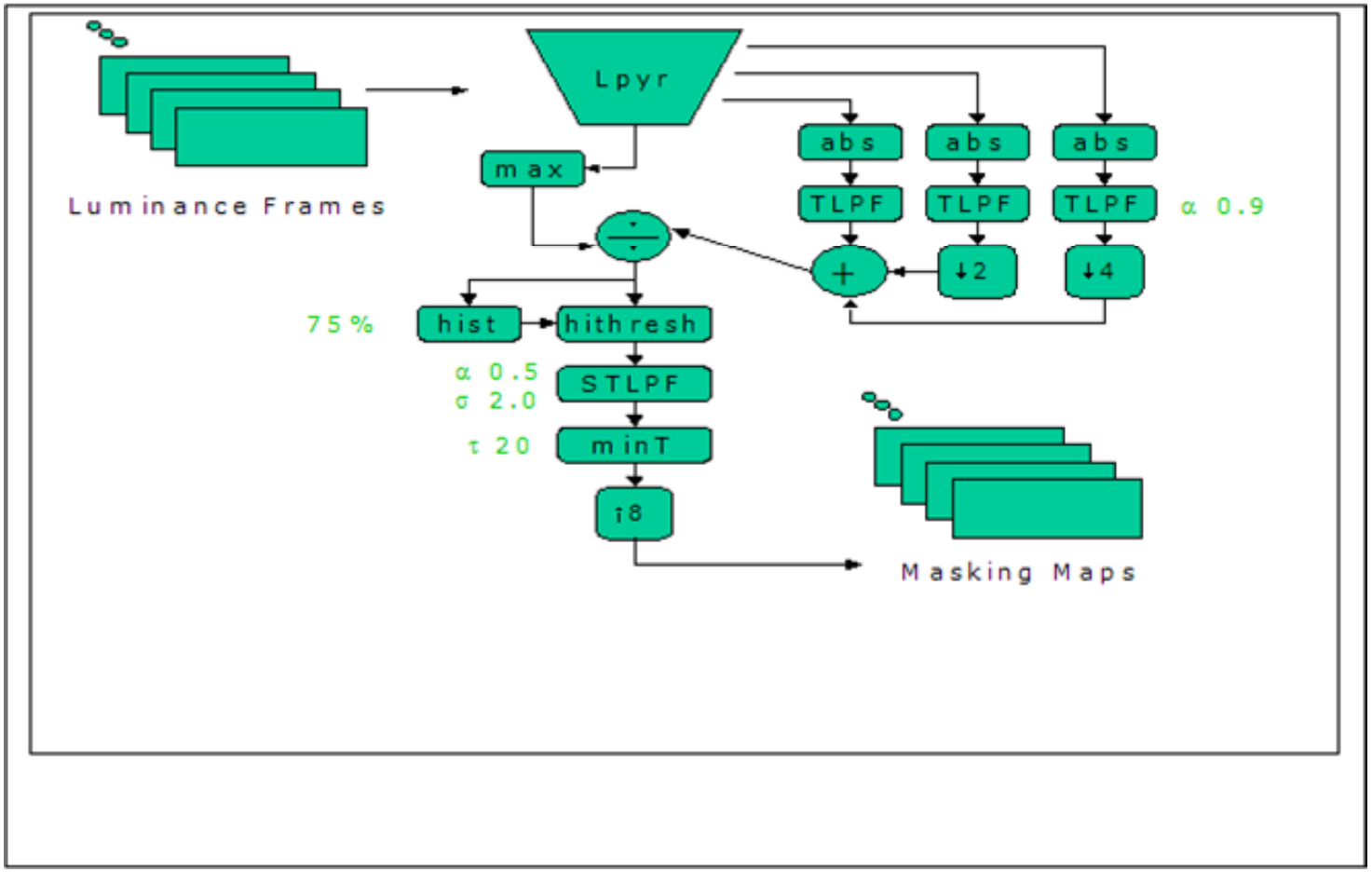
- Masking regions
  - High contrast edges
    - e.g., textured area such as grass
  - Occasionally around visually salient features with broad frequency coverage
    - Some salient features attract more visual attention
      - Therefore, special treatment required to minimize visibility of artifacts

© CRC Press, Wu & Rao Eds, 2006

- Spatio-temporal masking
  - Laplacian pyramid decomposition of frames
    - Into different resolutions, i.e., ½ and ¼
  - Sum of energy across local spatio-temporal frequency and orientation bands
    - Regions of **high** saliency removed from calculations
  - Calculate the local contrast
    - Sum of local energy normalized by the maximum luminance

© CRC Press, Wu & Rao Eds, 2006

Maskability model with salient feature rejection

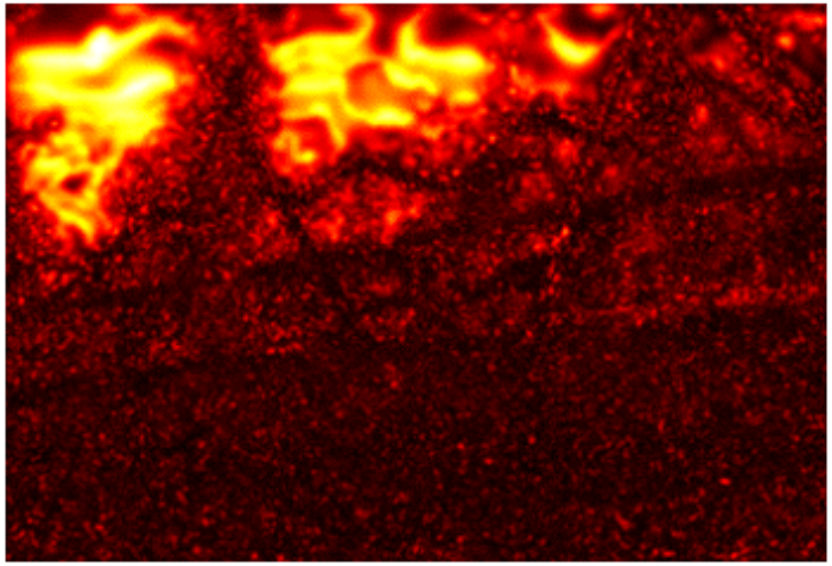


© CRC Press, Wu & Rao Eds, 2006

- Saliency rejection
  - Construct histogram of local contrast energy for each frame
  - Region above percentile threshold are rejected
    - Threshold is a free parameter set to 75%
- Temporal saliency considerations
  - Object based with motion tracking

© CRC Press, Wu & Rao Eds, 2006

Flowergarden sequence (*Left*) and its masking map (*right*). Bright area indicates low masking.



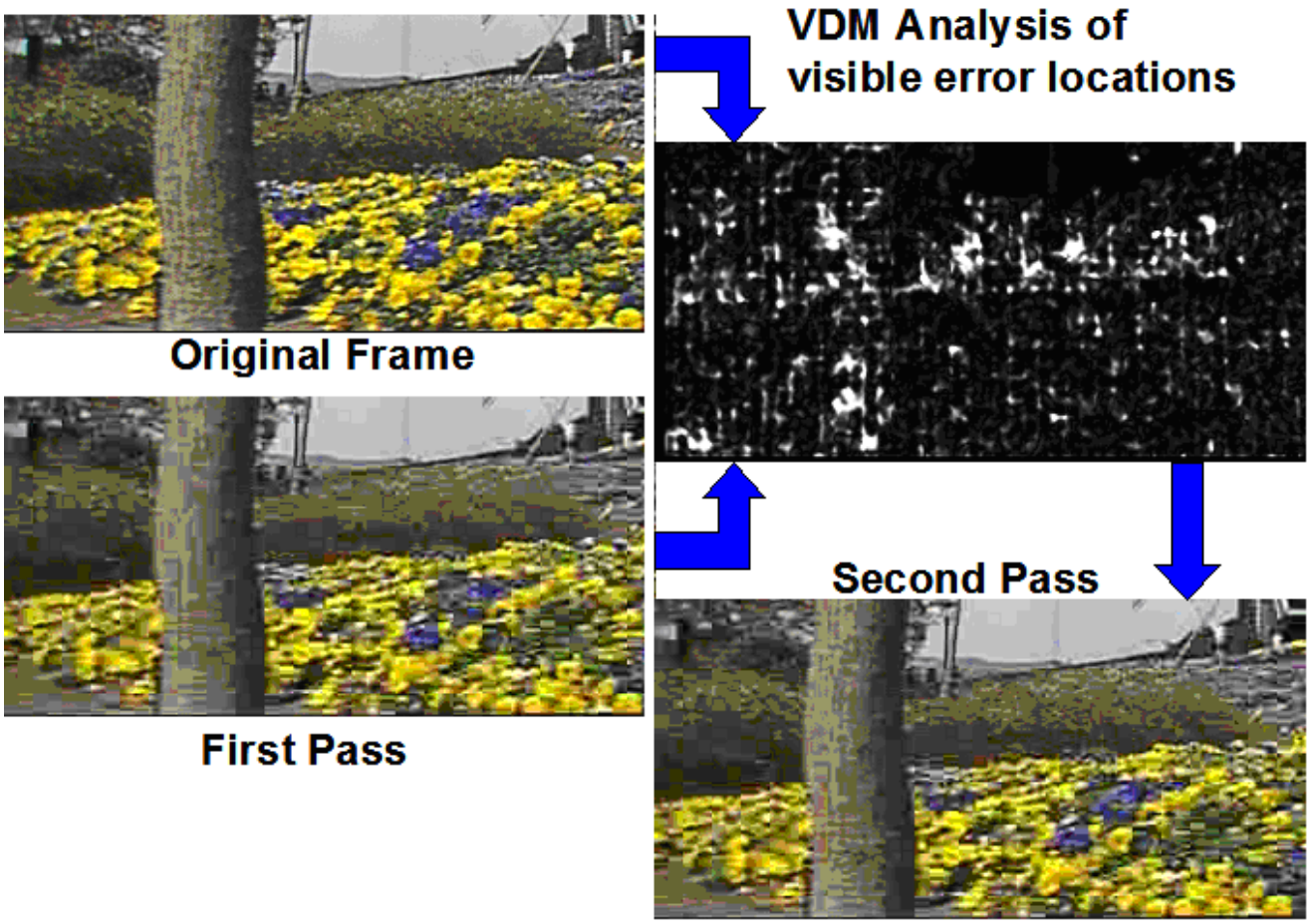
© CRC Press, Wu & Rao Eds, 2006

12.3.2 Perceptual Based Rate Control

- Perceptual metric in rate control at:
  - Macroblock level (block size limitation)
    - Only provides relative estimates of distortions
  - Picture level (pre-assigned bitrate for frame)
    - Good account of visible distortion
    - Uniform distribution of distortion over frame
  - GOP level (pre-assigned bitrate for GOP)
    - Utilize full vision model, i.e., temporal mechanism
    - Uniform distortion distribution over GOP
    - With VBR enables constant quality video

© CRC Press, Wu & Rao Eds, 2006

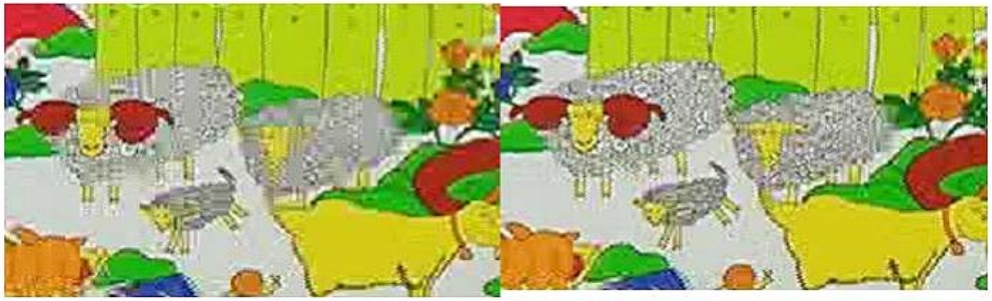
Macroblock bit allocation with the *Visual Discrimination Model* (VDM)



© CRC Press, Wu & Rao Eds, 2006

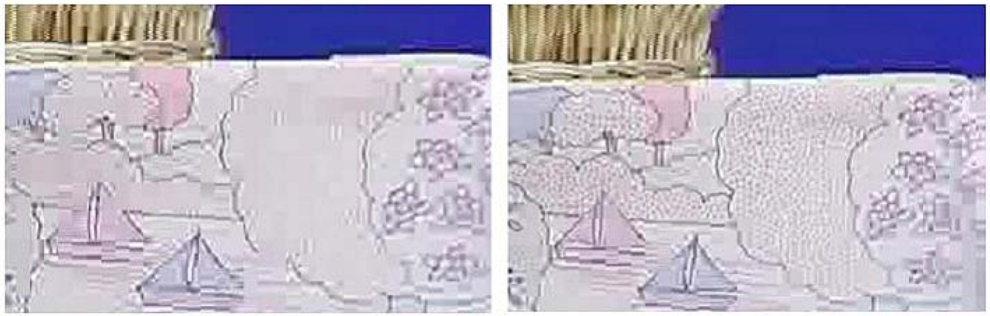
Example of vision based coding with DVC (incorporating VDM) compared to MPEG TM5

Baseline Encoder-TM5



DVC Encoder

Baseline Encoder-TM5



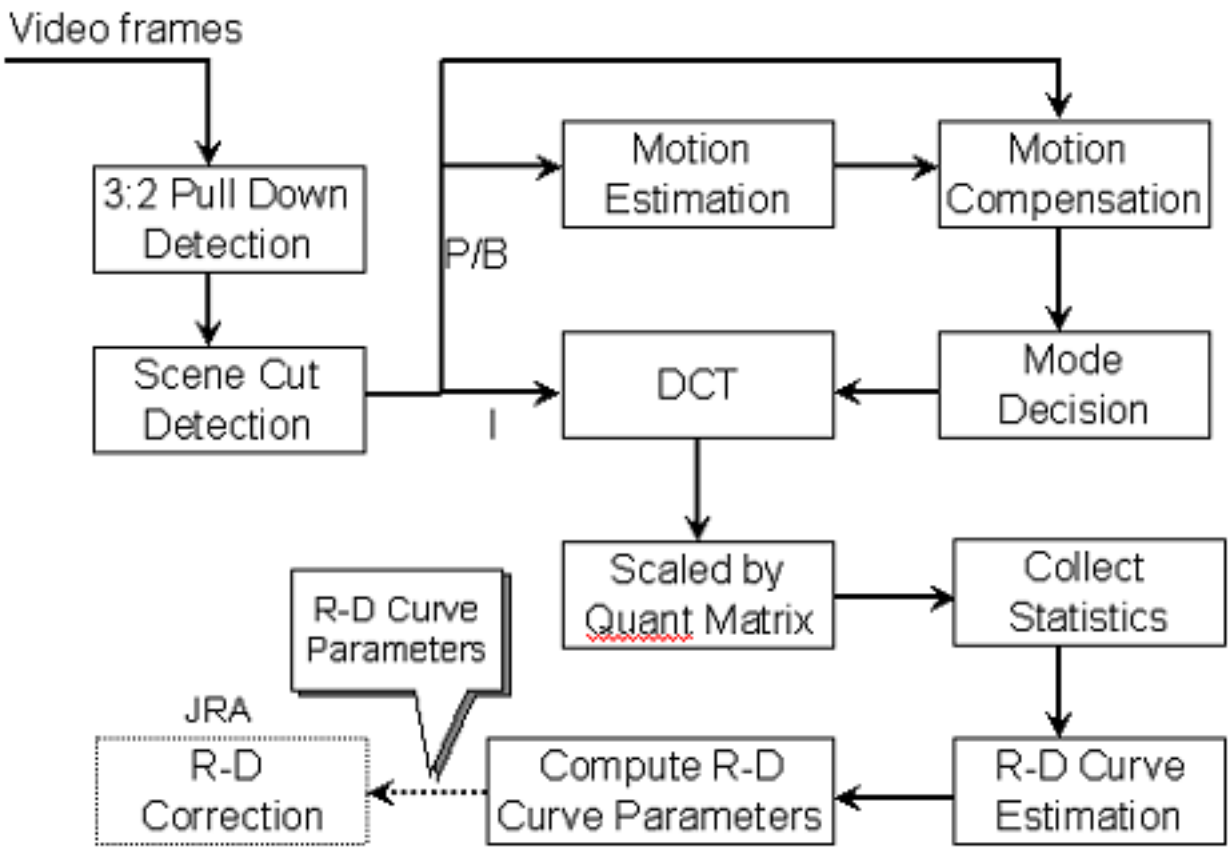
DVC Encoder

© CRC Press, Wu & Rao Eds, 2006

- Look Ahead R-D curve
  - Estimate R-D curves based on sample videos
    - Acquire distortions of sample videos at various bitrates (quantization level)
    - Estimate R-D curve for different video profiles
      - i.e., different GOP
  - Interpolate points in curve for R-D estimation
  - Generally used as the first pass for a multi-pass coding scheme

© CRC Press, Wu & Rao Eds, 2006

Illustration of Sarnoff look-ahead R-D curve estimation (JRA: Joint Rate Allocation)



© CRC Press, Wu & Rao Eds, 2006

- Sarnoff Look-Ahead rate allocation
  - Real-time implementation
  - Picture level rate allocation
  - Operation depends on frame type
    - i.e., I-frame, P-frame or B-frame
  - Histogram approach
  - Utilise Visual MSE to compute error
    - Visual MSE = MSE/(Visual mask)

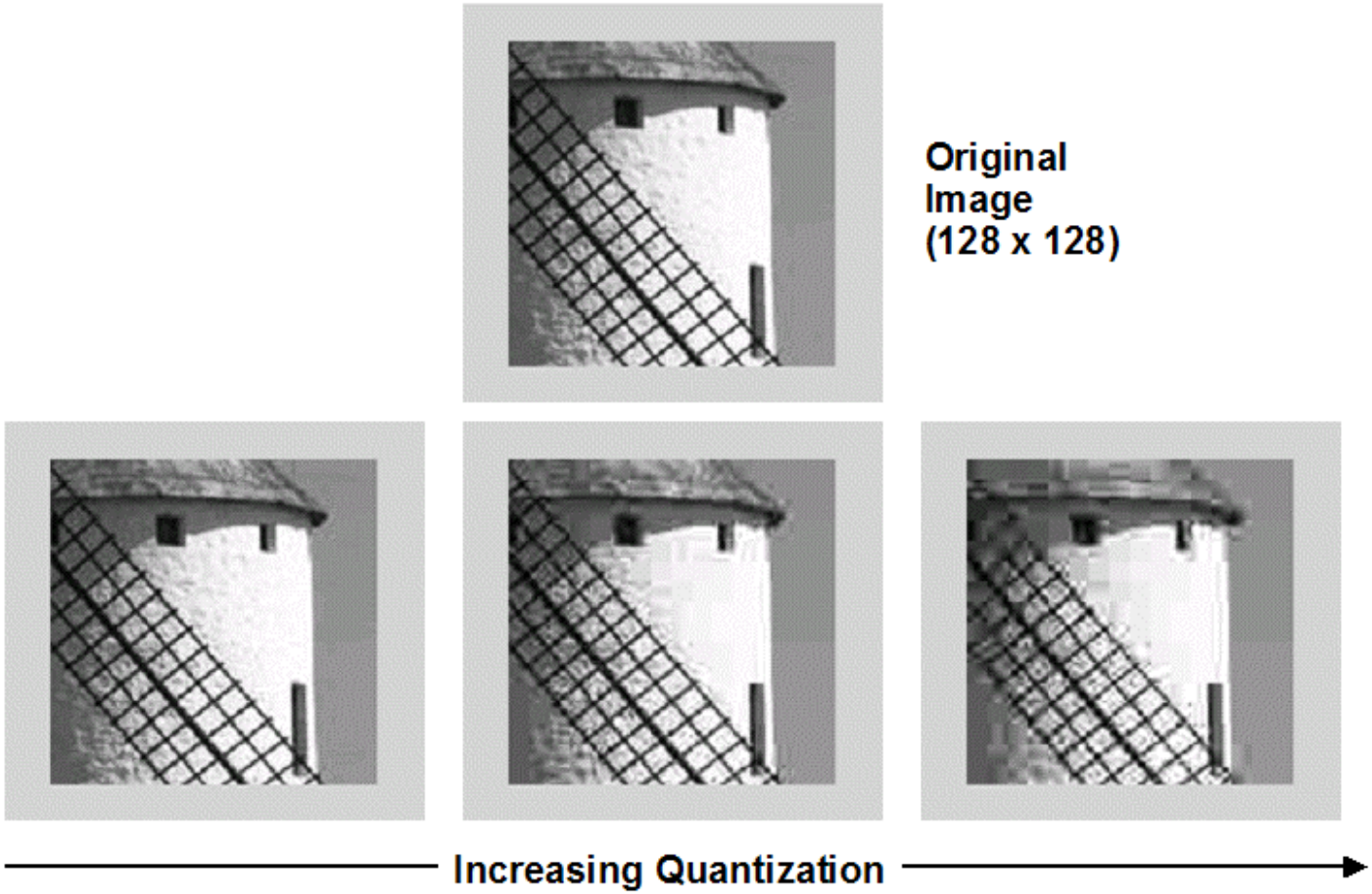
© CRC Press, Wu & Rao Eds, 2006

12.4 Standard-Specific Features

- Block sizes (block based coding)
  - Blocking artifacts
    - Due to coarse quantization
    - Dependent on picture content
      - More visible in smooth areas than highly textured areas
    - Dependent on motion activity and prediction
    - Affected by viewing distance (visual acuity)
      - Blocking diminishes as viewing distance increases
    - Block size affects visibility of blocking effect
      - H.264/MPEG-4 AVC uses 4x4 DCT block

© CRC Press, Wu & Rao Eds, 2006

Impact of quantization on blocking distortion



© CRC Press, Wu & Rao Eds, 2006

Visibility of blocking artifacts in smooth and textured areas

© CRC Press, Wu & Rao Eds, 2006

12.4.2 In-Loop Filtering

- For de-blocking
  - Reduces visibility of block
  - Eliminates the need for post-filtering at decoder
- De-blocking filters used in
  - H.261, H.263 and H.264/MPEG-4 AVC
- Reduction of blocking in H.264/MPEG-4 AVC attributed (in part) to
  - 4x4 DCT and de-blocking filter

© CRC Press, Wu & Rao Eds, 2006

12.4.3 Perceptual-Based Scalable Coding Schemes

- Scalable coding
  - Multi-layered with one base and additional enhancement layers
  - Scalable modes
    - Temporal/spatial resolutions and SNR
  - MPEG-4 Fine Granularity Scalability (FGS)
    - DCT coefficients encoded in layers
      - Allows precise rate control
    - Applications in video streaming

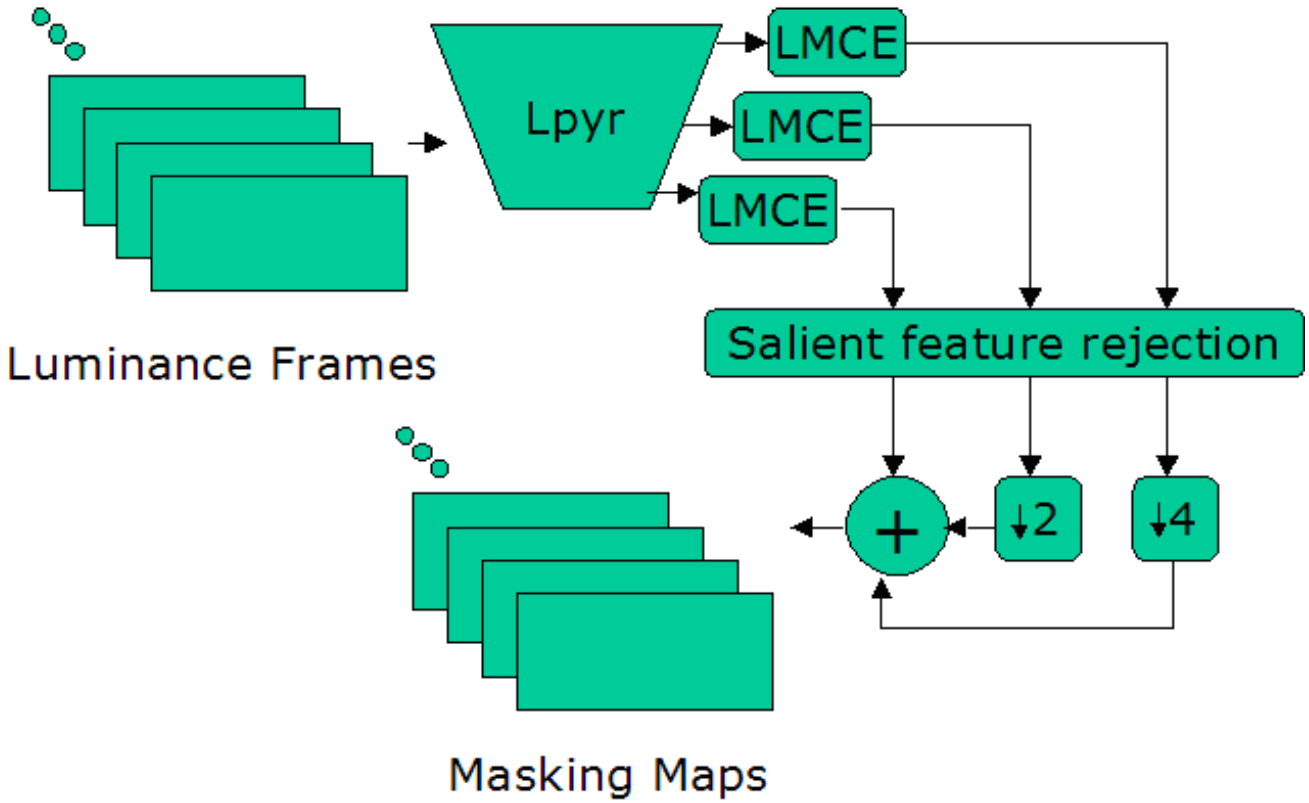
© CRC Press, Wu & Rao Eds, 2006

12.5 Salience/Maskability Pre-Processing

- Pre-filtering can be adapted to any coding system
- Saliency rejection
  - Histogram strategy
  - Coherence of local motion
    - Targets local regions with same/similar motion speed and direction
      - indicative of coherent moving objects

© CRC Press, Wu & Rao Eds, 2006

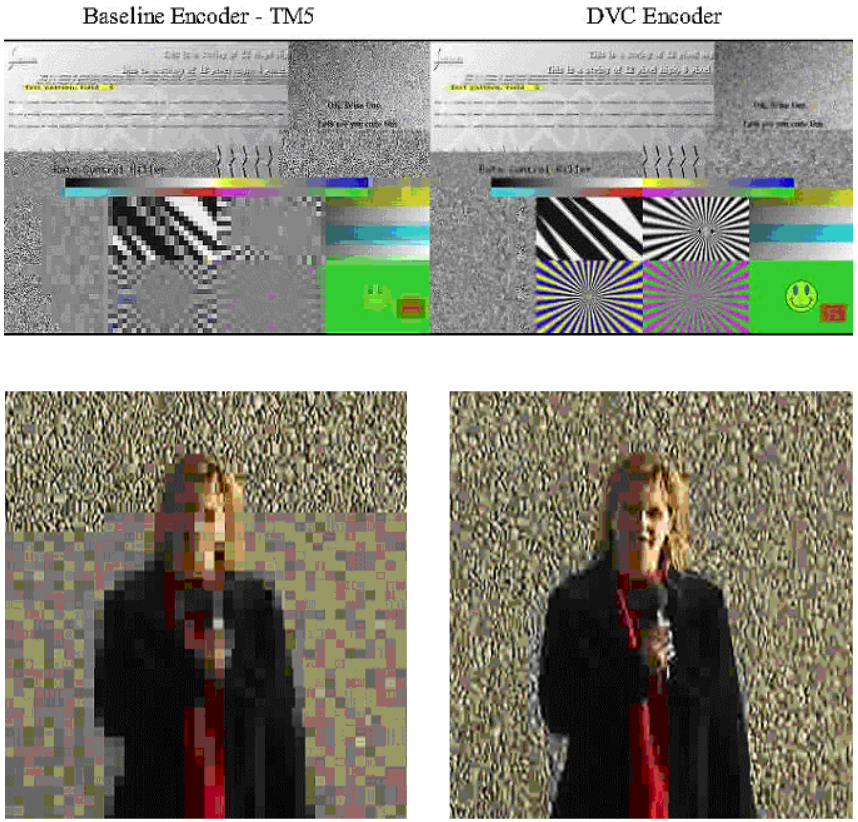
Pre-processor based maskability model with salient feature rejection



LME: local motion contrast energy detectors

© CRC Press, Wu & Rao Eds, 2006

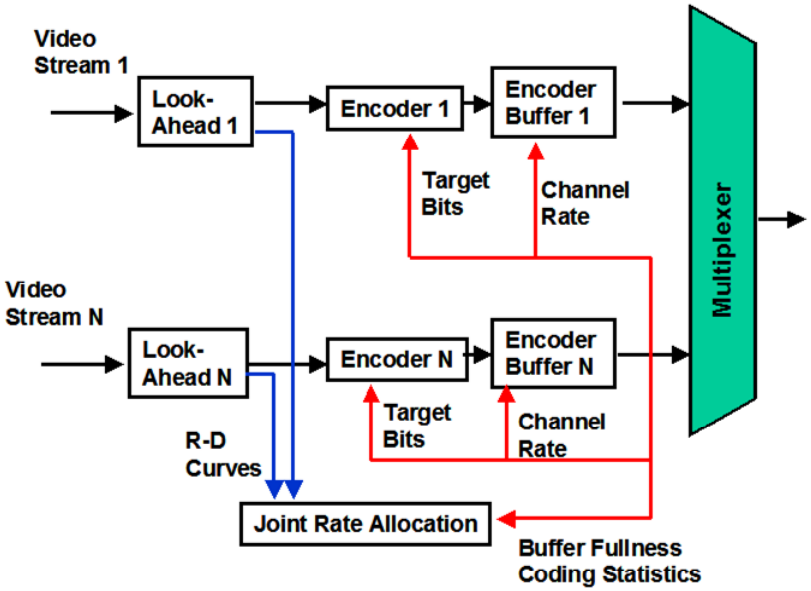
Illustration of pre-filtering. *Left*: MPEG-2 TM5 coded images. *Right*: DVC coded with pre-filtering



© CRC Press, Wu & Rao Eds, 2006

12.6 Application to Multi-Channel Encoding

- Multiplexed video streams
  - 8-12 programs in a 30 Mbps transport stream
  - Suited for Look-Ahead rate control scheme
    - *Join Rate Allocation* (JRA) equalizes distortions over all video streams



Multi-channel encoder

© CRC Press, Wu & Rao Eds, 2006

CP02	S. Cho and W. Pearlman. A Full-Featured, Error-Redilient, Scfalable Wavelet Video Codec Based on the Set Partitioning in Hierarchical Trees (SPIHT) Algorithm. *IEEE Transaction on Circuits and Systems for Video Technology*, 12(3):157-171, Mar. 2002.
CSE00	C. Christopoulos, A. Skodras, and T. Ebrahimi. The JPEG2000 Still Image Coding System: An Overview. *IEEE Trans. on Consumer Electronics*, 46(4):1103-1127, Nov. 2000
CZ97	T. Chiang and Y. Zhang. A New Rate Control Scheme using Quadratic Rate Distortion Model. *IEEE Transaction on Circuits and Systems for Video Technology*, 7(1):246-256, 1997
Gha03	M. Ghanbari. *Standard Codecs: Image Compression to Advanced Video Coding*. U K: IEE, 2003
HW03	C. Ho and J. Wu. Toward User Oriented Scalable Video by Using Foveated FGS Bitstreams. In *ICCE 2003 Conference Proceedings*, 46-47, June 2003
HWW03	C. Ho,J. Wu and S. Wang. A user Adaptive Perceptual Rate Control Scheme for FGS Videos. In *ICCE 2003 Conference Proceedings*, 42-43, June 2003

© CRC Press, Wu & Rao Eds, 2006

KHM01	Y.K. Kim, Z. He and S.K. Mitra. A Novel Linear Source Model and a Unified Rate Control Algorithm for H.263 / MPEG-2 / MPEG-4. In *Proc. of ICASSP*, 3:1777-1780, Salt Lake City, Utah, May 2001
Lee03	J.W. Lee. *Apparatus and Methods for Allocating Bits between Frames in a Coding Systems*. US Patent, 2003. Application 20020168007
LJL+03	P. List, A. Joch, J. Lainema, G. Bjøntegaard and M. Karczewicz. Adaptive Deblocking Filter. *IEEE Trans. on Circuits and Systems for Video Technology*, 13:614-619, July 2003
PE02	F. Pereira and T. Ebrahimi (eds). *The MPEG-4 Book*, Upper Saddle River, New Jersey: IMSC Press, 2002
PL03	H. Peterson and J. Lee. *Apparatus and Method for Optimizing Encoding and Performing Automated Steerable Image Compression in an Image Coding System using a Perceptual Metric*, US Patent, March 2003. 6,529,631 B1.
PM93	W.B. Pennebaker and J. Mitchell. *JPEG Still Image Data Compression Standard*. Norwell, MA: Van Nostrand Reinhold, 1993

© CRC Press, Wu & Rao Eds, 2006

Ric03	I. Richardson, *H.264 and MPEG-4 Video Compression*. Hoboken, NJ: Wiley, 2003
RJ02	M. Rabbani and R. Joshi. An Overview of the JPEG 2000 Still Image Compression Standard. *Signal Processing: Image Communincation*, 17:3-48, 2002
RY90	K.R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advances, Applications*. New York: Academic Press, 1990
SCE01	A. Skodras, C. Christopoulos and T. Ebrahimi. The JPEG 2000 Still Image Compression Standard. *IEEE Signal Processing Magazine*, 18:36-58, Sept. 2001
TM02	D. Taubman and M. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards and Practices*. Boston, MA: Kluwer, 2002

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

13. Perceptual Image Coding

D.M. Tan and H.R. Wu

Royal Melbourne Institute of Technology, Australia

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 13 Outline

13.1 Introduction

13.2 A Perceptual Distortion Metric Based Image Coder

13.3 Model Calibration

13.4 Performance Evaluation

13.5 Perceptual Lossless Coder

13.6 Summary

© CRC Press, Wu & Rao Eds, 2006

13.1 Introduction

- Redundancies in picture data
  - Statistical
    - Correlation between pixel data
  - Psychovisual
    - Information which the eye and mind cannot see
    - Removal does not impede image quality but
    - Leads to better compression performance

© CRC Press, Wu & Rao Eds, 2006

Human Visual System (HVS) Based Coders

- Common implementation through
  - Quantization (most common)
  - Coefficient weighting
    - Performed prior to quantization
  - Post- and/or pre-filtering
  - *Region Of Interest* (ROI) coding
    - Identify visually significant region and encode at higher quality

© CRC Press, Wu & Rao Eds, 2006

- Some HVS Based coders
  - Watson's *DCTune*
  - Safranek and Johnton's Subband Image Coder
  - Höntsch and Karam's *APIC*
  - Chou and Li's Perceptually Tuned Subband Image Coder
  - Taubman's EBCOT with visual masking
  - Zeng et al's Point-Wise Extended Visual Masking

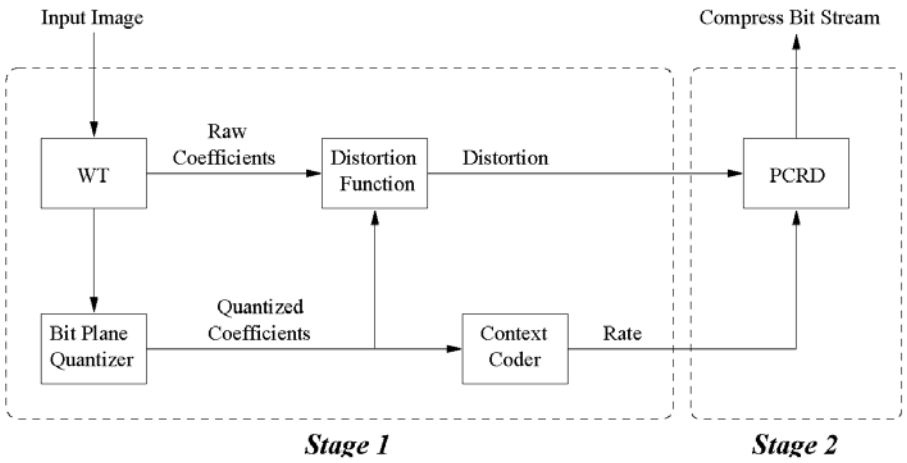
© CRC Press, Wu & Rao Eds, 2006

13.2 A Perceptual Distortion Metric Based Image Coder

- Utilize EBCOT (*Embedded Block Coding with Optimized Truncation*) architecture
  - Modular *Rate-Distortion* (R-D) component
    - Any distortion metric may be embedded in the R-D function
      - Perceptual distortion metric replaces MSE
  - Bitstream compliance maintained
    - No additional side information
    - No other alteration to bitstream

© CRC Press, Wu & Rao Eds, 2006

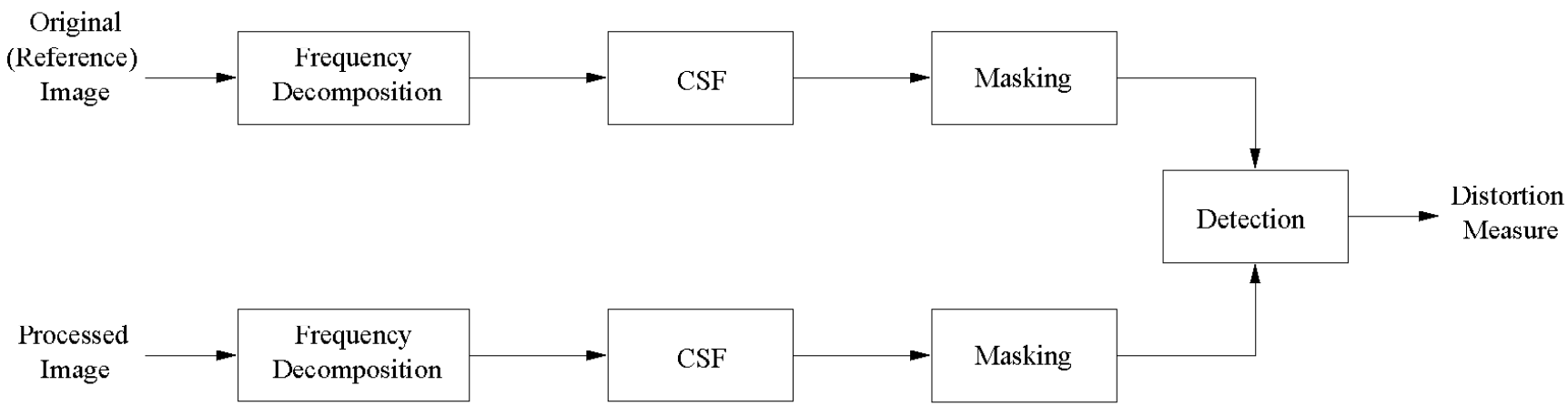
13.2.1 Coder Structure



- Basic EBCOT framework (*below*)
  - transform
    - 5 level transform
    - Bi-orthogonal 9/7 filters

© CRC Press, Wu & Rao Eds, 2006

13.2.2 Perceptual Image Distortion Metric (PIDM)



- Contrast Gain Control (CGC) (*above*)
  - Frequency transform (with 9/7 filter)
  - CSF weighting
  - Masking + detection

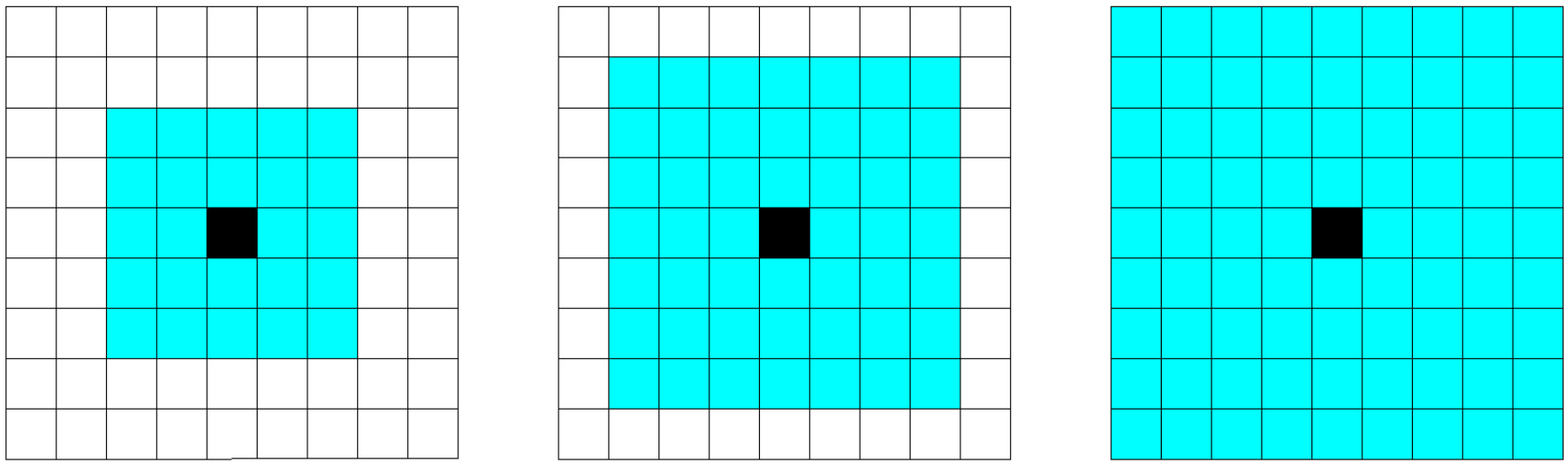
© CRC Press, Wu & Rao Eds, 2006

13.2.3 EBOCT Adaptation

- Perceptual distortion calculation
  - Localized to individual block
  - Each bitplane has *n* number of coding passes
    - Hence up to *n* number of quality layers
    - Perceived distortion calculated between layers
- Neighbourhood size for masking
  - Variable size neighbourhood
    - Dependent on resolution level
      - The higher the resolution, the larger the window

© CRC Press, Wu & Rao Eds, 2006

Masking neighbourhood size. Left to right: level 2, level 3 and level 4.



© CRC Press, Wu & Rao Eds, 2006

13.3 Model Calibration

- Calibrate vision model responses to subjective data
- Test material
  - 11 source images
  - Distorted derived from bitplane filtering of source images in the transform domain
    - Covers a wide range of quality levels

© CRC Press, Wu & Rao Eds, 2006

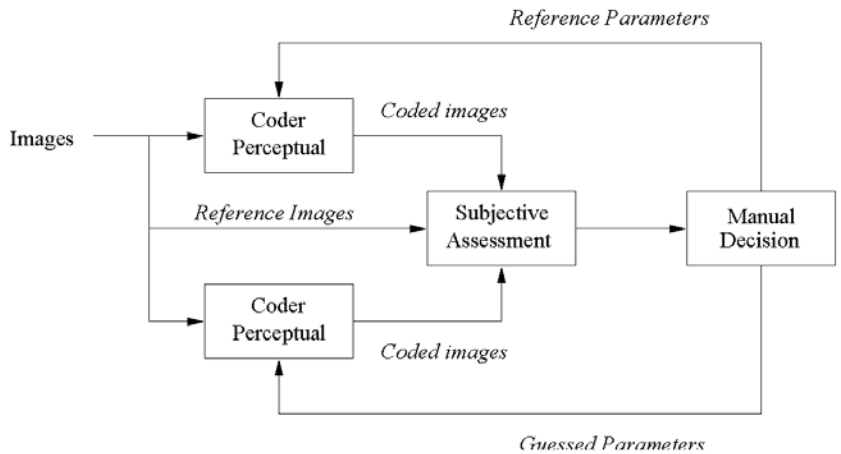
- Subjective test
  - DSCQS method from ITU-R bt.500
  - 3H viewing distance
  - 14 to 19 subjects
- Optimization of model paramters
  - In three stages
    - Parametric, algorithmic and manual optimization
  - Correlation measure
    - pearson’s correlation coefficient

© CRC Press, Wu & Rao Eds, 2006

- Parametric
  - Brute force search
  - Coarse approximation of model parameters
- Algorithmic
  - More detail search for optimum parameters
- Manual tuning at coder level
  - Parameters are tuned relative to performance of encoded images
    - Requires subjective input

© CRC Press, Wu & Rao Eds, 2006

Illustration of manual optimization



- Necessity of manual optimization
  - Subjective data provide overall quality of images
  - PIDM operates at coefficient/block level and require more refined subjective data

© CRC Press, Wu & Rao Eds, 2006

13.4 Performance Evaluation

- Comparison with benchmark coders
  - EBCOT with MSE
  - EBCOT-CVIS
    - Taubmans's spatial masking model
  - EBCOT-XMASK
    - Zeng et al's spatial masking model
- Assessment material
  - 8 images of which 5 have been used for model calibration

© CRC Press, Wu & Rao Eds, 2006

- Methods of evaluation
  - Objective
    - PSNR and *Universal Image Quality Index* (UIQI)
  - Subjective
    - Randomized comparative test
      - Dichotomous 2 images with 2 choices: *left* or *right*
      - Trichotomous 2 images with 3 choices: *left*, *right* or *either*
- Range of evaluation
  - Each image is coded to 4 bitrates
    - 0.0625 bpp, 0.125 bpp, 0.25 bpp and 0.5 bpp

© CRC Press, Wu & Rao Eds, 2006

Objective Results

- Overall
  - PSNR
    - Best performer EBCOT-MSE (no surprise)
    - EBCOT-PIDM 2nd and EBCOT-CVIS 3rd
    - Worst performer EBCOT-XMASK
  - UIQI
    - All metrics about the same except for EBCOT-XMASK
      - Worst performer below 0.25 bpp
      - Best performer at 0.5 bpp

© CRC Press, Wu & Rao Eds, 2006

Subjective Results

- Overall
  - PIBM vs CVIS (13 subjects)
    - Dichotomous: 81% of test images for PIBM
    - Trichotomous: 66% of test images for PIBM,
  - PIBM vs XMASK (7 subjects)
    - Dichotomous: 78% of test images for PIBM
    - Trichotomous: 78% of test images for PIBM,
  - PIBM vs MSE (15 subjects)
    - Dichotomous: 72% of test images for PIBM
    - Trichotomous: 57% of test images for PIBM

© CRC Press, Wu & Rao Eds, 2006

Barbara

512x512

© H.R. Wu

0.25bpp

MSE

PC

CVIS

XMASK

© H.R. Wu

0.25bpp

MSE

PC

CVIS

XMASK

Wu &

© H.R. Wu

0.25bpp

MSE

PC

CVIS

XMASK

Wu &

© H.R. Wu

0.25bpp

MSE

PC

CVIS

XMASK

Wu &

© H.R. Wu

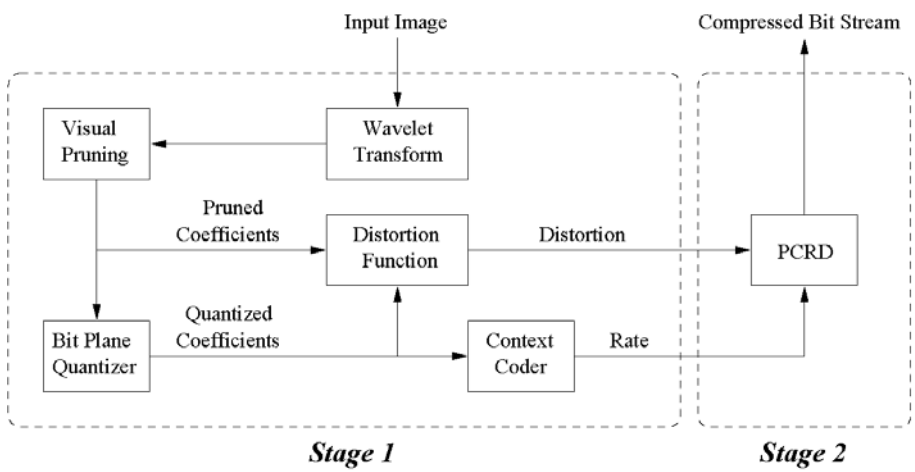
13.5 Perceptual Lossless Coder

- Encodes images at imperceptible level
  - Between coded and original
  - Constant quality coder at the *Just Not-Noticeable Level* (JNND)
- Visual Pruning Function (VPF)
  - Equivalent pre-filtering operator that removes visually redundant information
  - Operates in the spectral domain

© CRC Press, Wu & Rao Eds, 2006

13.5.1 Coding Structure (*below*)

- EBCOT with embedded VPF
  - Utilize similar vision model as the PIDM
  - Perceptual distortion of vision model calibrated to JNND threshold



© CRC Press, Wu & Rao Eds, 2006

13.5.2 Performance Analysis

- Compression performance
  - Compared with 2 benchmark coders
    - LOCO (LOw COmplexity coder)
      - JPEG-LS core
    - nLOCO (near-lossless LOCO)
      - With maximum pixel difference of 2 (d=2)
- Subjective Test for indifference
  - Trichotomous test
    - Randomized presentation of images
    - 2 images, 3 choices: *left*, *right* or *identical*

© CRC Press, Wu & Rao Eds, 2006

Compression performance of PLC and benchmark coders

Image	Dimensions	bpp		
		VLGI Coder	LOCO (LS)	LOCO (D2)
Actor	512x512	2.990	5.267	3.012
Barbara	512x512	1.986	4.733	2.552
Boat	512x512	1.866	4.250	2.149
Bridge	512x512	3.256	5.500	3.278
Butterfly	512x512	1.500	3.822	1.888
Couple	512x512	2.039	4.262	2.141
Crowd	512x512	1.910	3.914	2.065
Elaine	512x512	2.633	4.898	2.668
Goldhill	512x512	2.424	4.712	2.495
House	512x512	1.904	4.097	2.115
Lake	512x512	2.677	4.983	2.755
Lena	512x512	1.796	4.244	2.097
Mandrill	512x512	3.556	6.036	3.727
Peppers	512x512	2.152	4.489	2.270
F16 Jet	512x512	1.582	3.793	1.830
Sail	512x512	2.810	5.212	2.968
Tank	512x512	2.577	4.808	2.565
Tiffany	512x512	2.309	4.672	2.469
Tulip	512x512	1.841	4.176	2.127
Washington	512x512	2.047	4.129	2.112

- PLC is ahead
  - Up to 50% more compression than LOCO
  - Up to 10% more compression than nLOCO

© CRC Press, Wu & Rao Eds, 2006

Subjective results. *Left*: PLC vs Original. *Right*: PLC vs nLOCO

Image	Selection counts			No. of Subjects
	PLC	ORG	NEU	
Actor	10	5	2	17
Barbara	8	1	1	10
Boats	4	4	2	10
Bridge	8	6	2	16
Butterfly	3	4	6	13
Couple	5	4	2	11
Crowd	2	1	2	5
Elaine	4	2	1	7
Goldhill	4	0	5	9
House	5	3	2	10
Lake	1	3	4	8
Lena	7	6	2	15
Mandrill	6	1	9	16
Peppers	4	4	3	11
F16 Jet	5	7	3	15
Sail	4	1	7	12
Tank	4	4	2	10
Tiffany	2	4	2	8
Tulip	2	1	2	5
Washington	4	1	2	7
Total	92	62	61	215

Image	Selection counts			No. of Subjects
	PLC	nLC	NEU	
Actor	1	5	3	9
Barbara	8	3	1	13
Boats	5	4	3	12
Bridge	6	5	2	13
Butterfly	5	5	2	12
Couple	8	3	2	13
Crowd	4	4	8	16
Elaine	5	8	2	15
Goldhill	5	6	2	13
House	3	3	0	6
Lake	6	3	3	12
Lena	2	7	3	12
Mandrill	1	7	2	10
Peppers	2	4	3	9
F16 Jet	8	2	0	10
Sail	2	3	5	10
Tank	2	4	3	9
Tiffany	3	4	2	9
Tulip	8	2	3	13
Washington	1	8	6	15
Total	85	90	55	230

© CRC Press, Wu & Rao Eds, 2006

- Raw subjective results
  - Apparently PLC is better than original
    - Neutral indicative of differences
    - Preferred selection favours PLC
    - Perhaps PLC has an enhancement effect
  - PLC and nLOCO are similar
    - Neutral selection is not indicative similarity, but
    - Preferred selection is split almost evently
      - Raw preference is 5.5% apart

© CRC Press, Wu & Rao Eds, 2006

Original

PLC

© CRC Press, Wu & Rao Eds, 2006

PLC

nLOCO

© CRC Press, Wu & Rao Eds, 2006

13.6 Summary

- Perceptual Image coding
  - Enables further compression gains over pure statistical coders
  - Allows for graceful degradation of perceived quality relative to bitrate constrains
  - Enables constant (perceived) quality coding
    - e.g. constant quality at the JNND level (PLC)

© CRC Press, Wu & Rao Eds, 2006

- PIDM Coding structure
  - Start-of-the-art EBCOT architecture
    - Hierarchical progressive bitplane coder
    - Scalable in rate and resolution
  - PIDM used for R-D function for rate control
    - Measures perceived distortion resulting from bitplane encoding
  - Fully bitstream compliant with EBCOT/JPEG2000
    - No additional side information needed

© CRC Press, Wu & Rao Eds, 2006

ABMD92	M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies. Image Coding Using Wavelet Transform. *IEEE Trans. on Image Processing*, 1(2):205-220, April 1992
AP92	A.J. Ahumada and H.A. Peterson. Luminance-model-based DCT quantization for color image compression. In *B.E. Rogowitz, Ed., Human Vision, Visual Processing and Digital Display III, Proc. SPIE*, 1666:365-374, 1992
BB99	C.H. Brase and C.P Brase. *Understandable Statistics: Concepts and Methods*. Boston: Houghton Mifflin Company, 6th ed., 1999
BB01	V. Baroncini and M. Buxton. *Preliminary Results of Assessment of Responses to Digital Cinema Call for Proposals*, July 2001. ISO/IEC JTC 1/SC 29/WG 11 N4241
BL97	M.H. Brill and J. Lubin. Methods and Apparatus for Assessing the Visibility of Differences between Two Image Sequences, Dec. 1997. US Patent No. 5,694,491
Bor57	E.G. Boring. *A History of Experimental psychology*. New York: Appleton Century Crofts, Inc., 1957

© CRC Press, Wu & Rao Eds, 2006

CL95	C.H. Chou and Y.C. Li. A perceptually tuned subband image coder based on the measure of Just-Noticeable-Distortion Profile. *IEEE Trans. Circuits Systems for Video Technology*, 5(6):467-476, Dec. 1995
Cla85	T.J. Clarke. *Transform Coding of Images*. New York: Academic Press, 1985
Cor90	T.N. Cornsweet. *Visual Perception*. Orlando, FL: Academic Press, 1990
CP84	W.H. Chen and W.K. Pratt. Scene Adaptive Coder. *IEEE Trans. on Communications*, 32(3):225-232, Mar. 1984
CZ01	E.K.P. Chong and S.H. Żak. *An Introduction to Optimization*, Hoboken, NJ: Wiley, 2nd Ed., 2001
DAT82	R.L. De Valois, D.G. Albrecht and L.G. Thorell. Spatial Frequency Selectivity of Cells in Macaque Visual Cortex. *Vision Research*, 22(5):545-559, 1982
DKMM99	Y. Deng, C. Kenney, M.S. Moore and B.S. Manjunath. Peer Group Filtering and Perceptual Color Image Quantization. In *Proc. of IEEE International Symposium on Circuits and Systems*, 4:21-24, 1999

© CRC Press, Wu & Rao Eds, 2006

DS89	J.E. Dennis Jr. and R.B. Schnabel. A View of Unconstrained Optimization. In G.L. Nembause, A.H.G. Rinnooy Kan and M. J. Todd, Eds., *Handbooks in Optimization Research and Management Science: Optimization*, 1:1-72. Amsterdam: Elsevier Science Publishers, 1989
DSKK58	R.L. De Valois, C.J. Smith, A.J. Karoly and S.T. Kitai. Electrical Responses of Primate Visual System. I. Different Layers of macaque Lateral Geniculate Nucleus. *Journal of Comparative Physiology and Psychology*, 51:622-668, 1958
DZLL00	S. Daly, W. Zeng, J. Li and S. Lei. Visual Masking in Wavelet Compression for JPEG2000. In *IS&T/SPIE Conf. on Image and Video Communication and Processing*, 3974:66-80, Jan. 2000
FB94	J.M. Foley and G.M. Boynton. A New Model of Human Luminance Pattern Vision Mechanisms: Analysis of the Effects of Pattern Orientation Spatial Phase and Temporal Frequency. In T.A. Lawton, Ed., *Computational Vision Based on Neurobiology, Proceedings of SPIE*, 2054:32-42, 1994

© CRC Press, Wu & Rao Eds, 2006

FCH97	W.C. Fong, S.C. Chan and K.L. Ho. Determination of Frequency Sensitivities for Perceptual Subband Image Coding. *IEE Electronics Letters*, 33(7):581-582, Mar. 1997
Fle81	R. Fletcher. *Practical Methods of Optimization*. Hoboken, NJ: Wiley, 2nd ed., 1981
Fol94	J.M. Foley. Human luminance pattern-vision mechanisms: Masking experiments require a new model. *J. Opt. Soc. Am*, 11:1710-1719, 1994
Fri95	L. Friedenberg. *Psychological Testing: Design, Analysis and Use*. New York: Allyn & Bacon, A Simon & Schuster Company, 1995
GMW81	P.E. Gill, W. Murray and M.H. Wright. *Practical Optimization*. Orlando, FL: Academic Press, 1981
Hem97	S.S. Hemami. Visual Sensitivity Consideration for Subband Coding. In *Proceedings of the 31st Asilomar Conf. on Signals, Systems and Computers*, 1:622-656, 1997

© CRC Press, Wu & Rao Eds, 2006

HK97	I. Höntsch and L.J. karam. APIC: Adaptive Perceptual Image Coding Based on Subband Decomposition with Locally Adaptive Perceptual Weighting, In *Proc. of IEEE ICIP*, 1:37-40, 1997
HK00	I. Höntsch and L.J. Karam. Locally Adaptive Perceptual Image Coding. *IEEE Trans. on Image Processing*, 9(9):1472-1483, Sept. 2000
HK02	I. Höntsch and L.J. Karam, Adaptive Image Coding with Perceptual Distortion Control. *IEEE Trans on Image Processing*, 11(3):213-222, Mar. 2002
HKS97	I. Höntsch, L.J. Karam and R.J. Safranek. A Perceptually Tuned Embedded Zerotree Image Coder. In *Proc. of IEEE ICIP*, 1:41-44, 1997
HN97	M.J. Horowitz and D.L. Neuhoff. Image Coding by Matching Pursuit and Perceptual Pruning. In *Proc. of IEEE ICIP*, 3:654-657, 1997
ISO00a	ISO/IEC JTC 1/SC 29. *Information Technology – JPEG 2000 image coding system – Part 2: Extension*, 2000. ISO/IEC 15444-1:2000
ISO00b	ISO/IEC JTC1/SC9/WG1. JPEG2000 Verification Model 8.0 Software, 2000. ISO/IEC JTC1/SC9/WG1 B18191

© CRC Press, Wu & Rao Eds, 2006

ITU94a	ITU-R Recommendation BT.814-1. Specifications and Alignment Procedures for Setting of Brightness and Contrast of Displays. Technical report, ITU-R, July 1994
ITU94b	ITU-R Recommendation BT.815-1. Specification of a Signal for Measurement of the Contrast Ratio of Displays. Technical Report, ITU-R, July 1994
ITU00	ITU-R Recommendation BT.500-10. Methodology for the Subjective Assessment of the Quality of Television Pictures. Technical report, ITU-R, Mar. 2000
A.K. Jain	*Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989
JB87	R.A. Johnson and G.K. Bhattacharyya. *Statistics: Principles and Methods*. Hoboken, NJ: Wiley, 2nd ed., 1987
JN84	N.S. Jayant and P. Noll. *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1984

© CRC Press, Wu & Rao Eds, 2006

KS99	I. Kopilovic and T. Sziranyi. Non-Linear Scale-Selection for Image Compression Improvement Obtained by Perceptual Distortion Criteria. In *Proc. of IEEE International Conference on Image Analysis and Processing*, 197-202, 1999
KS01	R.M. Kaplan and D.P. Saccuzzo. *Psychological Testing: Principles, Applications and Issues*. Belmont, CA: Thomson Learning Inc., 5th ed., 2001
LF80	G.E. Legge and J.M. Foley. Contrast Masking in Human Vision. *Journal of the Optical Society of America*, 70(12):1458-1471, Dec. 1980
Lim78	J. Limb. On the Design of Quantiser for DPCM Coders – A Functional Relationship Between Visibility, Probability and Masking. *IEEE Trans. on Communications*, 26:573-578, 1987
LJ97	J. Li and J.S. Jin. Structure-Related Perceptual Weighting: A Way to Improve Embedded Zerotree Wavelet Image Coding. *IEE Electronics Letters*, 33(15):1305-1306, July 1997
LK98	Y.K. Lai and C.C. Kuo. Perceptual Image Compression with Wavelet Transform. In *Proc. Of IEEE Int. Symp. On Circuits and Systems*, 4:29-32, 1998

© CRC Press, Wu & Rao Eds, 2006

LTK99	T.H. Lan, A.H. Tewfik and C.H. Kuo. Sigma Filtered Perceptual Image Coding at Low Bit Rates. In *Proc. of IEEE ICIP*, 2:371-375, 1999
LW98	C.M. Liu and C.N. Wang. On the Perceptual Interband Correlation for Octave Subband Coding. In *Proc. of IEEE ICIP*, 4:166-169, 1998
LZFW00	K.T. Lo, X.D. Zhang, J. Feng and D.S. Wang. Universal Perceptual Weighted Zerotree Coding for Image and Video Compression. *IEE Proc. on Vision, Image and Signal Processing*, 147(3):261-265, June 2000
MH98	M. Miloslavski and Y.S. Ho. Zerotree Wavelet Image Coding Base on the Human Visual System Model. In *Proc. of IEEE Asia-Pacific Conference on Circuits and Systems*, 57-60, 1998
Mic64	A.A. Michelson. *Experimental Determination of the Velocity of Light*. Minneapolis, MN: Lund Press, 1964
ML95	A. Mazzarri and R. Leonardi. Perceptual Embedded Image Coding using Wavelet Transforms. In *Proc. of IEEE ICIP*, 1:586-589, 1995
Nad00	M.J. Nadenau. *Integration of Human Color Vision Models into Hight Quality Image Compression*. Ph.D. thesis, EPFL, Lausanne, 2000

© CRC Press, Wu & Rao Eds, 2006

NL89	K.N. Ngan, K.S. Leong and H. Singh. Adaptive Cosine Transform Coding of Images in Perceptual Domain. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 37(11):1743-1750, Nov 1989
NP77	A. Netravali and B. Prasada. Adaptive Quantisation of Picture Signals Using Spatial Masking. *Proceedings of the IEEE*, 65(4): 536-548, April 1977
NR77	A. Netravali and C. Rubinstein. Quantisation of Colour Signals. *Proceedings of the IEEE*, 65(8):1177-1187, Aug. 1977
OS95	T.P. O'Rourke and R.L. Stevenson. Human Visual System Based Wavelet Decomposition for Image Compression. *Journal of Visual Communication and Image Representation*, 6(2):109-121, June 1995
PAW93	H.A. Peterson, A.J. Ahumada and A.B. Watson. An Improved Detection Model for DCT Coefficient Quantization. *SPIE Proceedings*, 1913:191-201, 1993
Pel90	E. Peli. Contrast in Complex Images. *Journal of the Optical Society of America A*, 7(10):2032-2040, Oct. 1990

© CRC Press, Wu & Rao Eds, 2006

PH71	J.M. Parkinson and D. Hutchinson. An Investigation into the Efficiency of Variants on the Simplex Method. In F.A. Lootsma, Ed., *Numerical Methods for Non-Linear Optimization*. Orlando, FL: Academic Press, 1971
PMH96	T.N. Pappas, T.A. Michel and R.O. Hinds. Supra-Threshold Perceptual Image Coding. In *Proc. of IEEE ICIP*, 1:237-240, 1996
RH99	M.G. Ramos and S.S. Hemami. Activity Selective SPIHT Coding. In *Proc. of SPIE Conf. on Visual Communications and Image Processing*, 3653:315-326, 1999
RH00	M.G. Ramos and S.S. Hemami. Perceptual Quantization for Wavelet-Based Image Coding. In *Proc. of IEEE ICIP*, 1:645-648, 2000
RLCW00	A.M. Rohaly, J. Libert, P. Corriveau and A. Webster. Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment. Technical report, VQEG, Mar. 2000
Roh00	A.M. Rohaly *et al*. Video Quality Experts Group: Current Results and Future Directions. In *Proc. of SPIE 2000 Conference on Visual Communications and Image Processing*, 4067:742-753, June 2000

© CRC Press, Wu & Rao Eds, 2006

RW96	R. Rozenholtz and A.B. Watson. Perceptual Adaptive JPEG Coding. In *Proc. of IEEE ICIP*, 1:901-904, 1996
SFAH92	E.P. Simoncelli, W.T. Freeman, E.H. Adelson and D.J. Heeger. Shiftable Multiscale Transform. *IEEE Trans. on Information Theory*, 38(2):587-607, Mar. 1992
SGE02	D. Santa-Cruz, R. Grosbois and T. Ebrahimi. JPEG 2000 Performance Evaluation and Assessment. *Signal Processing: Image Communication*, 17(1):113-130, Jan. 2002
SHH62	W. Spendley, G.R. Hext and F.R. Himsworth. Sequential Applications of Simplex Designs in Optimisation and Evolutionary Operation. *Technometrics*, 4(4):441-461, Nov. 1962
SJ89	R.J. Safranek and J.D. Johnston. A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression. *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing*, 1945-1948, 1989

© CRC Press, Wu & Rao Eds, 2006

Tau00	D. Taubman. High Performance Scalable Image Compression with EBCOT. *IEEE Trans. on Image Processing*, 7(9):1158-1170, July 2000
TH94a	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. SPIE*, 2179:127-141, 1994
TH94b	P.C. Teo and D.J. Heeger. Perceptual image distortion. In *Proc. ICIP*, 2:982-986, 1994
TNH97	T.D. Tran, T.Q. Nguyen and Y.H. Hu. A Perceptually-Tuned Block-Transform-Based Progressive Transmission Image Coder. In *Proc. of 31st Asilomar Conf. on Signal, Systems and Computers*, 2:1000-1004, 1997
TPN96	S.H. Tan, K.K. Panf and K.N. Ngan. Classified Perceptual Coding with Adaptive Quantization. *IEEE Trans. on Circuits and Systems for Video Technology*, 6(4):375-388, Aug. 1996
Tur00	Turbolinux. Enfuzion Manual, Mar. 2000, (http://www.turbolinux.com)
usc	USC-SIPI Image Database. (http://sipi.usc.edu/services/database/)

© CRC Press, Wu & Rao Eds, 2006

van96	C.J. van den Branden Lambrecht. *Perceptual Models and Architectures for Video Coding Applications*. Ph.D thesis, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, 1996
Wal91	G.K. Wallace. The JPEG Still Picture Compression Standard. *Communication of ACM*, 34:30-44, April 1991
Wan95	B.A. Wandell. *Foundations of Vision*. Sunderland, MA: Sinauer Associates, Inc., 1995
Wat93	A.B. Watson. DCTune: A technique for visual optimization of DCT quantization matrices for individual images. *Society for Information Display Digest of Technical Papers XXIV*, 946-949, 1993
Wat94	A.B. Watson. Perceptual Optimization of DCT Colour Quantization Matrices. In *Proc. of IEEE ICIP*, 100-104, 1994
Wat00	A.B. Watson. Visual Detection of Spatial Contrast Patterns: Evaluation of Five Simple Models. *Optics Express*, 6(1):12-33, 2000
WB01	Z. Wang and A.C. Bovik. Embedded foveation image coding. *IEEE Trans. Image Processing*, 10(10):1397-1410, Oct. 2001

© CRC Press, Wu & Rao Eds, 2006

WB02	Z. Wang and A.C. Bovik, A universal image quality index. *IEEE Signal Proc. Letters*, 9(3):81-84, Mar. 2002
WHMM99	A. Watson, J. Hu, J.F. McGowen III and J.B. Mulligan. Design and Performance of a Digital Video Quality Metric. In *Proc. of the SPIE Conference on Human Vision and Electronic Imaging IV*, 3644:168-174, 1999
Win98	S. Winkler. A perceptual distortion metric for digital color images. In *Proc. ICIP*, 3:399-403, Chicago, IL, Oct. 4-7, 1998
Win99	S. Winkler. A Perceptual Distortion Metric for Digital Color Video. In *Proc. of the SPIE Conference on Human Vision and Electronic Imaging IV*, 3644:175-184, 1999
Win00	S. Winkler. *Vision Models and Quality Metrics for Image Processing Applications*. Ph.D. thesis, EPFL, Lausanne, Switzerland, Dec. 2000
WLB95	S.J.P. Westen, R.L. Lagendijk and J. Biemond. Perceptual Image Quality Based on a Multiple Channel HVS Model. In *Proc. of IEEE ICASSP*, 4:2351-2354, 1995

© CRC Press, Wu & Rao Eds, 2006

WS97	A.B. Watson and J.A. Solomon. A model of visual contrast gain control and pattern masking. *J. Opt. Soc. Am. A*, 14:2379-2391, 1997
WSS96	M. Weinberger, G. Seroussi and G. Sapiro. LOCO-I: A Low Complexity, Context Based, Lossless Image Compression Algorithm. In *Proc. of IEEE Data Compression Conference*, 140-149, 1996
WWT+04	C. White, H.R. Wu, D. Tan, R.L. Martin and D. Wu. *Experimental Design for Digital Image Quality Assessment*. Zeltand, NSW, Australia: Epsilon Publishing, 2004
YW98	M. Yuen and H.R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing*, 70:247-278, 1998
YWWC02	Z. Yu, H.R. Wu, S. Winkler and T. Chen. Objective assessment of blocking artifacts for digital video with a vision model. *Proc. IEEE*, 90(1):154-169, jan. 2002.
YZF99	K.H. Yang, W. Zhu and A.F. Faryar. Perceptual Quantization for Predictive Coding in Images. In *Proc. of IEEE ICIP*, 2:381-385, 1999

© CRC Press, Wu & Rao Eds, 2006

ZDL00	W. Zeng, S. Daly and S. Lei. Point-Wise Extended Visual Masking for JPEG-2000 Image Compression. In *Proc. of IEEE ICIP*, 1:657-660, 2000
Zen99	Y. Zeng. Perceptual Segmentation Algorithm and its Application to Image Coding. In *Proc. of IEEE ICIP*, 2:820-824, 1999
ZZQ97	W. Zheng, H. Zhu and Z. Quan. Lattice Vector Quantization of Wavelet Coefficients with Perceptual Visibility Model and Hybrid Entropy Coding. In *IEEE International Symposium on Circuits and Systems*, 1117-1120, 1997

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

14. Foveated Image and Video Coding

Z. Wang[†] and A. Bovik[‡]

† University of Texas at Arlington, USA

‡ University of Texas at Austin, USA

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 14 Outline

14.1 Foveated Human Vision and Foveated Image Processing

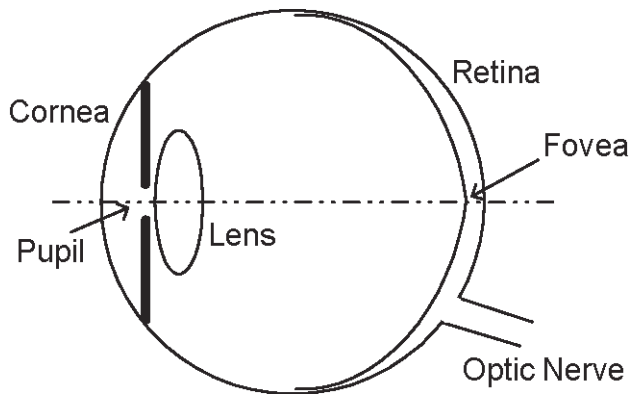
14.2 Foveation Methods

14.3 Scalable Foveated Image and Video Coding

14.4 Discussions

© CRC Press, Wu & Rao Eds, 2006

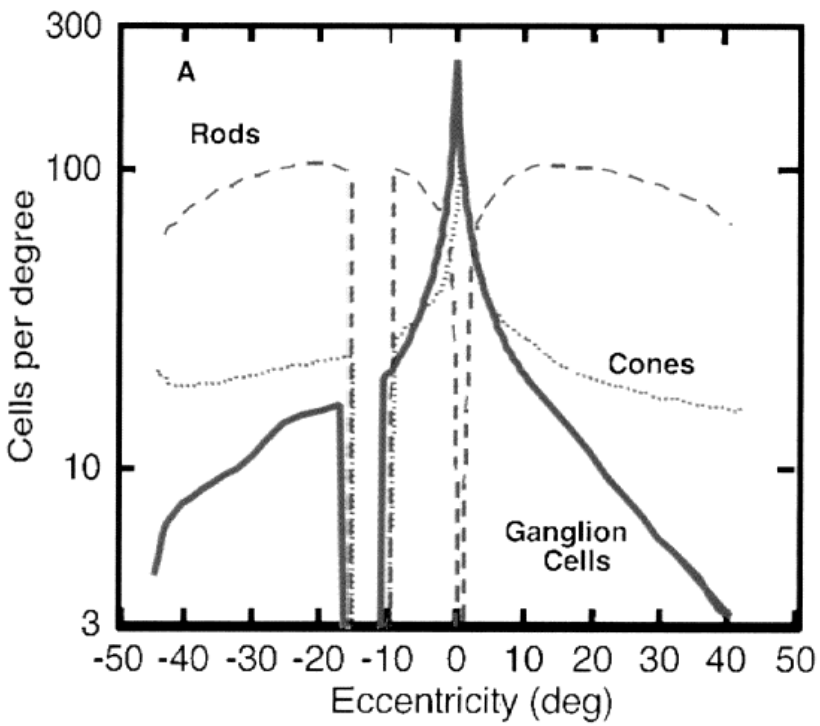
14.1 Foveated Human Vision and Foveated Image Processing



- Anatomy of the human eye
  - Photoreceptor: *Cones* and *Rods*
    - Populates the retina (focal point fovea)
    - Non-uniform distribution of photoreceptors in retina

© CRC Press, Wu & Rao Eds, 2006

Distribution of rods, cones and ganglion cells relative to eccentricity (central point of fovea)



- Cone concentrates falls off sharply from fovea
- No rods in fovea
- No photoreceptor in the area occupied by the optic nerve (blind spot)

© CRC Press, Wu & Rao Eds, 2006

- Basic operation of the eye
  - Light is captured by photoreceptors
    - Cones responsible for photopic vision
    - Rods handle scotopic vision
  - Ganglion cells integrates visual stimuli captured by photoreceptors
    - Some preliminary processing of visual signals
  - Integrated signal transmitted to cortex via visual pathways

© CRC Press, Wu & Rao Eds, 2006

- Foveated image processing and coding
  - Focused on the fovea
    - Point of fixation
  - Major redundancies in peripheral vision
    - Improve compression

© CRC Press, Wu & Rao Eds, 2006

Goldhill Image. a) original. b) foveated

© CRC Press, Wu & Rao Eds, 2006

14.2 Foveation Methods

- Geometric methods
  - Three approaches
    - Direct application of to uniform resolution image
      - Underlying image is mapped onto new coordinate system (1 to 1 mapping)
    - Superpixel method
      - Pixels are grouped and mapped into superpixel block (many to 1 mapping)
    - Foveated retinal geometry
      - Image mapped to foveal sensor grid

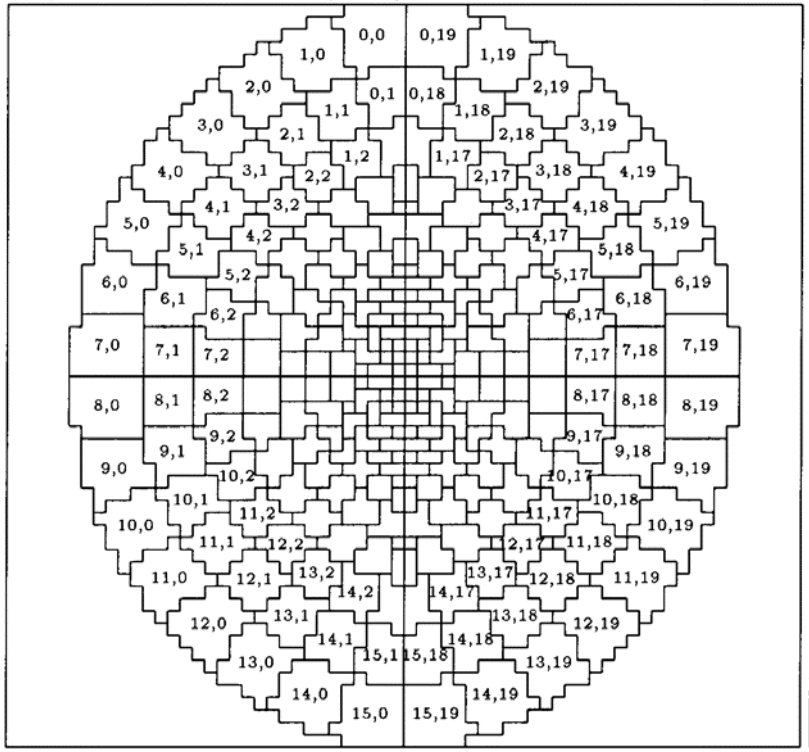
© CRC Press, Wu & Rao Eds, 2006

Foveation coordinate transform. a) original images. b) transformed image

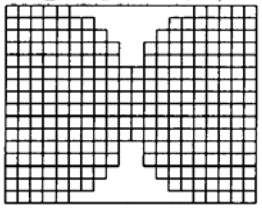
© CRC Press, Wu & Rao Eds, 2006

Longmap superpixel representation

Logmap lookup tables (S, R) for 60 by 64 TV image.



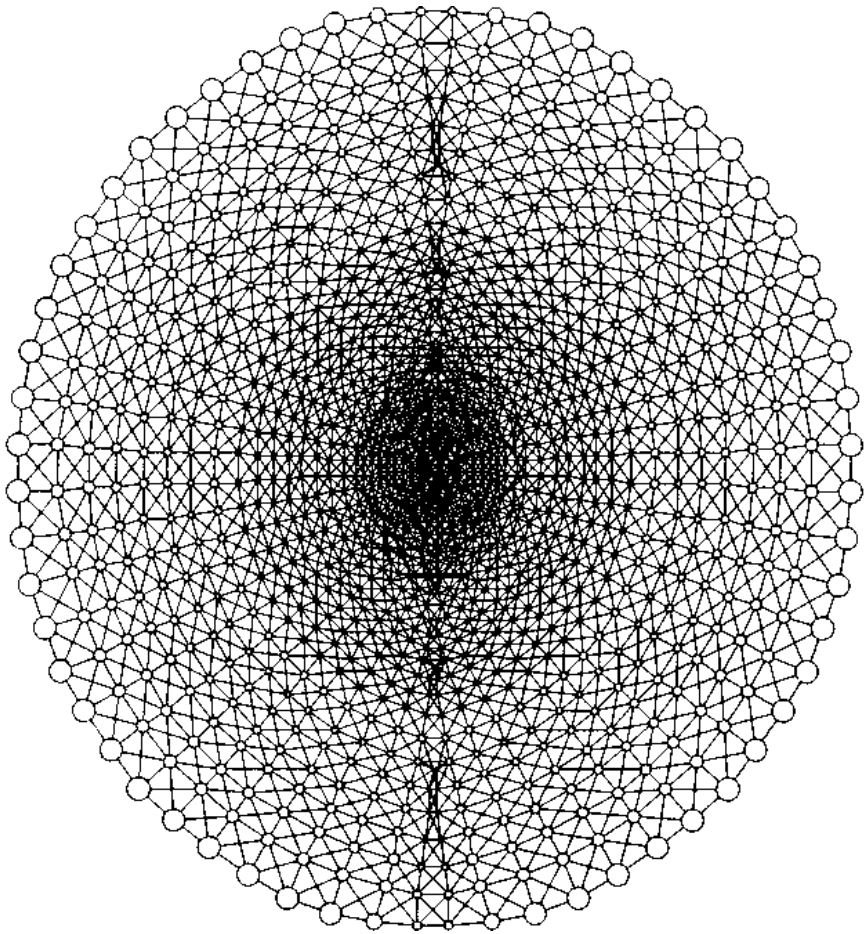
Logmap array (16 by 20)



(15,19)

© CRC Press, Wu & Rao Eds, 2006

Foveated sensor distribution



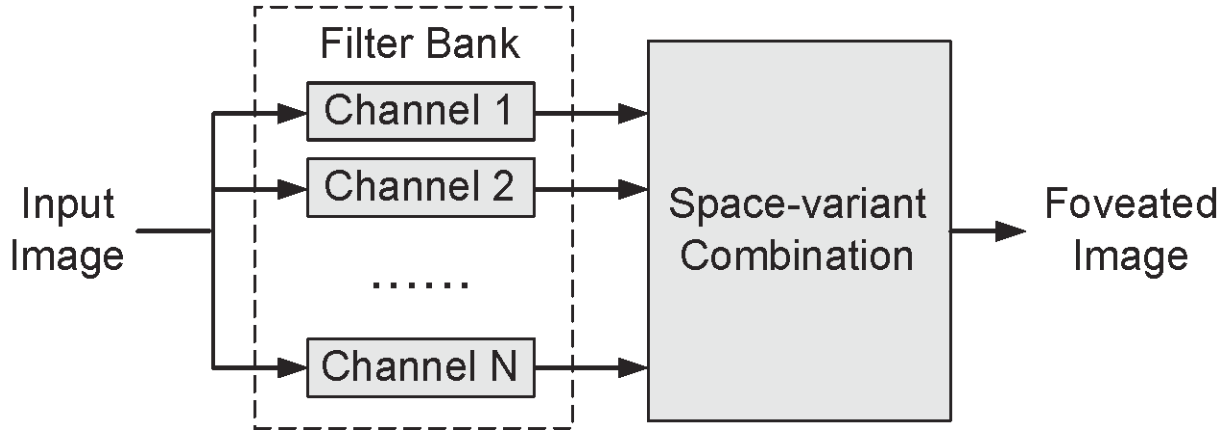
© CRC Press, Wu & Rao Eds, 2006

14.2.2 Filtering Based Methods

- Foveal processing
  - Shift-invariant low-pass filtering
    - Cut-off frequency determined by retinal sampling density
- Foveation implementation
  - Multi-channel Filter banks
    - Decomposition into different spectral bands
    - Foveal image obtained with space-invariant combination of spectral images
      - dependent on retinal sampling density

© CRC Press, Wu & Rao Eds, 2006

Filter bank foveation method

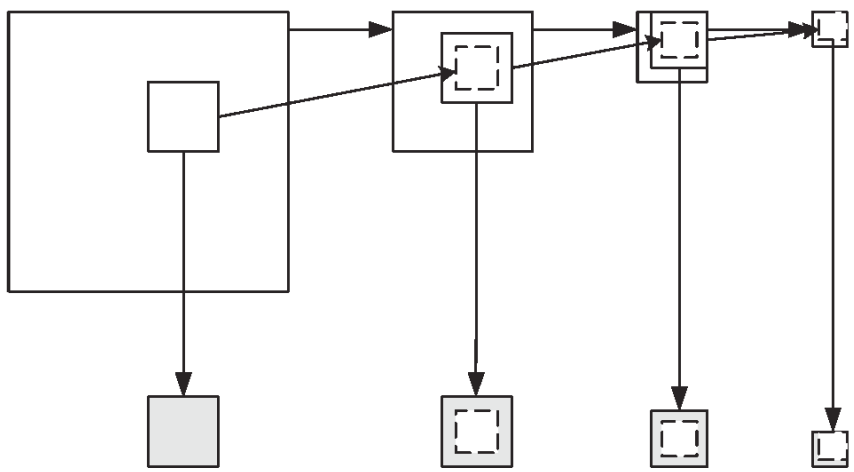


- Filter design considerations
  - Low-pass, band-pass or both
  - Filter type: IIR or FIR
  - Characteristics
    - Band size, ripple size

© CRC Press, Wu & Rao Eds, 2006

14.3.2 Multiresolution Methods

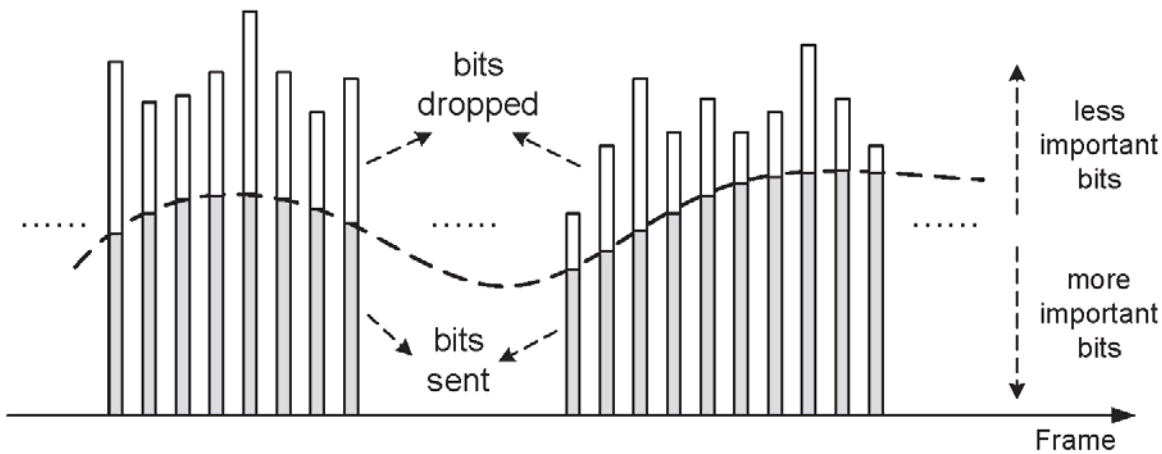
- Image projected into different resolutions
  - Combines both geometric and filtering-based methods



© CRC Press, Wu & Rao Eds, 2006

14.3 Scalable Foveated Image and Video Coding

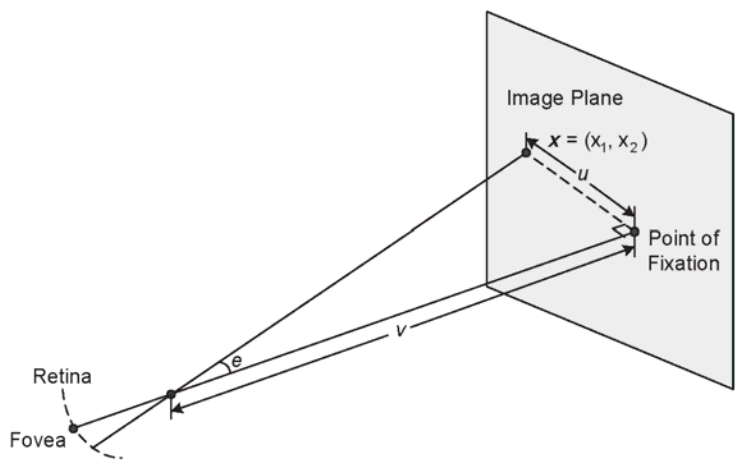
- Scalable video coding
  - Compressed stream layered to different quality levels
  - Quality tailored to user requirements



© CRC Press, Wu & Rao Eds, 2006

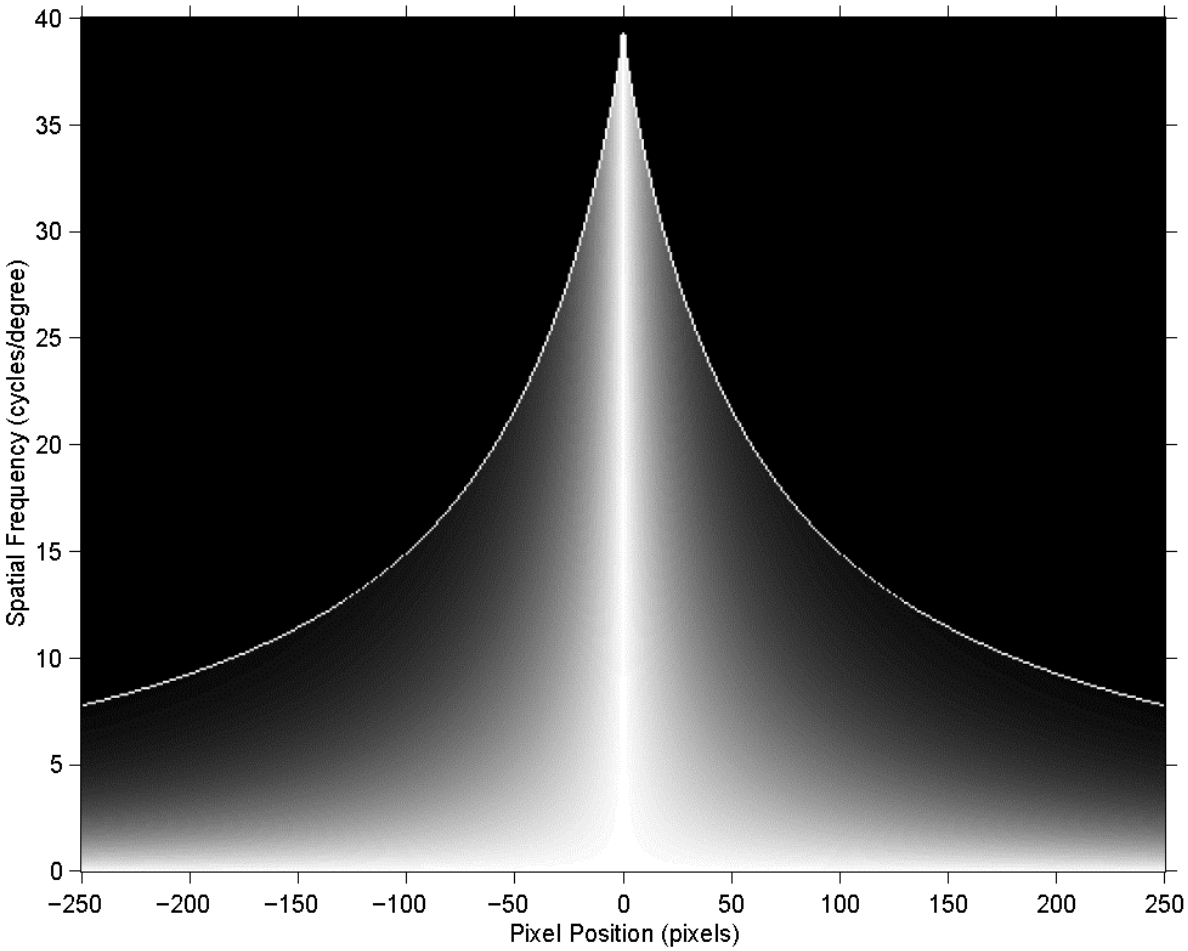
14.3.1 Foveated Perceptual Weighting Model

- Contrast sensitivity measured as a function of retinal eccentricity
  - Dependent of viewing distance (*below*)
    - Assuming constant fixation point (*u*), as viewing distance increases (*v*), angle of eccentricity (*e*) decreases



© CRC Press, Wu & Rao Eds, 2006

Normalized contrast sensitivity

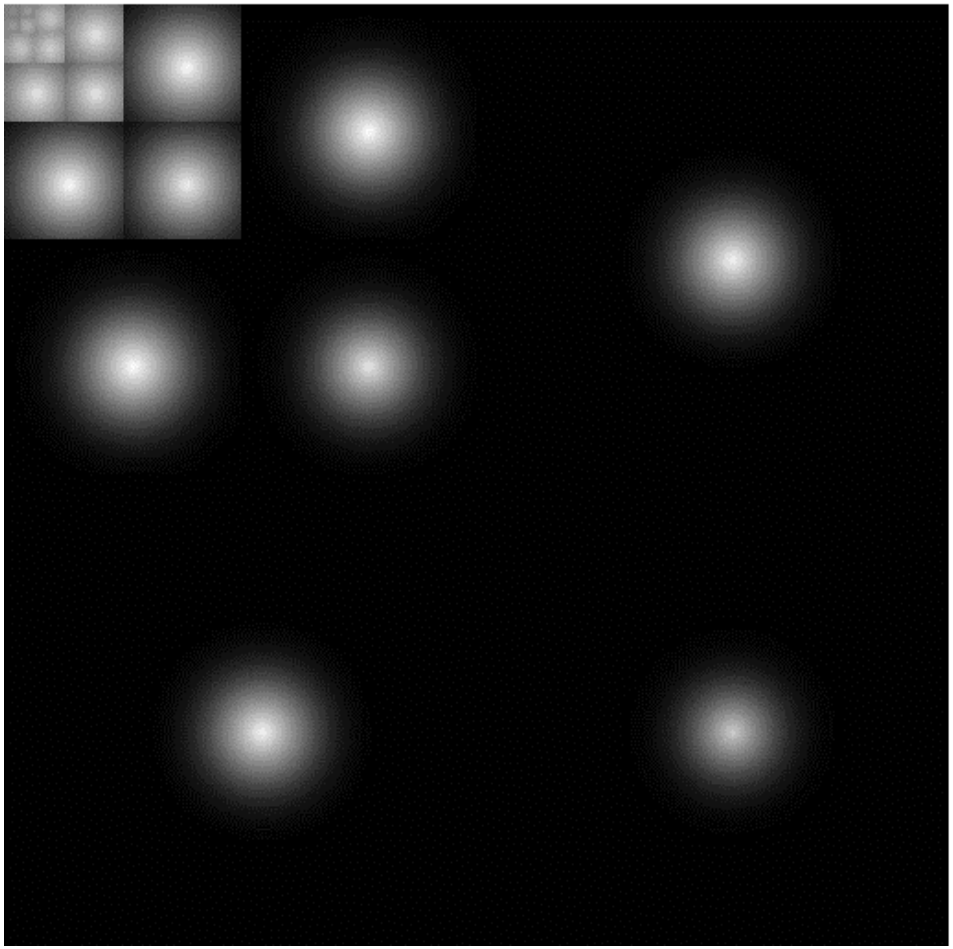


© CRC Press, Wu & Rao Eds, 2006

- Foveal weights in the wavelet domain
  - Determine visible threshold in the wavelet domain
    - Devise error model from data fitting of threshold level noise experiments in the wavelet domain
  - Formulate error detection function from error model
  - Adjust error detection function according to contrast sensitivity

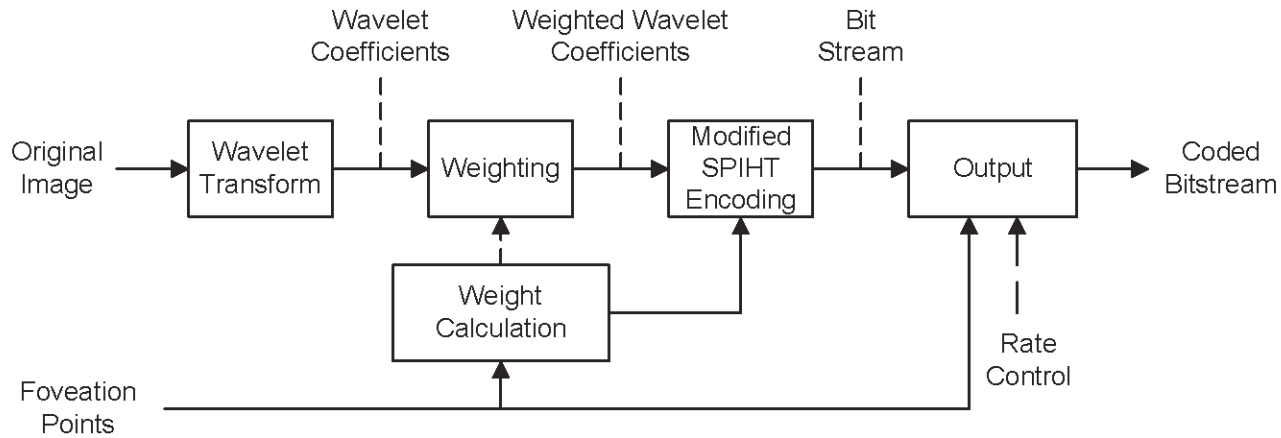
© CRC Press, Wu & Rao Eds, 2006

Wavelet domain weighting mask of a signal foveation point



© CRC Press, Wu & Rao Eds, 2006

14.3.2 Embedded Foveation Image Coding (EFIC)



- EFIC components
  - Wavelet transform
  - Foveal weighting (transmitted to decoder)
  - SPHIT encoding (modified)

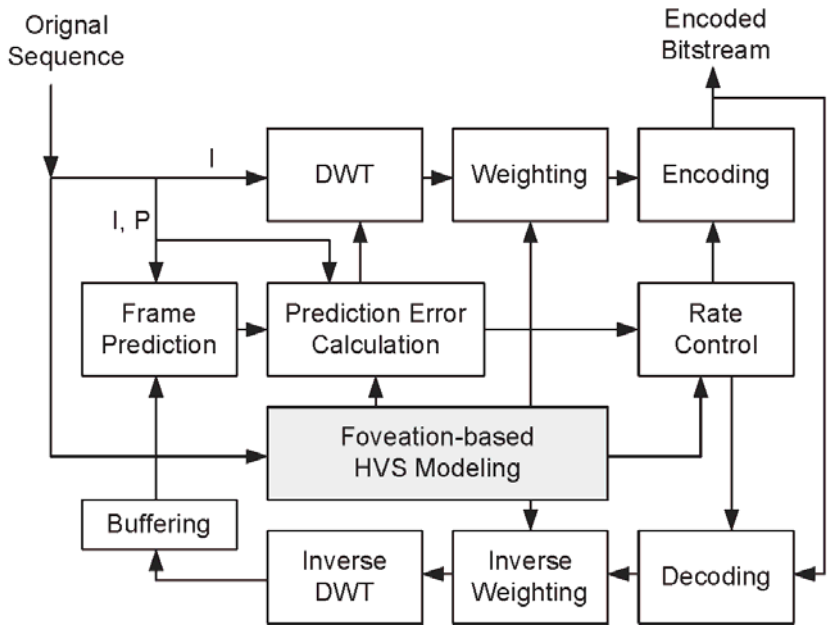
© CRC Press, Wu & Rao Eds, 2006

Zelda Image

a) SPIHT @ 512:1
b) EFIC @ 512:1
c) SPIHT @ 128:1
d) EFIC @ 128:1
e) SPHIT @ 32:1
f) EFIC @ 32:1

© CRC Press, Wu & Rao Eds, 2006

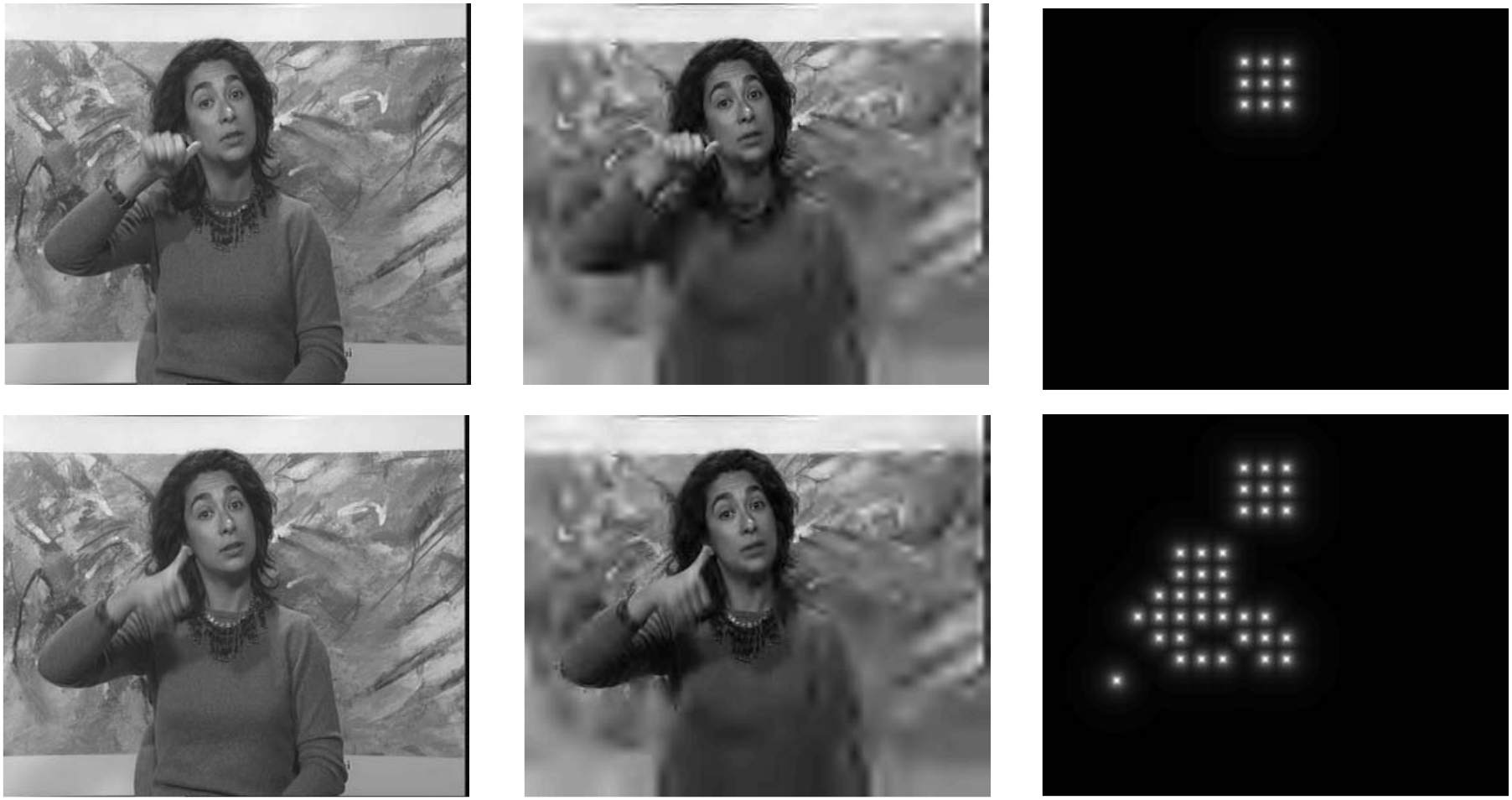
14.3.3 Foveation Scalable Video Coding (FSVC)



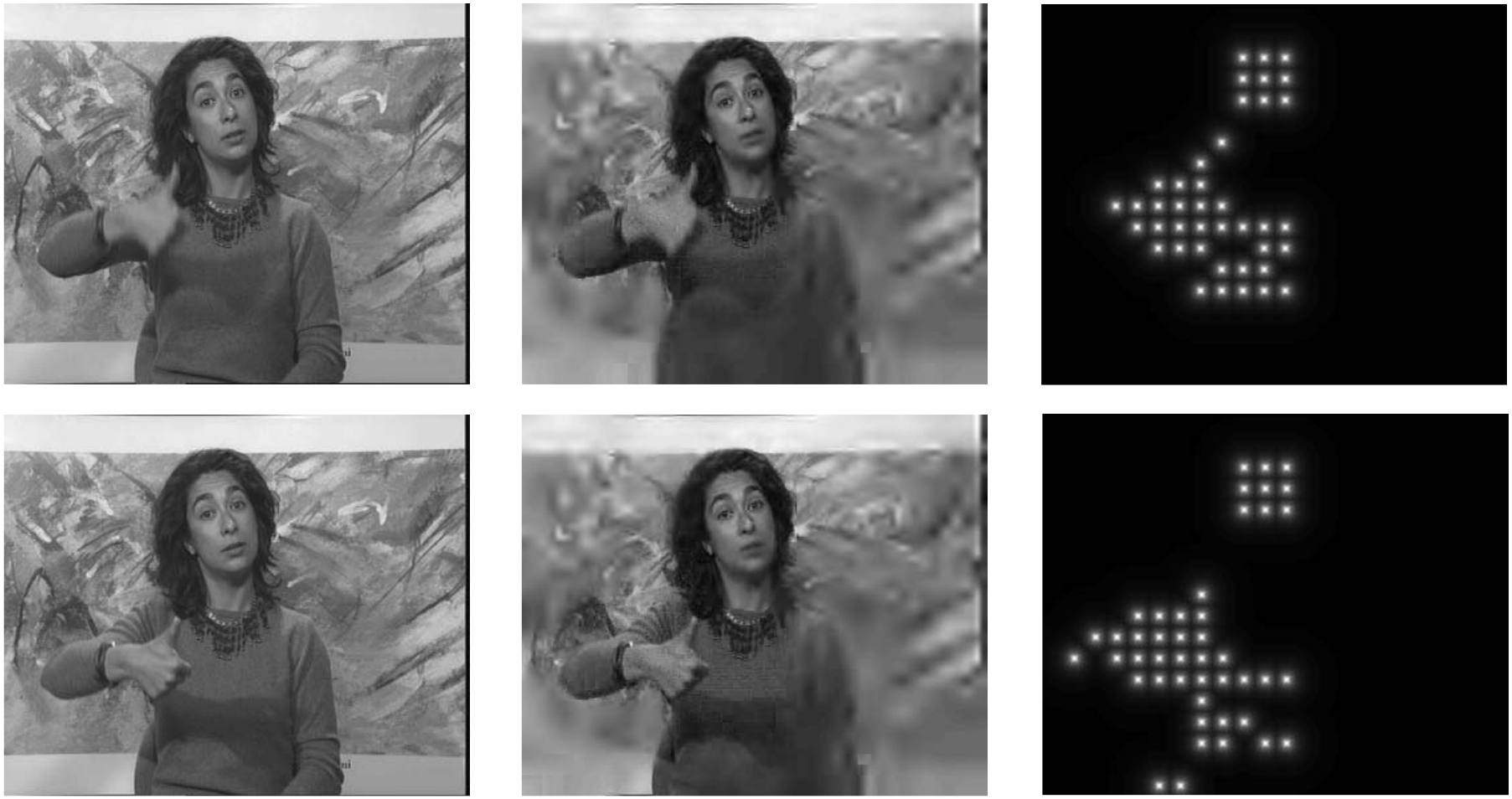
- Scalable coding
- I- and P-frame only
- EFIC for I-frame
- MC/DPCM for P-frame
  - Used previous original and decoded frame for prediction

© CRC Press, Wu & Rao Eds, 2006

Left: Consecutive frames of *Silence* sequence. One I-frame followed by P-frames. *Middle*: FSVC @ 200kbps. *Right*: Foveation points.



© CRC Press, Wu & Rao Eds, 2006



© CRC Press, Wu & Rao Eds, 2006

Frame 32 of salesman sequence coded with FSVC. a) original. b) 200 kbps. c) 400 kbps. d) 800 kbps

Original

200 Kbps

800 Kbps

400 Kbps

© CRC Press, Wu & Rao Eds, 2006

14.4 Discussions

- Principles of foveal operation
- Foveation methods
  - Geometric, filtering-based and mulitresolution
- Scalable foveated image/video coding
  - Foveated perceptual weighting model
  - EFIC (embedded foveation image coding)
  - FSVC (Foveation scalable video coding)

© CRC Press, Wu & Rao Eds, 2006

AG96	T.L. Arnow and W.S. Geisler. Visual detection following retinal damage: Prediction of an inhomogeneous retino-cortical model. *Proc. SPIE: Human Vision and Electronic Imaging*, 2674:119-130, 1996
AS00	S. Aramvith and M.T. Sun. *Handbook of Image and Video Processing (MPEG-1 and MPEG-2 video standards)*. San Diego, CA: Academic Press, May 2000
BA83	P.J. Burt and E.H. Adelson. The Laplacian Pyramid as a compact image code. *IEEE Trans. Communications*, 31:532-540, April 1983
BS89	C. Bandera and P. Scott. Foveal machine vision system. *IEEE International Conference on Systems, Man. and Cybernatics*, 569-599, Nov. 1989
BSA81	M.S. Bands, A.B. Sekuler and S.J. Anderson. Peripheral spatial vision: Limits imposed by optics, photoreceptors and receptor pooling. *Journal of Optical Society of America*, 8:1175-1787, 1981
BSW93	A. Basu, A. Sullivan and K.J. Wiebe. Variable resolution teleconferencing. *IEEE International Conference on Systems, Man. and Cybernatics*, 170-175, Oct. 1993

© CRC Press, Wu & Rao Eds, 2006

Bur88	P.J. Burt. Smart sensing within a pyramid vision machine. *Proc of the IEEE*, 76:1005-1015, Aug. 1988
BW98	A. Basu and K.J. Wiebe. Videoconferencing using spatially varying sensing with multiple and moving fovea. *IEEE Trans. on Systems, Man. and Cybernatics*, 28(2):137-148, 1998
CAS96	P. Camacho, F. Arrebola and F. Sandoval. Shifted fovea multiresolution geometries. *IEEE ICIP*, 1:307-310, 1996
CMY00	E.C. Chang, S. Mallat and C. Yap. Wavelet foveation. *Journal of Applied and Computational Harmonic Analysis*, 9(3):312-335, Oct. 2000
CW99	S.J. Choi and J.W. Woods. Motion-compensated 3-D subband coding of video. *IEEE Trans. on Image Processing*, 8(2):155-167, Feb. 19999
CY97	E.C. Chang and C. Yap. A Wavelet approach to foveating images. *ACM Symposium on Computational Geometry*, 397-399, June 1997
DMRC01	S. Daly, K. Matthews and J. Ribas-Corbera. As plain as the noise on you face: Adaptive video compression using face detection and visual eccentricity models. *Journal of Electronic Imaging*, 10:30-46, Jan. 2001

© CRC Press, Wu & Rao Eds, 2006

E+00	R. Etienne-Cimmings, J. van der Spiegel. P. Mueller and M.Z. Zhang. A foveated silicon retina or two-dimensional tracking. *IEEE Trans. Circuit and Systems II*, 47(6):504-517, June 2000
GB95	W.S. Geisler and M.S. Banks. *Handbook of Optics: Visual performance*. New York: McGraw-Hill, 1995
GP98	W.S. Geisler and J.S. Perry. A real-time foveated multiresolution system for low-bandwidth video communication. *Proc. SPIE: Human Vision and Electronic Imaging*, 3299(2):294-305, July 1998
GP02	W.S. Geisler and J.S. Perry. Space variant imaging. *Centre for Perceptual Systems, The University of Texas at Austin*, 2002 (http://fi.cvis.psy.utexas.edu)
GT99	C. Farcia and F. Tziritas. Face detection using quantized skin color regions merging and wavelet packet analysis. *IEEE Trans. Multimedia*, 1(3):264-277, Sept. 1999
Itt00	L. Itti. *Models of bottom-up and top-down visual attention*. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2000

© CRC Press, Wu & Rao Eds, 2006

KA02	A. Koz and A. Alatan. Foveated image watermarking. *IEEE ICIP*, 3:661-664, Sept. 2002
KB98	W.N. Karquist and A.C. Bovik. FOVEA: A foveated vergent active stereo vision system for dynamic three-dimensional scene recovery. *IEEE Trans. Robotics and Automation*, 14(5):755-770, Oct. 1998
KG96	P. Kortum and W.S. Geisler. Implementation of a foveal image coding system for image bandwidth reduction. *Proc. SPIE: Human Vision and Electronic Imaging*, 2657:350-360, 1996
KGG99	T. Kuyel, W. Geisler and J. Ghosh. Retinally reconstructed images: digital images having a resolution match with the human eyes. *IEEE Trans. Systems, Man. and Cybernetics, Part A: Systems and Humans*, 29(2):235-243, March 1999
Kin99	J.M. Kinser. Foveation from pulse images. In Proc. *IEEE International Conference on Information Intelligence and Systems*, 96-89, 1999
LB00	S. Lee and A.C. Bovik. Foveated video demonstration, 2000. Department of Electrical and Computer Engineering. The University of Texas at Austin. (http://live.ece.utexas.edu/research/foveated_video/demo.html)

© CRC Press, Wu & Rao Eds, 2006

LB03	S. Lee and A.C. Bovik. Fast algorithms for foveated video processing. *IEEE Trans. Circuits and Systems for Video Tech.*, 13(2):148-162, Feb. 2003
Lee00	S. Lee. *Foveated video compression and visual communications over wireless and wireline networks*. Ph.D. thesis, Dept. of ECE, University of Texas, Austin, May 2000
Liu02	S. Liu. *DCT domain video foveation and transcoding for heterogeneous video communication*, Ph.D. thesis, Dept. of ECE, University of Texas, Austin, 2002
LPB01	S. Lee, M.S. Pattichis and A.C. Bovik. Foveated video compression with optimal rate control. *IEEE Trans. Image Processing*, 10(7):977-992, July 2001
LPB02	S. Lee, M.S. Pattichis and A.C. Bovik. Foveated video quality assessment. *IEEE Trans. on Multimedia*, 4(1):129-132, Mar. 2002
LPKB03	S. Lee, M.S. Pattichis, V. Krishnan and A.C. Bovik. Foveation-based error resilience and unequal error protection over mobile networks. *Journal of VLSI Signal Processing*, 34(1/2):149-166, May 2003

© CRC Press, Wu & Rao Eds, 2006

LW90	M. Levoy and E. Whitaker. Gaze-Directed Volume Rendering. *Computer Graphics*, 24(2):217-223, 1990
Mal99	C.M. Privitera and L.W. Stark. Algorithms for defining visual regions-of-interest: comparison with eye fixations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(9):970-982, Sept. 2000
RCB03	U. Rajashekar, L. K. Cormack and A.C. Bovik. Image features that draw fixations. *IEEE ICIP*, 3:313-316, Sept. 2003
RG81	J.G. Robson and N. Graham. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21:409-418, 1981
RR96	T.H. Reeves and J.A. Robinson. Adaptive foveation of MPEG video. *ACM Multimedia*, 231-241, 1996
RS90	A.S. Rojer and E.L. Schwartz. Design considerations for a space-variant visual sensor with complex-logarithmic geometry. *IEEE International Conference on Pattern Recognition*, 2:278-285, 1990

© CRC Press, Wu & Rao Eds, 2006

Rud96	D.L. Ruderman. The statistics of natural images. *Network: Computation in Neural Systems*, 5:517-548, 1996
RWB01	U. Rajashekar, Z. Wang and A.C. Bovik. Real-time foveation: An eye tracker-driven imaging system, 2001. (http://live.ece.utexas.edu/research/realtime_foveation/)
Sab02	M.F. Sabir. Unequal error protection for scalable foveated image communication. Master's thesis, Dept. of ECE, University of Texas, Austin, May 2002
SBD93	P.L. Silsbee, A.C. Bovik and D. Chen. Visual pattern image sequence coding. *IEEE Trans. Circuits and Systems for Video Tech.*, 3(4):291-301, Aug. 1993
Sch77	E.L. Schwartz. Spatial mapping in primate sensory projection: Analytic structure and relevance to perception. *Biological Cybernetics*, 25:181-194, 1997
Sch80	E.L. Schwartz. Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. *Vision Research*, 20:645-669, 1980

© CRC Press, Wu & Rao Eds, 2006

SD99	K.S. Shen and E.J. Delp. Wavelet based rate scalable video compression. *IEEE Trans. Circuits and Systems for Video Technology*, 9(1):109-122, Feb, 1999
SEB02	H.R. Sheikh, B.L. Evans and A.V. Bovik. Real-time foveation techniques for low bit rate video coding. *Real-time Imaging*, 9(1):27-40, Feb. 2003
Sha93	J.M. Shapiro. Embedded image coding using zerotrees of wavelets coefficients. *IEEE Trans. Signal Processing*, 41:3445-3462, Dec. 1993
SKZ96	H. Sun, W. Kwok and J.W. Zdepski. Architectures for MPEG compressed bitstream scaling. *IEEE Trans. Circuits and Systems for Video Technology*, 6(2):191-199, April 1996
SLWB02	H.R. Sheikh, S. Liu, Z. Wang and A.C. Bovik. Foveated multipoint video conferencing at low bit rates. *IEEE ICASSP*, 2:2069-2072, May 2002
SP96	A. Said and W.A. Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circuits and Systems for Video Technology*, 6(3):243-250, June 1996

© CRC Press, Wu & Rao Eds, 2006

TEHM96	N. Tsumura, C. Endo, H. Haneishi and Y. Miyake. Image compression and decompression based on gazing area. *Proc. SPIE: Human Vision and Electronic Imaging*, 2657:361-367, 1996
TM01	D.S. Taubman and M.W. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Norwell, MA: Kluwer Academic Publishers, Nov. 2001
TSS+97	D.L. Tennenhouse, J.M. Smith, W.D. Sincoskie, D.J. Wetherall and G.J. Mindenk. A survey of active network research. *Proc. of the IEEE*, 35, Jan. 1997
TZ94	D. Taubman and A. Zakhor. Multirate 3-D subband coding of video. *IEEE Trans. Image Processing*, 3:572-588, Sept. 1994
VZW99	J.D. Villasenor, Y.W. Zhang and J. Wen. Robust video coding algorithms and systems. *Proc. of the IEEE*, 87:1724-1733, Oct. 1999
Wan95	B.A. Wandell,. *Foundations of Vision*. Sunderland, MA: Sinauer Associates Inc., 1995

© CRC Press, Wu & Rao Eds, 2006

Wan01	Z. Wang. *Rate scalable foveated image and video communications*. Ph.D. thesis, Dept. of ECE, The University of Texas at Austin, Dec. 2001
WB01a	Z. Wang and A.C. Bovik. Demo iamges for ‘embedded foveation image coding’, 2001 (http://www.cns.nyu.edu/~zwang/files/research/efic/demo.html)
WB01b	Z. Wang and A.C. Bovik. Embedded foveation image coding. *IEEE Trans. on Image Processing*, 10(10):1397-1410, Oct. 2001
WBL01	Z. Wang, A.C. Bovik and L. Lu. Wavelet-based foveated image quality measurement for region of interest image coding. *IEEE ICIP*, 2:89-92, Oct. 2001
WC97	H. Wang and S.F. Chang. A highly efficient system for automatic face region detection in MPEG video. *IEEE Trans. Circuits and Systems for Video Technology*, 7(4):615-628, 1997
Wei90	C. Weiman. Video compression via a log polar mapping. *Proc. SPIE: Real Time Image Processing II*, 1295:266-277, 1990

© CRC Press, Wu & Rao Eds, 2006

WLB03	Z. Wang, L. Lu and A.C. Bovik. Foveation scalable video coding with automatic fixation selection. *IEEE Trans. Image Processing*, 11(2):243-254, Feb. 2003
WOBS94	R.S. Wallace, P.W. Ong, B. Bederson and E.L. Schwartz. Space variant image processing. *International Journal of Computer Vision*, 13(1):71-90, 1994
WYSV97	A.B. Watson, G.Y. Yang, J.A. Solomon and J. Villasenor. Visibility of wavelet quantization noise. *IEEE Trans. Image Processing*, 6(8):1164-1175, 1997
ZS93	Y.Y. Zeevi and E. Shlomot. Nonuniform sampling and anitaliasing in image representation. *IEEE Trans. Signal Processing*, 41(3):1223-1236, March 1993

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

15. Artifact Reduction by Post-Processing in Image Compression

T. Chen† and H.R. Wu‡

† Panasonic Hollywood Laboratory, USA

‡ Royal Melbourne Institute of Technology, Australia

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 15 Outline

15.1 Introduction

15.2 Image Compression and Coding Artifacts

15.3 Reduction of Blocking Artifacts

15.4 Reduction of Ringing Artifacts

15.5 Summary

© CRC Press, Wu & Rao Eds, 2006

15.1 Introduction

- Proliferation of applications that rely on compression systems
  - Video-conferencing, video over internet, etc
- Coding systems introduces distortions
  - Visible artifacts degrades quality of services
  - Prevalent at low bitrates
- Picture processing for distortion reduction
  - Pre-processing at encoding end
  - Post processing at decoding end

© CRC Press, Wu & Rao Eds, 2006

15.2 Image Compression and Coding Artifacts

- Transform based coding
  - Employ data transformation
    - Reduces redundancies
  - Popular due to
    - Near-optimal decorrelation and energy compaction
    - Fast implementation
  - Recommended in international standards
    - DCT in JPEG, MPEG, H.26x
    - DWT in JPEG2000

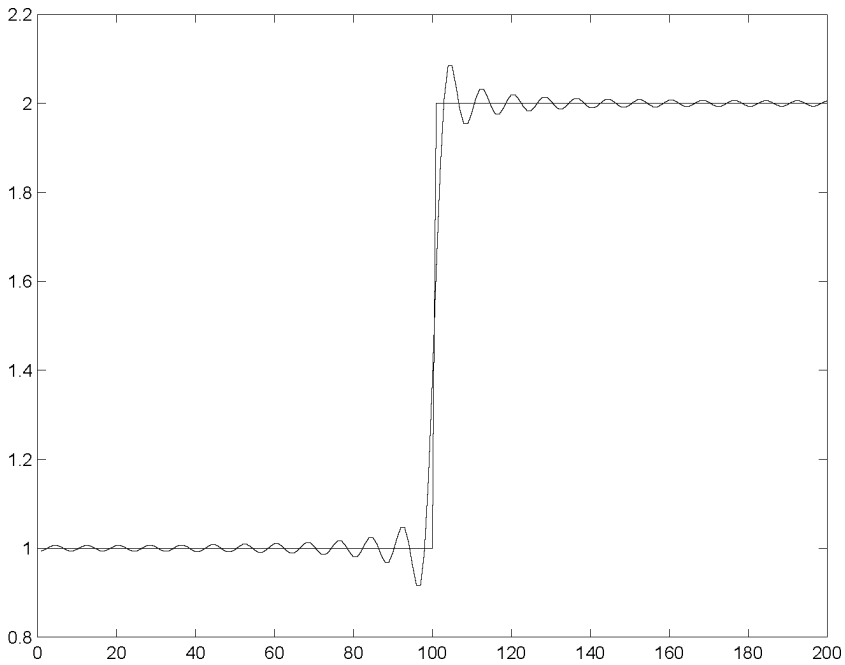
© CRC Press, Wu & Rao Eds, 2006

- Distortions in transform based coders
  - Appears primarily at low bitrates
  - Blocking artifact most prominent in DCT
  - Ringing and blurring artifacts for DWT
- Blocking Artifact (grid noise)
  - Discontinuities between adjacent blocks
  - Appears on block boundaries
  - Side effect of quantization operations
    - When mid and high frequency DCT coefficients are nullified

© CRC Press, Wu & Rao Eds, 2006

Coded *Lena* image at 0.22 bpp. *Left*: JPEG. *Right*: JPEG2000

© CRC Press, Wu & Rao Eds, 2006



- Ringing artifacts
  - Spurious oscillation in the vicinity of major edges (at low bitrates)
  - Similar to Gibb's phenomenon

Left: example of ringing for 1-D signal

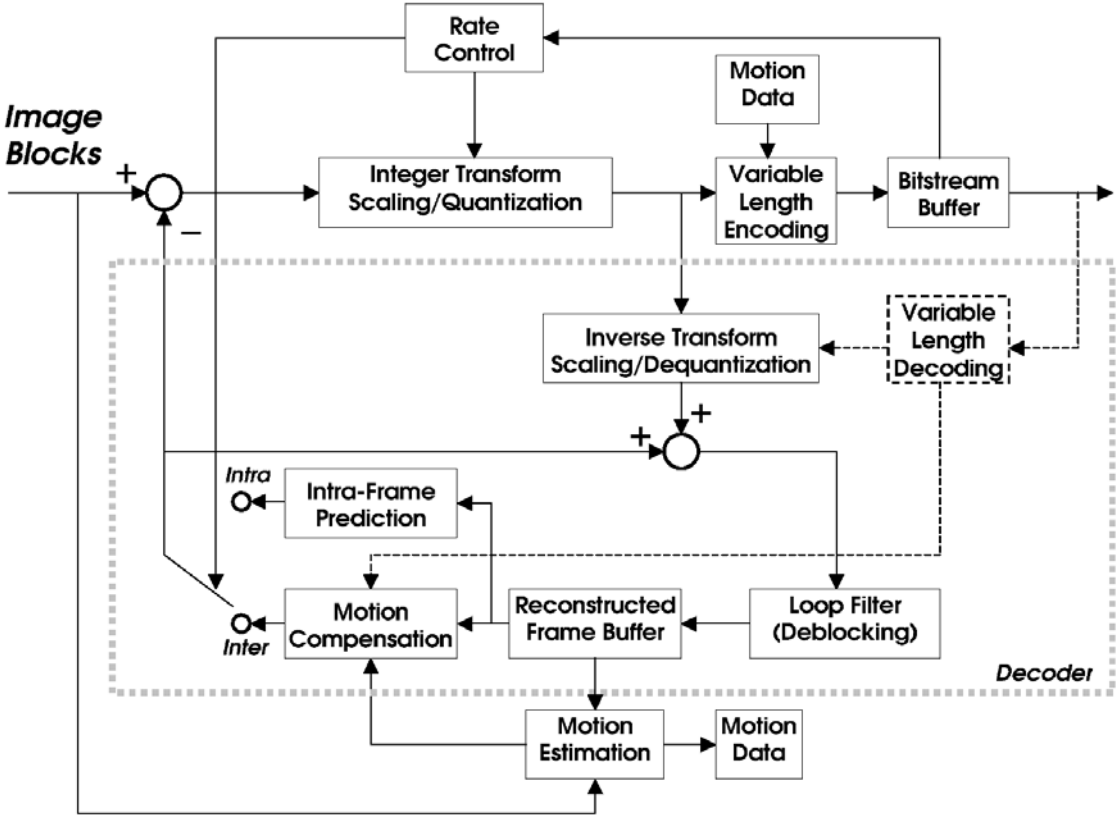
© CRC Press, Wu & Rao Eds, 2006

15.3 Reduction of Blocking Artifacts

- Important part of coding system
  - Implemented in H.264/MPEG-4 AVC
- Application of de-blocking filters
  - Mostly at post-processing end
  - May also be employed at the encoding end in video coders, e.g., H.264/MPEG-4 AVC
- Simple low pass filtering in spatial domain
  - Reduces blocking, may lead to strong blurring

© CRC Press, Wu & Rao Eds, 2006

Simplified block diagram of H.264/MPEG-4 AVC

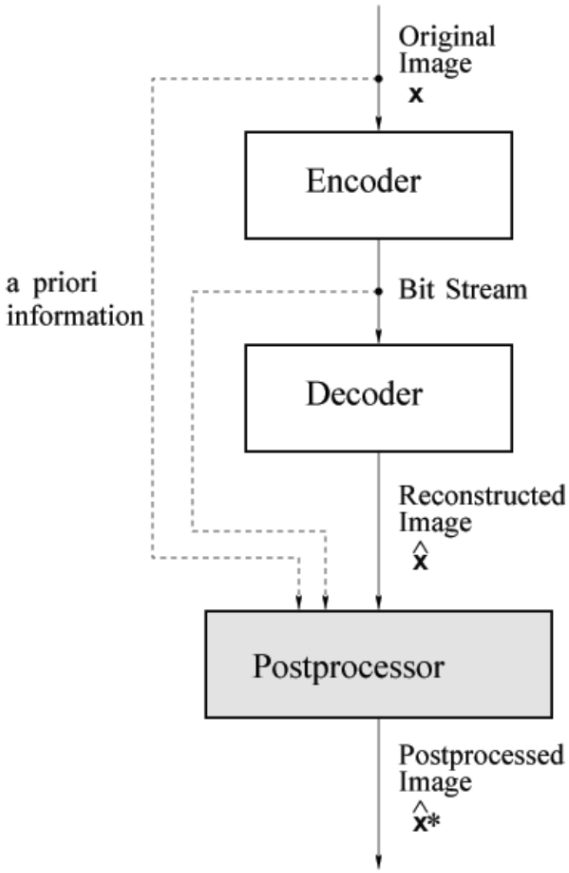


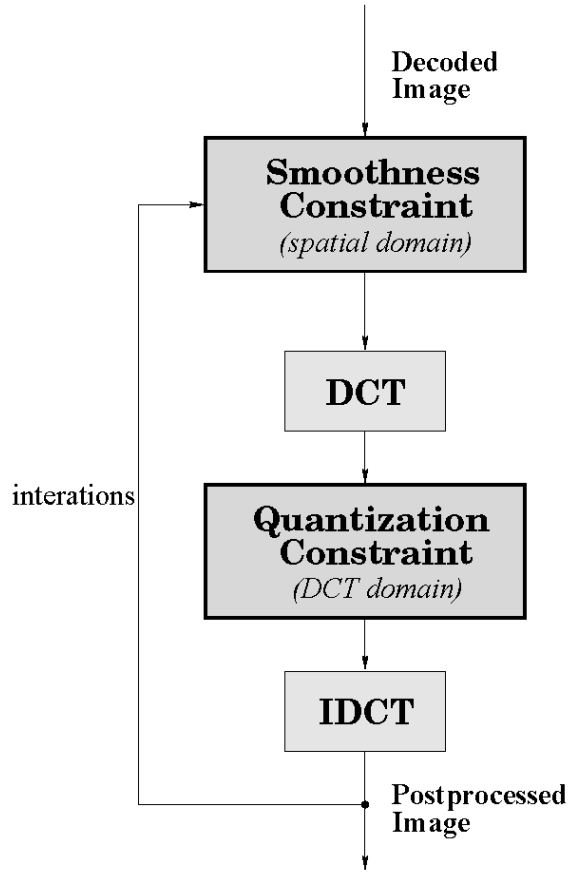
© CRC Press, Wu & Rao Eds, 2006

- Projection Onto Convex Set (POCS) filter
  - Applied at the decoding end
  - Iterative filter
  - Blurs images excessively
  - Deals with pixel straddling the block boundary
    - May lead to new discontinuity between each altered pixel and adjacent pixels

© CRC Press, Wu & Rao Eds, 2006

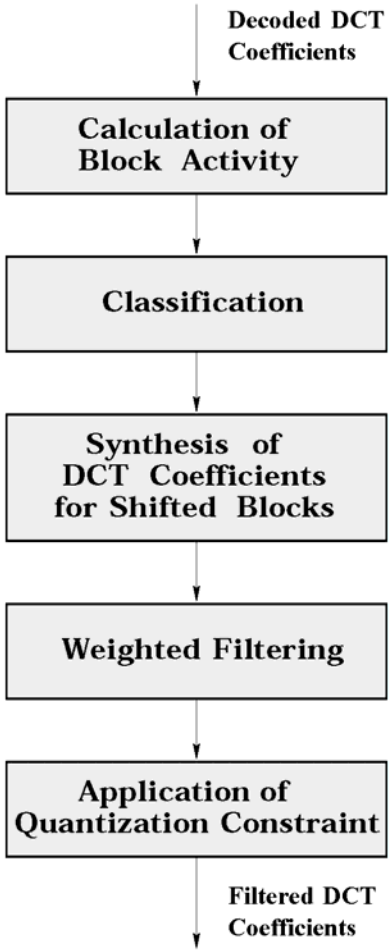
Left: General arrangement of picture coding systems. *Right*: General structure of POCS filters





© CRC Press, Wu & Rao Eds, 2006

15.3.1 Adaptive Postfiltering of Transform Coefficients

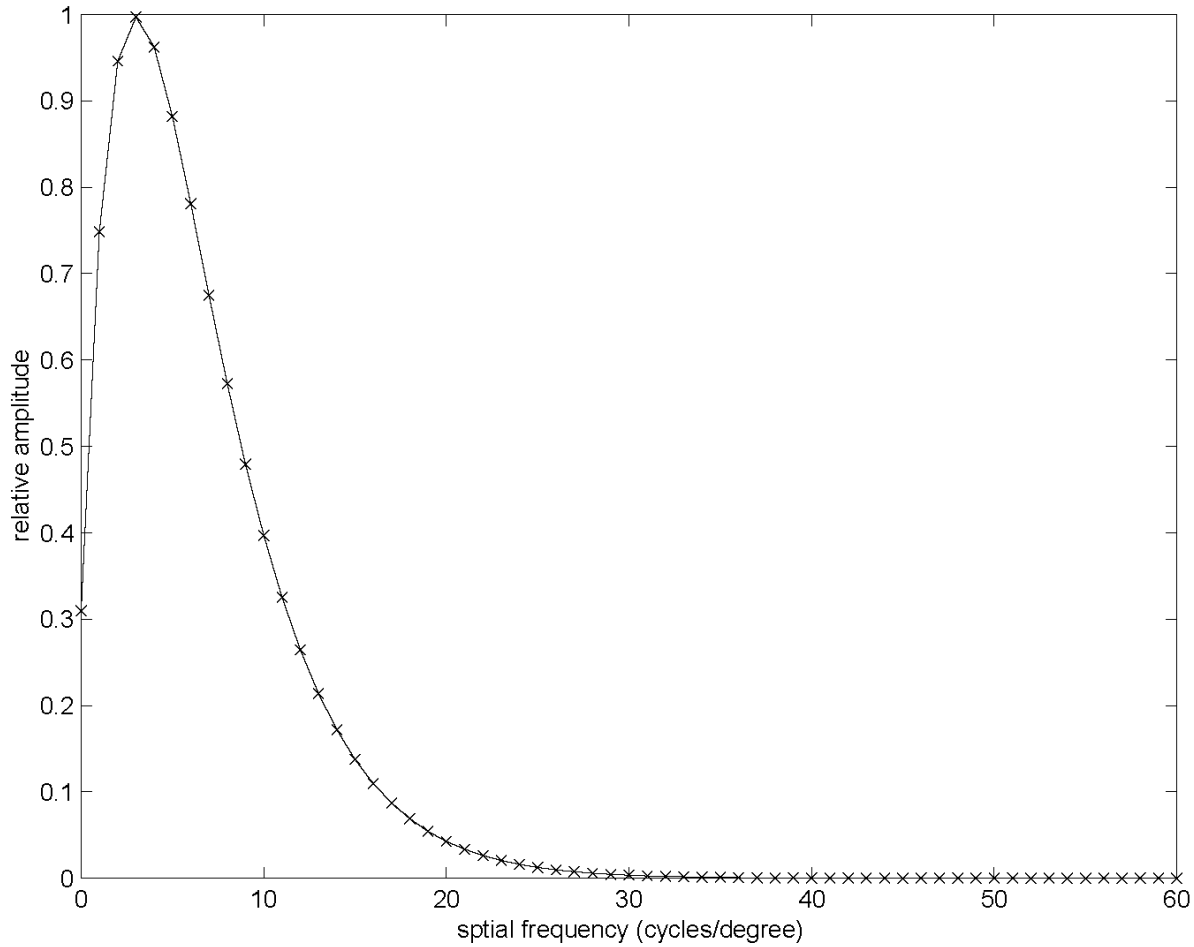


Left: Illustration of the implementation of adaptive postfiltering of DCT coefficients

- HVS considered and modelled by the MTF
  - Operationally, AC coefficients weighted by MTF

© CRC Press, Wu & Rao Eds, 2006

Modulation Transfer Function (MTF). Measures visual sensitivity relative to spatial frequencies

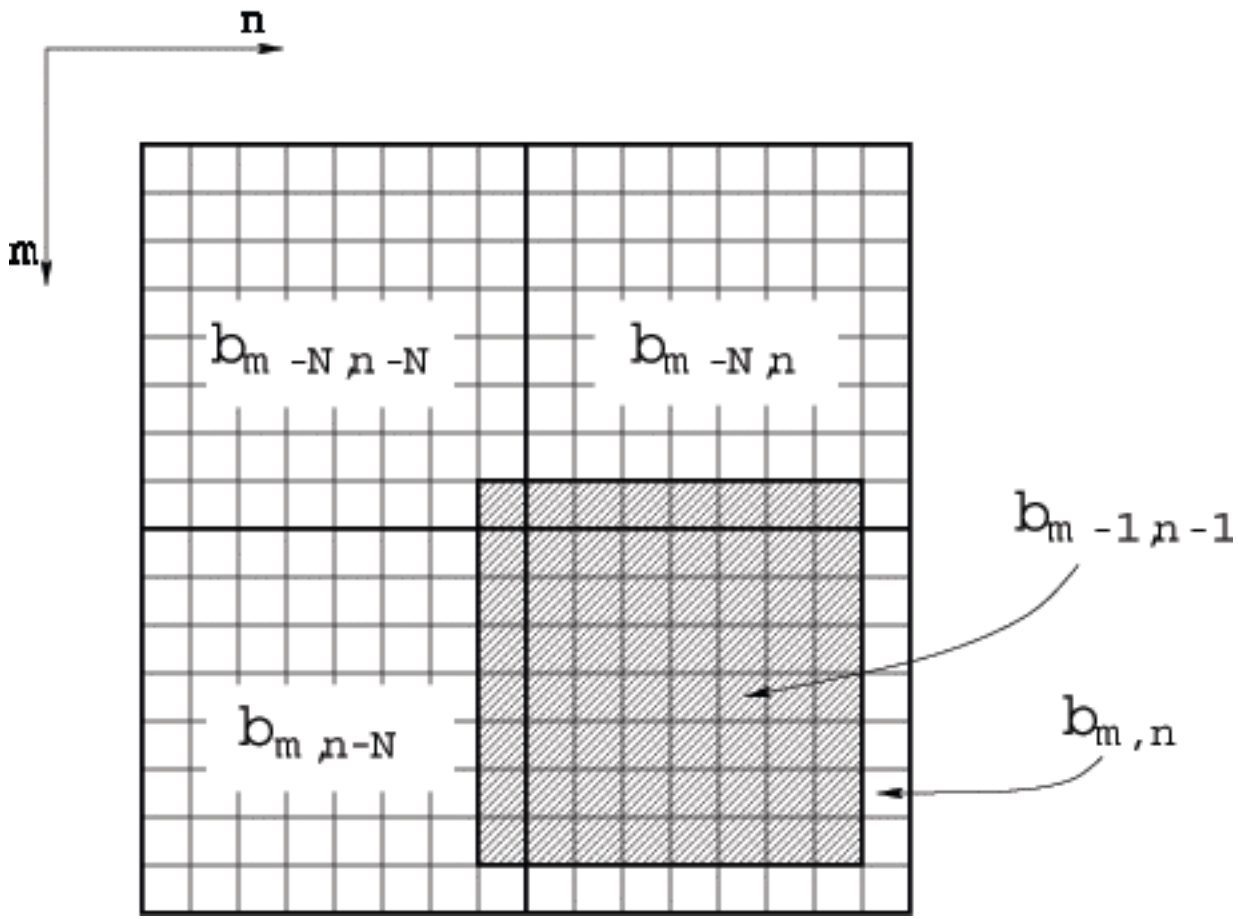


© CRC Press, Wu & Rao Eds, 2006

- Calculation of Block activity
  - Sum of AC coefficients normalized by local background luminance (Weber's Law)
- Classification of blocking activity
  - Classified into low or high activity blocks
    - Based on some pre-defined threshold value
- DCT block shifting
  - Forms new stationary estimates of DCT coefficients for weighted filtering

© CRC Press, Wu & Rao Eds, 2006

Illustration of block shifting



© CRC Press, Wu & Rao Eds, 2006

- Adaptive weighted filtering
  - Weighting relative to local activity
    - Low activity
      - 5x5 filtering window
      - Uniform filter weights
    - High activity –
      - 3x3 filtering window
      - Filter weights: 3 for central coefficient and 1 for all others
      - Avoids excessive blurring of high activity block

© CRC Press, Wu & Rao Eds, 2006

- Quantization constrains
  - Projects filtering to de-quantized domain
    - i.e. ensures quantizer output stays within the post-filtering boundary

© CRC Press, Wu & Rao Eds, 2006

15.3.3 Simulation Results

Image	Bitrate (bpp)	PSNR (dB)						
		Decoded	H.263	MPEG4	POCS-Z	POCS-Y	POCS-P	Proposed
Airplane	0.241	28.72	29.34	29.57	27.91	29.34	29.07	29.39
Barbara	0.295	25.08	25.31	25.52	25.06	25.52	25.34	25.62
Boats	0.235	28.52	29.07	29.30	28.01	29.16	28.76	29.25
Goldhill	0.230	27.91	28.50	28.53	27.72	28.46	28.12	28.54
Lena	0.219	29.47	30.20	30.43	29.02	30.23	29.88	30.39
Peppers	0.221	29.21	30.02	30.15	28.82	29.85	29.65	29.95
Sailboat	0.286	27.00	27.51	27.68	26.46	27.57	27.26	27.66
Splash	0.178	31.23	32.00	32.29	30.61	31.77	31.75	31.89
Tiffany	0.188	29.59	30.10	30.38	29.36	30.11	29.97	30.15

- Evaluations
  - 9 images
  - Compared with 5 de-block algorithms
    - POCS and 2 POCS variations
    - H.263 and MPEG-4 standard de-block filters

© CRC Press, Wu & Rao Eds, 2006

Image	M_{GBIM}							
	Original	JPEG	H.263	MPEG4	POCS-Z	POCS-Y	POCS-P	Proposed
Airplane	1.1177	4.0615	1.3917	1.4209	1.5470	1.7439	2.0292	1.2864
Barbara	1.0415	3.0789	1.7398	1.1465	1.3447	1.6753	1.7578	1.2404
Boats	1.0737	4.1674	1.9803	1.4785	1.5707	1.4402	2.1040	1.3387
Goldhill	1.0351	4.7412	1.5288	1.5091	1.5285	1.4169	2.2451	1.3979
Lena	1.1009	4.9905	1.5514	1.5931	1.5702	1.4780	2.4191	1.3777
Peppers	1.1368	5.6746	1.7197	1.7831	1.5581	1.8250	2.2613	1.3986
Sailboat	1.1414	3.8388	1.5359	1.4880	1.6358	1.4654	2.1186	1.3321
Splash	1.0435	5.7207	1.8760	1.4846	1.4191	2.3544	1.8109	1.2102
Tiffany	1.0139	6.0632	1.8979	1.5240	1.7656	2.2406	1.9322	1.3784

- Evaluation methods
  - PSNR
  - GBIM (*Generalized Block Impairment Metric*)
- Results
  - Proposed method superior to other algorithms

© CRC Press, Wu & Rao Eds, 2006

Original | JPEG coded | H.263 post-filter | MPEG-4 post-filter

POCS-Z | POCS-Y | POCSP | Proposed

© CRC Press, Wu & Rao Eds, 2006

15.4 Reduction of Ringing

- Problem with de-ringing
  - Blurring of image content
  - No enhancement edge sharpness
    - Due, in part, to the difficulty of distinguishing between content edges and ringing distortion
- Manifestation in DCT and DWT coders
  - More prominent on DWT coders

© CRC Press, Wu & Rao Eds, 2006

Left: Lena image coded by MPEG-4 coder. *Right*: left image after de-ringing

© CRC Press, Wu & Rao Eds, 2006

15.5 Summary

- Post-processing for applications in picture coding and processing
  - Reduction of coding distortions
    - Blocking for DCT
      - De-blocking filters
        - » Adaptive postfiltering of transform coefficients
    - Ringing primarily for DWT

© CRC Press, Wu & Rao Eds, 2006

Cla85	T.J. Clarke. *Transform Coding of Images*. Orlando, FL: Academic Press, 1985
CM95	S.F. Chang and D.G. Messerschmitt. Manipulation and compositing of MC-DCT compressed video. *IEEE J. Selected Areas of Communication*, 13(1):1-11, Jan. 1995
CR90	B. Chitprasert and K.R. Rao. Human visual weighting progressive image transmission. *IEEE Trans.on Communications*, 38(7):1040-1044, July 1990
CWQ01	T. Chen, H.R. Wu and B. Qiu. Adaptive postfiltering of transform coefficients for the reduction of blocking artifacts. *IEEE Trans. on Circuits and System for Video Technology*, 11:594-602, May 2001
dNR96	R.L. de Queiroz, T.Q. Nguyen and K.R. Rao. GenLOT: generalized linear-phase lapped orthogonal transform. *IEEE Trans. Signal Processing*, 44:497-507, Mar. 1996
FC00	G. Fan and W.K. Cham. Model-based edge reconstruction for low bit-rate wavelet compressed images. *IEEE Trans. on Circuits and Systems for Video Technology*. 10(1):120-132, Jan. 2000

© CRC Press, Wu & Rao Eds, 2006

FE96	Z. Fan and R. Eschbach. JPEG decompression with reduced artifacts. In *Proc. SPIE Symposium on Image and Video Compression*, 2186:50-55, Feb. 1996
FJ86	P. Farrelle and A.K. Jain. Recursive block-coding – a new approach to transform coding. *IEEE Trans. on Communications*, COM—34:161-179, Feb. 1986
Gha03	M. Ghanbari. *Standard Codecs: Image Compression to Advanced Video Coding*, London, UK:IEE, 2003
GK02	W. Gao and Y. Kim. A de-blocking algorithm and a blockiness metric for highly compressed images. *IEEE Trans. Circuits Sys. Video Tech.*, 12(12):1150-1159, Dec. 2002
HBS84	B. Hinman, J. Bernstein and D. Staelin. Short-space Fourier transform image processing. In *Proc. IEEE ICASSP*, 481-484, San Diego, CA, Mar. 1984
IEE03	IEEE Trans. Circuits Sys. Video Tech. *Special issue on H.264/MPEG-4 Part 10*, July 2003
ISO99a	ISO/IEC. *JPEG 2000 Requirements and Profile*, Mar. 1999

© CRC Press, Wu & Rao Eds, 2006

ISO99b	ISO/IEC JTC1/SC29/WG11. MPEG4 Video Verification Model Version 15.0 (VM 15.0)/ N3093, Dec. 1999
ISO00	ISO/IEC. *JPEG 2000 Part 1: Final Draft International Standard (ISO/IEC FDIS15444-1)*, Aug. 2000
ITU97	ITU-T Video Coding Experts Group. Video Codec Test Model, Near-term, version 8 (TMN8). Release 0, June 1997
ITU98	ITU-T Video Coding Experts Group. Video coding for low bit rate communication. ITU-T recommendation H.263, Jan. 1998
J.-92	J.H. Chen *et al*. A low delay CELP coder for the CCITT 16 kbps speech coding standard. *IEEE J. Select. Areas Commun.*, 10:830-849, June 1992
J. 96	J.L. Mitchell *et al*. *MPEG Video: Compression Standard*. New York, NY: Chapman & Hall, 1996
Jai89	A.K. Jain. *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989

© CRC Press, Wu & Rao Eds, 2006

JJ98	B. Joen and J. Jeong. Blocking artifacts reduction in image compression with blcok boundary discontonuity criterion. *IEEE Trans. Circuits Sys. Video Tech.*, 8(3):345-357, June 1998
JJS93	N. Jayant, J. Johnston and R. Safranek. Signal compression based on model of human perception. *Proc. IEEE*, 81(10):1385-1422, Oct. 1993
Joi03	Joint Video Team of ISO MPEG and ITU-T VCEG. Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264 and ISO/IEC 14496-10 AVC). Approved Output Document of JVT, JVT-G050r1 (ftp:standards.polycom.com), May 2003
KH95	C.J. Kuo and R.J. Hsieh. Adaptive prostprocessing for block encoded images. *IEEE Trans. Circuits Sys. Video Tech.*, 5:298-304, Aug. 1995
KK95	S.A. Karunasekera and N.G. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transaction on Image Processing*, 4(6):713-724, June 1995
Kun95	A. Kundu. Enhancement of JPEG coded images by adaptive spatial filtering. In *Proc. of IEEE ICIP*, 187-190, Oct. 1995

© CRC Press, Wu & Rao Eds, 2006

LB02	S. Liu and A.C. Bovik. Efficient DCT-domain blind measurement and reduction of blocking artifacts. *IEEE Trans. on Circuits System for Video Technology*, 12(12):1139-1149, Dec. 2002
LCP97	J. Luo, C. Chen and K.J. Parker. Image enhancement for low bit rate wavelet-based image compression. In *Proc. of IEEE ISCAS*, 1081-1084, June 1997
LK97	J. Li and C.C.J. Kuo. Coding artifact removal with multiscale postprocessing. In *Proc. IEEE ICIP*, 529-532, 1997
LRL95	W.E. Lynch, A.R. Reibman and B. Liu, Postprocessing transform coded images using edges. In *Proc. IEEE ICASSP*, 2323-2326, May 1995
LW03	Y. Luo and R.K. Ward. Removing the blocking artifacts of block-based DCT compressed images. *IEEE Trans. Image Processing*, 1(7):838-842, July 2003
Mal92	H.S. Malvar. *Signal Processing with Lapped Transforms*. Boston, MA: Artech House, 1992

© CRC Press, Wu & Rao Eds, 2006

MB97	N. Merhav and V. Bhaskaran. Fast algorithm for DCT-domain image down-sampling and for inverse motion compensation. *IEEE Trans. Circuits Sys. Video Tech.*, 7(3):468-476, June 1997
MS74	J.L. Mannos and D.J. Sakrison. The effects of a visual fidelity criterion on the encoding of images. *IEEE Trans. Information Theory*, 20(4):525-536, July 1974
MS82	H.S. Nalvar and D.H. Staelin. The LOT: Transform coding without blocking effects. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 37:553-559, April 1982
MSF97	J.D. McDonnell, R.N. Shortern and A.D. Fagan. Postprocessing of transform coded images via histogram-based edge classification. *J. Electronic Imaging*, 6(1):114-124, Jan. 1997
NH94	A.N. Netravali and B.G. Haskell. *Digital Pictures: Representation, Compression and Standards*. New York, NY: Plenum Press, 2nd ed., 1994
Nil85	N.B. Nill. A visual model weighted cosine transform for image compression and quality assessment. *IEEE Trans. on Communications*, COM-33(6):551-557, June 1985

© CRC Press, Wu & Rao Eds, 2006

NLS89	K.N. Ngan, K.S. Loeng and H. Singh. Adaptive cosine transform coding of images in perceptual domain. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 37(11):1743-1750, Nov. 1989
Nos99	A. Nosratinia. Embedded post-processing for enhancement of compressed images. In *Proc. IEEE Data Compression Conference*, 62-71, March 1999
OS95	T.P. O'Rouke and R.L. Stevenson. Improved image decompression for reduced transform coding artifacts. *IEEE Trans. on Circuits and Systems for Video Technology*, 5:490-499, Dec. 1995
OST93	M. Ozkan, M. Sezan and A. Tekalp. Adaptive motion-compensated filtering of noisy image sequences. *IEEE Trans. Circuits Sys. Video Tech.*, 3(4):277-290, Aug. 1993
P. 03	P. List *et al*. Adaptive deblocking filter. *IEEE Trans,. Circuits Sys. Video Tech.*, 13(7):614-619, July 2003
PLK98	H. Paek, R.C. Kim and S.U. Lee. On the POCS-based postprocessing technique to reduce the blocking artifacts in transform coded images. *IEEE Trans. Circuits Sys. Video Tech.*, 8(3):358-367, June 1998

© CRC Press, Wu & Rao Eds, 2006

PM93	W.B. Pennebaker and J.L. Mitchell. *JPEG Still Image Data Compression Standard*. New York, NY: Van Nostrand Reinhold, 1993
PW84	D. Pearson and M. Whybray. Transform coding of images using interleaved blocks. *Proc. IEE*, 131:466-472, August 1984
RG86	B. Ramaurthi and A. Gersho. Nonlinear space-variant postprocessing of block coded images. *IEEE Trans. on Acoustic, Speech and Signal Processing*, ASSP-34:1258-1267, Oct. 1986
Ric03	I.E.G. Richardson. *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia*. Hoboken, NJ: Wiley, 2003
RJ02	M. Rabbani and R. Joshi. An overview of the JPEG 2000 still image compression standard. *Signal Processing: Image Communication*, 17:3-48, Jan, 2002
RL84	H. Reeve and J. Lim. Reduction of blocking effects in image coding. *Optical Engineering*, 23(1):34-37, Jan, 1984
RY90	K.R. Rao and P. Yip. *Discrete Cosine Transform – Algorithms, Advantages, Applications*. Boston, MA: Academic Press, 1990

© CRC Press, Wu & Rao Eds, 2006

S. 01	S. Yang *et al*. Maximum-likelihood parameter estimation for image ringing-artifact removal. *IEEE Trans. Circuits Sys. Video Tech.*, 11(5):963-973, Aug. 2001
Sau91	K. Sauer. Enhancement of low bit-rate coded images using edge detection and estimation. *Computer Vision, Graphics, Image Processing*, 53:52-62, Jan. 1991
SCE01	A. Skodras, C. Christopoulos and T. Ebrahimi. The JPEG 2000 still image compression standard. *IEEE Signal Processing Magazine*, 18(5):36-58, Dec. 2001
Sha93	J.M. Shapiro. Embedded image coding using zerotress of wevelet coefficients. *IEEE Trans. Signal Processing*,, 41(12):3345-3462, Dec. 1993
SK98	M.Y. Shen and C.C.J. Kuo. Aritfact reduction in low bit-rate wavelet coding with robust nonlinear filtering. In *Proc. IEEE Workshop on Multimedia Signal Processing*, 480-485, 1998
SK99	M.Y. Shen and C.C.J. Kuo. Real-time compression artifact reduction via robust nonlinear filter. In *Proc. of IEEE ICIP*, 565-569, 1999

© CRC Press, Wu & Rao Eds, 2006

SP96	A. Said and W.A. Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circuits and Systems for Video Technology*, 6(3):243-250, June 1996
Ste93	R.L. Stevenson. Reduction of coding artifacts in transform image coding. In *Proc IEEE ICASSP*, 5:401-404, Mar. 1993
SY00	J. Song and B.L. Yeo. A fast algorithm for DCT-domain inverse motion compensation based on shared information in a macroblock. *IEEE Trans. Circuits Sys. Video Tech.*, 10(5):767-775, Aug. 2000
TH93	C.N. Tien and H.M. Hang. Transform-domain prostprocessing of DCT-coded images. In *Proc. SPIE Visual Communication and Image Processing*, 2094:1627-1638, 1993
TM01	D.S. Taubman and M.W. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Norwell, MA: Kluwer Academic Publishers, Nov. 2001
TPN96	S.H. Tan, K.K. Pang and K.N. Ngan. Classified perceptual coding with adaptive quantization. *IEEE Trans. Circuits Sys. Video Tech.*, 6(4):375-388, Aug. 1996

© CRC Press, Wu & Rao Eds, 2006

TTS02	G.A. Triantafyllidis, D. Tzovaras and M.G. Strintzis. Blocking artifact detection and reduction in compressed data. *IEEE Trans. Circuits Sys. Video Tech.*, 12(10):877-890, Oct. 2002
Tzo83	K.H. Tzou. Post-filtering of transform-coded images. In *Proc. SPIE Applications of Digital Image Processing XI*, 121-126, April 1983
WSA94	A.B. Watson, J.A. Solomon and A. Ahumada. DCT basis function visibility: Effects of viewing distance and contrast masking. In *Proc. SPIE Human Vision, Visual Processing and Digital Display*, 2178:99-108, May 1994
WY97	H.R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 4(11):317-320, Nov. 1997
YGK93	Y. Yang, N.P. Galastanos and A.K. Katsaggelos. Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images. *IEEE Trans. Circuits Sys. Video Tech.*, 3:421-432, Dec. 1993
YW98	M. Yuen and H.R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing (Special Issue on Image Quality Assessment)*, 70(3):247-278, Nov. 1998

© CRC Press, Wu & Rao Eds, 2006

Zak92	A. Zakhor. Iterative procedures for reduction of blocking effects in transform image coding. *IEEE Trans. on Circuits Sys. for Video Tech.*, 2:91-95, Mar. 1992
ZPL93	Y. Zhang, R. Pickholtz and M. Loew. A new approach to reduce the blcok effect of transform coding. *IEEE Trans. on Communications*, 41:299-302, Feb. 1993

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

16. Reduction of Color Bleeding in DCT Block-Coded Video

F.-X. Coudoux and M.G. Gazalet

University of Valenciennes, France

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 16 Outline

16.1 Introduction

16.2 Analysis of the Color Bleeding Phenomenon

16.3 Description of the Post-Processor

16.4 Experimental Results – Concluding Remarks

© CRC Press, Wu & Rao Eds, 2006

16.1 Introduction

- Color bleeding in DCT coders
  - Appears as a smearing of color between highly contrasting chrominance
  - May lead to secondary color ringing effect
  - Visually objectionable

Left: Illustration of color bleeding

© CRC Press, Wu & Rao Eds, 2006

16.2 Analysis of the Color Bleeding Phenomenon

- Manifestation of color bleeding
  - Occurs when boundaries of chroma channels overlaps inappropriately
    - Resulting in chrominance mismatch
  - Initiated by decimation and quantization of chrominance components
    - Coarse quantization of AC bands

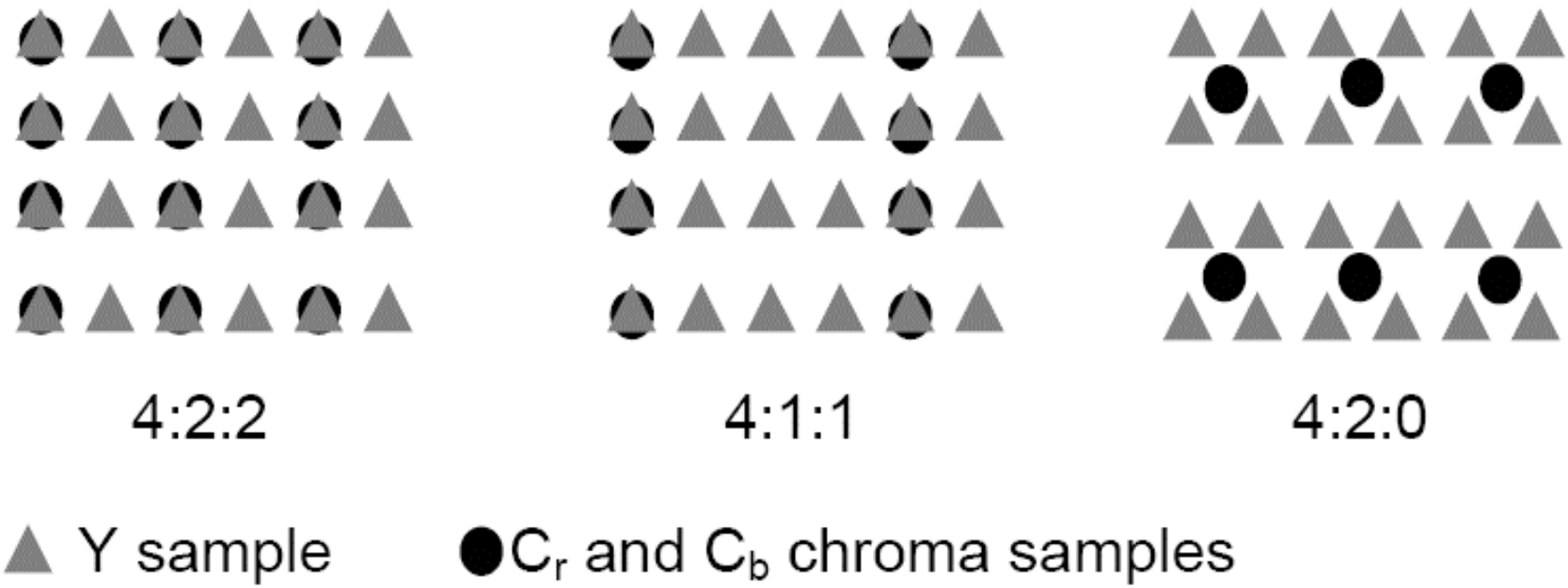
© CRC Press, Wu & Rao Eds, 2006

16.2.1 Digital Color Video Formats

- YUV colour space ($Y_dC_bC_r$ digital form)
  - Often employed in color picture coding
  - Y contains the luminance component
  - U and V the chrominance channels
  - Human vision is more sensitive to luminance than chrominance channel
    - Little or no quality with sub-sampling chroma channels

© CRC Press, Wu & Rao Eds, 2006

Chrominance sub-sampling strategies in $Y_dC_bC_r$ color space



© CRC Press, Wu & Rao Eds, 2006

16.2.2 Color Quantization

- Quantization account for visual sensitivity
  - Sensitivity varies with spatial frequency
  - Different sensitivities for luminance and chrominance channels
    - Chroma channels less sensitive
      - Higher quantization for C_b and C_r

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Luminance (*left*) and chrominance (*right*) quantization matrices

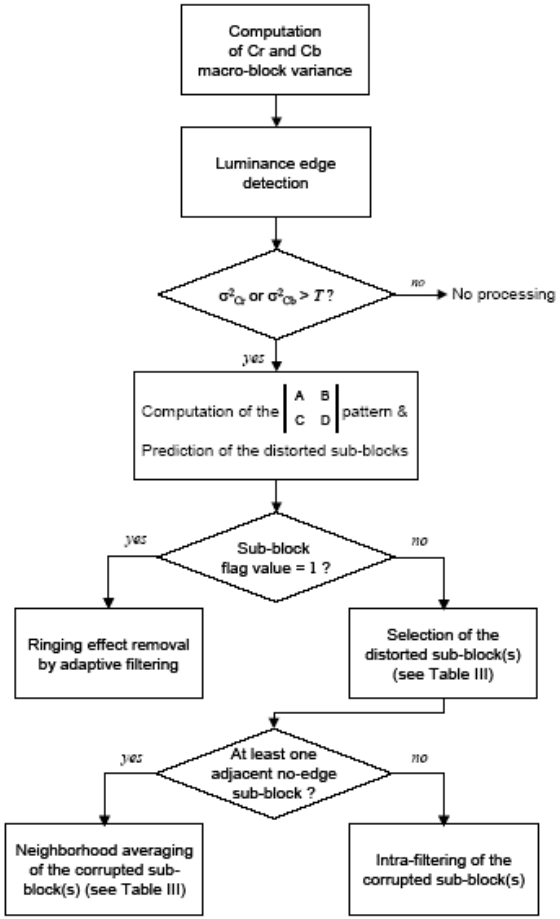
© CRC Press, Wu & Rao Eds, 2006

16.2.3 Analysis of Color Bleeding Distortion

- Ringing in sub-sampled chroma channels propagates when up-sampled to full chroma resolution

© CRC Press, Wu & Rao Eds, 2006

16.3 Description of the Post-Processor



- Major components
  - Detection
  - Classification
  - Filtering

Left: Illustration of post-filtering process

© CRC Press, Wu & Rao Eds, 2006

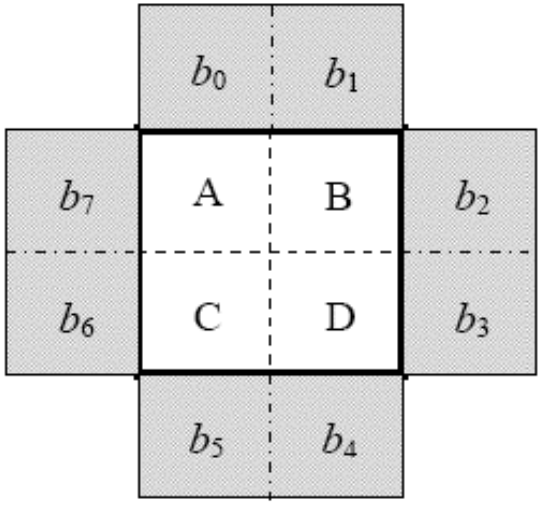
- Color bleeding detection
  - Variance computation in C_b and C_r
    - 4 x 4 windowed block
    - Computes activity level in chroma block
      - For chroma edge detection
  - Luminance edge detection
    - For vertical and horizontal edges
    - Used as additional reference for chrominance edges

© CRC Press, Wu & Rao Eds, 2006

- Computation of code block pattern
  - Classification of block pattern for filtering
  - Partitioning of 8 x 8 code block
    - To 4 quadrants (4x4 sub-block)
  - Binary flag for each block set if
    - Strong chrominance edge exist
    - Luminance edge detected
  - Pattern generated according to 4 flags
    - 16 possible patterns

© CRC Press, Wu & Rao Eds, 2006

Left: 4 quadrants of 8x8 code block. *Right*: Pattern configurations for code block classification



CONFIGURATION OF THE CURRENT 8x8 CHROMINANCE BLOCK				4x4 CHROMINANCE BLOCK TO BE FILTERED	4x4 NEIGHBORING BLOCKS TO BE ACCOUNTED FOR IN THE AVERAGING PROCESS
A	B	C	D		
0	0	0	0	–	–
0	0	0	1	A	b_0, b_7
0	0	1	0	B	b_1, b_2
0	0	1	1	A and B	b_0, b_1
0	1	0	0	C	b_5, b_6
0	1	0	1	A and C	b_6, b_7
0	1	1	0	–	–
0	1	1	1	A	b_0, b_7
1	0	0	0	D	b_3, b_4
1	0	0	1	–	–
1	0	1	0	B and D	b_2, b_3
1	0	1	1	B	b_1, b_2
1	1	0	0	C and D	b_4, b_5
1	1	0	1	C	b_5, b_6
1	1	1	0	D	b_3, b_4
1	1	1	1	–	–

© CRC Press, Wu & Rao Eds, 2006

- Filtering process
  - Ringing related color bleeding
    - Smoothing pixels around the target pixel (target pixel untouched)
    - Smoothing operation applied to all color channels
  - Blocking related color bleeding
    - Neighbourhood averaging (chroma channels)
    - Intra-filtering
      - Averaging of corrupted pixels identified by pattern configuration

© CRC Press, Wu & Rao Eds, 2006

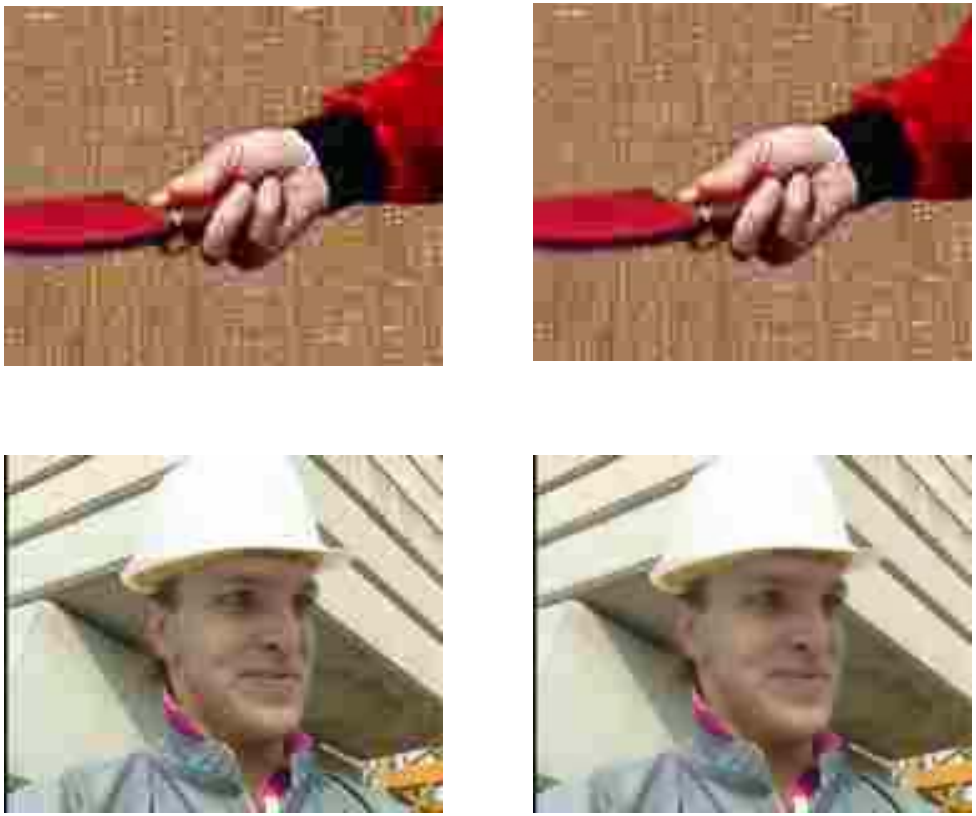
16.4 Experimental Results – Concluding Remarks

- Noticeable improvement with post-processing
  - Visual reduction of color bleeding
  - supported by PSNR results

Sequence (frame)	Δ PSNR (dB)	
	C_r component	C_b component
Table tennis (#001)	+0.32	+0.10
Foreman (#225)	+0.41	+0.43

© CRC Press, Wu & Rao Eds, 2006

Left: decode picture frames.
Right: post-filtered frames



© CRC Press, Wu & Rao Eds, 2006

AFR01	A. Al-Fahoum and M. Reza. Combined Edge Crispiness and Statistical Differencing dor Deblocking JPEG Compressed Images. *IEEE Trans. on Image Processing*, 10(9):1288-1298, Sept. 2001
BK95	V. Bhaskaran and K. Konstantinides. *Image and Video Compression Standards*. Norwell, MA: Kluwer Academic Publishers, 1955
CGC04	F.X. Coudoux, M. Gazalet and P. Corlay. An adaptive Post-processing Technique for the Reduction of Color Bleeding in DCT Coded Images. *IEEE Trans. on Circuits and Systems for Video Technology*, 14(1):114-121, Jan. 2004
Gle93	W. Glenn. Digital Image Compression Based on Visual POerception and Scene Properties. *SMPTE Journal*, 102(5):392-397, 1993
ISO94	ISO/IEC JTC1 110918-1. ITU-T Rec. T.8.1. Information Technology – Digital Compression and Coding of Continuous-tone Still Images: Requirements and Guidelines. Technical report, ISO/IEC, 1994
MH98	H.R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 4(11):317-320, Nov. 1997

© CRC Press, Wu & Rao Eds, 2006

MPFL96	J. Mitchell, W. Pennebaker, C. Fogg and D. LeGall. *MPEG Video Compression, 2nd Ed.*, New York and London: Plenum Press, 1995
NH95	A.N. Netravali and B.G. Haskell. *Digital Pictures-Representation, Compression, and Standards*. New York and London: Plenum Press, 2nd Ed., 1995
PL99	H. Park and Y. Lee. A Postprocessing Method for Reducing Quantization Effects in Low Bit-Rate Moving Picture Coding. *IEEE Trans. on Circuits and Systems for Video Technology*, 9(1):161-171, Feb. 1999
RY90	KK.R. Rao and P. Yip. *Discrete Cosine Transform – Algorithms, Advantages, Applications*. Boston, MA: Academic Press, 1990
vdBL96	C.J. van den Branden Lambrecht. *Perceptual Models and Architectures for Video Coding Applications*. Ph.D thesis, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland, 1996
Wan95	B.A. Wandell. *Foundations of Vision*. Sunderland, MA: Sinauer Associates Inc., 1995

© CRC Press, Wu & Rao Eds, 2006

Web96	J. Webb. Postprocessing to Reduce Blocking Artifacts for Low Bit-rate Video Coding Using Chrominance Information. In *Proc. of IEEE ICIP*, 2:9-12, Sept. 1996

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

17. Error Resilience for Video Coding Service

J. Zhang

National Information Communication Technology Australia (NICTA)

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 17 Outline

17.1 Introduction to Error Resilient Coding Technique

17.2 Error Resilient Coding Methods Compatible with MPEG-2

17.3 Methods for Concealment of Cell Loss

17.4 Experimental Procedure

17.5 Experimental Results

17.6 Conclusions

© CRC Press, Wu & Rao Eds, 2006

17.1 Introduction to Error Resilient Coding Technique

- Network transmission errors
  - Due network congestion, mis-routing and channel noise
    - Results in packet corruption and/or packet loss
  - Has severe impact on picture quality of video stream
  - Minimize transmission errors with error resilient coding

© CRC Press, Wu & Rao Eds, 2006

- Error resilient coding
  - Error detection
    - Know when and where errors occurs
  - Resynchronization
    - when video streams lose synchronization due to errors
  - Data recovery and error concealment
    - Estimate lost data
    - Minimize estimation errors

© CRC Press, Wu & Rao Eds, 2006

17.2 Error Resilient Coding Methods Compatible with MPEG-2

- Temporal localization
  - Minimizes propagation errors
- Spatial localization
  - Minimizes the effects of errors within picture frames
- Concealment processes
  - Minimizes effects of errors within picture frames

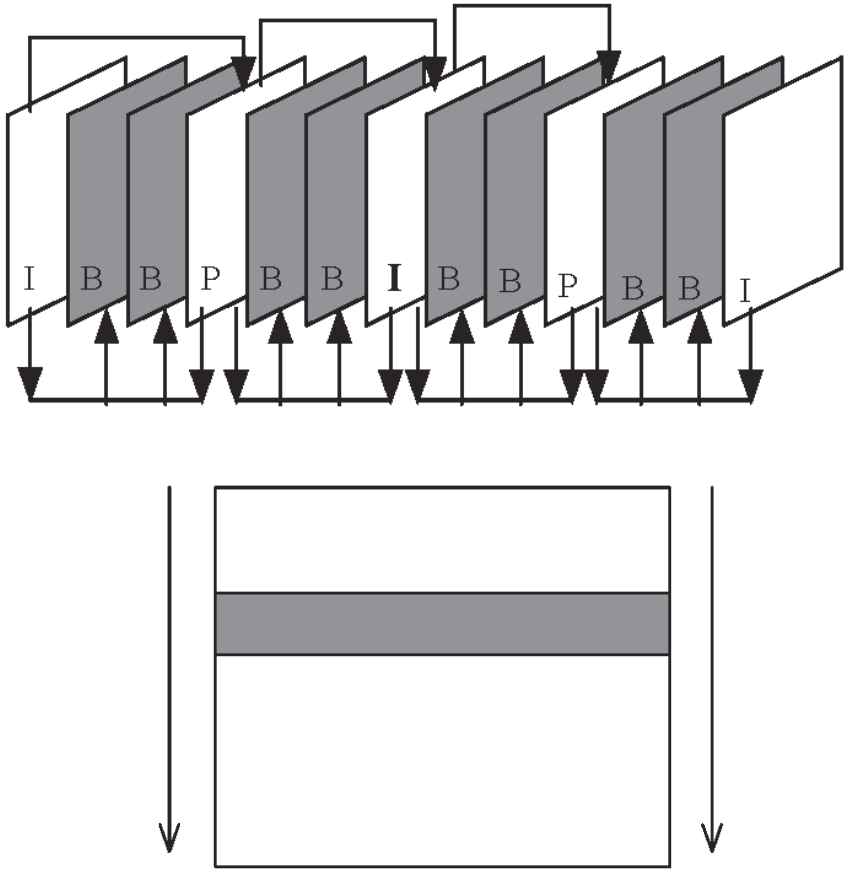
© CRC Press, Wu & Rao Eds, 2006

17.2.1 Temporal Localization

- Stop propagation error with early re-synchronization pictures
  - Cyclic intra-coded pictures
    - Replace damaged P or B frames with I frames
    - High cost in bitrate
  - Cyclic intra-coded slices
    - Intra-frame split into a number of row slices
    - Slices are embedded into a number of P-frames
      - One slice per P-frame
    - This approach leads to “window screen wipe”

© CRC Press, Wu & Rao Eds, 2006

Top: Cyclic intra-coded picture. *Bottom*: Cyclic intra-coded slices



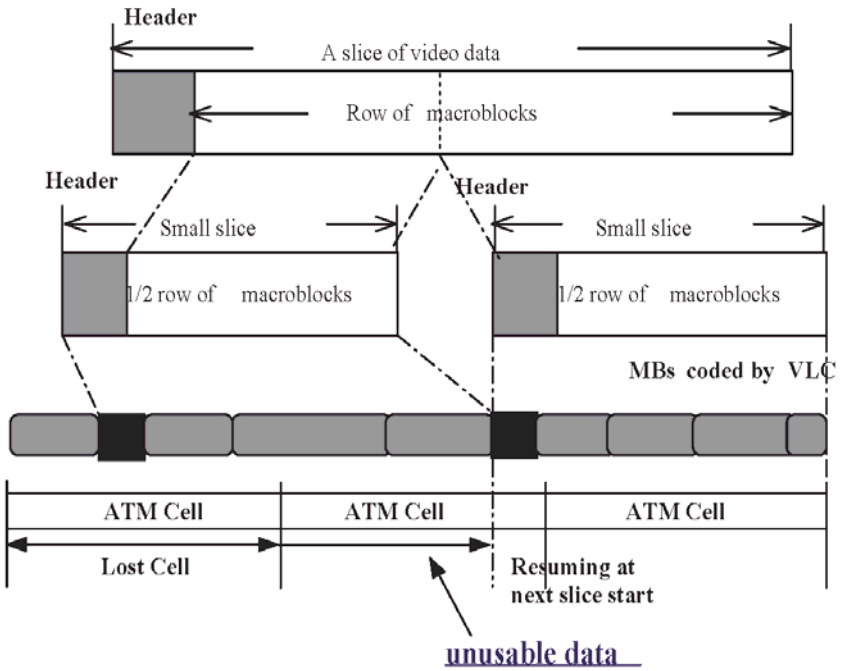
© CRC Press, Wu & Rao Eds, 2006

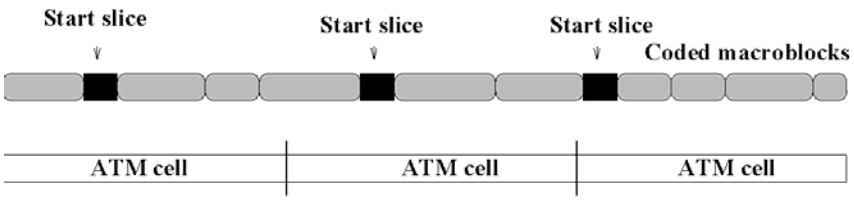
17.2.2 Spatial Localization

- Re-synchronization for differential coded macroblock (MB)
  - Small slice method
    - Minimize errors due to cell lost by
      - Reducing slice to cell ratio, i.e., having few cells per slice
  - Adaptive slice scheme
    - Slice size changes according to spacing of cells

© CRC Press, Wu & Rao Eds, 2006

Illustration of small slice (*left*) and adaptive slice (*right*) approaches





© CRC Press, Wu & Rao Eds, 2006

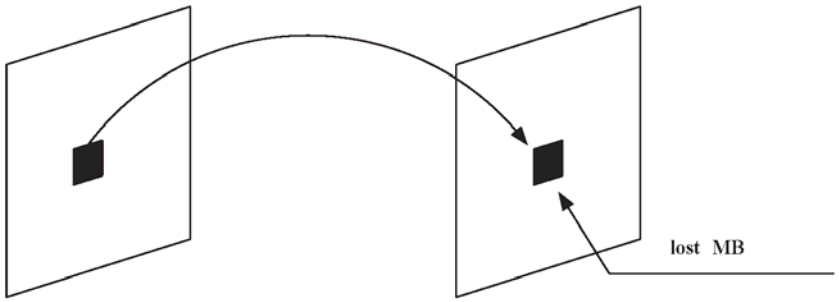
17.2.3 Concealment

- Methods for error concealment
  - Temporal Predictive Concealment
    - Block replacement from previous frame
  - Spatial Interpolation
    - Estimate missing MB with adjacent MBs
  - Motion Compensated Temporal Concealment
    - Motion compensated block replacement

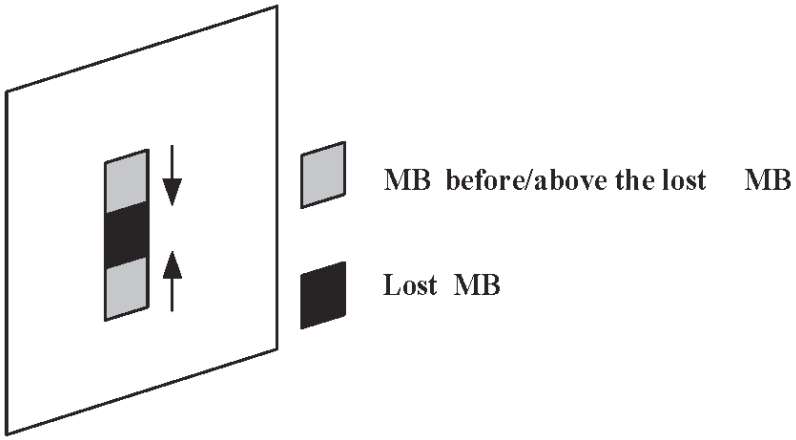
© CRC Press, Wu & Rao Eds, 2006

Concealment methods

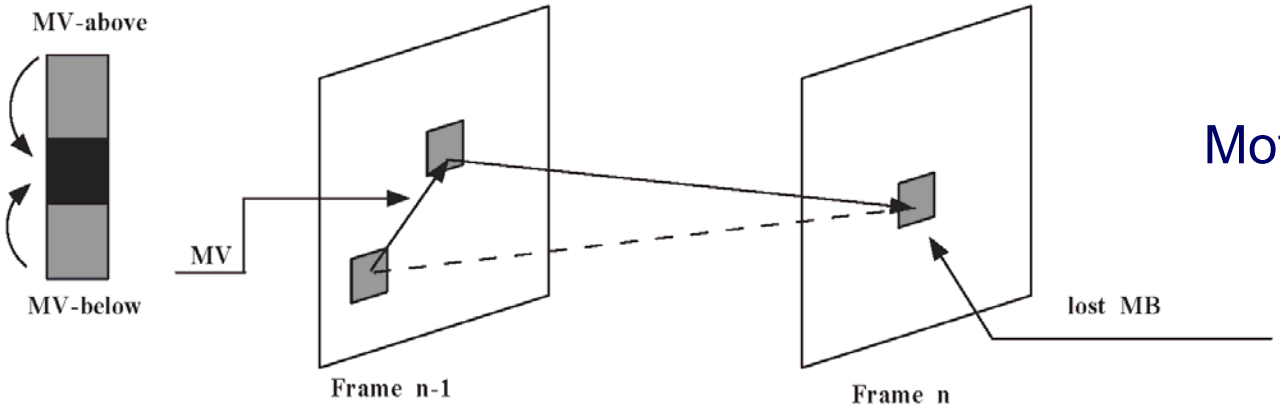
Temporal Predictive



Spatial Interpolation



Motion Compensated Concealment



© CRC Press, Wu & Rao Eds, 2006

Macro-block replacement (*right*) of the *Bus* sequence with cell loss (left)

© CRC Press, Wu & Rao Eds, 2006

Spatial interpolation error concealment (*right*) of the *Flower Garden* sequence with cell loss (left)

© CRC Press, Wu & Rao Eds, 2006

17.2.4 Scalability

- Multiple layers of information
  - Base layer has the most important information
  - Other layers carry additional information
- Applicable for ATM networks
  - Cells loss from congestion can be controlled
    - Network management can decide what cell to drop
- Ineffective for wireless networks
  - Cells loss uncontrollable
    - When cause by signal propagation

© CRC Press, Wu & Rao Eds, 2006

Scalability method

- Data partitioning
  - Split data into layers relative to importance
    - e.g. motion vectors, DC coefficient, AC coefficient
  - May have motion drift
    - When only base layer is available
      - Missing enhancement layers needed for MC
- SNR Scalability
  - Split into multiple quality layers
  - Motion compensation based on all layers
    - Motion drift occurs in the lower layer

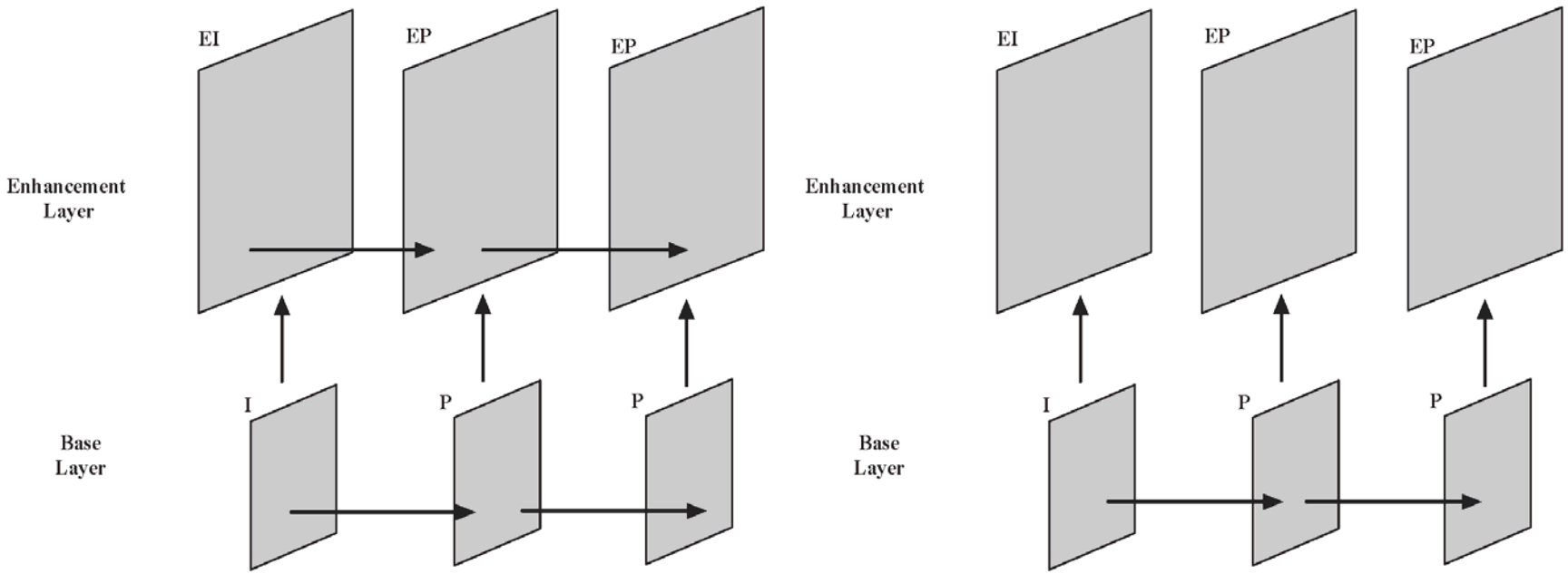
© CRC Press, Wu & Rao Eds, 2006

Scalability method

- Spatial Scalability
  - Motion compensation with separate picture store for each layer
    - Resolves the low layer motion drift
- Spatial Scalability without Temporal Prediction
  - Prediction in upper layers derived from base layer
    - Eliminates drift in upper layers

© CRC Press, Wu & Rao Eds, 2006

Illustration of spatial scalability with (*left*) and without (right) temporal prediction



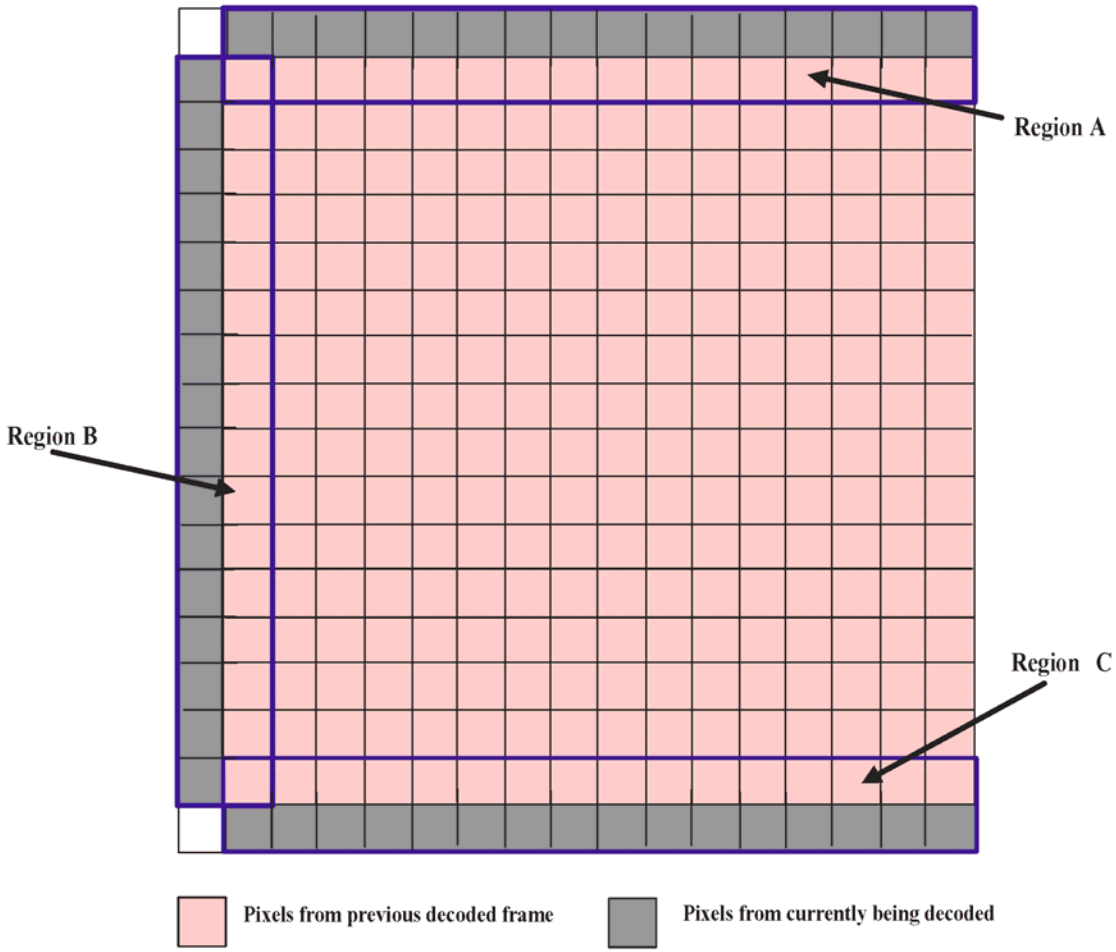
© CRC Press, Wu & Rao Eds, 2006

17.3 Methods for Concealment of Cell Loss

- *Boundary Matching Algorithm* (BMA)
  - Replacement block from previous frame is selected based on the best boundary match of current frame
  - 3 boundaries used for matching calculation
    - Top, left and bottom (of the lost block)
  - May employ full search or fast search
    - Fast searches relative to
      - Motion vector (MV) of same MB in previous frame
      - MV of a neighbouring MB
      - Average or median MV of neighbouring MBs

© CRC Press, Wu & Rao Eds, 2006

Illustration of boundary matching of BMA

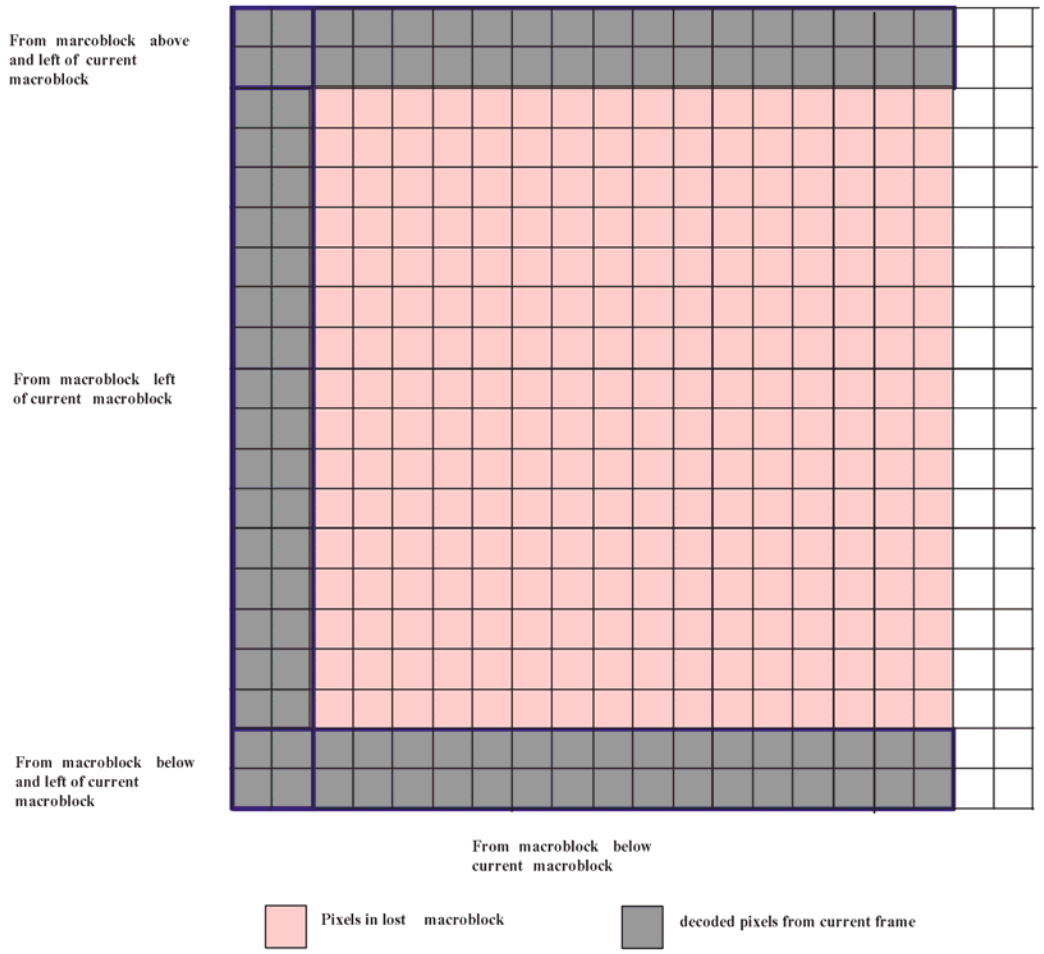


© CRC Press, Wu & Rao Eds, 2006

- *Decoder Motion Vector Estimation* (DMVE)
  - An improvement of the BMA
  - Uses additional boundaries for matching calculation
    - top right and bottom left
  - Performs full search for boundary matching

© CRC Press, Wu & Rao Eds, 2006

DMVE boundary matching



© CRC Press, Wu & Rao Eds, 2006

- Extension of DMVE
  - Utilize fast search as in BMA
    - *Optional Candidate Search* (OCS)
      - Based on MV of previous frame MB and neighbouring MB
  - Bi-directional frames prediction
    - If loss MB comes from B-frame
  - Two Line Search
    - Use only 2 lines surrounding lost MB for search
      - Reduces computation

© CRC Press, Wu & Rao Eds, 2006

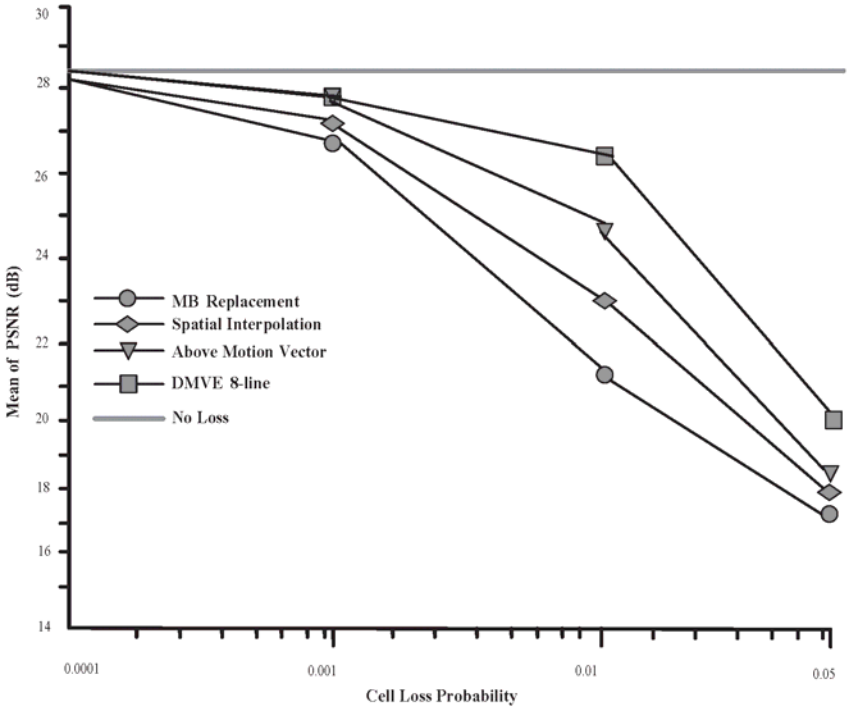
17.4 Experimental Procedure

- 2 MPEG-2 coded sequences @ 4 Mbps used for test
  - *Flower Garden* and *Bus*
- Cell loss probabilities set at 10^{-4}, 10^{-3}, 10^{-2} and 0.5
- 7 error concealment techniques compared

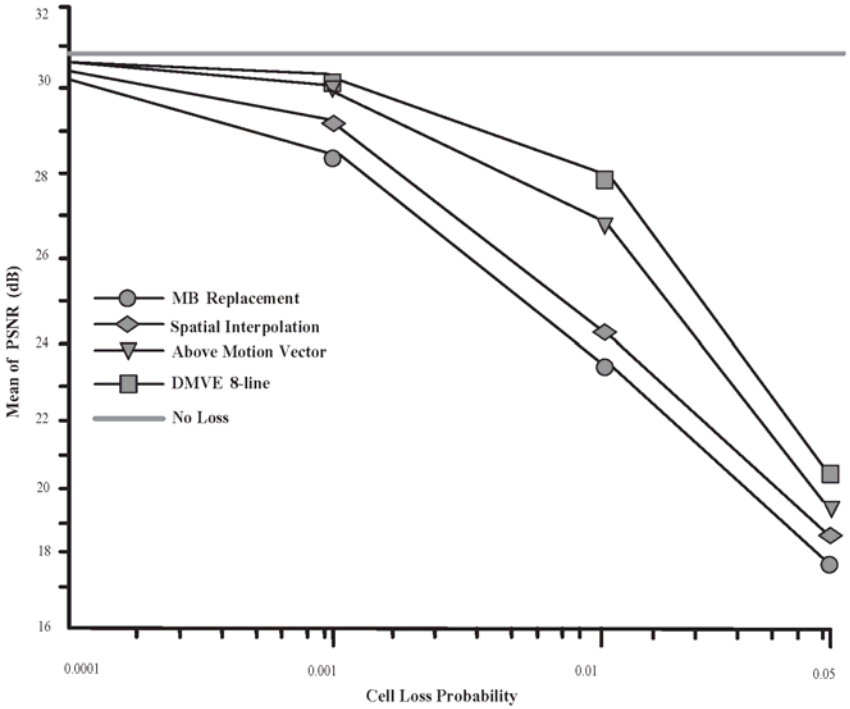
© CRC Press, Wu & Rao Eds, 2006

17.5 Experimental Results

PSNR vs Cell Loss



Flower Garden



Bus

© CRC Press, Wu & Rao Eds, 2006

- PSNR Observations
  - Low cell loss (10^{-4})
    - Little difference between various methods
  - Medium cell loss (10^{-3})
    - DMVE and Above Motion Vector noticeably better
  - High cell loss (10^{-2})
    - DMVE clearly the superior method
      - 6-7dB better than MB replacement
  - Very high cell loss (0.5)
    - Performance dropped substantially for all

© CRC Press, Wu & Rao Eds, 2006

Experimental results with cell loss of 10^{-4} (*left*) and 10^{-3} (*right*)

Measurements for *Flower Garden* sequence					
Mean of total damaged MBs = 92, PSNR with no loss = 28.38 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	28.16	28.30	27.88	0.13	1.0
Above Motion Vector	28.34	28.36	28.31	0.01	1.0
BMA	28.32	28.35	28.26	0.03	1.0
DMVE-OCS (8 line)	28.34	28.36	28.32	0.01	1.0
DMVE (2 line)	28.34	28.36	28.31	0.01	1.0
DMVE (8 line)	28.34	28.36	28.32	0.01	1.0
DMVE-Bi-Dir (8 line)	28.35	28.36	28.32	0.01	1.0

Measurements for Bus sequence					
Mean of total damaged MBs = 106, PSNR with no loss = 30.71 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	30.49	30.70	29.93	0.22	1.0
Above Motion Vector	30.67	30.71	30.57	0.04	1.0
BMA	30.64	30.70	30.51	0.06	1.0
DMVE-OCS (8 line)	30.68	30.71	30.60	0.03	1.0
DMVE (2 line)	30.67	30.70	30.57	0.04	1.0
DMVE (8 line)	30.68	30.71	30.61	0.03	1.0
DMVE-Bi-Dir (8 line)	30.69	30.71	30.61	0.03	1.0

Measurements for *Flower Garden* sequence					
Mean of total damaged MBs = 864, PSNR with no loss = 28.38 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	26.90	27.59	26.25	0.39	1.0
Above Motion Vector	28.11	28.26	27.77	0.13	1.0
BMA	28.03	28.21	27.69	0.13	1.0
DMVE-OCS (8 line)	28.13	28.25	27.96	0.09	1.0
DMVE (2 line)	28.16	28.26	27.91	0.10	1.0
DMVE (8 line)	28.17	28.26	27.98	0.08	1.0
DMVE-Bi-Dir (8 line)	28.18	28.27	27.99	0.07	1.0

Measurements for Bus sequence					
Mean of total damaged MBs = 865, PSNR with no loss = 30.71 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	29.24	29.72	28.80	0.32	1.0
Above Motion Vector	30.37	30.52	30.18	0.11	1.0
BMA	30.21	30.43	30.00	0.15	1.0
DMVE-OCS (8 line)	30.42	30.55	30.25	0.10	1.0
DMVE (2 line)	30.43	30.54	30.15	0.11	1.0
DMVE (8 line)	30.44	30.55	30.25	0.09	1.0
DMVE-Bi-Dir (8 line)	30.46	30.56	30.26	0.09	1.0

© CRC Press, Wu & Rao Eds, 2006

Experimental results with cell loss of 10^{-2} (*left*) and 0.5 (*right*)

Measurements for *Flower Garden* sequence					
Mean of total damaged MBs = 864, PSNR with no loss = 28.38 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	26.90	27.59	26.25	0.39	1.0
Above Motion Vector	28.11	28.26	27.77	0.13	1.0
BMA	28.03	28.21	27.69	0.13	1.0
DMVE-OCS (8 line)	28.13	28.25	27.96	0.09	1.0
DMVE (2 line)	28.16	28.26	27.91	0.10	1.0
DMVE (8 line)	28.17	28.26	27.98	0.08	1.0
DMVE-Bi-Dir (8 line)	28.18	28.27	27.99	0.07	1.0

Measurements for Bus sequence					
Mean of total damaged MBs = 865, PSNR with no loss = 30.71 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	29.24	29.72	28.80	0.32	1.0
Above Motion Vector	30.37	30.52	30.18	0.11	1.0
BMA	30.21	30.43	30.00	0.15	1.0
DMVE-OCS (8 line)	30.42	30.55	30.25	0.10	1.0
DMVE (2 line)	30.43	30.54	30.15	0.11	1.0
DMVE (8 line)	30.44	30.55	30.25	0.09	1.0
DMVE-Bi-Dir (8 line)	30.46	30.56	30.26	0.09	1.0

Measurements for *Flower Garden* Sequence					
Mean of total damaged MBs = 36374, PSNR with no loss = 28.38 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	17.45	17.55	17.36	0.05	1.0
Above Motion Vector	18.59	18.81	18.46	0.11	1.0
BMA	18.84	18.97	18.61	0.09	1.5
DMVE-OCS (8 line)	19.23	19.40	19.04	0.10	1.6
DMVE (2 line)	19.68	19.95	19.38	0.16	3.9
DMVE (8 line)	20.13	20.55	19.71	0.27	14.1
DMVE-Bi-Dir (8 line)	20.21	20.60	19.84	0.22	17.8

Measurements for Bus Sequence					
Mean of total damaged MBs = 38100, PSNR with no loss = 30.71 dB					
Technique	Single Slice Header per Row				
	Mean	Max	Min	Std. Dev	CPU Time
MB Replacement	17.75	17.95	17.59	0.11	1.0
Above Motion Vector	19.45	19.61	19.28	0.12	1.0
BMA	19.59	19.89	19.36	0.15	1.6
DMVE-OCS (8 line)	20.12	20.53	19.70	0.24	1.7
DMVE (2 line)	20.04	20.73	19.32	0.39	4.3
DMVE (8 line)	20.30	20.84	19.72	0.34	15.0
DMVE-Bi-Dir (8 line)	20.38	20.89	19.75	0.40	18.8

© CRC Press, Wu & Rao Eds, 2006

17.6 Conclusions

- Error resilient coding
  - Important for combating cell loss in networks
  - Several approaches
    - Spatial localization, MB resynchronization, concealment
  - Error concealment method
    - DVME demonstrates excellent performance at various cell loss probabilities
      - Further improvements if this method is combined with spatial localization and resynchronization methods

© CRC Press, Wu & Rao Eds, 2006

Ad-96	Ad-hic group on core experiments on error resilience aspect in MPEG 4 video. Description of Error Resilient Core Experiments. *ISO/IEC JTC1/SC29/WG11 MPEG-96*, Aug. 1996
ISO93	ISO/IEC JTC1/SC29/ WG11. MPEG 2 Test Model 4. *ISO/IEC*, 1993
ISO95	ISO/IEC JTC1/SC29/ WG11. Generic coding of moving pictures and associated audio information: Video. *ISO/IEC International Standard 13818-2*, 1995
J. 97	J. Zhang, M.R. Frater, J.F. Arnold and T.M. Percival. MPEG-2 Video Service for Wireless ATM Networks. *IEEE Journal on Selected Areas in Communications*, 15(1):119-128, Jan. 1997
J. 99	J. Zhang. *Error Resilience for Video Coding Services over Packer-based Networks*. Ph.D. thesis, University of New South Wales, 1999
M. 93a	M.. Ghanbari. Standard codecs: Image Compression to Advanced Video Coding. *IEE UK*, 1993
M. 93b	M. Ghanbari and V. Seferidis. Cell-loss concealment in ATM Video Codecs, *IEEE Trans. on Circuits and Systems for Video Technology*, 3:238-248, 1993

© CRC Press, Wu & Rao Eds, 2006

M. 96	M.D. Prycker. *Asynchronous Transfer Mode Solution for Broadband ISDN*. Upper Saddle River, NJ: Prentice-Hall, 3rd ed., 1996
R. 96	R. Aravind, M.R. Civanlar and A.R Reibman. Packet Loss Resilience of MPEG-2 Scalable Video Coding Algorithms. *IEEE Trans. on Circuits & Systems for Video Technology*, 6:426-435, Oct. 1996
Rec90	Recommendations and Report of the CCIR Recommendation of 601. Encoding Parameters of Digital Television of Studios: Broadcasting Service (television) Rec. 601-2. *CCIR*, 1990
W. 93	W.M. Lam, A.R. Reibman and B. Liu. Recovery of lost or erroneously received motion vectors. *Proc. ICASSP*, 5:417-420, 1993

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006

18. Critical Issues and Challenges

H.R. Wu† and K.R. Rao‡

‡ *Royal Melbourne Institute of Technology, Australia*

§ *University of Texas at Arlington, USA*

Slides for this presentation are not for general distribution without the expression permission of **CRC Press**

Chapter 18 Outline

18.1 Picture Coding Structures

18.2 Vision Modelling Issues

18.3 Spatio-Temporal Masking in Video Coding

18.4 Picture Quality Assessment

18.5 Challenges in Perceptual Coder Design

18.6 Codec System Design Optimization

18.7 Summary

© CRC Press, Wu & Rao Eds, 2006

18.1 Picture Coding Structures

- Picture coding structures
  - Not just a mathematical framework for reducing statistical redundancies
  - Need to consider human vision in the coding framework
    - To reduce visual redundancies
    - and help push the compression boundary further

© CRC Press, Wu & Rao Eds, 2006

- Coding systems often driven by applications
  - In turn driven by some practical limitations
    - Bitrate constrain is often the driving point in the development efforts for practical coding systems
    - Insufficient consideration given to picture quality in the development of coding systems
      - e.g. limited progress in addressing certain distortions in high definition digital TV
        - » Granular noise
        - » Stationary area temporal fluctuation
        - » Jerky motion

© CRC Press, Wu & Rao Eds, 2006

- Four directions for picture coding
  - Push boundaries of existing waveform and entropy coding techniques
  - Incorporating various HVS aspects into waveform coding techniques
  - Pursuing mathematical alternatives for waveform coding
  - Derive new coding framework and structures propelled by vision science and research

© CRC Press, Wu & Rao Eds, 2006

18.1.1 Performance Criteria

- Framework of coding systems
  - Determined by a set of performance criteria
  - Different criteria for different components of a coder, e.g.
    - Selection of transform is dependent on decorrelation and energy packing efficiency
    - Coding performance has traditionally been measured by MSE/PSNR for a given coding rate
  - Criterions are tailored towards reduction of statistical errors or redundancies

© CRC Press, Wu & Rao Eds, 2006

- Consequently, coding frameworks are determined
  - According to their mathematical properties
  - Performance in statistical redundancy reduction
- Choices of coding framework
  - Affects compression performance
  - Exhibit specific type of distortions
    - e.g. Blocking for DCT coders or ringing/edge rippling for DWT coders

© CRC Press, Wu & Rao Eds, 2006

- An assumption for lossy coding
  - Primarily removes statistical redundancy
  - Leads to a coding structure that reflects assumption
  - Strict definition of statistical redundancy
    - The removal of redundant data not the removal of information
  - Lossy coding removes both redundant data and information
- So assumption not longer holds!

© CRC Press, Wu & Rao Eds, 2006

- Paradigm shift may be necessary
  - From pure statistical model to a combination of statistical and psychovisual models
  - Such an approach could address more effectively issues concerning
    - Various types of picture impairments
    - Overall picture quality

© CRC Press, Wu & Rao Eds, 2006

18.1.2 Complete vs. Over-Complete Transform

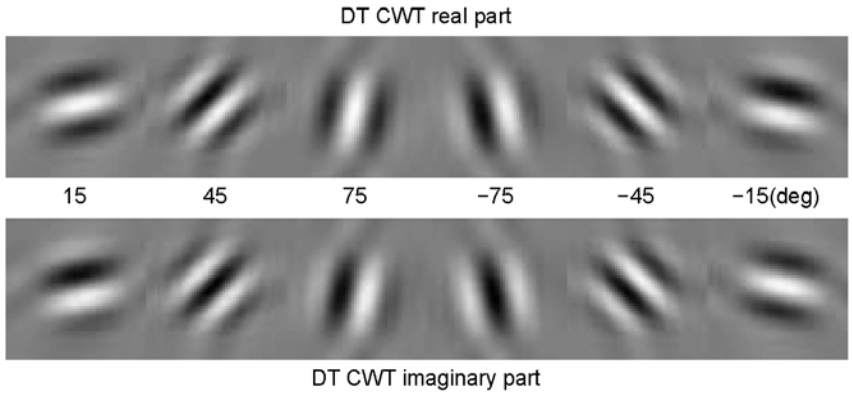
- Complete transform (common DWT)
  - Perfect reconstruction (reversible)
  - Good energy packing
  - Short filter support (FIR implementation)
  - Low complexity
  - Critically sample
    - Number of inputs = number of outputs
  - Lacks shift invariance
  - Limited directional selectivity

© CRC Press, Wu & Rao Eds, 2006

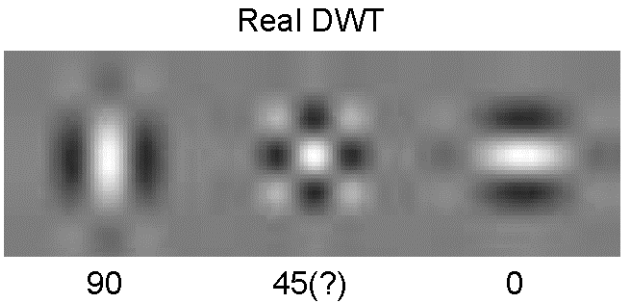
- Over-complete transform (CWT)
  - Good orientation selectivity
    - Relative to filter kernel
  - Over sampled
    - Number of Outputs > number of inputs
  - Shift invariant
  - More complex than complete transform

© CRC Press, Wu & Rao Eds, 2006

Basis function CWT and 2-D real DWT



2-D Q-shift Complex Wavelet Transform (CWT)

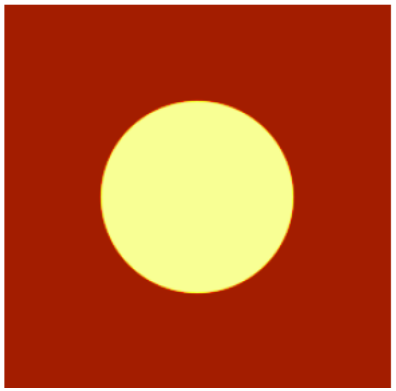


2-D real Discrete Wavelet Transform (DWT)

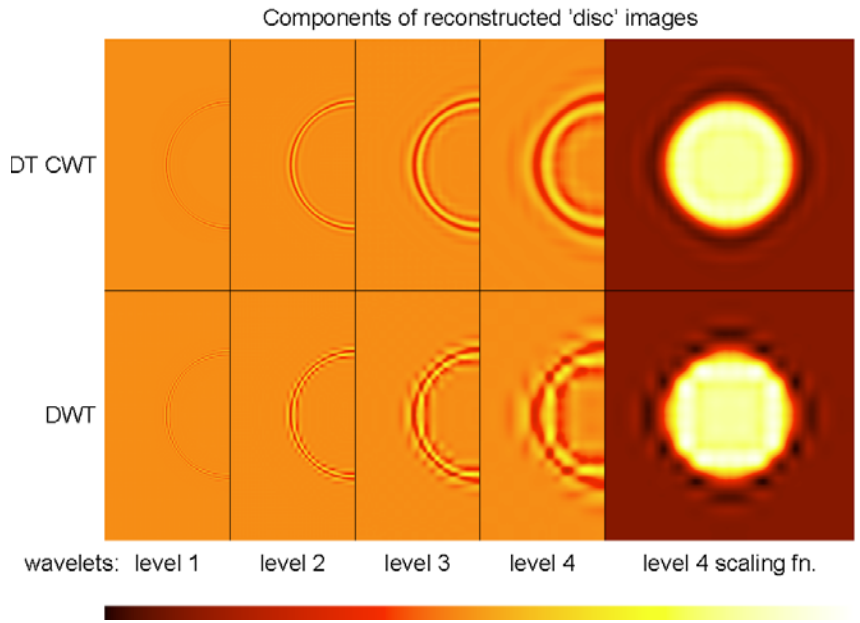
© CRC Press, Wu & Rao Eds, 2006

Illustration of wavelet and scaling functions component for CWT and DWT. Aliasing is visible in the DWT domain

Input (256 x 256)



Disc Image



© CRC Press, Wu & Rao Eds, 2006

- Aliasing in DWT domain
  - Due to shift invariance
    - Resulting from decimation of DWT coefficients
- In compression terms
  - CWT has demonstrated performance gains over DWT
    - ≈ 1dB PSNR gain and better visual quality
- What is the criteria for selecting coder transform?

© CRC Press, Wu & Rao Eds, 2006

DWT (*left*) and CWT (*right*) coded images

0.1994 bpp (PSNR: 33.47 dB)

0.1992 bpp (PSNR: 34.12 dB)

© CRC Press, Wu & Rao Eds, 2006

18.1.3 Decisions Decisions

- Each approach has it strengths and weaknesses
- How to quantify overall strengths over weaknesses?
  - Possibly joint performance criteria
    - Multiple criterions
    - Evaluate coding components individually and jointly
  - New criterions? Perhaps HVS based?

© CRC Press, Wu & Rao Eds, 2006

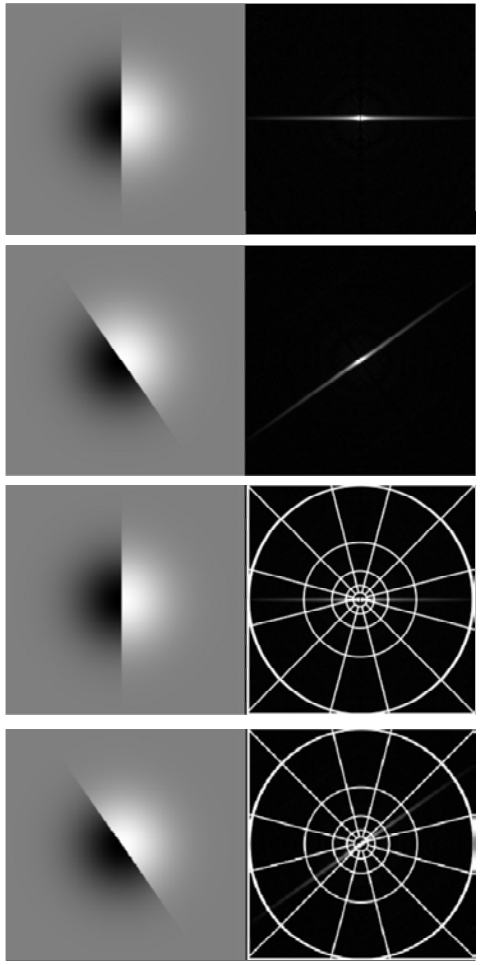
Example of 2-D aliasing and blurring tradeoffs

Original vertical disembodied edge

Original edge at 25° from vertical axis

Original vertical edge

Original edge at 45° from vertical axis



2-D Fourier spectrum of vertical edge

2-D Fourier spectrum of 25° edge

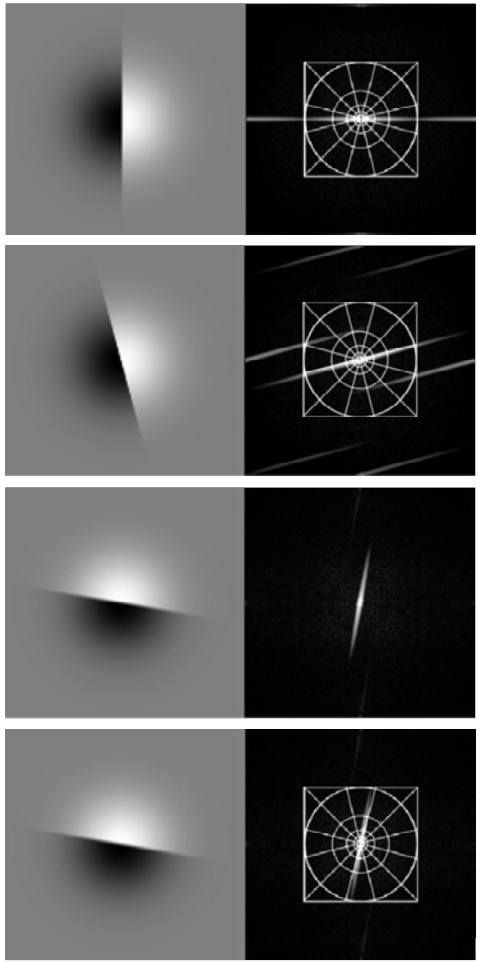
2-D Fourier spectrum of vertical edge showing Cortex channels

2-D Fourier spectrum of 25° edge showing Cortex channels

© CRC Press, Wu & Rao Eds, 2006

Example of 2-D aliasing and blurring tradeoffs

Sub-sampled vertical edge without anti-aliasing filtering



2-D Fourier spectrum edge showing Cortex channels

Sub-sampled edge at non-orthogonal orientation without anti-aliasing filter

2-D Fourier spectrum of edge

2-D Fourier spectrum edge showing Cortex channels

© CRC Press, Wu & Rao Eds, 2006

18.2 Vision Modelling Issues

- Limited understanding of overall human vision
  - Need to build up knowledge to model more complex higher level vision
- Better modelling of supra-threshold vision
  - Since considerable number of coding applications operate above the threshold level
- Optimizing vision models for applications
  - Devise more robust/comprehensive methods?

© CRC Press, Wu & Rao Eds, 2006

18.3 Spatio-Temporal Masking in Video Coding

- Transient and sustain channels in human vision
  - Segregates temporal information into low pass (sustain) and band pass (transient) bands
- Current standards have yet to employ temporal masking
  - To address temporal artifacts
    - e.g. Temporal granular noise and fluctuation

© CRC Press, Wu & Rao Eds, 2006

18.4 Picture Quality Assessment

- Picture quality metrics design approaches
  - MSE and PSNR insufficient for qualify visible picture impairments and picture quality
  - HVS based metric more suitable
    - But has a number of challenges
      - Computation and model complexity issue
      - Model accuracy
      - Optimization of models
      - Initial VQEG study indicated little gain for HVS based metric over PSNR

© CRC Press, Wu & Rao Eds, 2006

- Alternative assessment methods and issues
  - Measure impairments and quality
    - Can it measure enhancement as well?
  - FR, NR and RR objective metrics
    - Application specific
      - NR and RR have restricted range of applications
  - Subjective assessment
    - Impractical for many applications

© CRC Press, Wu & Rao Eds, 2006

- More challenges in picture quality assessment
  - Improvement needed for current picture assessment techniques to meet new applications, e.g., 3-D video/TV
  - Factors for consideration in assessment
    - e.g. picture accompanied by audio

© CRC Press, Wu & Rao Eds, 2006

18.5 Challenges in Perceptual Coder Design

- Pre- and post-processing
  - For removal of distortions
  - Enhancement of image details and overall visual quality
- Error detection and concealment
  - Embedded into coders for improve resilience from network induced noise

© CRC Press, Wu & Rao Eds, 2006

- Incorporating HVS into existing coders
  - Pre- and post-filtering
  - Quantization
  - Distortion metric for R-D optimized coders
    - EBCOT with self-masking function (Taubman)
    - PLC: EBCOT with PIDM
- Qualifying the lower bound for perceptually lossless picture coding
  - Perceptually lossless coding

© CRC Press, Wu & Rao Eds, 2006

18.6 Codec System Design Optimization

- Considerations for future generation picture coding framework
  - Pre- and/or post-filtering
  - Overall encoding and decoding system should be optimization
    - To better address existing coding issues
    - To improve compression performance and picture quality

© CRC Press, Wu & Rao Eds, 2006

18.7 Summary

- Goal of picture compression
  - Rate-Distortion minimization
    - Equivalently Rate-Quality optimization
- Most picture coders are rate driven
- Quality driven coders are rarities
- However, this may can if the lower bound for perceptually lossless coding is determined

© CRC Press, Wu & Rao Eds, 2006

Lena image coded at 0.125 bpp

Original

JPEG2000
MSE

JPEG2000
CVIS
(self-masking)

JPEG2000
PIDM

© CRC Press, Wu & Rao Eds, 2006

ABMD92	M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies. Image Coding Using Wavelet Transform. *IEEE Trans. on Image Processing*, 1(2):205-220, April 1992
AC96	R. Aravind and T. Chen. On Comparing MPEG-2 intraframe coding with a wavelet based codec. In M. Bigger, Ed., *Proc. of Picture Coding Symposium*, Melbourne, Australia, 1996
BCM00	M. Boliekm, C. Christopoulos and E. Majani. JPEG 2000 Part 1 Final Committee Draft Version 1.0. Technical report, ISO/IEC JTC1/SC9/WG1 N1646, Mar. 2000
Ber71	T. Berger. *Rate-Distortion Theory*, Englewood Cliffs, NJ: Prentice Hall 1971
Bud72	Z.L. Budrikis. Visual fidelity criterion and modeling. *Proc. of the IEEE*, 60(7):771-779, July 1972
Cav02	Special session on objective video quality metrics. In *Proc. of IEEE ICIP*, Rochester, New York, Sept. 2002
Cae99	P. Corriveau *et al*, All subjective scales are not created equal: The effect of context on different scales. *Signal Processing*, 77:1-9, 1999

© CRC Press, Wu & Rao Eds, 2006

Che81	M.J. Cheng. *Lao-Tzu: My words are very easy to understand*. Translated by T.C. Gibbs. Richmond, California: North Atlantic Books, 1981. Chinese ed., Taipei, Taiwan: Chung Hwa Book Company, Ltd., 1971
Cla85	R.J. Clarke. *Transform Coding of Images*. New York: Academic Press, 1985
Cla95	R.J. Clarke. *Digital Compression of Still Images and Video*. San Diego, CA: Academic Press, 1995
Cut52	C.C. Cutler. *Differential Quantization of Communication Signals*. US Patent No. 2,605,361, July 1952
Dal93	S. Daly. The visible difference predictor: An algorithm for the assessment of image fidelity. In A.B. Watson, Ed., *Digital Images and Human Vision*, 179-206. Cambridge, MA: The MIT Press, 1993
Dal02	S. Daly. Using spatio-temporal-chromatic visual models for perceived quality optimization. In *International Congress of Imaging Science*, Tokyo, Japan, May 2002

© CRC Press, Wu & Rao Eds, 2006

Dau85	J.G. Daugman. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A*, 2(7):1160-1169, 1985
Dau88	J. Daugman. Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 36(7):1169-1179, July 1988
DFS02	S. Daly. X. Feng and J. Speigle. A few practical applications that require some of the more advanced features of current visual models. In *SPIE/IS&T Electronic Imaging Conference*, 4662:70-83, San Jose, CA, Jan. 2002
DV01	M.N. Do and M. Vetterli. Pyramidal directional filter banks and curvelets. In *Proc. of IEEE ICIP*, 3:158-161, Oct. 2001
DV02	M. Do and M. Vetterli. Countourlets: A directional multiresolution image representation. In *Proc. of IEEE ICIP*, 1:357-360, Sept, 2002
DV03	M. Do and M. Vetterli. The finite ridgelet transform for image representation. *IEEE Trans. on Image Processing*, 12:16-28, Jan. 2003

© CRC Press, Wu & Rao Eds, 2006

DVV02	P.L. Dragotti, M. Vetterli and V. Velisavljevic. Directional wavelets and wavelet footprints for compression and denoising. In *Proc. International Workshop of Advanced Methods for Multimedia Signal Processing*, Capri, Italy, Sept. 2002
DZLL00	S. Daly, W. Zeng, J. Li and S. Lei. Visual Masking in Wavelet Compression for JPEG2000. In *IS&T/SPIE Conf. on Image and Video Communication and Processing*, 3974, Jan. 2000
EB93	M.P. Eckert and G. Buchsbaum. The significance of eye movement and image acceleration for coding television image sequences. In A.B. Watson, Ed., *Digital Images and Human Vision*, 89-98. Cambridge, MA: The MIT Press, 1993
EB98	M.P. Eckert and A.P. Bradley. Perceptual quality metrics applied to still :, 1998image compression. *Signal Processing*,
Ebr05	T. Ebrahimi, Ed., Proceedings of Workshop on Future Directions in Video Compression. Busan, Korea, April 2005. MPEG/JVT

© CRC Press, Wu & Rao Eds, 2006

F. 05	F. Yang, S. Wan, Y. Chang and H.R. Wu. A Novel Objective No-Reference Metric for Digital Video Quality Assessment. *IEEE Signal Processing Letters*, 2005
FG03	M. Flierl and B. Girod. Generalized B pictures and the draft H.264/AVC video compression standard. *IEEE Trans. on Circuits and Systems for Video Technology*, 13(7):587-597, July 2003
FvdBL02	J.E. Farrel and C. van den Branden Lambrecht, eds., Special issue on translating human vision research into engineering technology. *Proc. of the IEEE*, 90(1):3-169, Jan. 2002
FWG02	M. Fierl, T. Wiegand and B. Girod. Rate-constrained multihypothesis prediction for motion-compensated video compression. *IEEE Trans. on Circuits and Systems for Video Technology*, 12(11):957-969, Nov. 2002
Gir88	B. Girod. Eye movements and coding of video sequences. In T.R. Hsing, Ed., *SPIE Proc. Visual Communications and Image Processing*, 1001:398-405. Cambridge, MA:SPIE, Nov. 1988
Gir03	B. Girod. Video coding for compression and beyond. In *Proc. of IEEE ICIP*, Barcelona, Spain, Sept. 2003

© CRC Press, Wu & Rao Eds, 2006

GJR96	N. García, F. Jaureguizar and J.R. Ronda. Pixel-based video compression schemes. In L. Torres and M. Kunt, Eds., *Video Coding-The Second Generation Approach*, 31-78. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996
Goo51	W.M. Goodall. Television by pulse code modulation. *Bell System Technical Journal*, 33-49, Jan. 1951
HDZG03	C. He, J. Dong, Y.F. Zheng and Z. Gao. Optimal 3-D coefficient tree structure for 3-D wavelet video coding. *IEEE Trans. Circuits and Systems for Video Technology*, 13(10):961-972, Oct. 2003
HS92	R.F. Hess and R.J. Snowden. Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research*, 32(1):47-59, 1992
Hua65	T.S. Huang. PCM picture transmission. *IEEE Spectrum*, 5:57-63, Dec. 1965
IL96	A.F. Inglis and A.C. Luther. *Video Engineering*, New York, NY:McGraw-Hill, 2nd ed., 1996

© CRC Press, Wu & Rao Eds, 2006

ISO93a	ISO/IEC JTC1/SC29. ISO 11172-2 Information Technology: Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s. Technical report, ISO/IEC, 1993
ISO93b	ISO/IEC JTC1/SC29/WG11. CD 13818, Generic Coding of Moving Pictures and Associated Audio. Technical report, ISO/IEC, Nov. 1993
ITU96	ITU-T. ITU-T Recommendation H.264, Video Coding for Low Bit Rate Communication. Technical report, ITU-T, 1996
ITu97	ITU-R. *Investigation of Contextual Effects*. ITU-R, Mar. 1997, ITU-R Document 11E/34-E, 24
ITU03	ITU-T. ITU-T Recommendation H.264, Advance Video Coding for Generic Audiovisual Services. Technical report, ITU-T, 2003
JJS93	N. Jayant, J. Johnston and R. Safranek. Signal compression based on models of human perception. *Proc. of the IEEE*, 81(10):1385-1422, Oct. 1993

© CRC Press, Wu & Rao Eds, 2006

JN84	N.S. Jayant and P. Noll. *Digital Coding of Waveforms, Principles and Applications to Speech and Video*. Upper Saddle River, NJ: Prentice Hall, 1984
Joh88	J.D. Johnston. Transform coding of audio signals using perceptual noise criteria. *IEEE Journal on Selected Areas in Communications*, 6(2):314-323, Feb. 1988
Kin01	N.G. Kingsbury. Complex wavelets for shift invariant analysis and filtering of signals. *Journal of Applied and Computational Harmonic Analysis*, 10(3):234-253, May 2001
Kle93	S.A. Klein. Image quality and image compression: A psychophysicist's viewpoint. In A.B. Watson. Ed., *Digital Images and Human Vision*, 73-88. Cambridge, MA: The MIT Press, 1993
KR03	N.G. Kingsbury and T.H Reeves. Iterative image coding with overcomplete complex wavelet transform. In *Proc. of SPIE Conference on Visual Communications and Image Processing*, Lugano, Switzerland, July 2003
KTR05	S.K. Kwon, A. Tamhankar and K.R. Rao. Overview of H.264/MPEG-4 Part 10. *J. VCIR*, 2005

© CRC Press, Wu & Rao Eds, 2006

LSW03	A. Luthra, G. Sullivan and T. Wiegand. Special issue on the H.264/AVC video coding standard. *IEEE Trans. Circuits Systems for Video Technology*, 13(7):557-725, July 2003
LYC97	N. Lin, T. Yu and A.K. Chan. Perceptually lossless wavelet-based compression for medical images. *Proc. SPIE Medical Imaging: Image Display*, 3031:763-770, 1997
MDMS05	R. Mantiuk, S. Daly, K. Myszkowski and H. Seidel. Predicting visible difference sin high-dynamic-range images. Model and its calibration. In *SPIE Electronic Imaging Conference: Human Vision and Electronic Imaging*, 5666:204-214, 2005
NZ97	R. Neff and A. Zakhor. Very low bit-rate video coding based on matching pursuits. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):158-171, July 1997
NZ02a	R. Neff and A. Zakhor. Matching pursuit video coding I. Dictionary approximation. *IEEE Trans. Circuits and Systems for Video Technology*, 12(1):13-26, Jan. 2002

© CRC Press, Wu & Rao Eds, 2006

NZ02b	R. Neff and A. Zakhor. Matching pursuit video coding II. Operational models for rate and distortion. *IEEE Trans. on Circuits and Systems for Video Technology*, 12(1):27-39, Jan. 2002
NZV94	R. Neff, A Zakhor and M. Vetterli. Very low bitrate video coding using matching pursuits. In *Proc. of SPIE Visual Communications and Image Processing*, 2308:47-60, Chicago, IL, Sept. 1994
OS89	A.V. Oppenheim and R.W. Schafer. *Discrete-Time Signal Processing*. Englewood Cliffs., NJ: Prentice-Hall, 1989
Ree93	T.R. Reed. Local frequency representations for image sequence processing and coding. In A.B. Watson, ed., *Digital Images and Human Vision*, 3-12, Cambridge, MA: The MIT Press, 1993
RH96	K.R. Rao and J.J. Hwang. *Techniques and Standards for Image, Video and Audio Coding*. Upper Saddle, NJ: Prentice Hall, 1996
RY90	K.R. Rao and P. Yip. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. San Diego, CA: Academic Press, 1990

© CRC Press, Wu & Rao Eds, 2006

SB59	A.J. Seyler and Z.L. Budrikis. Measurements of temporal adaptation to spatial detain vision. *Nature*, 184:1215-1217, Oct. 1959
SB65	A.J. Seyler and Z.L. Budrikis. Detail perception after scene changes in television iamge presentations. *IEEE Trans. Information Theory*, 11(1)”61-43, 1965
SF95	E.P. Simoncelli and W.T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proc. IEEE ICIP*, 175-184, 1995
SFAH92	E.P. Simoncelli, W.T. Freeman, E.H. Adelson and D.J. Heeger. Shiftable Multiscale Transform. *IEEE Trans. on Information Theory*, 38(2):587-607, 1992
SFW+00	R.M. Slone, D.H. Foos. B.R. Whiting, E. Muka, D.A. Rubin, T.K. Pilgram, K.S. Kohm, S.S. Young, P. Ho and D.D. Hendrickson. Assessment of visually lossless irreversible image compression: Comparison of thress methods by using an image-comparison workstation. *Radiology-Computer Applications*, 215(2):543-553, 2000

© CRC Press, Wu & Rao Eds, 2006

Sha48	C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379-623, 1948
Sha93	J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing*, 41(12):3445-3462, 1993
SP96	A. Said and W.A. Pearlman. A new fast and efficient image codec based in set partitioning in hierarchical trees. *IEEE Trans. on Circuits and Systems for Video Technology*, 6:243-250, June 1996
SR87	P. Santago and S.A. Rajala. Using convex set techniques for combined pixel and frequency domain coding of time-varying images. *IEEE Jouranl on Selected Areas in Communications*, SAC-5(7):1127-1139, Aug. 1987
STM+00	L.B. Stelmach, W.J. Tam, D.V. Meegan, A. Vincent and P. Corriveau. Human perception of mismatched stereoscopic 3D inputs. In *Proc. ICIP*, Vancouver, BC, Canada, 2000
Tau00	D. Taubman. High performance scalable image compression with EBCOT. *IEEE Trans. on Image Processing*, 9:1158-1170, July 2000

© CRC Press, Wu & Rao Eds, 2006

THD84	K. Tzou, R. Hsing and J.G. Dunham. Applications of physiological human visual system model to image compression. In *SPIE: Applications of Digital Image Processing VII*, 504:419-424, 1984
TK96	L. Torres and M. Kunt. *Video Coding – The Second Generation Approach*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1996
TTW04	C.S. Tan, D.M. Tan and H.R. Wu. Perceptual coding of digital colour images based on a vision model. In *Proc. of IEEE International Symposium on Circuits and Systems*, 441-444, Vancouver, Canada, May 2004
TW99a	D.M. Tan and H.R. Wu. Multi-dimensional discrete cosine transform for image compression. In *Proc. of IEEE International Conference in Electronics, Communications and Systems*, T1E2.1-2.5, Sept. 1999
TW99b	D.M. Tan and H.R. Wu. Quantisation design for multi-dimensional dct for image coding. In *Proc. of IEEE Region Ten Conference*, 1:198-201, Cheju, Korea, Sept. 1999
TW03a	D. Tan and H. Wu. Perceptual lossless coding of digital monochrome images. In *Proc. of International Symposium on Intelligent Signal Processing and Communication Systems*, Awaji Island, Japan, Dec. 2003

© CRC Press, Wu & Rao Eds, 2006

TW03b	D.M. Tan and H.R. Wu. Adaptation of visually lossless colour coding to JPEG2000 in the composite colour space. In *Proc. of the 4th International Conference on Information, Communications & Signal Processing and 4th Pacific Rim Conference on Multimedia*, 1B2.4.1-7, Singapore, Dec. 2003
YWY04	D.M. Tan, H.R. Wu and Z. Yu. Perceptual coding of digital monochrome images. *IEEE Signal Processing Letters*, 11(2):239-242, Feb. 2004
van96	C. J. van den Branden Lambrecht. *Testing Digital Video System and Quality Metrics based on Perceptual Models and Architecture*, Ph.D. thesis, EPFL, Lausanne, Switzerland, May 1996
van98	C. J. van den Branden Lambrecht. Editorial. *Signal Processing*, 70:153-154, 1998
vdBL98	C. J. van den Branden Lambrecht, ed., Special issue on imaeg and video quality metrics. *Signal Processing*, 70(3):153-294, Nov. 1998
Vet01	M. Vetterli. Wavelets approximation and compression. *IEEE Signal Processing Magazine*, 18(5):59-73, Sept. 2001

© CRC Press, Wu & Rao Eds, 2006

VK95	M. Vetterli and J. Kovacevic. *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice Hall, 1995
VQE00	VQEG. Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment, Mar. 2000.
Wan95	B.A. Wandell. *Foundations of Vision*. Sunderland, MA: Sinauer Associates, Inc., 1995
Wat87	A.B. Watson. The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics and Image Processing*, 39(3):311-327, 1987
Wat89	A.B. Watson. Receptive fields and visual representations. In *SPIE Proceedings*, 1077:190-197, 1989
Wat93a	A.B. Watson. DCTune: A technique for visual optimization of DCT quantization matrices for individual images. *Society for Information Display Digest of Technical Papers XXIV*, 946-949, 1993
Wat93b	A.B. Watson, ed., *Digital Images and Human Vision*, Cambridge, MA: The MIT Press, 1993

© CRC Press, Wu & Rao Eds, 2006

Wat98	A.B. Watson. Toward a perceptual video quality metric. In *Human Vision and Electronic Imaging III*, 3299:139-147. Bellingham WA: SPIE Press, 1998
WTB97	A.B. Watson, M. Taylor and R. Borthwick. DCTune perceptual optimization of compressed dental X-rays. In Y. Kim, ed., *Proceedings of SPIE Medical Imaging*, 3031:358-371, 1997
WTB+04	D. Wu, D. M. Tan, M. Baird, J. DeCampo, C. White and H.R. Wu. Perceptually lossless medical image coding. In *Proc. of the International Symposium on Intelligent Multimedia, Video & Speech Processing*, HK, Oct. 2004
WTW03	D. Wu, D.M. Tan and H.R. Wu. Visually lossless adaptive compression of medical images. In *Proc. of the 4th International Conference on Information, Communications and Signal Processing and 4th Pacific Rim Conference on Multimedia*, Singapore, Dec. 2003
Wuv00	Special Session on testing and quality metrics for digital video services. In K. N. Ngan, T. Sikora and M.T. Sun, eds., *SPIE Proceedings, Visual Communications and Image Processing*, 4067:741-809, Peth, Australia, 2000

© CRC Press, Wu & Rao Eds, 2006

WYQ02	H.R. Wu, Z. Yu and B. Qiu. Multiple reference impairment scale subjective assessment method for digital video. In *Proc. of International Conference on Digital Signal Processing*, 1:185-189, Santorini, Greece, July 2002
YLL+03	X. Yang, W. Lin, Z. Lu, E.P. Ong and S. Yao. Perceptually adaptive hybrid video encoding based on just-noticeable-distortion profile. In *SPIE Proc. of VCIP*, 1448-1459, 2003
YWWC02	Z.Yu, H.R. Wu, S. Winkler and T. Chen. Vision model based impairment metric to evaluate blocking artifacts in digital video. *Proc. of the IEEE*, 90(1):154-169, Jan. 2002
Zha04	X. Zhang. Just-noticeable distortion estimation for images. Masters thesis. School of Electrical and Electronic Engineering, Nanyang Technological University. Singapore, 2004

© CRC Press, Wu & Rao Eds, 2006

© CRC Press, Wu & Rao Eds, 2006